IEEE JOURNAL ON

SELECTED AREAS IN COMMUNICATIONS

JUNE 1990 VOLUME 8 NUMBER 5 (ISSN 0733-8716)

A PUBLICATION OF THE IEEE COMMUNICATIONS SOCIETY

SPREAD SPECTRUM COMMUNICATIONS II

Guest Editors—D. L. Schilling, R. L. Pickholtz, and L. B. Milstein

PAPERS

Spread-Spectrum Design and Analysis

Jamming and Interception

CALLS FOR PAPERS

IEEE COMMUNICATIONS SOCIETY

The field of interest of the IEEE Communications Society consists of all telecommunications including telephone, telegraphy, facsimile, and point-to-point television, by electromagnetic propagation including radio; wire; aerial, underground coaxial, and submarine cables; waveguides, communication satellites, and lasers; in marine, aeronautical, space, and fixed station services; repeaters, radio relaying, signal storage, and regeneration; telecommunication error detection and correction; multiplexing and carrier techniques; communication switching systems; data communications; and communication theory.

In addition to the above, this JOURNAL or the IEEE TRANSACTIONS ON COMMUNICATIONS contains papers pertaining to analog and digital signal processing and modulation, audio and video encoding techniques, the theory and design of transmitters, receivers, and repeaters for communications via optical and sonic media, the design and analysis of computer communication systems, and the development of communication software. Contributions of theory enhancing the understanding of communication systems and techniques are included, as are discussions of the social implications of the development of communication technology. All members of the IEEE are eligible for membership in the Society upon payment of the annual Society membership fee of $12.00. Members may receive this JOURNAL or the IEEE TRANSACTIONS ON COMMUNICATIONS upon payment of an additional $10.00 ($22.00 total), or both publications upon payment of an additional $20.00 ($32.00 total). For information on joining, write to the IEEE at the address below.

IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS is published nine times a year in January, February, April, May, June, August, September, October, and December by The Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Headquarters:** 345 East 47 Street, New York, NY 10017. **NY Telephone:** 212-705 + extension: Information -7900; General Manager -7910; Controller -7748; Public Information -7867; Publishing Services -7560; Spectrum -7556; Technical Activities -7890. **Telecopiers:** NY (Headquarters) 212-752-4929, NY (Publications/Standards) 212-682-6005; **NY Telex:** 236-411 (international messages only). **IEEE Service Center** (for orders, subscriptions, address changes, Educational Activities, Region/Section/Student Services, Standards): 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ 08855-1331. **NJ Telephone:** 201-981-0060. **IEEE Washington Office** (for U.S. professional activities): 1828 L Street, NW, Suite 1202, Washington, DC 20036-5104. **Washington Telephone:** 202-785-0017. **Price/Publication Information:** Individual copies: IEEE members $10.00 (first copy only), nonmembers $20.00 per copy. (Note: Add $4.00 postage and handling charge to any order from $1.00 to $50.00, including prepaid orders.) Member and nonmember subscription prices available on request. Available in microfiche and microfilm. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limits of the U.S. Copyright Law for private use of patrons: 1) those post-1977 articles that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 29 Congress Street, Salem, MA 01970; 2) pre-1978 articles without fee. Instructors are permitted to photocopy isolated articles for noncommercial classroom use without fee. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publishing Services, 345 East 47th Street, New York, NY 10017. Copyright © 1990 by The Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Second-class postage paid at New York, NY, and at additional mailing offices. **Postmaster:** Send address changes to IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE, 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ 08855-1331.

New Radio Networks for Tactical Communication

JOHN ERIK RUSTAD, REIDAR SKAUG, AND ANDREAS AASEN

Abstract—**Future military tactical communication networks must be highly mobile, survivable, and reconfigurable. Basically, the existing systems consist of a digital trunk network of switches mounted in vehicles and interconneted by multichannel radio relays. Subscriber access takes place directly by field lines through multiplexer equipment which cosite the switches or indirectly through radio relay tails.**

Some form of node access is necessary in order to service the most mobile units and to make extensive use of trunk bearer capacity. This paper presents a packet-switched radio system designed and developed for the Norwegian Army. In particular, the paper discusses the development of an overall network concept, the network and access protocols, and the routing strategy. Network synchronization and radio access protocols are discussed and simulation results are given. Finally, the radio design allowing integration of ECM protected data and voice is presented.

I. Introduction

TACTICAL communication should support the operational units in the field, and must therefore reflect the strategy of the forces. A flexible threat reaction demands very mobile units which may be spread over a large geographical area. If the forces are to operate under a centralized management and at the same time retain their mobility, heavy demands are put on the communication system. These demands will be in the form of security, survivability, and protection against electronic warfare.

The trunk network is today the major tactical communication facility. Its backbone consist of switches mounted in vehicles and interconnected by multichannel radio relays.

Subscriber access takes place directly by field lines through multiplexer equipment which cosite the switches or indirectly through radio relay tails. Present systems mix analog and digital technology where the needs for sophisticated data traffic are generally not considered. However, the trunk network may provide, where implemented, for interconnection of packed nodes by initializing allocated channels within the multiplex structure.

Current trunk systems are not appropriate for extending services to the most mobile units or to allow connections of a large number of subscribers through tail links. A large number of tail links is costly and impractical, and will reduce the mobility and reconfigurability of the network due to difficulties in obtaining sites with a sufficient number of line-of-sight (LOS) paths.

Some form of radio access is therefore seen to be necessary in order to service the most mobile units and to make effective use of trunk bearer capacity without severely degrading the mobility.

At present, radio access has been solved by single-channel radio access (SCRA) systems or by providing special access for combat net radios (CNR) [11]; however, this paper will present a packet radio system which may both provide radio access to the trunk network and also work as an autonomous radio area network.

Such a network carrying a major part of information exchange in the battlefield has a number of mandatory operational requirements. A number of them will directly influence system parameters such as type of signaling, routing procedures, type of protocols, and synchronization.

First of all, the system characteristics must be based on a homogeneous concept allowing information to flow both in a hierarchical manner as well as to pass across chains of command. It should be transparent to the user, whether he is directly connected to the trunk network or must access this network through the radio system.

A number of important requirements for the radio system are:

1) ESM and ECM resistance
2) integrated voice and data to the user
3) performance to meet the projected user demand, i.e., error detection/correction, quality, and delays
4) effective use of transmission medium
5) interoperability
6) flexibility in deployment
7) survivability
8) provision of user mobility (e.g., manpack option, easy available access, etc.).

A number of these requirements are, at present, difficult to achieve simultaneously for any particular technical realization.

Techniques to give the system ECM resistance such as spread spectrum frequency hopping or direct sequence may oppose requirements regarding delays, effective use of the transmission medium, and interoperability. Frequency hopping systems do generally have long synchronization times, and direct sequence systems make use of extensive bandwidths and lack interoperability with, for example, conventional FM radios.

Cositing problems of frequency hopping radios makes the design of access points difficult for any reasonable voice and data capacity which may require several radios collocated.

In addition, present SCRA systems or CNR's with ECM resistance are designed primarily for voice services, and

Manuscript received February 24, 1989; revised December 11, 1989.

The authors are with the Department of Electronics, Norwegian Defence Research Establishment (NDRE), P.O. Box 25, N-2007 Kjeller, Norway.

IEEE Log Number 9035800.

0733-8716/90/0600-0713$01.00 © 1990 IEEE

it is difficult to meet the projected user demand with respect to data capacity and performance.

The survivability issue related to a graceful degradation of services when the backbone trunk system starts to deteriorate asks for a decentralized radio access system, which is difficult to achieve in systems depending on centralized synchronization procedures.

The packet radio system to be presented here does, through a unique radio and network design, meet the listed requirements. The packet radio system as well as the trunk network are, in particular, designed to match new requirements in tactical communication. It is expected that these requirements in the near future must carry digitized voice and various forms of formatted as well as unstructured data with widely differing data rates, traffic characteristics, and ''real-time'' demands.

II. Main Features of the Norwegian Tactical Communication System TADKOM

TADKOM is a service-integrated tactical telecommunication system for the Norwegian Army. The system will consist of a radio relay trunk system (TRS) and a radio access system.

The trunk system is defined as the system components necessary to build up a mesh network of digital switches and the wire extension to wire-connected subscribers. TADKOM is based on EUROCOM D/0 specifications and fulfills the specified interoperability requirements of the D/1.

By use of radio relay stations in the UHF and SHF band and line terminating equipment, the nodes are interconnected to a mesh network giving high flexibility and survivability in the battlefield environment. Each node is able to connect either to a multiplexer or another node. The access point for the wire-connected subscribers is the multiplexer.

TADKOM performs all services which are considered necessary in a battlefield infrastructure communication system, such as automatic rerouting, packet-switched data (X.25), and an advanced control and maintenance system. The maintenance system allows total control from brigade or division headquarters, as well as local control at each node.

To achieve the required mobility for the Norwegian brigades, there is a demand for a high-performance radio access system in the form of a packet radio network. The main services of the radio system will be voice traffic (semi-duplex) and data services such as circuit-switched and packet-switched data according to the CCITT X.25 standard. Such a packet radio network will be made up of two major system functions: the switching and network function, and the radio function. The switching and network part of the mobile access system would perform all necessary operations to make the required services available to the mobile subscribers, including calls and information transfer to and from subscribers in the trunk system and other mobile subscribers. The radio function is basically the radio used to convert the network services and received information to a format suitable for accessing the medium. However, the radio function is important because it represents the restriction on performance in the system.

III. Network Concepts

A. Overview

The overall network concept is divided into two main levels. One is the global network which includes all packet radio units in the brigade. The TADKOM trunk system, as discussed in the previous section, is the backbone of this global network. The other main level was created to meet the user requirements of smaller entities of the field army. A thorough study on user traffic done by simulation models showed that it was necessary to divide the global network into several packet radio subnetworks (PRNET's). Fig. 1 shows how local PRNET's are connected through the TADKOM trunk system to form a global network. This total network will cover the communication needs for the mobile terminals (MOT's) in a brigade.

B. User Requirements

Special care has been taken to develop a uniform solution for the use of radio in the Army, and the concept fulfills the user requirements by only using one type of packet radio unit. The following user requirements have been specified.

Voice and Data: The PRNET shall provide high-quality voice (16 kb/s delta modulated) and packet-switched ECM resistant data transfer.

Throughput and Transit Delay: The maximum end-to-end transit delay for given throughput/packet length constraints. The PRNET is required to handle even short packets (user data as low as 8 bytes/packet) with good efficiency.

Priority: The PRNET shall handle the user traffic in accordance with a signaled priority.

Residual Error Probability: Values for probabilities of packet loss, duplication, and out-of-sequence delivery have been specified. This requirement requires the PRNET to support reliable network connections for data transfer.

Secure Transfer of Data: The PRNET shall prevent the users from suffering from modification of user data or data to be revealed. This requires encryption of the user data, the packet headers, and the control traffic at the intranetwork level.

Coverage: Requirement to the average area from which packets can be received or transmitted (may be solved by performing multihop store-and-forwarding at the intranetwork level).

Resistance to Foreign Detection of the Radio Signal: Resistance against enemy detection of any radio transmissions (may be solved by spread-spectrum signaling and power control).

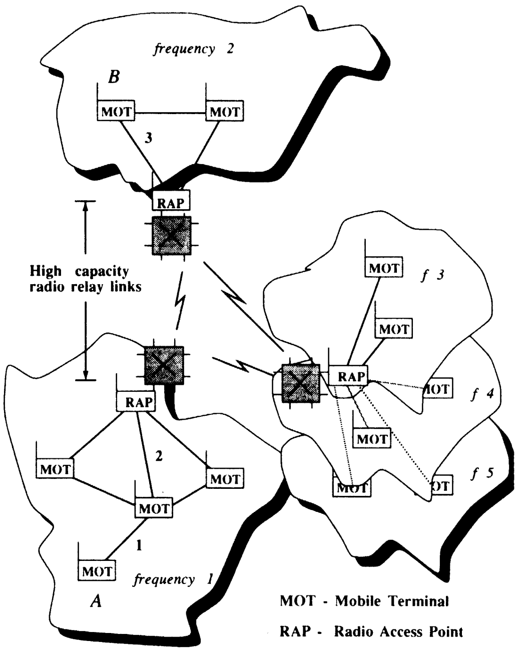


Fig. 1. Global network showing several PRNET's and internetwork communication.

Use of International Standards: If possible, international standards (ISO, CCITT, EUROCOM, etc.) shall be confirmed to at all levels. A definite requirement is that the PRNET shall facilitate attachment of X.25 (1984) terminals.

Autonomy: The PRNET shall not rely on any central control unit.

Collocation: Three packet radio units shall independently and simultaneously be operational in the same vehicle.

C. Network Topology Development

A great challenge in the overall network design process is to integrate the various needs of the different user groups in the brigade. Studies so far have shown that typically a few hundred users in the brigade will need the services and the mobility of a MOT unit. These users may be grouped into three categories; automated fire and control users, i.e., data communication between sensor and weapon systems; command users, i.e., voice and data communication between observer post and cannon commander in the field artillery; and finally, tactical information exchange users.

Studies of the application traffic of these users have shown that great channel capacity savings may be obtained when functional grouping is applied. Hence, the single radio channel of one particular PRNET is dedicated to a certain user group.

The number of nodes deployed in each of these PRNET's can be adjusted to meet the demand for various traffic loads. Each PRNET operates on one frequency at any instant in time. With a 2.4 kb/s channel capacity, each PRNET is designed to contain no more than a few tens of nodes (MOT's) to prevent traffic congestion. An important aspect of the network topology then becomes the deployment of users. Each user is preassigned to a dedicated PRNET (frequency) to minimize internetwork communication.

A PRNET that experiences traffic overloading may be divided into two or more PRNET's to solve the capacity problem. When this happens, it is possible to have geographically overlapping PRNET's, as illustrated in Fig. 1 by the PRNET's at frequencies *f* 3, *f* 4, and *f* 5.

A PRNET is connected to the TADKOM trunk system through a radio access point (RAP). Typically, less than ten RAP's will be sufficient to serve the overall brigade network. Each RAP can be equipped with a number of radio heads reflecting the traffic demands of the PRNET. In the case of internetwork communication, the RAP will serve as a gateway into the destination PRNET.

The data traffic is to be transmitted on a 2.4 kb/s channel. However, voice transmission takes place across a 16 kb/s delta modulated channel. There is a significant difference in the voice and data transmission in the sense that data are transferred by packet switching, whereas voice is circuit switched.

IV. Subnetwork Access Protocols

The terminal interface into the PR network needs to be in conformance with the standards generally employed by data communication networks. However, these standards are usually developed for wire-connected networks and fit poorly with the characteristics of a PRNET. A design objective is to let the type of network be transparent to the terminal to which it is connected.

The MOT units uses the X.25 standard for the terminal connection. A few deviations from this standard had to be employed in the data circuit terminating equipment (DCE) part of the node to be able to interface a standard X.25 terminal to the PR network. The services provided to the user are not affected by these deviations.

The Subnetwork Access Protocol is the layer 1–3 protocols between a data terminating equipment (DTE) and a DCE referred to the OSI Reference Model [1]. Our network shall provide two different Subnetwork Access Protocols (Fig. 2): a connection-mode ISO 8208 protocol [2], [4] that facilitates attachment of X.25 equipment and a connectionless-mode ISO 8473 protocol [3], [4].

A. The Connection-Mode Access Protocol

A major design objective for the mobile PRNET is to make possible an attachment of X.25 terminals. Even the inherent connectionless type of applications such as weapon control systems shall apply X.25. Problems related to the use of X.25 in connection with weapon control systems in a narrow-band PRNET have been identified. To circumvent these problems, some deviations from the X.25 are made. These deviations consist of making restrictions in the way a DTE may use the "X.25 Fast Select with Restriction on Response" and on the DCE operation. The recasted service is called *Reduced X.25*

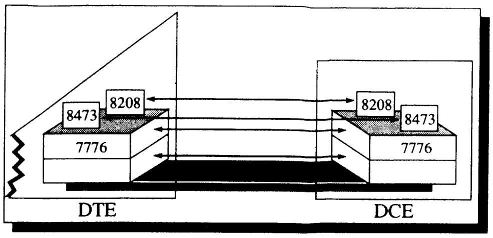


Fig. 2. The terminal interface.

Fast Select with Restriction on Response (Reduced X.25 FS w/RR).

B. Reduced X.25 FS w/RR

The reduced X.25 FS w/RR shall allow a call request packet to contain a call user data field up to 128 octets. This call user data field is transferred to the called DTE by the intranetwork layer protocols without end-to-end (DCE-to-DCE) control. The call user data may be lost or duplicated, and a sequence of X.25 FS w/RR with call user data may be received out of sequence, all without notification to the DTE's. After a time *L* (the maximum lifetime described below), the DCE will issue an X.25 clear indication packet containing no user data, and the DTE will respond with an X.25 DTE clear confirmation, Fig. 3. The receipt of an X.25 clear indication FS w/RR shall only be interpreted to have local significance, i.e., the call user data may or may not be delivered to the called DTE.

When a DTE receives an X.25 incoming call packet with FS w/RR, it will immediately issue an X.25 disconnect request on the same logical channel number. This packet will be stopped at the local DCE and not conveyed to the calling DTE by the intranetwork layer protocols.

In summary, the reduced X.25 FS w/RR has the following restrictions compared to the CCITT/X.25 FS w/RR:

- the receipt of the clear indication has only local significance.
- a called DTE may never include called user data in the clear request packet.

No change is made to the X.25 Fast Select without Restriction on Response.

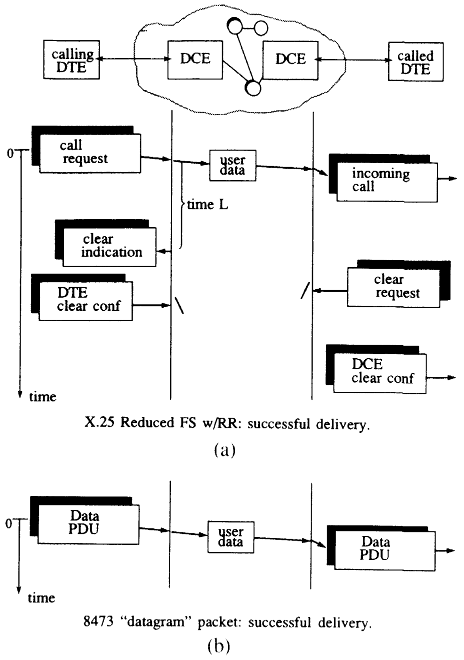
C. Enhancements to X.25 (1984)

Certain facilities not specified by ISO 8208 (see Table I) are required to fulfill military requirements. The nonstandard facilities have been specified such that they place no constraints on DTE's not making use of the facilities. Some services (e.g., fixed priority) can be agreed on for a period of time. In such cases, the services does not require any nonstandard call handling. The following nonstandard services will be provided:

- precedence and preemption
- maximum lifetime and
- semi-broadcast.

D. Precedence and Preemption

Priority is an essential service in the PRNET, and all traffic is to be marked with priority. There are four levels of priority (0–3), zero indicating the lowest priority. The priority mechanism is used in both call setup and traffic handling. Call attempts with high priority will be processed before call attempts with lower priority. If there are no available channels for setting up a logical connection, an existing connection with the lowest priority will be disconnected, provided that a connection with a priority lower than the new call exists. Both subscribers involved will be informed about the reason for the disconnection. This is done in the Clearing Cause and the Diagnostic code fields.



X.25 Reduced FS w/RR: successful delivery.
(a)

8473 "datagram" packet: successful delivery.
(b)

Fig. 3. Reduced X.25 FS w/RR.

E. Maximum Lifetime

As an optional user facility, it will be possible to select lifetime on a per-call basis. The lifetime denotes the maximum time a packet may live in the network. It will be possible to set the maximum lifetime to a maximum of 60 s with a resolution of 200 ms. If the DTE does not set a lifetime value, the network will add a default agreed upon in advance (i.e., at the subscription time).

To be able to provide X.25 VC's without corruption, every packet must be enforced a maximum lifetime at the intranetwork level. By allowing a DTE to select this value on a per-call basis (or per- "data packet" basis in conjunction with Reduced X.25 FS w/RR), network capacity is saved by discarding (i.e., no more store-and-forward operations are performed) packets that are too old to be usable for the application. Further, by including the remaining lifetime in an X.25 incoming call packet, DTE's have a mechanism to measure the network transit delay which is required by some applications.

F. Semi-Broadcast

In a PRNET where the network topology constitutes a fully connected graph, a broadcast facility may be implemented without the need to introduce additional network

TABLE I
X.25 OPTIONAL USER FACILITIES

Optional user facility	PRNET	Applies per call
Incoming Calls Barred	No	No
Outgoing Calls Barred	No	No
One-Way Logical Channel Outgoing	No	No
Flow Control Parameter Negotiation	Yes	Yes
Throughput Class Negotiation	Yes	Yes
Closed User Group (see note 1)	No	No
Closed User Group Selection	No	No
Fast Select	Yes	Yes
Fast Select Acceptance	Yes	No
Transit Delay Selection and Indication	Yes	Yes
Calling Address Extension	Yes	Yes
Called Address Extension	Yes	Yes
Minimum Throughput Class Negotiation	Yes	Yes
End-to-End Transit Delay Negotiation	Yes	Yes
Expedited Data Negotiation	Yes	Yes

Note: this facility requires too much bandwidth to be implemented.

functions. Generally, the network topology does not constitute a complete graph, and the broadcast protocol needed in such a case does not only introduce additional complexity, but also uses too much bandwidth in a narrow-band PRNET. For these reasons, only a semi-broadcast facility (i.e., a packet transmitted on the radio channel is only received by nodes one hop away) is required.

Packets transmitted as semi-broadcast packets are conveyed by means of the Reduced FS w/RR at the DTE/DCE interface (semi-broadcast cannot be used in the case of standard X.25 VC's!). The semi-broadcast facility is requested by setting a dedicated called DTE address in the X.25 call request packet. A packet conveyed by X.25 FS w/RR is identified as semi-broadcast by the special called DTE address value in the X.25 incoming call packet. The calling DTE address field will, as usual, contain the address of the DTE that issues the packet.

At the intranetwork level, no acknowledgment/retransmission procedure will be applied on semi-broadcast packets.

G. *The Connectionless Access Protocol*

From a technical point of view, a better solution to the requirement for a connectionless subnetwork service would be to implement the ISO 8473 protocol instead of changing the ISO 8208 X.25 Fast Select with Restriction on Response procedure. Unfortunately, so far we have not been able to get acceptance for implementing an ISO 8473 protocol in addition to the ISO 8208. Table II shows the 8473 protocol functions of most interest for our PRNET.

TABLE II
ISO 8473 NONSEGMENTING SUBSET

PDU Composition
PDU Decomposition
Header Format Analysis
PDU Lifetime Control
Route PDU
Forward PDU
Discard PDU
Error Reporting
Header Error Detection
Priority
Congestion Notification

V. INTRANETWORK PROTOCOLS

Fig. 4 shows our use of the Reference Model [1] for the PRNET architecture. The packet radio unit (PR unit) is logically divided into two components: one component, DCE, communicating with the subscriber equipment, DTE, and the other component, PSE (packet switching entity) communicating with other PR units. The connection between the DCE and the DTE is a 16 kb/s transmission wire, whereas two communicating PSE's are connected by a 2.4 kb/s radio channel. The intranetwork protocols describe the layered functionality of the PSE's, and hence how these entities communicate with each other.

It is mentioned that the DTE to DCE subnetwork access protocol follows the X.25 standard. It is very important to note that the packet format used in the intranctwork protocol is different from the packet format employed by the X.25 standard. It is therefore necessary to perform a conversion between the two protocols. The Local Protocol Mapping (LPM), Fig. 4, handles this conversion.

The DCE checks incoming X.25 packets for validity such as syntactical correctness, procedural behavior, and validity of parameters according to X.25. When accepted, the packet is delivered to the LPM. Furthermore, packets received by the 3b entity from its peer entity are checked for correctness before being delivered to the LPM. Thus, the LPM may be regarded as an unintelligent mapping function.

Before discussing the distributed routing functions [8], it is necessary to begin with the configuration of the PSE part of the node model. A detailed layout of the intranetwork architecture is given, Fig. 5 [1], indicating the protocols from the 3b (connection-oriented) layer down to the physical layer, as well as an overview of the management framework [9]. Hence, the PSE as described by the Basic Reference Model is divided into two parts in conformance with international standards.

The management framework standard [9] describes the facilities needed to control, coordinate, and monitor the resources which allow communication to take place in the OSI environment. More precisely, we can say that these functions will handle the establishment and the maintenance of the network topology in a PRNET.

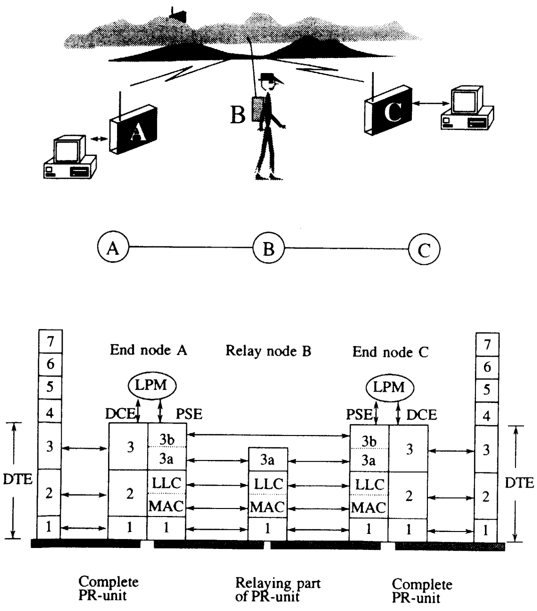


Fig. 4. Network architecture using the OSI Reference Model.

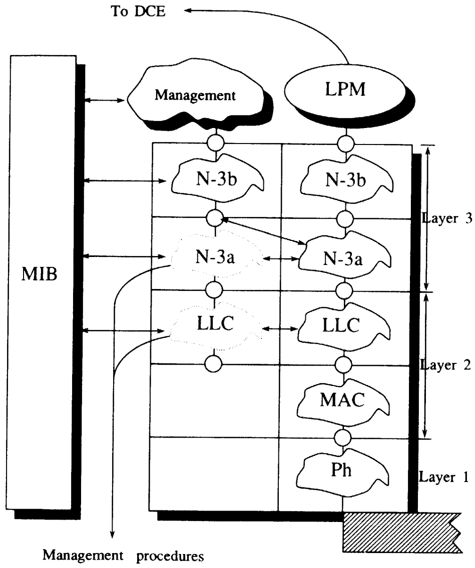


Fig. 5. Intranetwork protocols.

A. Layer Specification

The physical layer consist of a direct sequence spread-spectrum radio [10] with a function for changing the synchronization codes (preamble), PRNET synchronization, and cryptography.

The Medium Access Control (MAC) sublayer protocol must use a random access technique since all PR units in one PRNET operate on one frequency. Our MAC protocol is based on a preamble sense multiple-access (PSMA) technique. Simulations have shown that the PSMA protocol gives a better channel throughput with the predicted network connectivity and the dynamic topology in our system.

The Logical Link Control (LLC) sublayer provides the 3a entity with two different datagram services. One service is specified to transmit LLC service data units (SDU's) to the neighboring peer LLC entities with the possibility of retransmissions. The other service transmits LLC SDU's without any retransmission facility. The main function of the LLC entry is to transmit the LLC SDU's between neighboring PR units, establish and maintain the link factor value for routing purposes, and to perform MAC flow and congestion control.

The 3a sublayer provides a datagram service to the 3b entities. All 3a SDU's have a retransmission facility available if necessary in the 3a protocol. However, the retransmission may be performed at either the LLC sublayer or the 3a sublayer, depending on the acknowledgment procedure which is used. In PRNET, it is advantageous to employ passive acknowledgment, and when this is applicable, it is a function of the 3a protocol. The active acknowledgment is a function of the LLC protocol. Other main functions of the 3a connectionless protocol are the end-to-end segmenting and reassembly of packets, relaying, and routing functions.

The top layer of the intranetwork protocol is the 3b sublayer which is a connection-oriented protocol. The 3b sublayer uses the services from the connectionless 3a protocol which may provide out-of-sequence, duplication, or loss of packets. The PRNET must provide a reliable transfer of data on the X.25 virtual circuits. The main function of the 3b layer is hence to perform an error control on the unreliable services of the 3a sublayer.

B. Management Framework

Part 4 of the OSI Basic Reference Model [9] describes the structure of the management framework. Fig. 5 shows that all data elements needed for the different management functions are stored in a Management Information Base (MIB) which may be accessed from all layers in the OSI network architecture.

The management functions employed for our PRNET may use the same services as provided by the 3b service provider as indicated in Fig. 5. However, we notice that we have two 3b entities which are required for addressing purposes. A function in the 3a entity of the intranetwork protocol checks the address field of all packets arriving from its peer entities to distinguish between the management and application packets.

VI. Routing Algorithms

It has been mentioned that the overall network concept must be designed to withstand the threat of jamming or any other external activities that may influence the network topology and performance. The dynamic nature of the network suggests that the execution of all management functions should be done on a distributed basis. The philosophy greatly affects the routing strategy which is developed for the network. The discussion that follows includes the highlights of the PRNET algorithm [5], [6].

A. Link Factor

The challenge of the distributed routing algorithm is to calculate the optimum route for the *current transmission environment.* To be able to perform a selection of such a route, each node needs to know the current noise level in the surrounding transmission environment. In addition, it is necessary to have collected and tabulated, Fig. 6, a set of calculated link factors on all radio links in the PRNET. The link factor is a predetermined parameter that records the lowest signal strength received at either end of a radio link.

Establishment of Link Factor: Each radio link (radio hop) in a PRNET has a link factor assigned to it. The two nodes of a radio link do not induce symmetrical conditions for receiving a signal. Hence, a specific procedure must be executed at the PRNET establishment to determine the link factors.

Fig. 7 shows a timing sequence diagram of the link factor establishment. To establish a radio connection with the neighboring nodes, the entering node transmits semi-broadcast packets called quack packets.

After the end of a quack transmission, the entering node may start to determine its noise level (*N* level). This is done by integrating the detected noise level over a specified time window. Assuming that the neighboring node *B* received one of the quack packets, this node may calculate the SNR on the *A–B* link. Node *B* then transmits this value in a PROP (Packet Radio Organization Packet) packet, also containing node *B*'s routing status to node *A*.

At the reception of this PROP packet, node *A* may calculate the SNR on link *B–A*. By fetching the transmitted value of the SNR of link *A–B*, the symmetrical link factor is determined by taking the smallest of the two one-way SNR's. Finally, node *A* transmits a PROP packet to enable node *B* to do the same calculation.

Link Factor Values: There are introduced tree levels of link factors in the concept. Each of these levels specifies a quality of robustness on the radio link. Fig. 8 illustrates the three different levels of the link factor. It must be noted that two links rated with LF = 1 and LF = 3, respectively, may perform equally well when the current noise level in the environment is kept unchanged since the rating was done. However, an increase in the noise level of 15 dB would disqualify the LF = 1 link, but the LF = 3 link would be accepted.

The direct sequence spread-spectrum radio used for the PRNET has the ability of capturing signals which are 10 dB *below* the noise level (SNR = −10 dB). As indicated in Fig. 8, it must be noted that a link factor value gives the margin to the level needed for capturing the signal. So an SNR = 0 dB gives a margin (or robustness) of 10 dB, and hence an LF value of one.

The transmission environment is constantly changing, mainly due to jamming and the mobility of the PR units. Hence, a continuous update of the link factor is important to maintain the reliability of the route entries in the table. The PCI (protocol control information) field in the packet layout contains a field for the SNR value. Before each packet transmission, the source node includes the last incoming SNR recorded on the radio link to be used. The exchange of SNR values is a facility of the LLC protocol. Hence, the LLC sublayer reads the SNR field of the in-

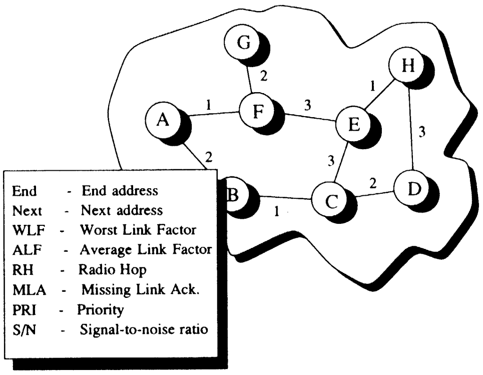


Table at Node A

End	Next	S/N	WLF	ALF	RH	MLA	PRI
H	*F*	*3*	*1*	*1.7*	*3*	*0*	*1*
H	***B***	***14***	***1***	***2***	***4***	***1***	***2***
E	*B*	*14*	*1*	*2*	*3*	*2*	*2*
E	***F***	***3***	***1***	***2***	***2***	***0***	***1***
D	*B*	*14*	*1*	*1.7*	*3*	*0*	*1*
D	***F***	***3***	***1***	***2.3***	***4***	***0***	***2***

Fig. 6. The routing table for a PRNET.

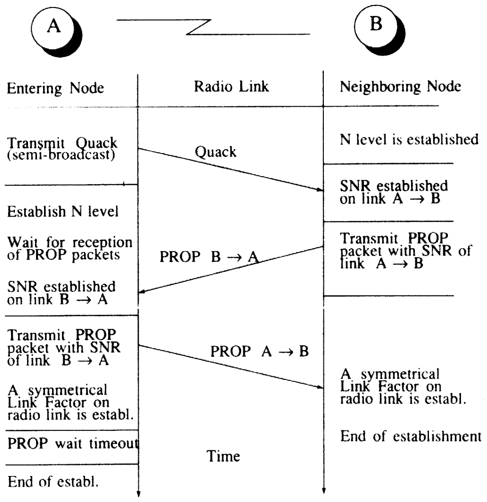

A
B
Entering Node
Radio Link
Neighboring Node
Transmit Quack (semi-broadcast)
Quack
N level is established
SNR established on link A → B
Establish N level
Wait for reception of PROP packets
PROP B → A
Transmit PROP packet with SNR of link A → B
SNR established on link B → A
Transmit PROP packet with SNR of link B → A
PROP A → B
A symmetrical Link Factor on radio link is establ.
A symmetrical Link Factor on radio link is establ.
PROP wait timeout
End of establishment
Time
End of establ.

Fig. 7. Time sequence diagram of LF establishment.

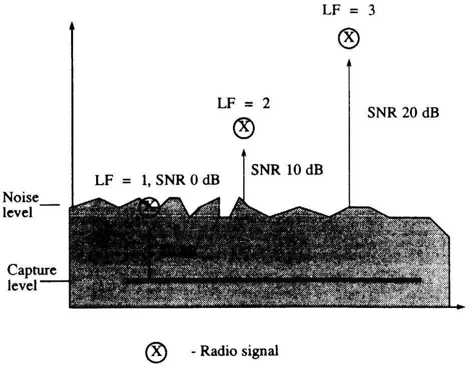


Fig. 8. The link factor.

coming packet, and the LLC management procedures may then adjust the link factors in the routing table.

B. Route Priority and Route Selection

Up to four route entries may be saved for the same destination. The routes with priority 1 and 2 are selected by the criteria which gives the smallest demand on the channel bandwidth. In Fig. 9, the route with only two radio hops is given first priority, regardless of link factor values as long as the minimum LF = 1. In general, this means the routes with the fewest number of radio hops are selected as the first- and second-priority routes, whereas the most resistant routes (highest LF values) are selected as the third- and fourth-priority routes. Here, a high value of the WLF parameter is important since only one poor radio link is necessary to congest the throughput on the complete route. Fig. 10 gives an example of selecting priorities to routes of equal number of radio hops.

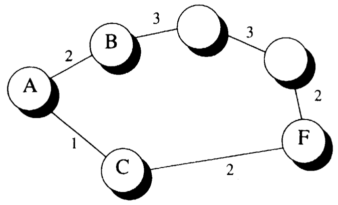


Fig. 9. Selecting the optimum route.

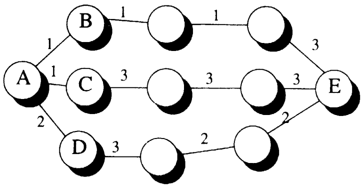


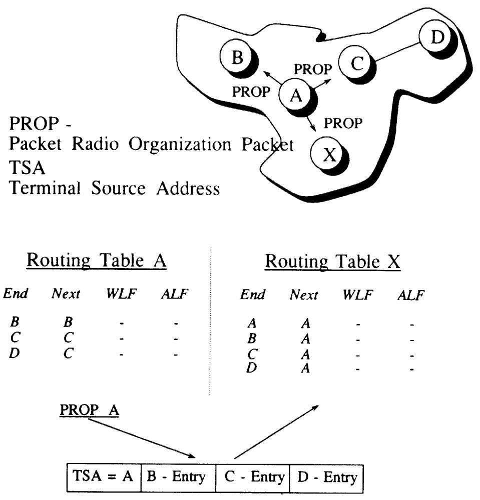
Fig. 10. Route selection of alternatives with equal number of radio hops.

C. Distribution of Routing Data

A packet transmitted on the radio channel which has *all* its neighboring nodes as the final destination is defined as a semi-broadcast. In our system, this service is available from the management entity and is used for management purposes in the event of distributing various routing data elements.

These particular semi-broadcast packets, Packet Radio Organization Packets (PROP) [6], distribute data elements which are used to determine the various route alternatives to the destination nodes. Each node transmits a PROP packet at regular intervals during the lifetime of the node. Each PROP packet will describe the complete routing status of the transmitting node to all its neighbors. As most management packets, the PROP packets will have a low priority in order to minimize the interference with the regular traffic.

The distributed control of the routing algorithm is dependent on the update and the exchange of routing data [8], [6]. Only the data needed for the route selection algorithm are discussed (WLF, ALF, incoming SNR, radio hop count, destination node address). This information is packed into the PROP packets and distributed as illustrated in Fig. 11. Here, the PROP packet distributed at *A* updates the new node *X* about the existing nodes in the PRNET. Node *X* may now reach all destinations via *A*.



Routing Table A

End	*Next*	*WLF*	*ALF*
B	*B*	-	-
C	*C*	-	-
D	*C*	-	-

Routing Table X

End	*Next*	*WLF*	*ALF*
A	*A*	-	-
B	*A*	-	-
C	*A*	-	-
D	*A*	-	-

Fig. 11. Distribution of routing information.

D. PROP Management

The PROP distribution facility is a part of the management entity, Fig. 5. This same figure shows that the management entity may fetch all the needed routing data information from the MIB. The management entity sends this management SDU containing a complete routing status (a PROP packet) to the 3a sublayer (via 3b) for semi-broadcast transmission.

There are two ways of receiving routing data elements. The PROP packets are transmitted for the only purpose of distributing this type of data. However, all packets transmitting either application or management data contain a PCI field reflecting the different parts of the intranetwork protocol. Any information in the PCI field such as addresses, SNR values, and radio hop count may be stripped off the packet as the packet travels upward through the layers. These data are directly handled by the management procedures at respective layers as indicated in Fig. 5.

The routing data distributed in the information field of the PROP packets are only handled by the management entity. The incoming routing data elements are evaluated so that only the changes can be stored in the routing table in the MIB.

E. Forwarding Algorithms

The radio unit constantly detects the noise level in the PRNET. A sudden change in the noise level may lead to poor link factor rating on various routes. The route finally selected for packet forwarding is required to fulfill the minimum link factor value (i.e., SNR = 0 dB or LF = 1). If such a route does not exist due to a severe change in the current transmission conditions, the route with the highest margin to the capture level (SNR = −10 dB) is selected.

The advantage of this forwarding algorithm is the ability of adjusting the selection of the optimum route to the current transmission environment. The current noise level is always sampled before each forwarding decision is made. If the transmission environment appears unchanged since the priority rating was done, the prioritized list of entries in the routing table is followed.

An important facility of the distributed control in the forwarding algorithm is being able to make individual forwarding decisions at each hop on a route. This flexibility may also lead to unwanted side effects such as oscillations and looping of packets. A maximum radio hop count mechanism is therefore included to make sure the packets does not hop ''forever'' and misuse the limited bandwidth.

Our routing algorithm is based on a deterministic strategy since it uses the knowledge obtained from the current transmission environment in the PRNET. It is predicted that there will not exist a route with a higher robustness than the optimum route obtained according to the link factor rating. However, a final attempt can be made by using a search or a broadcast strategy if all other alternatives have failed.

VII. Network Synchronization

The synchronization of the nodal clocks in the network is needed for two reasons. First, the network synchronization function must facilitate a synchronized change of spread-spectrum preamble codes, and in this way maintain a continuous radio communication channel. This facility is a main objective and has very strict requirements. Second, the synchronization function must provide network timing to support the real-time services. Network synchronization must not be confused with other types of synchronization such as bit or frame synchronization needed in the demodulation/decoding phase of a digital network.

Tactical systems operate in a hostile and unpredictable environment and with a changing network topology. Survivability and reliability therefore become extremely important issues in military communication systems, and depend heavily on the network synchronization performance. Network operation should not depend on the continued operation of any particular node, which leads to the requirement of a distributed network synchronization algorithm.

To ensure reliable synchronization, the independent clocks (plesiochronous) technique has been chosen where time information transmitted in the header of each packet is used to adjust the nodal clocks (times) in the network. This technique implies that the oscillator frequency of the clock is not changed/controlled; however, each node is equipped with a highly accurate oscillator satisfying the radio demodulation requirements.

There is a hierarchical structure inherent in the network synchronization. Under normal operating conditions, there will be a special node in the PRNET that is a radio access point (RAP) into the TADKOM trunk network. This node, when available, is used to distribute the network time reference.

A. The Network Synchronization Algorithm

The proposed algorithm requires a very limited amount of extra bandwidth (overhead), which is important in the narrow-band environment in which we are operating. Only 16 b, located in the packet header, are required for network synchronization, 14 b representing the node's local time and 2 b representing the node's status in the timing hierarchy, Fig. 12. The same figure also shows the status levels representing the node's position in the timing hierarchy. The status levels are obtained according to the number of radio hops the node is away from the RAP, with the highest status closest to the RAP. The nodes can now adjust their clocks towards the RAP's clock by adjusting their clock when receiving packets from a node with higher status. Simulation results have shown that the best performance is obtained when the receiving node also adjusts the clock to nodes with equal status. The algorithm is shown in Fig. 13.

If the RAP fails, all nodes in the PRNET will degrade their status to zero after a time-out period. Synchronization is now obtained by mutual influence, i.e., the nodes adjust their clocks each time a packet is received without knowing the ''quality'' of time information. Simulation results have shown that the network now will drift, relative to the time reference, according to the mean oscillator inaccuracy.

B. Simulation Examples

In this section, some of the simulation results are presented. The purpose of the simulations is to investigate the consequences when varying

- algorithm parameters
- packet traffic
- topology.

A program to simulate the synchronization algorithm has been developed using a discrete event simulation technique and implemented in Ada.

The packet traffic is kept the same for all the simulations; mean interarrival time in the network = 150 s. This corresponds to radios transmitting approximately one packet per hour. The graphs present the synchronization performance as a function of system time. The synchronization performance is represented by the mean time difference between nodes within each other's radio range (mean link deviation). In these simulations, the system time represents approximately 3.5 days.

1) Algorithm Parameters: From the algorithm, the receiver clock (*Rx* clock) is set equal to the transmitter clock (*Tx* clock) when the transmitter status is greater than the receiver status. When the receiver status is equal to the transmitter status, the receiver clock is adjusted according to

$$\text{clock adjustment} = P\,\Delta t, \qquad \Delta t = Tx\text{ clock} - Rx\text{ clock}$$

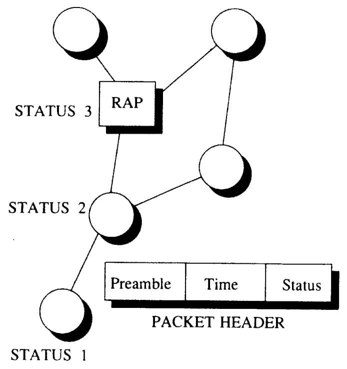


Fig. 12. PRNET with RAP.

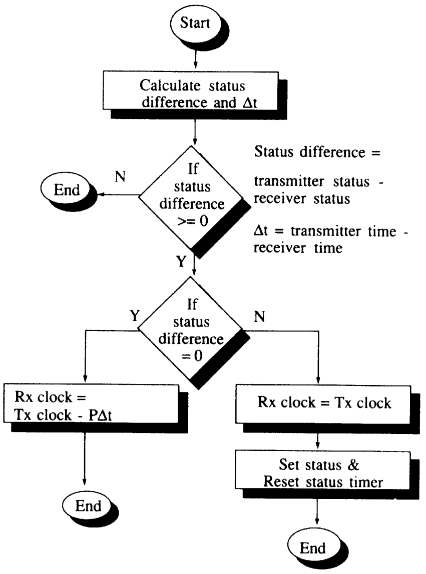

Start
Calculate status difference and Δt
If status difference >= 0
N
End
Y
Status difference = transmitter status - receiver status
Δt = transmitter time - receiver time
If status difference = 0
Y
N
Rx clock = Tx clock - PΔt
Rx clock = Tx clock
Set status & Reset status timer
End
End

Fig. 13. The synchronization algorithm.

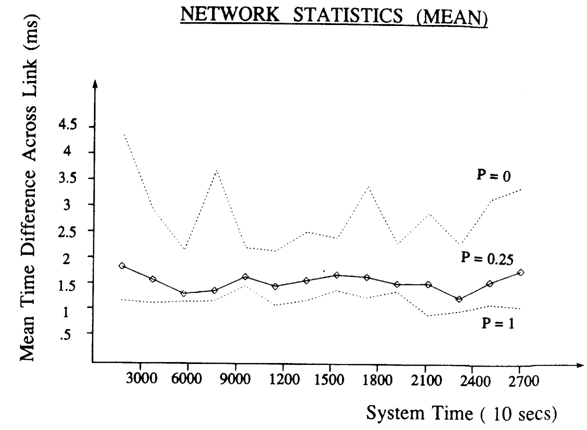


Fig. 14. Simulation results, varying the parameter *P*.

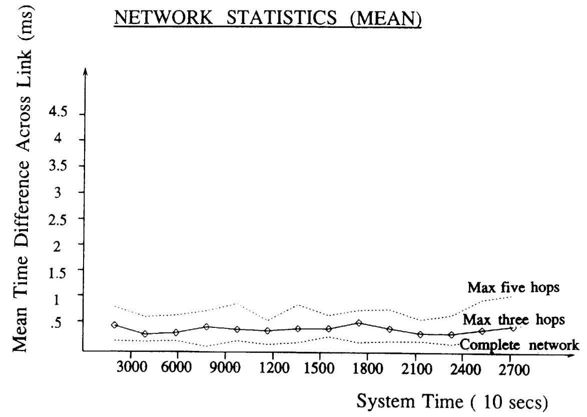


Fig. 15. Simulation results, varying the number of radio hops.

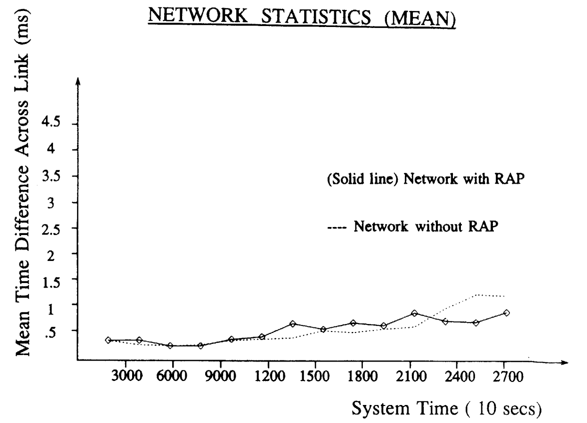


Fig. 16. Simulation results of topology with and without RAP.

and where the parameter *P* is a real number in the range [0 . . 1]. Fig. 14 shows the simulation results when varying *P* in a network with hidden nodes. $P = 0$ represents the situation when no clock adjustment is made upon reception of packets from nodes with equal status.

2) Synchronization Performance with Respect to Topology: In these simulations, the synchronization performance is illustrated with respect to different topologies. Fig. 15 shows the relationship between synchronization performance and number of hidden nodes in the network. (The networks used have a regular ring topology, and the number of hidden nodes is increased by adding more rings.)

The simulation result for networks with and without the RAP is shown in Fig. 16. The link deviation (time difference across the links) is increasing in both of the graphs due to new nodes entering the network during the simulation run, hence increasing the number of hidden nodes.

3) Comments on the Simulation Results: The simulation results in Fig. 14 show that choosing $P = 1$ is favorable, indicating that the best performance is obtained when the receiver clock always is set equal to the transmitter clock (unless the transmitter status is lower than the receiver status).

Fig. 15 shows that synchronization performance degrades with an increasing number of hidden nodes, and Fig. 16 illustrates the relative performance for a PRNET with and without the RAP. The synchronization performance with respect to the time difference across the links is nearly the same for PRNET's with and without the RAP.

Simulations of performance with respect to throughput have not been shown, but as expected, the synchronization performance is proportional to the packet throughput.

The average packet loss due to unsynchronized preamble codes is expressed as

$$\text{average packet loss} = \text{mean performance}/\text{preamble interval}$$

where

$$\text{mean performance} = 1/T \int (\text{mean time difference across links})\, dt, \quad \text{for a time period } T$$

and

$$\text{preamble interval} = \text{time between change of preamble codes.}$$

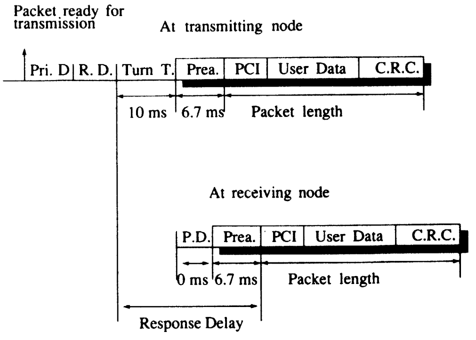
From the simulations, even with very low throughput (packet traffic), the mean performance is on the order of 1 ms. For a preamble interval of 10 s, the average packet loss then equals 1 ms/10 s = 0.01%. In a more typical network with higher throughput, the average packet loss due to unsynchronized preamble codes will be less than 0.01%.

VIII. Radio Access Protocol

The radio access protocol is the mechanism that controls the sharing of the channel capacity between the nodes in a PRNET. The communication system has to cover a wide range of applications, which requires great flexibility in the number of nodes and the traffic parameters in each PRNET. An important system requirement is autonomity, i.e., the system should not be operationally dependent on one central node. Based on these requirements, a preamble sense multiple-access (PSMA) protocol has been chosen for the communication system. Because of the spread-spectrum technique, a packet transmission is "sensed" by the detection of the synchronization preamble of the current preamble interval. Both the preamble codes and the spreading codes are changed at regular intervals to provide additional antijam protection.

The particular PSMA version we have chosen for our system is illustrated in Figs. 17 and 18 (flowchart). A node having a packet ready for transmission first waits a priority delay which corresponds to the priority of the packet. Second, it waits a random delay which is a time delay drawn from a uniform distribution. This time is introduced to decrease the probability of having two nodes transmitting simultaneously. During the time the node is waiting, it is (concurrently) also listening to the channel. If the node detects another transmission during this period, it resets the waiting time and starts the algorithm over again.

However, if the channel is still idle after both the priority delay and the random delay, the node will turn its radio from the receiving to the transmitting mode. The radio needs a finite turn time to switch to the transmitting mode, which greatly affects the performance of the access protocol.



a = Response Delay / Packet Length

Fig. 17. Time sequence diagram of the radio access protocol.

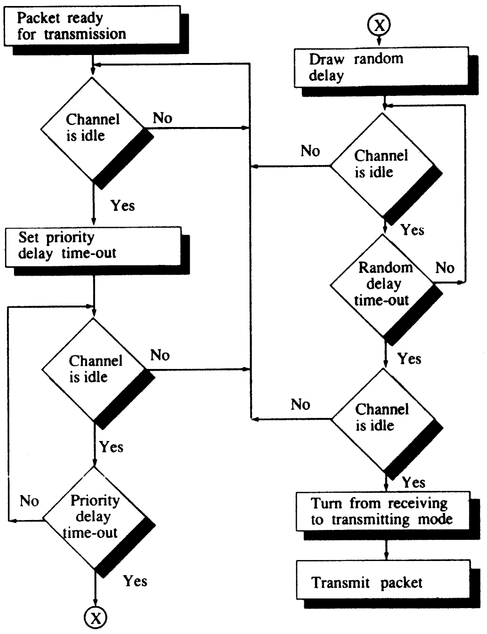


Fig. 18. Flowchart of the PSMA protocol.

The preamble is inserted in front of the protocol control information (PCI) field and the user data field of the packet. The propagation delay, i.e., the time from when the signal leaves the transmitting antenna to when it reaches the receiving node, is assumed to be zero compared to the turn time and the duration of the preamble. The period of time from when the node starts to turn its radio to the transmitting mode until the other nodes detect the transmission is defined as the response delay. Hence,

$$\text{response delay} = \text{turn time} + \text{propagation delay} + \text{preamble.}$$

During this time, it is possible for other nodes to start a transmission as well. There is therefore a finite probability that two or more packets will collide, and hence they might be lost. To make sure that the packet being transmitted is correctly received by the other node(s), an acknowledgment mechanism is used. The acknowledgment is transmitted without any priority delay. Furthermore, deadlock is avoided by using a random delay for all data and acknowledgment packets.

The probability for collisions may be specified by the parameter a as shown in Fig. 17. A small value of a will, in general, result in a better throughput performance; hence, it is important that the response delay be as small as possible [12], [13].

An important property of the direct sequence spread-spectrum radio used in our communication system is its ability to discriminate between two or more colliding packets. Hence, when two packets are transmitted nearly simultaneously (collision), the receiving radio will be able to detect and synchronize on the first packet if its signal strength relative to the second signal is larger than a certain value. However, if this signal relation is less than the same value, both packets are lost. As shown in the example on radio access protocol simulation, this property is very important in networks where not all the nodes have radio connectivity (i.e., the hidden terminal problem).

A. Radio Access Protocol Simulation

Consideration of the relative performance of the different protocols with respect to medium access, routing, and synchronization will inevitably involve the design of simulation models of the proposed protocols. The results from these simulations will provide the basis for determining optimized protocols for the given system. A large number of parameters affecting the PRNET performance need to be considered. Thus, the models will need to provide great flexibility in the setting of parameters to allow evaluation of optimized protocol configurations. This flexibility requirement contributes largely to the complexity of the models, and makes model design and verification nontrivial tasks.

At present, a simulation model based on the PSMA concept has been developed. The software model is designed and implemented using Ada due to its comprehensive structuring mechanisms, supporting reusability, and correctness. The simulation model has been used to study the special version of the PSMA protocol chosen for the PRNET. A large amount of parameters need to be set before simulating a PRNET configuration; priority delay, random delay, turn time, acknowledge method, processing delay, acknowledge wait time, node placement, radio range, message generator, and bit error probability are some of these important parameters.

Examples of simulation results are shown in Fig. 19. The examples are all from a scenario consisting of 16 nodes placed in a 4 × 4 matrix. The R^4 path loss approximation is assumed to be valid. The distance between all neighbor nodes in two units along a normalized length measure. The radio range is ten, five, and three units for the connectivity configurations shown in situations A, B, and C, respectively.

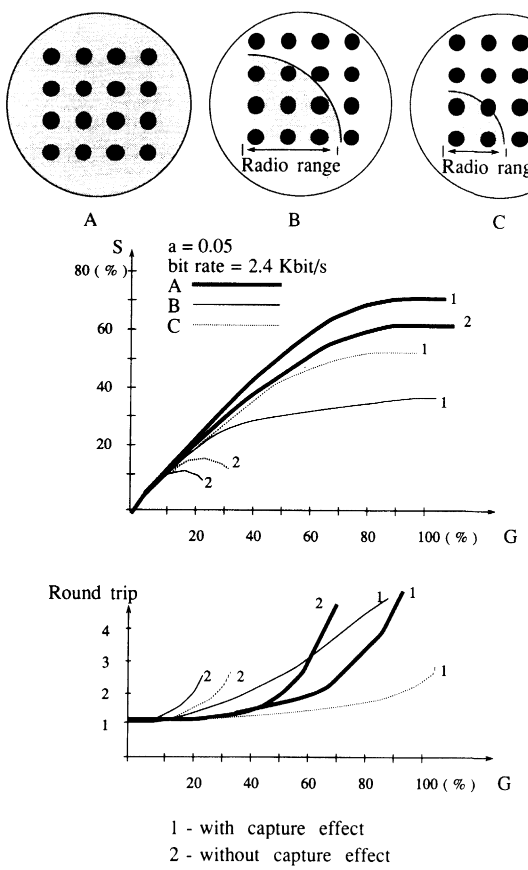


Fig. 19. Graph showing the results of simulating three different network connectivities.

The graph shown in Fig. 19 consist of two parts. The upper part shows the channel throughput (S) versus the offered channel traffic (G). The acknowledgment traffic is included in the offered traffic (G), but not included in the throughput (S). There are two throughput curves; the one marked with capture represents the use of the direct sequence spread-spectrum (DSSS) radio, and the curve marked without capture represents a radio where all packets are lost when collisions occur.

The lower part of the graphs shows the average packet delay versus the offered channel traffic. The unit roundtrip delay is defined as the time taken from when a node starts to send a packet to when the acknowledgment is received by the same node. This definition is based on successful packet delivery and that the random delay drawn is equal to the mean value. The outermost left axis of the lower graph shows the roundtrip delay converted to seconds.

IX. Radio Design

The radio is a direct sequence spread-spectrum (DSSS) radio, utilizing orthogonal coding. The conventional DSSS system represents each data bit with a code according to the modulation principle which, for example, can be differential phase-shift keying (DPSK). If the number of code bits is L, the bandwidth of the system then approximately increases L times, the jam resistance increases by a factor L, and the receiver needs to correlate over a length L. The correlator may be regarded as a coherent integrator, and increases the SNR at the output L times. The processing gain (PG) of the system is said to be L, which most often is associated with an increase in bandwidth and not the correlator length. The latter is the most correct way to regard it from a physical point of view.

The radio developed applies orthogonal signaling, which gives a significantly more efficient modulation. For an n-ary scheme, the bit rate may either be increased by a factor n or the bandwidth of the system may be reduced by a factor n compared to the original bandwidth if the bit rate is constant. If the last choice is preferred, the propagation range of the radio is increased since the amount of noise in the receiver is reduced by a factor n compared to the original noise. Furthermore, the necessary increase in signal power to detect the codes correctly is very small compared to the reduction in noise power. Note that the correlator length is constant, which means that the PG remains unchanged, except for the small loss due to the extra signal power needed to detect codes correctly. Against a broad-band jammer, the performance will increase by approximately 10 log (n).

DSSS systems are often said to have a low probability of intercept (LPI) since their output power spectral densities are lower. In a system employing orthogonal coding, the spectral density increases, and it is important to note that the range for the same output power increases as well. If the range is kept constant, the LPI performance is approximately the same as before, which is given by the correlator length.

The codes used for the orthogonal signal scheme should ideally be perfectly orthogonal, i.e., the output from the correlator should be zero when the reference code is not the same as the transmitted code. In a practical communication system, the necessary level of the cross correlation needs only to be approximately 20 dB down. This means that the number of codes available is more than adequate. The receiver will be slightly more complex in order to search for the codes in the alphabet used, but on the other hand, the bandwidth of the system is reduced. The developed radio is using orthogonal signaling at very high level, namely, 256-ary for 2.4 kb/s and 128-ary for 16 kb/s.

The data format is shown in Fig. 17. To establish chip synchronism, the data are preceded by a preamble consisting of two long codes. These are correlated in the receiver and the results are incoherently integrated. Synchronism is established in a very short time, less than 10 ms. The synchronizing codes are followed by codes for information to the radio, F codes, giving information on whether 16 kb/s voice/data or 2.4 kb/s data are to be received. Hence, the radio does not have to know in advance the type of information to be transmitted. The information symbols may be codes of variable length, and the total number of codes in a frame can vary with the application. The frame may end with Message End symbols if necessary.

The main specifications for the radio can be summarized as follows:

Radio Type:	Fixed frequency, direct sequence spread spectrum.
Frequency Range:	30–88 MHz in 25 kHz step.
Tx Bandwidth:	40 kHz.
Rx Bandwidth:	45 kHz.
Channel Separation:	50 kHz.
Data Rates:	2.4 kb/s–16 kb/s and low rate data 150 b/s.
Voice:	FM and 16 kb/s delta modulation.
Processing Gain:	24 and 15 dB plus coding gain for 150 b/s low rate.
Error Correction:	FEC, Reed Solomon.
Sensitivity:	Better than −120 dBm for 2.4 kb/s.
Modulation Type:	Noncoherent orthogonal, DSSS, MSK.
Additional ECCM:	Antenna null system.

A point worth commenting on is the channel separation. In conventional combat net radios, the bandwidth is approximately 25 kHz and the channel separation is 50 kHz. In the DSSS radio, the channel separation and the bandwidth are approximately the same. To accept this, one must bear in mind that the radio employs a DSSS system with processing gain. The tolerable noise power at the input of the receiver from a radio with a different code is at the center frequency more than 7 dB above the correct signal. In terms of channel separation, this means that one can define a "new" channel closer to a certain frequency than in a conventional system. The PG is reducing the necessary filtering. (For a high enough PG, one would have operated the system as true code division multiple access (CDMA), and have a channel separation of 0 Hz.)

When a stream of data is presented to the radio at its input port, the radio sorts out the data in blocks of 8 (7) b. These 8 (7) b represent a number between 0 and 255 (127). The code corresponding to this number is then read from a memory and put forward to the MSK modulator which is modulating some intermediate frequency according to the code. Then a conventional up converter follows with associated filtering and amplification. Extensive filtering is employed to improve collocation performance.

In the receiver chain, extensive preselection filtering is used to improve the collocation performance. The radio is double superheterodyne with linear IF amplifiers. After down conversion to the baseband, a correlator bank follows with associated synchronization and demodulation circuitry. The correlator is controlled to look for the synchronization codes, and then select the code that gives the highest correlation peak and declare this as the code transmitted. The associated number of the code is forwarded to the processor as the data. The data are either routed to the terminal or used internally to test the performance of the radio (i.e., bit error rate measurements).

A. The Radio Performance in View of Some of the Mandatory Requirements

Interception: It is extremely difficult to demodulate the data from the DSSS transmission without knowing the codes being used, even if one has a radio on hand. The number of codes that can be used for the alphabets is very large. Furthermore, it is perfectly feasible to change the synchronization codes and alphabets at a rather fast rate, making it virtually impossible to demodulate the data for anyone not knowing the key for changing the codes. As for detection of traffic going on, the LPI performance of the DSSS radio is approximately the same as for conventional direct sequence systems if the output power is adjusted according to the sensitivity gained in the receiver due to the multilevel signaling. The ability of the DSSS radio to transmit very short data messages is also very important with respect to interception.

Resistance to Jamming: The resistance to jamming is approximately the same as for conventional direct sequence systems. The PG is 24 dB in a 2.4 kb/s and 15 dB in a 16 kb/s mode. To further increase the resistance to narrow-band jammers, a system with automatic channel selection and slow frequency agility will be employed.

Because of the multilevel signaling, the performance against the broad-band barrage jammer is very good compared to other radios not using this technique. A comparison is given in Fig. 20 where the DSSS radio is compared to a frequency hopper using noncoherent FSK in a scenario limited by a white noise background. As can be seen, the range performance of the DSSS radio is superior.

Suitability for Data Applications: The ability of the DSSS radio to transfer data messages is very good. The reason is the short synchronization times needed, the fixed frequency format, and the modulation method. In addition, the radio's ability to automatically demodulate data or voice is important.

Suitability for Collocation: The collocation performance is, to a large extent, decided by the filtering employed in the transmitter chain. Here, the DSSS radio shows better performance than the frequency hoppers. A greater selectivity is obtained in the receiver because of the direct sequence principle involved. The fixed frequency format is also needed when it comes to avoiding spurious responses by frequency management both in the transmitter and the receiver.

Power Consumption: As a consequence of the multilevel signaling, low clocking speeds are employed, and the power consumption is equal to or lower than present ECCM radios.

Data Rate		FSK	DSSS Multi-level	Range Gain
16 kb/s	Eb/E0	11dB	5.5dB	36%
2.4 kb/s	Eb/N0	13.2dB	4.5dB	66%

Fig. 20. Comparisons of DSSS radio and FH techniques in the case of broad-band jamming.

Suitability for Antenna Null Steering: A fixed frequency DSSS radio is the ideal partner for an antenna null-steering system. The slightly increased bandwidth makes a shorter adaption time for the nulling possible, and the direct sequence principle working with negative signal-to-noise ratios avoids the problem of attacking a friendly signal.

Simulating Noncoherent, DSSS M-ary Receivers: To evaluate the performance of the radio, a simulation of the functions in the receiver has been done. In this simulation, the friendly signal is first generated as noncoherent MSK. Interference of various kinds is then added, such as sinusoidal jammers or simply white noise. A filtering process is acting upon both signal and noise to evaluate the influence of narrow-band filters or real-world filters with inaccuracies in phase linearity. The baseband signals are then sampled twice per chip interval. In this process, both RF offsets and timing offsets can be varied. The number of bits in the A/D process may also be varied.

In the model, many of the most important sources of implementation losses may be evaluated. These are, for instance, nonorthogonal codes, intersymbol interference caused by filtering, etc.

References

[1] "Basic reference model," ISO Int. Standard 7498.

[2] "X.25 packet level protocol for data terminal equipment," ISO/IEC JTC 1/SC 6 N 5039, 1988-06-13.

[3] "Protocol for providing the connectionless-mode network service," ISO/TC 97/SC 6 N 4542, 1987-05-12.

[4] G. C. Kessler, "A comparison between CCITT Recommendation X.25 and International Standards 8208 and 7776," *IEEE Trans. Commun.*, vol. 36, pp. 492–498, Apr. 1988.

[5] N. Shacham and J. Westcott, "Future directions in packet radio architectures and protocols," *Proc. IEEE*, and Special Issue on Packet Radio Networks," Jan. 1987.

[6] J. Jubin and J. Tomow, "The DARPA packet radio network protocols," in *Proc. IEEE*, Special Issue on Packet Radio Networks, Jan. 1987.

[7] J. M. McQuillan, G. Falk, and I. Richer, "A review of the development and performance of the ARPANET routing algorithm," *IEEE Trans. Commun.*, vol. COM-26, Dec. 1978.

[8] T. Berg, O. Stoeren, and J. E. Rustad, "Packet radio network concepts for the Norwegian field army," in *Proc. Agard EPP Conf.*, Paris, France, Oct. 1988.

[9] "Basic reference model," ISO Int. Standard 7498/4.

[10] T. Thorvaldsen, "CORA: A direct sequence spread spectrum radio for voice and data," in *Proc. Agard EPP Conf.*, Paris, France, Oct. 1988.

[11] B. H. Davis and T. R. Davis, "The application of packet switching techniques to combat net radio," *Proc. IEEE*, vol. 75, Jan. 1987.

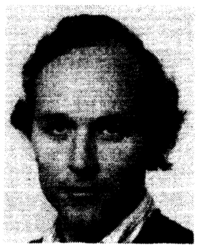
[12] L. Kleinrock and F. A. Tobagi, "Packet switching in radio channels: Part I—Carrier sense multiple-access modes and their throughput-de-

lay characteristics," *IEEE Trans. Commun.*, vol. COM-23, Dec. 1975.

[13] F. A. Tobagi, "Multiaccess protocols in packet communication systems," *IEEE Trans. Commun.*, vol. COM-28, Apr. 1980.

John Erik Rustad was born in Pontypool, England, on December 28, 1959. He received the B.Sc. degree magna cum laude in electrical engineering from the University of Utah, Salt Lake City, in 1984.

He served his National Service at the Norwegian Defence Research Establishment (NDRE), working with radio communications. He joined the NDRE as a Scientist in 1985. His work has been in the field of packet radio networks, with emphasis on intranetwork protocols. He is currently employed as a Senior Scientist, with special research interest in the evaluation of intranetwork protocols by discrete event simulation models.



Reidar Skaug was born in Norway on July 14, 1949. He received the B.Sc. degree in electrical and electronic engineering from Heriot-Watt University, England, in 1972 and the M.B. A. degree from the North European Management Institute, Norway, in 1973.

He served his National Service partly with the Navy as a Technical Instructor and partly with the Norwegian Defence Research Establishment (NDRE). He joined the NDRE as a Scientist in 1974. He was awarded a Canadian Defence Research Fellowship in 1980, and spent one year at the Communication Research Centre, Ottawa, Ont., Canada. He was promoted to Chief Scientist at NDRE in 1988, with responsibility for radar, electronic warfare, and communication activities. He has a number of technical publications in the field of communication, and has coauthored a textbook entitled *Spread Spectrum in Communication.*

Mr. Skaug is a Member of the Institution of Electrical Engineers and is a Chartered Engineer in the United Kingdom.

Andreas S. Aasen was born in Bonn, West Germany, on May 19, 1963. He received the B.Sc. degree (with honours) in electronics from the University of Manchester Institute of Science and Technology, Manchester, England, in 1987.

He joined the Norwegian Defence Research Establishment in 1987, and has been involved in the development of packet radio protocols. His research interest is performance analysis of packet radio networks.

Spread-Spectrum Multi-*h* Modulation

WILLIAM D. LANE, MEMBER, IEEE, AND AUBREY M. BUSH, SENIOR MEMBER, IEEE

Abstract—**Applying a direct random spreading sequence to a digital information sequence prior to multi-*h* modulation creates a new class of signals called spread-spectrum multi-*h* (SSMH) signaling. By spreading a known bandwidth efficient modulation scheme, the power spectral density is controlled so that the transmitted spectrum will have a wide flat mainlobe and rapid sidelobe rolloff. Coincidently, the power efficient modulation allows transmission at lower signal-to-noise ratios when the receiver knows *a priori* the spreading sequence and modulation index sequence. Optimal receiver structures are derived and numerically evaluated in an additive white Gaussian noise environment. It is shown that performance is dependent on the spreading sequence, the modulation indexes, and the possible phase states, and can exceed direct sequence binary phase-shift keying by 1–2 dB at bit error rates of 10^{-5}. Composite likelihood ratio analysis reveals a reduction of 50–70% in detectability of completely known SSMH signals vis-à-vis DS/BPSK at error rates of 10^{-4}. The spectral control, power efficiency, and reduced detectability make SSMH signals a viable low probability of intercept signaling technique.**

I. Introduction

The rapidly increasing demand for utilization of the frequency spectrum has led to extensive research and development of bandwidth efficient techniques such as continuous phase modulation (CPM) where the information is carried in the instantaneous carrier phase or frequency, while the envelope of the signal remains constant. One form of constant envelope CPM that is both power and bandwidth efficient is multi-*h* modulation. Coincident with the development of CPM, extensive development has taken place in communications techniques that are immune to intercept and jamming. Spread-spectrum techniques such as direct sequence (DS) pseudorandom spreading have been used to create signaling methods that are much more difficult to intercept by an unintended receiver or have a low probability of intercept (LPI).

The work reported here is an attempt to create a low probability of intercept signaling technique by applying direct sequence spreading to a digital information sequence prior to bandwidth and power efficient multi-*h* modulation.

By spreading a known bandwidth efficient modulation technique, a signal structure is created, and defined in Section II, which has a power spectral density that can be controlled by the signaling parameters to have a wide flat mainlobe and rapid sidelobe rolloff. At low signal-to-noise ratios, this spectrum will be difficult to detect, observe, and parameterize. Analytical expressions for the power spectral density of spread-spectrum multi-*h* signals are derived in Section III and numerically evaluated.

Coherent receiver structures to detect the transmitted information sequence are derived in Section IV where it is assumed that the intended receiver knows the spreading sequence, the modulation index sequence, the carrier phase, and the symbol and chip timing. Performance analysis (Section V) and numerical analysis (Section VI) are accomplished to validate the structure.

The issue of the detectability of the transmitted signal is addressed in Section VII from an optimal intercept standpoint where the best possible performance of an intercept receiver is considered.

Finally, a comparison is made to direct sequence binary phase-shift keying (DS/BPSK) to show that SSMH is a viable low probability of intercept signaling technique.

II. Signal Definition and Transmission

Spread-spectrum multi-*h* (SSMH) modulated signals are defined by a transmitted signal

$$s(t, \boldsymbol{\alpha}, \boldsymbol{c}, \boldsymbol{h}) = \sqrt{\frac{2E_c}{T_c}} \cos\left(2\pi f_0 t + \phi(t, \boldsymbol{\alpha}, \boldsymbol{c}, \boldsymbol{h}) + \theta_0\right) \tag{1}$$

where the information symbol α_i is multiplied by a random/pseudorandom spreading sequence of c_{ij} dibits, or chips, and modulated in the phase

$$\phi(t, \boldsymbol{\alpha}, \boldsymbol{c}, \boldsymbol{h}) = 2\pi \int_{-\infty}^{t} \sum_{i=-\infty}^{\infty} \sum_{j=0}^{N_c-1} \alpha_i h_{(i+j) \operatorname{mod} H} \times c_{ij} g\left(\tau - (j + iN_c) T_c\right) d\tau, \quad -\infty \le t \le \infty. \tag{2}$$

The data sequence $\boldsymbol{\alpha}$ is an M-ary (M is a power of 2) infinitely long sequence of uncorrelated equally likely data symbols, each with a value of $\alpha_i = \pm 1, \pm 3, \cdots, \pm(M - 1)$, with $i = 0, \pm 1, \pm 2, \cdots$, and symbol duration T_s. The carrier frequency is f_0 and θ_0 is an arbitrary constant initial phase which will be set to zero under the assumption of perfect coherency.

The spreading sequence $\boldsymbol{c}$ is assumed to be an infinitely long sequence of uncorrelated equally likely chips, each with a value of $c_{ij} = \pm 1$. This sequence is presumed to be known *a priori* by both the transmitter and the re-

Manuscript received February 17, 1989; revised October 20, 1989.

W. D. Lane is with the Department of Electrical Engineering and Computer Science, United States Military Academy, West Point, NY 10996.

A. M. Bush is with the Department of Electrical Engineering, Georgia Institute of Technology, Atlanta, GA 30332.

IEEE Log Number 9035801.

ceiver. Additionally, it will be assumed that there is an integer number of equal energy chips N_c per symbol interval such that $T_s = N_c T_c$.

The modulation index $\boldsymbol{h}$ is a finite length sequence of cyclically varying fixed frequency deviations such that

$$h_{i+j} = h_{i+j+H} = h_{(i+j)\,\mathrm{mod}\,H} = n_i/p \tag{3}$$

where the numerator n_i and denominator p are fixed integer numbers. As a result, a finite-state description for the signal structure can be made and can be reflected in finite-state trellis structures.

The phase transitions in the signal definition to this point have been indexed on a bit basis for the information sequence and on a chip basis for the chip sequence. The application of the frequency deviation ratio $h_{(i+j)\,\mathrm{mod}\,H}$ on a chip basis will be referred to as *conventional* spread-spectrum multi-h modulation. This leads to the expression of the phase of the transmitted signal as

$$\phi(t, \boldsymbol{\alpha}, \boldsymbol{c}, \boldsymbol{h}) = 2\pi \sum_{i=-\infty}^{\infty} \sum_{j=0}^{N_c-1} \alpha_i c_{ij} h_{(i+j)\,\mathrm{mod}\,H} \times q(t - (j + iN_c)T_c) \tag{4}$$

where $q(t)$ is the chip phase transition resulting from the frequency pulse shaping function $g(t)$.

The frequency deviations can also be applied on a bit basis resulting in the ratios h_i being indexed along with the bit sequence. This type of application will be referred to as *modified* spread-spectrum multi-h modulation. In this situation, the phase is expressed as

$$\phi(t, \boldsymbol{\alpha}, \boldsymbol{c}, \boldsymbol{h}) = 2\pi \sum_{i=-\infty}^{\infty} \sum_{j=0}^{N_c-1} \alpha_i c_{ij} h_{i\,\mathrm{mod}\,H} \times q(t - (j + iN_c)T_c). \tag{5}$$

The frequency smoothing pulse $g(t)$ acts on the product of the chip and the data symbol to define the shape of the phase response $q(t) = \int_{-\infty}^{t} g(\tau)\, d\tau$. Generally, $g(t)$ is a causal smoothing pulse of arbitrary form and length LT_c where L is a positive integer. The pulse is normalized so that $q(LT_c) = 1/2$. For the present work, full response ($L = 1$) rectangular signaling is utilized where $g(t)$ is assumed to be constant over a chip interval, resulting in a constant slope phase change. With the smoothing pulse $g(t)$ acting on each chip and data symbol, the characteristics of continuous phase systems are preserved by requiring that the phase transitions at chip intervals be a continuous function of time.

As shown by Anderson and others [1], the phase change during the lth bit and nth chip can be expressed (assuming rectangular pulse shaping and fixed rational modulation indexes) as

$$\phi(t, \boldsymbol{\alpha}, \boldsymbol{c}, \boldsymbol{h}) = 2\pi h_{(l+n)\mathrm{mod}\,H} \alpha_l c_{ln} q(t) + \theta_{ln} \tag{6}$$

$$= \theta(t, \alpha_l, c_{ln}, h_{(l+n)\mathrm{mod}\,H}) + \theta_{ln} \tag{7}$$

where

$$\theta_{ln} = \left[\pi \sum_{i=-\infty}^{l} \sum_{j=0}^{n-1} h_{(i+j)\,\mathrm{mod}\,H} \alpha_i c_{ij}\right] \mathrm{mod}\ 2\pi. \tag{8}$$

This expression describes a Markov state system with p states, with the result that the signaling waveform exhibits a periodic trellis structure. However, the inclusion of the cyclically changing modulation indexes will alter the periodicity of the trellis as compared to fixed deviation (single h) continuous phase systems. This extension of the trellis structure contributes to enhanced performance vis-à-vis single h systems [1]. The application of the dibits alters the periodicity on a bit basis of the trellis structure even further.

The conceptual structure of the SSMH transmitter follows from applying direct sequence spreading to a digital information sequence prior to multi-h modulation and is shown in Fig. 1. The binary (±1) information bit is multiplied by a much higher rate binary dibit to yield a digital sequence that is input to a conventional multi-h modulator.

Using the signal definitions above, the transmitter structure can be further refined following trigonometric expansion and substitution. With the initial arbitrary phase set to zero and rectangular pulse shaping, the signal definition for the interval $jT_c \leq t \leq (j + 1)T_c$ may be expressed as

$$s(t, \alpha, c, h) = \sqrt{\frac{2E_c}{T_c}}\left[I(t) \cos \omega_0 t - Q(t) \sin \omega_0 t\right] \tag{9}$$

where the in-phase term is

$$I(t) = \left[\cos\left(2\pi h_{(i+j)\,\mathrm{mod}\,H} \alpha_i c_{ij} q(t)\right)\right] \cos \theta_{ij} - \left[\sin\left(2\pi h_{(i+j)\,\mathrm{mod}\,H} \alpha_i c_{ij} q(t)\right)\right] \sin \theta_{ij} \tag{10}$$

and the quadrature term is

$$Q(t) = \left[\sin\left(2\pi h_{(i+j)\,\mathrm{mod}\,H} \alpha_i c_{ij} q(t)\right)\right] \cos \theta_{ij} + \left[\cos\left(2\pi h_{(i+j)\,\mathrm{mod}\,H} \alpha_i c_{ij} q(t)\right)\right] \sin \theta_{ij}. \tag{11}$$

These equations form the basis of the quadrature transmitter structure of Fig. 2.

The ease with which the fundamental nature of the transmitted signal can be changed is shown in the flexibility of the transmitter format. The spreading sequence and the pulse shaping function are produced external to the modulation process and can be easily altered. The frequency deviation ratio sequence can also be changed readily. Not only can the magnitude of the ratios be changed, but the method of application of the ratios can also be altered. In one instance, the index can change during each chip interval, thereby creating *conventional* SSMH signals. On the other hand, it is a simple matter to clock the indexes in accordance with the data sequence, and thereby produce *modified* SSMH signals. Thus, the

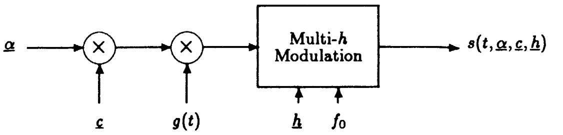


Fig. 1. Spread-spectrum multi-h transmitter.

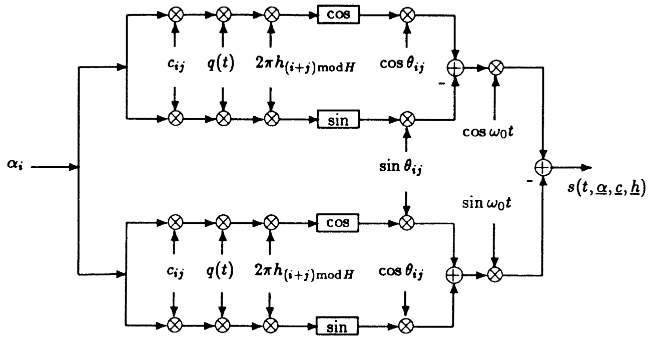


Fig. 2. Quadrature transmitter structure.

distinction between the two types of modulation is the timing of the application of the frequency deviation ratios. Subsequent sections address the attributes of the signals produced by either of these methods.

III. Spectral Analysis

One major purpose of SSMH is to create a signal waveform that presents a power density spectrum that has a wide flat mainlobe and rapid sidelobe rolloff. If transmitted at low signal-to-noise ratios, the broad flat spectrum becomes difficult to detect, observe, and parameterize by unintended receivers. The spectra for *conventional* and *modified* SSMH signals are derived in this section.

The general method of analysis for the power density spectrum is to obtain the autocorrelation of the signal expression, invoke the time-averaged autocorrelation approach of Papoulis [8], and implement the Wiener–Khintchine theorem to Fourier transform the time-averaged autocorrelation to obtain a "probabilistic representation" [7] of the power spectral density. The derivations of the spectra for both *conventional* and *modified* SSMH signals are shown in [6].

The final normalized expressions representing the *conventional* SSMH power spectral density with H deviation ratios become

$$G(f_n) = 2\left\{ G_1(f_n) + \frac{a}{c} G_2(f_n) - \frac{b}{c} G_3(f_n) \right\}$$

$$G_1(f_n) = \frac{1}{H} \int_0^H \int_0^H \prod_{j=0}^{2H+1} r_c(t, \tau) \cos(\hat{\tau})\, dt\, d\tau$$

$$G_2(f_n) = \frac{1}{H} \int_H^{2H} \int_0^H \prod_{j=0}^{2H+1} r_c(t, \tau) \cos(\hat{\tau})\, dt\, d\tau$$

$$G_3(f_n) = \frac{1}{H} \int_H^{2H} \int_0^H \prod_{j=0}^{2H+1} r_c(t, \tau) \sin(\hat{\tau})\, dt\, d\tau$$

$$r_c(t, \tau) = \cos\left[2\pi h_{j \bmod H}\left(q(t + \tau - j) - q(t - j)\right)\right]$$

$$\hat{\tau} = 2\pi f_n \tau$$

$$C_\alpha = \prod_{j=0}^{H-1} \cos(\pi h_{j \bmod H})$$

$$a = 1 - C_\alpha \cos(2\pi f_n H)$$

$$b = C_\alpha \sin(2\pi f_n H)$$

$$c = 1 + C_\alpha^2 - 2C_\alpha \cos(2\pi f_n H). \tag{12}$$

The spectrum of (12) was evaluated via Fortran implemented computer programs for selected values of modulation indexes. These indexes were selected based on the results from Hsu [4] and Lereim [7], which indicated the "best" modulation codes for bandwidth efficient modulations in terms of coding gain and bandwidth efficiency. Additionally, the codes were selected based on the desire to have a spectrum with a flat mainlobe and rapid sidelobe rolloff. An example of the numerical evaluation is shown in Fig. 3 where the *conventional* spectrum for modulation indexes 7/12, 8/12 and five chips per bit is shown. The spectrum has been normalized to unity chip energy and to 0 dB at the carrier frequency. Since the spectrum for real signals is being displayed, it is only necessary to display the single-sided density. It should be noted that due to the numerical integration routines utilized, it was necessary to restrict the number of chips per bit to low levels.

This analysis reflects the application of the methodology of Anderson and Lereim to this new class of spread spectrum signals. It confirms the expected results that when normalized to the chip rate, *conventional* SSMH signals have the same spectra as their parent multi-h signals.

By changing the application of the modulation indexes to a bit basis, the structure of the *modified* SSMH transmitted signal is much different. The analysis for the power spectral density is similar to the *conventional* signal case and is shown in [6]. The resulting spectra for *modified* SSMH signals with H indexes can be expressed as

$$G(f_n) = 2\left\{ \int_0^{HN_c} r_m(\tau) \cos(2\pi f_n \tau)\, d\tau + \frac{a}{c} \int_{HN_c}^{2HN_c} r_m(\tau) \cos(2\pi f_n \tau)\, d\tau - \frac{b}{c} \int_{HN_c}^{2HN_c} r_m(\tau) \sin(2\pi f_n \tau)\, d\tau \right\}$$

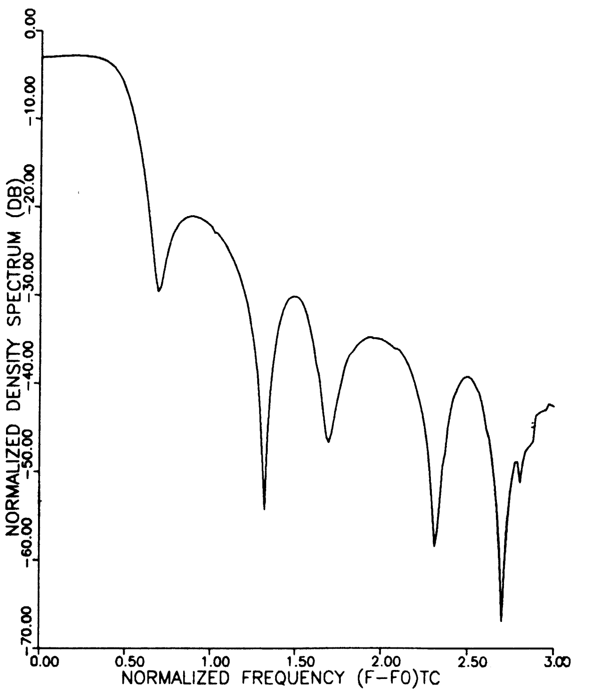


Fig. 3. *Conventional* 7/12, 8/12, $N_c = 5$ spectrum.

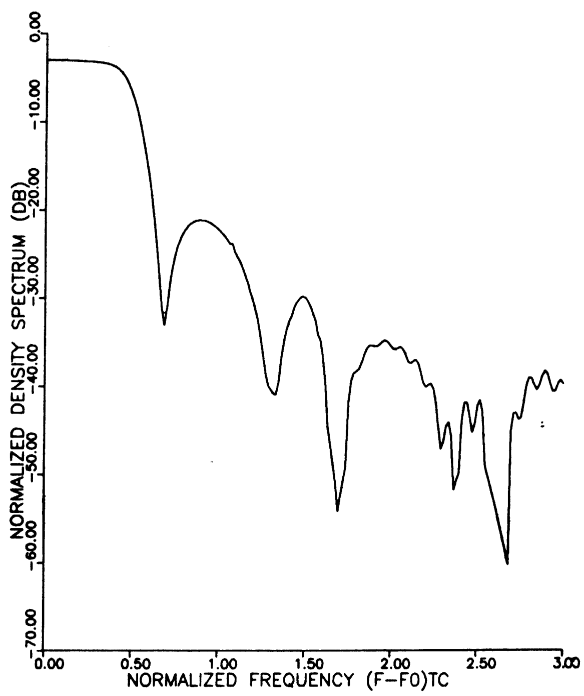


Fig. 4. *Modified* 7/12, 8/12, $N_c = 5$ spectrum.

with

$$
\begin{aligned}
C_\alpha &= \prod_{j=0}^{H-1} \left\{\cos\left(\pi h_{j \bmod H}\right)\right\}^{N_c}\\
a &= 1 - C_\alpha \cos\left(2\pi f_n H N_c\right)\\
b &= C_\alpha \sin\left(2\pi f_n H N_c\right)\\
c &= 1 + C_\alpha^2 - 2C_\alpha \cos\left(2\pi f_n H N_c\right)\\
r_m(\tau) &= \frac{1}{HN_c}\int_0^{HN_c} \prod_{m=0}^{2H+1} \prod_{n=0}^{N_c-1} r_1(t, \tau)\, dt\\
r_1(t, \tau) &= \cos\left[2\pi h_{m \bmod H}\left(q\left(t + \tau - (n + mN_c)\right)\right.\right.\\
&\quad \left.\left. - q\left(t - (n + mN_c)\right)\right)\right]. \qquad (13)
\end{aligned}
$$

These expressions are not the same as those for *conventional* spectra, but numerical analysis allows comparison. The spectra for selected *modified* SSMH were also numerically evaluated. For comparison purposes, Fig. 4 shows the spectrum for *modified* 7/12, 8/12 SSMH and five chips per bit.

Following the analysis and numerical evaluations, some conclusions and characterizations can be made regarding the spectra of this class of signals.

Foremost is the fact that the spectra have similar characteristics to their parent multi-h signals when normalized to the chip rate. This is seen by comparison of the evaluated spectra to previously derived spectra for multi-h signals as shown by Lereim [7], Wilson [17], and Anderson *et al.* [1]. Correspondingly, it also implies that the spectrum of the signals normalized to the bit rate will be spread by a factor of the number of chips per bit, while maintaining the shape characteristic of the multi-h signal with the given modulation indexes.

It is also apparent that the method of modulation has little effect on the resulting spectrum. That is to say, the spectra of *conventionally* modulated and *modified* modulated spread-spectrum signals with the same modulation indexes closely resemble each other. This is a new and exciting result in that it characterizes the results of a new spread-spectrum modulation technique, *modified* SSMH, in terms of existing spread-spectrum signaling results. After-the-fact analysis confirms this result since the averaging of the spectra is over the same segments, but in a different order. It is also apparent that considerable control over the spectrum is afforded by selection of the modulation indexes, the spreading rate, and the other signaling parameters.

With the transmitted spectrum characterized, it is appropriate to determine if the signal can be transmitted at sufficiently low signal-to-noise ratios to take advantage of the spectral characteristics and still be detectable by an intended receiver.

IV. Optimal Receiver Structure

Detecting spread-spectrum multi-h signals in an additive white Gaussian noise (AWGN) environment and decoding the transmitted information are addressed in this section. While significant research has been done on multi-h receiver structures as a generalization of CPM signaling, the addition of another level of detection and synchronization is novel. The bit and chip synchronization and timing are assumed to be known exactly, as well as the receiver having complete knowledge of the spreading code and the modulation index code. The receiver still must detect and decode the chip sequence and ultimately the transmitted data. For this work, it is assumed that complete coherency can be established and the phase offset is considered to be zero.

Assuming equally likely transmitted digital sequences, the received signal can be expressed as

$$r(t) = s(t) + n(t) \tag{14}$$

where $n(t)$ is white Gaussian noise with a two-sided variance of $N_0/2$. Following the derivations of Viterbi and Omura [16], Jackson showed [5] that the decision metric for a digital sequence of N symbols (bits) reduces to a log-likelihood metric of

$$l = \frac{2}{N_0} \int_0^{NT_s} r(t) \left[s(t, \boldsymbol{\alpha}, \boldsymbol{h}, \boldsymbol{c}) - s(t, \boldsymbol{\alpha}', \boldsymbol{h}, \boldsymbol{c}) \right] dt$$

$$l \gtrless 0 \tag{15}$$

which, for SSMH signals, can be expressed as

$$l = \sum_{i=0}^{N-1} \sum_{j=0}^{N_c-1} \int_{jT_c}^{(j+1)T_c} r(t - (j + iN_c)T_c)$$

$$\times\, s(t - (j + iN_c)T_c, \alpha_i, h_{(i+j) \bmod H}, c_{ij})\, dt. \tag{16}$$

Thus, the decision metric for a sequence of N bits is

$$l = \sum_{i=0}^{N-1} \lambda_i \tag{17}$$

where the individual bit metrics are

$$\lambda_i = \sum_{j=0}^{N_c-1} \int_{jT_c}^{(j+1)T_c} r(t - jT_c)$$

$$\times\, s(t - jT_c, \alpha_i, h_{(i+j) \bmod H}, c_{ij})\, dt. \tag{18}$$

These expressions indicate that the overall decision metric is the result of the sum of bit metrics, which in turn are the sum of "chip" metrics. Equation (18) also indicates that the metrics are the correlation of the received signal with the possible transmitted signals. Hence, the optimal receiver computes the maximum correlation over all possible received sequences. Forney [3] and others [1], [4] have shown that for sequences from a finite-state Markov process, such as the multi-h signals defined in Section II, the Viterbi algorithm is a recursive method to exhaustively search for the optimal sequence.

The final receiver structure is defined after determining the correlation filter structure. The chip metrics for rectangular signaling over the ith bit and jth chip can be written in the following manner:

$$\lambda(\alpha_i = +1, c_{ij}, \theta_{ij})$$

$$= \int_0^{T_c} r(t) \cos\left(2\pi f_0 t + \frac{\pi c_{ij} h_{ij} t}{T_c} + \theta_{ij}\right) dt \tag{19}$$

$$\lambda(\alpha_i = -1, c_{ij}, \theta_{ij})$$

$$= \int_0^{T_c} r(t) \cos\left(2\pi f_0 t - \frac{\pi c_{ij} h_{ij} t}{T_c} + \theta_{ij}\right) dt. \tag{20}$$

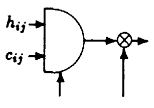


Fig. 5. Receiver multiplexer circuit.

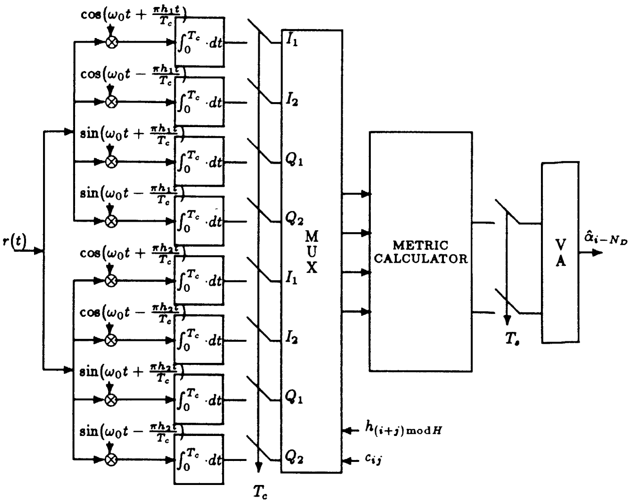

$r(t)$
$\cos(\omega_0 t + \frac{\pi h_1 t}{T_c})$
$\cos(\omega_0 t - \frac{\pi h_1 t}{T_c})$
$\sin(\omega_0 t + \frac{\pi h_1 t}{T_c})$
$\sin(\omega_0 t - \frac{\pi h_1 t}{T_c})$
$\cos(\omega_0 t + \frac{\pi h_2 t}{T_c})$
$\cos(\omega_0 t - \frac{\pi h_2 t}{T_c})$
$\sin(\omega_0 t + \frac{\pi h_2 t}{T_c})$
$\sin(\omega_0 t - \frac{\pi h_2 t}{T_c})$
$\int_0^{T_c} \cdot dt$
I_1
I_2
Q_1
Q_2
I_1
I_2
Q_1
Q_2
MUX
T_c
$h_{(i+j) \bmod H}$
c_{ij}
METRIC CALCULATOR
T_s
VA
$\hat{\alpha}_{i-N_D}$

Fig. 6. Maximum likelihood receiver structure.

Expanding these expressions on the initial phase angle yields quadrature forms of the correlators. That is,

$$\begin{bmatrix} \lambda(+1, c_{ij}, \theta_{ij}) \\ \lambda(-1, c_{ij}, \theta_{ij}) \end{bmatrix} = \begin{bmatrix} I_1 & -Q_1 \\ I_2 & -Q_2 \end{bmatrix} \begin{bmatrix} \cos \theta_{ij} \\ \sin \theta_{ij} \end{bmatrix} \tag{21}$$

where, over a chip interval,

$$I_1 = \int_0^{T_c} r(t) \cos\left(2\pi f_0 t + \frac{\pi c_{ij} h_{ij} t}{T_c}\right) dt \tag{22}$$

$$Q_1 = \int_0^{T_c} r(t) \sin\left(2\pi f_0 t + \frac{\pi c_{ij} h_{ij} t}{T_c}\right) dt \tag{23}$$

$$I_2 = \int_0^{T_c} r(t) \cos\left(2\pi f_0 t - \frac{\pi c_{ij} h_{ij} t}{T_c}\right) dt \tag{24}$$

$$Q_2 = \int_0^{T_c} r(t) \sin\left(2\pi f_0 t - \frac{\pi c_{ij} h_{ij} t}{T_c}\right) dt. \tag{25}$$

These equations imply a bank of $4H$ correlators and a phase rotation network to account for the allowable phases at the start of each chip interval.

The last requirement in defining the receiver architecture is to account for the changes in the modulation indexes and the chip sequence. Minor modification of the simple switching circuitry shown by Sadr [13] allows proper multiplexing of the quadrature elements to select the appropriate modulation index and the known chip sequence. Assuming a two-index modulation code, the multiplexing circuit is shown in Fig. 5 where h_{ij} = logic(+1) for modulation index h_1, h_{ij} = logic(0) for modulation index h_2, c_{ij} = logic(+1) for $c_{ij} = +1$, and c_{ij} = logic(0) for $c_{ij} = -1$.

The receiver architecture *in toto* is then shown in Fig. 6.

Prior to evaluating this receiver structure, the analytic performance bounds to the modulation technique are derived in the following section. The receiver derived in this section is then numerically evaluated for comparison to the bounds.

V. Performance Bounds

The error performance of the maximum likelihood receiver of the previous section is dependent on the distance properties and performance bounds for the signaling structure. The results for equally likely transmitted signals is the total error probability which is upper bounded by summing the probabilities of error for all possible error sequences against all possible transmitted sequences [16], [4]. This is expressed as

$$P_e \leq \sum_{\text{all errors}} Q\left(\sqrt{\frac{d_i^2}{2N_0}}\right) P(\alpha_N) \tag{26}$$

where d_i is the Euclidean distance separating two signals of a particular error event length and

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-(t^2/2)}\, dx. \tag{27}$$

Equation (26) reflects the fact that the probability of error for each error event is simply a pairwise comparison of two sequences over a given length. This implies the implementation of the "difference" sequence approach as shown by Hsu [4] and Anderson [1]. Since the phase changes of the signaling scheme correspond directly to the information sequences, the difference phase state approach can be invoked where

$$\Delta\theta(i+1) = \left[\Delta\theta(i) + \sum_{j=0}^{N_c-1} \pi h_{ij} \gamma_{ij}\right] \bmod 2\pi \tag{28}$$

and γ_{ij} is the difference,

$$\gamma_{ij} = c_{ij}\alpha_i - c_{ij}\hat{\alpha}_i. \tag{29}$$

The all-zero path then corresponds to error-free transmission.

Equation (29) illustrates a key and essential difference between the current analysis and that of Hsu and Anderson. Although the chip sequence is known *a priori* by the transmitter and receiver, the chip values help determine the phase changes on a chip basis, and hence contribute to the bit difference state transitions. The Markov process structure is not distributed, indicating that the phase changes over a bit interval are the sum of the changes over each chip interval. Hence, (28) can be written

$$\Delta\theta(i+1) = \left[\Delta\theta(i) + \sum_{j=0}^{N_c-1} \Delta\theta_{ij}\right] \bmod 2\pi \tag{30}$$

where

$$\Delta\theta_{ij} = \pi h_{(i+j)\bmod H}\, \gamma_{ij}. \tag{31}$$

The result of this analysis is that a "difference" state structure and analysis can be followed, but the state diagram must account for all possible chip sequences. Additionally, the analysis must include all possible transmitted phase states.

To perform the difference state analysis, (26) must be modified in two ways. The first reflects the fact that the bit error sequences $\gamma_i = \alpha_i - \hat{\alpha}_i$ can occur in different ways and must be appropriately weighted. Second, (26) must account for the different possibilities for the chip interval dibits. For bit interval differences of $\gamma_i = 0$, there is no effect since no error is made, but for $\gamma_i = \pm 2$, all of the possible chip values and sequences must be considered. These chip difference sequences have a significant effect on the phase transitions and on the distance characteristics, which is an important new finding in this research.

It should be noted that this result generalizes to have potential impact on other signaling schemes such as combined modulation and error-correcting coding techniques.

The result is (26) modified to be

$$P_e \leq \sum_{i=1}^{\infty} \nu_i Q\left(\sqrt{\frac{d_i^2}{2N_0}}\right) \tag{32}$$

where

$$\nu_i = \prod_{i=1}^{N} \prod_{j=1}^{N_c} \Pr(\gamma_{ij} \mid \gamma_i) \Pr(\gamma_i). \tag{33}$$

The final bit error probability bound is then

$$P_b \leq \sum_{i=1}^{\infty} \mu_i \nu_i Q\left(\sqrt{\frac{d_i^2}{2N_0}}\right) \tag{34}$$

where μ_i is the number of bit errors in each sequence with separation distance d_i.

The squared Euclidean distance between any two signals can be expressed as

$$d_{12}^2(N) = \int_0^{NT_s} (s_1(t) - s_2(t))^2\, dt, \tag{35}$$

but this distance is cumulative over bit and chip intervals, and for SSMH signals with rectangular signaling reduces to

$$d_{12}^2(N) = \begin{cases} 2E_c \displaystyle\sum_{i=1}^{N} \sum_{j=0}^{N_c-1} 1 - \frac{\sin(\Delta\theta_{i,j+1}) - \sin(\Delta\theta_{ij})}{\Delta\theta_{i,j+1} - \Delta\theta_{ij}}, \\ \quad \text{if } \Delta\theta_{i,j+1} \neq \Delta\theta_{ij}, \\ 2E_c \displaystyle\sum_{i=1}^{N} \sum_{j=0}^{N_c-1} 1 - \cos(\Delta\theta_{i,j+1}), \\ \quad \text{if } \Delta\theta_{i,j+1} = \Delta\theta_{ij}. \end{cases} \tag{36}$$

For large signal-to-noise ratios, and as the number of bits considered grows large, the probability of error is dominated by a few error events with small distances. Hence, a minimum distance between any two signals will dominate and be defined as the "free" distance where

$$d_{\text{free}}^2(N) = \lim_{N \to \infty} \min d_{mn}^2(N), \qquad m \neq n. \tag{37}$$

Since the bit differences are not constant over a bit interval due to the spreading sequence as discussed previously, the use of the Viterbi algorithm to find the minimum distance cannot be used, as suggested by Aulin [2] and Hsu [4]. However, assuming all signal paths in the trellis structure begin at some point nT_s, that is, assuming error-free transmission up to time nT_s, the state transition diagrams can be used to determine the shortest distance to return to the all-zero path. This distance will become the minimum distance for the code and determine the constraint length.

The determination of the minimum distance is important for several reasons. First, from (34), the free distance term will dominate the error probability for high signal-to-noise ratios. Second, the depth or number of bits required to obtain the free distance will determine the minimum decoding delay for a Viterbi algorithm processor to obtain maximal performance. Finally, if the entire distance distribution for a given code is known, the performance in relation to the minimum distance can be determined. For example, if a code has a significant number of error events at distances that are close to the minimum distance, its performance may be worse than a code with a single, but rarely occurring minimum distance.

From (34) and (36), the probability of bit error is upper bounded by

$$P_b \leq \sum_{i=1}^{\infty} \mu_i \nu_i Q\left(\sqrt{\frac{d_i^2 E_c}{N_0}}\right). \tag{38}$$

Invoking the signal flow graph, and difference state analysis of Hsu [4] and Anderson [1], and assuming a two modulation index code with an infinite length decoder, this bit error bound reduces to (see [6] for details)

$$P_{b|1} \leq Q\left(\sqrt{\frac{d_f^2 E_c}{N_0}}\right) e^{d_f^2 E_c/2N_0} \left.\frac{\partial T_1}{\partial E}\right|_{D = e^{-E_c/2N_0}, E=1, L=1} \tag{39}$$

or

$$P_{b|2} \leq Q\left(\sqrt{\frac{d_f^2 E_c}{N_0}}\right) e^{d_f^2 E_c/2N_0} \left.\frac{\partial T_2}{\partial E}\right|_{D = e^{-E_c/2N_0}, E=1, L=1}, \tag{40}$$

depending on the starting modulation index, and with T_1 and T_2 representing the overall path gain, depending on the starting index. Since it is equally likely that the receiver will begin decoding on either modulation index, the total probability of bit error is the average of the individual code bit error probabilities. Hence,

$$P_b = \tfrac{1}{2}(P_{b|1} + P_{b|2}). \tag{41}$$

As a final note, (39) can be rewritten

$$P_{b|i} \leq \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{d_f^2 E_c}{2N_0}}\right) e^{d_f^2 E_c/2N_0} \times \left.\frac{\partial T_i}{\partial E}\right|_{D = e^{-E_c/2N_0}, E=1, L=1} \tag{42}$$

or

$$P_{b|i} \leq \frac{\dfrac{\partial T_i}{\partial E}}{\sqrt{\dfrac{2\pi d_f^2 E_c}{N_0}}}. \tag{43}$$

In order to implement the error bound expressions, it is first necessary to determine the bit transition diagram for the specific modulation scheme and code selected. A recursive Fortran implemented routine was used to determine the distance over a bit interval and difference phase transitions for all possible beginning difference phases and all possible chip difference sequences where there are 2^{N_c} possible sequences if the bit difference is not zero. The bit difference transition diagrams can also be used to determine the minimum required distance for an error event, and thus the free distance for the modulation code. Figs. 7–10 show the resulting bit difference transition diagrams for the 1/2, 5/8 code. This modulation index code was selected due to its flat spectral characteristics and coding gain as found by Lereim [7]. The free distances were determined from the difference state diagrams to be 6.23 for *conventional* modulation and 7.24 for *modified* modulation.

With the difference state transitions and the free distance, (42) or (43) can be solved to determine the upper bound to the bit error probability. Once again, these equations were numerically implemented via Fortran code, with the partial differentiation accomplished with numerical differencing. Fig. 11 shows the resultant upper bounds to the bit error probability for the selected modulation indexes and a spreading rate of three chips per bit, while Fig. 12 reflects the bit error probability bounds for seven chips per bit. The circled path reflects *conventional* 1/2, 3/4 coding, the triangle path is for *conventional* 1/2, 5/8 coding, the plus path delineates *modified* 1/2, 3/4 coding, and *modified* 1/2, 5/8 coding is shown by X's. The unmarked path corresponds to DS/BPSK.

At this point, it is appropriate to comment on the relative merits of these findings vis-à-vis more realistic spreading rates. The results above incorporated spreading rates of three and seven chips per bit, while normal spreading rates are on the order of 1000 chips per bit. However, the number of possible chip difference sequences is on the order of 2^{N_c}, implying that for a chip rate of 1000 chips per bit, an unreasonably large number of chip difference sequences would have to be considered for each starting phase. As a result, only small order chip rates were considered in this work to verify the procedure.

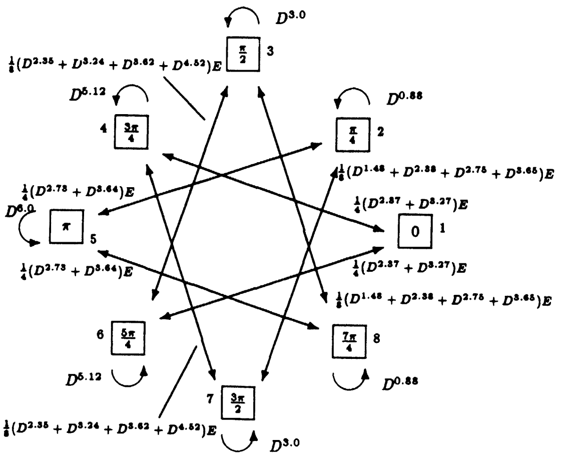

$D^{3.0}$
$\frac{\pi}{2}$ 3
$\frac{1}{8}(D^{2.35} + D^{3.24} + D^{3.62} + D^{4.52})E$
$D^{5.12}$
4 $\frac{3\pi}{4}$
$D^{0.88}$
$\frac{\pi}{4}$ 2
$\frac{1}{8}(D^{1.48} + D^{2.38} + D^{2.75} + D^{3.65})E$
$\frac{1}{4}(D^{2.73} + D^{3.64})E$
$D^{6.0}$
π 5
$\frac{1}{4}(D^{2.87} + D^{3.27})E$
0 1
$\frac{1}{4}(D^{2.73} + D^{3.64})E$
$\frac{1}{4}(D^{2.37} + D^{3.27})E$
$\frac{1}{8}(D^{1.48} + D^{2.38} + D^{2.75} + D^{3.65})E$
6 $\frac{5\pi}{4}$
$\frac{7\pi}{4}$ 8
$D^{5.12}$
$D^{0.88}$
7 $\frac{3\pi}{2}$
$\frac{1}{8}(D^{2.35} + D^{3.24} + D^{3.62} + D^{4.52})E$
$D^{3.0}$

Fig. 7. *Conventional* 1/2, 5/8 SSMH code; $N_c = 3$, difference states.

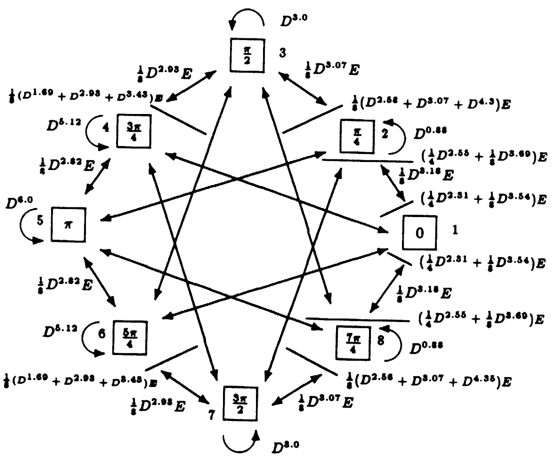

$D^{3.0}$
$\frac{\pi}{2}$ 3
$\frac{1}{8}D^{2.93}E$
$\frac{1}{8}D^{3.07}E$
$\frac{1}{8}(D^{1.69} + D^{2.93} + D^{3.43})E$
$\frac{1}{8}(D^{2.58} + D^{3.07} + D^{4.3})E$
$D^{5.12}$
4 $\frac{3\pi}{4}$
$\frac{\pi}{4}$ 2
$D^{0.88}$
$(\frac{1}{4}D^{2.55} + \frac{1}{8}D^{3.69})E$
$\frac{1}{8}D^{2.82}E$
$\frac{1}{8}D^{3.18}E$
$(\frac{1}{4}D^{2.31} + \frac{1}{8}D^{3.54})E$
$D^{6.0}$
5 π
0 1
$(\frac{1}{4}D^{2.31} + \frac{1}{8}D^{3.54})E$
$\frac{1}{8}D^{2.82}E$
$\frac{1}{8}D^{3.18}E$
$(\frac{1}{4}D^{2.55} + \frac{1}{8}D^{3.69})E$
$D^{5.12}$
6 $\frac{5\pi}{4}$
$\frac{7\pi}{4}$ 8
$D^{0.88}$
$\frac{1}{8}(D^{1.69} + D^{2.93} + D^{3.43})E$
$\frac{1}{8}(D^{2.56} + D^{3.07} + D^{4.35})E$
$\frac{1}{8}D^{2.93}E$
7 $\frac{3\pi}{2}$
$\frac{1}{8}D^{3.07}E$
$D^{3.0}$

Fig. 10. *Modified* 5/8 SSMH code; $N_c = 3$, difference states.

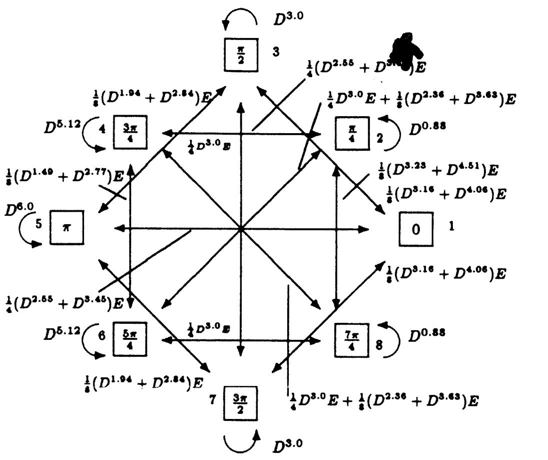

$D^{3.0}$
$\frac{\pi}{2}$ 3
$\frac{1}{4}(D^{2.55} + D)E$
$\frac{1}{8}(D^{1.94} + D^{2.84})E$
$\frac{1}{4}D^{3.0}E + \frac{1}{8}(D^{2.36} + D^{3.63})E$
$D^{5.12}$
4 $\frac{3\pi}{4}$
$\frac{1}{4}D^{3.0}E$
$\frac{\pi}{4}$ 2
$D^{0.88}$
$\frac{1}{8}(D^{1.49} + D^{2.77})E$
$\frac{1}{8}(D^{3.23} + D^{4.51})E$
$\frac{1}{8}(D^{3.16} + D^{4.06})E$
$D^{6.0}$
5 π
0 1
$\frac{1}{8}(D^{3.16} + D^{4.06})E$
$\frac{1}{4}(D^{2.55} + D^{3.45})E$
$D^{5.12}$
6 $\frac{5\pi}{4}$
$\frac{1}{4}D^{3.0}E$
$\frac{7\pi}{4}$ 8
$D^{0.88}$
$\frac{1}{8}(D^{1.94} + D^{2.84})E$
7 $\frac{3\pi}{2}$
$\frac{1}{4}D^{3.0}E + \frac{1}{8}(D^{2.36} + D^{3.63})E$
$D^{3.0}$

Fig. 8. *Conventional* 5/8, 1/2 SSMH code; $N_c = 3$, difference states.

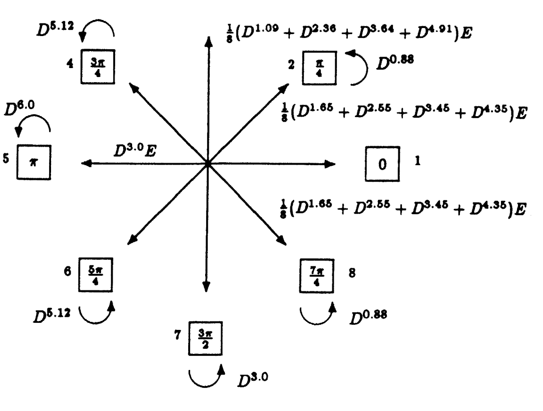

$D^{5.12}$
4 $\frac{3\pi}{4}$
$\frac{1}{8}(D^{1.09} + D^{2.36} + D^{3.64} + D^{4.91})E$
2 $\frac{\pi}{4}$
$D^{0.88}$
$D^{6.0}$
$\frac{1}{8}(D^{1.65} + D^{2.55} + D^{3.45} + D^{4.35})E$
5 π
$D^{3.0}E$
0 1
$\frac{1}{8}(D^{1.65} + D^{2.55} + D^{3.45} + D^{4.35})E$
6 $\frac{5\pi}{4}$
$\frac{7\pi}{4}$ 8
$D^{5.12}$
$D^{0.88}$
7 $\frac{3\pi}{2}$
$D^{3.0}$

Fig. 9. *Modified* 1/2 SSMH code; $N_c = 3$, eight difference states.

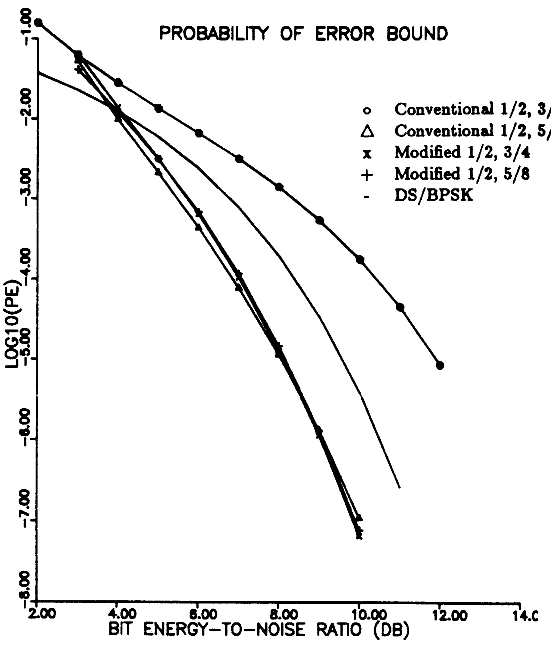


Fig. 11. Probability of bit error bounds, $N_c = 3$.

Numerical evaluations of the receiver structure consider more reasonable figures of 127 and 1023 chips per bit.

With the analytical bounds to the performance of the modulation scheme established, it is appropriate to consider the receiver structure of Section IV. The following section discusses the numerical evaluation of this receiver, and comparisons can be made to the established bounds.

VI. Receiver Evaluation

This section numerically evaluates the analytically derived receiver in an additive white Gaussian noise environment. The individual chip metrics are represented as sampled outputs from matched correlation filters, with correlated noise samples added to the noise-free filter outputs. This arrangement represents the transmitted signal corrupted by white noise and the received signal processed by correlation filters. The remainder of the receiver consists of a Viterbi algorithm processor, which uses the sum of chip metrics over a bit interval to form the bit branch metrics, and uses a sufficient delay to allow proper bit decoding.

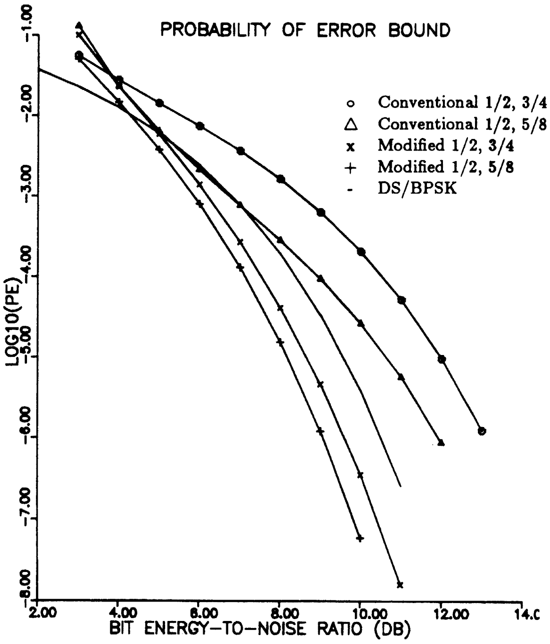


Fig. 12. Probability of bit error bounds, $N_c = 7$.

With the groundwork of Section IV, the signal components of the chip metric calculators are derived first. The correlated noise samples are then determined using uncorrelated white noise samples as a source and a linear transformation to achieve the appropriate correlated noise samples. With proper multiplexing to ensure correct chip and modulation index alignment, the chip metrics are then summed to form individual bit metrics over a bit interval. A state traceback implementation of the Viterbi algorithm then forms the decision processing. Throughout the evaluation, perfect phase synchronization, bit and chip timing, and modulation index synchronization are assumed.

Equations (18), (19), and (20) form the basis of the maximum likelihood bit and chip metrics for the decision making process. Expanding the expressions for the chip metrics on the initial phase angle yields quadrature forms of the correlators as shown in (21)–(25).

The received signal is

$$r(t) = s(t) + n(t) \tag{44}$$

where $n(t)$ is AWGN with variance $N_0/2$ and $s(t)$ is the transmitted signal

$$s(t) = \sqrt{\frac{2E_c}{T_c}} \cos\left(2\pi f_0 t + \frac{\pi h_{ij} \alpha_i c_{ij} t}{T_c} + \psi_{ij}(t)\right). \tag{45}$$

This expression represents a signal transmitted over any bit interval i and chip interval jT_c to $(j + 1)T_c$, with chip multiplier c_{ij}, data bit α_i, modulation index $h_{(i+j)\,\mathrm{mod}\,H} = h_{ij}$, and initial phase angle $\psi_{ij}(t)$.

After substitution, the expected value of the correlation expressions yields the signal components of the correlation filter output for a given α_i, c_{ij}, and h_{ij} [6]; hence, for

$\alpha_i c_{ij} = +1$,

$$\mathcal{E}\begin{Bmatrix} I_1 \\ Q_1 \\ I_2 \\ Q_2 \end{Bmatrix}_{|\alpha_i c_j = +1, h_{ij}} = \sqrt{\frac{E_c}{2}} \begin{Bmatrix} \cos(\psi(0)) \\ -\sin(\psi(0)) \\ \hat{\psi} \\ \check{\psi} \end{Bmatrix} \tag{46}$$

where

$$\hat{\psi} = \frac{\sin(2\pi h_{ij})}{2\pi h_{ij}} \cos(\psi(0)) - \frac{[1 - \cos(2\pi h_{ij})]}{2\pi h_{ij}} \sin(\psi(0)) \tag{47}$$

and

$$\check{\psi} = -\frac{[1 - \cos(2\pi h_{ij})]}{2\pi h_{ij}} \cos(\psi(0)) - \frac{\sin(2\pi h_{ij})}{2\pi h_{ij}} \sin(\psi(0)), \tag{48}$$

and for $\alpha_i c_{ij} = -1$,

$$\mathcal{E}\begin{Bmatrix} I_1 \\ Q_1 \\ I_2 \\ Q_2 \end{Bmatrix}_{|\alpha_i c_j = -1, h_{ij}} = \sqrt{\frac{E_c}{2}} \begin{Bmatrix} \acute{\psi} \\ \grave{\psi} \\ \cos(\psi(0)) \\ -\sin(\psi(0)) \end{Bmatrix} \tag{49}$$

where

$$\acute{\psi} = \frac{\sin(2\pi h_{ij})}{2\pi h_{ij}} \cos(\psi(0)) + \frac{[1 - \cos(2\pi h_{ij})]}{2\pi h_{ij}} \sin(\psi(0)) \tag{50}$$

and

$$\grave{\psi} = \frac{[1 - \cos(2\pi h_{ij})]}{2\pi h_{ij}} \cos(\psi(0)) - \frac{\sin(2\pi h_{ij})}{2\pi h_{ij}} \sin(\psi(0)). \tag{51}$$

The determination of the correlated noise samples begins with the quadrature equations above. Now, we seek the covariance matrix $\boldsymbol{C}$ of the four Gaussian signals I_1, Q_1, I_2, and Q_2.

The result is the covariance matrix below for the multivariate Gaussian random variables for the outputs of the correlation filters [6]:

$$C = \frac{N_0}{4}\begin{bmatrix} 1 & 0 & \dfrac{\sin(2\pi h_{ij})}{2\pi h_{ij}} & \dfrac{-1+\cos(2\pi h_{ij})}{2\pi h_{ij}} \\ 0 & 1 & \dfrac{1-\cos(2\pi h_{ij})}{2\pi h_{ij}} & \dfrac{\sin(2\pi h_{ij})}{2\pi h_{ij}} \\ \dfrac{\sin(2\pi h_{ij})}{2\pi h_{ij}} & \dfrac{1-\cos(2\pi h_{ij})}{2\pi h_{ij}} & 1 & 0 \\ \dfrac{-1+\cos(2\pi h_{ij})}{2\pi h_{ij}} & \dfrac{\sin(2\pi h_{ij})}{2\pi h_{ij}} & 0 & 1 \end{bmatrix} \tag{52}$$

Examination of this covariance matrix reveals that the values are constant over a chip interval and depend on the modulation index. As a result, the evaluation need only maintain proper index synchronization, and this covariance can be used to determine the correlated noise samples.

This covariance matrix gives the relationship that must exist between samples of the random quadrature components at the output of the correlation filters. The noise-free signal components that would appear at the outputs were discussed above. Now, using the covariance matrix, a method can be established to add correlated noise samples the noise-free signal components so that the random variable signal components will have the appropriate covariance.

As shown by Papoulis [8] and others, including Stark and Woods [14], if selected properly, a linear transformation on a correlated Gaussian process could yield an uncorrelated Gaussian process. The inverse approach is taken here where four uncorrelated Gaussian random variables are transformed by linear transformation to a multivariate Gaussian random variable with a known covariance.

Applying these results to the evaluation at hand implies that a linear transformation applied to four uncorrelated equal variance Gaussian random noise variables will yield four Gaussian random variables of known covariance if the linear transformation is the inverse of the matrix product of the eigenvalues and eigenvectors of the covariance matrix. Since (52) delineates the desired covariance matrix, the necessary transformation can be obtained from the eigenvalues and eigenvectors.

A Fortran program was used to determine the eigenvalues and eigenvectors from each of the covariance matrices for the selected modulation indexes. Then, using a random Gaussian generator with unity variance (N_0 = 2.0), the transformation matrices were determined. The correlated noise samples $N1C$, $N1S$, $N2C$, and $N2S$ were then added to the noise-free signal outputs of the correlation filters to produce noise corrupted correlation metrics.

With the noise-corrupted outputs of the correlation filters, the chip and bit metrics can be calculated from (18)–(25) and must be accomplished for each possible phase state. A phase rotation network accomplished this calculation for each state θ_{ij}.

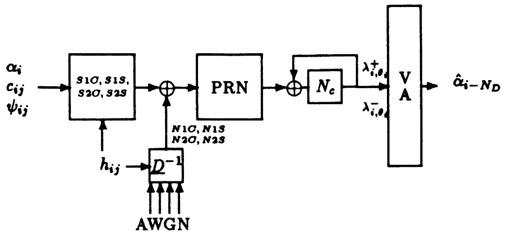


Fig. 13. Evaluation receiver.

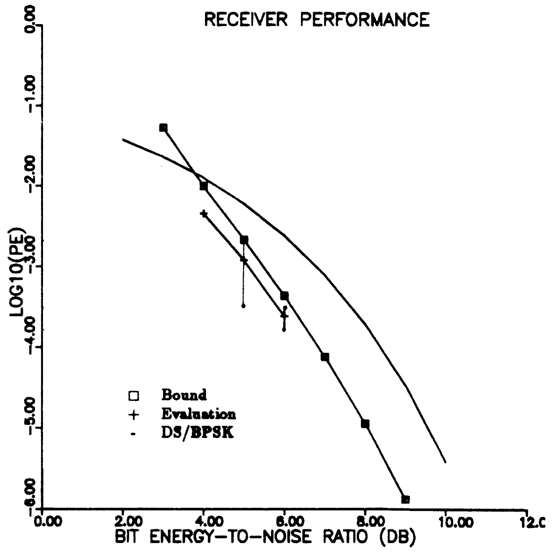


Fig. 14. *Conventional* 1/2, 5/8 SSMH code; N_c = 3 evaluation.

The overall bit metrics from (18) are just the sum of the chip metrics over a bit interval. A Viterbi algorithm processor then makes a bit decision after a delay of N_D bits. The algorithm was implemented using the state traceback method in Fortran to make a majority logic decision after observation over the predetermined bit delay. This delay was selected to exceed the minimum distance for optimal performance for the modulation indexes in use. An overview diagram of the evaluation receiver is shown in Fig. 13.

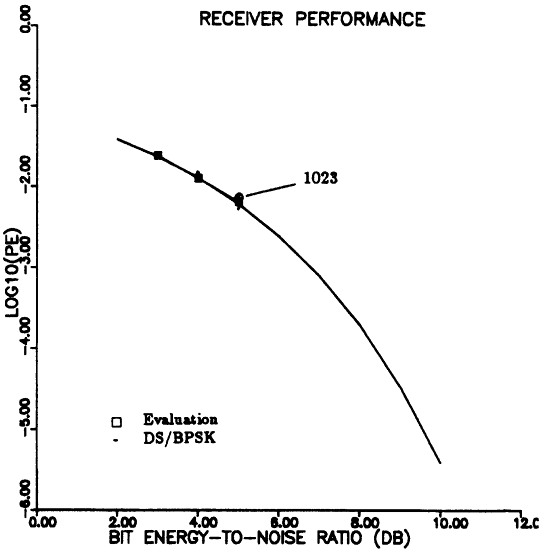


Fig. 15. *Conventional* 1/2, 5/8 SSMH code; N_c = 127 evaluation.

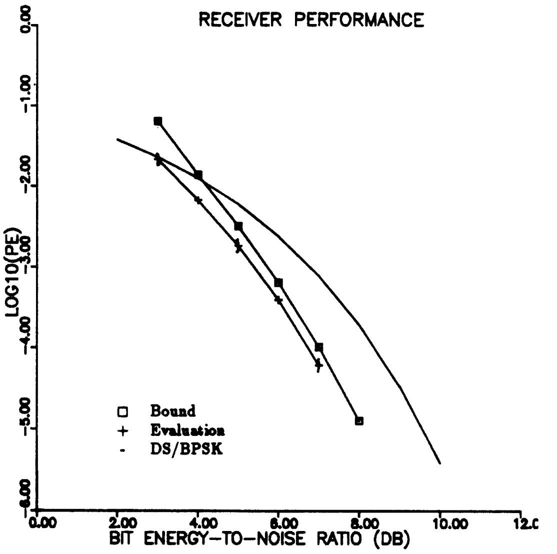


Fig. 16. *Modified* 1/2, 5/8 SSMH code; N_c = 3 evaluation.

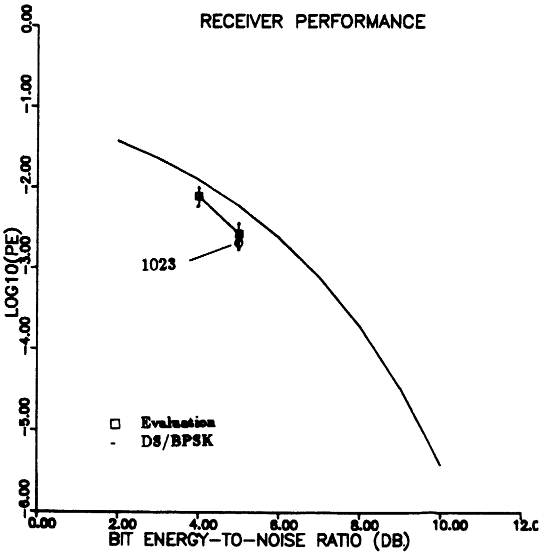


Fig. 17. *Modified* 1/2, 5/8 SSMH code; N_c = 127 evaluation.

Figs. 14–17 show the evaluation receiver bit error rate performance and the analytically determined performance bound for the indicated modulation criteria. The curves reflect 90% confidence intervals on the observed data that were obtained using the weak law of large numbers, and depict the exact bit error performance of DS/BPSK as a baseline comparison. As mentioned previously, evaluations for 127 chips per bit are shown without bounds. Additionally, isolated samples for 1023 chips per bit are shown with the evaluations for 127 chips per bit.

VII. Signal Detectability

The previous sections considered the issues of signal definitions and transmission characteristics, as well as detection by an intended receiver under nearly ideal conditions. Since the purpose of the signaling scheme is to create a signal structure with a low probability of intercept, it is appropriate to address the issue of detectability by an unintended as well as an intended receiver.

The detectability of SSMH signals in an AWGN environment is considered in this section where the unintended receiver must make a signal/no signal detection decision based on two hypotheses:

$$H1\colon r(t) = s(t) + n(t) \tag{53}$$

$$H0\colon r(t) = n(t). \tag{54}$$

It should be apparent that the longer the observation interval, the better the detectability would be; however, so that comparisons can be made to DS/BPSK, it is assumed that the observation interval is exactly one bit interval.

In order to obtain "best possible" capability, it is also assumed that the receiver knows everything about the transmitted signal, except the information bits and the spreading code sequence. This implies complete knowledge of

- the modulation index code and timing,
- the pulse shaping function,
- the bit and chip interval timing, and
- the carrier frequency and phase.

It is unlikely that the receiver will have knowledge of these modulation parameters, but for this analysis, every benefit is given so that an upper bound is attained.

For the signals in AWGN, the expected value of the generalized likelihood decision statistic reduces to [10], [15], [6]

$$\mathcal{E}\Lambda = \sum_{j=0}^{N_c-1} \frac{1}{2} e^{L_{cj}+L_{sj}} + \frac{1}{2} e^{L_{cj}-L_{sj}} \gtrless e^{E_b/N_0} \tag{55}$$

where, over a given bit and chip interval,

$$L_{cj} = \frac{2}{N_0}\sqrt{\frac{2E_c}{T_c}} \int_0^{T_c} r(t) \cos(2\pi f_0 t) \times \cos\left(\frac{\pi h_j t}{T_c}\right) dt \tag{56}$$

$$L_{sj} = \frac{2}{N_0}\sqrt{\frac{2E_c}{T_c}}\int_0^{T_c} r(t)\sin(2\pi f_0 t) \times \sin\left(\frac{\pi h_j t}{T_c}\right) dt, \tag{57}$$

and therefore,

$$\mathcal{E}\Lambda = \prod_{j=0}^{N_c-1} e^{L_{cj}}\cosh(L_{sj}) \gtrless e^{E_b/N_0}. \tag{58}$$

After taking the logarithm, the log-likelihood ratio becomes

$$\ln\Lambda = \sum_{j=0}^{N_c-1} L_{cj} + \ln\left[\cosh(L_{sj})\right] \gtrless \frac{E_b}{N_0}. \tag{59}$$

If N_c is large, which is typically the case, the probability distribution of the left side of (59) will approach a Gaussian distribution by the Central Limit Theorem. The mean value would be

$$\mu = \sum_{j=1}^{N_c} \mu_j \tag{60}$$

where

$$\mu_j = \mathcal{E}\left\{L_{cj} + \ln\left[\cosh(L_{sj})\right]\right\}. \tag{61}$$

Similarly, the variance would be

$$\sigma^2 = \sum_{j=1}^{N_c} \sigma_j^2 \tag{62}$$

where

$$\sigma_j^2 = \mathcal{E}\left\{\left(L_{cj} + \ln\left[\cosh(L_{sj})\right]\right)^2\right\} - \mu_j^2. \tag{63}$$

With this background, consideration can be given to determining the probability of false alarm and probability of detection.

The probability of false alarm is the probability of selecting hypothesis one, signal present, when in fact no signal is present. Thus,

$$P_F = \int_{\gamma}^{\infty} g(n)\, dn \tag{64}$$

where $g(n)$ is the distribution of the decision ratio when no signal is present and $\gamma = E_b/N_0$ is the decision threshold. As shown above, $g(n)$ will have a Gaussian distribution with statistics

$$\mu_g = N_c \mu_{gj} \tag{65}$$

$$\sigma_g^2 = N_c \sigma_{gj}^2. \tag{66}$$

The probability of false alarm can now be obtained from (64) since $g(n)$ is known to be Gaussian with mean μ_g and variance σ_g^2. To accomplish the sum of chip interval values and account for the changing modulation indexes (again assuming two indexes), these expressions become

$$\mu_g = \frac{N_c}{2}\mu_{gj|h_1} + \frac{N_c}{2}\mu_{gj|h_2} \tag{67}$$

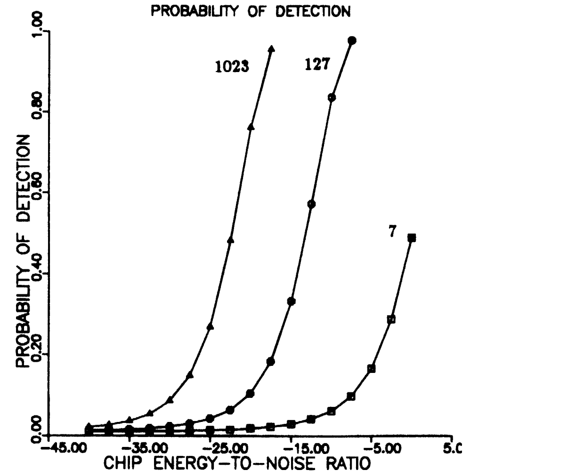


Fig. 18. *Conventional* 1/2, 5/8 code SSMH detectability.

$$\sigma_g^2 = \frac{N_c}{2}\sigma_{gj|h_1}^2 + \frac{N_c}{2}\sigma_{gj|h_2}^2 \tag{68}$$

and

$$P_F = \int_{E_b/N_0}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_g} e^{-[(n-\mu_g)^2/2\sigma_g^2]}\, dn. \tag{69}$$

A similar analysis is applied to finding the probability of detection P_D where now the likelihood ratio is considered under hypothesis $H1$ and there is a signal present. Thus,

$$P_D = \int_{E_b/N_0}^{\infty} N(\mu_s, \sigma_s^2)\, dn \tag{70}$$

where μ_s is the mean value sum of μ_{sj}'s, which are calculated from (61), but now considering the presence of the transmitted signal. Similarly, σ_s^2 is obtained from the second-order statistics with the signal present.

These expressions form the foundation for calculation of probability of detection and probability of false alarm. However, as shown by Polydoros [10], the two probabilities are interrelated, and a simpler expression can be formed to determine the probability of detection for a given false alarm rate. Polydoros and Weber [9], [10] showed that

$$P_D = Q\left[\frac{\sigma_{H0}Q^{-1}(P_{FA}) + \mu_{H0} - \mu_{H1}}{\sigma_{H1}}\right]. \tag{71}$$

This equation can now be used along with the means and variances under the two hypotheses to determine the probability of detection for any given value of false alarm rate. At the same time, the receiver operating characteristics or the expression of probability of detection versus probability of false alarm can be determined.

Equation (71) was implemented numerically in Fortran code to determine the probability of detection for given false alarm rates and for selected SSMH signaling schemes. The probability of detection and receiver oper-

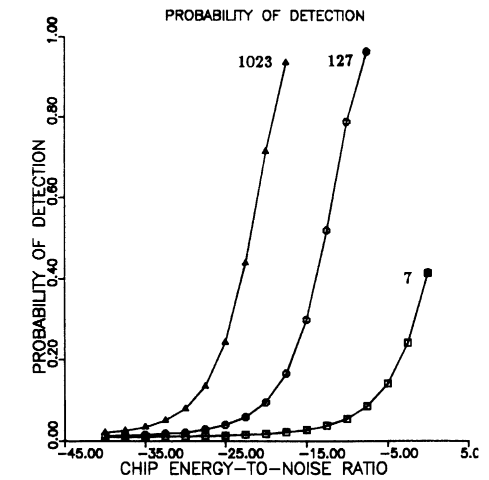


Fig. 19. *Modified* 1/2, 5/8 code SSMH detectability.

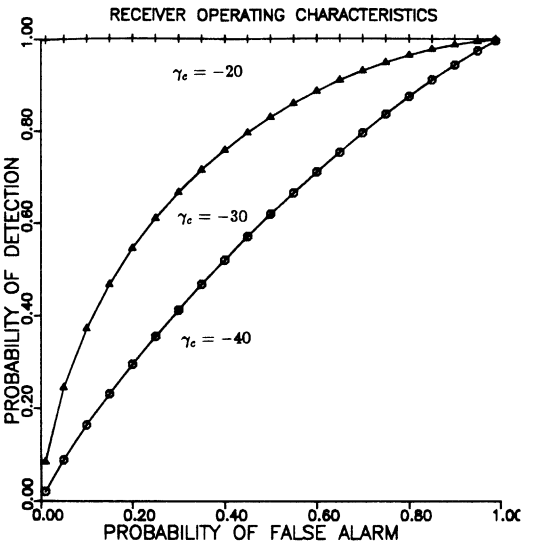


Fig. 20. *Conventional* 1/2, 5/8 code receiver operating characteristics.

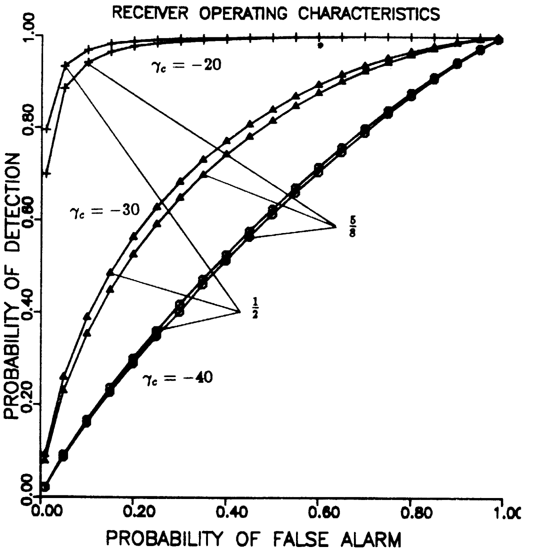


Fig. 21. *Modified* 1/2, 5/8 code receiver operating characteristics.

ating characteristics (P_D versus P_F) for selected modulation coding are shown in Figs. 18–21. For comparison purposes, the probability of detection curves are plotted for a consistent value of $P_F = 0.01$, $N_c = 7$, 127, and 1023, and $N_c = 1000$ chips per bit. It should also be noted that receiver operating curves for *modified* SSMH reflect the operation over two bit intervals to account for both modulation indexes.

VIII. Results and Conclusions

This section summarizes the results from the current work by comparing the results to existing spread-spectrum systems exemplified by direct sequence binary phase-shift-keying systems.

Following the definition of the new class of signals called spread-spectrum multi-h signals in Section II, the spectra for SSMH signals were analytically derived and evaluated. It was shown that the spectra were spread replicas of the spectra of the parent multi-h schemes. Thus, significant control over the spectra is gained through the modulation index selection in terms of the amount of spread, the relative flatness of the mainlobe, and the sidelobe rolloff. It was also shown that there would be no distinguishing features to the spectra, such as nulls or discrete components. Conversely, the spectra for DS/BPSK signals, by necessity through the $\text{sinc}^2(x)$ nature of the spectra, would have nulls at the chip rate. Additionally, square law detection would reveal many parameters of a DS/BPSK signals (carrier frequency and chip rate), while no information is readily apparent on SSMH signals until the signal is raised to the power of the denominator of the modulation indexes. This comparison indicates that SSMH signals allow much more control over the transmitted spectral shape, and the spectra will be much less susceptible to parameterization than DS/BPSK signals.

Turning to performance comparisons, the performance bounds of Section V are evoked, especially Figs. 11 and 12. These performance curves reflect the first inevitable path merger [1] and free distance, depending on the modulation code and the spreading rate. It can be seen that at an error rate of 10^{-5}, SSMH signals are upper bounded at approximately 8 dB, while DS/BPSK are bounded at approximately 10 dB. From this comparison, it appears that SSMH can perform more reliably than DS/BPSK for a given signal-to-noise ratio. It should be noted though, that at very low signal-to-noise ratios, the discrepancy in performance decreases. For example, at 4 dB of signal-to-noise ratio, the bit error rates are essentially the same. Thus, in the range of 2–4 dB, DS/BPSK may perform better, but above 4 dB, SSMH would be a wiser choice.

With these bounds in mind, the detectability of these signals can be considered. From Section VII, the detectability curves for 1/2, 5/8 *conventional* and *modified* (Figs. 18 and 19) are evoked. It should be remembered that these measures of detectability were based on an unintended receiver having almost complete knowledge of the transmitted signal.

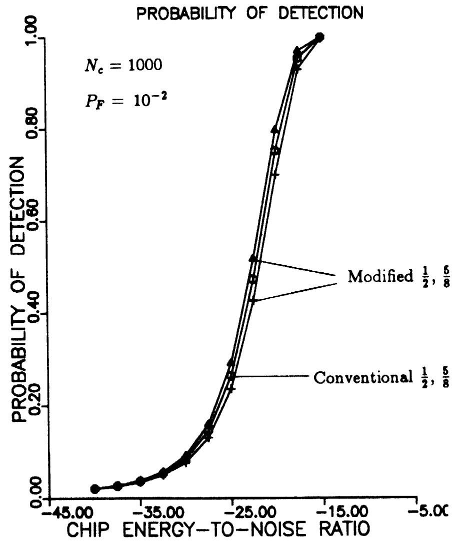


Fig. 22. *Conventional* and *modified* 1/2, 5/8 code SSMH detectability.

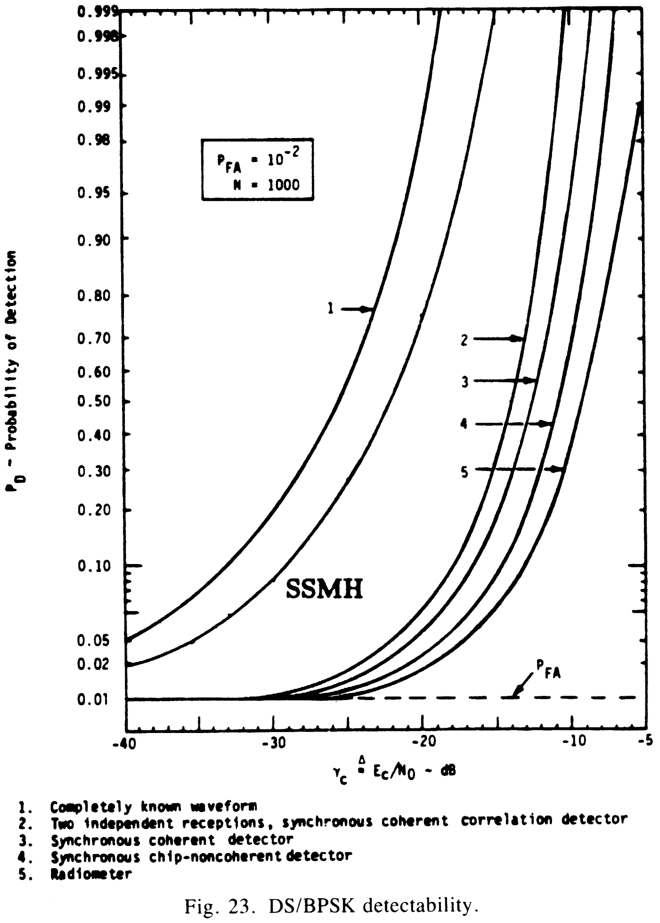


Fig. 23. DS/BPSK detectability.

The probability of detection curves for equivalent parameter $N_c = 1000$, $P_F = 10^{-2}$, *conventional* and *modified* 1/2, 5/8 code SSMH systems are shown in Fig. 22.

Last, the detectability curves for DS/BPSK are reproduced from the Axiomatix Corporation study on LPI waveforms by Polydoros and Weber [9], [10] in Fig. 23. The detectability curve for a completely known 1/2, 5/8 code SSMH signal has been added for comparison to the completely known DS/BPSK signal, as shown by curve 1.

Recalling that the original purpose of the design of SSMH waveforms was to create a signaling format that had a low probability of intercept, the performance capabilities can now be reconciled with the signal detectability. The performance curves show that for an intended receiver operating at a received signal-to-noise ratio of 7 dB, the probability of error performance is bounded at approximately 10^{-4}. After conversion of this bit energy-to-noise ratio to chip energy-to-noise ratio (-23 dB at 1000 chips per bit), it is clear from Fig. 22 that even with knowledge of the signal and the high false alarm rate, the unintended receiver would have unacceptably low probabilities of detection of less than one half.

The comparison of the detectability results for SSMH to the detectability of DS/BPSK reveals that a completely known (except for information sequence and spreading sequence) SSMH signal is less detectable than the corresponding completely known DS/BPSK signal. This is a new and significant result showing the potential for this class of continuous phase spread-spectrum signals.

It is worthwhile to mention the performance of SSMH in a stressed environment. This study has assumed complete coherency and synchronization, and has not considered the effects of jamming, interference, or multipath fading. It must be remembered that continuous phase modulated systems rely on the fact that the information is contained in the memory of the phase transitions. As a result, complete coherency is not absolutely required. The implication in this statement is that noncoherent systems may perform with little degradation compared to coherent systems. This has been verified by Premji and Taylor [11], [12]. For the same reason, the loss of individual chips or strings of chips due to fading or interference will not have the deleterious effect that a similar loss would have on DS/BPSK systems. All of these factors and effects are fertile areas of potential research.

In summary, this work has introduced a new method of spread-spectrum signaling known as spread-spectrum multi-*h* modulation. This technique has been shown to be a viable alternative for low probability of intercept communications.

References

[1] J. B. Anderson *et al.*, *Digital Phase Modulation.* New York: Plenum, 1986.

[2] T. Aulin and C. Sundberg, "On the minimum Euclidean distance for a class of signal space codes," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 43-55, Jan. 1982.

[3] G. Forney, "The Viterbi algorithm," *Proc. IEEE*, vol. 61, pp. 268-278, Mar. 1973.

[4] C.-D. Hsu, "Multi-*h* phase coding—Its theory and design," Ph.D. dissertation, Univ. Virginia, Charlottesville, 1981.

[5] D. E. Jackson, "Bandwidth efficient communication and coding," Ph.D. dissertation, Univ. California, Los Angeles, 1980.

[6] W. D. Lane, "Spread spectrum multi-*h* modulation," Ph.D. dissertation, Georgia Inst. Technol., Atlanta, 1988. This reference is available from University Microfilms Inc., Ann Arbor, MI, publ. 89-04817.

[7] A. Lereim, "Spectral properties of multi-h phase codes," Commun. Res. Lab., McMaster Univ., Hamilton, Ont., Canada, Rep. CRL-57, July 1978.

[8] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 2nd ed. New York: McGraw-Hill, 1984.

[9] A. Polydoros *et al.*, "Advanced LPI intercept detector research," Axiomatix Corp., Tech. Rep. R8403-1, Mar. 1984.

[10] A. Polydoros and C. L. Weber, "Detection performance considerations for direct-sequence and time-hopping LPI waveforms," *IEEE J. Select. Areas Commun.*, vol. SAC-3, pp. 727-744, Sept. 1985.

[11] A. Premji and D. Taylor, "A practical receiver structure for multi-h CPM signals," *IEEE Trans. Commun.*, vol. COM-35, pp. 901-908, Sept. 1987.

[12] ——, "Receiver structures for multi-h signaling formats," *IEEE Trans. Commun.*, vol. COM-35, pp. 439-451, Apr. 1987.

[13] R. Sadr, "Receiver design and analysis for generalized minimum shift keying modulation techniques," Ph.D. dissertation, Univ. California, Los Angeles, 1983.

[14] H. Stark and J. M. Woods, *Probability, Random Processes, and Estimation Theory for Engineers.* Englewood Cliffs, NJ: Prentice-Hall, 1986.

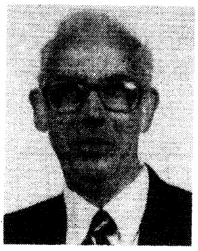
[15] H. L. VanTrees, *Detection Estimation, and Modulation Theory: Part I, Detection, Estimation, and Linear Modulation Theory.* New York: Wiley, 1968.

[16] A. J. Viterbi and J. K. Omura, *Digital Communication and Coding.* New York: McGraw-Hill, 1979.

[17] S. Wilson and R. Gaus, "Power spectra of multi-h phase codes," *IEEE Trans. Commun.*, vol. COM-29, pp. 250-256, Mar. 1981.

William D. Lane (S'85-M'88) was born in Denver, CO, on February 25, 1948. He received the B.S. degree from the U.S. Military Academy, West Point, NY, in 1970, and the M.S. degree in electrical engineering in 1978 and the Ph.D. degree in 1988, both from Georgia Institute of Technology, Atlanta. Additionally, he received the M.B.A. degree in management from C. W. Post College, NY, in 1981.

Following commissioning in the U.S. Army in 1970, he has served in various assignments worldwide including division, corps, joint task force, and special operations signal units, as well as with the Defense Communications Agency. He has served as an Instructor, Assistant Professor, and currently serves as an Associate Professor of Electrical Engineering in the Department of Electrical Engineering and Computer Science, U.S. Military Academy. He currently holds the rank of Lieutenant Colonel. His interests include communications systems and spread spectrum communications.



Aubrey M. Bush (S'58-M'60-SM'74) received the B.E.E. degree in 1959 and the M.S.E.E. degree in 1961, both from the Georgia Institute of Technology, Atlanta, and the Sc.D. degree in electrical engineering in 1965, from the Massachusetts Institute of Technology, Cambridge.

From 1955 to 1959 he was employed by the Lockheed-Georgia Company. During 1960-1961 he was an Instructor at the Georgia Institute of Technology. He was employed by Honeywell EDP, Needham, MA, over the summer of 1961. During the summers of 1965 and 1966 he was employed by Radiation, Inc., Melbourne, FL (now Harris Corporation). Since 1965 he has been again on the faculty of the School of Electrical Engineering, Georgia Institute of Technology, where he is currently Professor of Electrical Engineering. He has worked and taught in the areas of communications theory, systems theory, and digital signal processing. His current interests include mobile radio, digital transmission, and telecommunication networks.

Dr. Bush served as Chairman of the Atlanta Section of the IEEE in 1973-1974, as General Chairman of NTC'73, and is currently Technical Program Chairman for ICC/SUPERCOMM'90.

A Method of a Spread-Spectrum Radar Polyphase Code Design

MIROSLAV L. DUKIĆ, MEMBER, IEEE, AND ZORAN S. DOBROSAVLJEVIĆ, MEMBER, IEEE

Abstract—**In this paper, a new method of pulse compression polyphase code synthesis based on the aperiodic autocorrelation function properties and the assumption of coherent processing in the radar receiver is presented. Properties of the proposed code, including spectrum, ambiguity function, precompression bandlimiting effects, and the noise/clutter influence are analyzed. The main features of the synthesized code are the absence of sidelobes in the compressed pulse and its tolerance of precompression bandlimiting limitations. The comparison to the known classes of pulse compression codes is made.**

I. Introduction

In designing a radar system which utilizes the advantages of pulse compression, choice of the appropriate waveform is the most important. Numerous methods of radar pulse modulation, which make it possible to obtain the advantages of the pulse compression techniques, are known and widely analyzed [1]–[4]. Among these, the polyphase codes offer some convenience in comparison to the analog techniques, such as chirp modulation. Polyphase codes offer lower sidelobes in the compressed signal and easier digital processing techniques implementation.

In this paper, a new method of a uniform amplitude, polyphase pulse compression code synthesis is presented. The code synthesis is based on the properties of the aperiodic autocorrelation function and the assumption of coherent radar pulse processing in the receiver. The autocorrelation function and amplitude spectrum are analyzed. It is shown that in the absence of Doppler shift, the compressed and coherently demodulated pulse has the shape of a δ function.

It is shown, by computer simulation of Butterworth and Gaussian filters, that the precompression bandlimiting does not cause the appearance of sidelobes in the compressed pulse.

The distributed clutter environment is also simulated, and the influence of the clutter on the average sidelobe level in the compressed pulse is analyzed. Finally, properties of the proposed code are compared to other well-known codes, such as Huffman, Frank, and P codes.

The subject of this paper is divided into six sections. Theoretical analysis of the code proposed is presented in Section II, while in Section III, the analysis of its spectrum, ambiguity function, Doppler sensitivity, and clutter effects is established. In Section IV, a comparison of the proposed code to the other polyphase codes is presented. Section V provides one possible realization of the radar transmitter and receiver on the basis of pulse-by-pulse coherence. Section VI provides concluding remarks.

II. Theoretical Analysis

There are two main criteria that polyphase codes used in pulse compression systems should satisfy [1], [2], [4]:

1) the uncompressed radar pulse should have constant amplitude, i.e., it only has to be phase modulated, and

2) the aperiodic autocorrelation function should have very low sidelobes.

