AN ANALYSIS AND DEMONSTRATION OF CLOCK SYNCHRONIZATION BY VLBI

William J. Hurd
Jet Propulsion Laboratory

ABSTRACT

A prototype of a semireal-time system for synchronizing the DSN station clocks by radio interferometry was successfully demonstrated on August 30, 1972. The system utilized an approximate maximum likelihood estimation procedure for processing the data, thereby achieving essentially optimum time synchronization estimates for a given amount of data, or equivalently, minimizing the amount of data required for reliable estimation. Synchronization accuracies as good as 100 nsec rms were achieved between DSS 11 and DSS 12, both at Goldstone, California. The accuracy can be improved by increasing the system bandwidth until the fundamental limitations due to position uncertainties of baseline and source and atmospheric effects are reached. These limitations are under ten nsec for transcontinental baselines.

I. INTRODUCTION

It is well known that the clocks at widely separated antenna ground stations can be synchronized by the techniques of very long baseline interferometry (VLBI). The objectives of this work are to optimize the signal processing of VLBI data and, utilizing the processing techniques developed, to demonstrate an operationally feasible time-synchronization system for the Deep Space Net (DSN). Although the results are discussed with application to the 26-m and 64-m antennas of the Deep Space Station (DSS) of the DSN, the analysis and techniques are applicable to any similar networks.

There are two reasons that an operational VLBI time-synchronization system may be desirable for the DSN. First, accuracies an order of magnitude better than currently attained by the moon-bounce system may be attainable with little initial investment and with operational costs which should be no higher than for the existing system. Second, VLBI may be the only operationally feasible method for achieving the 10- to 20-nanosecond (ns) accuracies required for two-station tracking of deep space probes. [1, 2]

The time-synchronization accuracy attainable by interferometry over very long baselines is fundamentally limited by the uncertainties in the differential time delay from the radio source to the antennas. These uncertainties, which increase with baseline length, are caused by errors in the estimates of the source positions and antenna location and by the variable propagation delays in the atmosphere. It is anticipated that the antenna locations will soon be known to within about one meter, and source position errors can be reduced to this same level by interferometry. The atmospheric effects depend on frequency in a

known manner, and can be calibrated by receiving on two frequencies, say S- and X-band. The fundamental limitation of accuracy can probably thus be reduced to ten ns or less for intercontinental baselines.

Until the fundamental limit is approached, the synchronization accuracy depends primarily on the utilized bandwidth, provided that the signal-to-noise ratio is high enough for reliable detection. The experiment reported on here confirms the two most important analytical results: First, that reliable estimates can be achieved with a small enough amount of data, about 1 million bits, so that semireal-time processing is feasible; and second, that with this amount of data, the rms errors are less than 0.1 times the inverse system bandwidth, so that rms errors of less than ten ns can be achieved with system bandwidths of only about ten MHz.

II. DESCRIPTION OF EXPERIMENT

As a first step in demonstrating the feasibility of an operational system for DSN clock synchronization by VLBI, an experiment was conducted on August 30, 1972, between the 26-m antennas at DSS 11 and 12, both at Goldstone. The experiment was implemented using a minimum of special interfacing hardware in addition to standard DSN station equipment. The data were acquired and processed following the approximate maximum-likelihood method described in the Appendix and in Reference 3.

A simplified block diagram of the experiment is shown in Figure 1. At each station, the received signals were demodulated in two-phase quadrature channels, filtered, quantized to one bit, and buffered into an XDS 920 TCP computer. Besides the receivers, the TCP computers are the major portion of the system. The special equipment for the experiment consisted of the two-channel demodulators, the filters, limiters, and samplers; and the buffers from the sampler to the TCP computers. This was all contained in one small chassis for each station, plus cables to interface to the computers.

The experiment procedure was to initiate sampling at the same time at each receiver according to the station master clocks, and to fill the TCP computer memories with data at the highest possible sampling rate. The computer speeds limited the data rate to 500 kbps, or 250 kbps per channel, so that the system bandwidth was limited to 250 kHz. Furthermore, the maximum number of samples which could be taken at this rate was limited by the memory sizes to approximately 320,000 bits. In an operational system, the data could be transmitted directly from the computers to JPL over the high speed data lines and processed within a few minutes in the Network Control System (NCS) or other computers. In the experiment, however, real-time operation was not required, but instead it was desired to make a number of independent estimates of time synchronization using each of several radio sources. Therefore, the data were written onto magnetic tape and processed later on a Sigma-5 computer at JPL. Five different radio sources were observed, with a total of 504 batches of data taken at ten-second intervals.