These criteria are used to estimate properties of the code presented in this paper.

A generalized form of a phase-modulated radar pulse can be written as

$$u(t) = \mathrm{Re}\left\{U_0\Pi(t - T)\exp\left[j\omega_0 t + j\varphi(t)\right]\right\} \quad (1)$$

where U_0 and ω_0 are the amplitude and the angular frequency of the carrier and $\varphi(t)$ is the phase modulation function. $\Pi(t)$ is the rectangular pulse of unit amplitude and of duration T.

Let us assume now that digital polyphase modulation is performed. According to (1), the complex low-frequency equivalent of the signal $u(t)$ is given by

$$C(t) = \exp\left[j\varphi(t)\right] = \exp\left[j\sum_{i=0}^{N-1}\varphi_i\Pi(t - iT_c)\right] \quad (2)$$

where φ_i is the ith element of the code, N is the code length, and T_c is the chip or code element duration.

The z transform of the time sequence $C[(i + 1/2)T_c]$ may be expressed as

$$C(z) = \sum_{i=0}^{N-1} z^{-i}\exp(j\varphi_i). \quad (3)$$

If a matched filter is used in the receiver, the low-frequency equivalent at the output of that filter, in the ideal case, is equal to the aperiodic autocorrelation function of the radar pulse. The z transform of the autocorrelation of

Manuscript received March 8, 1989; revised February 9, 1990. This paper was presented at the URSI International Symposium on Signals, Systems, and Electronics (ISSSE'89), Erlangen, West Germany, September 1989.

The authors are with the Department of Communications, Faculty of Electrical Engineering, University of Belgrade, Bul. Revolucije 73, Belgrade, Yugoslavia.

IEEE Log Number 9035802.

the signal $C(t)$ will be

$$\begin{aligned} R(z) &= z^{-(N-1)} \cdot C(z) \cdot C^*(1/z^*) \\ &= z^{-(N-1)} \sum_{i=0}^{N-1} \sum_{k=0}^{N-1} z^{-(i-k)} \exp\left[j(\varphi_i - \varphi_k)\right]. \end{aligned} \tag{4}$$

In the synthesis of the polyphase code, it is presumed that the received radar pulse is coherently demodulated. The condition that the real part of the compressed pulse should be a δ function can be expressed as

$$\sum_{\substack{i=0 \\ i \neq k}}^{N-1} \sum_{k=0}^{N-1} z^{-(i-k)} \operatorname{Re}\left\{\exp\left[j(\varphi_i - \varphi_k)\right]\right\} = 0 \tag{5}$$

where Re (·) stands for "real part of."

This condition gives $(N - 1)$ nonlinear equations, each with N variables, $\varphi_0, \varphi_1, \cdots, \varphi_{N-1}$. Replacing the substitution $\Delta_i = \varphi_i - \varphi_{i-1}$ in (5), we obtain the following system of $(N - 1)$ equations, with the same number of variables, Δ_i, $i = 1, \cdots, N - 1$:

$$\begin{aligned} &\cos(\Delta_1 + \Delta_2 + \cdots + \Delta_{N-1}) = 0, \\ &\cos(\Delta_1 + \cdots + \Delta_{N-2}) \\ &\quad + \cos(\Delta_2 + \cdots + \Delta_{N-1}) = 0, \\ &\quad \vdots \\ &\cos(\Delta_1 + \Delta_2 + \Delta_3) + \cos(\Delta_2 + \Delta_3 + \Delta_4) \\ &\quad + \cdots + \cos(\Delta_{N-3} + \Delta_{N-2} + \Delta_{N-1}) = 0, \\ &\cos(\Delta_1 + \Delta_2) + \cos(\Delta_2 + \Delta_3) \\ &\quad + \cdots + \cos(\Delta_{N-2} + \Delta_{N-1}) = 0, \\ &\cos(\Delta_1) + \cos(\Delta_2) + \cdots + \cos(\Delta_{N-1}) = 0. \end{aligned} \tag{6}$$

The solution to these equations can be used to obtain the elements of the new polyphase code.

This system is inconvenient for numerical solution. It is easy to solve the system analytically for small N. Thus, for $N = 2$, 3, 4, and 6, the solutions are written as follows:

$$\begin{aligned} &N = 2\colon \Delta_1 = \pi/2, \\ &N = 3\colon \Delta_1 = \pi/4, \quad \Delta_2 = 5\pi/4, \\ &N = 4\colon \Delta_1 = 0, \quad \Delta_2 = \pi/2, \quad \Delta_3 = \pi, \\ &N = 6\colon \Delta_1 = 0, \quad \Delta_2 = \pi/4, \quad \Delta_3 = \pi/2, \\ &\qquad \Delta_4 = 3\pi/4, \quad \Delta_5 = \pi. \end{aligned} \tag{7}$$

These small values of N are not of practical interest. However, they can be used for the synthesis of longer codes. With one of the solutions $\{\Delta_i^{(1)}\}$, $i = 1, \cdots, N - 1$, given by (7) for some small N (superscript (1) indicates the first step of longer code synthesis), it is possible to obtain the vector $\{\Delta_i^{(2)}\}$, $i = 1, \cdots, 4N - 1$, which is the solution of system (6) with $4N - 1$ equations. Variables in vector $\{\Delta_i^{(2)}\}$, $i = 1, \cdots, 4N - 1$ are given by the following relations, as shown in the Appendix:

$$\begin{aligned} &\Delta_i^{(2)} = 0, && i = 1, 3, 5, \cdots, 2N - 1 \\ &\Delta_{2i}^{(2)} = \Delta_i^{(1)}, && i = 1, N - 1 \\ &\Delta_{2N}^{(2)} = \pi/2, && \\ &\Delta_{4N-i}^{(2)} = \pi - \Delta_i^{(2)}, && i = 1, 2N - 1. \end{aligned} \tag{8}$$

By repeated use of (8), it is possible to obtain the solution of (6) for very long code sequences. Once the $\{\Delta_i\}$, $i = 1, \cdots, N - 1$ is gained, the elements of the polyphase code that satisfy conditions 1) and 2) will be given from the following substitutions:

$$\begin{aligned} &1)\ \text{we accept } \varphi_1 = 0, \\ &2)\ \varphi_{i+1} = \varphi_i + \Delta_i, \qquad i = 1, \cdots, N - 1. \end{aligned} \tag{9}$$

Example: If we want to obtain a code with length $N = 64$, we can start the code synthesis from the vector of differences $\{\Delta_i^{(1)}\}$ for $N = 4$:

$$\Delta_1^{(1)} = 0, \quad \Delta_2^{(1)} = \pi/2, \quad \Delta_3^{(1)} = \pi \tag{10}$$

as we can see from expression (7).

Using relations (8) for $N = 4$, we obtain the following elements of the vector $\{\Delta_i^{(2)}\}$:

$$\begin{aligned} &\Delta_1^{(2)} = 0, \quad \Delta_3^{(2)} = 0, \quad \Delta_5^{(2)} = 0, \quad \Delta_7^{(2)} = 0; \\ &\Delta_2^{(2)} = \Delta_1^{(1)} = 0, \quad \Delta_4^{(2)} = \Delta_2^{(1)} = \pi/2, \\ &\qquad \Delta_6^{(2)} = \Delta_3^{(1)} = \pi; \\ &\Delta_8^{(2)} = \pi/2; \\ &\Delta_{15}^{(2)} = \pi - \Delta_1^{(2)} = \pi, \quad \Delta_{14}^{(2)} = \pi - \Delta_2^{(2)} = \pi, \cdots, \\ &\qquad \Delta_9^{(2)} = \pi - \Delta_7^{(2)} = \pi. \end{aligned} \tag{11}$$

In the above expressions, the superscript (2) indicates the second step in the proposed code synthesis.

The obtained vector $\{\Delta_i^{(2)}\}$, $i = 1, \cdots, 15$ is

$$\begin{aligned} \{\Delta_i^{(2)}\} = \{&0, 0, 0, \pi/2, 0, \pi, 0, \pi/2, \pi, 0, \pi, \\ &\pi/2, \pi, \pi, \pi\}. \end{aligned} \tag{12}$$

Now, using relations (8) with vector $\{\Delta_i^{(2)}\}$ for $N = 16$, we obtain the following vector $\{\Delta_i^{(3)}\}$:

$$\begin{aligned} &\Delta_1^{(3)} = 0, \quad \Delta_3^{(3)} = 0, \cdots, \Delta_{31}^{(3)} = 0; \\ &\Delta_2^{(3)} = \Delta_1^{(2)} = 0, \quad \Delta_4^{(3)} = \Delta_2^{(2)} = 0, \cdots, \\ &\qquad \Delta_{30}^{(3)} = \Delta_{15}^{(2)} = \pi; \\ &\Delta_{32}^{(3)} = \pi/2; \\ &\Delta_{63}^{(3)} = \pi - \Delta_1^{(3)} = \pi, \quad \Delta_{62}^{(3)} = \pi - \Delta_2^{(3)} = \pi, \cdots, \\ &\qquad \Delta_{33}^{(3)} = \pi - \Delta_{31}^{(3)} = \pi. \end{aligned} \tag{13}$$

Vector $\{\Delta_i^{(3)}\} = \{0, 0, \cdots, \pi\}, i = 1, \cdots, 63$ is the solution of system (6) for $N = 64$ so it can be used to obtain the code for $N = 64$. Using relations (9), we have the proposed code $\{\varphi_i\}, i = 1, \cdots, 64$ that satisfies conditions 1) and 2):

$$\varphi_1 = 0;$$

$$\varphi_2 = \varphi_1 + \Delta_1^{(3)} = 0, \quad \varphi_3 = \varphi_2 + \Delta_2^{(3)} = 0, \cdots,$$

$$\varphi_{64} = \varphi_{63} + \Delta_{63}^{(3)} = \frac{3\pi}{2}. \tag{14}$$

Finally, the whole set of code elements of the proposed code, for $N = 64$, is given by the following relation:

$$\begin{aligned}\{\varphi_i\} = \{&0, 0, 0, 0, 0, 0, 0, 0, \pi/2, \pi/2, \pi/2, \pi/2,\\ &3\pi/2, 3\pi/2, 3\pi/2, 3\pi/2, 0, 0, \pi, \pi, \pi, \pi,\\ &0, 0, \pi/2, \pi/2, 3\pi/2, 3\pi/2, \pi/2, \pi/2, 3\pi/2,\\ &3\pi/2, 0, \pi, \pi, 0, 0, \pi, \pi, 0, \pi/2, 3\pi/2,\\ &\pi/2, \pi/2, 3\pi/2, 3\pi/2, \pi/2, 3\pi/2, 0, \pi,\\ &0, \pi, \pi, 0, \pi, 0, \pi/2, 3\pi/2, \pi/2, 3\pi/2,\\ &\pi/2, 3\pi/2, \pi/2, 3\pi/2\}.\end{aligned} \tag{15}$$

III. Performance Analysis

A. *Autocorrelation Function and Code Spectrum*

One of the main properties of the proposed code is that the real part of the aperiodic autocorrelation function has the shape of a δ function. For the code length of $N = 64$, the real part of the compressed pulse is shown in Fig. 1.

The amplitude spectrum of the uncompressed pulse is shown in Fig. 2. The overall shape of the spectrum is $\sin(\pi f T)/(\pi f T)$ where T is the duration of the modulated radar pulse. The fine structure of the spectrum is caused by the phase modulation of the pulse. This fine structure is not symmetrical, which makes this code different from the P code and makes it similar to the Huffman code spectrum. This is the result of the lack of symmetry in the time domain of the radar pulse modulated with the proposed code. Symmetry of P codes in the time domain results in a symmetry in the frequency domain [9].

B. *Ambiguity Function and Doppler Properties*

Woodward's ambiguity function is one of the most useful tools for determining the resolution properties of pulse compression signals. Based on the assumption of coherent demodulation of the received radar pulse, the real part of the ambiguity function $\chi(\tau, f_D)$ of the proposed code, obtained by computer simulation, is shown in Fig. 3. This figure shows only one half of the delay τ and Doppler frequency shift f_D plane because of the symmetry properties of the ambiguity function.

The central part of the ambiguity function is shown in Fig. 4. The analysis of this diagram clearly indicates that there are no sidelobes on the τ axis, but that time sidelobes become a significant part of the compressed pulse when a Doppler shift is introduced.

Polyphase codes generally have rather poor behavior in the presence of Doppler frequency shift. Greater Doppler shifts result in a rapid decrease of the compressed pulse amplitude and the appearance of secondary peaks in a

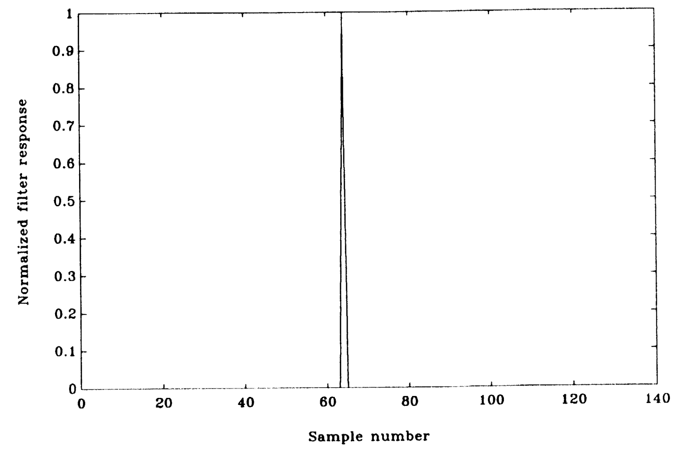


Fig. 1. The real part of the aperiodic autocorrelation function of the proposed code. The code length is 64 chips.

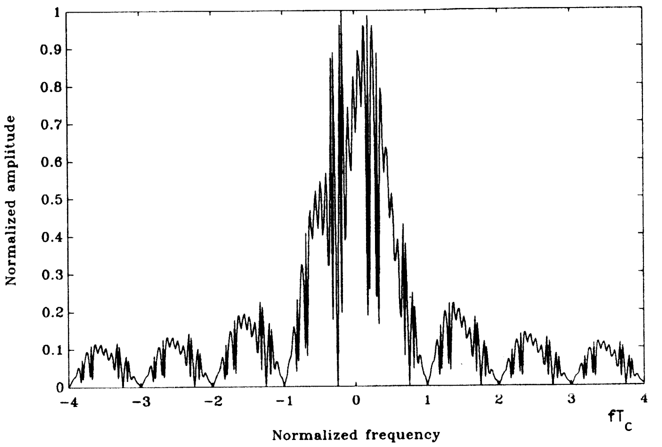


Fig. 2. The amplitude spectrum of the uncompressed radar pulse. The amplitude is normalized to the maximum value in the spectrum. The code length is 64, and the frequency is normalized to the chip duration T_c.

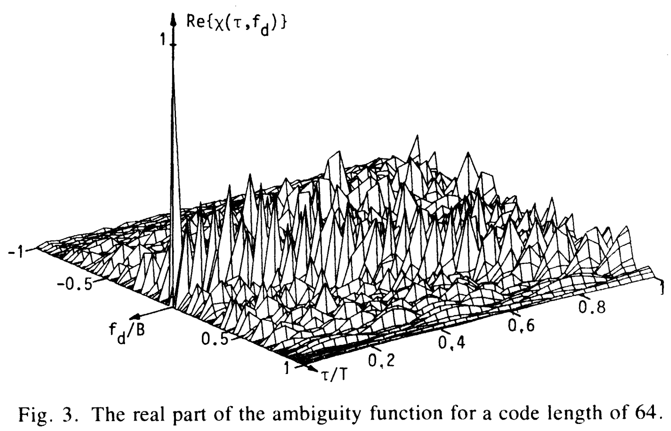


Fig. 3. The real part of the ambiguity function for a code length of 64.

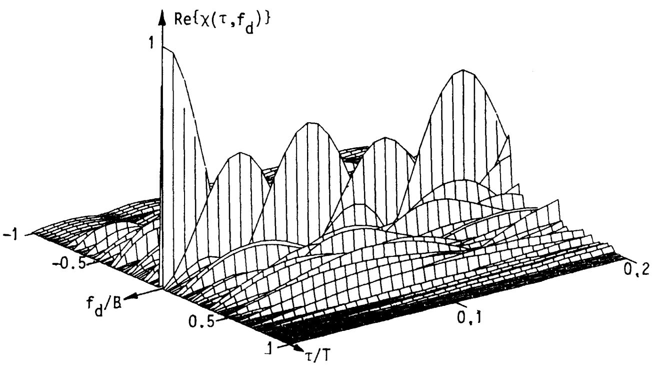


Fig. 4. The central part of the ambiguity function from Fig. 3.

compressed pulse, modulated with the proposed code. However, only small frequency shifts are of practical interest.

Figs. 5 and 6 present a comparison between a Frank polyphase code, length $N = 16$, the proposed code with the same length, and a Huffman code with length $N = 14$ and sidelobe amplitude $s = -22$ dB. Results are obtained by computer simulation.

Decrease of the compressed pulse amplitude versus Doppler frequency shift, normalized to chip duration T_c, is shown in Fig. 5. This decrease is identical for the Frank and proposed codes, and depends only on the code length N; it has the shape of a $\sin(x)/x$ function.

In Fig. 6, the maximum sidelobe, normalized to the compressed pulse amplitude, is shown as a function of the normalized Doppler frequency shift. One can see that the proposed code has a lower maximum sidelobe compared to the Frank and Huffman codes; this difference is distinctive for very small frequency shifts where the sidelobe level of the proposed code is decreasing to zero, while the level of the sidelobes for the Frank and Huffman codes keeps some constant level (accepted in a design for Huffman codes).

C. Precompression Bandlimiting Effects

The influence of the precompression filtering is analyzed on the code of length $N = 64$. Butterworth filters of first and fourth order and a Gaussian filter of fourth order are simulated. After bandlimiting, the polyphase signal is compressed in a filter simulated as a tapped-delay line.

The influence of the cutoff frequency f_g of the Butterworth and Gaussian filters on the peak height and duration of the compressed radar pulse is shown in Figs. 7 and 8.

The analysis of these diagrams indicates that the effect of the cutoff frequency is stronger for the higher order of the filter. Decreasing the peak amplitude is faster for the Gaussian and Butterworth filter of the first order. For the Butterworth filter of the fourth order, for the product $f_g T_c > 0.8$, the peak amplitude is almost constant, but for $f_g T_c < 0.8$, the peak is linearly decreasing. Based on Fig. 7, we can accept the $f_g T_c = 0.8$ as the narrowest band for the precompression bandlimiting.

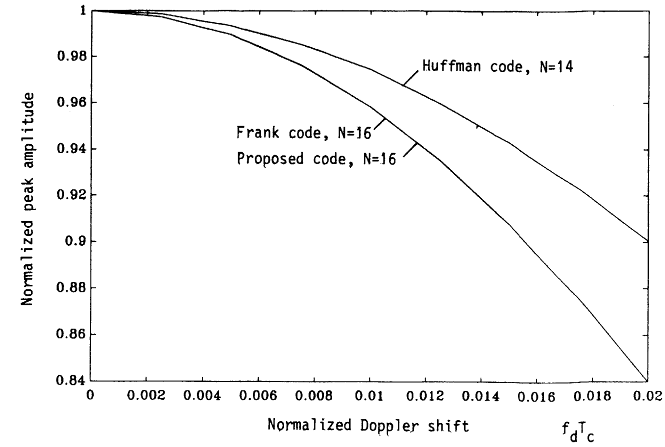


Fig. 5. The compressed pulse amplitude versus Doppler frequency shift f_D normalized to the chip duration T_c for Huffman code, length $N = 14$, Frank code, $N = 16$, and the proposed code, $N = 16$.

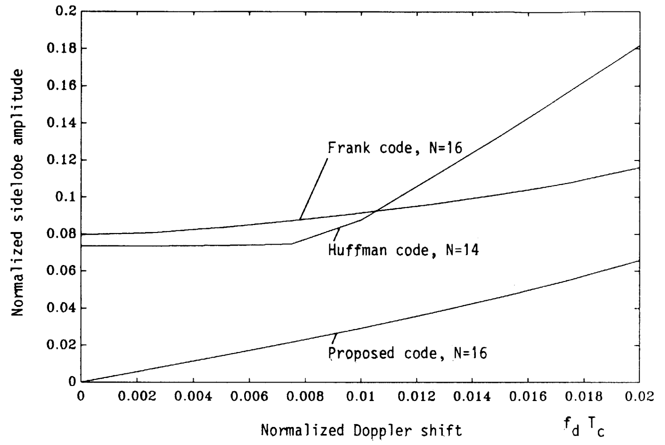


Fig. 6. The maximum sidelobe versus Doppler frequency shift f_D normalized to the chip duration T_c for Huffman code, length $N = 14$, Frank code, $N = 16$, and the proposed code, $N = 16$.

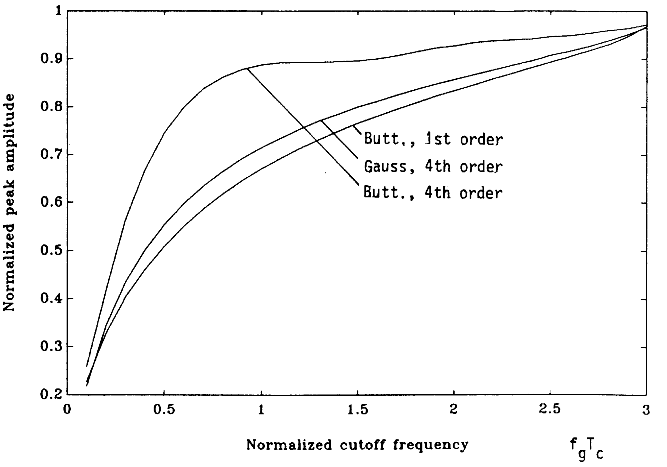


Fig. 7. The peak amplitude versus cutoff frequency f_g of the simulated filters used. The peak amplitude is normalized to the amplitude of the compressed pulse in the absence of the filtering. Code length is 64.

If the peak duration is defined at one half of the maximum value, results shown in Fig. 8 point out that the Butterworth filter of the first order has a slightly stronger influence on the compressed pulse prolonged duration than

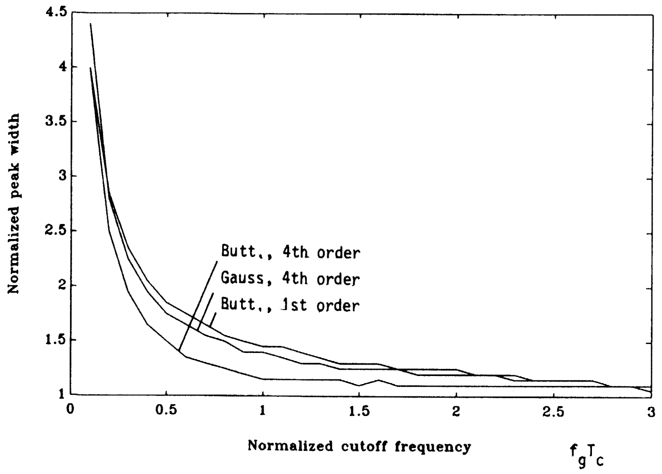


Fig. 8. The peak width versus cutoff frequency f_g of the filters used. The radar pulse width is measured at 0.5 of the peak amplitude. Code length is 64.

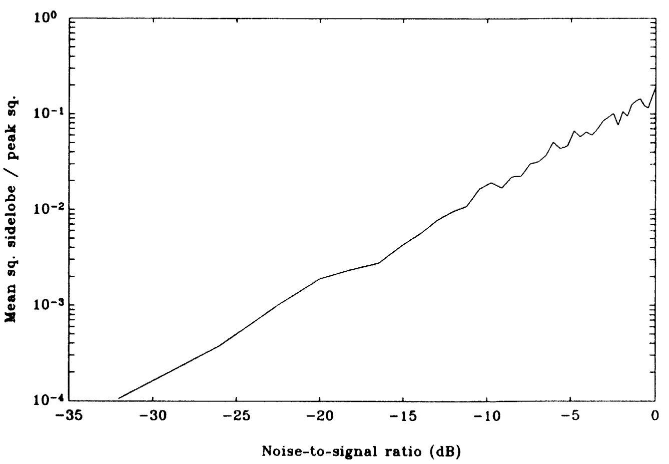


Fig. 9. The average sidelobe power of the compressed pulse versus noise-to-signal power ratio. The sidelobe power is normalized to the squared peak of the compressed pulse. Code length is 64.

the Gaussian filter does. The higher order of the filter causes the narrowing of the main spike in the compressed signal.

Having in mind all of the above mentioned, we can conclude that the precompression filtering does not cause the appearance of sidelobes, but broadens the compressed pulse.

D. Distributed Clutter Effects

Distributed clutter, such as rain or chaff, is modeled as Gaussian white noise added to the received radar pulse [4]. After coherent demodulation and compression, performed by computer simulation, the average power of the sidelobes is plotted in Fig. 9 as a function of the noise-to-signal power ratio. The average sidelobe power is normalized to the squared peak of the compressed pulse.

The relationship between the sidelobe power and the noise-to-signal ratio is practically linear in the log-log scale. The variations in Fig. 9 are due to the relatively small noise sample (equal to the duration of the uncompressed pulse) used in the simulation.

IV. Comparison of the Proposed Polyphase Code to Other Polyphase Codes

It was shown in the previous analysis that the assumption of coherent processing gives us the possibility to generate a polyphase code for which the compressed pulse has the shape of a δ function. This characteristic, together with the constant amplitude of the uncompressed pulse, make this code much more preferable compared to Huffman, Frank, or P codes [4]–[8]. When the small Doppler frequency shift is introduced, the excellent sidelobe suppression of this code is decreased, but the peak-to-sidelobe ratio is still better compared to that of the Huffman and Frank codes. The absence of sidelobes in the filtered and compressed pulse (in the absence of frequency shift) makes the proposed code favorable compared to the above-mentioned polyphase codes.

V. Implementation Technique of the Proposed Code

Coherent demodulation of the received radar pulse is not practical. Synthesis of the proposed code is based on the assumption of coherent demodulation. It is possible, however, to transmit the radar pulses modulated with the proposed code and complex conjugated pulses alternately. The radar pulse modulated with the proposed code is given by (2). The complex conjugate pulse is

$$C^*(t) = \exp\left[-j\varphi(t)\right] = \exp\left[-j\sum_{i=0}^{N-1}\varphi_i\Pi(t-iT_c)\right] \tag{16}$$

and the z transform of the compressed conjugated pulse $C^*(t)$ will be

$$R^{(*)}(z) = z^{-(N-1)}\sum_{i=0}^{N-1}\sum_{k=0}^{N-1} z^{-(i-k)}\exp\left[-j(\varphi_i-\varphi_k)\right]. \tag{17}$$

If it is possible to obtain pulse-by-pulse coherence in the receiver, pulses $C(t)$ and $C^*(t)$ can be summed after compression, and the z transform of this sum is using (4) and (17),

$$\begin{aligned} R_{\text{tot}}(z) &= R(z) + R^{(*)}(z) \\ &= 2z^{-(N-1)}\sum_{i=0}^{N-1}\sum_{k=0}^{N-1} z^{-(i-k)}\cos(\varphi_i-\varphi_k), \end{aligned} \tag{18}$$

which gives the same condition (5), system of equations (6), and the same result—ideal shape of the compressed pulse [10].

The configuration of the radar transmitter and receiver is shown in Fig. 10. The radar transmitter uses the proposed code $\{\varphi_i\}$, $i = 1, \cdots, N$ together with the code $\{-\varphi_i\}$, $i = 1, \cdots, N$ as a complementary pair. The ideal shape of the compressed pulse in the receiver is ob-

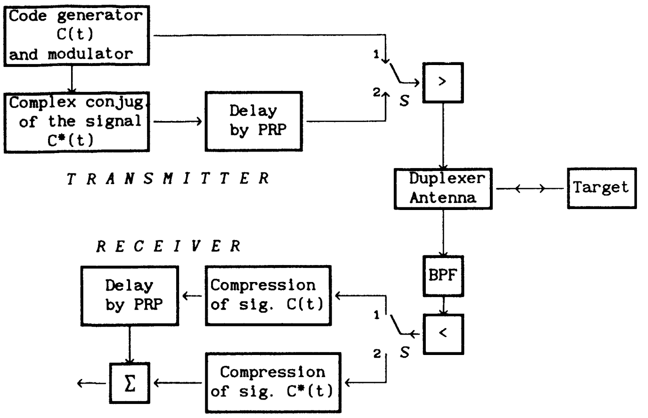


Fig. 10. Configuration of the low-frequency equivalent of the radar transmitter and receiver which uses the proposed code. Abbreviations are: PRP—pulse repetition period, BPF—bandpass filter. Switch S is turned from the position 1 to position 2 with PRP.

tained without coherent demodulation on the basis of the summation of the pairs of the received pulses.

VI. Conclusions

In this paper, a new method of polyphase pulse compression code synthesis has been proposed. The proposed code performance analysis is carried out on the basis of the autocorrelation and the ambiguity functions, and its sensitivity to the precompression bandlimiting and the presence of the distributed clutter.

Comparing this code to the well-known classes of Frank, P, and Huffman codes, we have shown that this code offers the ideal shape of the compressed pulse, and is more tolerant to limiting prior to pulse compression. The other properties are comparable to those of the other pulse compression codes.

The waveform generated from the proposed code has constant amplitude. As differences between consecutive elements of the proposed code can have values that are integer multiples of $\pi/4$ or $\pi/2$, obtained code elements also can have only four or eight different phase values on the unit circle, depending on the code length, as can be seen from example (15). This means that signal processing is very efficient in digital implementation techniques.

A weaknesses of the proposed code might be in the possibility of error in coherent processing in the radar receiver.

Appendix

In order to determine the recurrent relation (8) for the long proposed code synthesis, we note, first, that the system of equations (6) for $(4N - 1)$ can be written as

$$\sum_{i=1}^{k} \cos\left(\Delta_i + \Delta_{i+1} + \cdots + \Delta_{i+(4N-1)-k}\right) = 0, \qquad k = 1, 2, \cdots, 4N-1. \tag{A.1}$$

Equations of this system for odd values of k can be expressed as

$$\sum_{i=1}^{(k-1)/2} \cos\left(\Delta_i + \Delta_{i+1} + \cdots + \Delta_{i+(4N-1)-k}\right) + \cos\left(\Delta_{(k+1)/2} + \Delta_{(k+1)/2+1} + \cdots + \Delta_{4N-(k+1)/2}\right) + \sum_{j=1}^{(k-1)/2} \cdot \cos\left(\Delta_{k-j+1} + \Delta_{k-j+2} + \cdots + \Delta_{4N-j}\right). \tag{A.2}$$

If we accept that

$$\Delta_{4N-i} = \pi - \Delta_i, \qquad i = 1, 2, \cdots, 2N-1, \tag{A.3}$$

relation (A.2) becomes

$$\sum_{i=1}^{(k-1)/2} \cos\left(\Delta_i + \Delta_{i+1} + \cdots + \Delta_{i+(4N-1)-k}\right) + \cos\left(\frac{4N-1-k}{2}\pi + \Delta_{2N}\right) - \sum_{j=1}^{(k-1)/2} \cos\left(\Delta_{j+(4N-1)-k} + \Delta_{j+(4N-1)-k+1} + \cdots + \Delta_j\right) = \cos\left(\frac{4N-1-k}{2}\pi + \Delta_{2N}\right) \tag{A.4}$$

and it is zero only for

$$\Delta_{2N} = \pi/2. \tag{A.5}$$

When the system of equations (A.1) for odd values of k is satisfied, these equations for even values of k can be expressed in the following form:

$$\sum_{i=1}^{k/2} \cos\left(\Delta_i + \Delta_{i+1} + \cdots + \Delta_{i+(4N-1)-k}\right) + \sum_{i=k/2+1}^{k} \cos\left(\Delta_i + \Delta_{i+1} + \cdots + \Delta_{i+(4N-1)-k}\right). \tag{A.6}$$

With substitution of $j = k - i + 1$ in the second sum, and using solutions (A.3) and (A.5), the system of equations (A.6) can be reexpressed as

$$\sum_{i=1}^{k/2} \cos\left(\Delta_i + \Delta_{i+1} + \cdots + \Delta_{i+(4N-1)-k}\right) + \sum_{j=1}^{k/2} \cos\left(\pi - \Delta_{j+(4N-1)-k} + \cdots + \pi - \Delta_j\right) = 2\sum_{i=1}^{k/2} \cos\left(\Delta_i + \Delta_{i+1} + \cdots + \Delta_{i+(4N-1)-k}\right). \tag{A.7}$$

If we accept, now,

$$\Delta_i = 0, \qquad i = 1, 3, 5, \cdots, 2N-1, \tag{A.8}$$

and using the substitution

$$\Delta_{2i} = \Delta'_i, \qquad i = 1, 2, \cdots, (N - 1), \qquad (A.9)$$

(A.7) finally are given by

$$4\left[\sum_{i=1}^{1} \cos\left(\Delta'_i + \Delta'_{i+1} + \cdots + \Delta'_{i+(N-1)-1}\right)\right],$$

$$l = 1, 2, \cdots, N - 1. \qquad (A.10)$$

We would like to point out that this system of equations is equal to the system given by (A.1), but now only with $(N - 1)$ equations and variables. Thus, it is proven that the solution of the system (A.1) with $(4N - 1)$ equations can be derived from the solution of the same system with $(N - 1)$ equations by use of the substitutions (A.3), (A.5), (A.8), and (A.9), which are summarized in the text as the expression (8).

Acknowledgment

The authors wish to thank the reviewers for their very valuable comments.

References

[1] K. J. Kelley, "Principles of spread-spectrum radar," Ph.D. dissertation, Univ. Southern California, Los Angeles, 1985.

[2] D. R. Wehner, *High Resolution Radar*. Dedham, MA: Artech House, 1987.

[3] A. Farina and G. Galatti, "An overview of current and advanced signal processing techniques for surveillance radars," in *Proc. IEEE Int. Radar Conf.*, Arlington, VA, 1985, pp. 175-183.

[4] B. L. Lewis, F. F. Kretschmer, Jr., and W. W. Shelton, *Aspects of Radar Signal Processing*. Dedham, MA: Artech House, 1986.

[5] ——, "A new class of polyphase pulse compression codes and techniques," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, May 1981.

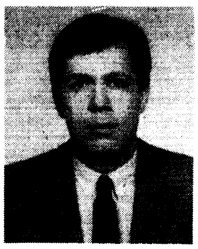
[6] F. F. Kretschmer and B. L. Lewis, "Polyphase pulse compression waveforms," NRL Rep. 8540, Jan. 5, 1982.

[7] B. L. Lewis and F. F. Kretschmer, "Linear frequency modulation derived polyphase pulse compression codes," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-18, Sept. 1982.

[8] F. F. Kretschmer and F. C. Lin, "Huffman coded pulse compression waveforms," NRL Rep. 8894, May 23, 1985.

[9] B. L. Lewis, F. F. Kretschmer, and F. C. Lin, "Effects of bandwidth limitation on polyphase coded pulse compression systems," NRL Rep. 8625, Sept. 20, 1982.

[10] Z. S. Dobrosavljević and M. L. Dukić, "A new method of a polyphase pulse compression code synthesis," in *Proc. ISSSE'89*, Erlangen, W. Germany, Sept. 1989.



Miroslav L. Dukić (M'85) was born in Novi Pazar, Yugoslavia. He received the Dipl.Eng. degree in communication and electronics in 1973, the M.Sc. degree in 1975, and the Doctor of Science degree in 1981, all from the Faculty of Electrical Engineering, University of Belgrade, Belgrade, Yugoslavia.

In 1973 he joined the Faculty of Electrical Engineering, University of Belgrade, where he currently holds the position of Assistant Professor in the Department of Communications. Since 1973 he has taught several courses in communication theory, radar system design, and modern electronic warfare systems. He also participated as a Project Engineer in several communication projects in Yugoslavia, and he has published a number of papers in the field of communications. His research interests lie in the areas of spread-spectrum communications, radar, and electronic warfare systems.

Dr. Dukić is a member of the Yugoslav Committee on Electronics, Telecommunications, Automation, and Nuclear Engineering.

Zoran S. Dobrosavljević (M'90) was born in Velika Plana, Yugoslavia. He received the B.S. degree in 1986 and the M.Sc. degree in 1990, both from the Faculty of Electrical Engineering, University of Belgrade, Belgrade, Yugoslavia.

In 1986 he joined the Faculty of Electrical Engineering, University of Belgrade, where he currently holds the position of Teaching Assistant in the Department of Communications. His current research interests are in the areas of digital signal processing, radar, and electronic warfare systems.

Capacity and Cutoff Rate of Coded FH/MFSK Communications with Imperfect Side Information Generators

HYUCK M. KWON, MEMBER, IEEE

Abstract—**In a coded frequency-hopped *M*-ary frequency shift-keying (FH/MFSK) communication system, it is known that if perfect side information about the jamming state is available to the decoder, substantial improvement in performance is possible. However, thermal noise, which is present in practical communication systems, can corrupt the jamming state information (JSI). In this paper, when JSI is imperfect due to thermal noise, we calculate the capacities and cutoff rates of the channels as a function of the signal-to-jamming-noise ratio for memoryless, noncoherent FH/MFSK systems under partial-band noise jamming (PBNJ). We consider both soft and hard-decision metrics with perfect, imperfect, and no JSI. Also, we introduce three imperfect JSI generators. The first imperfect JSI generator uses the maximum *a posteriori* (MAP) decision rule based on the energy from an FH tone frequency which is near the *M*-signaling FH tone frequencies. The second decision rule utilizes the MAP rule, but it is based on the total energy received in the *M*-signaling FH tone frequencies. The third generator has the same decision statistics as the second generator, but its decision rule is an easily implementable suboptimum rule. If hard decisions are made and code rates are high, e.g., ≥0.7, then the differences between the imperfect JSI generators and perfect JSI generator can be larger than 1 dB in the signal-to-jamming-noise ratio required to achieve the given capacity or cutoff rate, even though the thermal noise is quite small, e.g., a 25 dB signal-to-thermal noise ratio. If soft decisions are made, then the differences between the imperfect and perfect JSI cases are negligible.**

I. Introduction

IT IS known that if perfect jamming state information (JSI) about the partial-band noise jamming (PBNJ) [to be defined in Section II] is available to the decoder of coded frequency-hopped *M*-ary frequency shift keying (FH/MFSK) systems, then substantial improvement in performance is possible, e.g., more than 5 dB improvement in signal-to-jamming-noise ratio required to achieve a 0.65 cutoff rate for a coded FH/binary FSK system with hard decisions and a 25 dB signal-to-thermal noise ratio [1]. However, thermal noise, which is present in practical communication systems, can corrupt the JSI. Thus, a JSI generator in practical environments cannot be perfect. For example, the probabilities of detection and false alarm are 0.81022 and 0.00963, respectively, for the JSI generator II (defined later) when the signal-to-thermal noise ratio is 25 dB and the cutoff rate of a coded FH/BFSK system with hard decisions is 0.65 under PBNJ with a worst case jamming fraction. We need 2 dB more in the signal-to-jamming-noise required to achieve a 0.65 cutoff rate for this JSI generator, compared to that for the perfect JSI generator, even though the thermal noise is quite small.

In this paper, when JSI is imperfect due to the presence of thermal noise, we calculate the capacities and cutoff rates of the channels as a function of the signal-to-jamming-noise ratio for the memoryless, noncoherent FM/MFSK systems under PBNJ. We consider both soft- and hard-decision metrics with perfect, imperfect, and no JSI. Also, we introduce three imperfect JSI generators. The first imperfect JSI generator uses the maximum *a posteriori* (MAP) decision rule based on the energy from an FH tone frequency which is near to the *M*-signaling FH tone frequencies. The second decision rule utilizes the MAP rule, but it is based on the total energy received in the *M*-signaling FH tone frequencies. The third generator has the same decision statistics as the second generator, but its decision rule is an easily implementable suboptimum rule. There are studies that describe other approaches to the generation of JSI in FH communications, in which test symbols are inserted into a packet which consists of data and test symbols [2]–[5].

The paper is organized as follows. Section II describes the channel model, and Section III presents three imperfect JSI generators. Section IV derives the channel capacities of memoryless FH/MFSK channels with imperfect JSI, and two capacity theorems about perfect, imperfect and no JSI cases are stated. (The proofs are shown in the Appendix.) Section V gives the corresponding analyses when the cutoff rate is used as the performance criterion. Finally, Section VI gives numerical results and discussions.

II. Channel Model

The block diagram of the communication system we consider is shown in Fig. 1. Existing papers regarding coding and spread spectrum communication systems, e.g., [6, vol. I, ch. 4, and vol. II, ch. 2], [7]–[9], describe

Manuscript received February 15, 1989; revised October 20, 1989. This paper was presented in part at the IEEE International Symposium on Information Theory, Kobe, Japan, June 1988, and the IEEE Military Communications Conference, Boston, MA, October 1989.

The author is with the Department of Electrical Engineering and Computer Science, University of Wisconsin, Milwaukee, WI 53211, on leave at the Lockheed Engineering and Sciences Company, Houston, TX 77058.

IEEE Log Number 9035803.

0733-8716/90/0600-0750$01.00 © 1990 IEEE

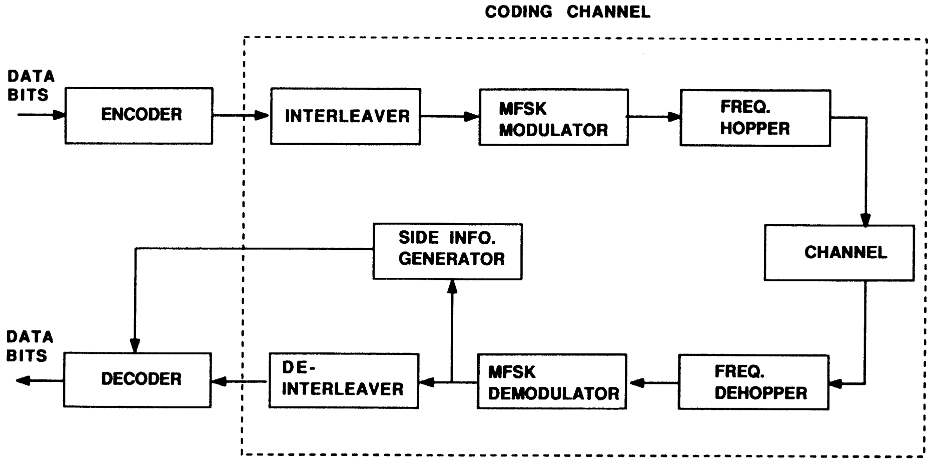


Fig. 1. Communication system block diagram.

each block well. As shown in Fig. 1, the code symbols are transmitted over the coding channel outlined by dotted lines. Let X denote an M-ary coding channel input symbol which takes values in the alphabet $\{1, 2, \cdots, M\}$. Assume that the interleaver and deinterleaver make the coding channel memoryless. The system uses the FH/MFSK signaling, and one MFSK symbol is transmitted per each hopping time interval. Let B denote the bandwidth used by an FH tone frequency, which is the inverse of a hopping time interval. The frequency-hopped symbols are transmitted over a waveform channel which is affected by thermal noise and PBNJ. Assume that thermal noise is a white Gaussian process with two-sided power spectral density $N_0/2$ and that PBNJ is a bandlimited white Gaussian process. The receiver uses a synchronized version of the spreading sequence to dehop the received signal. The receiver contains noncoherent square-law envelope detectors. [Refer to Fig. 2(b).] In addition, let $Y = (Y_1, Y_2, \cdots, Y_M)$ denote a coding channel output observation vector, based on energy detections from M-ary signaling channels after frequency dehopping. Let the event $\{S = 1\}$ mean that a transmitted code symbol is truly jammed, the event $\{S = 0\}$ truly not jammed, $\hat{S}$ an imperfect JSI from a JSI generator, P_D the conditional probability that $\hat{S} = 1$ given $S = 1$ which is called the probability of detection, P_{FA} the conditional probability that $\hat{S} = 1$ given $S = 0$ which is called the probability of false alarm, and ρ the fraction of the full spread-spectrum bandwidth W jammed by the PBNJ. Then the probability distribution of S is $P\{S = 1\} = \rho$ and $P\{S = 0\} = 1 - \rho$. The probabilities of detection and false alarm depend on ρ, $N_0/2$ and the uniform noise jamming power spectral density $N_J/2 = J/2W$ where J is the total jamming power received. We assume that N_0 and N_J are constant parameters, and that the PBNJ minimizes the channel capacities or cutoff rates of coded FH/MFSK systems by varying ρ for a given jamming power J. Then, the joint probabilities of S and $\hat{S}$ are $P(S = 0, \hat{S} = 0) = (1 - \rho)(1 - P_{FA}(\rho))$, $P(S = 0, \hat{S} = 1) = (1 - \rho)P_{FA}(\rho)$, $P(S = 1, \hat{S} = 1) = \rho P_D(\rho)$, and $P(S = 1, \hat{S} = 0) = \rho(1 - P_D(\rho))$.

The conditional densities of Y_i, given X and $\hat{S}$, are then

$$\begin{aligned} p(y_i|X = j, \hat{S} = 1) &= \alpha p(y_i|X = j, S = 1) \\ &\quad + (1 - \alpha)p(y_i|X = j, S = 0) \end{aligned} \tag{1}$$

$$\begin{aligned} p(y_i|X = j, \hat{S} = 0) &= \beta p(y_i|X = j, S = 1) \\ &\quad + (1 - \beta)p(y_i|X = j, S = 0), \\ &\qquad i = 1, \cdots, M, j = 1, \cdots, M \end{aligned} \tag{2}$$

where

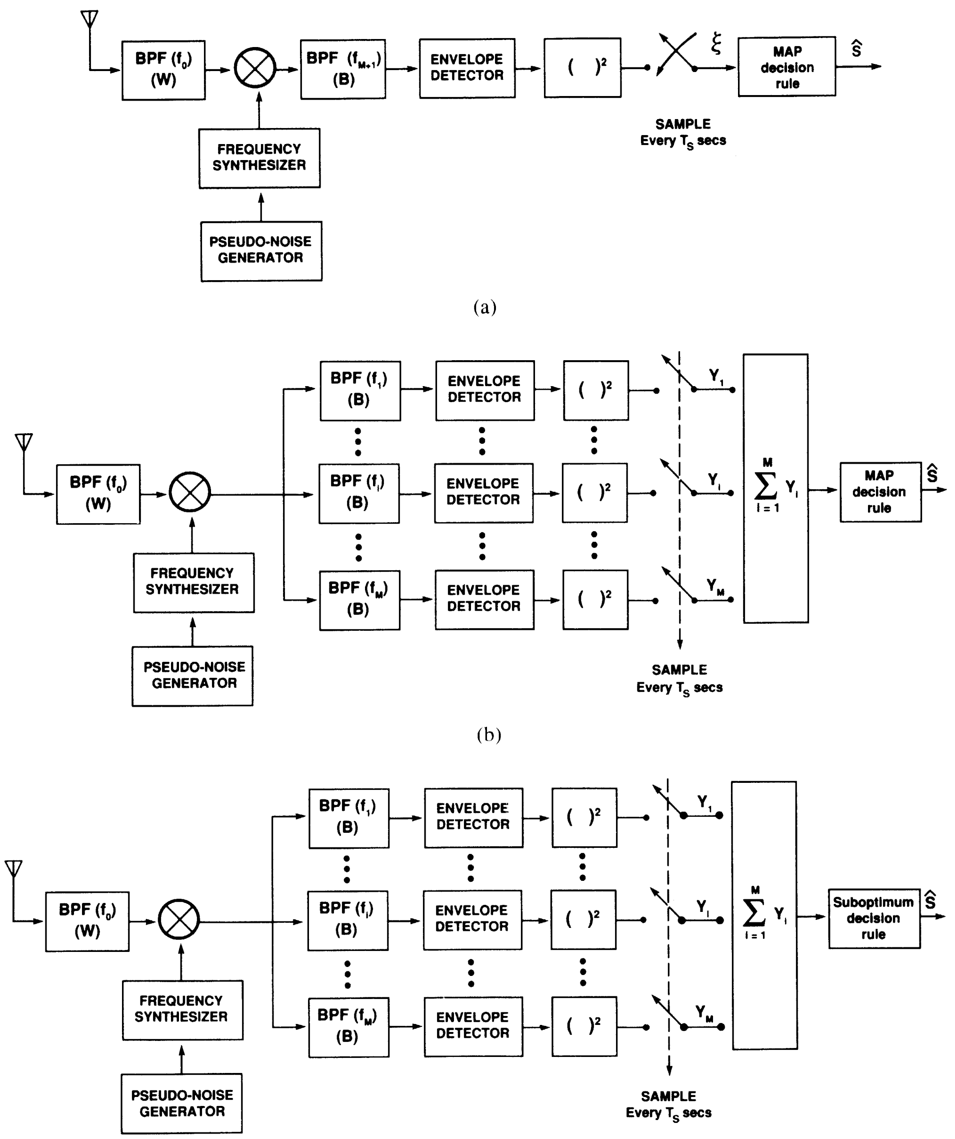
$$\alpha = \frac{\rho P_D(\rho)}{(1 - \rho)P_{FA}(\rho) + \rho P_D(\rho)},$$
$$\beta = \frac{\rho(1 - P_D(\rho))}{(1 - \rho)(1 - P_{FA}(\rho)) + \rho(1 - P_D(\rho))}. \tag{3}$$

Note that if $P_D(\rho) = 1$ and $P_{FA}(\rho) = 0$ for any ρ (i.e., perfect JSI is available), then $p(y_i|X = j, \hat{S} = k) = p(y_i|X = j, S = k)$. If $P_D(\rho) = 0.5$ and $P_{FA}(\rho) = 0.5$ for any ρ (i.e., no side information is available), then $p(y_i|X = j, \hat{S} = k) = p(y_i|X = j)$ for $k = 0$ or 1. If i is equal to j (matched channels), then $p(y_i|X = j, S = k)$ is a noncentral chi-square density of two degrees of freedom with the noncentral parameter $2E_sB/\sigma_k^2$ [12, eq. (4)-(77)], [1, eq. (17)] where E_s is the received code symbol energy, $\sigma_1^2 = (N_0 + N_J/\rho)B$, and $\sigma_0^2 = N_0B$. If i is not equal to j (unmatched channels), then $p(y_i|X = j, S = k)$ is a central chi-square density of two degrees of freedom.

III. Imperfect JSI Generators

A. JSI Generator I

Assume that, among $M + 1$ FH tone frequencies, one FH tone frequency is dedicated to generating JSI, and M FH tone frequencies are assigned to transmit an M-ary



(a)

(b)

(c)

Fig. 2. (a) Jamming state information generator I. (b) Jamming state information generator II. (c) Jamming state information generator III.

coding channel symbol. Assume also that the dedicated FH tone frequency for JSI generation and the M signaling FH tone frequencies are hopped with the same frequency-hopping pattern. In addition, we can reasonably assume that if PBNJ jams any FH tone frequency in the $M + 1$ contiguous FH tone frequencies, then all $M + 1$ FH tone frequencies are jammed because an $M + 1$ contiguous frequency-hopping band is very small compared to W. Other authors [1], [6]-[10] have made similar assumptions. The JSI generator I (as well as JSI generators II and III discussed later) is not extremely vulnerable to the PBNJ because the frequency-hopping pattern is not available to the jammer.

As shown in Fig. 2(a), after the frequency dehopper, JSI generator I measures energy received in the dedicated FH tone frequency. The measured energy is due only to thermal noise if the channel was not jammed or due to thermal noise plus jamming noise if the channel was jammed. If the measured energy, denoted by ξ, is larger than or equal to the threshold value, then $\hat{S} = 1$; otherwise, $\hat{S} = 0$. We assume that all JSI generators in this paper have knowledge of the jamming fraction ρ, and that JSI generator I uses the MAP rule to determine the threshold value. Using $\sigma_0^2 = N_0 B$ and $\sigma_1^2 = (N_0 + N_J/\rho)B$, the optimal threshold value is $2\{\lambda_0 + \ln[\sigma_1^2/\sigma_0^2]\}\sigma_0^2\sigma_1^2/(\sigma_1^2 - \sigma_0^2)$ where $\lambda_0 = \ln((1-\rho)/\rho)$. (Refer to the MAP rule in [12].) The probabilities of detection and false alarm of JSI generator I are

$$P_D(\rho) = \exp\left[-(\lambda_0 + \ln(\sigma_1^2/\sigma_0^2))/(\sigma_1^2/\sigma_0^2 - 1)\right] \tag{4a}$$

$$P_{FA}(\rho) = \exp\left[-(\lambda_0 + \ln(\sigma_1^2/\sigma_0^2))/(1 - \sigma_0^2/\sigma_1^2)\right]. \tag{4b}$$

For the discussion of numerical results in Section VI, we express our $P_D(\rho)$ and $P_{FA}(\rho)$ in terms of the symbol energy-to-jamming-noise ratio E_s/N_J and symbol energy-to-thermal noise ratio E_s/N_0 using $\sigma_1^2/\sigma_0^2 = 1 + \{E_s/N_0\}/\{\rho E_s/N_J\}$. For our JSI generator I, the $P_D(\rho) \geq P_{FA}(\rho)$ since $\sigma_1^2 \geq \sigma_0^2$. When there is no thermal noise ($\sigma_0^2 = 0$), our JSI generator I becomes the perfect JSI generator with $P_D(\rho) = 1$ and $P_{FA}(\rho) = 0$ for any ρ between 0 and 1. Note that $P_D(\rho)$ and $P_{FA}(\rho)$ of JSI generator I do not depend on M. In addition, note that the threshold of JSI generator I is easily implementable. It is, however, not bandwidth-efficient, especially for small M. For example, for $M = 2$, $1/3$ of the total bandwidth is dedicated to obtaining side information. A more bandwidth-efficient method of obtaining JSI from energy measurements would be to base the decision on the total energy received in the M signaling FH tone frequencies instead of setting aside a dedicated tone frequency.

B. JSI Generator II

Fig. 2(b) illustrates JSI generator II, which is the same as the noncoherent FH/MFSK envelope-square demodulator, except for the energy summing device and the MAP decision rule device. Still, we assume a contiguous M-ary symbol per hop as in the previous JSI generator I. Let y denote the total energy (from M signaling FH tone frequencies) divided by the total power of the bandpassed thermal noise $\sigma_0^2 = N_0 B$. In addition, let $p_1(y)$ denote the conditional probability density of y given the channel jammed (i.e., block of all M contiguous frequency tone slots is jammed), and $p_0(y)$ the conditional probability density of y given the channel unjammed (i.e., block of all M contiguous frequency tone slots is unjammed). Then, for $y \geq 0$,

$$p_1(y) = \frac{\sigma_0^2}{2\sigma_1^2}\left(\frac{y}{a_0^2}\right)^{(M-1)/2} \cdot \exp\left(-\frac{a_1^2}{2} - \frac{\sigma_0^2}{\sigma_1^2}\frac{y}{2}\right) I_{M-1}\left(\sqrt{\frac{\sigma_0^2}{\sigma_1^2}a_1^2 y}\right) \tag{5a}$$

$$p_0(y) = \frac{1}{2}\left(\frac{y}{a_0^2}\right)^{(M-1)/2} \exp\left(-\frac{a_0^2}{2} - \frac{y}{2}\right) I_{M-1}(\sqrt{a_0^2 y}) \tag{5b}$$

where $a_1^2 = 2E_s/(N_0 + N_J/\rho)$, $a_0^2 = 2E_s/N_0$, and $I_{M-1}(x)$ is the modified Bessel function of order $M - 1$. Let $R(y)$ denote the optimum decision region based on the MAP rule for the jamming signal detection, which is $R(y) = \{y | p_1(y)/p_0(y) \geq (1 - \rho)/\rho\}$. If the total energy y belongs to $R(y)$, then $\hat{S} = 1$; otherwise, $\hat{S} = 0$. From the numerical analysis, we find that there are two types of decision regions, which depend on ρ, E_s/N_J, E_s/N_0, and M. One is $R(y) = \{y \geq 0 | y \leq y_1^* \text{ or } y \geq y_2^*\}$. The other is $R(y) = \{y \geq 0 | y \geq y_2^*\}$, which is a special case of the first type with $y_1^* = 0$. The probabilities of detection and false alarm of JSI generator II are then

$$P_D(\rho) = 1 - Q_M(a_1, b_{11}) + Q_M(a_1, b_{12}) \tag{6a}$$

$$P_{FA}(\rho) = 1 - Q_M(a_0, b_{01}) + Q_M(a_0, b_{02}) \tag{6b}$$

where $b_{11}^2 = y_1^*\sigma_0^2/\sigma_1^2$, $b_{12}^2 = y_2^*\sigma_0^2/\sigma_1^2$, $b_{01}^2 = y_1^*$, $b_{02}^2 = y_2^*$, and $Q_M(x, y)$ is the Marcum Q function. Note that the difference between the conditional average of y given the channel jammed and the conditional average of y given the channel unjammed is $2M(\sigma_1^2 - \sigma_0^2)$. As we increase M, the difference also increases. This implies that $P_D(\rho)$ increases and $P_{FA}(\rho)$ decreases as we increase M for given N_0, N_J, ρ, and average signal power. Hence, the performance of JSI generator II becomes better as M increases. In addition, note that JSI generator II is bandwidth-efficient. It is, however, difficult to implement JSI generator II because we need the modified Bessel function of order $M - 1$.

C. JSI Generator III

Fig. 2(c) shows JSI generator III, which is the same as JSI generator II except for the decision device. JSI generator III uses a suboptimum decision rule: $\hat{S} = 1$ if the total energy y is larger than the threshold Th; otherwise, $\hat{S} = 0$. The primary difference between JSI generator III and JSI generator II is that the threshold of generator III is chosen to be half the sum of the conditional average of y given the channel jammed, plus the conditional average of y given the channel unjammed, while the threshold of generator II is determined from the MAP rule. Then, the threshold of generator III Th is equal to two times average signal power plus $M(\sigma_1^2 + \sigma_0^2)$. The probabilities of detection and false alarm of JSI generator III are then

$$P_D(\rho) = Q_M(a_1, b_1) \tag{7a}$$

$$P_{FA}(\rho) = Q_M(a_0, b_0) \tag{7b}$$

where $b_1^2 = Th/\sigma_1^2$ and $b_0^2 = Th/\sigma_0^2$. Note that, as we increase M, the performance of JSI generator III also improves for the same reasons mentioned in the case of JSI generator II. In addition, note that JSI generator III is bandwidth-efficient and easily implementable if it has knowledge of the jamming fraction ρ.

IV. Channel Capacity

In this section, we compute the channel capacity for memoryless, noncoherent FH/MFSK channels with PBNJ and thermal noise when JSI is imperfect. Since the coding channel is symmetric, the best probability distribution of X to maximize the mutual information between X and Y, $I(X; Y, \hat{S})$, is the uniform distribution on the alphabet [11, p. 64]. The channel capacity is

$$C = \min_{0 < \rho \leq 1} \Big\{ \big(\rho P_D(\rho) + (1 - \rho)P_{FA}(\rho)\big) I(X; Y | \hat{S} = 1) + \big(\rho(1 - P_D(\rho)) + (1 - \rho)(1 - P_{FA}(\rho))\big) \cdot I(X; Y | \hat{S} = 0)\Big\}. \tag{8}$$

A. Capacity for Soft Decisions

If soft decisions are made at the demodulator, then the conditional mutual information $I(X;\ Y|\hat{S} = k)$ in (8) is $I(X;\ Y|\hat{S} = k) = \int_y p(y|X = x, \hat{S} = k) \log_M \{p(y|X = x, \hat{S} = k)/p(y|\hat{S} = k)\}\, dy$. In the above integral, we can calculate the integrand, which is a function of the conditional probability densities, using (1)-(3). The above integral is, however, an M-dimensional integral. For $M > 2$, numerical evaluation becomes a long and complex computation. We do not include results on the capacity for soft decisions in this paper.

B. Capacity for Hard Decisions

If hard decisions are made at the demodulator, then the probability that the receiver makes an error on a symbol during a hop time interval is given as

$$p_1 = \frac{1}{M}\sum_{j=2}^{M} (-1)^j \binom{M}{j} \cdot \exp\left\{-\frac{E_s}{N_0 + N_J/\rho}(1 - 1/j)\right\} \quad \text{for } S = 1$$

and

$$p_0 = \frac{1}{M}\sum_{j=2}^{M} (-1)^j \binom{M}{j} \exp\left\{-\frac{E_s}{N_0}(1 - 1/j)\right\} \quad \text{for } S = 0$$

[1, eq. (19)]. The coding channel with hard decisions and perfect JSI is an M-ary symmetric channel with crossover probabilities p_0 and p_1. Also, the coding channel with hard decisions and imperfect JSI is an M-ary symmetric channel with crossover probabilities $\bar{p}_1$ and $\bar{p}_0$ where

$$\bar{p}_1 = \alpha p_1 + (1 - \alpha)p_0 \quad \text{if } \hat{S} = 1, \quad \bar{p}_0 = \beta p_1 + (1 - \beta)p_0 \quad \text{if } \hat{S} = 0. \tag{9}$$

We know that the conditional mutual information for the symmetric channel is given by

$$I(X;\ Y)|\hat{S} = 1) = 1 + (1 - \bar{p}_1)\log_M(1 - \bar{p}_1) + \bar{p}_1 \log_M (\bar{p}_1/(M - 1)) \tag{10a}$$

$$I(X;\ Y)|\hat{S} = 0) = 1 + (1 - \bar{p}_0)\log_M(1 - \bar{p}_0) + \bar{p}_0 \log_M (\bar{p}_0/(M - 1)). \tag{10b}$$

We can calculate the capacity for the channel with hard decisions and imperfect JSI by substitution of (10) into (8) in Section VI.

C. Inequalities Between Capacities

With imperfect JSI, the capacity is always less than or equal to the capacity with perfect JSI and always greater than or equal to the capacity without JSI, as summarized in the following theorems.

Theorem 1: Assume that soft decisions are made. Let $C_{\text{soft}}^{\text{without}}$ and $C_{\text{soft}}^{\text{perfect}}$ represent the channel capacity without and with perfect JSI, respectively, and let $C_{\text{soft}}^{\text{imperfect}}$ represent the channel capacity with any imperfect JSI. Then

$$C_{\text{soft}}^{\text{without}} \le C_{\text{soft}}^{\text{imperfect}} \le C_{\text{soft}}^{\text{perfect}}. \tag{11}$$

Proof: See the Appendix.

Theorem 2: Assume that hard decisions are made. Let $C_{\text{hard}}^{\text{without}}$ and $C_{\text{hard}}^{\text{perfect}}$ represent the channel capacity without and with perfect JSI, respectively, and let $C_{\text{hard}}^{\text{imperfect}}$ represent the channel capacity with any imperfect JSI. Then

$$C_{\text{hard}}^{\text{without}} \le C_{\text{hard}}^{\text{imperfect}} \le C_{\text{hard}}^{\text{perfect}}. \tag{12}$$

Proof: See the Appendix.

V. Cutoff Rate

In this section, we compute the computational cutoff rate for memoryless, noncoherent channels with PBNJ and thermal noise when JSI is imperfect. The cutoff rate is defined as

$$R_0 = 1 - \log_M \{1 + (M - 1)D\} \frac{\text{information symbols}}{M\text{-ary channel symbol}} \tag{13}$$

where parameter D is a worst case (with respect to ρ) Chernoff bound on the probability that the metric value corresponding to the transmitted symbol x is smaller than the metric value corresponding to a nontransmitted symbol $\hat{x}$ [6]-[9]:

$$\begin{aligned} D &= \max_{0<\rho\le 1} D(\rho) = \max_{0<\rho\le 1} \min_{\lambda\ge 0} D(\rho, \lambda) \\ &= \max_{0\le\rho\le 1} \min_{\lambda\ge 0} E\Big\{\exp\big(\lambda[m(y, \hat{x};\ \hat{S}) \\ &\quad - m(y, x;\ \hat{S})]\big)\Big|\hat{x} \text{ not equal to } x\Big\}. \end{aligned} \tag{14}$$

In (14), "E" represents the expectation over the observed random vector $\boldsymbol{Y}$ and $\hat{S}$. The $m(y, x;\ \hat{S})$ is an additive metric, and λ is the Chernoff bound parameter [6]. Given the channel output sequences, the decoder uses the metric to decide the sequence of the transmitted symbols.

The general relation of the parameter D to the coded bit error probability is $P_b \le G(D)$ where $G(\cdot)$ is a polynomial function determined solely by the specific code, whereas the parameter D depends only on the coding channel and the decoder metric [6, vol. I, pp. 194, 199].

A. Cutoff Rate for Soft Decision Metrics

When soft decisions are made, the metric we are considering is $m(y, x;\ \hat{S}) = c_{\hat{S}} y_x$, which is a weighted version of the observed energy y_x corresponding to input symbol x. In Section VI, we numerically optimize the weighting factors $c_{\hat{S}}$ (which depend on the imperfect JSI) and λ. The parameter $D(\rho, \lambda)$ in (14) can be calculated as below:

$$\begin{aligned} D(\rho, \lambda) = {} & (1-\rho)(1-P_{FA}(\rho)) \frac{1}{1-(\lambda c_0)^2} \cdot \exp\left[-\frac{E_s}{N_0} \frac{\lambda c_0}{(1+\lambda c_0)} \right] \\ & + (1-\rho)P_{FA}(\rho) \frac{1}{1-(\lambda c_1)^2} \cdot \exp\left[-\frac{E_s}{N_0} \frac{\lambda c_1}{(1+\lambda c_1)} \right] \\ & + \rho(1-P_D(\rho)) \frac{1}{1-(\lambda c_0)^2} \cdot \exp\left[-\frac{E_s}{N_0 + N_J/\rho} \frac{\lambda c_0}{(1+\lambda c_0)} \right] \\ & + \rho P_D(\rho) \frac{1}{1-(\lambda c_1)^2} \cdot \exp\left[-\frac{E_s}{N_0 + N_J/\rho} \frac{\lambda c_1}{(1+\lambda c_1)} \right] \qquad (15) \end{aligned}$$

for $0 < \lambda c_0 < 1$ and $0 < \lambda c_1 < 1$. The $D(\rho, \lambda)$ in (15) becomes $D(\rho, \lambda)$ in [6, vol. I, eq. (4.91)] for the special case of perfect JSI, and the $D(\rho, \lambda)$ in (15) becomes $D(\rho, \lambda)$ in [6, vol. I, eq. (4.93)] for the special case of no JSI.

B. Cutoff Rate for Hard-Decision Metrics

When hard decisions are made, the metric we are considering is $m(y, x; \hat{S}) = c_{\hat{S}}$ if $y_x > y_{\hat{x}}$ for all $\hat{x}$ not equal to x and 0 otherwise. The metric chooses the index of the largest component as the transmitted symbol and applies a weight depending on the imperfect JSI. Let $A_i = p_i/(M-1)$ be a branch crossover probability in the M-ary symmetric channel with $S = i$, $i = 0, 1$. Then, after the minimizations over λ and $c_{\hat{S}}$, parameter D in (14) can be written as

$$D = \max_{0<\rho\le 1} D(\rho), \qquad D(\rho) = 2\sqrt{\alpha_0\beta_0} + 2\sqrt{\alpha_1\beta_1} + \gamma \qquad (16)$$

where

$$\begin{aligned} \alpha_0 &= (1-\rho)(1-P_{FA}(\rho))(1-(M-1)A_0) \\ &\quad + \rho(1-P_D(\rho))(1-(M-1)A_1) \\ \alpha_1 &= (1-\rho)P_{FA}(\rho)(1-(M-1)A_0) \\ &\quad + \rho P_D(\rho)(1-(M-1)A_1) \\ \beta_0 &= (1-\rho)(1-P_{FA}(\rho))A_0 + \rho(1-P_D(\rho))A_1, \\ \beta_1 &= (1-\rho)P_{FA}(\rho)A_0 + \rho P_D(\rho)A_1 \\ \gamma &= (1-\rho)(M-2)A_0 + \rho(M-1)A_1. \end{aligned}$$

If perfect JSI is available, then $D(\rho)$ in (16) becomes a known formula [6, vol. I, eq. (4.81)], $D(\rho) = \rho[2\sqrt{A_1(1-(M-1)A_1)} + (M-2)A_1] + (1-\rho)[2\sqrt{A_0(1-(M-1)A_0)} + (M-2)A_0]$, and if no JSI is available,

$$\begin{aligned} D(\rho) = {} & 2\sqrt{\{\rho A_1 + (1-\rho)A_0\}\{\rho(1-(M-1)A_1) + (1-\rho)(1-(M-1)A_0)\}} \\ & + \rho(M-2)A_1 + (1-\rho)(M-2)A_0. \end{aligned}$$

From a convexity argument, we can state the following inequalities about cutoff rates for hard decisions:

$$R_{0,\text{hard}}^{\text{wihtout}} \le R_{0,\text{hard}}^{\text{imperfect}} \le R_{0,\text{hard}}^{\text{perfect}}. \qquad (17)$$

VI. Numerical Results and Discussion

In this section, we present numerical results for the imperfect JSI generators discussed in Section III. We compare our results for these imperfect JSI cases to the results for both the perfect JSI case and no JSI case.

For channel capacity C in (8)–(10), we can determine the minimum value of the bit energy-to-jamming-noise ratio E_b/N_J for reliable communication as a function of the code rate r from the relation as $E_b/N_J > C^{-1}(r)/(r \log_2 M)$ where r is the code rate in M-ary units [1]. In Fig. 3, the E_b/N_J required to achieve the channel capacity is shown for BFSK with hard decisions when $E_s/N_0 = 15$ dB. The general behavior of the required E_b/N_J to achieve the channel capacity for other MFSK and E_s/N_0 is similar to that of the cutoff rate (to be shown later). Hence, only cutoff results will be presented below. If we use the cutoff rate as a measure of the channel performance, the analysis yields $E_b/N_J > R_0^{-1}(r \log_2 M)$ as the necessary condition for practical and reliable communications [6]–[9].

A. Numerical Results for Soft-Decision Metrics Based on Cutoff Rate

In Fig. 4, the E_b/N_J required to achieve the cutoff rate is shown for binary and 8-ary FSK with soft decisions. For the case of soft decisions and no JSI, we know that the performance is poor [6]. However, notice that for soft decisions and imperfect JSI, results are very nearly as good as those for the perfect JSI case. This is explained as follows, and is a key result of this paper. Let μ denote λc_0 and let ν denote λc_1 in (15). To minimize $D(\rho, \lambda)$ in (15) with respect to both μ and ν, we can minimize the sum of the first term plus the third term with respect to μ, and we can separately minimize the sum of the second term plus the fourth term with respect to ν. Suppose that $\mu_{\min}$ and $\nu_{\min}$ minimize $D(\rho, \lambda)$ for given $P_{FA}(\rho)$, $P_D(\rho)$, E_s/N_0, and E_s/N_J. From the numerical results, we found that for any pair $(P_{FA}(\rho), P_D(\rho))$, the value of exp

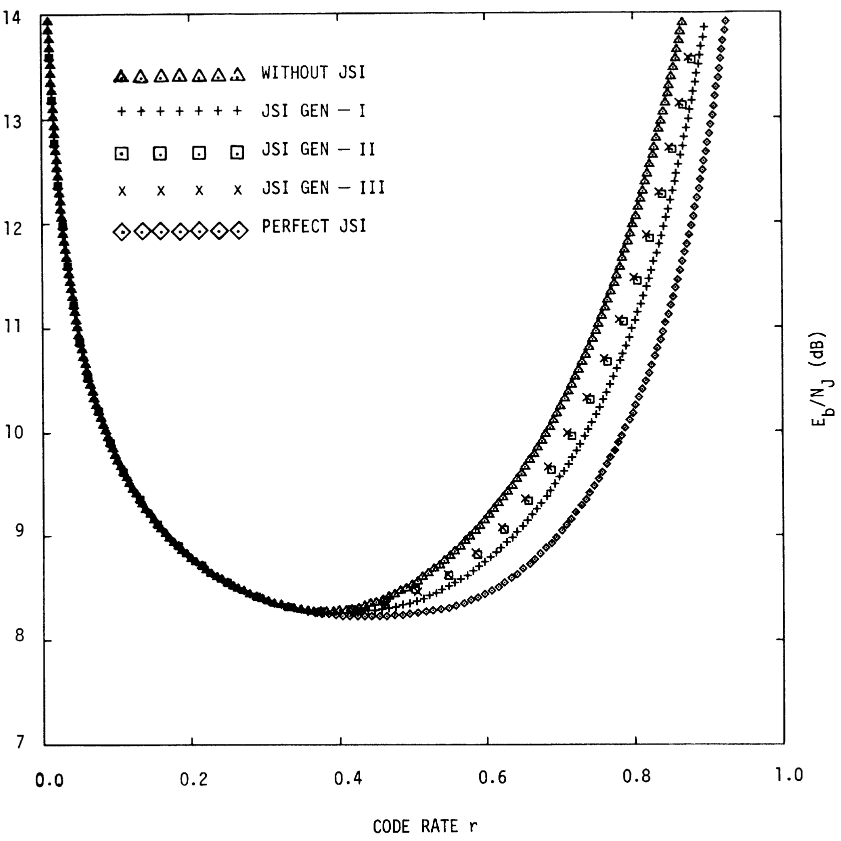


Fig. 3. E_b/N_J needed to achieve channel capacity for frequency-hopped binary FSK with hard-decision metrics when E_s/N_0 = 15 dB.

$[-E_s/N_0\,\mu/(1+\mu)]/(1-\mu^2)$ at $\mu=\mu_{\min}$ is almost equal to the value of exp $[-E_s/N_0\,\nu/(1+\nu)]/(1-\nu^2)$ at $\nu=\nu_{\min}$, and the value of exp $[-E_s/(N_0+N_J/\rho)\mu(1+\mu)]/(1-\mu^2)$ at $\mu=\mu_{\min}$ is almost equal to the value of exp $[-E_s/(N_0+N_J/\rho)\nu/(1+\nu)]/(1-\nu^2)$ at $\nu=\nu_{\min}$. Hence, the sum of the first and the second terms in (15) is almost independent of $P_{FA}(\rho)$, and the sum of the third and the fourth terms in (15) is almost independent of $P_D(\rho)$. This implies that any imperfect JSI generator can achieve almost the same performance as the perfect JSI generator.

B. Numerical Results for Hard-Decision Metrics Based on Cutoff Rate

In Fig. 5, when hard decisions are made for BFSK, the E_b/N_J needed to achieve the cutoff rate are shown for two different values of E_s/N_0. The results for the 8-ary FSK case are shown in Fig. 6. Three main results are observed from Figs. 5 and 6.

First, we observe in Fig. 5 that as thermal noise power becomes weaker, the required E_b/N_J versus r for the imperfect JSI cases approaches that for perfect JSI cases. Second, Figs. 5 and 6 show that the difference between the perfect JSI generator and imperfect JSI generators can be significant. For example, in Fig. 5(b), even if thermal noise is very small (e.g., E_s/N_0 = 25 dB), the imperfect JSI generator I is 1.11 dB worse than the perfect JSI generator, the JSI generator II is 2.04 dB worse, and the JSI generator III is 3.03 dB worse at a code rate of 0.7. (Under the conditions in the above example, the optimum jamming fractions ρ^* are 0.3234 for JSI generator I, 0.2224 for JSI generator II, and 0.1464 for JSI generator III. The P_D and P_{FA} at those optimum jamming fractions are 0.9408 and 0.5223×10^{-2} for JSI generator I, 0.8016 and 0.6849×10^{-2} for JSI generator II, and 0.5852 and 0.4241×10^{-5} for JSI generator III, respectively.)

Third, in Fig. 5 for BFSK, we observe that JSI generator I is the best among the three JSI generators, JSI generator II is the second best, and JSI generator III is the worst in the sense of the bit energy-to-jamming-noise ratio required to achieve the cutoff rate. However, in Fig. 6, we observe that if M is large, then JSI generators II and III can be better than JSI generator I (see Fig. 6(a) for a code rate larger than 0.6). This is because JSI generators II and III, which are based on the total energy in M signaling FH tones, improve as M increases, as we expect from (6) and (7), while generator I does not.

In Table I (M = 2 case) and Table II (M = 8 case), the E_b/N_J needed to achieve the cutoff rate is listed for eight different values of E_s/N_0 and for four different values of cutoff rate. An interesting observation in Table II is that JSI generator III (a suboptimum rule) can be better than JSI generator II (a MAP rule) in the E_b/N_J required to achieve the cutoff rate for 8FSK with hard decisions when E_s/N_0 = 12.8 dB (column 1) and 14.8 dB (column 2). We know that a detector based on a MAP rule is an

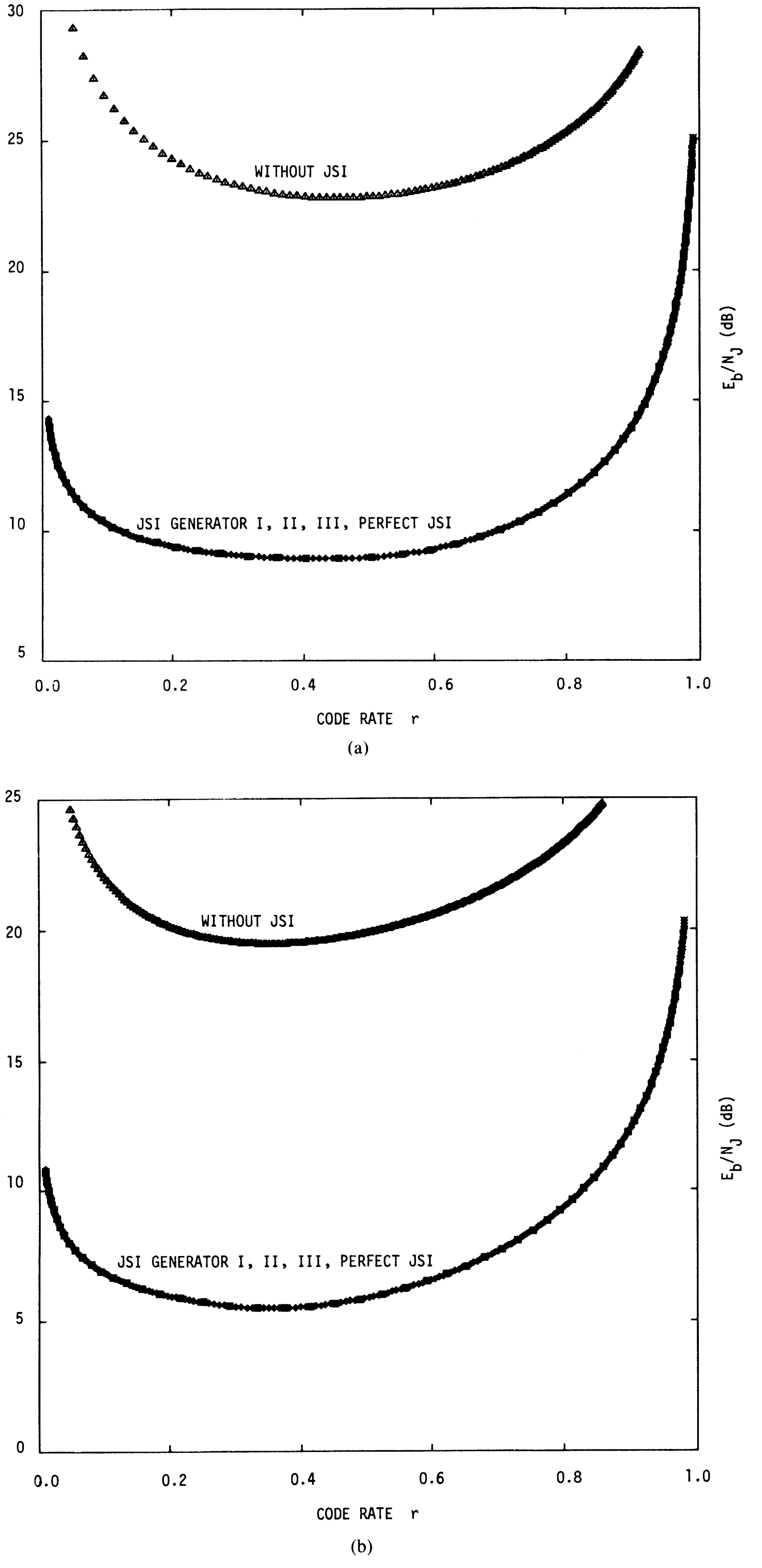
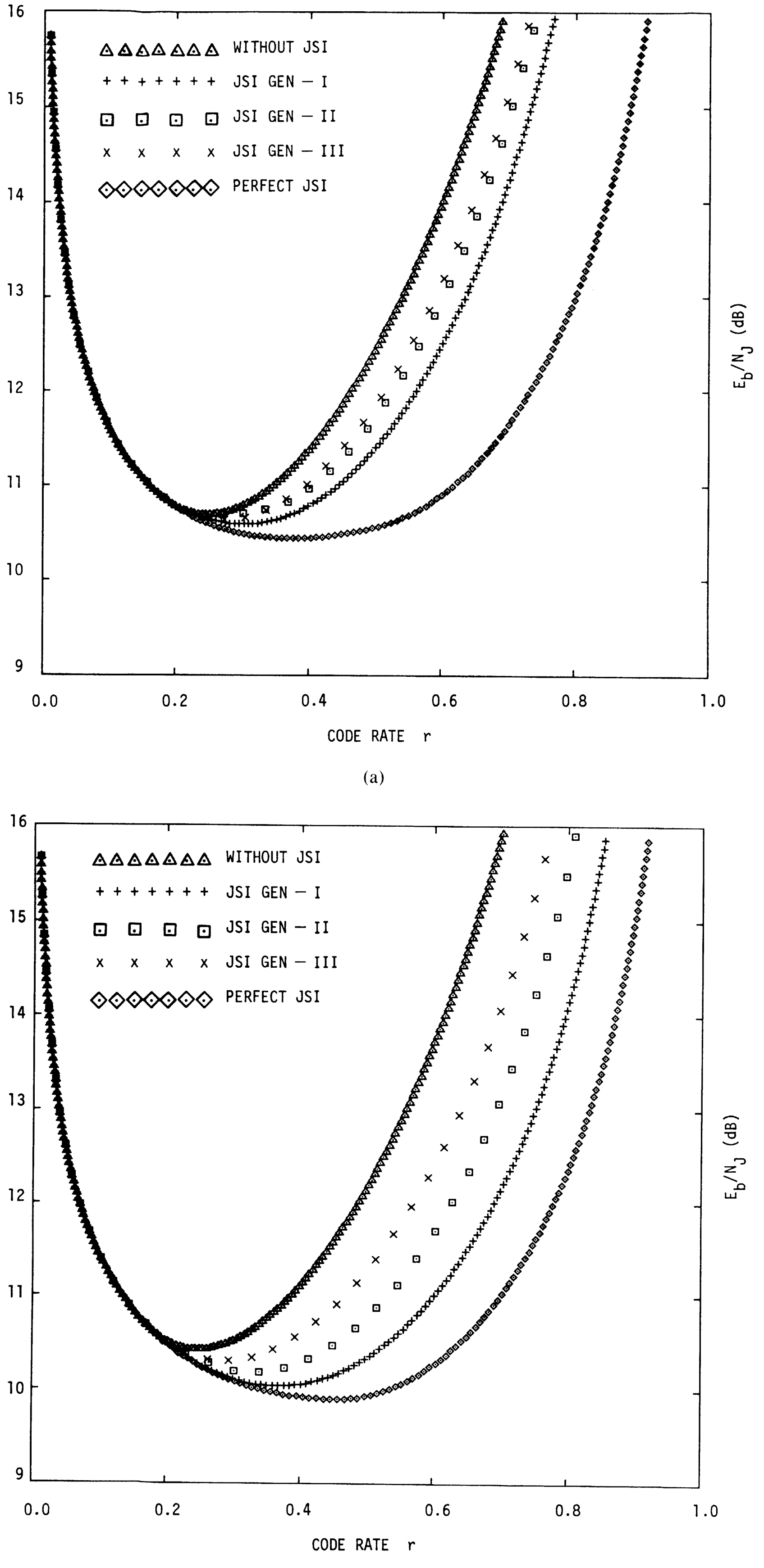


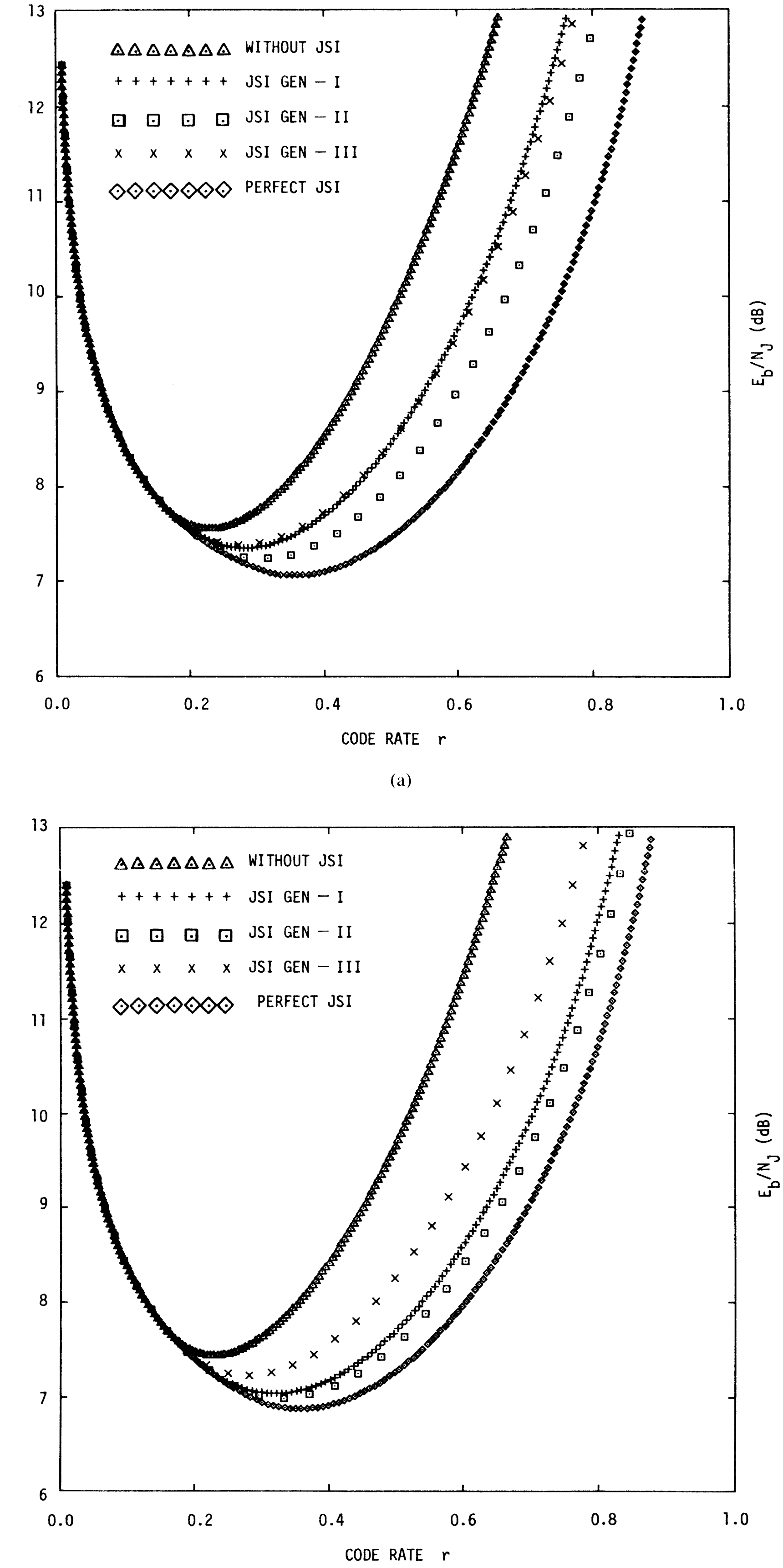
Fig. 4. (a) E_b/N_J needed to achieve cutoff rate for frequency-hopped binary FSK with soft-decision metrics when E_s/N_0 = 15 dB. (b) E_b/N_J needed to achieve cutoff rate for frequency-hopped 8-ary FSK soft-decision metrics when E_s/N_0 = 19.8 dB.



(a)

(b)

Fig. 5. (a) E_b/N_J needed to achieve cutoff rate for frequency-hopped binary FSK with hard-decision metrics when $E_s/N_0 = 15$ dB. (b) E_b/N_J needed to achieve cutoff rate for frequency-hopped binary FSK with hard-decision metrics when $E_s/N_0 = 25$ dB.



(a)

(b)

Fig. 6. (a) E_b/N_J needed to achieve cutoff rate for frequency-hopped 8-ary FSK with hard-decision metrics when E_s/N_0 = 19.8 dB. (b) E_b/N_J needed to achieve cutoff rate for frequency-hopped 8-ary FSK with hard-decision metrics when E_s/N_0 = 29.8 dB.

TABLE I

E_b/N_J (dB) NEEDED TO ACHIEVE THE CUTOFF RATE FOR BFSK WITH HARD DECISIONS AND PERFECT JSI, JSI GENERATORS I, II, III AND NO JSI WHEN E_s/N_0 = 8, 10, 12, 14, 16, 20, 30, AND 40 (dB) CODE RATE r = 0.6, 0.7, 0.8, AND 0.9 (INFORMATION SYMBOLS/M-ARY CHANNEL SYMBOLS)

r	JSI	E_s/N_0							
		8 *dB*	10 *dB*	12 *dB*	14 *dB*	16 *dB*	20 *dB*	30 *dB*	40 *dB*
0.6	Perfect	21.816	13.810	11.933	11.122	10.729	10.412	10.227	10.217
	JSI-I	22.514	14.912	13.502	12.799	12.292	11.584	10.703	10.401
	JSI-II	22.515	15.100	13.802	13.208	12.803	12.203	. . .	. . .
	JSI-III	22.515	15.196	13.916	13.403	13.006	12.586	12.401	12.398
	Without	22.516	15.396	14.400	14.105	13.991	13.808	13.713	13.711
0.7	Perfect	>>1	15.738	12.846	11.859	11.459	11.147	10.957	10.951
	JSI-I	>>1	17.743	15.536	14.634	13.939	12.941	11.725	11.237
	JSI-II	>>1	17.939	15.935	15.232	14.649	13.928	. . .	. . .
	JSI-III	>>1	17.949	16.046	15.438	14.941	14.330	14.133	14.131
	Without	>>1	18.238	16.636	16.334	16.143	16.036	15.939	15.938
0.8	Perfect	>>1	20.569	14.557	13.262	12.856	12.472	12.278	12.273
	JSI-I	>>1	23.967	18.768	17.568	16.668	15.267	13.453	12.751
	JSI-II	>>1	24.068	19.263	18.362	17.668	16.661	. . .	. . .
	JSI-III	>>1	24.166	19.366	18.566	17.969	17.067	16.860	16.858
	Without	>>1	24.269	19.967	19.568	19.462	19.269	19.259	19.259
0.9	Perfect	>>1	>>1	18.356	15.959	15.458	15.155	14.958	14.956
	JSI-I	>>1	>>1	25.257	23.355	22.155	20.053	16.855	15.653
	JSI-II	>>1	>>1	25.753	24.157	23.453	21.957	. . .	. . .
	JSI-III	>>1	>>1	25.754	24.354	23.657	22.455	21.957	21.957
	Without	>>1	>>1	26.355	25.454	25.257	25.154	25.056	25.055

TABLE II

E_b/N_J (dB) NEEDED TO ACHIEVE THE CUTOFF RATE FOR 8-ARY FSK WITH HARD DECISIONS AND PERFECT JSI, JSI GENERATORS I, II, III AND NO JSI WHEN E_s/N_0 = 12.8, 14.8, 16.8, 18.8, 20.8, 24.8, 34.8, AND 44.8 dB, CODE RATE r = 0.6, 0.7, 0.8, AND 0.9 (INFORMATION SYMBOLS/M-ARY CHANNEL SYMBOLS)

r	JSI	E_s/N_0 (*dB*)							
		12.8	14.8	16.8	18.8	20.8	24.8	34.8	44.8
0.6	Perfect	9.428	8.665	8.364	8.234	8.135	8.036	7.956	7.954
	JSI-I	11.536	10.829	10.319	9.844	9.531	9.031	8.408	8.128
	JSI-II	10.920	10.040	9.529	9.143	8.926	8.625	. . .	. . .
	JSI-III	10.842	10.218	9.846	9.718	9.614	9.445	9.422	9.420
	Without	12.233	11.925	11.732	11.631	11.541	11.524	11.441	11.440
0.7	Perfect	10.668	9.779	9.475	9.286	9.187	9.088	9.070	9.068
	JSI-I	13.866	12.964	12.267	11.676	11.262	10.562	9.654	9.268
	JSI-II	12.964	11.867	11.167	10.675	10.367	9.964	. . .	. . .
	JSI-III	12.771	11.877	11.561	11.361	11.256	11.074	11.057	11.055
	Without	14.670	14.277	14.164	14.063	13.970	13.877	13.870	13.870
0.8	Perfect	12.690	11.493	11.187	10.993	10.894	10.794	10.703	10.702
	JSI-I	17.392	16.195	15.296	14.588	13.894	12.892	11.582	11.079
	JSI-II	16.196	14.696	13.789	13.093	12.596	12.081	. . .	. . .
	JSI-III	15.796	14.497	14.089	13.885	13.694	13.588	13.491	13.490
	Without	18.290	17.796	17.688	17.587	17.491	17.395	17.391	17.391
0.9	Perfect	16.884	14.488	14.089	13.983	13.883	13.783	13.690	13.687
	JSI-I	24.085	22.185	21.082	19.982	18.982	17.286	15.079	14.183
	JSI-II	22.786	20.282	18.784	17.685	16.886	15.885	. . .	. . .
	JSI-III	21.985	19.483	18.686	18.481	18.283	18.085	17.985	17.985
	Without	24.883	23.883	23.685	23.585	23.581	23.483	23.481	23.386

optimum detector for the side information generation. However, a MAP rule may not be the optimal strategy for the communicator since the decision threshold that results in a MAP rule may not maximize the cutoff rate of the communication channel. In Tables I and II, columns 7 and 8 for JSI generator II were not completed because of the excessive computing time to find the optimum decision region $R(y)$ from (5). In addition, column 1 (E_s/N_0 = 8 dB) and column 2 (E_s/N_0 = 10 dB) in Table I were not completed because the nonnegligible thermal noise, e.g., E_s/N_0 = 8 dB and 10 dB, does not permit the BFSK communicator to operate in a code rate larger than 0.7 (column 1) or 0.9 (column 2) with zero error probability.

From Tables I and II, we observe that the difference between the perfect JSI generator and JSI generator III can be larger than the 2 dB in E_b/N_J required to achieve the cutoff rate larger than 0.7, even as the thermal noise power approaches zero. (See column 8.) Hence, it can be concluded that if an implementable and bandwidth-efficient side information generator (e.g., generator III) is used, then we should pay attention to the performance difference between the theoretically perfect JSI generator and the practical, imperfect JSI generator, even if the thermal noise is negligible.

APPENDIX

Proof of Theorem 1

The transition probabilities with imperfect JSI, $p(y|x, \hat{S} = 1)$ and $p(y|x, \hat{S} = 0)$, are convex combinations of the transition probabilities with perfect JSI, $p(y|x, S = 1)$ and $p(y|x, S = 0)$, as shown in (1) and (2). We can use a theorem [11, p. 29] which says the mutual information $I(X; Y)$ is convex ∪ in the transition probabilities $p(y|x)$. Thus, the mutual information between X and Y given imperfect JSI is

$$I(X; Y|\hat{S} = 1) \le \alpha I(X; Y|S = 1) + (1 - \alpha) I(X; Y|S = 0) \quad \text{(A-1)}$$

and

$$I(X; Y|\hat{S} = 0) \le \beta I(X; Y|S = 1) + (1 - \beta) I(X; Y|S = 0). \quad \text{(A-2)}$$

Applying the above two inequalities to (8) gives us

$$C_{\text{soft}}^{\text{imperfect}} \le \min_{0 < \rho \le 1} \{\rho I(X; Y|S = 1) + (1 - \rho) I(X; Y|S = 0)\}. \quad \text{(A-3)}$$

The right-hand side of the above inequality is just the channel capacity for soft decisions and perfect JSI, $C_{\text{soft}}^{\text{perfect}}$.

When soft decisions are made, the channel capacity without JSI is given by

$$C_{\text{soft}}^{\text{without}} = \min_{0 < \rho \le 1} I(X; Y) = \int_y p(y|x) \log_M \frac{p(y|x)}{p(y)} \, dy \quad \text{(A-4)}$$

where

$$p(y|x) = \rho p(y|x, S = 1) + (1 - \rho) p(y|x, S = 0)$$

and

$$p(y) = \rho p(y|S = 1) + (1 - \rho) p(y|S = 0).$$

The imperfect JSI generator becomes a worst case generator when the probability of detection is 0.5 and the probability of false alarm is 0.5 for ρ. Assume a worst case imperfect JSI generator. Then the conditional density of y given X and $\hat{S}$ becomes the conditional density of y given X. This implies that the mutual information between X and Y, given $\hat{S}$, becomes the mutual information between X and Y. Hence, from (8), the channel capacity for a worst case imperfect JSI is

$$C_{\text{soft}}^{\text{imperfect}} = \min_{0 < \rho \le 1} \{\tfrac{1}{2} I(X; Y|\hat{S} = 1) + \tfrac{1}{2} I(X; Y|\hat{S} = 0)\} = \min_{0 < \rho \le 1} I(X; Y). \quad \text{(A-5)}$$

The right-hand side of the above equality is just equal to the channel capacity for soft decisions and no JSI. This completes the proof of Theorem 1.

Proof of Theorem 2

Assume the demodulator makes hard decisions. Let $C(x)$ represent the capacity of an M-ary symmetric channel with crossover probability x. Then

$$C(x) = 1 + (1 - x)\log(1 - x) + x \log_M \frac{x}{M - 1}. \tag{A-6}$$

The channel capacity for hard decisions and imperfect JSI is then from (10):

$$\begin{aligned} C_{\text{hard}}^{\text{imperfect}} = \min_{0<\rho\le 1} \Big\{ &\big(\rho P_D(\rho) + (1-\rho)P_{FA}(\rho)\big)C(\bar{p}_1) \\ &+ \big(\rho(1 - P_D(\rho)) \\ &+ (1-\rho)(1 - P_{FA}(\rho))\big)C(\bar{p}_0)\Big\}. \end{aligned} \tag{A-7}$$

Since the capacity is convex ∪ in the transition probabilities p_1 or p_0 [11, theorem, p. 29], the capacity is then

$$\begin{aligned} C(\bar{p}_1) = C(\alpha p_1 + (1-\alpha)p_0) \le \alpha C(p_1) \\ + (1-\alpha)C(p_0) \end{aligned} \tag{A-8}$$

and

$$\begin{aligned} C(\bar{p}_0) = C(\beta p_1 + (1-\beta)p_0) \le \beta C(p_1) \\ + (1-\beta)C(p_0). \end{aligned} \tag{A-9}$$

Applying the above two inequalities to (A-7) yields

$$\begin{aligned} C_{\text{hard}}^{\text{imperfect}} &\le \min_{0<\rho\le 1} \{\rho C(p_1) + (1-\rho)C(p_0)\} \\ &= C_{\text{hard}}^{\text{perfect}}. \end{aligned} \tag{A-10}$$

When no JSI is available, an equivalent optimum jammer's strategy is to choose ρ, maximizing the error probability instead of minimizing the mutual information since $C(x)$ is a decreasing function of the error probability when the error probability is less than $(M - 1)/M$. The average of the error probabilities for $S = 1$ and $S = 0$ gives the error probability of the coding channel. Thus, the channel capacity without JSI is then

$$\begin{aligned} C_{\text{hard}}^{\text{without}} &= C\Big(\max_{0<\rho\le 1}\,[\rho p_1 + (1-\rho)p_0]\Big) \\ &\le \min_{0<\rho\le 1} C(\rho p_1 + (1-\rho)p_0). \end{aligned} \tag{A-11}$$

The above inequality follows because $C(x)$ is a decreasing function. The right-hand side of the above inequality is just equal to the channel capacity for hard decisions and imperfect JSI when $P_D(\rho) = 0.5$ and $P_{FA}(\rho) = 0.5$. We know that a worst case imperfect JSI generator produces $P_D(\rho) = 0.5$ and $P_{FA}(\rho) = 0.5$ for any ρ. This completes the proof of Theorem 2.

References

[1] W. E. Stark, "Coding for frequency-hopped spread-spectrum communication with partial-band interference—Part I: Capacity and cutoff rate," *IEEE Trans. Commun.*, vol. COM-33, pp. 1036-1044, Oct. 1985.

[2] M. B. Pursley, "Packet error probabilities in frequency-hop radio networks—Coping with statistical dependence and noisy side information," in *IEEE Global Telecommun. Conf. Rec.*, Dec. 1986, pp. 5.2.1-5.2.6.

[3] ——, "Multiple-access capability of frequency-hop transmission with noisy side information," in *Proc. IEEE Int. Symp. Inform. Theory*, Ann Arbor, MI, Oct. 1986, pp. 142-143.

[4] ——, "Tradeoffs between side information and code rate in slow-frequency-hop packet radio networks," in *IEEE Global Telecommun. Conf. Rec.*, Dec. 1987, pp. 26.4.1-26.4.6.

[5] A. J. Viterbi, "A robust ratio-threshold technique to mitigate tone and partial-band jamming in coded MFSK system," in *Proc. IEEE MILCOM*, Oct. 1982, pp. 22.4.1-22.4.5.

[6] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread-Spectrum Communications, Vol. I, II.* Rockville, MD: Computer Science Press, 1985.

[7] A. J. Viterbi and I. M. Jacobs, "Advances in coding and modulation for noncoherent channels affected by fading, partial band, and multipath access interference," in *Advances in Communication Systems, Vol. 4.* New York: Academic, 1975, pp. 279-308.

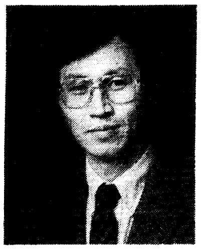
[8] B. K. Levitt, "FH/MFSK performance in multitone jamming," *IEEE J. Select. Areas Commun.*, vol. SAC-3, pp. 627-643, Sept. 1985.

[9] J. K. Omura and B. K. Levitt, "Coded error probability evaluation for antijam communication systems," *IEEE Trans. Commun.*, vol. COM-30, pp. 896-903, May 1982.

[10] J. S. Lee, L. E. Miller, and R. H. French, "The analysis of uncoded performances for certain ECCM receiver design strategies for multi-hops/symbol FH/MFSK waveforms," *IEEE J. Select. Areas Commun.*, vol. SAC-3, pp. 611-621, Sept. 1985.

[11] R. J. McElice, *The Theory of Information and Coding.* Reading, MA: Addison-Wesley, 1977.

[12] A. D. Whalen, *Detection of Signals in Noise.* New York: Academic, 1971.



Hyuck M. Kwon (S'82–M'84) was born in Korea on May 9, 1953. He received the B.S. and M.S. degrees in electrical engineering from Seoul National University, Seoul, Korea, in 1978 and 1980, respectively, and the Ph.D. degree in computer, information, and control engineering from the University of Michigan, Ann Arbor, in 1984.

In 1985 he joined the faculty of the University of Wisconsin, Milwaukee, as an Assistant Professor of Electrical Engineering and Computer Science. Since 1989 he has been with Lockheed Engineering & Sciences, Houston, TX, as a Principal Engineer, on leave from the University of Wisconsin, Milwaukee. In addition, since 1985 he has served as a consultant for J. S. Lee Associates, Inc., Arlington, VA. His research interests are in the area of digital communications, with special emphasis on spread spectrum communications, space station and shuttle communications, information theory, coded modulation, and neural networks.

Dr. Kwon received a national scholarship from the Minister of the Education Department of the Republic of Korea for his Ph.D. degree.

Optimum Transmission Ranges in a Direct-Sequence Spread-Spectrum Multihop Packet Radio Network

ELVINO S. SOUSA, MEMBER, IEEE, AND JOHN A. SILVESTER, SENIOR MEMBER, IEEE

Abstract—**In this paper, we obtain the optimum transmission ranges to maximize throughput for a direct-sequence spread-spectrum multihop packet radio network. In the analysis, we model the network self-interference as a random variable which is equal to the sum of the interference power of all other terminals plus background noise. The model is applicable to other spread spectrum schemes where the interference of one user appears as a noise source with constant power spectral density to the other users. The network terminals are modeled as a random Poisson field of interference power emitters. The statistics of the interference power at a receiving terminal are obtained and show to be the stable distributions of a parameter that is dependent on the propagation power loss law. The optimum transmission range in such a network is of the form CK^{α} where C is a constant, K is a function of the processing gain, the background noise power spectral density, and the degree of error-correction coding used, and α is related to the power loss law. The results obtained can be used in heuristics to determine optimum routing strategies in multihop networks.**

I. Introduction

In a large distributed packet radio network, it is not always desirable for a terminal with a packet to send to attempt to transmit directly to the destination. It may be the case that the destination terminal is out of the transmitter's range, in which case it is impossible to transmit directly to the destination, or that the destination is within range, but the transmission protocol dictates that the packet take a series of short hops so as to achieve "space reuse" [1]–[3] which results in a higher network throughput. If a packet is transmitted directly to the destination, then there is no "store and forward" delay, but due to a larger number of potentially interfering terminals, the probability of a successful transmission is smaller than that for short-hop transmission. Using a model where terminals are assumed to be randomly distributed on the plane, Kleinrock and Silvester [1]–[2] were able to show that, to maximize overall network throughput, a terminal should transmit with a power so that the average number of terminals within range is six. Subsequent refinements to this analysis [4]–[5] which also include different routing strategies have resulted in the conclusion that the number of neighbors of a terminal should be approximately six–eight.

These results were derived for narrow-band radio networks where at most one successful transmission at a time can occur in a given region of space. With spread spectrum signaling [6], multiple simultaneous successful transmissions are possible, and the above results do not apply. Moreover, the model used in these analyses assumes that the reception of a packet is independent of the distance from the transmitter to the receiver, as long as this distance is less than a critical radius. Also, if another terminal is undergoing a reception just outside of the transmission range of node *X*, then according to these models, the reception is unaffected by the interference from a transmission by node *X*. In this paper, we are concerned with the solution of the above problem (i.e., determining the optimum transmission ranges) for the direct sequence spread spectrum (DSSS) case. In the work reported in [1]–[5], the main issue is the choice of a transmitting power which defines a transmission radius that depends on the background noise. The choice of transmission power is a compromise between network connectivity and network interference which occurs in the form of packet collisions from terminals transmitting within the receiving radius. In such a model, the probability of packet success is assumed not to be a strong function of the distance between the transmitter and receiver if this distance is less than the transmission radius. In the case of spread spectrum transmission, a receiver will pick up some noise from each transmitter, and the packet error probability will be strongly dependent on the signal strength, even within the transmission radius. If the source of interference is mostly from other users, the probability of packet success will not depend as much on the absolute transmission power that each terminal uses since scaling the signal power also scales the interference by the same amount, but more on the transmission range selection. The transmission range should be expressed relative to the average distance between terminals. A related measure is the average number of terminals that are closer to the receiver than the transmitter. The transmission range must, of course, be less than the link distance or transmission radius that is defined by the transmitter power used. In general, the greater the transmission range, the lower is the probability of the transmission being successful. If the objective function is expected forward

Manuscript received February 10, 1989; revised November 3, 1989. This work was supported in part by the U.S. Army Research Office under Contract DAAG29-84-K-0084 and in part by the Natural Sciences and Engineering Research Council of Canada under a URF Grant. This paper was presented in part at the IEEE Military Communications Conference, Boston, MA, October 1985.

E. S. Sousa is with the Department of Electrical Engineering, University of Toronto, Toronto, Ont., Canada M5S 1A4.

J. A. Silvester is with the Communication Sciences Institute, Department of Electrical Engineering—Systems, University of Southern California, Los Angeles, CA 90089.

IEEE Log Number 9035804.

0733-8716/90/0600-0762$01.00 © 1990 IEEE

progress per transmission, then there is an optimum distance that the packet should attempt to travel per hop, and this distance is not necessarily the full distance to the destination. We assume that the transmitted power and processing gain are fixed, and that the main source of interference is multiaccess interference, and we find the expected forward progress per transmission, defined as the product of the probability of success times the link distance, in terms of the expected number of interferers that are closer to the receiver than the transmitter.

We derive the statistics of the received interference power at a terminal for a class of signal propagation laws (i.e., how the strength of a propagating signal varies with distance). For the inverse fourth power law (commonly used for ground radio), we show that the optimum transmission range is such that on the average, the number of terminals closer to the transmitter than the receiver is proportional to the square root of the processing gain.[1] Even though the analysis presented in this paper is for the direct sequence form of spread spectrum, depending on the detection scheme, the methdology may also be applicable to other spread spectrum schemes.

It is hoped that the technique presented here to analyze the interference at a terminal will have wider ranging implications in the analysis of routing strategies, adaptive techniques involving the variation of transmitter power, and the impact of jammers whose positions are randomly varying in space [7].

II. Multihop Networks

In most multihop network models employed by researchers thus far, each transmitter is assumed to use the same transmitting power. This power is assumed to determine a circle such that each terminal lying within the circle hears the given transmission, and any terminal lying outside the circle is completely unaffected by the transmission. If we denote the strength of the transmitted signal as a function of the distance from the transmitter by $g(r)$ (where g stands for gain), then the above model corresponds to a g of the form

$$g(r) = \begin{cases} c & 0 \le r \le r_0 \\ 0 & r > r_0 \end{cases} \tag{1}$$

where c is some constant and r_0 is the critical radius determined by the transmitted power. With this model, the network can be represented by a graph with vertices corresponding to the terminals and an edge present between two vertices if and only if the distance between them is less than r_0. According to the model, a transmitted packet is successful if and only if no other terminal adjacent to the destination transmits at the same time.

A weakness of the above model can be illustrated with the aid of Fig. 1 which depicts a network of four terminals that use a constant (equal) power to communicate. If the transmitting power is such that the critical radius is determined to be r_0, then according to the model, terminals $\boldsymbol{a}_1$ and $\boldsymbol{b}_1$ can successfully transmit to terminals $\boldsymbol{a}_2$ and $\boldsymbol{b}_2$, respectively, at the same time. However, if the powers are increased so that the critical radius becomes r_0', then the above two communications cannot occur simultaneously. On the other hand, from a communication theory point of view, we know that the important parameter determining the success of a transmission is the signal-to-noise ratio at the receiver, which is not strongly dependent on the transmission power as long as all terminals transmit with the same power and the background noise is much smaller than the interference power.

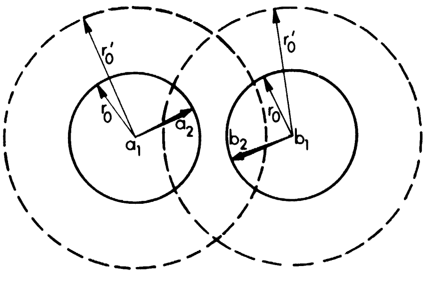


Fig. 1. The effect of transmission radius on interference.

The main drawback to the above model is that it does not discriminate between the differences in distance to a receiver of two transmitting terminals as long as the terminals are within the critical radius. An improved model is the capture model studied in [8]–[10], which is suited to FM transmission. If a given transmitting terminal, at distance r_1 from the receiver, is the closest transmitting terminal to the receiver, and if the next closest transmitting terminal is at a distance r_2, then the given transmission will be successful if $r_1 < r_0$ and the ratio r_2/r_1 exceeds a threshold called the capture ratio.

The above models have been used for the nonspread spectrum case. A straightforward extension of these models to spread spectrum would result by setting a threshold on the maximum possible number of successful simultaneous transmissions and declaring that any time the number of transmissions exceeds the threshold, all transmissions are lost. However, since the powers of the various interferers vary greatly due to differences in their distances to the receiver, the number of transmitting terminals is not a good variable to work with in accounting for network interference at a particular receiver. This is especially the case with DSSS signaling. The analysis presented in this paper is an enhancement of that in [11] where we use the sum of the interference powers to model multiuser interference. This approach to interference modeling has been used in many analyses of cellular radio systems (e.g., [12]), spread spectrum multiple-access systems (e.g., [13]–[14]), and has recently also been adopted for throughput analysis of packet radio networks with fixed topologies and prespecified routing schemes in [15].

[1] The processing gain is equal to the ratio of the system bandwidth to the uncoded data rate (see [6, p. 138]).

III. System Model

We assume a multihop packet radio network operating under heavy traffic conditions. The system is slotted, and in each slot, a terminal transmits a packet with probability p. The slot duration is assumed to be sufficiently large so as to allow a preamble for spreading code and carrier synchronization. The traffic matrix is assumed to be uniform. We are interested in calculating performance over many different changing topologies rather than for a specific terminal configuration. As a result, we obtain statistical performance values over a set of topologies. To do this, we model the positions of the terminals as a Poisson point process in the plane with parameter λ. If A is the area of a given region $\boldsymbol{R}$ in the plane, then the probability of finding k terminals in $\boldsymbol{R}$ is given by

$$P[k \text{ in } \boldsymbol{R}] = \frac{e^{-\lambda A}(\lambda A)^k}{k!} \tag{2}$$

where λ is the average number of terminals per unit area. We assume that during each slot, the network topology is constant; hence, the interference level will be constant over a packet transmission time. We also assume that the interference is independent from slot to slot. This is the assumption that was made in [1]–[2] and other work that followed. With this assumption, we may apply the results to a network with dynamically changing topology or to obtain average performance results for a collection of random networks. In this work, we are interested in obtaining optimum transmission ranges; hence, we assume a prespecified link in the network with a given distance R. The objective is to optimize the expected forward progress for a packet transmission as a function of the distance of the link in the presence of unknown network interferers that are modeled as a Poisson process. This philosophy is also consistent with the approach taken in [13] where a centralized system is analyzed.

IV. Interference Modeling

The collision model for channel interference is not applicable in the case of DSSS signaling or when interfering signals have small powers. An alternative model that is not usually used in the narrow-band packet radio network literature, but has been used in many analysis of DSSS systems such as in [12]–[14] is that of summing the interference powers and treating the total interference as Gaussian noise. At the network analysis level, many spread spectrum schemes may be modeled this way. In this paper, we assume a direct sequence scheme with binary phase shift keying (DS/BPSK), although many other schemes such as DS/DPSK and the DPSK schemes that have been considered for cellular radio (e.g., [12]) allow the same type of analysis.

From a network analysis viewpoint, we are usually interested in calculating the probability of packet success given that the receiver is idle. This calculation is usually conditioned on some state of the network, and a tractable state model cannot usually contain much more than information on which terminals are transmitting and what are their received powers. A desirable model to work with in spread spectrum is the threshold model where we assume that the packet is successful if the signal to interference power ratio is greater than some threshold. To apply the threshold model, the variance of the interference at the detector must be directly related to the received interference power. In DS/BPSK systems with a large processing gain, it can indeed be shown that the noise at the detector due to one interferer is approximately Gaussian; however, the variance depends on the relative chip phase of the signal to the interferer. For the case of many interfering signals with approximately equal levels of interference, the chip phases average out and the noise variance is constant for a prespecified location of interferers. However, in some DS/BPSK system models with one strong interferer, the variance of the noise at the detector is a random variable that depends on the chip phase of the signal relative to the interferer. For a given total received power, the noise at the detector will have a variance that will vary from packet to packet. This variation can, however, be made small if there is an offset between the clocks of the various signals. We will assume this to be the case in this paper.

Threshold models in digital communications are ultimately dependent on the degree of error-correction coding. If the interference over a packet can be modeled as Gaussian noise, then the probability of packet success as a function of the signal-to-interference ratio is a smooth curve with a slope in the transition region that depends on the degree of error correction; for good long codes, the curve approaches a step function. In this paper, we model the actual smooth curve.

To calculate the packet success probability, we assume that the level of interference is constant over the transmission of a packet. The noise at the detector is due to interference from other users and to a constant background noise with power spectral density $N_0/2$. We denote the symbol energy-to-noise ratio at the detector by $E_b/N_{0\text{eff}}$ where $N_{0\text{eff}}/2$ is an equivalent white noise power spectral density for the same SNR at the detector. If the received signal has power P_0 and the interferers have powers $P_1, P_2, \cdots, P_q$, the average symbol energy-to-noise ratio at the detector in the case of DS/BPSK with rectangular chip pulse is then (see [16, eq. (17)] for a similar result with equal powers)

$$\mu \triangleq \frac{E_b}{N_{0\text{eff}}} = \left(\frac{2Y}{3LP_0} + \frac{1}{\mu_0}\right)^{-1} \tag{3}$$

where L is the processing gain, $Y = \Sigma_{i=1}^{q} P_i$, and $\mu_0 = E_b/N_0$. The parameter μ_0 is the SNR at the detector in the absence of interferers.

For a given μ, the probability of symbol error is

$$p_e = \tfrac{1}{2}\operatorname{erfc}(\sqrt{\mu}). \tag{4}$$

The probability of packet success is dependent on the coding scheme. We denote the probability of packet success conditioned on the SNR μ as $s(\mu)$. As an example, for a t-error-correcting block code of length n, and under our assumption that symbol errors are independent given μ, we have

$$P[\text{success}/\mu] \triangleq s(\mu) = \sum_{i=0}^{t} \binom{n}{i} \left(\frac{1}{2} \operatorname{erfc} \sqrt{\mu}\right)^{i} \left(1 - \frac{1}{2} \operatorname{erfc} \sqrt{\mu}\right)^{n-i} \tag{5}$$

From packet to packet, the parameter μ is a random variable with density function $f_\mu(\cdot)$. The unconditional probability of packet success can then be written as

$$P_s = \int_0^\infty s(x) f_\mu(x)\, dx = \int_0^\infty [1 - F_\mu(x)]\, s'(x)\, dx \tag{6}$$

where the second expression is obtained after an integration by parts and $F_\mu(\cdot)$ is the probability distribution function of μ.

V. Interference at a Given Terminal

To evaluate the above probability of packet success, we need to obtain the probability density function $f_\mu(\cdot)$. We prefer, though, to obtain the probability density $f_Y(\cdot)$ first, i.e., the probability density function for the multiuser interference power. Towards this end, we may assume that the terminal at which we are interested in obtaining the interference power is located at the origin.

Let $g(r)$ be the power of a given signal at a distance r from the transmitter of the signal. In general, the exact form of g will depend on the environment; however, it will always be very large for small r, and will approach zero as $r \to \infty$. For the moment, though, so as not to restrict ourselves to any particular environment, we merely assume that $g(r)$ satisfies the following two conditions:

1) $g(r)$ is monotone decreasing, $\lim_{r \to 0} g(r) = \infty$,

$$\text{and } \lim_{r \to \infty} g(r) = 0 \tag{7a}$$

2) $$\lim_{r \to \infty} r^2 g(r) = 0. \tag{7b}$$

Without condition (7b), it can be shown that the interference power at a given terminal would be infinite for an infinite network. We will see later that in order for the characteristic function of the interference power to exist, we require that condition (7b) hold. The above expression for the power loss law is a far-field approximation that does not hold close to the transmitter where the transmitted signal attains a maximum. We assume that this maximum is sufficient to cause a transmission error, even in the case of one interferer; hence, the above function will result in the same probability of error while being easier to handle analytically.

Let X be a Poisson process in the plane with the average number of points per unit area equal to λ. A sample function of X is determined by a set of points in the plane which will correspond to locations of terminals. The probability law for X is determined by (2). We assume that the probability that a terminal is transmitting is p. The set of transmitting terminals also forms a Poisson process X' with parameter $\lambda_t = \lambda p$. Now, with each sample function of X', we can associate the random variable

$$Y = \sum g(r_i) \tag{8}$$

where the summation is over all the points of the sample function, and r_i is the distance of the ith point to the origin. We assume that each terminal (i.e., each point of the sample function) is transmitting. Thus, Y is the total interference power at the origin, and we wish to find its probability density.

Let Y_a be the interference power received from those terminals which are in a disk or radius a, i.e.,

$$Y_a = \sum_{r_i \leq a} g(r_i). \tag{9}$$

Thus, we have $\lim_{a \to \infty} Y_a = Y$. We work with Y_a and then let $a \to \infty$ to obtain the characteristic function of Y. The probability density function is then the inverse Fourier transform. Let ϕ_{Y_a} be the characteristic function of Y_a, i.e.,

$$\phi_{Y_a}(\omega) = E(e^{i\omega Y_a}). \tag{10}$$

Using conditional expectations, this may be evaluated as

$$E(e^{i\omega Y_a}) = E(E(e^{i\omega Y_a}/k \text{ in } D_a)) = \sum_{k=0}^{\infty} \frac{e^{-\lambda_t \pi a^2}(\lambda_t \pi a^2)^k}{k!} E(e^{i\omega Y_a}/k \text{ in } D_a) \tag{11}$$

where "k in D_a" is the event that there are k terminals in the disk of radius a, and the expectation is over the random variable Y_a.

Now, given that there are k points in D_a, and due to the nature of the Poisson process, the distribution of their locations is that of k independent and identically distributed points with uniform distribution. If R is the distance to the origin of a point that is uniformly distributed in D_a, then the probability density of R is

$$f_R(r) = \begin{cases} \dfrac{2r}{a^2} & r \leq a \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

Also, since the characteristic function of the sum of a number of independent random variables is the product of the individual characteristic functions, we have

$$E(e^{i\omega Y_a}/k \text{ in } D_a) = \left(\int_0^a \frac{2r}{a^2} e^{i\omega g(r)}\, dr\right)^k. \tag{13}$$

Note that in (13), we are considering Y_a to be the sum of k random variables which are functions of the random variables R_i, and each R_i has the density given by (12). Thus, substituting (13) in (11) and summing the series, we obtain

$$\phi_{Y_a}(\omega) = \exp\left(\lambda_t \pi a^2\left(\int_0^a \frac{2r}{a^2} e^{i\omega g(r)}\, dr - 1\right)\right). \quad (14)$$

Integrating by parts, the exponent of (14), letting $a \to \infty$, and using condition (7a), we obtain, after some simplification, the characteristic function of Y:

$$\phi_Y(\omega) = \exp\left(i\lambda_t \pi\omega \int_0^\infty \left[g^{-1}(t)\right]^2 e^{i\omega t}\, dt\right) \quad (15)$$

where $g^{-1}(\cdot)$ denotes the inverse of $g(\cdot)$.

A. Statistics for a Class of Propagation Laws

To proceed further, we must now specify $g(r)$. We specify it only up to a multiplicative constant since the results that we obtain will be independent of scaling factors. In free space, $g(r)$ would be $1/r^2$; however, this does not satisfy (7b). This means, as we will see shortly, that if we are going to assume the ideal law for g, then we cannot assume an infinite network, for the interference power would be strongly dependent on the network size. On the ground, $g(r)$ takes the form $1/r^4$ [17]; thus, to work out this case and any other cases where the dependency on r is not exactly an inverse fourth power, we consider the following class of propagation laws:

$$g(r) = \frac{1}{r^\gamma} \qquad \gamma > 2. \quad (16)$$

For this class of propagation laws, (15) becomes

$$\phi_Y(\omega) = \exp\left(i\lambda_t \pi\omega \int_0^\infty t^{-\alpha} e^{i\omega t}\, dt\right) \quad (17)$$

where $\alpha = 2/\gamma$. The integral in (17) may be evaluated to obtain the following:

$$\phi_Y(\omega) = \exp\left(-\pi\lambda_t \Gamma(1-\alpha) e^{-\pi\alpha/2}\omega^\alpha\right) \qquad \omega \geq 0 \quad (18)$$

where $\Gamma(\cdot)$ is the gamma function and $\phi_Y(\omega) = \phi_Y^*(-\omega)$. The probability laws with characteristic functions given by (18) are the stable laws of exponent α with the restriction $0 < \alpha < 1$ [18]. For $\alpha = 1/2$, (18) becomes

$$\phi_Y(\omega) = \exp\left(-\pi\sqrt{\pi/2}(1-i)\lambda_t\sqrt{\omega}\right). \quad (19)$$

This probability law ($\alpha = 1/2$) is the inverse Gaussian probability law, and is the only one of the stable laws, for the case $0 < \alpha < 1$, which is known to have a density given by a closed-form expression. The density for $\alpha = 1/2$ is given by

$$f_Y(y) = \frac{\pi}{2}\lambda_t y^{-3/2} e^{-\pi^3\lambda_t^2/4y} \quad (20)$$

and the distribution function is

$$F_Y(y; 1/2) = \operatorname{erfc}\left(\frac{\pi^{3/2}\lambda_t}{2\sqrt{y}}\right). \quad (21)$$

In general, for $0 < \alpha < 1$, the densities can be found as infinite series (see [18, pp. 581-583]). Let $\rho = \pi\lambda_t\Gamma(1-\alpha)$; then the density, obtained by taking the inverse Fourier transform of (18), is

$$f_Y(y; \alpha) = \frac{1}{\pi y}\sum_{k=1}^{\infty} \frac{\Gamma(\alpha k + 1)}{k!}\left(\frac{\rho}{y^\alpha}\right)^k \sin k\pi(1-\alpha) \quad (22)$$

which, for $\alpha = 1/2$, results in a series expansion for (20). The general distribution function is

$$F_Y(y; \alpha) = \frac{1}{\pi}\sum_{k=1}^{\infty} \frac{\Gamma(\alpha k)}{k!}\left(\frac{\rho}{y^\alpha}\right)^k \sin k\pi(1-\alpha). \quad (23)$$

VI. Probability of Packet Success

We assume a $1/r^4$ propagation power loss law. Using the above distribution function for the interference, the probability of packet success is now computed. Let the distance between the transmitter and receiver be R. The signal power is then $1/R^4$. The random variables Y and μ are related through (3); hence, the probability distribution function for μ may be obtained from that of Y in (21) as

$$F_\mu(\mu) = \begin{cases} 1 - \operatorname{erfc}\left(\frac{p\lambda\pi R^2}{2}\sqrt{\frac{\pi}{K(\mu)}}\right) & \mu < \mu_0 \\ 0 & \mu > \mu_0 \end{cases} \quad (24)$$

where

$$K(\mu) = \frac{3L}{2}\left(\frac{1}{\mu} - \frac{1}{\mu_0}\right).$$

Substituting (24) in (6), we obtain the probability of packet success

$$P_s = \int_0^{\mu_0} \operatorname{erfc}\left(\frac{p\lambda\pi R^2}{2}\sqrt{\frac{\pi}{K(\mu)}}\right) s'(\mu)\, d\mu. \quad (25)$$

In the above equation, $K(\mu) + 1$ may be interpreted as a multiple-access capability [14], in the case of equal interference powers, given the required SNR μ at the detector [as can be seen from (3)]. The function $s'(\cdot)$ depends on the level of coding. For the best long codes, this function approaches a delta function at some value of μ, μ_c. The integral can then be easily evaluated, and the result corresponds to working with a reception model based on a threshold assumption. In any case, it can be seen that for the purpose of probability of packet success calculations, we can always assume a threshold model. The above equation gives the means to obtain the effective threshold, defined as the threshold, which when used with the threshold model, gives the same results as (25).

VII. Optimum Transmission Ranges

We are now ready to apply the above statistics of the interference power to the determination of the optimum transmission range in a multihop network. First, we find an expression for nodal throughput as a function of the link distance R, the average number of terminals per unit area λ, the transmission probability p, and the processing gain L.

A. Local Throughput

The local throughput is the rate at which a terminal successfully transmits packets. For a network with uniform traffic (as we assume here) and assuming that the routing is "balanced," then the local throughput will be the same for all terminals. In terms of packets per slot, the local throughput will simply be the probability of success. Let ζ be the local throughput from terminal $\boldsymbol{A}$, and let terminal $\boldsymbol{B}$ designate a generic destination terminal; then we have

$$\zeta = P[\boldsymbol{A} \text{ transmits}] \cdot P[\boldsymbol{B} \text{ does not transmit}] \cdot P[\text{packet received}/\boldsymbol{B} \text{ does not transmit}]. \tag{26}$$

Given that $\boldsymbol{B}$ does not transmit, it may not receive $\boldsymbol{A}$'s transmission for one of the following two reasons: 1) $\boldsymbol{B}$ may receive a transmission from another terminal, or 2) the interference power at $\boldsymbol{B}$ may exceed the threshold. These two events are not strictly independent, and the exact calculation of the last factor of (26) is a difficult task. The lack of strict independence is due to the fact that if the interference power is large, then there is a greater probability of a large number of terminals in the vicinity of $\boldsymbol{B}$. However, we will see shortly that the probability that $\boldsymbol{B}$ receives another transmission is weakly dependent on the number of terminals in the vicinity of $\boldsymbol{B}$, and we may assume that the above two events are independent. If all terminals transmit with probability p (heavy traffic case), then the first two factors of (26) are given by $p(1-p)$. Now, due to the memoryless property of the Poisson distribution, if we fix a transmitter and receiver, terminals $\boldsymbol{A}$ and $\boldsymbol{B}$, the remaining terminals are still Poisson distributed with parameter λ; thus, the probability of the second of the above events is simply P_s and (26) becomes

$$\zeta = p(1-p) \cdot P\left[\begin{array}{c} \boldsymbol{B} \text{ chooses} \\ \boldsymbol{A}\text{'s transmission} \end{array}\right] \cdot P_s. \tag{27}$$

To obtain the probability that $\boldsymbol{B}$ chooses $\boldsymbol{A}$'s transmission, we need to know how many terminals are transmitting to $\boldsymbol{B}$. If k terminals (including $\boldsymbol{A}$) are transmitting to $\boldsymbol{B}$, then we assume that the probability that $\boldsymbol{B}$ receives $\boldsymbol{A}$'s transmission (assuming, of course, that $\boldsymbol{B}$ does not transmit and the interference power is less than the threshold) is $1/k$. The exact calculation that $\boldsymbol{B}$ chooses $\boldsymbol{A}$'s transmission is very difficult. We do not know exactly how many terminals are potential transmitters to $\boldsymbol{B}$; and of the potential transmitters to $\boldsymbol{B}$, we do not know the probability of a transmission to $\boldsymbol{B}$ from a given one of them. These parameters are tied to the results that we are trying to obtain. If the number of potential transmitters to $\boldsymbol{B}$ is n, and if we assume that each of these can transmit to n terminals (hence, the probability of a transmission to $\boldsymbol{B}$ is p/n), i.e., local traffic is uniform, then we have

$$\begin{aligned} P\left|\begin{array}{c} \boldsymbol{B} \text{ chooses} \\ \boldsymbol{A}\text{'s transmission} \end{array}\right| &= \sum_{i=0}^{n-1} \binom{n-1}{i} \left(\frac{p}{n}\right)^{i} \left(1 - \frac{p}{n}\right)^{n-1-i} \cdot \frac{1}{i+1} \\ &= \frac{1}{p}\left(1 - \left(1 - \frac{p}{n}\right)^{n}\right) \end{aligned} \tag{28}$$

where the summation results from conditioning on the number of transmissions addressed to $\boldsymbol{B}$. Substituting (28) in (27), we obtain

$$\zeta = (1-p)\left(1 - \left(1 - \frac{p}{n}\right)^{n}\right) \cdot P_s \triangleq \tau_n(p) \cdot P_s. \tag{29}$$

We have plotted the factor $\tau_n(p)$ versus p for different values of n in Fig. 2. From the plots, we see that for $n > 2$, $\tau_n(p)$ is not too sensitive to n, and we verify our independence assumption. We therefore assume that $n = \infty$. Letting $n \rightarrow \infty$ in (29), we obtain

$$\zeta = \tau(p) \cdot P_s \tag{30}$$

where

$$\tau(p) \triangleq \tau_\infty(p) = (1-p)(1-e^{-p}).$$

In [19], we discussed two factors which affect the throughput of a spread spectrum network and referred to them as the *tendency to pair up* and the *availability of a channel*. In terms of our notation here, τ is the *tendency to pair up* (given per terminal) and P_s is the *availability of a channel*. From the above assumption that the two events given by 1) and 2) are independent, we have gained the factorization of the throughput into these two factors.

B. Expected Progress per Slot

Using the system model previously described, we determine the optimum transmission range by using the expected forward progress of a packet per slot as the performance criterion. In a multihop network, the probability of packet success increases as the link distance decreases. However, in choosing a small link distance, the number of hops that the packet must take is increased and the network internal traffic is artificially increased. This, in turn, causes the probability of packet success to decrease. A performance measure is required which increases with an increase in the packet success probability and decreases as the number of hops increases. The expected forward

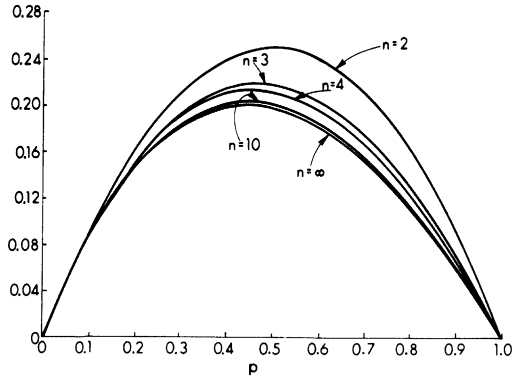


Fig. 2. Tendency to pair up $\tau_n(p)$ versus the probability of transmission p.

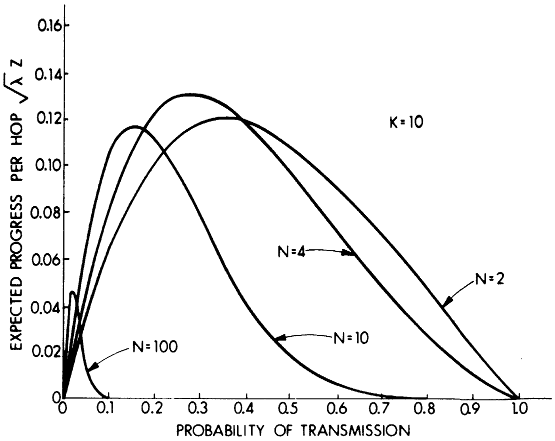


Fig. 3. Expected progress per hop versus probability of transmission.

progress per slot [1]-[5] is such a measure. The expected forward progress per slot Z is

$$Z = \zeta \cdot R. \tag{31}$$

To express the following results in terms of dimensionless quantities, we let $N = \lambda \pi R^2$. N will be the average number of terminals which are closer to $\boldsymbol{A}$ than $\boldsymbol{B}$, and in terms of N, $R = \sqrt{N/\pi\lambda}$. For the propagation law given by $\gamma = 4$ (i.e., $\alpha = 1/2$), we obtain the following [it results from substituting (30) in (31) and using (28) and (25)]:

$$\sqrt{\lambda} Z = \sqrt{N/\pi}(1-p)(1-e^{-p}) \cdot \int_0^{\mu_0} \operatorname{erfc}\left(\frac{pN}{2}\sqrt{\frac{\pi}{K(\mu)}}\right) s'(\mu)\, d\mu. \tag{32}$$

We have multiplied Z by $\sqrt{\lambda}$ in (32) so that the right-hand side is dimensionless. Note that contrary to the analysis in [1]-[5], N is not to be interpreted as the average number of neighbors of a given terminal (A in our case). Rather, N is the average number of terminals that are within the *chosen* range.

To compute (32), we need the coding function $s(\mu)$. As an example, we assume a highly coded system so that $s(\cdot)$ is close to a step function at some critical SNR which we denote as the parameter μ_c. The derivative is then approximately given by $s'(\mu) = \delta(\mu - \mu_c)$ and (32) becomes

$$\sqrt{\lambda} Z = \sqrt{N/\pi}(1-p)(1-e^{-p}) \operatorname{erfc}\left(\frac{pN}{2}\sqrt{\frac{\pi}{K(\mu_c)}}\right). \tag{33}$$

In the above, K is a function of μ_c, the required SNR for the coding scheme used, and also implicitly a function of μ_0 and the processing gain L. K is directly proportional to the processing gain. The parameter μ_c is mainly dependent on the level of coding, and assuming a fixed background noise level, μ_0 is mainly dependent on the link distance R. Setting K equal to zero, we obtain the maximum link distance as $R_0 = (T_s/(\mu_c N_0))^{1/4}$. Even though the maximum link distance is R_0, the probability of reception becomes very small for link distances close to R_0, and an optimum transmission range will be considerably smaller than R_0.

VIII. Parameter Optimization

We have plotted (33) in Figs. 3-5 for the cases of $K(\mu_c)$ equal to 10, 100, and 1000 and for different values of the parameter N. From the figures, we note that in each case, there is an optimum value of N, and as N increases from this value, the expected forward progress decreases. Although we only show four plots for each case of K, the N yielding the maximum expected progress for the given K and at the optimum p has been chosen as the value over all N which yields the maximum expected progress at its optimum probability of transmission p.

In Fig. 6, we show plots of the maximum expected progress versus N with K as a parameter, that is, the peaks in the plots of Figs. 3-5 versus N. An interesting observation is the relation between the optimum N, that is, the N yielding the maximum expected progress in Fig. 6, and the corresponding value of K.

To find the parameters yielding the maximum expected forward progress, we may optimize (32) over the parameters p and N. Setting the partial derivatives with respect to p and N to zero yields the following two equations:

$$\int_0^{\mu_0} \left(\frac{4\psi(\mu)}{\sqrt{\pi}} e^{-\psi(\mu)^2} - \operatorname{erfc}\left[\psi(\mu)\right]\right) s'(\mu)\, d\mu = 0 \tag{34a}$$

$$\int_0^{\mu_0} \frac{e^{\psi(\mu)^2}}{\psi(\mu)} \operatorname{erfc}\left[\psi(\mu)\right] s'(\mu)\, d\mu = \frac{2(1-p)(1-e^{-p})}{\sqrt{\pi} p\left[e^{-p}(2-p) - 1\right]} \tag{34b}$$

where $\psi(\mu) = (pN/2)\sqrt{\pi/K(\mu)}$. To solve the above, we assume a fixed L and μ_0, and then solve (34a) for the function $\psi(\mu)$. For a fixed L and μ_0, such a solution amounts to finding a value for the product pN. We then

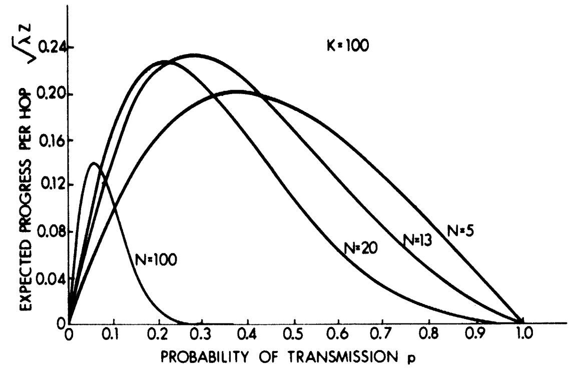


Fig. 4. Expected progress per hop versus probability of transmission.

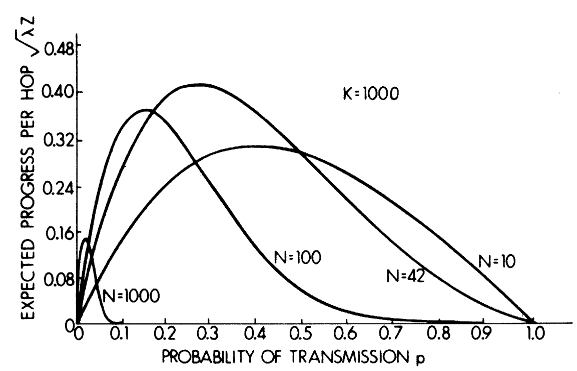


Fig. 5. Expected progress per hop versus probability of transmission.

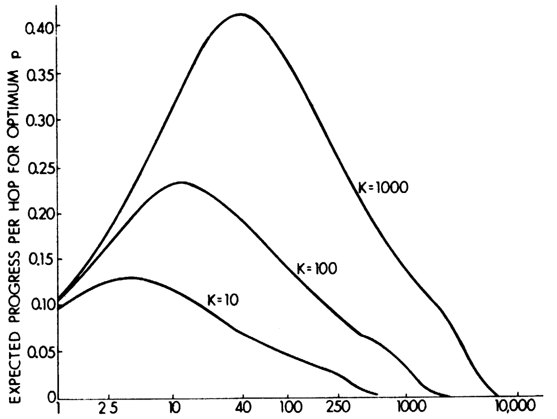


Fig. 6. Expected progress per hop for optimum probability of transmission versus the expected number of interferers that are closer to the receiver than the transmission N.

substitute for $\psi(\mu)$ in the left side of (34b), evaluate the integral, and solve the resulting equation for p. N is then obtained from ψ and p. As an example, for $s(\cdot)$ equal to a step function at $\mu = \mu_c$, we obtain

$$N_0 = 1.33\sqrt{K(\mu_c)} \tag{35a}$$

and

$$p_0 = 0.271. \tag{35b}$$

IX. Generalizations

In the previous analysis, we were concerned with a network on the plane, and we determined optimum transmission ranges only for the case of an inverse fourth power propagation loss law. It turns out that the results obtained for the interference statistics can be easily generalized to d-dimensional space, and similar results for the optimum transmission range can be obtained. Although only the additional cases of one dimensional and three dimensional have any physical significance, the following expressions hold for any dimension.

A. Interference Statistics

We assume, as previously, that we have an infinite network. The terminals are distributed according to a Poisson law on a d-dimensional space. The meaning here is that if V is the volume of a region $\boldsymbol{R}$ in d-dimensional space, then the number of terminals in $\boldsymbol{R}$ has the following distribution:

$$P[k \text{ in } \boldsymbol{R}] = \frac{e^{-\lambda V}(\lambda V)^k}{k!} \tag{36}$$

where λ is the average number of terminals per unit volume.

We may follow the same procedure that led to (15) where, instead of working with a sphere in two-dimensional space (i.e., a circle), we work with a sphere in d-dimensional space. The generalized version of (15) is

$$\phi_Y(\omega) = \exp\left(iK_d\lambda\omega \int_0^\infty [g^{-1}(t)]^d e^{i\omega t}\, dt\right) \tag{37}$$

where K_d is the volume of the unit sphere in d-dimensional space ($K_1 = 2$, $K_2 = \pi$, $K_3 = 4\pi/3$, etc.). Assuming the class of propagation laws given by (16), (37) becomes

$$\phi_Y(\omega) = \exp\left(-K_d\lambda\Gamma(1-\alpha)e^{-i\pi\alpha/2}\omega^{a}\right) \qquad \omega > 0 \tag{38}$$

where $\alpha = a/\gamma$. We note that for the interference to be finite, we must have $\gamma > d$.

The probability distribution function in the general case is given by (23) where $\rho = K_d\lambda\Gamma(1-\alpha)$.

B. Optimum Transmission Ranges

In a similar manner as for (32), we may calculate the expected forward progress and obtain

$$\begin{aligned}(\lambda)^{1/d}Z = {} & (N/K_d)^{1/d}(1-p)(1-e^{-p}) \\ & \cdot\left(1 - \frac{1}{\pi}\sum_{n=1}^{\infty}\frac{\Gamma(\alpha n)}{n!}\sin n\pi(1-\alpha)\right. \\ & \left.\cdot\int_0^{\mu_0}\left(\frac{pN\Gamma(1-\alpha)}{K(\mu)^{\alpha}}\right)^n s'(\mu)\, d\mu\right).\end{aligned} \tag{39}$$

As previously, Z is multiplied by $\lambda^{1/d}$ so as to make the expected forward progress dimensionless. Also, as pre-

TABLE I
SOME OF THE CONSTANTS IN (40)

	$C(d, \alpha)$				
d \ α	1/4	1/3	1/2	2/3	3/4
2	1.73	1.62	1.34	0.94	0.71
3	0.90	0.88	0.77	0.58	0.46

viously, we would like to find the optimum N as a function of K and for the optimum transmission probability p. We may optimize over the parameters p and N as before. For the case of a step coding function $s(\cdot)$, we summarize the results as follows: for a given d and α, the optimum transmission range N_0 is given by the relation

$$N_0(d, \alpha) = C(d, \alpha) K^{\alpha}(\mu_c). \tag{40}$$

As previously, N_0 is proportional to a power of the parameter $K(\mu_c)$. We give a few values of the constant C in Table I.

We summarize the above results. In obtaining the interference statistics, the important parameters are d, the dimension of the network, and γ, the exponent of the propagation loss law. Given these two parameters, we define $\alpha = d/\gamma$. The interference statistics are then given by the stable law of exponent α, and the optimum N_0 is proportional to $K^{\alpha}(\mu_c)$ where $K(\cdot)$ is defined as in (24) and Table I gives a few of the proportionality constants.

X. SUMMARY AND CONCLUSIONS

Multihop packet radio models used in the past have used the concept of a transmission radius where within a given radius of a transmitter, a packet has an "equal" probability of being received. This model essentially assumes a step function for the signal strength versus distance from the transmitter. In this paper, we have taken a new approach: a signal is assumed to decay in strength according to a gradually decreasing function of the distance from the transmitter. By assuming a random distribution of the terminals, we were able to obtain the statistics of the interference power from all other transmissions at a particular receiver. Assuming an inverse power law for the signal strength versus distance from the transmitter, we showed that the probability laws of the interference power are the stable laws with parameter α restricted to $0 < \alpha < 1$. The case of an inverse fourth power propagation law, which results from ground wave propagation, corresponds to the stable law with $\alpha = 1/2$, which has a density known as the inverse Gaussian probability density.

In a multihop packet radio network, there is usually a tradeoff between the distance covered in one hop and the probability of a successful transmission. Using the above interference analysis, we proceeded to obtain the optimum transmission range for a DSSS network. For a narrow-band packet radio network, previous results have concluded that the range of a transmission should be such that on the average there are approximately six–eight terminals closer to the transmitter than the receiver. For a spread spectrum system, we expect different results since multiple simultaneously successful transmissions are possible. In the spirit of the analysis of Kleinrock and Silvester [1]–[3], we have concluded that in a direct sequence spread spectrum network, the range of a transmission should be chosen so that on the average there are 1.33 $\sqrt{K(\mu_c)}$ terminals closer to the transmitter than the receiver where $K(\mu_c)$ is a parameter that is proportional to the processing gain, which can be interpreted as the *effective* maximum number of simultaneously successful transmissions possible in some region as a function of the effective SNR μ_c required for a particular coding scheme.

REFERENCES

[1] L. Kleinrock and J. A. Silvester, "Optimum transmission radii for packet radio networks or why six is a magic number," in *Conf. Rec., Nat. Telecommun. Conf.*, Dec. 1978, pp. 4.3.1–4.3.5.

[2] J. A. Silvester, "On the spatial capacity of packet radio networks," Dep. Comput. Sci., Univ. California, Los Angeles, Eng. Rep. UCLA-ENG-8021, May 1980.

[3] L. Kleinrock, and J. A. Silvester, "Spatial reuse in multihop packet radio networks," *Proc. IEEE*, vol. 75, pp. 116–134, Jan. 1987.

[4] H. Takagi and L. Kleinrock, "Optimal transmission ranges for randomly distributed packet radio terminals," *IEEE Trans. Commun.*, vol. COM-32, pp. 264–257, Mar. 1984.

[5] T. C. Hou and V. O. K. Li, "Transmission range control in multihop packet radio networks," *IEEE Trans. Commun.*, vol. COM-34, pp. 38–44, Jan. 1986.

[6] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications, Vol. I.* Rockville, MD: Computer Science Press, 1984.

[7] W. C. Peng, "Some communication jamming games," Ph.D. dissertation, Dep. Elec. Eng., Univ. Southern California, Los Angeles, Jan. 1986.

[8] L. G. Roberts, "ALOHA packet system with and without slots and capture," ARPA Network Inform. Cen., Stanford Res. Inst., Menlo Park, CA, ASS Note 8 (NIC 11290), June 1972; reprinted in *Comput. Commun. Rev.*, vol. 5, pp. 28–42, Apr. 1975.

[9] L. Fratta and D. Sant, "Some models of packet radio networks with capture," in *Proc. Int. Conf. Comput. Commun.*, Oct. 1980, pp. 155–161.

[10] R. Nelson and L. Kleinrock, "The spatial capacity of a slotted ALOHA multihop packet radio network with capture," *IEEE Trans. Commun.*, vol. COM-32, pp. 684–694, June 1984.

[11] E. S. Sousa and J. A. Silvester, "Determination of optimum transmission ranges in a multi-hop spread spectrum network," in *Proc. MILCOM*, Boston, MA, Oct. 1985, pp. 449–454.

[12] G. R. Cooper and R. W. Nettleton, "A spread-spectrum technique for high capacity mobile communications," *IEEE Trans. Vehic. Technol.*, vol. VT-27, Nov. 1978.

[13] S. A. Musa and W. Wasylkiwskyj, "Co-channel inference of spread spectrum systems in a multiple user environment," *IEEE Trans. Commun.*, vol. COM-26, pp. 1405–1413, Oct. 1978.

[14] C. L. Weber, G. K. Huth, and B. H. Batson, "Performance consideration of code division multiple-access systems," *IEEE Trans. Vehic. Technol.*, vol. VT-30, pp. 3–9, Feb. 1981.

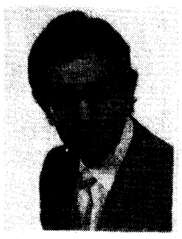
[15] O. DeSouza, P. Sen, and R. R. Boorstyn, "Performance analysis of spread spectrum packet radio networks," in *Proc. MILCOM*, San Diego, CA, Oct. 1988, pp. 599–604.

[16] M. B. Pursley, "Performance evaluation for phase-coded spread-spectrum multiple-access communication—Part I: System analysis," *IEEE Trans. Commun.*, vol. COM-25 pp. 795–799, Aug. 1977.

[17] J. J. Egli, "Radio propagation above 40 Mc over irregular terrain," *Proc. IRE*, pp. 1383–1391, Oct. 1957.

[18] W. Feller, *An Introduction to Probability Theory and Its Applications, Vol. II.* Wiley, 1966, pp. 554–583.

[19] E. S. Sousa and J. A. Silvester, "Spreading code protocols for distributed spread spectrum packet radio networks," *IEEE Trans. Commun.*, vol. 36, pp. 272–281, Mar. 1988.



Elvino S. Sousa (S'80–M'86) was born in the Azores, Portugal, on December 28, 1956. He received the B.A.Sc. degree in engineering science and the M.A.Sc. degree in electrical engineering from the University of Toronto, Toronto, Ont., Canada, in 1980 and 1982, respectively, and the Ph.D. degree in electrical engineering from the University of Southern California degree in 1985.

He was a Teaching Assistant at the University of Toronto, a Research Assistant at the University of Southern California, and is currently with the Department of Electrical Engineering, University of Toronto. He has performed research in the areas of data transmission and spread spectrum packet radio networks. He is also interested in optical communications and the performance of distributed computer systems, on which he has consulted for Technology Transfer Institute. In the past three years at the University of Toronto, he has taught graduate courses in error-correcting codes and mobile communications.

Dr. Sousa is a member of Etta Kappa Nu and the Association of Professional Engineers of Ontario. He has held a Natural Sciences and Engineering Research Council of Canada (NSERC) Postgraduate Scholarship, was a recipient of the IEEE Communications Society Student Scholarship, and is currently an NSERC University Research Fellow.

John A. Silvester (M'79–SM'85) was born in Kent, England, in 1950. He received the B.A. and M.A. degrees in mathematics and operations research from the University of Cambridge, Cambridge, England, in 1971 and 1975, the M.S. degree in statistics and computer science from West Virginia University, Morgantown, in 1973, and the Ph.D. degree in computer science from the University of California, Los Angeles, in 1980.

Since 1979 he has been at the University of Southern California, Los Angeles, where he is an Associate Professor in the Department of Electrical Engineering and Director of the Computer Engineering Division. He teaches courses in computer networks, queueing theory, computer system performance evaluation, and computer architecture. His current research interests are the performance evaluation and modeling of computer communication networks and distributed systems, packet radio and spread spectrum radio networks, high-speed networks, local area networks, and multiple-access protocols. He is the author of over 60 technical papers and has lectured both in the United States and abroad. He has consulted for many of the leading networking companies, the U.S. Army, and was Director of Consulting for Technology Transfer Institute from 1984 to 1986 where he is still a leading consultant. He also teaches short courses on packet radio, network design tools, and data communications fundamentals.

Dr. Silvester is a member of the ACM and has been active in the Technical Committee on Computer Communications of the IEEE Communications Society since 1981, acting as Conference Coordinator from 1983 to 1985 and Chairman from 1985 to 1987. He was Program Chairman of the First IEEE Computer Communications Workshop (1986), is on the INFOCOM conference board, was Vice Chairman of INFOCOM'89, and General Chair for INFOCOM'90. He has served on the Program Committees of INFOCOM, the Data Communications Symposium, and the International Conference on Computer Communications.

Microcellular Direct-Sequence Spread-Spectrum Radio System Using *N*-Path RAKE Receiver

URS GROB, MEMBER, IEEE, ARNOLD L. WELTI, STUDENT MEMBER, IEEE, ERNST ZOLLINGER, MEMBER, IEEE, ROLAND KÜNG, MEMBER, IEEE, AND HANS KAUFMANN, MEMBER, IEEE

Abstract—**A microcellular local area network (LAN) for indoor communications is proposed using code division multiple access (CDMA) and DPSK for data modulation. The pseudonoise (PN) codes in the transmitters of the base station are mutually synchronized. For this purpose, sets of Gold code sequences having low cross correlation have been found by an exhaustive computer search. Together with wide-band measurements of the indoor radio channel at 900 MHz, a five-path RAKE receiver was designed to combat fading effects and to process the time diversity by using multipath signal reception. Each receiver path is demodulated independently. Several methods of diversity combining of these paths have been investigated. Acquisition and tracking of the spreading code in the receiver are controlled by a digital signal processor (DSP). Experimental results of the CDMA system are presented, showing the behavior in a multipath environment.**

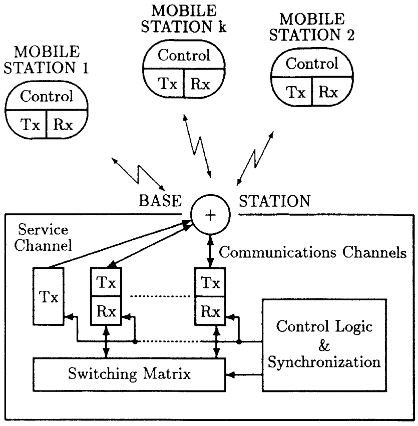


Fig. 1. Spread-spectrum radio system for wireless indoor communications (*Tx*: transmitter, *Rx*: receiver).

I. Introduction

We have investigated and realized a direct-sequence spread-spectrum (DSSS) microcellular system for indoor radio communications [1], [2]. A microcell consists of *k* mobile users who can communicate with each other through a base station as shown in Fig. 1.

The base station is able to provide for several simultaneous bidirectional links to the mobile stations using code division multiple access (CDMA). The system uses differential phase-shift keying (DPSK) modulation for ease of implementation and a RAKE receiver to reduce multipath fading effects [3]. The theoretical performance of such a system has neen analyzed in [4]–[6]. In this paper, we describe the practical implementation and experimental results of parts of our system.

In a CDMA system, the number of simultaneous users depends on the cross-correlation properties of the spreading codes. In [5], [6], the codes were optimized only with respect to low sidelobes of their autocorrelation function (ACF) for good synchronization behavior. In a star-connected system like ours, chip-synchronous processing at the base station makes it also possible to optimize the codes for very low values of the even and odd cross-correlation function (CCF) between individual codes. Therefore, we use codes with optimized ACF and CCF, resulting in an interference term at the mobile stations which is no longer random, but has a deterministic and low value.

To serve a big number of users over a large area, several microcells can be combined in a cellular arrangement. Each cell is assigned a set of codes for a number of simultaneous communication links, whereas adjacent cells use different code sets.

The code length $L = 1023$ chips at a rate $R_c = 16 \times 10^6$ chip/s was chosen to accommodate a large enough number of users per cell. With these parameters, it is not possible to realize matched filters or correlators using surface acoustic wave (SAW) or charge-coupled device (CCD) technology because of the restricted number of taps, the device length, or the chip rate [2], [7]. Therefore, the RAKE receiver was realized with a bank of seven active correlators which do not suffer from these limitations.

In Section II, the typical indoor channel characteristics as they have been measured in Swiss buildings are described. Section III examines the rules for designing the microcellular system and gives optimization criteria for choosing the code sequences. Section IV describes the transmitter and the receiver design. Special emphasis is given to the desciption of the acquisition and tracking strategies. Section V provides some experimental results, followed by a summary and conclusions in Section VI.

Manuscript received February 24, 1989; revised December 8, 1989.

U. Grob was with the Institute for Communication Technology, Swiss Federal Institute of Technology, CH-8092 Zürich, Switzerland. He is now with the European Telecommunications Standards Institute (ETSI), Sophia Antipolis, 06561 Valbonne, France.

A. L. Welti and E. Zollinger are with the Institute for Communication Technology, Swiss Federal Institute of Technology, CH-8092 Zürich, Switzerland.

R. Küng and H. Kaufmann are with the Research Department, Ascom Zelcom AG, CH-8634, Hombrechtikon, Switzerland.

IEEE Log Number 9035805.

0733-8716/90/0600-0772$01.00 © 1990 IEEE

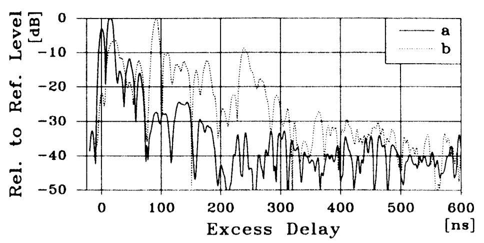


Fig. 2. Magnitude of the complex envelope of the impulse response for (a) a line-of-sight and (b) an obstructed wave propagation path in the laboratory buildings.

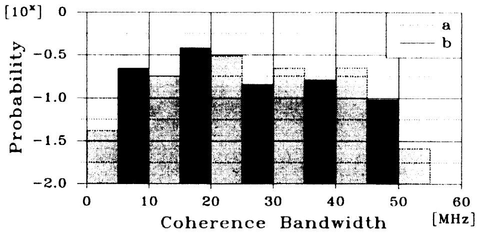


Fig. 3. Probability of the coherence bandwidth falling into an interval of 10 MHz (a) for the line-of-sight and (b) for the obstructed wave propagation path in the laboratory buildings.

II. Indoor Radio Channel Characteristics

Among the considerable literature on radio wave propagation in general, some recent papers concentrate on the characteristics of the indoor radio channel [8]–[10]. The statistics of the mean excess delay, the rms delay spread, and the propagation path loss have been measured in many indoor situations, but only a few report measurements of the coherence bandwidth. Therefore, this section concentrates only on the statistics of the coherence bandwidth and the rms delay spread at a carrier of 900 MHz in our laboratory buildings.

We measured the complex envelope of the time-variant impulse response at several distinct locations. Two vertical polarized conical λ/4-monopole antennas were used at the transmitter and receiver sites. A PN sequence of 127 chips was radiated periodically every 1.4 μs with a power of +30 dBm and a bandwidth of 360 MHz. We obtained a delay resolution of 11 ns.

The multipath signals, which are depicted in Fig. 2 for two typical situations, cause frequency selective fading seen as notches.

A possible characterization of the average width of these notches is the coherence bandwidth W_c. It is the minimal necessary frequency separation of two harmonic signals such that the correlation of the two amplitudes decreases below a given value (here, 0.5). Fig. 3 shows the probability that W_c falls into the specified intervals of 10 MHz.

To avoid large variations in the propagation path loss in the obstructed paths case, the signal bandwidth has to be at least 40 MHz, as shown in Fig. 3(b), in order to cover 90% of the situations. Because of practical purposes, we have chosen a signal bandwidth of 32 MHz, resulting in a coverage of 77% of the W_c's observed.

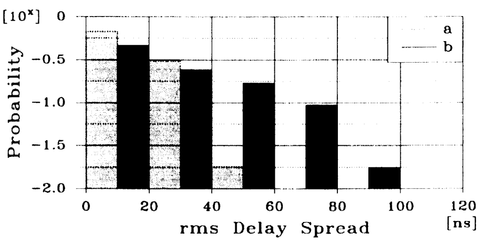


Fig. 4. Probability of the rms delay spread falling into an interval of 20 ns (a) for line-of-sight and (b) for the obstructed wave propagation path in the laboratory buildings.

In the time domain, the excess delay time of the multipath signal components are expected to be mainly in the range of ≈ 2× the maximal rms delay spread (2 × 100 ns, as shown in Fig. 4).

Our RAKE receiver is designed such that this delay range is taken care of.

III. Microcellular System

We consider an indoor radio system using DSSS with code division multiple access organized in a microcellular structure with one base station per cell. Each cell uses a different set of spreading codes. If the code sequences are multiplied by the data bit value and transmitted exactly once per bit, the interference from other CDMA links can be described by the even and odd cross-correlation function (CCF) [11], [12] of their spreading code sequences.

First, we consider the link from the base station (BS) to the mobile station (MS). If we use chip- and bit-synchronous operation at the BS and assuming an ideal channel, the interferences from other links at any MS only depend on the CCF values at zero delay [5]. For indoor multipath channels, a small region around this origin becomes important. Assuming stationary conditions, these interferences can be minimized for optimal CDMA performance. Therefore, in contrast to [5], we started a code search with the goal to find 20 code sets having optimal CCF's around the origin between all codes of the same set.

The communications protocol is organized as follows. The base station continuously transmits synchronization and signaling information by a special sequence, the service channel. All mobile stations are permanently synchronized to this channel in the idle state. For data transmission, they can switch to a communications sequence without loss of synchronism. The mobile stations use their locally generated despreading sequence for spreading the transmitted data.

The cell radius is very small in an indoor microcellular system, e.g., about 50 m. Therefore, at the base station the maximum time difference between the transmitted and received PN sequence is only twice the propagation time, e.g., 330 ns. This fact enables the base station to synchronize to the received code in a very short time, even at high processing gains. With our code rate $R_c = 16 \times 10^6$ chips/s and code length $L = 1023$, the acquisition circuit has to search within six code chips only. There-

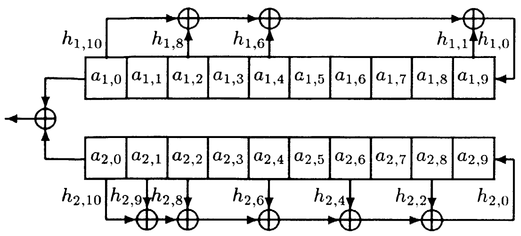


Fig. 5. LFSR to generate the Gold sequences with $h_1 = 3025$ and $h_2 = 2527$.

TABLE I
INITIAL STATES OF THE SHIFT REGISTER WITH $h_1 = 3025$ AND $h_2 = 2527$ (OCTAL NOTATION)

a_1	a_2	a_1	a_2	a_1	a_2	a_1	a_2	a_1	a_2
0700	0000	0000	1060	1127	0645	0756	0077	1473	1443
1454	0643	1043	0346	1363	1177	1177	1740	1021	0355
1075	1216	1614	1027	0746	0751	1505	1373	1514	1545
0556	1571	0740	0037	0076	1577	1546	0644	1363	1444
1324	1145	0315	0215	1677	1554	0347	1203	1622	1350
1136	0774	1252	0242	0550	1114	0256	0662	1407	0176
1407	1564	0073	0723	1104	0765	0116	1561	1236	1360
0761	0714	1350	0562	0765	0076	1355	1023	0732	1221
1077	0115	0751	0324	0642	1277	0412	1566	1727	0556
0040	0207	0345	0205	0260	1640	1424	0542	0737	0562
0236	1700	0171	1436	0041	1431	1122	0523	1436	1674
0703	1164	0200	1243	0726	0177	1664	1677	0276	1657
0030	0764	1245	1030	1524	1263	1765	1654	0261	1162
0241	0256	0744	0733	1033	1530	0116	1152	0053	1255
1555	1002	0765	0416	0054	1662	0724	0524	1240	1621
1304	0367	1531	1257	1104	0155	0571	1447	0732	1755
0256	0166	0133	0374	0720	0623	0515	1163	1355	1504
1675	0764	0011	0064	1147	1213	1653	1732	0545	0451
1024	1105	1646	0117	0052	1705	0442	1606	1774	0247
1273	0740	0551	1324	1352	0552	0742	0766	1661	1156
1302	0311	0750	0532	1772	1120	1136	0504	1146	0716
0543	1340	0532	0250	1072	1023	1405	1472	0662	0573
1013	1224	1202	1304	0545	0112	0636	0705	0730	0431
0373	1335	1050	1056	0542	1655	1164	1425	0553	1242
1156	1362	0351	1772	1175	1311	1744	0565	0142	1103
1325	0743	0536	0064	1505	1323	0672	0703	0554	0021
0232	1365	0475	0072	1472	0636	1714	1327	1436	0605
1633	0230	0013	0336	1271	1764	1527	1151	1455	0601
1563	0543	0075	0655	0467	0076	1104	1777	1443	0121
0303	0606	1132	1744	1220	1455	0113	1606	1756	0334
1714	0001	1450	0205	0317	1412	0335	0455	1017	0160
0242	1270	0760	0564	1116	1210	1545	0671	1154	0172
0312	1403	1137	0124	1273	0731	0157	0051	0550	0276
0575	0051	0402	0307	0745	0241	0442	0466	1310	1705
1322	0540	1736	0234	0661	1376	0476	0257	1604	0156
1075	1146	0037	1434	0551	1215	0436	1705	0216	1354
0642	1477	0342	1112	0333	1275	0115	1223	1157	1417
1246	1370	1531	1102	0323	1443	0276	0605	1355	1541
1303	0244	0520	1163	1264	1700	1310	0673	0266	1472
0704	0714	0274	0352	1203	1402	1267	1562	0740	1270
0420	1677	1410	0060	1271	1074	0173	0061	0715	0637
0674	0667	0054	1041	1446	0451	0725	1173	1275	1067
1244	0250	0357	1226	0376	0536	1027	1124	0575	0736
0001	1407	0321	0676	1617	1236	1522	0417	0103	0242
1126	0231	0713	0127	0414	0060	0047	1437	1204	1420
0532	0350	0667	0163	1243	0344	0063	0424	0030	0114
1124	1452	0705	0105	0317	0662	0114	0672	1452	1631
1066	1175	1324	0740	1210	0072	1442	0665	1643	1216
1445	1610	0672	1341	0214	1117	1471	1147	1561	0211
1712	0640	1626	1523	1010	0211	1070	0276	1027	1370
0454	1631	1177	0053	0356	1526	1716	0361	1122	1746
1632	1703	1623	1711	0145	1743	1053	0602	1303	1435
0137	1771	0656	1244	1142	0372	1557	1221	1720	1504
0526	1110	0222	1453	0030	0503	1605	1041	1554	0256
0725	0315	1742	1046	1331	0734	0202	1300	1367	1425
0561	1301	0455	0076	1533	1230	0142	1106	0475	0246
1475	1613	0747	0351	0300	0323	1535	0713	1707	1070
1225	0147	1352	0033	0044	1257	1272	1146	1557	0603
1042	0262	0420	0607	0452	0513	0633	0345	0051	0167
0664	1655	0754	0630	0606	1032	0451	0714	1651	1532
0702	1030	0764	1313	0470	0356	0742	0517	1360	1202
1231	0463	0220	0157	0730	1121	0343	0030	0654	1674
0514	0066	0356	0337	0237	0034	1227	0432	1126	0576
1174	1337	0176	1352	1023	0467	1724	0747	0717	1060
1301	1435	1760	0466	1547	1367	0212	0640	1415	0163
0033	0656	1060	1271	0010	1031	1705	1164	1700	0764
0770	0706	0453	1163	0660	0675	0114	1764	0406	1442
0327	1464	1315	1745	0254	0416	1471	0570	1771	0772

fore, the link from the mobile to the base station partly benefits from the low CCF values around the origin.

Kasami and Gold sequences of length $L = 1023$ have been evaluated as spreading functions. The even CCF of these sequences, which is defined as

$$E_{ij}(n) = \sum_{l=0}^{L-1} c_i[l]\, c_j[(l+n) \bmod L] \qquad (1)$$

where $c_i[l]$ is the lth code chip of the sequence c_i, only takes on values out of the set $\{-65, -33, -1, 31, 63\}$. The class of Gold sequences consists of 1025 members per family, but only 769 are balanced. For synchronization purposes, it is advantageous that the autocorrelation function (ACF) takes on the same low value ($E_{ii}(1) = -1$) near the origin. Selecting the codes according to this condition, 577 sequences are left.

To optimize the CCF around the origin, the even CCF $E_{ij}(n)$ of two sequences c_i and c_j from the same set should fulfill the condition

$$E_{ij}(n) = -1; \qquad n \in \{-1, 0, 1\}. \qquad (2)$$

This condition implies that the absolute value of the odd CCF over the same range will be less than three. It can be shown that no advantage results in taking the large set of Kasami sequences for this kind of optimization [13]. The size of the code sets will be maximum when searching within one Gold sequence family only.

All 90 different families of Gold sequences were examined to find the family containing the 20 best of all balanced sequences which are autooptimal and have the least sidelobe energy (AO/LSE) [14]. These sequences will be used for the service channel in each cell because they perform best with respect to initial synchronization. The feedback polynomials obtained for the Gold sequence generator are (octal notation) $h_1 = 3025$ and $h_2 = 2527$, Fig. 5.

By an exhaustive search, 17 was found to be the greatest possible number of sequences per set. Table I shows the corresponding initial states of the shift registers for the 20 cells, the first line of each set being AO/LSE phase.

IV. Experimental System

An experimental system was designed to measure the behavior of a spread spectrum system in a microcellular indoor environment with CDMA, as well as to study various system parameters. Two identical transmitters and receivers (Fig. 6) have been built for this purpose.

A. *Transmitter Design*

Each transmitter (Fig. 7) consists of a differential encoder, a pseudonoise (PN) code generator, a binary phase-shift keying (BPSK) modulator, a radio frequency (RF) converter, and an amplifier.

The incoming data bit stream with a rate $R_b = 16 \times 10^3$ b/s is differentially encoded. The spreading code ($R_c = 16.368 \times 10^6$ chips/s) is multiplied with the information bit and fed to the BPSK modulator that operates at an intermediate frequency (IF) of 70 MHz. After filtering, the modulated signal is up-converted to an RF carrier frequency of 959 MHz and passed through a bandpass filter with 26 MHz bandwidth. The amplified signal with a power level of about 10 mW is radiated by a vertical broad-band λ/4-monopole antenna.

To generate the different Gold sequences, a PN-code generator was implemented in a CMOS gate array. It contains two programmable 12-stage linear feedback shift

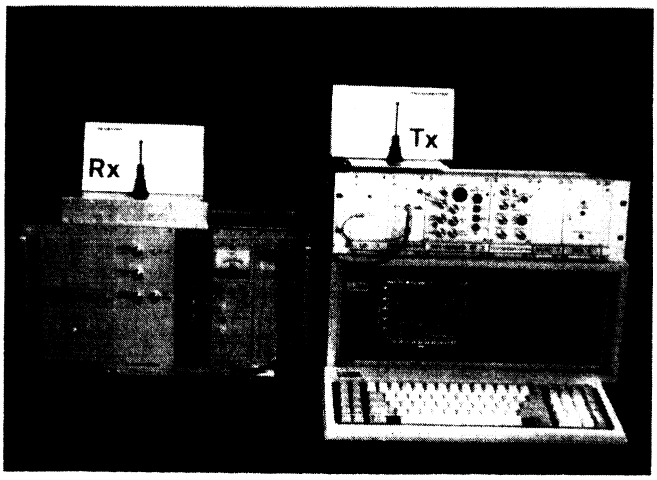


Fig. 6. Photograph of spread-spectrum receiver (left), transmitter (right), and personal computer for displaying the measured data.

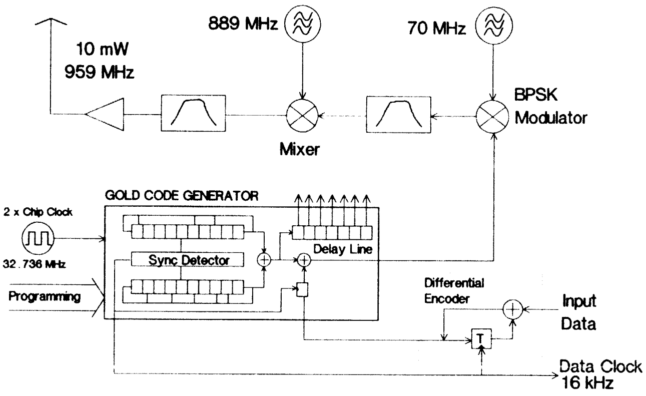


Fig. 7. Block diagram of the spread-spectrum transmitter.

registers whose outputs are modulo-2 added. The initial values of the two shift registers can be arbitrarily chosen. For despreading in a RAKE receiver, a digital delay line delivers seven time-shifted versions of this code. The data bits are synchronized to the spreading code by a short pulse which is generated at the beginning of every code period by a built-in synch detector.

B. Receiver Design

The receiver design is based on the RAKE principle. We have chosen five diversity paths spaced one chip (61 ns) apart. Thus, all received signals with delays falling into a window of about 300 ns can be processed. This seems to be a reasonable compromise between receiver complexity and performance in multipath environments. The block diagram of the receiver is shown in Fig. 8.

The RF down-converter consists of bandpass filters, amplifiers with automatic gain control (AGC), a local oscillator, and a mixer to convert the received signal to a first IF at 70 MHz. The signal is then fed to the seven despreaders of the RAKE receiver. Their output signals are passed through crystal bandpass filters with 34 kHz bandwidth and then coverted to a second IF at 460 kHz.

The digital delay line in the code generator provides the delayed samples of the required Gold code. Each signal from the taps located at -2, -1, 0, $+1$, and $+2$ chips delay, respectively, is up-converted to a third IF of 80.7 MHz. For synchronization purposes only, taps with delays of -0.5 and $+0.5$ chips are provided. These signals serve as references for despreading.

The signal strength of each despreader output is obtained by the envelope detectors (ED). These detector output voltages are integrated and A/D converted. A digital signal processor (DSP) controls the acquisition of the received code by adjusting the local code generator in steps of 0.5 chip. A tracking loop with an adaptive discriminator characteristic and loop gain tracks the code. The acquisition and tracking strategies will be described in the next paragraph.

The signals with relative delays of -2, -1, 0, $+1$, and $+2$ chips are used for differential data demodulation. This procedure eliminates the need for a carrier recovery circuit for each RAKE arm, but requires the data bit duration $T_b = 62.5$ μs to be an integer multiple of the IF carrier cycle. With an RF-oscillator stability of 1-2 ppm, the IF frequency may differ from the ideal value in a range up to ± 3 kHz. The resulting loss of up to 8 dB can be avoided by choosing an I-Q structure of the DPSK demodulator.

The input signal of each arm and its 90° phase-shifted version are multiplied with the hard-limited and delayed signal of the preceding bit interval. This delay of T_b is achieved by a charge-coupled device (CCD) delay line. The baseband signal of all five in-phase (I) channels are integrated over $T_i = 0.8$ T_b and summed. The quadrature (Q) channels are processed in the same way. The decision logic determines the bit value from the sums of the I and Q channels. The weight of each channel is adjusted by the DSP using a variable attenuator at the demodulator input.

C. Synchronization

The code synchronization in the receiver is attained in two steps: by code acquisition, and by code tracking. The acquisition is based on the measuring of the magnitude of the complex envelope of the channel impulse response. Because of the RAKE-type structure, the receiver processes any signal component that falls into a window of width W and uses it for data demodulation. If the receiver has an odd number of N despreading correlator paths with a delay of one code chip between them, then the receiver window has a normalized width of $W = (N - 1)$. Fig. 9 shows the acquisition procedure for $N = 5$.

The receiver window $w(n)$ of length $W = 4$ is given in Fig. 9(a). The magnitude of the impulse response $|x(n)|$ is shown in Fig. 9(b). The distribution $y(k)$ of the summed signal magnitudes that fall into the receiver window is plotted as a function of the window position relative to the impulse response in the Fig. 9(c). This function can approximately be measured by sliding the local PN-code generator with respect to the received signal. Sliding can be accomplished in two ways, either the clocks of the PN-

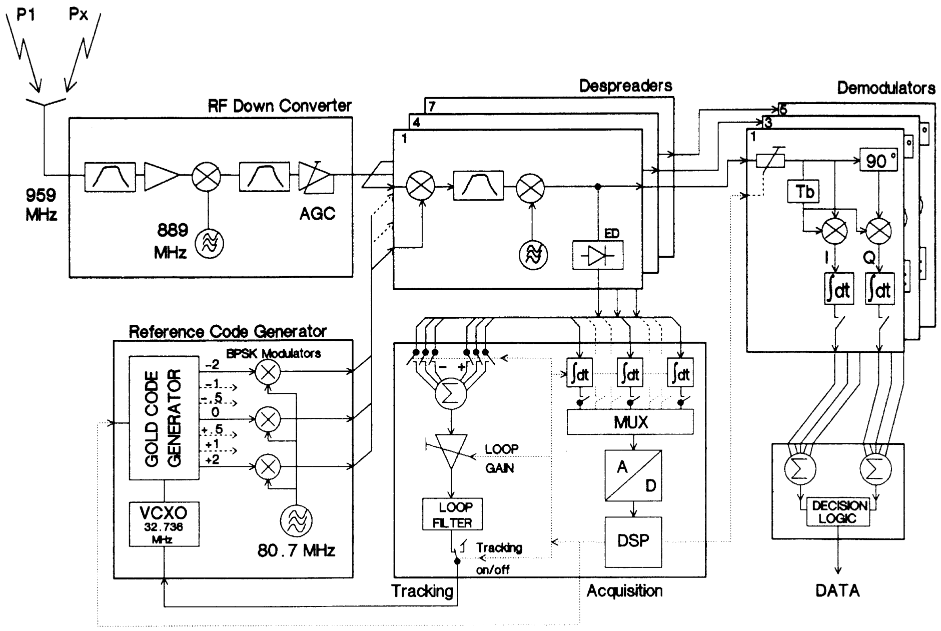


Fig. 8. Block diagram of the five-path RAKE receiver.

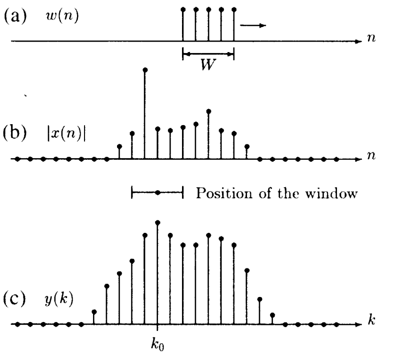


Fig. 9. Acquisition algorithm. (a) Receiver window. (b) Magnitude of the complex envelope of the channel impulse response. (c) Distribution of the summed signal magnitudes that fall into the receiver window.

code generators have a constant frequency offset (continuous sliding correlation) or the code of the local generator can be shifted in discrete-time intervals ϵ with respect to the code in the transmitter (discrete sliding correlation). After despreading, the second solution represents a time-discrete version of the impulse response of the channel. The timing resolution depends on the time shift ϵ, which is assumed to be one code chip in Fig. 9. During the sliding procedure, $y(k)$ is calculated at every location k of the window. After sliding over one period of the PN code, a maximum at some position k_0 will be found. The receiver window is then centered around this position by shifting the local code, and the code tracking loop is switched on. For better noise suppression, the outputs of the envelope detectors can be integrated during a variable time interval.

After acquisition, it is necessary to verify that the real maximum was detected. Our verification strategy averages the signal magnitude outside the receiver window and then compares it to the average signal magnitude inside the window after one more integration interval. Averaging over several intervals will increase the robustness of the verification algorithm.

The acquisition algorithm can be described mathematically by a correlation between the receiver window $w(n) = \mathrm{rect}\,(n/W)$ and the magnitude of the complex envelope of the channel impulse response $|x(n)|$:

$$\begin{aligned} y(k) &= \sum_{n=-\infty}^{\infty} |x(n)| \ \mathrm{rect}\left(\frac{n+k}{W}\right) \\ &= \sum_{n=k-W/2}^{k+W/2} |x(n)| \qquad (3) \end{aligned}$$

$$y(k)\big|_{\max} \Rightarrow k_0. \qquad (4)$$

This algorithm is implemented by using a digital signal processor.

A tracking loop similar to the classical delay-lock loop (DLL) [15] is used to perform code tracking, except that the discriminator characteristic (or S curve) is expanded. The receiver has N despreaders which can be used to make the S curve as wide as the receiver window [15]. The various correlator paths can be switched on and off to change the characteristic. This allows us to adapt the loop behavior to the time-varying impulse response of the channel. In the case of line-of-sight transmission, a narrow S curve is advantageous since the outer correlator paths only process noise. If there is no line-of-sight condition and several delay paths are present, a broad S curve makes the tracking loop more robust since there is more signal power available in the loop.

Calculating the mean time to lose lock (MTLL) of tracking loops shows that their performance degradates very rapidly in the presence of long-term frequency instability (code clock frequency mismatch) [16]. It has been shown that the loop bandwidth can be optimized. There is also a large degradation in the MTLL if the bandwidth is not properly chosen. The loop bandwidth can be adjusted in the realized system by means of the signal processor.

While tracking, the signal processor is continuously sampling the output signal of the despreaders. Analyzing the power of each path allows for weight adjustments of the demodulator input signals and the implementation of an out-of-lock detector.

During acquisition at low signal-to-noise conditions, i.e., when it is difficult to detect the maximum of $y(k)$, several impulse responses can be low-pass filtered. That may give a detectable maximum. A similar result is obtained by extending the integration time. The DSP adaptively controls this duration using the signal-to-noise ratio measured during the acquisition procedure.

D. Diversity Combining

The RAKE receiver structure provides five time-diversity signals from different propagation paths which are combined after demodulation. Because of the hard-limited reference signal, the demodulators (Fig. 8) have a linear input/output characteristic. The DSP controls the weightings of the individual channels in 3 dB steps according to the signal strength which is extracted by the envelope detectors. Basically, two effects influence the transmission quality: interferences and multipath propagation. Controlling the attenuators with respect to power level (signal and noise power together) and summing several paths to accomplish the decision causes combining losses as soon as interference is present. Assuming ideal conditions for propagation, but very strong interference, a simple single-path receiver is optimum. This receiver type is simulated by providing maximum attenuation to all diversity signals except the strongest one. The other extreme is the situation where no interference influences the transmission, but the radio channel shows a wide multipath profile. In this case, the optimum receiver takes the decision from the weighted sum of all signals.

Five different strategies for diversity combining have been worked out and experimentally compared.

1) Only the signal of the zero chip delay path is demodulated (single-path receiver).

2) Only the strongest signal is taken (best-of receiver).

3) The two strongest signals are added (diversity combining).

4) Signals with a level $> 30\%$ of the strongest are added (diversity combining).

5) Normalizing the strongest signal level to 1, the other signals are attenuated in function of their levels:

- 3 dB for levels 0.4–0.7
- 6 dB for levels 0.25–0.4
- maximum attenuation for levels < 0.25 (approximation of maximum ratio combining).

V. Experimental Results

A. Receiver Performance with AWGN Channel

The overall performance of the receiver with an additive white Gaussian noise (AWGN) channel is specified by the probability of error (P_e) as a function of the signal-to-noise ratio E_b/N_0. The measurements included different frequency offsets with respect to the optimum IF carrier frequency for the chosen CCD delay line (Fig. 10).

The implementation loss compared to the theoretical performance of DPSK is 4 dB at the nominal IF carrier frequency. This loss results from several compromises made to keep the five demodulators as simple as possible. The reduced integration time of 0.8 T_b, the use of a bandpass filter instead of a matched filter in the despreader, and the hard-limited reference signal in the demodulator each contribute about 1 dB loss. At low E_b/N_0, the performance is even worse due to offset voltages. Deviations from the nominal IF frequency result in additional 3 dB loss due to the simple quantization of the decision bounds (45° sectors in the I/Q plane).

B. Acquisition and Tracking with Multipath Propagation

Measurements have been made to verify the proper operation of the acquisition algorithm and tracking loop. For this purpose, we moved the transmitter along several paths in the laboratory building. Therefore, the receiver had to cope with the time-varying multipath profile. Situations were encountered where line-of-sight transmission was dominant. In some other extreme cases, up to four strong reflection paths appeared in the profile. A second identical receiver with special acquisition software was operated in parallel to the data receiver. It continuously scanned the magnitude of the impluse response. This receiver is capable of observing the profile in a window of 20 chips (1.2 μs), while the data receiver tracks only paths within a five chip window.

In Fig. 11 (a), we recognize a typical profile measured by the receiver configuration mentioned above.

The shaded area shows the position of the demodulator taps in the data receiver. The tracking loop is able to process both of the appearing strong echos because of its wide S curve.

By processing the two strongest paths instead of only one, more energy is made available to the demodulator. For this situation, the combining strategy 3) (see Section IV-D) for demodulation resulted in satisfactory low probability of error.

Fig. 11(b) and (c) show extreme multipath situations. The delay between different paths is larger than the window of the demodulator. Fig. 11(b) shows three strong echos fed to the demodulator (shaded area). Moving the transmitter 15 cm resulted in a change of profile such that only one path remained in the range of the demodulator window, Fig. 11(c). This situation indicates that scanning a wider part of the multipath delay profile will be useful.

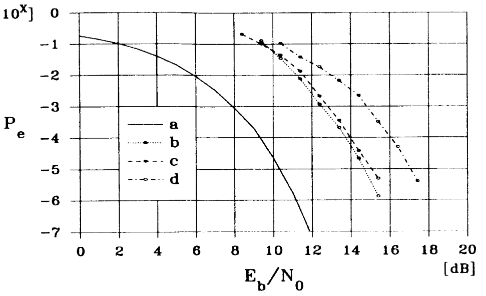


Fig. 10. Measured probability of error for experimental DPSK demodulator with AWGN channel and pseudorandom data. (a) Theoretical DPSK. (b) Frequency offset 0 kHz. (c) Frequency offset 1 kHz. (d) Frequency offset 2 kHz.

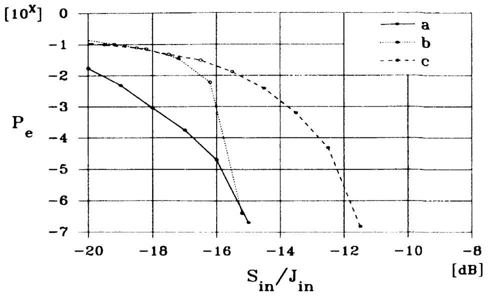


Fig. 12. Measured probability of error versus signal-to-interference ratio at the receiver input for (a) one user and AWGN, (b) two users, synchronous CDMA, (c) two users, asynchronous CDMA.

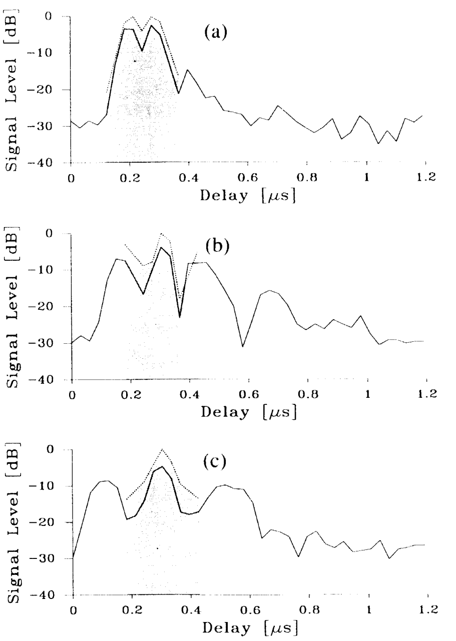


Fig. 11. Tracking in multipath situations showing demodulator window (shaded) and multipath profile. (a) Typical situation. (b), (c) Extreme cases.

C. Synchronous and Asynchronous CDMA

The CDMA environment was investigated by using two transmitters operating with different code sequences. During asynchronous operation, both transmitters used their own free-running chip clocks. In the synchronous case, the chip and data clocks of all base station transmitters were synchronized for exploiting the optimized CCF properties as described in Section III. The analysis and performance of a DPSK matched filter receiver is given in [5]. In our receiver, we use a correlator with a bandpass filter which results in a somewhat different analysis.

For a certain pair of code sequences, we therefore estimated the signal and interference power in the frequency domain using the Wiener–Khintchine theorem for calculating the spectral components. The result shows that additional frequency components from the interfering link pass the bandpass filter and add to the noise obtained by a matched filter receiver. Without data modulation at the transmitters, the signal-to-jamming ratio S_d/J_d at the DPSK detector input of the experimental system becomes

$$\frac{S_d}{J_d} = \frac{P_1}{P_2} \cdot \frac{L}{2} \tag{5}$$

with P_1 and P_2 denoting the power of the data and interfering link, respectively. Therefore, the jammer suppression ratio is on the order of the processing gain:

$$\frac{J_{in}}{J_d} = \frac{2}{L} = \frac{2}{1023} \quad \text{or} \quad \left(\frac{J_{in}}{J_d}\right)_{\text{dB}} = -27 \text{ dB}. \tag{6}$$

For a reasonable probability of error $P_e = 10^{-4}$, our system needs a signal-to-noise ratio $S/N = 11$ dB (i.e., $E_b/N_0 = 14$ dB) in Fig. 10. With (5) and (6), the required signal-to-jamming ratio at the receiver input is estimated as

$$\frac{S_{in}}{J_{in}} = \frac{P_1}{P_2} = -16 \text{ dB} \qquad \text{at } P_e = 10^{-4}. \tag{7}$$

In Fig. 12, the probability of error is plotted as a function of the signal-to-jamming ratio at the receiver input.

If there is only one synchronous interfering user, the jamming magnitude is deterministic and not Gaussian. Therefore, a threshold effect occurs. At low S/J, the despreaded signal magnitude is reduced because the AGC is mainly acting on the jammer. In the present receiver, no AGC is provided in the demodulator. Therefore, its dynamic range limits the maximum usable jammer suppression ratio.

In contrast to [5], the performance of asynchronous CDMA of our system is worse than synchronous operation. This behavior is because of the optimized CCF properties of the code sequences for synchronous CDMA which are useless in the other case. The asynchronous performance is determined by cross-correlation peaks at any possible code chip offset due to sliding of the user codes. From the system point of view, it is even more severe that at low signal-to-jamming ratio, loss of lock in the tracking loop arises whenever local maxima of the CCF occur. In the synchronous case, the loop remains locked even at the lowest allowable S/J.

VI. Summary and Conclusions

We have described the design of a microcellular spread spectrum radio system using CDMA. The system bandwidth and the design parameters of the code tracking loop were derived from measurements of the indoor radio channel. Field tests confirmed the proper choice of these parameters. A RAKE-type receiver structure with five demodulator paths was implemented to mitigate the multipath fading effects. The receiver complexity was reduced by using DPSK modulation. An *I–Q* structure of the demodulators minimizes losses due to carrier frequency offsets.

For an optimal performance of a CDMA system, a synchronous operation of the base station in each cell was suggested. In this case, the synchronous timing operation between individual users allowed optimization of the spreading codes with respect to a low CCF near the origin. This resulted in lower probabilities of error and in more robust tracking behavior than in the asynchronous case.

A digital signal processor adaptively controlled the acquisition and tracking of the received signal as well as the weighting of the demodulator paths. Measurements proved the advantage of the realized receiver structure under various multipath propagation conditions.

The realization of a receiver with a more flexible reference code generator will be further investigated. The aim is to reduce the number of demodulator paths without losing performance quality.

Acknowledgment

The authors are very grateful to P. Straub for implementing the synchronization part.

References

[1] K. Pahlavan, "Wireless communications for office information networks," *IEEE Commun. Mag.*, vol. 23, pp. 19–27, June 1985.

[2] M. Kavehrad and G. E. Bodeep, "Design and experimental results for a direct sequence spread spectrum radio using DPSK modulation for indoor, wireless communications," *IEEE J. Select. Areas Commun.*, vol. SAC-5, pp. 815–23, June 1987.

[3] J. S. Lehnert and M. B. Pursley, "Multipath diversity reception of spread-spectrum multiple-access communications," *IEEE Trans. Commun.*, vol. COM-35, pp. 1189–1198, Nov. 1987.

[4] G. L. Turin, "The effects of multipath and fading on the performance of direct-sequence CDMA systems," *IEEE J. Select. Areas Commun.*, vol. SAC-2, pp. 597–603, July 1984.

[5] E. A. Geraniotis, "Performance of noncoherent direct-sequence spread-spectrum multiple-access communications," *IEEE J. Select. Areas Commun.*, vol. SAC-3, pp. 687–694, Sept. 1985.

[6] H. Xiang, "Binary code-division multiple-access systems operating in multipath fading, noisy channels," *IEEE Trans. Commun.*, vol. COM-33, pp. 775–784, Aug. 1985.

[7] S. S. Rappaport and D. M. Grieco, "Spread spectrum signal acquisition: Methods and technology," *IEEE Commun. Mag.*, vol. 22, pp. 6–21, June 1984.

[8] R. J. C. Bultitude, S. A. Mahmoud, and W. A. Sullivan, "A comparison of indoor radio propagation characteristics at 910 MHz and 1.75 GHz," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 20–30, Jan. 1989.

[9] D. M. J. Devasirvatham, "Multipath time delay jitter measured at 850 MHz in the portable radio environment," *IEEE J. Select. Areas Commun.*, vol. SAC-5, pp. 855–861, June 1987.

[10] E. Zollinger, "A statistical model for wideband inhouse radio channels," in *Proc. MELECON '89*, Lisbon, Portugal, Apr. 1989, pp. 429–432.

[11] J. L. Massey and J. J. Uhran, "Sub-baud coding," in *Proc. 13th Annu. Allerton Conf. Circuit Syst. Theory*, Oct. 1975, pp. 539–547.

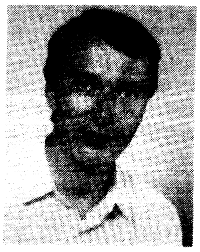
[12] M. B. Pursley, "Performance evaluation for phase-coded spread-spectrum multiple-access communication—Part I: System analysis," *IEEE Trans. Commun.*, vol. COM-25, pp. 795–799, Aug. 1977.

[13] U. Grob, "Über die Anwendung der Phasenhüpfertechnik in zellularen Digitalfunknetzen," Ph.D. dissertation, Swiss Fed. Inst. Technol., ETH, Zürich, dissertation ETH 9014, 1989.

[14] M. B. Pursley and H. F. A. Roefs, "Numerical evaluation of correlation parameters for optimal phases of binary shift-register sequences," *IEEE Trans. Commun.*, vol. COM-27, pp. 1597–1604, Oct. 1979.

[15] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications, Vol. III.* Rockville, MD: Computer Science Press, 1985.

[16] A. L. Welti and B. Z. Bobrovsky, "Long-term frequency stability requirements and optimal design of a code tracking delay-lock loop," in *Proc. IEEE Global Telecommun. Conf., GLOBECOM' 88*, vol. I, Hollywood, FL, Nov. 1988, pp. 0551–0555.

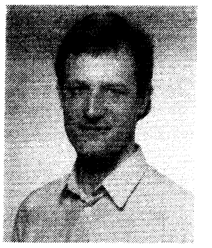


Urs Grob (M'89) was born in Zürich, Switzerland, in 1957. He received the Dipl. El. Ing. ETH degree from the Swiss Federal Institute of Technology, Zürich, in 1982. In 1988 he completed his postgraduate studies in communications technology.

Since 1982 he has been a Research Assistent with the Institute for Communication Technology, Swiss Federal Institute of Technology. Since 1984 his primary concern has been the application of the direct sequence spread spectrum technique for indoor and outdoor mobile communication systems. During his work towards the Ph.D. degree, he became leader of a group involved in this area since 1987. In the Fall 1989, he joined the ETSI (European Telecommunications Standards Institute), Sophia Antipolis, France.

Arnold L. Welti (S'84) was born in Berikon, Switzerland, in 1958. He received the Dipl. Ing. HTL degree from the Brugg College of Engineering, Brugg, Switzerland, in 1981 and the M.S. degree from Syracuse University, Syracuse, NY, in 1985, both in electrical engineering.

From 1982 to 1983 he was an R&D Engineer at Zellweger Uster, Uster, Switzerland, involved in the area of industrial electronics and signal processing. In 1983 he became a Teaching Assistant at Syracuse University. Since 1985 he has been a Research Assistant at the Institute for Communication Technology, Swiss Federal Institute of Technology (ETH), Zürich, Switzerland, where he is working towards the Ph.D. degree in the area of spread spectrum communications. Currently, his primary interests are in spread spectrum communications, synchronization, and signal processing.



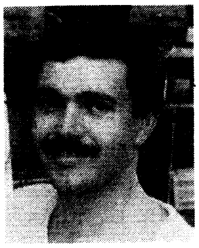
Ernst Zollinger (M'89) was born in Uster, Switzerland, in 1957. He received the Dipl. El. Ing. ETH degree from the Swiss Federal Institute of Technology (ETH), Zürich, in 1982 and completed his postgraduate studies in communications technology in 1990.

He then joined the Institute for Communication Technology, Swiss Federal Institute of Technology, where he is actively engaged in different projects on indoor radio wave propagation for wide-band communication systems, which are supported by the Swiss National Foundation and the Swiss PTT. Since 1989 he has been a Swiss delegate in the European projects COST 231 and RACE MOBILE (R 1043) where he is actively involved as an expert in the corresponding radio wave propagation groups. In 1989 he became head of a research group on spread spectrum techniques within the Institute for Communication Technology. Currently, he is working towards the Ph.D. degree in this field.

Roland Küng (M'86) was born in Frauenfeld, Switzerland, in 1954. He received the Dipl. Ing. ETH degree from the Swiss Federal Institute of Technology (ETH), Zürich, in 1978.

He was active in the design and the development of RF-communications equipment from 1979 to 1983 at Zellweger Telecommunications AG. In 1984 he founded a research group at this company with interests in RF communications, wireless inhouse communications, digital signal processing, and VLSI design. His own areas of interests are communication technology and RF design including surface acoustic wave (SAW), digital signal processing, and algorithm engineering. He holds several patents on these topics. In 1988 Zellweger Telecommunications AG joined the new European telecommunications company, Ascom, where he is now head of the Research Department of the Ascom Zelcom AG Division.

Mr. Küng is member of the SEV and the ITG Switzerland.



Hans Kaufmann (M'89) was born in Sursee, Switzerland, in 1958. He received the Dipl. Ing. ETH degree from the Swiss Federal Institute of Technology, Zürich, in 1982.

From 1982 to 1986 he was involved in the development of integrated optic devices on GaAs at the Institute of Quantum Electronics, Swiss Federal Institute of Technology. In 1986 he joined the Communication Group of the Research Department of Ascom Zelcom AG. His current area of interest is signal processing for spread spectrum systems, including surface acoustic wave (SAW) devices, and digital and optical signal processing applications.

Security Evaluation of a New Analog Speech Privacy/Scrambling Device Using Hopping Filters

ALEX GONIOTAKIS, STUDENT MEMBER, IEEE, AND AHMED K. ELHAKEEM, SENIOR MEMBER, IEEE

Abstract—A new concept in analog speech scrambling is introduced. The technique can be looked upon as processing the analog signal in two dimensions of time and frequency. A parallel bank of time variant filters and a dither random pulse amplitude signal, both generated from multinonlinear direct-sequence spread-spectrum codes, operate on the analog signal to destroy its intelligibility. At the receiver, a reversed sequence of time and frequency processing takes place, and with appropriate synchronization through a preample, the analog received signal is recovered, with a minimum amount of distortion. In this paper we introduce the building blocks, and conduct a statistical analysis to evaluate the security of the scrambling system, in terms of the autocorrelation and cross covariance. Finally, the security of the hybrid system is compared to other one-dimensional scrambling schemes, and derivations thereof.

I. Introduction

COMMUNICATION, an invaluable service needed in all aspects of society (i.e., business, military), can prove very destructive to some if abused by others (via intrusion to private interactions).

The vulnerability to intrusion sometimes even outweighs the benefits, depending on the level of secrecy of the information communicated. To remedy such situations, various algorithms have been developed for destroying the intelligibility of speech before transmitting, as well as the reverse techniques for recovering the original speech after receiving. These will be collectively referred to as secure algorithms.

Of the many secure algorithms for speech proposed so far, all can be classified either as analog or digital depending on the form of the input signal. Digital secure algorithms, which encrypt and correspondingly decrypt speech, are known to achieve a higher degree of security, but they are still not quite compatible with today's technical environment. The most important channels (i.e., telephone and radio) used for speech and data communication today are analog channels, and hence cannot support the high bit rate needed for proper digital transmission [1]. This paper is concerned with a new class of analog secure algorithms.

Analog secure algorithms, which scramble and correspondingly descramble speech, can be classified as one dimensional or two dimensional. One-dimensional algorithms are those that manipulate a signal in the time domain, frequency domain, or with respect to amplitude (amplitude scrambling), while two-dimensional algorithms are combinations of two or more one-dimensional algorithms.

A pair of analog secure algorithms (one scrambling and one descrambling), together with some circuitry needed for controlling their operation, are usually implemented on one device, which will be referred to as a secure device, or more precisely a speech privacy/scrambling device. Section II of this paper describes three one-dimensional secure algorithms, and two proposed combinations of these as two-dimensional secure algorithms (one a slight variation of the other), all of which are possible candidates for the basis of a secure device. In Section III of this paper, the security of each algorithm is examined by first finding expressions for the cross covariance, autocorrelation, and power spectral density of each, assuming the input signal is speech, and then plotting the autocorrelation and power spectral densities. The results of the cross covariance, and the various plots of the autocorrelations and power spectral densities, are the criteria used to choose the best secure algorithm among the few candidates proposed. This algorithm is then used as the basis of a proposed secure device, whose general operation is described in Section IV of this paper.

II. Secure One-Dimensional and Two-Dimensional Algorithms Proposed

The first of three one-dimensional algorithms to be described, and a possible candidate for the secure device to be proposed, is a new frequency domain algorithm. This algorithm consists of choosing a particular hopped filter from a group of n hopped filters, and passing the incoming source analog signal through this filter for a certain period of time, before hopping to another filter at a rate to be discussed later. Each hopped filter is a piecewise combination of seven unity-gain bandpass filters, which are mutually exclusive and collectively exhaust the frequency band 200–2400 Hz. Following each bandpass filter is a gain control device, which enhances or attenuates the incoming signal by one of several built-in factors or

Manuscript received July 29, 1989; revised January 12, 1990. This work was supported by a grant from A. K. Investment Inc., Montreal, P.Q., Canada. This work was presented at the 12th Symposium on Information Theory and Its Applications (SITA '89), Inuyama, Japan, Dec. 6–9, 1989. This work is also known as U.S. Disclosure Document 192628, 1988 and Canada Caveat 33503, 1988, A. K. Elhakeem.

A. Goniotakis is at 5937 Park Avenue, Montreal, P.Q., Canada H2V 4H4.

A. K. Elhakeem is with the Department of Electrical and Computer Engineering, Concordia University, 1455 de Maisonneuve Blvd. W., Montreal, P.Q., Canada H3G 1M8.

IEEE Log Number 9035806.

levels, as chosen from by a short binary codeword. The outputs of all seven gain control devices are then added to produce the final output of this algorithm (Fig. 1). The seven short binary codewords are, respectively, parts of seven pseudorandomly generated secret codes, which are produced independently from, but simultaneously with, each other.

The time duration for which a particular hopped filter will be used is practically upper-bounded by $T_F = 1.25$ ms for reasons to be explained later, and lower-bounded by the time needed to generate seven 1-bit codewords in parallel (the smallest codewords that can be generated). The lower bound thus depends on the clock frequency of each code generator, which is selected as 3200 Hz for reasons to be explained later. As a result, the fastest available timing signal has a period of $T_c = 0.3125$ ms. This provides the lower bound on the time duration for which a particular hopped filter will be used. These bounds restrict the number of bits in a particular short codeword (which is the same for all seven such short codewords), from 1 to 4. As a result, the number of gain factors or levels, by which each unity-gain bandpass filter output can be enhanced or attenuated, as chosen by the seven short codewords, range from 2^1 (= 2) to 2^4 (= 16). This leads to the number of possible piecewise hopped filters n to range from 2^7 to 16^7 (= 2^{28}). At this point, it might seem that the more the hopped filters to choose from (i.e., 2^{28}) the higher the security of the algorithm. This is true if the cryptanalytical attack consists of trying to find the particular hopped filter used, through exhaustively applying all possible inverse filters to the scrambled signal. Since this number of hopping filters is inversely proportional to the hopping rate, the residual intelligibility will be increased. On the other hand, if the cryptanalytical attack consists of simply listening to the scrambled signal, then a lower residual intelligibility is required, which implies a faster hopping rate. Since the hopping rate is inversely proportional to the number of hopping filters, then fewer hopping filters will be necessary. Since the cryptanalyst will follow the attack which best serves his/her purposes, it is up to the designer to provide a compromise between the two conflicting factors of maximum possible number of hopping filters, and maximum possible hopping rate, thus providing overall security. This compromise will be resolved later with a design choice as to the actual value of T_F.

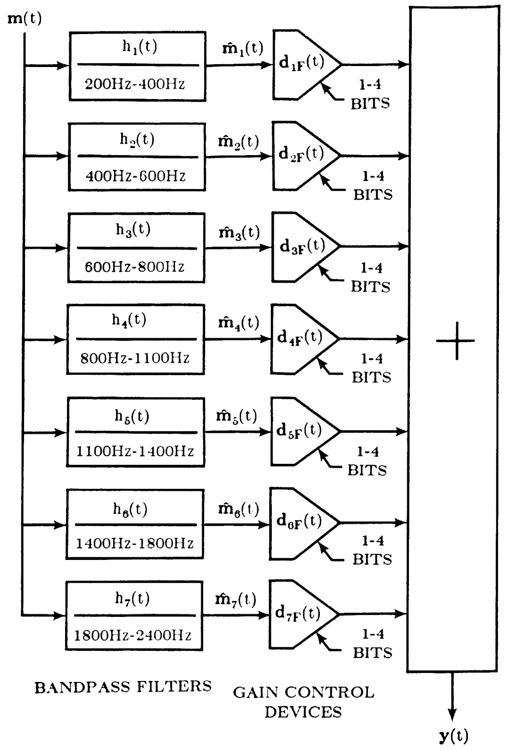


Fig. 1. Details of the hopping filters algorithm.

Let the source input analog signal to the hopping filters algorithm be a band-limited speech signal between 200 and 3200 Hz, represented by the stochastic process $\boldsymbol{m}(t)$. At any particular time t, the probability density of $\boldsymbol{m}(t)$ "is characterized in general by a very high probability of zero and near-zero amplitudes (related to pauses and low-energy segments of the speech waveform), by a significant probability of very high amplitudes, and by a monotonically decreasing function of amplitudes in between those extremes" [2]. Furthermore, let the output signals of the seven bandpass filters in this algorithm, which are filtered versions of $\boldsymbol{m}(t)$, be represented by the stochastic processes $\hat{\boldsymbol{m}}_1(t) \cdots \hat{\boldsymbol{m}}_7(t)$, respectively. Finally, let the control inputs of the hopping filters algorithm be seven identical but independent PAM processes. The levels of each of these processes are the gain factors by which a particular frequency band, represented by a particular stochastic process $\hat{\boldsymbol{m}}_1(t) \cdots \hat{\boldsymbol{m}}_7(t)$, can be enhanced or attenuated, as chosen by a particular pseudorandomly generated codeword. The seven PAM processes are represented as $\boldsymbol{d}_{1F}(t), \cdots, \boldsymbol{d}_{jF}(t), \cdots, \boldsymbol{d}_{7F}(t)$ such that

$$\boldsymbol{d}_{jF}(t) = \sum_{k=-\infty}^{+\infty} g_F(t - kT_F)\, \boldsymbol{D}_{jF}(k). \tag{1}$$

$\boldsymbol{D}_{1F}(k), \cdots, \boldsymbol{D}_{7F}(k)$ are seven identical but independent uniformly distributed discrete-time stochastic processes, and $g_F(t - kT_F)$ is a gate function defined as

$$g_F(t - kT_F) = \begin{cases} 1 & \text{if } kT_F \le t \le (k+1)T_F \\ 0 & \text{if otherwise,} \end{cases} \tag{2}$$

where 0.3125 ms $\le T_F \le$ 1.25 ms depending on the design choice to resolve the compromise outlined above. Refer to Fig. 1 for the details of the hopping filters algorithm (HF) as described above. The final output of this algorithm is denoted by the stochastic process $\boldsymbol{y}(t)$, which based on the above description is

$$\boldsymbol{y}(t) = \sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t)\, \boldsymbol{d}_{iF}(t). \tag{3}$$

The second and third one-dimensional algorithms to be described are known algorithms of amplitude scrambling. One of them consists of linearly adding an analog signal to the input source analog signal (Fig. 2), while the other

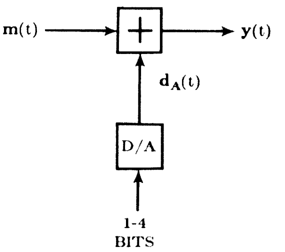


Fig. 2. Details of the amplitude (addition) scrambling algorithm.

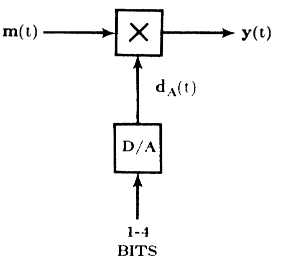


Fig. 3. Details of the amplitude (multiplication) scrambling algorithm.

consists of multiplying the input source analog signal by an analog signal (Fig. 3). In both these cases, the analog signal is made up of various dc levels, which are the outputs of a D/A converter, to which the inputs are short codewords produced from the code generator. Using the same reasoning as in the timing of the hopping filters algorithm explained previously, the time duration for which a particular dc level will multiply, or be added to the input source analog signal, ranges from $T_A = 0.3125$ ms to $T_A = 1.25$ ms. Furthermore, the length of the codeword choosing a particular level will range from 1 to 4 bits, and as a result the number of possible codewords at the input of the D/A, or number of possible levels at the output of the D/A, will range from 2^1 (= 2) to 2^4 (= 16). Similarly, as before, a compromise will have to be reached between the maximum possible number of levels to multiply, or be added to the incoming speech signal, and the maximum possible rate of change between these levels. These are two conflicting factors both of which contribute in different ways to the security of the algorithm. This compromise will also be resolved later with a design choice as to the actual value of T_A.

As before, let the source input analog signal to these two amplitude scrambling algorithms be a band-limited speech signal between 200-3200 Hz, represented by the stochastic process $\boldsymbol{m}(t)$. Furthermore, let the control input to these algorithms be a PAM process represented by $\boldsymbol{d}_A(t)$, such that

$$\boldsymbol{d}_A(t) = \sum_{k=-\infty}^{+\infty} g_A(t - kT_A)\, \boldsymbol{D}_A(k). \tag{4}$$

$\boldsymbol{D}_A(k)$ is a uniformly distributed discrete-time stochastic process, and $g_A(t - kT_A)$ is a gate function defined as

$$g_A(t - kT_A) = \begin{cases} 1 & \text{if } kT_A \le t \le (k+1)T_A \\ 0 & \text{if otherwise,} \end{cases} \tag{5}$$

where 0.3125 ms $\le T_A \le$ 1.25 ms, depending on the design choice to resolve the compromise outlined above. The levels of the PAM process $\boldsymbol{d}_A(t)$ are the levels which are added to, or multiply, the input speech signal $\boldsymbol{m}(t)$, as shown in Figs. 2 or 3, respectively.

The input to both of these candidate algorithms is $\boldsymbol{m}(t)$, while their outputs are

$$y(t) = \boldsymbol{m} + \boldsymbol{d}_A(t), \tag{6}$$

or

$$y(t) = \boldsymbol{m}(t)\, \boldsymbol{d}_A(t) \tag{7}$$

for the amplitude (addition) scrambling (AAS) algorithm, or the amplitude (multiplication) scrambling (AMS) algorithm, respectively.

The first of the two-dimensional algorithms, and a possible candidate for the secure device to be proposed, consists of the hopping filters algorithm described above, followed by amplitude (addition) scrambling in that order (HF-AAS), as shown in Fig. 4. The output of the frequency domain algorithm (3) is used as input to the amplitude scrambling algorithm (6). Maintaining the input of this overall two-dimensional algorithm as $\boldsymbol{m}(t)$, and referring to the output as $z(t)$, then

$$z(t) = \left(\sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t)\, \boldsymbol{d}_{iF}(t) \right) + \boldsymbol{d}_A(t). \tag{8}$$

The second of the two-dimensional algorithms, and a possible candidate for the secure device to be proposed consists of the hopping filters algorithm described above, followed by amplitude (multiplication) scrambling in that order (HF-AMS), as shown in Fig. 5. The output of the frequency domain algorithm (3) is used as input to the amplitude scrambling algorithm (7). Maintaining the input of this overall two-dimensional algorithm as $\boldsymbol{m}(t)$ and referring to the output as $z(t)$, then

$$z(t) = \left(\sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t)\, \boldsymbol{d}_{iF}(t) \right) \boldsymbol{d}_A(t). \tag{9}$$

III. Security Evaluation of the Proposed Secure Algorithms

The first objective of this analysis is to find the cross covariance

$$C_{mz}(\tau) = E[\boldsymbol{m}(t)\, z(t + \tau)] - E[\boldsymbol{m}(t)]\, E[z(t + \tau)] \tag{10}$$

[3] between the input process of each scrambling algorithm $\boldsymbol{m}(t)$, and the output process $z(t)$ of the same algorithm shifted by an amount τ, thus, $z(t + \tau)$ with respect to this shift. This is carried out so as to examine the degree of correlation between the input and output of a particular algorithm, with respect to a time difference of τ. The smaller this correlation, the lesser the amount of information resulting at the output of the algorithm, with regard to the input of the algorithm. Thus, the ideal value of cross covariance for a scrambling algorithm is zero.

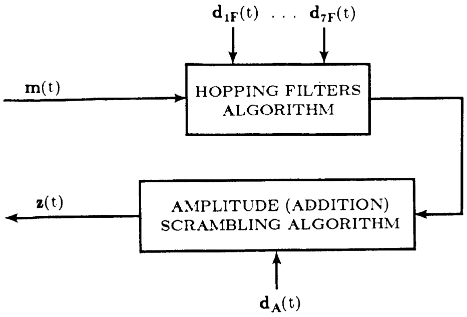

d1F(t) . . . d7F(t)
m(t)
HOPPING FILTERS ALGORITHM
z(t)
AMPLITUDE (ADDITION) SCRAMBLING ALGORITHM
dA(t)

Fig. 4. Hopping filters—amplitude (addition) scrambling.

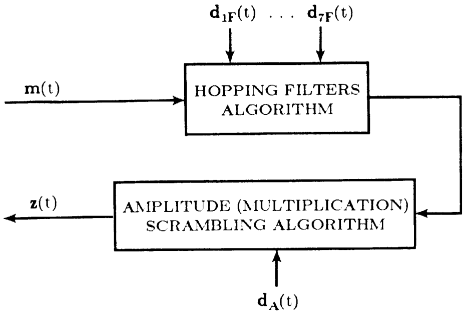

d1F(t) . . . d7F(t)
m(t)
HOPPING FILTERS ALGORITHM
z(t)
AMPLITUDE (MULTIPLICATION) SCRAMBLING ALGORITHM
dA(t)

Fig. 5. Hopping filters—amplitude (multiplication) scrambling.

This would imply that the input and output are statistically uncorrelated, and hence the statistics of the output signal would convey no information as to the statistics of the input signal. This is a slightly weaker condition than, but nevertheless resulting from, the idea of perfect secrecy, which implies that the output of a scrambling algorithm is statistically independent from its input [4].

The first algorithm to be examined in terms of cross covariance is the two-dimensional hopping filters amplitude (addition) scrambling algorithm. Substituting the output process $z(t)$ (8) in the definition of cross covariance (10),

$$\begin{aligned}E[\boldsymbol{m}(t)\, z(t+\tau)] &= E\left[\boldsymbol{m}(t)\left(\left(\sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t+\tau)\, \boldsymbol{d}_{iF}(t+\tau)\right) + \boldsymbol{d}_A(t+\tau)\right)\right] \\ &= E\left[\sum_{i=1}^{7} \boldsymbol{m}(t)\, \hat{\boldsymbol{m}}_i(t+\tau)\, \boldsymbol{d}_{iF}(t+\tau)\right] + E[\boldsymbol{m}(t)\, \boldsymbol{d}_A(t+\tau)] \\ &= \sum_{i=1}^{7} E[\boldsymbol{m}(t)\, \hat{\boldsymbol{m}}_i(t+\tau)\, \boldsymbol{d}_{iF}(t+\tau)] + E[\boldsymbol{m}(t)\, \boldsymbol{d}_A(t+\tau)].\end{aligned}$$

The PAM processes $\boldsymbol{d}_{1F}(t+\tau), \cdots, \boldsymbol{d}_{7F}(t+\tau)$ and $\boldsymbol{d}_A(t+\tau)$ are independent of the speech process $\boldsymbol{m}(t)$ and its filtered versions $\hat{\boldsymbol{m}}_1(t+\tau), \cdots, \hat{\boldsymbol{m}}_7(t+\tau)$; hence,

$$\begin{aligned}E[\boldsymbol{m}(t)\, z(t+\tau)] &= \left(\sum_{i=1}^{7} E[\boldsymbol{m}(t)\, \hat{\boldsymbol{m}}_i(t+\tau)]\, E[\boldsymbol{d}_{iF}(t+\tau)]\right) \\ &\quad + E[\boldsymbol{m}(t)]\, E[\boldsymbol{d}_A(t+\tau)].\end{aligned}$$

At this point, a constraint is put on the PAM processes $\boldsymbol{d}_{1F}(t+\tau), \cdots, \boldsymbol{d}_{7F}(t+\tau)$ such that their levels have a constant mean $\mu_{1F}, \cdots, \mu_{7F}$, respectively. Since the levels of these PAM processes are to be the same for each, then $\mu_{1F} = \cdots = \mu_{7F} = \mu_F$. Furthermore, the same constraint is applied to the PAM process $\boldsymbol{d}_A(t+\tau)$ and its mean is μ_A. As a result, $E[\boldsymbol{d}_{1F}(t+\tau)] = \cdots = E[\boldsymbol{d}_{7F}(t+\tau)] = \mu_F$ and $E[\boldsymbol{d}_A(t+\tau)] = \mu_A$ for all t.

$$\begin{aligned}E[\boldsymbol{m}(t)\, z(t+\tau)] &= \mu_F \sum_{i=1}^{7} E[\boldsymbol{m}(t)\, \hat{\boldsymbol{m}}_i(t+\tau)] + \mu_A E[\boldsymbol{m}(t)] \\ &= \mu_F E\left[\boldsymbol{m}(t) \sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t+\tau)\right] + \mu_A E[\boldsymbol{m}(t)]. \qquad (11)\end{aligned}$$

Also,

$$\begin{aligned}E[\boldsymbol{m}(t)]\, E[z(t+\tau)] &= E[\boldsymbol{m}(t)]\, E\left[\sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t+\tau)\, \boldsymbol{d}_{iF}(t+\tau) + \boldsymbol{d}_A(t+\tau)\right] \\ &= E[\boldsymbol{m}(t)]\left(\sum_{i=1}^{7} E[\hat{\boldsymbol{m}}_i(t+\tau)\, \boldsymbol{d}_{iF}(t+\tau)] + E[\boldsymbol{d}_A(t+\tau)]\right).\end{aligned}$$

The PAM processes $\boldsymbol{d}_{1F}(t+\tau), \cdots, \boldsymbol{d}_{7F}(t+\tau)$ are independent of the filtered speech processes $\hat{\boldsymbol{m}}_1(t+\tau), \cdots, \hat{\boldsymbol{m}}_7(t+\tau)$; hence,

$$\begin{aligned}E[\boldsymbol{m}(t)]\, E[z(t+\tau)] &= E[\boldsymbol{m}(t)]\left(\sum_{i=1}^{7} E[\hat{\boldsymbol{m}}_i(t+\tau)]\, E[\boldsymbol{d}_{iF}(t+\tau)] + E[\boldsymbol{d}_A(t+\tau)]\right).\end{aligned}$$

Putting the same constraints on the PAM processes as before $E[\boldsymbol{d}_{1F}(t+\tau)] = \cdots = E[\boldsymbol{d}_{7F}(t+\tau)] = \mu_F$ and $E[\boldsymbol{d}_A(t+\tau)] = \mu_A$ for all t.

$$\begin{aligned}E[\boldsymbol{m}(t)]\, E[z(t+\tau)] &= \mu_F E[\boldsymbol{m}(t)] \sum_{i=1}^{7} E[\hat{\boldsymbol{m}}_i(t+\tau)] + \mu_A E[\boldsymbol{m}(t)] \\ &= \mu_F E[\boldsymbol{m}(t)]\, E\left[\sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t+\tau)\right] + \mu_A E[\boldsymbol{m}(t)]. \qquad (12)\end{aligned}$$

The cross covariance then becomes

$$C_{mz}(\tau) = \mu_F\left(E\left[\boldsymbol{m}(t) \sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t+\tau)\right] - E[\boldsymbol{m}(t)]\, E\left[\sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t+\tau)\right]\right). \quad (13)$$

Since the bandpass filters are of unity gain, mutually exclusive, and collectively exhaust the band 200-2400 Hz, then

$$\sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t+\tau) = \tilde{\boldsymbol{m}}(t+\tau),$$

where $\tilde{\boldsymbol{m}}(t+\tau)$ is $\boldsymbol{m}(t+\tau)$ filtered by a bandpass filter of unity gain and bandwidth 200-2400 Hz. Since speech $\boldsymbol{m}(t+\tau)$ ranges from 200 to 3200 Hz, for all intents and purposes $\boldsymbol{m}(t+\tau) \approx \tilde{\boldsymbol{m}}(t+\tau)$, and hence,

$$\sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t+\tau) \approx \boldsymbol{m}(t+\tau).$$

The cross covariance then becomes

$$C_{mz}(\tau) = \mu_F(E[\boldsymbol{m}(t)\, \boldsymbol{m}(t+\tau)] - E[\boldsymbol{m}(t)]\, E[\boldsymbol{m}(t+\tau)]), \quad (14)$$

which is the product of the mean of the seven identical PAM processes of the hopping filters algorithm, and the autocovariance of the input process $\boldsymbol{m}(t)$.

The second algorithm to be examined in terms of cross covariance is the two-dimensional hopping filters amplitude (multiplication) scrambling algorithm. Substituting the output process $z(t)$ (9) in the definition of cross covariance (10),

$$\begin{aligned} E[\boldsymbol{m}(t)\, z(t+\tau)] &= E\left[\boldsymbol{m}(t)\left(\sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t+\tau)\, \boldsymbol{d}_{iF}(t+\tau)\right) \cdot \boldsymbol{d}_A(t+\tau)\right] \\ &= \sum_{i=1}^{7} E[\boldsymbol{m}(t)\, \hat{\boldsymbol{m}}_i(t+\tau)\, \boldsymbol{d}_{iF}(t+\tau) \cdot \boldsymbol{d}_A(t+\tau)]. \end{aligned}$$

The PAM processes $\boldsymbol{d}_{1F}(t+\tau), \cdots, \boldsymbol{d}_{7F}(t+\tau)$ and $\boldsymbol{d}_A(t+\tau)$ are independent of each other as well as independent of the speech processes $\hat{\boldsymbol{m}}_1(t+\tau), \cdots, \hat{\boldsymbol{m}}_7(t+\tau)$ and $\boldsymbol{m}(t)$; hence,

$$\begin{aligned} E[\boldsymbol{m}(t)\, z(t+\tau)] &= \sum_{i=1}^{7} E[\boldsymbol{m}(t)\, \hat{\boldsymbol{m}}_i(t+\tau)]\, E[\boldsymbol{d}_{iF}(t+\tau)] \cdot E[\boldsymbol{d}_A(t+\tau)]. \end{aligned}$$

Again, the constraint of constant mean $\mu_{1F}, \cdots, \mu_{7F}$ is put on the PAM processes $\boldsymbol{d}_{1F}(t+\tau), \cdots, \boldsymbol{d}_{7F}(t+\tau)$, respectively. Since they are identical processes (i.e., same levels) their mean is $\mu_{1F} = \cdots = \mu_{7F} = \mu_F$. Furthermore, the same constraint is applied to the PAM process $\boldsymbol{d}_A(t+\tau)$ and its mean is μ_A. As a result, $E[\boldsymbol{d}_{1F}(t+\tau)] = \cdots = E[\boldsymbol{d}_{7F}(t+\tau)] = \mu_F$ and $E[\boldsymbol{d}_A(t+\tau)] = \mu_A$ for all t. Then

$$\begin{aligned} E[\boldsymbol{m}(t)\, z(t+\tau)] &= \mu_A\mu_F \sum_{i=1}^{7} E[\boldsymbol{m}(t)\, \hat{\boldsymbol{m}}_i(t+\tau)] \\ &= \mu_A\mu_F E\left[\boldsymbol{m}(t) \sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t+\tau)\right]. \end{aligned} \quad (15)$$

Also,

$$\begin{aligned} &E[\boldsymbol{m}(t)]\, E[z(t+\tau)] \\ &\quad= E[\boldsymbol{m}(t)]\, E\left[\left(\sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t+\tau)\, \boldsymbol{d}_{iF}(t+\tau)\right) \cdot \boldsymbol{d}_A(t+\tau)\right] \\ &\quad= E[\boldsymbol{m}(t)]\left(\sum_{i=1}^{7} E[\hat{\boldsymbol{m}}_i(t+\tau)\, \boldsymbol{d}_{iF}(t+\tau) \cdot \boldsymbol{d}_A(t+\tau)]\right). \end{aligned}$$

The PAM processes $\boldsymbol{d}_{1F}(t+\tau), \cdots, \boldsymbol{d}_{7F}(t+\tau)$ and $\boldsymbol{d}_A(t+\tau)$ are independent of each other as well as of the speech processes $\hat{\boldsymbol{m}}_1(t+\tau), \cdots, \hat{\boldsymbol{m}}_7(t+\tau)$, and hence,

$$\begin{aligned} &E[\boldsymbol{m}(t)]\, E[z(t+\tau)] \\ &\quad= E[\boldsymbol{m}(t)]\left(\sum_{i=1}^{7} E[\hat{\boldsymbol{m}}_i(t+\tau)]\, E[\boldsymbol{d}_{iF}(t+\tau)] \cdot E[\boldsymbol{d}_A(t+\tau)]\right). \end{aligned}$$

Putting the same constraints on the PAM processes as before, $E[\boldsymbol{d}_{1F}(t+\tau)] = \cdots = E[\boldsymbol{d}_{7F}(t+\tau)] = \mu_F$, and $E[\boldsymbol{d}_A(t+\tau)] = \mu_A$ for all t. Then

$$\begin{aligned} &E[\boldsymbol{m}(t)]\, E[z(t+\tau)] \\ &\quad= \mu_A\mu_F E[\boldsymbol{m}(t)] \sum_{i=1}^{7} E[\hat{\boldsymbol{m}}_i(t+\tau)] \\ &\quad= \mu_A\mu_F E[\boldsymbol{m}(t)]\, E\left[\sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t+\tau)\right]. \end{aligned} \quad (16)$$

The cross covariance $C_{mz}(\tau)$ then becomes

$$C_{mz}(\tau) = \mu_A \mu_F \left(E\left[\boldsymbol{m}(t) \sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t+\tau) \right] - E[\boldsymbol{m}(t)]\, E\left[\sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t+\tau) \right] \right). \quad (17)$$

Using the same reasoning as before,

$$\sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t+\tau) \approx \boldsymbol{m}(t+\tau).$$

Hence,

$$C_{mz}(\tau) = \mu_A \mu_F \big(E[\boldsymbol{m}(t)\, \boldsymbol{m}(t+\tau)] - E[\boldsymbol{m}(t)]\, E[\boldsymbol{m}(t+\tau)] \big), \quad (18)$$

which is the product of the mean of one of the seven PAM processes of the hopping filters algorithm, the mean of the process of the amplitude scrambling algorithm, and the autocovariance of the input speech process $\boldsymbol{m}(t)$.

The last three algorithms to be examined in terms of cross covariance are the one-dimensional hopping filters algorithm, and the one-dimensional amplitude (addition and multiplication) scrambling algorithms. Referring to $\boldsymbol{y}(t)$ as $z(t)$ in (3), (6), and (7), substituting these in the definition of cross covariance (10), and following similar steps as in the previous calculations of cross covariance results in the cross covariances of these algorithms being

$$C_{mz}(\tau) = \mu_F \big(E[\boldsymbol{m}(t)\, \boldsymbol{m}(t+\tau)] - E[\boldsymbol{m}(t)]\, E[\boldsymbol{m}(t+\tau)] \big), \quad (19)$$

$$C_{mz}(\tau) = E[\boldsymbol{m}(t)\, \boldsymbol{m}(t+\tau)] - E[\boldsymbol{m}(t)]\, E[\boldsymbol{m}(t+\tau)], \quad (20)$$

and

$$C_{mz}(\tau) = \mu_A \big(E[\boldsymbol{m}(t)\, \boldsymbol{m}(t+\tau)] - E[\boldsymbol{m}(t)]\, E[\boldsymbol{m}(t+\tau)] \big), \quad (21)$$

respectively. The code generators choosing between the various levels of the PAM processes discussed so far are not biased in any way. As a result, the levels of each PAM process at any particular time have a uniform distribution. Given this fact, and tightening the constraint of constant mean to a mean of 0 for all the PAM processes, implies that the sum of the levels should be 0. In the amplitude scrambling algorithm, the single PAM process requires the same levels for the scrambling, as well as the descrambling procedures, with a one-to-one correspondence. In the hopping filters algorithm, the seven PAM processes require that their levels in the scrambling procedure are multiplicative inverses of their levels in the descrambling procedure, and vice versa, again with a one-to-one correspondence. Choosing the levels of all the PAM processes accordingly results in $\mu_F = \mu_A = 0$; and making use of the fact that the speech process has zero mean amplitudes for any particular time t ($E[\boldsymbol{m}(t)] = E[\boldsymbol{m}(t + \tau)] = 0$) [2], then the cross covariances derived so far all become 0, with the exception of that of the amplitude (addition) scrambling algorithm. The latter cross covariance is $R_m(\tau)$ which is equal to $E[\boldsymbol{m}(t)\, \boldsymbol{m}(t + \tau)]$ and is the autocorrelation of the speech input process $\boldsymbol{m}(t)$. Thus, for all the algorithms except that of amplitude (addition) scrambling, there is no correlation between input and output. As a result, when dealing with any of these algorithms, a cryptanalyst cannot obtain information as to the statistics of the input signal to the algorithm by observing the statistics of the output signal from the algorithm. In the case of the one-dimensional amplitude (addition) scrambling algorithm, the above does not hold since the autocorrelation of the speech process is nonzero, as will be shown later. Thus, this particular algorithm does not measure up to the others for scrambling speech, and is thus not examined further.

The second objective of this part of the paper is to find the autocorrelation between the output process $z(t)$ of each algorithm, and the shifted version of the same process $z(t + \tau)$, with respect to the amount of shift τ. This is carried out so as to examine how close the autocorrelation of each algorithm is to the autocorrelation of white noise, which is

$$R_w(\tau) = \frac{N_0}{2}\,\delta(\tau) \quad (22)$$

[5]. This autocorrelation implies that no matter how close in time two samples are taken, they are uncorrelated. This is the ideal autocorrelation desired for the output of a scrambling algorithm. It would imply that the information contained in one sample of the output signal would never appear again in any future output signals, no matter how close they are taken in time to the original signal. Thus, the cryptanalyst would not be able to confirm any information obtained, by relating to each other a few close samples in time, of the output scrambled signal. Furthermore, the closer the autocorrelation of the output signal of a scrambling algorithm is to a delta function, the more the output will sound like white noise.

The first of the two-dimensional algorithms to be examined in terms of autocorrelation is the hopping filters amplitude (addition) scrambling algorithm. The output of this algorithm is as shown in (8), and the shifted version of the output shifted by τ is

$$z(t+\tau) = \left(\sum_{i=1}^{7} \hat{\boldsymbol{m}}_i(t+\tau)\, \boldsymbol{d}_{iF}(t+\tau) \right) + \boldsymbol{d}_A(t+\tau). \quad (23)$$

Autocorrelation is defined as

$$R_{zz}(\tau) = E[z(t)\, z(t+\tau)] \quad (24)$$

[3]. Using j as a dummy variable in (23), and substituting (8) and (23) in (24), results in

$$\begin{aligned} R_{zz}(\tau) = E\Bigg[&\Bigg(\Bigg(\sum_{i=1}^{7} \hat{m}_i(t)\, d_{iF}(t)\Bigg) + d_A(t)\Bigg) \\ &\cdot \Bigg(\Bigg(\sum_{j=1}^{7} \hat{m}_j(t+\tau)\, d_{jF}(t+\tau)\Bigg) \\ &+ d_A(t+\tau)\Bigg)\Bigg] \\ = &\sum_{i=1}^{7}\sum_{j=1}^{7} E[\hat{m}_i(t)\, \hat{m}_j(t+\tau)\, d_{iF}(t)\, d_{jF}(t+\tau)] \\ &+ \sum_{j=1}^{7} E[\hat{m}_j(t+\tau)\, d_{jF}(t+\tau)\, d_A(t)] \\ &+ \sum_{i=1}^{7} E[\hat{m}_i(t)\, d_{iF}(t)\, d_A(t+\tau)] \\ &+ E[d_A(t)\, d_A(t+\tau)]. \end{aligned}$$

As before, all the PAM processes are independent of each other, as well as independent from the speech process and its filtered versions; hence,

$$\begin{aligned} R_{zz}(\tau) = &\sum_{\substack{i=1 \\ i=j}}^{7}\sum_{j=1}^{7} \big(E[\hat{m}_i(t)\, \hat{m}_j(t+\tau)] \\ &\cdot E[d_{iF}(t)\, d_{jF}(t+\tau)]\big) \\ &+ \sum_{\substack{i=1 \\ i\neq j}}^{7}\sum_{j=1}^{7} \big(E[\hat{m}_i(t)\, \hat{m}_j(t+\tau)]\, E[d_{iF}(t)] \\ &\cdot E[d_{jF}(t+\tau)]\big) \\ &+ \sum_{j=1}^{7} \big(E[\hat{m}_j(t+\tau)]\, E[d_{jF}(t+\tau)] \\ &\cdot E[d_A(t)]\big) \\ &+ \sum_{i=1}^{7} \big(E[\hat{m}_i(t)]\, E[d_{iF}(t)]\, E[d_A(t+\tau)]\big) \\ &+ E[d_A(t)\, d_A(t+\tau)]. \end{aligned}$$

Applying the constraint of constant mean to the PAM processes, and following the same reasoning as before, $E[d_{1F}(t)] = E[d_{1F}(t+\tau)] = \cdots = E[d_{7F}(t)] = E[d_{7F}(t+\tau)] = \mu_F$, and $E[d_A(t)] = E[d_A(t+\tau)] = \mu_A$. Then

$$\begin{aligned} R_{zz}(\tau) = &\sum_{i=1}^{7} R_{\hat{m}_i}(\tau)\, R_{d_{iF}}(\tau) + \mu_F^2 \sum_{\substack{i=1 \\ i\neq j}}^{7}\sum_{j=1}^{7} R_{\hat{m}_i\hat{m}_j}(\tau) \\ &+ \mu_F\mu_A \sum_{j=1}^{7} E[\hat{m}_j(t+\tau)] \\ &+ \mu_F\mu_A \sum_{i=1}^{7} E[\hat{m}_i(t)] + R_{d_A}(\tau), \end{aligned}$$

where $R_{\hat{m}_i}(\tau)$ is the autocorrelation of the ith filtered speech process, $\underset{i\neq j}{R_{\hat{m}_i\hat{m}_j}(\tau)}$ is the cross correlation of the ith and jth filtered processes, $R_{d_{iF}}(\tau)$ is the autocorrelation of the ith out of seven PAM processes in the hopping filters algorithm, and $R_{d_A}(\tau)$ is the autocorrelation of the PAM process $d_A(t)$ from the amplitude (addition) scrambling algorithm. Since the seven PAM processes from the hopping filters algorithm are identical, their autocorrelations are equal, and hence, $R_{d_{1F}}(\tau) = \cdots = R_{d_{7F}}(\tau) = R_{d_F}(\tau)$. Thus,

$$\begin{aligned} R_{zz}(\tau) = &R_{d_F}(\tau) \sum_{i=1}^{7} R_{\hat{m}_i}(\tau) + \mu_F^2 \sum_{\substack{i=1 \\ i\neq j}}^{7}\sum_{j=1}^{7} R_{\hat{m}_i\hat{m}_j}(\tau) \\ &+ \mu_F\mu_A E\Bigg[\sum_{j=1}^{7} \hat{m}_j(t+\tau)\Bigg] \\ &+ \mu_F\mu_A E\Bigg[\sum_{i=1}^{7} \hat{m}_i(t)\Bigg] + R_{d_A}(\tau). \end{aligned}$$

Using the same reasoning as before,

$$\sum_{j=1}^{7} \hat{m}_j(t+\tau) \approx m(t+\tau) \text{ and } \sum_{i=1}^{7} \hat{m}_i(t) \approx m(t);$$

hence,

$$\begin{aligned} R_{zz}(\tau) = &R_{d_F}(\tau) \sum_{i=1}^{7} R_{\hat{m}_i}(\tau) + \mu_F^2 \sum_{\substack{i=1 \\ i\neq j}}^{7}\sum_{j=1}^{7} R_{\hat{m}_i\hat{m}_j}(\tau) \\ &+ \mu_F\mu_A\big(E[m(t+\tau) + m(t)]\big) + R_{d_A}(\tau). \end{aligned} \tag{25}$$

The second of the two-dimensional algorithms to be examined in terms of autocorrelation is the hopping filters amplitude (multiplication) scrambling algorithm. The output of this algorithm is as shown in (9), and the shifted version of this output shifted by τ is

$$z(t+\tau) = \Bigg(\sum_{i=1}^{7} \hat{m}_i(t+\tau)\, d_{iF}(t+\tau)\Bigg) d_A(t+\tau). \tag{26}$$

Using j as a dummy variable in (26), and substituting (9) and (26) in (24), results in

$$\begin{aligned} R_{zz}(\tau) = E\Bigg[&\Bigg(\sum_{i=1}^{7} \hat{m}_i(t)\, d_{iF}(t)\Bigg)(d_A(t)) \\ &\cdot \Bigg(\sum_{j=1}^{7} \hat{m}_j(t+\tau)\, d_{jF}(t+\tau)\Bigg)(d_A(t+\tau))\Bigg] \\ = &\sum_{i=1}^{7}\sum_{j=1}^{7} E[\hat{m}_i(t)\, \hat{m}_j(t+\tau)\, d_{iF}(t)\, d_{jF}(t+\tau) \\ &\cdot d_A(t)\, d_A(t+\tau)]. \end{aligned}$$

As before, all the PAM processes are independent of each other, as well as independent of the speech process and its filtered versions; hence,

$$\begin{aligned} R_{zz}(\tau) = \sum_{\substack{i=1\\ i=j}}^{7} \sum_{j=1}^{7} \big(E[\hat{\boldsymbol{m}}_i(t)\, \hat{\boldsymbol{m}}_j(t+\tau)] \\ \cdot E[\boldsymbol{d}_{iF}(t)\, \boldsymbol{d}_{jF}(t+\tau)]\, E[\boldsymbol{d}_A(t)\, \boldsymbol{d}_A(t+\tau)]\big) \\ + \sum_{\substack{i=1\\ i\neq j}}^{7} \sum_{j=1}^{7} \big(E[\hat{\boldsymbol{m}}_i(t)\, \hat{\boldsymbol{m}}_j(t+\tau)]\, E[\boldsymbol{d}_{iF}(t)] \\ \cdot E[\boldsymbol{d}_{jF}(t+\tau)]\, E[\boldsymbol{d}_A(t)\, \boldsymbol{d}_A(t+\tau)]\big). \end{aligned}$$

Applying the constraint of constant mean to the PAM processes, and following the same reasoning as before, $E[\boldsymbol{d}_{1F}(t)] = E[\boldsymbol{d}_{1F}(t+\tau)] = \cdots = E[\boldsymbol{d}_{7F}(t)] = E[\boldsymbol{d}_{7F}(t+\tau)] = \mu_F$, and $E[\boldsymbol{d}_A(t)] = E[\boldsymbol{d}_A(t+\tau)] = \mu_A$. Then

$$\begin{aligned} R_{zz}(\tau) = R_{d_A}(\tau) \sum_{i=1}^{7} R_{\hat{m}_i}(\tau)\, R_{d_{iF}}(\tau) \\ + \mu_F^2 R_{d_A}(\tau) \sum_{\substack{i=1\\ i\neq j}}^{7} \sum_{j=1}^{7} R_{\hat{m}_i\hat{m}_j}(\tau), \end{aligned}$$

where $R_{\hat{m}_i}(\tau)$, $\underset{i\neq j}{R_{\hat{m}_i\hat{m}_j}(\tau)}$, $R_{d_{iF}}(\tau)$, and $R_{d_A}(\tau)$ are as defined previously. As before, $R_{d_{1F}}(\tau) = \cdots = R_{d_{7F}}(\tau) = R_{d_F}(\tau)$; thus,

$$\begin{aligned} R_{zz}(\tau) = R_{d_A}(\tau)\, R_{d_F}(\tau) \sum_{i=1}^{7} R_{\hat{m}_i}(\tau) \\ + \mu_F^2 R_{d_A}(\tau) \sum_{\substack{i=1\\ i\neq j}}^{7} \sum_{j=1}^{7} R_{\hat{m}_i\hat{m}_j}(\tau). \qquad (27) \end{aligned}$$

The two one-dimensional algorithms remaining for discussion (after eliminating the amplitude (addition) scrambling algorithm, because of its poor cross covariance in relation to the other algorithms) are the hopping filters algorithm and the amplitude (multiplication) scrambling algorithm. If each of these is used alone such that their input is $\boldsymbol{m}(t)$ and their output is $z(t)$, following similar steps as in the cases of the two two-dimensional algorithms, their autocorrelations become

$$R_{zz}(\tau) = R_{d_F}(\tau) \sum_{i=1}^{7} R_{\hat{m}_i}(\tau) + \mu_F^2 \sum_{\substack{i=1\\ i\neq j}}^{7} \sum_{j=1}^{7} R_{\hat{m}_i\hat{m}_j}(\tau), \qquad (28)$$

and

$$R_{zz}(\tau) = R_m(\tau)\, R_{d_A}(\tau), \qquad (29)$$

respectively. Respecting the constraint of zero mean for all the PAM processes as discussed before, $\mu_F = 0$ and $\mu_A = 0$. Since speech amplitudes have zero mean, and since the correlations that exist among these amplitudes depend on the time difference between them, then according to [3] the process representing speech, in this case $\boldsymbol{m}(t)$, is wide sense stationary (WSS). According to [6], if a WSS process is applied to the input of a time-invariant linear network with transfer function $H(f)$, then the output power spectral density is the product of $|H(f)|^2$ and the input power spectral density. In this case, each bandpass filter is a linear time-invariant network. The autocorrelation $R_{\hat{m}_j}(\tau)$, and the power spectral density $S_{\hat{m}_j}(f)$ at the output of the jth bandpass filter, are related as follows:

$$\boldsymbol{F}[R_{\hat{m}_j}(\tau)] = S_{\hat{m}_j}(f).$$

For all bandpass filters,

$$\begin{aligned} \sum_{j=1}^{7} \boldsymbol{F}[R_{\hat{m}_j}(\tau)] &= \sum_{j=1}^{7} S_{\hat{m}_j}(f) \\ \boldsymbol{F}\left[\sum_{j=1}^{7} R_{\hat{m}_j}(\tau)\right] &= \sum_{j=1}^{7} |H_j(f)|^2 S_m(f) \\ &= S_m(f) \sum_{i=1}^{7} |H_j(f)|^2. \end{aligned}$$

Since the bandpass filters are of unity gain, mutually exclusive, and collectively exhaust the band 200-2400 Hz, then

$$\boldsymbol{F}\left[\sum_{i=1}^{7} R_{\hat{m}_j}(\tau)\right] = S_m(f)\, |X(f)|^2,$$

where $X(f)$ is a unity gain bandpass filter with bandwidth 200-2400 Hz. Since the voice bandwidth is 200-3200 Hz, if speech signals above 2400 Hz are disregarded, then $S_m(f)\,|X(f)|^2 \approx S_m(f)$. Thus,

$$\begin{aligned} \boldsymbol{F}\left[\sum_{j=1}^{7} R_{\hat{m}_j}(\tau)\right] &\approx S_m(f) \\ \sum_{j=1}^{7} R_{\hat{m}_j}(\tau) &\approx \boldsymbol{F}^{-1}[S_m(f)] \\ &\approx R_m(\tau). \qquad (30) \end{aligned}$$

The autocorrelations of the four algorithms examined above can then be summarized as shown below.

Case 1) Hopping Filters—Amplitude (Addition) Scrambling:

$$R_{zz}(\tau) = R_{d_F}(\tau)\, R_m(\tau) + R_{d_A}(\tau). \qquad (31)$$

Case 2) Hopping Filters—Amplitude (Multiplication) Scrambling:

$$R_{zz}(\tau) = R_{d_A}(\tau)\, R_{d_F}(\tau)\, R_m(\tau). \qquad (32)$$

Case 3) Hopping Filters:

$$R_{zz}(\tau) = R_{d_F}(\tau)\, R_m(\tau). \qquad (33)$$

Case 4) Amplitude (Multiplication) Scrambling:

$$R_{zz}(\tau) = R_m(\tau)\, R_{d_A}(\tau). \qquad (34)$$

Refer to Fig. 6 for the long time (55 s) averaged autocorrelation of speech denoted $C(n)$, in terms of Nyquist

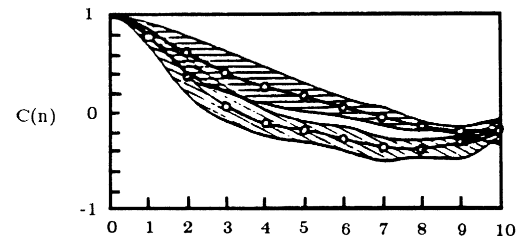


Fig. 6. Long time autocorrelation function of speech [2].

samples n (125 μs), as a result of speech from two males and two females. The top curve represents low-pass-filtered (0-3400 Hz) speech, and the bottom curve represents bandpass-filtered (200-3400 Hz) speech. Both curves are normalized by the expected (time-averaged) value of squared amplitudes, or the average power of the voice. For purposes of analysis, points from the bottom curve were used to obtain a polynomial representing this curve. This polynomial denoted by $R_p(\tau)$ is of degree 10, and represents the autocorrelation of speech (200-3400 Hz) in the interval $|\tau| \leq 1.25$ ms, since autocorrelation is an even function. The product of this polynomial, and the average power of the voice $R_m[0](= E[m^2(t)])$, is the autocorrelation of speech $R_m(\tau)$ used in the analysis of the various scrambling algorithms. Hence,

$$R_m(\tau) = R_m[0]\, R_p(\tau) \qquad \text{for } |\tau| \leq 1.25 \text{ ms.} \tag{35}$$

An upper bound had to be chosen for the hopping rate in the hopping filters algorithm, or for the time taken to change levels in the amplitude scrambling algorithms. This is because the longer it takes for the changes to occur, the easier the human ear may adapt to the current situation and begin to understand what is being said. Since this limit could be at most 5 ms as seen from experiment, and since the autocorrelation curve available ranged up to 1.25 ms, for analysis purposes the upper bound was chosen to be 1.25 ms.

In general, a PAM process changes levels every T units of time, but maintains a certain level for c units of time (the duty cycle factor), such that $c \leq T$. The autocorrelation of such a process consists of triangles each of base $2c$, and centered at multiples of T with height $cR[s]/T$, where s is the time in discrete time units between the PAM process and its shifted version (Fig. 7). Applying this general form of the autocorrelation to the PAM processes at hand, where $c = T$, the autocorrelation of each of these processes consists of triangles each of base $2T$ and centered at multiples of T with height $R[s]$. Thus, for the seven identical PAM processes $\boldsymbol{d}_{1F}(t), \cdots, \boldsymbol{d}_{7F}(t)$ of the hopping filters algorithm, and the one PAM process $\boldsymbol{d}_A(t)$ of the amplitude scrambling algorithm, the autocorrelation functions are

$$R_{d_F}(\tau) = \sum_{s=-\infty}^{+\infty} g(\tau - sT_F)\Bigg((s+1)R_{d_F}[s] - sR_{d_F}[s+1] + \frac{\tau}{T_F}\left(R_{d_F}[s+1] - R_{d_F}[s]\right)\Bigg), \tag{36}$$

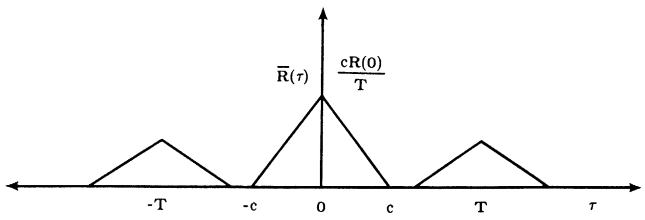


Fig. 7. Autocorrelation of a PAM process [3].

and

$$R_{d_A}(\tau) = \sum_{s=-\infty}^{+\infty} g(\tau - sT_A)\Bigg((s+1)R_{d_A}[s] - sR_{d_A}[s+1] + \frac{\tau}{T_A}\left(R_{d_A}[s+1] - R_{d_A}[s]\right)\Bigg), \tag{37}$$

respectively, where $R_{d_F}[s]$, $R_{d_F}[s+1]$, $R_{d_A}[s]$, and $R_{d_A}[s+1]$ depend on the levels of the PAM processes, and

$$g(\tau - sT_F) = \begin{cases} 1 & \text{if } sT_F \leq \tau \leq (s+1)T_F \\ 0 & \text{if otherwise} \end{cases}$$

and

$$g(\tau - sT_A) = \begin{cases} 1 & \text{if } sT_A \leq \tau \leq (s+1)T_A \\ 0 & \text{if otherwise,} \end{cases}$$

where T_F and T_A are between 0.3125 and 1.25 ms, as discussed previously.

From the conditions and constraints imposed on the levels of the PAM processes so far, the levels of the seven equivalent PAM processes $\boldsymbol{d}_{1F}(t), \cdots, \boldsymbol{d}_{7F}(t)$ can be defined as $\alpha = (\pm p_1, \pm p_2, \pm p_3, \cdots, \pm p_{k_F/2})$, where k_F is the number of levels. The mean is then defined as

$$\mu_F = \frac{1}{k_F}\sum_{i=1}^{k_F} \alpha_i = 0. \tag{38}$$

The autocorrelation in discrete time units s (multiples of T_F) was found to be

$$R_{d_F}[s] = E[D_{F_0} D_{F_s}] = \begin{cases} \dfrac{1}{k_F}\displaystyle\sum_{i=1}^{k_F} \alpha_i^2 & \text{if } s = 0 \\ \dfrac{1}{k_F^2}\left(\displaystyle\sum_{i=1}^{k_F} \alpha_i\right)^2 & \text{if } s \neq 0. \end{cases} \tag{39}$$

Similarly, the levels of the amplitude scrambling PAM process can be defined as $\beta = (\pm q_1, \pm q_2, \pm q_3, \cdots, \pm q_{k_A/2})$, where k_A is the number of levels. The mean is then

$$\mu_A = \frac{1}{k_A}\sum_{i=1}^{k_A} \beta_1 = 0. \tag{40}$$

The autocorrelation in discrete time units s (multiples of T_A) was found to be

$$R_{d_A}[s] = E[D_{A_0} D_{A_s}] = \begin{cases} \dfrac{1}{k_A} \displaystyle\sum_{i=1}^{k_A} \beta_i^2 & \text{if } s = 0 \\ \dfrac{1}{k_A^2} \left(\displaystyle\sum_{i=1}^{k_A} \beta_i \right)^2 & \text{if } s \neq 0. \end{cases} \tag{41}$$

Combining (38) and (39), and substituting in (36), results in

$$R_{d_F}(\tau) = \begin{cases} R_{d_F}[0] \left(1 - \dfrac{|\tau|}{T_F}\right) & \text{if } |\tau| < T_F \\ 0 & \text{if otherwise.} \end{cases} \tag{42}$$

Similarly, combining (40), (41), and substituting in (37), results in

$$R_{d_A}(\tau) = \begin{cases} R_{d_A}[0] \left(1 - \dfrac{|\tau|}{T_A}\right) & \text{if } |\tau| < T_A \\ 0 & \text{if otherwise.} \end{cases} \tag{43}$$

Substituting (35), (42), and (43) in (31), (32), (33), and (34) results in the following autocorrelations.

Case 1a) Hopping Filters—Amplitude (Addition) Scrambling ($T_F \leq T_A$):

$$R_{zz}(\tau) = \begin{cases} R_{d_F}[0]\, R_m[0] \left(1 - \dfrac{|\tau|}{T_F}\right) (R_p(\tau)) \\ \quad + R_{d_A}[0] \left(1 - \dfrac{|\tau|}{T_A}\right) & \text{if } |\tau| \leq T_F \\ R_{d_A}[0] \left(1 - \dfrac{|\tau|}{T_A}\right) & \text{if } T_F \leq |\tau| \leq T_A \\ 0 & \textit{if } |\tau| \geq T_A. \end{cases} \tag{44}$$

Case 1b) Hopping Filters—Amplitude (Addition) Scrambling ($T_F \geq T_A$):

$$R_{zz}(\tau) = \begin{cases} R_{d_F}[0]\, R_m[0] \left(1 - \dfrac{|\tau|}{T_F}\right) (R_p(\tau)) + R_{d_A}[0] \\ \quad \cdot \left(1 - \dfrac{|\tau|}{T_A}\right) & \text{if } |\tau| \leq T_A \\ R_{d_F}[0]\, R_m[0] \left(1 - \dfrac{|\tau|}{T_F}\right) (R_p(\tau)) \\ & \text{if } T_A \leq |\tau| \leq T_F \\ 0 & \text{if } |\tau| \geq T_F. \end{cases} \tag{45}$$

Case 2a) Hopping Filters—Amplitude (Multiplication) Scrambling ($T_F \leq T_A$):

$$R_{zz}(\tau) = \begin{cases} R_{d_F}[0]\, R_{d_A}[0]\, R_m[0] \left(1 - \dfrac{|\tau|}{T_F}\right)\left(1 - \dfrac{|\tau|}{T_A}\right) \\ \quad \cdot (R_p(\tau)) & \text{if } |\tau| \leq T_F \\ 0 & \text{if } |\tau| \geq T_F. \end{cases} \tag{46}$$

Case 2b) Hopping Filters—Amplitude (Multiplication) Scrambling ($T_F \geq T_A$):

$$R_{zz}(\tau) = \begin{cases} R_{d_F}[0]\, R_{d_A}[0]\, R_m[0] \left(1 - \dfrac{|\tau|}{T_F}\right)\left(1 - \dfrac{|\tau|}{T_A}\right) \\ \quad \cdot (R_p(\tau)) & \text{if } |\tau| \leq T_A \\ 0 & \text{if } |\tau| \geq T_A. \end{cases} \tag{47}$$

Case 3) Hopping Filters:

$$R_{zz}(\tau) = \begin{cases} R_{d_F}[0]\, R_m[0] \left(1 - \dfrac{|\tau|}{T_F}\right) (R_p(\tau)) \\ \quad \text{if } |\tau| \leq T_F \\ 0 \quad \text{if otherwise.} \end{cases} \tag{48}$$

Case 4) Amplitude (Multiplication) Scrambling:

$$R_{zz}(\tau) = \begin{cases} R_{d_A}[0]\, R_m[0] \left(1 - \dfrac{|\tau|}{T_A}\right) (R_p(\tau)) \\ \quad \text{if } |\tau| \leq T_A \\ 0 \quad \text{if otherwise.} \end{cases} \tag{49}$$

The third objective of this part of the paper is to find the power spectral density of the scrambled signal in each case. This is carried out so as to examine how close the power spectral density of the scrambled signal in each case is to the power spectral density of white noise, which is

$$S_W(f) = \frac{N_0}{2} \tag{50}$$

[5], which is the Fourier transform of the autocorrelation noise is independent of frequency. A similar power spectral density would be ideal for the scrambled signal at the output of a scrambling algorithm, since it would hide the frequency characteristics of the input signal, as well as any frequency domain manipulations performed on this input. The cryptanalyst would see all frequency components as equal (in terms of power) all the time. In general, the power spectral density is defined as

$$S_{zz}(f) = \int_{-\infty}^{+\infty} R_{zz}(\tau) e^{-j2\pi f\tau}\, d\tau \tag{51}$$

[5], which is the Fourier transform of the autocorrelation $R_{zz}(\tau)$. In this particular situation, let T_x be such that

$R_{zz}(\tau)$ is nonzero for $|\tau| \leq T_x$, and zero otherwise. Then

$$S_{zz}(f) = \int_{-T_x}^{T_x} R_{zz}(\tau) e^{-j2\pi f\tau}\, d\tau.$$

Replacing $e^{-j2\pi f\tau}$ by $(\cos 2\pi f\tau - j \sin 2\pi f\tau)$, and noting that the autocorrelation is an even function (i.e., $R_{zz}(\tau) = R_{zz}(-\tau)$), then

$$S_{zz}(f) = 2 \int_0^{T_x} R_{zz}(\tau) \cos 2\pi f\tau\, d\tau, \tag{52}$$

where $T_x = T_F$ or T_A accordingly. Substituting the various autocorrelation functions (44)-(49) for $R_{zz}(\tau)$ in (52) results in the functions (53)-(58), respectively, as shown below, which are the power spectral densities of the scrambled signals at the output of each scrambling algorithm, having speech $\boldsymbol{m}(t)$ as input.

Case 1a) Hopping Filters—Amplitude (Addition) Scrambling ($T_F \leq T_A$):

$$\begin{aligned} S_{zz}(f) = 2 \int_0^{T_F} &\left(R_{d_F}[0]\, R_m[0]\left(1 - \frac{|\tau|}{T_F}\right)(R_p(\tau)) \right. \\ &\left. + R_{d_A}[0]\left(1 - \frac{|\tau|}{T_A}\right)\right)(\cos 2\pi f\tau)\, d\tau \\ &+ 2 \int_{T_F}^{T_A} R_{d_A}[0]\left(1 - \frac{|\tau|}{T_A}\right)(\cos 2\pi f\tau)\, d\tau. \end{aligned} \tag{53}$$

Case 1b) Hopping Filters—Amplitude (Addition) Scrambling ($T_F \geq T_A$):

$$\begin{aligned} S_{zz}(f) = 2 \int_0^{T_A} &\left(R_{d_F}[0]\, R_m[0]\left(1 - \frac{|\tau|}{T_F}\right)(R_p(\tau)) \right. \\ &\left. + R_{d_A}[0]\left(1 - \frac{|\tau|}{T_A}\right)\right)(\cos 2\pi f\tau)\, d\tau \\ &+ 2 \int_{T_A}^{T_F} R_{d_F}[0]\, R_m[0]\left(1 - \frac{|\tau|}{T_F}\right) \\ &\cdot (R_p(\tau))(\cos 2\pi f\tau)\, d\tau. \end{aligned} \tag{54}$$

Case 2a) Hopping Filters—Amplitude (Multiplication) Scrambling ($T_F \leq T_A$):

$$\begin{aligned} 2 \int_0^{T_F} &R_{d_F}[0]\, R_{d_A}[0]\, R_m[0]\left(1 - \frac{|\tau|}{T_F}\right)\left(1 - \frac{|\tau|}{T_A}\right) \\ &\cdot (R_p(\tau))(\cos 2\pi f\tau)\, d\tau. \end{aligned} \tag{55}$$

Case 2b) Hopping Filters—Amplitude (Multiplication) Scrambling ($T_F \geq T_A$):

$$\begin{aligned} S_{zz}(f) = 2 \int_0^{T_A} &R_{d_F}[0]\, R_{d_A}[0]\, R_m[0]\left(1 - \frac{|\tau|}{T_F}\right) \\ &\cdot \left(1 - \frac{|\tau|}{T_A}\right)(R_p(\tau))(\cos 2\pi f\tau)\, d\tau. \end{aligned} \tag{56}$$

Case 3) Hopping Filters:

$$\begin{aligned} S_{zz}(f) = 2 \int_0^{T_F} &R_{d_F}[0]\, R_m[0]\left(1 - \frac{|\tau|}{T_F}\right) \\ &\cdot (R_p(\tau))(\cos 2\pi f\tau)\, d\tau. \end{aligned} \tag{57}$$

Case 4) Amplitude (Multiplication) Scrambling:

$$\begin{aligned} S_{zz} = 2 \int_0^{T_A} &R_{d_A}[0]\, R_m[0]\left(1 - \frac{|\tau|}{T_F}\right) \\ &\cdot (R_p(\tau))(\cos 2\pi f\tau)\, d\tau. \end{aligned} \tag{58}$$

Having the functions for the autocorrelations and power spectral densities of the outputs of the various scrambling algorithms, a program was written to compute the points necessary to plot these functions. In plotting the autocorrelation and power spectral density of the outputs of the various one-dimensional and two-dimensional algorithms, the objective is to find the best algorithm and the best possible combination of number of PAM levels for the PAM process(es) in this algorithm. The criteria in this case defining the best algorithm are, as mentioned previously, that the autocorrelation and power spectral density of its output scrambled signal resemble those of white noise.

The first comparison is carried out between the two two-dimensional algorithms. To simplify this comparison, let λ_i and λ_j as a pair (λ_i, λ_j) denote the number of levels in each of the seven identical PAM processes of the hopping filters algorithm, and the number of levels in the single PAM process of the amplitude scrambling algorithm, respectively. Since the number of levels of a particular PAM process can range from 2 to 16, the two two-dimensional algorithms are compared for the four extreme cases of $(\lambda_i, \lambda_j) = (2, 2)$, $(2, 16)$, $(16, 2)$, and $(16, 16)$ as shown in Figs. 8-11, respectively, while keeping the actual values of the levels in each case the same between the two algorithms. Furthermore, the autocorrelation functions are normalized by the average power of the output scrambed signals, and the power spectral density functions are normalized by the power spectral densities at $f = 0$, in each case accordingly. As a result, it is found that in all of the four cases outlined above, the statistics of the hopping filters amplitude (multiplication) scrambling algorithm approximate those of white noise more closely than do the statistics of the hopping filters amplitude (addition) scrambling algorithm. In fact, the same thing is found to apply irrespective of the actual PAM levels used in each case. The hopping filters amplitude (multiplication) scrambling algorithm is thus chosen as the better of the two two-dimensional algorithms compared.

Next, the various possible combinations of the number of levels for the better of the two two-dimensional algorithms chosen above are compared. Keeping the number of levels in the PAM processes of the hopping filters algorithm constant, while varying the levels of the amplitude scrambling PAM process, results in the plots of Figs. 12-15. In any case, it is seen that the autocorrelation and

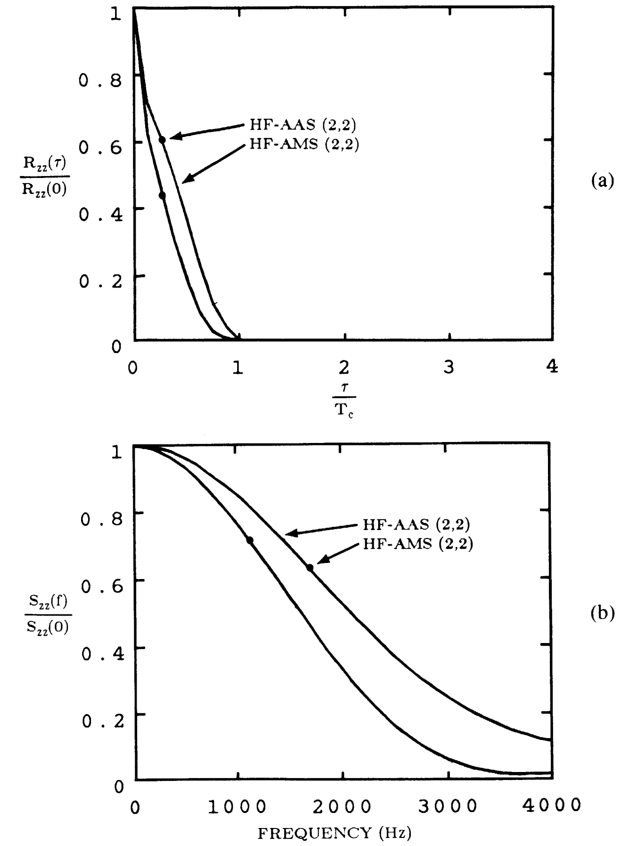


Fig. 8. Comparison of the two-dimensional algorithms using (2, 2). (a) Autocorrelation. (b) Power spectral density.

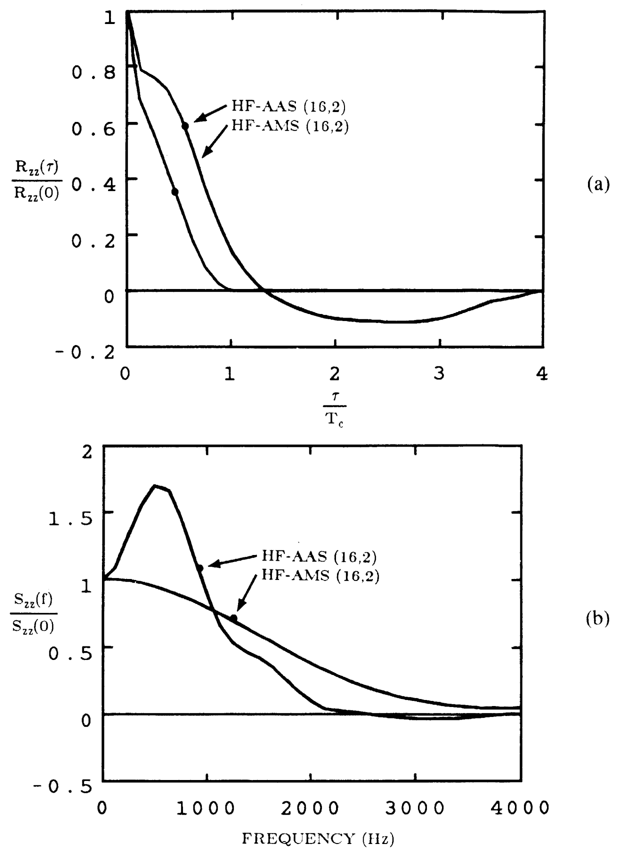


Fig. 10. Comparison of the two-dimensional algorithms using (16, 2). (a) Autocorrelation. (b) Power spectral density.

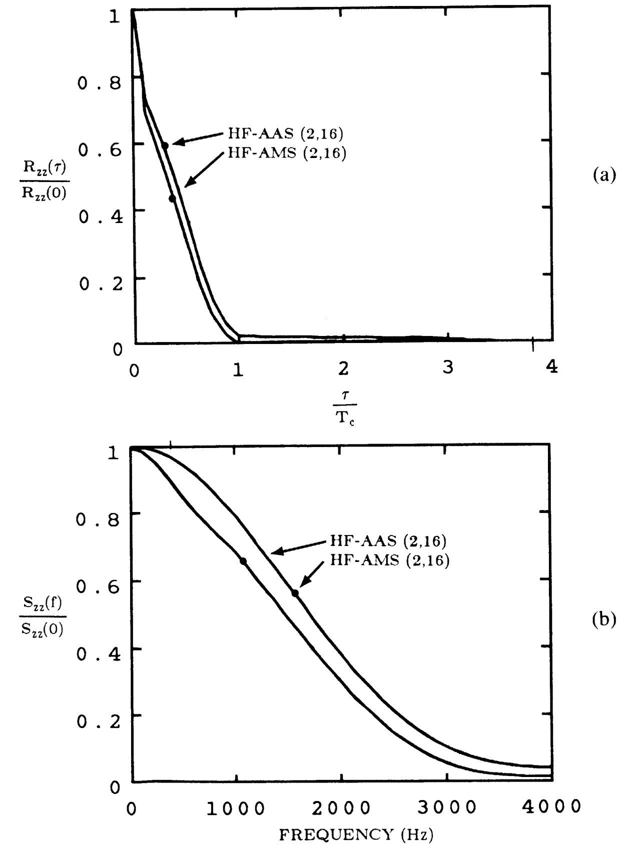


Fig. 9. Comparison of the two-dimensional algorithms using (2, 16). (a) Autocorrelation. (b) Power spectral density.

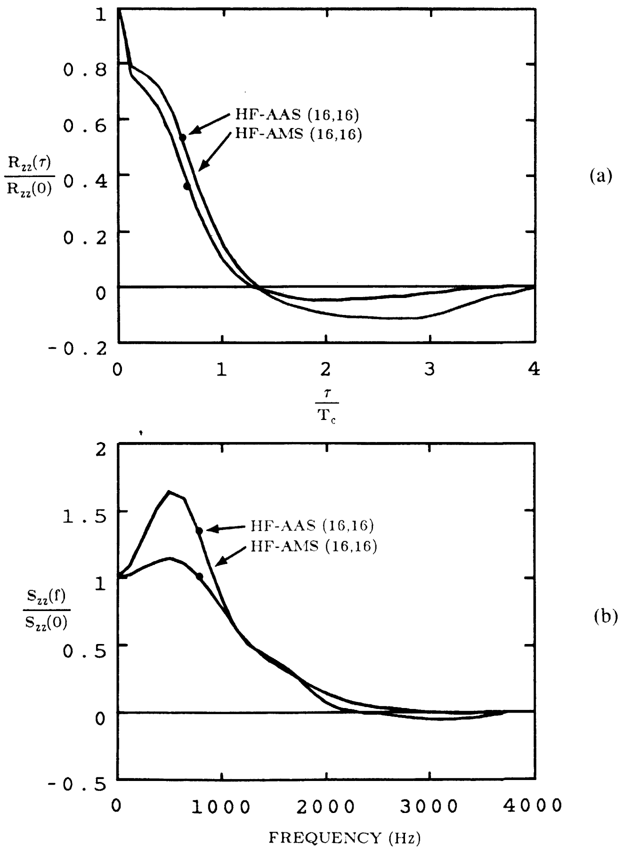


Fig. 11. Comparison of the two-dimensional algorithms using (16, 16). (a) Autocorrelation. (b) Power spectral density.

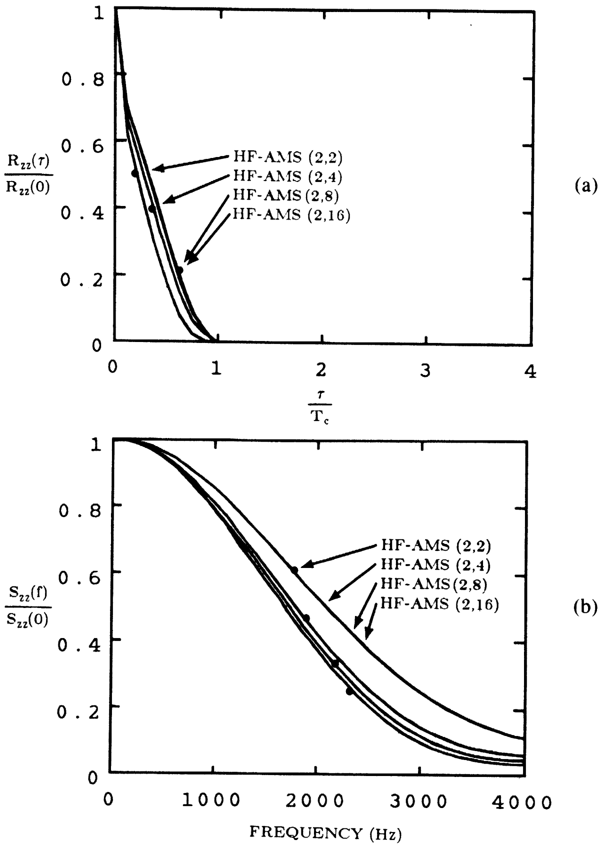


Fig. 12. Hopping filters—amplitude (multiplication) scrambling using (2, 2), (2, 4), (2, 8), (2, 16). (a) Autocorrelation. (b) Power spectral density.

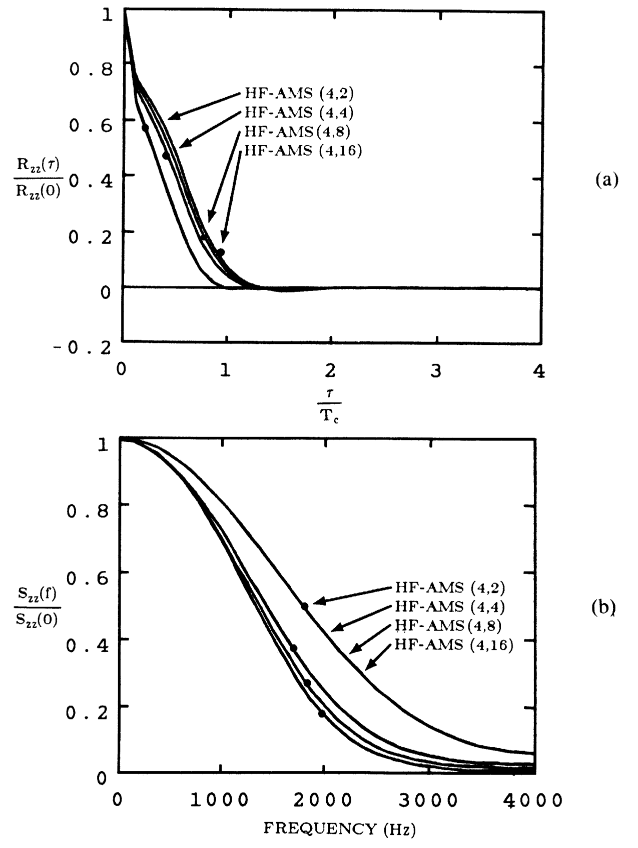


Fig. 13. Hopping filters—amplitude (multiplication) scrambling using (4, 2), (4, 4), (4, 8), (4, 16). (a) Autocorrelation. (b) Power spectral density.

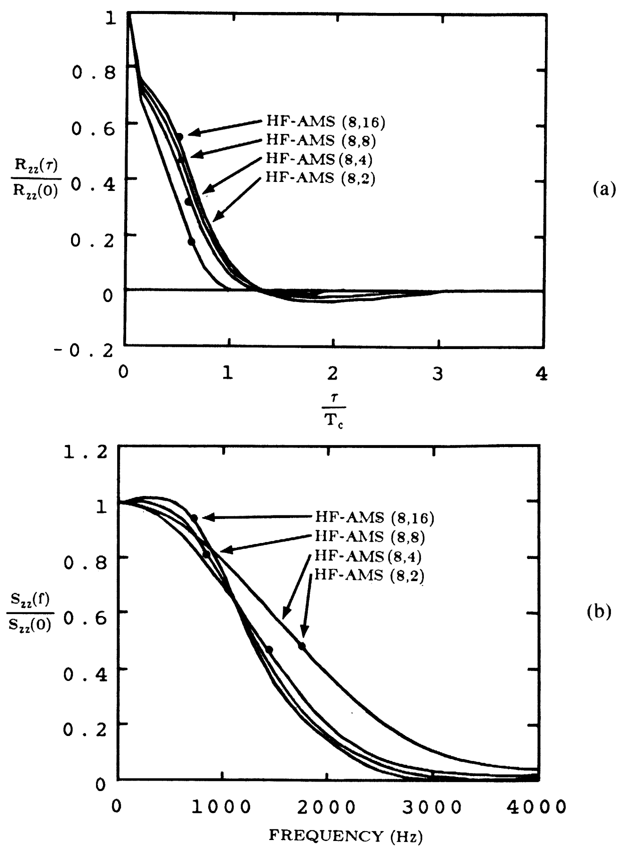


Fig. 14. Hopping filters—amplitude (multiplication) scrambling using (8, 2), (8, 4), (8, 8), (8, 16). (a) Autocorrelation. (b) Power spectral density.

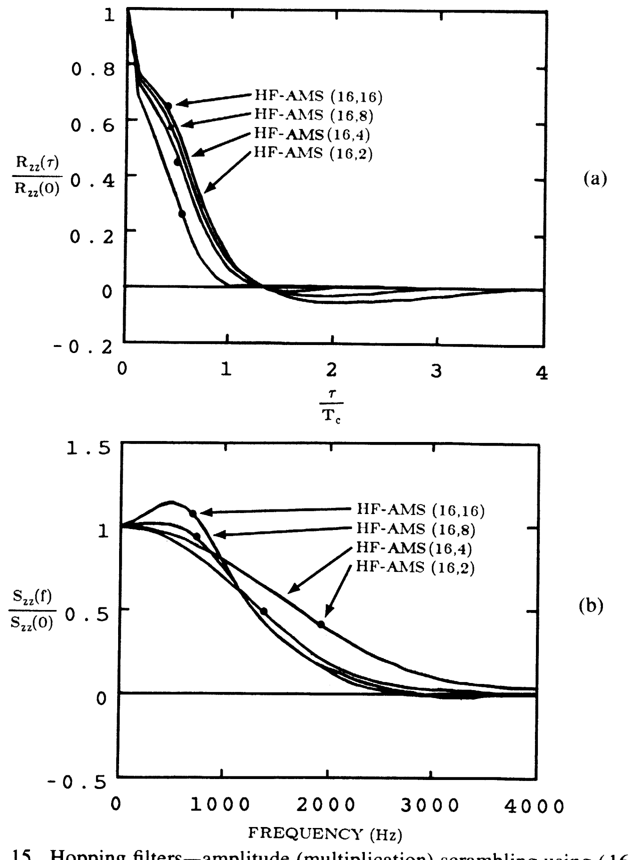


Fig. 15. Hopping filters—amplitude (multiplication) scrambling using (16, 2), (16, 4), (16, 8), (16, 16). (a) Autocorrelation. (b) Power spectral density.

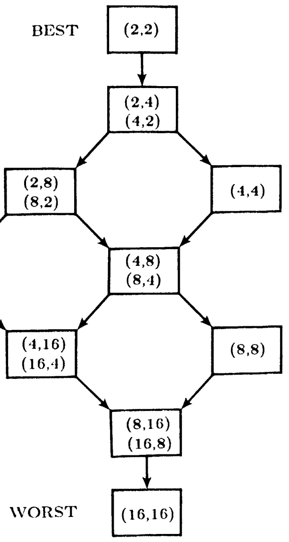


Fig. 16. Combinations resulting in best to worst statistics.

power spectral density of the output scrambled signal deviate more and more from those of white noise, as the number of levels of the amplitude scrambling PAM process increase from 2 to 16. Furthermore, this two-dimensional algorithm is such that the plots of (λ_i, λ_j) are identical to those of (λ_j, λ_i). As a result, keeping the number of levels of the amplitude scrambling PAM process constant, and varying the number of levels in the seven identical PAM processes of the hopping filters algorithm from 2 to 16, results in the statistics of the output scrambled signal deviating more and more from those of white noise.

The above results, concerning the two-dimensional hopping filters amplitude (multiplication) scrambling algorithm can be summarized by a directed graph (Fig. 16), where each combination is a node, and each directed edge joins two combinations pointing in the direction of the worst combination (i.e., combination which results in statistics furthest from those of white noise). From the directed graph, it is seen that the statistics of the combinations (2, 16) or (16, 2) and (4, 4) have to be further compared, and the statistics of the combinations (4, 16) or (16, 4) and (8, 8) have to be further compared. These comparisons are shown in Figs. 17 and 18, respectively. As a result, two more edges are added to the directed graph of Fig. 16, and redundant edges are removed resulting in the directed graph of Fig. 19. This shows all the possible distinct combinations for the number of levels of the PAM processes in the hopping filters amplitude (multiplication) scrambling algorithm, in the direction from best to worst, as governed by how the statistics of each compares to those of white noise.

Another arrangement of the possible combinations of Fig. 19 is from best to worst, with respect to the number of combined possibilities of hopping filters and amplitude scrambling levels (i.e., the more the better from the point of view of the designer), as shown in Fig. 20. Considering both of these arrangements, the only combinations which are among the top (half) best are (16, 2) and (8, 2). Since (16, 2) is statistically identical to (2, 16), (8,

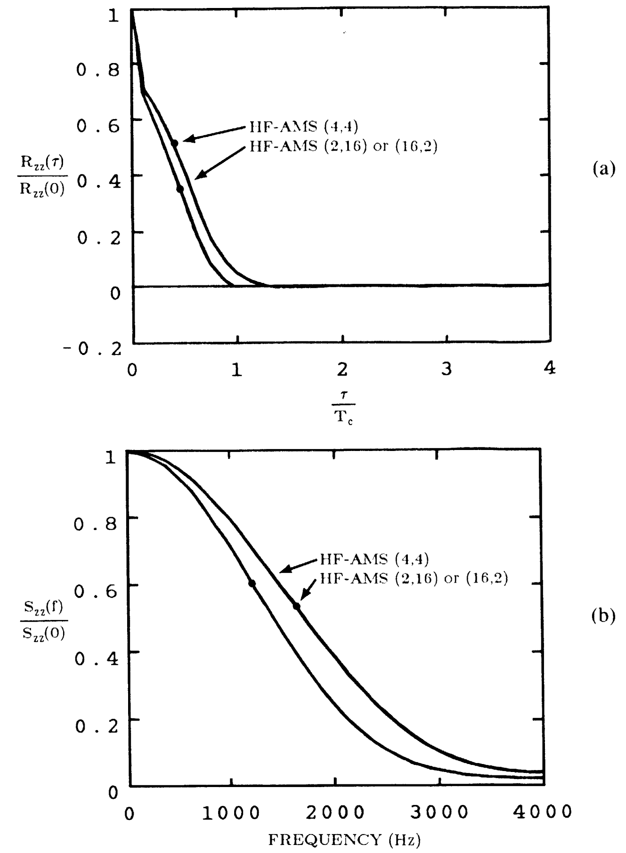


Fig. 17. Hopping filters—amplitude (multiplication) scrambling using (2, 16), (16, 2), (4, 4). (a) Autocorrelation. (b) Power spectral density.

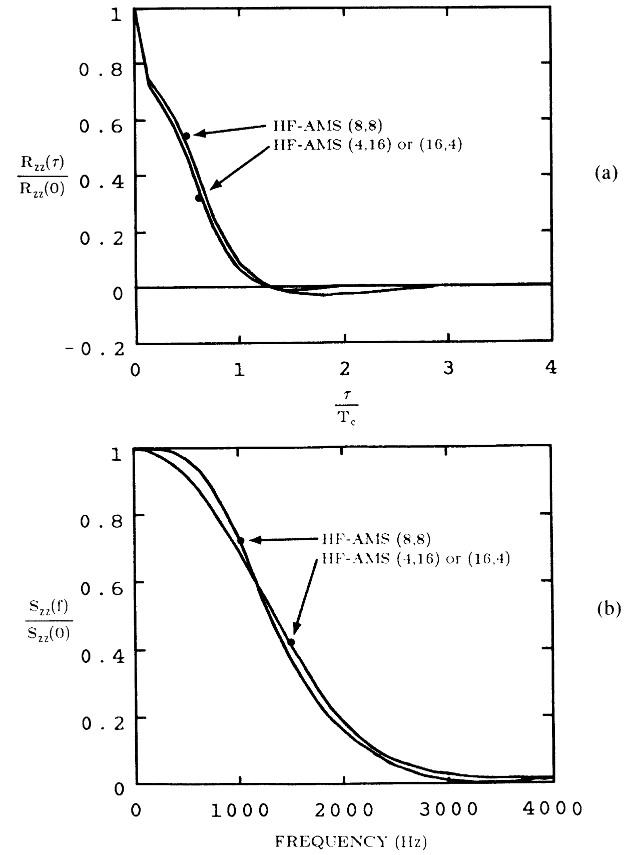


Fig. 18. Hopping filters—amplitude (multiplication) scrambling using (4, 16), (16, 4), (8, 8). (a) Autocorrelation. (b) Power spectral density.

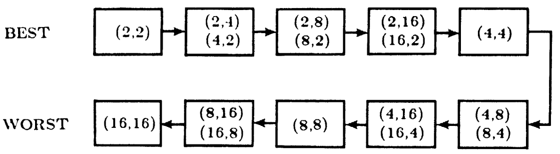


Fig. 19. Strictly directed graph of best to worst statistics.

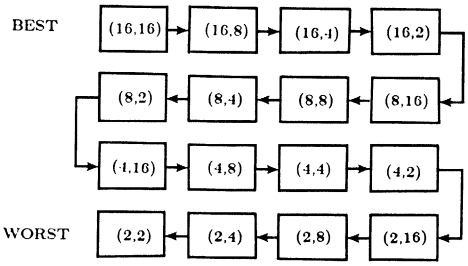


Fig. 20. Best to worst number of combined possibilities.

2) is statistically identical to (2, 8), and from Fig. 12 it is seen that the combinations (2, 16) and (2, 8) are statistically very close, then (16, 2) and (8, 2) are also considered statistically very close. Furthermore, since the combination (16, 2) provides more hopping filter possibilities, this is chosen as the best compromise combination of number of levels to define the two-dimensional algorithm already chosen.

In comparing the two one-dimensional algorithms (Fig. 21), it is seen that these are statistically identical. On the other hand, the hopping filters algorithm provides more possible hopping filters than the amplitude (multiplication) scrambling algorithm provides possible levels, for all cases of 2, 4, 8, and 16 PAM process levels. As a result, the hopping filters algorithm is chosen as the best one-dimensional algorithm. For this algorithm, it is seen (Fig. 21) that the smaller the number of PAM levels, the better the statistics of its output scrambled signal. On the other hand, it is known that the larger the number of PAM levels, the more the available hopping filters.

Finally, in comparing the hopping filters amplitude (multiplication) scrambling algorithm using the previously chosen combination (16, 2), with the hopping filters algorithm alone using any possible number of PAM levels, it is seen that the two-dimensional algorithm has better statistics than the best statistics possible (i.e., 2 PAM levels), with the hopping filters algorithm alone, as shown in Fig. 22. Furthermore, the two-dimensional algorithm with the chosen combination provides more combined possibilities of hopping filters, and amplitude scrambling levels, than the one-dimensional algorithm provides hopping filters with the maximum number of possible PAM levels (i.e., 16). As a result, the two-dimensional algorithm involving hopping filters with 16 PAM levels, and amplitude (multiplication) scrambling with 2 PAM levels, is found to be the best algorithm from those examined, providing a compromise between the desired statistics of its output scrambled signal, and the maximum number of possible hopping filters and amplitude scrambling level combinations. This algorithm is thus

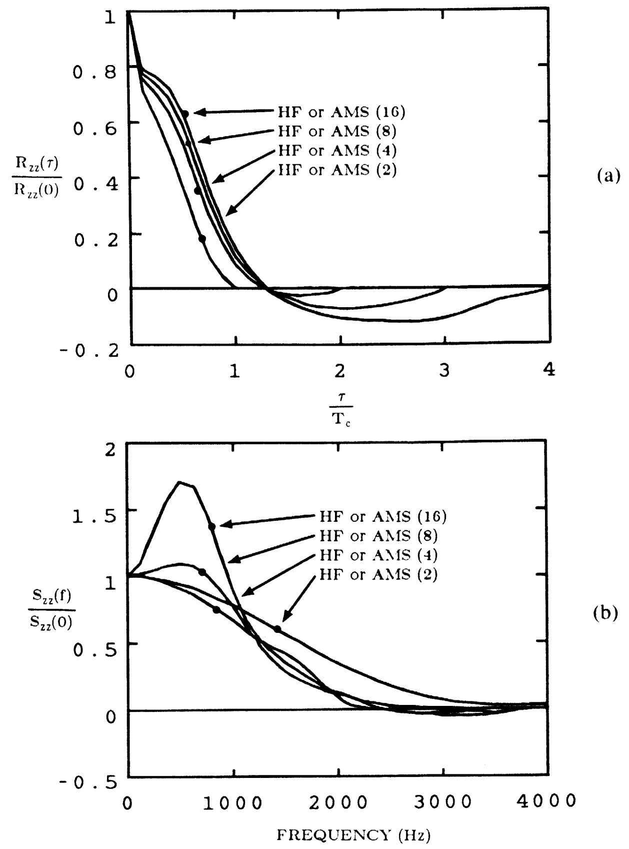


Fig. 21. Comparison of the one-dimensional algorithms using (2), (4), (8), (16). (a) Autocorrelation. (b) Power spectral density.

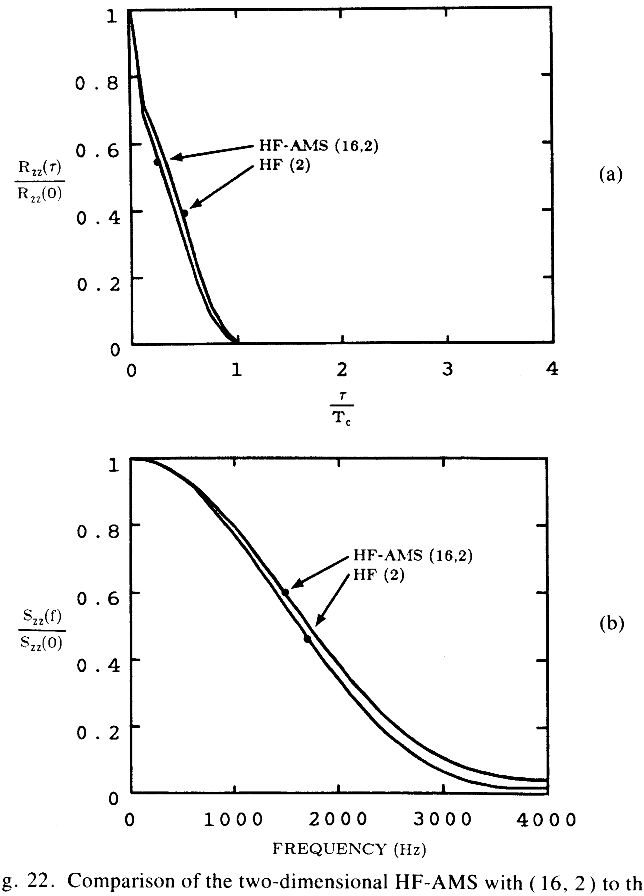


Fig. 22. Comparison of the two-dimensional HF-AMS with (16, 2) to the one-dimensional HF with (2). (a) Autocorrelation. (b) Power spectral density.

used as the basis of the secure device to be proposed in the next section.

IV. Hardware Design of a Voice Secure Device

The secure device proposed is one that will allow voice from a speaker (in analog form) to be scrambled before entering the telephone transmitter or the IF stage of an RF module, and conversely, voice coming from the telephone receiver (in analog form) to be descrambled before reaching the listener as shown in Fig. 23. This device will have both a normal mode and a secure mode, with the choice between the two given to the user through a manual switch located on the device. A conversation between two users is always initiated in normal mode. The two users select *a priori* two pairs of secret keys (one pair for transmission and one pair for reception) for one device, and the alternate combination of pairs of keys for the other device, such that the combination of two pairs of keys is unique for the combination of two users involved in the communication process.

The entire device (Fig. 24) consists of a transmitter unit and a receiver unit. Each such unit is further divided into three sections.

The first section consists of the scrambling and descrambling secure algorithms for the transmitter unit and receiver unit, respectively. The scrambling secure algorithm is the two-dimensional algorithm chosen in Section III. It consists of a hopping filters algorithm cascaded with an amplitude (multiplication) scrambling algorithm. Each of the seven PAM processes in the hopping filters algorithm has 16 levels resulting in 16^7 (or 268 435 456) possible hopped filters. This algorithm manipulates the incoming speech signal in the frequency domain, by passing this signal through 1 out of 268 435 456 hopped filters for a duration of $4T_c = 1.25$ ms, before possibly hopping to another filter. The resulting signal is then fed into an amplitude scrambling algorithm, which multiplies the incoming signal by 1 of 2 levels for a duration of $T_c = 0.3125$ ms, before possibly switching to the other level. The output of this algorithm constitutes the final output of the two-dimensional algorithm, which is the original input speech signal with its amplitude manipulated in both the frequency domain and time domain, thus scrambled. The descrambling secure algorithm is similar to the two-dimensional scrambling secure algorithm, with the exception that the order of the hopping filters and amplitude scrambling algorithms is reversed. Furthermore, in this case, the amplitude scrambling algorithm employs the inverse of multiplication, or division.

The second section consists of two arrays of 8 independent code generators (one array for each transmitter, and one array for each receiver unit of the device). These generators generate a pseudorandom secret code for each of the 8 PAM processes, respectively, of the two-dimensional scrambling and descrambling secure algorithms. Each individual generator (Fig. 25) is a multilayer feedforward generator as proposed in [7]. Basically, it consists of an eight-stage linear feedback shift register (LFSR), the taps of which are chosen so that its output sequence is a maximum length sequence of length $L = 2^8 - 1$ (= 255). The output of the LFSR is then fed into a shift register, which is tapped at certain positions to provide input to a stage of nonlinear logic, the output of which is fed into another shift register, and so on. There are alternately three shift registers and three stages of nonlinear logic needed, as a result of the 8 stages of the LFSR. It is imperative that there is at least a 2-input multiplier in each of the three nonlinear logic stages, such that

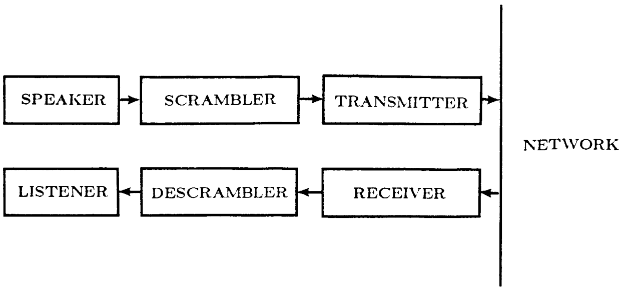


Fig. 23. Secure speech communication environment.

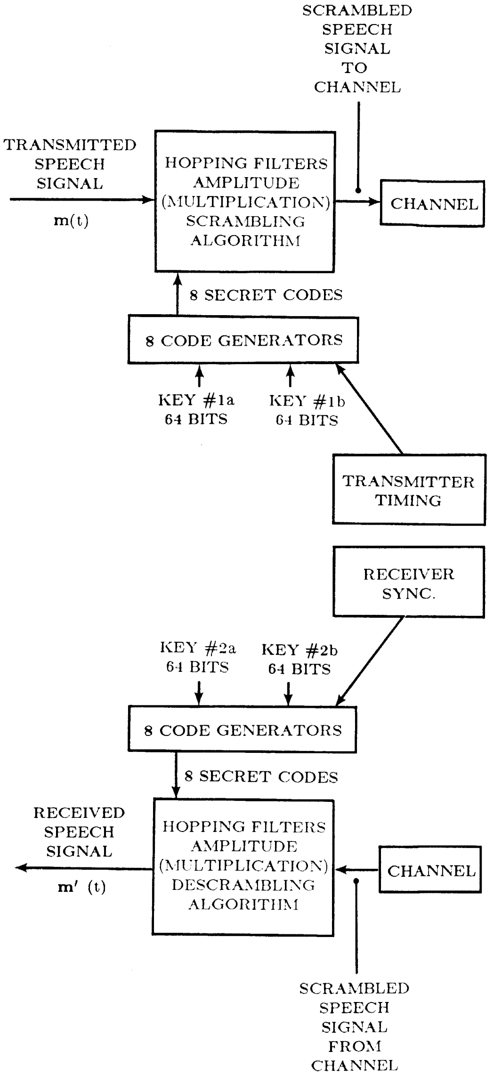

SCRAMBLED SPEECH SIGNAL TO CHANNEL
TRANSMITTED SPEECH SIGNAL
m(t)
HOPPING FILTERS AMPLITUDE (MULTIPLICATION) SCRAMBLING ALGORITHM
CHANNEL
8 SECRET CODES
8 CODE GENERATORS
KEY #1a 64 BITS
KEY #1b 64 BITS
TRANSMITTER TIMING
RECEIVER SYNC.
KEY #2a 64 BITS
KEY #2b 64 BITS
8 CODE GENERATORS
8 SECRET CODES
RECEIVED SPEECH SIGNAL
m' (t)
HOPPING FILTERS AMPLITUDE (MULTIPLICATION) DESCRAMBLING ALGORITHM
CHANNEL
SCRAMBLED SPEECH SIGNAL FROM CHANNEL

Fig. 24. Block diagram of a secure device.

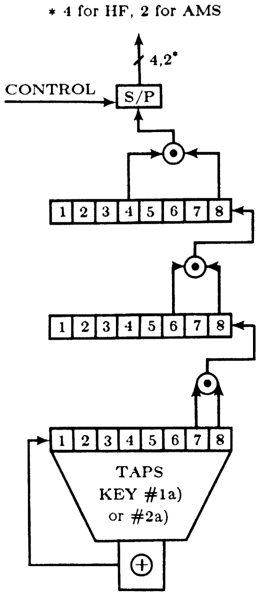


Fig. 25. Sample of a secure secret code generator.

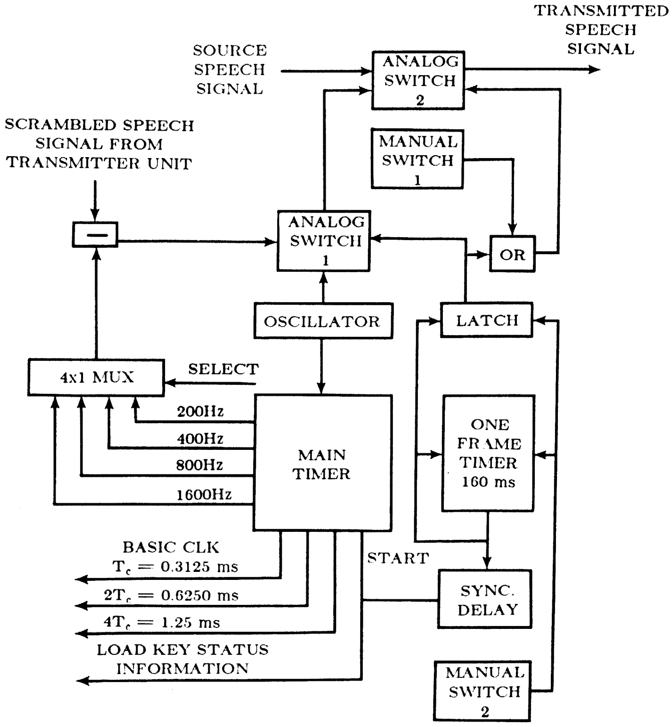


Fig. 26. Timing circuit in the transmitter unit.

the spans of these multipliers are 1, 2, and 4, for the first, second, and third nonlinear logic stages, respectively. From [7], it is known that this particular scheme increases the linear complexity or, more practically, the number of simultaneous equations that the cryptanalyst would need to solve, in order to find the configuration of one or more linear feedback shift registers, which would generate the same sequence as a particular multilayer feedforward generator. This increase in linear complexity goes from 8 (using only the basic LFSR) to 255 (using the multilayer feedforward generator). Furthermore, although the linear complexity of 255 can also be achieved by putting a two-input multiplier of span 1 at every layer, this would require a total of 7 layers as opposed to 3. The serial outputs of the first seven code generators are fed through 7 S/P converters of length 4, respectively. This accumulation of 4 bits is needed to choose among the 16 PAM levels of each of the 7 PAM processes. The output of the last code generator is fed as such to the amplitude scrambling PAM process it controls, since this process has only 2 levels, and thus requires only 1 bit to choose between these levels. This output through a 2-bit S/P converter is also used as control in the third section of this device. The set of code generators in the transmitter unit of one device, and the set of code generators in the receiver unit of another device communicating with the former, operate synchronously with a time delay between them. Each set of 8 generators on the alternate units of two secure devices make use of the same pair of 64-bit keys. The one key of the pair is used to set the 64 taps of the 8 combined 8-stage generators in each set, while the second key of the pair is used to initiate the eight combined code generators in each set.

The third section consists of the timing and synchronization circuitry needed to keep the operation of each receiver unit synchronized with the operation of each transmitter unit, on opposite secure devices. The successful communication of these two devices depends heavily on whether the secret code bit(s), which manipulate(s) the secure scrambling algorithm of one device at a particular instance in time, are exactly the same as the secret code bit(s), which manipulate(s) the secure descrambling algorithm of the other device a certain number of time units later, the time difference being due to processing delays and the delay for the signal to travel through the communications channel.

As shown in Fig. 26, the part of the third section which controls the timing in the transmitter unit consists of a subtractor, a 4 × 1 multiplexer, an oscillator, a main timer, a one-frame timer, two analog switches, and two manual switches. Part of this section manipulates the already scrambled signal, resulting from the scrambling algorithm, by subtracting from it one of four possible binary signals of various frequencies produced by the main timer. The frequencies of these signals are multiples of two from each other, and the choice between them is controlled through the 4 × 1 multiplexer, using two bits from an S/P converter of length 2 at the output of the code generator operating the amplitude scrambling PAM process. The first manual switch allows (through the second analog switch) the choice between the source analog signal, when the select is low, or its scrambled version when the select is high, to pass to the receiver. The second manual switch is pressed at the beginning, once an ordinary telephone circuit connection has been established between the two communicating parties. When pressed, it acts as a clock to the latch, whose input is wired high, and thus latches a high which selects the oscillator output through the first analog switch. The latched high disables the operation of the first manual switch through the OR gate,

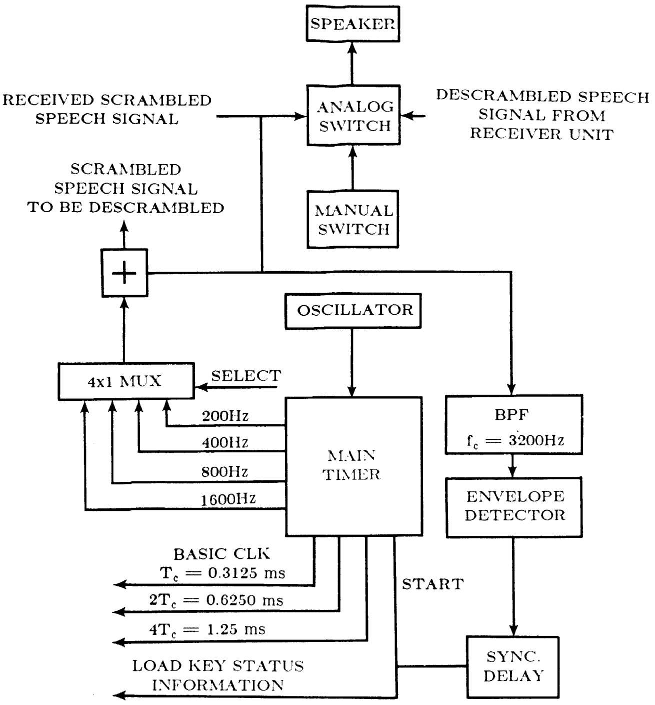


Fig. 27. Synchronization circuit in the receiver unit.

and selects the output of the first analog switch, which is the oscillator output in this case, to pass through the second analog switch as output of the transmitter unit, thus resulting in the transmission of a (SYNC) clock to the receiver. When pressed, the second manual switch also sends a pulse to the one-frame timer, which in turn produces an end-of-frame pulse after 160 ms. The latter pulse resets the latch to low, thus choosing to pass the scrambled analog signal through the first analog switch, and stopping the oscillator signal from being sent to the output of the transmitter unit after 160 ms. Furthermore, resetting the latch to low enables the first manual switch once again. From this point on, the first analog switch will remain with its select low until the manual switch is pressed once again. The end-of-frame pulse, besides resetting the latch, loads through a sync delay the initial 64-bit key on the 8 code generators, and starts or restarts the main timer. The oscillator output from the transmitter unit tells the receiver unit on the opposite secure device when to load its initial 64-bit key on its 8 code generators, and when to start or restart its main timer, so that both units on opposite devices are synchronized.

As shown in Fig. 27, the part of the third section which controls the timing and synchronization in the receiver unit mainly consists of an adder, a 4 × 1 multiplexer, a main timer, an oscillator, one manual and one analog switch, a bandpass filter, and an envelope detector. The manual switch is used to control the analog switch, in choosing between the scrambled signal, and descrambled signal to appear at the output of the receiver unit. Part of this section performs the descrambling procedure, corresponding to the scrambling procedure performed by the corresponding section in the transmitter unit. The rest of this section provides the synchronization for the receiver unit, by obtaining the oscillator output from the transmitter unit of the opposite device, bandpass filtering it at 3200 Hz, and passing it through an envelope detector and sync delay. This is then used to load the initial 64-bit key on the 8 code generators at the receiver, and start or restart the main timer (controller). Since the oscillator output from the transmitter unit of each device needs to be sent to the receiver unit of the opposite device, through the analog channel (i.e., speech band), it has to be restricted to the highest allowable possible frequency in the speech band which is 3200 Hz. Thus, the maximum frequency in either transmitter or receiver unit is bounded at 3200 Hz by the frequency of the oscillator.

The main timer in both transmitter and receiver units provides the basic clock $T_c = 0.3125$ ms (resulting from the maximum allowable frequency 3200 Hz) to operate the code generator. It also provides the clock of $2T_c = 0.6250$ ms and $4T_c = 1.25$ ms used to control the 2-bit and 4-bit S/P converters for the selection of timing signals in the third section of the device, and the selection of hopping filters in the first section of the device, respectively.

V. Conclusion

The security of various analog scrambling algorithms has been evaluated based on how statistically uncorrelated the output of each scrambling algorithm was to the input, and furthermore, on the degree to which the statistics of the output of each scrambling algorithm resembled those of white noise.

As a result, a new hybrid hopping filters/random amplitude (multiplication) scrambling algorithm was chosen as most secure. The security of this algorithm is valid under the constraint that all of its 8 PAM processes have zero mean. Under this constraint, it was found that the input and output of the algorithm are totally statistically uncorrelated. This constraint implies that the seven PAM processes of the hopping filters portion of this algorithm should be such that for every PAM level in a particular process, the additive inverse of that level should exist within the same process, and the multiplicative inverse of that level should exist within the corresponding descrambling PAM process. The actual values of these levels is not important.

Furthermore, it was found that this algorithm is best operated with 268 435 456 possible hopped filters and 2 possible levels for amplitude scrambling. This operation provides a compromise between the maximum possible security against an exhaustive attack (i.e., needing the most combinations of hopping filters and amplitude scrambling levels possible), and the maximum possible security against a direct statistical attack on the output of the algorithm (i.e., needing the lowest residual intelligibility). The latter is the result of a faster hopping rate between filters and between amplitude scrambling levels making the statistics of the output of the algorithm resemble those of white noise, which is totally unintelligible.

The two-dimensional algorithm resulting from the security evaluation as most secure, and operated as necessary, was used as the basis of a new proposed speech privacy/scrambling device, which scrambles and correspondingly descrambles speech signals in analog form. Furthermore, it uses the analog channel (i.e., a small part at the end of the speech band) so as to synchronize the descrambling receiver unit at one end of the channel with the scrambling transmitter unit at the other end.

A simplified prototype of the above algorithm was built and is currently tested. Phase linearity of the hopping filters, which is difficult to maintain, had some effect on the scrambled voice quality. Maintaining this phase linearity for all hopping filters is a challenging problem theoretically, and practically worth looking at.

References

[1] N. S. Jayant, "Analog scramblers for speech privacy," *Comput. Security*, vol. 1, pp. 275-289, 1982.

[2] J. L. Flanagan, M. R. Schroeder, B. S. Atal, R. E. Crochiere, N. S. Jayant, and J. M. Tribolet, "Speech coding," *IEEE Trans. Commun.*, vol. COM-27, pp. 710-737, Apr. 1979.

[3] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.* New York: McGraw-Hill, 1984, pp. 205-262.

[4] A. Gersho, "Perfect secrecy encryption of analog signals," *IEEE J. Select. Areas Commun.*, vol. SAC-2, pp. 460-466, May 1984.

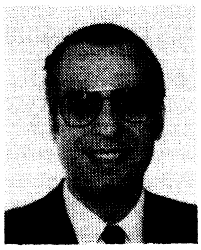
[5] S. Haykin, *Communication Systems.* New York: Wiley, 1983, pp. 229-314.

[6] L. W. Couch, II, *Digital and Analog Communication Systems.* New York: Macmillan, 1983, p. 297.

[7] E. J. Groth, "Generation of binary sequences with controllable complexity," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 288-296, May 1971.

Alex Goniotakis (S'89) was born in Montreal, Canada, on May 17, 1964. He received the B.Eng. degree in computer engineering from Concordia University, Montreal, Canada, in 1987, and is currently pursuing the M.Eng. degree in electrical engineering at Concordia University, to be completed by May 1990.

His interests are in secure speech communications and information security.



Ahmed K. Elhakeem (S'75-S'79-M'79-SM'86) received the Ph.D. degree in 1979 from Southern Methodist University (SMU), Dallas, TX.

At SMU his research interests were in the acquisition of spread spectrum signals, and jamming resistance by new test coding techniques. He spent the subsequent two years working overseas as a Visiting Professor. In 1982 he joined the Department of Electrical Engineering, University of Manitoba, Winnipeg, Man., Canada, and in 1983 moved to Concordia University, Montreal, P.Q., Canada, where he is now Associate Professor in the Department of Electrical and Computer Engineering. His current research interests are performance and reliability of integrated services (MAN's), acquisitionless spread spectrum systems, and secure speech communications.

Dr. Elhakeem was Technical Program Chair for IEEE Montech '86.