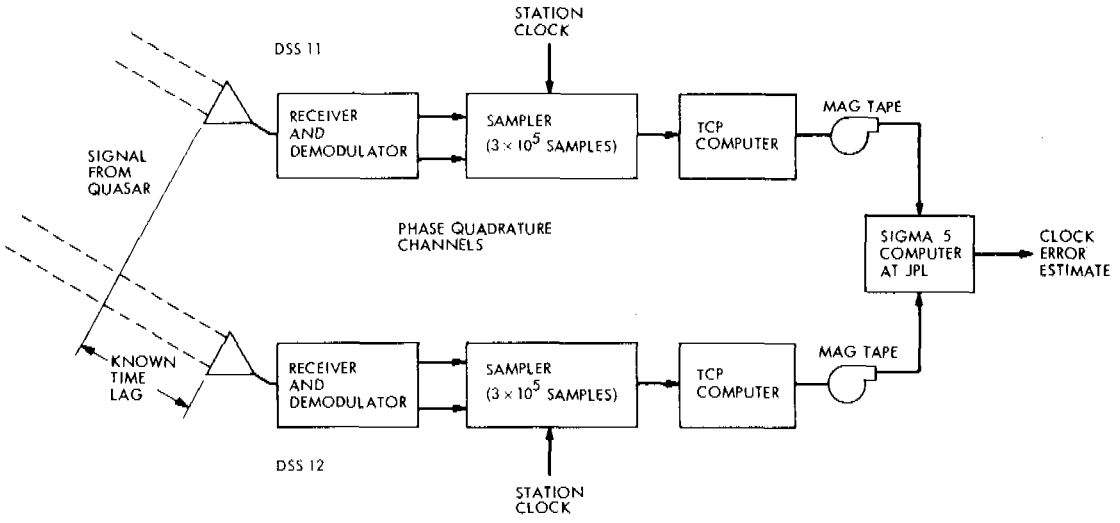

Figure 1. VLBI time-synchronization experiment block diagram.

III. PRINCIPAL RESULTS AND SYSTEM IMPLICATIONS

The desirability of an interferometry time synchronization system for the DSN depends on the ability to achieve reliable results with a reasonable amount of data. This, in turn, depends on the availability of radio sources with enough correlated flux, that is, with enough electro-magnetic flux which appears to be from an ideal point-source when viewed by the long baseline interferometer. In this section, we set a standard for the required source intensity for various system configurations based on experimental and analytical results, and show that adequate sources are available to result in an operationally feasible system.

The experimental results were limited by the system parameters of two 26-m antennas with temperatures of 16.3K and 37K, 250-kHz bandwidth, and 3.2×10^5 bit buffer size. The theoretical and experimental results are compared in Figure 2. Also shown are the theoretical results for a 2.5-MHz bandwidth, which could be realized by removing the sampling rate restriction from the current (Block III receiver) system, and for a 25-MHz bandwidth, which can be realized with the future DSN Block IV receivers. For the three strongest radio sources, rms processing errors of 96, 228, and 403 ns were achieved, in close agreement with theory. The results for the weakest of these sources, with an estimated correlated flux of 4.6 fu, are most significant for two reasons: First, the estimates were reliable even though the signal-to-noise ratio was somewhat lower than the desirable minimum, and second, the results were in close agreement with theory, indicating that the theory does not break down until the signal-to-noise ratio is reduced below this level. Based on both the theoretical and experimental results, we conclude that a source intensity of 5.5 fu would have been adequate to reliably achieve an rms error of less than 0.4 μsec, or less than one-tenth of the inverse of the system bandwidth. Whenever possible, higher accuracies should be achieved by increasing the system bandwidth and not the amount of

data or the signal-to-noise ratio, both because few sources have more than two to three fu of correlated flux over long baselines, and because increasing the amount of data is expensive in terms of buffer storage, computer time, and ground communications facility (GCF) usage.

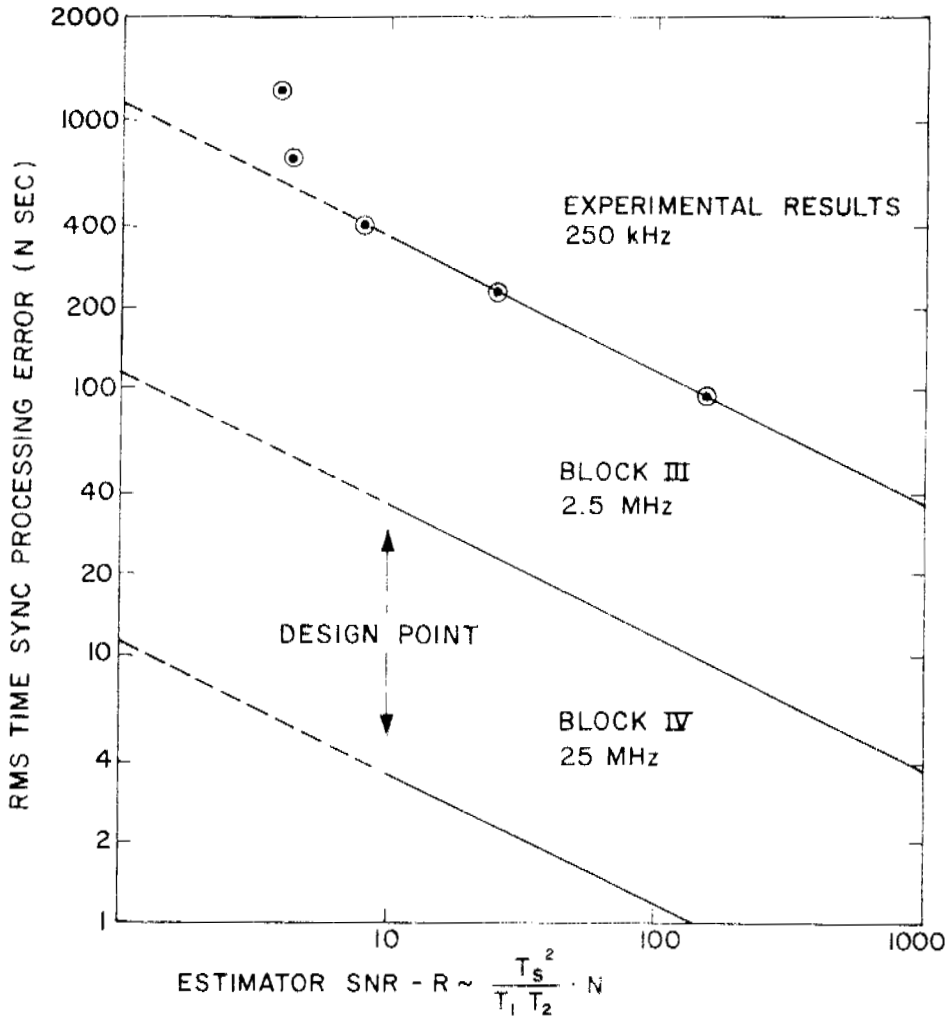

Figure 2. Theoretical and experimental time-synchronization error versus signal-to-noise ratio and bandwidth.

Table 1 presents the source intensities required to achieve the reliable performance level of one-tenth of the inverse system bandwidth for various antenna sizes and receiver noise temperatures in the DSN. Two buffer sizes are considered, the 0.32 megabit usable in the TCP computers, and 1.0 megabit, which is a practical size to consider if special-purpose memories are used for wider bandwidths. In utilizing Table 1, one should keep in mind that the system temperatures increase at low elevation angles, so that the required fluxes might increase by a factor of about 1.5.

Table 1
Source Intensity Required for Various System Parameters

Antenna Diameters (m)		System Temperatures (k)		Amount of Data or Buffer Size (10^6 bits)	Source Intensity (fu of Correlated Flux)
26	26	17	37	0.32	5.50
26	26	17	37	1.0	3.48
26	26	17	17	0.32	3.70
26	26	17	17	1.0	2.34
64	26	17	37	0.32	2.24
64	26	17	37	1.0	1.26
64	26	17	17	0.32	1.51
64	26	17	17	1.0	0.86
64	64	17	17	0.32	0.61
64	64	17	17	1.0	0.35

The availability of known radio sources was surveyed using Reference 4 and a computer program for mutual visibility devised by J.G. Williams of the JPL Tracking and Orbit Determination Section. Considering sources to be jointly visible only when the elevation from both stations is 10 degrees or greater, there is always at least one source of 1.3 fu or stronger visible by the station pairs at Goldstone and Spain, Goldstone and Australia, and Spain and South Africa. Sources of 2.0 fu are available for most of the day, and sources of 3 to 6 fu are normally visible for at least a few hours each day. The source 3C-454.3, which is sometimes as strong as 6.38 fu,[4] is visible to each of the above pairs for at least three hours a day, but unfortunately it has at other times been observed to be considerably weaker, and similar variations occur with some of the other strong sources. It is therefore not desirable to base a system on the strongest few sources.