Acquisition Time Distribution for Spread-Spectrum Receivers

SI-MING PAN, DAVID E. DODDS, MEMBER, IEEE, AND SURINDER KUMAR, MEMBER, IEEE

Abstract—**This paper analyzes serial search spread-spectrum code acquisition. A modified flow graph is presented which permits the use of a generalized lock strategy and allows the false alarm time to be treated as a random variable. The distribution of acquisition time is obtained directly by using an extended generalization of Bernoulli trials. Compared to recent work which uses a fixed penalty false alarm time, the analysis is more general and requires fewer assumptions and approximations. The analysis is suitable for both single dwell and multidwell systems. Examples have been presented which show the effect on the acquisition time distribution of sequence length and normalized detection threshold.**

I. Introduction

In spread-spectrum communication systems, direct-sequence modulation is commonly used. To despread the received pseudonoise (PN) code, the direct-sequence spread-spectrum receiver generates a local replica of the PN code which is synchronized to the received PN code. The synchronization process operates in two modes: search mode and lock mode. When transmission is first received, the search mode brings the two PN codes into a coarse alignment. Once the system declares the coarse alignment found, the system enters the lock mode where the fine synchronization is achieved and maintained by code-tracking loops. Loss of synchronization is monitored by checking the condition of coarse alignment. Search mode will be reentered if this check fails repeatedly. The transition between search mode and lock mode is logically controlled by the receiver's search/lock (SL) strategy. To describe the operation in more detail, a typical search/lock strategy [1] is used as an example and illustrated in Fig. 1.

When a cell is tested in search mode and the integrate-and-dump output fails to exceed the detection threshold, the cell is rejected and the next cell is searched. On the other hand, the cell is accepted and retested if the integrate-and-dump output exceeds the detection threshold. If the second test (hit verification) is accepted, the system enters the lock mode. If the second test is a failure, the cell is dismissed as being incorrect and search continues at the next cell. After the lock mode has been entered, the system maintains the alignment. The system returns to search mode if three successive misses occur. To obtain rapid acquisition and robust true lock, the dwell time in the search mode (T_{d1}) is shorter than the dwell time in the hit verification (T_{d2}) and the dwell time in lock mode (T_{d3}). In practice [1]–[3], T_{d3} could be the same as T_{d2}. The time period of T_{d2} is called "verification/lock" time in this paper. The presence of a verification mode is usually termed as "multidwell-time detectors" as compared to "single-dwell-time detectors" which do not have a verification state [4]–[6]. For convenience, the search/lock (SL) strategy is redefined as the search/verification/lock (SVL) strategy in this paper to differentiate search and hit verification.

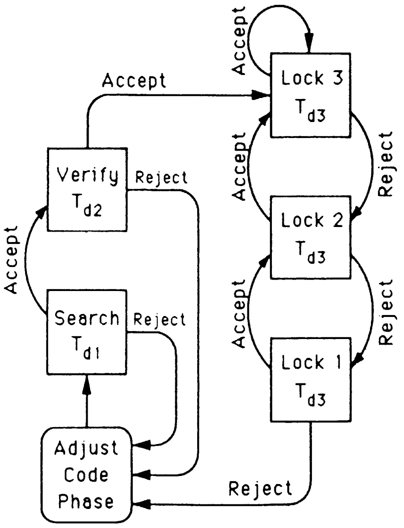


Fig. 1. An example of search/verification/lock strategy [1].

This paper evaluates the acquisition performance of direct-sequence spread-spectrum receiver with fixed dwell times and straight-serial-search strategies. The results of the paper may be extended to other search strategies, such as *Z* search and expanding-window search, which share the serial feature [2], [4], [7]. The acquisition performance is measured by the acquisition time (T_{ACQ}), which is defined as the time from starting of the search to finding of a correct code phase and then ending of the search. In the example shown in Fig. 1, the minimum value of T_{ACQ} is $T_{d1} + T_{d2}$, which occurs when search starts at the correct cell position and two successive tests are accepted. Acquisition time is a random variable which depends on

Manuscript received March 17, 1989; revised January 19, 1990. This work was presented in part at the International Conference on Communications, Boston, MA, June 1989.

The authors are with the Communication System Research Group, Department of Electrical Engineering, University of Saskatchewan, Saskatoon, Sask. S7N 0W0, Canada.

0733-8716/90/0600-0800$01.00 © 1990 IEEE

the starting position, the number of false alarms, and the number of times the true code phase is missed. The distribution of the acquisition time is developed in this paper.

Although the mean and variance of the acquisition time may provide a description of the acquisition process [1], [8], [9], a complete evaluation of the acquisition performance must utilize the probability distribution of the acquisition time. Some example distributions are shown at the end of the paper. For communication systems where catastrophe occurs when synchronization delay exceeds a certain value, the failure probability can be determined from the distribution.

Two approaches have been previously used to obtain statistics of a random time variable: time-domain techniques and transform-domain techniques. The flow graph (or characteristic function) technique, which was suggested by Holmes and Chen [1] and unified by Hopkins [10], Polydoros and Weber [4], is a transform-domain technique. When transform-domain techniques are used to obtain the distribution of the acquisition time, the Inverse Fourier Transform is generally required and only the numerical solution for a specific problem is provided. Combining the idea of algebraic characterization of the search with the transform-domain methods, a "direct" approach to obtain statistics of the acquisition time was recently presented by Jovanovic [2], [7]. He used the binomial approximation to the distribution of the partial acquisition time, instead of the Gaussian approximation used previously [11], [12]. In the previous works mentioned above, false alarm time was treated as a fixed "penalty" time and some approximations were necessary in the calculation of T_{ACQ}.

In this paper, the probability density function of acquisition time is obtained directly by using an extended generalization of Bernoulli trails. After the distribution is determined, the generating function (the discrete form of the characteristic function) is used to determine the mean and variance of the acquisition time. Even though the distribution will be directly developed, a modified flow graph is still used to illustrate the acquisition process and to help the analysis. Compared to the traditional flow graph techniques [1], [3], [8], our modified flow graph has two features: 1) a random variable x is introduced to present the starting cell position of the acquisition process and 2) a variety of different verification/lock (VL) strategies can be accommodated by introducing the concept of a "detour process." Throughout this paper, false alarm time is treated as a random variable instead of the fixed "penalty" time used previously. A random false alarm time is closer to reality and provides a possibility to precisely evaluate the effect of VL strategies to the acquisition time.

II. Flow Graph

In previous works [1], [3], [8]–[10], a Markov chain model was used to describe the acquisition process and the model is represented by a generating function flow graph. A simplified flow graph is used in this paper to characterize the Markov chain model. A uniformly distributed random variable x is introduced to represent the starting cell position of the acquisition process for straight-serial-search strategy. When no prior information about the position of the correct cell is available, this assumption is reasonable. For the SVL strategy mentioned above, a simplified flow graph is shown in Fig. 2.

In the flow graph, q denotes the number of possible code alignments and equals the number of cells to be tested in the serial search. The full code search time (T_q) is qT_{d1}. During q possible alignment positions, there must be one true code phase position (designated by $x = 0$) and ($q - 1$) positions at which the true code phase is not present and a false sync (false alarm) may occur. In the search mode, the false code phase may be accepted with probability P_{F1} or rejected with probability $1 - P_{F1}$. In the hit verification and lock states, the false sync (false alarm) may be maintained with probability P_{F2} or rejected with $1 - P_{F2}$. When the correct cell is tested in search mode, the true code phase is accepted with probability P_{D1} and accepted in hit verification with probability P_{D2}. The true code phase may be rejected with probability $1 - P_{D1}$ or fail verification with probability $1 - P_{D2}$, in which case the search process continues from the cell position furthest from the next true code phase position ($x = q - 1$). The probabilities P_{F1}, P_{F2}, P_{D1}, and P_{D2} may be calculated in terms of system parameters [1]. In the acquisition process flow graph, decisions are made after delay T_{d1} and T_{d2} which are indicated by $z^{T_{d1}}$ and $z^{T_{d2}}$, respectively. For the straight-serial-search strategy, the starting state is defined to be a random search state x, which is uniformly distributed from 0 to $q - 1$, while the absorbing state is defined to be the true lock state in Fig. 2.

At each of ($q - 1$) incorrect cell positions shown in Fig. 2, a false sync (or false alarm) may occur. In the flow graph, the false sync process is represented by a detour which is defined as a transition from a search mode cell position to one or more false sync states followed by a subsequent return to search mode at the next cell position. The detour process may include both hit verification and lock phases. Each detour is represented by subgraph starting at state 1 and exiting to state S in Fig. 2. The detour process is independent of the search process. The detour time, which is a component of the acquisition time, is defined to be the time from the incorrect acceptance of a false code phase through the failed verification and finally back to the search mode. Once the system enters a detour, false sync may be maintained for a random number of T_{d2} times before returning to search. Designating the random number by m, the detour time is represented as

$$T_{DE} = mT_{d2}, \qquad m = 1, 2, \cdots . \tag{1}$$

For a number of search/verification/lock strategies [2], [8], [9], the detour process may be fully characterized by the random variable m. The distribution, mean, and variance of m, designated by $P(m)$, $\overline{m}$, and Var (m), may be developed independently of the search process for a given

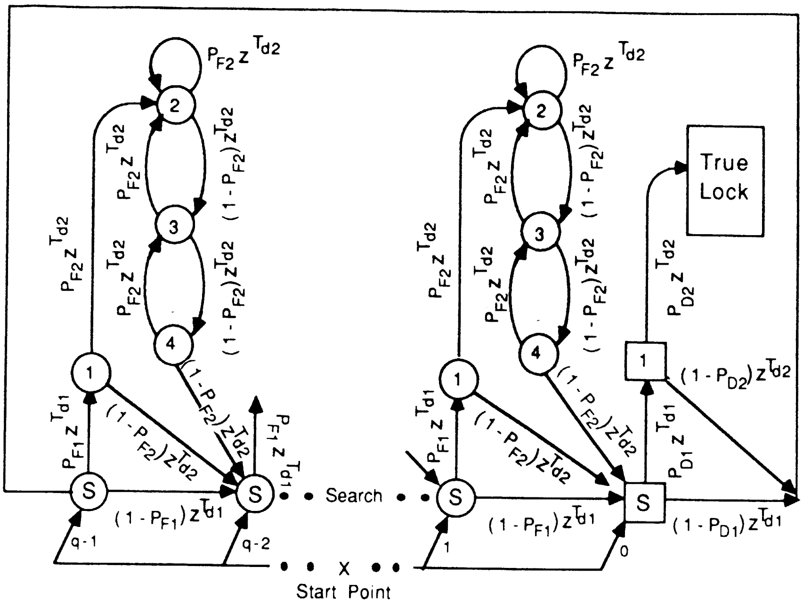


Fig. 2. Simplified flow graph for acquisition process.

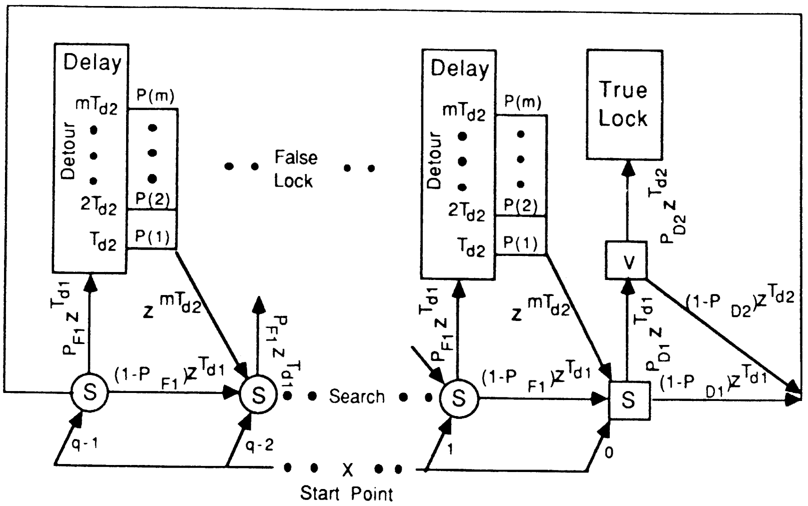


Fig. 3. Modified flow graph for acquisition process.

verification/lock strategy. This problem will be studied in Section IV.

Several different SVL strategies are currently used or have been proposed [3], [8]. These SVL strategies differ only in hit verification and lock strategy (VL). The analysis of the acquisition process is expected to accommodate a variety of different VL strategies. To handle this problem, no specific VL strategy is assumed and the distribution of the acquisition time will be developed in terms of a general detour time distribution $P(m)$. It is assumed that $P(m)$, $\overline{m}$, and Var (m) are known in the development of the acquisition performance. The resulting equations of acquisition time will be suitable for any VL strategy as long as the statistical characteristic of the detour time can be derived for that strategy.

Based on this idea, a modified flow graph of acquisition process is developed as shown in Fig. 3. A "white" box, which is characterized by $P(m)$, represents a detour process. The branches labeled $P(m)\ z^{mT_{d2}}$ indicate that the detour time of mT_{d2} delay occurs with probability $P(m)$.

III. The Acquisition Time Distribution

During the acquisition process, the system will spend $(x + 1)$ dwell times in search plus a random number (n_2) of verification/lock times in detour processes plus a random number (n_q) of full code search times which result from the true code phase being skipped. Consequently, the acquisition time is represented as: for single-dwell systems,

$$T_{ACQ} = n_2 T_{d2} + n_q T_q + (x + 1)T_{d1}; \qquad (2a)$$

and for multidwell systems,

$$T_{ACQ} = (n_2 + I)T_{d2} + n_q T_q + (x + 1)T_{d1}; \qquad (2b)$$

where

$$n_2 = 0, 1, 2, \cdots;$$
$$n_q = 0, 1, 2, \cdots;$$
$$x = 0, 1, 2, \cdots, q - 1.$$

Define $P(n_2, n_q, x)$ to be the joint probability function for n_2, n_q, and x. Because the acquisition time, T_{ACQ}, takes on discrete values depending on the combination of n_2, n_q, and x in (2), the probability distribution at T_{ACQ} will be known if $P(n_2, n_q, x)$ is found.

Since the density function of x is assumed to be uniform and equal to $1/q$, the joint probability function may be expressed in terms of the conditional probability function $P[(n_2, n_q) \mid x]$ as

$$P(n_2, n_q, x) = P[(n_2, n_q) \mid x] \frac{1}{q}. \tag{3}$$

The search process can be considered as three subprocesses: the first one, with probability $P(n_2 \mid x)$, starts at position x and ends at $x = 0$; while the second one, with probability $Q(n_2, n_q)$, starts at $x = 0$ and ends at $x = 0$ (the correct cell position) after n_q rejections of the true code phase; and the third one starts at $x = 0$ and ends at the true lock state. Corresponding to the first subprocess, the conditional probability function of n_2 verification/lock time slips starting at position x and reaching the first code phase position is defined as $P(n_2 \mid x)$. In the second subprocess, the probability of n_2 verification time slips occurring in the second subprocess with n_q rejections of the true code phase is defined as $Q(n_2, n_q)$. The third subprocess is simply given by the transition probability $P_{D1}P_{D2}$ and a fixed delay $(T_{d1} + T_{d2})$. This has been considered in the basic expression (2a) and (2b). The number of verification/lock time slips in the whole acquisition process is the sum of the verification/lock time slips resulting from both first two independent subprocesses. The probability function $P(n_2, n_q \mid x)$ can therefore be calculated by using a discrete convolution as

$$P[(n_2, n_q) \mid x] = P(n_2 \mid x) \otimes Q(n_2, n_q) P_{D1} P_{D2}. \tag{4}$$

The probability function $P(n_2 \mid x)$ is now developed. The development here is based on a common feature of the serial-search techniques, namely, they do not account for any additional information gathered during the past search time, and thus the tests in different cells are independent [3]. The probability is expressed in terms of the conditional probability, designated by $R[n_2 \mid k, P(m)]$, of n_2 verification/lock time slips given that k detours occur in the first subprocess. For a known detour time distribution, $P(m)$, and given k detours, the probability $R[n_2 \mid k, P(m)]$ may be developed by using an extended generalization of Bernoulli trails. Each detour is a trial whose outcome can be partitioned into an infinite number of events designated at $A_1, A_2, \cdots, A_m, \cdots$, with probabilities of occurrence $P(1), P(2), \cdots, P(m), \cdots$, which sum to unity.

If the number of detours with m verification/lock time slips is designated by l_m, the number of time that each event occurs in k detours is given by the series l_1, l_2, $\cdots$, l_m, $\cdots$, and the total number of detours is thus $k\ \Sigma_{m=1}^{\infty} l_m$. The number of verification/lock time slips in a detour is given by m, and thus the total number of slips is equal to the sum of the series $1l_1$, $2l_2$, $\cdots$, ml_m, $\cdots$, which is given by $n_2 = \Sigma_{m=1}^{\infty} ml_m$.

For a particular set of l_m's which satisfies the above two summations, the probability of observing n_2 verification/lock time slips equals

$$\frac{k!}{l_1! l_2! \cdots l_m! \cdots} P(1)^{l_1} P(2)^{l_2} \cdots P(m)^{l_m} \cdots.$$

For given values of n_2 and k, f designates all sets of l_m's which satisfy the previous two summations. Thus, the probability of observing n_2 verification time slips with k detours is the sum of the probability for all sets which satisfy the constraint f resulting in

$$R[n_1/k, P(m)] = \sum_f \frac{k!}{l_1! l_2! \cdots l_m! \cdots} \cdot P(1)^{l_1} P(2)^{l_2} \cdots P(m)^{l_m} \cdots. \tag{5}$$

For illustration, the conditional probability $R[n_2 \mid k, P(m)]$ is shown in Table I for a number of values of n_2 and k. An example of $n_2 = 7$ and $k = 3$ is given in [13].

Finally, the conditional probability of n_2 verification/lock time slips during a search process given starting position x and reaching position 0 (true code phase position), $P(n_2 \mid x)$, can be determined. To develop $P(n_2 \mid x)$, the conditional probabilities $R[n_2 \mid k, P(m)]$ must be weighted by the probability of k detours occurring during the first subprocess and summed over all $k \leq x$. For any synchronization path starting at position x and including k detours, the probability of k detours, occurring at any particular set of k cell positions (out of a possible x positions) is $(1 - P_{F1})^{x-k} P_{F1}^k$. Since the k cell positions can be chosen in $\binom{x}{k}$ different distinguishable ways, the required probability function $P(n_2 \mid x)$ is

$$P(n_2 \mid x) = \sum_{k=0}^{x} \binom{x}{k} (1 - P_{F1})^{x-k} \cdot P_{F1}^k R[n_2/k, P(m)]. \tag{6}$$

In the second subprocess, we define a recycle as a process which first rejects the true code phase at $x = 0$, subsequently searches all $q - 1$ cell positions, and finally reaches the correct cell position again at $x = 0$. The second subprocess can have number n_q of recycles ($n_q = 0$, $1, 2, \cdots$) as illustrated in Fig. 4. A recycle can be considered in two parts: the first one, with probability $T(n_2)$, starts at $x = 0$ and ends at the cell position furthest from the next true code phase position $(x = q - 1)$; while the second part, with probability $P[n_2 \mid (q - 1)]$, starts at $x = q - 1$ and ends at $x = 0$. Since these two parts are

TABLE I

Probability R of Slipping n_2 Verification Times as a Function of k Entries to Detour

	k				
n_2	0	1	2	3	4
0	1				
1	0	$P(1)$			
2	0	$P(2)$	$P(1)^2$		
3	0	$P(3)$	$2P(1)P(2)$	$P(1)^3$	
4	0	$P(4)$	$2P(1)P(3) + P(2)^2$	$3P(1)^2P(2)$	$P(1)^4$

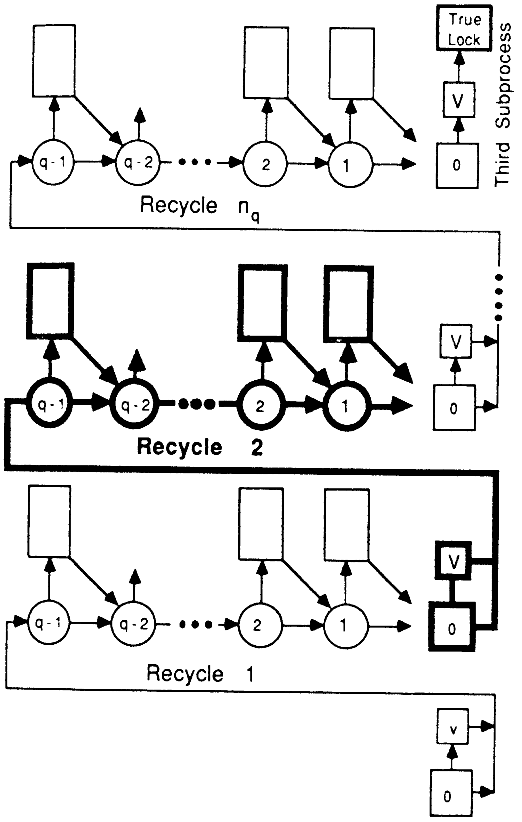


Fig. 4. Illustration of the second subprocess.

independent, the probability of n_2 verification/lock time slips occurring in a recycle is the discrete convolution of $T(n_2)$ and $P[n_2 \mid (q-1)]$. Clearly, probability $T(n_2)$ is equal to $1 - P_{D1}$ for $n_2 = 0$ and is equal to $P_{D1}(1 - P_{D2})$ for $n_2 = 1$. With a single verification state, other values of n_2 are not possible. This discrete convolution, which is designated by $G(n_2)$, is

$$G(0) = (1 - P_{D1}) P[0 \mid (q-1)]$$

$$G(1) = (1 - P_{D1}) P[1 \mid (q-1)] + P_{D1}(1 - P_{D2}) \cdot P[0 \mid (q-1)]$$

$$\cdots$$

$$G(n_2) = (1 - P_{D1}) P[n_2 \mid (q-1)] + P_{D1}(1 - P_{D2}) \cdot P[(n_2 - 1) \mid (q-1)]. \quad (7)$$

The second subprocess may consist of several recycles, each of which includes one full code search time (T_q) and a random number of verification/lock time slips (T_{d2}) in the false sync states. A single recycle is indicated by the bold portion of Fig. 4. By the definition, the probability of n_2 verification/lock time slips occurring in the second subprocess with n_q recycles is $Q(n_2, n_q)$. The probability $Q(n_2, n_q)$ may be developed using the same mathematical method as the conditional probability $R\{n_2 \mid k, P(m)\}$, as long as we imagine the recycle as a detour and replace $P(m)$ with $G(n_2)$. The probability function $Q(n_2, n_q)$ can be developed as

$$Q(n_2, n_q) = R[n_2 \mid n_q, G(n_2)], \quad (8)$$

where the form of R is given by (5).

The joint probability function $P(n_2, n_q, x)$ has been developed in terms of (3)–(8). The probability of the acquisition time is determined by $P(n_2, n_q, x)$ along with (2). Note that when T_{d2} is an integer multiple of T_{d1}, several combinations of n_2, n_q, and x may correspond to the same T_{ACQ}.

The mean of acquisition time has been derived directly from the developed distribution functions [5]. The mean, which is given by [5, eq. (A-15)], is consistent with [1, eq. (1.288)]. This partly verifies the above development of the distribution.

IV. Detour Time Distribution

The distribution of acquisition time has been determined assuming that the detour time distribution is known. For some verification/lock (VL) strategies, the distribution of detour time can be developed in closed form. For some complex verification/lock (VL) strategies, it is not simple to determine distributions of detour time in closed form. Fortunately, for most of the complex VL strategies which have been proposed, such as the 5 state strategy described in [9], [14], and generating function of detour time distribution can be easily written using Mason's formula. Then the distribution may be obtained by dividing out the generating function by long division, and the mean can be developed by differentiating the generating function and evaluating at $z = 1$.

For a VL strategy in which the number of the lock states is only 2 or 1, the distribution and mean of detour time have been previously developed in closed form [13]. For the VL strategy having 3 lock states as illustrated in Fig. 1, the flow graph of the detour process is shown as a subgraph in Fig. 2. Using a method similar to one used in [13], the distribution of the detour time is developed as

$$P(1) = (1 - p)$$

$$P(2) = 0$$

$$P(3) = 0$$

$$P(4) = p(1 - p)^3$$

$$\cdots \quad (9)$$

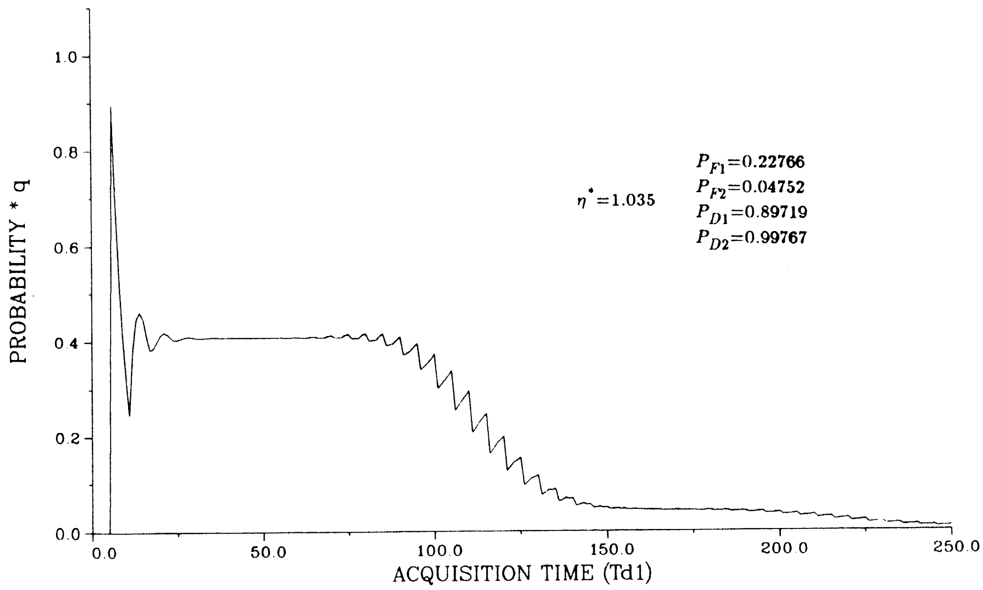


Fig. 5. Envelope of discrete pdf of T_{ACQ} for $\eta^* = 1.035$, $q = 50$, and one-state VL.

Note that the symbol p equals P_{F2} in all equations of this section.

Using Mason's formula on the detour process flow graph, the generating function of detour time is

$$P(z) = (1-p)z + \frac{p(1-p)^3 z^4}{1 - pz - 2p(1-p)z^2 + p^2(1-p)z^3}. \tag{10}$$

By differentiating and evaluating the first derivative of the above generating function, the mean of detour time has been found to be

$$\overline{T}_{DE} = \frac{1 - p^2 + p^3}{(1-p)^3} T_{d2}. \tag{11}$$

V. Calculation Example

Two specific examples are presented which illustrate the preceding theoretical development. The computational results indicate certain criteria to be used in system design. Given transition probabilities (P_{F1}, P_{F2}, P_{D1}, and P_{D2}) and the sequence length q, the discrete probability density function (pdf) of the acquisition time can be calculated. The cumulative distribution function of the acquisition time can easily be obtained by summing the discrete density function values.

The acquisition time discrete density functions have been computed for $q = 50$ and $q = 200$, and the results are presented in Figs. 5 and 6. Calculations for the case of $q = 50$ were performed according to the equations presented above. In the calculations for $q = 200$, a Gaussian approximation of the accumulated detour time component was used for acquisition time values greater than $50T_{d1}$. This was because the computational effort became excessive at larger acquisition time (see the Appendix). Figs. 5 and 6 present the envelope of the discrete density function where the horizontal axis represents the acquisition time discrete units of T_{d1}, while the vertical axis represents the probability, normalized by q, of each discrete acquisition time. The curves start at $T_{d1} + T_{d2} = 6T_{d1}$, which corresponds to the starting cell being in the correct code position. The normalized probability for this point is $P_{D1}P_{D2}$.

In Fig. 5, the "transient" at the beginning of the curve occurs because low values of T_{ACQ} preclude events which combine false alarm with straight serial search. Thus, the probability first decreases then increases to a stable value as more combinations of false alarm and straight search are possible. The zigzag behavior in the sloping part of the probability curve occurs when the true code phase is missed and results from the same factors which cause the "transient" at the start of the curve.

Fig. 6 presents the envelope of the discrete pdf for the case when $q = 200$. Each curve begins with a "transient" then reaches a discontinuity at $T_{ACQ} = 50T_{d1}$ after which the effect of the Gaussian approximation is evident. This results in the nonsmooth variations in flat portions of the curves. These approximation effects diminish and are barely observable above $500T_{d1}$. In Fig. 6, the normalized detection threshold η^* was considered as a variable parameter. For given system parameters of dwell times, T_{d1} and T_{d2}, and system bandwidth, the normalized detection threshold becomes an important design parameter which determines the transition probabilities [1], [6]. In the past, a "minimum mean" has often been used as the design criterion for selecting the normalized threshold. While this may be a good criterion for certain applications, it is not the correct choice if a certain cumulative probability (such as 0.99) is a system requirement. This important conclusion is drawn by considering the "tail" of the curves in Fig. 6 [6].

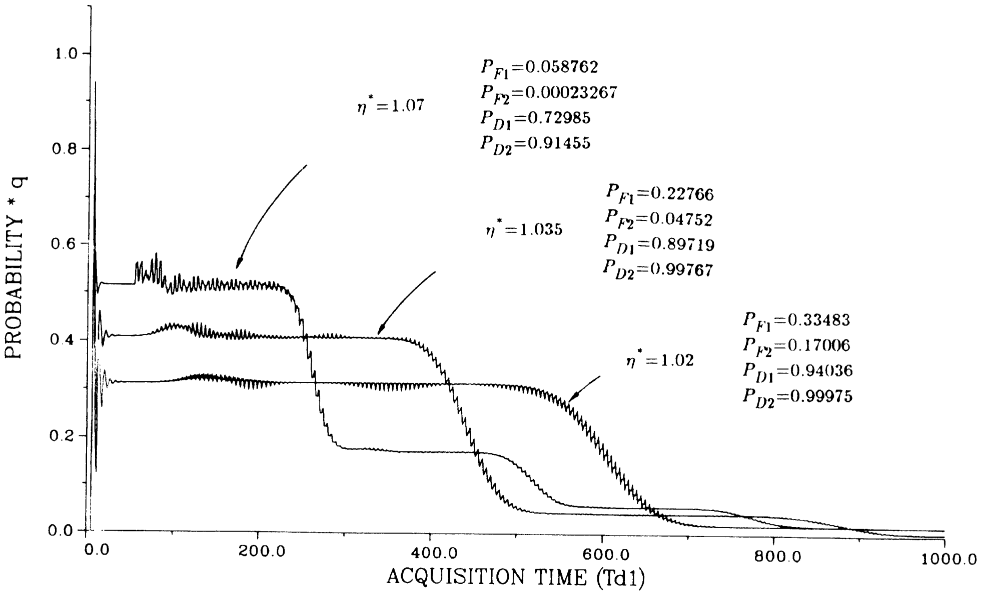


Fig. 6. Envelopes of discrete pdf of T_{ACQ} for $q = 200$ and one-state VL.

Although the example results are for a single verification/single lock strategy, theoretical developments given in the paper may be used for other VL strategies. Also, practical systems would use a higher value of q such as 1022 or 4094. To save computation time, smaller values of q were selected for our examples. By comparing curves for $q = 50$ and $q = 200$ when $\eta^* = 1.035$, it can be seen that a similar shape exists with only a change in scale as the value of q increases. Mean acquisition times for several VL strategies and $q = 4094$ have been presented in [5].

VI. Conclusion

This paper has developed a simplified flow graph technique to analyze the acquisition process of direct-sequence spread spectrum receiver. Based on the flow graph, the probability density function of the acquisition time has been developed in terms of a general verification/lock strategy. A "detour" time has been defined as the time spent in the verification and/or lock states at an incorrect code phase position. The analysis requires only knowledge of detour time distribution. Accordingly, the acquisition performance of a direct-sequence spread spectrum receiver can be fully evaluated. Examples have been presented which show the effect on the acquisition time distribution of sequence length and normalized detection threshold.

Appendix
Problems Associated with the Computation

The key step in the development of the acquisition time distribution is calculation of the conditional probability $R[n_2 \mid k, P(m)]$ in (5) and $R[n_2 \mid n_q, G(n_2)]$ in (8). These calculations require finding all nonnegative integer solutions of the following equation:

$$l_1 + l_2 + l_3 + \cdots = k$$
$$l_1 + 2l_2 + 3l_3 + \cdots = n_2 (n_2 \geq k). \quad \text{(A-1)}$$

Equation (A-1) is equivalent to the following two equations:

$$l_1 + l_2 = k - a_0$$
$$l_2 = n_2 - k - b_0, \quad \text{(A-2)}$$

where

$$a_0 = l_3 + l_4 + l_5 + \cdots$$
$$b_0 = 2l_3 + 3l_4 + 4l_5 + \cdots . \quad \text{(A-3)}$$

The solution requires that all pairs of a_0 and b_0 be selected subject to the following conditions:

$$0 \leq a_0 \leq k,$$
$$0 \leq b_0 \leq n_2 - k,$$
$$b_0 \geq 2a_0$$
$$b_0 \geq n_2 - 2k + a_0. \quad \text{(A-4)}$$

Each pair of a_0 and b_0 will indicate potential values of l_1 and l_2 in the solution. Values of l_1 and l_2 are determined by $l_1 = 2k - n_2 - a_0 + b_0$ and $l_2 = n_2 - k - b_0$. The terms a_0 and b_0 in (A-3) may be written as

$$l_3 + l_4 = a_0 - a_1$$
$$l_4 = b_0 - 2a_0 - b_1 \quad \text{(A-5)}$$

where

$$a_1 = l_5 + l_6 + l_7 + \cdots$$
$$b_1 = 2l_5 + 3l_6 + 4l_7 + \cdots . \qquad \text{(A-6)}$$

For each pair of a_0 and b_0 (and potential l_1 and l_2), candidates for l_3 and l_4 are found by selecting all pairs of a_1 and b_1 subject to the condition (A-7):

$$0 \le a_1 \le a_0,$$
$$0 \le b_1 \le b_0 - 2a_0,$$
$$b_1 \ge 2a_1$$
$$b_1 \ge b_0 - 3a_0 + a_1. \qquad \text{(A-7)}$$

Note that (A-6) has the identical form as (A-3). This means that the procedure can be identically performed until a solution can be obtained, or until it is certain that no solution exists. Since the algorithm presented here is a recursion, a program was written in Pascal to obtain all nonnegative integer solutions of (A-1), however, the computation time rapidly increases when the value of k increases. Fortunately, a simple closed-form approximation for $R[n_2 \mid k, P(m)]$ with large values of k can be developed as follows.

As shown by [5, eq. (A-6)], the generating function of $R[n_2 \mid k, P(m)]$ is

$$\sum_{n_2=0}^{\infty} z_2 n_2 R[n_2 \mid k, P(m)] = \left[\sum_{m=0}^{\infty} z_2^m P(m) \right]^k = P(z_2)^k.$$

where $P(z_2)$ is the generating function of $P(m)$.

Since the generating function of the conditional probability $R[n_2 \mid k, P(m)]$ is the product of k terms all equal to $P(z_2)$, the distribution of $R[n_2 \mid k, P(m)]$ will be the convolution of k distributions all equal to $P(m)$. The distribution $R[n_2 \mid k, P(m)]$ will be identical to the distribution of a random variable composed of the sum of k random independent variables each with distribution $P(m)$.

According to the Central Limit Theorem, the conditional probability $R[n_2 \mid k, P(m)]$ tends toward a nominal distribution as $k \to \infty$. Namely, for large k,

$$R[n_2 \mid k, P(m)] \approx \frac{1}{\sigma \sqrt{2\pi}} e^{-[(n_2 - \eta)^2 / 2\sigma^2]} \qquad \text{(A-8)}$$

where $\eta = k\overline{m}$ and $\sigma^2 = k \operatorname{Var}(m)$.

Similarly, the Gaussian approximation of the conditional probability $R[n_2 \mid n_q, G(n_2)]$ for large value of n_q may be developed if required. Note that the Gaussian approximation for the conditional probability $P(n_2 \mid x)$ may also be developed based on [5, eq. (A-8)]. The computation indicates that this relative error is quite small for normal system parameters when $k \ge 20$. This corresponds to $T_{ACQ} \ge 100 T_{d1}$. For the case of k less than 20, the probability $R[n_2 \mid k, P(m)]$ can be computed by using the recursive method developed in (A-1) to (A-7), or alternatively by using successive convolutions.

The successful Gaussian approximation of $R[n_2 \mid k, P(m)]$ indicates that, for large acquisition times, different distributions of detour time, $P(m)$, have little effect on the acquisition time distribution, as long as they have the same mean and variance. The exact name of verification/lock strategy has little effect on the tail of the acquisition time distribution curve.

References

[1] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications.* Boca Raton, FL: Computer Science Press, 1985.

[2] V. M. Jovanovic, "On the distribution of the acquisition time in spread-spectrum serial-search code acquisition," in *Proc. IEEE MILCOM'88*, New York, Oct. 1988, pp. 997–1001.

[3] J. K. Holmes and C. C. Chen, "Acquisition time performance of PN spread-spectrum systems," *IEEE Trans. Commun.*, vol. COM-25, pp. 778–783, Aug. 1977.

[4] A. Polydoros and C. L. Weber, "A unified approach to serial search spread spectrum acquisition—Part 1: General theory," *IEEE Trans. Commun.*, vol. COM-32, pp. 542–549, May 1984.

[5] S. M. Pan, D. E. Dodds, and S. Kumar, "Statistical distribution of PN acquisition time in direct-sequence spread spectrum receivers," in *Proc. ICC'89*, Boston, MA, June 1989, pp. 950–1000.

[6] D. E. Dodds, S. Kumar, and S. M. Pan, "Factors affecting serial search performance in spread spectrum systems," in *Proc. Canadian Conf. Elec. Comput. Eng.*, Montreal, Sept. 1989, pp. 388–391.

[7] V. M. Jovanovic, "Analysis of strategies for serial search spread-spectrum code acquisition—Direct approach," *IEEE Trans. Commun.*, vol. 36, pp. 1208–1220, Nov. 1988.

[8] J. K. Holmes, *Coherent Spread Spectrum Systems.* New York: Wiley, 1982.

[9] R. E. Ziemer and R. L. Peterson, *Digital Communications and Spread Spectrum Systems.* New York: Macmillan, 1985.

[10] P. M. Hopkins, "A unified analysis of pseudonoise synchronization by envelope correlation," *IEEE Trans. Commun.*, vol. COM-25, pp. 770–777, Aug. 1977.

[11] W. R. Braun, "Performance analysis for the expanding search spread spectrum PN acquisition algorithms," *IEEE Trans. Commun.*, vol. COM-30, pp. 424–435, Mar. 1982.

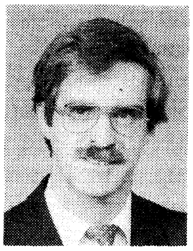
[12] A. Weinberg, "Generalized analysis for the evaluation of search strategies effects on PN acquisition performance," *IEEE Trans. Commun.*, vol. COM-31, pp. 37–49, Jan. 1983.

[13] D. E. Dodds, S. M. Pan, and A. G. Wacker, "Statistical distribution of PCM framing times," *IEEE Trans. Commun.*, vol. 36, pp. 1236–1241, Nov. 1988.

[14] W. K. Alem *et al.*, "Spread spectrum acquisition and tracking performance for shuttle communication links," *IEEE Trans. Commun.*, vol. COM-26, pp. 1689–1703, Nov. 1978.

Si-Ming Pan was born in Peking, China, in 1945. He graduated from the University Quing Hua, Peking, in 1968. He received the M.Sc. degree in electromagnetic compatibility from the Post and Telecommunication Academy, Peking, in 1981, and the M.Sc. and the Ph.D. degrees in digital transmission from the University of Saskatchewan, Canada, in 1985 and 1988, respectively. He is now a Postdoctoral Fellow at the University of Saskatchewan.

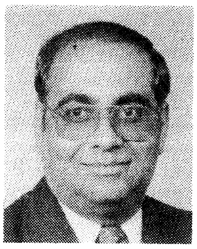
From 1969 to 1973 he worked in the Jie Lin Electrical Instrumentation Factory, and from 1973 to 1978 he worked in the Peking General Petrochemical Works. He was a Research Engineer of the Post and Telecommunication Academy in Peking before coming to Canada. He is interested in digital transmission, voice processing, spread-spectrum communication, and electromagnetics.



David E. Dodds (M'70) was born in Saskatoon, Canada, in 1945. He received the B.Eng. and M.Sc. degrees from the University of Saskatchewan in 1966 and 1968, respectively.

He worked with Bell Northern Research (Ottawa) on PBX design in 1969, 1972, and 1978. Since 1969 he has been with the University of Saskatchewan where he is currently Professor of Electrical Engineering. He teaches courses in electronics and communications, and also presents short courses on the fundamentals of telephony. Current research interests are in data transmission, signal processing architecture, and frame synchronization for PCM and spread spectrum systems. He has been granted five patents relating to commercial products in delta-modulation, telephone line interfacing, and computer data transmission. In 1986 he was on sabbatical leave at BNR Inc. and worked with an ANSI committee on the standardization of FDDI-II, a combined voice and data system.

Mr. Dodds is a Registered Consulting Engineer in Saskatchewan.



Surinder Kumar (M'88) received the B.E. degree from the Indian Institute of Science, Bangalore, India, and M.Tech. degree from the Indian Institute of Technology, Kanpur, India. He received the Ph.D. degree from Carleton University, Ottawa, Canada, where he was a Commonwealth Scholar.

He worked for about ten years with a government research lab in India. From 1982–1987 he was with SED Systems, Saskatoon, Canada, where as Vice President Research he lead teams involved in design of a wide variety of satellite two-way and TVRO earth stations. At present he is with the Department of Electrical Engineering, University of Saskatchewan, Saskatoon, Canada, where he is a Research Professor appointed to a Chair in Communications. The Chair is sponsored by Sask Tel and NSERC Canada. He is involved in setting up a Center for Communication Studies. His interests are in microwave-communications systems and circuits, as well as in digital modulation methods and hardware. He has a number of publications in these areas, and is a consultant to a number of companies.

Mean Time to Lose Lock for a Coherent Second-Order PN-Code Tracking Loop—The Singular Perturbation Approach

ARNOLD L. WELTI, STUDENT MEMBER, IEEE, AND BEN-ZION BOBROVSKY, MEMBER, IEEE

***Abstract*—The singular perturbation method is used to approximate the mean time to lose lock (MTLL) for a second-order coherent pseudonoise code tracking delay-lock loop (DLL). Approximate expressions for the MTLL are given. The influence of loop "offset" due to Doppler rate is studied, and optimal loop parameters (namely, natural frequency and damping factor) which maximize the MTLL are presented. Upper bounds to the maximum allowable Doppler rates for various operating conditions are given.**

I. Introduction

The mean time to lose lock (MTLL) is well known to be an important design objective for various tracking loops [12], [14], [23]. The MTLL of coherent and noncoherent pseudonoise code tracking loops was studied extensively in the literature [8], [16], [20]. While the calculation of the MTLL yields explicit, although complicated results for first-order loops, the second-order case is by far more difficult to analyze. We propose in this paper a different approach from those used in the engineering literature to approximate the MTLL. Following the observation made in [2], the problem at hand is shown to be a stable dynamical system excited by only small intensity noise. This is true even near the threshold. Thus, the singular perturbation method—originally used by Kramers [11] for similar problems—can be applied. Extensions of Kramers' results related to our case were given in [13], [18], [19]; see also [5]–[7]. The advantage of this approach is that there is no need to find a solution to the general second-order case (which is unknown today), once recognizing that all systems of interest can be characterized as "small noise" systems (using proper time scaling). There is no need to use ad hoc approximations. Instead, an asymptotic expansion of the solution is given. Applications of a singular perturbation method to similar engineering systems were given in [1], [2], [10], [24]. In this paper, we give exact analytical expressions for the leading order term of the MTLL of a coherent tracking loop.

The equations of the coherent loop under Doppler rate are given in Section II. The exit problem, the domain of attraction, and its boundary are discussed in Section III. The computation of the MTLL using singular perturbation is given in Section IV, where an alternative simple method is also considered. Optimal loop design is suggested in Section V. Discussion and comparison to a case study of a GPS (global positioning system) tracker are given in Section VI.

II. Loop Equations

The received signal is modeled as

$$r(t) = \sqrt{P}c(t - T) + \sqrt{\frac{N_0}{2}}\, n(t) \tag{1}$$

where P is the average signal power, $c(t)$ is the spreading code with code rate R_c, and T is the instantaneous propagation delay time of the channel. The additive noise term has a two-sided power spectral density of $N_0/2$ and $n(t)$ is standard white Gaussian noise.

It is well known [28] that the coherent delay-lock loop of Fig. 1 can be modeled by the "baseband equivalent," Fig. 2, where ϵ is the normalized timing error given by

$$\epsilon(t) = \frac{T(t) - \hat{T}(t)}{T_c},$$

$\hat{T}(t)$ is the loop estimate of $T(t)$, and $T_c = 1/R_c$ is the code-chip time. Note that $T(t)/T_c$ is the loop input which is being tracked by the loop output $\hat{T}(t)/T_c$. The loop filter is described by $F(s)$, and the nonlinearity $S(\epsilon)$ (S-curve or discriminator characteristic) is given in Fig. 3.

It can be shown [28] that

$$\dot{\epsilon}(t) = \frac{\dot{T}(t)}{T_c} - \sqrt{P}KF(s)\left\{S(\epsilon) + \sqrt{\frac{N_0}{P}}\,\dot{w}(t)\right\} \tag{2}$$

where $\dot{w}(t)$ is an additive standard white Gaussian noise and K represents the combined discriminator and VCO gains. In this paper, we consider $F(s) = 1 + K_1/s$. Thus, the loop is of second order with two perfect integrators. As is well known [23], such a loop is capable of tracking

Manuscript received February 24, 1989; revised December 8, 1989. A preliminary version of this paper was presented in part at the IEEE International Conference on Systems Engineering, Dayton, OH, August 1989.

A. L. Welti is with the Institute for Communication Technology, Swiss Federal Institute of Technology, ETH Zentrum, 8092 Zürich, Switzerland.

B. Z. Bobrovsky was on leave at the Institute for Communication Technology, Swiss Federal Institute of Technology, ETH Zentrum, 8092 Zürich, Switzerland. He is with the Department of Electrical Engineering, Tel Aviv University, Tel Aviv, 69978 Israel.

IEEE Log Number 9035808.

0733-8716/90/0600-0809$01.00 © 1990 IEEE

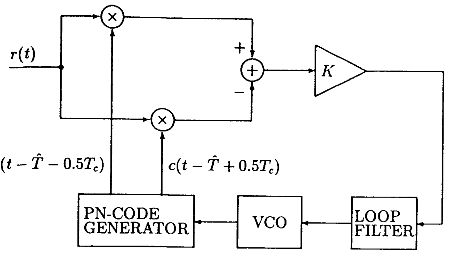


Fig. 1. Coherent PN-code tracking loop.

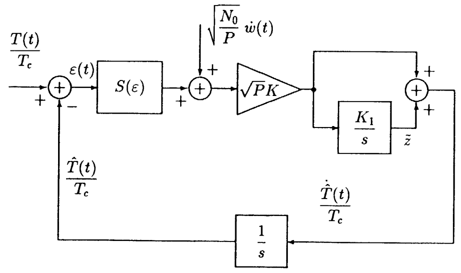


Fig. 2. Second-order baseband equivalent model.

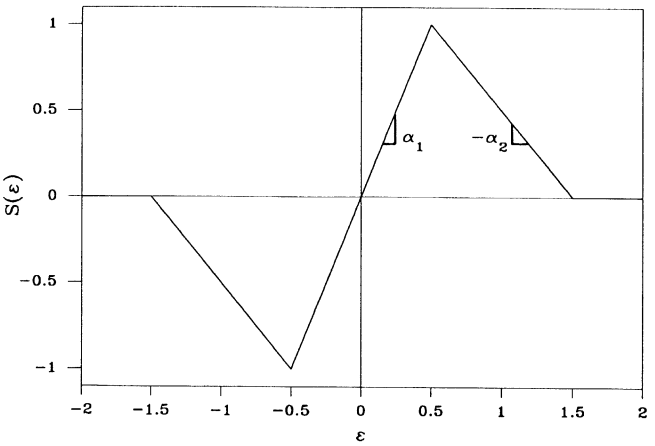


Fig. 3. The discriminator characteristic (S curve).

delays of the type

$$\frac{T(t)}{T_c} = a_0 + a_1 t + \frac{a_2}{2} t^2. \tag{3}$$

The delay T can vary in time, for example, due to a velocity $v(t)$ of the receiver relative to the transmitter (Doppler shift). Let $v(t) = v_0 + \dot{v}t$; in that case,

$$T(t) = a_0 T_c + \frac{v_0}{c} t + \frac{\dot{v}}{2c} t^2 \tag{4}$$

where c is the velocity of light, so

$$a_2 = \frac{\dot{v}R_c}{c}. \tag{5}$$

The loop equations can be written in state-space form (the auxiliary variable $\tilde{z}$ is denoted in Fig. 2) as

$$\left.\begin{aligned} \dot{\epsilon} &= -\tilde{z} + \frac{\dot{T}}{T_c} - \sqrt{P}KS(\epsilon) - K\sqrt{N_0}\dot{w}(t) \\ \dot{\tilde{z}} &= \sqrt{P}KK_1S(\epsilon) + KK_1\sqrt{N_0}\dot{w}(t) \end{aligned}\right\} \tag{6}$$

or, using $z = \tilde{z} - (\dot{T}/T_c) = \tilde{z} - a_1 - a_2 t$,

$$\left.\begin{aligned} \dot{\epsilon} &= -z - \sqrt{P}KS(\epsilon) - K\sqrt{N_0}\dot{w}(t) \\ \dot{z} &= -a_2 + \sqrt{P}KK_1S(\epsilon) + KK_1\sqrt{N_0}\dot{w}(t). \end{aligned}\right\} \tag{7}$$

Let

$$\omega_n = \sqrt{S'(0)KK_1\sqrt{P}}$$

$$\zeta = \frac{1}{2}\sqrt{S'(0)K\frac{\sqrt{P}}{K_1}}.$$

For operating conditions when linearization is valid, ω_n is the system natural radian frequency and ζ is its damping factor. The maximal $|a_2|$ of the deterministic system ($N_0 = 0$) such that steady state is still possible is given by

$$0 = \dot{z} = -|a_2|_{\max} + \sqrt{P}KK_1|S(\epsilon)|_{\max};$$

hence, since $|S(\epsilon)|_{\max} = 1$ (Fig. 3),

$$|a_2|_{\max} = \sqrt{P}KK_1 = \frac{\omega_n^2}{S'(0)}. \tag{8}$$

Equations (7) represent a dynamical system excited by a white Gaussian noise. Strictly speaking, the "white noise" in (1) represents "physical noise" (thus having a "wide but finite" bandwidth); therefore (7) should be understood in the Stratonovich sense [27]. However, since the coefficients of the noise in this coherent case are constant, the Wong-Zakai correction [27] vanishes (this is not true for the noncoherent case).

In general, analysis of nonlinear systems like (7) might be very difficult. An observation [2] that simplifies the case by far is that using proper scaling, (7) can be viewed as a dynamical system excited by only "small noise." This seems to be quite surprising, especially for systems operating near the threshold. On second thought, however, it does make sense, since for high noise excitation, the drift term in (7) can almost be neglected compared to the noise term $\dot{w}(t)$, resulting in a system without restoring force—a situation of no practical interest.

Following [1], [2], we scale the time

$$t^* = \alpha t = \sqrt{P}Kt = \frac{2\omega_n\zeta}{S'(0)} t, \tag{9}$$

choose a new variable $y = z/K_1$, and denote the signal power-to-channel noise intensity by P/N_0. Define

$$\left.\begin{aligned} \beta &= \frac{K_1}{\sqrt{P}K} = \frac{S'(0)}{4\zeta^2} \\ a &= \frac{a_2}{|a_2|_{\max}} = \frac{a_2}{\sqrt{P}KK_1} = \frac{\dot{v}R_cS'(0)}{\omega_n^2 c} \\ 2\rho &= \frac{N_0K}{\sqrt{P}} = \frac{2\omega_n\zeta}{S'(0)}\frac{N_0}{P}. \end{aligned}\right\} \tag{10}$$

Note that for the Brownian motion $w(t)$, we have [2] $(dw(t))^2 \sim dt$, due to the quadratic variation of the

Brownian motion [9]. Hence, $(dw^*(t^*))^2 \sim dt^* = \alpha dt \sim \alpha(dw(t))^2$ so that the "white noise" is scaled by

$$\dot{w}(t) = \frac{d}{dt} w(t) = \frac{1/\sqrt{\alpha}}{1/\alpha} \frac{d}{dt^*} w^*(t^*) = \sqrt{\alpha}\dot{w}^*(t^*). \tag{11}$$

Thus, we finally have from (7), (9), (10), and (11),

$$\left.\begin{aligned} \dot{\epsilon}(t^*) &= -\beta y - S(\epsilon) - \sqrt{2\rho}\dot{w}^*(t^*) \\ \dot{y}(t^*) &= -a + S(\epsilon) + \sqrt{2\rho}\dot{w}^*(t^*). \end{aligned}\right\} \tag{12}$$

III. The Exit Problem

From a mathematical point of view, for small ρ, (12) represents a small stochastic perturbation of the dynamical system

$$\left.\begin{aligned} \dot{\epsilon}(t^*) &= -\beta y - S(\epsilon) \\ \dot{y}(t^*) &= -a + S(\epsilon) \end{aligned}\right\} \tag{13}$$

which has a stable equilibrium at the point

$$\left.\begin{aligned} \epsilon_a &= \frac{a}{S'(0)} \\ y_a &= -\frac{a}{\beta} \end{aligned}\right\} \tag{14}$$

and nonstable (saddle) point(s) at

$$\left.\begin{aligned} \epsilon_b &= -1.5 + |a|, \qquad a < 0 \\ y_b &= -\frac{a}{\beta} \end{aligned}\right\} \tag{15}$$

where for $a = 0$, $\epsilon_b = \pm 1.5$.

The stable equilibrium is an attractor to all trajectories starting in some neighborhood to it. This neighborhood is called the domain of attraction D and its boundary is denoted by ∂D. From the engineering point of view, when the state vector of (12) is inside D, the loop is locked.

For $a = 0$, the attractor lies at the origin, and the domain D is symmetric about it; see Fig. 4. This is well known, although for a different choice of the state vector, namely, $(\epsilon, \dot{\epsilon})$; see, e.g., [8], [21].

We prefer our choice of the state vector since there is no need to differentiate $S(\epsilon)$; thus, the deterministic trajectories (and the boundaries ∂D) are smooth. For different damping factors ζ, the shape of ∂D changes—smaller ζ results in smaller D. For $a \neq 0$ (Doppler rate), the attractor is moved away from the origin according to (14), and furthermore, the domain D is not symmetric anymore (Figs. 5–7). Here again, the domain monotonically decreases for smaller ζ. Intuitively, we feel, as we will see later, that a smaller D is related to a higher probability to lose lock. These shapes of D indicate that small damping should be avoided. Similarly, the domain decreases monotonically for larger $|a|$ (Fig. 6) and disappears for $|a| = 1$.

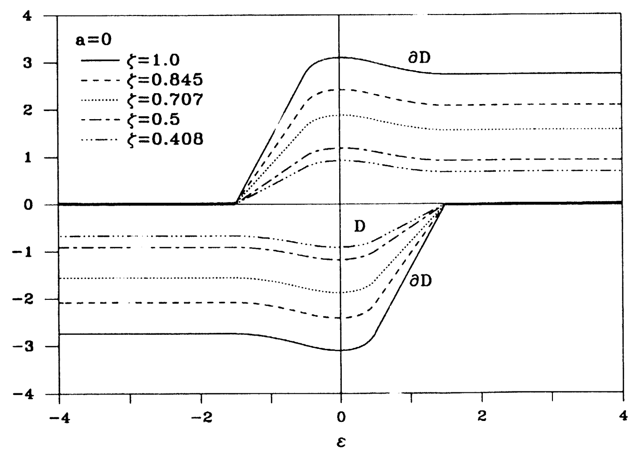


Fig. 4. The domain of attraction D and its boundaries ∂D for $a = 0$.

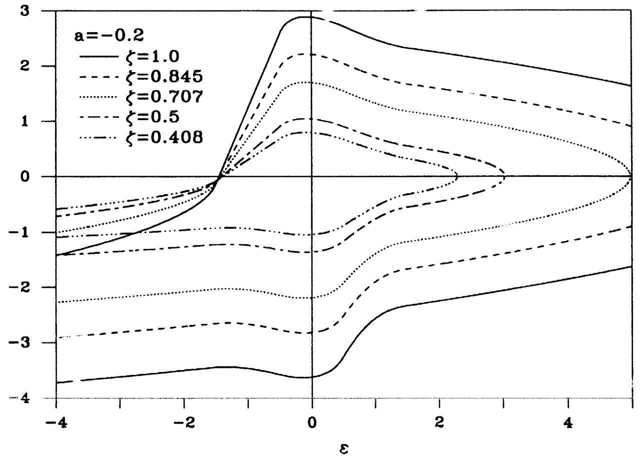


Fig. 5. The domain of attraction D for various damping factors ζ. The dimensionless loop offset is $a = -0.2$.

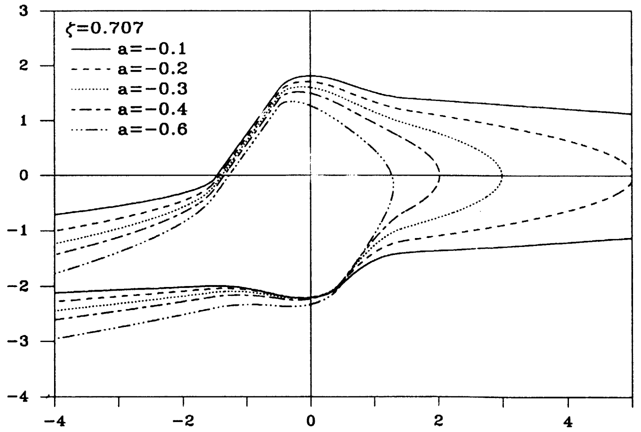


Fig. 6. The domain of attraction D for various dimensionless loop offsets a for a damping factor $\zeta = 0.707$.

For the deterministic system (13), all trajectories that begin inside D will converge to the attractor, and will never cross the boundary ∂D. In other words, a locked loop will remain locked. However, even the slightest stochastic perturbation is sure to cause a crossing in some finite time [18], resulting in loss of lock. Let τ be the first

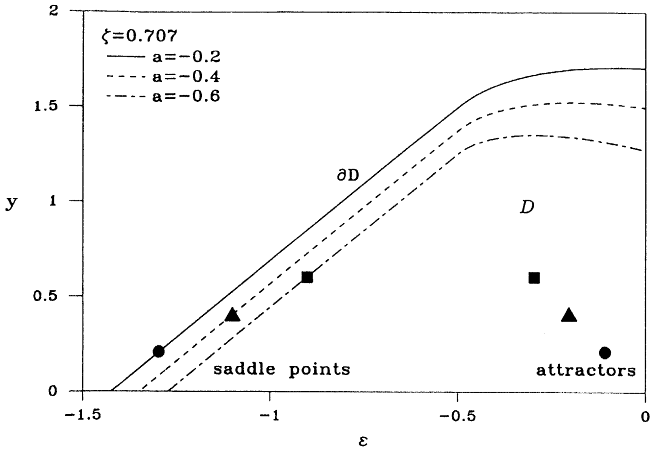


Fig. 7. The attractors (stable equilibria) and the saddle points (nonstable equilibria) for various dimensionless loop offsets a. The damping factor is $\zeta = 0.707$.

time that a locked trajectory reaches the boundary:

$$\tau = \inf \left\{ t \middle| [\epsilon(t), y(t)] \in \partial D \right\}. \tag{16}$$

We will analyze the system (12) and consider the scaled time to reach the boundary τ^*. The desired unscaled time is then given by $\tau = \tau^*/\alpha$.

IV. Computation of the Mean Time to Lose Lock

In general, the random time τ^* to reach the boundary of (12) depends on the initial conditions (ϵ, y). The mean time to reach the boundary conditioned on the initial state is

$$\bar{\tau}^*(\epsilon, y) = E[\tau^* | \epsilon, y]. \tag{17}$$

The mean *exit* time $\bar{t}_L^*$ is $\bar{t}_L^* = 2\bar{\tau}^*$ since once a trajectory hits the boundary, it crosses it or returns with equal probabilities [2], [23]. (It was pointed out to us [17] that technically speaking, the lock detector might indicate loss of lock when the trajectory is close to ∂D, and in that case, $\bar{t}_L^* \sim \bar{\tau}^*$.)

It is well known [18] that $\bar{\tau}^*(\epsilon, y)$ is the solution of the Kolmogorov–Dynkin equation

$$\left.\begin{aligned} \mathcal{L}\bar{\tau}^*(\epsilon, y) &= -1 \quad &\text{in } D \\ \bar{\tau}^*(\epsilon, y) &= 0 \quad &\text{on } \partial D \end{aligned}\right\} \tag{18}$$

where $\mathcal{L}$ is the backward Kolmogorov operator of (12),

$$\begin{aligned} \mathcal{L}(\cdot) \equiv & \left\{-\beta y - S(\epsilon)\right\} \frac{\partial}{\partial \epsilon}(\cdot) \\ & + \left\{-a + S(\epsilon)\right\} \frac{\partial}{\partial y}(\cdot) \\ & + \rho \left\{ \frac{\partial^2}{\partial \epsilon^2}(\cdot) - 2\frac{\partial^2}{\partial \epsilon \partial y}(\cdot) + \frac{\partial^2}{\partial y^2}(\cdot) \right\}. \end{aligned} \tag{19}$$

While for first-order systems the Kolmogorov–Dynkin equation can be solved analytically, for second- or higher order systems, this cannot be done in general. Thus, we propose to use the fact that ρ is a small parameter and construct an asymptotic solution to (18). We will give a brief description of the method. More details can be found in [1], [2], and [18].

In order to get some insight into the structure of our solution, we begin by discussing first a simpler case, namely, a dynamical system having a potential Φ, excited by a "small" white Gaussian noise. The mean time $\bar{\tau}$ to exit from the domain D was given by Arrhenius; see, e.g., [18]. Let the potential on the boundary be $\Phi(\partial D)$, and let $\Phi_a = \Phi(\text{attractor})$. Thus, the potential barrier between the attractor and the boundary is $\Delta\Phi = \Phi(\partial D) - \Phi_a$. Define $\Phi_{\min}$ to be the minimal value of the potential Φ on the boundary ∂D: $\Phi_{\min} = \min \Phi(\partial D)$. The minimal (lowest) potential barrier is thus $\Delta\Phi_{\min} = \Phi_{\min} - \Phi_a$. Let the noise intensity 2ρ be small compared to $\Delta\Phi_{\min}$. The mean time to reach ∂D was shown to be

$$\bar{\tau} \propto \exp(\Delta\Phi_{\min}/\rho). \tag{20}$$

Furthermore, if $\Phi(\partial D)$ is minimal at only one point, the trajectories will hit (and exit) the boundary mainly in the neighborhood of this point. This agrees with our intuition: it is most likely to exit from the lowest potential barrier; and since exits occur mainly there, only the lowest potential barrier should influence the mean exit time.

Potential stochastic dynamical systems are not common in tracking systems (except, of course, of all first-order loops). For nonpotential systems, a quasi-potential function Ψ will be defined, and will play a role similar to Φ in the expression for $\bar{\tau}$.

For small ρ, the system fluctuates most of the time around the attractor, and exits are rare events. Thus, $\bar{\tau}^*(\epsilon, y)$ is almost constant over most of D (since a trajectory would rather return to the attractor than exit) and has a boundary layer near ∂D. We thus assume that $\bar{\tau}^*$ is of the form [13], [19]

$$\bar{\tau}^*(\epsilon, y) = C(\rho)\, U(\epsilon, y) \exp\left(\frac{\hat{\Psi}}{\rho}\right) \tag{21}$$

where $\hat{\Psi}$ is shown (see the Appendix) to be the minimal value of Ψ on the boundary ∂D, thus playing a similar role to $\Delta\Phi_{\min}$ in (20), (we choose $\Psi(\text{attractor}) = 0$). The function $U(\epsilon, y)$ is almost constant around the attractor, $U(\epsilon_a, y_a) = 1$, and vanishes rapidly towards ∂D. Hence, for a locked loop and for small ρ, it is enough to know $\bar{\tau}^*(\epsilon_a, y_a)$. Furthermore, since the dominant contribution to $\bar{\tau}^*$ comes from the exponential term, our main concern is to find $\hat{\Psi}$ where

$$\hat{\Psi} = \Psi_{\min} \triangleq \min \Psi(\partial D). \tag{22}$$

Let

$$\left.\begin{aligned} \mathcal{L}^*W(\epsilon, y) &= 0 \quad \text{in } D \\ W(\epsilon_a, y_a) &= 1 \end{aligned}\right\} \tag{23}$$

where $\mathcal{L}^*$ is the forward Kolmogorov (Fokker–Planck) operator of (12), namely,

$$\begin{aligned} \mathcal{L}^*(\cdot) \equiv{} & \frac{\partial}{\partial \epsilon}\left\{[\beta y + S(\epsilon)](\cdot)\right\} \\ & + \frac{\partial}{\partial y}\left\{[a - S(\epsilon)](\cdot)\right\} \\ & + \rho\left\{\frac{\partial^2}{\partial \epsilon^2}(\cdot) - 2\frac{\partial^2}{\partial \epsilon \partial y}(\cdot) + \frac{\partial^2}{\partial y^2}(\cdot)\right\}. \end{aligned} \tag{24}$$

Applying the "ray method" [18], we assume $W(\epsilon, y)$ to have the form

$$W(\epsilon, y) = g(\rho, \epsilon, y) \exp\left(-\frac{\Psi(\epsilon, y)}{\rho}\right). \tag{25}$$

This assumption is motivated by the fact that if the dynamical system (12) had a potential $\Phi(\epsilon, y)$, $W(\epsilon, y)$ would have been given by $W(\epsilon, y) \propto \exp(-\Phi(\epsilon, y)/\rho)$. Substitute (25) in (23); we have

$$\begin{aligned} \frac{1}{\rho}\Bigg\{ & -[\beta y + S(\epsilon)]\frac{\partial \Psi}{\partial \epsilon} + [-a + S(\epsilon)]\frac{\partial \Psi}{\partial y} \\ & + \left[\frac{\partial \Psi}{\partial \epsilon} - \frac{\partial \Psi}{\partial y}\right]^2\Bigg\} W(\epsilon, y) + O(1) = 0 \end{aligned} \tag{26}$$

where $O(1)$ represents terms of "order 1" and smaller. For (26) to hold for small ρ, i.e., $\rho < 0.1$, the terms of leading order in ρ (namely, those multiplied by $1/\rho$) must be zero, resulting in the "eikonal" equation

$$\begin{aligned} & -[\beta y + S(\epsilon)]\frac{\partial \Psi}{\partial \epsilon} + [-a + S(\epsilon)]\frac{\partial \Psi}{\partial y} \\ & + \left[\frac{\partial \Psi}{\partial \epsilon} - \frac{\partial \Psi}{\partial y}\right]^2 = 0. \end{aligned} \tag{27}$$

Next we calculate $\hat{\Psi}$ analytically. For $|\epsilon| \le 1/2$, it is easy to verify by direct substitution that $\Psi(\epsilon, y) = \varphi(\epsilon, y)$ where

$$\begin{aligned} \varphi(\epsilon, y) = \frac{\alpha_1^2}{2\beta}\Bigg[& \left(\epsilon - \frac{a}{\alpha_1}\right)^2 + 2\left(\epsilon - \frac{a}{\alpha_1}\right)\left(y + \frac{a}{\beta}\right) \\ & + \left(1 + \frac{\beta}{\alpha_1}\right)\left(y + \frac{a}{\beta}\right)^2\Bigg] \end{aligned} \tag{28}$$

and $\alpha_1 = S'(0) = 2$; see Fig. 3. For $|\epsilon| \ge 1/2$, the eikonal equation (27) can be solved by the characteristic equations [4]

$$\left.\begin{aligned} \dot{\epsilon} &= -\beta y - S(\epsilon) + 2\{p - q\} \\ \dot{y} &= -a + S(\epsilon) - 2\{p - q\} \\ \dot{p} &= S'(\epsilon)\{p - q\} \\ \dot{q} &= \beta p \end{aligned}\right\} \tag{29}$$

$$\dot{\Psi} = \{p - q\}^2 \tag{30}$$

where $p = \partial \Psi/\partial \epsilon$ and $q = \partial \Psi/\partial y$. Since $\Psi(\epsilon, y)$ is known for $|\epsilon| \le 1/2$, the characteristics can be initialized at $\epsilon = \pm 1/2$. The initial conditions for (29), (30) are

$$\left.\begin{aligned} \epsilon_0 &= -0.5 \qquad \text{for } a < 0 \\ y &= y_0 \\ p_0 &= \left.\frac{\partial \varphi}{\partial \epsilon}\right|_{-0.5, y_0} \\ q_0 &= \left.\frac{\partial \varphi}{\partial y}\right|_{-0.5, y_0} \\ \Psi(-0.5, y_0) &= \varphi(-0.5, y_0). \end{aligned}\right\} \tag{31}$$

Note that for $|\epsilon| > 1/2$, $S'(\epsilon) = -\alpha_2$ in (29); see Fig. 3. For every y_0, we get a different characteristic, which may cross the boundary ∂D in the neighborhood of the saddle point; see Fig. 8. The value of Ψ at the crossing point versus the arc length s of ∂D ($s = 0$ at the saddle point) is plotted in Fig. 9. The minimal value of $\Psi(s)$, $\hat{\Psi}$ is the desired constant.

This method of "shooting" characteristics to get $\hat{\Psi}$ is very useful for a general $S(\epsilon)$ [2]. For the specific shape of our S curve, an analytical solution can be derived. There is one trajectory, let us call it "critical," that hits precisely the saddle point (Fig. 8). The initial condition of this trajectory is $\hat{y}_0$, and can be found analytically since the system (29) is linear. Once this $\hat{y}_0$ is known, (30) can be integrated to yield $\hat{\Psi} = \Psi(\epsilon_b, y_b)$. After some manipulations [1], [26], we have for the case $\alpha_2 = 1$

$$\left.\begin{aligned} \lambda_1 &= -\frac{1}{2}[1 + \sqrt{1 + 4\beta}] \\ \epsilon_0 &= -\frac{1}{\alpha_1} \\ \hat{y}_0 &= \frac{-\alpha_1\lambda_1\epsilon_0 + \dfrac{a}{\lambda_1\beta}(\beta^2 - \alpha_1\lambda_1^2)}{\alpha_1\lambda_1 - \beta} \\ p_0 &= \frac{\alpha_1}{\lambda_1}\left(\hat{y}_0 + \frac{a}{\beta}\right) \\ \hat{\Psi} &= \varphi(\epsilon_0, \hat{y}_0) - \frac{\lambda_1}{2}p_0^2 \end{aligned}\right\} \tag{32}$$

where φ is given by (28).

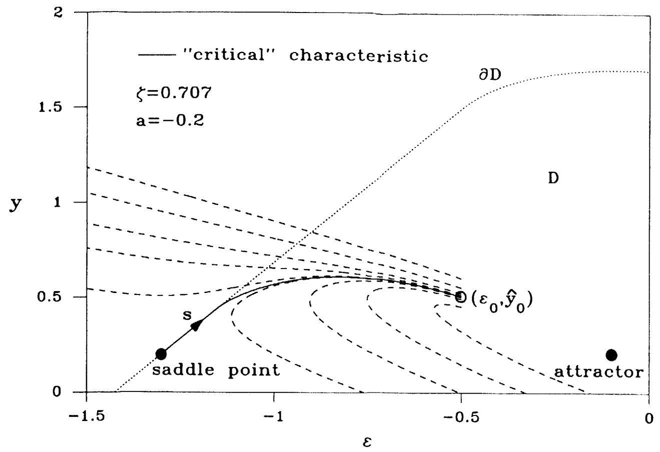


Fig. 8. Characteristics to determine the quasi-potential on the boundary $\Psi(s)$ and the "critical" characteristic for $\hat{\Psi}$.

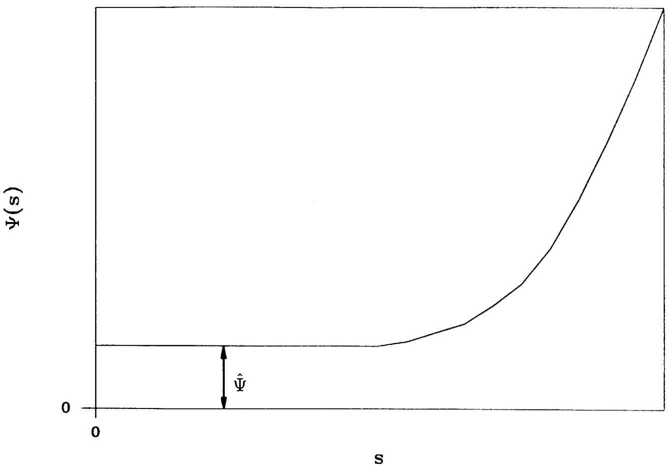


Fig. 9. The quasi-potential $\Psi(s)$ as a function of the arc length s on the boundary ∂D (the saddle is at $s = 0$).

The scaled MTLL $\bar{t}_L^*$ is

$$\bar{t}_L^* = 2C \exp\left(\frac{\hat{\Psi}}{\rho}\right)$$

and the MTLL is

$$\bar{t}_L = \frac{2C}{\alpha} \exp\left(\frac{\hat{\Psi}}{\rho}\right). \tag{33}$$

At this point, we present an alternative approach to describe the exit phenomenon which is not as accurate, but is much easier to follow. The global loop behavior is approximated by: 1) a linearized loop which contributes to the random fluctuations; and 2) the boundary ∂D that is related to the nonlinear deterministic system (13). Thus, we define "loss of lock" to occur when a trajectory of the linearized loop hits the boundary ∂D. It is easy to calculate the corresponding MTLL. In [18], it is shown that

$$\bar{t}_L \propto \frac{2}{\alpha} \exp\left(\frac{\hat{\varphi}}{\rho}\right) \tag{34}$$

where $\hat{\varphi}$ is the value of $\varphi(\epsilon, y)$ (28) which just hits the

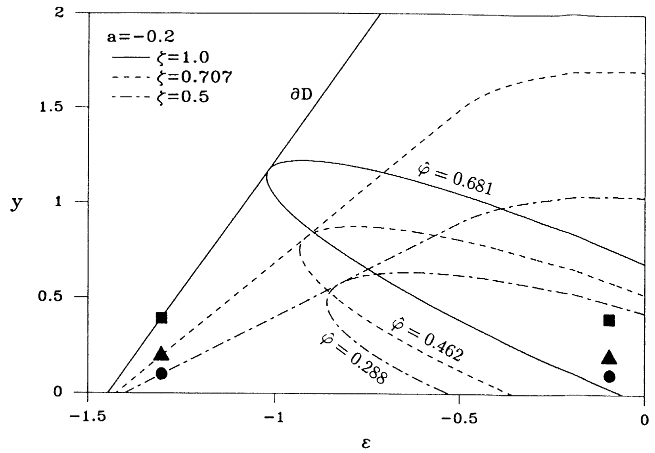


Fig. 10. Contour lines of constant φ hitting the corresponding boundaries ∂D for various damping factors ζ. The dimensionless loop offset is $a = -0.2$.

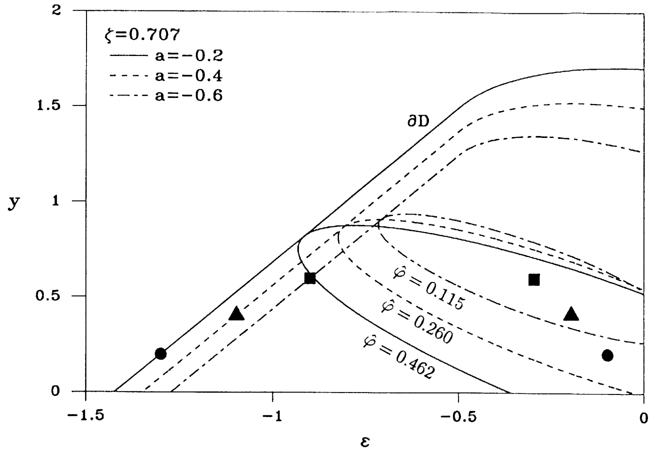


Fig. 11. Contour lines of constant φ hitting the corresponding boundaries ∂D for various dimensionless loop offsets a. The damping factor is $\zeta = 0.707$.

boundary ∂D; Figs. 10 and 11. It is given by

$$\left.\begin{aligned}
\hat{\varphi} &= \varphi(\epsilon_1, y_1)\\
\lambda_2 &= -\frac{1}{2}\left[1 - \sqrt{1 + 4\beta}\right]\\
y_1 &= \frac{\alpha_2}{\lambda_2}(\epsilon_1 - \epsilon_b) + y_b\\
\epsilon_1 &= \frac{A\dfrac{a}{\alpha_1} + \dfrac{\alpha_2}{\lambda_2}\left(1.5 + \dfrac{a}{\alpha_2}\right)}{A - \dfrac{\alpha_2}{\lambda_2}}\\
A &= -\frac{1 + \dfrac{\alpha_2}{\lambda_2}}{1 + \dfrac{\alpha_2}{\lambda_2}\left(1 + \dfrac{\beta}{\alpha_1}\right)}.
\end{aligned}\right\} \tag{35}$$

Consider now the exit points. From Figs. 10 and 11, it is clear that φ has a minimum on ∂D. Hence, exit points will concentrate in the neighborhood of ∂D where φ is minimal [18]. Note that these neighborhoods are a little away to the right of their corresponding saddle points, Figs. 10 and 11.

As we can see (Fig. 9), Ψ on ∂D is flat near the saddle point. It was shown in [2], [3], using a detailed analysis, that because of this flatness, the exit neighborhood is shifted away from the saddle. It is interesting that our simplified approximation predicts a very similar shift. (Maybe the observation that the exit points concentrate in a small neighborhood could lead to improvements in tracking loops.)

V. Optimal Second-Order Loop

We propose an optimal choice of the loop natural radian frequency ω_n and its damping factor ζ that maximizes the MTLL, and thus extends the threshold to the maximum for various P/N_0 and Doppler rates due to $\dot{v}$. This choice is thus recommended near the threshold; for high P/N_0 and small $\dot{v}$ where $\bar{t}_L$ is high enough anyway, other choices, like minimization of the "worst case" tracking error [22], [26], for example, should be considered.

In order to give simple design rules, we make use of the observation made in [24] (see also Section VI) that the main contribution to $\bar{t}_L$ comes from the exponential term, which is obviously dominant for small ρ. Note that ρ is small even near the threshold since under all operating conditions, $\bar{t}_L$ must be sufficiently large. Using (9), (10), and (33), we have

$$\bar{t}_L \propto \frac{1}{\zeta\omega_n} \exp\left(\frac{S'(0)\hat{\Psi}}{\zeta\omega_n}\frac{P}{N_0}\right) \tag{36}$$

where $\hat{\Psi} = \hat{\Psi}(\dot{v}, \omega_n, \zeta)$ according to (10) and (32). Maximization of

$$Q = \frac{S'(0)\hat{\Psi}}{\zeta\omega_n}\frac{P}{N_0} \tag{37}$$

with respect to ω_n, ζ leads to the following [25]:

$$\zeta^{\text{opt}} = 0.599 \approx 0.6 \tag{38}$$

$$|a^{\text{opt}}| = 0.2, \qquad \dot{v} \neq 0 \tag{39}$$

$$\hat{\Psi}^{\text{opt}} = 0.194, \qquad \dot{v} \neq 0. \tag{40}$$

Using (10), we have from (39)

$$\omega_n^{\text{opt}} = \sqrt{\frac{S'(0)R_c}{0.2c}|\dot{v}|}, \qquad \dot{v} \neq 0. \tag{41}$$

Equations (38) and (41) are our proposed design rules, and the (almost) optimal MTLL is given by

$$\bar{t}_L^{\text{opt}} \propto \frac{1}{0.6\omega_n^{\text{opt}}} \exp\left(\frac{0.647}{\omega_n^{\text{opt}}}\frac{P}{N_0}\right). \tag{42}$$

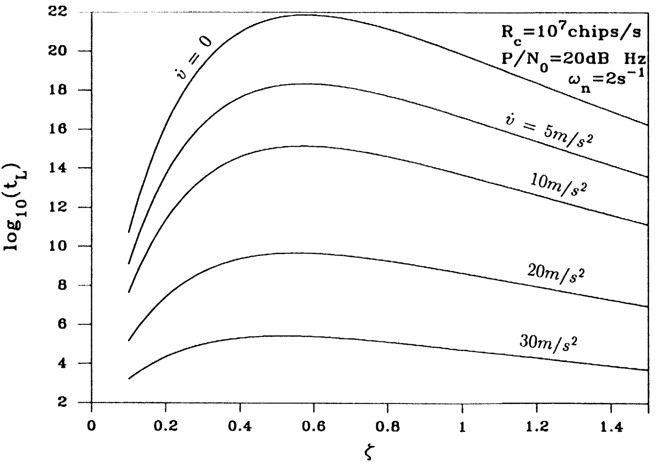


Fig. 12. The mean time to lose lock (MTLL) $\bar{t}_L$ (in seconds) versus the damping factor ζ for various Doppler rates $\dot{v}$.

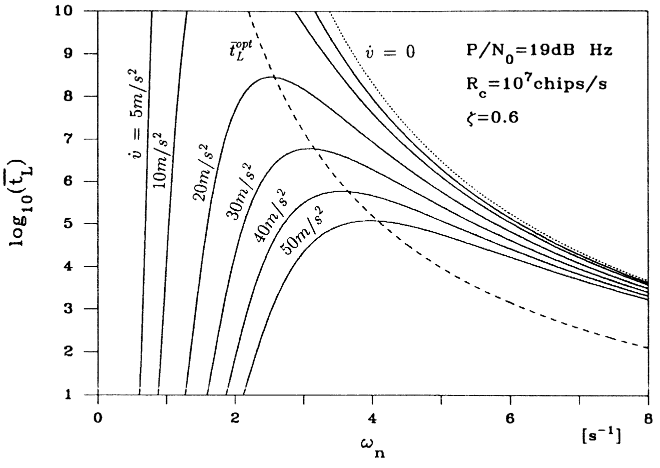


Fig. 13. The mean time to lose lock (MTLL) $\bar{t}_L$ and the optimal MTLL $\bar{t}_L^{\text{opt}}$ (in seconds) versus the natural frequency ω_n for various Doppler rates $\dot{v}$, $P/N_0 = 19$ dB · Hz.

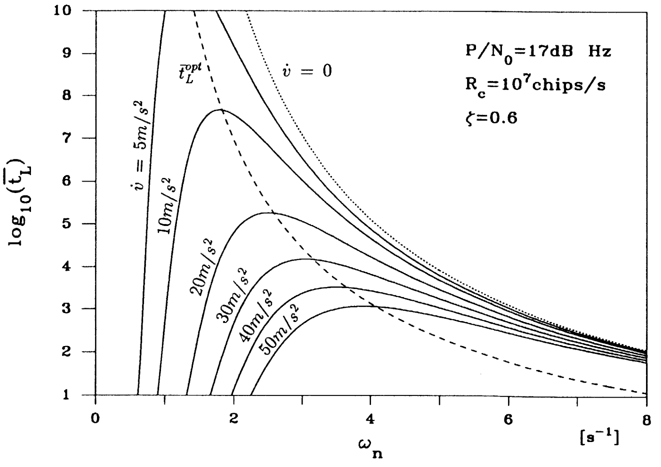


Fig. 14. The mean time to lose lock (MTLL) $\bar{t}_L$ and the optimal MTLL $\bar{t}_L^{\text{opt}}$ (in seconds) versus the natural frequency ω_n for various Doppler rates $\dot{v}$, $P/N_0 = 17$ dB · Hz.

In Figs. 12–16, $\bar{t}_L$ and $\bar{t}_L^{\text{opt}}$ are plotted for various ω_n, ζ, $\dot{v}$, and P/N_0. It is seen that (38), (41), and (42) are really very close to the optimum. Note that for $\dot{v} = 0$, (39) and (41) do not hold. In this case, ω_n can be chosen to be arbitrary small, so $\bar{t}_L$ can be made arbitrary large; Figs. 13 and 14. Obviously, for the case $\dot{v} \equiv 0$, a first-order loop is sufficient; however, such a loop has a nonzero steady-state error. Note that for the second-order loop, the performance is upper bounded for the case $\dot{v} = 0$.

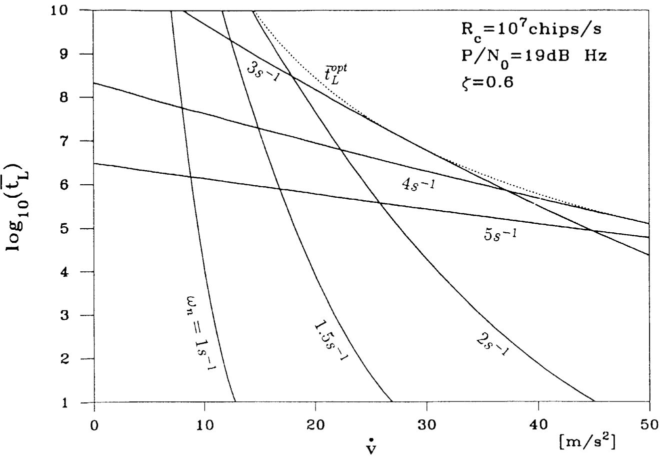


Fig. 15. The mean time to lose lock (MTLL) $\bar{t}_L$ and the optimal MTLL $\bar{t}_L^{\rm opt}$ (in seconds) versus the Doppler rate $\dot{v}$ for various natural frequencies ω_n, $P/N_0 = 19$ dB · Hz.

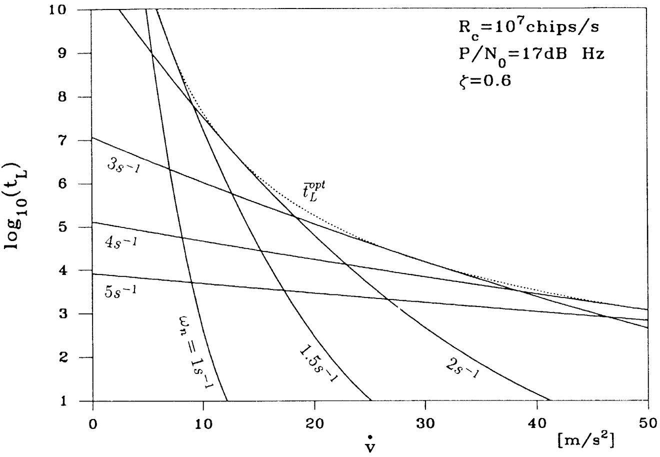


Fig. 16. The mean time to lose lock (MTLL) $\bar{t}_L$ and the optimal MTLL $\bar{t}_L^{\rm opt}$ (in seconds) versus the Doppler rate $\dot{v}$ for various natural frequencies ω_n, $P/N_0 = 17$ dB · Hz.

As was pointed out, the parameter ρ must be small enough for (26) to hold. Indeed, in Figs. 12–14, ρ is smaller than 0.05, which is sufficiently small for our derivations.

Equation (42) can be written in another useful form. For small ρ and conditioned that the loop is locked, the jitter, i.e., the error standard deviation σ_ϵ, can be given approximately by the Gaussian approximation $\sigma_\epsilon^2 \approx 0.5 B_L N_0/P$ where $B_L = 0.5\omega_n\{\zeta + 1/(4\zeta)\}$ is the equivalent loop noise bandwidth. Using this, (10), (38), and (41) in (42) yields

$$\bar{t}_L^{\rm opt} \propto \frac{0.85}{B_L} \exp\left(\frac{0.164}{\sigma_\epsilon^2}\right). \tag{43}$$

Turning now to the simplified approach (34), (35), it is interesting to note that here we also get $|a^{\rm opt}| = 0.2$; hence, $\omega_n^{\rm opt}$ is like in (41), and the optimal damping factor $\zeta^{\rm opt}$ is ~ 1.0, which is not too far from (38). However, $\hat{\varphi}/\hat{\Psi} \sim 2 \div 3$; thus, $\hat{\varphi}$ cannot be used to calculate the actual MTLL, but can be used to indicate a reasonable choice of the loop parameters. Since extensions of this method for a general $S(\epsilon)$ and for higher order loops are straightforward, and since in such cases no solutions to the MTLL are known, this simplified method might still be found useful.

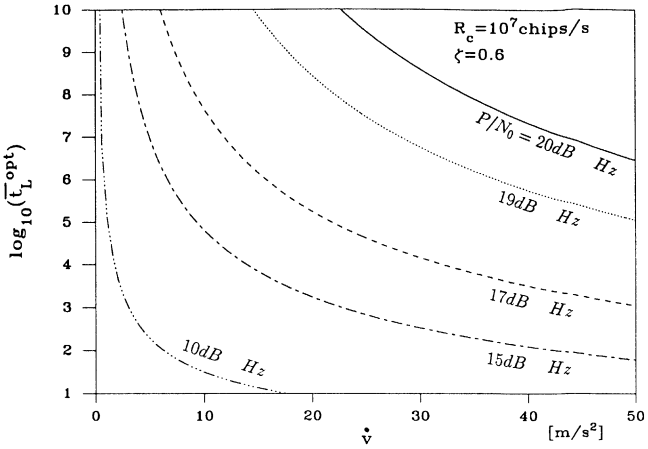


Fig. 17. Upper limit to loop performance under Doppler rate $\dot{v}$: maximal mean time to lose lock (MTLL) versus $\dot{v}$ for various P/N_0.

VI. Discussion

To get some insight into our proposed design, we consider the detailed GPS case study in [22, sect. 3.3], [26]. In Figs. 13–17, we plot the MTLL for parameters similar to [22]. It is interesting that our choice of ω_n is smaller, but close to the choice made in [22]. The optimal MTLL, and thus the upper limit to loop performance under Doppler rate, is given in Fig. 17 which, we believe, is an important result of our derivation.

If we take the minimum allowable MTLL as a definition for the threshold, we can get from Fig. 17 the required P/N_0 for any given $\dot{v}$. For example, let us take the minimum allowable MTLL to be $\sim 10^3$ s. We see that for $\dot{v} = 30$ m/s^2, the threshold is at $P/N_0 \sim 16$ dB · Hz; since the actual P/N_0 is much higher ($\sim$38 dB · Hz) [22], such acceleration can easily be tracked. Using the above MTLL $= 10^3$ s in (43), we have σ_ϵ(threshold) $=$ 0.15, which is again rather close, although somewhat smaller than the estimated value given in [22].

The method presented here to calculate the MTLL was further derived to give the preexponential term [C in (33)] and the probability density of the exit points on the boundary ∂D, following [1] and [2]. It can be shown that taking the preexponential C into account results in optimal parameters which are very close to those given in Section V. Since C can be given only numerically, in contrast to the analytical expressions for $\hat{\Psi}$, (28), (32), the approach suggested here seems both elegant and practical. Monte Carlo simulations were carefully performed to further confirm our computations of the MTLL, the exit probability density, and the loop parameters optimization. The simulation turned out to agree very well with all our theoretical calculations. These results will be reported elsewhere.

We hope that the approach presented here will help in gaining more insight into the exit phenomenon in tracking loops.

Appendix

We give an outline to the proof of (22), namely, that $\hat{\Psi}$ in (21) should be the minimal value of Ψ on the boundary ∂D [2], [18].

From Green's identity, we have

$$\int_D \int f \mathcal{L} h \, d\epsilon \, dy = \int_D \int h \mathcal{L}^* f \, d\epsilon \, dy + \rho \oint_{\partial D} f \left[\frac{\partial h}{\partial \epsilon} - \frac{\partial h}{\partial y} \right] (\nu_1 - \nu_2) \, ds \tag{A1}$$

where f, h are any "well-behaved" functions, $\mathcal{L}$ and $\mathcal{L}^*$ are the adjoint operators defined in (19), (24), $\mathbf{v} = (\nu_1, \nu_2)^T$ is the local unit outer normal vector to ∂D, and ds is the infinitesimal arc length on the boundary ∂D. Next, choose

$$h = \overline{\tau}^*(\epsilon, y)$$
$$f = W(\epsilon, y).$$

From (18), $\mathcal{L}\overline{\tau}^* = -1$ in D; hence, the left-hand side of (A1) is just

$$-\int_D \int W(\epsilon, y) \, d\epsilon \, dy,$$

and from (23), the first integral on the right-hand side of (A1) is zero since

$$\mathcal{L}^* W = 0.$$

Hence,

$$-\int_D \int W(\epsilon, y) \, d\epsilon \, dy = \rho \oint_{\partial D} W(\epsilon, y) \left[\frac{\partial \overline{\tau}^*(\epsilon, y)}{\partial \epsilon} - \frac{\partial \overline{\tau}^*(\epsilon, y)}{\partial y} \right] \cdot (\nu_1 - \nu_2) \, ds. \tag{A2}$$

Now, substitute the assumed forms of $\overline{\tau}^*$, (21), and W, (25), into (A2). We have

$$-\int_D \int g(\rho, \epsilon, y) \exp\left(-\frac{\Psi(\epsilon, y)}{\rho} \right) d\epsilon \, dy = C(\rho) \exp\left(\frac{\hat{\Psi}}{\rho} \right) \oint_{\partial D} g(\rho, s) \exp\left(-\frac{\Psi(s)}{\rho} \right) \cdot \left[\left(\frac{\partial U}{\partial \epsilon} - \frac{\partial U}{\partial y} \right) (\nu_1 - \nu_2) \right] ds \tag{A3}$$

where the expression in the bracket comes from the boundary layer of $U(\epsilon, y)$ near ∂D [2] and $g(\rho, s)$, $\Psi(s)$ are $g(\rho, \epsilon, y)$, $\Psi(\epsilon, y)$ on the boundary ∂D. Since ρ is a small parameter, we can use the Laplace integration formula [15]. Hence,

$$\int_D \int g(\rho, \epsilon, y) \exp\left(-\frac{\Psi(\epsilon, y)}{\rho} \right) d\epsilon \, dy \simeq g(\rho, \epsilon_a, y_a) \frac{2\pi\rho}{\sqrt{J}} \tag{A4}$$

where

$$J = \left| \det \begin{bmatrix} \Psi_{yy} & \Psi_{\epsilon y} \\ \Psi_{\epsilon y} & \Psi_{\epsilon\epsilon} \end{bmatrix} \right| \Bigg|_{\epsilon = \epsilon_a, y = y_a}.$$

Since $\Psi(\epsilon, y) = \varphi(\epsilon, y)$ for $|\epsilon| \leq 1/2$, using (28), we get $J = \alpha_1^3/\beta$. We turn now to the line integral. It can be shown that a portion of the boundary ∂D in the neighborhood of the saddle is a straight line. We denote this portion by Γ (Fig. 7). Since $\Psi(s)$ is growing rather rapidly on Γ, Fig. 9 (see also [1]), the main contribution to the line integral comes from this linear portion. Obviously, $\mathbf{v}$ is not changing with s on Γ. It turns out [1] that the boundary layer U is also independent of s near Γ. Hence,

$$I = \oint_{\partial D} g(\rho, s) \exp\left(-\frac{\Psi(s)}{\rho} \right) \cdot \left[\left(\frac{\partial U}{\partial \epsilon} - \frac{\partial U}{\partial y} \right) (\nu_1 - \nu_2) \right] ds \simeq C_1(\rho) \int_\Gamma g(\rho, s) \exp\left(-\frac{\Psi(s)}{\rho} \right) ds.$$

We again use the Laplace method to evaluate this integral. Let

$$\Psi_{\min} \triangleq \min_{s \in \Gamma} \{ \Psi(s) \};$$

from [15], we have

$$I = \oint_{\partial D} \cdots ds = C_2(\rho) \exp\left(-\frac{\Psi_{\min}}{\rho} \right) \tag{A5}$$

where $C_2(\rho)$ is a polynomial in ρ. Substitute (A4) and (A5) into (A3); we see that the exponentially large number $\exp(\hat{\Psi}/\rho)$ multiplies the exponentially small number $\exp(-\Psi_{\min}/\rho)$ (recall that ρ is small) where the product is only polynomial in ρ. Hence, for equality to hold for the leading order term in ρ, we must have

$$\exp\left(\frac{\hat{\Psi}}{\rho} \right) \cdot \exp\left(-\frac{\Psi_{\min}}{\rho} \right) \simeq 1;$$

hence,

$$\Psi_{\min} = \hat{\Psi}. \qquad \text{Q.E.D.}$$

References

[1] B. Z. Bobrovsky and Z. Schuss, "Jamming and maneuvering induced loss of lock in a range tracking loop," Swiss Fed. Inst. Technol., Zürich, IKT Tech. Rep., Mar. 1988.

[2] ——, "Singular perturbation in filtering theory," *SIAM J. Appl. Math.*, vol. 42, pp. 174–187, 1982.

[3] B. Z. Bobrovsky and O. Zeitouni, "Some results on the problem of exit from a domain," submitted to *Stochastics*.

[4] C. R. Chester, *Techniques in Partial Differential Equations.* New York: McGraw-Hill, 1971.

[5] M. V. Day, "Recent progress on the small parameter exit problem," *Stochastics*, vol. 20, pp. 121-150, 1987.

[6] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems.* New York: Springer, 1984.

[7] C. W. Gardiner, *Handbook of Stochastic Methods.* New York: Springer Ser. Synergetics, 1983.

[8] J. K. Holmes, *Coherent Spread Spectrum Systems.* New York: Wiley, 1982.

[9] K. Itô and H. P. McKean, *Diffusion Process and Their Sample Paths.* New York: Academic, 1964.

[10] R. Karmy, B. Z. Bobrovsky, and Z. Schuss, "Loss of lock induced by Doppler or code rate mismatch in code tracking loops," in *Proc. MILCOM '87*, Washington, DC, Oct. 1987, pp. 294-298.

[11] H. A. Kramers, "Brownian motion in field of force and the diffusion model of chemical reaction," *Physica*, vol. 7, pp. 284-304, 1940.

[12] W. C. Lindsey, *Synchronous Systems in Communications and Control.* Englewood Cliffs, NJ: Prentice-Hall, 1972.

[13] B. J. Matkowsky and Z. Schuss, "The exit problem for randomly perturbed dynamical systems," *SIAM J. Appl. Math.*, vol. 33, no. 2, pp. 365-382, 1977.

[14] H. Meyr, "Nonlinear analysis of correlative tracking systems using renewal process theory," *IEEE Trans. Commun.*, vol. COM-23, pp. 192-203, Feb. 1975.

[15] F. W. J. Olver, *Asymptotics and Special Functions.* New York: Academic, 1974.

[16] A. Polydoros and C. L. Weber, "Analysis and optimization of correlative code tracking loops in spread spectrum systems," *IEEE Trans. Commun.*, vol. COM-33, pp. 30-34, Jan. 1985.

[17] A. Polydoros, private communication.

[18] Z. Schuss, *Theory and Applications of Stochastic Differential Equations.* New York: Wiley, 1980.

[19] Z. Schuss and B. J. Matkowsky, "The exit problem: A new approach to diffusion across potential barriers," *SIAM J. Appl. Math.*, no. 35, pp. 604-623, 1979.

[20] M. K. Simon, J. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications*, Vols. I-III. Rockville, MD: Computer Science Press, 1985.

[21] J. J. Spilker, *Digital Communications by Satellite.* Englewood Cliffs, NJ: Prentice-Hall, 1977.

[22] —, "GPS signal structure and performance characteristics," *Navigation: J. Inst. Navigation*, vol. 25, no. 2, pp. 121-146, 1978.

[23] A. J. Viterbi, *Principles of Coherent Communications.* New York: McGraw-Hill, 1966.

[24] A. L. Welti and B. Z. Bobrovsky, "Optimal design of a modified code tracking loop," in *Proc. 8th European Conf. Electrotechnics, EUROCON '88*, Stockholm, Sweden, June 1988, pp. 80-83.

[25] —, "Optimization of a second-order PN-code tracking loop using the mean exit time criterion," in *Proc. IEEE Int. Conf. Syst. Eng.*, Dayton, OH, Aug. 1989, pp. 487-490.

[26] —, "Doppler acceleration influence on code tracking in direct sequence spread spectrum systems: Threshold calculation and AGC algorithms," presented at GLOBECOM '89, Dallas, TX, Nov. 1989, pp. 1624-1628.

[27] E. Wong and M. Zakai, "On the convergence of ordinary integrals to stochastic integrals," *Ann. Math. Statist.*, vol. 36, pp. 1560-1564, 1965.

[28] R. E. Ziemer and R. L. Peterson, *Digital Communications and Spread Spectrum Systems.* New York: Macmillan, 1985.

Arnold L. Welti (S'84), for a photograph and biography, see this issue, p. 779.

Ben-Zion Bobrovsky (S'73-M'74-M'87) was born in Jerusalem, Israel, in 1943. He received the B.Sc. and M.Sc. degrees in mechanical engineering and the D.Sc. degree in electrical engineering from the Technion, Israel Institute of Technology, Haifa, in 1968, 1970, and 1974, respectively.

Since 1974 he has been with the Electrical Engineering Faculty, Tel Aviv University, Tel Aviv, Israel. During 1981-1982 and 1987-1989 he was on sabbatical at the Swiss Federal Institute of Technology (ETH), Zürich. His research interests include filtering theory, singular perturbation methods, synchronization, and tracking.

Maximal Length Sequences for Frequency Hopping

JOHN J. KOMO, SENIOR MEMBER, IEEE, AND SHYH-CHANG LIU

Abstract—Normally, frequency hopping sequences for spread spectrum communication systems are obtained by selecting groups of elements of binary m sequences. An alternative to using groups of binary m-sequence elements that is developed here is to obtain nonbinary m sequences with the number of desired hopping frequencies equal to the number of symbols in the finite field which the nonbinary m sequence is over. The grouping of elements of binary m sequences does not necessarily have the m-sequence balanced property or the run property of m sequences as do the nonbinary m sequences. In addition, the autocorrelation function of the nonbinary m sequence has a maximal period, whereas the frequency hopping sequences obtained from the grouping of elements of binary m sequences may not have a maximal period.

Introduction

An m sequence (maximal length sequence) over GF(p^m) (Galois field of p^m elements where p is a prime and m is an integer) of length $N = p^{mn} - 1$ is generated by a degree n primitive polynomial over GF(p^m) [1]-[3]. Let $a_0 a_1 a_2 \cdots$ denote an m sequence over GF(p^m) where the a_i's are elements of GF(p^m). Letting the primitive polynomial be

$$p(x) = x^n - c_{n-1}x^{n-1} - \cdots - c_1 x - c_0 \qquad (1)$$

where the c_i's are elements of GF(p^m), the m sequence $a_0 a_1 a_2 \cdots$ satisfies the recursion relationship

$$a_r = \sum_{i=1}^{n} c_{n-i} a_{r-i}. \qquad (2)$$

For $m = 1$, GF(p) is a prime finite field, and the elements of GF(p) are expressed as the integers 0, 1, $\cdots$, $p - 1$ ($p = 2$ is the binary finite field of two elements 0 and 1). Likewise, for $m \neq 1$, GF(p^m) is a prime extension field [an extension of GF(p)]. Since GF(p^m) is an extension of GF(p), the elements of GF(p^m) can be expressed as a vector representation of elements of GF(p). Now, GF(p^m) is an m-dimensional vector space over GF(p), and each element of GF(p^m) can be represented as [1], [2]

$$a = b_0 + b_1\gamma + b_2\gamma^2 + \cdots + b_{m-1}\gamma^{m-1} \qquad (3)$$

where the b_i's are elements of GF(p) for all $i = 0, 1, 2, \cdots, m - 1$ and $\{1, \gamma, \gamma^2, \cdots, \gamma^{m-1}\}$ forms a basis for GF(p^m) over GF(p). The elements of GF(p^m) are represented in terms of the basis elements using a degree m primitive polynomial over GF(p).

As an example, consider the degree 2 primitive polynomial over GF$(2^2)x^2 + x + \gamma$, which yields the recursion $a_r = a_{r-1} + \gamma a_{r-2}$ and generates the length $N = 2^4 - 1 = 15$ m sequence

$$11\gamma^2 10\gamma\gamma 1\gamma 0\gamma^2\gamma^2\gamma\gamma^2 0 \text{ (repeat)} \qquad (4)$$

where $\{1, \gamma\}$ is the basis for GF(2^2) over GF(2) obtained from the degree 2 primitive polynomial $x^2 + x + 1$ over GF(2). The elements of GF(2^2) are expressed as $0 = 0$, $\gamma^0 = 1$, $\gamma^1 = \gamma$, and $\gamma^2 = 1 + \gamma$. The vector m sequence of (4) can be represented in terms of its component m sequences as

$$\begin{aligned} 1 &\,|\, 111100010011010 \text{ (repeat)} \\ \gamma &\,|\, 001001101011110 \text{ (repeat).} \end{aligned} \qquad (5)$$

It can be noted that the component m sequence aligned with γ is a phase shift of 5 (left shift of 5) with respect to the component m sequence aligned with 1. Also, the component m sequence is a binary m sequence generated by the degree 4 primitive polynomial $x^4 + x + 1$ over GF(2), which yields the recursion $b_r = b_{r-3} + b_{r-4}$. The degree 2 primitive polynomial $x^2 + x + \gamma$ over GF(2^2) is a factor of the degree 4 primitive polynomial $x^4 + x + 1$ over GF(2), i.e., $x^4 + x + 1 = (x^2 + x + \gamma)(x^2 + x + \gamma^2)$. The other degree 2 primitive polynomial $x^2 + x + \gamma^2$ over GF(2^2) yields an m sequence that is different from (4). When this m sequence is represented in terms of its component m sequence as in (5), it has the same binary m sequence aligned with 1, but the m sequence aligned with γ is a phase shift of 10 with respect to the component m sequence aligned with 1.

The m sequence of (4) can be used for a frequency hopping system using four hopping frequencies with the correspondence $0 \rightarrow f_0$, $1 \rightarrow f_1$, $\gamma \rightarrow f_2$, and $\gamma^2 \rightarrow f_3$. The binary m sequence obtained from the degree 4 binary primitive polynomial $x^4 + x + 1$ over two periods is given as

$$111100010011010\ 111100010011010 \text{ (repeat).} \qquad (6)$$

Representing pairs of these elements as the representations of four hopping frequencies (the first of the pair is the coefficient of 1 and the second is the coefficient of γ), the following sequence over GF(2^2) is obtained in component form as

$$\begin{aligned} 1 &\,|\, 101111000100110 \text{ (repeat)} \\ \gamma &\,|\, 010011010111100 \text{ (repeat)} \end{aligned} \qquad (7)$$

Manuscript received February 24, 1989; revised January 12, 1990.
The authors are with the Department of Electrical and Computer Engineering, Clemson University, Clemson, SC 29634.
IEEE Log Number 9035809.

0733-8716/90/0600-0819$01.00 © 1990 IEEE

or in terms of elements of $GF(2^2)$ as

$$1\gamma 1 1\gamma^2\gamma^2 0\gamma 0\gamma^2\gamma\gamma\gamma^2 10 \text{ (repeat)}. \tag{8}$$

Even though (8) has many of the properties of an *m* sequence, it is not an *m* sequence.

Properties of *m* Sequences Over $GF(p^m)$

For an *m* sequence of period $q^n - 1$, with $q = p^m$, each of the $q^n - 1$ distinct nonzero sequences of length *n* over $GF(q)$ is initiated exactly once per period. The balanced property of *m* sequences is the statement that there is an approximately even distribution of each of the *q* elements of $GF(q)$ in one period of an *m* sequence. Since each possible length *n* nonzero sequence is initiated exactly once per period and q^{n-1} length *n* distinct sequences begin with each of the *q* nonzero symbols, each nonzero element occurs q^{n-1} times. Zero occurs once less, $q^{n-1} - 1$ times, since the all-zero sequence never occurs.

Let the term "run" denote a contiguous subsequence of identical symbols. The distribution of runs property is the statement that short runs of the same symbol are more numerous than longer runs of the same symbol. There are no runs longer than *n* of any symbol; otherwise, the generating register would contain those *n* symbols during successive iterations. Also, a single run of *n* digits is initiated once per period for each nonzero element in $GF(q)$ and a run of *n* zeros cannot occur. For runs of length $n - 1$, there are $2(q - 1)$ sequences of the form $x(b)^{n-1}$ or $(b)^{n-1}x$ where $x \neq b$ and $(b)^{n-1}$ indicates a run of $n - 1$ *b*'s. Each run of length *n* accounts for one of each form, and each run of $n - 1$ symbols also accounts for one of each form. Therefore, for each nonzero symbol, there are $q - 2$ runs of length $n - 1$ initiated per period. There is no length *n* runs of zeros, thus giving $q - 1$ runs of $n - 1$ zeros initiated per period. The runs of length $1 \leq r < n - 1$ are of the form $(A_i)^{n-r-2}x(b)^r y$ where $x \neq b$, $y \neq b$, and $(A_i)^{n-r-2}$ indicates a sequence of any $n - r - 2$ elements of $GF(q)$. Thus, there are $(q - 1)^2 q^{n-r-2}$ runs of length $1 \leq r < n - 1$ initiated per period for each element of $GF(q)$. These results are summarized in Table I.

The total number of runs of all lengths initiated in one period of an *m* sequence is given as

$$\begin{aligned}(q - 1) &+ \sum_{i=1}^{n-1} (q - 1)^2 q^{n-i-1} \\ &= (q - 1)\left[1 + \sum_{i=0}^{n-2} (q^{i+1} - q^i)\right] \\ &= (q - 1)q^{n-1}.\end{aligned} \tag{9}$$

The fraction of runs of length $1 \leq r < n - 1$ is obtained as

$$\frac{(q - 1)^2 q^{n-r-1}}{(q - 1)q^{n-1}} = \frac{q - 1}{q^r}. \tag{10}$$

TABLE I
Number of Runs Initiated per Period in an *m* Sequence of Length $q^n - 1$

Run length	Number per symbol $a_i \neq 0$	$a_i = 0$	Total number
r=n	1	0	q-1
r=n-1	q-2	q-1	$(q-1)^2$
1≤r<n-1	$(q-1)^2q^{n-r-2}$		$(q-1)^2q^{n-r-1}$

TABLE II
Number of Runs in an *m* Sequence with *q* a Power of 2

q	n	Total	Length	Number per symbol $a_i \neq 0$	$a_i = 0$	Total	Fraction
2^2	2	12	2	1	0	3	1/4
			1	2	3	9	3/4
	3	48	3	1	0	3	1/16
			2	2	3	9	3/16
			1	9		36	3/4
	4	192	4	1	0	3	1/64
			3	2	3	9	3/64
			2	9		36	3/16
			1	36		144	3/4
	5	768	5	1	0	3	1/256
			4	2	3	9	3/256
			3	9		36	3/64
			2	36		144	3/16
			1	144		576	3/4
2^3	2	56	2	1	0	7	1/8
			1	6	7	49	7/8
	3	448	3	1	0	7	1/64
			2	6	7	49	7/64
			1	49		392	7/8
	4	3584	4	1	0	7	1/512
			3	6	7	49	7/512
			2	49		392	7/64
			1	392		3136	7/8
	5	28672	5	1	0	7	1/4096
			4	6	7	49	7/4096
			3	49		392	7/512
			2	392		3136	7/64
			1	3136		25088	7/8
2^4	2	240	2	1	0	15	1/16
			1	14	15	225	15/16
	3	3840	3	1	0	15	1/256
			2	14	15	225	15/256
			1	225		3600	15/16
	4	61440	4	1	0	15	1/4096
			3	14	15	225	15/4096
			2	225		3600	15/256
			1	3600		57600	15/16
2^5	2	992	2	1	0	31	1/32
			1	30	31	961	31/32
	3	31744	3	1	0	31	1/1024
			2	30	31	961	31/1024
			1	961		30752	31/32
	4	1015808	4	1	0	31	1/32768
			3	30	31	961	31/32768
			2	961		30752	31/1024
			1	30752		984064	31/32
2^6	2	4032	2	1	0	63	1/64
			1	62	63	3969	63/64
	3	258048	3	1	0	63	1/4096
			2	62	63	3969	63/4096
			1	3969		254016	63/64

For the previous example with $q = 2^2$ and $n = 2$, it can be observed from (4) that there are $q = 4$ of the symbols 1, γ, and γ^2 and $q - 1 = 3$ zeros. Also, there is one run of length 2 of the symbols 1, γ, and γ^2, $q - 2 = 2$ runs of length 1 of the symbols 1, γ, and γ^2, and $q - 1 = 3$ runs of length 1 zero. Table II gives some results for the number of runs in an *m* sequence with *q* a power of 2. A power of 2 is the most common number of frequencies in a frequency hopping communication system.

A final property of m sequences is the subsequence length property. One period of an m sequence over GF(q) of length $q^n - 1$ can be divided into $q - 1$ equal length subsequences of length $(q^n - 1)/(q - 1)$ [4], [5]. These $q - 1$ subsequences are related to one another through multiplication by a nonzero constant in GF(q).

From the previous example with $q = 2^2$ and $n = 2$, the length of the subsequences is $(q^n - 1)/(q - 1) = 5$. It can be observed from (4) that the first subsequence is $11\gamma^2 10$, and that the second subsequence $\gamma\gamma 1 \gamma 0$ can be obtained from the first by multiplying by γ. Likewise, the third subsequence $\gamma^2\gamma^2\gamma\gamma^2 0$ can be obtained from the first by multiplying by γ^2.

Now, observing the sequence of (8), which is over GF(2^2), it can be seen that there are $q = 4$ of the symbols 1, γ, and γ^2 and $q - 1 = 3$ zeros. This sequence satisfies the balanced property of m sequences. Also, from (8), it can be seen that there is one run of length 2 of the symbols 1, γ, and γ^2, $q - 2 = 2$ runs of length 1 of the symbols 1, γ, and γ^2, and $q - 1 = 3$ runs of length 1 zero. Thus, this sequence satisfies the distribution of runs property of m sequences. Even though (8) satisfies the balanced property and the distribution of runs property, it does not satisfy the subsequence length property of m sequences, and (8) is not an m sequence. Since (8) is not an m sequence, it cannot be generated by a degree 2 primitive polynomial over GF(2^2).

Autocorrelation of m Sequences Over GF(p^m)

Let the vector m sequence over GF(p^m) be represented as $\boldsymbol{C} = \boldsymbol{c}_0\boldsymbol{c}_1\boldsymbol{c}_2 \cdots$ where the $\boldsymbol{c}_i$'s are row vectors with $\boldsymbol{c}_i = (c_{i0}, c_{i1}, \cdots, c_{im-1})$ and c_{ij} is the component aligned with the jth basis element γ^j. Now, define the vector mapping $\boldsymbol{\theta}(\boldsymbol{c}_i) = [\theta(c_{i0}), \theta(c_{i1}), \cdots, \theta(c_{i,m-1})]$ where

$$\theta(c_{ij}) = \exp\frac{j2\pi c_{ij}}{p} \qquad j = \sqrt{-1} \tag{11}$$

which yield the pth roots of unity. The autocorrelation function relative to $\boldsymbol{\theta}$ is then defined as [6].

$$R(\tau) = \frac{1}{m}\sum_{i=1}^{N-1} \boldsymbol{\theta}(\boldsymbol{c}_i)\left\{[\boldsymbol{\theta}(\boldsymbol{c}_{i+\tau})]^*\right\}^T \tag{12}$$

where $N = q^n - 1 = p^{mn} - 1$ and $*$ indicates complex conjunction. This autocorrelation function can be expressed in terms of the components as

$$R(\tau) = \frac{1}{m}\sum_{j=0}^{m-1}\sum_{i=0}^{N-1} \theta(c_{ij})\,\theta^*(c_{i+\tau,j}). \tag{13}$$

With $p = 2$, the mapping θ reduces to $\theta(0) = 1$ and $\theta(1) = -1$.

When the vector sequence $\boldsymbol{C}$ is an m sequence, the component $c_{0j}c_{1j}c_{2j} \cdots$ is an m sequence of length $q^n - 1$ and the inner sum of (13) is the bilevel autocorrelation function of an m sequence over GF(p). Thus, (13) reduces to

$$\begin{aligned} R(\tau) &= q^n - 1 \qquad \tau = k(q^n - 1), \quad k \text{ an integer} \\ &= -1 \qquad \text{otherwise.} \end{aligned} \tag{14}$$

For the previous example, it can be seen that the vector m sequence of (4), with $q = 2^2$ and $n = 2$, has the autocorrelation function

$$\begin{aligned} R(\tau) &= 15 \qquad \tau = 15k, \quad k \text{ an integer} \\ &= -1 \qquad \text{otherwise.} \end{aligned} \tag{15}$$

Also, using (13), it can be seen that the sequence of (8) yields the same bilevel autocorrelation function of (15).

Frequency Hopping Sequences

Frequency hopping sequences of 2^m distinct frequencies can be obtained as a direct correspondence with the elements of the finite field GF(2^m) [3] as they occur in an m sequence over GF(2^m). An m sequence over GF(2^m), of length $N = 2^{mn} - 1$, can be generated directly with a degree n primitive polynomial over GF(2^m) or with a degree mn primitive polynomial over GF(2), along with the amount of the shifts for the binary component m sequences [7]. Table III lists primitive polynomials over GF(2) and their corresponding shifts, along with the equivalent primitive polynomial over GF(2^m) for various values of m and n. Also included is the basis polynomial for expressing the elements of GF(2^m) as a vector with elements from GF(2). In Table III, the primitive polynomial over GF(2) and the basis polynomials [also over GF(2)] are given in an octal representation of the binary coefficients of the polynomials with the highest power on the left. The primitive polynomial over GF(2^m) is listed with the coefficients of the power of γ plus 1, and 0 indicates that there is no γ. The amount that the component m sequences are shifted to the left relative to the first component is listed as a base shift and a factor in parantheses, where the product of the factor and the base shift indicates the left shift of that particular component m sequence. For the case where $m = 3$ and $n = 3$ ($N = 2^9 - 1 = 511$), the primitive polynomial over GF(2) is $x^9 + x^4 + 1$ (since 1021_8 equals 1000010001_2), the basis polynomial for expressing elements of GF(2^3) over GF(2) is $x^3 + x + 1$ (since $13_8 = 1011_2$), the primitive polynomial over GF(2^3) is $x^3 + \gamma x^2 + \gamma^5 x + \gamma$, and the component m sequences are left shifted 146 and 73, with respect to the first component, for the second and third components, respectively.

When the frequency hopping sequences are obtained by grouping elements of binary m sequences and then represented by component sequences, the component sequences are the original m sequence decimated by m (taking every mth bit of the given m sequence). Each component is a shifted version of the first component. For $m = 2$ and $m = 4$, the decimation by m yields a shifted version of the given m sequence. For $m = 3$, 5, and 6, the decimation by m yields a different sequence. In these cases, the sequence is not an m sequence unless m and $N = 2^{mn} - 1$ are relatively prime. When m divides N, the length of the sequence reduces to N/m, and the sequence does not have the balance property nor the run property of an m sequence. In addition, the bilevel autocorrelation does not have the maximal period.

TABLE III
GENERATORS FOR FREQUENCY HOPPING *m* SEQUENCES

m	n	generator GF(2)	shift(factors)	generator GF(2^m)	basis
2	2	23	5(1)	112	7
	3	103	21(1)	1132	7
	4	435	85(1)	11222	7
	5	2011	341(1)	113202	7
	6	10123	1365(1)	1000112	7
	7	42103	5461(1)	10013132	7
	8	210013	21845(1)	101011112	7
3	2	147	9(2 1)	172	13
	3	1021	73(2 1)	1262	13
	4	10123	585(2 1)	12412	13
	5	100003	4681(2 1)	103312	13
	6	1000201	37449(2 1)	1362242	13
	7	10000005	299593(2 1)	15252262	13
4	2	435	17(3 2 1)	1 3 2	23
	3	10123	273(3 2 1)	1 11 14 2	23
	4	210013	4369(3 2 1)	1 1 0 1 2	23
5	2	2033	33(11 25 2 1)	1 19 2	57
	3	100003	1057(11 25 2 1)	1 0 21 2	57
6	2	10123	65(62 61 60 59 1)	1 31 2	141
	3	1000201	4161(62 61 60 59 1)	1 39 28 2	141

TABLE IV
DISTRIBUTION OF RUNS FOR GROUPING OF BINARY ELEMENTS WITH $m = 3$ AND $n = 4$

run length	zero	nonzero elements						
1	408	366	381	384	390	390	420	381
2	42	51	57	60	42	33	57	57
3	9	12	3	12	6	0	6	3
4	0	0	0	3	3	0	0	0

TABLE V
FREQUENCY HOPPING SEQUENCES THAT ARE NOT *m* SEQUENCES

m	n
3	2i
5	4i
7	3i
9	2i
i an integer ≥1	

As an example, for $m = 3$ and $n = 2$, $N = 2^6 - 1 = 3^2 \cdot 7$ and m and N are nor relatively prime. For this sequence, the zero element occurs nine times, three nonzero elements occur 12 times, three nonzero elements occur six times and one nonzero element does not occur. Also, the zero element does not have a run of length 2, two nonzero elements have three runs of length 2, and five nonzero elements do not have a run of length 2. The autocorrelation function is bilevel, but has the reduced period of length 21 (maximal length 63). Table IV lists the number of runs of the elements of the sequence obtained from the grouping method for $m = 3$ and $n = 4$. From this table, it can be seen that this sequence does not have the m-sequence run property or the m-sequence balanced property.

Table V lists several values of m and n for which the frequency hopping sequences obtained by grouping elements of binary m sequences does not have the balance property nor the run property of an m sequence and the bilevel autocorrelation does not have the maximal period.

CONCLUSION

A procedure for obtaining frequency hopping sequences with good random properties and good autocorrelation properties has been presented. This procedure was shown to be superior to selecting groups of elements of binary m sequences. These frequency hopping sequences are obtained directly from m sequences over the finite field GF(2^m) which is an extension field of GF(2). These sequences have the balanced property and run property of m sequences. They also have the bilevel autocorrelation values of m sequences with the maximal period.

REFERENCES

[1] R. J. McEliece, *Finite Fields for Computer Scientists and Engineers*. Boston, MA: Kluwer, 1987.

[2] R. Lidl and H. Niederreiter, *Introduction to Finite Fields and Their Applications*. New York: Cambridge Univ. Press, 1986.

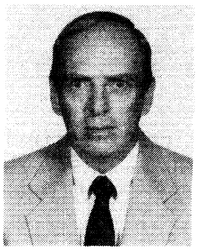
[3] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications, Vol. I*. Rockville, MD: Computer Science Press, 1985.

[4] W. J. Park, Jr., "An investigation of maximum length linear recursive sequences over finite fields," Ph.D. dissertation, Dep. Elec. Comput. Eng., Clemson Univ., Clemson, SC, May 1986.

[5] N. Zierler, "Linear recurring sequences," *J. Soc. Indust. Appl. Math.*, vol. 7,'pp. 31-48, 1959.

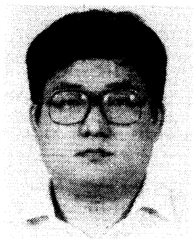
[6] W. J. Park, Jr. and J. J. Komo, "The autocorrelation of m-sequences over nonprime finite fields," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, pp. 459-461, July 1988.

[7] M. S. Lam, "Generation of m-sequences and Gold codes over GF(q)," Ph.D. dissertation, Dep. Elec. Comput. Eng., Clemson Univ., Clemson, SC, Dec. 1988.



John J. Komo (S'60-SM'87) was born in St. Louis, MO, on January 1, 1942. He received the B.S., M.S., and Ph.D. degrees from the University of Missouri, Rolla, in 1962, 1963, and 1966, respectively.

He was on the faculty of St. Louis University from 1965 to 1968, the faculty of the University of Missouri, Columbia, from 1968 to 1973, and the faculty of Clemson University from 1973 to the present. He has been a consultant to the McDonnell-Douglas Corporation, Harris Corporation, and Warner Robbins Logistic Center, and is currently a Professor of Electrical and Computer Engineering at Clemson University. His current research interests are centered on sequences and finite field applications, and he is the author of the book, *Random Signal Analysis in Engineering Systems* (New York: Academic, 1987).



Shyh-Chang Liu was born in Taipei, Taiwan, on May 27, 1959. He received the B.S. degree from the National Tsing Hua University of Taiwan in 1981 and the M.S. degree from Clemson University, Clemson, SC, in 1986.

From 1981 to 1983 he served as an Officer in the Chinese Army. He is currently pursuing the Ph.D. degree in electrical engineering at Clemson University. His current research is in digital communications and coding theory applications.

Convolutionally Coded Frequency-Hopping Communications with Nonideal Interleaving

SHAUL LAUFER, MEMBER, IEEE, AND ARIE REICHMAN, MEMBER, IEEE

Abstract—**The interleaving span of coded frequency-hopped (FH) systems is often constrained to be smaller than the decoder memory length, i.e., nonideal interleaving is performed. An upper bound on the performance of a Viterbi decoder of a convolutional code with nonideal interleaving is presented. Also, a soft decision diversity combining technique is introduced, and the performance of combined convolutional and diversity coding subject to worst case partial band noise jamming is investigated. Optimization of the FH system performance subject to constraints of allowed delay and synthesizer settling time provides the best combination of interleaving span and hopping rate. The FH system considered employs *M*-ary frequency-shift key (MFSK) modulation and noncoherent demodulation with 2 b soft decision based on Viterbi's ratio-threshold technique.**

I. Introduction

The need for interleaving in coded frequency-hopped (FH) systems is well known [1], [2]. Interleaving converts a channel with memory to one that can be treated as memoryless. Previous works, e.g., [2], [9], and many others, dealing with analysis of the performance of coded FH communications, assumed ideal interleaving, meaning that each code symbol which belongs to a specific codeword is transmitted in a different hop. However, in many cases, the interleaving span is constrained to be small, e.g., the allowed delay introduced by the interleaving system is limited or short messages are transmitted in a slow FH system (SFH). In these cases, there is a degradation in the performance of the decoder due to the nonideal interleaving.

Interleaving is a form of time diversity achieved at a cost of buffer storage and transmission delay, but does not require additional coding redundancy. Nevertheless, additional diversity may be implemented when larger coding gain is required or when a single rate modem is used for transmitting multirate data. Diversity is added to an existing coding system by transmitting each *M*-ary symbol *m* times on separate hops. Effectively, we get a concatenated coding system with an *m*-fold repetition code as the innermost code.

In this paper, we shall investigate the effect of nonideal interleaving on the performance of a convolutionally encoded FH *M*-ary frequency-shift key (MFSK) modulation system with noncoherent demodulation and hard or soft decision Viterbi decoding. Soft decision metrics are generated by Viterbi's ratio-threshold anti-jamming (AJ) technique [5], [6]. Also, a soft decision diversity combining method is presented, and the performance of the combined convolutional and diversity coding is investigated subject to nonideal interleaving.

The performance evaluation is based on computation of the superchannel (which includes the repetition encoder and the soft decision diversity combiner) transition probabilities and an upper bound on the bit error rate (BER) of the Viterbi decoder. This technique of upper bound evaluation is based on the code transfer function and the channel transition probabilities [3]. There are two criteria for performance evaluation: the average BER, and the availability. The former is useful for both ideal and nonideal interleaving, while the latter is applicable only for nonideal interleaving and provides more information on the BER statistic.

A realizable FH system is subject to the constraints of allowed delay and synthesizer settling time, leading to a tradeoff between interleaving span and hopping rate. We shall present an optimization problem of the FH system performance which provides the best combination of the interleaving span and hopping rate subject to these constraints.

In this paper, the jammer performs worst case partial band noise jamming (WCPBNJ). It is a white stationary Gaussian process with zero mean and two-sided spectral density $N_J/2$. Also, an additive white Gaussian noise (AWGN) environment is assumed, with two-sided spectral density $N_0/2$. The jammer has *a priori* knowledge of all FH system parameters, except for the pseudonoise (PN) code, and devises an optimum strategy of restricting its total power J to a fraction ρ of the full spread spectrum bandwidth W which causes the worst decoded BER. According to the definition of a partial-band jammer, we have $(W_{\text{MFSK}}/W) \leq \rho \leq 1$ where W_{MFSK} is the bandwidth of the MFSK signal and depends on M. However, throughout this paper, we assume $W_{\text{MFSK}} << W$ as applicable for practical FH systems. The processing gain PG is defined as $\text{PG} = W/R_b$ where R_b is the information rate. This definition is independent of specific design parameters such as M and the code rate R. Hence, the parameter

Manuscript received February 24, 1989; revised February 8, 1990. This paper was presented in part at the 1988 IEEE International Conference on Communications, Philadelphia, PA, June 1988.

S. Laufer is with the Communication Group, Tadiran Ltd., Holon 58102, Israel and the Department of Electronic Systems, Tel Aviv University, Tel Aviv 69978, Israel.

A. Reichman is with the Communication Group, Tadiran Ltd., Holon 58102, Israel.

IEEE Log Number 9035810.

$E_b/N_J = (P/R_b)/(J/W) = (P/J)(W/R_b)$, where P is the received signal power, may be used for comparison of systems with different M and R, but with the same information rate R_b and total bandwidth available W.

The outline of this paper is as follows. In Section II, the system model and the channel characterization are presented. In Section III, an upper bound on the performance of a Viterbi decoder of a nonideally interleaved convolutional code is developed and the notion of availability is introduced. In Section IV, a method of soft decision decoding of concatenated convolutional and diversity coding is presented, and a technique for evaluating its performance subject to nonideal interleaving is given. The corresponding results follow in Section V. In Section VI, the optimization problem is presented, and finally, a summary and conclusions are given in Section VII.

II. System Description and Definitions

A. System Model

Fig. 1 depicts the FH/MFSK system. The source produces independent, equally likely, binary input symbols. The binary data are convolutionally encoded by an encoder with constraint length $K = 7$ and rate $R = 1/2$. The convolutional code symbols are repeated m times and interleaved. The MFSK modulator chooses one of M orthogonal frequencies based on $\log_2 M$ binary symbols from the interleaved sequence. In the receiver, the M-ary noncoherent demodulator performs both a conventional hard decision and a ratio-threshold test [5], [6] which generates a quality bit q per each M-ary symbol. The quality bit q is attached to each of the $\log_2 M$ binary symbols associated with the M-ary symbol. Soft decision diversity combining on the m-fold repetition code symbols is done after deinterleaving. The diversity combiner also generates a quality bit according to a specific algorithm described in Section IV, and provides a 2 b soft decision per each binary convolutional code symbol to the Viterbi decoder.

B. Coding Channel Characterization

We consider an FH/MFSK system which transmits I binary code symbols per hop ($I/\log_2 M$ should be an integer). The binary code symbols are interleaved by a pseudorandom block interleaver [10] with dimensions NI where I is the interleaver depth and N is the interleaver span. The product NI is assumed to be very large in comparison to the convolutional code constraint length. The channel (for the code symbols) is a binary-input, output-symmetric channel, shown in Fig. 2, with four-level output according to the following ratio-threshold test [5]. Let $Q_1, Q_2, \cdots, Q_M$ be the M matched filter outputs. Then the $\log_2 M$ binary decision symbols are extracted from the index of Q_j where

$$Q_j = \max_i Q_i. \tag{1}$$

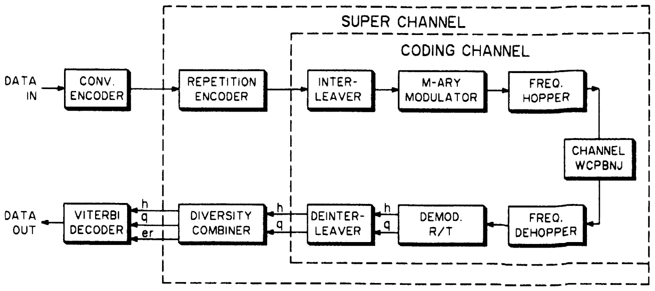


Fig. 1. System block diagram.

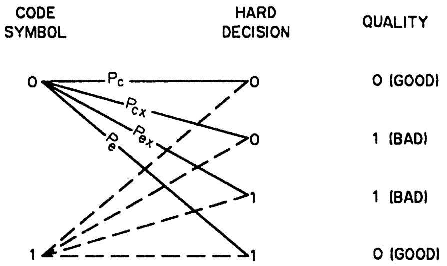


Fig. 2. Coding channel model.

In addition, a quality bit q is derived by forming $M - 1$ ratios as follows:

$$q = \begin{cases} 0\ (\text{good}) & \text{if } Q_j/Q_k \geq \Theta \text{ for all } k \neq j \\ 1\ (\text{bad}) & \text{otherwise} \end{cases} \tag{2}$$

where the ratio threshold $\Theta \geq 1$ is a parameter chosen by the communicator. In the sequel, a fixed value of Θ is used. It is expected that a randomly varying threshold [6] will further improve the performance.

The discrete transition probabilities of the channel are given by [8]

$$P_c = F_c(\Theta) + (M/2 - 1)F_e(\Theta) \tag{3}$$

$$P_{cx} = [F_c(1) - F_c(\Theta)] + (M/2 - 1)[F_e(1) - F_e(\Theta)] \tag{4}$$

$$P_{ex} = (M/2)[F_e(1) - F_e(\Theta)] \tag{5}$$

$$P_e = (M/2)F_e(\Theta) \tag{6}$$

where the F's are the exceeding probabilities given by [8]

$$F_c(\Theta) = \Pr\{Q_j \geq \Theta Q_i \quad \text{for all } i \neq j \,|\, j \text{ sent}\}$$

$$F_e(\Theta) = \Pr\{Q_n \geq \Theta \max_{i \neq n} Q_i \,|\, j \text{ sent and } n \neq j\}.$$

Let a partial band jammer jam a fraction ρ of the frequency hopping bandwidth W, with a total power J. Then, if the frequency slot selected by the communicator is within the jammed fraction, it confronts a noise level $N_0 + N_J/\rho$ where N_j is defined by $N_J = J/W$ watts/Hz. Assuming that the hopping is a random process, the occurrence of this event is with probability ρ, while with probability $1 - \rho$, only background thermal noise is encountered with noise level N_0.

The probabilities $F_c(\Theta)$, $F_e(\Theta)$ may be written as [8]

$$F_c(\Theta) = 1 + \sum_{k=1}^{M-1} \binom{M-1}{k} \frac{(-1)^{k\Theta 2}}{(k+\Theta^2)} \cdot \exp\left(-\frac{k}{(k+\Theta^2)} \frac{E_s}{N_I}\right) \tag{7}$$

$$F_e(\Theta) = \frac{1}{M-1} \sum_{k=1}^{M-1} \binom{M-1}{k} \cdot \frac{(-1)^{k+1\Theta 2}}{(k+\Theta^2)} \frac{k}{(k+\Theta^2-1)} \cdot \exp\left(-\frac{k+\Theta^2-1}{(k+\Theta^2)} \frac{E_s}{N_I}\right) \tag{8}$$

where E_s is the energy of an M-ary modulation symbol. E_s is given by

$$E_s = \frac{r}{m} E_b \log_2 M \tag{9}$$

where R is the convolutional code rate and E_b is the bit energy. N_I is the noise level confronted in the specific frequency slot:

$$N_I = \begin{cases} N_0 + N_J/\rho & \text{for a jammed hop} \\ N_0 & \text{for a nonjammed hop.} \end{cases} \tag{10}$$

The resulting transition probabilities are denoted by $P_{c,j}$, $P_{cx,j}$, $P_{ex,j}$, $P_{e,j}$ for a jammed hop, and by $P_{c,o}$, $P_{cx,o}$, $P_{ex,o}$, $P_{e,o}$ for a nonjammed hop.

Assuming that ideal interleaving is performed (each code symbol is transmitted in a different hop), the average transition probabilities are given by

$$\overline{P}_c = (1-\rho) P_{c,o} + \rho P_{c,j} \tag{11}$$

$$\overline{P}_{cx} = (1-\rho) P_{cx,o} + \rho P_{cx,j} \tag{12}$$

$$\overline{P}_{ex} = (1-\rho) P_{ex,o} + \rho P_{ex,j} \tag{13}$$

$$\overline{P}_e = (1-\rho) P_{e,o} + \rho P_{e,j}. \tag{14}$$

C. Upper Bound on BER

The upper bound on the bit error probability P_b for a Viterbi decoder of a convolutional code is given by [3]

$$P_b \leq \sum_{d=df}^{\infty} b(d) P_d \tag{15}$$

where d is the Hamming distance between two code vectors over their unmerged segment d_f (the free distance) and the coefficients $b(d)$ are determined solely by the specific convolutional code. P_d is the pairwise error probability for two code vectors with Hamming distance d over the unmerged segment. P_d depends only on the coding channel and the decoder metric. For the $K = 7$, $R = 1/2$ commonly used convolutional code found by Odenwalder [4], the bit error bound is

$$P_b \leq 36P_{10} + 211P_{12} + 1404P_{14} + 11633P_{16} + 77433P_{18} + 502690P_{20} + \cdots . \tag{16}$$

The pairwise error probability P_d is calculated assuming the Viterbi decoder uses only integer branch metrics and a binary input, output-symmetric channel, including an erasure output. The technique for the exact calculation of P_d was given in [3, problem (4.20)] and is extended for the case of a channel with an erasure output. The vector $\{\pi_k, k = 0, \pm 1, \pm 2, \cdots, \pm\Psi/2\}$ represents the conditional probabilities given that a "zero" was transmitted where k are the symmetric integer output levels. Let

$$\pi(z) = \sum_{k=-\Psi/2}^{\Psi/2} \pi_k z^k$$

be the corresponding generating function of the transition probabilities conditioned on a "zero" being sent. Then P_d is the sum of the coefficients of the negative powers and a half of the coefficient of the zero power of the series $[\pi(z)]^d$.

III. Nonideal Interleaving

Binary convolutional codes are not capable of dealing directly with error bursts. Even practical nonbinary codes cannot deal directly with long error bursts. Therefore, in a slow FH system, interleaving is required to disperse the error bursts caused by a partial band jammer.

Ideal interleaving means that the decoder sees a memoryless channel where the impact of jamming or other channel interferences is independent from any received code symbol to another. Hence, for a block code, it means that no two code symbols of any codeword are transmitted in the same hop. Therefore, the interleaver span N should be made at least equal to the block length. For a convolutional code, N should be larger than the decoder memory length, which is not as well defined as for block codes.

In many cases, the interleaving span is constrained to be smaller than the block length or the decoder memory length, e.g., 1) when the allowed delay introduced by the interleaving system ($2NI$ for the block interleaver [3]) is limited, or 2) when short messages, which should be interleaved separately, are transmitted in a slow FH system (N is constrained to be equal to L/I where L is the message length). In these cases of nonideal interleaving, a performance degradation is expected in the decoding process because error bursts are not fully dispersed.

For nonideal interleaving with a span of N hops, we shall assume that separate encoding and decoding are performed on each NI block of interleaving. This assumption enables the calculation of the decoder's BER for each block of interleaving separately according to the average transition probabilities in the specific block, which depend on the number of jammed hops n out of the N interleaved hops. To minimize any degradation which may result from operating a convolutional code on blocks of data, we shall assume that $NI >> K$ and that there is appropriate code synchronization, state initialization, and finalization. This method transforms the problem to one of selecting a channel, out of $N + 1$ different channels, for a period of NI code symbols. Due to pseudorandom interleaving, each of these channels may be considered memoryless (in a very close approximation for $I >> 1$) as the channel model given in Fig. 2. Each channel is characterized by a specific set of transition probabilities according to n.

A. Upper Bound on Average BER

The average decoded BER P_b for interleaving of span N is

$$P_b = \sum_{n=0}^{N} P_n P_b(n) \tag{17}$$

where P_n is the probability of having n jammed hops out of N interleaved hops given by

$$P_n = \binom{N}{n} \rho^n (1 - \rho)^{N-n} \tag{18}$$

and $P_b(n)$ is the decoded BER in the deinterleaved sequence of length NI with n jammed hops out of N.

For the computation of the upper bound on $P_b(n)$, we have to calculate the average transition probabilities for the specific case of n jammed hops out of N:

$$\overline{P}_c(n) = \left(1 - \frac{n}{N}\right) P_{c,o} + \frac{n}{N} P_{c,j} \tag{19}$$

$$\overline{P}_{cx}(n) = \left(1 - \frac{n}{N}\right) P_{cx,o} + \frac{n}{N} P_{cx,j} \tag{20}$$

$$\overline{P}_{ex}(n) = \left(1 - \frac{n}{N}\right) P_{ex,o} + \frac{n}{N} P_{ex,j} \tag{21}$$

$$\overline{P}_e(n) = \left(1 - \frac{n}{N}\right) P_{e,o} + \frac{n}{N} P_{e,j}. \tag{22}$$

The calculation of the upper bound on $P_b(n)$ is based on (15) and (16) as described in Section II. This upper bound is tight for decoded BER less than 10^{-3} [3]. Therefore, an upper bound on P_b based on (17), which averages upper bounds on $P_b(n)$, some of them in the nontight region of a decoded BER higher than 10^{-3}, may suffer from untightness, larger than that of the upper bound on $P_b(n)$. The tightness of the bound was investigated by simulation for the case of BFSK with $K = 7$ rate $1/2$ convolutional code, hard decision Viterbi decoding. The results are demonstrated in Fig. 5. The difference between the bound and the simulation, at a decoded BER of 10^{-5} and $N = 20$, is 2.1 dB. The tightness depends on the interleaving span and on the specific code and modulation. It is expected that soft decision decoding will result in better tightness since the curves (in Figs. 6–10) are closer to the wide-band curve which is known to be a tight bound for a decoded BER under 10^{-3} [3].

B. Definition of Availability

The coding channel of FH communications with partial band jamming and nonideal interleaving is a channel with memory. The noise severity level differs from one block of interleaving to another according to the number of jammed hops chosen for a block of interleaving. Therefore, the BER level out of the decoder differs from one block of interleaving to another. When short messages are transmitted ($L = NI$) or when the terminal equipment frame synchronization algorithm requires BER $\le T_a$ (T_a is a specified threshold) for any block of interleaving, the upper bound on the average BER is not a sufficient measure. For these cases, a definition of probability of availability is required: $P_a(E_b/N_J) = \text{Prob}\,\{\text{BER per block of interleaving} \le T_a\}$. Let $^u P_b(n)$ denote the upper bound on $P_b(n)$ computed as described in Section III-A above. The probability of availability may be lower bounded by

$$P_a(E_b/N_J) \ge \min_{\rho} \sum_{n=0}^{n_a} P_n$$

where $n_a = \lfloor {}^u P_b^{-1}(T_a, E_b/N_J, \rho) \rfloor$ and P_n is given by (18). $\lfloor x \rfloor$ is the largest integer less than or equal to x. Hence, n_a is the largest number of jammed hops in a block of interleaving for which the threshold T_a is not exceeded, i.e., $^u P_b(n_a, E_b/N_J, \rho) \le T_a$. The aforementioned lower bound on P_a results from using the upper bound on $P_b(n)$ for evaluating n_a.

IV. Soft Decision Diversity Combining

Repetition code can be used in an FH system as the only error-correcting code [2], [7] or as the inner code concatenated with an outer convolutional code [2]. The repetition code provides coding gain against partial band jamming and flexibility in operating a single rate modem with multirate data by changing the code rate $R = 1/m$. However, there is an optimum m^* for achieving the best performance against partial band jamming. In this section, we analyze the performance of concatenated repetition and convolutional codes with nonideal interleaving. Nevertheless, we assume that the repetition code symbols are interleaved ideally, i.e., the repetition code symbols of any convolutional code symbol are transmitted in different hops. This condition can be fulfilled if $N \ge m$.

The soft decision diversity combiner processes the received soft decision (ratio-threshold test) repetition code symbols and outputs soft decision convolutional code symbols according to the following algorithm.

- Majority decoding is performed on the good quality ($q = 0$) repetition code symbols corresponding to a single convolutional code symbol, and a good quality bit is attached to the decision made.

- If all the repetition code symbols, corresponding to a single convolutional code symbol, were received with bad quality ($q = 1$) or a tie occurred among the good quality symbols, then majority decoding is performed on all bad quality symbols and a bad quality bit is attached to the decision made. If a tie occurred again (possible for m even), an erasure is declared.

Note that a tie means that the number of zeros is equal to the number of ones. Hence, when a tie occurs, majority decoding cannot produce a decision. If a tie occurs among good quality symbols, they can be excluded from further consideration since the good quality zeros neutralize the good quality ones.

The diversity combining algorithm described above, performed after deinterleaving, composes a superchannel with five-level output for the convolutional code symbols (see Fig. 3). We have to compute the transition probabilities of the five-level superchannel in order to calculate the upper bound on the BER performance of the Viterbi decoder.

Denote the transition probabilities of the superchannel by P_c^s, P_{cx}^s, P_{er}^s, P_{ex}^s, P_e^s where P_{er}^s is the probability of erasure. Assuming that the repetition code symbols are interleaved ideally and using the soft decision diversity combining algorithm, we get the following superchannel transition probabilities for n jammed hops out of N for $m = 2$ diversity. Similar expressions, obtained for $m = 3$ and $m = 4$, are excluded since they are very long and do not contribute to the understanding.

$m = 2$:

$$\begin{aligned} P_c^s(n) = {} & P_{20}\left[P_{c,o}^2 + 2P_{c,o}(P_{cx,o} + P_{ex,o})\right] \\ & + P_{21}\left[P_{c,o}(P_{c,j} + P_{cx,j} + P_{ex,j}) + P_{cx,o}P_{c,j}\right. \\ & \left. + P_{ex,o}P_{c,j}\right] \\ & + P_{22}\left[P_{c,j}^2 + 2P_{c,j}(P_{cx,j} + P_{ex,j})\right] \end{aligned} \tag{23}$$

$$P_{cx}^s(n) = P_{20}P_{cx,o}^2 + P_{21}P_{cx,o}P_{cx,j} + P_{22}P_{cx,j}^2 \tag{24}$$

$$\begin{aligned} P_{er}^s(n) = {} & P_{20}(2P_{c,o}P_{e,o} + 2P_{cx,o}P_{ex,o}) \\ & + P_{21}(P_{e,o}P_{c,\mathrm{j}} + P_{c,o}P_{e,j} + P_{ex,o}P_{cx,j} \\ & + P_{cx,o}P_{ex,j}) \\ & + P_{22}(2P_{c,j}P_{e,j} + 2P_{cx,j}P_{ex,j}) \end{aligned} \tag{25}$$

$$P_{ex}^s(n) = P_{20}P_{ex,o}^2 + P_{21}P_{ex,o}P_{ex,j} + P_{22}P_{ex,j}^2 \tag{26}$$

$$\begin{aligned} P_e^s(n) = {} & P_{20}\left[P_{e,o}^2 + 2P_{e,o}(P_{cx,o} + P_{ex,o})\right] \\ & + P_{21}\left[P_{e,o}(P_{e,j} + P_{ex,j}\right. \\ & \left. + P_{cx,j}) + P_{ex,o}P_{e,j} + P_{cx,o}P_{e,j}\right] \\ & + P_{22}\left[P_{e,j}^2 + 2P_{e,j}(P_{cx,j} + P_{ex,j})\right] \end{aligned} \tag{27}$$

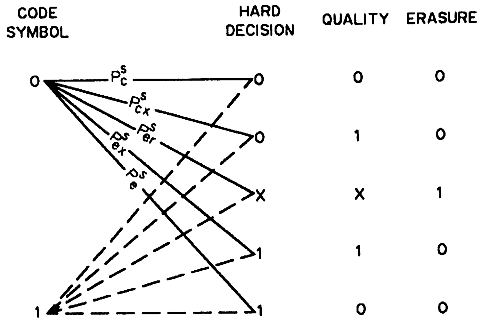


Fig. 3. Superchannel model.

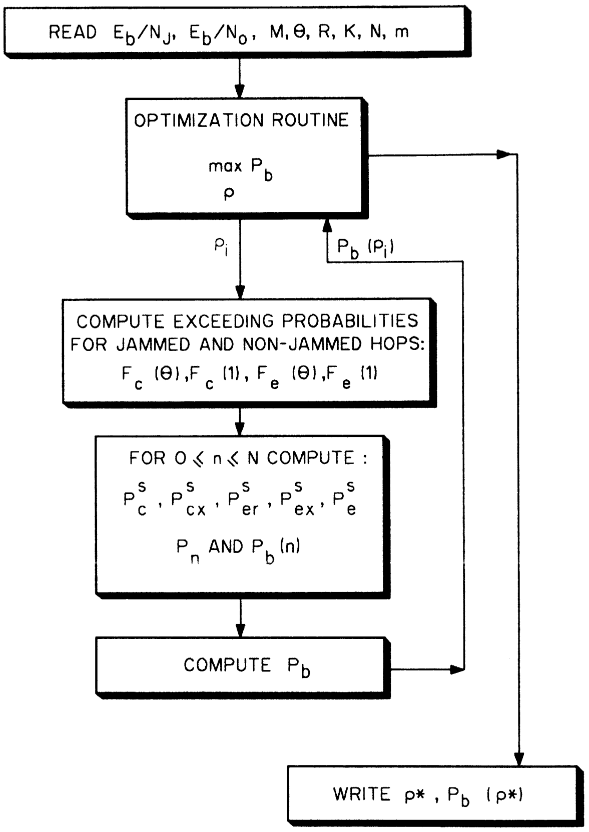

READ E_b/N_J, E_b/N_0, M, θ, R, K, N, m
OPTIMIZATION ROUTINE
max P_b
ρ
ρ_i
$P_b(\rho_i)$
COMPUTE EXCEEDING PROBABILITIES FOR JAMMED AND NON-JAMMED HOPS:
$F_c(\theta), F_c(1), F_e(\theta), F_e(1)$
FOR $0 \leqslant n \leqslant N$ COMPUTE :
$P_c^s, P_{cx}^s, P_{er}^s, P_{ex}^s, P_e^s$
P_n AND $P_b(n)$
COMPUTE P_b
WRITE ρ*, P_b (ρ*)

Fig. 4. A procedure for computation of the upper bound on the BER subject to worst case partial-band noise jamming (ρ_i and $P_b(\rho_i)$ corresponds to the ith optimization iteration).

where P_{20}, P_{21}, and P_{22} were substituted for:

$$P_{20} = P(\text{none is jammed}) = \left(1 - \frac{n}{N}\right)\left(1 - \frac{n}{N - 1}\right) \tag{28}$$

$$\begin{aligned} P_{21} = P(\text{one is jammed}) = {} & \left(1 - \frac{n}{N}\right)\left(1 - \frac{n}{N - 1}\right) \\ & + \left(\frac{n}{N}\right)\left(1 - \frac{n - 1}{N - 1}\right) \end{aligned} \tag{29}$$

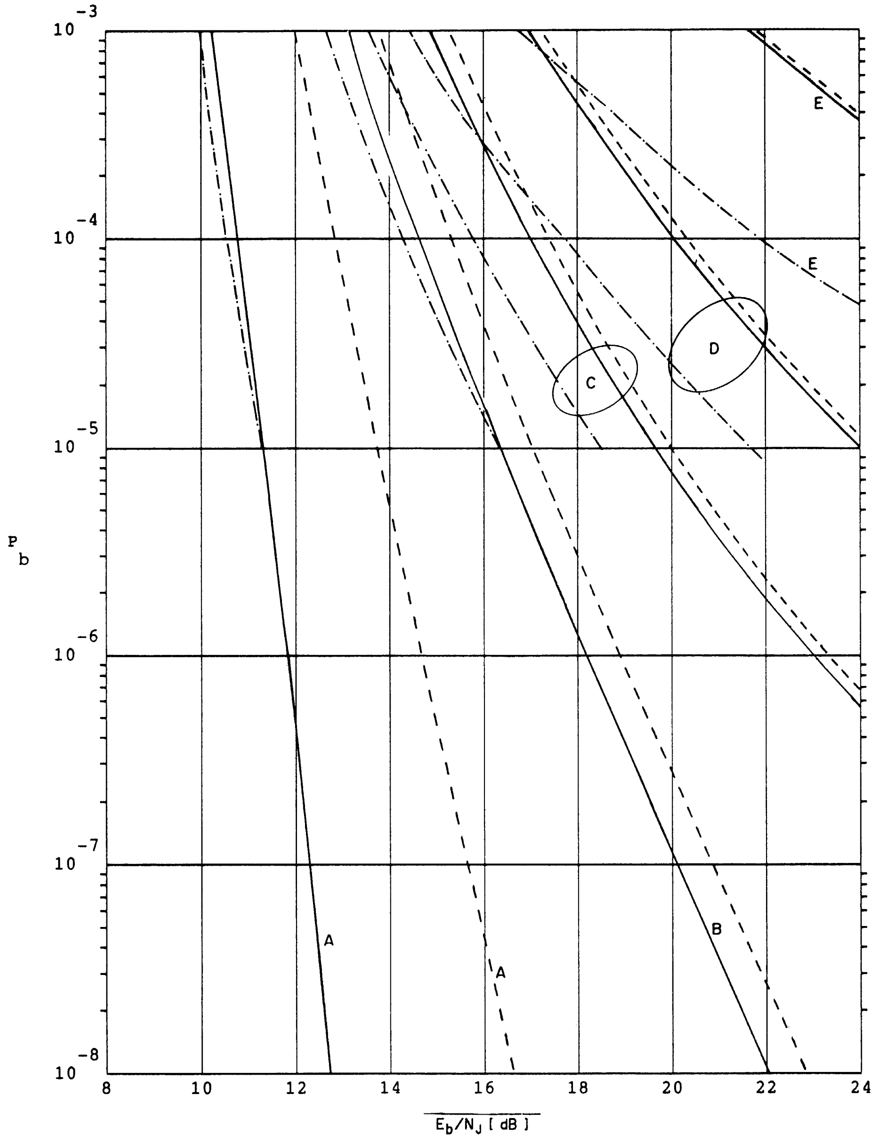


Fig. 5. BER performance of BFSK modulation with hard decision ($\Theta = 1$), $K = 7$, $R = 1/2$ convolutional code without diversity. (a) Wide-band jamming, (b) WCPBNJ, ideal interleaving, (c), (d), (e) WCPBNJ, nonideal interleaving, interleaving span of $N = 40$, 20, 10 hops, respectively. Continuous lines for negligible thermal noise ($N_0 = 0$); dashed lines for $E_b/N_0 = 15$ dB. Dot-dash lines for simulation results (with negligible thermal noise).

$$P_{22} = P(\text{both are jammed}) = \left(\frac{n}{N}\right)\left(\frac{n-1}{N-1}\right). \qquad (30)$$

The upper bound on $P_b(n)$ may be computed using the superchannel transition probabilities and (16) where the pairwise error probability $P_d(n)$ for n jammed hops out of N is used. Then, the upper bound on the average BER P_b is computed according to (17).

For ideal interleaving, P_{20}, P_{21}, and P_{22} in (23)-(27) are substituted for $(1-\rho)^2$, $2\rho(1-\rho)$, and ρ^2, respectively.