Considering both the source intensities required and their availability, it is safe to say that station pairs with at least one 64-m antenna can be synchronized at will to within one-tenth of the inverse system bandwidth with 1.0 megabit of data. That is, there would be little if any operational restriction as to time of day due to lack of mutual visibility of adequate sources. Synchronization of two 26-m antennas could be accomplished with some restrictions on time of day, or by using more data. It is important to note that the amount of data used is not restricted by the high-speed buffer size, but convenience is sacrificed if it becomes necessary to fill the buffer several times, store the data on magnetic tape between fills, and then transmit a larger amount of data to the central computer for processing.

We conclude that a system with 1.0-megabit buffers would be operationally feasible. It would be less restricted than the X-band moon-bounce system, for which moon visibility restricts the time of day, and even the time of year for two northern hemisphere stations.

IV. ESTIMATION PROCEDURE

The estimates of time-synchronization error were made using the approximate maximum-likelihood method derived in the Appendix and in Reference 3. This method is distinguished from normal cross-correlation methods in that the cross-products are multiplied by appropriate weighting functions before being summed or envelope detected. This weighting accounts for changes in clock offset during the measurement time, and provides an optimum method for resolving the time estimates to greater accuracy than the time between samples. Approximate maximum-likelihood estimates of fringe frequency, phase, and signal-to-noise ratio also result.

The demodulation, filtering, and sampling procedures are shown in Figure 3, and are described in detail in Section V and in the Appendix. The ith samples in the phase-quadrature channels after demodulation, filtering, and limiting are denoted by X_i and Y_i for DSS 11 and by Z_i and W_i for DSS 12. These signals have cross correlations which depend on i and on ρ, τ, δ, ω, and ϕ where

$$\rho^2 = \frac{T_c^{\,2}}{T_{11} T_{12}}$$

is the product of the input signal-to-noise ratios, and

T_c = increase in system temperatures due to correlated flux from source

T_{11}, T_{12} = total system noise temperatures at DSS 11 and 12, including total source, flux, correlated or otherwise

δ = difference in path length from source to the two stations, in seconds (often called τ_g)

τ = error in clocks, or actual time difference between first samples at the two stations.

ω = stopped fringe frequency, or apparent doppler difference after demodulation, in rad/sec

ϕ = stopped fringe phase

As shown in the Appendix, the cross correlations, that is the expected values of the cross-products, can be expressed as

$$E(X_i Z_j) = \frac{2}{\pi} \rho\, a_{ij}(\tau, \delta) \cos(j\Delta + \phi) \tag{2}$$

$$E(X_i W_j) = \frac{2}{\pi} \rho\, b_{ij}(\tau, \delta) \sin(j\Delta + \phi) \tag{3}$$

$$E(Y_i Z_j) = \frac{2}{\pi} \rho\, c_{ij}(\tau, \delta) \sin(j\Delta + \phi) \tag{4}$$

and

$$E(Y_i W_j) = \frac{2}{\pi} \rho\, d_{ij}(\tau, \delta) \cos(j\Delta + \phi) \tag{5}$$

where $\Delta = \omega \cdot 4\,\mu$sec and it is assumed that the timing is such that the cross products are uncorrelated expcpt for $i \approx j$. The factor $2/\pi$ arises due to the hard limiting, and the coefficients a_{ij}, b_{ij}, c_{ij}, and d_{ij} are determined by the particular filtering and sampling method.