General expressions for the superchannel transition probabilities for m-fold diversity combining, according to the algorithm given above, when ideal interleaving is performed are given by

$$P_c^s = \sum_{i=1}^{m} \binom{m}{i} (\overline{P}_{cx} + \overline{P}_{ex})^{m-i} \cdot \left(\sum_{k=\lceil (i+1)/2 \rceil}^{i} \binom{i}{k} \overline{P}_c^k \overline{P}_e^{i-k} \right) \qquad (31)$$

$$P_{cx}^s = \sum_{i=0}^{m} \binom{m}{i} \left(\left\lceil \frac{i+1}{2} \right\rceil - \left\lfloor \frac{i+1}{2} \right\rfloor \right) \binom{i}{i/2} (\overline{P}_c \overline{P}_e)^{i/2} \cdot \left(\sum_{k=\lceil (m-i+1)/2 \rceil}^{m-i} \binom{m-i}{k} \overline{P}_{cx}^k \overline{P}_{ex}^{m-i-k} \right) \qquad (32)$$

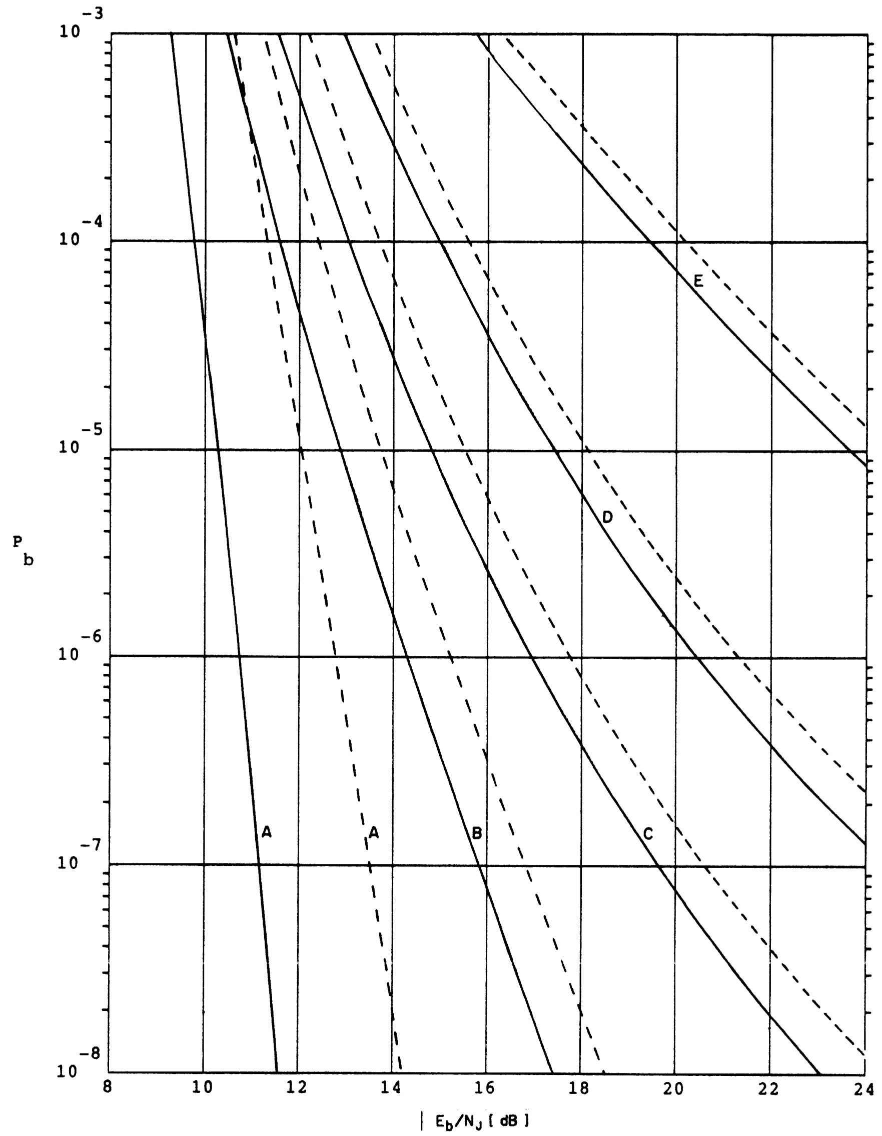


Fig. 6. BER performance of BFSK modulation with soft decision ($\Theta = 2$), $K = 7$, $R = 1/2$ convolutional code without diversity. (a) Wide-band jamming, (b) WCPBNJ, ideal interleaving, (c), (d), (e) WCPBNJ, nonideal interleaving, interleaving span of $N = 40, 20, 10$ hops, respectively. Continuous lines for negligible thermal noise ($N_0 = 0$); dashed lines for $E_b/N_0 = 15$ dB.

$$P_{er}^s = \sum_{i=0}^{m} \binom{m}{i} \left(\left\lceil \frac{i+1}{2} \right\rceil - \left\lfloor \frac{i+1}{2} \right\rfloor \right) \binom{i}{i/2} (\overline{P}_c \overline{P}_e)^{i/2} \cdot \left(\left\lceil \frac{(m-i)+1}{2} \right\rceil - \left\lfloor \frac{(m-i)+1}{2} \right\rfloor \right) \cdot \binom{m-i}{(m-i)/2} (\overline{P}_{cx} \overline{P}_{ex})^{(m-i)/2} \qquad (33)$$

$$P_{ex}^s = \sum_{i=0}^{m} \binom{m}{i} \left(\left\lceil \frac{i+1}{2} \right\rceil - \left\lfloor \frac{i+1}{2} \right\rfloor \right) \binom{i}{i/2} (\overline{P}_c \overline{P}_e)^{i/2} \cdot \left(\sum_{k=\lceil (m-i+1)/2 \rceil}^{m-i} \binom{m-i}{k} \overline{P}_{ex}^{k} \overline{P}_{cx}^{m-i-k} \right) \qquad (34)$$

$$P_e^s = \sum_{i=1}^{m} \binom{m}{i} (\overline{P}_{cx} + \overline{P}_{ex})^{m-i} \cdot \left(\sum_{k=\lceil (i+1)/2 \rceil}^{i} \binom{i}{k} \overline{P}_e^{k} \overline{P}_c^{i-k} \right) \qquad (35)$$

where $\overline{P}_c$, $\overline{P}_{cx}$, $\overline{P}_{ex}$, $\overline{P}_e$ are computed by (11)-(14). The index i counts the repetition code symbols which have good quality bit. $\lceil x \rceil$ is defined as the smallest integer greater than or equal to x; $\lfloor x \rfloor$ is defined as the largest integer less than or equal to x.

V. Numerical Results

The upper bound on the average BER performance of the concatenated ECC system was computed for several cases for hard ($\Theta = 1$) and soft ($\Theta = 2$) decisions and

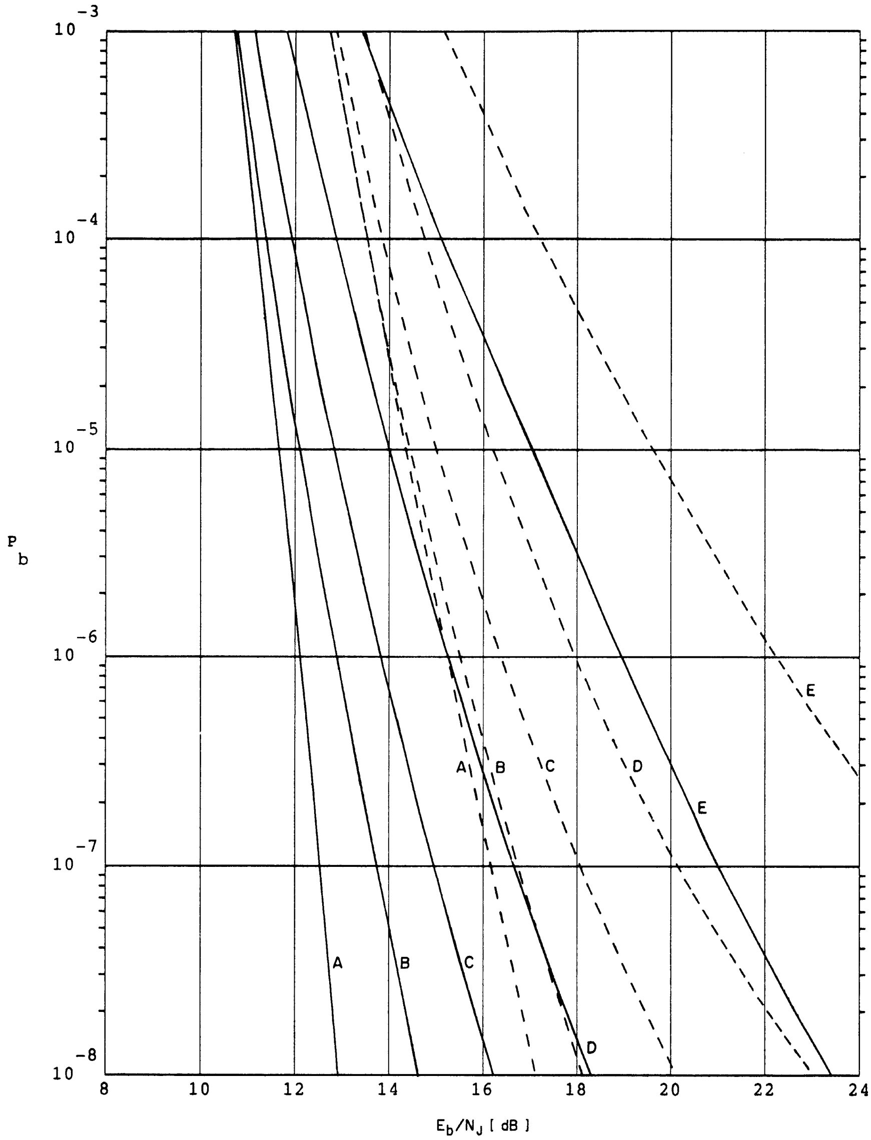


Fig. 7. BER performance of BFSK modulation with soft decision ($\Theta = 2$), $K = 7$, $R = 1/2$ convolutional code with $m = 2$ diversity. (a) Wide-band jamming, (b) WCPBNJ, ideal interleaving, (c), (d), (e) WCPBNJ, nonideal interleaving, interleaving span of $N = 40, 20, 10$ hops, respectively. Continuous lines for negligible thermal noise ($N_0 = 0$); dashed lines for $E_b/N_0 = 15$ dB.

various combinations of M (modulation signals), m (repetition coding), and N (interleaving span). The convolutional code was Odenwalder's best $K = 7$, $R = 1/2$ code [4]. A $K = 7$, $R = 1/4$ code was generated by repeating the rate $1/2$ code twice (the bound for this code is similar to the bound given in (16), but the indexes d of P_d are doubled). An optimization program was performed for evaluating the decoded BER performance subject to a worst case partial band jammer and the appropriate optimum fraction of band jammed ρ^*. The procedure for computation of the average BER is described in Fig. 4.

A sample of the results is presented in Figs. 5-10. The BER performance is given for binary and 4-ary modulation with hard ($\Theta = 1$) and soft decision ($\Theta = 2$) with a $K = 7$, rate $1/2$ or $1/4$ convolutional code, with ($m = 2$) and without ($m = 1$) a repetition code. Results are given for wide-band jamming and for WCPBNJ with ideal and nonideal interleaving.

In Fig. 5, we illustrate the performance for BFSK modulation, noncoherent detection with a hard decision, $K = 7$, rate $1/2$ convolutional code with no repetition code. The degradation due to nonideal interleaving, for negligible thermal noise, for a required average BER of 10^{-5} is 3.3 dB for an interleaving span of $N = 40$ and 7.7 dB for $N = 20$.

In Fig. 6, we illustrate the performance for a similar case, but with soft decision decoding. Comparing Figs. 5 and 6, we conclude that soft decision decoding with the ratio-threshold test quality bit achieves a significant gain over hard decision decoding.

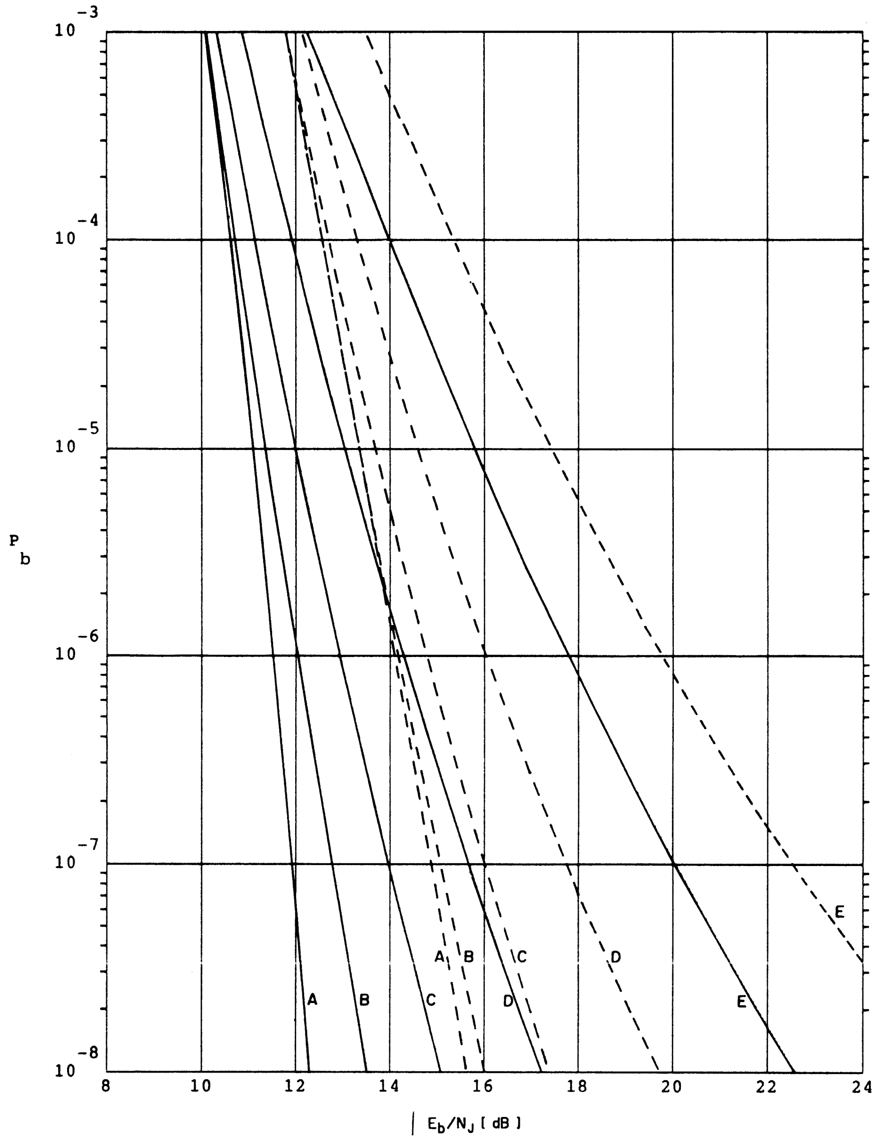


Fig. 8. BER performance of BFSK modulation with soft decision ($\Theta = 2$), $K = 7$, $R = 1/4$ convolutional code without diversity. (a) Wide-band jamming, (b) WCPBNJ, ideal interleaving, (c), (d), (e) WCPBNJ, nonideal interleaving, interleaving span of $N = 40, 20, 10$ hops, respectively. Continuous lines for negligible thermal noise ($N_0 = 0$); dashed lines for $E_b/N_0 = 15$ dB.

In Fig. 7, we illustrate the performance for a similar case as in Fig. 6, but with two-fold repetition code concatenated to the convolutional code in the transmitter and soft decision diversity combining in the receiver. Although there is a degradation in the performance against wide-band jamming, a significant gain is achieved against WCPBNJ when ideal interleaving is performed, and even a larger gain when nonideal interleaving is performed. However, it is clear from Fig. 7 that the coding system becomes more sensitive to wide-band thermal noise due to larger noncoherent combining loss for a lower code rate. The same code, with a code rate of 1/4, could be decoded directly by an appropriate Viterbi decoder. Fig. 8 presents the performance for this case, which is slightly better than the performance for the concatenated coding system with the same rate.

In Figs. 9 and 10, we illustrate the performance for 4-ary FSK modulation with soft decision decoding, with and without diversity. Comparing Figs. 9 and 6 or Figs. 10 and 7, it can be concluded that 4FSK achieves a gain of about 2 dB (at average BER of 10^{-5}) over BFSK for wide-band jamming, and even larger gain against WCPBNJ. Also, the sensitivity to wide-band thermal noise is reduced.

The optimum diversity against WCPBNJ is $m^* = 2$ for ideal interleaving [2]. When nonideal interleaving is performed, larger diversity is required to mitigate some of the additional degradation. See Table I.

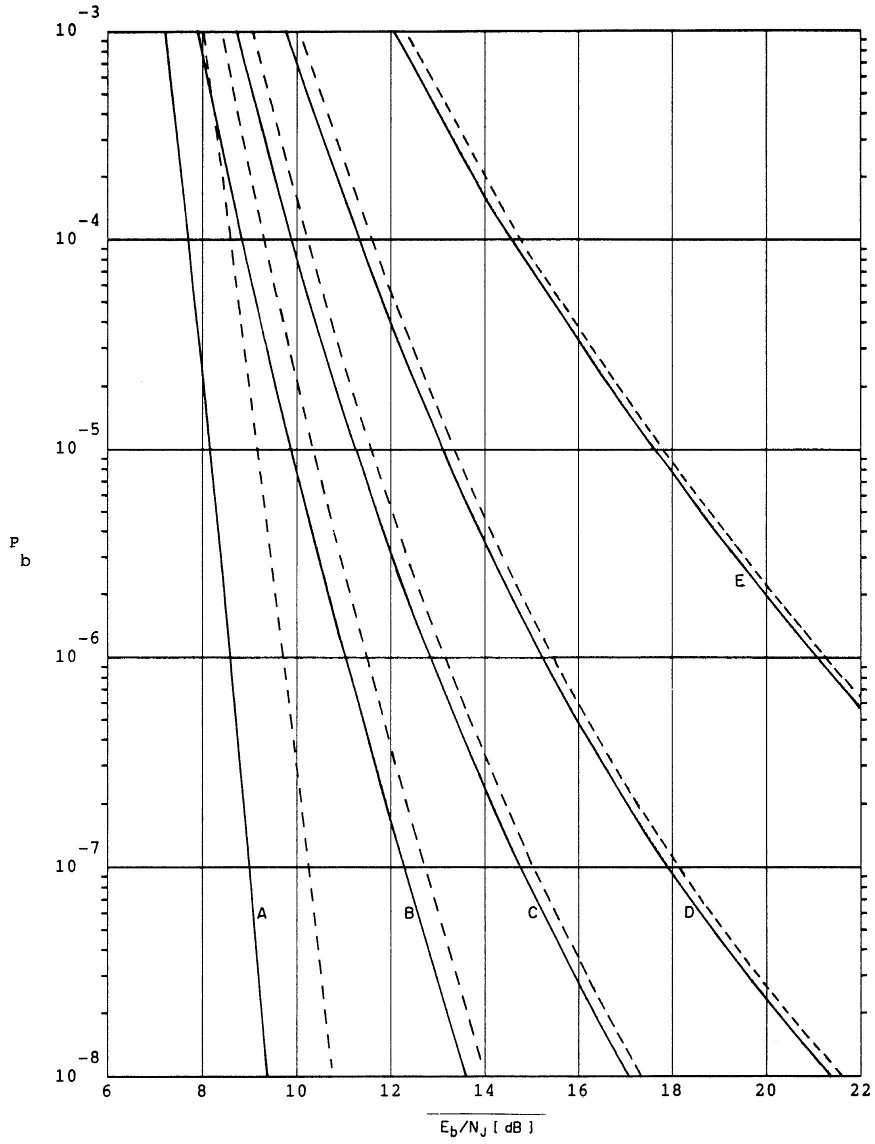


Fig. 9. BER performance of 4FSK modulation with soft decision ($\Theta = 2$), $K = 7$, $R = 1/2$ convolutional code without diversity. (a) Wide-band jamming, (b) WCPBNJ, ideal interleaving, (c), (d), (e) WCPBNJ, nonideal interleaving, interleaving span of $N = 40$, 20, 10 hops, respectively. Continuous lines for negligible thermal noise ($N_0 = 0$); dashed lines for $E_b/N_0 = 15$ dB.

Table II summarizes the performance at required average BER of 10^{-5}.

The concept of availability, which was introduced in Section III, is demonstrated in Table III for the requirements of $T_a = 10^{-4}$ and $P_a \geq 0.999$.

VI. Optimization of FH System Performance

A realizable FH system is subject to the constraints of allowed delay and synthesizer settling time. Increasing the interleaving span reduces the degradation due to nonideal interleaving. However, for a given allowed delay, the interleaving span can increase only by increasing the hopping rate. On the other side, increasing the hopping rate, while the absolute time overhead per hop (mainly due to synthesizer settling time) remains constant, requires increasing the instantaneous transmission rate, which reduces the energy per bit and causes a degradation. Therefore, a tradeoff has to be made between the interleaving span and the hopping rate. We shall present an optimization problem of the FH system performance which provides the best combination of the interleaving span and hopping rate subject to the constraints of allowed delay and synthesizer settling time.

Denote the overhead time per hop by T_0 seconds, the allowed delay by T_d seconds, and the hopping rate by $R_h = 1/T_h$ hops/s. For a block interleaving with a span of N hops, we have

$$T_d = 2 \cdot N \cdot T_h. \tag{36}$$

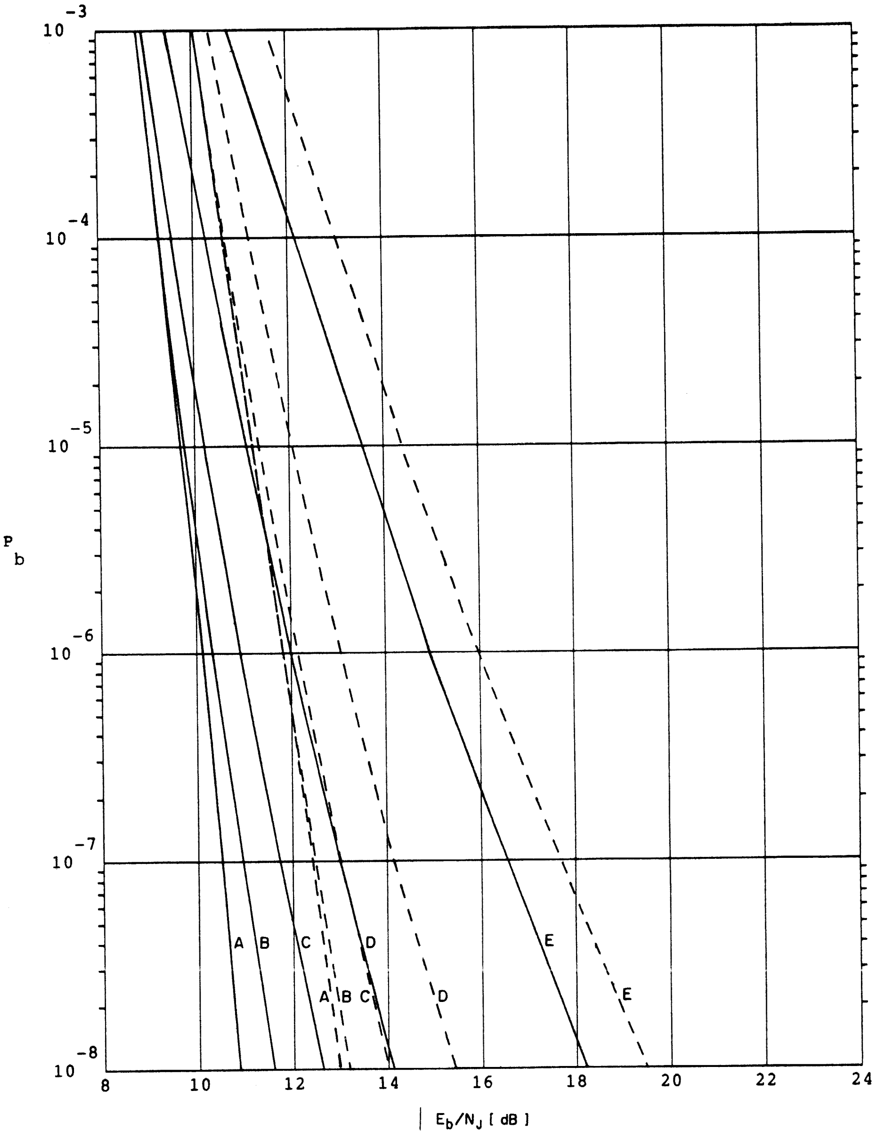


Fig. 10. BER performance of 4FSK modulation with soft decision ($\Theta = 2$), $K = 7$, $R = 1/2$ convolutional code with $m = 2$ diversity. (a) Wide-band jamming, (b) WCPBNJ, ideal interleaving, (c), (d), (e) WCPBNJ, nonideal interleaving, interleaving span of $N = 40, 20, 10$ hops, respectively. Continuous lines for negligible thermal noise ($N_0 = 0$); dashed lines for $E_b/N_0 = 15$ dB.

TABLE I
THE OPTIMUM DIVERSITY m^* AGAINST WCPBNJ AT AVERAGE BER OF 10^{-5} WITH NEGLIGIBLE THERMAL NOISE

		N=10	N=20	N=50	IDEAL
M=2	Θ=1	≥4	≥4	≥4	2
	Θ=2	≥4	3	2	2
M=4	Θ=1	≥4	≥4	≥4	2
	Θ=2	2	2	2	2

The efficiency of the data transmission denoted by η is

$$\eta(R_h) = \frac{T_h - T_0}{T_h} = 1 - \frac{T_0}{T_h} = 1 - T_0 \cdot R_h \quad (37)$$

resulting in a degradation in the bit energy E_b of $10 \cdot \log_{10} \eta$ (dB) (assuming the transmitter cannot increase its peak power). Substituting (36) into (37), we get η as a function of the interleaving span N:

$$\eta(N) = 1 - \frac{2 \cdot T_0}{T_d} \cdot N = 1 - \alpha \cdot N \quad (38)$$

where α is substituted for $2 \cdot T_0/T_d$.

TABLE II
REQUIRED E_b/N_J (dB) AND WORST CASE ρ (ρ^*) FOR $K = 7$ CONVOLUTIONALLY CODED FH/MFSK SYSTEM WITH NONIDEAL INTERLEAVING AT AVERAGE BER OF 10^{-5}

M	Θ	R	m	N	E_b/N_J (ρ*)	
					N_o=0	E_b/N_o=15 dB
2	1	1/2	1	10	33.9 (.0004)	34.0 (.0004)
				20	24.0 (0.005)	24.3 (0.005)
				40	19.6 (0.02)	20.0 (0.02)
				ID	16.3 (0.09)	17.0 (0.11)
2	2	1/2	1	10	23.7 (0.004)	24.3 (0.004)
				20	17.4 (0.03)	18.1 (0.02)
				40	14.8 (0.07)	15.5 (0.08)
				ID	12.9 (0.19)	13.7 (0.21)
2	2	1/2	2	10	17.0 (0.06)	19.6 (0.03)
				20	14.0 (0.20)	16.2 (0.13)
				40	12.9 (0.33)	15.0 (0.29)
				ID	12.1 (0.56)	14.4 (0.65)
2	2	1/4	1	10	15.8 (0.09)	17.4 (0.06)
				20	13.0 (0.24)	14.6 (0.21)
				40	12.0 (0.40)	13.7 (0.44)
				ID	11.4 (0.64)	13.3 (0.90)
4	2	1/2	1	10	17.6 (0.01)	17.8 (0.01)
				20	13.1 (0.07)	13.4 (0.07)
				40	11.3 (0.14)	11.6 (0.16)
				ID	9.9 (0.29)	10.3 (0.33)
4	2	1/2	2	10	13.6 (0.16)	14.4 (0.11)
				20	11.1 (0.33)	12.1 (0.35)
				40	10.2 (0.54)	11.4 (0.65)
				ID	9.8 (0.82)	11.2 (1.00)

TABLE III
REQUIRED E_b/N_J (dB) AND WORST CASE ρ (ρ^*) FOR $K = 7$ CONVOLUTIONALLY CODED FH/MFSK WITH NONIDEAL INTERLEAVING FOR PROBABILITY OF AVAILABILITY $P_a \geq 0.999$ AND $T_a = 10^{-4}$

M	Θ	R	m	N	E_b/N_J (ρ*) , n_a
					N_o=0
2	2	1/2	1	10	20.1 (0.054) , 3
				20	18.8 (0.054) , 5
				40	17.2 (0.055) , 7

The optimization problem definition is

$$P_{b,\text{opt}} = \min_{N} \max_{\rho} P_b[N, \rho, \eta(N) \cdot E_b/N_J] \quad (39)$$

subject to

$$\eta(N) = 1 - \alpha \cdot N$$

where $\eta(N) \cdot E_b$ is the effective energy per bit of information.

Solving this optimization problem, we get the best interleaving span (denoted by N^*) which minimizes the decoder's BER for WCPBNJ.

A solution for the optimization problem is illustrated in Fig. 11 for a specific example of an FH/BFSK system with a concatenated $K = 7$, rate 1/2 convolutional code and a twofold repetition code with soft decision decoding and negligible thermal noise (this case was illustrated in Fig. 7). Also, in this example, $\alpha = 1/50$ was selected. The solution is based on minimizing the total degradation $D_{\text{total}} = D_{\text{int}} + D_{\text{rate}}$ for a required average BER of 10^{-5} where D_{int} is the degradation due to nonideal interleaving for a required BER of 10^{-5} and $D_{\text{rate}} = 10 \cdot \log_{10} \eta$ is the degradation due to the inefficiency. It is clear from Fig. 11 that the optimum span for this example is $N^* = 20$, resulting in a surprisingly low efficiency $\eta^* = 60\%$. Compared to a higher efficiency of 80%, the optimum span provides a gain of 1.8 dB. The resulting optimal hopping rate is [according to (37)] $R_h^* = (1 - \eta^*)/T_0 = 0.40/T_0$.

VII. SUMMARY

In this paper, the effect of nonideal interleaving on the performance of convolutionally encoded FH/MFSK systems subject to WCPBNJ was investigated. Also, a soft decision diversity combining method was presented, and the performance of the combined convolutional and diversity coding subject to WCPBNJ was investigated. An

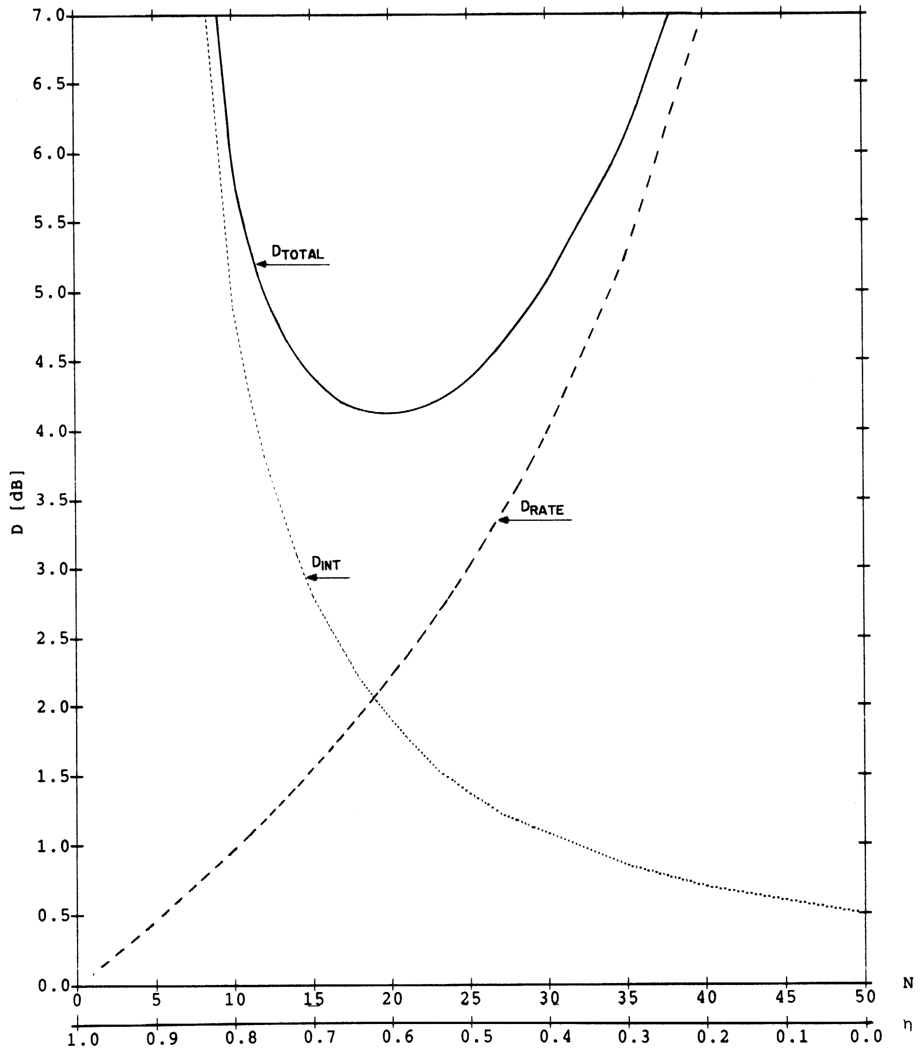


Fig. 11. The degradation D (dB) as a function of the interleaving span N or the efficiency η for BFSK modulation with soft decision ($\Theta = 2$), $K = 7$, $R = 1/2$ convolutional code, concatenated with $m = 2$ repetition code. The curves are plotted for a BER of 10^{-5} and assuming negligible thermal noise ($N_0 = 0$). Also, $\alpha = 1/50$ was selected. (a) Degradation due to inefficiency (D_{rate}). (b) Degradation due to nonideal interleaving (D_{int}). (c) The total degradation ($D_{\text{total}} = D_{\text{rate}} + D_{\text{int}}$).

optimization problem for the system performance in WCPBNJ was defined, leading to the best interleaving span (and the appropriate hopping rate).

The results lead to the following conclusions.

- Nonideal interleaving causes significant degradation. A convolutional code $K = 7$, $R = 1/2$, with an interleaving span less than 10, is almost useless.
- Concatenating a repetition code mitigates much of the degradation caused by nonideal interleaving.
- Soft decision decoding with the ratio-threshold test quality bit achieves a significant gain over hard decision decoding both with and without diversity.
- The optimum diversity against WCPBNJ is $m^* = 2$ for ideal interleaving. When nonideal interleaving is performed, larger diversity is required to mitigate some of the additional degradation.
- For achieving the best system performance, the optimal interleaving span and hopping rate should be used.

Acknowledgment

The authors wish to thank an anonymous referee for his useful comments that helped improve the presentation of this paper.

References

[1] A. J. Viterbi, "Spread spectrum communication—Myths and realities," *IEEE Commun. Mag.*, vol. 17, pp. 11-18, May 1979.

[2] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications, Vol. I & II.* Rockville, MD: Computer Science Press, 1985.

[3] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding.* New York: McGraw-Hill, 1979.

[4] J. P. Odenwalder, "Optimum decoding of convolutional codes," Ph.D. dissertation, School Eng. Appl. Sci., Univ. California, Los Angeles, 1970.

[5] A. J. Viterbi, "A robust ratio-threshold technique to mitigate tone and partial band jamming coded MFSK systems," in *Proc. MILCOM'82*, pp. 22.4.1-22.4.5.

[6] L.-F. Chang and R. J. McEliece, "A study of Viterbi's ratio-threshold AJ technique," in *Proc. MILCOM'84*, pp. 11.2.1-11.2.5.

[7] C. M. Keller and M. B. Pursley, "A comparison of diversity combining techniques for frequency-hop communications with partial-band interference," in *Proc. MILCOM'85*, pp. 33.1.1-33.1.5.

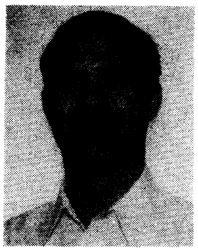
[8] A. J. Viterbi, "Robust decoding of jammed frequency hopped modulation," in R. E. Kalman *et al.*, Ed., *Recent Advances in Communication and Control Theory.* Optimization Software Inc., 1988, pp. 47-129.

[9] J. K. Omura and B. K. Levitt, "Coded error probability evaluation for antijam communication systems," *IEEE Trans. Commun.*, vol. COM-30, pp. 896-903, May 1982.

[10] G. C. Clark and J. B. Cain, *Error-Correction Coding for Digital Communications.* New York: Plenum, 1981.

Shaul Laufer (M'86) was born in Petah-Tikva, Israel. He received the B.Sc. and M.Sc. degrees in electrical engineering from Tel Aviv University, Tel Aviv, Israel, in 1979 and 1985, respectively. He is presently studying towards the Ph.D. degree at Tel Aviv University.

He is currently a Staff Scientist at the Communication Group, Tadiran Ltd., Holon, Israel, where he is involved in research and development in the fields of error-correction codes, spread spectrum systems, modulation techniques, and digital communications. From 1979 to 1984 he was involved in research and development in the field of communications at the Israeli Ministry of Defence. In 1985 he was involved in research and development of error-correction systems and secure data links at Linkabit Ltd., Israel. His research interests include spread spectrum and multiple-user communication, error control coding, and digital modulation techniques.



Arie Reichman (S'82-M'85) was born on August 2, 1949. He received the B.Sc. and M.Sc. degrees in electrical engineering from the Technion, Israel Institute of Technology, Haifa, in 1971 and 1975, respectively, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1984.

From 1971 to 1981 he served in the Israel Defence Forces as an Electrical Engineer where, in his last position, he was head of a research and development group. He was responsible for the analysis and development of advanced communication systems. In 1984 he joined Tadiran Ltd., Holon, Israel, as a Scientist and the head of the Applied Research Department of the Communication Group. He is currently responsible for the analysis and development of spread spectrum systems, and of advanced receivers based on adaptive methods and digital signal processing. He is also an Adjunct Professor at Tel Aviv University.

On the Power Spectral Density of Certain Digitally Modulated Signals with Applications to Code Despreading

NIKOS B. PRONIOS, MEMBER, IEEE, AND ANDREAS POLYDOROS, MEMBER, IEEE

Abstract—**The purpose of this paper is to demonstrate techniques for deriving the power spectral density (psd) of certain digitally modulated signals, which are more general and easier to use than other similar methods. These techniques are especially well suited for deriving the psd of signal-product waveforms, with arbitrary modulating pulse-shapes. This is done by decomposing the psd expression into two factors, one depending solely on the underlying sequences and the other depending only on the pulse-shapes. General formulas are derived and, for some cases, they are expressed in terms of the discrete Fourier transform of appropriately defined sequences and the Fourier transform of the modulating pulse-shape. Applications where these expressions would be useful include bit and code synchronizers, delay-and-multiply type of detectors, spread-spectrum and code-division multiple-access systems, either at radio or at optical frequencies.**

I. Introduction

CONSIDER the generic low-pass equivalent communication/radar system model of Fig. 1, which incorporates the baseband complex envelopes $\tilde{s}_i(t)$; $i = 1, 2$ of two signals of interest $s_i(t) = \text{Re}\,[\tilde{s}_i(t)e^{j2\pi f_c t}]$ around the frequency f_c, namely, the received and the local signals, as well as the frequency and phase offsets $\Delta\omega$ and ϕ, respectively, and the noise complex envelope $\tilde{\eta}(t)$. The aforementioned offsets can be attributed to transmitter-receiver motion, channel disturbances, etc. whereas the timing offset Δ at the receiver site can model unintentional epoch-synchronization error or deliberately introduced delay. The receiver front-end bandpass (BP) filter shown in Fig. 1 is typically wide enough to pass a significant portion of the energy in $s_1(t)$, plus any expected Doppler offset, while limiting the undesirable noise at the same time. The problem addressed in this paper concerns the evaluation of the spectral characteristics of either $\tilde{s}_1(t)$ or the product waveform $g(t) = \tilde{s}_1(t)\,\tilde{s}_2^*(t + \Delta)$, with the emphasis on the second. The desirability of such an evaluation is discussed at length below.

This model is general enough to accommodate a great variety of receiver operations, whose specific purpose is designated by the choise of the postmultiplier filter (PMF) and the final processor (FP). For example, in data demodulation/detection of completely known signals (i.e., with $\Delta\omega = \phi = 0$) in additive white Gaussian noise, $\tilde{s}_2(t)$ can represent any one of the possible transmitted signals, the PMF is an integrator and the FP takes the real part Re $[\cdot]$; it then performs appropriate comparisons with the other similar branches. This is the standard correlation demodulator [4], [21], [29]. For noncoherent demodulation (ϕ a random variable), the FP forms the appropriate envelopes and compares. Similarly, in radar target detection, $\tilde{s}_2(t)$ is the pulse matched to the one transmitted, the PMF is again an integrator (resulting in the standard matched-filtering operation) and the FP combines pulses and decides.

In another class of operations, the information sequence embedded in $\tilde{s}_1(t)$ (see elaborate description below) is of no interest and can be treated as random, either exactly or approximately. Problems in this category include random-signal detection/interception [28], [20], [2], bit synchronization [4], [29] and code synchronization [4], [18], [19], [14], [25], etc. When energy-detection of an unknown signal is desired (a problem very common in target detection, spread-spectrum interception and code acquisition), we can set $\tilde{s}_2(t) = 1$ and have the FP perform this energy detection with a quadratic nonlinearity. The model is also applicable to the "delay-and-multiply" operation, widely used as a bit synchronizer as well as an interceptor which is less immune to noise-level fluctuations. Here, the noise-corrupted received waveform $\tilde{r}(t)$ is multiplied with a version of itself, which is delayed by an appropriate amount Δ, then narrow-band filtered at an appropriate frequency (such as the bit rate) and, finally, phase-lock-loop tracked (bit sync) or energy-detected (interception). As far as the signal-times-signal ($s \times s$) portion of this operation is concerned, the model of Fig. 1 is applicable with the aforementioned appropriate choice of parameters/processors.

In the previous discussion, the information embedded in $\tilde{s}_1(t)$ was modeled as random. In the spread-spectrum category of applications, for which Fig. 1 is again directly applicable, $\tilde{s}_1(t)$ and $\tilde{s}_2(t)$ actually contain deterministic periodic sequences (codes). These are either pseudorandom (PN), maximal-length sequences employed for anti-

Manuscript received November 7, 1989; revised February 27, 1990. This work was supported in part by NSF-PYI under Contract NCR-8552527 and in part by ARO under Contract DAAL03-88-K-0050.

N. B. Pronios is with Bell Communications Research, Bldg. NVC-IC 202, Red Bank, NJ 07701.

A. Polydoros is with the Department of Electrical Engineering—Systems, Communication Sciences Institute, University of Southern California, Los Angeles, CA 90089.

IEEE Log Number 9036078.

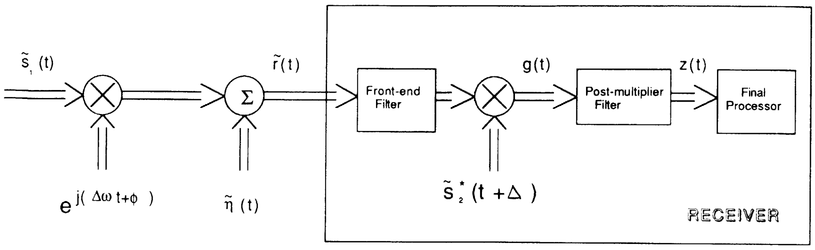


Fig. 1. Generic diagram of a communication/radar system.

jam (AJ) purposes, or combinations of these used in multiaccessing applications (plus, possibly, AJ). In the first case, we set $s_1(t) = s_2(t)$, whereas in multiple-access we can have $s_1(t) \neq s_2(t)$, signifying the cross-correlation between two different user-codes. Again, the specification of the PMF and the FP depends on the part of the spread-spectrum receiver we are addressing: despreading, code acquisition, code tracking, etc. If a random modeling of the inherent sequence(s) is adequate, we fall into the category of the previous paragraph. This is true, for instance, when the period of the code is very large and the PMF possesses a short impulse response, resulting in partial-period auto- or cross correlation. In many other designs this is not true, and the specific structure of the codes employed must be taken into account in order to avoid unpleasant phenomena such as, for instance, false-lock during acquisition due to partial correlation [4, sect. IX].

The above discussion was meant to point out the wide applicability of the model in Fig. 1, regardless of the modulation format (narrow-band versus spread) or the operating band (baseband, radio-frequencies, or optical). Of course, the physical generation and impact of impediments such as frequency, phase and timing offsets, additive and multiplicative noise, etc., will vary from one application to another. Notwithstanding, it is of considerable and continuing interest to find the spectral characteristics of either a linear noiseless signal $\tilde{s}_1(t)$ or of the noiseless product $g(t) = \tilde{s}_1(t)\,\tilde{s}_2^*(t + \Delta)$ because that, to a large extent, determines the choice of the front-end and post-multiplier filters. For spread spectrum in particular, when the PMF is not a perfect one-code-period integrator, the spectral portions of $g(t)$ passing through it, in addition to the desirable dc component, determine the so-called "*code self-noise*" level. In some operations this might degrade performance only mildly while in others, such as multiple-access, it is the most crucial performance factor.

The computation of the power spectral density (psd) of $g(t)$, which is the main subject of this paper, along with the frequency-domain description of the PMF allows one to separate the desirable and undesirable (e.g., signal-generated "noise") components of the process $z(t)$ in Fig. 1, working only with second-moment theory. From an analytical perspective, this is useful whenever the spectrum of $z(t)$ (or its correlation function with the associated sidelobes) suffices to characterize performance. This would be the case, for instance, if $z(t)$ can be approximately modeled as Gaussian due to, say, the narrow-band nature of the PMF plus a central-limit type of argument. Further linear or nonlinear functional operations on $z(t)$ can then be analyzed based on this (approximately) complete statistical description of $z(t)$. This would include phase-locking, quadratic transformations for energy measurements, etc. If, on the other hand, an application involves nonlinear operations directly on $g(t)$, and a more detailed joint statistical information between samples of $g(t)$ is required, then obviously the psd is by itself inadequate and can only serve, at best, as an intermediate step. Fortunately, most applications involve *some* filtering (even mild) on $g(t)$ before further processing, which increases the purposefulness of evaluating its psd.

Most of the literature regarding the spectral analysis of stochastic digital sequences and waveforms has concentrated on linear entities, like $\tilde{s}_1(t)$, and focuses on the impact that the detailed statistical dependence of modulator states (e.g., Markovian structure) has upon the final results (see the historical survey in [1, ch. 2], as well as [27], [4, ch. 2], [16], etc.). This dependence can be deliberate (partial-response signaling) or due to channel filtering (intersymbol interference) or the presence of a finite-state machine (e.g., convolutional codes, see [3]). As a byproduct, one gets the psd of purely random sources or more complicated PCM formats, for which the end result is affected by both the memory introduced between symbols, as well as the pulse-shape in use [13], [4], [1]. A simpler computational method can also be followed, if the autocorrelation function of the underlying symbol sequence is given or can be evaluated [4], [21]. However, this approach necessarily addresses a narrower class of problems (always of the linear-signal form) and, as expected, it yields the same answers as the previous Markovian modeling method when applied to simple PCM formats. Furthermore, a slight change in the definition of the autocorrelation function (from stochastic to deterministic) allows one to cover the deterministic, periodic-waveform case, such as the PN-code modulated waveform. Regardless of the particular method followed (depending on the problem complexity), the effect of the choice of pulse-shape (which can be of arbitrary length) is clearly manifested through its Fourier transform.

Contrary to the above, the literature for the psd of the product-waveform $g(t)$ is very sparse, presumably because the problem is harder. The two key available references are [7] and [6], which are repeated with minor changes in [4, sect. 8.12] and [29, App. G]. They compute the psd of the product between a BPSK/square-pulse (or nonreturn to zero, NRZ) modulated PN code and a displaced version of itself. The key idea there is an appropriate decomposition of the product-waveform, which we shall also adopt here. However, the computation of the individual spectra was strictly based on the biphase modulation/NRZ-pulse assumption. Furthermore, results were only provided for a displacement of less than a chip interval. This last restriction was relaxed in [8]; however, the modulation format and computational procedure were kept the same. This resulted in a restricted problem formulation, tedious computations, and long final expressions.

The goal of the present paper is to propose a systematic, straightforward and computationally tractable way for evaluating the psd of a wide variety of signals, especially of the product type where $\tilde{s}_i(t)$; $i = 1, 2$ can include arbitrary codes (that can be different), modulation formats (beyond BPSK) and pulse shapes (again, not necessarily the same or NRZ-specific). As a fringe benefit, the psd of linear waveforms $\tilde{s}_i(t)$ will be obtained, since it is a necessary intermediate step. We will illustrate the method by applying it to a new class of *on-off* optical spreading codes [23], to PN waveforms and their self-products of arbitrary pulse-shapes, and to cross-products of Gold codes for multiple access. We will also show how previous complicated expressions can be reconciled with the present, simpler ones. The key idea is to express the signals of interest as a convolution between the pulse-shape and an appropriately defined train of sequence-weighted Dirac-delta functions. The key advantage of this formulation is that it allows for a great degree of freedom in specifying the signal formats, while at the same time it detaches the procedure from the waveform specifics (of course, the final results depend on those). The price to be paid for this flexibility is that a) all memory in the symbol sequences must be expressible via their auto- and cross-correlation functions, thus excluding the richer stochastic details of Markovian formulations and b) the pulses must be confined within the symbol (or chip) interval, thus excluding any intersymbol interference (namely, any narrow channel or front-end filtering). We conjecture, however, that the theory can be extended in these directions also.

In order to proceed, we need to specify the class of complex envelopes under consideration. In general, $\tilde{s}_i(t)$; $i = 1, 2$ will have the standard form

$$\tilde{s}_i(t) = \sum_{n=-\infty}^{\infty} I_{i,n} p_i(t - nT_c) \tag{1}$$

where T_c is the symbol (or, depending on the case, baud or chip) interval, $p_i(t)$ is a pulse of duration T_c seconds, and $\{I_{i,n}\}$ is a sequence of symbols that can be stochastic (such as random data) or deterministic (such as a periodic spreading code). The simplest case arises when both $\{I_{i,n}\}$ and $p_i(t)$ are real as, for instance, in double sideband pulse amplitude modulation (DSB/PAM) where each $I_{i,n}$ takes on one of $M \geq 2$ real values (levels). The BPSK/NRZ case previously mentioned falls in this class, with $I_{i,n} \in [-1, 1]$ and $p_i(t)$ the standard NRZ pulse. An immediate generalization occurs when $I_{i,n} \triangleq I_{i,n}^{(c)} + jI_{i,n}^{(s)}$ where each of the real components $I_{i,n}^{(c)}$ and $I_{i,n}^{(s)}$ of the complex-valued $I_{i,n}$ takes on one of the $\sqrt{M}$ possible amplitudes or levels. This is the familiar QAM modulation, a special case of which is the QPSK modulation with $M = 4$. The latter can also be viewed as a special case of MPSK modulation where $I_{i,n} = e^{j\theta_n}$; $\theta_n \in [2\pi(m-1)/M;\ m = 1, 2, \cdots, M]$. A combination of PAM-PSK is also possible. Thus, by allowing a complex symbol sequence in the present model, we can cover a much wider class of interesting modulations such as, for example, quadriphase PN spreading with a different code per quadrature component.

Another kind of generalization is to keep $I_{i,n}$ real but allow $p_i(t)$ to be complex. Examples of interest in this case are the vestigial and single-sideband modulations, where an imaginary component is added to the real pulse in order to cancel part of the upper or lower sideband (for SSB, this component is the Hilbert transform of the original pulse). Yet another example in this category is when the transmitted pulse is real, but is filtered by an asymmetric BP channel or transmitter/receiver filter; in other words, by a filter whose low-pass equivalent contains an imaginary component [1]. In all of these cases, the received $\tilde{r}(t)$ will be complex.

In order to maintain the greatest degree of generality in our development, we will allow both $\{I_{i,n}\}$ and $p_i(t)$ to be simultaneously complex-valued because this covers the case where either one or both are real. The question arises as to whether the "simultaneously complex-valued" model represents any physical case or modulation of interest. The first obvious affirmative answer is whenever $\{I_{i,n}\}$ is complex by modulation design and the received $p_i(t)$ has become complex because of filtering asymmetries. Another, more subtle case regards modulation with nonidentical quadrature pulses. Let $p_i^{(c)}(t)$ and $p_i^{(s)}(t)$ represent the in-phase and quadrature pulse-shapes, respectively, so that $p_i(t) = p_i^{(c)}(t) + jp_i^{(s)}(t)$. Standard examples in this category are OQPSK and MSK modulations, for which $p_i^{(s)}(t) = p_i^{(c)}(t - T_c/2)$ with NRZ or sinusoidal pulse-shape, respectively. The appropriate complex envelope in this case is

$$\tilde{s}_i(t) = \sum_{n=-\infty}^{\infty} \left[I_{i,n}^{(c)} p_i^{(c)}(t - nT_c) + jI_{i,n}^{(s)} p_i^{(s)}(t - nT_c)\right]$$

which cannot be put into the compact form (1) with complex $I_{i,n}$ and $p_i(t)$. However, a little manipulation shows that $\tilde{s}_i(t)$ can be rewritten as

$$\tilde{s}_i(t) = \sum_{n=-\infty}^{\infty} e^{j(\pi/4)}\left[I_{i,n} p_i^*(t - nT_c) + I_{i,n}^* p_i(t - nT_c)\right] + e^{-j(\pi/4)}\left[I_{i,n} p_i(t - nT_c) + I_{i,n}^* p_i^*(t - nT_c)\right]$$

where the asterisk denotes conjugation. We conclude, then, that these cases can also be handled by our present method, at least in principle, except that the amount of bookkeeping will increase drastically. In what follows, we restrict attention to modulations obeying (1).

The paper is organized as follows. In Section II, the general expressions for the psd of modulated linear waveforms are derived, while in Section III the same is done for signal-product waveforms. We also derive an equivalent discrete-Fourier-transform (DFT) formulation which, from a practical standpoint, provides a convenient numerical method for evaluating the psd by computer, whenever long sequences are employed and analytical methods become too cumbersome. Applications of these general expressions can be found in Section IV. Some known expressions for psd's are shown to be special cases of the aforementioned general expressions. Finally, concluding remarks are given in Section V.

II. Linear Waveforms Modulated by Periodic Sequences

We now demonstrate and apply the general technique for deriving the psd of all signals of the form (1). The following definitions regarding correlation functions and the steps for their evaluation are, of course, well known (see, for instance, [13, sect. 4.3]). Their inclusion here serves to exhibit the methodology to be used in subsequent sections. Let us assume that $\{I_n\}$ represents a periodic sequence of symbols $\{c_n\}$ with period L, i.e., $I_j \triangleq c_j = c_{j+L}$ for every j. We can express a linear waveform $\tilde{s}(t)$ as the convolution (*)

$$\tilde{s}(t) = \sum_{n=-\infty}^{\infty} c_n \delta(t - nT_c) * p(t) = \tilde{s}_\delta(t) * p(t) \quad (2)$$

where $\tilde{s}_\delta(t)$ is defined as a $\{c_n\}$-weighted train of delta functions

$$\tilde{s}_\delta(t) \triangleq \sum_{n=-\infty}^{\infty} c_n \delta(t - nT_c). \quad (3)$$

Since the processes $\tilde{s}(t)$ and $\tilde{s}_\delta(t)$ are deterministic (periodic), their temporal autocorrelation functions $R_s(\tau)$ and $R_\delta(\tau)$, respectively, are defined via an integration over the period of LT_c seconds. Thus, $R_\delta(\tau)$ is given by[1]

$$\begin{aligned} R_\delta(\tau) &\triangleq \langle \tilde{s}_\delta(t), \tilde{s}_\delta^*(t-\tau)\rangle = \frac{1}{LT_c}\int_{0-}^{LT_c-} \sum_{n=-\infty}^{\infty}\sum_{n'=-\infty}^{\infty} \\ &\quad \cdot c_n c_{n'}^* \delta(t - nT_c)\, \delta(t - n'T_c - \tau)\, dt \\ &= \frac{1}{LT_c}\int_{0-}^{LT_c-} \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} c_n c_{n-m}^* \delta(t - nT_c) \\ &\quad \cdot \delta(t - (n-m)T_c - \tau)\, dt \\ &= \frac{1}{LT_c}\sum_{n=0}^{L-1}\sum_{m=-\infty}^{\infty} c_n c_{n-m}^* \delta(\tau - mT_c) \end{aligned}$$

so that

$$R_\delta(\tau) = \frac{1}{LT_c}\sum_{n=0}^{L-1}\sum_{m=-\infty}^{\infty} c_n c_{n-m}^* \delta(\tau - mT_c).$$

If we define the periodic autocorrelation sequence as

$$R_c(m) \triangleq \begin{bmatrix} \dfrac{1}{L}\displaystyle\sum_{n=0}^{L-1} c_n c_{n-m}^*; & m = 0, 1, \cdots, L-1 \\ R_c(m + nL); & n = \pm 1, \pm 2, \pm 3, \cdots \end{bmatrix} \quad (4)$$

we arrive at the (also periodic, with period LT_c) autocorrelation function of the sequence-modulated delta-train

$$R_\delta(\tau) = \frac{1}{T_c}\sum_{m=-\infty}^{\infty} R_c(m)\, \delta(\tau - mT_c). \quad (5)$$

A similar expression can be derived [21] for stochastic sequences, and the only difference will be in the definition of the autocorrelation function [17]. From (2) we formally obtain the psd of $\tilde{s}(t)$ as the Fourier transform (FT) of $R_s(\tau)$

$$\begin{aligned} S(f) &= F[R_s(\tau)] = F[R_\delta(\tau)]\,|P(f)|^2 \\ &= S_\delta(f)\,|P(f)|^2 \end{aligned} \quad (6)$$

where $S_\delta(f)$ is the spectral density of $\tilde{s}_\delta(t)$ in (3). We can identify the $S_\delta(f)$ component of (6) as the sequence-dependent spectral factor (or *sequence factor*), since it depends only on the correlation properties of the c_n sequence; on the other hand, $P(f)$, which depends only on the particular pulse-shape, will be designated as the pulse-shape-dependent spectral factor (or *pulse-shape factor*).

In the periodic-sequence case we can proceed further by utilizing the aforementioned periodicity of $R_\delta(\tau)$. Let us define the kth DFT coefficient of the periodic sequence $\{R_c(m)\}$ as

$$D_k = \begin{bmatrix} \displaystyle\sum_{m=0}^{L-1} R_c(m) e^{-j2\pi km/L}; & \\ & k = 0, 1, \cdots, L-1 \\ D_{k+nL}; & n = \pm 1, \pm 2, \pm 3, \cdots. \end{bmatrix} \quad (7)$$

Note that the Hermitian symmetry of $R_c(m)$, i.e., $R_c(-m) = R_c^*(m)$, implies that the DFT coefficients will be real numbers. Then, the sequence factor is

$$\begin{aligned} S_\delta(f) &= F[R_\delta(\tau)] \\ &= \frac{1}{LT_c}\sum_{k=-\infty}^{\infty}\left[\int_{0-}^{LT_c-} R_\delta(\tau) e^{-j2\pi f\tau}\, d\tau\right] \\ &\quad \cdot \delta\left(f - \frac{k}{LT_c}\right) \\ &= \frac{1}{LT_c}\sum_{k=-\infty}^{\infty}\left[\int_{0-}^{LT_c-} \frac{1}{T_c}\sum_{m=-\infty}^{\infty} R_c(m) \right. \\ &\quad \left. \cdot\, \delta(\tau - mT_c) e^{-j2\pi f\tau}\, d\tau\right] \delta\left(f - \frac{k}{LT_c}\right) \end{aligned}$$

[1] The limits of integration are from 0- to LT_c—in order to include all delta functions situated at nT_c; $n = 0, 1, 2, \cdots$.

$$= \frac{1}{LT_c^2}\sum_{k=-\infty}^{\infty}\left[\sum_{m=0}^{L-1} R_c(m)e^{-j2\pi f m T_c}\right]\delta\left(f-\frac{k}{LT_c}\right)$$

$$= \frac{1}{LT_c^2}\sum_{k=-\infty}^{\infty}\left[\sum_{m=0}^{L-1} R_c(m)e^{-j2\pi km/L}\right]\delta\left(f-\frac{k}{LT_c}\right)$$

$$= \frac{1}{LT_c^2}\sum_{k=-\infty}^{\infty} D_k\delta\left(f-\frac{k}{LT_c}\right). \qquad (8)$$

Finally, combining the pulse-shape factor from (6),

$$S(f) = L\sum_{k=-\infty}^{\infty} D_k\left|\frac{1}{LT_c}P\left(\frac{k}{LT_c}\right)\right|^2\delta\left(f-\frac{k}{LT_c}\right) \qquad (9)$$

which will be used in the applications section.

We now consider the more complicated scenario of waveforms created by product of other signals (waveforms).

III. Signal-Product Waveforms

Let the signal-product waveform be defined as

$$g(t) = \tilde{s}_1(t)\,\tilde{s}_2^*(t+\Delta) = \tilde{s}_1(t)\,\tilde{s}_2^*(t+kT_c+\rho T_c) \qquad (10a)$$

where $\Delta = (k+\rho)T_c$ with k a positive integer and $0 \leq \rho \leq 1$. The waveforms $\tilde{s}_i(t)$; $i = 1, 2$ are of the form described in (1). We can generalize the standard decomposition of [7] to rewrite $g(t)$ as

$$g(t) = \left[\sum_{n=-\infty}^{\infty} I_{1,n}p_1(t-nT_c)\right] \cdot \left[\sum_{n=-\infty}^{\infty} I_{2,n+k}^* p_2^*(t-(n+k)T_c+\rho T_c)\right]$$

$$= \sum_{n=-\infty}^{\infty} I_{1,n}I_{2,n+k}^* q(t-nT_c) + \sum_{n=-\infty}^{\infty} I_{1,n}I_{2,n+k+1}^* h(t-nT_c)$$

or

$$g(t) = \sum_{n=-\infty}^{\infty} a_n q(t-nT_c) + \sum_{n=-\infty}^{\infty} b_n h(t-nT_c)$$

$$= g_1(t) + g_2(t) \qquad (10b)$$

where we have defined

$$a_n = I_{1,n}I_{2,n+k}^*; \quad b_n = I_{1,n}I_{2,n+k+1}^* \qquad (11a)$$

$$g_1(t) = \sum_{n=-\infty}^{\infty} a_n q(t-nT_c) = a_\delta(t) * q(t);$$

$$g_2(t) = \sum_{n=-\infty}^{\infty} b_n h(t-nT_c) = b_\delta(t) * h(t) \qquad (11b)$$

$$a_\delta(t) = \sum_{n=-\infty}^{\infty} a_n\delta(t-nT_c);$$

$$b_\delta(t) = \sum_{n=-\infty}^{\infty} b_n\delta(t-nT_c) \qquad (11c)$$

$$q(t) = \begin{cases} p_1(t)\,p_2^*(t+\rho T_c); & 0 \leq t \leq (1-\rho)T_c \\ 0; & (1-\rho)T_c \leq t \leq T_c \end{cases} \qquad (12a)$$

$$h(t) = \begin{cases} 0; & 0 \leq t \leq (1-\rho)T_c \\ p_1(t)\,p_2^*(t-(1-\rho)T_c); & (1-\rho)T_c \leq t \leq T_c. \end{cases} \qquad (12b)$$

An example explaining the decomposition via the pulses of (12) appears in Fig. 2.

It can be observed that a_n and b_n depend on $I_{i,n}$ and k, while $q(t)$ and $h(t)$ depend on $p_i(t)$ and ρ.

We can express the psd of $g(t)$ as

$$S_g(f) = S_{g_1}(f) + S_{g_2}(f) + S_{g_1g_2}(f) + S_{g_2g_1}(f)$$

where $S_{g_1}(f) = F[R_{g_1}(\tau)]$, $S_{g_2}(f) = F[R_{g_2}(\tau)]$ and $S_{g_1g_2}(f) = F[R_{g_1g_2}(\tau)]$ is the cross-spectral density of $g_1(t)$ and $g_2(t)$. Since $S_{g_1g_2}(f) = S_{g_2g_1}^*(f)$ (see [21, Section 1.2.2]), $S_g(f)$ can be written as

$$S_g(f) = S_{g_1}(f) + S_{g_2}(f) + 2\,\mathrm{Re}\left[S_{g_1g_2}(f)\right]. \qquad (13)$$

The individual terms $S_{g_i}(f)$; $i = 1, 2$ can be derived using the methods of Section II. For the derivation of the cross-spectral density $S_{g_1g_2}(f)$ we can again use the technique of expressing $g_i(t)$; $i = 1, 2$ as appropriate convolutions of modulated delta-trains with the $q(t)$ and $h(t)$ pulses, respectively [see (11b), (11c)]. It follows after some manipulations that

$$S_{g_1g_2}(f) = F[R_{ab}^{(\delta)}(\tau)]\,Q(f)\,H^*(f)$$

$$= S_{ab}^{(\delta)}(f)\,Q(f)\,H^*(f) \qquad (14)$$

where $Q(f) = F[q(t)]$, $H(f) = F[h(t)]$, and $R_{ab}^{(\delta)}(\tau)$ is the cross correlation[2] of the delta-trains $a_\delta(t)$ and $b_\delta(t)$, modulated by the sequences $\{a_n\}$, $\{b_n\}$ of (11a).

In conclusion, (13), (14) provide the desired answer with

$$S_{g_1}(f) = \frac{|Q(f)|^2}{T_c}\sum_{m=-\infty}^{\infty} R_a(m)e^{-j2\pi f m T_c} \qquad (15a)$$

$$S_{g_2}(f) = \frac{|H(f)|^2}{T_c}\sum_{m=-\infty}^{\infty} R_b(m)e^{-j2\pi f m T_c} \qquad (15b)$$

[2] Cross correlation can be defined either in the temporal or statistical sense $\langle a_\delta(t), b_\delta^*(t-\tau)\rangle$, $E[a_\delta(t)\,b_\delta^*(t-\tau)]$).

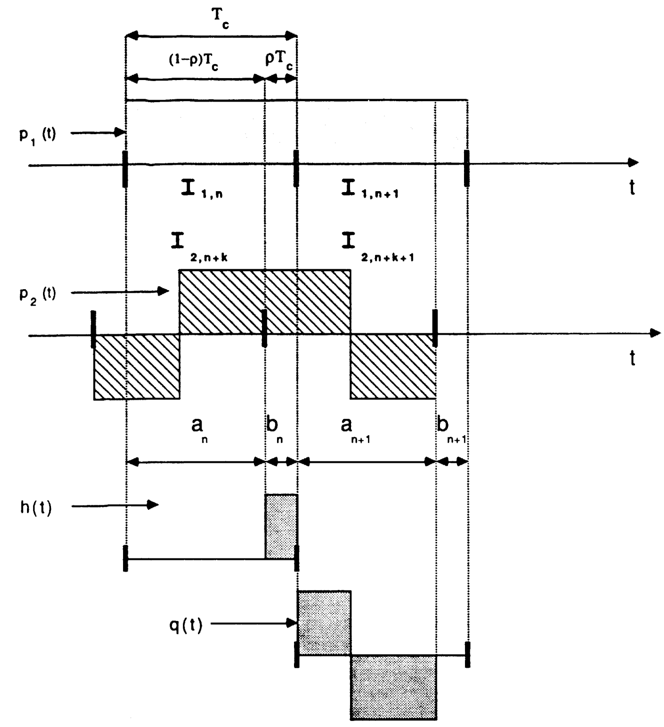


Fig. 2. Decomposition of signal-product waveform.

$$S_{g_1g_2}(f) = \frac{Q(f)\,H^*(f)}{T_c} \sum_{m=-\infty}^{\infty} R_{ab}(m)e^{-j2\pi fmT_c}. \qquad (15c)$$

Here, $R_{ab}(m)$ is either $R_{ab}(m) = E[a_n b^*_{n-m}]$ or $R_{ab}(m) = 1/L\,\Sigma_{n=0}^{L-1} a_n b^*_{n-m}$, assuming that the product $a_n b^*_{n-m}$ has period L. For the latter case, the DFT formulation can be used, as per the previous section. We note again that the sequences and their products have been separated from the actual shape of the modulating pulse. With regard to the impact of the parameters ρ and k, we note that ρ will affect $Q(f)$ and $H(f)$, while k will affect the various correlation sequences. Not much more can be said about (15), which is rather formalistic in appearance, without specifying further the quantities of interest. We undertake that below, in the applications section.

IV. Applications

Some specific examples of the general cases are now presented. In particular, we examine the PN sequence (Section IV-A), the orthogonal optical code (OOC) (Section IV-B), the self-product of PN waveforms (Section IV-C) and the cross product of Gold sequences (Section IV-D).

A. PN Waveforms of Arbitrary Pulse-Shape

For the specific case of a periodic (period L) pseudorandom code taking ± 1 values (antipodal signaling), we can derive the psd using (9). From the well-known fact that [11]

$$R_c(m) = \begin{bmatrix} 1 & m = kL \\ -\dfrac{1}{L} & m \neq kL \end{bmatrix}$$

we conclude that $D_k = \Sigma_{m=0}^{L-1} R_c(m)e^{-j2\pi km/L} = (1 + 1/L) - 1/L\,\Sigma_{m=0}^{L-1} e^{-j2\pi km/L}$. Therefore,

$$D_k = \begin{bmatrix} \dfrac{1}{L}; & k = 0 \\ \dfrac{L+1}{L}; & k = 1, 2, \cdots, L-1. \end{bmatrix} \qquad (16)$$

Thus, (9) gives

$$S(f) = \frac{1}{LT_c^2}\left\{\left(\frac{L+1}{L}\right)\sum_{k=-\infty}^{\infty}\left|P\left(\frac{k}{LT_c}\right)\right|^2 \delta\left(f - \frac{k}{LT_c}\right) - \sum_{m=-\infty}^{\infty}\left|P\left(\frac{m}{T_c}\right)\right|^2 \delta\left(f - \frac{m}{T_c}\right)\right\}. \qquad (17)$$

The pulse-shape $p(t)$ and its transform $P(f)$ in (17) can be arbitrary. In fact, for this development, $p(t)$ need not be confined to T_c seconds. Previous derivations have relied on the autocorrelation function of the modulated waveform, which is necessarily pulse-shape dependent.

As specific examples, we have that for a NRZ pulse of duration T_c,[3] with $P(f) = T_c$ sinc (fT_c), (17) yields

$$S(f) = \frac{1}{L^2}\delta(f) + \frac{L+1}{L^2}\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \text{sinc}^2\left(\frac{\text{k}}{\text{L}}\right)\cdot\delta\left(\text{f} - \frac{\text{k}}{\text{LT}_c}\right) \qquad (18)$$

(NRZ)

an expression[4] that also can be found in many other references (see for instance [4], [29]). Notice that the harmonics at m/T_c disappear for this case. As another example, consider the Manchester pulse of duration T_c with $|P(f)| = \pi f T_c^2/2$ sinc^2 $(f(T_c/2))$, for which we get

$$S(f) = \frac{L+1}{L^2}\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty}\left[\frac{k\pi}{2L}\right]^2 \text{sinc}^4\,(\text{k}/2L)\,\delta\left(f - \frac{k}{LT_c}\right) - \left[\frac{2}{\pi}\right]^2 \frac{1}{L}\sum_{\substack{m=-\infty \\ m\neq 0 \\ m=\text{odd}}}^{\infty}\frac{1}{m^2}\delta\left(f - \frac{m}{T_c}\right) \qquad (19)$$

(Manchester)

[3] We use the notation sinc $(x) = \sin(\pi x)/(\pi x)$.

[4] Equation (18) is commonly referred to as the "power spectrum of a PN sequence," as opposed to the more precise term "power spectrum of a specific PN waveform," corresponding to the particular choice of pulse-shape.

which includes the odd m/T_c harmonics. As a third example, we also determine the psd for the raised-cosine pulse. Specifically, for $p(t)$ given by [21]

$$p(t) = \frac{1}{2}\left[1 + \cos\left[\frac{2\pi}{T_c}\left(t - \frac{T_c}{2}\right)\right]\right]; \quad 0 \le t \le T_c \tag{20a}$$

which is a time-limited pulse, the FT is

$$P(f) = \frac{T_c}{2}\frac{\sin(\pi f T_c)}{\pi f T_c(1 - f^2T_c^2)}e^{-j\pi f T_c}. \tag{20b}$$

Then, from (17) we arrive at

$$S(f) = \frac{1}{4L^2}\delta(f) + \frac{1}{16L^2}\left(\delta\left(f + \frac{1}{T_c}\right) + \delta\left(f - \frac{1}{T_c}\right)\right) + \frac{L+1}{4L^2}\sum_{\substack{k=-\infty \\ k \ne 0, \pm L}}^{\infty}\left[\frac{\sin\left(\pi\frac{k}{L}\right)}{\pi\frac{k}{L}\left(1 - (k/L)^2\right)}\right]^2 \cdot \delta\left(f - \frac{k}{LT_c}\right)$$

$$\text{(time-limited raised-cosine).} \tag{21}$$

As a final example of a pulse extending beyond T_c, consider the raised-cosine family of pulses which satisfy the Nyquist criterion for zero intersymbol-interference. Here $p(t)$ is defined as [29]

$$p(t) = \frac{\cos(2\pi\beta t)}{1 - (4\beta t)^2}\,\mathrm{sinc}\,(\mathrm{t}/\mathrm{T}_c) \tag{22a}$$

and the corresponding FT is

$$P(f) = \begin{cases} T_c; & |f| \le \dfrac{1}{2T_c} - \beta \\[2ex] \dfrac{T_c}{2}\left[1 + \cos\left(\dfrac{\pi\left(|f| - \dfrac{1}{2T_c} + \beta\right)}{2\beta}\right)\right]; & \dfrac{1}{2T_c} - \beta < |f| \le \dfrac{1}{2T_c} + \beta \\[2ex] 0; & |f| > \dfrac{1}{2T_c} + \beta. \end{cases} \tag{22b}$$

Again, an expression similar to (21) can be derived, but we did not include it here for save of space. Instead, we chose to plot a particular example. The psd's for the PN waveforms and the aforementioned pulse-shapes are shown in Fig. 3(a)-(d) for codes of period $L = 31$. Although we are actually dealing with line spectra, we illustrate them with continuous envelopes for pictorial ease. Note that the pulses in Fig. 3(b), 3(c) contain distinct nulls at the harmonic equal to the code period, while the others do not.

B. Optical Orthogonal Sequences

The previous examples (where the binary ,equences take ±1 values) are widely used in spread-spectrum systems, whenever coherent processing is possible, with direct-sequence being the standard case. When frequency hopping is used as the spreading technique in conventional systems [25], or when optical code-division multiple-access (CDMA) is considered [22], [23], then the resultant processing consists of a summation of individual power-terms in appropriate slots (time or frequency). Sequences having good auto- or cross-correlation properties in an antipodal modulation may not retain their properties in an orthogonal (i.e., 0, 1) format [22]; thus, new designs are needed for these systems, such as the Costas arrays [12] or orthogonal sequences for optical CDMA. Focusing on the latter, which have been called *optical orthogonal codes* [23], we note that they can be described by a quadruple $(F, K, \lambda_a, \lambda_c)$ where F is the length, K is the weight (i.e., number of 1's), λ_a is the upper bound for the autocorrelation function (excluding the $R_c(0)$ value) and λ_c is the upper bound for the cross-correlation function. As of yet, neither these sequences nor their autocorrelation functions have been described in specific analytic form, as is the case for PN sequences. Subsequently, general closed-form expressions for the resultant psd cannot be derived at this point. Nonetheless, the proposed technique can be applied for any *given* sequence, since its autocorrelation sequence $\{R_c(m)\}$ can be computed. As an illustrative example of the method we refer to Fig. 4 where we show the psd of the NRZ waveform corresponding to a specific sequence.

C. PN Waveform Self-Product

As we mentioned before, this is the standard model which arises in code despreading, synchronization, etc. When the signal under consideration is a PN waveform multiplied by a shifted version of itself [i.e., $I_{1,n} = I_{2,n}$ and $p_1(t) = p_2(t)$ in (10a)], then (11a) results in $a_n = c_nc_{n+k} = c_{n+m}$, $b_n = c_nc_{n+k+1} = c_{n+l}$, $l \ne m$. This is because of the "cycle-and-add" property of PN sequences [11], i.e., the modulo-two sum of a (0, 1)-valued binary PN sequence with a displaced (shifted) replica is equal to the same sequence, dispaced by a different shift. We now examine the *cross-sequence factor* for two separate cases, namely,

1) ($k = 0$): If the relative shift between $s(t)$ and $s(t + \Delta)$ is only a fraction of a chip ρT_c, then, $a_n = c_n^2 = 1$ and $b_n = c_{n+j}$. Therefore, $g_1(t)$ of (11b) is now a purely periodic (i.e., not PN-modulated) waveform. We will indicate this periodic waveform (period T_c) as $g_p(t)$, whose fundamental pulse-shape is given by (12a); the remaining modulated waveform (by the same PN sequence) will be denoted by $g_c(t)$, whose pulse-shape is described by (12b)

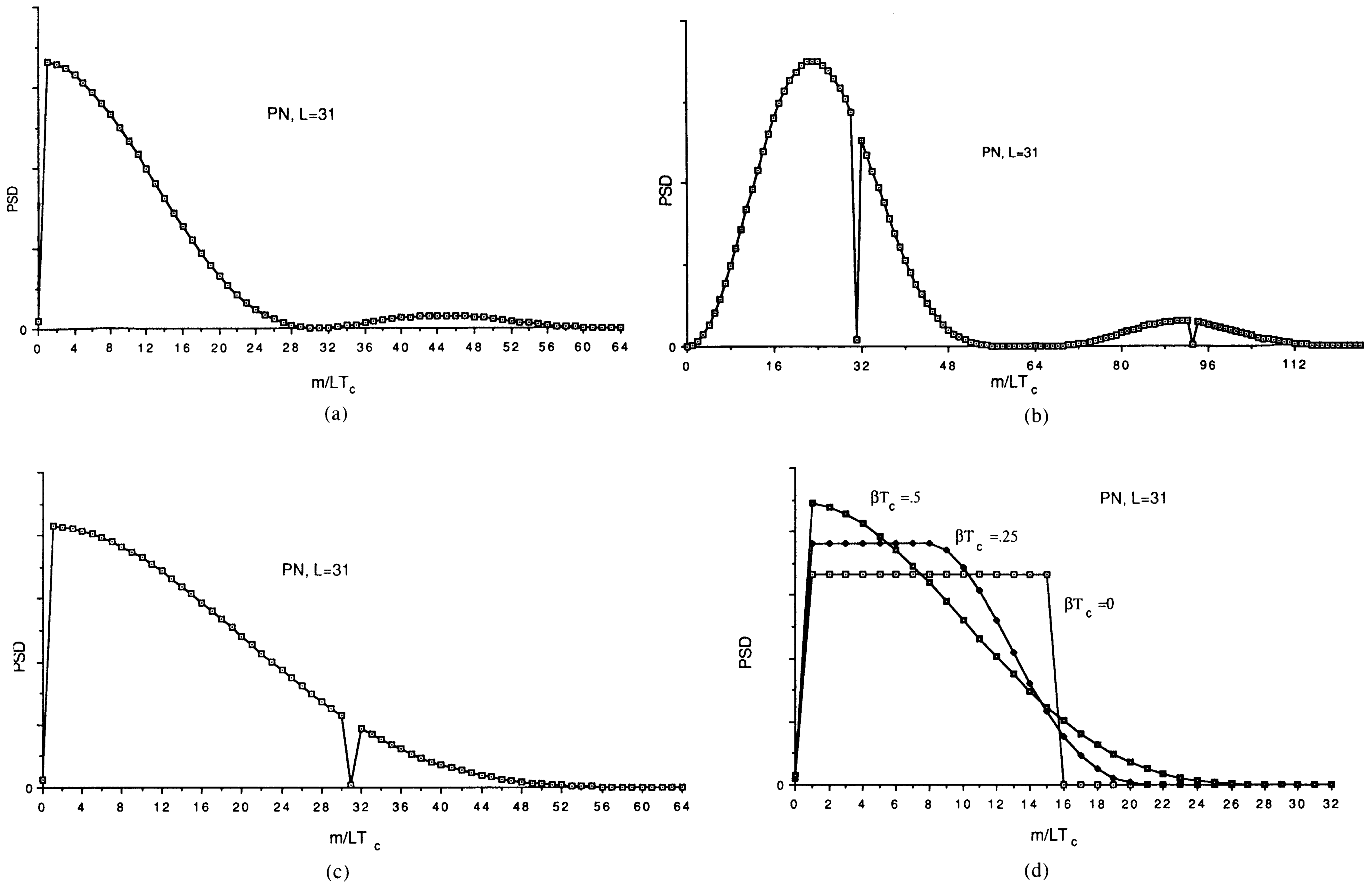


Fig. 3. (a) psd of a PN waveform with NRZ signaling. (b) psd of a PN waveform with Manchester signaling. (c) psd of a PN waveform with time-limited raised-cosine pulse-shape. (d) psd of a PN waveform with band-limited raised-cosine pulse-shape.

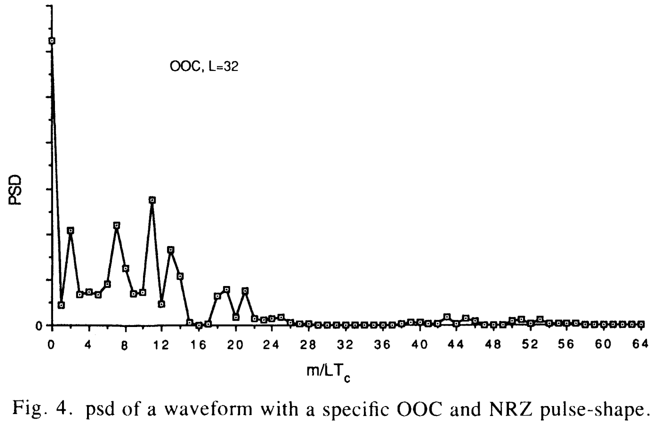


Fig. 4. psd of a waveform with a specific OOC and NRZ pulse-shape.

and whose period is LT_c. Thus, we can write

$$g(t) = g_p(t) + g_c(t) = a_\delta(t) * q(t) + b_\delta(t) * h(t) \tag{23a}$$

with

$$a_\delta(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_c),$$

$$b_\delta(t) = \sum_{n=-\infty}^{\infty} c_{n+j}\delta(t - nT_c). \tag{23b}$$

This decomposition was first proposed in [7], but it was only analyzed for the special case of a rectangular pulse-shape.

Due to the individual periodicity of $a_\delta(t)$, $b_\delta(t)$ in (11c), their product $a_\delta(t)\, b_\delta(t)$ is also periodic with period LT_c; thus,

$$R_{ab}^{(\delta)}(\tau) = \frac{1}{LT_c}\int_{0-}^{LT_c-} \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} c_{n+j-m}\delta(t - nT_c) \cdot \delta(t - (n-m)T_c - \tau)\, dt$$

$$= \frac{1}{LT_c}\sum_{n=0}^{L-1}\sum_{m=-\infty}^{\infty} c_{n+j-m}\delta(\tau - mT_c).$$

Since $c_n = c_{n+L}$, $R_{ab}^{(\delta)}(\tau)$ reduces to $R_{ab}^{(\delta)}(\tau) = 1/LT_c \sum_{n=0}^{L-1} c_{n+j} \sum_{m=-\infty}^{\infty} \delta(\tau - mT_c)$. Furthermore, because $\sum_{n=0}^{L-1} c_{n+j} = -1$, we arrive at $R_{ab}^{(\delta)}(\tau) = -(1/LT_c) \sum_{m=-\infty}^{\infty} \delta(\tau - mT_c)$. Finally, the *cross-sequence factor* is

$$S_{ab}^{(\delta)}(f) = -\frac{1}{LT_c^2}\sum_{m=-\infty}^{\infty} \delta\left(f - \frac{m}{T_c}\right). \tag{24}$$

We now turn our attention to the next case.

2) ($k > 0$): Here, $a_n = c_{n+j}$, $b_n = c_{n+j-l}$; in other words, the sequences of (11a) are shifted versions of the original c_n. Furthermore, the integer l represents *the number of shift positions between a_n and b_n*. Therefore, the corresponding expressions in (11c) are

$$a_\delta(t) = \sum_{n=-\infty}^{\infty} c_{n+j}\delta(t - nT_c),$$

$$b_\delta(t) = \sum_{n=-\infty}^{\infty} c_{n+j-l}\delta(t - nT_c). \tag{25}$$

For $R_{ab}^{(\delta)}(\tau)$, we have

$$R_{ab}^{(\delta)}(\tau) = \frac{1}{LT_c}\int_{0-}^{LT_c-}\left[\sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} a_n\delta(t - nT_c)\right.$$

$$\left.\cdot\, b_{n-m}\delta(t - (n - m)T_c - \tau)\right]dt$$

$$= \frac{1}{T_c}\sum_{m=-\infty}^{\infty}\left[\frac{1}{L}\sum_{n=0}^{L-1} c_{n+j}c_{n+j-m-l}\right]\delta(\tau - mT_c)$$

or

$$R_{ab}^{(\delta)}(\tau) = \frac{1}{T_c}\sum_{m=-\infty}^{\infty} R_c(m + l)\,\delta(\tau - mT_c) \tag{26}$$

which is a shifted (by lT_c seconds to the left) version of (5). Thus, we can either employ the DFT coefficients of (7), or the shift-property in conjunction with (16) to arrive at

$$D_k(l) = \begin{bmatrix} \dfrac{1}{L}; & k = 0 \\ \dfrac{L+1}{L} e^{j2\pi lk/L}; & k = 1, 2, \cdots, L - 1 \end{bmatrix} \tag{27}$$

where $D_k(l) = D_k(l + mL)$. Note that for $l = 0$, $D_k(0) = D_k$ of (16). Thus,

$$S_{ab}^{(\delta)}(f) = \frac{1}{LT_c^2}\left[\frac{L+1}{L}\sum_{k=-\infty}^{\infty} e^{j2\pi kl/L}\delta\left(f - \frac{k}{LT_c}\right)\right.$$

$$\left. - \sum_{m=-\infty}^{\infty}\delta\left(f - \frac{m}{T_c}\right)\right]. \tag{28}$$

Again, this determines the cross-sequence factor independently of the particular pulse-shape in use. As specific examples, consider the following.

Example 1) NRZ Pulse-Shape: The two cases ($k = 0$ and $k \geq 1$) result in the following.

1a) ($k = 0$): Using (24) and (14) for the specific pulse-shape factor, it follows that

$$S_{g_p g_c}(f) = -\frac{1}{LT_c^2}\sum_{m=-\infty}^{\infty}\left[\rho T_c \,\text{sinc}\,(f\rho T_c)e^{-j\rho\pi f T_c}\right]$$

$$\cdot\left[(1 - \rho)T_c \,\text{sinc}\,(f(1 - \rho)T_c)e^{j\pi f(1+\rho)Tc}\right]$$

$$\cdot\,\delta\left(f - \frac{m}{T_c}\right). \tag{29a}$$

Using (A.2), which is proven in the Appendix, (29a) reduces to

$$S_{g_p g_c}(f) = -\frac{\rho(1 - \rho)}{L}\delta(f)$$

$$+ \frac{\rho^2}{L}\sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \cdot\,\text{sinc}^2\,(m\rho)\,\delta\left(f - \frac{m}{T_c}\right). \tag{29b}$$

Finally, we can evaluate $S_g(f)$ using (13), (17) as

$$S_g(f) = \left[1 - \rho\frac{L+1}{L}\right]^2\delta(f)$$

$$+ \rho^2\frac{(L+1)}{L^2}\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \cdot\,\text{sinc}^2\left(\frac{k\rho}{L}\right)\delta\left(f - \frac{k}{LT_c}\right)$$

$$+ \rho^2\frac{L+1}{L}\sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \text{sinc}^2\,(m\rho)\,\delta\left(f - \frac{m}{T_c}\right) \tag{30}$$

which can also be found in [29], [4]. Note that for long codes (i.e., $L >> 1$),

$$S_g(f) \approx (1 - \rho)^2\delta(f)$$

$$+ \rho^2\sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \cdot\,\text{sinc}^2\,(m\rho)\,\delta\left(f - \frac{m}{T_c}\right).$$

1b) ($k > 0$): Substituting the expression of (28) in (14) we have

$$S_{g_1 g_2}(f) = \left[\frac{L+1}{L}\right]\left[\frac{1}{LT_c^2}\right]\sum_{k=-\infty}^{\infty} e^{j2\pi kl/L}\rho(1 - \rho)T_c^2$$

$$\cdot\,\text{sinc}\left(\frac{k\rho}{L}\right)\text{sinc}\left(\frac{k(1 - \rho)}{L}\right)e^{j\pi(k/L)}$$

$$\cdot\,\delta\left(f - \frac{k}{LT_c}\right) - \frac{1}{LT_c^2}\sum_{m=-\infty}^{\infty}\rho(1 - \rho)T_c^2$$

$$\cdot\,\text{sinc}\,(m\rho)\,\text{sinc}\,(m(1 - \rho))\,e^{j\pi m}\delta\left(f - \frac{m}{T_c}\right)$$

or

$$S_{g_1 g_2}(f) = \frac{\rho(1 - \rho)}{L^2}\delta(f) + \frac{L+1}{L^2}\rho(1 - \rho)$$

$$\cdot\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} e^{j\pi k[(2l+1)/L]}\,\text{sinc}\left(\frac{k\rho}{L}\right)$$

$$\cdot\,\text{sinc}\left(\frac{k}{L}(1 - \rho)\right)\cdot\delta\left(f - \frac{k}{LT_c}\right)$$

$$+ \frac{\rho^2}{L}\sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty}\text{sinc}^2\,(m\rho)\cdot\delta\left(f - \frac{m}{T_c}\right). \tag{31}$$

If we make use of (13), (17), and (A.2) of the Appendix, we arrive at

$$\begin{aligned} S_g(f) &= \frac{1}{L^2}\,\delta(f) + \frac{L+1}{L^2} \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \left[\rho^2 \operatorname{sinc}^2\left(\frac{k\rho}{L}\right) \right. \\ &\quad + (1-\rho)^2 \operatorname{sinc}^2\left(\frac{k}{L}(1-\rho)\right) \\ &\quad + 2\rho(1-\rho)\cos\left(k\pi\,\frac{(2l+1)}{L}\right) \operatorname{sinc}\left(\frac{k\rho}{L}\right) \\ &\quad \left. \cdot \operatorname{sinc}\left(\frac{k(1-\rho)}{L}\right)\right] \delta\left(f - \frac{k}{LT_c}\right). \qquad (32) \end{aligned}$$

For $\rho = 0$ or 1 (or, equivalently, any integer shift $\Delta = kT_c$) (32) reduces to (18), the psd of the modulated PN waveform, as expected. We note that an analogous expression was derived in [8] with the use of the autocorrelation function of the modulated waveform. It can be shown (see Appendix, part A.3) that the two expressions are equivalent. However, let us note that expression (32) is much more compact than the corresponding one in [8] and was derived in a more general manner. If we ignore the $1/L^2$ terms, then the dominant terms of (32) yield

$$\begin{aligned} S_g(f) &= \frac{1}{L} \sum_{\substack{k=-\infty \\ k \neq 0 \\ k \neq nL}}^{\infty} \left[\rho^2 \operatorname{sinc}^2\left(\frac{k\rho}{L}\right) + (1-\rho)^2 \right. \\ &\quad \cdot \operatorname{sinc}^2\left(\frac{k}{L}(1-\rho)\right) + 2\rho(1-\rho) \\ &\quad \cdot \cos\left(k\pi\,\frac{(2l+1)}{L}\right) \operatorname{sinc}\left(\frac{k\rho}{L}\right) \\ &\quad \left. \cdot \operatorname{sinc}\left(\frac{k(1-\rho)}{L}\right)\right] \delta\left(f - \frac{k}{LT_c}\right). \end{aligned}$$

Example 2) Manchester Signaling: We have to distinguish between the following three cases with respect to the values of ρ:

1) $\rho = 0, 1$. For this case, refer back to (30).

2) $\rho = 0.5$. The new modulating waveforms $h(t)$, $q(t)$, are as in the NRZ pulse-case for $\rho = 0.5$, but with reversed polarity; thus we can refer back to (30), (32) for either $k = 0$ or $k > 0$ of the previous example.

3) We observe that whenever $0 \leq \rho \leq 0.5$ the roles of $h(t)$ and $q(t)$ are the reverse of those in $0.5 \leq \rho \leq 1$. We will concentrate on the $0 \leq \rho \leq 0.5$ case (see Fig. 5). Again, we should distinguish between $k = 0$ and $k \geq 1$.

2a) (k = 0): Now, using the cross-sequence factor of (24) and the appropriate pulse-shape factor as per (14),

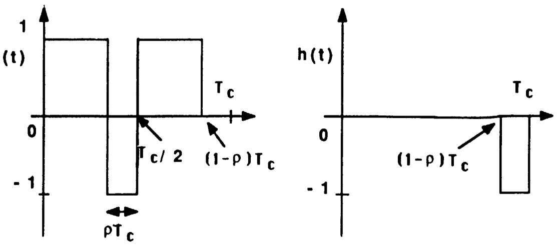


Fig. 5. Pulse-shape decomposition for Manchester signaling.

the cross-spectral density is

$$\begin{aligned} S_{g_p g_c}(f) &= \frac{\rho}{L}(1-3\rho)\delta(f) \\ &\quad + \frac{1}{L} \sum_{\substack{k=-\infty \\ k \neq 0 \\ k = \text{odd}}}^{\infty} \rho^2 \operatorname{sinc}^2(k\rho)\,\delta\left(f - \frac{k}{T_c}\right) \\ &\quad - \frac{3}{L} \sum_{\substack{k=-\infty \\ k \neq 0 \\ k = \text{even}}}^{\infty} \rho^2 \operatorname{sinc}^2(k\rho)\,\delta\left(f - \frac{k}{T_c}\right). \end{aligned} \qquad (33)$$

If we use (9) for the psd's of $g_i(t)$; $i = 1, 2$, with the appropriate pulse-shapes, and after a series of calculations required for combining these results along with (33) [as per (13)], we arrive at

$$\begin{aligned} S_g(f) &= \left(1 - 3\rho + \frac{\rho}{L}\right)^2 \delta(f) \\ &\quad + \rho^2\left(9 - \frac{7}{L}\right) \sum_{\substack{m=-\infty \\ m \neq 0 \\ m = \text{even}}}^{\infty} \operatorname{sinc}^2(m\rho)\,\delta\left(f - \frac{m}{T_c}\right) \\ &\quad + \frac{L+1}{L}\rho^2 \sum_{\substack{m=-\infty \\ m \neq 0 \\ m = \text{odd}}}^{\infty} \operatorname{sinc}^2(m\rho)\,\delta\left(f - \frac{m}{T_c}\right) \\ &\quad + \frac{L+1}{L}\rho^2 \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \operatorname{sinc}^2\left(\frac{k\rho}{L}\right)\delta\left(f - \frac{k}{LT_c}\right). \end{aligned} \qquad (34)$$

For long codes ($L >> 1$)

$$\begin{aligned} S_g(f) &\approx (1-3\rho)^2\delta(f) \\ &\quad + 9\rho^2 \sum_{\substack{m=-\infty \\ m \neq 0 \\ m = \text{even}}}^{\infty} \operatorname{sinc}^2(m\rho)\,\delta\left(f - \frac{m}{T_c}\right) \\ &\quad + \rho^2 \sum_{\substack{m=-\infty \\ m = \text{odd}}}^{\infty} \operatorname{sinc}^2(m\rho)\,\delta\left(f - \frac{m}{T_c}\right) \\ &\quad + \rho^2 \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \operatorname{sinc}^2\left(\frac{k\rho}{L}\right)\delta\left(f - \frac{k}{LT_c}\right). \end{aligned}$$

Note that, for $\rho = 1/3$ the dc term disappears, something which does not occur in the NRZ case.

2b) ($k > 0$): Again, the combination of (14), (28) and a good amount of algebraic manipulations yields

$$\begin{aligned} S_{g_1 g_2}(f) = {} & \frac{(3\rho - 1)\rho}{L^2}\,\delta(f) \\ & + \frac{L+1}{L^2}\,\rho \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \left[2\rho\,\mathrm{sinc}^2\left(\frac{k\rho}{L}\right) \right. \\ & - (1-\rho)\,\mathrm{sinc}\left(\frac{k\rho}{L}(1-\rho)\right) \\ & \left. \cdot\,\mathrm{sinc}\left(\frac{k\rho}{L}\right)\right] e^{j(k\pi/L)(2l+1)}\delta\left(f - \frac{k}{LT_c}\right) \\ & + \frac{\rho^2}{L} \sum_{\substack{m=-\infty \\ m \neq 0 \\ m = \text{odd}}}^{\infty} \mathrm{sinc}^2(m\rho)\,\delta\left(f - \frac{m}{T_c}\right) \\ & - \frac{3\rho^2}{L} \sum_{\substack{m=-\infty \\ m \neq 0 \\ m = \text{even}}}^{\infty} \mathrm{sinc}^2(m\rho)\,\delta\left(f - \frac{m}{T_c}\right). \end{aligned} \tag{35}$$

The final result combines (13), (17) with (35):

$$\begin{aligned} S_g(f) = {} & \left(\frac{1 - 4\rho}{L}\right)^2 \delta(f) \\ & - \frac{(4\rho)^2}{L} \sum_{\substack{m=-\infty \\ m \neq 0 \\ \mathrm{m = even}}}^{\infty} \mathrm{sinc}^2(m\rho)\,\delta\left(f - \frac{m}{T_c}\right) \\ & + \frac{L+1}{L^2} \sum_{\substack{m=-\infty \\ m \neq 0}}^{\infty} \left[\left[(1-\rho)\,\mathrm{sinc}\left(\frac{k}{L}(1-\rho)\right)\right.\right. \\ & \left. - 2\rho\,\mathrm{sinc}\left(k\frac{\rho}{L}\right)\right]^2 + \rho^2\,\mathrm{sinc}^2\left(k\frac{\rho}{L}\right) \\ & - 2\rho\left[(1-\rho)\,\mathrm{sinc}\left(\frac{k}{L}(1-\rho)\right)\mathrm{sinc}\left(k\frac{\rho}{L}\right)\right. \\ & \left. - 2\rho\,\mathrm{sinc}^2\left(k\frac{\rho}{L}\right)\right] \\ & \left.\cdot\,\cos\left(\frac{k\pi}{L}(2l+1)\right)\right]\delta\left(f - \frac{k}{LT_c}\right). \end{aligned} \tag{36}$$

We conclude that, in the case of self-products of PN waveforms, the psd *depends on both the possible offset ρ as well as the number of integer chips k.* The dependence on the number of integer shifts k is introduced either explicitly through the parameter l of (28), or implicitly in (24).

TABLE I
An Example of a PN Sequence and Associated Values of k and l

PN sequence: 111101011001000 (L=15)													
k	1	2	3	4	5	6	7	8	9	10	11	12	13
l	3	5	1	8	2	10	7	4	12	6	13	9	11

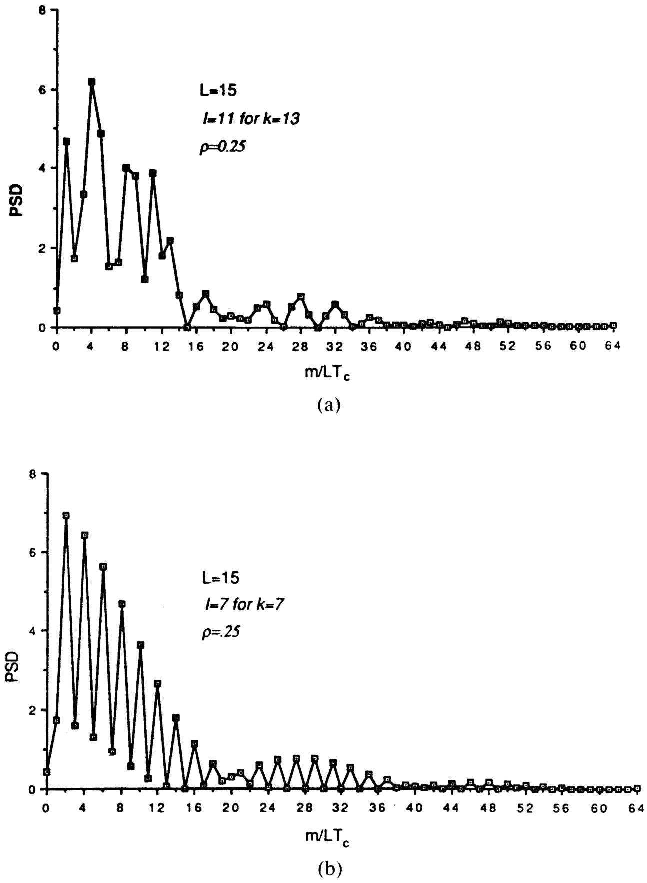


Fig. 6. (a) psd for the self-product of a PN waveform with NRZ signaling, for $k = 13$ and fixed ρ. (b) psd for the self-product of a PN waveform with NRZ signaling, for $k = 7$ and fixed ρ.

To demonstrate this dependence, we can look at Table I which shows the coupling between k and l, for a specific PN sequence. The values of Table I can be used in (32) and (36) or for any other pulse-shape. In Figs. 6(a), (b) and 7(a), (b) NRZ signaling is used. In Fig. 6(a), (b) the value of ρ is fixed ($\rho = 0.25$) and the variation in the resultant psd is depicted for different values of l. In Fig. 7(a), (b) the value of k (and thus l) is fixed, whereas ρ changes from 0.25 to 0.50. In Fig. 8, for the same sequence of Table I and fixed values of k and ρ, the psd is shown for Manchester signaling, which can be compared with the psd of Fig. 7(a) for NRZ pulses.

D. Cross Product of Gold Waveforms

A set of desired sequences in traditional antipodal CDMA are the Gold codes (for more details on PN, Gold and other related sequences see [9], [10], [24]; for other

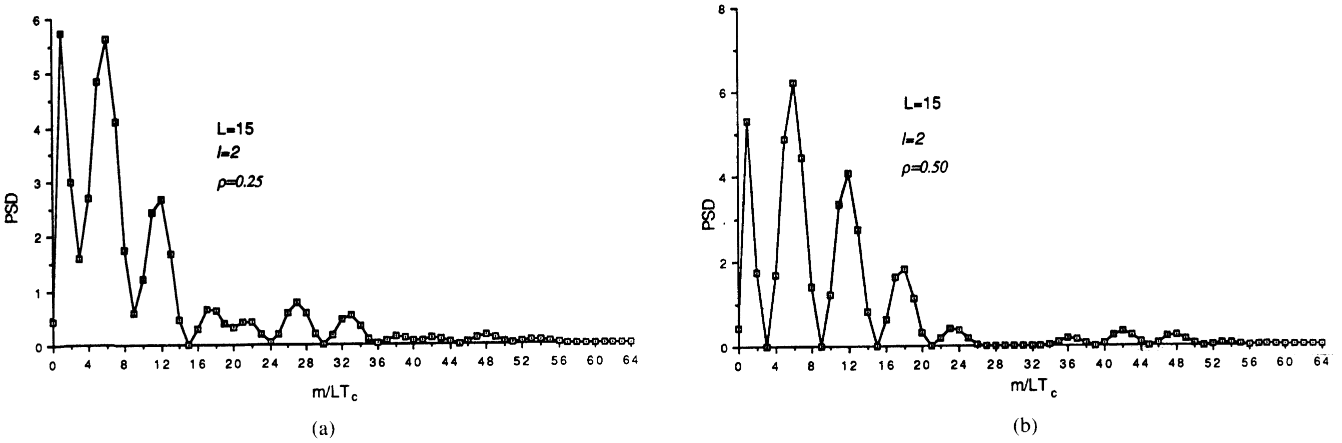


Fig. 7. (a) psd for the self-product of a PN waveform with NRZ signaling, for fixed k and fixed $\rho = 0.25$. (b) psd for the self-product of a PN waveform with NRZ signaling, for fixed k and $\rho = 0.50$.

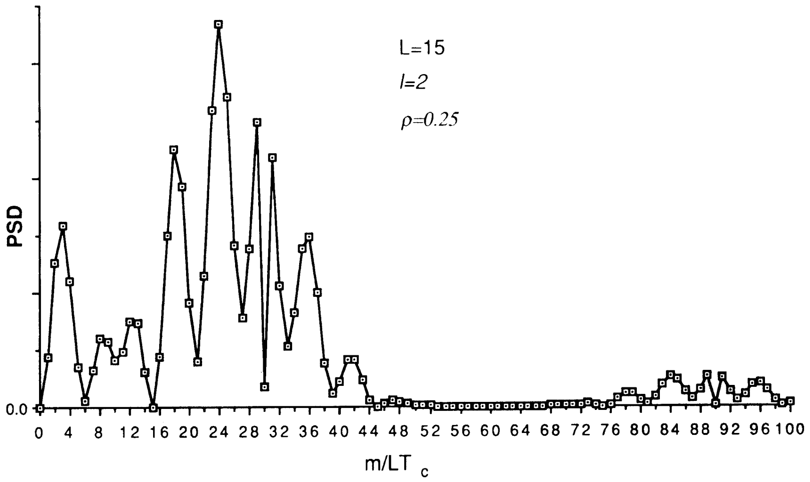


Fig. 8. psd for the self-product of a PN waveform with Manchester signaling, for fixed k and $\rho = 0.25$.

TABLE II
A Preferred Pair of PN Sequences with $L = 31$

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
$I_{1,n}$	1	0	1	0	1	1	1	0	1	1	0	0	0	1	1	1	1	1	0	0	1	1	0	1	0	0	1	0	0	0	0
$I_{2,n}$	1	0	1	1	0	1	0	1	0	0	0	1	1	1	0	1	1	1	1	1	0	0	1	0	0	1	1	0	0	0	0

TABLE III
Sequences a_n and b_n Resulting from Table II, for $k = 0$

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
a_n	0	0	0	1	1	0	1	1	1	1	0	1	1	0	1	0	0	0	1	1	1	1	1	1	0	1	0	0	0	0	0
b_n	1	1	0	0	0	1	0	0	1	1	1	1	1	1	0	0	0	0	1	0	1	0	0	1	1	1	1	0	0	0	1

sequences with good cross-correlation properties see also [15]). Each set of Gold codes of period $L = 2^n - 1$, consisting of $L + 2 = 2^n + 1$ members, can be constructed from an appropriate pair of PN sequences. These pairs of PN sequences are called *preferred pairs*.

Although there exist no closed-form expressions for the auto- or cross-correlation functions of Gold sequences, the present method can be employed for specific pairs. To illustrate, consider the example of Table II, where the two sequences $I_{i,n}$; $i = 1, 2$ under consideration comprise a

TABLE IV
CROSS- AND AUTOCORRELATION FUNCTIONS FOR a_n AND b_n OF TABLE III

m	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
$R_{ab}(m)$	b	b	b	a	a	a	b	b	c	c	c	c	c	c	b	a	b	a	b	b	a	b	a	b	c	c	c	b	c	c	b
$R_a(m)$	c	c	c	b	b	a	a	a	a	a	b	b	b	c	c	c	c	b	b	b	a	a	a	a	a	b	b	c	c	c	1
$R_b(m)$	c	b	a	a	b	b	b	b	a	b	b	c	c	b	b	b	b	c	c	b	b	a	b	b	b	b	ȧ	a	b	c	1

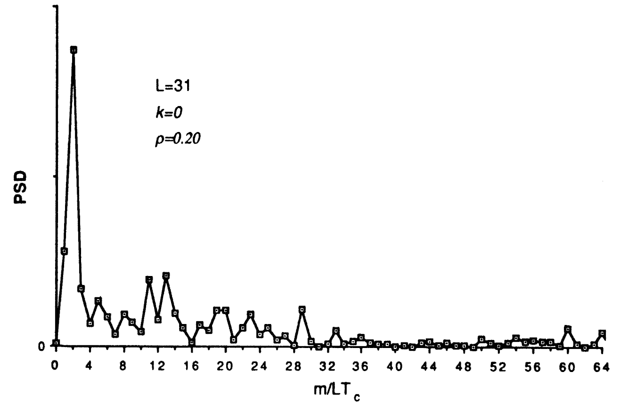


Fig. 9. psd for cross-product of two Gold waveforms with NRZ signaling, for $k = 0$ and $\rho = 0.20$.

preferred pair with $L = 31$. Direct application of the results in Section II for $k = 0$ yields the a_n and b_n sequences of (11a), shown in Table III. They are Gold sequences, and their cross- and autocorrelation functions are shown in Table IV. They take on the values $a = -9/31$, $b = -1/31$, $c = 7/31$, except for $R_a(0) = R_b(0) = 1$.

The psd of the cross-product with NRZ signaling is depicted in Fig. 9. It was determined using the DFT formulation in connection to (9), (13), and (14).

IV. CONCLUSION

A new technique for deriving the psd of digitally modulated signals has been proposed. It focuses on the auto- and cross-correlation functions of certain sequence-modulated delta-trains, and makes use of the DFT coefficients of these functions. This method provides general expressions for psd's and can be useful in a variety of situations, by virtue of a decomposition of the psd expressions into a *sequence factor* and a *pulse-shape factor*. As mentioned in the Introduction, these psd's can help determine certain choices of the system design (pre- and postmultiplication filters, delay parameters, sets of employed code sequences) as well as quantify performance in the presence of signal-induced noise. As examples, we have derived closed-form expressions for the psd's of PN waveforms, and the self-products of PN waveforms with arbitrary modulating pulse-shapes. Some known expressions were shown to be special cases of this general and easily adaptable approach. We have illustrated the dependence of the psd of the self-products on the relative shift between the two versions of the PN sequence, using a specific example. This could potentially have implications on the design of systems employing the standard correlation operation for such PN waveforms. Finally, we have considered the example of OOC's sequences and cross-product of Gold waveforms, following the general formulae developed here. For these cases, only specific examples were entertained.

APPENDIX

In this Appendix we prove some identities which were used in the derivation of the psd's in Section IV-C. We also prove the equivalence of the expression in (32) to the one in [8].

(A.1) We first show that

$$\rho(1-\rho)\,\mathrm{sinc}\,(k(1-\rho))\,\mathrm{sinc}\,(k\rho) = -(-1)^k\rho^2\,\mathrm{sinc}^2\,(k\rho); \qquad k \neq 0. \tag{A.1}$$

Proof: Starting from the left-hand side,

$$\begin{aligned}\rho(1-\rho)\,&\mathrm{sinc}\,(k(1-\rho))\,\mathrm{sinc}\,(k\rho)\\ &= \rho(1-\rho)\\ &\quad\cdot\left(\frac{\sin(k\pi)\cos(k\rho\pi) - \cos(k\pi)\sin(k\rho\pi)}{k(1-\rho)\pi}\right)\\ &\quad\cdot\,\mathrm{sinc}\,(k\rho)\\ &= \rho\left(\frac{-\cos(k\pi)\sin(k\rho\pi)}{k\pi}\right)\mathrm{sinc}\,(k\rho)\\ &= \rho^2\left(\frac{-(-1)^k\sin(k\rho\pi)}{k\rho\pi}\right)\mathrm{sinc}\,(k\rho)\\ &= -(-1)^k\rho^2\,\mathrm{sinc}^2\,(k\rho).\end{aligned}$$

Note that for $k \neq 0$ and $\rho = 0, 1$, (A.1) still holds, since both sides are equal to zero.

(A.2) With the same approach, it can be shown that

$$(1-\rho)^2\,\mathrm{sinc}^2\,(k(1-\rho)) = \rho^2\,\mathrm{sinc}^2\,(k\rho) \qquad k \neq 0 \tag{A.2}$$

Proof: Starting from the left-hand side,

$$\begin{aligned}(1-\rho)^2\,&\mathrm{sinc}^2\,(k(1-\rho))\\ &= (1-\rho)^2\\ &\quad\cdot\left(\frac{\sin(k\pi)\cos(k\rho\pi) - \cos(k\pi)\sin(k\rho\pi)}{k(1-\rho)\pi}\right)^2\\ &= (1-\rho)^2\frac{\sin^2(k\rho\pi)}{(k(1-\rho)\pi)^2} = \rho^2\frac{\sin^2(k\rho\pi)}{(k\rho\pi)^2}\\ &= \rho^2\,\mathrm{sinc}^2\,(k\rho).\end{aligned}$$

(A.3) Here we show the equivalence of (32) to the corresponding expression in [8].

Proof: To match the notation in [8] to the present one, we let

$$\rho = \frac{\epsilon_1}{T}; \quad 1 - \rho = \frac{\epsilon_2}{T}; \quad n_{pq} = l. \tag{A.3}$$

Also, the sinc (x) function here is equivalent to the sinc (πx) function in [8].

The corresponding expression for (32) in [8] is a combination of the (G27), (G28), and (G31) where the letter G is used when referring to equations in [8] in order to avoid confusion. The weight of $\delta(f)$ in (G27) is $\rho^2(\epsilon) = 1/L^2$, which agrees with the corresponding weight of $\delta(f)$ in (32). Because of (A.2), both terms of (G28) are canceling out. We consider next the expression in (G31); according to the present notation, the right-hand side of this equation (with the left-hand side being the $\Sigma_{\substack{k=-\infty \\ k \neq 0}}^{\infty}$ $N_s(f, \epsilon)\, \delta(f - k/T_c)$ term of (G27)) is equivalent to

$$\frac{L+1}{L^2} \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \left[\rho^2 \operatorname{sinc}^2\left(\frac{k\rho}{L}\right) + (1-\rho)^2 \cdot \operatorname{sinc}^2\left(\frac{k}{L}(1-\rho)\right)\right] \delta\left(f - \frac{k}{LT_c}\right)$$

$$+ \frac{1}{2}\frac{L+1}{L^2} \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \cos\left(k\pi \frac{(2l+1)}{L}\right) \cdot \left[\operatorname{sinc}^2\left(\frac{k}{2L}\right) - (1-2\rho)^2 \operatorname{sinc}^2\left(\frac{k(1-2\rho)}{2L}\right)\right] \cdot \delta\left(f - \frac{k}{LT_c}\right). \tag{A.4}$$

A comparison of (A.4) to (32) indicates that for the proof of equivalence, we only have to show that (for $k \neq 0$ and $\rho \neq 1, 0$)

$$4\rho(1-\rho) \operatorname{sinc}\left(\frac{k\rho}{L}\right) \operatorname{sinc}\left(\frac{k(1-\rho)}{L}\right) = \operatorname{sinc}^2\left(\frac{k}{2L}\right) - (1-2\rho)^2 \operatorname{sinc}^2\left(\frac{k(1-2\rho)}{2L}\right) \tag{A.5}$$

It is obvious that for $k \neq 0$ and $\rho = 0, 1$, (A.5) holds. To prove (A.5), we start from the right-hand side.

$$\operatorname{sinc}^2\left(\frac{k}{2L}\right) - (1-2\rho)^2 \operatorname{sinc}^2\left(\frac{k(1-2\rho)}{2L}\right) = \frac{\sin^2\left(\frac{k\pi}{2L}\right)}{\left(k\frac{\pi}{2L}\right)^2} - (1-2\rho)^2 \frac{\sin^2\left(\frac{k\pi(1-2\rho)}{2L}\right)}{\left(k\pi\frac{(1-2\rho)}{2L}\right)^2}$$

$$= \frac{\sin^2\left(\frac{k\pi}{2L}\right) - \left[\sin\left(\frac{k\pi}{2L}\right)\cos\left(k\pi\frac{\rho}{L}\right) - \sin\left(k\pi\frac{\rho}{L}\right)\cos\left(\frac{k\pi}{2L}\right)\right]^2}{\left(k\frac{\pi}{2L}\right)^2}$$

$$= \frac{4}{\left(k\frac{\pi}{L}\right)^2}\left[\sin^2\left(\frac{k\pi}{2L}\right) - \sin^2\left(\frac{k\pi}{2L}\right)\cos^2\left(k\pi\frac{\rho}{L}\right) - \sin^2\left(k\pi\frac{\rho}{L}\right)\cos^2\left(\frac{k\pi}{2L}\right)\right.$$
$$\left. + 2\sin\left(\frac{k\pi}{2L}\right)\cos\left(k\pi\frac{\rho}{L}\right)\sin\left(k\pi\frac{\rho}{L}\right)\cos\left(\frac{k\pi}{2L}\right)\right]$$

$$= \frac{4}{\left(k\frac{\pi}{L}\right)^2}\left[\sin^2\left(\frac{k\pi}{2L}\right)\left[1 - \cos^2\left(k\pi\frac{\rho}{L}\right)\right] - \sin^2\left(k\pi\frac{\rho}{L}\right)\cos^2\left(\frac{k\pi}{2L}\right)\right.$$
$$\left. + \sin\left(\frac{k\pi}{L}\right)\cos\left(k\pi\frac{\rho}{L}\right)\sin\left(k\pi\frac{\rho}{L}\right)\right]$$

$$= \frac{4}{\left(k\frac{\pi}{L}\right)^2}\left[\sin^2\left(\frac{k\pi}{2L}\right)\sin^2\left(k\pi\frac{\rho}{L}\right) - \sin^2\left(k\pi\frac{\rho}{L}\right)\cos^2\left(\frac{k\pi}{2L}\right)\right.$$
$$\left. + \sin\left(\frac{k\pi}{L}\right)\cos\left(k\pi\frac{\rho}{L}\right)\sin\left(k\pi\frac{\rho}{L}\right)\right]$$

$$= \frac{4}{\left(k\frac{\pi}{L}\right)^2}\left[\sin^2\left(k\pi\frac{\rho}{L}\right)\left[\sin^2\left(\frac{k\pi}{2L}\right) - \cos^2\left(\frac{k\pi}{2L}\right)\right]\right.$$

$$\left. + \sin\left(\frac{k\pi}{L}\right)\cos\left(k\pi\frac{\rho}{L}\right)\sin\left(k\pi\frac{\rho}{L}\right)\right]$$

$$= \frac{4}{\left(k\frac{\pi}{L}\right)^2}\left[-\sin^2\left(k\pi\frac{\rho}{L}\right)\cos\left(\frac{k\pi}{L}\right) + \sin\left(\frac{k\pi}{L}\right)\cos\left(k\pi\frac{\rho}{L}\right)\sin\left(k\pi\frac{\rho}{L}\right)\right]$$

$$= \frac{4}{\left(k\frac{\pi}{L}\right)^2}\sin\left(k\pi\frac{\rho}{L}\right)\left[\sin\left(\frac{k\pi}{L}\right)\cos\left(k\pi\frac{\rho}{L}\right) - \sin\left(k\pi\frac{\rho}{L}\right)\cos\left(\frac{k\pi}{L}\right)\right]$$

$$= \frac{4}{\left(k\frac{\pi}{L}\right)^2}\sin\left(k\pi\frac{\rho}{L}\right)\sin\left(\frac{k(1-\rho)\pi}{L}\right)$$

$$= 4\rho(1-\rho)\frac{\sin\left(k\pi\frac{\rho}{L}\right)}{\rho\left(k\frac{\pi}{L}\right)}\,\frac{\sin\left(\frac{k(1-\rho)\pi}{L}\right)}{(1-\rho)\left(k\frac{\pi}{L}\right)}$$

$$= 4\rho(1-\rho)\,\mathrm{sinc}\left(\frac{k\rho}{L}\right)\mathrm{sinc}\left(\frac{k(1-\rho)}{L}\right).$$

Acknowledgment

The authors wish to wholeheartedly thank a reviewer for a thorough, witty, and very useful reading of the draft. It reshaped the presentation significantly to the better.

References

[1] S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[2] R. A. Dillard and G. M. Dillard, *Detectability of Spread-Spectrum Signals*. Artech House, 1989.

[3] D. Divsalar and M. K. Simon, "Spectral characteristics of convolutionally coded digital signals," *IEEE Trans. Commun.*, vol. COM-28, pp. 173-186, Feb. 1980.

[4] J. K. Holmes, *Coherent Spread Spectrum Systems*. New York: Wiley, 1982.

[5] N. A. Ismail, "Comparison of sequences for local area networks using spread-spectrum multiple access," *Proc. IEEE*, vol. 76, pp. 87-88, Jan. 1988.

[6] W. J. Gill, "Effect of synchronization error in pseudorandom carrier communications," in *1st Ann. IEEE Conf. Rec.*, Denver, CO, June 7-9, 1965, pp. 187-191.

[7] W. J. Gill and J. J. Spilker, "An interesting decomposition property for the self-products of random or pseudorandom binary sequences," *IEEE Trans. Commun. Syst.*, June 1963.

[8] S. G. Glisic, "Power density spectrum of the product of the time displaced versions of maximum length binary PN signals," *IEEE Trans. Commun.*, vol. COM-31, pp. 281-286, Feb. 1983.

[9] R. Gold, "Optimal binary sequences for spread spectrum multiplexing," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 619-621, Oct. 1967.

[10] —, "Maximal recursive sequences with 3-valued recursive cross-correlation functions," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 154-156, Jan. 1968.

[11] S. W. Golomb, *Shift Register Sequences*. Agean Press, 1982, revised ed.

[12] S. W. Golomb and H. Taylor, "Constructions and properties of costas arrays," *Proc. IEEE*, vol. 72, Sept. 1984.

[13] D. Middleton, *An Introduction to Statistical Communication Theory*. New York: McGraw-Hill, 1960; see also reprint edition, Peninsula Publishing, 1987.

[14] R. S. Neto, A. Polydoros, and R. A. Scholtz, "Performance of standard code-tracking loops in the presence of dual tone interference," *IEEE Trans. Commun.*, vol. COM-34, pp. 999-1008, Oct. 1986.

[15] J.-S. No and P. Vijay Kumar, "A new family of binary pseudorandom sequences having optimal periodic correlation properties and large linear span," *IEEE Trans. Inform. Theory*, vol. 35, pp. 371-379, Mar. 1989.

[16] R. Padovani and G. L. Pierobon, "Spectral analysis of digital messages through finite memory transformations," *IEEE Trans. Commun.*, vol. COM-32, pp. 1214-1218, Nov. 1984.

[17] A. Polydoros and N. B. Pronios, "On the power spectral density of digitally modulated signals and applications in code despreading," in *Proc. IEEE MILCOM'88*, pp. 977-981.

[18] A. Polydoros and C. L. Weber, "A unified approach to serial search spread-spectrum code acquisition, Part I and II," *IEEE Trans. Commun.*, vol. COM-32, pp. 30-43, May 1984.

[19] —, "Analysis and optimization of correlative code-tracking loops in spread spectrum systems," *IEEE Trans. Commun.*, vol. COM-33, pp. 30-43, Jan. 1985.

[20] —, "Wideband detection performance considerations for direct-sequence and time-hopping LPI waveforms," *IEEE J. Select. Areas Commun.*, vol. SAC-3, pp. 727-744, Sept. 1985.

[21] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1983.

[22] P. R. Prucnal and M. A. Santoro, "Spread spectrum optic local area networks, using optical processing," in *Digital Communications*, E. Biglieri and G. Prati, Eds. New York: Elsevier Science, 1986, pp. 193-202.

[23] J. A. Salehi, "Code division multiple-access techniques in optical fiber networks—Part I: Fundamental principles," *IEEE Trans. Commun.*, vol. 37, pp. 824-833, Aug. 1989.

[24] D. V. Sarwate and M. B. Pursley, "Crosscorrelation properties of pseudorandom and related sequences," *Proc. IEEE*, vol. 68, pp. 5937-619, May 1980.

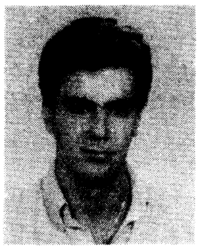
[25] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications*. Rockville, MD: Computer Science, 1985.

[26] C. T. Spracklen and C. Smythe, "Communication protocols for a spread spectrum local area network," in *Proc. FOC/LAN 84*, Las Vegas, NV, Sept. 19-21, 1984, pp. 70-74.

[27] R. C. Titsworth and L. R. Welch, "Modulation by random and pseudorandom sequences," Progress Rep. 20-387, JPL, June 1959.

[28] H. Urkovitz, "Energy detection of unknown deterministic signals," *Proc. IEEE*, vol. 55, pp. 523-531, Apr. 1967.

[29] R. E. Ziemer and R. L. Peterson, *Digital Communications and Spread Spectrum Systems*. New York: Macmillan, 1985.



Nikos B. Pronios (S'86-M'88) was born in Larissa, Greece, in 1960. He received the Diploma in electrical engineering from the University of Patras, Greece, in 1982, and the M.Sc. and the Ph.D. degrees in electrical engineering from the University of Southern California, Los Angeles, in 1984 and 1988 respectively.

He has been with Bell Communications Research since 1988. His current interests include spread-spectrum systems, HDTV, medical imaging, and PACS.

Dr. Pronios is a member of SPIE, Eta Kappa Nu, and the Technical Chamber of Greece.

Andreas Polydoros (S'76-M'82) was born in Athens, Greece, in 1954. He was educated at the National Technical University of Athens, Greece (Diploma in EE, 1977), State University of New York at Buffalo (MSEE, 1979) and the University of Southern California (Ph.D., E.E., 1982).

He has been a faculty member in the Electrical Engineering/Systems Department and the Communication Sciences Institute at U.S.C. since 1982, currently as an Associate Professor. He has also been affiliated with Axiomatrix Corporation since 1979, involved in a variety of digital telecommunication projects with current emphasis on spread spectrum systems. He is currently participating in the development of a textbook entitled, *Principles of Spread Spectrum Systems*.

Dr. Polydoros is the recipient of a 1986 NSF Presidential Young Investigator Award. He has also served as the Associate Editor for Communications of the IEEE TRANSACTIONS ON INFORMATION THEORY.

Interception of Frequency-Hopped Spread-Spectrum Signals

NORMAN C. BEAULIEU, WENDY L. HOPKINS, MEMBER, IEEE, AND PETER J. MCLANE, FELLOW, IEEE

Abstract—**A frequency-hopped spread-spectrum signal is modeled as a sinusoid that has one of N random frequencies. Coherent and noncoherent interception receiver structures based on Neyman–Pearson detection theory are determined. Under the assumption that there is a single hop per detection period, the optimum receiver structure is shown to consist of a bank of matched filters called the average likelihood (AL) receiver. A suboptimum structure called the maximum likelihood (ML) receiver is also analyzed. It is shown that AL and ML receivers have essentially the same performance. Simple formulas that relate the probability of detection P_D to the probability of false alarm P_F, and the signal-to-noise ratio (SNR) for large N are derived.**

Receiver structures are also derived and analyzed for the case where the signal hops a number of times in one detection interval. This may correspond to the detection of a multihop signal in one symbol interval or to detection based on integration over a number of symbol intervals. The relationships of P_D to P_F, for both coherent and noncoherent multiple hop receivers, are examined.

I. Introduction

During the last 15 years, spread-spectrum communications has been receiving increased interest in the open literature. In addition to military applications, numerous civilian uses are being developed. Examples can be found in [1]–[5].

Spread-spectrum signals have several forms: frequency hopping (FH), pseudonoise or direct sequence (PN), and time hopping. Time hopping is generally used with either FH or PN signaling, but will not be discussed in this paper. The first two forms are the most commonly used and are similar in that the frequency of the transmitted signal is altered by a pseudorandom sequence. For FH signals, this sequence selects the carrier frequency, while for PN signals, the sequence generates a phase that is used to modulate the message. Tutorials on spread-spectrum communications can be found in [1] and [2].

Among the many attractive properties of spread spectrum, from a user's point of view, it its low probability of interception [5]. Since the signal power for the PN signals is spread across the transmission bandwidth to resemble noise (hence the name), and since FH signals have their power transmitted in a narrow bandwidth, PN signals may have a lower probability of interception. However, FH signals may have a much wider transmission bandwidth and superior antijamming properties. Therefore, some systems will use FH signaling rather than PN.

The growing use of spread-spectrum communications has given rise to the question of interception by an unauthorized receiver. Since this question is not only of interest to the military, spread-spectrum signal interception is being addressed by communications regulatory boards, such as the CRTC in Canada, and the FCC in the United States.

In this paper, we model a frequency-hopped signal as a sinusoidal signal with a random frequency that has a discrete probability distribution. The number of possible frequencies in this distribution is denoted by N. Neyman–Pearson detection theory is used to establish the constant false alarm rate (CFAR) interception receiver. For a given false alarm probability, the detection probability of this receiver represents a fundamental detection performance limit for any receiver. The frequency-hopping rate is first assumed to be slow so that the modulated signal is not allowed to change in frequency over the detection interval. The analysis is then generalized to fast hopping where the signal hops M times in one detection interval.

The receiver structure derived for slow hopping consists of a summation of the outputs from a bank of correlators, one correlator for each discrete frequency in the random frequency distribution. Such a receiver can be realized using surface acoustic wave filters [6]. The resulting receiver is called the average likelihood (AL) receiver. The output from each correlator in the bank could also be tested. This receiver, the maximum likelihood (ML) receiver, is easily analyzed and has very similar performance to the AL receiver. This is fortunate, as the AL receiver is not as easily implemented or analyzed. The near equivalence of the two receivers in a radar detection problem with unknown Doppler frequency was established by Brennan *et al.* [7]. Herein, we establish the near equivalence of the AL and ML spread-spectrum receivers for small N, i.e., $N \leq 20$, and large N, i.e., $N > 10^4$. The latter case represents a key issue of the security of a slow, frequency-hopped spread-spectrum signal. Both coherent and noncoherent AL and ML receivers are derived and analyzed. It is shown that the detection probability P_D tends to the false alarm rate P_F as $N \to \infty$. Furthermore, a value of SNR is found such that $P_D \simeq P_F$ for all

Manuscript received February 17, 1989; revised October 23, 1989. This work was supported under a D.S.S. Contract based on funding from the Canadian Department of Communications and the Canadian Department of National Defence; and by the Natural Sciences and Engineering Research Council of Canada under Grants A-3986 and A-7389.

N. C. Beaulieu and P. J. McLane are with the Department of Electrical Engineering, Queen's University, Kingston, Ont., Canada.

W. L. Hopkins was with the Department of Electrical Engineering, Queen's University, Kingston, Ont., Canada. She is now with Bell-Northern Research, Ottawa, Ont., Canada.

IEEE Log Number 9035811.

SNR's smaller than this one. For such SNR's, signal interception is impractical as the receiver cannot distinguish detections from false alarms. This SNR increases with N but decreases with the false alarm rate. For a given N, this SNR represents the ultimate security of the frequency-hopped signal. This is because it is assumed that perfect chip synchronization is available to the interception receiver.

The fast-hopping signal receiver derived has the following conceptual structure. There is a bank of M AL receivers, where M is the number of hops per detection interval. Each AL receiver is an optimum receiver for detecting the presence or absence of the signal on one of the interval chips corresponding to one hop. The logarithms of the M chip receiver outputs are summed and thresholded. This may be implemented using one AL receiver and M consecutive outputs. That is, the AL receiver is matched to the chip interval and the output is used M times. This receiver, the multihop average likelihood (MAL) receiver, is approximated by a multihop maximum likelihood (MML) receiver. Coherent and noncoherent MML receivers are analyzed. Analogous results to those obtained for the slow-hop ML receiver are obtained.

We consider explicitly two situations. The first is the case where the signal hops once per detection interval. The second is when the signal hops several times per detection interval. The first case represents the situation where the detection interval and the symbol interval are the same. The second case models a more general situation. The transmitter may hop M_H times per symbol and the receiver may integrate over several, M_S, symbol periods with $M_H \cdot M_S = M$.

The main references on the problem we consider can be found in [7]–[13].

II. The Single-Hop AL and ML Receivers

The signal is assumed to change frequency once per detection interval. The detection problem is formulated as a binary hypothesis problem. Hypothesis H_1 is that a signal is present, and hypothesis H_0 is that no signal is present. That is,

$$\begin{aligned} H_1&: r(t) = s(t) + n(t) \\ H_0&: r(t) = n(t) \end{aligned} \qquad 0 \le t \le T$$

where

$$s(t) = A \cos(\omega t + \theta) \tag{1}$$

with $n(t)$, white Gaussian noise with spectral height, $N_o/2$ W/Hz. The signal, when present, has amplitude, angular frequency, and phase constant: A, ω, and θ, respectively. The signal duration is denoted as T. In the coherent case, $\theta = 0$, and for the noncoherent case it is uniform in $[0, 2\pi]$. The angular frequency ω is assumed to have the discrete distribution

$$p_\omega(\omega) = \frac{1}{N}\sum_{i=1}^{N} \delta(\omega - \omega_i) \tag{2}$$

where $\delta(x)$ is the Dirac-delta function. The frequencies are spaced multiples of $1/T$ or $1/2T$ apart for noncoherent and coherent systems, respectively, making the signals orthogonal.

The likelihood ratio for this problem is given by [14]

$$L[r(t)] = E_{\omega,\theta}\left\{\exp\left\{\frac{2}{N_o}\int_0^T r(t)s(t)\,dt - \frac{E}{N_o}\right\}\right\} \tag{3}$$

where E is the signal energy and $s(t)$ is given in (1). Here $E_{\omega,\theta}$ represents expectation with respect to frequency and phase. The signal energy is assumed to be independent of these signal parameters.

For the coherent case, the expectation in θ in (3) does not occur as the phase θ is assumed to be perfectly known. The unknown frequency has the pdf in (2), and for $s(t)$ in (1) the AL ratio test is

$$L = \sum_{i=1}^{N} l_i \underset{H_0}{\overset{H_1}{\gtrless}} \eta N e^{d^2/2} = \eta' \tag{4}$$

where $l_i = e^{\alpha_i}$, the signal-to-noise ratio (SNR) is $d^2 = 2E/N_o$, and

$$\alpha_i = \frac{2A}{N_o}\int_0^T r(t) \cos \omega_i t \, dt. \tag{5}$$

The constant η sets the false alarm rate, and η' is then the detection threshold. One notes that α_i is Gaussian and that the sufficient statistic is a sum of independent, lognormal random variables e^{α_i}. The independence follows from the frequency spacing assumed following (2) as a consequence of the orthogonality of the signals [25, p. 151]. The receiver structure is shown in Fig. 1.

For the noncoherent case, the phase is uniform in $[0, 2\pi]$. Performing the expectations in (3), the AL ratio test is

$$L = \sum_{i=1}^{N} I_o(q_i) \underset{H_0}{\overset{H_1}{\gtrless}} \eta N e^{d^2/2} = \eta' \tag{6}$$

where $I_o(x)$ is the modified Bessel function of zero order and

$$q_i^2 = L_{c_i}^2 + L_{s_i}^2 \tag{7}$$

$$L_{c_i} = \frac{2A}{N_o}\int_0^T r(t) \cos \omega_i t \, dt \tag{8}$$

and

$$L_{s_i} = \frac{2A}{N_o}\int_0^T r(t) \sin \omega_i t \, dt. \tag{9}$$

The independence of the summands in (6) again follows from the assumed frequency spacing [25, p. 210]. The receiver structure is shown in Fig. 2.

The coherent ML receiver is a suboptimum, hard-limited version of the AL receiver shown in Fig. 1. In words, the test is as follows. An optimal test to determine the presence or absence of signal at each of the N frequencies is performed ignoring the other frequencies. The overall

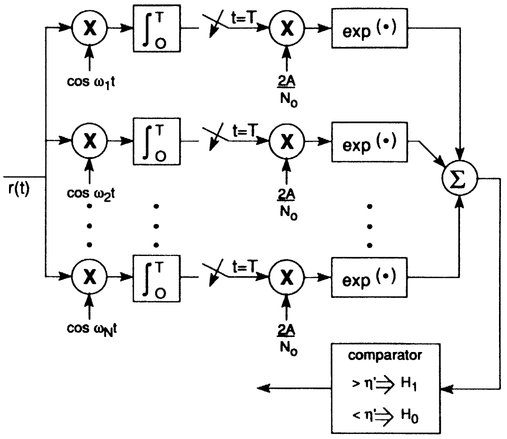


Fig. 1. Optimum coherent receiver.

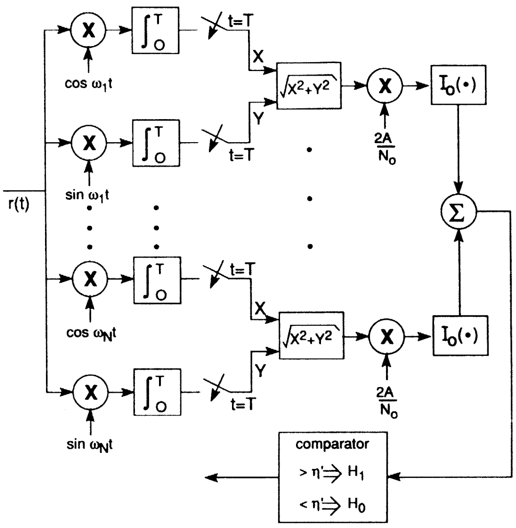


Fig. 2. Optimum noncoherent receiver.

decision as to whether signal is present at any of the N frequencies is made by declaring the signal present if the signal has been declared present for one or more of the N frequencies. The optimum AL receiver may then be viewed as a "soft" combining of the signal tests for the individual frequencies where the combining is done with exponential weighting. The optimal detection of the presence or absence of signal for the ith frequency is obtained by setting $N = 1$ in (4),

$$e^{\alpha_i} \underset{H_0}{\overset{H_1}{\gtrless}} \eta e^{d^2/2},$$

or

$$A\int_0^T r(t)\cos\omega_i t\,dt \underset{H_0}{\overset{H_1}{\gtrless}} \frac{N_o}{2}\ln\eta + \frac{E}{2} = \gamma.$$

The ML receiver statistic may be written in the form of (4) by letting l_i be a random variable (R.V.) that assumes the values 0 and 1 as signal is declared absent or present at the ith frequency. Then,

$$L = \sum_{i=1}^{N} l_i \underset{H_0}{\overset{H_1}{\gtrless}} 1/2 \qquad (10)$$

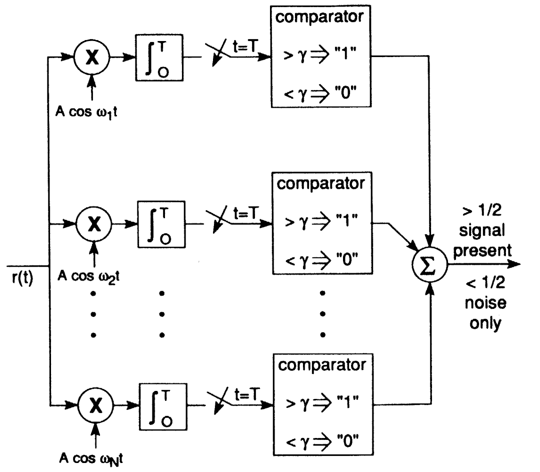


Fig. 3. Coherent maximum likelihood receiver.

with l_i determined by

$$A\int_0^T r(t)\cos\omega_i t\,dt \underset{l_i=0}{\overset{l_i=1}{\gtrless}} \frac{N_o}{2}\ln\eta + \frac{E}{2} = \gamma, \qquad (11)$$

and where the threshold for L is chosen arbitrarily to be $1/2$. Any value between 0 and 1 may be used. The ML receiver structure is shown in Fig. 3. The probability of detection of signal at the ith frequency can be shown to be

$$Q_D = Q\left(\frac{\ln\eta}{d} - \frac{d}{2}\right) \qquad (12)$$

and is independent of i. The probability of false alarm is

$$Q_F = Q\left(\frac{\ln\eta}{d} + \frac{d}{2}\right). \qquad (13)$$

Here $Q(x)$ is the area under the unit variance, zero mean, Gaussian probability density function from x to ∞. The ML receiver detection and false alarm probabilities are, respectively,

$$P_D = 1 - (1 - Q_D)(1 - Q_F)^{N-1} \qquad (14)$$

and

$$P_F = 1 - (1 - Q_F)^N. \qquad (15)$$

The noncoherent ML receiver is the analogous suboptimum version of the noncoherent AL receiver of Fig. 2. That is, l_i denotes an R.V. whose outcome depends upon an optimum test for signal at the ith frequency according to

$$q_i \underset{l_i=0}{\overset{l_i=1}{\gtrless}} I_o^{-1}(\eta e^{d^2/2}) = \gamma \qquad (16)$$

which follows from (6) by setting $N = 1$. The overall test is then

$$L = \sum_{i=1}^{N} l_i > 1/2$$

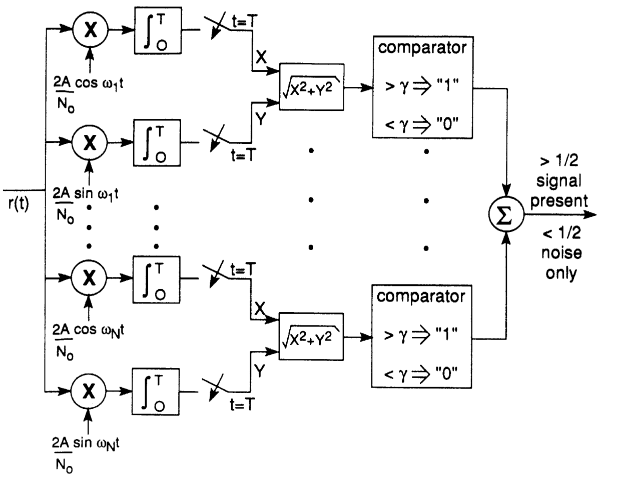


Fig. 4. Noncoherent maximum likelihood receiver.

in analogy to the coherent ML test. The receiver structure is shown in Fig. 4. It follows from [14] that

$$Q_F = \exp\left[-\gamma^2/2d^2\right] \tag{17}$$

and

$$Q_D = Q_M(d, \gamma/d) \tag{18}$$

where $Q_M(\alpha, \beta)$ is Marcum's Q-function,

$$Q_M(\alpha, \beta) = \int_\beta^\infty x \exp\left\{-\frac{x^2+\alpha^2}{2}\right\} I_o(\alpha x)\, dx. \tag{19}$$

The overall noncoherent ML detector performance is given by (14) and (15) with Q_F and Q_D given in (17)-(19).

III. Detection Performance for Small N

A. Coherent Receivers

For the ML case, the performance is given by (12)-(15). Computational results for 10 frequencies are given in Fig. 5 in the usual format for the receiver operating characteristic (ROC).

To analyze the AL receiver, it follows from (4) and (5) that the distribution of a sum of independent, lognormal R.V.'s is required. This is an unsolved problem. Accordingly, various approximations were attempted and are documented in [16]. The most useful approximations were taken from the work of Schwartz and Yeh [15] who considered the distribution of power sums in mobile radio communications.

The first approximation is due to Wilkinson. This approximation regards

$$L = e^z = \sum_{i=1}^N e^{\alpha_i} \tag{20}$$

and assumes that L is lognormal. An argument on the validity of this assumption is presented in [17]. In our detection problem, we need the mean, m_z, and variance, σ_z, of z as

$$P(L > \eta') = P(z > \ln \eta') = Q\left(\frac{\ln \eta' - m_z}{\sigma_z}\right). \tag{21}$$

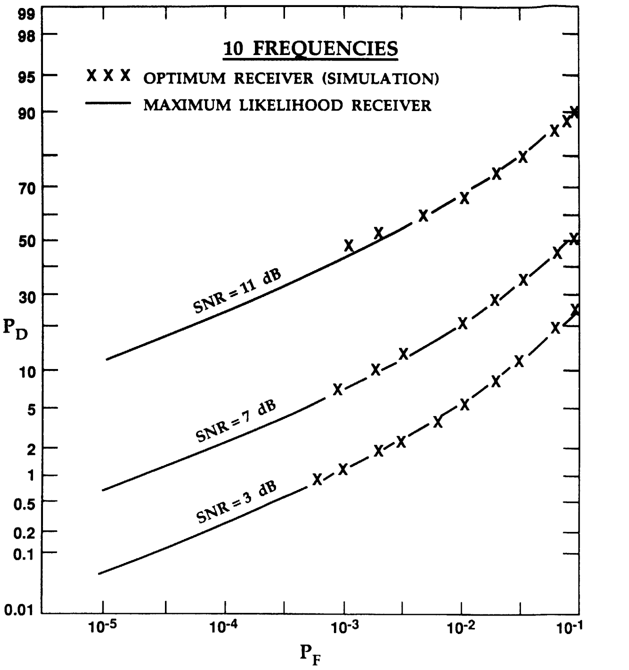


Fig. 5. Performance curves of optimum coherent and coherent maximum likelihood receivers for ten frequencies.

Denote the first and second moments of L as μ_1 and μ_2, respectively. Then as $L = e^z$,

$$\ln \mu_1 = m_z + \sigma_z^2/2$$
$$\ln \mu_2 = 2m_z + 2\sigma_z^2$$

and

$$m_z = 2 \ln \mu_1 - (\ln \mu_2)/2 \tag{22}$$

$$\sigma_z^2 = \ln \mu_2 - 2 \ln \mu_1. \tag{23}$$

The moments μ_1 and μ_2 are obtained from (20) using the independence of the summands

$$\mu_1 = \sum_{i=1}^N \exp\left(m_{\alpha_i} + \sigma_{\alpha_i}^2/2\right) \tag{24}$$

$$\mu_2 = \sum_{i=1}^N \exp\left(2m_{\alpha_i} + \sigma_{\alpha_i}^2\right) \cdot \left[\exp\left(\sigma_{\alpha_i}^2\right) - 1\right] + \mu_1^2 \tag{25}$$

where α_i is a normal R.V. with mean m_{α_i} and variance $\sigma_{\alpha_i}^2$. That is, α_i is $\eta(m_{\alpha_i}, \sigma_{\alpha_i}^2)$.

Another approximation to the probability distribution function of L, Farley's approximation, is discussed in Schwartz and Yeh. Let $P_N = 10 \log_{10} L$ where

$$L = \sum_{i=1}^N 10^{x_i/10}$$

with the x_i $\eta(m_x, \sigma_x)$ and independent. Then, Farley's approximation is

$$P\left\{\frac{P_N - m_x}{\sigma_x} < t\right\} \approx \left[\Phi(t)\right]^N \tag{26}$$

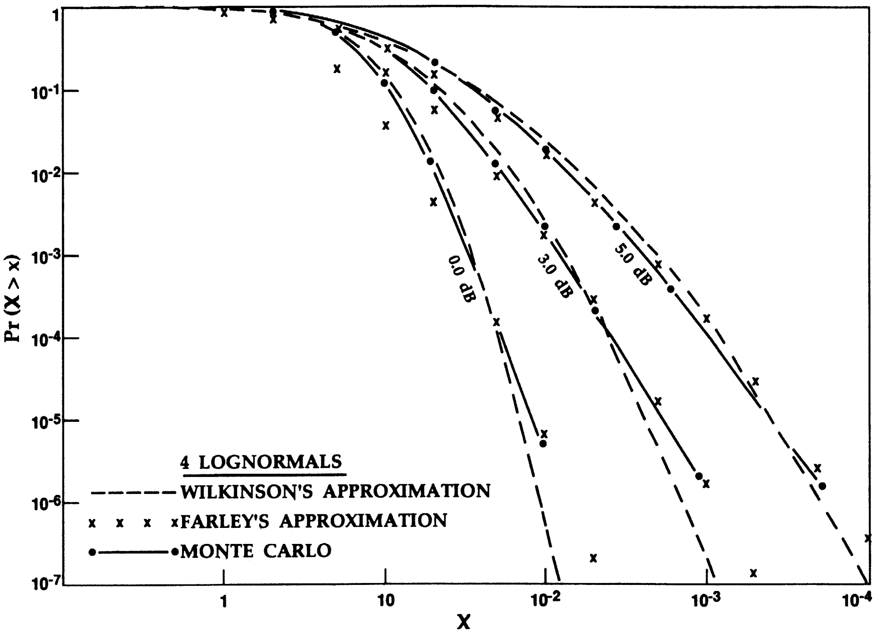


Fig. 6. The probability distribution function, and two approximations, of the sum of four lognormal random variables. The curves are distinguished by the common variance of the exponents. The exponents of the four random variables all have zero mean.

where $\Phi(t) = 1 - Q(t)$. The transformation $x_i = 10\ \alpha_i/\ln 10$ converts L to the required form

$$L = \sum_{i=1}^{N} e^{\alpha_i}.$$

Thus, in (26), setting $\beta = 10/\ln 10$ gives

$$P(10 \log_{10} L - m_{\alpha_i}\beta < \beta\sigma_{\alpha_i} t) \approx [\Phi(t)]^N$$

or as $e^x = 10^{x/\ln 10}$

$$P(L < \exp\{t\sigma_{\alpha_i} + m_{\alpha_i}\}) = [\Phi(t)]^N.$$

Finally, letting $\eta' = \exp\{t\sigma_{\alpha_i} + m_{\alpha_i}\}$

$$P(L < \eta') \approx \left[\Phi\left(\frac{\ln \eta' - m_{\alpha_i}}{\sigma_{\alpha_i}}\right)\right]^N. \qquad (27)$$

The Wilkinson [i.e., (21)-(25)] and Farley [i.e., (27)] approximations to the sum of 4 lognormals are presented in Figs. 6 and 7. Following Schwartz and Yeh's notation, $\sigma^2_{\alpha_i}$ is given in dB and distinguishes the curves. Fig. 6 shows the case where all four $m_{\alpha_i} = 0$, and Fig. 7 shows the case where three $m_{\alpha_i} = 0$ and one $m_{\alpha_i} = \sigma_{\alpha_i}$.[1] In all cases, σ_{α_i} is the same for all i. These two cases are those encountered in (4).

As a basis for comparison, the probability distribution function of $L = \Sigma_{i=1}^{4} e^{\alpha_i}$ was determined by Monte-Carlo simulation. It appears that Farley's approximation is best at low probability, whereas the Wilkinson approximation is best at high probability.

Schwartz and Yeh do not report on the basis of Farley's approximation. However, it may be related to approximating $\Sigma_{i=1}^{N} e^{\alpha_i}$ by $\Sigma_{i=1}^{N} l_i$ where $l_i = 0, 1$ if e^{α_i} exceeds a threshold. This follows since P_F in our detection law (4) is given by (4) and (27) as

$$P_{F,\mathrm{AL}} = P(L > \eta') \approx 1 - \left[\Phi\left(\frac{\ln \eta' - m_{\alpha_i}}{\sigma_{\alpha_i}}\right)\right]^N,$$

and $m_{\alpha_i} = 0$, $\sigma_{\alpha_i} = d$, $\Phi(t) = 1 - Q(t)$. Letting $\eta' = \eta \exp(d^2/2)$, we get, from (13) and (15), $P_{F,\mathrm{AL}} \approx P_{F,\mathrm{ML}}$. In other words, Farley's approximation to the P_F of the AL receiver yields the P_F of the ML receiver. Thus, as Farley's approximation is quite good for $P_F < 10^{-2}$, one expects the AL and ML receivers to have similar detection performance in this range. A similar result was found earlier for a Doppler radar detection problem by Brennan *et al.* [7]. For modest N, we are confident that the detection performance of the AL receiver can be closely estimated by the simple detection formulas for the ML receiver.

B. Noncoherent Receivers

For the ML case, the detection performance is given by (14)-(19). The ROC for 10 frequencies is shown in Fig. 8. Results are given for much larger values of N in Section IV-C.

The detection variable for the AL case is given by (6)-(9). To determine the false alarm probabilities, a sum of Bessel functions whose arguments are Rayleigh R.V.'s must be considered. This is difficult to treat analytically and, therefore, a Monte-Carlo simulation was used. These digital computer simulation results are also shown in Fig. 8. The ML and AL receivers have similar performances.

[1] These results were obtained by using the Wilkinson or Farley approximation for $N = 3$ and conditioning on the dissimilar R.V. [16], [21, p. 135].

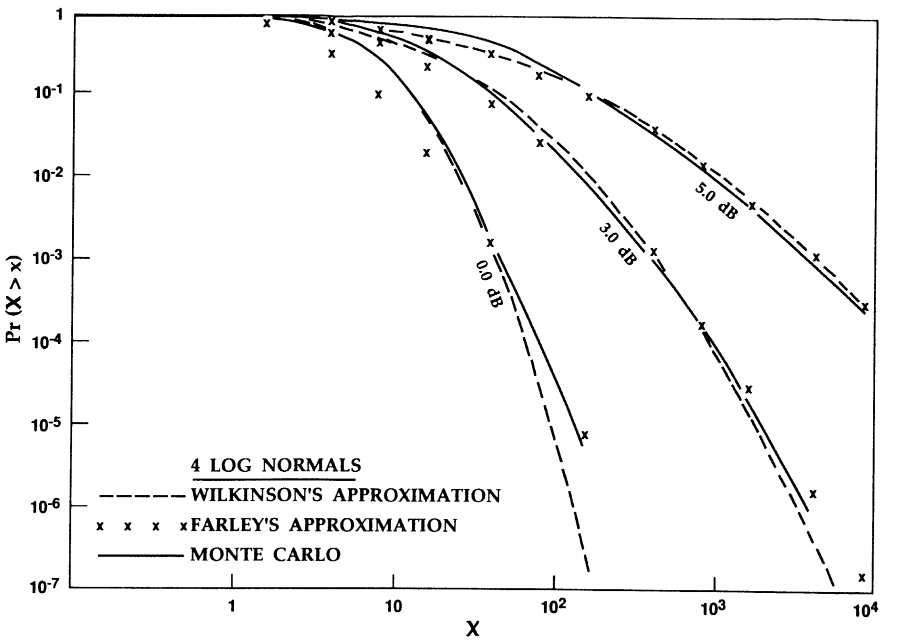


Fig. 7. The probability distribution function, and two approximations, of the sum of four lognormal random variables. The mean of the exponent of one of the random variables equals the SNR. The exponents of the other three have zero mean. The variances of the exponents are equal.

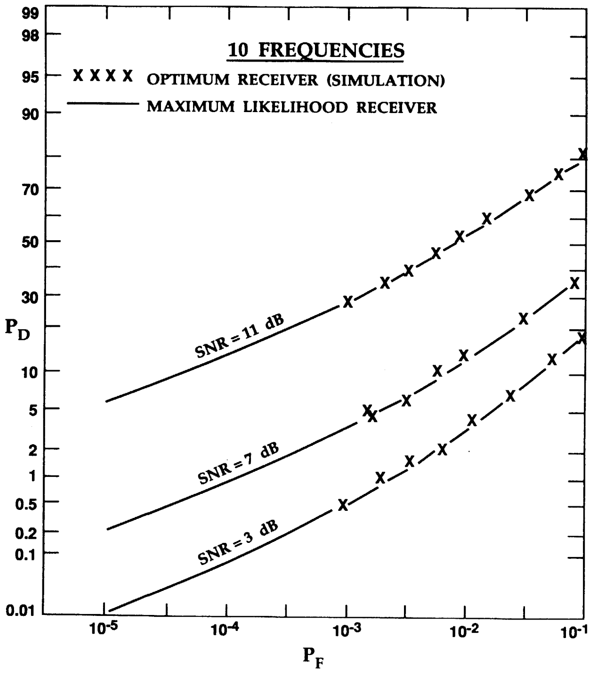


Fig. 8. Comparison of optimum noncoherent and noncoherent maximum likelihood receiver performances for ten frequencies.

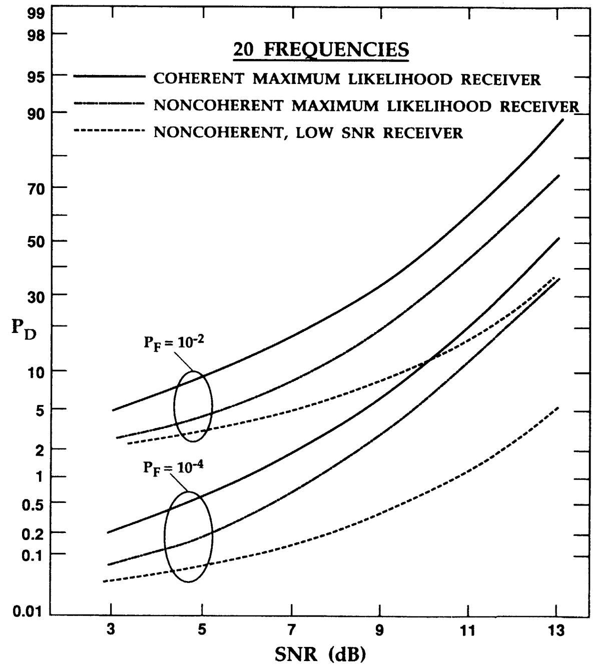


Fig. 9. Performance plots of noncoherent and coherent receivers for 20 frequencies.

As a final consideration of our moderate N results, the relative detection performances of coherent and noncoherent interception receivers are compared in Fig. 9. The comparison is for the ML case. For $P_F = 10^{-4}$, the noncoherent receiver loses between 0.9 and 2.2 dB to the coherent case. For the higher P_F, $P_F = 10^{-2}$, the loss is similar. The low SNR case in Fig. 9 is for an AL receiver based on using the first term in a series expansion of the Bessel function, $I_0(x)$. The derivation can be found in Hopkins' thesis [16] and is a well-known approximate receiver analysis. Note from Fig. 9 that the ML and AL noncoherent receiver performances approach one another as the SNR decreases.

IV. Frequency-Hopped Spread-Spectrum Considerations

A. Introduction

In the previous sections, a modest number of frequencies, N, in the discrete frequency distribution of the signal has been considered. In some frequency-hopped spread-spectrum systems, N would be in the range 10^3–10^6. Recall that N represents the processing gain of the frequency-hopped spread-spectrum system. In this section, we consider large N for both the ML and the AL receiver. We show that for a constant false alarm rate, the detection probability of both receivers goes to P_F as $N \to \infty$. For the maximum likelihood receiver, we establish N for a given SNR such that $P_D \approx P_F$. Such a processing gain implies that detections cannot be distinguished from false alarms. Hence, signal interception is impractical and a fundamental performance limit has been reached. We establish this performance limit for both the coherent and noncoherent cases. Some approximate results are also given for the AL receiver. In all cases, slow-frequency hopping is assumed as the modulation is not allowed to hop in the detection time period for signal interception.

B. Coherent Maximum Likelihood Receiver

We first note that $P_D \to P_F$, its lower bound, as $N \to \infty$. Thus, letting $N \to \infty$ effectively randomizes the signal, $A \cos \omega t$, when ω has the discrete probability density function (2). This result may be obtained as a special case ($M = 1$) of the more general MML analysis of Section V-B.

In Fig. 10 we plot P_D versus $2E/N_o = d^2$ in dB with P_F and N as parameters. The computations are based on (12)–(15). From Fig. 10, when $N = 10^6$ ($\lambda = \log_{10} N = 6$), we see that $P_D \approx P_F$ when $P_F = 10^{-6}$ for all SNR's less than 5.9 dB. Thus, below this limiting SNR, interception is impractical. For $P_F = 10^{-9}$ and $\lambda = 6$, the limiting SNR is 4.5 dB.

The dependence of P_D on P_F as N gets large is of interest. From (14) and (15),

$$P_D = \frac{1 - Q_D}{1 - Q_F} P_F + \frac{Q_D - Q_F}{1 - Q_F}. \tag{28}$$

It is established in Section V-B that $Q_D \to 0$ and $Q_F \to 0$ as $N \to \infty$ (Q_D and Q_F depend indirectly on N through η) for a constant false alarm detection probability P_F. Furthermore, using (12), (13), and an asymptotic expansion for the $Q(\cdot)$ function, it can be shown that $Q_D/Q_F \to \infty$ as $N \to \infty$. Therefore, for large N,

$$P_D \approx P_F + Q_D \tag{29}$$

or using (12)–(15),

$$P_D \approx P_F + Q\left\{Q^{-1}\left[1 - (1 - P_F)^{1/N}\right] - d\right\}. \tag{30}$$

A simple method for finding the approximate SNR, where P_D becomes close to P_F, is useful. Setting $P_D = 1.2\, P_F$ in (30) and using

$$(1 - P_F)^{1/N} = e^{(1/N)\ln(1 - P_F)} \approx 1 + \frac{\ln(1 - P_F)}{N} \tag{31}$$

yields, for large N,

$$d = \sqrt{2E/N_o} \approx Q^{-1}\left(\frac{-\ln(1 - P_F)}{N}\right) - Q^{-1}(P_F/5). \tag{32}$$

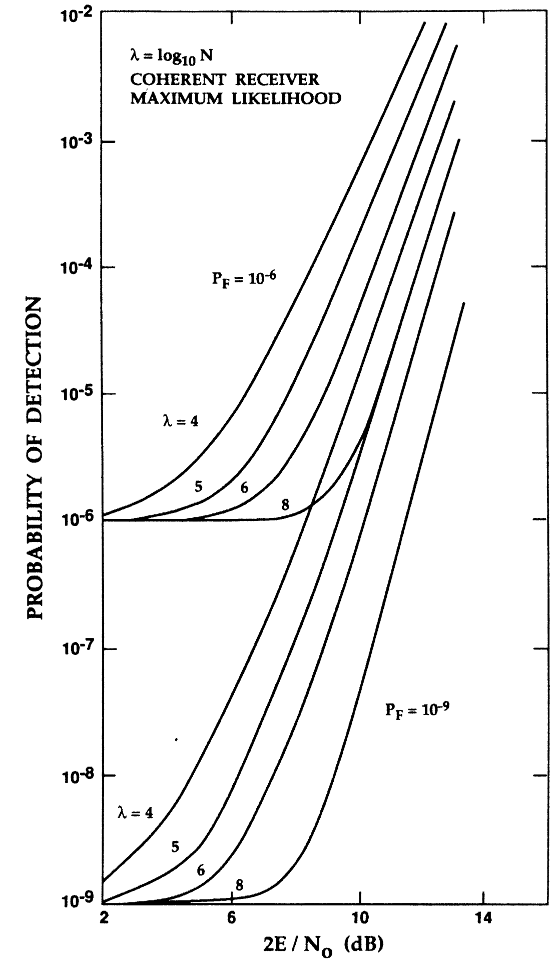


Fig. 10. Probability of detection as a function of $2E/N_o$ for $P_F = 10^{-6}$, 10^{-9} for the coherent ML receiver.

The inverse Q-function, $Q^{-1}(\cdot)$ may be conveniently computed with good accuracy for the arguments of interest here using a rational approximation given in [23, p. 933].

C. Noncoherent Maximum Likelihood Receiver

One has as a special case ($M = 1$) of the results of Section V-B that $P_D \to P_F$ as $N \to \infty$ for the single-hop noncoherent ML receiver. As in the coherent case, P_D and P_F are given by (14) and (15), respectively. Now, however, Q_D and Q_F are the detection and false alarm probability, respectively, for each individual noncoherent receiver in the parallel bank of receivers in Fig. 4.

Using (17) and (18), we plot in Fig. 11 P_D versus $2E/N_o = d^2$ in dB. The algorithm in [18] was used to compute Marcum's Q-function. The results are similar to those presented earlier in Fig. 10 for the coherent case. However, the noncoherent receiver does not perform as well as the coherent receiver. The performance is compared in Fig. 12 for two values of N and $P_F = 10^{-9}$. The SNR degradation of the noncoherent case, relative to the coherent case, decreases with N at fixed P_D. For $N = 10^5$ and $P_D = 10^{-6}$, it is approximately 0.8 dB; and for $N = 10^8$ and $P_D = 10^{-6}$, it is 0.6 dB.

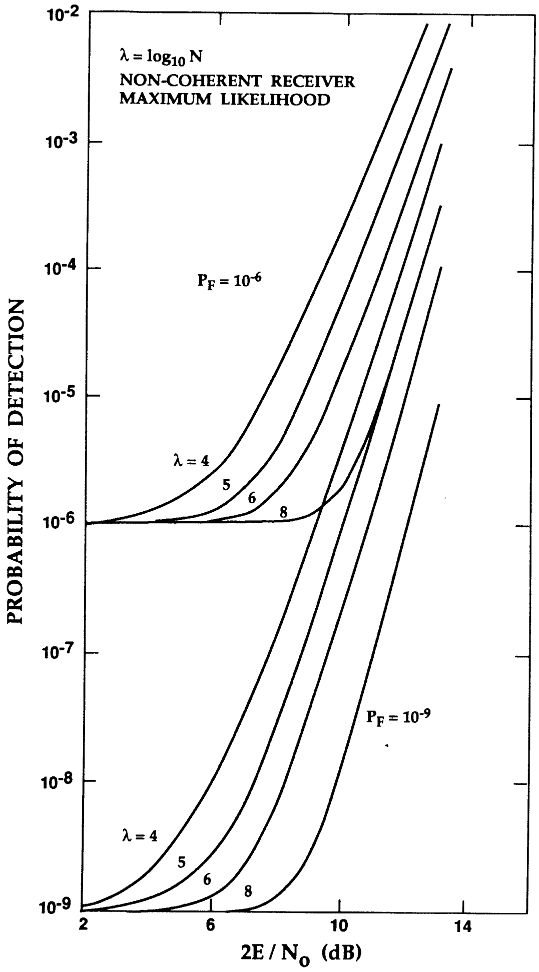


Fig. 11. Probability of detection as a function of $2E/N_o$ for $P_F = 10^{-6}$, 10^{-9} for the noncoherent ML receiver.

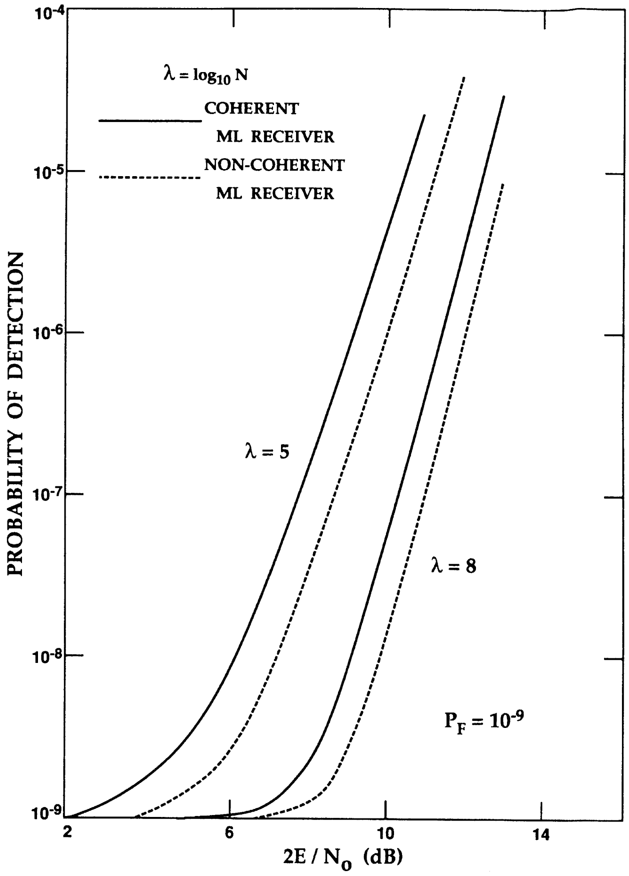


Fig. 12. Coherent versus noncoherent P_D for the ML receiver.

We also note that the SNR, where $P_D \approx P_F$, rendering interception impractical for SNR's smaller than this value, is larger in the noncoherent case. This result is as expected as the noncoherent receiver is inferior in performance to the coherent receiver. For instance, in Fig. 11 for $P_F = 10^{-9}$ and $\lambda = 6$, this SNR for the noncoherent case is 5.6 dB. In the coherent case, in Fig. 10, for the same parameters the SNR is 4.5 dB. As expected, interception will fail at a larger SNR for the noncoherent case.

Equation (28) is still valid for the present consideration. It is established in Section V-B that $Q_D \rightarrow 0$ and $Q_F \rightarrow 0$ as $N \rightarrow \infty$ for fixed P_F in the noncoherent case. Using an asymptotic expansion [19, F. 17] for the Marcum Q-function with (17) and (18), it can be shown that $Q_D/Q_F \rightarrow \infty$ as $N \rightarrow \infty$ and hence, for large N,

$$P_D \approx P_F + Q_D \tag{33}$$

or

$$P_D \approx P_F + Q_M\left\{d, \left\{-2 \ln\left[1 - (1 - P_F)^{1/N}\right]\right\}^{1/2}\right\}. \tag{34}$$

Equation (34) may also be used to find the approximate SNR where P_D becomes close to P_F for the noncoherent case. In the noncoherent case, this SNR is about 1 dB greater than that determined using (32) for the coherent case.

D. Optimum Single-Hop Coherent and Noncoherent Receivers

An exact analysis of the optimum AL receiver for large N appears intractable. In this section, some approximate analyses of coherent and noncoherent AL receivers are presented. Our treatment of the AL receivers is based on a Central Limit Theorem (CLT) [20, p. 58]. We consider the detection random variable (4), (6)

$$L = \sum_{i=1}^{N} l_i \tag{35}$$

where $l_i = e^{\alpha_i}$, and $l_i = I_o(q_i)$, in the coherent and noncoherent cases, respectively. We wish to determine

$$P_F = P(L > \eta')$$

given that H_0, the noise-only hypothesis, is active and

$$P_D = P(L > \eta')$$

given that H_1, the signal present hypothesis, is active. Let $\bar{L} = E[L]$ for notational convenience. Under certain conditions [20] which are satisfied here, the statistics

$$\frac{L_{H_0} - \bar{L}_{H_0}}{\sigma^2_{L_{H_0}}}$$

TABLE I
WLKN, CLT, AND CLT + 1 APPROXIMATIONS TO P_D FOR COHERENT OPTIMUM AL RECEIVER AND P_D FOR ML RECEIVER, SNR = 3 dB

	P_F					
$N = 10^4$	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
WLKN	2.26×10^{-2}	3.70×10^{-3}	6.09×10^{-4}	1.01×10^{-4}	1.67×10^{-5}	2.76×10^{-6}
CLT	1.09×10^{-2}	1.12×10^{-3}	1.15×10^{-4}	1.18×10^{-5}	1.21×10^{-6}	1.23×10^{-7}
ML	1.04×10^{-2}	1.08×10^{-3}	1.13×10^{-4}	1.23×10^{-5}	1.37×10^{-6}	1.60×10^{-7}
CLT + 1	1.09×10^{-2}	1.18×10^{-3}	1.47×10^{-4}	3.00×10^{-5}	1.29×10^{-5}	8.27×10^{-6}
$N = 10^5$						
WLKN	1.15×10^{-2}	1.24×10^{-3}	1.34×10^{-4}	1.43×10^{-5}	1.54×10^{-6}	1.65×10^{-7}
CLT	1.02×10^{-2}	1.03×10^{-3}	1.04×10^{-4}	1.04×10^{-5}	1.05×10^{-6}	1.05×10^{-7}
ML	1.01×10^{-2}	1.01×10^{-3}	1.02×10^{-4}	1.04×10^{-5}	1.06×10^{-6}	1.09×10^{-7}
CLT + 1	1.02×10^{-2}	1.03×10^{-3}	1.05×10^{-4}	1.09×10^{-5}	1.30×10^{-6}	2.65×10^{-7}
$N = 10^6$						
WLKN	1.03×10^{-2}	1.04×10^{-3}	1.05×10^{-4}	1.06×10^{-5}	1.07×10^{-6}	1.08×10^{-7}
CLT	1.01×10^{-2}	1.01×10^{-3}	1.01×10^{-4}	1.01×10^{-5}	1.01×10^{-6}	1.02×10^{-7}
ML	1.00×10^{-2}	1.00×10^{-3}	1.00×10^{-4}	1.01×10^{-5}	1.01×10^{-6}	1.01×10^{-7}
CLT + 1	1.01×10^{-2}	1.01×10^{-3}	1.01×10^{-4}	1.01×10^{-5}	1.02×10^{-6}	1.03×10^{-7}
$N = 10^7$						
WLKN	1.01×10^{-2}	1.01×10^{-3}	1.01×10^{-4}	1.01×10^{-5}	1.02×10^{-6}	1.02×10^{-7}
CLT	1.00×10^{-2}	1.00×10^{-3}	1.00×10^{-4}	1.00×10^{-5}	1.00×10^{-6}	1.01×10^{-7}
ML	1.00×10^{-2}	1.00×10^{-3}	1.00×10^{-4}	1.00×10^{-5}	1.00×10^{-6}	1.00×10^{-7}
CLT + 1	1.00×10^{-2}	1.00×10^{-3}	1.00×10^{-4}	1.00×10^{-5}	1.00×10^{-6}	1.01×10^{-7}

and

$$\frac{L_{H_1} - \bar{L}_{H_1}}{\sigma^2_{L_{H_1}}}$$

are asymptotically normal in the sense that their cumulative distributions approach the zero-mean, unit variance normal distribution. Thus, for large N,

$$P_F \approx Q\left(\frac{\eta' - \bar{L}_{H_0}}{\sigma_{L_{H_0}}}\right) \tag{36}$$

where

$$L_{H_0} = \sum_{i=1}^{N} l_{i0} \tag{37}$$

and

$$P_D \approx Q\left(\frac{\eta' - \bar{L}_{H_1}}{\sigma_{L_{H_1}}}\right) \tag{38}$$

where

$$L_{H_1} = \sum_{i=1}^{N} l_{i1}. \tag{39}$$

Combining (36) with (38) gives

$$P_D \approx Q\left\{\frac{\sigma_{L_{H_0}}}{\sigma_{L_{H_1}}} Q^{-1}(P_F) + \frac{\bar{L}_{H_0} - \bar{L}_{H_1}}{\sigma_{L_{H_1}}}\right\}. \tag{40}$$

Equation (40) can be used as an approximation to P_D for large N. We call this the CLT approximation. The parameters $\sigma_{L_{H_0}}$, $\sigma_{L_{H_1}}$, $\bar{L}_{H_0}$, and $\bar{L}_{H_1}$ will be specified in terms of the SNR, d^2, so that P_D is a function of d and P_F.

Consider first the coherent case where $l_i = e^{\alpha_i}$. Under H_0, the α_i are normal with zero mean and variance d^2, $d^2 = 2E/N_o$, i.e., $\eta(0, d^2)$. Under H_1, $N - 1$ of the α_i's are $\eta(0, d^2)$ and one, say α_k, is $\eta(d^2, d^2)$. It can be shown that [24]

$$\bar{l}_{i0} = e^{d^2/2} \tag{41a}$$

$$\sigma^2_{l_{i0}} = e^{d^2}(e^{d^2} - 1) \tag{41b}$$

$$\bar{l}_{i1} = e^{d^2/2}, \qquad i \neq k \tag{42a}$$

$$\bar{l}_{k1} = e^{3d^2/2} \tag{42b}$$

$$\sigma^2_{l_{i1}} = e^{d^2}(e^{d^2} - 1), \qquad i \neq k \tag{42c}$$

$$\sigma^2_{l_{k1}} = e^{3d^2}(e^{d^2} - 1). \tag{42d}$$

Combination of (41) and (42) with (37), (39), and (40) yields

$$P_D \approx Q\left\{Q^{-1}(P_F)\sqrt{\frac{N}{N + e^{2d^2} - 1}} - \sqrt{\frac{e^{d^2} - 1}{N + e^{2d^2} - 1}}\right\}. \tag{43}$$

TABLE II
WLKN, CLT, AND CLT + 1 APPROXIMATIONS TO P_D FOR COHERENT OPTIMUM AL RECEIVER AND P_D FOR ML RECEIVER, SNR = 13 dB

	P_F					
$N = 10^4$	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
WLKN	1×10^{-0}	9×10^{-1}	8×10^{-1}	7×10^{-1}	5×10^{-1}	3×10^{-1}
CLT	5×10^{-1}	5×10^{-1}	5×10^{-1}	5×10^{-1}	5×10^{-1}	5×10^{-1}
ML	4×10^{-1}	2×10^{-1}	1×10^{-1}	6×10^{-2}	3×10^{-2}	1×10^{-2}
CLT + 1	1×10^{-1}	1×10^{-1}	9×10^{-2}	9×10^{-2}	8×10^{-2}	8×10^{-2}
$N = 10^5$						
WLKN	9×10^{-1}	8×10^{-1}	7×10^{-1}	5×10^{-1}	3×10^{-1}	2×10^{-1}
CLT	5×10^{-1}	5×10^{-1}	5×10^{-1}	5×10^{-1}	5×10^{-1}	5×10^{-1}
ML	2×10^{-1}	1×10^{-1}	6×10^{-2}	3×10^{-2}	1×10^{-2}	5×10^{-3}
CLT + 1	8×10^{-2}	7×10^{-2}	6×10^{-2}	5×10^{-2}	5×10^{-2}	5×10^{-2}
$N = 10^6$						
WLKN	9×10^{-1}	7×10^{-1}	5×10^{-1}	3×10^{-1}	2×10^{-1}	1×10^{-1}
CLT	5×10^{-1}	5×10^{-1}	5×10^{-1}	5×10^{-1}	5×10^{-1}	5×10^{-1}
ML	1×10^{-1}	6×10^{-2}	3×10^{-2}	1×10^{-2}	5×10^{-3}	2×10^{-3}
CLT + 1	5×10^{-2}	4×10^{-2}	3×10^{-2}	3×10^{-2}	3×10^{-2}	3×10^{-2}
$N = 10^7$						
WLKN	8×10^{-1}	5×10^{-1}	4×10^{-1}	2×10^{-1}	1×10^{-1}	6×10^{-2}
CLT	5×10^{-1}	5×10^{-1}	5×10^{-1}	5×10^{-1}	5×10^{-1}	5×10^{-1}
ML	7×10^{-2}	3×10^{-2}	1×10^{-2}	5×10^{-3}	2×10^{-3}	7×10^{-4}
CLT + 1	4×10^{-2}	2×10^{-2}	2×10^{-2}	2×10^{-2}	2×10^{-2}	2×10^{-2}
$N = 10^8$						
WLKN	6×10^{-1}	4×10^{-1}	2×10^{-1}	1×10^{-1}	6×10^{-2}	3×10^{-2}
CLT	5×10^{-1}	5×10^{-1}	5×10^{-1}	5×10^{-1}	5×10^{-1}	5×10^{-1}
ML	4×10^{-2}	1×10^{-2}	5×10^{-3}	2×10^{-3}	7×10^{-4}	3×10^{-4}
CLT + 1	2×10^{-2}	1×10^{-2}	1×10^{-2}	9×10^{-3}	8×10^{-3}	8×10^{-3}

Table I shows that the CLT approximation to the coherent AL receiver performance is very close to the ML receiver performance when the SNR is 3 dB. Table II, however, shows that the CLT approximation is useless for 13 dB. A better approximation to P_D for larger SNR can be obtained by using the Gaussian distribution for α_k in (35) and the limit distribution, also Gaussian, for $\Sigma_{\substack{i=1\\ i\neq k}}^{N}$ in (35). We call this the CLT + 1 approximation. The mathematical details are given in the Appendix. Data from these approximations are given in Tables I and II. One sees that the estimated P_D using the CLT + 1 approximation is close to P_D for the ML receiver for P_F's greater than 10^{-3}. The CLT + 1 estimated P_D is within a factor of about 4 of P_D for the ML receiver for P_F between 10^{-3} and 10^{-5}. Also shown in Tables I and II are estimates of P_D obtained using Wilkinson's (WLKN) approximation. That is, the sum of lognormals is assumed to have a lognormal distribution. Note that for small SNR (3 dB), the CLT and CLT + 1 approximations are almost always closer to the ML performance than the lognormal approximation. For larger SNR (13 dB), the CLT + 1 approximation is always closer to the ML performance than the WLKN estimate. While the CLT + 1 estimate is closest to P_D of the ML receiver at 13 dB, none of the three estimates is close to the ML P_D for $P_F < 10^{-6}$. Recall that the ML receiver analysis gives the exact relationship among Q_F, Q_D, P_F, and P_D for the suboptimum ML receiver. It provides, therefore, a valid measure of performance for this receiver at all values of P_F. Considering the results of Section III-A and B and Figs. 5–8 as well as Tables I and II, one is tempted to conclude that, for large N, the AL (optimum) receiver performs nearly the same as the ML receiver. Our approximate analysis is better at low SNR than large SNR. This follows since a large SNR makes one term in (35) dominant, a situation that, as [21, p. 194] points out, is poor for CLT approximations. Finally, note that letting $N \to \infty$ in (43) for fixed SNR (d^2) yields $P_D \to P_F$.

We now consider the noncoherent case where $l_i = I_o(q_i)$ in (35). When $r(t) = n(t)$ in (6), i.e., the H_0 hypothesis, q_i is Rayleigh distributed. When the signal present hypothesis is active, $N - 1$, q_i's are Rayleigh distributed and one, say q_k, has a Rician distribution. To proceed with a CLT analysis, let $l_{i0} = I_o(q_{i0})$, where q_{i0} is the

TABLE III
CLT + 1 APPROXIMATIONS TO P_D FOR NONCOHERENT OPTIMUM AL RECEIVER AND P_D FOR ML RECEIVER, SNR = 3 dB

	P_F					
$N = 10^4$	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
ML	1.01×10^{-2}	1.02×10^{-2}	1.03×10^{-4}	1.06×10^{-5}	1.09×10^{-6}	1.13×10^{-7}
CLT + 1	1.03×10^{-2}	1.05×10^{-3}	1.07×10^{-4}	1.15×10^{-5}	1.53×10^{-6}	3.91×10^{-7}
$N = 10^5$						
ML	1.00×10^{-2}	1.00×10^{-3}	1.01×10^{-4}	1.01×10^{-5}	1.01×10^{-6}	1.02×10^{-7}
CLT + 1	1.00×10^{-2}	1.01×10^{-3}	1.02×10^{-4}	1.02×10^{-5}	1.03×10^{-6}	1.05×10^{-7}
$N = 10^6$						
ML	1.00×10^{-2}	1.00×10^{-3}	1.00×10^{-4}	1.00×10^{-5}	1.00×10^{-6}	1.00×10^{-7}
CLT + 1	1.00×10^{-2}	1.00×10^{-3}	1.00×10^{-4}	1.01×10^{-5}	1.01×10^{-6}	1.01×10^{-7}
$N = 10^7$						
ML	1.00×10^{-2}	1.00×10^{-3}	1.00×10^{-4}	1.00×10^{-5}	1.00×10^{-6}	1.00×10^{-7}
CLT + 1	1.00×10^{-2}	1.00×10^{-3}	1.00×10^{-4}	1.00×10^{-5}	1.00×10^{-6}	1.00×10^{-7}

TABLE IV
CLT + 1 APPROXIMATIONS TO P_D FOR NONCOHERENT OPTIMUM AL RECEIVER AND P_D FOR ML RECEIVER, SNR = 13 dB

	P_F					
$N = 10^4$	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
ML	2.54×10^{-1}	1.34×10^{-1}	6.60×10^{-2}	3.02×10^{-2}	1.29×10^{-2}	5.20×10^{-3}
CLT + 1	9.16×10^{-2}	7.16×10^{-2}	6.42×10^{-2}	5.99×10^{-2}	5.67×10^{-2}	5.42×10^{-2}
$N = 10^5$						
ML	1.42×10^{-1}	6.67×10^{-2}	3.03×10^{-2}	1.29×10^{-2}	5.20×10^{-3}	1.98×10^{-3}
CLT + 1	5.88×10^{-2}	4.24×10^{-2}	3.73×10^{-2}	3.44×10^{-2}	3.24×10^{-2}	3.08×10^{-2}
$N = 10^6$						
ML	7.54×10^{-2}	3.12×10^{-2}	1.30×10^{-2}	5.21×10^{-3}	1.98×10^{-3}	7.21×10^{-4}
CLT + 1	3.76×10^{-2}	2.39×10^{-2}	2.03×10^{-2}	1.86×10^{-2}	1.74×10^{-2}	1.64×10^{-2}
$N = 10^7$						
ML	4.00×10^{-2}	1.39×10^{-2}	5.30×10^{-3}	1.99×10^{-3}	7.21×10^{-4}	2.51×10^{-4}
CLT + 1	2.47×10^{-2}	1.29×10^{-2}	1.05×10^{-3}	9.47×10^{-3}	8.79×10^{-3}	8.27×10^{-3}
$N = 10^8$						
ML	2.28×10^{-2}	6.20×10^{-3}	2.08×10^{-3}	7.31×10^{-4}	2.52×10^{-4}	8.39×10^{-5}
CLT + 1	1.74×10^{-2}	6.83×10^{-3}	5.12×10^{-3}	4.53×10^{-3}	4.17×10^{-3}	3.90×10^{-3}
$N = 10^9$						
ML	1.52×10^{-2}	2.98×10^{-4}	8.21×10^{-4}	2.61×10^{-4}	8.48×10^{-5}	2.71×10^{-5}
CLT + 1	1.35×10^{-2}	3.68×10^{-3}	2.37×10^{-3}	$2.04 \times 10\text{-}3$	1.86×10^{-3}	1.73×10^{-3}

value of q_i under H_0. Similarly, $l_{i1} = I_o(q_{i1})$ for hypothesis H_1. It can be shown that [22], [24]

$$\bar{l}_{i0} = e^{d^2/2} \tag{44a}$$

$$\sigma_{l_{i0}}^2 = e^{d^2}\left(I_o(d^2) - 1\right) \tag{44b}$$

$$\bar{l}_{i1} = e^{d^2/2}, \qquad i \neq k \tag{45a}$$

$$\bar{l}_{k1} = e^{d^2/2} I_o(d^2) \tag{45b}$$

$$\sigma_{l_{i1}}^2 = e^{d^2}\left(I_o(d^2) - 1\right), \qquad i \neq k \tag{45c}$$

$$\sigma_{l_{k1}}^2 < \infty \text{ for } d^2 < \infty. \tag{45d}$$

Using (44) and (45) with (37), (39), and (40) yields that $P_D \rightarrow P_F$ as $N \rightarrow \infty$. A noncoherent CLT approximation analogous to (43) for the coherent case has not been pursued since the second moment of $I_o(q_k)$ is unknown.

We have computed the CLT + 1 approximation for the noncoherent case [this approximation for the coherent case was described following (43)]. To describe the CLT + 1 approximation for the noncoherent case, consider (39),

$$L_{H_1} = \sum_{\substack{i=1 \\ i \neq k}}^{N} I_o(q_{i1}) + I_o(q_{k1}) \tag{46}$$

where the q_{i1} are Rayleigh and q_{k1} is Rician. The first term in (46) is assumed Gaussian with the mean and variance computed using (44). Then P_D can be computed numerically using a Rician pdf for q_{k1} [21, p. 139]. More details can be found in the Appendix. The results of the computation are given in Tables III and IV. Over a considerable range of parameters, the CLT + 1 approximation gives results for P_D that exceed the ML case. The comparative results between the AL and ML cases are similar to the coherent case.

V. The Multiple-Hop AL and ML Receivers

In this and subsequent sections, the signal is assumed to change frequency a number of times, M, per detection interval. This may be the case where the detection interval is the same as the symbol interval and the signal hops M times per symbol. It may also be the case where the signal hops once per symbol interval and the receiver integrates over M symbol intervals before making a decision. In this way, an interceptor concerned only with the presence or absence of the signal may gain an advantage over a friendly receiver. In general, let M_s denote the number of symbol periods, and M_H the number of hops per symbol interval. Then in one detection interval, the signal is permitted to hop $M = M_s \cdot M_H$ times. That is, the MML receiver may integrate over M_s symbols a signal that hops M_H times per symbol.

Again, the SNR, $d^2 = 2E/N_o$. However, now it must be remembered that the energy E is to be interpreted as the signal energy per detection interval. Therefore, $E = M_s \cdot E_s$, where E_s is the energy per symbol.

A. Interception Receiver Structures

We again consider the binary hypothesis problem, (1) but now

$$s(t) = \sum_{j=1}^{M} s^j(t) \tag{47a}$$

where

$$s^j(t) = \begin{cases} A\cos(\omega t + \theta), & (j-1)T/M \le t \le jT/M \\ 0, & \text{elsewhere.} \end{cases} \tag{47b}$$

The angular frequency ω can assume one of N values ω_i, $i = 1, \cdots, N$, in each hop interval of width T/M. That is, there are N hopping frequencies for each chip. The discrete distribution of ω is again given by (2). Now, however, the frequencies are spaced multiples of $1/MT$ or $1/2MT$ apart for noncoherent and coherent systems, respectively. Let $\bar{\omega} = (\omega^1, \cdots, \omega^M)$ be the random vector whose component ω^j, $j = 1, \cdots, M$ is a random variable that represents the angular frequency during the jth hop. Then $\bar{\omega}$ can assume one of N^M values under hypothesis H_1. Each outcome of $\bar{\omega}$ is assumed to be equally likely. Similarly, let $\bar{\theta} = (\theta^1, \cdots, \theta^M)$ be a random vector whose component θ^j is a random variable that represents the phase constant for the jth hop. In the coherent case, θ^j is known, whereas in the noncoherent case, it is uniformly distributed on $[0, 2\pi]$. The components of $\bar{\theta}$ and $\bar{\omega}$ are assumed mutually independent and are also independent of the noise $n(t)$. As previously, $n(t)$ is additive white and Gaussian, having spectral density $N_o/2$ W/Hz and A is the signal amplitude.

The likelihood ratio is

$$\begin{aligned} L[r(t)] &= E_{\bar{\omega},\bar{\theta}}\left\{\exp\left(\frac{2}{N_o}\int_0^T r(t)s(t)\,dt - \frac{E}{N_o}\right)\right\} \\ &= E_{\bar{\omega},\bar{\theta}}\left\{\prod_{j=1}^{M}\exp\left(\frac{2}{N_o}\int_{(j-1)(T/M)}^{j(T/M)} r(t)s^j(t)\,dt\right)\right. \\ &\qquad \left.\cdot \exp\left(\frac{-E}{N_o}\right)\right\} \\ &= \exp\left(-\frac{E}{N_o}\right)\prod_{j=1}^{M} E_{\omega^j,\theta^j} \\ &\qquad \cdot \left\{\exp\left(\frac{2}{N_o}\int_{(j-1)(T/M)}^{j(T/M)} r(t)s^j(t)\,dt\right)\right\} \\ &= \exp\left(-\frac{E}{N_o}\right)\prod_{j=1}^{M} \\ &\qquad \cdot E_{\theta^j}\left\{\frac{1}{N}\sum_{i=1}^{N}\exp\left(\frac{2A}{N_o}\int_{(j-1)(T/M)}^{j(T/M)} r(t)\right.\right. \\ &\qquad \left.\left. \cdot \cos(\omega_i t + \theta^j)\,dt\right)\right\}. \end{aligned} \tag{48}$$

The logarithm likelihood ratio is thus

$$\begin{aligned} \ln\{L[r(t)]\} &= \sum_{j=1}^{M}\ln\left\{E_{\theta^j}\left\{\frac{1}{N}\sum_{i=1}^{N}\exp\left(\frac{2A}{N_o}\int_{(j-1)(T/M)}^{j(T/M)} r(t)\right.\right.\right. \\ &\qquad \left.\left.\left.\cdot \cos(\omega_i t + \theta^j)\,dt\right)\right\}\right\} - \frac{E}{N_o}. \end{aligned} \tag{49}$$

For the coherent case, the log AL ratio test is

$$\begin{aligned} \ln\{L\} &= \sum_{j=1}^{M}\ln\left\{\sum_{i=1}^{N}\exp\left(\frac{2A}{N_o}\int_{(j-1)(T/M)}^{j(T/M)} r(t)\right.\right. \\ &\qquad \left.\left.\cdot \cos(\omega_i t + \theta^j)\,dt\right)\right\} \underset{H_0}{\overset{H_1}{\gtrless}} \ln\eta + \frac{E}{N_o} \\ &\qquad + M\ln N = \eta', \end{aligned}$$

which may be written as

$$\ln\{L\} = \sum_{j=1}^{M}\ln\left\{\sum_{i=1}^{N} l_i^j\right\} \underset{H_0}{\overset{H_1}{\gtrless}} \eta' \tag{50}$$

where

$$l_i^j = e^{\alpha_i^j} \tag{51}$$

and

$$\alpha_i^j = \frac{2A}{N_o} \int_{(j-1)(T/M)}^{j(T/M)} r(t) \cos(\omega_i t + \theta^j)\, dt. \qquad (52)$$

The constant η sets the false alarm probability, and η' is then the detection threshold. The receiver statistic is a logarithmic sum of sums of lognormal random variables $e^{\alpha_i^j}$. Note the relation of the optimum coherent multiple-hop average likelihood (CMAL) receiver to the optimum coherent single-hop average likelihood (CAL) receiver. Conceptually, the optimum multihop receiver may be implemented by adding the logarithms of the outputs of M receivers. Each of these is the optimum AL receiver for detecting the presence of a single signal chip. This may be implemented using one AL receiver that is used M times, each consecutive time corresponding to one hop.

For the noncoherent case, performing the expections in (49), the log AL ratio test is

$$\ln\{L\} = \sum_{j=1}^{M} \ln\left\{\sum_{i=1}^{N} I_o(q_i^j)\right\} \underset{H_0}{\overset{H_1}{\gtrless}} \eta + \frac{E}{N_o} + M \ln N = \eta' \qquad (53)$$

where q_i^j is given by

$$q_i^j = (L_{c_i}^j)^2 + (L_{s_i}^j)^2 \qquad (54)$$

$$L_{c_i}^j = \frac{2A}{N_o} \int_{(j-1)(T/M)}^{j(T/M)} r(t) \cos \omega_i t\, dt \qquad (55)$$

$$L_{s_i}^j = \frac{2A}{N_o} \int_{(j-1)(T/M)}^{j(T/M)} r(t) \sin \omega_i t\, dt. \qquad (56)$$

The noncoherent multiple-hop AL (NCMAL) receiver has the same relationship to the noncoherent single-hop AL (NCAL) receiver as does the CMAL receiver to the CML receiver. That is, the NCMAL receiver is a logarithmic sum of M consecutive outputs of an NCAL chip receiver.

We have previously noted the similar performances of the single-hop AL and ML receivers. Motivated by this similarity, and by the forms of (50) and (53), we propose the following suboptimum statistic for detecting multihop FH signals:

$$L = \sum_{j=1}^{M} l^j \underset{H_0}{\overset{H_1}{\gtrless}} T_o. \qquad (57)$$

In (57), $l^j = 0, 1$ according to whether signal is declared absent or present during the jth hop by an ML receiver for the jth hop. That is, $l^j = 0, 1$ according to whether an ML receiver for the jth hop declares H_0 or H_1, respectively. Observe that the multihop statistic (57) has an extra parameter, T_o. Thus, there are M hops, and the multiple-hop ML (MML) receiver declares H_0 if signal is declared present for fewer than T_o hops. If signal is detected in T_o or more of the hops, the MML receiver declares H_1. The issue of choosing the value of T_o is treated in the next section.[2]

In the coherent case, each chip receiver is a CML receiver. Let

$$\beta_i^j = A \int_{(j-1)(T/M)}^{j(T/M)} r(t) \cos(\omega_i t + \theta^j)\, dt, \qquad (58)$$

then the random variable l_i^j assumes the values 0, 1 according to

$$\beta_i^j \underset{l_i^j=0}{\overset{l_i^j=1}{\gtrless}} \frac{N_o}{2} \ln \eta + \frac{E}{2M} = \gamma_M$$

and

$$l^j = \max_{i=1,\cdots,N} \{l_i^j\}. \qquad (59)$$

The probability of detection for each hop detector is

$$Q_D = Q\left(\frac{\ln \eta}{d_M} - \frac{d_M}{2}\right) \qquad (60)$$

where $d_M^2 = d^2/M = 2E/MN_o$. The probability of hop false alarm is

$$Q_F = Q\left(\frac{\ln \eta}{d_M} + \frac{d_M}{2}\right). \qquad (61)$$

In the noncoherent case, each hop receiver is an NCML receiver. One has

$$q_i^j \underset{l_i^j=0}{\overset{l_i^j=1}{\gtrless}} I_0^{-1}(\eta e^{d_M^2/2}) = \gamma_M \qquad (62)$$

and, again

$$l^j = \max_{i=1,\cdots,N} \{l_i^j\}. \qquad (63)$$

The noncoherent hop probabilities of detection and false alarm are

$$Q_D = Q_M(d_M, \gamma_M/d_M) \qquad (64)$$

and

$$Q_F = \exp\left[-\gamma_M^2/2d_M^2\right], \qquad (65)$$

respectively.

The MML receiver detection and false alarm probabilities are, respectively,

$$P_D = 1 - \sum_{j=0}^{T_o-1} \binom{M}{j} \left\{1 - (1 - Q_D)(1 - Q_F)^{N-1}\right\}^j \cdot \left\{(1 - Q_D)(1 - Q_F)^{N-1}\right\}^{M-j}, \qquad (66)$$

$$P_F = 1 - \sum_{j=0}^{T_o-1} \binom{M}{j} \left\{1 - (1 - Q_F)^N\right\}^j \cdot \left\{(1 - Q_F)^N\right\}^{M-j}. \qquad (67)$$

[2] A reviewer has pointed out that the MML receiver structure is similar to the channelized radiometer described in [12, p. 136].

In the next subsections, the performances of the CMML and NCMML receivers will be examined. The MAL receivers are not analyzed. However, the performances of the MML receivers may be viewed as lower bounds of the performances of the corresponding MAL receivers.

B. Performance of Coherent and Noncoherent MML Receivers

We begin our examination of the CMML receiver performance by showing that $P_D \rightarrow P_F$ as $N \rightarrow \infty$, regardless of the value of T_o. In order to show this, we first show that for a fixed P_F, $Q_D \rightarrow 0$ as $N \rightarrow \infty$. Equations (60) and (61) give

$$Q_D = Q(Q^{-1}(Q_F) - d_M) \tag{68}$$

so that $Q_D \rightarrow 0$ as $N \rightarrow \infty$ if $Q_F \rightarrow 0$ as $N \rightarrow \infty$. The latter is proved by contradiction. Consider $0 < P_F < 1$ and assume that $\lim_{N \rightarrow \infty} Q_F \neq 0$. Then (67) leads to

$$\lim_{N \rightarrow \infty} P_F = P_F = 1$$

which contradicts the original assumption that $P_F < 1$. Hence, $\lim_{N \rightarrow \infty} Q_F = 0$ and $Q_D \rightarrow 0$ as $N \rightarrow \infty$. Interpret now (66) and (67) in the following manner:

$$P_D = 1 - \sum_{j=0}^{T_o - 1} \binom{M}{j} \left\{ 1 - \frac{1 - Q_D}{1 - Q_F} (1 - Q_F)^N \right\}^j \cdot \left\{ \frac{1 - Q_D}{1 - Q_F} (1 - Q_F)^N \right\}^{M-j},$$

$$P_F = 1 - \sum_{j=0}^{T_o - 1} \binom{M}{j} \left\{ 1 - (1 - Q_F)^N \right\}^j \cdot \left\{ (1 - Q_F)^N \right\}^{M-j}.$$

Since $Q_D \rightarrow 0$ and $Q_F \rightarrow 0$ as $N \rightarrow \infty$, $P_D \rightarrow P_F$ as $N \rightarrow \infty$ for any choice of T_o.

It can also be shown that $P_D \rightarrow P_F$ as $N \rightarrow \infty$ for the NCMML receiver. The proof is similar to that for the CMML receiver, except that Q_D and Q_F are given by (64) and (65) rather than (60) and (61). Again, this is true for any value of T_o.

The intuitive tradeoff involved in the choice of T_o is the following. The greater the value of T_o, the smaller the likelihood of a false alarm, since more hop receivers must erroneously declare signal present. At the same time, the smaller the probability of detection, since T_o or more hop receivers must declare signal present. That is, a greater T_o leads to a smaller P_F, but at the same time leads to smaller P_D. One wants to select T_o to achieve the greatest P_D for a given P_F.

Fig. 13 for the CMML receiver shows the ratio P_D/P_F plotted as a function of P_F for $M = 3$, $T_o = 1 - 3$, and $N = 10\,000$. Analogous results for the NCMML receiver for $M = 3$ and SNR = 13 dB are presented in Fig. 14. Interestingly, the best performance is achieved for $T_o = 1$. We have examined a number of cases for $M \leq 10$, and $T_o = 1$ was the best choice every time [24].

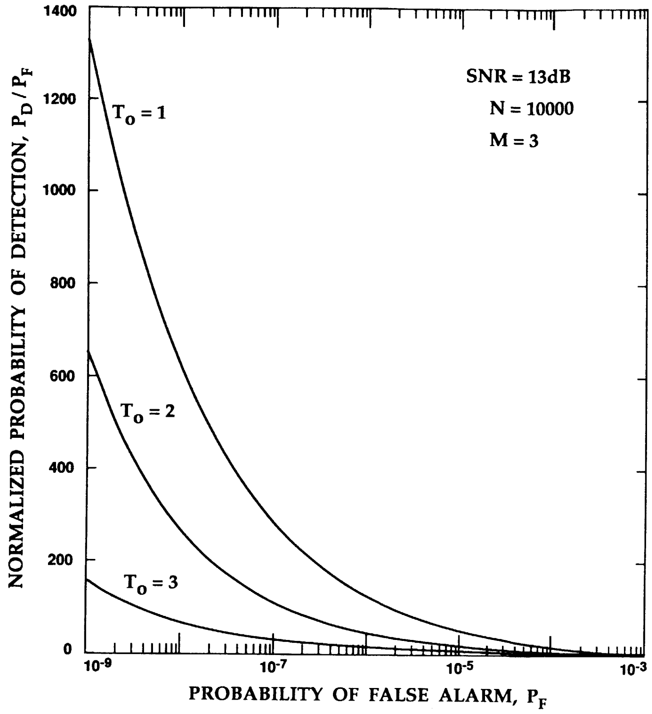


Fig. 13. The normalized detection probability P_D/P_F as a function of P_F for the coherent MML receiver with $M = 3$.

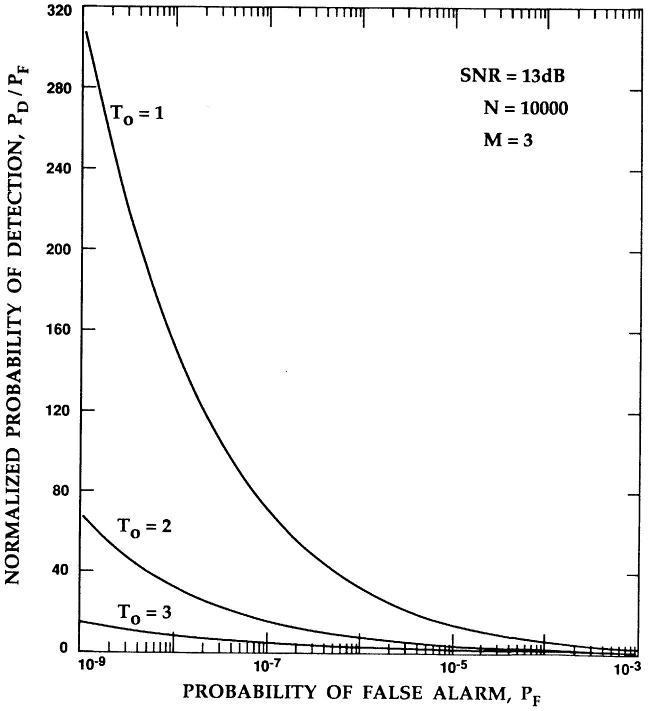


Fig. 14. The normalized detection probability P_D/P_F as a function of P_F for the noncoherent MML receiver with $M = 3$.

In the next two subsections, the results of Sections V-A and B are interpreted from the point of view of a signal that hops M times per symbol, and from the point of view of a receiver that integrates over several symbols. In all cases, $T_o = 1$ is used.

C. Multiple Hops per Symbol

The spread-spectrum user may elect to hop several times per symbol to make the jamming of each symbol more difficult. That is, in some cases the symbol may be recovered because the jamming has only wiped out a part of the symbol. Hopping several times during the duration of a symbol is also desirable from the point of view of defeating an unauthorized interceptor. This is shown in Figs. 15 and 16, where the ratio of the detection probability using $(N, M) = (10^4, M)$ to the detection probability using $(N, M) = (M \cdot 10^4, 1)$ is plotted as a function of P_F at SNR = 13 dB. The coherent case is shown in Fig. 15, and the noncoherent case in Fig. 16. Observe, for example, that hopping 10 times per bit results in a P_D that is 4.4×10^{-6} times the P_D that results from hopping once per bit over 10 times as many frequencies. This is for $P_F = 10^{-9}$, SNR = 13 dB, and $N = 10\,000$ in the coherent case. The noncoherent results are similar. For example, P_D for $M = 10$, SNR = 13 dB, $N = 10\,000$, and $P_F = 10^{-9}$ is 5.8×10^{-6} times the P_D obtained hopping once over 10 times as many frequencies. The comparisons are made on the basis of equal received energy over one detection interval of duration T. That is, the transmitter may, for the same bandwidth, put energy E in one of $M \cdot 10^4$ frequencies for the bit duration T. Alternatively, energy E/M is put in one of 10^4 frequencies for a duration of T/M corresponding to hopping M times per bit. These results are obtained from (66) and (67) by setting $T_o = 1$. One also notes from observation of these figures that P_D approaches P_F very rapidly as M is increased.

Corresponding to (29) and (33), one can show that for large N and M hops/symbol,

$$P_D(M\text{ hops/symbol}) = P_F + MQ_D \qquad (69)$$

for both coherent and noncoherent cases. Comparison of (69) with (29) and (33) shows why hopping M times per symbol is so much more effective in reducing P_D to P_F than hopping M times as many frequencies once per symbol. In (69), Q_D is given by a Q-function of an appropriate argument. The energy per hop enters into the argument. One can show that for a given P_F, Q_D in (69) is proportional to exp $(\text{constant})/M^{1/2})$. The details are lengthy and are omitted. This effect dominates the scaling of Q_D by M in (69). Note, for example, that hopping 10 times per symbol at $P_F = 10^{-6}$ and SNR = 13 dB reduces P_D from 2.9×10^{-2} to 1.6×10^{-6} when $N = 10\,000$. Hopping once per symbol and increasing N from 10 000 to 10^{10} reduces P_D to 8.7×10^{-5}.

In analogy to (32) for the single-hop ML receiver, one may also derive an expression for the SNR at which P_D becomes approximately equal to P_F for the M hop/symbol case. Setting $P_D = 1.2P_F$ in (69), $T_o = 1$ in (67), and using (60) and (31) gives, for large N,

$$d = \sqrt{2E/N_o} \approx \sqrt{M}\left[Q^{-1}\left(-\frac{\ln(1-P_F)}{NM}\right) - Q^{-1}(P_F/5M)\right]. \qquad (70)$$

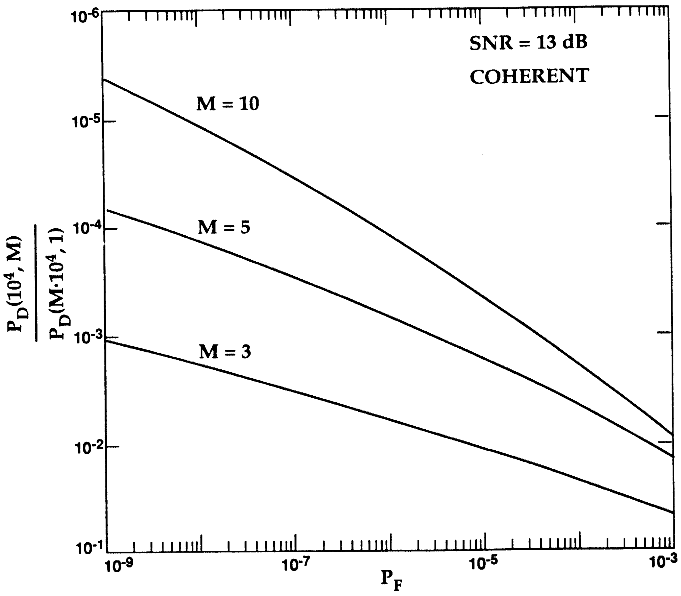


Fig. 15. The ratio of P_D for a coherent MML receiver with $(N, M) = (10^4, M)$ to P_D for a coherent ML receiver with $(N, M) = (M \cdot 10^4, 1)$.

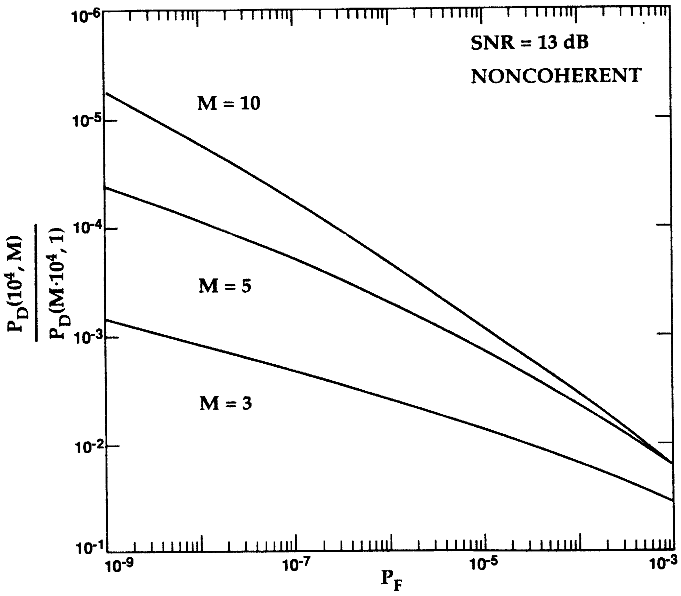


Fig. 16. The ratio of P_D for a noncoherent MML receiver with $(N, M) = (10^4, M)$ to P_D for a noncoherent ML receiver with $(N, M) = (M \cdot 10^4, 1)$.

The performances of the coherent and noncoherent MML receivers are compared in Fig. 17. The comparison is made for $N = 10\,000$, $P_F = 10^{-6}$, and $M = 3$ and 5. The noncoherent receiver performs ~ 1 dB poorer than the coherent receiver at high SNR. At lower SNR, the difference is ~2 dB.

D. Integration Over Several Symbols

The interceptor may gain an advantage by integrating over M_s symbols before deciding whether signal is present. In this subsection, the case where the transmitter hops M_H times per symbol and the MML receiver integrates

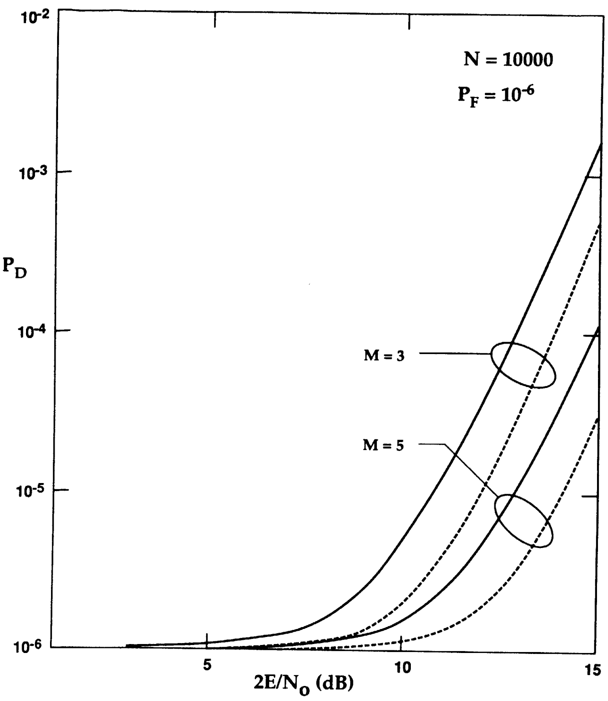


Fig. 17. The probability of detection P_D as a function of SNR for the coherent and noncoherent MML receivers. The false alarm probability P_F is 10^{-6} and the noncoherent case is represented by a broken line.

over M_s symbol periods is examined. In this case, the energy E per detection interval and symbol energy E_s are related by $E = M_s \cdot E_s$. The MML receiver performance is given by (66) and (67) with Q_d and Q_F given by (60), (61) and (64), (65) for the coherent and noncoherent cases, respectively. However, now

$$d_M^2 = \frac{2E}{MN_o} = \frac{2M_sE_s}{M_HM_sN_o} = \frac{2E_s}{M_HN_o}. \qquad (71)$$

Note that the energy per hop d_M in (71) does not depend on M_s. As a consequence, the probability of detection increases slowly as M_s increases. This is seen in Fig. 18 where P_D is plotted versus P_F for SNR = 5 dB and M_H = 1. That is, the signal hops once per symbol interval, and the receiver integrates over M_s symbols to improve the detection probability. Curves for M_s = 1, 10, 100, and 1000 are presented, and P_F ranges from 10^{-3} to 10^{-9}. Observe that P_D is increased by less than a factor of 4.4 by integrating over 1000 symbols. Contrast these results shown in Fig. 18 with those shown in Fig. 19 concerning the influence of M_H. Note that d_M in (71) is inversely proportional to M_H. One sees from Fig. 19 that for a receiver that integrates over 1000 symbols, the probability of detection is decreased by a factor of 4.4 if the transmitter hops three times per symbol rather than once per symbol. This is at a false alarm probability , $P_F = 10^{-9}$. Note that in the examples of Figs. 18 and 19, T_o = 1 has been used. While T_o = 1 is optimum for $M_s \leq 10$, a reviewer has pointed out that optimization of T_o for M_s = 100 and M_s = 1000 may yield a significant improvement in P_D. This optimization is laborious, so we present the values of P_D in Figs. 18 and 19 as lower bounds.

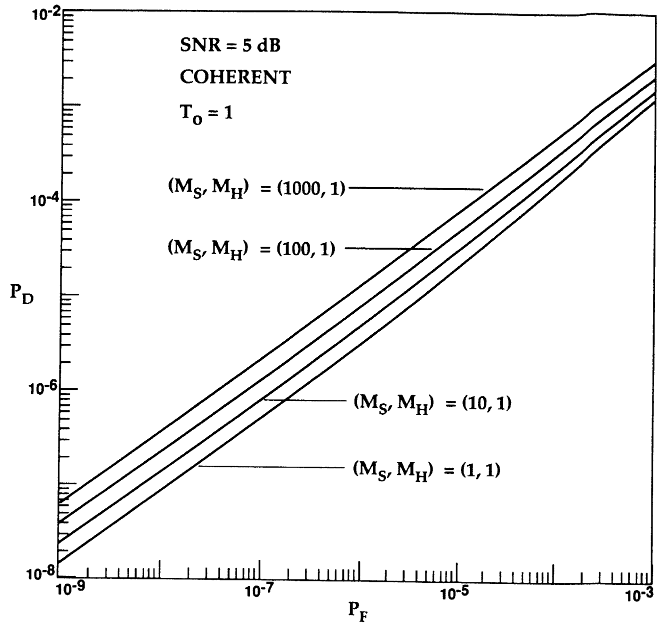


Fig. 18. The detection probability P_D as a function of the false alarm probability P_F. The signal hops once per symbol interval, and the coherent MML receiver integrates over M_s symbols before making a detection decision.

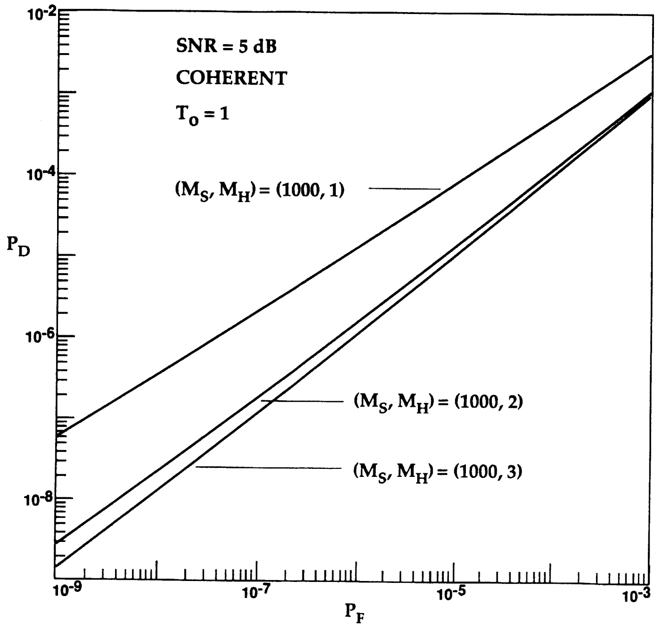


Fig. 19. The detection probability P_D as a function of the false alarm probability P_F for a coherent MML receiver that integrates over 1000 symbols. The signal hops M_H times per symbol.

VI. Conclusions

Modeling a slow-frequency-hopped spread-spectrum signal as a sinusoid with a random, discrete frequency distribution gives a summation of nonlinear transformations of the outputs from a bank of correlators as the optimum, coherent, Neyman-Pearson, interception receiver. One correlator is present for each frequency in the discrete distribution. In the noncoherent case, in-phase and quadrature correlators are followed by an envelope detector.

For both the coherent and noncoherent cases, the probability of detection of the optimum receiver tends to the probability of false alarm as N, the number of discrete frequencies, gets large. A related suboptimum receiver tests each correlator output against a threshold. This receiver, the ML receiver, is not optimum, but its detection performance is close to optimum. This receiver is easier to both realize and analyze. For a given N and the ML case, we have established the SNR below which interception is not practical as the detection probability is very close to the false alarm probability.

The optimum receiver for fast-frequency-hopped spread-spectrum signals is a logarithmic summation of optimum receivers for each chip. Fast frequency hopping assumes the carrier frequency is constant over a chip but changes a number of times over the detection interval. A less complex receiver, the MML receiver, has been proposed and analyzed. For both coherent and noncoherent MML receivers, the probability of detection tends rapidly to the probability of false alarm as M, the number of hops per interval, increases. Furthermore, for a given received energy per detection interval and a given bandwidth, P_D tends to P_F much faster by increasing M than by increasing N. For a given symbol energy, hopping several times per symbol decreases P_D much more rapidly than P_D can be increased by integrating over many symbols prior to making a detection decision.

Appendix

The following are the details for the CLT + 1 approximation for the coherent optimum receiver.

Consider (35) and the probability of detection,

$$P_D = P(L_{H_1} \geq \eta').$$

Now,

$$L_{H_1} = Z + e^{\alpha_k},$$

where

$$Z = \sum_{\substack{i=1 \\ i \neq k}}^{N} e^{\alpha_i}.$$

If N is large, Z is approximately

$$\eta(\bar{Z}, \sigma_Z^2) = \eta((N-1)e^{d^2/2}, (N-1)e^{d^2}(e^{d^2}-1))$$

and $P(L_{H_1} \geq \eta') = P(Z \geq \eta' - e^{\alpha_k})$. Therefore,

$$P_D = \int_{-\infty}^{\infty} \frac{e^{-(\alpha_k - \bar{\alpha}_k)^2/2\sigma_{\alpha_k}^2}}{\sigma_{\alpha_k}\sqrt{2\pi}} \cdot Q\left(\frac{\eta' - e^{\alpha_k} - \bar{Z}}{\sigma_Z}\right) d\alpha_k. \tag{A-1}$$

Note that P_F is given by (36), and hence η' is determined by solving this equation. We have

$$\eta' = Ne^{d^2/2} + e^{d^2/2}\sqrt{N(e^{d^2}-1)}\,Q^{-1}(P_F).$$

The result above for η' is substituted into (A-1), and numerical integration is used to compute P_D.

The CLT + 1 approximation for the noncoherent case is analogous to the coherent case.

Acknowledgment

We thank L. L. Campbell, Mathematics Department, Queen's University, for his assistance in obtaining the results on $E(I_o^n(x))$, $n = 1, 2$, where x is Rayleigh or Rician. Queen's undergraduates S. Graham and D. Rae have assisted with the preparation of the figures.

References

[1] R. C. Dixon, *Spread Spectrum Systems*, 2nd ed. New York: Wiley, 1984.

[2] A. J. Viterbi, "Spread spectrum communications—Myths and Realities," *IEEE Commun. Mag.*, vol. 17, pp. 11-18, May 1979.

[3] —, "When not to spread spectrum—A sequel," *IEEE Commun. Mag.*, vol. 23, pp. 12-17, Apr. 1985.

[4] M. Kavehrad and P. J. McLane, "Spread spectrum for indoor digital radio," *IEEE Commun. Mag.*, vol. 25, pp. 23-40, June 1987.

[5] A. B. Glenn, "Low probability of intercept," *IEEE Commun. Mag.*, vol. 21, pp. 26-33, July 1983.

[6] R. Rhodes, "SAW devices for modulation and demodulation," in *1979 EASCON Conf. Rec.*, Oct. 1979.

[7] L. E. Brennan, I. S. Reed, and W. Sollfrey, "A comparison of average likelihood and maximum likelihood ratio tests for detecting radar targets of unknown Doppler frequency," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 104-110, Jan. 1968.

[8] A. Polydoros and C. L. Nikias, "Advanced detection of unknown frequency sinusoids in broadband noise," in *Proc. ICC'86*, June 1986, pp. 266-270.

[9] R. A. Dillard, "Detectability of spread-spectrum signals," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-15, pp. 526-637, July 1979.

[10] G. R. Cooper, "Detection of frequency-hop signals," in *Proc. MILCOM'86*, Oct. 1986, pp. 10.2.1-10.2.5.

[11] A. Polydoros and K. T. Woo, "LPI detection of frequency-hopping signals using autocorrelation technique," *IEEE J. Select. Areas Commun.*, vol. SAC-3, Sept. 1985.

[12] D. J. Torrieri, *Principles of Military Communications Systems*. Dedham, MA: Artech House, 1981.

[13] W. E. Snelling and E. Geraniotis, "Sequential detection of unknown frequency-hopped waveforms," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 602-617, May 1989.

[14] H. L. Van Trees, *Detection, Estimation, and Modulation Theory—Part I*. New York: Wiley, 1968.

[15] S. C. Schwartz and Y. S. Yeh, "On the distribution function and moments of power sums with log-normal components," *Bell Syst. Tech. J.*, vol. 61, pp. 1441-1462, Sept. 1982.

[16] W. L. Hopkins, "Detection of a sinusoidal signal with a wideband discrete frequency distribution," M.Sc. thesis, Dep. Elec. Eng., Queen's Univ., Kingston, Canada, Sept. 1987.

[17] N. A. Marlow, "A normal limit theorem for power sums of independent random variables," *Bell Syst. Tech. J.*, vol. 46, pp. 2082-2089, Nov. 1967.

[18] W. F. McGee, "Another recursive method for computing the Q-function," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 500-501, July 1970.

[19] C. V. Helstrom, *Statistical Theory of Signal Detection*. Elmsford, NY: Pergamon, 1968.

[20] H. Cramer, *Random Variables and Probability Distributions*, 2nd ed. Cambridge, England: Cambridge University Press, 1970.

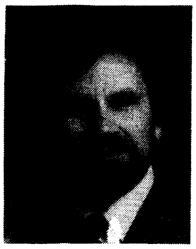
[21] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 2nd ed. New York: McGraw-Hill, 1984.

[22] G. N. Watson, *A Treatise on the Theory of Bessel Functions*. Cambridge, England: Cambridge Univ. Press, 1965.

[23] M. Abramowitz and I. E. Stegun, *Handbook of Mathematical Functions*. Washington, DC: National Bureau of Standards, 1972.

[24] P. J. McLane, W. L. Hopkins, and N. C. Beaulieu, "The study of space communications spread spectrum systems: Part IV—Detection of a sinusoidal signal with a wideband discrete frequency distribution," Queen's Univ. Res. Rep. 88-3, Feb. 1988.

[25] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1983.

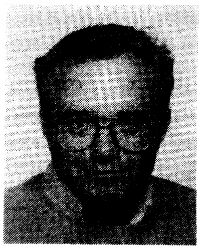


Norman C. Beaulieu received the B.A.Sc. (honors), M.A.Sc., and Ph.D. degrees in electrical engineering in 1980, 1983, and 1986, respectively.

He joined the Department of Electrical Engineering, Queen's University, Kingston, Ont., Canada, as Queen's National Scholar Assistant Professor in September 1986. Since June 1988 he has been an Associate Professor at Queen's. His current research interests include digital modulation, synchronization, importance sampling, and fading channels.

Wendy L. Hopkins (S'84–M'88) was born in Montreal, P.Q., Canada, in 1963. She received the Bachelor of Electrical Engineering degree from McGill University in 1986 and the M.Sc. degree in digital communications from Queen's University at Kingston in 1987.

During the summer of 1985, she worked for the Defense Research Establishment, Ottawa, where she developed communications and control interfaces between automatic test equipment and gyrospace testbeds for the Navigation Group there. She has been employed with Bell-Northern Research, Ottawa, Ont., since 1987. She initially was part of the customer support team, but in 1989 she became a member of the Radio Group, where her responsibilities include design of some of the software test tools for the Radio Control Unit. Her research interests are communication protocols, satellite communications, and mobile communications.



Peter J. McLane (S'68–M'69–SM'80–F'88) was born in Vancouver, B.C., Canada, on July 6, 1941. He received the B.A.Sc. degree from the University of British Columbia, Vancouver, in 1965, the M.S.E.E. degree from the University of Pennsylvania, in 1966, and the Ph.D. degree from the University of Toronto, Toronto, Ont., Canada, in 1969. He held a Ford Foundation Fellowship at the University of Pennsylvania and a National Research Council of Canada Scholarship at the University of Toronto.

From 1966 to 1967 he was a Junior Research Officer with the National Research Council, Ottawa, Ont. He held summer positions there in 1965 and 1966 and with the Defence Research Board of Canada in 1964. He joined the Department of Electrical Engineering, Queen's University, Kingston, Ont., in 1969 as an Assistant Professor, and since 1978 he has held the rank of Professor. His research interests are in signal processing for digital communication systems and radar. Usually this involves computer-aided analysis, but of late he has been involved in experimental work involving microprocessors and LSI-based implementation. He has served as a consultant to the Canadian Department of Communication; the Canadian Institute of Guided Ground Transport; Canadian Astronautics Ltd. of Ottawa, Ont.; Spar Aerospace of Toronto, Ont.; and AT&T Bell Laboratories and Technology Group of Los Angeles, CA. During 1984–1985 he was on leave at AT&T Bell Laboratories, Crawford Hill Laboratory, Holmdel, NJ.

Dr. McLane has been active in the IEEE Communications Society. He is a member of the Communication Theory Committee and served as its representative on the Technical Program Committee of the 1978 International Conference on Communications. He is a former Associate Editor for the IEEE COMMUNICATIONS MAGAZINE and is currently an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS. In addition, he was a Co-Editor of a special issue of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He is also a member of the Association of Professional Engineers of Ontario and Eta Kappa Nu, and is listed in *American Men and Women in Science.*

Combined Tone and Noise Jamming Against Coded FH/MFSK ECCM Radios

HYUCK M. KWON, MEMBER, IEEE, AND PIL JOONG LEE, SENIOR MEMBER, IEEE

Abstract—**It is known that partial-band tone jamming (PBTJ) is generally the worst form of jamming for frequency-hopped *M*-ary FSK (FH/MFSK) communication systems. Recent studies show that for some coded systems, full-band noise jamming (FBNJ) is more effective than worst-case PBTJ if a receiver is able to utilize jamming state information (JSI) for decoding, when the symbol energy-to-uniform noise jamming power spectral density ratio (E_s/N_J) is small. In this paper, we conjecture that a proper combination of PBTJ and FBNJ under a given total jamming power constraint may be more effective than PBTJ alone, not only for the case with low E_s/N_J but also for the case with high E_s/N_J, since the FBNJ can corrupt the JSI. Assuming this combination of PBTJ and FBNJ jamming, we consider three cases of receiver processing—the hard decision (HD) metric without JSI, the HD metric with perfect JSI, and the maximum likelihood (ML) metric using Viterbi's ratio threshold (VRT) to generate a 1-bit symbol decision quality indicator. System performance is evaluated in terms of the Chernoff bound on the probability of symbol error. From extensive numerical analysis we conclude the following. For the case of the HD metric without JSI, PBTJ-only jamming is the worst form of jamming as expected since the receiver does not use JSI at all; for the other cases, a combination of PBTJ and FBNJ is the worst, with the worst ratio of PBTJ power to FBNJ power a function of the values of *M* and E_s/N_J.**

I. Introduction

In an electronic warfare environment, where a battle is waged between the communicating party and a jammer who is intent on disrupting the communicator's link, strategy plays an important and fundamental role for the opposing parties. In this paper, we analyze and evaluate the effectiveness of sophisticated jamming waveforms in degrading the performance of a frequency-hopped *M*-ary frequency shift keying (FH/MFSK) communications system which utilizes various ECCM (electronic counter countermeasures) techniques to mitigate the jamming effects.

The FH/MFSK system considered in this paper is a scheme which can provide the communicator with jam-resistant radio capabilities. The most successful type of jamming against FH/MFSK radios has been shown to consist of placing equal power jamming tones such that, at most, one of the *M*-ary signaling frequencies is jammed on a given hop, called "$n = 1$ band multitone jamming" [1, vol. II, p. 113], [7]. In this paper, we call this jamming simply "tone jamming" or "partial-band tone jamming" (PBTJ).

Recent studies show that for some coded systems, full-band noise jamming (FBNJ)[1] is more effective than worst-case PBTJ if a receiver is able to utilize jamming state information (JSI) for decoding, especially when the symbol energy-to-uniform noise jamming power spectral density ratio (E_s/N_J) is small [1, vol. II, p. 178], [7]. In this paper, we conjecture that a proper combination of PBTJ and FBNJ under a given total jamming power constraint may be worse than PBTJ alone, not only for the case with low E_s/N_J but also for the case with high E_s/N_J, since the FBNJ can corrupt the JSI.

Because the tone jamming power is concentrated at certain frequencies, a JSI generator in the FH/MFSK receiver can easily detect whether a jamming signal is present or not during a symbol transmission. Such JSI can be exploited by the ECCM receiver's decoder to emphasize the unjammed symbol and to deemphasize the jammed symbol [1]–[3]. As the jamming power becomes stronger and stronger, the JSI about the PBTJ is more and more reliable. Sometimes, a strong PBTJ may actually improve the communication link performance if perfect JSI is available. Hence, a more intelligent jammer will not use PBTJ only against an ECCM receiver with a JSI generator. Instead, the jammer may choose to combine PBTJ with FBNJ because the ECCM radio cannot easily detect FBNJ, and therefore FBNJ can cause the ECCM radio receiver to generate inaccurate JSI about the PBTJ. It is the objective of this paper to investigate how well a coded FH/MFSK system with JSI that has been fortified against PBTJ alone withstands a simultaneous onslaught of both FBNJ and PBTJ. We assume the total jamming power is fixed and the portion of total jamming power used in FBNJ can be controlled to maximize the jamming effects.

Assuming this combination of PBTJ and FBNJ jamming, we consider three cases of receiver processing—the hard decision (HD) metric without JSI, the HD metric with perfect JSI, and the maximum likelihood (ML) metric using Viterbi's ratio threshold (VRT) to generate a 1-bit symbol decision quality indicator [4], [5], [8]. We consider a VRT receiver in this paper because it is a simple and practical antitone-jamming receiver. In addition,

Manuscript received February 3, 1989; revised November 15, 1989.

H. M. Kwon is with the Lockheed Engineering & Sciences Company, Houston, TX 77058 on leave from the Department of Electrical Engineering and Computer Science, University of Wisconsin-Milwaukee, Milwaukee, WI 53211.

P. J. Lee is with Bell Communications Research, 445 South Street, Morristown, NJ 07960.

IEEE Log Number 9035812.

[1] This is identical to the effect of classical additive Gaussian noise (AWGN), except that the channel corruption is caused by a broadband jammer.

0733-8716/90/0600-0871$01.00 © 1990 IEEE

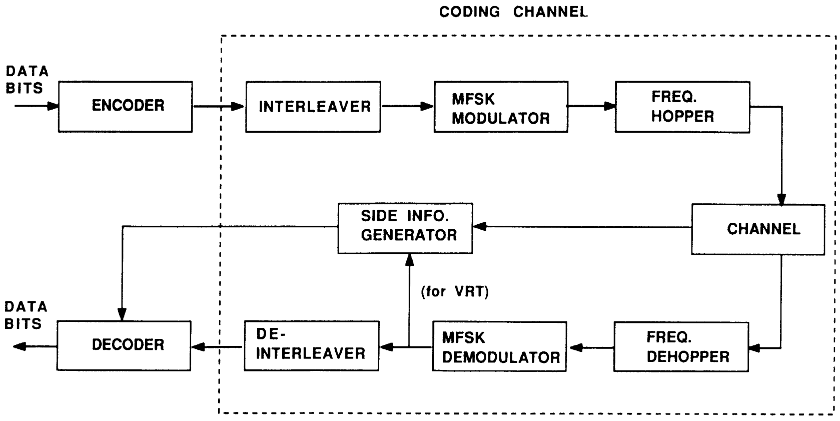


Fig. 1. A coded frequency-hopped M-ary frequency shift keying communication system overview.

the VRT techniques can be regarded as imperfect side information[2] generators. Therefore, the ML metric with VRT is a good candidate for the purpose of comparison to the HD metric for the cases of perfect JSI and no JSI. System performance is evaluated in terms of the Chernoff bound on the probability of symbol error [1, vol. I, p. 193]. Many who are familiar with coding will appreciate the value of the Chernoff bound's D parameter in the computation of the error probability bounds for given code structures and in the comparison of cutoff rates, which are practically achievable code rates for the memoryless coding channel [1]-[5]. Here the meaning of coding channel is the effective channel as seen by the encoder and decoder system (see Fig. 1). Once the D parameter is obtained, the cutoff rate and bit error probability bound can be calculated for given specific codes [1, vol. I, ch. 4]. Hence, we will evaluate D parameters for coded FH/MFSK systems. First, we will maximize D parameters with respect to the fraction of hopping slots jammed by the PBTJ when a portion of the total jamming power is assigned to FBNJ and the rest of total jamming power is used by PBTJ. Second, the worst-case combined jamming with respect to the portion of total jamming power used in FBNJ against the three cases of receivers mentioned above will be numerically calculated for some specific E_s/N_J. Finally, we plot D versus E_s/N_J with the portion of total jamming power used in FBNJ as a parameter.

The paper is organized as follows. A general system description is given in Section II. In Section III, the HD metric cases are analyzed. Section IV examines the ML metric case with imperfect JSI generated by the VRT. In Section V, numerical results are discussed, and conclusions are given in Section VI.

II. System Description

Our analysis neglects thermal noise, which is assumed to be dominated by the effects of the jammer. We assume contiguous MFSK modulation and one hop per transmitted symbol. In addition, we consider "$n = 1$ partial-band multitone jamming" [1, vol. II, p. 79] which jams at most one of the M-ary signaling frequencies. The PBTJ chooses randomly the jamming portion of the total frequency-hopped system bandwidth W.

Let J be the total jamming power and ϵ be a portion of J used by the FBNJ, $0 \le \epsilon \le 1$. Then the PBTJ uses power $(1-\epsilon)J$. Let $N_J = J/W$ denote the uniform noise jamming power spectral density over the total frequency-hopped system bandwidth W, with $N = \epsilon N_J/T_h$ as the FBNJ power measured in a frequency hopping slot, whose bandwidth is the inverse of a hopping time interval T_h. Let $N_{fh} = WT_h$ denote the total number of frequency hopping slots, q the number of jamming tones chosen by the PBTJ, I the power in a single tone (equal to $(1-\epsilon)J/q$), S the received signal power, E_s the symbol energy, and $K = \log_2 M$ the number of bits in a channel symbol. Then, the fraction of hopping symbols jammed by the PBTJ, ρ, is given by

$$\rho = \frac{Mq}{N_{fh}}, \qquad 0 < \rho \le 1. \tag{1}$$

This ρ corresponds to the probability that any symbol in a contiguous M-ary band is tone-jammed. (Our ρ is μ in [1, vol. II, eq. (2.28), p. 80].) Note that this ρ is also analogous to the fraction of the full spread-spectrum bandwidth W jammed in the case of partial-band noise jamming. The worst-case PBTJ chooses ρ to make the FH/MFSK system have the worst performance for given PBTJ power, $(1-\epsilon)J$. The signal-to-FBNJ power ratio S/N, the single jamming tone-to-FBNJ power ratio I/N, and the signal-to-single tone power ratio S/I are then

$$\frac{S}{N} = \frac{1}{\epsilon}\frac{E_s}{N_J}, \quad \frac{I}{N} = \frac{(1-\epsilon)M}{\rho\epsilon}, \quad \frac{S}{I} = \frac{\rho E_s/N_J}{(1-\epsilon)M}. \tag{2}$$

The ratios in (2) will be frequently used in the analysis and numerical computations of D parameters in Sections III-V for a given M and E_s/N_J.

[2] Actually, this side information is not restricted to JSI, that is, the quality bit indicates the general quality of the channel itself.

Let the vector z_i denote an M-dimensional PBTJ state vector whose $(M - 1)$ components are 0 for frequency slots unjammed, and 1 for the ith frequency slot jammed, for an M-ary channel orthogonal symbol transmission, $i = 1, \cdots, M$ (see Fig. 2). And let z_0 mean that no M-ary symbols are tone jammed. Since at most one of the M-ary signaling frequencies is tone jammed, there are $(M + 1)$ possible PBTJ states. The probabilities of possible PBTJ states are then

$$\Pr[z_1] = \Pr[z_2] = \cdots = \Pr[z_M] = \frac{\rho}{M},$$
$$\Pr[z_0] = 1 - \rho. \tag{3}$$

Without loss of generality, we can assume the transmitted channel symbol is the first symbol because the M-ary symbols are equiprobable.

Under a PBTJ state vector z_1, the transmitted signal and the jamming tone fall in the same frequency slot, and the conditional probability density functions of envelope detector outputs, $R_1, \cdots, R_M$, after dehopping [1, vol. I, p. 206], are

$$p(R_1|z_1) = \frac{R_1}{N} I_0\left(\frac{A_c R_1}{N}\right) \exp\left(-\frac{R_1^2 + A_c^2}{2N}\right), \tag{4a}$$

for the signal channel, and

$$p(R_k|z_1) = \frac{R_k}{N} \exp\left(-\frac{R_k^2}{2N}\right), \qquad k = 2, \cdots, M, \tag{4b}$$

for the nonsignal channels, where I_0 is the modified Bessel function of the first kind, $A_c^2 = 2S + 2I + 2\sqrt{2S}\sqrt{2I} \cos \phi$, and ϕ is the relative tone phase to signal, uniformly distributed over $[0, 2\pi]$.

Under a PBTJ state vector z_2, the signal and the tone fall in different frequency slots, and the conditional probability densities [1, vol. I, p. 206] are

$$p(R_1|z_2) = \frac{R_1}{N} I_0\left(\frac{\sqrt{2S}R_1}{N}\right) \exp\left(-\frac{R_1^2 + 2S}{2N}\right) \tag{5a}$$

for the unjammed signal channel,

$$p(R_2|z_2) = \frac{R_2}{N} I_0\left(\frac{\sqrt{2I}R_2}{N}\right) \exp\left(-\frac{R_2^2 + 2I}{2N}\right) \tag{5b}$$

for the jammed nonsignal channel, and

$$p(R_k|z_2) = \frac{R_k}{N} \exp\left(-\frac{R_k^2}{2N}\right), \qquad k = 3, \cdots, M \tag{5c}$$

for the unjammed nonsignal channels.

Under z_0, no channels are tone jammed, and the probability density of R_1 is that shown in (5a), and the probability densities for nonsignal channels are given in (4b).

For a memoryless M-ary orthogonal coding channel, and given jammer's choices of ϵ and ρ, the cutoff rate R_0 is a function of the parameter D [1, vol. I, pp. 193–195], given by

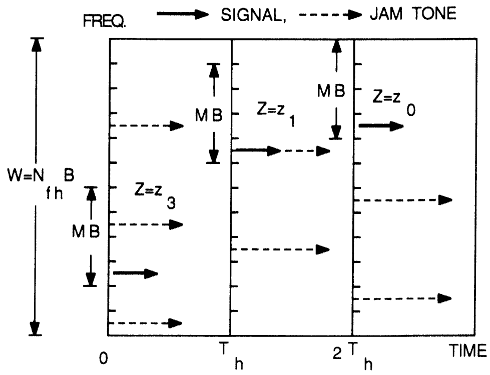


Fig. 2. Examples of partial-band tone jamming events $z_3 = (0, 0, 1, 0)^T$, $z_1 = (1, 0, 0, 0)^T$, and $z_0 = (0, 0, 0, 0)^T$, for a frequency-hopped M = 4-ary FSK modulation, where T is the transpose of a vector. Without loss of generality, we assume the transmitted symbol is the first symbol because the M-ary symbols are equiprobable. A frequency hopping slot bandwidth B is equal to the inverse of a frequency hopping time inteval T_h. We consider one hop per symbol in this paper.

$$R_0(\epsilon, \rho) = 1 - \log_M[1 + (M - 1)D(\epsilon, \rho)], \tag{6}$$

where

$$0 \leq D(\epsilon, \rho) = \min_{\lambda \geq 0} D(\epsilon, \rho, \lambda) \leq 1, \tag{7}$$

where λ is the Chernoff probability bound parameter and

$$D(\epsilon, \rho, \lambda) = E\Big\{\exp\left(\lambda \cdot [\mu(Y, X = m', Z) - \mu(Y, X = m, Z)]\right)\Big| X = m, m' \neq m\Big\}. \tag{8}$$

In (8), "E" means the expectation over the random received vector Y (here $Y = R$) and PBTJ state vector Z. The metric $\mu(y, x = m, z)$ is the receiver's decision metric when the transmitted symbol is m, $1 \leq m \leq M$, and the PBTJ state vector is z. We assume $m = 1$ for the transmitted symbol, and $m' = 2$ for the nontransmitted symbol, without loss of generality. The D parameter is the Chernoff bound on the probability that the decision metric for the nontransmitted symbol, $m' = 2$, is larger than that for the transmitted symbol, $m = 1$, on a single symbol transmission. The average of conditional D parameters over possible tone jamming conditions is then

$$\begin{aligned} D(\epsilon, \rho, \lambda) = {} & \Pr(z_1) D(\epsilon, \rho, \lambda|z_1) \\ & + \Pr(z_2) D(\epsilon, \rho, \lambda|z_2) \\ & + (M - 2)\Pr(z_3) D(\epsilon, \rho, \lambda|z_3) \\ & + \Pr(z_0) D(\epsilon, \rho, \lambda|z_0). \end{aligned} \tag{9}$$

Jammer's goal is to maximize $D(\epsilon, \rho)$ by choosing the best ϵ and ρ. The general relation of the parameter D to the coded bit error probability is $P_b \leq G(D)$, where $G(\cdot)$ is a function determined solely by the specific code, whereas the parameter D depends only on the coding channel and the decoder metric [1, vol. I, pp. 194, 199].

III. Hard Decision Metrics

When hard decisions are made and JSI about PBTJ is available, the metric is

$$\mu(y, x = m, z_i) = \begin{cases} c_i & \text{if } y_m > y_{m'} \quad \text{for all } m' \neq m \\ 0 & \text{otherwise,} \end{cases} \tag{10}$$

for an M-dimensional observation vector y of envelopes where the weighting coefficients are

$$c_i = \begin{cases} 1 & \text{for } i = 0, 1, \\ c & \text{for } i = 1, 2, \cdots, M. \end{cases} \tag{10a}$$

A practically implementable JSI generator with weights (10a) can be built as follows [1, vol. II, last paragraph, p. 113]. When only one energy detector output among M detector outputs is high on a given transmission, use relatively large weight 1, and when two or more energy detector outputs are high, use relatively small weight c. In this weighting, the decoder makes use of the PBTJ event that helps the communication link, and discards the PBTJ events that disturb the communications, together with the transmitted information. If there is only PBTJ, then such JSI becomes perfect. But if FBNJ shares the total jamming power with PBTJ, then such JSI cannot be perfect because FBNJ can cause the JSI generator to generate wrong JSI about PBTJ. In this paper we assume that this JSI with weights (10a) is perfect for an ideal case. Using the HD metric with perfect JSI in (10) and (10a), we can express the parameter D in (9) as follows:

$$\begin{aligned} D(\epsilon, \rho, \lambda, c) = {} & \frac{\rho}{M}\left[e^{-\lambda}(1 - (M-1)A) + e^{\lambda}A + (M-2)A\right] \\ & + \frac{(M-1)\rho}{M}\left[e^{-\lambda c}B + e^{\lambda c}\frac{1-B}{M-1} + \frac{(M-2)(1-B)}{M-1}\right] \\ & + (1-\rho)\left[e^{-\lambda}(1-(M-1)C) + e^{\lambda}C + (M-2)C\right] \end{aligned} \tag{11}$$

where

$$\begin{aligned} A &= \Pr\left[R_{m'} > R_j \quad \text{for all } j \neq m',\ m' \neq 1 \,|\, z_1\right] \\ &= \sum_{k=0}^{M-2}\binom{M-2}{k}(-1)^k \frac{1}{(1+k)(2+k)} \\ &\quad \cdot \exp\left[-\frac{S+I}{N}\frac{1+k}{2+k}\right] I_0\left(\frac{2\sqrt{SI}}{N}\frac{1+k}{2+k}\right), \end{aligned} \tag{11a}$$

$$\begin{aligned} B &= \Pr\left[R_1 > R_j \quad \text{for all } j \,|\, z_2\right] = \sum_{k=0}^{M-2}\binom{M-2}{k} \\ &\quad \cdot \frac{(-1)^k}{1+k}\exp\left[-\frac{S}{N}\frac{k}{1+k}\right] \\ &\quad \times \left\{1 - Q(\sqrt{2a}, \sqrt{2b}) + \frac{1}{2+k}\right. \\ &\quad \left.\cdot \exp\left[-(a+b)\right]I_0(2\sqrt{ab})\right\}, \end{aligned} \tag{11b}$$

$$\begin{aligned} C &= \Pr\left[R_{m'} > R_j \quad \text{for all } j \neq m',\ m' \neq 1 \,|\, z_0\right] \\ &= \sum_{k=0}^{M-2}\binom{M-2}{k}(-1)^k \frac{1}{(1+k)(2+k)} \\ &\quad \cdot \exp\left[-\frac{S}{N}\frac{1+k}{2+k}\right], \end{aligned} \tag{11c}$$

$Q(x, y)$ is the Marcum Q function, $a = (I/N)(1 + k)/(2 + k)$ and $b = (S/N)(1/(1 + k)(2 + k))$. In derivations of (11a)–(11c), binomial expansions were used.

A. Hard Decision Metric Receiver with Perfect JSI

With perfect JSI, the receiver can optimize c. After minimizing $D(\epsilon, \rho, \lambda, c)$ with respect to λ and c, we obtain

$$D(\epsilon, \rho) = D_1(\epsilon, \rho) + D_0(\epsilon, \rho) \tag{12}$$

where

$$D_1(\epsilon, \rho) = \begin{cases} 2\sqrt{\alpha_1\beta_1} + \gamma_1 & \text{if } \alpha_1 \geq \beta_1 \\ (M-1)\rho/M & \text{otherwise,} \end{cases}$$

$$D_0(\epsilon, \rho) = \begin{cases} 2\sqrt{\alpha_0\beta_0} + \gamma_0 & \text{if } \alpha_0 \geq \beta_0 \\ 1 - (M-1)\,\rho/M & \text{otherwise,} \end{cases} \tag{13}$$

$$\begin{aligned} \alpha_1 &= \frac{\rho}{M}(M-1)B, \\ \alpha_0 &= \frac{\rho}{M}(1 - (M-1)A) + (1-\rho)(1-(M-1)C), \end{aligned} \tag{14}$$

$$\beta_1 = \frac{\rho}{M}(1 - B), \qquad \beta_0 = \frac{\rho}{M}A + (1-\rho)C, \tag{15}$$

$$\begin{aligned} \gamma_1 &= \frac{\rho}{M}(M-2)(1-B), \\ \gamma_0 &= \frac{\rho}{M}(M-2)A + (1-\rho)(M-2)C. \end{aligned} \tag{16}$$

As an extreme case, assume that there is no FBNJ and only PBTJ is active, i.e., $\epsilon = 0$, and that the perfect JSI is available. From (11), probabilities A and C are zero and probability B is

$$B = \begin{cases} 1 & \text{if } S/I \geq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

Then, the parameter D in (12) can be simplified to the well-known form [1, vol. II, eq. (2.138)]

$$D = \begin{cases} \dfrac{M-1}{E_s/N_j} & \text{if } \dfrac{E_s}{N_J} \geq M \\ \dfrac{M-1}{M} & \text{otherwise.} \end{cases} \tag{18}$$

As another extreme case, suppose only FBNJ is active, i.e., $\epsilon = 1$. From (11), probabilities A and B are equal to zero, and the parameter D in (12) becomes

$$D = \begin{cases} 2\sqrt{\alpha\beta} + \gamma & \text{if } \alpha \geq \beta \\ 1 & \text{otherwise;} \end{cases} \tag{19}$$

where

$$\alpha = \left(1 - (M-1)C\right), \qquad \beta = C, \qquad \gamma = (M-2)C. \tag{20}$$

In (19) and (20), the probability C is as in (11c) with $\epsilon = 1$.

B. Hard Decision Metric Receiver without JSI

Suppose the receiver cannot derive JSI. Then the metric must be independnet of the PBTJ states, which implies $c = 1$ in (11). After minimizing $D(\epsilon, \rho, \lambda)$ with respect to λ, we have

$$D(\epsilon, \rho) = \begin{cases} 2\sqrt{\alpha\beta} + \gamma & \text{if } \alpha \geq \beta \\ 1 & \text{otherwise,} \end{cases} \tag{21}$$

where

$$\begin{aligned} \alpha &= \frac{\rho}{M}\left(1 - (M-1)A\right) + \frac{\rho}{M}(M-1)B \\ &\quad + (1-\rho)\left(1 - (M-1)C\right), \\ \beta &= \frac{\rho}{M}A + \frac{\rho}{M}(1-B) + (1-\rho)C, \\ \gamma &= \frac{\rho}{M}(M-2)A + \frac{\rho}{M}(M-2)(1-B) \\ &\quad + (1-\rho)(M-2)C. \end{aligned} \tag{22}$$

When $\epsilon = 0$ (tone jamming) and JSI is not available, $A = C = 0$ in (11) and the parameter D in (21) becomes

$$\begin{aligned} D = \max_{0 < \rho \leq \min((1, M/E_s/N_J)} &\left[2\sqrt{(\rho/M + 1 - \rho)\rho/M}\right. \\ &\left. + (M-2)\rho/M\right]. \end{aligned} \tag{23}$$

If $M = 2$, then (23) becomes well-known result [2]

$$D = \begin{cases} 1 & \text{if } E_s/N_J < 2 \\ 2\sqrt{\left\{1 - 1/(E_s/N_J)\right\}/(E_s/N_J)} & \text{otherwise.} \end{cases} \tag{24}$$

When the FBNJ-only is active against FH/MFSK with the HD metric without JSI, the D parameter is the same as that for the HD metric with perfect JSI in (19), since JSI about PBTJ is meaningless.

IV. ML Metric with Viterbi's Ratio Threshold Technique

For the purpose of comparison to the HD cases discussed in Section III, we consider the VRT receiver [4], [5], [8] which can be regarded as a receiver with imperfect side information about the channel (not just PBTJ state information). In [5], the performance of the VRT receiver against PBTJ in the presence of background noise, which is not under the control of jammer, was investigated. In this section, we take the same model as in [5], with the following three main differences. First, here the FBNJ power can be controlled by the jamming in contrast to [5]. We will numerically try to find the worst-case value of ϵ, the portion of total jamming power used in FBNJ. Second, in [5], a Gaussian quadrature numerical integration method was used for computation of the parameter D, while in this paper we avoid numerical integrations as much as possible by expanding the corresponding probability expression in binomial expansion form, and by using Massaro's results [6, eq. (16)]. Third, our analysis is given for the original M-ary channel, which is in contrast to the binary decomposed channel analyzed in [5].

A coding channel with VRT techniques takes M-ary input symbols interleaved (e.g., 1, 2, 3, or 4 in Fig. 3 for $M = 4$) and produces $2M$-ary soft decision symbols. The $2M$-ary soft decision symbol consists of a hard decision symbol (from MFSK maximum envelope detector) and a quality bit Q (from VRT side information generator) (e.g., $1G$, $1B$, $\cdots$, $4B$ in Fig. 3 for $M = 4$). After deinterleaving, the quality bit is exploited to predict which bits are jammed, and hence discount potentially bad decisions in the decoder. The quality bit Q is derived as follows. If the ratio of the largest of the filter output envelopes, $R_1, \cdots, R_M$, to the next largest is bigger than a threshold θ, then $Q = G$ (good), otherwise $Q = B$ (bad). The threshold value is one of the parameters that can be controlled by the receiver. Notice that there are many parameters, i.e., ϵ, ρ, and θ, for our worst-case analysis. To make the problem solvable in a reasonable time period, we fix the threshold value, θ, in the numerical analysis of the next section.

The coding channel with M-ary inputs and $2M$-ary outputs can be characterized by discrete transition probabilities, P_C (correct, $Q = G$), P_{CX} (correct, $Q = B$), P_{EX} (error, $Q = B$), and P_E (error, $Q = G$). In Fig. 3, a coding channel with 4-ary input and 8-ary output is shown as an illustration similar to the analysis in [5]. These tran-

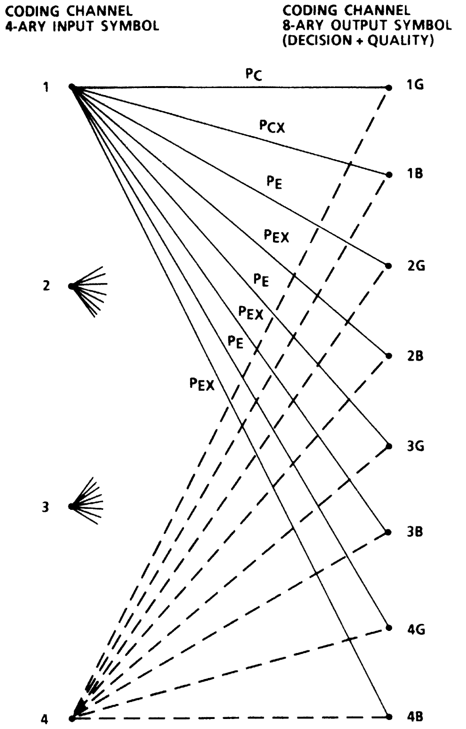

CODING CHANNEL 4-ARY INPUT SYMBOL
CODING CHANNEL 8-ARY OUTPUT SYMBOL (DECISION + QUALITY)
1
2
3
4
PC
PCX
PE
PEX
PE
PEX
PE
PEX
1G
1B
2G
2B
3G
3B
4G
4B

Fig. 3. M-ary input $2M$-ary output coding channel with ratio-threshold quality measure (here $M = 4$).

sition probabilities will be used in computations of the parameter D. For the ML metric with VRT techniques, the parameter D is

$$D(\epsilon, \rho, \theta) = 2\sqrt{P_C P_E} + 2\sqrt{P_{CX} P_{EX}} + (M-2)(P_E + P_{EX}). \tag{25}$$

The transition probabilities are functions of the average exceeding probabilities $\overline{F}_c$ and $\overline{F}_e$,

$$P_C = \Pr[R_1 \geq \theta R_j, \quad j = 2, \cdots, M | R_1] = \overline{F}_c(\theta, \rho), \tag{26}$$

$$P_E = \Pr[R_2 \geq \theta R_j, \quad j = 1, 3, \cdots, M | R_1] = \overline{F}_e(\theta, \rho), \tag{27}$$

$$\begin{aligned} P_{EX} &= \Pr[R_j < R_2 < \theta R_j, \quad j = 1, 3, \cdots, M | R_1] \\ &= \Pr[R_2 \geq R_j | R_1] - \Pr[R_2 > \theta R_j | R_1] \\ &= \overline{F}_e(1, \rho) - \overline{F}_e(\theta, \rho), \end{aligned} \tag{28}$$

$$\begin{aligned} P_{CX} &= 1 - \left(P_C + (M-1)P_{EX} + (M-1)P_E\right) \\ &= \overline{F}_c(1, \rho) - \overline{F}_c(\theta, \rho), \end{aligned} \tag{29}$$

where $\overline{F}_c(\theta)$ is the probability that the correct detector output is θ greater than the other detector outputs, and $\overline{F}_e(\theta)$ is the probability that an incorrect detector output is θ greater than the other detector outputs. By using a binomial expression and (16) of [6], we express the average exceeding probabilities $\overline{F}_c$ and $\overline{F}_e$ over the PBTJ states for a combined jamming of PBTJ and FBNJ as given below:

$$\begin{aligned} \overline{F}_c(\theta, \rho) &= \overline{F}_c(\theta, \rho | z_1) \Pr(z_1) \\ &\quad + (M-1)\overline{F}_c(\theta, \rho | z_2) \Pr(z_2) \\ &\quad + \overline{F}_c(\theta, \rho | z_0) \Pr(z_0) \\ &= \frac{\rho}{M} \times (31) + (M-1)\frac{\rho}{M} \times (32) \\ &\quad + (1-\rho) \times (33). \end{aligned} \tag{30}$$

From (5) and (6), the conditional correct exceeding probabilities (31)–(33) given the jamming states can be derived as

$$\begin{aligned} \overline{F}_c(\theta, \rho | z_1) &= \Pr\left[\bigcap_{j=2}^{M} (R_1 \geq \theta R_j) | z_1\right] \\ &= \sum_{k=0}^{M-1} (-1)^k \binom{M-1}{k} \frac{\theta^2}{\theta^2 + k} \\ &\quad \cdot \exp\left[-\frac{k}{k+\theta^2} \frac{S+I}{N}\right] \\ &\quad \times I_0\left[\frac{2\sqrt{SI}}{N} \frac{k}{k+\theta^2}\right], \end{aligned} \tag{31}$$

$$\begin{aligned} \overline{F}_c(\theta, \rho | z_2) &= \Pr\left[\bigcap_{j=2}^{M} (R_1 \geq \theta R_j) | z_2\right] \\ &= \sum_{k=0}^{M-2} (-1)^k \binom{M-2}{k} \frac{\theta^2}{\theta^2 + k} \\ &\quad \cdot \exp\left[-\frac{k}{k+\theta^2} \frac{S}{N}\right] \\ &\quad \times \left\{1 - Q(\sqrt{2a}, \sqrt{2b}) + \frac{1}{1+\theta^2+k} \right. \\ &\quad \left. \cdot \exp[-(a+b)] I_0(2\sqrt{ab})\right\}, \end{aligned} \tag{32}$$

and

$$\overline{F}_c(\theta, \rho | z_0) = \Pr\left[\bigcap_{j=2}^{M} (R_1 \geq \theta R_j) | z_0\right] = (31) \quad \text{with } I = 0, \tag{33}$$

where $a = (I/N)(\theta^2 + k)/(1 + \theta^2 + k)$ and $b = (S/N)\,\theta^2/((\theta^2 + k)(1 + \theta^2 + k))$. Note that if $\theta = 1$, then (32) becomes (11b) for ordinary HD metric. Similarly, we can derive a binomial expansion form for $\bar{F}_e$ as

$$\bar{F}_e(\theta, \rho) = \frac{\rho}{M} \times (35) + (M-1)\frac{\rho}{M} \times (37) + (1-\rho) \times (36), \tag{34}$$

$$\begin{aligned}\bar{F}_e(\theta, \rho|z_1) &= \Pr\left[\bigcap_{j=1}^{M} (R_2 > \theta R_j), \quad j \neq 2|z_1\right] \\ &= \sum_{k=0}^{M-2} \binom{M-2}{k} (-1)^k \frac{\theta^2}{\theta^2+k} \\ &\quad \cdot \frac{1}{1+\theta^2+k} \exp\left[-\frac{S+I}{N}\frac{\theta^2+k}{1+\theta^2+k}\right] \\ &\quad \times I_0\left(\frac{2\sqrt{SI}}{N}\frac{\theta^2+k}{1+\theta^2+k}\right) \\ &= \frac{-1}{M-1}\sum_{l=1}^{M-1}\binom{M-1}{l} \\ &\quad \cdot (-1)^l l \frac{\theta^2}{\theta^2+l-1}\frac{1}{\theta^2+l} \\ &\quad \cdot \exp\left[-\frac{S+I}{N}\frac{\theta^2+l-1}{\theta^2+l}\right] \\ &\quad \times I_0\left(\frac{2\sqrt{SI}}{N}\frac{\theta^2+l-1}{\theta^2+l}\right),\end{aligned} \tag{35}$$

$$\bar{F}_e(\theta, \rho|z_0) = \Pr\left[\bigcap_{j=1}^{M} (R_2 \geq \theta R_j), j \neq 2|z_0\right] = (35) \quad \text{with } I = 0, \quad \text{and} \tag{36}$$

$$\bar{F}_e(\theta, \rho|z_2) = \frac{1}{M-1} \times (38) + (M-2)\frac{1}{M-1} \times (39), \tag{37}$$

where

$$\Pr\left[\bigcap_{j=1}^{M} (R_2 \geq \theta R_j), j \neq 2|z_2\right] = (32) \quad \text{replaced } S \text{ by } I, \text{ and } I \text{ by } S, \tag{38}$$

$$\begin{aligned}\Pr&\left[\bigcap_{j=1}^{M} (R_3 \geq \theta R_j), j \neq 3|z_2\right] \\ &= \sum_{k=0}^{M-3}\binom{M-3}{k}(-1)^k\frac{\theta^2}{\theta^2+k} \\ &\quad \times \left\{\frac{1}{1+\theta^2+k}\left[\exp\left(-\frac{S}{N}\frac{\theta^2+k}{1+\theta^2+k}\right)\right.\right. \\ &\quad \left. + \exp\left(-\frac{I}{N}\frac{\theta^2+k}{1+\theta^2+k}\right)\right] - 1 \\ &\quad + \int_0^\infty te^{-t^2/2}Q\left(\sqrt{2S/N}, \frac{t}{\sqrt{\theta^2+k}}\right) \\ &\quad \left.\cdot Q\left(\sqrt{2I/N}, \frac{t}{\sqrt{\theta^2+k}}\right)dt\right\}.\end{aligned} \tag{39}$$

As an extreme case, suppose the PBTJ-only is active, i.e., $\epsilon = 0$. Then (30) and (34) can be simplified as

$$\bar{F}_c(\theta, \rho) = \begin{cases} 1 - \rho + \rho/M & \text{if } \sqrt{S/I} \leq \theta \\ 1 & \text{otherwise,} \end{cases}$$

$$\bar{F}_e(\theta, \rho) = \begin{cases} \rho/M & \text{if } \sqrt{S/I} \leq 1/\theta \\ 0 & \text{otherwise.} \end{cases} \tag{40}$$

Furthermore, for the PBTJ-only case, by defining $D_1(\rho) \equiv 2\sqrt{(1-\rho+\rho/M)\rho/M} + (M-2)\rho/M$, $D_2(\rho) \equiv (M-2)\rho/M$, $\rho_1 \equiv M/\theta^2 E_s/N_J$, and $\rho_2 \equiv M/E_s/N_J$, we can simplify the parameter D in (25) from (40) and (26)-(29) as follows:

$$D = \begin{cases} \max\left(D_1(\rho_1), D_2(\rho_2)\right) \\ \qquad \text{for } E_s/N_J > M, \quad \text{i.e., } \rho_2 < 1, \\ \max\left(D_1(\rho_1), D_2(1)\right) \\ \qquad \text{for } M/\theta^2 \leq E_s/N_J \leq M, \\ \qquad \quad \text{i.e., } \rho_1 < 1 \leq \rho_2 \\ 1 \quad \text{for } E_s/N_J \leq M/\theta^2, \quad \text{i.e., } 1 \leq \rho_1. \end{cases} \tag{41}$$

The closed form in (41) says that for $E_s/N_J \leq M/\theta^2$, i.e., low E_s/N_J, $D = 1$; and, for $E_s/N_J > M$, i.e., high E_s/N_J, D is proportional to the inverse of the square root of $\theta^2 E_s/N_J$. As another extreme case, suppose only FBNJ is active, i.e., $\epsilon = 1$. Then (35) = (38) = (39) = (37) = (36), and (31) = (32) = (33). Our general equations (30) and (34) become (2) of [5] as a special case.

V. Numerical Results

In this section, for a given M and receiver type, we compute the D parameter numerically for the three receivers considered. To see the intermediate results, we first compute $D(E_s/N_J, \epsilon, \rho)$ versus ρ, the fraction of hopping slots jammed by PBTJ, $0 < \rho \leq 1$, for a given E_s/N_J and ϵ. Second, after maximizing D with respect to ρ, we observe $D(E_s/N_J, \epsilon, \rho^*)$ versus ϵ, the portion of total jamming power used in FBNJ, $0 \leq \epsilon \leq 1$, with E_s/N_J as a parameter. The ρ^* is the worst-case jamming fraction for a given E_s/N_J and ϵ. Finally, we plot $D(E_s/N_J, \epsilon, \rho^*)$ versus E_s/N_J with ϵ as a parameter.

A. Hard Decision Metric with Perfect JSI

In Fig. 4, for the HD metric with perfect JSI against an FH/4FSK system, plots of $D(E_s/N_J, \epsilon, \rho)$ versus ρ are shown, with ϵ as a parameter, for four different values of E_s/N_J. For BFSK, the behavior of the D versus ρ curves is similar to that for 4FSK.

We observe the following three facts in Fig. 4(a) for a low E_s/N_J example. First, the worst-case jamming fraction ρ of the PBTJ is 1, for any ϵ. The PBTJ tries to hit as many M-ary bands as possible when E_s/N_J is small. Second, the FBNJ-only (i.e., $\epsilon = 1$ extreme case) is seen to be more effective jamming than the PBTJ-only (i.e.,

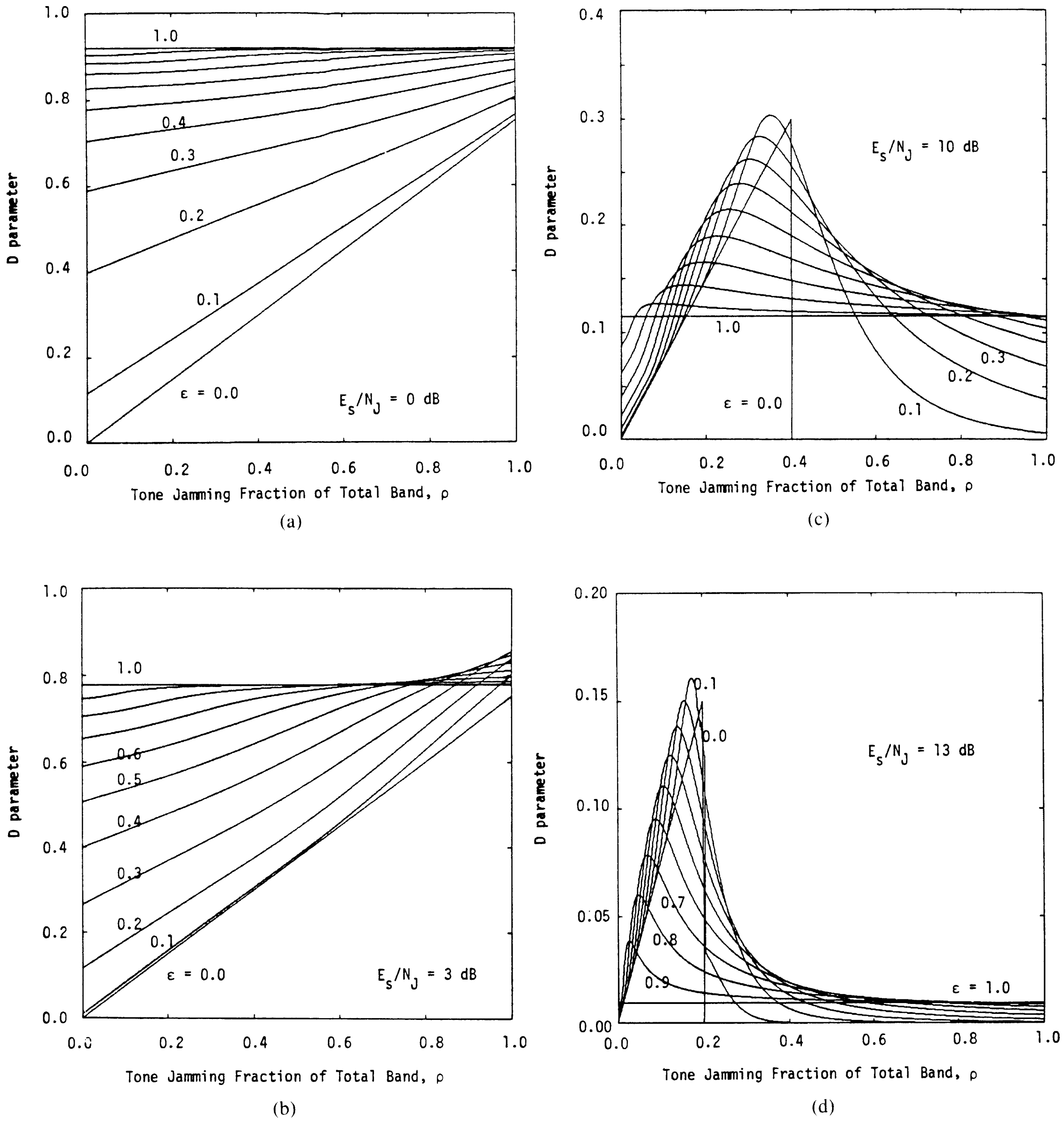


Fig. 4. D versus ρ for 4-ary FSK, hard decision with perfect tone jamming state information and ϵ as a parameter (step 0.1) when E_s/N_J = 0 dB (a), 3 dB (b), 10 dB (c), 13 dB (d).

$\epsilon = 0$ extreme case), if perfect JSI is available to the decoder. The reason for this result is that PBTJ-only can cause the receiver to have a maximum D equal to $(M - 1)/M$ [see (18)] while the FBNJ-only can cause D to be equal to 1 if E_s/N_J is small. Finally, as ϵ changes from zero to one, the combined jamming becomes the more effective jamming, because the FBNJ becomes more effective with increasing FBNJ power, while the PBTJ effect does not change much with decreasing PBTJ power.

We observe the following three facts in Fig. 4(d) for a high E_s/N_J example. First, $D(E_s/N_J, \epsilon, \rho)$ increases and then decreases as ρ varies from zero to one for a given E_s/N_J and ϵ. The peaks occur at ρ less than $M/(E_s/N_J)$ [= 0.20047 in Fig. 4(d)] at which D is maximum for the PBTJ-only. [See (2), (17), and (18).] Second, the PBTJ-only is the more effective jamming than the FBNJ-only, which is a known result [1, vol. II, p. 178], if perfect JSI is available to the decoder. The reason for this is that the D for the PBTJ-only with the worst-case jamming fraction ρ is a linearly inverse function of E_s/N_J [see (18)] while D for the FBNJ-only is an exponentially decreasing function of E_s/N_J [see (19)], and, after the crossing point of the two curves, the exponentially decreasing function drops faster than the linearly inverse one. Finally, we observe in Fig. 4(d) that $D(E_s/N_J, \epsilon, \rho^*)$, for some combined jamming (for example, $\epsilon = 0.1$) with the worst jamming fraction, denoted by ρ^*, can be larger than that of the PBTJ-only, if E_s/N_J is high.

For intermediate E_s/N_J examples, $D/(E_s/N_J, \epsilon, \rho)$ versus ρ with ϵ as a parameter is shown for a given E_s/N_J in Fig. 4(b) and (c). We can see that there is a tradeoff between FBNJ power and PBTJ power in order to make the HD metric with perfect JSI have the worst performance.

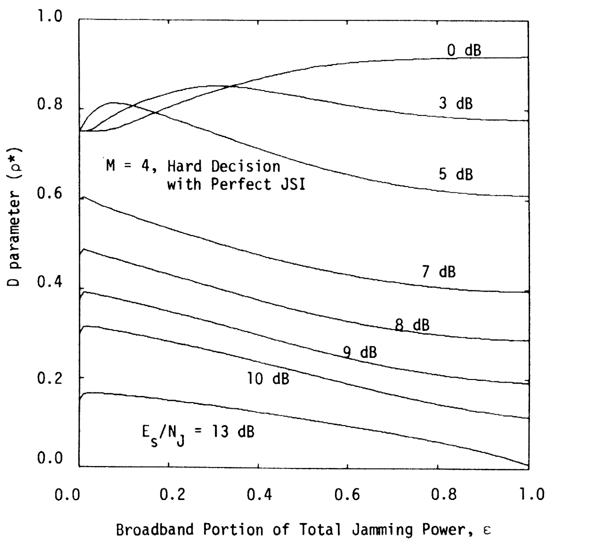


Fig. 5. D versus ϵ for 4-ary FSK, hard decision with perfect tone jamming state information and E_s/N_J as a parameter.

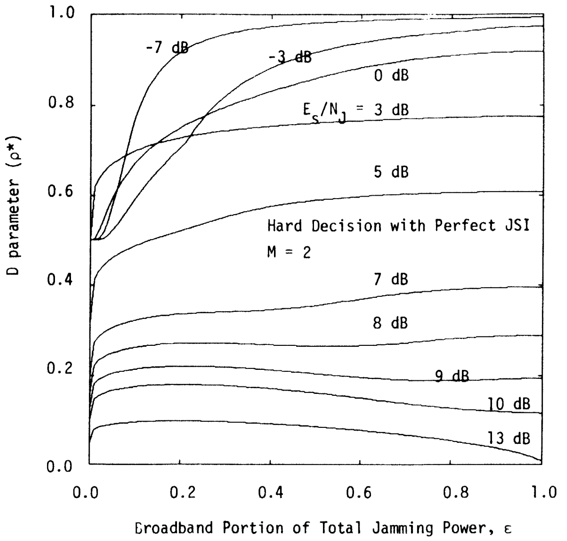


Fig. 7. D versus ϵ for binary FSK, hard decision with perfect tone jamming state infromation and E_s/N_J as a parameter.

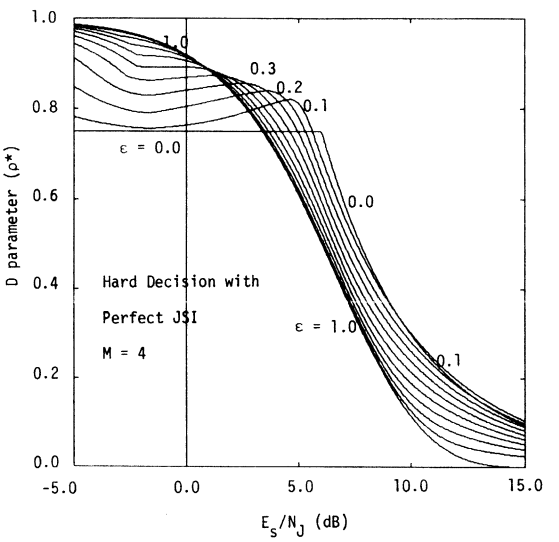


Fig. 6. D versus E_s/N_J for binary FSK, hard decision with perfect tone jamming state information and ϵ as a parameter (step 0.1).

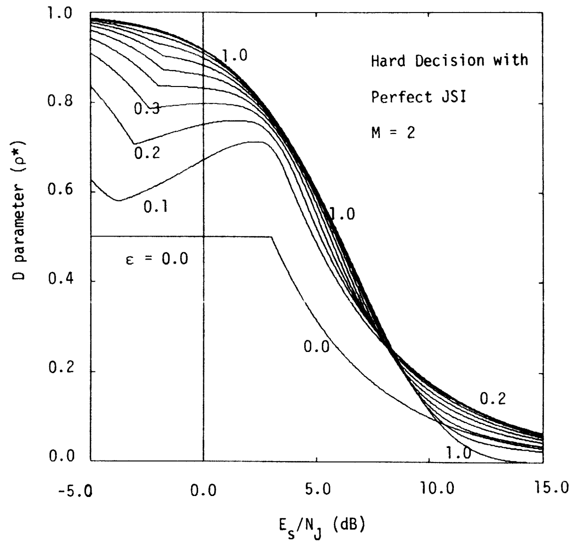


Fig. 8. D versus E_s/N_J for binary FSK, hard decision with perfect tone jamming state information and ϵ as a parameter (step 0.1).

In Fig. 5, we draw $D(E_s/N_J, \epsilon, \rho^*)$ versus ϵ with E_s/N_J as a parameter to see the worst case ϵ for $M = 4$. $D(E_s/N_J, \epsilon, \rho^*)$ versus E_s/N_J with ϵ as a parameter is shown in Fig. 6. The corresponding results for $M = 2$ are shown in Figs. 7 and 8. The FBNJ-only is the worst-case jamming if E_s/N_J is less than 1.14 dB for $M = 4$ (see Fig. 6), and the FBNJ-only is the worst case jamming if E_s/N_J is less than 8.23 dB for $M = 2$ (see Fig. 8). For other E_s/N_J, as shown in Figs. 6 and 8, we can design more effective combined jamming than the PBTJ-only or the FBNJ-only, to be active against the HD metric with perfect JSI.

B. Hard Decision Metric Receiver without JSI

For the HD metric without JSI, $D(E_s/N_J, \epsilon, \rho)$ is maximized by some worst-case value of the jamming fraction, denoted ρ^*, for given ϵ and E_s/N_J. The maximized value of D, $D(E_s/N_J, \epsilon, \rho^*)$, is inversely proportional to both ϵ and E_s/N_J, as shown in Fig. 9. Furthermore, as ϵ increases from 0 to 1, the combined jamming scheme uniformly becomes less effective, which says that PBTJ alone ($\epsilon = 0$) is the most efficient jamming, and FBNJ alone is ($\epsilon = 1$) the least, when the JSI is not available to the HD metric.

C. Maximum Likelihood Metric Receiver with VRT

The receiver with VRT techniques can change the threshold value to minimize the combined jamming effect. In this paper, however, we fix the threshold value equal to 2.5 for our numerical analysis because this threshold value was shown to be near optimum for the performance of evaluation of ratio threshold for FH/MFSK under PBTJ plus background noise [5]. In Fig. 10, for an FH/4FSK system using an ML metric with VRT,

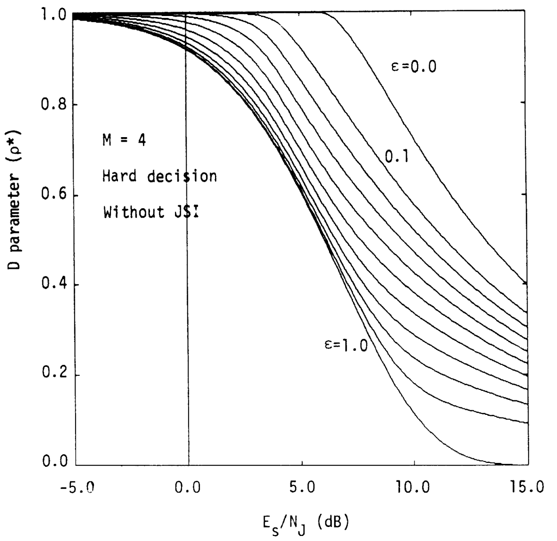
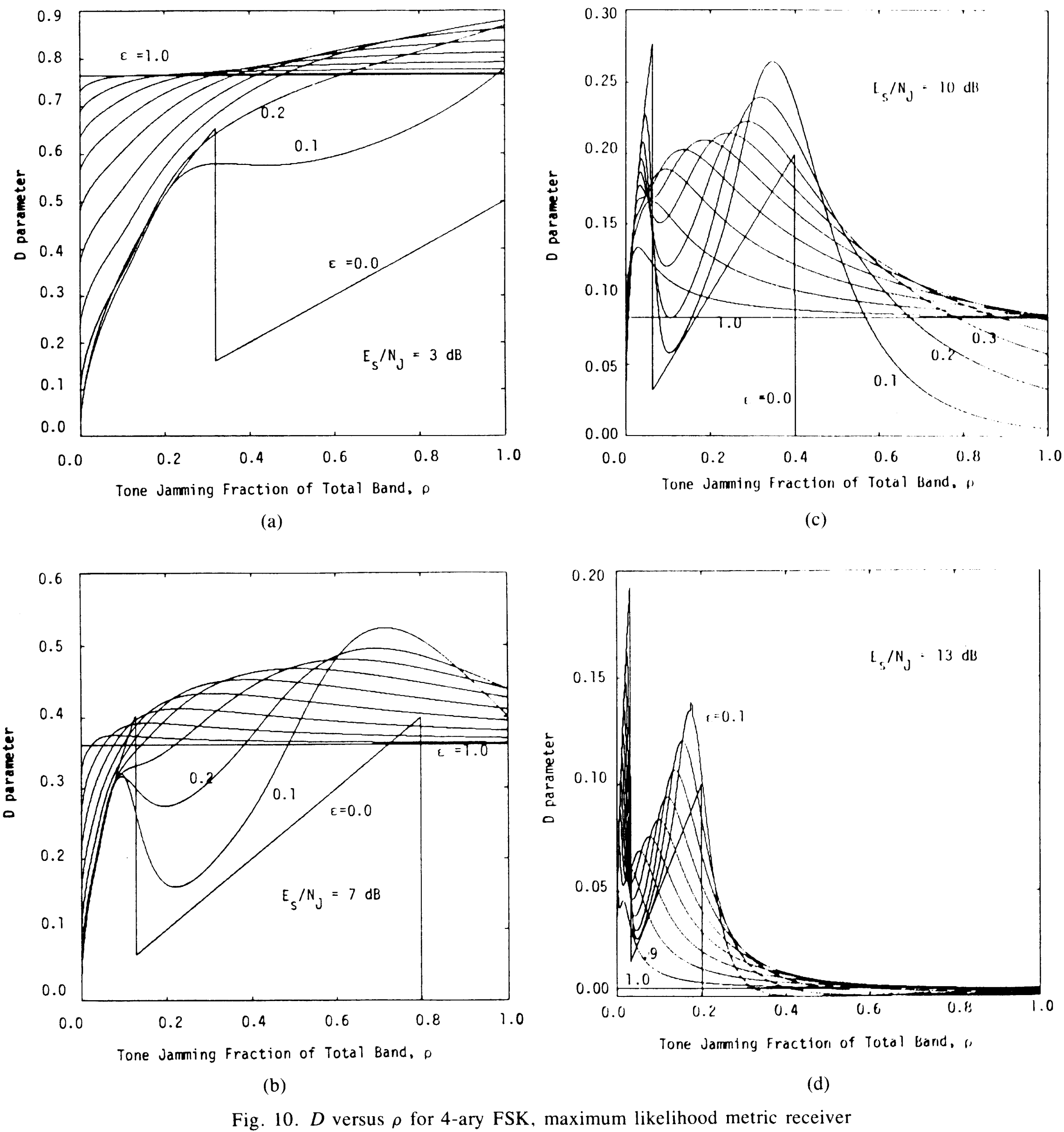


Fig. 9. D versus E_s/N_J for 4-ary FSK, hard decision without tone jamming state information and ϵ as a parameter (step 0.1).



(a)

(c)

(b)

(d)

Fig. 10. D versus ρ for 4-ary FSK, maximum likelihood metric receiver with Viterbi's ratio threshold techniques of threshold 2.5 and ϵ as a parameter (step 0.1) when E_s/N_J = 3 dB (a), 7 dB (b), 10 dB (c), 13 dB (d).

curves of $D(E_s/N_J, \epsilon, \rho)$ versus ρ are shown with ϵ as a parameter, for four different values of E_s/N_J.

For the extreme case $\epsilon = 0$ (the PBTJ-only case), we observe that there are two disconnected functions if $\rho_1 = M/(\theta^2 E_s/N_J) \le 1$, i.e., $E_s/N_J \ge M/\theta^2$. The first function is $D_1(\rho)$ in (41) for $0 \le \rho \le \rho_1$, and the second one is $D_2(\rho)$ for $\rho_1 < \rho \le$ minimum $(1, M/E_s/N_J)$ [see (41)]. These two pieces are monotonically increasing with ρ, and two local maxima occur at the ends of each piece. If $\rho_1 > 1$, i.e., $E_s/N_J < M/\theta^2$ (low E_s/N_J), then for the PBTJ-only with the worst-case ρ, D is equal to 1 [see (41)].

When the FBNJ begins to share total jamming power with the PBTJ (i.e., $\epsilon > 0$), the behavior of the $D(E_s/N_J, \epsilon, \rho)$ versus ρ curves in Fig. 10 becomes smooth, although we can still observe two local peaks for small ϵ. As ϵ approaches 1, we generally observe one peak.

$D(E_s/N_J, \epsilon, \rho^*)$ is plotted against ϵ with E_s/N_J as a parameter in Fig. 11, and versus E_s/N_J with ϵ as a parameter in Fig. 12, for 4FSK with the ML metric. In Fig. 12, we observe that PBTJ-only is the most efficient jamming, with $D = 1$, if $E_s/N_J \le M/\theta^2 = -1.93$ dB (i.e., low E_s/N_J), for $M = 4$ and $\theta = 2.5$ [see (41)]. We observe also that, for high E_s/N_J (for example, $E_s/N_J \ge 9.43$ dB in Fig. 12), the PBTJ alone ($\epsilon = 0$) is again the most efficient jamming against the ML metric with VRT. We explain this result as follows. Assume that only PBTJ is active at first. Now, let the FBNJ take a little jamming power of the total jamming power, the PBTJ loses that amount of jamming power. Then, the FBNJ contribution to the jammer is less sensitive to the type of receiver, while the PBTJ effect due to the reducing of PBTJ power is more sensitive if E_s/N_J is high. We can observe this sensitivity depending on the receiver type of comparing Fig. 12 to Figs. 6 and 9. From the comparison of the gap between the $\epsilon = 0$ and $\epsilon = 1$ extreme cases for high E_s/N_J in Figs. 6, 9, and 12, the HD metric without JSI is seen to be the most sensitive to an incremental PBTJ power change, the ML metric with VRT less sensitive, and the HD metric with perfect JSI is the least sensitive. Also, for high E_s/N_J, the ML metric with VRT behaves as the HD metric without JSI with reduced signal power. (Compare the first equation of (41) with (23) for $\epsilon = 0$ extreme case. They are equivalent.) This is why the PBTJ alone is the most efficient jamming against the ML metric with VRT for high E_s/N_J.

In the medium range of E_s/N_J, the behavior of the ML metric with VRT is similar to that of the HD metric with perfect JSI (see Figs. 6 and 12). The most efficient jamming involves a combination of tone and noise.

Notice that the D parameter is constant, 0.5, for the interval of 5 dB to 6 dB, when only the PBTJ is active (see Fig. 12). The reason for this result is that the maximum of two peaks in the middle equation of (41) is $D_2(1) = (M-2)/M$ (equal to 0.5 if $M = 4$) for this interval.

The corresponding results for $M = 2$ are shown in Figs. 13 and 14. The general behavior of the $M = 2$ case is similar to that of $M = 4$.

Many days of CPU time were used to obtain one $D(E_s/N_J, \epsilon, \rho^*)$ versus E_s/N_J curve for a given ϵ in the

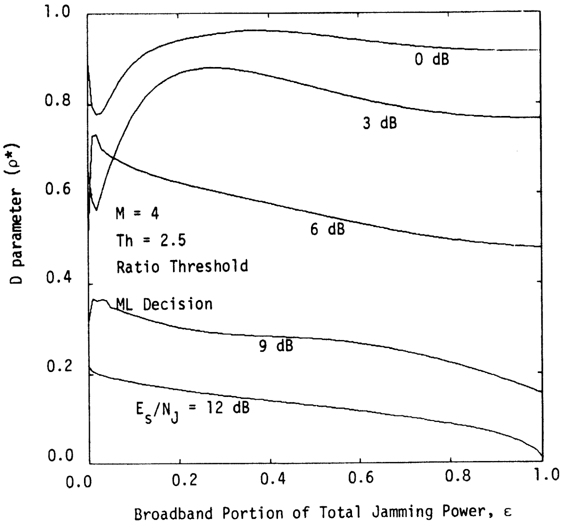


Fig. 11. D versus ϵ for 4-ary FSK, ratio threshold receiver with threshold 2.5 and E_s/N_J as a parameter.

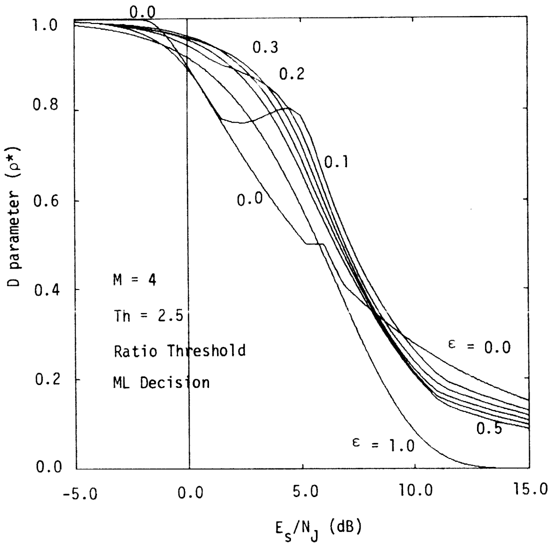


Fig. 12. D versus E_s/N_J for 4-ary FSK, ratio threshold receiver with threshold 2.5 and ϵ as a parameter (step 0.1).

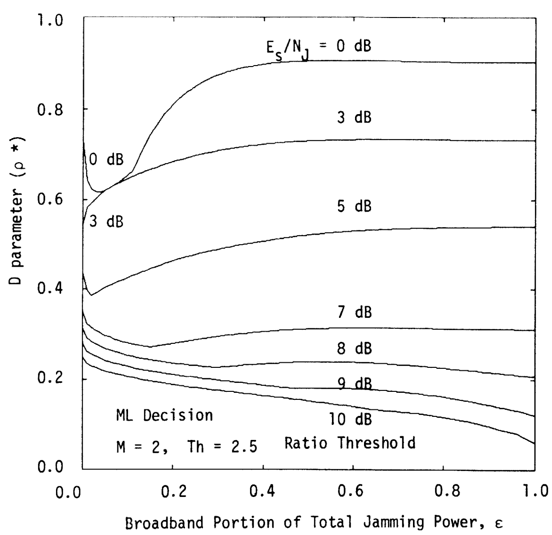


Fig. 13. D versus ϵ for binary FSK, ratio threshold receiver with threshold 2.5 and E_s/N_J as a parameter.

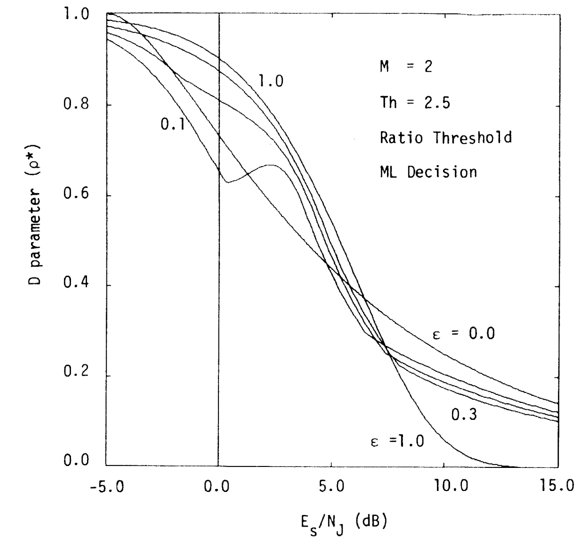


Fig. 14. D versus E_s/N_J for binary FSK, ratio threshold receiver with threshold 2.5 and ϵ as a parameter (step 0.1).

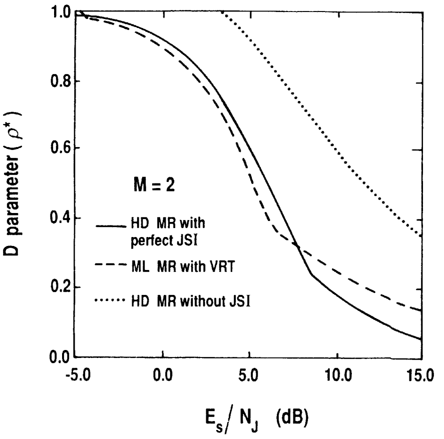


Fig. 15. Comparison of three receivers—hard decision metric receiver with perfect tone jamming state information, HD MR without JSI, and maximum likelihood MR with Viterbi's ratio threshold techniques in terms of the worst-case D versus E_s/N_J, when M = 2-ary FSK.

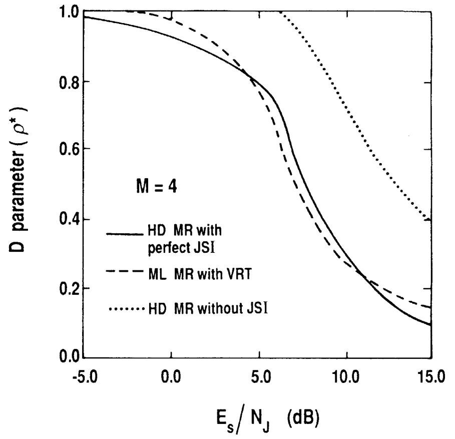


Fig. 16. Comparison of three receivers—hard decision metric receiver with perfect tone jamming state information, HD MR without JSI, and maximum likelihood MR with Viterbi's ratio threshold techniques in terms of the worst-case D versus E_s/N_J, when M = 4-ary FSK.

ML metric with VRT receiver when a VAX-780 machine was used with a time-sharing load of 2–3. We could not produce sufficient data for many ϵ's because of the long computation. With our data, it is, however, interesting to try to draw the upper envelopes of $D(E_s/N_J, \epsilon, \rho^*)$ versus E_s/N_J curves for the comparison of three receivers. They are shown in Figs. 15 and 16.

In Fig. 15 for BFSK, we observe that the ML metric with VRT is the best among three receivers for the medium range of E_s/N_J, -4.57 dB $< E_s/N_J <$ 5.73 dB, and the HD metric with perfect JSI is the best for $E_s/N_J \leq -4.75$ dB or $E_s/N_J \geq$ 5.73 dB. The difference between these two receivers is less than 1 dB in the E_s/N_J for a given D if E_s/N_J is less than 5.73 dB, and the difference increases for $E_s/N_J \geq$ 5.73 dB. The HD metric without JSI is the worst among the three receivers for any E_s/N_J, and 4 to 7 dB (or more) worse in E_s/N_J for a given D than the receiver with the second worst performance.

In Fig. 16, for 4-ary FSK, we observe that for the medium range of E_s/N_J, 5 dB $< E_s/N_J <$ 10.86 dB, the ML metric with VRT is the best, the HD metric with perfect JSI is the second best, and the HD metric without JSI is the worst. For $E_s/N_J \leq$ 5 dB or $E_s/N_J \geq$ 10.86 dB, the HD metric with perfect JSI is the best, the ML metric with VRT is second, the HD metric without JSI is the worst. The difference in E_s/N_J between the HD metric with perfect JSI and the ML metric with VRT is less than 2.3 dB for low or high E_s/N_J, and less than 1 dB for the medium range of E_s/N_J. The HD metric without JSI is 3.42 dB to 6.29 dB worse for a given D than the receiver with the second worst performance.

VI. Conclusions

We have considered combined tone and noise jamming against coded FH/MFSK systems under given total jamming power. The performances of three receivers were examined when combined jamming was used against them and evaluated in terms of the D parameter.

From the numerical analysis, we observe the following. As expected for the HD metric without JSI, PBTJ alone is the worst-case jamming from the communicator's view, as expected, since the receiver does not use JSI at all. For the HD metric with perfect JSI, if E_s/N_J is low, then FBNJ-only is the worst-case and PBTJ-only the least efficient jamming; otherwise, the most efficient portion of the total jamming power used in FBNJ is determined by the values of M and E_s/N_J. For the ML metric with VRT, if E_s/N_J is high, or low as $E_s/N_J \leq M/\theta^2$, then PBTJ-only is the worst-case; otherwise, the most efficient portion of the total jamming power used in FBNJ depends on M and E_s/N_J. If the three receivers are compared to each other when the worst-case combined jamming is used against them, then the HD metric without JSI is the worst receiver for any E_s/N_J, the HD metric with perfect JSI is the best receiver for high or low E_s/N_J, and the ML metric with VRT is the best receiver for the medium range of

E_s/N_J. The difference in E_s/N_J between the case of the HD metric with perfect JSI and the case of the ML metric with VRT is less than 1 dB for the medium range of E_s/N_J to achieve the same D.

As even more powerful hybrid jamming scenario is the combination of PBTJ and two-level partial-band noise jamming in place of the less effective FBNJ. However, the corresponding analysis is much more difficult, and this major extension of the authors' current paper is reserved for a follow-up effort, which will also include $M \geq$ 8-ary FSK results.

Acknowledgment

The authors would like to thank Dr. L. E. Miller for his valuable comments and English corrections in editing the paper.

References

[1] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications*, Vol. I, II, and III. Rockville, MD: Computer Science Press, 1985.

[2] J. K. Omura and B. K. Levitt, "Coded error probability evaluation for antijam communication systems," *IEEE Trans. Commun.*, vol. COM-30, pp. 896-903, May 1982.

[3] W. E. Stark, "Coding for frequency-hopped spread spectrum communication with partial-band interference—Part I: Capacity and cutoff rate," *IEEE Trans. Commun.*, vol. COM-33, pp. 1036-1044, Oct. 1985.

[4] A. J. Viterbi, "A robust ratio-threshold technique to mitigate tone and partial-band jamming in coded MFSK system," in *Proc. IEEE MILCOM*, Oct. 1982, pp. 22.4-1-22.4-5.

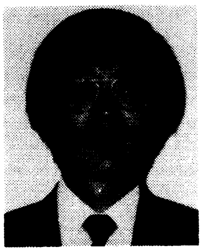
[5] K. M. Clifford and T. A. Schonhoff, "Performance evaluation of ratio threshold for frequency hopped MFSK communication in partial-band tone jamming plus background noise," in *Proc. IEEE MILCOM*, Oct. 1986, pp. 12.3.1-12.3.5.

[6] M. J. Massaro, "Error performance of *M*-ary noncoherent FSK in the presence of CW tone interference," *IEEE Trans. Commun.*, vol. COM-23, pp. 1367-1369, Nov. 1975.

[7] B. K. Levitt, "FH/MFSK performance in multitone jamming," *IEEE J. Select. Areas Commun.*, vol. SAC-3, pp. 627-643, Sept. 1985.

[8] L. Chang and R. J. McEliece, "A study of Viterbi's ratio-threshold AJ technique," in *IEEE MILCOM'84 Conf. Rec.*, Oct. 1984, pp. 11.2.1-11.2.5.

Hyuck M. Kwon (S'82-M'84), for a photograph and biography, see this issue, p. 761.



Pil Joong Lee (S'79-M'82-SM'87) was born in Masan, Korea, in 1951. He received the B.S. and M.S. degrees in electronics engineering from Seoul National University, Seoul, Korea, in 1974 and 1977, respectively, and the Engineer degree in system science in 1982 and the Ph.D. degree in electrical engineering in 1985, both from the University of California, Los Angeles.

He worked as a Member of Technical Staff at the Communications Systems Research Section of the Jet Propulsion Laboratory in Pasadena, CA, from 1980 to 1985, where he mainly worked for coded coding systems analyses and developments. In 1985 he joined the Multimedia Communications Research Division of the Bell Communications Research, where he is currently responsible for research on providing privacy and security for telephone users. His research interests include cryptology, computer and network security, error correction coding, spread spectrum communications, modulations, and information theory.

Dr. Lee is a member of Eta Kappa Nu, the International Association of Cryptologic Research, Office Automation Society International, and Korean Scientists and Engineers Association in America.

Performance of FH/BFSK with Generalized Fading in Worst Case Partial-Band Gaussian Interference

PAUL J. CREPEAU, SENIOR MEMBER, IEEE

Abstract—For FH/BFSK on a worst case partial-band Gaussian interference channel, the bit error probability results are well known for the extreme cases where the signal is either nonfading or Rayleigh fading. This paper fills in the region between these extremes by considering the general Nakagami-m fading model. Here the worst case partial-band Gaussian interference results are given by a one-parameter family which for $m \rightarrow \infty$ gives the Viterbi-Jacobs nonfading result, and for $m = 1$ gives the Rayleigh fading result. In the latter case, a broadband interference strategy is optimal. Thus, the Nakagami-m results provide a smooth one-parameter bridge between the Viterbi–Jacobs channel and the Rayleigh fading channel. The results show that the worst case interference fraction ρ increases as the fading variance increases, up to Rayleigh fading. Any fading less severe than Rayleigh, however slight the departure from Rayleigh, requires a partial-band strategy for sufficiently large E_b/N_I.

I. Introduction

FOR frequency-hopped noncoherent binary frequency shift keying (FH/BFSK) on a worst case partial band Gaussian interference channel, the bit error probability performance results are well known for the extreme cases where the signal is either nonfading or Rayleigh fading. In this paper, we construct a smooth one-parameter bridge between the nonfading and Rayleigh fading channels by using the Nakagami-m fading generalization. With the Nakagami-m fading model, we varying a single parameter m from one to infinity to obtain a continuous transition from a Rayleigh fading channel to a worst case partial band Gaussian interference channel.

In the next section we establish the basic channel definitions and review the known results. In Section III, we give new results for the generalized channel. Finally, in Section IV, we consider extension of the results to other modulation types.

II. Background and Definitions

Viterbi and Jacobs [1] derived the probability of error for worst case partial-band Gaussian interference of FH/BFSK when the transmitted signal is received with constant amplitude. The partial-band interference model is characterized by additive Gaussian noise interference with flat spectral density N_I/ρ over a fraction ρ (where $0 < \rho \leq 1$) of the total hopping bandwidth and negligible interference over the remaining fraction $(1 - \rho)$ of the band. Background thermal noise is neglected. It is assumed that the two candidate tone slots are either both interfered with, or they are both noise free. As a worst case condition, the parameter ρ is chosen so as to maximize the probability of error. Finally, it is assumed that the channel is memoryless, which can be achieved by hopping once per tone symbol or by interleaving. With these assumptions, the well-known probability of error result of Viterbi and Jacobs is

$$P_b = \begin{cases} \dfrac{1}{2}\exp(-E_b/2N_I); & \dfrac{E_b}{N_I} \leq 2, \quad \rho = 1 \\ \dfrac{e^{-1}}{E_b/N_I}; & \dfrac{E_b}{N_I} \geq 2, \quad \rho = \dfrac{2}{E_b/N_I}. \end{cases} \tag{1}$$

Although worst case partial-band Gaussian noise interference is highly detrimental to nonfading signals (changing the exponential dependence of P_b on E_b/N_I to an inverse linear dependence), it has been shown by Omura [2] that if the FH/BFSK signal has Rayleigh amplitude fading, then the worst case partial-band Gaussian interference is always full band interference, independent of E_b/N_I. Here the result becomes that of the standard Rayleigh fading channel with additive white Gaussian interference, that is,

$$P_b = \frac{1}{2 + E_b/N_I}; \quad \text{all } \frac{E_b}{N_I}, \quad \rho = 1 \tag{2}$$

where E_b is the average energy per bit. Omura's result shows that, for Rayleigh fading, the performance is so poor for full band interference that there is no advantage for the interferer to concentrate on a fraction of the hopping bandwidth. In this paper, it will be shown that a partial-band strategy is required for any fading less severe than Rayleigh fading. That is, we show that Rayleigh fading is the limiting case where a full band strategy is optimal for the interferer regardless of signal-to-noise ratio.

The Viterbi–Jacobs result in (1) and the Rayleigh fading channel result in (2) have the common characteristic that the probability of bit error, P_b, is inversely proportional to the signal-to-noise ratio, E_b/N_I, for large values of E_b/N_I. This similarity of behavior for these channels has been noted frequently, but the exact relationship be-

Manuscript received February 6, 1989; revised October 10, 1989. This paper was presented at MILCOM '89, Boston, MA, October 1989.

The author is with the Information Technology Division, Naval Research Laboratory, Washington, DC 20375.

IEEE Log Number 9035813.

0733-8716/90/0600-0884$01.00 © 1990 IEEE

tween these channels has been only vaguely understood. Imprecise statements are sometimes made that a partial-band interference can make a nonfading channel behave like a Rayleigh fading channel. In this paper we clarify this relationship and give a precise connection between these two well-known channels.

III. Worst Case Interference for Generalized Fading

There are two fading channel generalizations which enable us to traverse continuously the range of channels from Rayleigh fading to nonfading. These are the Rician (noncentral chi with two degrees of freedom) and the Nakagami-m (central chi with $2m$ degrees of freedom, generalized to noninteger m). In this paper we choose the Nakagami-m generalization because it leads to results of greater mathematical simplicity. This simplicity of results is characteristic of the Nakagami-m channel and has been observed in a variety of situations starting with the original work of Barrow [3] (see also [4]) and continuing in several recent papers [5]–[7].

For orthogonal BFSK on a Nakagami-m fading channel with partial-band interference, the received signal in the interval $(0, T)$ has the form

$$r(t) = R\sqrt{2E_b/T}\cos(\omega_i t + \theta) + n_I(t) \tag{3}$$

where E_b is the average received bit energy (under a normalization described below), $n_I(t)$ is the partial-band Gaussian interference process, θ is a uniform random variable $(0, 2\pi)$, ω_i is one of two orthogonally spaced frequencies, and R is a Nakagami-m random variable whose pdf is given by

$$p(R) = \frac{2m^m R^{2m-1}}{\Gamma(m)\Omega^m}\exp(-mR^2/\Omega), \qquad R \geq 0 \tag{4}$$

where

$$\Omega = \overline{R^2},$$

and

$$m = \frac{\Omega^2}{\operatorname{var}(R^2)} \geq \frac{1}{2}$$

are the two parameters of the distribution. For $m = 1$ we have a Rayleigh fading channel, and as $m \to \infty$ we have a channel that becomes nonfading (as the pdf tends to an impulse function).

We restrict our discussion to slow nonselective fading. With no loss of generality, we adopt the convenient normalization $\Omega = 1$ so that the received energy in the fading channel R^2E_b has an average value $\overline{R^2E_b} = \Omega E_b = E_b$. With $\Omega = 1$, the pdf in (4) is reduced to a one-parameter distribution so that all of the results can be expressed in terms of the single parameter m.

The partial-band interference process $n_I(t)$ in (3) is the same as that used in the Viterbi–Jacobs channel model. That is, for a fraction ρ of the hops, the interference is Gaussian with flat spectral density N_I/ρ, and the interference is zero over the remaining fraction $1 - \rho$ of the band. Furthermore, the channel is assumed memoryless from bit to bit.

Following Viterbi and Jacobs, we start with the broadband results and then proceed to derive the partial band results. For broad-band Gaussian interference with constant spectral density N_I, the performance of noncoherent BFSK on a Nakagami-m fading channel was found by Barrow [3], [4] to be

$$P_b = \frac{1}{2}\left[\frac{m}{m + \frac{1}{2}\frac{E_b}{N_I}}\right]^m. \tag{5}$$

For partial-band interference with arbitrary interference fraction ρ, this can be modified to become

$$P_b(\rho) = \frac{\rho}{2}\left[\frac{m}{m + \frac{\rho}{2}\frac{E_b}{N_I}}\right]^m. \tag{6}$$

As a worst case condition, we choose ρ so as to maximize the probability of bit error, subject to the constraint $(0 < \rho \leq 1)$. By differentiating the right-hand side of (6) with respect to ρ and setting it equal to zero (along with verifying that the second derivative is negative), we find the worst case probability of error and associated worst case interference fraction to be given by

$$P_b = \begin{cases} \dfrac{1}{2}\left[\dfrac{m}{m + \frac{1}{2}\frac{E_b}{N_I}}\right]^m; & \dfrac{E_b}{N_I} \leq 2\left(\dfrac{m}{m-1}\right), \quad \rho = 1 \\[2ex] \dfrac{\left(\frac{m-1}{m}\right)^{m-1}}{E_b/N_I}; & \\[2ex] \dfrac{E_b}{N_I} \geq 2\left(\dfrac{m}{m-1}\right), & \rho = \dfrac{2}{E_b/N_I}\left(\dfrac{m}{m-1}\right). \end{cases} \tag{7}$$

The results of (7) are presented graphically in Fig. 1 for selected values of m. On each curve we indicate by a dot the transition point between the two functional forms in the right-hand side of (7). This point marks the onset of inverse linear behavior of P_b as we increase E_b/N_I. In the inverse linear region, the multiplicative constant $[(m - 1)/m]^{m-1}$ decreases monotonically from 1 to e^{-1} as m increases from 1 to ∞. For $m \to \infty$, the result in (7) becomes identical to the Viterbi-Jacobs result of (1), and for $m = 1$ it is identical to the Rayleigh result of (2). Thus, the Nakagami-m solutions provide a smooth one-parameter bridge between the nonfading Viterbi-Jacobs channel and the Rayleigh fading channel. The results show that the worst case interference fraction increases as the

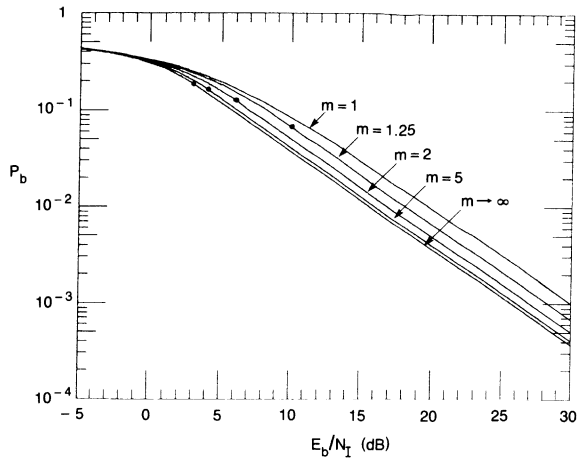


Fig. 1. Performance of FH/BFSK with worst-case partial band Gaussian interference in a Nakagami-*m* fading channel.

fading variance increases, up to Rayleigh fading. Any fading less severe than Rayleigh, however slight the departure from Rayleigh, requires a partial-band strategy for sufficiently large E_b/N_I.

IV. Related Problems

The probability of error for frequency-hopped differentially coherent phase shift keying (FH/DPSK) on a memoryless Nakagami-*m* channel with worst case partial-band Gaussian interference is given by (7), replacing E_b by $2E_b$ everywhere. This follows from Barrow's original DPSK result which is given by (5) with E_b replaced by $2E_b$. Graphically, the new DPSK worst case interference result for Nakagami-*m* fading is given by Fig. 1 with all curves shifted 3 dB to the left.

The basic approach used for FH/BFSK can be used for the more general case of *M*-ary orthogonal signaling alphabets. For FH/MFSK in Nakagami-*m* fading with additive broad-band Gaussian interference, the exact probability of bit error expression is given in [7]. By applying the Viterbi–Jacobs procedure to this result, we conjecture that worst case partial band results can be found which form a smooth one-parameter transition from the Rayleigh fading to the nonfading channels. For the *M*-ary case, the level of computational difficulty increases greatly with increasing alphabet size, and the numerical results remain to be determined.

References

[1] A. J. Viterbi and I. M. Jacobs, "Advances in coding and modulation for noncoherent channels affected by fading, partial-band and multiple-access interference," in *Advances in Communication Systems*, Vol. 4. New York: Academic, 1975.

[2] J. K. Omura, "Variable data bit rates with a fixed hop rate noncoherent FH/MFSK system," in *Proc. Int. Telem. Conf.*, Oct. 13–15, 1981, pp. 1445–1463.

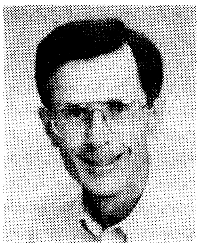
[3] B. B. Barrow, "Error probabilities for data transmission over fading radio paths," Tech. Memo. TM-26, SHAPE Air Defence Center, The Hague, The Netherlands, 1962.

[4] P. F. Panter, *Communication Systems Design*. New York: McGraw-Hill, 1972, Section 18.3.2.

[5] A. H. Wojnar, "Unknown bounds on performance in Nakagami channels," *IEEE Trans. Commun.*, vol. COM-34, pp. 22–24, 1986.

[6] P. J. Crepeau, "Asymptotic performance of *M*-ary orthogonal modulation in generalized fading channels," *IEEE Trans. Commun.*, vol. 36, pp. 1246–1248, 1988.

[7] ——, "Coding performance on generalized fading channels," in *Conf. Proc. MILCOM'88*, San Diego, CA, Oct. 1988.



Paul J. Crepeau (M'58–SM'84) was born in Woonsocket, RI, on October 17, 1935. He received the B.S.E.E. degree from the University of Rhode Island, Kingston, in 1957, and the M.E.E. and Ph.D. degrees in electrical engineering from the Polytechnic Institute of Brooklyn, Brooklyn, NY, in 1960 and 1967, respectively.

From 1957 to 1961 he was employed by the Sperry Gyroscope Company, Great Neck, NY, and from 1961 to 1972 he was a member of the Department of Electrical Engineering at the Polytechnic Institute of Brooklyn. In 1972 he was appointed Presidential Intern at the U.S. Army Electronics Command, Ft. Monmouth, NJ. Since 1973 he has been a member of the Information Technology Division of the U.S. Naval Research Laboratory, Washington, DC. He has also taught many courses in the Washington, DC, area, and he is currently an Adjunct Professor of Electrical Engineering at the George Washington University.

On the Symbiotic Nature of Key Antijamming and Antiscintillation Functions for MFSK and DPSK Channels

RICHARD A. YOST, MEMBER, IEEE

Abstract—**Differentially coherent phase shift keying (DPSK) and *M*-ary frequency shift keying (MFSK) are modulation formats widely used in frequency hopping applications. These formats, coupled with diversity techniques such as multiple transmissions and convolutional coding with Viterbi decoding and interleaving, prove to be very proficient antijamming (AJ) and antiscintillation (AS) techniques. It is interesting to note that the relationships between the detector, diversity combiner, and soft decision generator are unique in the way they respectively contribute to providing the AJ and AS capability. This paper places its primary focus on describing this symbiotic relationship and does so with an approach that permits possible architectures for detection, diversity combining, and soft decision generation to surface as its secondary focus. Considerable time will be spent on the soft decision generators for DPSK and MFSK channels since such generators present a variety of implementation problems related to automatic gain control and jamming sensitivity.**

Introduction

DIFFERENTIALLY coherent phase shift keying (DPSK) and *M*-ary frequency shift keying (MFSK) are modulation formats widely used in frequency hopping applications. The attribute that makes these formats so appropriate for frequency hopping systems is that their individual detections may be carried out in a differentially coherent or noncoherent manner. However, simply combining one of these formats with a frequency hopping synthesizer does not guarantee a successful antijamming (AJ) or antiscintillation (AS) system. It has been previously shown [1] that an optimized jammer can force the bit-error-rate (BER) dependence on signal-to-noise ratio to one that is inversely linear as opposed to one that is inversely exponential. An optimized jammer that causes such a dependency change can reduce the system jamming margin by 30 dB or more. Likewise, it has been shown [2] that a similar fate awaits the BER performance in a scintillating (or fading) channel as well. To counter such jamming and fading effects, frequency hopping systems usually incorporate diversity that may take the form of repetitive transmissions of the same symbol and subsequent combining of said transmissions at the receiver, or forward error correction coding. The latter is preferred from a performance basis because of its greater efficiency in diversity combining [3], although in practice a hybrid scheme combining the attributes of both is usually more feasible.

In these frequency hopping types of systems, there exists a symbiotic relationship between the elements of those systems, particularly the channel characteristics, the diversity approach, the modulation format, the detector implementation, and, for the case of the error correction coding, the generator of the decoder soft decisions. Previous literature, examples of which include [7], [8], [10], [12], [13], has considered various subsets of the above elements and identified unique relationships between the elements within those subsets. For example, in [7] various diversity combining receivers were evaluated in partial-band jamming, or in [10] an analysis of different soft decision metrics was conducted in a pulse jamming channel. In contrast, the purpose of this paper is to focus on the relationship encompassing all of the above elements simultaneously, to provide comparative results for both the jamming and fading channel, and to draw conclusions that highlight the symbiotic nature of this relationship. Specific attention is given to those systems incorporating DPSK and 8-ary FSK modulation formats and rate $1/2$, $K = 7$ convolutional coding with Viterbi decoding.

Throughout this paper, key functional architectures for diversity combining, detection, and soft decision generation will be offered as examples and potential candidates for implementation. In some cases, these architectures, or techniques, will be no different from those previously reported in [7], [8], [12], and [13]; however, the primary focus is not on the techniques themselves, but on the relationships between the elements comprising a frequency hopping system. Because soft decision decoding presents a variety of implementation problems related to such areas as automatic gain control (AGC) and jamming sensitivity, the paper addresses these problems specifically and proposes soft decision techniques that minimize their effect. It is not the intent of this paper to provide an exhaustive tradeoff analysis of the key functional architectures, and so the conclusions drawn from the limited data provided herein should be consumed accordingly. However, arguments will be presented to demonstrate the utility of the architectures and the credibility they offer with respect to optimality.

Manuscript received February 17, 1989; revised November 8, 1989.

The author is with the Government Communications Systems Division, Harris Corporation, Melbourne, FL 32902.

IEEE Log Number 9035814.

DPSK Channel

There are two elements of a DPSK receiving system that will be addressed here. They are the detector and the soft decision generator.

DPSK Detector

Detection of DPSK corresponds to determining the presence or absence of a phase change from one bit to the next. This can be accomplished in the traditional matched filter approach by passing the incoming signal through two filters having impulse responses as those shown in Fig. 1. A power-dependent functional calculation is performed on each filter output, and in the uncoded situation a comparison is made to determine the bit polarity. This is shown in Fig. 2 for the case of an envelope detector. For the coded situation, which is of greatest interest herein, the two outputs are subsequently processed by a soft decision generator prior to the decoding function and instead of the comparison process.

Note in Fig. 1 that the filters perform the detection over a two-channel-symbol time window. When combined with the power-dependent calculation, these filters essentially perform a noncoherent detection of two orthogonal waveforms, which is an accurate signal representation of DPSK over the duration of two symbols. It is this DPSK orthogonality over two symbols that gives rise to the 3 dB performance advantage DPSK has over noncoherently detected binary FSK in an additive white Gaussian noise channel.

An implementation architecture does evolve directly from the above functional description; the most straightforward approach is shown in Fig. 3(a). In this architecture, the IF-carried DPSK signal is downconverted to baseband, both I and Q branches are integrated-and-dumped (I&D) and sampled, a matched filter envelope detection is performed for each hypothesis, and a difference-comparison with zero is conducted. It is assumed in Fig. 3(a), and without a loss of generality, that the bit "0" corresponds to the lack of a phase transition, and the bit "1" to the occurrence of a phase transition.

The difference-comparison process motivates a second architecture that is mathematically equivalent and considerably less arithmetically intense; it is shown in Fig. 3(b). Assuming the bit "0" is transmitted, the mathematical equivalence can be demonstrated as follows:

$$R_0 < R_1$$

$$R_0^2 < R_1^2$$

$$I_2^2 + 2I_2I_1 + I_1^2 + Q_2^2 + 2Q_2Q_1 + Q_1^2$$
$$< I_2^2 - 2I_2I_1 + I_1^2 + Q_2^2 - 2Q_2Q_1 + Q_1^2$$
$$I_2I_1 + Q_2Q_1 < 0.$$

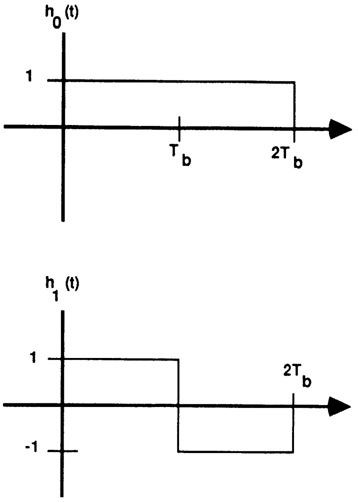


Fig. 1. Impulse response of matched filters for DPSK.

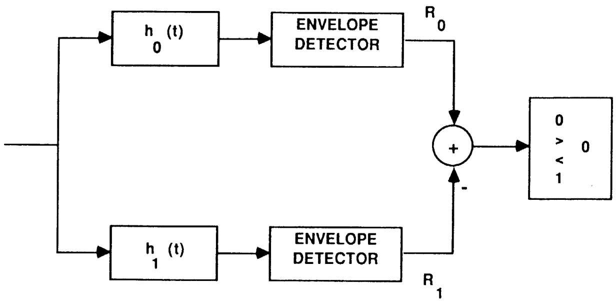
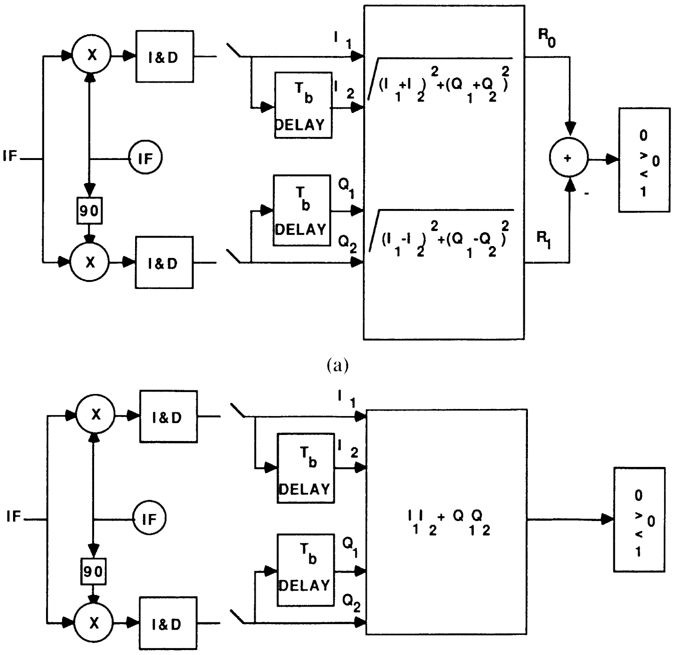


Fig. 2. Generic matched filter detector for DPSK.



(a)

(b)

Fig. 3. Two implementation architectures for the DPSK matched filter detector.

When consideration is given only to the detection process, and since both of the above architectures produce the same output statistics and thus the same BER, the latter architecture is preferred because of the lesser complexity. However, once coding is included into the system, the desired detection architecture is dependent upon the soft decision generator which in turn affects performance. Thus, the first evidence of the interrelationship spoken of earlier has surfaced. In order to see this more quantitatively, the soft decision generator for DPSK will be presented.

DPSK Soft Decision Generation

In lieu of the difference and comparison process in the functional architecture of Fig. 2, consider a 3-bit soft decision generator. There are at least two methods of producing reliability information from the two matched filter outputs. One technique would form the difference $R_0 - R_1$ (or $R_0^2 - R_1^2$) and then quantize the result, while the second technique would form the ratio R_0/R_1 and then quantize the result. Note that the former technique is made possible with both of the implementation architectures presented since both are capable of generating a difference. On the other hand, the specific values of R_0 and R_1 are necessary to generate the ratio R_0/R_1. Therefore, only the architecture in Fig. 3(a) can be used for generating a ratio soft decision.

Since the statistics out of the soft decision generators are different, BER performance can possibly be different. In order to assess the performance of these two soft decision generators, a detector/soft decision generator/decoder simulation was constructed. The simulation is a direct implementation based on Monte Carlo principles. The code used is a rate 1/2, constraint length 7 (generators: 171,133 [9]) convolutional code. Fig. 4 presents the BER results for the two generators (3 bit quantization is used) in an additive white Gaussian noise environment. Here it can be seen that the difference soft decision generator is approximately 0.5 dB more efficient than the ratio approach in this benign channel.

A similar result is experienced when the coded, DPSK, signal is transmitted through a Rayleigh fading channel, which is typical of certain scintillating environments. Fig. 5 illustrates the BER results for the difference and ratio soft decision generators in this Rayleigh fading channel where each DPSK bit is assumed to be independently faded in the computer simulation. In this instance, the performance difference between the two soft decision generators can be attributed to the optimality of a linear diversity combiner [11, p. 538] in a Rayleigh fading channel. Since the maximum likelihood decoder is essentially a diversity combiner, one would assume that a soft decision generator which could preserve the linearity of the input would produce superior performance. That is exactly what has been illustrated here since the ratio technique is a nonlinear operation in contrast to the linear operation performed by the difference technique. As will be shown next, the ratio does exhibit powerful capabilities in a jamming environment.

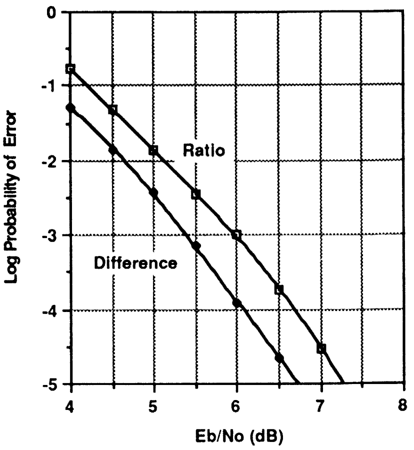


Fig. 4. Ratio versus difference soft decisions for DPSK in additive white Gaussian noise.

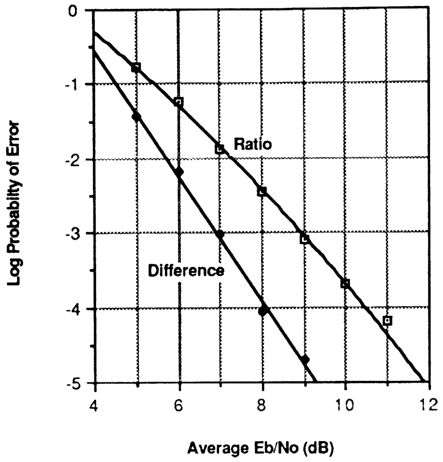


Fig. 5. Ratio versus difference soft decisions for DPSK in a Rayleigh fading channel.

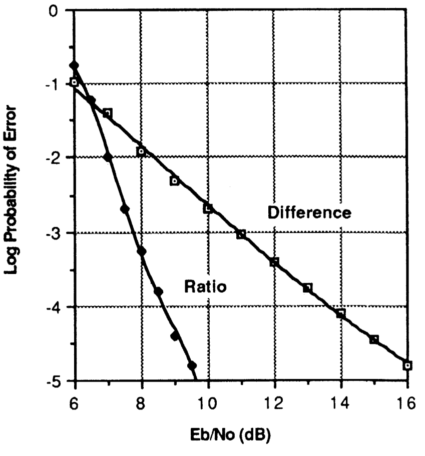


Fig. 6. Ratio versus difference soft decisions for DPSK in optimized partial band noise jamming.

Fig. 6 illustrates the difference and ratio soft decision comparison for worst case jamming threats with a maximum jamming-power-to-thermal-noise-power of −3 dB. This value was chosen for illustrative purposes and should not be considered the worst case instantaneous ratio; the jammer is permitted to trade average for peak power and full band for partial band strategies. The performance curves reflect the optimum jamming strategy in that the jammer is permitted to select the percentage of the band jammed for greatest degradation. Note that the difference technique suffers significantly under jamming while the ratio tends to neutralize the jamming optimization through its inherent normalization [12], [13]. Whereas the difference statistics vary significantly enough to "deoptimize" the quantization thresholds, the normalization of the ratio approach decouples the jamming strategy from the threshold settings.

What these three performance comparisons indicate is that a system designer must first decide which channel environment is the most critical within the system under consideration. Once that is decided, the soft decision generator is known and, because of the close relationship between detection and soft decision generation, the designer also knows which of the two detectors must be used. For example, if performance in a jamming channel is the most critical performance element of a system, the system designer is led to the use of the ratio soft decision technique and subsequently to the more complex detector implementation.

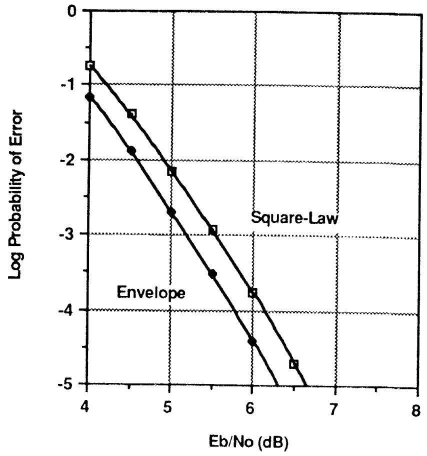


Fig. 7. Square-law versus envelope detector for 8-ary FSK in optimized partial band tone jamming (second-order diversity).

MFSK Channel

The symbiotic relationship between the receiving elements for an 8-ary FSK channel will be discussed in a manner similar to that just conducted for DPSK. In contrast, however, it is over the MFSK channel that multiple transmissions of the same symbol can be made and subsequent diversity combining performed. As a result, this section will add an additional element into the interrelationship of the receiving elements, that is, the diversity combiner. Initial focus will be placed on the detector.

8-ary FSK Detector

Detectors for 8-ary FSK signals traditionally incorporate a bank of filters, each of which has a center frequency matched to one of the 8 frequencies and whose duration is equal to the signal duration, followed by a device that evaluates either the power or some function related to the power out of each filter. Practically, this filter bank can be accomplished with SAW filters [4], fast Fourier transforms [5], or separately with lumped-element filters. In this section we will not dwell on the benefits of these detailed implementation techniques for the filter bank, but rather on the power related device and what function it needs to perform.

In the DPSK receiver, it made little difference what power-related function was used because the soft decision thresholds could be optimized for each type with very little performance variation. However, the diversity combining that takes place subsequent to the MFSK detector does have an impact on the type of detector used in the MFSK receiver. The major tradeoff in the detector area is whether to use a magnitude (envelope) or a magnitude-squared (square-law) type of detector. It was felt that the envelope detector, being more compressive than the square-law detector, would be more beneficial in the suppression of high-powered jammers in a diversity receiver. These benefits are supported in Fig. 7, where it can be seen that the envelope detector is approximately 0.6 dB better at 10^{-5} BER than the square-law detector. In Fig. 7, second-order time diversity (2 hops per 8-ary symbol), worst case tone jamming with $J/N = -3$ dB and optimum soft decision thresholds were assumed in a Monte Carlo computer simulation that included the FSK detector, diversity combiner, and decoder. The actual soft decision algorithm will be discussed later. Although the amount of variation may change as J/N or the diversity order change, no evidence surfaced to indicate that the envelope detector was any more or less effective than that shown in Fig. 7.

A similar simulation comparison was conducted for these detectors in a scintillating channel that is described as a slow, independent, Rayleigh fading environment. Again the FSK detector, diversity combiner, and decoder (rate $1/2$, $K = 7$) are integrated into the simulation. With a diversity order of 8, Fig. 8 illustrates the comparison between a square-law and envelope detector. In contrast to the jamming results, the square-law detector appears to be slightly superior in performance over the envelope detector in the fading environment. While this difference is minor, the superior square-law performance can again be traced back to the optimum receiver being of the square-law type [11, p. 538].

While the specific examples cited above cannot substantiate general conclusions regarding the "best" detector, they have pointed out that unlike the DPSK receiver, there does appear to be a performance difference between different types of power-related detectors, and that this difference has been caused primarily by the need to perform diversity combining. Furthermore, the examples have pointed out that the channel through which the diversity transmissions must pass can significantly influence the performance and choice of the "best" detector: envelope detector for jammed channels [10] and square-law detector for fading channels.

Diversity Combining

A simple, multihop, diversity receiver for 8-ary FSK is shown in Fig. 9. It is extremely sensitive to high-powered jamming signals dominating the accumulated output even when the jammer "hits" only one of the N hops. Some

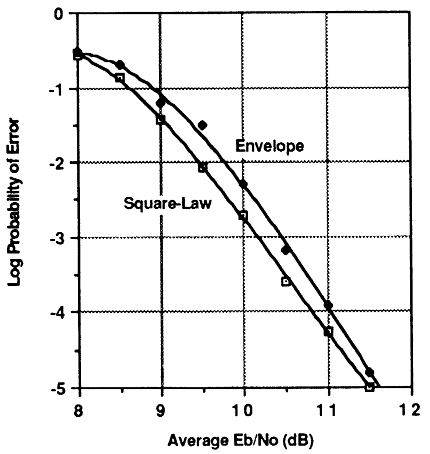


Fig. 8. Square-law versus envelope detector for 8-ary FSK in Rayleigh fading (8th-order diversity).

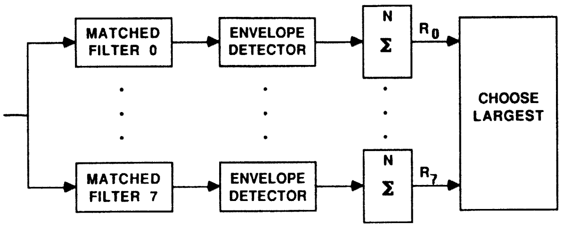


Fig. 9. Linear diversity combiner (receiver 1).

method is needed to mitigate this dominance and force the jammer to "hit" more hops than just one and to do so with less power.

One such method is shown in Fig. 10, where a signal weighting has been applied to the received signal. The weighting is the instantaneous S/N that exists during each diversity element (or hop). When the S/N is low, as during jamming, the weight is correspondingly low so that the influence on the accumulated output is minimized. When the S/N is high, as during nonjamming conditions, the weight is correspondingly high so that the influence on the accumulated output is maximized. Such a weighting process has been shown to be optimum [6] for partial band noise jamming. And while not proven to be optimum for tone jamming, a similar weighting would seem appropriate. In either case, the generation of an instantaneous and accurate S/N weighting is quite difficult, if not impossible. Nonetheless, the consideration of this method is appropriate for it establishes the performance benchmark for comparison with practical implementations [12].

A number of "weighting" techniques have been evaluated in the literature [7]. The weighting function used in this paper is known as signal normalization and is shown in Fig. 11. In this method, and prior to the accumulators, the largest 8-ary FSK detector output on each hop is selected and used to normalize to it all the outputs on that hop. As a result, the largest value into the accumulators is the value one, originating from that detector output which was the largest on that hop. Moreover, a high-powered jammer that "hits" just one hop is limited in its effect on the accumulated output to a value of unity.

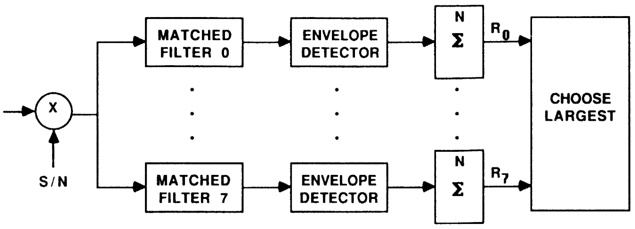


Fig. 10. S/N weighted diversity combiner (receiver 2).

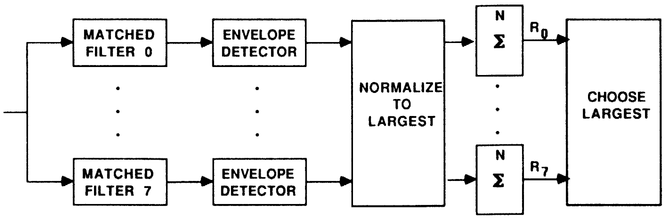


Fig. 11. Normalized diversity combiner (receiver 3).

Figs. 12 and 13 compare these three diversity receivers (no coding) in *worst case* partial band noise jamming and independent multitone tone jamming [15] environments, respectively. The results are worst case in that the percentage of the band jammed has been optimized in favor of the jammer. Exact expressions [8] were used for the noise jamming performance for receivers 1 and 2, whereas the nonlinear signal normalization of receiver 3 required computer simulations. Furthermore, instead of relying on loose Chernoff bounding techniques to assess the tone jamming performance, computer simulations of all three receivers were conducted for this specific jammer. Under tone jamming conditions, the weighting in receiver 2 is the signal-to-tone jamming ratio when jammed and signal-to-thermal-noise ratio when not jammed. For illustrative purposes, fourth-order diversity has been used and no error correction coding is assumed.

Even though no error correction coding is used in this comparison, the point at which the comparison is extracted is however based on coding; that point being the channel error rate required to achieve a decoded BER of 10^{-5} with a rate $1/2$, $K = 7$ Viterbi decoder using hard decisions. From [9], that channel error rate is 0.0175. From the noise jamming case of Fig. 12, receiver 1 requires an E_b/N_o of 5.7 dB whereas receivers 2 and 3 require 5.2 dB and 5.4 dB, respectively. From the tone jamming case of Fig. 13, the respective E_b/N_o values are 11.0 dB, 4.8 dB, and 5.0 dB. Under noise jamming, the optimum S/N weighting receiver (receiver 2) betters the simple linear diversity receiver (receiver 1) by only 0.5 dB. However, under independent multitone jamming, the difference jumps to 6.2 dB which clearly substantiates the claim made earlier regarding the sensitivity of receiver 1 to high-powered jammers. Upon comparing receivers 2 and 3, it is evident that the normalization receiver (receiver 3) is only 0.2 dB inferior to the optimum S/N weighting receiver (receiver 2). Therefore, the signal normalization technique is a valid weighting function for diversity combining receivers in a jamming channel.

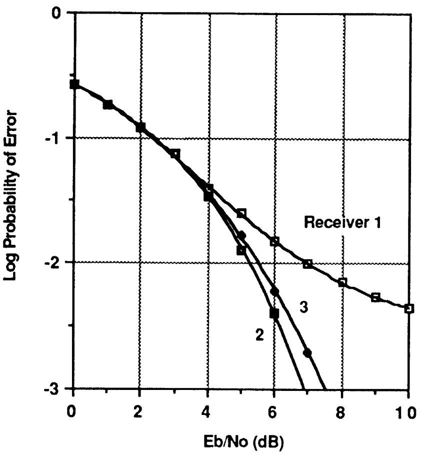


Fig. 12. Fourth-order diversity receiver comparison for 8-ary FSK in optimized partial band noise jamming.

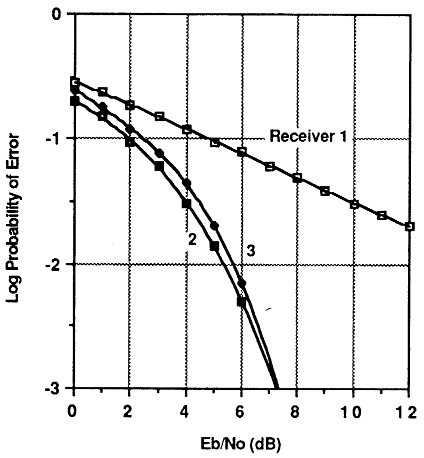


Fig. 13. Fourth-order diversity receiver comparison for 8-ary FSK in optimized independent multitone jamming.

The optimum diversity combining receiver in an independent, slowly Rayleigh fading channel is none other than receiver 1 [11]. And although this linear diversity combining receiver is practical to implement, it should not be used for receiving a signal through a combined jamming and fading channel since its performance against specialized jammers has just been shown to be severely degraded. Fortunately, as Fig. 14 illustrates, the suboptimum weighting function presented earlier (receiver 3) performs only slightly worse (about 1 dB) than the optimum linear diversity combiner in a fading channel. Therefore, in light of its excellent jamming performance and acceptable fading channel performance, the signal normalization receiver is an appropriate diversity combining receiver, and as such is appropriate for the next discussion on soft decision generators for MFSK channels.

MFSK Soft Decision Generation

The rate 1/2, $K = 7$ convolutional code with generators 171,133 [9] is a very often used error correction coding scheme. Its use with 8-ary FSK is somewhat structurally unusual when soft decisions are required because a mapping is necessary to generate soft decisions representing bit statistics from the 8-ary symbol statistics at the output of the FSK detector. Certainly the decoder could be fashioned theoretically into an optimum processor that uses symbol statistics, but the complexity of generating symbol branch statistics and performing branch decisions makes its use less practical [8]. A mapping such as this can further be avoided if a code more appropriately matched to the 8-ary modulation format were used. The class of Trumpis codes [14] is one example. However, the Trumpis code [14], which is optimized for 8-ary modulation, is one selected from a class of rate 1/3 codes, and as such is not guaranteed to be superior to a code selected from a larger class of codes whose use might require a symbol-to-bit mapping. In fact, while the Trumpis code [14] has a greater free distance on a symbol basis than the rate 1/2 code considered here (7 versus 5), this additional coding gain is considerably offset by the increased noncoherent combining loss in a multiple diversity system and by the fact that the rate 1/2 code actually has 1.5 bits per symbol versus the 1 bit per symbol of the rate 1/3 code. As such, the mapping is potentially worth what amounts to approximately 0.5 dB improvement in performance of the rate 1/2 code over the Trumpis code [14].

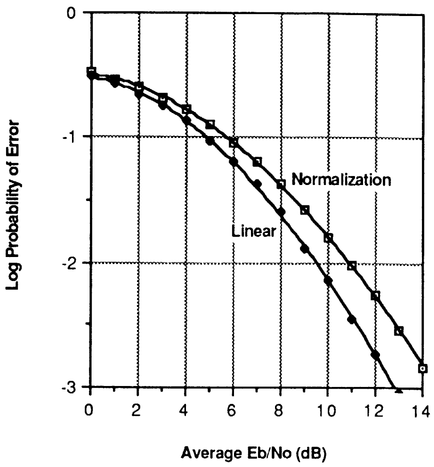


Fig. 14. Linear versus normalization combiner comparison for 8-ary FSK with fourth-order diversity in a Rayleigh fading channel.

For the 8-ary format, there exist three bits for each 8-ary symbol. Generation of soft decisions for each bit requires information pertaining to the possibility of the bit being a "1" or a "0." The example mapping to be presented here was first presented in [8] and is described mathematically below:

Bit 1

$$M_{01} = \max\,(R_0, R_1, R_2, R_3)$$

$$M_{11} = \max\,(R_4, R_5, R_6, R_7)$$

Bit 2

$$M_{02} = \max\,(R_0, R_1, R_4, R_5)$$

$$M_{12} = \max\,(R_2, R_3, R_6, R_7)$$

Bit 3

$$M_{03} = \max\ (R_0, R_2, R_4, R_6)$$

$$M_{13} = \max\ (R_1, R_3, R_5, R_7).$$

Here, M_{0j} is related to the probability that the jth bit of the symbol is equal to a 0, and M_{1j} is related to the probability that the jth bit of the symbol is equal to a 1. Furthermore, the parameters $\{R_j\colon j = 0, 1, \cdots, 7)$ correspond to the eight outputs of the diversity accumulators after all diversity transmissions are accumulated. This mapping assumes the following relationships between the three encoded bits and the matched filters which respectively generate $R_0, \cdots, R_7$:

Matched Filter	Encoded Bits
R_0	000
R_1	001
R_2	010
R_3	011
R_4	100
R_5	101
R_6	110
R_7	111

The algorithm generating M_{01} is motivated by the fact that if the first bit were a 0, it could have been sent only by a tone matched to one of the filters generating R_0, R_1, R_2, R_3. In general, the algorithm generating M_{ij} is motivated by the fact that if the jth bit (out of 3) were an "i" (0 or 1), it could have been sent only by a tone matched to one of the filters generating R_k, where R_k can be determined by selecting it if there is an "i" in the jth column of the above matched-filter/encoded-bit relationship.

Once the bit statistics are generated, the respective M_{0j} and M_{1j} values ($j = 1, 2, 3$) can either be subtracted or used to form a ratio, from which the actual soft decisions could be determined. This is exactly the same tradeoff performed for DPSK, however, the result turns out to be different. In the ratio case (M_{0j}/M_{1j}), problems begin to appear as the diversity order increases. Fig. 15 presents two sets of performance curves in an additive white Gaussian noise channel. The first set, curves A and B, are for a second-order diversity system with the difference generator performance reflected through curve A and that of the ratio generator through curve B. The second set of curves, C and D, are for an eighth-order diversity system with the difference and ratio soft decision generators indicated by curves C and D, respectively. Note that as the diversity order increases, so does the relative difference between the two soft decision generators. These problems are manifested in the ever-increasing compression of the ratio statistic near the value of unity as the diversity order increases. Such a compression requires ever-increasing resolution of the quantizer, and ultimately results in a soft decision BER performance that is limited to that obtained with hard decisions. On the other hand, this compression does not occur if the difference, $M_{0j} - M_{1j}$, is used instead of the ratio. And furthermore, the difference technique used here does not suffer the degradation caused by signal level variation as in the DPSK case because the signal normalization has removed all signal fluctuations.

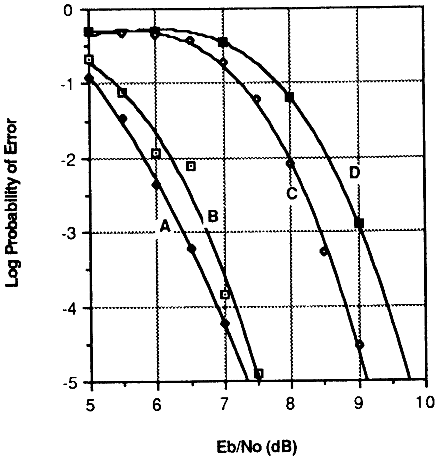


Fig. 15. Difference versus ratio soft decisions in AWGN for 8-ary FSK with second- and eighth-order diversity. A: Second-order diversity, difference. B: Second-order diversity, ratio. C: Eighth-order diversity, difference. D: Eighth-order diversity, ratio.

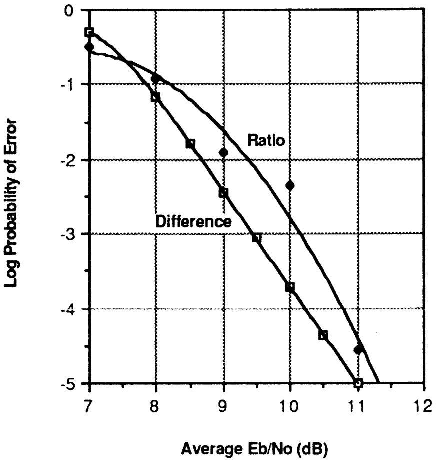


Fig. 16.

In a jamming and fading channel environment, there does not exist much difference in the comparative performance between the difference and ratio soft decision generators from that which was just presented in an additive white Gaussian noise channel. In fact, a partial-band noise jamming signal is mitigated from any partial-band advantage once the diversity order reaches about four. And, at that point, the comparative performance is identical to that for the benign, additive white Gaussian noise channel since the jammer is forced to be full band. In contrast to the DPSK soft decision generation, where the ratio technique decoupled the jamming strategy from the threshold settings, the decoupling in the FSK case is achieved by the signal normalization process in the diversity combining receiver. Fig. 16 illustrates the comparison between the difference and ratio soft decision generators in a Rayleigh fading environment. Again, the difference technique shows a superior performance over that of the ratio generator, but not much more than that experienced in an additive white Gaussian noise channel. In both the jamming and fading channel cases, it is suspected that the compressive nature of the ratio statistic is the reason behind this performance differential as well.

Summary

Detection and soft decision generation routines have been described for frequency hopped, DPSK and MFSK channels experiencing optimized jamming strategies and scintillating signals. Results have been presented to substantiate the utility of these routines, while primary discussions have been offered to establish the symbiotic relationship that exists among the routines, and the critical role each routine plays in mitigating the jamming and fading effects. In the case of DPSK, it was shown that a soft decision generator formed by taking the ratio of the two detector outputs was preferred over a generator taking the difference of the two outputs. The preference was based on the ability of the ratio to neutralize the jamming strategy more effectively; however, the ratio approach required a more complex detection implementation. For 8-ary FSK, it was found that the combination of an envelope detector, a signal normalizer and diversity combiner, and a symbol-to-bit statistical mapping with difference soft decisions formed a powerful architecture to combat jamming. Some results were shown that indicated the same architecture, but with the envelope detector replaced by a square-law detector provided improved performance in a fading channel although the differences were slight. It was finally shown that the channel and the existence of the signal normalization/combiner forced the acceptance of the detector type, while the signal normalization/combiner itself allowed the difference soft decision to be once again acceptable since the ratio approach became too compressive in its statistics as the diversity order increased.

References

[1] S. W. Houston, "Modulation techniques for communication, Part I: Tone and noise jamming performance of spread spectrum *M*-ary FSK and 2, 4-ary DPSK waveforms," in *NAECON 1975 Rec.*, 1975, pp. 51-58.

[2] J. G. Proakis, *Digital Communications.* New York: McGraw-Hill, 1983.

[3] A. J. Viterbi and I. M. Jacobs, "Advances in coding and modulation for noncoherent channels affected by fading, partial band, and multiple-access interference," in *Advances in Communication Systems, Vol. 4.* New York: Academic, 1975.

[4] D. M. Boroson, R. R. Rhodes, and R. C. Williamson, "A surface acoustic wave demodulator for MFSK signals," in *Nat. Telecommun. Conf. 1979 Rec.*, pp. 33.6.1-33.6.5.

[5] R. A. Yost, "On nonuniform windowing *M*-ary FSK data in a DFT-based detector," *IEEE Trans. Commun.*, vol. COM-28, pp. 2014-2019, Dec. 1980.

[6] C. J. Waylan, "Detection of fast, noncoherent, frequency-hopped FSK," *IEEE Trans. Commun.*, vol. COM-23, pp. 543-546, May 1975.

[7] J. S. Lee, L. E. Miller, and Y. K. Kim, "Error performance analyses of linear and nonlinear combining square-law receivers of L-hops per bit FH/BFSK waveforms in worst-case partial-band jamming," in *1983 MILCOM Rec.*, pp. 22-28.

[8] R. W. Boyd, "Diversity transmission of *M*-ary orthogonal signals in a hostile environment," in *1983 MILCOM Rec.*, pp. 12-16.

[9] G. C. Clark and J. B. Cain, *Error-Correction Coding for Digital Communications.* New York: Plenum, 1981.

[10] D. Torrieri, "The performance of five different metrics against pulse jamming," *IEEE Trans. Commun.*, vol. COM-34, pp. 200-204, Feb. 1986.

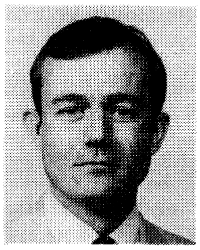
[11] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering.* New York: Wiley, 1965.

[12] A. J. Viterbi, "A robust ratio-threshold technique to mitigate tone and partial band jamming in coded MFSK systems," in *1982 MILCOM Rec.*, pp. 22.4-1-22.4-5.

[13] L. E. Miller, J. S. Lee, and A. P. Kadrichu, "Probability of error analyses of a BFSK frequency-hopping system with diversity under partial-band jamming interference—Part III: Performance of a square-law self-normalizing soft decision receiver," *IEEE Trans. Commun.*, vol. COM-34, pp. 669-675, July 1986.

[14] B. D. Trumpis, "Convolutional coding for *M*-ary channels," Ph.D. dissertation, UCLA, 1975.

[15] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications, Vol. II.* Rockville, MD: Computer Science Press, 1985.



Richard A. Yost (S'73-M'78) received the B.S.E.E. degree in 1974 from Purdue University, West Lafayette, IN, and the M.S.E.E. and Ph.D. degrees in 1975 and 1978, respectively, from the Georgia Institute of Technology, Atlanta.

Since 1978 he has been employed by the Government Communications Systems Division (GCSD), Harris Corporation, Melbourne, FL, in the Signal Processing and Software Department. His areas of principal interest are spread spectrum communications and DSP-based modular processors. He is currently the Manager of GCSD's IR&D program.

Probability Distribution of DPSK in Tone Interference and Applications to SFH/DPSK

QIANG WANG, MEMBER, IEEE, T. AARON GULLIVER, AND VIJAY K. BHARGAVA, SENIOR MEMBER, IEEE

Abstract—**Much recent attention has been focused on slow frequency hopped differential PSK (SFH/DPSK). The main reason for this interest in SFH/DPSK is that it can sustain a much higher data rate than a fast frequency hopped system while having the same hop rate.**

In this paper we present a study of the probability distribution of the received DPSK signal under tone jamming. These results will facilitate the analysis of the SFH/DPSK system.

The results given in this paper are more general than those previously published, in several aspects. First, the differential phase of the transmitted DPSK signal can assume any value. Second, probability distributions are derived instead of a set of probabilities calculated over certain symmetrical regions. This allows analyzing performance for arbitrarily selected decision regions as well as determining optimum decision regions for demodulating the DPSK signal. Third, the joint probability distribution of both the magnitude and differential phase of the jammed DPSK signal is given. This can be used in the analysis where both tone jamming and Gaussian noise are considered.

Applying these results, we analyze the error probability performance of a general uncoded SFH/DPSK signal under worst case tone jamming and Gaussian noise.

I. Introduction

In much previous work, the performance of fast frequency hopped (FFH) systems has been considered (see, e.g., [1, ch. 2] and [2]-[4] and their references). In an FFH system, the information bit rate R_b is low enough to require that an M-ary symbol be transmitted over one or more hops. If R_b exceeds the hop rate R_h, FFH is impossible. For example, it may be required to transmit 1.5 Mbit/s information at $R_h = 20$ khop/s. In this case, slow frequency hopping (SFH) must be used.

In slow frequency hopping, several symbols are transmitted during each hop. This paper is concerned with SFH/DPSK where the transmitted symbols are modulated in the form of differential PSK. Differential PSK is used because the hop period in SFH is usually not long enough to allow the receiver to recover the carrier phase, and to maintain the phase coherence between different hops at the transmitter. Thus, genuine coherent detection is usually impossible. However, since there are many symbols transmitted over one hop, differential coherent detection is possible. Because differential coherent detection outperforms noncoherent detection such as that used in FFH/MFSK, it is a logical choice for an SFH system.

In this paper, both binary and nonbinary (4-ary, 8-ary, etc.) DPSK, i.e., general M-ary DPSK, are considered. Anticipated interference may be both Gaussian noise and tone jamming. Unlike FFH/MFSK, little has been published on SFH/DPSK in the literature. There are many basic questions yet to be answered. In this paper we focus on the effects of jamming. Coded systems are not considered. The intent is to study the effects of jamming against SFH/DPSK and to provide some tools for the analysis of such a system.

At the receiving end of an uncoded SFH/DPSK system, the differential phase between two consecutive received symbols is detected, and this is used to decide which information symbol was transmitted. Houston [5] and Simon [6] (which is also a part of [1, ch. 4]) have analyzed the performance of SFH/DPSK under multiple continuous tone jamming for a specific set of signal phases and equally spaced decision regions. Recently, Gong analyzed the performance of a specific binary SFH/DPSK scheme in both tone and noise interference [7].

If the jamming tone over a jammed hop is continuous, i.e., the amplitude and initial phase are constant over a particular hop, then the received symbols over that hop are subject to an interference which is highly correlated from symbol to symbol. Recently, Winters has suggested that in correlated *noise*, the performance of DPSK depends on the set of signal phases and decision regions [8]. In order to minimize the demodulator output symbol error rate, we must consider the dependence of the performance of SFH/DPSK, under highly correlated tone jamming, on the signal phases and decision regions.

While the probability distribution of a received differential phase in Gaussian noise has been widely studied and well documented [9], [10], no general results have been published on the probability distribution of DPSK in tone interference. Therefore, we will derive in the next section the general probability distribution of a received DPSK signal corrupted by continuous tone jamming. By "continuous tone jamming" we mean that a jamming tone interferes with two consecutively transmitted DPSK symbols (with the same amplitude, frequency, and initial

Manuscript received February 17, 1989; revised November 21, 1989. This work was supported by the Department of Communications of Canada under DSS Contract 27ST.36001-6-3539. This work was presented at the 1990 IEEE International Symposium on Information Theory, San Diego, CA, January 14-19, 1990.

Q. Wang and V. K. Bhargava are with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, B.C., Canada V8W 2Y2.

T. A. Gulliver was with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, B.C., Canada V8W 2Y2. He is now with the Defence Research Establishment, Ottawa, Ont., Canada K1A 0Z4.

IEEE Log Number 9035815.

0733-8716/90/0600-0895$01.00 © 1990 IEEE

phase). When DPSK symbols are jammed by a single tone, the jamming tone is assumed to have the same frequency as the DPSK carrier frequency. In Section III, we apply the results obtained in Section II to the evaluation of SFH/DPSK systems.

II. Probability Distributions of the Received DPSK Under Tone Jamming

In complex form, the transmitted DPSK signal in the ith signaling interval is represented by

$$S^{(i)} = Ee^{j(2\theta + \theta_T^{(i-1)})},$$

where $\theta_T^{(i-1)}$ is the total accumulated phase in the $(i-1)$th signaling interval and 2θ is the differential phase transmitted in the ith signaling interval with $0 \leq \theta < \pi$. The jamming tone is represented by

$$J = Ie^{j\theta_J'},$$

where θ_J' is a random phase uniformly distributed in an interval of length 2π. Let β denote the ratio of the amplitude of the jamming tone to that of the signal tone,

$$\beta = \frac{I}{E}.$$

The received signals (on which a decision on the transmitted differential phase in the ith signaling interval is to be based) are represented by $Y^{(i-1)}$ and $Y^{(i)}$. The received differential phase is then

$$\Psi = \arg\left(Y^{(i)}Y^{(i-1)*}\right),$$

where the phase angle function arg has a main value in the range $(-\pi, \pi]$ and the asterisk denotes complex conjugation.

Simon [6] derived the probability of $\Psi - 2\theta$ within equally spaced decision regions for a *specific* set of θ. The final results are very complicated. This might seem to suggest that for *any* θ, the derivation of the probability of Ψ or $\Psi - 2\theta$ (or equivalently the probability distribution) over *any* region would be prohibitively complex. However, we have found that unlike Ψ or $\Psi - 2\theta$, $\Gamma = \Psi - \theta$ has some symmetry that can be utilized to simplify the derivation significantly, as is shown below.

Under continuous tone jamming, we have

$$Y^{(i-1)} = Ee^{j\theta_T^{(i-1)}} + Ie^{j\theta_J'},$$

and

$$Y^{(i)} = Ee^{j(2\theta + \theta_T^{(i-1)})} + Ie^{j\theta_J'}.$$

We define

$$R_1 = \left|Y^{(i-1)}\right|,$$

and

$$R_2 = \left|Y^{(i)}\right|.$$

To analyze the bit error rate (BER) performance under strong tone jamming and negligibly low system thermal noise, only the probability distribution of Γ is required. Otherwise, we must consider the joint probability distribution of Γ, R_1, and R_2, as will be seen later. To clarify the derivation procedure, we first derive the probability distributions of Γ, R_1, and R_2 separately, and then consider the joint probability distribution.

A. Probability Distribution of Differential Phase Under Continuous Tone Jamming

We first consider the probability distribution of Γ, and those of Ψ and $\Psi - 2\theta$. For Γ, we have

$$\begin{aligned}\Gamma &= \arg\left[Y^{(i)}\left(Y^{(i-1)}\right)^* e^{-j\theta}\right] \\ &= \arg\left[\left(e^{j(2\theta + \theta_T^{(i-1)})} + \beta e^{j\theta_J'}\right)\right. \\ &\qquad \left.\cdot \left(e^{-j\theta_T^{(i-1)}} + \beta e^{-j\theta_J'}\right)e^{-j\theta}\right] \\ &= \arg\left[\left(e^{j2\theta} + \beta e^{j(\theta_T^{(i-1)} + 2\theta - \theta_J')}\right.\right. \\ &\qquad \left.\left. + \beta e^{j(\theta_J' - \theta_T^{(i-1)})} + \beta^2\right)e^{-j\theta}\right] \\ &= \arg\left[e^{j\theta} + \beta^2 e^{-j\theta}\right. \\ &\qquad \left. + \beta\left(e^{j(\theta_T^{(i-1)} + \theta - \theta_J')} + e^{-j(\theta_T^{(i-1)} + \theta - \theta_J')}\right)\right] \\ &= \arg\left[e^{j\theta} + \beta^2 e^{-j\theta} + 2\beta\cos\theta_J\right] \end{aligned} \tag{1}$$

where

$$\theta_J = \theta_J' - \theta_T^{(i-1)} - \theta.$$

Since θ_J' can be assumed to be uniformly distributed over $(\theta_T^{(i-1)} + \theta - \pi, \theta_T^{(i-1)} + \theta + \pi]$, θ_J is uniformly distributed over $(-\pi, \pi]$. Suppose $\beta > 0$ and denote

$$U = \frac{(1 + \beta^2)\cos\theta}{2\beta} + j\frac{(1 - \beta^2)\sin\theta}{2\beta},$$

and

$$V = U + \cos\theta_J.$$

Then we have

$$\Gamma = \arg(V).$$

It is clear that θ_J does not change the value of the imaginary part of V. Consequently, $\operatorname{Im}(V)$ is equal to $\operatorname{Im}(U)$. Let

$$\Phi = \arg(U).$$

Obviously, if $\theta = 0$, Γ is always equal to 0 and the probability density function (pdf) of Γ is

$$p_\Gamma(\gamma) = \delta(\gamma),$$

where $\delta(x)$ is the Dirac delta function.

If $\beta = 1$, $U = \cos\theta$ and $V = \cos\theta + \cos\theta_J$. Consider $\cos\theta + \cos\theta_J > 0$, i.e., $\cos\theta_J > -\cos\theta$. Then we have $|\theta_J| < \arccos(\cos(\pi - \theta)) = \pi - \theta$. Thus, Γ equals 0 with probability $1 - \theta/\pi$ and π with probability θ/π. For $0 < \theta < \pi$,

$$p_\Gamma(\gamma) = \left(1 - \frac{\theta}{\pi}\right)\delta(\gamma) + \frac{\theta}{\pi}\delta(\gamma - \pi).$$

Now we assume $\theta \neq 0$ and $\beta \neq 1$, i.e., $\sin(\Phi) \neq 0$. Suppose $\beta < 1$ (i.e., $\mathrm{Im}(\boldsymbol{V}) > 0$). Then $0 < \Phi < \pi$ and

$$\mathrm{Prob}\left\{-\pi < \Gamma < 0\right\} = 0.$$

Now we calculate

$$\mathrm{Pr}(\gamma) = \mathrm{Prob}\left\{0 < \Gamma \leq \gamma\right\},$$

where $0 < \gamma < \pi$. As shown in Fig. 1, the intermediate variable d is defined as

$$d = |\boldsymbol{U}|\left(\sin\left(|\Phi|\right)\cot|\gamma| - \cos\Phi\right). \tag{2}$$

Noting that $|\cos\theta_j| \leq 1$, we have a symmetric region of θ_j centered at 0 in which $0 < \Gamma \leq \gamma$. Specifically, for $-1 < d < 1$, the corresponding θ_j is in $[-\theta_\Gamma, \theta_\Gamma]$, where

$$\theta_\Gamma(\gamma) = \arccos(d), \qquad |d| \leq 1.$$

For $d \leq -1$, the corresponding θ_j can be anywhere in $(-\pi, \pi]$. For $d \geq 1$, there is no such θ_j that may result in $0 < \Gamma < \gamma$. Then we have

$$\mathrm{Pr}(\gamma) = \begin{cases} 1, & d \leq -1; \\ \dfrac{\theta_\Gamma}{\pi}, & -1 < d < 1; \\ 0, & d \geq 1. \end{cases} \tag{3}$$

The cumulative distribution function (cdf) of Γ is then

$$P_\Gamma(\gamma) = \begin{cases} 0, & \gamma \leq 0; \\ \mathrm{Pr}(\gamma), & 0 < \gamma < \pi; \\ 1, & \gamma \geq \pi. \end{cases} \tag{4}$$

Note that, from (2), for $\gamma \geq 0$, we have

$$\cot\gamma = \frac{d}{|\boldsymbol{U}|\sin|\Phi|} + \cot|\Phi|,$$

and

$$\frac{\partial d}{\partial\gamma} = -|\boldsymbol{U}|\sin\left(|\Phi|\right)\csc^2\gamma.$$

Then the pdf of Γ is

$$p_\Gamma(\gamma) = \begin{cases} \dfrac{|\boldsymbol{U}|\sin\left(|\Phi|\right)\csc^2\gamma}{\pi\sqrt{1-d^2}}, & \operatorname{arc\,cot}\left(\dfrac{1}{|\boldsymbol{U}|\sin|\Phi|} + \cot|\Phi|\right) \\ & \quad < \gamma < \operatorname{arc\,cot}\left(\dfrac{-1}{|\boldsymbol{U}|\sin|\Phi|} + \cot|\Phi|\right); \\ 0, & \text{elsewhere.} \end{cases} \tag{5}$$

Suppose $\beta > 1$ (i.e., $\mathrm{Im}(\boldsymbol{V}) < 0$). Then $-\pi < \Phi < 0$ and

$$\mathrm{Prob}\left\{0 < \Gamma \leq \pi\right\} = 0.$$

By symmetry, and noting the term $|\Phi|$ in (2), we have

$$\mathrm{Prob}\left\{-\gamma \leq \Gamma < 0\right\} = \mathrm{Pr}(\gamma), \tag{6}$$

where $0 < \gamma < \pi$ (note Fig. 1). Then the cdf of Γ is

$$P_\Gamma(\gamma) = \begin{cases} 0, & \gamma \leq -\pi; \\ 1 - \mathrm{Pr}(-\gamma), & -\pi < \gamma < 0; \\ 1, & \gamma \geq 0, \end{cases} \tag{7}$$

and the pdf of Γ for $0 < \theta < \pi$ is

$$p_\Gamma(\gamma) = \begin{cases} \dfrac{|\boldsymbol{U}|\sin\left(|\Phi|\right)\csc^2\gamma}{\pi\sqrt{1-d^2}}, \\ \quad -\operatorname{arc\,cot}\left(\dfrac{-1}{|\boldsymbol{U}|\sin|\Phi|} + \cot|\Phi|\right) \\ \quad < \gamma < -\operatorname{arc\,cot}\left(\dfrac{1}{|\boldsymbol{U}|\sin|\Phi|} + \cot|\Phi|\right); \\ 0, \quad \text{elsewhere.} \end{cases} \tag{8}$$

Note that using the absolute value of γ in (2) is only for conciseness in (8), where we actually have

$$d = |\boldsymbol{U}|\left(\sin\left(|\Phi|\right)\cot(-\gamma) - \cos\Phi\right).$$

Using the cdf's or pdf's given above, we can calculate the arbitrary probability

$$\mathrm{Pr}_\Gamma(\gamma_1, \gamma_2) = \mathrm{Prob}\left\{\gamma_1 < \Gamma \leq \gamma_2\right\},$$

where $\gamma_1 < \gamma_2$, and both are main-valued bounds (both are in the main value interval of the argument). To use Pr_Γ to calculate the probability distribution of Ψ or $\Psi - 2\theta$, all we need to do is shift the specified region and convert it into one or two pairs of main-valued bounds for use in Pr_Γ. For example, for main-valued bounds b_1 and b_2, we want to calculate

$$\begin{aligned} P_1 &= \mathrm{Prob}\left\{b_1 < \Psi - 2\theta \leq b_2\right\} \\ &= \mathrm{Prob}\left\{b_1 + \theta < \Psi - \theta \leq b_2 + \theta\right\} \\ &= \mathrm{Prob}\left\{b_1 + \theta < \Gamma \leq b_2 + \theta\right\}. \end{aligned} \tag{9}$$

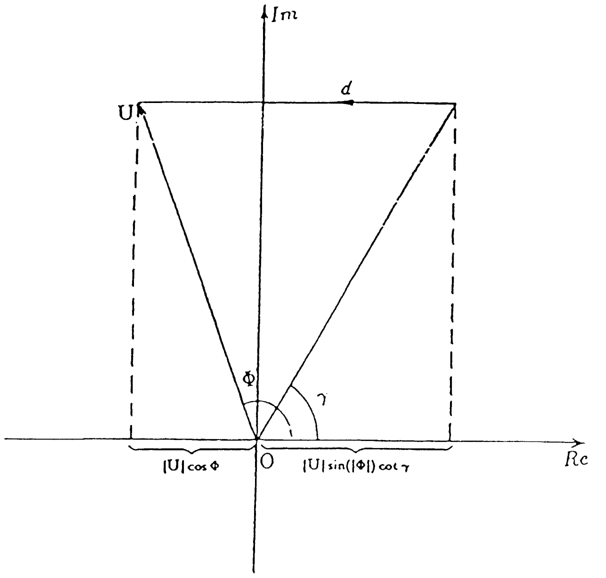


Fig. 1. Illustration of the relation between the intermediate variable d and U and γ. Note that $|U| \sin(|\Phi|) \cot|\gamma|$ and $|U| \cos \Phi$ are real coordinates that can assume both positive and negative values.

By adding multiples of 2π to $b_1 + \theta$ and $b_2 + \theta$, respectively, we can obtain a pair of bounds bm_1 and bm_2 (corresponding to $b_1 + \theta$ and $b_2 + \theta$, respectively) such that both bm_1 and bm_2 are in $(-\pi, \pi]$. If $bm_1 < bm_2$, they are main-valued bounds and

$$P_1 = \Pr_\Gamma(bm_1, bm_2).$$

If $bm_1 > bm_2$,

$$P_1 = \Pr_\Gamma(bm_1, \pi) + \Pr_\Gamma(-\pi, bm_2).$$

In terms of pdf's, we can obtain the pdf's of Ψ and $\Psi - 2\theta$ by periodically extending p_Γ and then shifting by $\pm\theta$. The periodic extension of $p_\Gamma(\gamma)$ (with period 2π) is

$$\tilde{p}_\Gamma(\gamma) = \sum_{l=-\infty}^{+\infty} p_\Gamma(\gamma - l2\pi).$$

Then the pdf of Ψ is

$$p_\Psi(\psi) = \begin{cases} \tilde{p}_\Gamma(\psi - \theta), & -\pi < \psi \le \pi; \\ 0, & \text{elsewhere,} \end{cases} \tag{10}$$

and the pdf of $\Psi - 2\theta$ is

$$p_{\Psi-2\theta}(\psi_2) = \begin{cases} \tilde{p}_\Gamma(\psi_2 + \theta), & -\pi < \psi_2 \le \pi; \\ 0, & \text{elsewhere.} \end{cases} \tag{11}$$

For later use in deriving the joint probability distribution, we define

$$H_\Gamma(\theta_J) = \begin{cases} 1, & \text{if } \Gamma \le \gamma; \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

For $\theta = 0$,

$$H_\Gamma(\theta_J) = \begin{cases} 1, & \text{if } \gamma \ge 0; \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

For $\beta = 1$ and $0 < \theta < \pi$,

$$H_\Gamma(\theta_J) = \begin{cases} 1, & \text{if } \gamma \ge \pi; \\ \Pi\left(\dfrac{\theta_J}{\pi - \theta}\right), & \text{if } 0 \le \gamma < \pi; \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

Here $\Pi(x)$ is the rectangular function which is equal to 1 if $|x| \le 1$ and 0 otherwise. In the above equation, it is implied that $\arg(\boldsymbol{X}) = 0$ if $|\boldsymbol{X}| = 0$. In this case, we have

$$\frac{\partial H_\Gamma}{\partial \gamma} = \begin{cases} \delta(\gamma), & \text{if } |\theta_J| \le \pi - \theta; \\ \delta(\gamma - \pi), & \text{otherwise.} \end{cases} \tag{15}$$

For $\beta \ne 1$ and $0 < \theta < \pi$, from (3)-(8), it is clear that for $|d| < 1$ (or $\gamma_1 < \gamma < \gamma_2$ where $\gamma_1 = \text{arc cot}((1/|\boldsymbol{U}| \sin|\Phi|) + \cot|\Phi|)$, $\gamma_2 = \text{arc cot}((-1/|\boldsymbol{U}| \sin|\Phi|) + \cot|\Phi|)$ for $\beta < 1$, and $\gamma_1 = -\text{arc cot}((-1/|\boldsymbol{U}| \sin|\Phi|) + \cot|\Phi|)$, $\gamma_2 = -\text{arc cot}((1/|\boldsymbol{U}| \sin|\Phi|) + \cot|\Phi|)$ for $\beta > 1$), we have

$$H_\Gamma(\theta_J) = \begin{cases} \Pi\left(\dfrac{\theta_J}{\theta_\Gamma}\right), & \text{if } \beta < 1; \\ \Pi^-\left(\dfrac{\theta_J}{\theta_\Gamma}\right), & \text{if } \beta > 1. \end{cases} \tag{16}$$

Here $\Pi^-(x)$ is a function which is equal to 1 if $|x| \ge 1$ and 0 otherwise. Then we have

$$\frac{\partial H_\Gamma}{\partial \gamma} = (\delta(\theta_j + \theta_\Gamma) + \delta(\theta_j - \theta_\Gamma))\frac{\partial \theta_\Gamma}{\partial \gamma} c(\beta), \tag{17}$$

where

$$c(\beta) = \begin{cases} 1, & \text{if } \beta < 1; \\ -1, & \text{if } \beta > 1. \end{cases}$$

We can also write the inverse function of θ_Γ

$$\theta_{\Gamma^{-1}}(\theta_\gamma) = \begin{cases} \text{arc cot}\left(\dfrac{\cos\theta_\gamma}{|\boldsymbol{U}| \sin|\Phi|} + \cot|\Phi|\right), \\ \text{if } \beta < 1; \\ -\text{arc cot}\left(\dfrac{\cos\theta_\gamma}{|\boldsymbol{U}| \sin|\Phi|} + \cot|\Phi|\right), \\ \text{if } \beta > 1. \end{cases} \tag{18}$$

B. Probability Distribution of the Amplitude Under Continuous Tone Jamming

In this section we consider the probability distributions of R_1 and R_2 as functions of θ_J. They will be used in deriving the joint probability distribution.

Note that R_1 and R_2 are nonnegative. For R_1 we have

$$\begin{aligned} R_1^2 &= \left| \boldsymbol{Y}^{(i-1)} e^{-j(\theta + \theta_T^{(i-1)})} \right|^2 \\ &= \left| E\left(e^{j\theta_T^{(i-1)}} + \beta e^{j\theta_J'}\right) e^{-j(\theta + \theta_T^{(i-1)})} \right|^2 \\ &= \left| E\left(e^{-j\theta} + \beta e^{j(\theta_J' - \theta_T^{(i-1)} - \theta)}\right) \right|^2 \\ &= \left| E\left(e^{-j\theta} + \beta e^{j\theta_J}\right) \right|^2 \\ &= E^2\left[(\cos\theta + \beta\cos\theta_J)^2 + (-\sin\theta + \beta\sin\theta_J)^2\right] \\ &= E^2\left[1 + \beta^2 + 2\beta\cos(\theta + \theta_J)\right]. \end{aligned} \tag{19}$$

If $r_1 \geq (1 + \beta)E$, Prob $\{R_1 \leq r_1\} = 1$. If $r_1 \leq |1 - \beta| E$, Prob $\{R_1 \leq r_1\} = 0$. If $|1 - \beta| E < r_1 < |1 + \beta| E$, then we have

$$\text{Prob}\,\{R_1 \leq r_1\} = \text{Prob}\,\{\cos(\theta_J + \theta) \leq \cos\theta_{R_1}\},$$

where

$$\theta_{R_1} = \arccos\left[\left(\frac{r_1^2}{E^2} - 1 - \beta^2\right)\frac{1}{2\beta}\right].$$

This implies a symmetrical region of θ_J centered at $\pi - \theta$ with a width of $\pi - \theta_{R_1}$ on either side in which $R_1 \leq r_1$. Thus, the cdf of R_1 is

$$P_{R_1}(r_1) = \begin{cases} 1, & r_1 \geq (1 + \beta)E; \\ 1 - \dfrac{\theta_{R_1}}{\pi}, & |1 - \beta| E < r_1 < |1 + \beta| E; \\ 0, & r_1 \leq |1 - \beta| E. \end{cases} \tag{20}$$

The pdf of R_1 is

$$p_{R_1}(r_1) = \begin{cases} -\dfrac{1}{\pi}\dfrac{\partial\theta_{R_1}}{\partial r_1} = \dfrac{2r_1}{\pi E^2\sqrt{4\beta^2 - \left(\dfrac{r_1^2}{E^2} - 1 - \beta^2\right)^2}}, \\ \qquad |1 - \beta| E < r_1 < (1 + \beta)E; \\ 0, \quad \text{elsewhere.} \end{cases} \tag{21}$$

Similar to $H_\Gamma(\theta_J)$, we define

$$H_{R_1}(\theta_J) = \begin{cases} 1, & \text{if } R_1 \leq r_1; \\ 0, & \text{otherwise.} \end{cases} \tag{22}$$

For $|1 - \beta| E < r_1 < (1 + \beta)E$ and $-\pi < \theta_J \leq \pi$, we have

$$H_{R_1}(\theta_J) = \sum_{t=-\infty}^{+\infty} \Pi\left(\frac{\theta_J - (\pi - \theta) - l2\pi}{\pi - \theta_{R_1}}\right), \quad -\pi < \theta_J \leq \pi; \tag{23}$$

and

$$\begin{aligned} \frac{\partial H_{R_1}}{\partial r_1} = \sum_{l=-\infty}^{+\infty} &\left[\delta\left(\theta_J - (\pi - \theta) - l2\pi + \pi - \theta_{R_1}\right)\right. \\ &\left. + \delta\left(\theta_J - (\pi - \theta) - l2\pi - (\pi - \theta_{R_1})\right)\right] \\ &\times \left(-\frac{\partial\theta_{R_1}}{\partial r_1}\right), \qquad -\pi < \theta_J \leq \pi. \end{aligned} \tag{24}$$

We can also write the inverse function of θ_{R_1},

$$\theta_{R_1^{-1}}(\theta_{r_1}) = E\sqrt{1 + \beta^2 + 2\beta\cos(\theta_{r_1})}. \tag{25}$$

Similarly, for R_2 we have

$$\begin{aligned} R_2^2 &= \left| \boldsymbol{Y}^{(i)} e^{-j(\theta + \theta_T^{(i-1)})} \right|^2 \\ &= \left| E\left(e^{j(\theta_T^{(i-1)} + 2\theta)} + \beta e^{j\theta_J'}\right) e^{-j(\theta + \theta_T^{(i-1)})} \right|^2 \\ &= \left| E\left(e^{j\theta} + \beta e^{j(\theta_J' - \theta_T^{(i-1)} - \theta)}\right) \right|^2 \\ &= \left| E\left(e^{j\theta} + \beta e^{j\theta_J}\right) \right|^2 \\ &= E^2\left[(\cos\theta + \beta\cos\theta_J)^2 + (\sin\theta + \beta\sin\theta_J)^2\right] \\ &= E^2\left[1 + \beta^2 + 2\beta\cos(\theta - \theta_J)\right]. \end{aligned} \tag{26}$$

If $r_2 \geq (1 + \beta)E$, Prob $\{R_2 \leq r_2\} = 1$. If $r_2 \leq |1 - \beta| E$, Prob $\{R_2 \leq r_2\} = 0$. If $|1 - \beta| E < r_2 < |1 + \beta| E$, then we have

$$\text{Prob}\,\{R_2 \leq r_2\} = \text{Prob}\,\{\cos(\theta_J - \theta) \leq \cos\theta_{R_2}\},$$

where

$$\theta_{R_2} = \arccos\left[\left(\frac{r_2^2}{E^2} - 1 - \beta^2\right)\frac{1}{2\beta}\right].$$

This implies a symmetrical region of θ_J centered at $\theta + \pi$ with a width of $\pi - \theta_{R_2}$ on either side in which $R_2 \leq r_2$. It can be seen that the cdf and pdf of R_2 are the same as those of R_1. That is, the cdf of R_2 is

$$P_{R_2}(r_2) = P_{R_1}(r_2). \tag{27}$$

The pdf of R_2 is

$$\mathrm{p}_{R_2}(\mathrm{r}_2) = \mathrm{p}_{R_1}(\mathrm{r}_2). \tag{28}$$

As was done above, we define

$$H_{R_2}(\theta_J) = \begin{cases} 1, & \text{if } R_2 \leq r_2; \\ 0, & \text{otherwise.} \end{cases} \tag{29}$$

For $|1 - \beta| E < r_2 < (1 + \beta)E$ and $-\pi < \theta_J \leq \pi$, we have a slightly different result from H_{R_1},

$$H_{R_2}(\theta_J) = \sum_{t=-\infty}^{+\infty} \Pi\left(\frac{\theta_J - (\pi + \theta) - l2\pi}{\pi - \theta_{R_2}}\right), \quad -\pi < \theta_J \leq \pi, \tag{30}$$

and

$$\frac{\partial H_{R_2}}{\partial r_2} = \sum_{l=-\infty}^{+\infty} \left[\delta(\theta_J - (\pi + \theta) - l2\pi + \pi - \theta_{R_2}) + \delta(\theta_J - (\pi + \theta) - l2\pi - (\pi - \theta_{R_2}))\right] \times \left(-\frac{\partial \theta_{R_2}}{\partial r_2}\right), \quad -\pi < \theta_J \leq \pi. \tag{31}$$

We can also write the inverse function of θ_{R_2},

$$\theta_{R_2^{-1}}(\theta_{r_2}) = E\sqrt{1 + \beta^2 + 2\beta \cos(\theta_{r_2})}. \tag{32}$$

C. The Joint Distribution and the Expectation

Using H_Γ, H_{R_1}, and H_{R_2} defined previously, we have the joint cdf of Γ, R_1, and R_2,

$$P_{\Gamma,R_1,R_2}(\gamma, r_1, r_2) = \frac{\int_{-\pi}^{+\pi} H_\Gamma(\theta_J)\, H_{R_1}(\theta_J)\, H_{R_2}(\theta_J)\, d\theta_J}{2\pi}. \tag{33}$$

The joint pdf, $p_{\Gamma,R_1,R_2}(\gamma, r_1, r_2)$, has a somewhat unconventional form. Considering

$$p_{\Gamma,R_1,R_2}(\gamma, r_1, r_2)\, d\gamma\, dr_1\, dr_2 = \text{Prob}\,\{\gamma < \Gamma \leq \gamma + d\gamma,\ r_1 < R_1 \leq r_1 + dr_1,\ r_2 < R_2 \leq r_2 + dr_2\},$$

we can see that $p_{\Gamma,R_1,R_2}(\gamma, r_1, r_2)$ is nonzero only over a line (or several lines) in the three-dimensional space which consists of values of γ, r_1, and r_2. In fact, it can be shown that over these lines, the pdf assumes infinite values.

For the analysis of the BER performance, all we need to know is the expectation of $G(\Gamma, R_1, R_2)$ where G is assumed to be an arbitrary continuous function. This will be shown later. When $0 < \theta < \pi$ and $\beta \neq 1$, the expectation is given by

$$\begin{aligned}
\overline{G} &= \int\int\int_{-\infty}^{+\infty} G(\gamma, r_1, r_2)\, p_{\Gamma,R_1,R_2}(\gamma, r_1, r_2)\, d\gamma\, dr_1\, dr_2 \\
&= \int\int_{|1-\beta|E}^{(1+\beta)E} dr_1\, dr_2 \int_{\gamma_1}^{\gamma_2} G(\gamma, r_1, r_2)\, \frac{\partial^3 P_{\Gamma,R_1,R_2}}{\partial\gamma\, \partial r_1\, \partial r_2}\, d\gamma \\
&= \frac{1}{2\pi} \int\int_{|1-\beta|E}^{(1+\beta)E} dr_1\, dr_2 \int_{\gamma_1}^{\gamma_2} d\gamma\, G(\gamma, r_1, r_2) \\
&\quad \cdot \int_{-\pi}^{\pi} \frac{\partial H_\Gamma(\theta_J)}{\partial\gamma}\, \frac{\partial H_{R_1}(\theta_J)}{\partial r_1}\, \frac{\partial H_{R_2}(\theta_J)}{\partial r_2}\, d\theta_J \\
&= \frac{1}{2\pi} \int\int_{|1-\beta|E}^{(1+\beta)E} dr_1\, dr_2 \int_{-\pi}^{\pi} \frac{\partial H_{R_1}(\theta_J)}{\partial r_1}\, \frac{\partial H_{R_2}(\theta_J)}{\partial r_2}\, d\theta_J \\
&\quad \times \int_{\gamma_1}^{\gamma_2} d\gamma\, G(\gamma, r_1, r_2)\, (\delta(\theta_J + \theta_\Gamma) + \delta(\theta_J - \theta_\Gamma)) \\
&\quad \cdot \left(\frac{\partial \theta_\Gamma}{\partial \gamma}\right) c(\beta) \\
&= \frac{1}{2\pi} \int\int_{|1-\beta|E}^{(1+\beta)E} dr_1\, dr_2 \\
&\quad \cdot \int_{-\pi}^{\pi} \frac{\partial H_{R_1}(\theta_J)}{\partial r_1}\, \frac{\partial H_{R_2}(\theta_J)}{\partial r_2}\, d\theta_J \\
&\quad \times \int_0^{\pi} d\theta_\Gamma\, G(\theta_{\Gamma^{-1}}(\theta_\Gamma), r_1, r_2) \\
&\quad \cdot (\delta(\theta_J + \theta_\Gamma) + \delta(\theta_J - \theta_\Gamma)) \\
&= \frac{1}{2\pi} \int\int_{|1-\beta|E}^{(1+\beta)E} dr_1\, dr_2 \int_{-\pi}^{\pi} \frac{\partial H_{R_1}(\theta_J)}{\partial r_1}\, \frac{\partial H_{R_2}(\theta_J)}{\partial r_2} \\
&\quad \cdot G(\theta_{\Gamma^{-1}}(|\theta_J|), r_1, r_2)\, d\theta_J \\
&= \frac{1}{2\pi} \int_{|1-\beta|E}^{(1+\beta)E} dr_2 \int_{-\pi}^{\pi} d\theta_J \int_0^{\pi} d\theta_{R_1} \\
&\quad \times G(\theta_{\Gamma^{-1}}(|\theta_J|), \theta_{R_1^{-1}}(\theta_{R_1}), r_2)\, \frac{\partial H_{R_2}}{\partial r_2} \\
&\quad \cdot (\delta(\arg(e^{j(\theta_J + \theta)}) + \theta_{R_1}) \\
&\quad + \delta(\arg(e^{j(\theta_J + \theta)}) - \theta_{R_1})) \\
&= \frac{1}{2\pi} \int_{|1-\beta|E}^{(1+\beta)E} dr_2 \int_{-\pi}^{\pi} d\theta_J\, G(\theta_{\Gamma^{-1}}(|\theta_J|), \\
&\quad \theta_{R_1^{-1}}(|\arg(e^{j(\theta_J + \theta)})|), r_2)\, \frac{\partial H_{R_2}}{\partial r_2} \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} G(\theta_{\Gamma^{-1}}(|\theta_J|), \\
&\quad \theta_{R_1^{-1}}(|\arg(e^{j(\theta_J + \theta)})|), \theta_{R_2^{-1}}(|\arg(e^{j(\theta_J - \theta)}|))\, d\theta_J.
\end{aligned} \tag{34}$$

From (18), (25), and (32) we can slightly simplify (34) to

$$\overline{G} = \frac{1}{2\pi} \int_{-\pi}^{\pi} G(\theta_{\Gamma^{-1}}(\theta_J), \theta_{R_1^{-1}}(\theta_J + \theta), \theta_{R_2^{-1}}(\theta_J - \theta))\, d\theta_J. \tag{35}$$

If $\theta = 0$, it is clear that Γ is independent of R_1 and R_2. Thus,

$$p_{\Gamma,R_1,R_2}(\gamma, r_1, r_2) = \delta(\gamma)\, p_{R_1,R_2}(r_1, r_2)$$

where $p_{R_1,R_2}(r_1, r_2)$ is the joint pdf of R_1 and R_2. Note that the joint cdf of R_1 and R_2 is given by

$$P_{R_1,R_2}(r_1, r_2) = \frac{\int_{-\pi}^{+\pi} H_{R_1}(\theta_J)\, H_{R_2}(\theta_J)\, d\theta_J}{2\pi}.$$

Then, following a procedure similar to that of deriving (35), we have

$$\overline{G} = \frac{1}{2\pi} \int_{-\pi}^{\pi} G(0, \theta_{R_1^{-1}}(\theta_J), \theta_{R_2^{-1}}(\theta_J))\, d\theta_J. \tag{36}$$

When $\theta = 0$, $\Phi = 0$ and $|U| = (1 + \beta^2)/2\beta$, and $\lim_{x \to 0} \cot x - 1/\sin x = 0$ (for $\beta \neq 1$), so (18) gives

$$\theta_{\Gamma^{-1}}(\theta_\gamma) \equiv 0. \tag{37}$$

When $\beta = 1$, (37) is also valid except for $\theta_\gamma = (2k + 1)\pi$, where k is an integer. Since G is a continuous function, (35) also applies when $\theta = 0$. From the point of view of numerical computation, (36) provides a good approximation when $\theta \approx 0$.

For $\beta = 1$ and $0 < \beta < \pi$ using (15), we similarly have

$$\begin{aligned}
\overline{G} &= \frac{1}{2\pi} \int_0^\pi d\gamma \int_{-\pi}^{\pi} G(\gamma, \theta_{R_1^{-1}}(\theta_J + \theta), \\
&\quad \theta_{R_1^{-1}}(\theta_J - \theta)) \frac{\partial H_\Gamma(\theta_J)}{\partial_\gamma} d\theta_J \\
&= \frac{1}{2\pi} \Bigg[\int_{-(\pi-\theta)}^{\pi-\theta} d\theta_J \, G(0, \theta_{R_1^{-1}}(\theta_J + \theta), \\
&\quad \theta_{R_2^{-1}}(\theta_J - \theta)) \\
&\quad + \int_{-\pi}^{-(\pi-\theta)} d\theta_J \, G(\pi, \theta_{R_1^{-1}}(\theta_J + \theta), \\
&\quad \theta_{R_2^{-1}}(\theta_J - \theta)) \\
&\quad + \int_{\pi-\theta}^{\pi} d\theta_J \, G(\pi, \theta_{R_1^{-1}}(\theta_J + \theta)\, \theta_{R_2^{-1}}(\theta_J - \theta)) \Bigg] \\
&= \frac{1}{2\pi} \Bigg[\int_{-(\pi-\theta)}^{\pi-\theta} d\theta_J \, G(0, \theta_{R_1^{-1}}(\theta_J + \theta), \\
&\quad \theta_{R_2^{-1}}(\theta_J - \theta)) \\
&\quad + \int_{-\theta}^{\theta} d\theta_J \, G(\pi, \theta_{R_1^{-1}}(\theta_J - \pi + \theta), \\
&\quad \theta_{R_2^{-1}}(\theta_J - \pi - \theta)) \Bigg].
\end{aligned} \tag{38}$$

Now we examine the relation between (35) and (38). When $\beta = 1$, $\boldsymbol{U} = \cos\theta$ and $\Phi = 0$ or π. If $\cos\theta \geq 0$, $\Phi = 0$ and from (18),

$$\begin{aligned}
&\frac{\cos\theta_\gamma}{|\boldsymbol{U}| \sin|\Phi|} + \cot|\Phi| \\
&= \left(\frac{\cos\theta_\gamma}{\cos\theta} + 1\right) \frac{1}{\sin|\Phi|} + \cot|\Phi| - \frac{1}{\sin|\Phi|}.
\end{aligned}$$

Then, except for the single point $\theta_\gamma = \pi - \theta$, we have when $\beta \to 1$

$$\lim_{\Phi \to 0} \theta_{\Gamma^{-1}}(\theta_\gamma) = \begin{cases} 0, & \text{if } |\theta_\gamma| < \pi - \theta; \\ \pi, & \text{if } \pi > |\theta_\gamma| > \pi - \theta. \end{cases} \tag{39}$$

If $\cos\theta < 0$, $\Phi = \pi$ and from (18),

$$\begin{aligned}
&\frac{\cos\theta_\gamma}{|\boldsymbol{U}| \sin|\Phi|} + \cot|\Phi| \\
&= \left(\frac{\cos\theta_\gamma}{-\cos\theta} - 1\right) \frac{1}{\sin|\Phi|} + \cot|\Phi| + \frac{1}{\sin|\Phi|}.
\end{aligned}$$

Note $\lim_{x \to \pi} \cot x + 1/\sin x = 0$. Then we can get (39) again. In conclusion, (38) is the limit form of (35) when $\beta \to 1$, and when $\beta = 1$ we have a continuous point. For numerical computation, (38) may be used to provide a good approximation for $\beta \approx 1$.

III. Performance of SFH/*M*DPSK Under Multitone Jamming

In this section we consider the performance of uncoded SFH/MDPSK under multitone jamming. The transmitted M-ary DPSK signal has M possible differential phases $2\theta_i$ for $i = 1, \cdots, M$, with equal probability of transmission. The signal is hopped over N frequencies and is jammed with probability ρ. When the signal is jammed, it has the probability distribution as calculated in Section II. We assume that there are enough symbols per hop so that the energy loss due to the first dummy symbol of each hop is negligible. We assume that all jamming tones have equal power $I^2/2$. With a total jamming power J available, the number of jammed frequency slots is

$$Q = \frac{J}{I^2/2} = \frac{J}{S\beta^2},$$

where $S = E^2/2$ is the signal power. Suppose the hop frequency spacing is $1/T_s$, where T_s is the M-ary symbol period. Then the total number of hop frequency slots with total spread spectrum bandwidth W_{ss} is

$$N = \frac{W_{ss}}{1/T_s} = W_{ss} T_b \log_2 M,$$

where T_b is the bit period. Then

$$\rho = \frac{Q}{N} = \frac{J/(S\beta^2)}{W_{ss} T_b \log_2 M} = \frac{1}{\log_2 M\beta^2 E_b/J_O}, \tag{40}$$

where J_O is the equivalent broadband jamming power spectral density given by

$$J_O = J/W_{ss}.$$

E_b is the signal energy per bit. Therefore, β is defined as

$$\beta = \frac{1}{\sqrt{\log_2 M\rho E_b/J_O}}. \tag{41}$$

Note that $\rho \leq 1$ is a constraint, which implies $\beta \geq (1/\sqrt{\log_2 ME_b/J_O})$. The above result is derived in [1]. Since Q and N are integers, ρ is not continuous, as it appears to be. Nevertheless, when N is large we may *assume* that ρ is continuous for computational simplicity.

A decision region is specified for each of the M phases representing the M-ary signal. The probability that the received phase falls outside the decision region is the symbol error probability conditioned on the transmission of that signal. The sum of all M such conditional probabilities divided by M and averaged over the jamming state (whether a hop is jammed or not) is the average symbol error probability P_s. Specifically, when there is AWGN with one-side spectral density N_O (system thermal noise) which is not negligible,

$$\begin{aligned} P_s &= \rho P_{s1} + (1 - \rho) P_{s2} \\ &= \rho(P_{s1} - P_{s2}) + P_{s2} \\ &= \frac{1}{\log_2 M E_b / J_O \beta^2} (P_{s1} - P_{s2}) + P_{s2}, \\ \beta &\geq \frac{1}{\sqrt{\log_2 M E_b / J_O}}, \end{aligned} \tag{42}$$

where P_{s1} and P_{s2} are the symbol error rates (SER) conditional on that hop being jammed or not, respectively. P_{s1} is a function of both E_b/J_O and E_b/N_O, and P_{s2} is a function of E_b/N_O only. P_{s1} can be calculated by considering the signal as first being jammed and then further contaminated by the additive noise. Then we can use (35) and (38) to compute P_{s1}. Let G_i be the SER conditional on $\Gamma = \gamma$, $R_1 = r_1$, and $R_2 = r_2$. Let b_{i1} and b_{i2}, with $b_{i1} < b_{i2}$, be the bounds determining the decision region for differential phase $2\theta_i$. b_{i1} and b_{i2} lie within the particular 2π interval of interest (not necessarily $(-\pi, \pi]$). Then if $2\theta_i$ is transmitted, we have the conditional SER [9]

$$G_i(\gamma, r_1, r_2) = \begin{cases} F(b_{i1}) - F(b_{i2}), \\ \qquad b_{i1} - \theta_i < \gamma < b_{i2} - \theta_i; \\ 1 - F(b_{i2}) + F(b_{i1}), \\ \qquad b_{i1} - \theta_i > \gamma \text{ or } \gamma > b_{i2} - \theta_i, \end{cases} \tag{43}$$

where

$$F(b) = \frac{W \sin(\gamma + \theta_i - b)}{4\pi} \int_{-\pi/2}^{\pi/2} dt \cdot \frac{e^{-[U - V\sin t - W\cos(\gamma + \theta_{\mathrm{i}} - b)\cos t]}}{U - V \sin t - W \cos(\gamma + \theta_i - b) \cos t}, \tag{44}$$

and

$$U = \tfrac{1}{2}(\eta_2 + \eta_1), \qquad V = \tfrac{1}{2}(\eta_2 - \eta_1) \qquad W = \sqrt{\eta_1 \eta_2},$$

and

$$\eta_1 = \frac{r_1^2 T_b \log_2 M}{2N_O}, \qquad \eta_2 = \frac{r_2^2 T_b \log_2 M}{2N_O}.$$

Then we can write

$$G_i(\gamma, r_1, r_2) = G_i^*\left(\gamma, \frac{r_1^2 T_b}{2N_O}, \frac{r_2^2 T_b}{2N_O}\right). \tag{45}$$

For $\beta \neq 1$, from (35), we have

$$\begin{aligned} P_{s1} = P_{s1}\left(\beta, \frac{E_b}{N_O}\right) &= \frac{1}{M 2\pi} \sum_{i=1}^{M} \int_{-\pi}^{\pi} d\theta_J \\ &\times G_i^*\Bigg(\theta_{\Gamma^{-1}}(\theta_J)\Big|_{\theta = \theta_i}, \frac{E_b}{N_O}(1 + \beta^2 + 2\beta \\ &\cdot \cos(\theta_J + \theta_i)), \\ &\frac{E_b}{N_O}(1 + \beta^2 + 2\beta\cos(\theta_J - \theta_i))\Bigg). \end{aligned} \tag{46}$$

For $\beta = 1$, from (38), we have

$$\begin{aligned} P_{s1} = \frac{1}{M 2\pi} \sum_{i=1}^{M} \Bigg[&\int_{-(\pi - \theta_i)}^{\pi - \theta_i} d\theta_J \, G_i^* \\ &\cdot \Bigg(0, \frac{E_b}{N_O}(1 + \beta^2 + 2\beta\cos(\theta_J + \theta_i)), \\ &\frac{E_b}{N_O}(1 + \beta^2 + 2\beta\cos(\theta_J - \theta_i))\Bigg) \\ &+ \int_{-\theta_i}^{\theta_i} d\theta_J \, G_i^*\Bigg(\pi, \frac{E_b}{N_O}(1 + \beta^2 - 2\beta \\ &\cdot \cos(\theta_J + \theta_i)), \\ &\frac{E_b}{N_O}(1 + \beta^2 - 2\beta\cos(\theta_J - \theta_i))\Bigg)\Bigg]. \end{aligned} \tag{47}$$

Similarly, we have

$$P_{s2} = P_{s2}\left(\frac{E_b}{N_O}\right) = \frac{1}{M} \sum_{i=1}^{M} G_i^*\left(\theta_i, \frac{E_b}{N_O}, \frac{E_b}{N_O}\right). \tag{48}$$

Note that for a given E_b/N_O, P_{s2} is a constant and P_{s1} is the function of β.

To determine the worst case ρ, ρ_{wc}, which maximizes P_s for a given E_b/J_O and E_b/N_O, we rewrite (42) as

$$\begin{aligned} \frac{E_b}{J_O}\left(P_s - P_{s2}\left(\frac{E_b}{N_O}\right)\right) &= \frac{1}{\log_2 M \beta^2}\left(P_{s1}\left(\beta, \frac{E_b}{N_O}\right) - P_{s2}\left(\frac{E_b}{N_O}\right)\right) \\ &= \zeta\left(\beta, \frac{E_b}{N_O}\right). \end{aligned} \tag{49}$$

Suppose that, with the constraint $\beta \geq 1/\sqrt{\log_2 M E_b / J_O}$, $\beta = \beta_{wc}$ gives the maximum $\zeta(\beta, (E_b/N_O)) = \zeta_{\max}$. Then

$$\rho_{wc} = \frac{1}{\log_2 M E_b / J_O \beta_{wc}^2}, \tag{50}$$

and the worst case SER is

$$P_{s_{wc}} = \frac{\zeta_{\max}}{E_b/J_O} + P_{s2}\left(\frac{E_b}{N_O}\right). \tag{51}$$

Note that, in general, β_{wc} and $\zeta_{\max}$ are functions of both E_b/N_O and E_b/J_O. In two special cases which are commonly encountered, β_{wc} and $\zeta_{\max}$ are functions of E_b/N_O only.

If $\beta_{wc} = 1/\sqrt{\log_2 ME_b/J_O}$ for a *range* of E_b/J_O (which may occur for small E_b/J_O when, e.g., ζ is a function of β with a single maximum), then $\rho_{wc} = 1$ which corresponds to full-band multitone jamming and

$$P_{s_{wc}} = P_{s1}\left(\frac{1}{\sqrt{\log_2 ME_b/J_O}}, \frac{E_b}{N_O}\right). \qquad (52)$$

If $\beta_{wc} = \beta_{wc}(E_b/N_O)$ for a *range* of E_b/J_O (which may occur for large E_b/J_O when, e.g., ζ is a function of β with a single maximum), then

$$\rho_{wc} = \frac{1}{\log_2 M\beta_{wc}^2}\frac{1}{E_b/J_O}, \qquad (53)$$

which corresponds to an inverse linear function of E_b/J_O with a slope (or the vertical shift in the logarithmic scale) dependent on E_b/N_O, and

$$P_{s_{wc}} = \frac{\zeta_{\max}}{E_b/J_O} + P_{s2}\left(\frac{E_b}{N_O}\right) \qquad (54)$$

which corresponds to a similar inverse linear function *plus* a floor SER due to the AWGN.

If the system thermal noise can be neglected, then $P_{s2} = 0$ and

$$P_{s1} = 1 - \frac{1}{M}\sum_{i=1}^{M}\int_{b_{i1}}^{b_{i2}} \tilde{p}_\Gamma(\gamma)\big|_{\theta=\theta_i}\, d\gamma. \qquad (55)$$

The function $P_{s_{wc}}$ may be optimized with respect to the signal phases and decision regions. For example, for binary DPSK ($M = 2$), we can have: 1) phase $2\theta_1 = \pi/2$ corresponding to 0 and phase $2\theta_2 = 3\pi/2$ corresponding to 1; reasonable decision regions in this case are $[0, \pi]$ for 0 and $(-\pi, 0)$ for 1; or 2) phase $2\theta_1 = 0$ corresponding to 0 and phase $2\theta_2 = \pi$ corresponding to 1; reasonable decision regions are $[-\psi_1, \psi_1]$ for 0 and the rest of the phasor plane for 1. In scheme 2), without thermal noise, $\psi_1 = 0$ would make $P_s = 0$ because when 0 is transmitted, the continuous jamming tone could ever alter the transmitted differential phase. Thus, in this case, inclusion of the thermal noise in the analysis is indispensable and the desirable ψ_1 is greater than 0. The peculiarity in scheme 2) does not exist for similar signal phase schemes with $M > 2$.

To compare system performance for different M, we must convert P_s into an equivalent bit error rate (BER) P_b. For a small signal-to-noise ratio, we can use an orthogonal model which results in

$$P_b \approx \frac{M}{2(M-1)}P_s.$$

If a Gray code is used, so that the Hamming distance of the binary representation of adjacent differential signal phases is 1, we may have, for a large signal-to-noise ratio [11]

$$P_b \approx \frac{1}{\log_2 M}P_s.$$

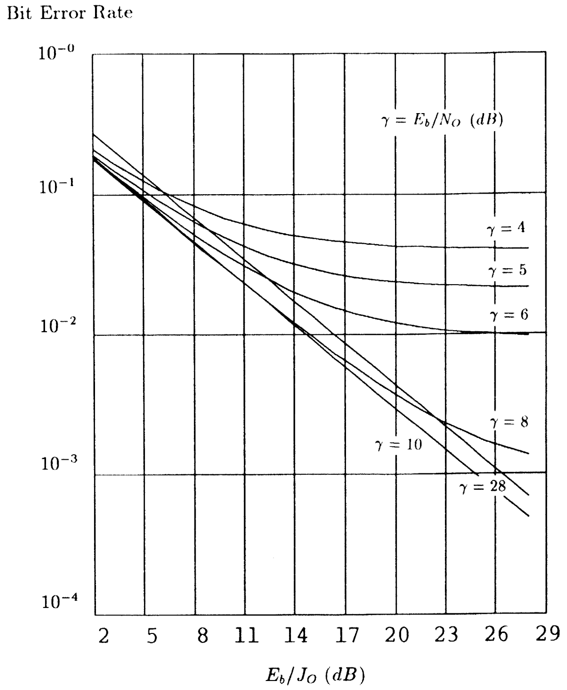


Fig. 2. Worst case BER versus E_b/J_O(dB) for E_b/N_O(dB) = 4, 5, 6, 8, 10, 28, for binary DPSK with $2\theta_1 = 0$ and $2\theta_2 = \pi$. Decision regions are equal and symmetric.

A. Performance Analysis

In this section we present some numerical results based on the analysis in the preceding sections. As examples, we give results for $M = 2$ and 4. Performance is measured by the worst case BER for the DPSK signaling scheme, and specific E_b/J_O and E_b/N_O. Two configurations were given in the preceding section for binary DPSK. One had $2\theta_1 = 0$ and $2\theta_2 = \pi$, and the other $2\theta_1 = \pi/2$ and $2\theta_2 = 3\pi/2$. Fig. 2 shows the worst case BER performance of the first of these schemes for E_b/J_O(dB) from 2 to 28, and E_b/N_O(dB) = 4, 5, 6, 8, 10, 28. This figure is the same as Fig. 3 in [7]. Note the threshold effect due to the noise level. The corresponding plot of E_b/N_O(dB) versus BER for E_b/J_O(dB) = 0, 2, 4, 6, 10, 15, 20 is given in Fig. 3. This plot also shows the threshold effect, this time due to the fixed E_b/J_O. The worst case jamming parameter versus E_b/J_O is given in Fig. 4. This is the same as Fig. 4 in [7]. The second signaling scheme is $2\theta_1 = \pi/2$ and $2\theta_2 = 3\pi/2$. The BER performance of this scheme, versus E_b/J_O, is given in Fig. 5. Comparison to the first scheme shows that using $2\theta_1 = 0$ is best, as indicated in the previous section. Asymptotically, they are identical in performance, but for large E_b/J_O, the first scheme is better, i.e., when $E_b/J_O >> E_b/N_O$. The two previous schemes were evaluated with symmetric decision regions. If we now take the first scheme and distort the decision

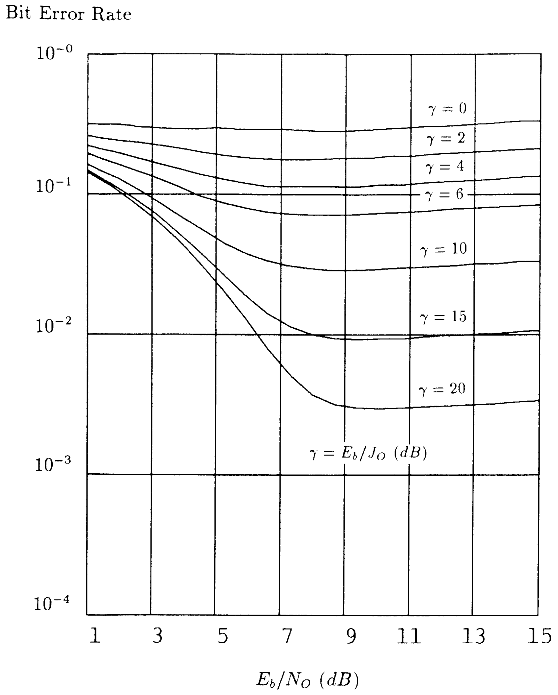


Fig. 3. Worst case BER versus E_b/N_O(dB) for E_b/J_O(dB) = 0, 2, 4, 6, 10, 15, 20, for binary DPSK with $2\theta_1 = 0$ and $2\theta_2 = \pi$. Decision regions are equal and symmetric.

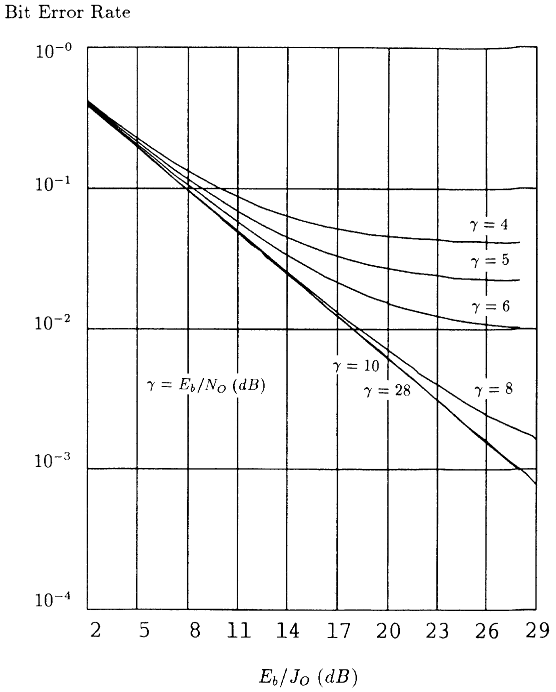


Fig. 5. Worst case BER versus E_b/J_O(dB) for E_b/N_O(dB) = 4, 5, 6, 8, 10, 28, for binary DPSK with $2\theta_1 = \pi/2$ and $2\theta_2 = 3\pi/2$. Decision regions are equal and symmetric.

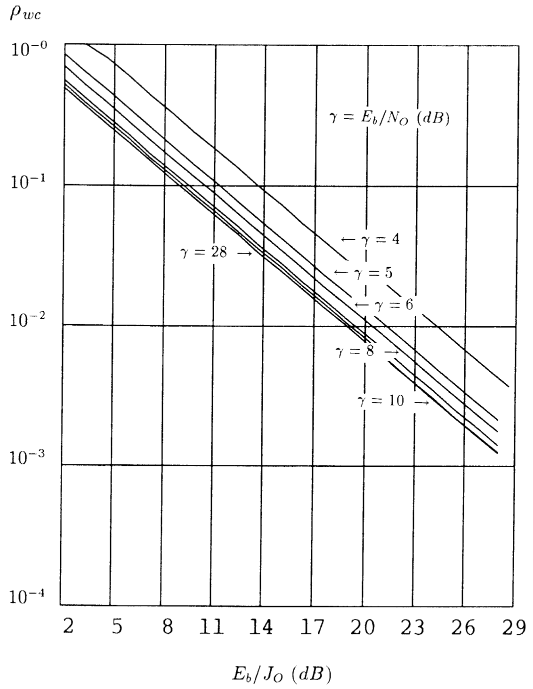


Fig. 4. ρ_{wc} versus E_b/N_O(dB) = 4, 5, 6, 8, 10, 28, for binary DPSK with $2\theta_1 = 0$ and $2\theta_2 = \pi$. Decision regions are equal and symmetric.

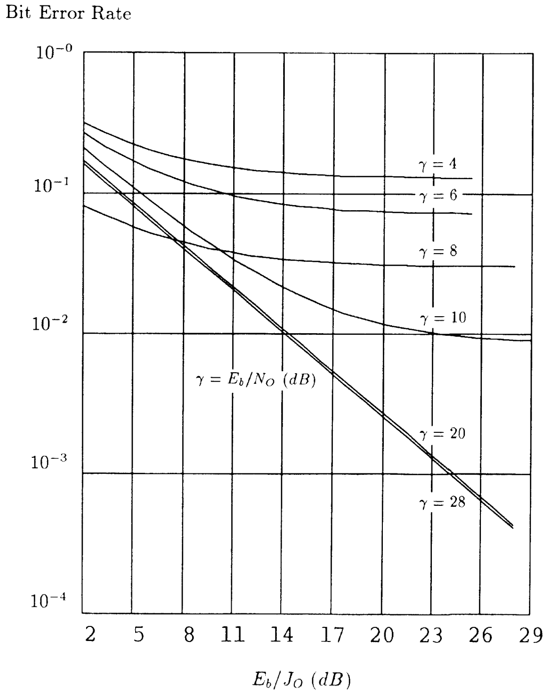


Fig. 6. Worst case BER versus E_b/J_O(dB) for E_b/N_O(dB) = 4, 5, 6, 8, 10, 28, for binary DPSK with $2\theta_1 = 0$ and $2\theta_2 = \pi$. The decision region boundaries are $\pi/4$ and $-\pi/4$.

regions so that $b_{11} = \pi/4$ and $b_{12} = -\pi/4$, and $b_{21} = \pi/4$ and $b_{22} = 7\pi/4$, we get the result shown in Fig. 6. The performance of this scheme is worse than those with equal decision regions when E_b/N_O is large, but superior when E_b/N_O is small. Thus, the choice of the best DPSK signaling scheme is dependent upon the relative strength of the noise and tone jamming.

For $M = 4$, we look at two symmetric signaling schemes with equal decision regions. The first has $2\theta_1 = 0$, and the second has $2\theta_1 = \pi/4$. Figs. 7 and 8 give the

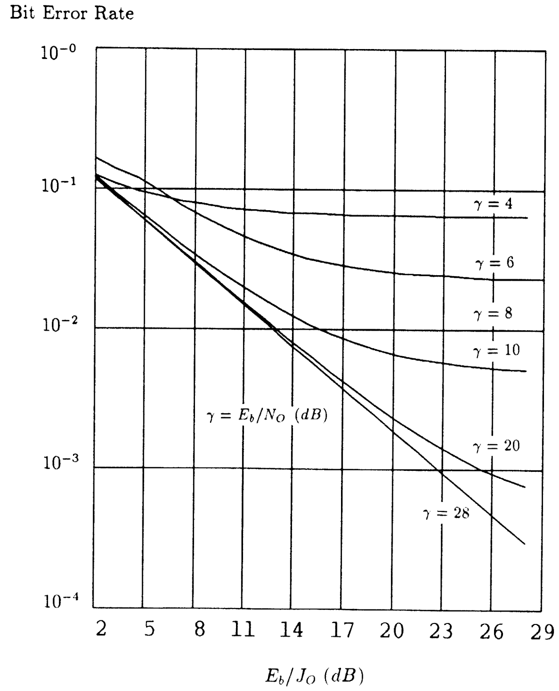


Fig. 7. Worst case BER versus E_b/J_O(dB) for E_b/N_O(dB) = 4, 5, 6, 8, 10, 28, for 4-ary DPSK with $2\theta_1 = 0$. The decision regions are equal and symmetric.

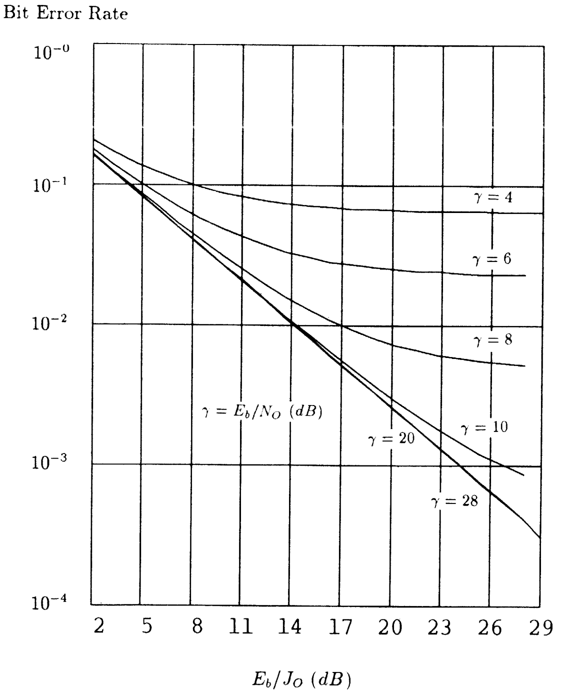


Fig. 8. Worst case BER versus E_b/J_O(dB) for E_b/N_O(dB) = 4, 5, 6, 8, 10, 28, for 4-ary DPSK with $2\theta_1 = \pi/4$. The decision regions are equal and symmetric.

BER performance of these schemes, respectively, for $E_b/J_O = 2$ to 28, and E_b/N_O(dB) = 4, 6, 8, 10, 20, 28. From these figures, it is clear that choosing $2\theta_1 = 0$ is the best when E_b/J_O is large, as was the case for binary DPSK. As expected, the performance of 4-ary DPSK is better than binary DPSK.

IV. Concluding Remarks

This paper has addressed some basic problems associated with SFH/DPSK. General probability distributions are derived for arbitrary DPSK signals. Applying these distributions, we have evaluated the performance of SFH/DPSK under both tone jamming and system thermal noise for $M = 2$ and 4. The performance results indicate that choosing $2\theta_1 = 0$ is best under tone jamming, and equal and symmetrical decision regions are best when the noise is predominant, with the choice of $2\theta_1$ arbitrary.

References

[1] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications*, Vol. II. Rockville, MD: Computer Science Press, 1985.

[2] J. S. Lee, R. H. French, and L. E. Miller, "Error correcting codes and nonlinear diversity combining against the worst case partial-band noise jammer of frequency-hopping MFSK systems," *IEEE Trans. Commun.*, vol. 36, pp. 471-478, Apr. 1988.

[3] J. S. Bird and E. B. Felstead, "Antijam performance of fast frequency-hopped *M*-ary NCFSK—An overview," *IEEE J. Select. Areas Commun.*, vol. SAC-4, pp. 216-233, Mar. 1986.

[4] Q. Wang, T. A. Gulliver, V. K. Bhargava, and E. B. Felstead, "Coding for fast frequency hopped noncoherent MFSK spread spectrum communications under worst case jamming," in *Proc. IEEE MILCOM*, 1988, pp. 15.4.1-15.4.7.

[5] S. W. Houston, "Modulation techniques for communication, Part 1: Tone and noise jamming performance of spread spectrum *M*-ary FSK and 2, 4-ary DPSK waveforms," in *NAECON'75 Rec.*, pp. 51-58.

[6] M. K. Simon, "The performance of *M*-ary DPSK/FH in the presence of partial-band multitone jamming," *IEEE Trans. Commun.*, vol. COM-30, pp. 953-958, May 1982.

[7] K. S. Gong, "Performance analysis of FH/DPSK in additive white Gaussian noise (AWGN) and multitone jamming," in *Proc. IEEE MILCOM*, 1988, pp. 53.4.1-53.4.7.

[8] J. K. Winters, "On differential detection of *M*-ary DPSK with intersymbol interference and noise correlation," *IEEE Trans. Commun.*, vol. COM-35, pp. 117-120, Jan. 1987.

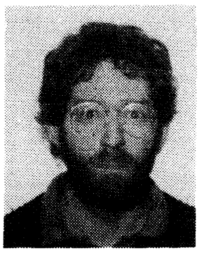
[9] R. F. Pawula, S. O. Rice, and J. H. Roberts, "Distribution of the phase angle between two vectors perturbed by Gaussian noise," *IEEE Trans. Commun.*, vol. COM-30, pp. 1828-1841, Aug. 1982.

[10] R. F. Pawula, "On the theory of error rates for narrow-band digital FM," *IEEE Trans. Commun.*, vol. COM-29, pp. 1634-1643, Nov. 1981.

[11] R. E. Ziemer and R. L. Peterson, *Digital Communications and Spread Spectrum Systems*. New York: Macmillan, 1985.

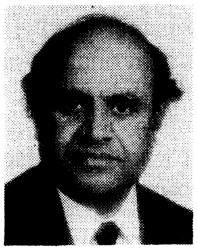
Qiang Wang (S'87-M'88) was born in Beijing, China, in 1961. He received the B.Sc. and M.Sc. degrees from the Nanjing Communications Engineering Institute, Nanjing, China, in 1982 and 1985, respectively, and the Ph.D. degree from the University of Victoria, B.C., Canada, in 1988, all in electrical engineering.

From 1987 to March 1990 he was a Member of the Technical Staff at Microtel Pacific Research, Burnaby, B.C., Canada. Since then, he has been an Assistant Professor in the Department of Electrical and Computer Engineering at the University of Victoria. His current research interests include error control coding, spread spectrum communications, and some aspects of neural networks.



T. Aaron Gulliver was born in Bathurst, N.B., Canada, in 1960. He received the B.Sc. and M.Sc. degrees in Electrical Engineering from the University of New Brunswick, Fredericton, N.B., in 1982 and 1984, respectively. He received the Ph.D. degree in electrical and computer engineering from the University of Victoria, Victoria, B.C., Canada, in 1989.

He is currently with the Defence Research Establishment, Ottawa, Ont., Canada, where he is employed as a Defence Scientist. His current interests include coding theory and its application to spread spectrum communications.



Vijay K. Bhargava (S'70-M'74-M'76-SM'82) received the B.Sc. degree (Hons.) from the University of Rajasthan in 1966, and the B.Sc. degree in mathematics and engineering, and the M.Sc. and Ph.D. degrees, both in electrical engineering, from Queen's University in Kingston, Ont., Canada, in 1970, and 1972, and 1974, respectively.

After brief stays at the Indian Institute of Science and the University of Waterloo, he joined Concordia University in Montreal and was promoted to Professor in 1984. For 1982–1983 he was on sabbatical leave at Ecole Polytechnique de Montréal. In August 1984 he joined the newly formed Faculty of Engineering at the University of Victoria as a Professor of Electrical Engineering. In July 1988 he was appointed a Fellow of the BC Advanced Systems Institute. His research interest is in the area of digital communications with special emphasis on error control coding techniques, cryptography, and spread spectrum communications. He has been a consultant to Bell Northern Research, Microtel Pacific Research, Mobile Data International, GE Mobile Communication Group, Transport Canada, Revenue Canada, Communications Research Centre (CRC) of the Federal Department of Communications and Ministére des Communications, Gouvernement du Québec. He is a coauthor of the book *Digital Communications by Satellite* (New York: Wiley, 1981). A Japanese translation of the book was published in 1984, while a Chinese translation was published in 1987.

Dr. Bhargava is a Past Chairman of the IEEE Montreal Section (1981–1982) and a Past Director of the IEEE Montreal Conferences Inc. (1980–1982). He was a Co-Vice Chairman of the 1983 IEEE International Symposium on Information Theory and served as the Founding Chairman of the IEEE Information Theory Group Chapter in the IEEE Montreal Section. He served as the Chairman of IEEE Victoria Section during 1986 to 1987 and as the Founding Chairman of the 1987 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing. In January 1988 he assumed the Editorship of the *Canadian Journal of Electrical and Computer Engineering*. Currently he is serving as Chairman of IEEE Western Canada Council. He is a recipient of the IEEE Centennial Medal, EIC Centennial Medal, and the 1987 A. F. Bulgin Premium awarded by the Institute of Radio and Electronic Engineers (UK). In March 1988 he was elected a Fellow of the Engineering Institute of Canada.

Performance of Direct-Sequence Spread-Spectrum Receiver Using Decision Feedback and Transversal Filters for Combatting Narrowband Interference

MIROSLAV L. DUKIĆ, MEMBER, IEEE, ZORKA D. STOJANOVIĆ, MEMBER, IEEE, AND ILIJA S. STOJANOVIĆ, SENIOR MEMBER, IEEE

Abstract—**A direct-sequence spread-spectrum (DSSS) receiver using both decision feedback (DFB) and two-sided transversal filters for combatting narrowband interference (NBI) is proposed. The receiver is made up of two branches. In the first branch, the conventional demodulator is followed by a DFB filter, while in the second auxiliary branch, a demodulator with the carrier in quadrature is followed by a two-sided adaptive transversal (AT) filter. Performance of this receiver has been analyzed on the basis of the calculated mean-square error and the probability of error at the output of the receiver. Special attention was paid to the effects caused by the propagation of errors in the DFB filter. The results obtained show that NBI rejection is fairly high, and practically does not depend upon the difference of frequencies of the desired and interfering carriers nor upon the interfering carrier level.**

I. Introduction

The inherent property of DSSS systems is their increased immunity against interference. This increase is proportional to the system processing gain. However, in real conditions, mainly in those characterized by an adverse electromagnetic environment, the processing gain is often insufficient to ensure the desired interference rejection. Then, special means are applied to protect the system against the interference.

Recently, several papers in the literature have been devoted to the rejection of NBI in DSSS systems [1]–[12]. They dealt with different proposals aiming at an additional interference reduction besides that inherent to DSSS systems. This is of particular importance in the case of a narrowband interferer which can be considered as a very efficient jammer of the DSSS systems, especially when the processing gain is not as high as desired. However, none of these solutions is completely satisfactory.

In his review paper [13], Milstein has described different interference rejection techniques, concluding that there is still much to be learned in the area of interference rejection. Following this idea, the attention of the authors of this paper has been paid to the improvement of those NBI rejection techniques which only make use of adaptive transversal (AT) filters without and with decision feedback.

In this paper, a new DSSS receiver is proposed in which good features of the receivers described in earlier papers [2], [5], [6], [10], [11] are combined. The receiver is made up of two parallel branches. In the first of them, the conventional demodulator is followed by a DFB filter, while in the second, auxiliary branch, a demodulator with the carrier in quadrature is followed by a two-sided AT filter. Thus, using the DFB filter instead of the AT filter [10], the distortion of the desired signal has been avoided. This is particularly important when carrier frequencies of the desired and interfering signals are equal.

It may be worthwhile to mention that the idea of using a DFB filter in OQPSK systems for combatting the NBI was treated in [15]. The role of this filter was only to decrease the useful signal distortion, while in this paper, the DFB filter is applied just to reduce the NBI.

The analysis of the receiver performance is presented in detail, and the explicit expressions for the average error probability and the mean-square error are deduced. Special attention has been paid to the analysis of the decision error propagation through the DFB filter.

The results obtained are improved in respect to those previously known, and show that NBI rejection practically is not dependent on the carrier frequency offsets, nor upon the interfering signal level. In addition, the effect of the error propagation through the DFB filter on the total error probability can be neglected.

This paper is divided into seven sections. In Section II, a brief discussion of the DSSS receiver, previously proposed [11], using AT interference rejection filters without decision feedback, is presented. Theoretical analysis of the receiver proposed in this paper is presented in Section III, while in Section IV, the signal-to-interference ratio improvement is evaluated. In Section V, the probability of error analysis is presented. Some examples and discussion of the results obtained are given in Section VI. The last section summarizes the conclusion.

Manuscript received March 8, 1989; revised December 17, 1989. This work was presented at the Mediterranean Electrotechnical Conference, Lisboa, Portugal, April 11–13, 1989.

The authors are with the Department of Communications, Faculty of Electrical Engineering, Bul. Revolucije 73, Belgrade, Yugoslavia.

IEEE Log Number 9035816.

II. Background of the DSSS Receivers Employing Interference Suppression Filters Without Decision Feedback

The process of NBI reduction using AT filters is based in its very nature on the absence of correlation between adjacent chips of the desired signal. Namely, on the assumption that the pseudorandom sequence used is long enough, it can be assumed that the values of the desired signal in different taps of the AT filter are mutually uncorrelated, the delay of each tap being equal to the chip duration. Therefore, by using such a filter, the estimation of signal parameters cannot be done. On the other hand, when considering the NBI signal, the situation is quite different, for the interfering signal values at the tap outputs of the AT filter are mutually correlated. That means that in the process of the linear mean-square estimation of the total signal entering the AT filter, only the estimation of the NBI signal can be performed. This yields the reference output interfering signal, which can be subtracted from the input signal thus reducing the interference.

The structure of the DSSS receiver based on the above-mentioned idea, using two two-sided AT filters, is shown in Fig. 1 [11]. In the upper branch of this receiver, the conventional coherent demodulator is followed by one of the transversal filters, TF1, while the other, TF2, has been placed in the parallel quadrature branch. The performance of such a receiver is much better than that of those proposed earlier [2], [5], [6]. The average probability of error when both filters, or only TF1 or only TF2, are used, as a function of the interference-to-signal power ratio Γ and the product ΩT for this receiver, is shown in Figs. 2 and 3, respectively [11].

Analysis of these diagrams clearly indicates that the effect of the filter TF2 is the quadrature branch on the rejection of the interference is much more significant than that of TF1 in those cases in which $\Omega T > 5°$. For $\Omega T = 0$, the filter TF2 contributes almost nothing to the interference rejection.

However, from Fig. 3 it can be seen that even this receiver, while having a very small error probability determined practically by the additive white Gaussian noise (AWGN) for frequency offsets $\Omega T > 5°$, does not give satisfactory error probability in the vicinity of the desired signal carrier frequency, i.e., when ΩT is small ($<5°$). Just this region of frequency offsets is the most critical. Namely, the desired signal together with the interference and the AWGN are transmitted through TF1. This filter acts as a notch filter against the interferer, and a part of desired signal energy is excised. The consequence is signal distortion manifested as intersymbol interference (ISI), thus causing the increase of the error probability. The reason for such distortion lies in the residual correlation between the adjacent chips of the desired signal, a phenomenon which has not drawn due attention in the literature as yet.

In order to emphasize the impact of this distortion on the performance of the receiver from Fig. 1, in Fig. 3 the dashed line shows average error probability under the assumption that there is no signal distortion and that the tap weights have the optimum values. These results show clearly that when the processing gain is not sufficient, the distortion of the desired signal due to the filter TF1 significantly decreases the protection against co-channel NBI.

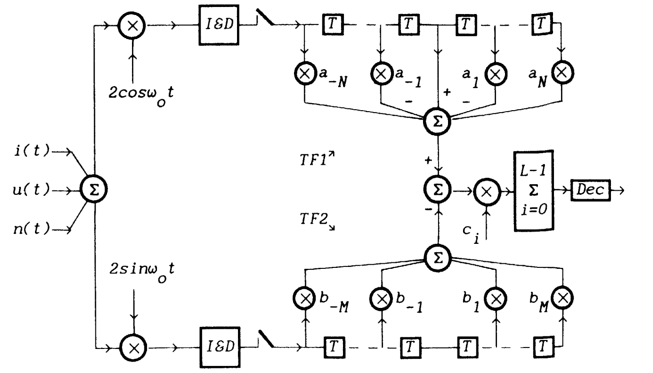


Fig. 1. The block diagram of the receiver [11]; *I&D* is integrate and dump, $u(t)$ and $i(t)$ are desired and interfering signal, respectively, $n(t)$ is AWGN, T is the chip interval, L is processing gain, N and M are the number of AT filters taps, and *Dec* is decision circuit.

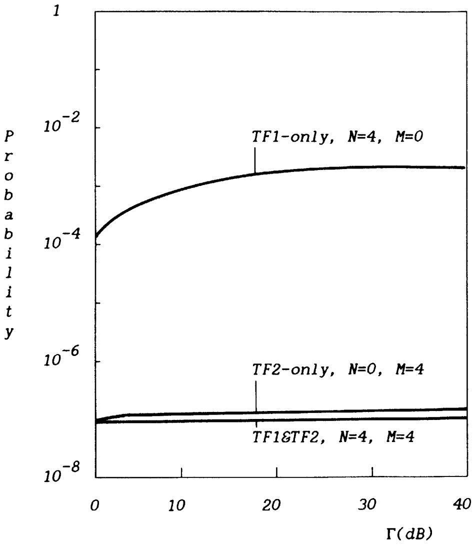


Fig. 2. The average error probability at the output of the receiver [11], as a function of the ratio Γ of the interfering and desired signal powers. M and N are the numbers of taps on each side of the filters TF1 and TF2, $L = 7$ is the processing gain, $\Omega T = 2\pi/L$, where Ω is the interfering carrier offset from the DSSS signal carrier, and T is the chip interval, and signal-to-noise ratio is $A_n = 12$ dB.

In the receiver proposed in this paper, the filter TF1 is replaced by a DFB filter, as in [10]. Now, assuming that the signal at the output is error-free, when fed back at the DFB filter input, it is subtracted from the sum of the signal, interference, and AWGN. In this way, the desired signal is not transmitted through the DFB filter and, as a result, there is no signal distortion.

III. Theoretical Analysis

A. Description of the Proposed Receiver

The block diagram of the proposed receiver is shown in Fig. 4. The received signals are demodulated to baseband, integrated, and then sampled at the chip rate of the

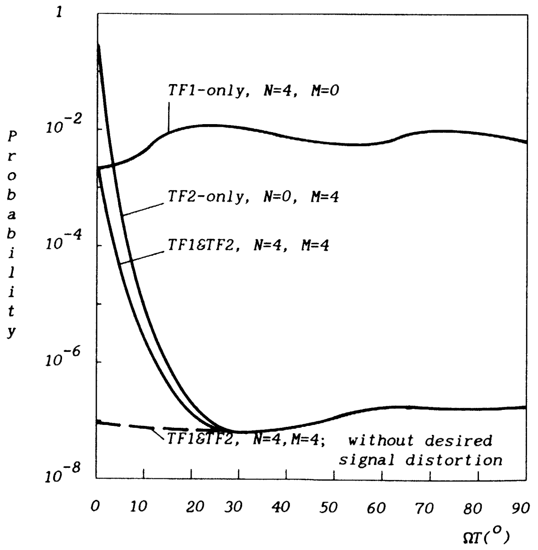


Fig. 3. The average error probability at the output of the receiver [11], as a function of the product ΩT. M and N are the numbers of taps on each side of the filters TF1 and TF2, $L = 7$ is the processing gain, signal-to-noise ratio is $A_n = 12$ dB, and $\Gamma = 20$ dB is interference-to-signal ratio.

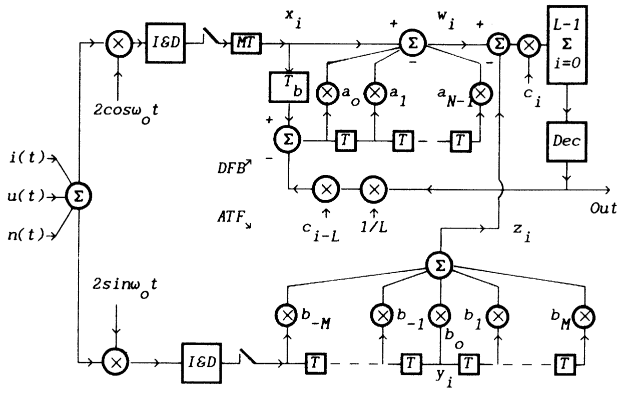


Fig. 4. The block diagram of the proposed receiver; *I&D* is integrate and dump and *Dec* is a decision circuit.

pseudorandom sequence (PRS). The DFB AT filter in the upper branch has N taps, while the number of taps in the two-sided AT filter in the lower branch is $2M + 1$. The duration of a chip in the PRS is T, while the duration of a bit is denoted with $T_b = LT$, where L is the system processing gain.

The desired signal at the receiver input is

$$u(t) = Ud(t)\ \mathrm{PRS}\ (t) \cos \omega_o t, \tag{1}$$

where U and ω_o are the amplitude and the angular frequency of the carrier, respectively, and $d(t)$ is a random binary signal representing the message given by

$$d(t) = \sum_i d_i \Pi(t - iT_b), \qquad d_i \in \{1, -1\}. \tag{2}$$

$\Pi(t)$ is the rectangular pulse of the unit amplitude and of the duration T_b.

The PRS is described by

$$\mathrm{PRS}\ (t) = \sum_i c_i \Pi(t - iT), \qquad c_i \in \{1, -1\}. \tag{3}$$

The NBI signal at the receiver input is:

$$i(t) = \mathrm{Re}\left\{I(t) \exp\left[j\varphi(t)\right] \exp\left(j\omega_o t\right)\right\}, \tag{4}$$

where $I(t) \exp [j\phi(t)]$ is the complex low-frequency equivalent of the signal $i(t)$ and Re $[\cdot]$ stands for the real part.

In addition to the signals given by (1) and (4), AWGN is present also at the receiver input:

$$n(t) = n_c(t) \cos \omega_o t + n_s(t) \sin \omega_o t. \tag{5}$$

The in-phase and quadrature noise component $n_c(t)$ and $n_s(t)$, respectively, are assumed to be independent of each other. The one-sided power spectral density of the AWGN is denoted by η.

It is assumed that all the signals at the receiver input are mutually statistically independent.

B. Analytical Analysis

The signal samples x_i in the upper branch and y_i in the lower branch of the receiver in Fig. 4, at the ith sampling instant in the jth bit interval considered, are given by

$$x_i = Ud_j c_i T + I_{c,i} + n_{c,i}, \tag{6}$$

$$y_i = I_{s,i} + n_{s,i}. \tag{7}$$

The time dependence of the terms in (6) and (7) has been omitted for the sake of simplicity. The first term in (6), $Ud_j c_i T$, is desired signal sample, where d_j is the transmitted bit, $I_{c,i}$ and $I_{s,i}$ are the NBI signal samples in the corresponding branches, and $n_{c,i}$ and $n_{s,i}$ are the samples of the noise $n_c(t)$ and $n_s(t)$, respectively.

Since during the decision process an error can be made, the signal at the receiver output is appropriately described as

$$r = Ud_j LT + q_I + q_N + q_D, \tag{8}$$

where

$$q_I = \sum_{i=0}^{L-1} c_i \left[I_{c,i} - \sum_{k=0}^{N-1} a_k I_{c,i-L-k} - \sum_{k=-M}^{M} b_k I_{s,i-k} \right], \tag{9}$$

represents the output signal component due to the interfering signal, and

$$q_N = \sum_{i=0}^{L-1} c_i \left[n_{c,i} - \sum_{k=0}^{N-1} a_k n_{c,i-L-k} - \sum_{k=-M}^{M} b_k n_{s,i-k} \right], \tag{10}$$

describes the output signal component caused by the AWGN.

The last term in (8) is the ISI at the receiver output, caused by the erroneous decisions in the previous bit intervals, which propagated through the DFB filter. If a reasonable assumption is made that the number of taps in the DFB filter is less than the processing gain, then the signal value in the jth bit interval can be affected only by the decision errors in the two foregoing bit intervals. If ξ_{j-1}

and ξ_{j-2} are two decision errors at the input of the DFB filter in $(j-1)$th and $(j-2)$th bit intervals, respectively, the following sets of events are possible:

$$\text{(i) } \xi_{j-2} = 0, \quad \text{(ii) } \xi_{j-2} = 2d_{j-2}, \quad \text{(iii) } \xi_{j-2} = 2d_{j-2}$$
$$\xi_{j-1} = 2d_{j-1}, \qquad \xi_{j-1} = 0 \qquad\qquad \xi_{j-1} = 2d_{j-1}. \tag{11}$$

Consequently, the distortion q_D at the receiver output is given by

$$q_D = \sum_{i=0}^{L-1} UTc_i \left[\sum_{k=0}^{i} a_k \xi_{j-1} c_{i-L-k} + \sum_{k=i+1}^{N-1} a_k \xi_{j-2} c_{i-L-k} \right]. \tag{12}$$

Now, the total error at the receiver output is

$$e = \sum_{i=0}^{L=1} e_i = q_I + q_N + q_D. \tag{13}$$

In accordance to the principle of linear mean-square estimation [14], the mean-square error at the receiver output in the jth bit interval will have a minimal value when the error e_i is orthogonal both to the signal at the DFB filter input in the upper branch, and to the signal y_i in the lower branch of the receiver. These conditions are written as follows:

$$E\{e_i [UTc_{i-L-l} \xi_{j-1} + I_{c,i-L-l} + n_{c,i-L-l}]\} = 0 \mid l = 0, 1, \cdots, N-1, \tag{14}$$

$$E\{e_i [I_{s,i-m} + n_{s,i-m}]\} = 0 \mid m = -M, \cdots, M. \tag{15}$$

Using these conditions, expressions (6)–(13), and if it is assumed that the PRS is sufficiently long, the following expression for the minimum mean-square error at the receiver output is obtained:

$$\overline{e^2_{m,DFB}} = \sum_{i=0}^{L-1} E\left[I^2_{c,i} + n^2 - \sum_{k=0}^{N-1} a_k I_{c,i} I_{c,i-L-k} - \sum_{k=-M}^{M} b_k I_{c,i} I_{s,i-k} + \sum_{k=i+1}^{N-1} (a_k UT\xi_{j-2})^2 \right], \tag{16}$$

where

$$E(n^2_{c,i}) = E(n^2_{s,i}) = n^2 \quad \text{and}$$
$$E(n_{c,i}) = E(n_{s,i}) = 0. \tag{17}$$

C. Determination of the Tap Weights Under Single-Tone Interference

The optimal values of the tap weights of filters in the receiver proposed are those for which the mean-square value of the error at the receiver output reaches its minimum. They are determined by means of [14] and [15]. In stationary operating conditions, when no decision errors at the output of the receiver have occurred, the error probability is very small, say $< 10^{-6}$. Then, it is justified to assume that the optimal tap weights of filters can be determined for the case when there are no decision errors. This assumption is also justified furthermore by the speed of adaptation of the filter tap weights [13].

When a single-tone interfering signal is present at the receiver input,

$$i(t) = I \cos[(\omega_o + \Omega)t + \theta], \tag{18}$$

where I is the amplitude of the interferer and Ω is its carrier frequency offset from the DSSS signal carrier. The random phase θ is uniformly distributed over the interval $[0, 2\pi)$.

The interfering signal samples will be

$$I_{c,i} = \int_{iT}^{(i+1)T} I \cos(\Omega t + \theta)\, dt$$
$$I_{s,i} = \int_{iT}^{(i+1)T} -I \sin(\Omega t + \theta)\, dt. \tag{19}$$

Replacing these expressions in (14) and (15), under assumption that there is no error propagation, the system of the following $(2M + N + 1)$ equations is obtained:

$$\cos(L + l)\Omega T - a_l(L/\Gamma A_n \mu^2) - \sum_{k=0}^{N-1} a_k \cos(l - k)\Omega T - \sum_{k=-M}^{M} b_k \sin(k - L - l)\,\Omega T = 0, \tag{20}$$

$$\sin m\Omega T - b_m(L/\Gamma A_n \mu^2) - \sum_{k=0}^{N-1} a_k \sin(m - L - k)\Omega T - \sum_{k=M}^{M} b_k \cos(m - k)\Omega T = 0, \tag{21}$$

where $l = 0, \cdots, N-1$, $m = -M, \cdots, M$.

Their solutions give the following values for the optimal tap weights:

$$a_k = \frac{(\lambda + \alpha + \nu)\cos(k + L)\Omega T - \gamma \sin(k + L)\Omega T}{(\lambda + \beta + \chi)(\lambda + \alpha + \nu) - \gamma^2}, \tag{22}$$

$$b_k = \frac{\gamma \cos k\Omega T + (\lambda + \alpha + \nu)\sin k\Omega T}{(\lambda + \beta + \chi)(\lambda + \alpha + \nu) - \gamma^2}. \tag{23}$$

In these expressions,

$$\chi = \sum_{k=0}^{N-1} \cos^2(L + k)\Omega T,$$
$$\nu = \sum_{k=0}^{N-1} \sin^2(L + k)\Omega T, \tag{24}$$

$$\gamma = \sum_{k=0}^{N-1} \sin(k+L)\Omega T \cos(k+L)\Omega T,$$

$$\mu = \sin(\Omega T/2)/(\Omega T/2), \tag{25}$$

$$\alpha = \sum_{k=-M}^{M} \cos^2 k\Omega T, \qquad \beta = \sum_{k=-M}^{M} \sin^2 k\Omega T,$$

$$\lambda = L/\Gamma A_n \mu^2, \tag{26}$$

$$A_n = U^2 T/\eta, \qquad \Gamma = I^2/U^2. \tag{27}$$

A_n is the signal-to-noise ratio and Γ is the ratio of the mean power of the interfering and desired signals at the receiver input.

IV. Signal-to-Interference Ratio Improvement

The signal-to-interference ratio improvement gained by the proposed receiver is defined as the ratio G_{DFB} of the signal-to-interference ratio at its output $(S/I)_{\text{DFB}}$ to the signal-to-interference ratio at the output of the conventional DSSS receiver $(S/I)_{\text{CON}}$, in which the AT filters for NBI rejection are not used.

If the assumption is made that only the single-tone interferer is present and that there is no error propagation, the signal-to-interference improvement G_{DFB} is given by

$$G_{\text{DFB}} = \frac{(S/I)_{\text{DFB}}}{(S/N)_{\text{CON}}} = \frac{\overline{e^2_{\text{CON}}} - L\eta T}{\overline{e^2_{m,\text{DFB}}} - L\eta T}$$

$$= \frac{\lambda + \beta + \chi}{\lambda} - \frac{\gamma^2}{\lambda(\lambda + \alpha + \nu)}. \tag{28}$$

In this expression, $\overline{e^2_{\text{CON}}}$ is the mean-square error at the conventional DSSS receiver output given by (16) when $a_k = b_k = 0$.

Let us assume now the AWGN can be neglected, i.e., that the signal-to-noise ratio $A_n >> 1$. Then, according to (26) $\lambda \cong 0$, and the signal-to-interference improvement from (28) $G_{\text{DFB}} \rightarrow \infty$. In other words, it can be considered that the NBI is completely reduced regardless of the level and the frequency of the carrier.

The improvement G_{DFB} as a function of the product ΩT is shown in Fig. 5 together with the improvement of the receiver from Fig. 1, according to [11], and is defined in the same way as in (28). It is to be noted that the improvement of the proposed receiver with the DFB is much higher than that of the receiver from Fig. 1 for the case when the carrier frequencies of the desired and interfering signals are close, i.e., when $\Omega T < 5°$. This is due to the signal distortion in the AT filter in the upper branch of the receiver from Fig. 1.

Having in mind jamming, it can be concluded that the case $\Omega T \cong 0$ is particularly important. Then, it is evident that the proposed receiver with the DFB offers much higher protection against NBI than the receiver with conventional AT filters.

According to the block diagram from Fig. 4, the transfer function of the filters in the proposed receiver are given by

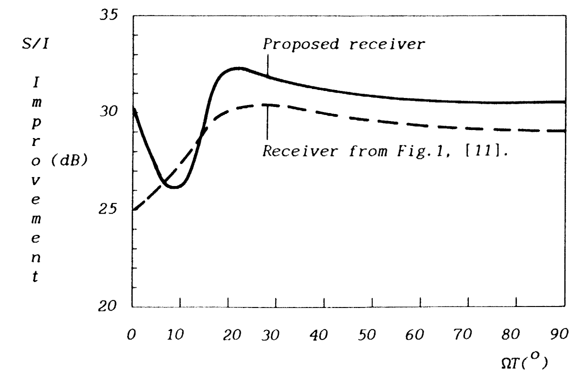


Fig. 5. The signal-to-interference ratio improvement as a function of the product ΩT; $L = 7$, $\Gamma = 20$ dB, $A_n = 12$ dB, $N = 4$, and $M = 4$.

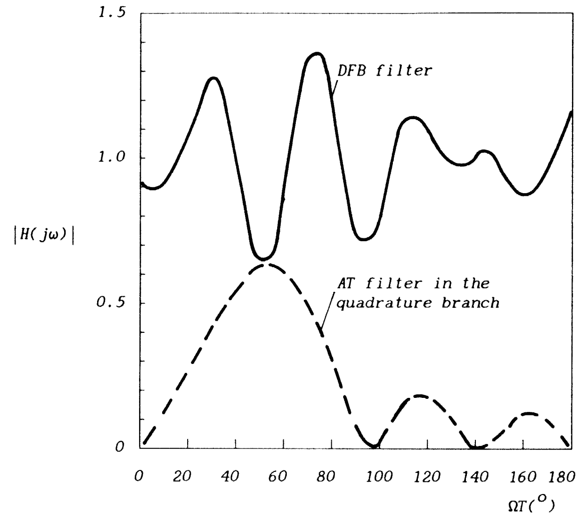


Fig. 6. Amplitude characteristics of the filters in the proposed receiver; $A_n = 12$ dB, $\Gamma = 20$ dB, $\Omega T = 2\pi/L$, $N = 4$, and $M = 4$.

$$H_{\text{DFB}}(j\omega) = 1 - \sum_{k=0}^{N-1} a_k \exp\left[-j(L+k)\omega T\right], \tag{29}$$

$$H_{TF}(j\omega) = \sum_{k=-M}^{M} b_k \exp\left[-j(k+M)\omega T\right], \tag{30}$$

where $H_{\text{DFB}}(j\omega)$ and $H_{TF}(j\omega)$ correspond to the DFB and AT filters, respectively. For the case of the single-tone interferer, their amplitude characteristics are shown in Fig. 6, while the influence of the number of taps of these filters on the signal-to-interference improvement is given in Fig. 7.

V. The Error Probability Analysis

If the binary symbols are equally probable, the conditional error probabilities in the jth bit interval at the receiver output are given by

$$\begin{aligned} P_n(c, \theta, \xi) = {} & 0.5P[UTL + q_I + q_N + q_D \\ & < 0 \mid d_j = 1] + 0.5P[-UTL + q_I + q_N \\ & + q_D > 0 \mid d_j = -1], \end{aligned} \tag{31}$$

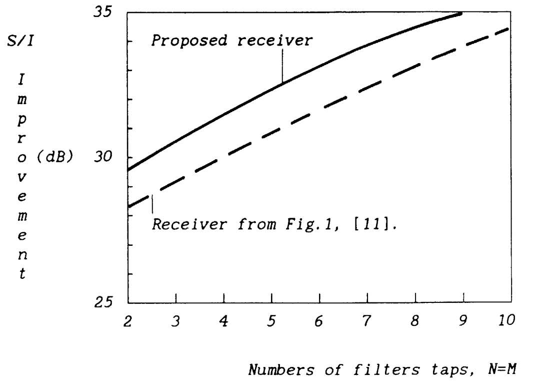


Fig. 7. The signal-to-interference ratio improvement at the output of the proposed receiver as a function of the number of taps; $L = 7$, $A_n = 12$ dB, $\Gamma = 20$ dB, $\Omega T = 2\pi/L$.

where the (c, θ, ξ) denotes the dependence of the probability upon the PRS structure, the phase shift, and the polarity of the decision errors.

In accordance with (11), the index n has the following meaning:

$$\begin{aligned} &n = 0, && \text{no propagation errors;} && q_D = 0, && \text{state } s_o, \\ &n = 1, && \xi_{j-2} = 0, && \xi_{j-1} = \pm 2, && \text{state } s_1, \\ &n = 2, && \xi_{j-2} = \pm 2, && \xi_{j-1} = 0, && \text{state } s_2, \\ &n = 3, && \xi_{j-2} = \pm 2, && \xi_{j-1} = \pm 2, && \text{state } s_3. \end{aligned} \tag{32}$$

However, the computation of the error probability according to (31) is rather difficult. Therefore, the problem has been simplified so that the worst case as a function of θ has been solved. For the worst case of θ, the error probability will be

$$P_n(c, \xi) = 0.25[\text{erfc}\,(Z_+) + \text{erfc}\,(Z_-)] \tag{33}$$

where

$$Z_\pm = \frac{\sqrt{A_n}}{\sqrt{1 + \sum_{k=0}^{N-1} a_k^2 + \sum_{k=-M}^{M} b_k^2}} \cdot \Bigg\{ 1 \mp \sum_{i=0}^{L-1} \frac{c_i}{L} \Bigg[\sum_{k=0}^{i} a_k \xi_{j-1} c_{i-L-k} + \sum_{k=i+1}^{N-1} a_k \xi_{j-2} c_{i-L-k} \Bigg] \pm \frac{\mu\sqrt{\Gamma}}{L} \Bigg(\Bigg[\Bigg(1 - \sum_{k=0}^{N-1} a_k \cos\,(L+k)\Omega T - \sum_{k=-M}^{M} b_k \sin k\Omega T \Bigg)^2 + \Bigg(\sum_{k=0}^{N-1} a_k \sin\,(L+k)\Omega T - \sum_{k=-M}^{M} b_k \cos k\Omega T \Bigg)^2 \Bigg] \Bigg)^{1/2} \cdot \Bigg(\Bigg[\Bigg(\sum_{i=0}^{L-1} c_i \cos i\Omega T \Bigg)^2 + \Bigg(\sum_{i=0}^{L-1} c_i \sin i\Omega T \Bigg)^2 \Bigg] \Bigg)^{1/2} \Bigg\}. \tag{34}$$

The tap weights of the filters are given by the expressions (22) and (23). The average values of the conditional error probabilities at the receiver output, including each case in (32), are given by the expression

$$P_n = E[P_n(c, \xi)] \,\big|\, n = 0, 1, 2, 3. \tag{35}$$

The operator E stands for the averaging over the decision error polarities and different structures of the PRS over the $(j-2)$th, $(j-1)$th, and jth bit.

A. The Effect of Error Propagation on the Error Probability

The nonlinear system in the proposed receiver consisting of the decision circuit and the DFB filter is modeled as an irreducible ergodic Markov's chain of length 2^2 [15]. That means that the number of taps in the DFB filter is less than the processing gain, $N < L$, so that only the errors originated from the two preceding bit intervals can affect the decision process in the interval under consideration.

The state transition diagram of the DFB filter is shown in Fig. 8, assuming $N = 4$, $L = 7$, and that its states are defined by means of (32). The tap content of the tapped delay line is denoted by "0" when there is no error propagation, while the errors in the $(j-2)$th and $(j-1)$th bit intervals with respect to the jth bit interval, which is under consideration, are denoted by ξ_2 and ξ_1, respectively.

As an example, the first block corresponds to the state s_o when there is no propagation through the DFB filter. It means that in all $L = 7$ chips in the frame of a bit of the useful signal, the state in delay line will be "0000," as shown in Fig. 8, in which the symbol "0" stands for the case of no error propagation. Insofar as in the previous decision instant there was an error at the receiver output ξ_1, then the process of the propagation of this error through the delay line in the DFB filter is as in the second block in Fig. 8 (state s_1). It can be noticed that the error ξ_1 after the first tact interval, $i = 0$, appears at the output of the first memory element in the delay line, so that at the end of the last tact interval, $i = 6$, this error in the frame of the considered bit is present in the whole delay line. Also, for the two remaining states, s_2 and s_3, the analogous process is developed.

In the diagram, the state-transition probabilities are separately indicated.

Consequently, the total error probability at the receiver output, taking into account the possible error propagation,

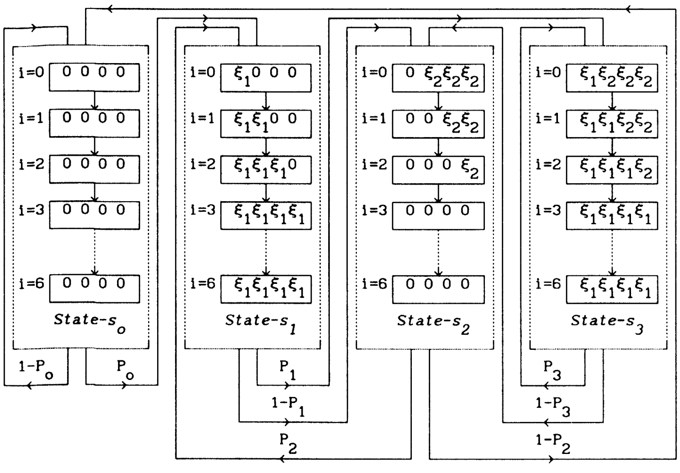


Fig. 8. The state-transition diagram.

is

$$P_e = \sum_{n=0}^{3} P_n P(s_n) \,|\, N < L, \tag{36}$$

where the conditional error probabilities P_n are given by (35). The probabilities $P(s_n)$ represent the state probabilities in the two preceding bit intervals, the combinations of which are given by (32), and they are

$$P(s_n) = \sum_{k=0}^{3} P(s_k)\, P_{k \to n}, \qquad n = 0, 1, 2, 3, \tag{37}$$

where $P_{k \to n}$ is the state transition probability from the kth state to the nth state, as shown in Fig. 8. Adding the equation

$$\sum_{n=0}^{3} P(s_n) = 1, \tag{38}$$

and solving (37) for the $P(s_n)$, the total error probability at the output of the proposed receiver, including the possible error propagation, is given by

$$P_e = \frac{P_0(1 + P_1 - P_3)}{P_0 P_1 + 2P_0(1 - P_3) + (1 - P_2)(1 - P_3)}. \tag{39}$$

VI. Numerical Examples and Discussion of the Results Obtained

The results of computer simulation of the proposed receiver are shown in Figs. 9 and 10. They give the conditional error probabilities according to (35) and the total error probability according to (39), as a function of the ratio of the interfering and desired carrier powers at the receiver input, and as a function of the product of the frequency difference between the desired and interfering carriers and a chip duration, respectively.

The following conclusions can be drawn from these results.

1) In the case of perfect decision, the error probability is practically constant, of the order of 10^{-7} for the assumed value of A_n, and it does not depend upon the frequency difference between the desired and interfering carriers nor upon the relative level of the latter.

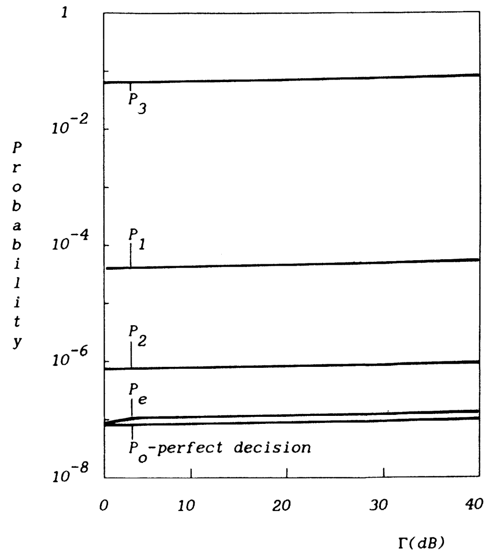


Fig. 9. The conditional error probabilities $P_0 - P_3$ and the total error probability P_e at the output of the proposed receiver, as functions of the ratio Γ of the interfering and desired signal powers; $A_n = 12$ dB, $L = 7$, $N = 4$, $M = 4$, and $\Omega T = 2\pi/L$.

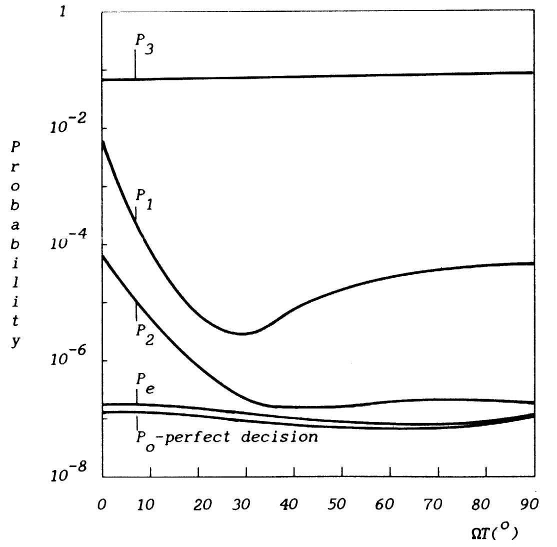


Fig. 10. The conditional error probabilities $P_0 - P_3$ and the total error probability P_e at the output of the proposed receiver, as functions of the product ΩT; $A_n = 12$ dB, $\Gamma = 20$ dB, $L = 7$, $N = 4$, and $M = 4$.

2) The proposed receiver shows much better protection against NBI than the receivers [2], [5], [6], [11], especially for the case when $\Omega T \cong 0$.

3) The role of the DFB filter is dominant for the case $\Omega T = 0$. Then, the tap weights of the AT filter in quadrature branch of the receiver are equal to zero.

4) Having in mind the values of the transition probabilities from Figs. 9 and 10, it follows that P_0, P_1, P_2 $<< P_3$, so that the total error probability is $P_e \cong P_0$. This leads to the conclusion that the propagation of the error in the proposed receiver can be neglected.

VII. Conclusion

In this paper, a new DSSS receiver using combined effects of DFB and AT filters in combatting NBI has been proposed. The receiver performance analysis is carried out on the basis of the error probability at the receiver output and the obtained signal-to-interference ratio improvement in the presence of the single-tone interferer.

Comparing this receiver to those using either the AT filter (or filters) or the DFB filter only, it can be noted that the proposed receiver offers a higher protection against NBI, particularly when the carrier frequencies of the desired and interfering signals are equal or close to each other. Having in mind that this case is particularly important in the presence of the NBI jammer, the results obtained appear to be important.

A weakness of the receiver proposed might be sought in the possibility of the propagation of decision error through the DFB filter. However, if the assumption can be made that the nonlinear system—decision circuit and DFB filter—can be considered as an irreducible Markoff chain, it has been shown that the effects of the error propagation can be neglected.

In conclusion, it can be said that the receiver proposed is one of the most efficient in combatting NBI in which the AT filters are used.

References

[1] R. L. Pickholtz, D. L. Schilling, and L. B. Milstein, "Theory of spread-spectrum communications—A tutorial," *IEEE Trans. Commun.*, May 1982.

[2] D. Shklarsky, P. Das, and L. B. Milstein, "Adaptive narrowband interference suppression," in *Proc. Nat. Telecommun. Conf.*, 1979, pp. 15.2.1-3.

[3] C. K. Pauw, "Application of decision feedback equaliser in an environment with narrowband interfering signals," *Electron Lett.*, no. 20, Sept. 1980.

[4] J. W. Ketchum and J. G. Proakis, "Adaptive algorithms for estimating and suppressing narrowband interference in PN spread-spectrum systems," *IEEE Trans. Commun.*, May 1982.

[5] L. Li and L. B. Milstein, "Rejection of narrowband interference in PN spread-spectrum systems using transversal filters," *IEEE Trans. Commun.*, May 1982.

[6] R. A. Iltis and L. B. Milstein, "Performance analysis of narrowband interference rejection in DS spread-spectrum systems," *IEEE Trans. Commun.*, Nov. 1984.

[7] J. W. Ketchum, "Decision feedback techniques for interference cancellation in PN spread spectrum communication systems," in *IEEE Proc. Military Commun. Conf.*, 1984, pp. 39.5.1-5.

[8] G. J. Saulnier, P. Das, and L. B. Milstein, "An adaptive digital suppression filter for direct sequence spread-spectrum communications," *IEEE J. Select. Areas Commun.*, Sept. 1985.

[9] E. Masry and L. B. Milstein, "Performance of direct spread-spectrum receiver employing interference-suppression filters under a worst case jamming condition," *IEEE Trans. Commun.*, Jan. 1986.

[10] F. Takawira and L. B. Milstein, "Narrowband interference rejection in PN spread-spectrum communication systems using decision feedback filters," in *IEEE Proc. Military Commun. Conf.*, 1986, pp. 20.4.1-5.

[11] Z. D. Stojanović, M. L. Dukić, and I. S. Stojanović, "A new method for the narrow-band interference rejection in the direct sequence spread-spectrum systems using transversal filters," presented at the IEEE MELECON Conf., Rome, Mar. 1987.

[12] ——, "Analysis of the methods proposed for the narrow-band interference rejection in the direct sequence spread-spectrum systems using transversal filters," presented at the Int. Symp. Spread-Spectrum Techniques, Belgrade, May 1988.

[13] L. B. Milstein, "Interference rejection techniques in spread-spectrum communications," *Proc. IEEE*, vol. 76, pp. 657-671, June 1988.

[14] A. Papoulis, *Probability, Random Variables and Stochastic Processes.* New York: McGraw-Hill, 1965.

[15] L. M. Li and L. B. Milstein, "Rejection of CW interference in QPSK systems using decision-feedback filters," *IEEE Trans. Commun.*, vol. COM-31, pp. 473-483, Apr. 1983.

[16] M. L. Dukić, Z. D. Stojanović, and I. S. Stojanović, "A new direct sequence spread-spectrum receiver using decision feedback and transversal filters for rejection of the narrow-band interference and errors caused by signal distortion," in *Proc. IEEE MELECON'89*, Lisboa, Apr. 1989, pp. 395-398.

Miroslav L. Dukić (M'85), for a photograph and biography, see this issue, p. 749.

Zorka D. Stojanović (M'85) was born in Belgrade, Yugoslavia. She received the Dipl.Eng. degree in communication and electronics engineering in 1958 and the M.Sc. degree in 1968, both from the Faculty of Electrical Engineering University of Belgrade, Yugoslavaia. She received the Doctor of Science degree from the Faculty of Electrical Engineering, University of Zagreb, Yugoslavia.

Since 1958 she has been with the Department of Communications of the Faculty of Electrical Engineering, University of Belgrade, where, as a Professor, she taught several courses in communication theory, communication system design, and stochastic theory applied in telecommunications. She has published a number of papers and is the author of three books in the field of communications. She also participated as a Project Engineer in several radio relay system projects and in TV network design Yugoslavia.

Dr. Stojanović is a member of the Yugoslav Committee on Electronics, Telecommunications, Automation and Nuclear Engineering.

Ilija S. Stojanović (M'85–SM'89) was born in Otočac, Lika, Yugoslavia. He received the Dipl.Eng. degree in 1951 and the Doctor of Science degree in 1957, both from the Faculty of Electrical Engineering, University of Belgrade, Yugoslavia.

Since 1951 he has been with the Faculty of Electrical Engineering, University of Belgrade, Yugoslavia. He served for 20 years as Head of the Department of Communications. He has taught a number of courses in communication theory and communication system design. He is the author of a number of papers and has written five books in the field of communications. He was also a leading engineer in many radio-relay projects, TV networks, and other communication systems in Yugoslavia. He participated in many ITU conferences and was the Chairman of the First (1985) and the Second (1988) Sessions of the World Administrative Conference on the use of the geostationary-satellite orbit and the planning of the space services using it; Chairman of the XVIth CCIR Plenary Assembly (Dubrovnik, 1986); and the Chairman of the Panel of Experts on the long-term future of the IFRB. He also participated in the work of the Committee on the Peaceful Users of Outer Space.

Dr. Stojanović is a member of the Serbian Academy of Science and Arts.

Presence Detection of Binary-Phase-Shift-Keyed and Direct-Sequence Spread-Spectrum Signals Using a Prefilter-Delay-and-Multiply Device

JOHN F. KUEHLS, MEMBER, IEEE, AND EVAGGELOS GERANIOTIS, SENIOR MEMBER, IEEE

Abstract—The specific problem of detecting the presence of either binary-phase-shift-keyed (BPSK) signals or BPSK direct-sequence spread-spectrum (DS/SS) signals with a prefilter-delay-and-multiply (PFDM) device is considered. Using stationary process theory and Fourier analysis, the optimum PFDM structures for signal presence detection of BPSK signals with known bit rates and carrier frequencies, and BPSK DS/SS signals with known chip rates and carriers, in additive colored Gaussian noise are derived. The structures are optimum in the sense that they maximize the spectral signal-to-noise ratio (SNR) of an output periodic waveform which has fundamental frequency equal to the bit or chip rate of the signal. Two of the optimum structures that are derived and analyzed herein are the optimal prefilter-square device, and the optimal PFDM with delay set to one-half of the signal's bit or chip duration T. The latter structure has not been reported as of yet, and it is significant because it specifies the optimum prefilter for a $T/2$ delay-and-multiply detector.

Exploiting a general expression for the output spectral SNR that was needed to derive the optimal structures for known bit or chip rates, a robust structure for the presence detection of BPSK or BPSK DS/SS signals with unknown bit or chip rates is also found. Finally, the spectral SNR is related to true performance measures when probabilities of detection and false alarm for both known and unknown bit or chip rates are derived, and the tradeoffs between SNR, length of observation interval, and time or bandwidth mismatch are studied. Additionally, the detection probability for an optimal PFDM is compared to that for a standard ad hoc configuration.

I. Introduction

HOW does one detect the presence (presence detection differs from what is usually called detection in that a presence detector only seeks to determine whether or not a signal is on the air, whereas the usual detector seeks to determine the transmitted message) of a signal in Gaussian noise? Standard techniques are available, and for one class of signals—the known signal—they are well established. A known signal may be detected by either passing the received waveform through a filter matched to the signal, or by correlating the received waveform with a reference that is proportional to the signal. These approaches are equivalent mathematically and optimum (follow from the likelihood ratio test) when the additive Gaussian noise is white (AWGN) [1]. When the signal is not known, the matched filter/correlator approach is not particularly useful because it requires knowledge of the signal. With regard to signal presence detection, any digitally modulated signal cannot be considered known unless the sequence which modulates the signal is known. Matched filters/correlators cannot therefore be used to detect the presence of a binary-phase-shift-keyed (BPSK) signal unless the underlying message sequence, which shifts the phase, is known. BPSK direct-sequence spread-spectrum signals are BPSK signals with a pseudonoise (PN) spreading sequence overlaying the message sequence; hence, matched filters/correlators cannot be used to presence detect BPSK DS/SS signals unless the PN spreading sequence is known.

Because of the random nature of the sequence which shifts their phase, BPSK signals are not periodic; hence, they have continuous Fourier spectra (DS/SS signals have a *pseudocontinuous* spectrum—spectral lines separated in frequency by the reciprocal of the duration of the PN sequence). This, consequently, makes them difficult to detect using a conventional analog spectrum analyzer or Fast Fourier Transform (FFT) [2], which are the optimal devices for detecting unknown signals that have discrete spectra in AWGN [1]. It is known, however, that when certain nonlinear operations are applied to BPSK signals, discrete spectral components arise. These components are then often detectable using spectrum analyzer/FFT techniques. A nonlinear operation can thus serve to map BPSK signal presence detection from the detection of an unknown continuous-spectrum signal to the detection of a discrete-spectrum signal (with known or unknown spectral line frequencies, depending upon the extent of *a priori* knowledge of the BPSK signal's parameters).

In practical systems currently used for the detection of BPSK signals, nonlinear operations are employed. In general, the nonlinearity is quadratic, as is the case with the delay-and-multiply detector, which will be discussed in this work. Another feature that characterizes these nonlinear transformations is that they are typically ad hoc. Little has been done with regards to determining optimal transformations. The work of Gardner [3], [4], however, has provided some insight. He has applied the theory of *cyclostationary* processes [5], exploiting the *spectral correlation* [6] properties of PSK signals to determine *locally*

Manuscript received March 7, 1989; revised October 19, 1989.

J. F. Kuehls is with the Department of Defense, Fort Meade, MD, and George Washington University, Washington, DC.

E. Geraniotis is with the Department of Electrical Engineering and Systems Research Center, University of Maryland, College Park, MD 20742.

IEEE Log Number 9035817.

optimum (LO) (low signal-to-noise ratio (SNR), long observation time) detectors. In particular, he has established that, under the LO conditions for BPSK, the optimized quadratic continuous-to-discrete-spectrum-transformation detector can be realized by a filter followed by a squarer, and that this configuration has a relation to the optimum detector, which is based upon the likelihood ratio.

In this paper, we investigate the detection of BPSK signals using a detector that employs a particular quadratic transformation known as a prefilter-delay-and-multiply (PFDM). The PFDM will be configured to generate discrete spectral components at the signal's bit frequency and the harmonics of the signal's bit frequency. (It is understood that for DS/SS, we are dealing with the *chip* frequency.) On the basis of a very practical figure of merit, we will derive the optimum form of the PFDM for presence detection of signals with *known bit rates*. The tools used in the derivation will be Fourier analysis and stationary process theory.

It will be shown that there is a countably infinite number of possible realizations for the optimum PFDM. One of these structures—the optimized prefilter-squarer (mentioned above)—has been previously reported by Krasner [7] and Gardner [3]. We will establish, however, that the prefilter-squarer is not as robust with respect to small uncertainties in the bit rate as another of the optimum realizations, the optimized prefilter-delay-and-multiplier with delay equal to one-half of the signal's bit duration.

With regard to the detection of BPSK signals with *unknown bit rates*, this problem has not received much attention in the literature. Here, however, we will find a very simple robust quadratic structure, the rectangular prefilter-delay-and-multiply with delay equal to the reciprocal of the prefilter bandwidth.

Our investigation concludes with an analysis of the detection of the discrete-spectrum signals at the output of the quadratic devices. For three practical cases of interest (known bit rates with known and unknown bit timing, and unknown bit rates), we arrive at expressions for signal presence false alarm and presence probabilities P_F and P_D. The central limit theorem is used to approximate the distribution of the detection statistics. The long observation times required for the low input SNR's that are of practical interest make these asymptotic results valid. We then use the P_F and P_D expressions to study the impact of the tradeoff between SNR, length of observation interval, and time or bandwidth mismatch on the detection and false alarm probabilities. No study of this type has ever appeared in the literature. This paper provides such a study and sheds light on the true performance of PFDM detectors for BPSK signals.

We emphasize here what this paper provides, an analysis of how a specific detector (the PFDM) will work on a specific signal (BPSK). Although our analysis is fairly complete, we do not address the more general problems of arbitrary detection structures for arbitrary PSK signals and the relation of continuous-to-discrete-spectrum-transformation detectors to the likelihood ratio. These problems have been considered by Gardner [3] and Krasner [7].

The paper is organized in the following manner. In Section II, we describe the model under which the analysis was performed. The PFDM, noise process, signaling format, and the figure of merit are all discussed. In Section III, the noise and signal components of the figure of merit are derived by using stationary process theory (for the noise) and Fourier analysis (for the signal). After maximizing the figure of merit, the criteria for realizable optimum detection structures are established; this is done in Section IV. In Section V, we use the figure of merit to compare the performance of the optimal systems against that of some suboptimal ad hoc approaches. It is here that the simple robust structure for presence detection of BPSK signals with unknown bit rates is reported. Section VI contains the derivations for signal presence detection and false alarm probabilities, as well as numerical results illustrating the performance of the various systems considered. Finally, in Section VII, we summarize our work and make concluding remarks.

II. Model Description

A block diagram of the system considered in this analysis is shown in Fig. 1. The system consists of a filter with real impulse response $h(t)$ followed by a delay-and-multiply circuit with delay d. We choose the PFDM system for two reasons. First of all, the optimum general quadratic structure can be realized with a PFDM since the optimized prefilter-square is a special case corresponding to $d = 0$. Second, most of the ad hoc detectors employ the PFDM structure. It is therefore of interest to optimize that structure.

The noise process $\boldsymbol{n}(t)$ is, by assumption, zero-mean wide-sense stationary Gaussian noise with power spectral density $S_{nn}(f)$ and autocorrelation $R_{nn}(\tau) = E[\boldsymbol{n}(t)\boldsymbol{n}(t + \tau)]$, where $E[\cdot]$ is the expectation operator. A special case to be considered later is the *white* case; here

$$S_{nn}(f) = \frac{N_0}{2}$$

and

$$R_{nn}(\tau) = \frac{N_0}{2}\delta(\tau)$$

where $\delta(\cdot)$ is the Dirac delta function.

The signal $s(t)$ is BPSK or BPSK DS/SS. We assume in this paper that the carrier frequency is known so that we can use the baseband representation

$$s(t) = \sum_{n=-\infty}^{\infty} \boldsymbol{a}_n p(t - \Delta - nT) \tag{1}$$

where $\{\boldsymbol{a}_n\}$ is a sequence of binary random variables with states $\pm \boldsymbol{a}$, T is the bit (or chip, for DS/SS) duration, Δ is a constant on $(0, T)$, and $p(\cdot)$ is a baseband pulse. Both BPSK and DS/SS signals can be written as in (1). For

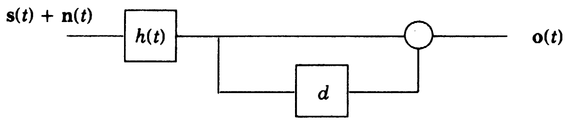


Fig. 1. Prefilter-delay-and-multiply device.

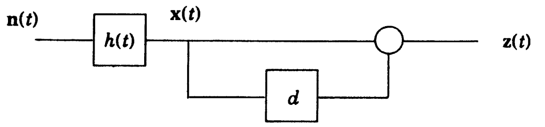


Fig. 2. Prefilter-delay-and-multiply driven by Gaussian noise.

BPSK signals, $\{a_n\}$ is the random message sequence, while for DS/SS, a_n is the product of the random message bit and the random chip.

The figure of merit will be defined now. Remember that our detector is to be configured so as to generate tones at the signal's bit frequency f_b ($= 1/T$) and the harmonics of the signal's bit frequency. Therefore, our figure of merit is the ratio of the following two quantities: 1) the amplitude squared (intensity) of the kth coefficient of the Fourier series expansion of the prefilter-delay-and-multiply output $o(t)$ about the fundamental frequency f_b, and 2) the amplitude of the power spectral density of $o(t)$ in the immediate vicinity of the frequency kf_b. This figure can be interpreted as a spectral SNR since it is a measure of the relative strengths of a composite (discrete plus continuous) spectrum, the discrete spectrum being the "signal" and the continuous spectrum the "noise." As shown in the sequel, it directly effects the detectability of the output harmonic signal in the output process $o(t)$. Spectral bin SNR has also been proposed and used as a figure of merit by Gardner [8] in his work on synchronizers.

III. Derivation of the Figure of Merit

The output $o(t)$ of the PFDM contains three terms: a signal · signal term, a noise · noise term, and a signal · noise cross-product term. The signal · signal term contains a deterministic component, which is periodic, and a random (or pseudorandom, for shift register generated (SRG) sequences) component. We can separate the four output terms into those which have a discrete spectrum, and those which have a continuous (or pseudocontinuous) spectrum. The periodic component of the signal · signal term has a discrete spectrum. All other terms have continuous (or pseudocontinuous) spectra. To derive the numerator (signal component) of the figure of merit, we only need consider the periodic part of the signal · signal term. To derive the denominator term (noise component), we must consider each of the other three terms. However, in the derivation that follows, we assume that the noise · noise term dominates so that we can safely neglect the other two terms. This assumption implies a low input signal-to-noise ratio. (Imbeaux [9] has expressions which may be used if one wants to include the contributions of the signal · noise cross product and signal · signal self-noise terms.)

There are two parameters which completely describe a PFDM, the delay d, and the filter response $h(t)$. So that the optimal PFDM can be specified, the expressions for the signal and noise components of the figure of merit are derived in terms of the delay d and the filter response $h(t)$ [or $H(f)$, the Fourier transform of $h(t)$].

A. Noise Component

Given Fig. 2, our problem is to determine the power spectrum of $z(t)(S_{zz}(f))$ in terms of d and $h(t)$. It turns out that complete derivation of $S_{zz}(f)$ is fairly straightforward, so we table it to Appendix A. Here, we obtain $S_{zz}(f)$ by applying an underived formula in [10] for the power spectrum of the delay-and-multiply product of a Gaussian process. Note that, because $n(t)$ is Gaussian, $x(t)$ is also Gaussian, so that $z(t) = x(t)\,x(t-d)$ is the delay-and-multiply product of a Gaussian process. From [10, p. 237], we have then

$$S_{zz}(f) = R_{xx}^2(d)\,\delta(f) + \int_{-\infty}^{\infty} [1 + \cos(4\pi\lambda d)] \cdot S_{xx}\left(\lambda + \frac{f}{2}\right) S_{xx}\left(\lambda - \frac{f}{2}\right) d\lambda. \qquad (2)$$

The power spectrum of the process $x(t)$ can be written in terms of the filter response via

$$S_{xx}(f) = |H(f)|^2 S_{nn}(f)$$

so that (2) becomes

$$S_{zz}(f) = R_{xx}^2(d)\,\delta(f) + 2\int_{-\infty}^{\infty} \cos^2(2\pi\lambda d) \cdot \left|H\left(\lambda + \frac{f}{2}\right)\right|^2 \left|H\left(\lambda - \frac{f}{2}\right)\right|^2 \cdot S_{nn}\left(\lambda + \frac{f}{2}\right) S_{nn}\left(\lambda - \frac{f}{2}\right) d\lambda \qquad (3)$$

after applying a simple trigonometric identity. Thus, we have an expression which exhibits the power spectrum of $z(t)$ in terms of the filter response $H(\cdot)$ and the delay parameter d. The noise component of the figure of merit is obtained by evaluating (3) at $f = kf_b$.

B. Signal Component

In this subsection, we derive the amplitudes of the discrete spectral components at frequencies kf_b in the PFDM output waveform. Fig. 3 depicts the model to be used. The filtered signal $c(t)$ is expressed as

$$c(t) = \sum_{n=-\infty}^{\infty} a_{ln} q(t - \Delta - nT) \qquad (4)$$

where

$$q(t) = \int_{-\infty}^{\infty} p(\zeta)\, h(t - \zeta)\, d\zeta$$

is the filtered pulse. The output $b(t)$ is the product of $c(t)$ and $c(t-d)$. As will be shown below, regardless of the properties of the binary random sequence $\{a_n\}$, the wave-

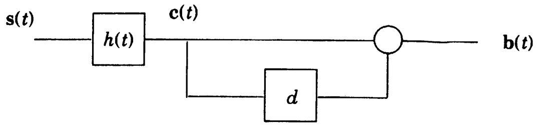


Fig. 3. Prefilter-delay-and-multiply driven by baseband binary-modulated pulse train.

form $\boldsymbol{b}(t)$ contains a nonrandom component that is periodic with fundamental frequency $f_b = 1/T$. (For the special case of $\{\boldsymbol{a}_n\}$, an SRG sequence [11] shows the spectrum of $\boldsymbol{b}(t)$.)

From (4) we see that

$$\begin{aligned}\boldsymbol{b}(t) &= \sum_{n=-\infty}^{\infty} \boldsymbol{a}_n q(t - \Delta - nT)\\ &\quad\cdot \sum_{m=-\infty}^{\infty} \boldsymbol{a}_m q(t - d - \Delta - mT)\\ &= \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} \boldsymbol{a}_n\boldsymbol{a}_m q(t - \Delta - nT)\\ &\quad\cdot q(t - d - \Delta - mT)\\ &= \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}{}' \boldsymbol{a}_n\boldsymbol{a}_m q(t - \Delta - nT)\\ &\quad\cdot q(t - d - \Delta\, mT)\\ &\quad+ \sum_{n=-\infty}^{\infty} \boldsymbol{a}_n^2 q(t - \Delta - nT)\\ &\quad\cdot q(t - d - \Delta - nT)\\ &\equiv \boldsymbol{b}_r(t) + b_p(t), \qquad (5)\end{aligned}$$

where the prime in the summation indicates exclusion of the $m = n$ term, and $\boldsymbol{b}_p(t)$ is nonrandom because $\boldsymbol{a}_n^2 = \boldsymbol{a}^2$, for each n.

We now write

$$b_p(t) = \sum_{k=-\infty}^{\infty} b_k e^{jk2\pi f_b t}.$$

Orthogonality and (5) give us the amplitude of the coefficient b_k:

$$\begin{aligned}b_k &= \frac{1}{T}\int_{-T/2}^{T/2} \sum_{n=-\infty}^{\infty} a^2 q(t - \Delta - nT)\\ &\quad\cdot q(t - d - \Delta - nT) e^{-jk2\pi f_b t}\, dt. \qquad (6)\end{aligned}$$

The filtered pulse $q(t)$ can be expressed in terms of its Fourier transform

$$q(t) = \int_{-\infty}^{\infty} Q(f) e^{j2\pi ft}\, df.$$

Substituting this expression into (5) and isolating the terms that depend upon n leads to

$$\begin{aligned}b_k &= \frac{a^2}{T}\int_{-T/2}^{T/2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left[j2\pi(f_1 t - f_1\Delta + f_2 t\right.\\ &\quad\left. - f_2 d - f_2\Delta - kf_b t)\right] Q(f_1)\, Q(f_2)\\ &\quad\cdot \sum_{n=-\infty}^{\infty} e^{-j2\pi nT(f_1 + f_2)}\, df_1\, df_2\, dt.\end{aligned}$$

Due to the periodicity of the complex exponential function, the infinite sum of phasors can be expressed as an infinite sum of impulses [12] giving

$$\begin{aligned}b_k &= \frac{a^2}{T^2}\int_{-T/2}^{T/2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left[j2\pi(f_1 t - f_1\Delta + f_2 t\right.\\ &\quad\left. - f_2 d - f_2\Delta - kf_b t)\right] Q(f_1)\, Q(f_2)\\ &\quad\cdot \sum_{i=-\infty}^{\infty} \delta(f_1 + f_2 - if_b)\, df_1\, df_2\, dt\\ &= \frac{a^2}{T^2}\int_{-T/2}^{T/2}\int_{-\infty}^{\infty}\sum_{i=-\infty}^{\infty} \exp\left[j2\pi(if_b - f_2)t\right.\\ &\quad\left. - (if_b - f_2)\Delta + f_2 t - f_2\Delta - f_2 d - kf_b t)\right]\\ &\quad\cdot Q(-f_2 + if_b)\, Q(f_2)\, df_2\, dt\\ &= \frac{a^2}{T^2}\int_{-\infty}^{\infty} e^{-j2\pi fd} Q(f) \sum_{i=-\infty}^{\infty} Q(-f + if_b) e^{-j2\pi if_b\Delta}\\ &\quad\cdot \int_{-T/2}^{T/2} e^{j2\pi((i-k)f_b t)}\, dt\, df\\ &= \frac{a^2}{T}\int_{-\infty}^{\infty} e^{-j2\pi(fd + kf_b\Delta)} Q(f)\, Q(kf_b - f)\, df,\end{aligned}$$

where we integrate with respect to f_1 to eliminate the impulses, isolate the terms that depend upon i, and integrate with respect to t, remembering that $f_b = 1/T$ and i is an integer.

Assuming that the filtered pulse has an even spectrum, and after making a change of variables, we are left with

$$\begin{aligned}b_k &= \frac{a^2}{T} e^{-j\pi f_b(d + 2k\Delta)} \int_{-\infty}^{\infty} \cos(2\pi fd) Q\left(f + k\frac{f_b}{2}\right)\\ &\quad\cdot Q\left(f - k\frac{f_b}{2}\right) df.\end{aligned}$$

Expressing $Q(\cdot)$ in terms of the filter response,

$$\begin{aligned}b_k &= \frac{a^2}{T} e^{-j\pi f_b(d + 2k\Delta)} \int_{-\infty}^{\infty} \cos(2\pi fd) H\left(f + k\frac{f_b}{2}\right)\\ &\quad\cdot P\left(f + k\frac{f_b}{2}\right) H\left(f - k\frac{f_b}{2}\right) P\left(f - k\frac{f_b}{2}\right) dt\end{aligned} \qquad (7)$$

where $P(\cdot)$ is the Fourier transform of the unfiltered pulse. The intensity of the kth coefficient, written in terms of the delay parameter d and the filter response $H(\cdot)$, is therefore

$$\begin{aligned}|b_k|^2 &= \frac{a^4}{T^2}\left|\int_{-\infty}^{\infty} \cos(2\pi fd) H\left(f + k\frac{f_b}{2}\right) P\left(f + k\frac{f_b}{2}\right)\right.\\ &\quad\left.\cdot H\left(f - k\frac{f_b}{2}\right) P\left(f - k\frac{f_b}{2}\right) df\right|^2. \qquad (8)\end{aligned}$$

The above expression can also be derived by applying a general result [8, eq. (15)] for the strength of regenerated spectral lines at the output of arbitrary quadratic devices.

The figure of merit is obtained by forming the ratio $|b_k|^2/S_{zz}(kf_b)$. From (3) and (8) we have

$$\frac{|b_k|^2}{S_{zz}(kf_b)} = \frac{\frac{a^4}{T^2}\left|\int_{-\infty}^{\infty} \cos(2\pi f d) H\left(f + k\frac{f_b}{2}\right) P\left(f + k\frac{f_b}{2}\right) H\left(f - k\frac{f_b}{2}\right) P\left(f - k\frac{f_b}{2}\right) df\right|^2}{2\int_{-\infty}^{\infty} \cos^2(2\pi f d)\left|H\left(f + k\frac{f_b}{2}\right)\right|^2 S_{nn}\left(f + k\frac{f_b}{2}\right)\left|H\left(f - k\frac{f_b}{2}\right)\right|^2 S_{nn}\left(f - k\frac{f_b}{2}\right) df}. \tag{9}$$

Note the similar structure of the definite integrals in the numerator and denominator.

IV. Criteria for Optimization

To specify the form of the optimum PFDM, we maximize (9). This is done with the Cauchy-Schwarz inequality, which establishes an upper bound on $|b_k|^2/S_{zz}(kf_b)$ as well as the condition required to achieve the upper bound. These are given in (10) and (11), respectively:

$$\frac{|b_k|^2}{S_{zz}(kf_b)} \le \frac{a^4}{2T^2}\int_{-\infty}^{\infty} \frac{\left|P\left(f + k\frac{f_b}{2}\right) P\left(f - k\frac{f_b}{2}\right)\right|^2}{S_{nn}\left(f + k\frac{f_b}{2}\right) S_{nn}\left(f - k\frac{f_b}{2}\right)} df \tag{10}$$

with equality when

$$H\left(f + k\frac{f_b}{2}\right) H\left(f - k\frac{f_b}{2}\right) \propto \frac{P\left(f + k\frac{f_b}{2}\right)^* P\left(f - k\frac{f_b}{2}\right)^*}{\cos(2\pi f d) S_{nn}\left(f + k\frac{f_b}{2}\right) S_{nn}\left(f - k\frac{f_b}{2}\right)} \tag{11}$$

where $\propto$ denotes proportionality and * denotes complex conjugation. The maximum spectral SNR (10) that we have obtained here, assuming a PFDM structure, is the same as that obtained in [8] where an arbitrary quadratic structure was assumed. This confirms that the optimum general quadratic structure can be realized with a PFDM.

Equation (10) gives the maximum value of the ratio of the kth coefficient's intensity to the noise power spectrum, at frequency $f = kf_b$, that is attainable with any PFDM. Equation (11) specifies the condition which must be satisfied in order to attain the maximum. Notice that, because the maximum is achieved whenever (11) is satisfied, the optimum PFDM is not unique. Also, note from (11) that the same PFDM maximizes the figure of merit for each value of k.

The criterion for optimization (11) is a nonlinear functional equation. We have been unable to obtain a general solution for the filter response $H(\cdot)$. We can, however, treat a few specific cases by fixing the delay d. When $d = 0$, the optimum filter is given by

$$H(f) \propto \frac{P(f)^*}{S_{nn}(f)}. \tag{12}$$

(This is the prefilter-square structure that has been derived by Gardner [3].) Some other cases for the delay d can be treated by noting that

$$\cos(2\pi f d) = 2\cos(\pi f d + m\pi/4)\cos(\pi f d - m\pi/4), \quad m = \pm 1, \pm 3, \pm 5, \cdots.$$

For $d = mT/2k$, $m = \pm 1, \pm 3, \pm 5, \cdots$, we have then

$$H(f) \propto \frac{P(f)^*}{\cos(m\pi fT/2) S_{nn}(f)}. \tag{13}$$

We see that the optimum prefilter-delay-and-multiply structure can be realized for delays equal to zero and for delays not equal to zero. The zero-delay version maximizes the spectral SNR for each harmonic of $b_p(t)$. The nonzero-delay version, however, maximizes the spectral SNR of only the kth harmonic where k is obtained from $d = mT/2k$. The example system that we will consider in the sequel is the $m = 1$, $d = T/2$ PFDM, which maximizes the spectral SNR of the fundamental component. Finally, note that (12) is really only a special case of (13), provided we allow $m = 0$. Equation (13), with $m = 0, \pm 1, \pm 3, \pm 5, \cdots$, is therefore a general expression which specifies the form for all optimal PFDM structures.

The noise power spectrum $S_{nn}(f)$ in the denominator of (13) can be interpreted as a factor which whitens the spectrum. This is intuitively pleasing because classic optimal receivers for colored Gaussian noise channels always incorporate whitening filters [1].

In the special case of white Gaussian noise, (13) becomes

$$H(f) \propto \frac{P(f)^*}{\cos(m\pi fT/2)}, \quad (d = mT/2k, \; m = 0, \pm 1, \pm 3, \pm 5, \cdots). \tag{14}$$

The $m = 0$ version of (14) states that in the case of white noise, one of the optimum quadratic structures is simply a filter that is matched to the signal's pulse, followed by a squarer, a result that has also been reported by Krasner [7] and Gardner [3]. The $m \neq 0$ versions of (14) give the other possible realizations for the optimum PFDM when the noise is white.

Notice that, in the derivation of the optimum PFDM detectors, we assumed knowledge of two signal parameters, the bit duration T and, as implied by the baseband analysis, the carrier frequency. In practice, these parameters may in fact be unknown, thus limiting the utility of the optimum PFDM detectors. In the following, we discuss a reasonable PFDM detection scheme for the case of unknown bit durations. Unknown carrier frequencies, however, are not considered, although it is readily apparent that a possible such scheme would involve adjusting the PFDM's prefilter center frequency.

One signal parameter that was not assumed to be known *a priori* is the bit timing Δ. This did not influence our results (12) and (13) because Δ effects only the phase of the coefficient b_k [cf. (7)], whereas our spectral SNR figure of merit was based upon $|b_k|^2$, which is phase independent. The bit timing, however, will be important when we discuss detection of the discrete spectral components in Section VI.

V. Optimum versus Ad Hoc Quadratic Schemes: Figure of Merit Comparison

In Section III, we derived an expression for the spectral SNR at the output of a PFDM when a baseband BPSK signal plus Gaussian noise is input. The generality of the derivation gives (9) considerable utility. It can be used to compute figures of merit for any BPSK signal (specifically, we consider those with known and unknown bit rates) and any detector in the PFDM class. Applications include: 1) assessing the performance of currently used or proposed detection schemes, and 2) determining the robustness of different systems to variations in their parameters, such as filter bandwidth and delay time d. Our analysis here is for rectangularly pulsed signals in AWGN and PFDM's which employ both optimal and rectangular prefilters.

The spectral SNR at the output of an optimal PFDM is given by (10). This is the maximum value that the figure of merit can assume. For white Gaussian noise, it becomes

$$\max\left\{\frac{|b_k|^2}{S_{zz}(kf_b)}\right\} = \frac{2a^4}{T^2N_0^2}\int_{-\infty}^{\infty}\left|P\left(f+k\frac{f_b}{2}\right)P\left(f-k\frac{f_b}{2}\right)\right|^2 df. \tag{15}$$

The above integral is most readily evaluated in the time domain. We apply Parseval's identity and the Fourier convolution theorem as follows:

$$\begin{aligned}\int_{-\infty}^{\infty}\left|P\left(f+k\frac{f_b}{2}\right)\right|^2\left|P\left(f-k\frac{f_b}{2}\right)\right|^2 df &= \int_{-\infty}^{\infty} p_2(t)e^{j2\pi k(f_b/2)t}\left(p_2(t)e^{-j2\pi k(f_b/2)t}\right)^* dt\\ &= \int_{-\infty}^{\infty} p_2^2(t)e^{j2\pi kf_bt}\,dt\end{aligned} \tag{16}$$

where $p_2^2(\cdot)$ is the inverse Fourier transform of $|P(\cdot)|^2$. Since $p_2^2(\cdot)$ may be computed by convolving $p(\cdot)$ with itself, and since the pulses under consideration are time-limited to $(-T/2, T/2)$, $p_2^2(t)$ is time-limited to $(-T, T)$. Furthermore, because $|P(\cdot)|^2$ is real, $p_2^2(\cdot)$ possesses even symmetry. Equation (16) thus reduces to

$$\int_{-\infty}^{\infty}\left|P\left(f+k\frac{f_b}{2}\right)\right|^2\left|P\left(f-k\frac{f_b}{2}\right)\right|^2 df = 2\int_0^T p_2^2(t)\cos(2\pi kf_bt)\,dt. \tag{17}$$

A rectangular pulse can be expressed as $p(t) = \mathrm{rec}(2t/T) \equiv p_r(t)$, where $\mathrm{rec}(\cdot) = 1$ if $|\cdot| \leq 1$ and zero otherwise. The Fourier transform of the pulse is

$$P_r(f) = T\,\mathrm{sinc}\,(\pi fT) \tag{18}$$

where $\mathrm{sinc}\,(\cdot) = \sin\,(\cdot)/(\cdot)$. Noting that $p_{r2}^2(t) = T\,\mathrm{tri}(t/T)$, where $\mathrm{tri}(\cdot) = 1 - |\cdot|$ for $|\cdot| \leq 1$ and zero otherwise, by (15), the maximum spectral SNR at a PFDM output is

$$\begin{aligned}\frac{|b_k|^2}{S_{zz}(kf_b)} &= \frac{2a^4}{T^2N_0^2}\int_0^T (t^2 - 2tT + T^2)\cos(2\pi kf_bt)\,dt\\ &= \frac{2a^4T}{k^2\pi^2N_0^2}, \qquad k = 1, 2, 3, \cdots.\end{aligned} \tag{19}$$

Making the definition $\beta \equiv a^2T/N_0\ (= E_b/N_0)$, the normalized maximum $|b_k|^2/S_{zz}(kf_b)$ can be expressed as

$$\max\left\{T\frac{|b_k|^2}{S_{zz}(kf_b)}\right\} = \frac{2}{k^2\pi^2}\beta^2, \qquad k = 1, 2, 3, \cdots \tag{20}$$

which, we parenthetically note, is proportional to the square of the input bit energy-to-white noise ratio. Also note that the fundamental ($k = 1$) component exhibits the largest figure of merit and the harmonics decrease as $1/k^2$.

A standard approach to BPSK signal detection is to use a PFDM with a rectangular "null-to-null" bandwidth prefilter and delay set to one-half of the signal's bit duration. For this system, using (9) and assuming that the noise spectrum is white, we see that

$$T\frac{|b_k|^2}{S_{zz}(kf_b)} = \frac{\dfrac{a^4}{T^2}\left|\displaystyle\int_{-(f_b/2)}^{f_b/2} \cos(\pi f T) P_r\left(f + k\frac{f_b}{2}\right) P_r\left(f - k\frac{f_b}{2}\right) df\right|^2}{\dfrac{N_0^2}{2}\displaystyle\int_{-(f_b/2)}^{f_b/2} \cos^2(\pi f T)\, df}$$

where the integrals have finite limits because $H(f) = \mathrm{rec}(fT)$ for null-to-null filtering. Setting $k = 1$, we obtain

$$T\frac{|b_1|^2}{S_{zz}(f_b)} = 2\beta^2\left[\frac{16}{\pi^2}\int_0^{1/2}\frac{\cos^3(\pi f)}{1 - 4f^2}\,df\right]^2$$
$$\doteq 0.1547\beta^2$$

by numerical integration. This represents a degradation of 1.17 dB from the optimum figure—$2\beta^2/\pi^2$. Equation (9) can be further exploited to determine the ideal bandwidth for a rectangular prefilter when $d = T/2$. It can be shown that $H(f) = \mathrm{rec}(1.45fT)$ results in a higher value of the figure of merit than any other rectangular prefilter. The degradation factor for the ideal bandwidth rectangular prefilter is 0.74 dB, so it performs marginally better than the null-to-null bandwidth system.

Analysis of the $d = 0$ rectangular PFDM is also of practical interest. Null-to-null bandwidth prefiltering leads to

$$T\frac{|b_1|^2}{S_{zz}(f_b)} = \frac{\dfrac{a^4}{T^2}\left|\displaystyle\int_{-(f_b/2)}^{f_b/2} P_r\left(f + \frac{f_b}{2}\right) P_r\left(f - \frac{f_b}{2}\right) df\right|^2}{\dfrac{N_0^2}{2}\displaystyle\int_{-(f_b/2)}^{f_b/2}}$$
$$= 2\beta^2\left[\frac{8}{\pi^2}\int_0^{1/2}\frac{\cos^2(\pi f)}{1 - 4f^2}\,df\right]^2$$
$$\doteq 0.1220\beta^2$$

where we again assume a flat $(N_0/2)$ noise spectrum and the last step was the result of a numerical integration. Here, the degradation factor is 2.20 dB. As with the $d = T/2$ case treated previously, the null-to-null bandwidth is not the ideal bandwidth for a rectangular prefilter. It can be shown that the ideal rectangular filter for a prefilter-square device is $H(f) = \mathrm{rec}(0.85fT)$ with corresponding degradation 1.56 dB.

We see that the $d = T/2$ rectangular PFDM detector in general outperforms the rectangular prefilter-square ($d = 0$) detector. Since a square is easier to implement than a delay-and-multiply, it is a situation where an increase in complexity leads to enhanced performance. Recall that when optimum prefilters are employed, this is not the case at all. The delay-and-multiply detector performs no better than the square detector there.

Equation (9) can be further exploited to exhibit the robustness of the detection schemes with respect to system parameters. Considering rectangular filters of the form $H(f) = u(f + B) - u(f - B)$ for white Gaussian noise, we can rewrite (9) as

$$\frac{|b_1|^2}{S_{zz}(f_b)} = \frac{16\left(B - \dfrac{f_b}{2}\right)a^4}{N_0^2}\,\frac{\left|\displaystyle\int_0^{B-(f_b/2)} \cos(2\pi f d)\,\mathrm{sinc}\left(\pi\left(f + \frac{f_b}{2}\right)T\right)\mathrm{sinc}\left(\pi\left(f - \frac{f_b}{2}\right)T\right) df\right|^2}{1 + \mathrm{sinc}\left(4\pi\left(B - \dfrac{f_b}{2}\right)d\right)} \tag{21}$$

where we have taken $P(f) = T\,\mathrm{sinc}(\pi f T)$ and $k = 1$. By varying the parameters b, d, and f_b (remember $T = 1/f_b$) and then evaluating the right side of (21) (a computer was used to do this, the finite-limit definite integral was summed by the trapezoidal rule), we can determine the sensitivity of the rectangular PFDM to variations in the above parameters.

First, we investigate the effect of variations in the prefilter bandwidth B. In Fig. 4, we plot the figure of merit $(|b_1|^2/S_{zz}(f_b))$ versus B for $d = 0$ and $d = T/2$. Note that the delay-and-multiply device is more robust to prefilter bandwidth than the square device. (The 0 dB reference on these plots and those that follow is $(2/\pi^2)\beta^2$, the maximum value that the figure of merit can take on [cf. (20)].)

It is of some practical interest to fix both B and d. Then (21) can be used to determine the effectiveness of such a system with respect to the bit frequency f_b. Close inspection of (21) immediately gives us an upper limit—$f_b = 2B$—since the right-hand side vanishes at that point. In Fig. 5, the figure of merit is plotted versus f_b for $d = 0$, $d = 1/B$, $d = 1.45/B$, and $d = 2/B$. It appears that the $d = 1/B$ system performs the best over a broad range of signal bit frequencies. A $d = 1/B$ rectangular PFDM can therefore be used to detect the presence of a BPSK signal with unknown bit rate f_b. The fact that the flat region in the curve goes from ~0.6B to 1.4B indicates, however, that the uncertainty in f_b can at most be ~one octave.

We will now investigate the robustness of two of the

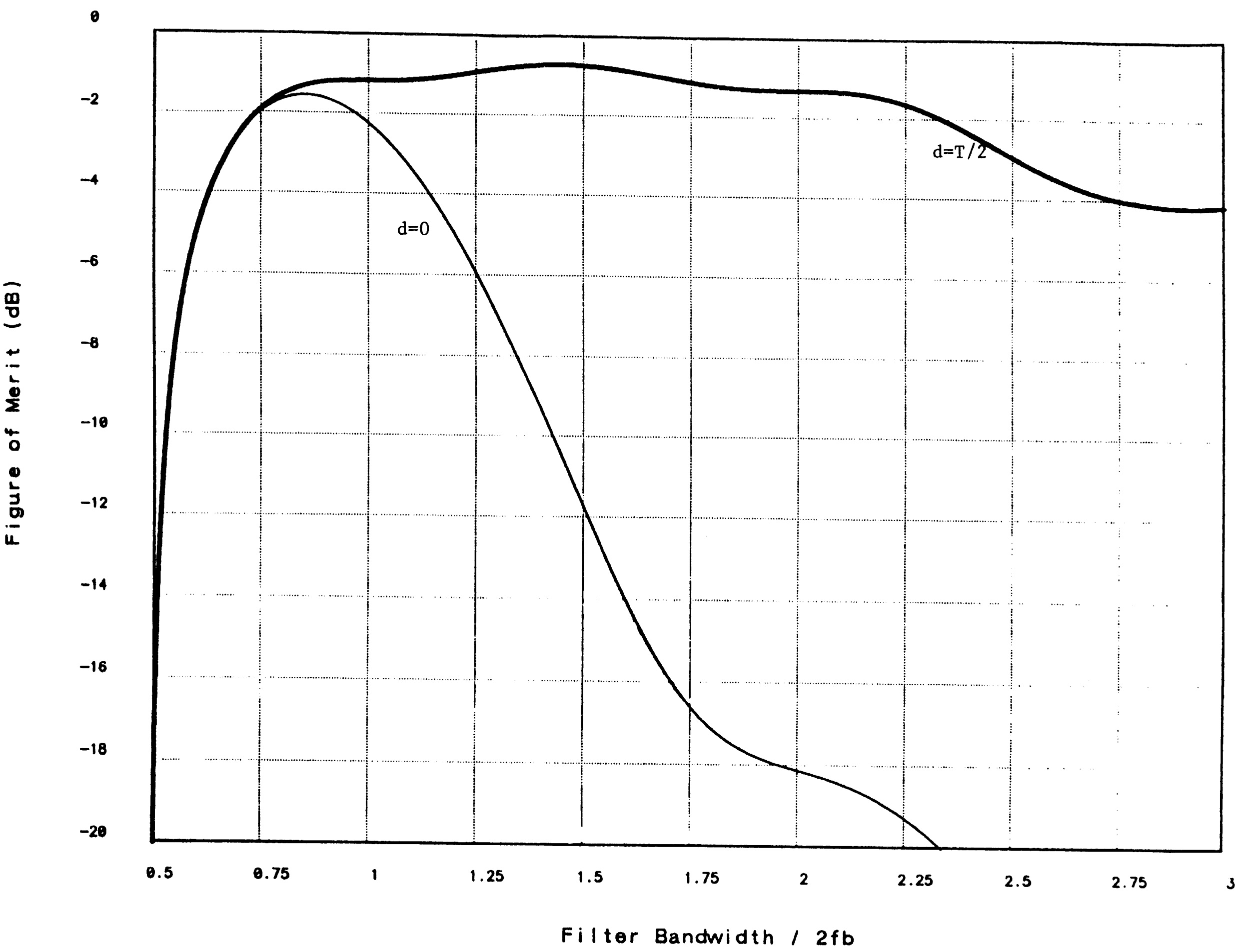


Fig. 4. Performance of rectangular prefilter-delay-and-multiply detector with respect to prefilter bandwidth.

optimal PFDM detectors with respect to uncertainty in the knowledge of the bit rate f_b. If we let $H(f) = \text{sinc}\,(\pi f l T)$, where l controls the bandwidth of the filter (inversely), (14) tells us that for rectangular signaling and white Gaussian noise, the $d = 0$ optimum prefilter has $l = 1$, and (15) indicates that the $d = T/2$ optimum prefilter has $l = 1/2$ since $\text{sinc}\,(\pi f T)/\cos\,(\pi f T/2) \propto \text{sinc}\,(\pi f T/2)$. The objective here is to determine the effect of the parameter l on the figure of merit. Substituting $H(f) = \text{sinc}\,(\pi f l T)$ and $P(f) = T\,\text{sinc}\,(\pi f T)$ into (9) and taking $k = 1$, we obtain

$$\frac{|b_1|^2}{S_{zz}(f_b)} = \frac{2a^4}{N_0^2}$$

$$\cdot \frac{\left| \int_{-\infty}^{\infty} \cos\,(2\pi f d)\, \text{sinc}\left(\pi\left(f + \frac{f_b}{2}\right) l T\right) \text{sinc}\left(\pi\left(f - \frac{f_b}{2}\right) l T\right) \text{sinc}\left(\pi\left(f + \frac{f_b}{2}\right) T\right) \text{sinc}\left(\pi\left(f - \frac{f_b}{2}\right) T\right) df \right|^2}{\int_{-\infty}^{\infty} \cos^2\,(2\pi f d)\, \text{sinc}^2\left(\pi\left(f + \frac{f_b}{2}\right) l T\right) \text{sinc}^2\left(\pi\left(f - \frac{f_b}{2}\right) l T\right) df}. \tag{22}$$

Equation (22) gives the spectral SNR at the output of a PFDM that is optimized for the detection of BPSK signals with bit duration T when, in fact, the true bit duration is lT. In Fig. 6, with the aid of Appendix B, (22) is plotted as a function of l for $d = 0$ and $d = T/2$. Note that the $d = T/2$ system is more robust. Its range of effectiveness, however, does not cover ~one octave as did the rectangular filter PFDM discussed above.

VI. Performance Analysis

In this section, we derive expressions for signal presence false alarm and detection probabilities (P_F and P_D)

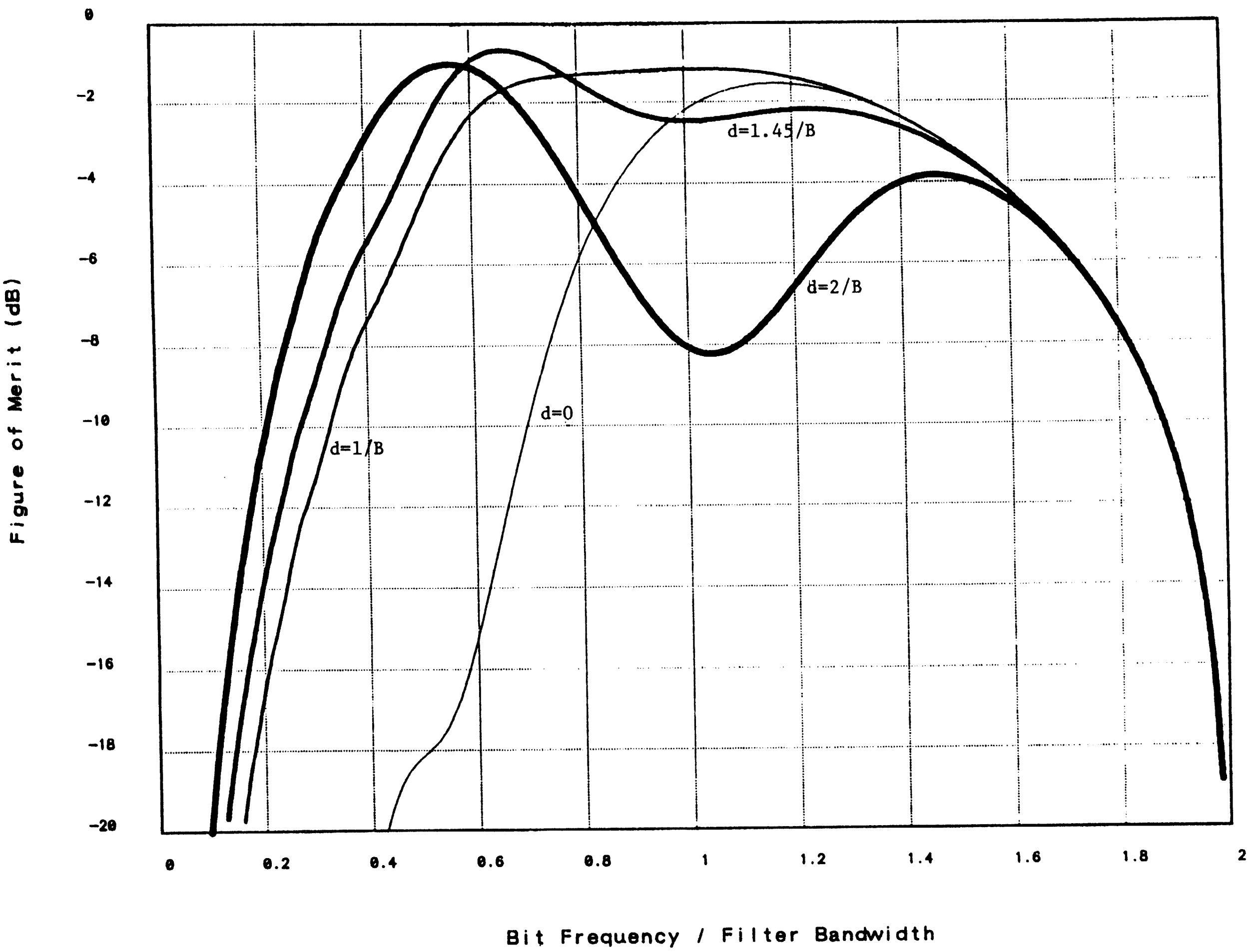


Fig. 5. Performance of various fixed delay, fixed bandwidth, rectangular prefilter-delay-and-multiply detectors with respect signal bit rate frequency.

for BPSK signals in AWGN. Both known and unknown bit rates are considered. In the known bit rate case, we assume that a correlator is used to detect the discrete spectral component at the PFDM output, while for unknown bit rates, an FFT is used. We choose these devices because they are simple yet powerful since they are the optimal detectors for discrete-spectrum signals in AWGN.

The resultant expressions for P_F and P_D will be used to study the performance tradeoffs related to the variation of a number of pertinent parameters. Several interesting points are brought out by these numerical results.

A. Known Bit Rates

1) Derivation of P_F and P_D: Since the bit rate is known, we use an optimal PFDM. We choose the $H(f) = \text{sinc}(\pi f T/2)$, $d = T/2$, optimal PFDM. The observation process at the output is given by

$$o(t) = \sum_{k=1}^{\infty} b_k \cos(2\pi k f_b t + \theta) + z(t) \qquad (23)$$

where b_k is given by (8) [$b_1 = 2a^2/\pi^2$, and the harmonics fall off at least as $1/k^2$ (if the optimum prefilter square system has been chosen, $b_k = 2a^2/\pi^2 k^2$)], θ is $-\pi/2$ or random, depending upon *a priori* knowledge of the bit timing Δ, and $z(t)$ is a non-Gaussian process with, by (A0),

$$R_{zz}(\tau) = \frac{4N_0^2}{T^4}\,\text{parab}(2\tau/T) \qquad (24)$$

where $\text{parab}(\cdot)$ is even symmetric with $\text{parab}(\cdot) = (\cdot - 1)^2$ for $\cdot \in [0, 1]$ and $\text{parab}(\cdot) = 0$ for $\cdot \in (1, \infty)$. The $-\pi/2$ phase factor for known bit timing can be seen from (7) since here $f_b d = 1/2$.

If somehow the bit timing is known at the receiver, $\theta = -\pi/2$ and the signal presence question is answered by comparing the coherent correlator test statistic

$$O_C = \frac{1}{T_o}\int_0^{T_o} o(t - T/4) \sum_{k=1}^{\infty} \frac{1}{k^2} \cos(2\pi k f_b t)\, dt \qquad (25)$$

to some threshold γ. The $T/4$ delay is introduced for phase synchronization purposes. In the more practical case, when the bit timing is unknown, θ is a uniform ran-

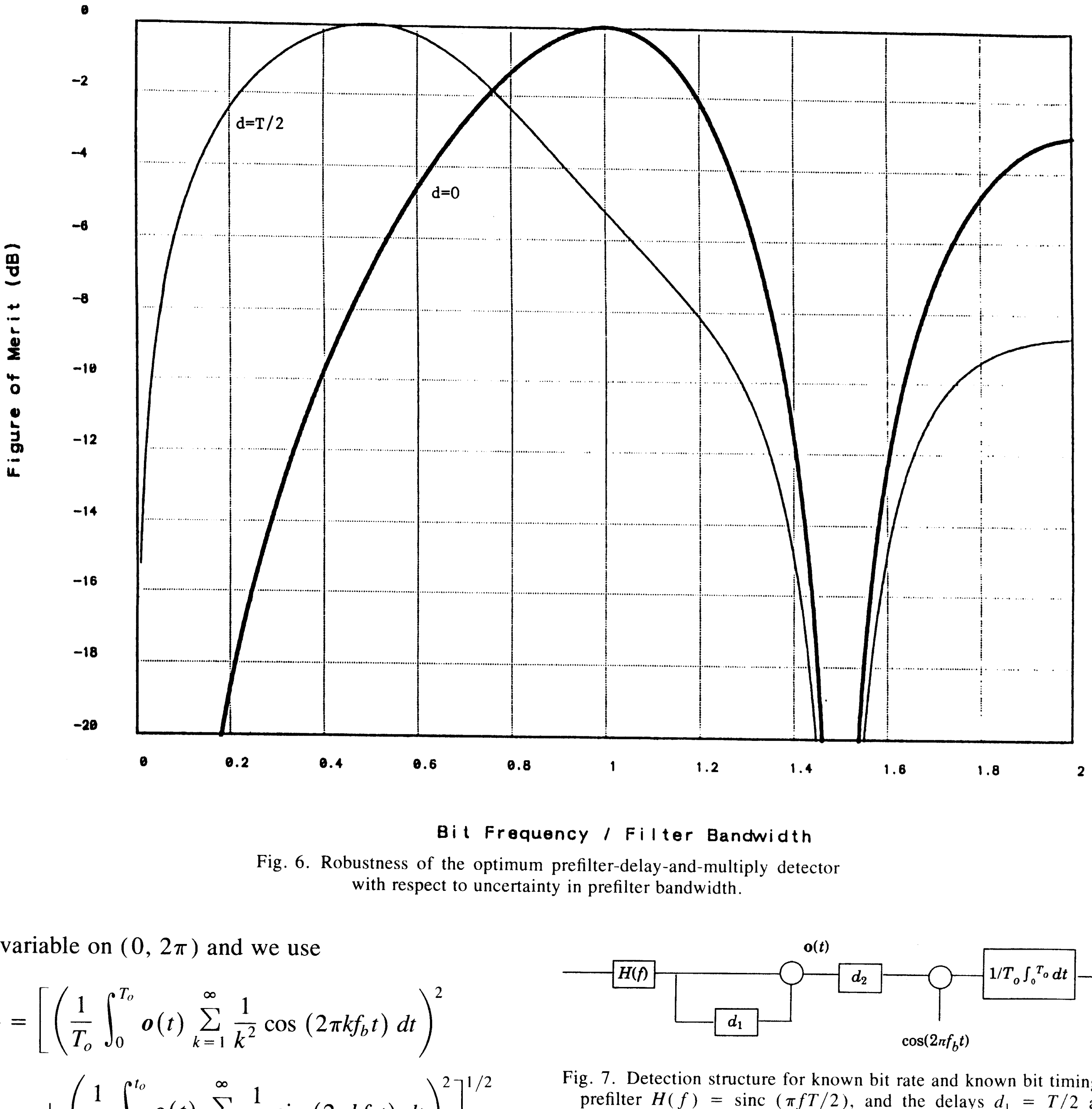


Fig. 6. Robustness of the optimum prefilter-delay-and-multiply detector with respect to uncertainty in prefilter bandwidth.

dom variable on $(0, 2\pi)$ and we use

$$\boldsymbol{O}_I = \left[\left(\frac{1}{T_o} \int_0^{T_o} o(t) \sum_{k=1}^{\infty} \frac{1}{k^2} \cos\,(2\pi k f_b t)\, dt \right)^2 + \left(\frac{1}{T_o} \int_0^{t_o} o(t) \sum_{k=1}^{\infty} \frac{1}{k^2} \sin\,(2\pi k f_b t)\, dt \right)^2 \right]^{1/2} \tag{26}$$

the output of an in-phase/quadrature correlator as a test statistic. In (25) and (26), T_o is the observation time. For both cases, our decision rule is

$$\boldsymbol{O} > \gamma \quad \text{decide signal present } (a > 0) \tag{27a}$$

$$\boldsymbol{O} < \gamma \quad \text{decide signal absent } (a = 0). \tag{27b}$$

We note here a similarity to the detection of coherent and noncoherent on-off-keyed (OOK) signals. Block diagrams of the complete detection structures that serve as the models for this analysis are shown in Figs. 7 and 8.

To derive false alarm and detection probabilities, it is necessary to know the distribution of the statistics $\boldsymbol{O}$ un-

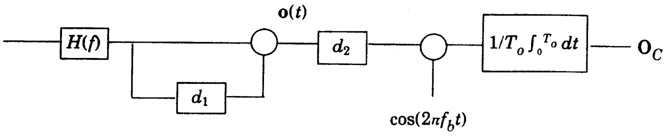


Fig. 7. Detection structure for known bit rate and known bit timing. The prefilter $H(f)$ = sinc $(\pi f T/2)$, and the delays $d_1 = T/2$ and $d_2 = T/4$.

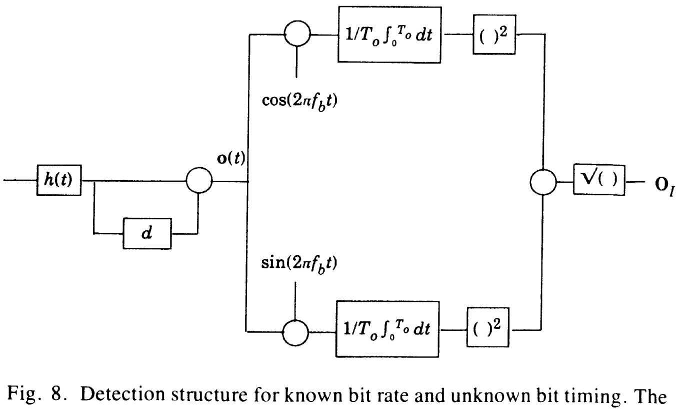


Fig. 8. Detection structure for known bit rate and unknown bit timing. The prefilter $H(f)$ = sinc $(\pi f T/2)$, and the delay $d = T/2$.

der the hypotheses $a > 0$ and $a = 0$. We examine the random variable

$$Z = \frac{1}{T_o} \int_0^{T_o} z(t)\, dt. \tag{28}$$

It is shown in Appendix C that provided $T_o \gg T$ (it will be shown presently that for the low input SNR's that are of practical interest, the observation time must in fact be very large compared to the bit duration in order to achieve reasonable values for P_F and P_D), $\boldsymbol{Z}$ is Gaussian by a version of the central limit theorem (CLT) for "m-dependent" sequences [13].

The statistics $\boldsymbol{O}_C$ are therefore Gaussian under both hypotheses with

$$E[\boldsymbol{O}_C | a = 0] = 0, \tag{29}$$

$$\frac{a^2}{\pi^2} < E[\boldsymbol{O}_C | a > 0] < \frac{a^2}{\pi^2} \frac{\pi^4}{90} \tag{30}$$

and

$$\begin{aligned} &\operatorname{var}(\boldsymbol{O}_C) \\ &= \frac{1}{4T_o} \int_0^{T_o} \int_0^{T_o} R_{zz}(t - s) \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \frac{1}{k^2 l^2} \\ &\quad \cdot \left[\cos\left(2\pi f_b(kt + ls)\right) + \cos\left(2\pi f_b(kt + ls)\right)\right] \\ &\quad \cdot dt\, ds. \end{aligned} \tag{31}$$

Equation (31) is difficult to deal with. Noting, however, that

$$\max\left\{E[\boldsymbol{O}_C | a > 0]\right\} \propto \sum_{k=1}^{\infty} \frac{1}{k^4} = \frac{\pi^4}{90} \doteq 1.082$$

and

$$\operatorname{var}(\boldsymbol{O}_C) \propto \sum_{k=1}^{\infty} \frac{1}{k^6} = \frac{\pi^6}{945} \doteq 1.017,$$

we see that correlating with all of the harmonics offers only slight improvement over correlating with just the fundamental ($k = 1$). This is why we include only the fundamental terms in our detection structure figures (Figs. 7 and 8). With just the fundamental terms we have

$$E[\boldsymbol{O}_C | a > 0] = \frac{a^2}{\pi^2} \equiv E_{\text{opt}} \tag{32}$$

and, from (24) and (31), it can be shown in a straightforward manner that

$$\operatorname{var}(\boldsymbol{O}_C) = \frac{1}{T_o} \frac{2N_0^2}{T\pi^2} \equiv \sigma_{\text{opt}}^2. \tag{33}$$

This being established, it is well known that the density functions for $\boldsymbol{O}_I$ follow the Rayleigh and Rician laws [14]:

$$p_{O_I}(\lambda | a = 0) = \frac{\lambda}{\sigma_{\text{opt}}^2} \exp\left\{-\frac{\lambda^2}{2\sigma_{\text{opt}}^2}\right\} \quad \lambda \geq 0 \tag{34}$$

and

$$p_{O_I}(\lambda | a > 0) = \frac{\lambda}{\sigma_{\text{opt}}^2} \exp\left\{-\frac{\lambda^2 + E_{\text{opt}}^2}{2\sigma_{\text{opt}}^2}\right\} I_0\left(\frac{\lambda E_{\text{opt}}}{\sigma_{\text{opt}}^2}\right)$$
$$\lambda \geq 0. \tag{35}$$

Using the decision rule (27), the false alarm and detection probabilities for the known bit timing case are

$$P_F = \Pr[\boldsymbol{O}_C > \gamma | a = 0]$$
$$\text{and } P_D = \Pr[\boldsymbol{O}_C > \gamma | a > 0]$$

which, because $\boldsymbol{O}_C$ is Gaussian, can be written as

$$P_F = Q\left(\frac{\gamma}{\sigma_{\text{opt}}}\right) \quad \text{and } P_D = Q\left(\frac{\gamma - E_{\text{opt}}}{\sigma_{\text{opt}}}\right) \tag{36}$$

where

$$Q(b) = \int_b^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\lambda^2/2}\, d\lambda.$$

When the bit timing is unknown, the corresponding probabilities are

$$P_F = \Pr[\boldsymbol{O}_I > \gamma | a = 0]$$
$$\text{and } P_D = \Pr[\boldsymbol{O}_I > \gamma | a > 0]$$

which, by (34) and (35), can be written as

$$P_F = \exp\left\{-\frac{\gamma^2}{2\sigma_{\text{opt}}^2}\right\} \quad \text{and } P_D = Q\left(\frac{E_{\text{opt}}}{\sigma_{\text{opt}}}, \frac{\gamma}{\sigma_{\text{opt}}}\right) \tag{37}$$

where

$$Q(b, c) = \int_c^{\infty} \lambda e^{-(\lambda^2 + b^2)/2} I_0(b\lambda)\, d\lambda.$$

2) Numerical Results: Equations (36) and (37) are plotted in Figs. 9 and 10. Comparing these curves gives some indication of the improvements afforded by knowledge of the bit timing Δ.

We now determine the observation time required for a given input SNR (β), P_D and P_F operating point. A parameter of the P_D curves is the ratio $E_{\text{opt}}/\sigma_{\text{opt}}$. With some minor manipulations of (32) and (33), we find that

$$T_o = 2(\pi/\beta)^2 (E_{\text{opt}}/\sigma_{\text{opt}})^2 T. \tag{38}$$

Note that for input SNR's less than unity, the observation time increases parabolically. That is, for every 3 dB (factor of 2) drop in β, the required observation time increases by a factor of 4.

As an example of the application of (38), we consider an ideal operating point for signal presence detection [16], namely, $P_F = 10^{-6}$, $P_D = 0.9$. Assuming that Δ is unknown, we find from (37) that for $P_F = 10^{-6}$, our normalized threshold $\gamma/\sigma_{\text{opt}} \doteq 5.3$. Inserting this into the (37) expression for P_D, we find that for $P_D = 0.9$, $E_{\text{opt}}/\sigma_{\text{opt}} \doteq 6.5$. Equation (38) then indicates that if $\beta = 1$ (0 dB),

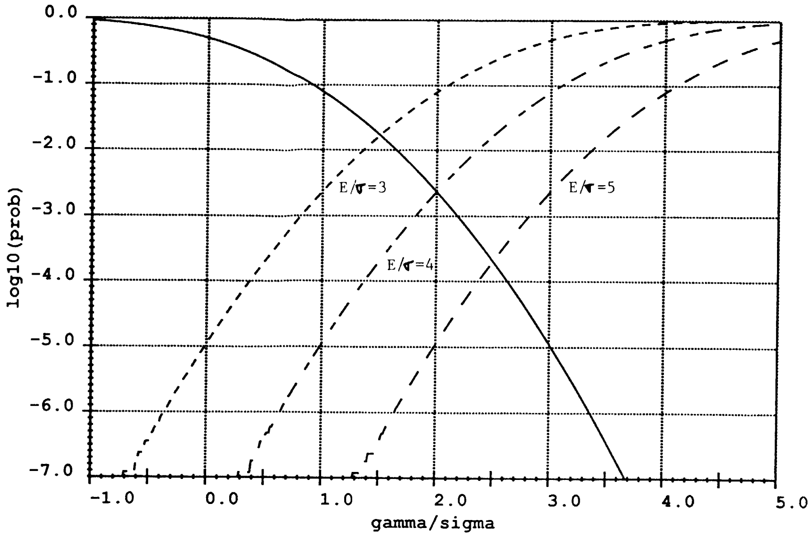


Fig. 9. Detection performance for known bit rates and known bit timing. Solid curve: P_F; dashed curves: $1 - P_D$.

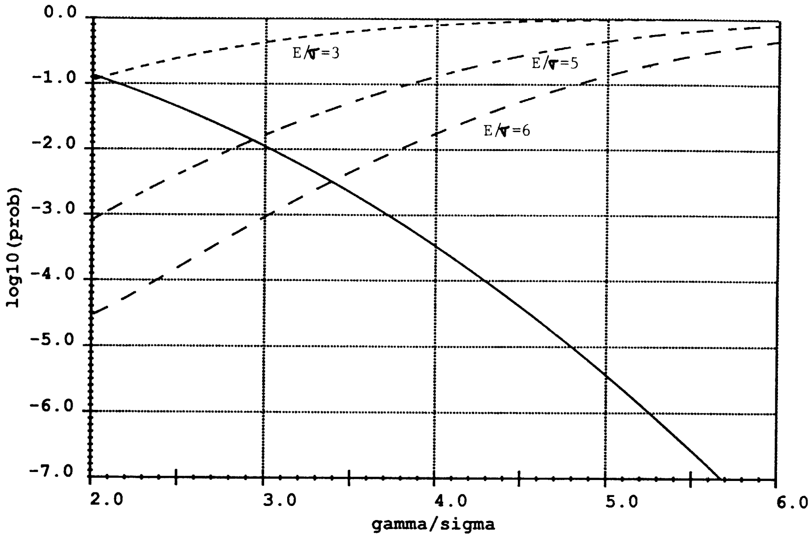


Fig. 10. Detection performance for known bit rates and unknown bit timing. Solid curve: P_F; dashed curves: $1 - P_D$.

the observation time required to achieve $P_F = 10^{-6}$, $P_D = 0.9$ is $T_o \doteq 834T$. Relaxing the constraint on false alarms to $P_F = 10^{-3}$, we obtain $T_o \doteq 473T$. Thus, for input SNR's $\beta \leq 0$ dB, the required observation time T_o is indeed large compared to the bit duration T. The detection probability is sketched as a function of observation time (for $P_F = 10^{-6}$, and $\beta = 0$ dB) in Fig. 11.

Another interesting situation is when the threshold and observation time are both fixed—the threshold by a P_F requirement and the observation time by a P_D requirement for a hypothesized input SNR β_{hypo}. The issue is: how does P_D degrade if the true value of β (β_{true}) is less than the hypothesized value? Noting from (38) that if $P_F = 10^{-6}$, $P_D = 0.9$ is the set operating point for β_{hypo}, a β_{true} 1 dB less than β_{hypo} results in $P_D \doteq 0.49$. If β_{true} is 3 dB less than β_{hypo}, the detection probability falls to $P_D \doteq 0.04$. The detection performance is thus very sensitive to the input SNR β.

In Section V we found that the spectral SNR at the output of an optimal PFDM is 1.17 dB greater than that at the output of a standard ad hoc configuration. Since $(E_{\text{opt}}/\sigma_{\text{opt}})^2$ is proportional to spectral SNR, (37) and (38) can be used to assess the performance degradation of the ad hoc detector with respect to the optimum detector. We see immediately that to achieve any specific operating point, the ad hoc detector will require $10^{0.117}$, which equals 1.31 times the observation time required by the optimal detector. On the other hand, if the observation time is fixed and β is such that the optimal detector gives $P_F = 10^{-6}$ and $P_D = 0.9$, it can be shown with (37) that use of the ad hoc detector reduces the detection probability to $P_D \doteq 0.66$.

B. Unknown Bit Rates

1) Derivation of P_F and P_D: To detect the presence of a BPSK signal with unknown bit rate in AWGN, we as-

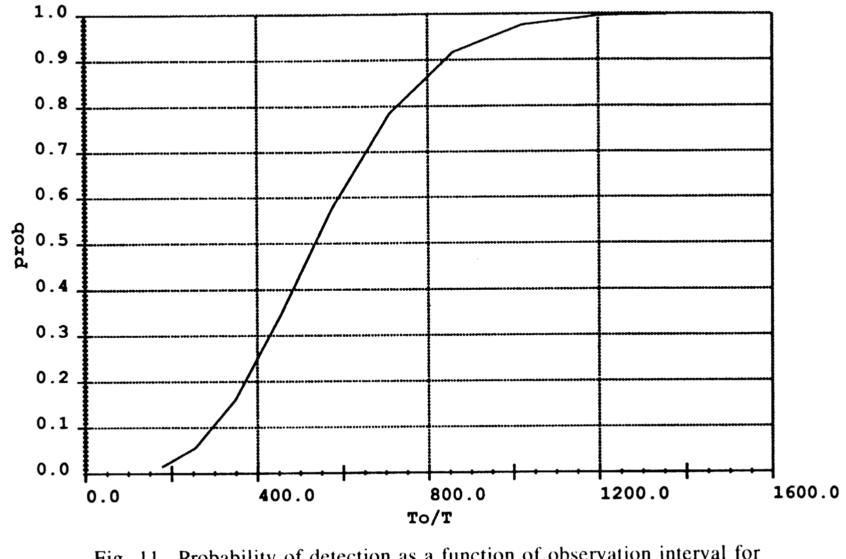


Fig. 11. Probability of detection as a function of observation interval for $\beta = 1$, and $P_F = 10^{-6}$.

sume that a $d = 1/B$ rectangular PFDM is employed to generate the bit rate line. As discussed in Section V and shown in Fig. 5, this PFDM is good for detecting the presence of signals with bit rates on $(0.6B, 1.4B)$. Since an FFT is to be used to detect the bit rate line in the PFDM output spectrum, our test statistics are

$$\boldsymbol{O}_i = \left| \frac{1}{T_o} \int_0^{T_o} \boldsymbol{o}(t) \exp\{-j2\pi f_i t\}\, dt \right| \qquad i = 1, 2, 3, \cdots, M \tag{39}$$

where $\{f_i\}_{i=1}^{M}$ are uniformly spaced frequencies on $(0.6B, 1.4B)$. An N-point FFT with sampling rate $f_s = \mu B$ can compute $M = T_o(0.8f_s)/\mu$ test statistics. The frequency resolution of the satistics is $f_{i+1} - f_i = 1/T_o$. (Because the signal to be detected has a discrete spectrum, it is important that the frequency resolution of the FFT be small. This implies that the observation time must be long and, hence, M large. If M is constrained in size by hardware or software limitations and the P_D that results from the current value of M is unacceptable, a technique often used is to average several FFT's to create M new statistics $\boldsymbol{O}_{i\text{avg}}$.) Note that for $f_i = f_b$, the right-hand sides of (26) and (38) are equivalent. The FFT can therefore be interpreted as M in-phase/quadrature correlators implemented in parallel.

The observation process $\boldsymbol{o}(t)$ given in (39) is as in (23); however, by (8),

$$b_1 \doteq 0.540a^2\sqrt{1 - f_b/2B},$$

and by (A0), $z(t)$ has correlation function

$$R_{zz}(\tau) = (N_0B)^2 \operatorname{sinc}^2(2\pi B\tau). \tag{40}$$

As shown in Appendix C, for $T_o \gg T$, the CLT again applies, albeit for this correlation function we use a version for "ρ-mixing" sequences [16]. The test statistics $\boldsymbol{O}_i$ will therefore be either Rayleigh or Rician distributed. The mean and variance terms that become parameters of these distributions are

$$E_i = E\left[\frac{1}{T_o}\int_0^{T_o} \boldsymbol{o}(t)\cos(2\pi f_i t)\, dt \,\middle|\, a > 0 \right.$$

$$\left. f_b = f_i,\ \theta = 0\right]$$

$$\doteq 0.270a^2\sqrt{1 - f_i/2B}$$

and

$$\sigma_i^2 = \frac{1}{T_o^2}\int_0^{T_o}\int_0^{T_o} R_{zz}(t - s)\cos(2\pi f_i t)\cos(2\pi f_i t)\, dt\, ds$$

$$\simeq \frac{N_0^2 B}{T_o}[1 - f_i/2B].$$

We note that, because E_i and σ_i^2 depend upon the frequency f_i, the test statistics $\boldsymbol{O}_i$ are not identically distributed, which makes writing down a signal presence decision rule complicated. It is possible, however, to scale the statistics $\boldsymbol{O}_i$ by a factor $A_i = 1/\sqrt{(2 - f_i/B)}$ in order to create a new sequence $\boldsymbol{O}_i' = A_i\boldsymbol{O}_i$ of identically distributed statistics with parameters

$$E_{\text{rec}} \equiv 0.191a^2 \tag{41}$$

and

$$\sigma_{\text{rec}}^2 \equiv \frac{N_0^2 B}{2T_o}. \tag{42}$$

The test statistics are also independent. This is illustrated in Appendix D. A block diagram of the complete detection structure is illustrated in Fig. 12.

The signal presence decision rule is described now. We compare the largest of the scaled statistics to a threshold.

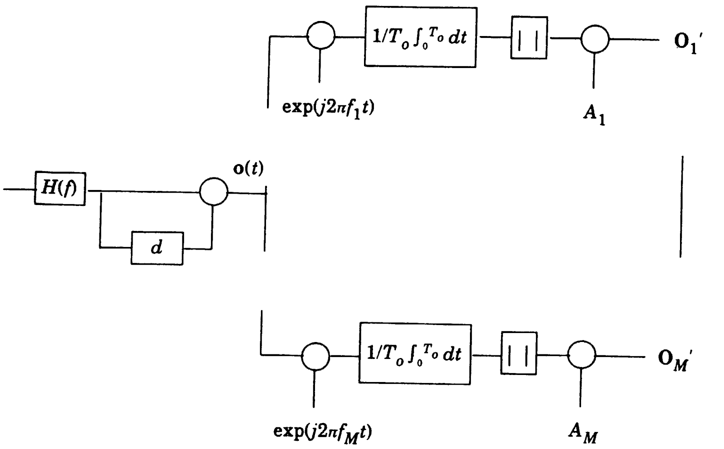


Fig. 12. Detection structure for unknown bit rates. The prefilter $H(f) = \text{rec}(f/B)$, the delay $d = 1/B$, the frequencies $f_i = 0.6B + (i - 1)/T_o$, and the scale factors $A_i = 1/\sqrt{(2 - f_i/B)}$.

Let $O_l' = \max \{O_i'\}_{i=1}^{M}$; the rule is

$$O_l' > \gamma \quad \text{decide signal present } (a > 0, f_b = f_l) \tag{43a}$$

$$O_l' < \gamma \quad \text{decide no signal present } (a = 0). \tag{43b}$$

Note the similarity of the above rule to the one for noncoherent detection of M-ary frequency-shift-keyed (MFSK) signals.

The probability of false alarm is derived by noting that

$$\begin{aligned} P_F &= \Pr [O_l' > \gamma | a = 0] \\ &= 1 - \Pr [O_l' \leq \gamma | a = 0] \\ &= 1 - \Pr [O_i' \leq \gamma \quad \text{for every } i | a = 0] \\ &= 1 - \left[\int_0^{\gamma} \frac{\lambda}{\sigma_{\text{rec}}^2} \exp\left\{ -\frac{\lambda^2}{2\sigma_{\text{rec}}^2} \right\} d\lambda \right]^M \\ &= 1 - \left[1 - \exp\left\{ -\frac{\lambda^2}{2\sigma_{\text{rec}}^2} \right\} \right]^M \end{aligned} \tag{44}$$

where the fourth step follows from the independence of the statistics. The detection probability is derived similarly:

$$\begin{aligned} P_D &= \Pr [O_l' > \gamma | a > 0, f_b = f_l] \\ &= \int_0^{\infty} \Pr [O_i' < \alpha \quad \text{for every } i \neq l \\ &\qquad O_l' > \gamma | a > 0, f_b = f_l] p_{O_l'}(\alpha) \, d\alpha \\ &= \int_{\gamma}^{\infty} \left[1 - \exp\left\{ -\frac{\alpha^2}{2\sigma_{\text{rec}}^2} \right\} \right]^{M-1} \frac{\alpha}{\sigma_{\text{rec}}^2} \\ &\qquad \cdot \exp\left\{ -\frac{\alpha^2 + E_{\text{rec}}^2}{2\sigma_{\text{rec}}^2} \right\} I_0\left(\frac{\alpha E_{\text{rec}}}{\sigma_{\text{rec}}^2} \right) d\alpha \\ &= \sum_{n=0}^{M-1} (-1)^n \binom{M-1}{n} \frac{1}{n+1} \\ &\qquad \cdot \exp\left\{ -\frac{E_{\text{rec}}^2}{2\sigma_{\text{rec}}^2} \frac{n}{n+1} \right\} \\ &\qquad \cdot Q\left(\frac{E_{\text{rec}}}{\sigma_{\text{rec}} \sqrt{n+1}}, \frac{\gamma \sqrt{n+1}}{\sigma_{\text{rec}}} \right). \end{aligned} \tag{45}$$

In deriving (44) and (45), examples in [17] and [18] have provided help.

2) Numerical Results: Equations (44) and (45) are plotted in Fig. 13 for the case $M = 512$. It can be seen by comparing the $P_F = 10^{-6}$ abscissas and the corresponding P_D in Figs. 10 and 13 that the performance of the FFT is not significantly different from the incoherent correlator. This is due to the fact that with the correlator, the output statistic is compared to a threshold and, with the FFT, the largest of the M output statistics is compared to a threshold. The only difference in the detection probability is consequently the probability that the largest of the M FFT statistics does not contain the bit rate line. This factor is given by the $[1 - \exp\{-\alpha^2/2\sigma_{\text{rec}}^2\}]^{M-1}$ term in (45). The ramification here is that it may be prudent in some known bit rate situations to use an FFT instead of a correlator, just in case the bit rate is slightly mismatched from its assumed known value. If the bit rate is in fact equal to its assumed value, the detection performance degradation due to the FFT is slight.

VII. Summary and Conclusions

In this paper, we applied Fourier analysis and stationary process theory to derive the optimum prefilter-delay-and-multiply structure for the detection of BPSK (or BPSK DS/SS) signals with known bit (or chip) rates in additive colored Gaussian noise. The structure is optimum in that it maximizes the spectral SNR of an output periodic waveform with a fundamental frequency equal to the bit (or

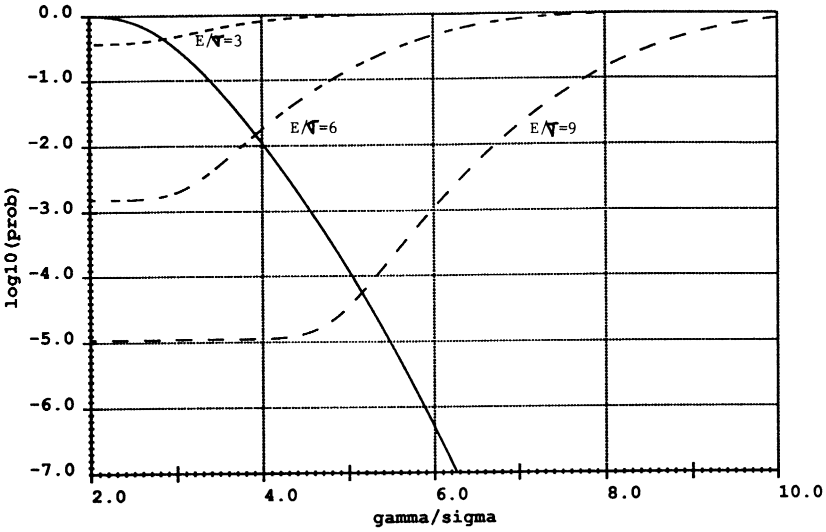


Fig. 13. Detection performance for unknown bit rates and unknown bit timing. Solid curve: P_F; dashed curves: $1 - P_D$.

chip) rate of the signal. We found that the optimum PFDM structure is not unique; it occurs whenever the prefilter, delay, and signal pulse shape satisfy a particular nonlinear functional equation. Two of the optimum structures were examined in some detail since they closely resembled two ad hoc schemes commonly used for PSK signal detection. These were the optimal prefilter-square ($d = 0$) and the optimal PFDM with one-half bit delay ($d = T/2$). The optimal prefilter-square was particularly interesting as it turned out that the optimal prefilter was one that was matched to BPSK signal's pulse. It was also interesting because its existence implied that for known bit rates, a delay-and-multiply device can perform no better than a square device, which is easier to implement. The optimal $d = T/2$ PFDM was found to be less sensitive to uncertainties in the bit rate, however.

The analytic methodology that we adopted—defining a figure of merit (the spectral SNR), deriving it for any general PFDM, and then maximizing it to specify the form of the optimum PFDM—proved useful in the analysis of suboptimum structures, as well as optimum ones. We were able to characterize the performance of some ad hoc detection schemes. It was determined that in terms of the spectral SNR, the optimum structures perform only marginally better (op ~ 1-2 dB) than the ad hoc schemes. The generality of our analysis also enabled us to find a robust PFDM capable of detecting BPSK signals with unknown bit rates that range over slightly more than one octave of frequencies. This was the rectangular-prefilter PFDM with prefilter bandwidth B and delay duration d related through $d = 1/B$.

Expressions for signal presence false alarm and detection probabilities (P_F and P_D) were derived for both the known and unknown bit rate cases. The derivation relied upon CLT arguments to obtain Gaussian statistics and exploited analogies to OOK and MFSK signaling. The expressions showed that while P_F is independent of signal strength, for a given P_F P_D increases with spectral SNR (because $E/\sigma \propto$ spectral SNR and the Q-functions increase with that parameter). Therefore, spectral SNR was a good choice for the figure of merit since the PFDM structures that maximize spectral SNR also maximize P_D.

The false alarm and detection probability expressions were then used to investigate how parameters such as input SNR and observation time effect P_F and P_D. Among other things, two significant points were brought out by this study. The first of these points is that although the optimum PFDM structures exhibit spectral SNR's that are only slightly higher than the ones possible with an ad hoc structure, this slight increase in spectral SNR can mean a dramatic increase in P_D. The second point is that in lieu of a correlator, use of an FFT for detecting bit lines of known frequency does not significantly reduce P_D, while on the other hand it offers a measure of robustness for bit frequency uncertainty.

In conclusion, we reiterate that the presence of a BPSK or BPSK DS/SS signal can be detected through the use of a continuous-to-discrete-spectrum-transformation operation. Restricting ourselves to PFDM transformations which generate discrete spectral components at the bit frequency of the signal, what we accomplished here was 1) derived the optimum transformation when the bit frequency is known, 2) found a robust transformation when the bit frequency is unknown, and 3) developed expressions which characterized the signal presence detection performance (via P_D and P_F) of these transformations with regard to SNR, observation time, threshold settings, and whether or not the signal's bit rate or bit timing is known *a priori* at the receiver.

Appendix A
Derivation of $S_{zz}(f)$

The approach we adopt here is to find the autocorrelation function of $z(t)$ (cf. Fig. 2) and then apply the

Wiener–Khintchine theorem. Now,

$$z(t) = x(t)\,x(t-d)$$

so the autocorrelation

$$\begin{aligned} R_{zz}(\tau) &\equiv E[z(t)\,z(t+\tau)] \\ &= E[x(t)\,x(t-d)\,x(t+\tau)\,x(t-d+\tau)] \\ &= R_{xx}^2(d) + R_{xx}^2(\tau) + R_{xx}(\tau+d)\,R_{xx}(\tau-d) \end{aligned} \tag{A0}$$

where $R_{xx}(\tau) = E[x(t)\,x(t+\tau)]$, and the last step, which is valid for Gaussian processes, follows from Karlin and Taylor [19]. (This result is sometimes called Price's theorem.) Notice that since $R_{zz}(\tau)$ depends upon τ only and not on t, $z(\cdot)$ is wide-sense stationary.

The power spectrum of $z(t)$ is

$$S_{zz}(f) = \int_{-\infty}^{\infty} R_{zz}(\tau) e^{-j2\pi f\tau}\, d\tau$$

where $j = \sqrt{(-1)}$. Exploiting the Price theorem result and invoking the convolution theorem, we obtain

$$\begin{aligned} S_{zz}(f) = R_{zz}^2(d)\,\delta(f) &+ \int_{-\infty}^{\infty} S_{xx}(\zeta)\,S_{xx}(f-\zeta)\,d\zeta \\ &+ \int_{-\infty}^{\infty} S_{xx}(\zeta) e^{j2\pi\zeta d} S_{xx}(f-\zeta) e^{-j2\pi(f-\zeta)d}\, d\zeta \end{aligned}$$

where

$$S_{xx}(f) = \int_{-\infty}^{\infty} R_{xx}(\tau) e^{-j2\pi f\tau}\, d\tau.$$

The power spectrum of $x(t)$ can be written in terms of the filter response via

$$S_{xx}(f) = |H(f)|^2 S_{nn}(f)$$

so

$$\begin{aligned} S_{zz}(f) = R_{xx}^2(d)\,\delta(f) &+ \int_{-\infty}^{\infty} |H(\zeta)|^2 |H(f-\zeta)|^2 \\ &\cdot S_{nn}(\zeta)\,S_{nn}(f-\zeta)\,d\zeta \\ &+ \int_{-\infty}^{\infty} |H(\zeta)|^2 |H(f-\zeta)|^2 S_{nn}(\zeta) \\ &\cdot S_{nn}(f-\zeta) e^{-j2\pi(f-2\zeta)d}\, d\zeta. \end{aligned}$$

After collecting terms, making a variable substitution, exploiting the symmetry of the filter, and applying a trigonometric identity, we see that

$$\begin{aligned} S_{zz}(f) = R_{xx}^2(d)\,\delta(f) + 2 \int_{-\infty}^{\infty} &\cos^2(2\pi\lambda d) \\ &\cdot \left|H\left(\lambda + \frac{f}{2}\right)\right|^2 \left|H\left(\lambda - \frac{f}{2}\right)\right|^2 \\ &\cdot S_{nn}\left(\lambda + \frac{f}{2}\right) S_{nn}\left(\lambda - \frac{f}{2}\right) d\lambda. \end{aligned}$$

Appendix B
Evaluation of (22)

Equation (22) is rewritten and renamed here for convenience:

$$\frac{|b_1|^2}{S_{zz}(f_b)} = \frac{2a^4}{N_0^2} \cdot \frac{\left| \int_{-\infty}^{\infty} \cos(2\pi fd)\, \text{sinc}\left(\pi\left(f + \frac{f_b}{2}\right)lT\right) \text{sinc}\left(\pi\left(f - \frac{f_b}{2}\right)lT\right) \text{sinc}\left(\pi\left(f + \frac{f_b}{2}\right)T\right) \text{sinc}\left(\pi\left(f - \frac{f_b}{2}\right)T\right) df \right|^2}{\int_{-\infty}^{\infty} \cos^2(2\pi fd)\, \text{sinc}^2\left(\pi\left(f + \frac{f_b}{2}\right)lT\right) \text{sinc}^2\left(\pi\left(f - \frac{f_b}{2}\right)lT\right) df}. \tag{A1}$$

It is of interest to determine the impact of the parameter l on $|b_1|/S_{zz}(f_b)$ above. The infinities in the integration limits inhibit the utility of numerical approaches so we must seek a closed form expression.

The numerator will be attacked first. Defining

$$\begin{aligned} X_l(f) \equiv X_{l+}(f)\,X_{l-}(f) \equiv\ &\text{sinc}\left(\pi\left(f + \frac{f_c}{2}\right)lT\right) \\ &\cdot \text{sinc}\left(\pi\left(f - \frac{f_c}{2}\right)lT\right) \end{aligned}$$

we can evaluate the integral by invoking both Parseval's identity and the Fourier convolution theorem to obtain

$$\begin{aligned} &\int_{-\infty}^{\infty} \cos(2\pi fd)\, \text{sinc}\left(\pi\left(f + \frac{f_b}{2}\right)lT\right) \\ &\quad \cdot \text{sinc}\left(\pi\left(f - \frac{f_b}{2}\right)lT\right) \text{sinc}\left(\pi\left(f + \frac{f_b}{2}\right)T\right) \\ &\quad \cdot \text{sinc}\left(\pi\left(f - \frac{f_b}{2}\right)T\right) df \\ &= \int_{-\infty}^{\infty} \cos(2\pi fd) X_l(f)\, X_1(f)\, df \\ &= \int_{-\infty}^{\infty} \frac{1}{2}\,[\delta(t+d) + \delta(t-d)] \\ &\quad \cdot \int_{-\infty}^{\infty} x_l(\lambda)\, x_1(t-\lambda)\, d\lambda\, dt \\ &= \frac{1}{2} \int_{-\infty}^{\infty} x_l(t)[x_1(t+d) + x_1(t-d)]\, dt \end{aligned} \tag{A2}$$

where x is the inverse Fourier transform of X and we have exploited its symmetry. Note that for $d = 0$, (A2) reduces to a standard form of Parseval's identity.

Since $X_l(f)$ is the product of two frequency domain functions, $x_l(t)$ can be determined via the convolution theorem. From the tables and by the frequency shift property, we have

$$x_{l\pm}(t) = \frac{1}{lT} e^{\pm i\pi f_b t}\left[u(t + lT/2) - u(t - lT/2)\right]$$

where $i = \sqrt{(-1)}$. Applying the convolution theorem, we obtain

$$\begin{aligned} x_l(t) &= \int_{-\infty}^{\infty} x_{l+}(\lambda)\, x_{l-}(t - \lambda)\, d\lambda \\ &= \frac{1}{\pi l^4 T^2}\Big\{\sin\big(\pi f_b(t + lT)\big)\left[u(t + lT) - u(t)\right] \\ &\quad - \sin\big(\pi f_b(t - lT)\big)\left[u(t) - u(t - lT)\right]\Big\} \end{aligned} \tag{A3}$$

where the second step follows after considerable, but straightforward, effort—convolving gated phasors.

At this point, we isolate the two cases $d = 0$ and $d = T/2$. Taking $d = 0$, substituting the above expression for $x_l(t)$ into (A2), and using symmetries, we have

$$\begin{aligned} &\int_{-\infty}^{\infty} X_l(f)\, X_1(f)\, df \\ &= \int_{-\infty}^{\infty} x_l(t)\, x_1(t)\, dt \\ &= 2\int_0^{\min\{T, lT\}} \frac{1}{\pi^2 l^2 T^2} \sin\big(\pi f_b(t - lT)\big) \\ &\quad \cdot \sin\big(\pi f_b(t - T)\big)\, dt \\ &= \frac{1}{\pi^2 l^2 T}\Big[l \cos\big(\pi(1 - l)\big) - \frac{1}{2\pi}\Big(\sin\big(\pi(1 + l)\big) \\ &\quad + \sin\big(\pi(1 - l)\big)\Big)\Big], \qquad l < 1 \\ &= \frac{1}{\pi^2 l^2 T}\Big[\cos\big(\pi(1 - l)\big) - \frac{1}{2\pi}\Big(\sin\big(\pi(1 + l)\big) \\ &\quad - \sin\big(\pi(1 - l)\big)\Big)\Big], \qquad l \geq 1. \end{aligned} \tag{A4}$$

In the case of $d = T/2$, it is instructive to note that

$$\begin{aligned} &x_1(t + d) + x_1(t - d) \\ &\quad = -\cos(\pi f_b t)\left[u(t + 3T/2) - u(t + T/2)\right] \\ &\qquad + 2\cos(\pi f_b t)\left[u(t + T/2) - u(t - T/2)\right] \\ &\qquad - \cos(\pi f_b t)\left[u(t - T/2) - u(t - 3T/2)\right]. \end{aligned}$$

Upon substitution into (A2), we obtain [analogous to (A4)]

$$\begin{aligned} &\int_{-\infty}^{\infty} \cos(\pi f T) X_l(f)\, X_1(f)\, df \\ &= \frac{1}{2}\int_{-\infty}^{\infty} x_l(t)\left[x_1(t + T/2) + x_1(t - T/2)\right] dt \\ &= \frac{1}{\pi^2 l^2 T^2}\Bigg[\Bigg[-\int_0^{lT} 2\sin\big(\pi f_b(t - lT)\big) \\ &\quad \cdot \cos(\pi f_b t)\, dt\Bigg] \cdot \left[u(l) - u(l - 1/2)\right] \\ &\quad + \Bigg[-\int_0^{T/2} 2\sin\big(\pi f_b(t - lT)\big)\cos(\pi f_c t)\, dt \\ &\quad + \int_{T/2}^{lT} \sin\big(\pi f_b(t - lT)\big)\cos(\pi f_b t)\, dt\Bigg] \\ &\quad \cdot \left[u(l - 1/2) - u(l - 3/2)\right] \\ &\quad + \Bigg[-\int_0^{T/2} 2\sin\big(\pi f_b(t - lT)\big)\cos(\pi f_c t)\, dt \\ &\quad + \int_{T/2}^{3T/2} \sin\big(\pi f_b(t - lT)\big)\cos(\pi f_b t)\, dt\Bigg] \\ &\quad \cdot u(l - 3/2)\Bigg] \\ &= \frac{1}{\pi^2 lT}\sin(\pi l), \qquad 0 < l \leq 1/2 \\ &= \frac{1}{2\pi^2 l^2 T}\Big\{(3/2 - l)\sin(\pi l) \\ &\quad - \frac{3}{2\pi}\left[\cos(\pi l) - \cos\big(\pi(1 - l)\big)\right]\Big\}, \\ &\qquad 1/2 < l \leq 3/2 \\ &= \frac{1}{2\pi^3 l^2 T}\Big\{\frac{3}{2}\cos\big(\pi(1 - l)\big) - \cos(\pi l) \\ &\quad - \frac{1}{2}\cos\big(\pi(3 - l)\big)\Big\}, \qquad 3/2 < l. \end{aligned} \tag{A5}$$

We have finished the numerator.

The denominator of (A1) can be evaluated similarly. Note that

$$\begin{aligned} &\int_{-\infty}^{\infty} \cos^2(2\pi f d)\, \mathrm{sinc}^2\left(\pi\left(f + \frac{f_b}{2}\right)lT\right) \\ &\quad \cdot \mathrm{sinc}^2\left(\pi\left(f - \frac{f_b}{2}\right)lT\right) df \\ &= \int_{-\infty}^{\infty} \cos^2(2\pi f d) X_l^2(f)\, df \\ &= \int_{-\infty}^{\infty} \frac{1}{2}\left[\delta(t) + \frac{1}{2}\delta(t + 2d) + \frac{1}{2}\delta(t - 2d)\right] \\ &\quad \cdot \int_{-\infty}^{\infty} x_l(\lambda)\, x_l(t - \lambda)\, d\lambda\, dt \\ &= \frac{1}{2}\int_{-\infty}^{\infty}\Big\{x_l^2(t) \\ &\quad + \frac{1}{2}x_l(t)\left[x_l(t + 2d) + x_l(t - 2d)\right]\Big\} dt. \end{aligned} \tag{A6}$$

Again, we isolate the cases $d = 0$ and $d = T/2$. For $d = 0$, (A3) and (A6) give

$$\begin{aligned}\int_{-\infty}^{\infty} X_l^2(f)\,df &= \int_{-\infty}^{\infty} x_l^2(t)\,dt \\ &= 2\int_0^{lT} \frac{1}{\pi^2 l^4 T^2}\sin^2\left(\pi f_b(t - lT)\right)dt \\ &= \frac{1}{\pi^2 l^4 T}\left[l - \frac{1}{2\pi}\sin(2\pi l)\right]. \qquad (A7)\end{aligned}$$

When $d = T/2$, we must consider regions for the parameter l. It can be shown by substituting (A3) into (A6) that

$$\begin{aligned}&\int_{-\infty}^{\infty} \cos^2(\pi f T) X_l^2(f)\,df \\ &= \frac{1}{2\pi^2 l^4 T}\left[l - \frac{1}{2\pi}\sin(2\pi l)\right] \\ &\qquad 0 < l \leq 1/2 \\ &= \frac{1}{2\pi^2 l^4 T}\left[\left[l - \frac{1}{2\pi}\sin(2\pi l)\right] - \frac{1}{2}\left[(2l - 1)\right.\right. \\ &\quad \left.\left.\cdot \cos\left(\pi(1 - 2l)\right) + \frac{1}{\pi}\sin\left(\pi(1 - 2l)\right)\right]\right], \\ &\qquad 1/2 < l \leq 1 \\ &= \frac{1}{2\pi^2 l^4 T}\left[\left[l - \frac{1}{2\pi}\sin(2\pi l)\right]\right. \\ &\quad - \frac{1}{2}\left[\cos\left(\pi(1 - 2l)\right) + 2(l - 1)\right. \\ &\quad \left.\left. - \frac{1}{\pi}\sin\left(\pi(1 - 2l)\right)\right]\right], \qquad 1 < l. \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad (A8)\end{aligned}$$

This completes the denominator.

Equations (A4), (A5), (A7), and (A8) can be used to write a closed-form expression for (A1). Parenthetically, we remark that although this was a laborious task, we can now plot $|b_l|/S_{zz}(f_b)$ versus l with must less computational effort than by numerically integrating (A1) for each value of l.

Appendix C
Central Limit Theorem Arguments for Detection Test Statistics

For both the known and unknown bit rate detection statistics, we must investigate the random variable

$$\boldsymbol{Z} = \frac{1}{T_o}\int_0^{T_o} z(t)\,dt \qquad (A9)$$

where $z(t)$ is a non-Gaussian process with

$$R_{zz}(\tau) = \frac{4N_0}{T^4}\,\text{parab}(2\tau/T) \qquad (A10)$$

for known bit rate statistics, and

$$R_{zz}(\tau) = (N_0 B)^2\,\text{sinc}^2(2\pi B\tau) \qquad (A11)$$

for the unknown bit rate statistics. Since the integral in (A9) can be thought of as the limit of a Riemann sum, we treat the distribution of the random variable $\boldsymbol{Z}$ as the distribution of the sum of a sequence of random variables $\{z(n\Delta t)\}$ where Δt is arbitrarily small. We cannot simply apply the standard version of the CLT because the variables $z(n\Delta t)$ are not independent. We must apply versions for dependent variables.

When (A10) holds, we see that $z(n\Delta t)$ and $z(l\Delta t)$ are uncorrelated random variables, provided $|n - l|\Delta t > T/2$. Since the variates $z(\cdot)$ are non-Gaussian, this does not imply independence. Glancing at Fig. 2, however, we note that we can write $z(n\Delta t)$ as the product of two uncorrelated, hence independent, Gaussian random variables:

$$z(n\Delta t) = \boldsymbol{x}(n\Delta t)\,\boldsymbol{x}(n\Delta t - T/2)$$

$\boldsymbol{x}(n\Delta t)$ and $\boldsymbol{x}(n\Delta t - T/2)$ are Gaussian because $h(t)$ ($= (2/T)$ rec $(t/4T)$) is a linear system, and independent because $R_{xx}(\tau) = (2N_0/T)$ tri $(\tau/2T)$ vanishes for $|\tau| \geq T/2$. Consequently, notice that $z(n\Delta t) = \boldsymbol{x}(n\Delta t)\,\boldsymbol{x}(n\Delta t - T/2)$ and $z(n\Delta t + T) = \boldsymbol{x}(n\Delta t + T)\,\boldsymbol{x}(n\Delta t + T/2)$ are independent random variables. If we let $|m|\Delta t = T$, then $z(n\Delta t)$ and $z((n + l)\Delta t)$ are independent for all $|l| > |m|$, and $\{z(n\Delta t)\}$ is termed an "m-dependent" sequence. Billingsley [13, Theorem 27.5] can therefore be invoked to argue that $\boldsymbol{Z}$ in (A9) is Gaussian when $T_o \gg T$, and (A10) holds.

Since the correlation function in (A11) is not time-limited, we cannot hope for an m-dependent sequence here. We look for a dependent variable CLT version, which relates directly to the sequence correlation function. The dependencies in a stochastic sequence are often characterized by so-called "mixing conditions" [20]. Versions of the CLT have been derived for a number of these conditions [14]. The mixing condition that relates to a sequence correlation function is "p-mixing" with the characteristic parameter p_l [20]. For our sequence $\{z(n\Delta t)\}$, we have $p_l = E[z(n\Delta t)\,z((n + l)\Delta t)]$. Note that from (A11), $p_l \propto 1/l^2$. Peligrad [16, Theorem 2.3] gives the conditions for a p-mixing sequence to satisfy the CLT. The condition on p_l is $\Sigma p_{2^l} < \infty$, which is clearly met here; so $\boldsymbol{Z}$ in (A9) is Gaussian when $T_o \gg T$, and (A11) holds.

Appendix D
Independence of Unknown Bit Rate Detection Test Statistics

We examine the correlation between the two Gaussian random variables

$$\boldsymbol{Z}_{C_i} = \frac{1}{T_o}\int_0^{T_o} z(t)\cos(2\pi f_i t)\,dt$$

and

$$Z_{C_l} = \frac{1}{T_o}\int_0^{T_o} z(t)\cos(2\pi f_l t)\,dt.$$

Now,

$$\begin{aligned}\left|E[Z_{C_i}Z_{C_l}]\right| &= \left|\frac{1}{T_o^2}\int_0^{T_o}\int_0^{T_o} R_{zz}(t-s)\cos(2\pi f_i t)\right.\\ &\qquad \left.\cdot \cos(2\pi f_l s)\,dt\,ds\right|\\ &= \left|\frac{1}{T_o^2}(N_0B)^2\int_0^{T_o}\int_0^{T_o}\operatorname{sinc}^2(2\pi B(t-s))\right.\\ &\qquad \left.\cdot\cos(2\pi f_i t)\cos(2\pi f_l s)\,dt\,ds\right|\\ &< \left|\frac{1}{T_o^2}(N_0B)^2\int_0^{T_o}\int_0^{T_o}\cos(2\pi f_i t)\right.\\ &\qquad \left.\cdot\cos(2\pi f_l s)\,dt\,ds\right|\\ &< \frac{1}{T_o^2}(N_0B)^2\frac{1}{(2\pi)^2}\frac{1}{f_if_l}.\end{aligned}$$

Since $f_if_l \sim B^2$, $E[Z_{C_i}Z_{C_l}] \propto 1/T_o^2$. Recalling that [cf. (42)] $E[Z_{C_i}^2] \propto B/T_o$, we see that

$$\frac{E[Z_{C_i}Z_{C_l}]}{E[Z_{C_i}^2]} \propto \frac{1/B}{T_o}.$$

By assumption, $T_o \gg 1/B$, and thus the correlation between the two Gaussian random variables is negligible. The detection statistics O_i and O_l are therefore effectively independent.

Acknowledgment

The authors would like to thank J. C. Smith for writing the computer programs that were used to generate numerical results in Section VI.

References

[1] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Vol. I.* New York: Wiley, 1968.

[2] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Math. Computation*, vol. 19, p. 297, 1965.

[3] W. A. Gardner, "Signal interception: A unifying theoretical framework for feature detection," *IEEE Trans. Commun.*, vol. 36, pp. 897-906, Aug. 1988.

[4] —, "Structural characterization of locally optimum detectors in terms of locally optimum estimators and correlators," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 924-932, Nov. 1982.

[5] W. A. Gardner and L. E. Franks, "Characterization of cyclostationary random signal processes," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 4-14, 1975.

[6] W. A. Gardner, W. A. Brown, and C.-K. Chen, "Spectral correlation of modulated signals: Part II—Digital modulation," *IEEE Trans. Commun.*, vol. COM-35, p. 595, June 1987.

[7] N. F. Krasner, "Optimal detection of digitally modulated signals," *IEEE Trans. Commun.*, vol. COM-30, pp. 885-895, May 1982.

[8] W. A. Gardner, "The role of spectral correlation in the design and analysis of synchronizers," *IEEE Trans. Commun.*, vol. COM-34, pp. 1089-1095, Nov. 1986.

[9] J. Imbeaux, "Performances of the delay-line multiplier circuit for clock and carrier synchronization in digital satellite communications," *IEEE J. Select. Areas Commun.*, vol. SAC-1, pp. 82-95, Jan. 1983.

[10] W. A. Gardner, *Introduction to Random Processes with Applications to Signals and Systems.* New York: Macmillan, 1986.

[11] R. E. Ziemer and R. L. Peterson, *Digital Communications and Spread Spectrum Systems.* New York: Macmillan, 1985, p. 396.

[12] D. Middleton, *An Introduction to Statistical Communication Theory*, reprint ed. Los Altos, CA: Peninsula, 1987, p. 205.

[13] P. Billingsley, *Probability and Measure.* New York: Wiley, 1979, p. 316.

[14] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.* New York: McGraw-Hill, 1984, p. 138.

[15] M. K. Simon *et al.*, *Spread Spectrum Communications, Vol. III.* Rockville, MD: Computer Science Press, 1985, p. 287.

[16] M. Peligrad, "Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random variables (A survey)," in *Dependence in Probability and Statistics.* Boston, MA: Birkhauser, 1985, p. 193.

[17] J. G. Proakis, *Digital Communications.* New York: McGraw-Hill, 1983, p. 210.

[18] A. J. Viterbi and J. K. Omura, *Principles of Digital Communications and Coding.* New York: McGraw-Hill, 1979, p. 124.

[19] S. Karlin and H. M. Taylor, *A First Course in Stochastic Processes.* New York: Academic, 1975, p. 479.

[20] R. Bradley, "Basic properties of strong mixing conditions," in *Dependence in Probability and Statistics.* Boston, MA: Birkhauser, 1985, p. 165.

John F. Kuehls (S'81-M'87) was born in Cleveland, OH, on September 12, 1959. He received the B.S. (high honors) and M.S. degrees in electrical engineering from the University of Akron, Akron, OH, in 1982 and 1983, respectively, and the Ph.D. degree in electrical engineering from the University of Maryland, College Park, in 1989.

In 1980 he joined the Department of Defense, Ft. Meade, MD, as a co-op student, and in 1984, he converted to a full-time employee. Since 1984 he has been involved with various research activities there, primarily in the areas of communications and signal processing. In 1987 he became a member of the part-time Faculty of George Washington University, Washington DC, where he has since taught graduate courses in communications theory, detection and estimation theory, and electromagnetic field theory.

Dr. Kuehls is a member of Tau Beta Pi, Eta Kappa Nu, Pi Mu Epsilon, and Mortar Board.

Evaggelos Geraniotis (S'76-M'82-SM'88), for a photograph and biography, see the May 1990 issue, p. 502.

Performance of Coded Direct Sequence Spread Spectrum in a Fading Dispersive Channel with Pulsed Jamming

BRANIMIR R. VOJČIĆ, MEMBER, IEEE, AND RAYMOND L. PICKHOLTZ, FELLOW, IEEE

Abstract—**This paper is concerned with the performance of a binary phase-shift-keyed (PSK) direct sequence spread spectrum (DS/SS) communication system in a fading dispersive channel with the direct or specular component of the signal present. Two types of multipath intensity profiles are analyzed. Large values of the multipath spread (larger than a symbol duration) are considered, and a useful approximation to the symbol error probability, based on random spreading sequences, is derived. The effect of pulsed jamming on the performance is analyzed. Both the uncoded and coded case are considered. In the coded case, hard and soft decision decoding metrics, with and without jammer-state information (JSI), are analyzed and discussed.**

I. Introduction

Communications over fading dispersive channels have been an area of interest for a long time. In the present paper, we analyze the performance of a DS/SS BPSK communication system, dispersive only in time (usually referred to as a frequency-selective Rician fading channel [8]–[10]). It is assumed that the direct or specular component of the signal exists.

The performance of a digital communication system is limited by the multipath spread, i.e., by the extent of the intersymbol interference imposed. It is usually assumed that the multipath spread is smaller than a symbol duration (see [2], [4], [7], and [9]) because larger values of the multipath spread cause unbearable intersymbol interference, which limits the achievable data rate. A DS/SS communication system can tolerate a higher degree of intersymbol interference due to the multipath interference rejection capability of the direct sequence decorrelation process in the receiver. It was assumed in [1] that the value of the multipath spread was an integer multiple of a symbol duration. Here we relax that constraint and allow for arbitrary values of the multipath spread.

In the present paper, we further extend our previous results in [1], in several aspects. First, in order to investigate the influence of the multipath intensity profile shape on the performance, we consider the triangular multipath intensity profile (as in [1]) and the rectangular one. It was shown in [7] that for a differential PSK communication system, the form of the multipath intensity profile alone is not relevant to the performance, but that the rms (root mean square) value of the multipath spread does effect the performance. We feel, however, that for smaller values of the multipath spread, a DS/SS system can exhibit different performances for different multipath profiles, due to the sensitivity of the system to the energy of the multipath components received within the first chip period.

It was noted in [2] that the average signal-to-noise ratio (SNR), based on the random sequence assumption, is a close approximation to the actual SNR of a DS/SS system. However, the SNR is not a very good measure of the performance; the error probability is usually required. In this paper we derive an expression for the expected value of the noise variance due to the multipath interference. From the fact that this expected value represents the asymptotic behavior for a large number of chips per symbol [3], we give arguments whether it can be used in the error probability estimation or not.

The previous performance analysis of DS/SS communication systems in fading dispersive channels were limited to short sequences [2], [4], [7]. In [1] we assumed, although it was not stated explicitly, that the period of the spreading sequence corresponded to T_m (the maximum dispersion). Here we consider spreading sequences of arbitrary lengths. In the former case, it was possible to get the exact expression for the error probability. The purpose of the random sequence assumption was to simplify the calculation and to enable an estimation of the performance independently of the type of the spreading sequence used. In the latter case, which is of interest in military communications, modeling the spreading sequence as a random process is practically the only viable alternative, in order to keep the computational burden at a reasonable level.

For the coded performance analysis, under pulsed jamming, the Chernoff bound technique is usually used. We noted in [1] that this approach gave a meaningless result regarding soft decision decoding without JSI. Namely, the bound on the probability of error was increasing uniformly with the decrease of the duty cycle of the jammer. As an attempt to overcome that problem, we use the approach suggested by Torrieri in [5], where an exact pair-

Manuscript received February 26, 1990.

B. R. Vojčić is with the Electronics and Communications Department, Ministry of Defence, Yugoslavia.

R. L. Pickholtz is with the Department of Electrical Engineering and Computer Science, George Washington University, Washington, DC 20052.

IEEE Log Number 9035818.

0733-8716/90/0600-0934$01.00 © 1990 IEEE

wise error probability was calculated instead of its Chernoff bound. We extend results of [1] and apply ideas in [5] to a consistent analysis of hard decision decoding with and without side information in fading dispersive channels. A comparison of hard and soft decision decoding is presented, as well as a comparison of the present approach with the Chernoff bound technique.

II. Evaluation of the Symbol Error Probability

The model of the communication system is taken from [1]. The data signal may be written as

$$d(t) = \sum_{-\infty}^{\infty} d_n p_T(t - nT) \tag{1}$$

where d_n is a sequence of independent, identically distributed (i.i.d.) random variables such that $\Pr\{d_n = +1\} = \Pr\{d_n = -1\} = 1/2$. $P_T(t)$ is a unit amplitude, rectangular pulse of duration T. A coded signal is formed by adding redundant symbols to the data bit sequence and by the appropriate adjustment of the coded symbol duration (if necessary). It should be noted that this does not necessarily mean an increase in the bandwidth which is determined by the chip duration T_c. If it is required to keep both the data rate and the bandwidth constant, the processing gain per symbol must be reduced. A coded signal can be modeled as

$$c(t) = \sum_{-\infty}^{\infty} c_n p_{T_s}(t - nT_s) \tag{2}$$

where c_n has the same probability distribution as d_n. T_s is the coded symbol duration.

It is required in military communications, for security reasons, that a pseudorandom (PR) spreading sequence should not repeat for a long period of time. Also, good cryptographic characteristics of spreading sequences for use in an antijamming system are important to prevent the efficiency of repeat-back jammers. If a repeat-back jammer does not know *a priori* a current symbol spreading signal (even assuming the possibility of detecting all incoming chips with a zero processing delay), its effectiveness is constrained by the chip duration T_c and the geometry of the problem (relative position of the jammer to the communicators). If the jamming signal is delayed by more than T_c, relative to the desired signal, decorrelation in the receiver takes place.

The only viable alternative to evaluate the performance of a DS/SS system with a very long period of PR sequence, at least in the initial system design, is to model this sequence as a random process. For that purpose, we assume N chips of duration T_c per coded symbol and the overall sequence as a set of M different random sequences of length N. Namely, we define a set $B = \{\boldsymbol{b}_0, \boldsymbol{b}_1, \cdots, \boldsymbol{b}_{M-1}\}$, where $\boldsymbol{b}_i$ is a sequence of i.i.d. random variables $b_i(j)$, such that $\Pr\{b_i(j) = +1\} = \Pr\{b_i(j) = -1\} = 1/2, j = 0, 1, \cdots, N-1$. Hence, the ith spreading waveform (corresponding to the ith coded symbol) may be written as

$$b_i(t) = \sum_{j=0}^{N-1} b_i(j)\, p_{T_c}(t - iT_s - jT_c). \tag{3}$$

The transmitted signal is

$$s(t) = Ac(t)\, b(t) \cos \omega_0 t \tag{4}$$

where A is the signal amplitude, ω_0 is the carrier frequency, and $c(t)$ is the coded bit sequence. The received signal may be written as in [1] and [4]

$$r(t) = \mathrm{Re}\left\{\left[\alpha A c(t)\, b(t) + \xi A \int_{-\infty}^{\infty} \beta(\tau)\, c(t-\tau) \cdot b(t-\tau)\, d\tau\right] \exp(j\omega_0 t)\right\} + j(t) \tag{5}$$

where α and ξ are real constants representing the attenuation of the channel for the direct/specular path and the average attenuation of the channel for the continuum of multipath components, respectively. The background noise is neglected and $j(t)$ represents the noise jamming signal. The assumed WSSUS (wide-sense stationary, uncorrelated scattering) fading channel dispersive only in time is used in [1]–[3] and described in more detail in [8]–[10]. The low-pass equivalent impulse response of the channel, $\beta(\tau)$, is assumed to be a zero-mean, complex, wide-sense stationary Gaussian random process, such that

$$\tfrac{1}{2}E[\beta(\tau_1)\, \beta^*(\tau_2)] = \zeta(\tau_1)\, \delta(\tau_1 - \tau_2) \tag{6}$$

where $\delta(t)$ is the Dirac delta function and $\zeta(t)$ is the multipath intensity profile.

To obtain specific results, two types of multipath intensity profiles are considered, the rectangular one

$$\zeta(\tau) = \frac{1}{2T_s}, \qquad \text{for } |\tau| \le T_m \tag{7}$$

and the triangular one

$$\zeta(\tau) = \frac{1}{T_s}\left(1 - \frac{|\tau|}{T_m}\right), \qquad \text{for } |\tau| \le T_m \tag{8}$$

where T_m denotes the maximum dispersion. Sometimes, the rms value of the multipath spread is used, and for a given intensity profile is monotonically referred to T_m. For the accepted definitions of $\zeta(\tau)$ in (7) and (8), the corresponding rms values of the multipath spread are $\sqrt{T_m^3/3T_s}$ and $\sqrt{T_m^3/6T_s}$, respectively.

The output of the integrate-and-dump filter is given by

$$g(T_s) = S(T_s) + S_f(T_s) + J(T_s) \tag{9}$$

where $S(T_s) = A_0 c_0 T_s$, with $A_0 = \alpha A$ and c_0 representing the current coded symbol. Modeling the jamming signal as white Gaussian noise with two-sided power spectral density $N_j/2$ in the SS bandwidth, $J(T_s)$ is a zero-mean Gaussian random variable with the variance $N_j T_s$. The component of the detected signal due to fading, $S_f(T_s)$, is a zero-mean Gaussian random variable given by

$$S_f(T_s) = \mathrm{Re}\left[A_1 \int_0^{T_s} \int_{-\infty}^{\infty} \beta(\tau)\, c(t-\tau) \cdot b(t-\tau)\, b(t)\, d\tau\, dt\right] \tag{10}$$

where $A_1 = \xi A$. Following the approach taken in [1], the variance of $S_f(T_s)$, conditioned on the coded symbol sequence and the corresponding set of spreading sequences, is given by

$$\sigma_f^2(\boldsymbol{c}_0, \boldsymbol{B}_0) = A_1^2 \left\{ \sum_{k=0}^{K-1} \sum_{n=0}^{N-1} \int_{kT_s+nT_c}^{kT_s+(n+1)T_c} \zeta(\tau)\, \Gamma_{kn}(\tau)\, d\tau + \sum_{n=0}^{L-1} \int_{KT_s+nT_c}^{KT_s+(n+1)T_c} \zeta(\tau)\, \Gamma_{Kn}(\tau)\, d\tau \right\} \tag{11}$$

where we assumed $T_m = KT_s + LT_c$. K and L are assumed to be positive integers such that $K \geq 0$ and $0 \leq L < N$.

$$\Gamma_{kn}(\tau) = \Xi_{kn}(\tau) + c_{-k}c_{-k-1}\Upsilon_{kn}^{-}(\tau) + c_k c_{k+1}\Upsilon_{kn}^{+}(\tau). \tag{12}$$

$\boldsymbol{c}_0 = (c_{-K-1}, \cdots, c_{-1}, c_0, c_1, \cdots, c_{K+1})$ in (11) corresponds to the window of the coded symbol sequence relative to the current symbol c_0, and $\boldsymbol{B}_0$ stands for the corresponding subset of B. We use the subscript 0 to denote the current coded symbol; actually the ith coded symbol during a communication session. Ξ, Υ^-, and Υ^+ are defined by

$$\Xi_{kn}(\tau) = R_{0,-k-1}^2(\tau) + \hat{R}_{0,-k}^2(\tau) + R_{0,k}^2(\tau) + \hat{R}_{0,k+1}^2(\tau) \tag{13}$$

$$\Upsilon_{kn}^{-}(\tau) = 2R_{0,-k-1}(\tau)\, \hat{R}_{0,-k}(\tau) \tag{14}$$

$$\Upsilon_{kn}^{+}(\tau) = 2R_{0,k}(\tau)\, \hat{R}_{0,-k+1}(\tau) \tag{15}$$

for $\tau = kT_s + nT_c + \gamma$, where $0 < \gamma < T_c$. Further, the continuous-time partial cross-correlation functions are defined by

$$R_{0,k}(\tau) = (T_c - \gamma)C_{0,k}(n) + \gamma C_{0,k}(n+1)$$
$$\hat{R}_{0,k}(\tau) = (T_c - \gamma)\hat{C}_{0,k}(n) + \gamma \hat{C}_{0,k}(n+1) \tag{16}$$

where the discrete aperiodic cross-correlation functions are defined by

$$C_{0,k}(n) = \sum_{i=0}^{n-1} b_0(i)\, b_k(N-n+i) \tag{17}$$

$$\hat{C}_{0,k}(n) = \sum_{i=n}^{N-1} b_0(i)\, b_k(i-n). \tag{18}$$

The symbol error probability conditioned on the $\boldsymbol{c}_0$ and the corresponding window of B, say, $\boldsymbol{B}_0 = (\boldsymbol{b}_{-K-1}, \cdots, \boldsymbol{b}_0, \cdots, \boldsymbol{b}_{K+1})$, is

$$P_e(\boldsymbol{c}_0, \boldsymbol{B}_0) = Q\left(\sqrt{\frac{(A_0 T_s)^2}{N_j T_s + \sigma_f^2(\boldsymbol{c}_0, \boldsymbol{B}_0)}}\right) \tag{19}$$

with $\sigma_f^2(\boldsymbol{c}_0, \boldsymbol{B}_0)$ given in [11]. $Q(x)$ denotes the Gaussian probability integral. Averaging (19) over all possible spreading sequences $i = 0, 1, \cdots, M-1$ and all possible combinations of coded symbols, we get

$$\overline{P}_e = \overline{P_e(\boldsymbol{c}_0, \boldsymbol{B}_0)} \tag{20}$$

as an average symbol error probability, which can be evaluated numerically. In the uncoded case, (20) represents the average bit-error-rate (BER).

As stated previously, M, the cardinality of the set B, is a very large number in military communications, and the evaluation of (20) is impractical. In order to circumvent this problem and get an estimate of the performance, we suggest an approximation to P_e based on the mathematical expectation of σ_f^2. Strictly speaking, it is mathematically incorrect because $Q(x)$ is a nonlinear function of x, i.e., $E[Q(x)] \neq Q[E(x)]$. However, in [2] it was verified numerically for some "short sequences" (the period of the spreading sequence corresponds to the symbol duration) that the difference between the actual SNR and the average SNR (based on $E[\sigma_f^2]$) lies within 0.05 dB. Both the multipath and multiple-access interference were considered. Also, it was claimed that for some sequences and channel parameters, the discrepancy can be as large as 3 dB.

It is well known that the multipath interference rejection capability can be intensified by choosing spreading sequences with good discrete aperiodic autocorrelation function properties [2]. For the range of the maximum dispersion values we consider, it is desired that every sequence in the set B has good discrete aperiodic cross-correlation function properties with $K + 1$ preceding and $K + 1$ succeeding sequences, too.

To evaluate the mathematical expectation of (11), we recall the expressions for the mathematical expectation of the discrete aperiodic autocorrelation function given in [1] and [2]. Similarly, for the discrete aperiodic cross-correlation function, we have

$$E[C_{0,k}(n)] = E[\hat{C}_{0,k}(n)] = 0, \quad \text{for any } n \tag{21}$$

$$E[C_{0,k}^2(n)] = \begin{cases} n, & \text{for } n \leq N-1 \\ 0, & \text{for } n \geq N \end{cases} \tag{22}$$

$$E[\hat{C}_{0,k}^2(n)] = \begin{cases} N-n, & \text{for } n \leq N-1 \\ 0, & \text{for } n \geq N. \end{cases} \tag{23}$$

Now, after some straightforward calculations, the expression for $E[\sigma_f^2]$ for the rectangular multipath intensity profile becomes

$$E[\sigma_{fr}^2] = A_1^2 T_s^2 I_r = A_1^2 T_s^2 \frac{2}{3N}\left[K + \frac{L}{N} + \frac{1}{2}\left(1 - \frac{1}{N}\right)\right] \cong A_1^2 T_s^2 \frac{2}{3N}\left(K + \frac{L}{N} + \frac{1}{2}\right) \tag{24}$$

and

$$E[\sigma_{ft}^2] = A_1^2 T_s^2 I_t = A_1^2 T_s^2 \frac{2}{3N} \cdot \left[K + \frac{L}{N} + 1 - \frac{(1 - 1/N)}{4(KN + L)} \right] \cong A_1^2 T_s^2 \frac{2}{3N} \left(K + \frac{L}{N} + 1 \right) \qquad (25)$$

for the triangular multipath intensity profile. The operator $\cong$ in (24) and (25) is valid for $T_m >> T_c$ and $N >> 1$. Hence, the approximation to the average symbol error probability is

$$P_A = Q\left(\sqrt{\frac{2a}{1 + 2ayI}} \right) \qquad (26)$$

where $a = E_s/N_j$ and $E_s = A_0^2 T_s/2$ represents the direct path symbol energy. $y = A_1^2/A_0^2$ corresponds to the ratio of received energy due to the dispersive component to received energy due to the direct signal path. Actually, (26) represents an asymptotic behavior of the symbol error probability for large N.

To illustrate the usefulness of P_A as an approximation to the average symbol error probability, a set of maximum length shift register sequences or m-sequences is considered—of lengths 31, 255, and 4095—generated by shift registers with feedback connections to the first stage at stages (5, 2), (8, 4, 3, 2), and (12, 6, 4, 1), respectively [16]. In Fig. 1 we compare the performance of the system which uses short m-sequences for spreading to P_A for random sequences. It can be noted that the discrepancy between them increases with the increase of the maximum dispersion, as noted in [2], and the m-sequences chosen at chance exhibit better performances.

It is also seen that the difference between P_A and the actual error probability becomes negligible with the increase of the processing gain. In Fig. 2 we evaluate numerically the average BER using (11)-(20), for $N = 31$, where averaging is over 130 subsequences (of length 31) of the m-sequence of the length 4095. It is evident that P_A is a close approximation to (20), in this particular case. The approximation improves with the increase of T_m. Also, the actual average BER is again smaller than P_A.

Although the number of subsequences is not comparable to the one of practical interest, the considered example gives us an indication of the BER behavior. In effect, for a very large set of subsequences, using the arguments mentioned previously, we can expect performance even closer to P_A, due to the averaging process which takes place. Namely, the contribution of a few subsequences in the set, with poor discrete aperiodic correlation function properties, to the average BER, would become negligible. In Figs. 1 and 2, the rectangular multipath intensity profile is considered, with $a = 10$ dB and $y = 1$.

In Fig. 3 we compare the irreducible BER (the BER in the limit as $E_b/N_j \rightarrow \infty$) versus the maximum dispersion for two types of the multipath intensity profiles defined by (7) and (8). It is evident that the difference can be one order of magnitude for smaller values of T_m, and that it becomes negligible for larger values of T_m. The same conclusion follows from (24) and (25) where we see that the difference between the corresponding variances can be as large as 3 dB, for $K = 0$ and $L << N$. From Fig. 3, we also see how the limiting performance, i.e., the irreducible BER, can be improved by doubling the processing gain. Again, the direct and specular components have the same amplitudes, i.e., $y = 1$.

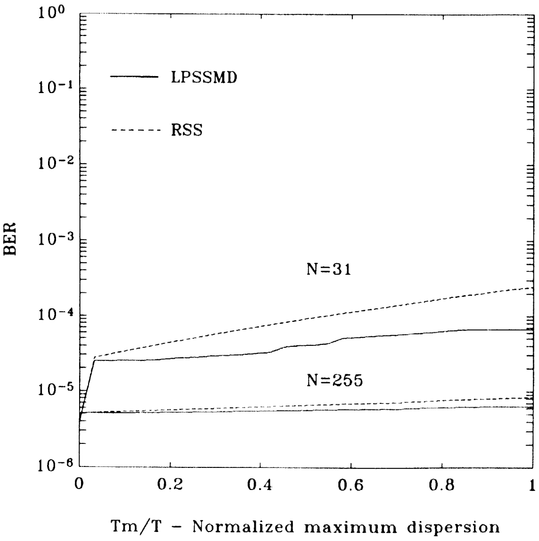


Fig. 1. Comparison of BER for random spreading sequences and m-sequences; $a = 10$ dB.

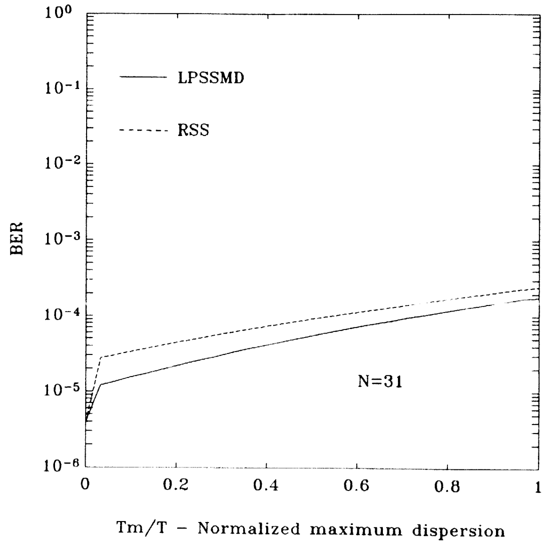


Fig. 2. Comparison of actual BER and P_A, for long spreading sequences; $a = 10$ dB, $y = 11$.

For pulsed jamming analysis, we assume an average power constrained jammer. Hence, the one-sided power density of the jamming signal is defined by

$$N_j = \begin{cases} N_j/\rho & \text{with probability } \rho \\ 0 & \text{with probability } (1 - \rho) \end{cases} \qquad (27)$$

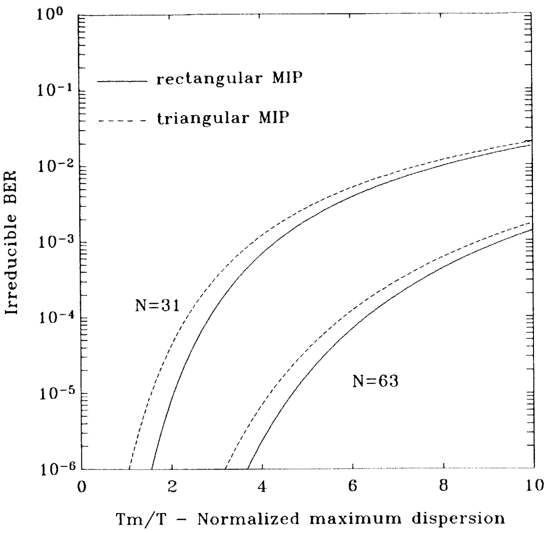


Fig. 3. Effect of the multipath intensity profile on irreducible BER; $y = 1$.

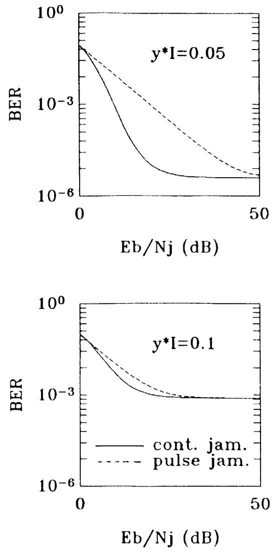


Fig. 4. Comparison of the pulsed jamming effects on BER with the increase of the multipath interference.

where ρ represents the duty cycle of the jammer, i.e., the probability that a symbol is jammed. From (26) and (27), the probability of error may be written as

$$P_A = \rho Q\left(\sqrt{\frac{2\rho a}{1 + 2\rho a y I}}\right) + (1 - \rho) Q\left(\sqrt{\frac{1}{yI}}\right) \quad (28)$$

which is of the same form as (21) in [1], except for $2I$ instead of I in the denominator because of the normalization in (6).

The value of ρ which maximizes (28) is approximately the one which maximizes the first term on the right-hand side. From [11], it follows that the optimal value of ρ, ρ^*, is given by

$$\frac{\rho^* a}{1 + 2\rho^* a y I} = 0.709 \quad (29)$$

i.e.,

$$\rho^* = \begin{cases} \dfrac{0.709}{a(1 - 1.418yI)}, & \text{for } yI < 1/1.418 \\ 1, & \text{otherwise.} \end{cases} \quad (30)$$

Equation (30) gives the conditions when a pulse jammer can optimize its strategy against communicators. It is a similar, but more precise, result as compared to the one obtained in [1], where we used an exponential bound on $Q(x)$ instead. From Fig. 4 we see that the difference in the performance for continual and pulsed jamming becomes negligible with the increase of yI.

III. Coded Performance

In the decoding process, a metric is used to decide which coded sequence $\boldsymbol{c}$ was transmitted, given the channel output sequence $\boldsymbol{g}$ and possibly side information $\boldsymbol{z}$. Assuming, for the moment that there are only two possible sequences which differ in i symbols, say $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$, and that $\boldsymbol{c}$ has been transmitted, the probability of incorrectly deciding $\hat{\boldsymbol{c}}$ is

$$P(\boldsymbol{c} \rightarrow \hat{\boldsymbol{c}} \mid i) = \Pr\left\{ \sum_{n=1}^{i} \left[m(g_n, \hat{c}_n, z_n) - m(g_n, c_n, z_n) \right] \geq 0 \mid \boldsymbol{c}, i \right\} \quad (31)$$

which is called the pairwise-error-probability [6]. In what follows, we use the maximum likelihood metric defined as the logarithm of the channel likelihood function [6]

$$m(\boldsymbol{g}, \boldsymbol{c}, \boldsymbol{z}) = \log p(\boldsymbol{g} \mid \boldsymbol{c}, \boldsymbol{z}) \quad (32)$$

from which various hard and soft decision decoding metrics could be derived. It is assumed that the channel is memoryless, which can be provided through adequate interleaving and deinterleaving. A straightforward evaluation of (32) gives, for soft decision decoding [6, vol. I], [10]

$$m(g_n, c_n, z_n) = w_s(z_n)\, g_n(T_s) c_n \quad (33)$$

and for hard decision decoding [6, vol. I]

$$m(g_n, c_n, z_n) = \begin{cases} w_h(z_n) & \text{if } \bar{g}_n = c_n \\ 0 & \text{if } \bar{g}_n = \hat{c}_n \end{cases} \quad (34)$$

where $\bar{g}$ denotes a binary quantized output of the integrate-and-dump filter and $w(z)$ corresponds to the weighting function to be defined. If a receiver has no available side information, in our case the jammer state information, without the loss of generality, we assume

$$w(z = 0) = w(z = 1) = 1, \quad (35)$$

i.e., that the jammed and unjammed metric values are weighted equally. $z = 0$ and $z = 1$ correspond to the jammer "off" and "on" states, respectively. With JSI available, the unjammed metric values are weighted more than the jammed ones. Taking the logarithm of the likelihood function, it turns out that for soft decision decoding [6], [10]

$$w_s(z) = \begin{cases} 1/\sigma_f^2, & z = 0 \\ 1/(\sigma_f^2 + N_j T_s/\rho), & z = 1 \end{cases} \tag{36}$$

and for hard decision decoding [6, vol. II]

$$w_h(z) = \begin{cases} \log\,(1 - \epsilon_0)/\epsilon_0, & z = 0 \\ \log\,(1 - \epsilon_1)/\epsilon_1, & z = 1 \end{cases} \tag{37}$$

where ϵ_0 and ϵ_1 correspond to the channel transition probabilities [10] for a binary symmetric channel, for $z = 0$ and $z = 1$ states, respectively. It should be noted that the weighting functions depend on the knowledge of the fading channel parameters. The background noise is neglected. While the temporal variations of σ_f^2 depend on both the multipath profile and the coded symbol sequence, for the purpose of the present analysis we use $E[\sigma_f^2]$ which depends only on the channel and can be estimated by means of the channel probe pulse measurements. Besides, for the pulsed jamming model we consider, it is appropriate to assume that the temporal variations of σ_f^2 are negligible as compared to the jamming power level variations. Hence, the approximation to the pairwise error probability, based on $E[\sigma_f^2]$, can be considered more confident than the approximation in (26).

We consider a system which uses an interleaving scheme to prevent burst errors. Consequently, it is appropriate to assume that the probability that a symbol is jammed is independent of the rest of the coded symbol sequence. Also, the jamming model defined by (27) assumes no partial jamming of a symbol. With these assumptions, the approximation to the pairwise error probability may be written as [5]

$$\overline{P_A(c \rightarrow \hat{c} \mid i)} = \sum_{j=0}^{i} \binom{i}{j} \rho^j (1 - \rho)^{i-j}\, \mathrm{P}(c \rightarrow \hat{c} \mid i, j) \tag{38}$$

where $P_A(c \rightarrow \hat{c} \mid i, j)$ corresponds to the pairwise error probability, given j out of i coded symbols were jammed. Now, we derive $P_A(c \rightarrow \hat{c} \mid i, j)$ for the metrics defined.

1) Hard Decision Decoding Without Side Information: From (31), (34), and (35)

$$P_A(c \rightarrow \hat{c} \mid i, j) = \Pr\left\{ \sum_{n=1}^{j} x_{n|z_n=1} + \sum_{n=j+1}^{i} x_{n|z_n=0} \geq 0 \,\middle|\, \boldsymbol{c}, i, j \right\} \tag{39}$$

where $x_n = m(g_n, \hat{c}_n, z_n) - m(g_n, c_n, z_n)$ is a random variable, such that

$$x_{n|z_n=1} = \begin{cases} +1 & \text{with probability } \epsilon_1 \\ -1 & \text{with probability } 1 - \epsilon_1 \end{cases} \tag{40}$$

and similarly for $x_{n|z_n=0}$. Using the change of variables $x_n = 1 - 2y_n$, where y_n is a binomially distributed random variable, such that

$$y_{n|z_n=1} = \begin{cases} +1, & \text{with probability } 1 - \epsilon_1 \\ 0, & \text{with probability } \epsilon_1. \end{cases} \tag{41}$$

Further, (39) reduces to

$$P_A(c \rightarrow \hat{c} \mid i, j) = \sum_{k=0}^{j} \binom{j}{k} (1 - \epsilon_1)^k \epsilon_1^{j-k} P' \tag{42}$$

where the probability P' is given by

$$\begin{aligned} P' &= \Pr\left\{ \sum_{n=j+1}^{i} x_{n|z_n=0} \geq 2k - j \,\middle|\, \boldsymbol{c}, i, j, k \right\} \\ &= \Pr\left\{ \sum_{n=j+1}^{i} y_{n|z_n=0} \leq \tfrac{1}{2}(i - 2k) \,\middle|\, \boldsymbol{c}, i, j, k \right\}. \end{aligned} \tag{43}$$

When $\Sigma_{n=1}^{i} [m(g_n, \hat{c}_n, z_n) - m(g_n, c_n, z_n)] = 0$, in (32), the decoder decides in favor of either sequence with equal probability, and it follows

$$P' = \begin{cases} \displaystyle\sum_{m=0}^{lm} \binom{i-j}{m} (1 - \epsilon_0)^m \epsilon_0^{(i-j-m)}, \\ \qquad \text{for } lm < q < lm + 1 \\ \displaystyle\sum_{m=0}^{lm-1} \binom{i-j}{m} (1 - \epsilon_0)^m \epsilon^{(i-j-m)} \\ \displaystyle\quad + 0.5 \binom{i-j}{l_m} (1 - \epsilon_0)^{lm} \epsilon_0^{(i-j-lm)}, & q = lm \end{cases} \tag{44}$$

where $q = 0.5(i - 2k)$ and lm is an integer. For this particular case, the first and the second row of (44) correspond to i odd and even, respectively.

2) Hard Decision Decoding with Side Information: From (31) and (34),

$$P_A(c \rightarrow \hat{c} \mid i, j) = \Pr\left\{ w_h(1) \sum_{n=1}^{j} x_{n|z_n=1} + w_h(0) \sum_{n=j+1}^{i} x_{n|z_n=0} \geq 0 \,\middle|\, \boldsymbol{c}, i, j \right\} \tag{45}$$

which reduces to (42) and (44) with

$$q = \frac{1}{2}\left[i - j + \frac{w_h(1)}{w_h(0)} (j - 2k) \right]. \tag{46}$$

3) Soft Decision Decoding Without Side Information: From (9), (31), and (33),

$$P_A(c \rightarrow \hat{c} \mid i, j) = \Pr\left\{ -\sum_{n=1}^{j} y_{n|z_n=1} - \sum_{n=j+1}^{i} y_{n|z_n=0} \geq iA_0 T_s \,\middle|\, \boldsymbol{c}, i, j \right\} \tag{47}$$

where $E[y_n] = 0$ and

$$E[y_n^2] = \begin{cases} \sigma_f^2, & \text{for } z = 0 \\ \sigma_f^2 + N_j T_s/\rho, & \text{for } z = 1. \end{cases} \qquad (48)$$

Hence, (47) reduces to

$$P_A(c \rightarrow \hat{c} \mid i, j) = Q\left(\sqrt{\frac{2i^2 a\rho}{j + 2ia\rho\gamma I}}\right). \qquad (49)$$

4) Soft Decision Decoding with Side Information: From (9), (31), and (34),

$$\begin{aligned} P_A(c \rightarrow \hat{c} \mid i, j) &= \Pr\left\{ -w_s(1) \sum_{n=1}^{j} y_{n|z_n=0} - w_s(0) \sum_{n=j+1}^{i} y_{n|z_n=0} \right. \\ &\left. \geq A_0 T_s [j w_s(1) + (i - j) w_s(0)] \mid c, j \right\} \\ &= Q\left(A_0 T_s \sqrt{j w_s(1) + (i - j) w_s(0)}\right) \qquad (50) \end{aligned}$$

and substituting (36) it reduces to

$$P_A(c \rightarrow \hat{c} \mid i, j) = Q\left(\sqrt{\frac{1}{\gamma I}\left[i - j(1 + 2a\rho\gamma I)^{-1}\right]}\right). \qquad (51)$$

The union bound on the bit error probability, using convolutional codes [5], is

$$P_{Ab} = \sum_{i=d_f}^{\infty} a(i)\, \overline{P_A(c \rightarrow \hat{c} \mid i)} \qquad (52)$$

where d_f corresponds to the minimum free distance of the code. $a(i)$'s are the coefficients depending on the code used. For the usual Odenwalder code of the constraint length 7 and rate 1/2, $a(i)$'s are given by [13]

$$a(10) = 36; \quad a(12) = 211; \quad a(14) = 1404; \quad a(16) = 11633 \cdots . \qquad (53)$$

It should be noted that in the computation of (52), we used only the first 8 terms. It gives a bit more optimistic result for higher values of the error probability (say, BER $> 10^{-3}$) because of the truncation error. On the other hand, the union bound partly compensates for the truncation effect because it shifts the BER curve in the opposite direction [5]. To get a more reliable quantitative measure of the performance, a larger set of the coefficients $a(i)$ can be used (e.g., 18 $a(i)$'s are listed in [15]).

In Fig. 5, we compare the uncoded performance for two values of the processing gain per information bit. The continual jamming case is considered. It is assumed that the information rate and SS bandwidth are kept fixed. The irreducible BER is based on P_A given by (26) for the rectangular multipath intensity profile. It is evident that coding approximately doubles the multipath spread which a system can withstand, without increasing the SS bandwidth or decreasing the information rate, for the error rates of practical interest. However, with the increase of T_m, after some point, the uncoded system outperforms the coded one. This is a consequence of the increased sensitivity of the system to the reduced symbol energy and the relative increase of the intersymbol interference. If larger values of T_m are expected, and higher error rates can be tolerated, say higher than 10^{-2}, the uncoded system may be preferred.

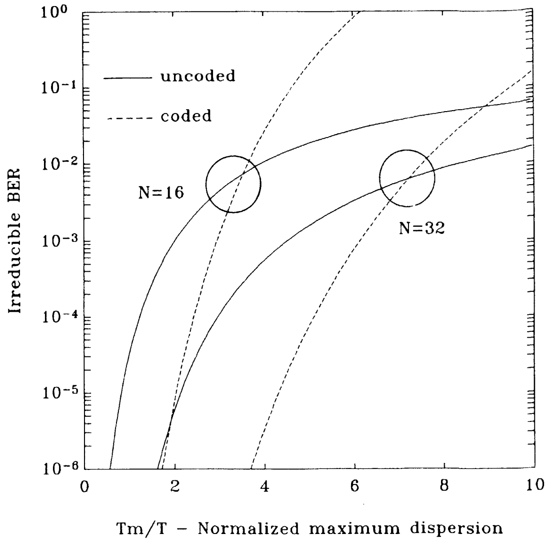


Fig. 5. Effect of coding on irreducible BER; $\gamma = 1$, fixed data rate, $r = \frac{1}{2}$, and $K = 7$ conv. code.

To analyze how a pulse jammer can optimize its strategy against the coded system in the assumed channel, we evaluate (52) for various metrics considered so far. In Fig. 6 we compare the performance of hard decision decoding with and without side information. It is a bit surprising that for values of ρ close to 1, hard decision decoding with side information exhibits a very poor performance. This is a consequence of the increased sensitivity of the metric defined by (37) to the variations of the jamming power levels. Similarly, it was noted in [5] that erasure decoding was not very useful for the values of duty cycle ρ close to 1. Actually, the metric in (37) underweights too many symbols for ρ close to 1 (e.g., $\rho = 0.5$), although the decoder can extract true information from a symbol, with some probability, without using the side information. As a consequence, the decoder cannot decide the true sequence. For smaller values of ρ, hard decision decoding with side information outperforms hard decision decoding without side information. However, the improvement does not justify the implementation complexity, and generally, hard decision decoding without side information exhibits more robust behavior.

The performance of soft decision decoding against pulsed jamming is examined in Fig. 7 for various SNR ratios. As noted in [1] and elsewhere (e.g., in [6]), soft decision decoding without JSI is very vulnerable to pulsed jamming because a few jammed symbols may dominate the metric. Soft decision decoding with JSI significantly

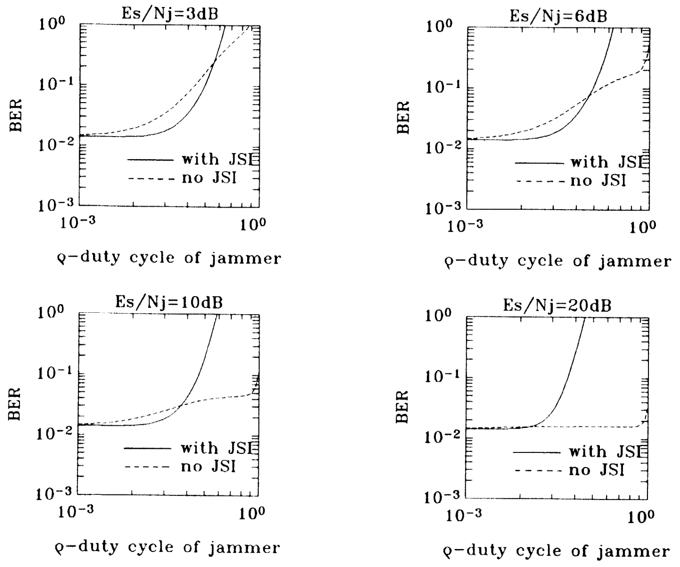


Fig. 6. Comparison of hard decision decoding with and without JSI, against pulsed jamming; $\gamma l = 0.4$.

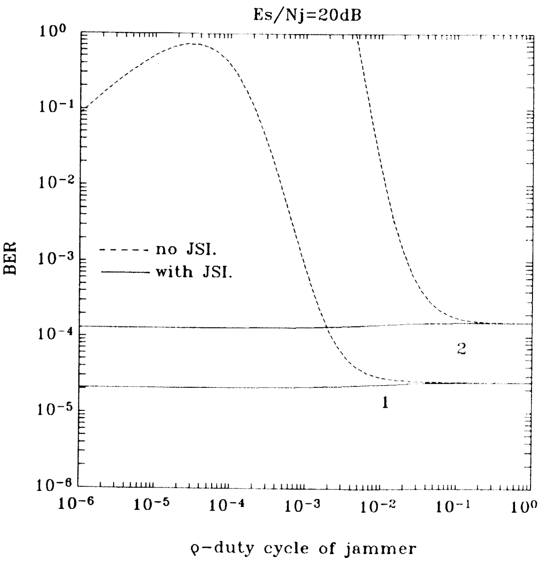


Fig. 8. Soft decision decoding with and without JSI, against pulsed jamming; $\gamma l = 0.4$. 1—present approach; 2—Chernoff bound.

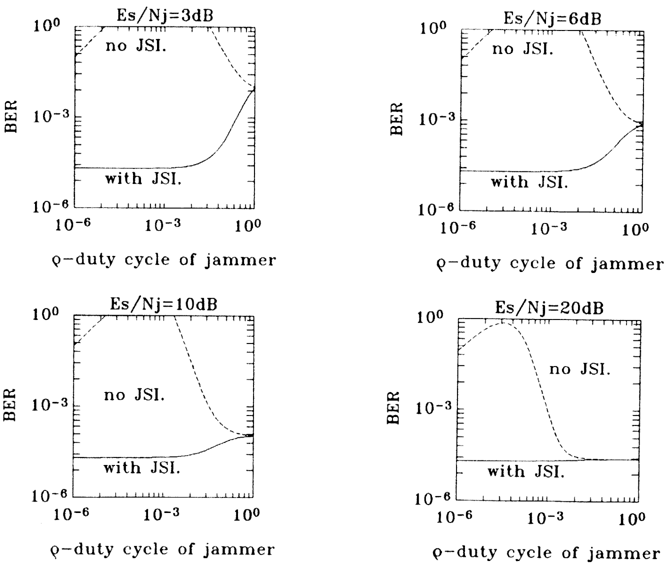


Fig. 7. Comparison of soft decision decoding with and without JSI, against pulsed jamming; $\gamma l = 0.4$.

outperforms both soft decision decoding without JSI and hard decision decoding.

Soft decision decoding without JSI exhibits a maximum BER for a certain value of ρ, and then shows an improvement with a further decrease of ρ. It is a logical result because $\rho \rightarrow 0$ corresponds to the no-jamming case. Thus, the approach taken in this paper gives a more reliable estimate of the performance under pulsed jamming than the one used in [1]. To further justify the usefulness of the suggested approach, in Fig. 8 we compare the results regarding soft decision decoding without JSI, as given by (49) and (52), to the estimate of the performance based on the Chernoff bound technique. Although both techniques give similar general behavior trends, the Chernoff bound technique gives very pessimistic results, which can be quite misleading for some values of ρ. More specifically, at $\rho = 10^{-2}$, for example, the present approach shows little degradation due to pulsed jamming, while the Chernoff bound indicates useless performance.

IV. Conclusions

In this paper, a simple closed-form expression for the expected value of the noise variance due to multipath fading is derived. The approximation to the error probability, based on the expected value of σ_f^2, represents an asymptotic behavior for large N or achievable performance for smaller values of N. The approximation derived is useful in the initial system design, particularly for long sequences which are of interest in military communications. Also, this approximation is especially appropriate for pulse jamming analysis, assuming that jamming power level variations are more dominant than the noise variations due to the multipath interference.

It is evident that the shape of the multipath intensity profile, for larger values of the maximum dispersion, does not effect the performance, which is in the agreement with the results in [7]. For smaller values of the maximum dispersion, however, a DS/SS BPSK communication system exhibits sensitivity to the energy of the multipath components received within the first chip period. Consequently, corresponding BER curves, for the forms of the multipath profiles considered, can differ by more than one order of magnitude.

An extensive and consistent analysis and comparison of different decoding metrics under pulsed jamming is presented. The pairwise error probability calculation for hard decision decoding is original to our best knowledge. It is a bit surprising that hard decision decoding without JSI outperforms hard decision decoding with JSI. A discussion of this curiosity is offered. For smaller values of the jammer's duty cycle, there is a small advantage of hard decision decoding with JSI. Although these two decoding metrics could be combined to make the optimal use of each one, the implementation simplicity favors hard decision decoding without JSI. Further, it is demonstrated that soft decision decoding with JSI offers a significant performance advantage over other analyzed metrics. It is also shown that the Chernoff bound technique can give very pessimistic estimates of the performance as compared to the present approach.

REFERENCES

[1] B. R. Vojčić and R. L. Pickholtz, "Performance of direct sequence spread spectrum in a fading dispersive channel with jamming," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 561-568, May 1989.

[2] D. E. Borth and M. B. Pursley, "Analysis of direct-sequence spread-spectrum multiple-access communication over Rician fading channels," *IEEE Trans. Commun.*, vol. COM-27, pp. 1566-1577, Oct. 1979.

[3] H. F. A. Roefs and M. B. Pursley, "Correlation parameters of random binary sequences," *Electron. Lett.*, vol. 13, pp. 488-489, Aug. 1977.

[4] L. B. Milstein and D. L. Schilling, "Performance of a spread spectrum communication system operating over a frequency-selective fading channel in the presence of tone interference," *IEEE Trans. Commun.*, vol. COM-30, pp. 240-247, Jan. 1982.

[5] D. Torrieri, "The performance of five different metrices against pulsed jamming," *IEEE Trans. Commun.*, vol. COM-34, pp. 200-204, Feb. 1986.

[6] M. K. Simon *et al.*, *Spread Spectrum Communications.* Rockville, MD: Computer Science Press, 1985.

[7] F. D. Garber and M. B. Pursley, "Performance of differentially coherent digital communications over frequency-selective fading channels," *IEEE Trans. Commun.*, vol. 36, pp. 21-31, Jan. 1988.

[8] R. S. Kennedy, *Fading Dispersive Communication Channels.* New York: Wiley, 1969.

[9] M. Schwartz *et al.*, *Communication Systems and Techniques.* New York: McGraw-Hill, 1966.

[10] J. G. Proakis, *Digital Communications.* New York: McGraw-Hill, 1983.

[11] R. J. McEliece and W. E. Stark, "An information theoretic study of communication in the presence of jamming," in *Proc. ICC'81*, pp. 45.3.1-45.3.5.

[12] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding.* New York: McGraw-Hill, 1979.

[13] J. P. Odenwaldwer, "Optimal decoding of convolutional codes," Ph.D. dissertation, UCLA, 1970.

[14] G. R. Cooper and C. D. McGillem, *Modern Communications and Spread Spectrum.* New York: McGraw-Hill, 1986.

[15] J. Conan, "The weight spectra of some short low-rate convolutional codes," *IEEE Trans. Commun.*, vol. COM-32, pp. 1050-1053, Sept. 1984.

[16] R. C. Dixon, *Spread Spectrum Systems.* New York: Wiley, 1976.

Branimir R. Vojčić (M'87) was born in Belgrade, Yugoslavia, in 1955. He received the Diploma in electrical engineering, and the M.S. and Ph.D. degrees in 1980, 1986, and 1989, respectively, from the University of Belgrade, Yugoslavia. He spent one year at the George Washington University, Washington, DC, through a continuing education program.

He currently works in the Electronics and Communications Department, Ministry of Defence, Yugoslavia, being involved in various aspects of military communications. His current research interests include communication theory, spread spectrum communications, communication networks, and frequency management.

Mr. Vojčić is a member of the IEEE Communications Society and is active in the Yugoslav Committee for CCIR.

Raymond L. Pickholtz (S'54-M'60-SM'77-F'82), for a photograph and biography, see the May 1990 issue, p. 488.

Call for Papers

IEEE Journal on Selected Areas in Communications

Congestion Control in High-Speed Packet Switched Networks

THERE has recently been substantial interest in Congestion Control in High-Speed Packet Switched Networks. In a high-speed environment, the high rate of the communication links combined with varied nature of the carried traffic make the traditional schemes inappropriate. Therefore, simpler and more efficient schemes have been proposed to fully exploit the large available bandwidth.

An issue of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS is being prepared to address the issues directly related to the above-mentioned topic. The issue will deal with performance analysis, algorithm design, and implementation methods associated with congestion control in high-sped packet switched networks. Topics of interest include (but not limited to) the following:

- Congestion Avoidance
  —admission control;
  —bandwidth enforcement;
  —routing;
  —scheduling (bandwidth allocation, buffer management).
- Reactive Congestion Control
  —network control based on feedback information;
  —effect of delayed feedback;
- Fairness and resource allocation;
- Load prediction and estimation.

Prospective authors should send five (5) copies of their manuscripts to either Dr. K. Sohraby or Prof. L. Fratta (addresses listed below) by *September 1, 1990.* The relevant dates are as follows:

Manuscript Submission Date:	September 1, 1990
Acceptance Notification:	March 1, 1991
Final Manuscript Due:	April 1, 1991
Publication Date:	4th Quarter 1991

Guest Editors

Dr. Khosrow Sohraby
IBM T. J. Watson Research Center
P.O. Box 704
Yorktown Heights, NY 10598
FAX: (914)784-6205
e-mail: sohraby@ibm.com

Dr. Inder S. Gopal
IBM T. J. Watson Research Center
P.O. Box 704
Yorktown Heights, NY 10598
FAX: (914)784-6205

Prof. Aurel A. Lazar
Department of Electrical Engineering
and
Center for Telecommunication Research
Columbia University
New York, NY 10027-6699
FAX: (212)932-9421

Prof. Luigi Fratta
Politecnino di Milano
Departimento di Elettronica
Piazza Leonardo Da Vinci 32
20133 Milano
Italy
FAX: +39-2-2399-3587

Call for Papers
IEEE Journal on Selected Areas in Communications
Signal Processing and Coding for Recording Channels

The IEEE Journal on Selected Areas in Communications announces a forthcoming issue on Signal Processing and Coding for Recording Channels. Prospective authors are encouraged to submit papers on disk systems (magnetic and optical) and tape systems. The focus of the issue is on signal processing and coding methods that can be used to improve recording channels, by increasing density, reducing manufacturing costs, increasing reliability, or otherwise. Both tutorial and original contributions in this area are solicited. A more detailed list of encouraged topics (all as related to recording channels) includes:

- channel characterization
- detection techniques and performance analysis
- run length limited and modulation codes
- error-correcting codes, combined error protection and modulation codes
- equalization and filtering, adaptive equalization
- write precompensation and write equalization
- sequence detection and partial response methods
- high-speed implementation of controller/channel electronics
- VLSI implementation of read/write processing circuitry
- timing recovery
- signal processing or coding for servicing
- coding bounds, density, and capacity
- data compression for digital and audio storage
- ac-bias or FM linearization techniques
- combined equalization and coding
- parallel access with multiple heads

Prospective authors should send 5 copies of their manuscript to one of the guest editors listed below. Deadline for initial paper submission is *September 20, 1990*. Acceptance notification by *January 15, 1991*. Final papers are due *April 1, 1991*, with an anticipated publication date in the third quarter of 1991.

Prof. John M. Cioffi
Information Systems Laboratory
Stanford University
Stanford, CA 94305
(415) 723-2150
fax: 415-723-8473

Prof. Jack K. Wolf
Center for Magnetic Recording Research, R001
University of California, San Diego
La Jolla, CA 92093
(619) 534-6218
fax: 619-534-2720

Dr. Hemant K. Thaper
IBM General Products Division, F84/0281
5600 Cottle Road
San Jose, CA 05193
(408) 284-0308
fax: 408-256-6577

Dr. Jan Bergmans
Philips Research Laboratories
P.O. Box 80.000
5600 JA Eindhoven
The Netherlands
31-40-743689
fax:31-40-743783