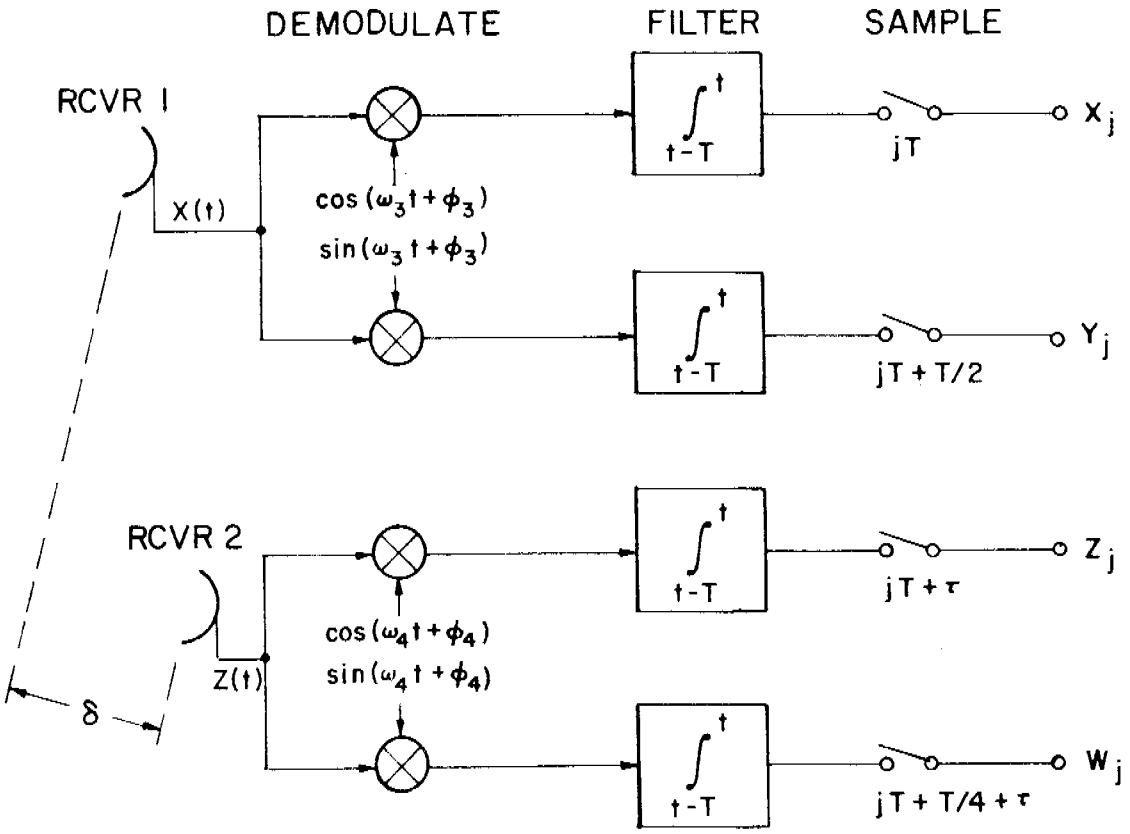

Figure 3. Demodulation, filtering, and sampling.

In general, for long baselines and measurement times, the fringe rate cannot be assumed constant, and $j\Delta$ must be replaced by a phase angle $\theta(j)$ which is known by the geometry. There is no significant difference in the estimation procedure. We assume here that the fringe rate is constant, for convenience and because this is valid for the short baseline of this experiment.

The approximate maximum-likelihood estimation procedure is derived in the Appendix. The implementation is to maximize the estimator function, G, over assumed values for τ and ω, for the actual received data samples. In calculating G, the stopped fringe rate ω is first normalized by subtracting out known quantities. Thus the frequency variable becomes

$$f = \frac{1}{2\pi}(\omega - \omega_0) \tag{6}$$

where ω_0 is the a priori estimate of the stopped fringe frequency. Two factors contribute to ω_0: the fringe rate as calculated from the geometry, and the difference in local oscillator frequencies, or effective receiver-center frequencies, at the two stations. The frequency f is the sum of the errors due to geometry and to oscillator instabilities, and the estimate of f is the estimate of these errors.

The steps in the estimation procedure for τ and f are:

(1) Assume a value of τ, say τ_k

(2) Form all cross-products whose cross correlations are nonzero for $\tau = \tau_k$

(3) Multiply the cross-products by the cross correlations for $\tau = \tau_k$ neglecting the sinusoidal terms, that is, form $X_i Z_j a_{ij}(\tau_k, \delta)$, and so on

(4) Assume a value for f, say f_j

(5) Evaluate $G(\tau_k, f_j)$

(6) Maximize $G(\tau_k, f_j)$ over the region of uncertainty in f by looping back to step 4

(7) Maximize $G(\tau_k, f_j)$ over the region of uncertainty in τ by looping back to step 1

(8) The estimates $\hat{\tau}$ and $\hat{f}$ of τ and f are the values of τ_k and f_j which maximize G

The distinguishing feature of this procedure is in weighting the cross-products by their assumed τ-dependence before envelope detecting. This gives a natural and optimum method for resolving the estimate of τ to greater resolution than the time between samples, and for accounting for the filtering and sampling methods and for the changes in τ over the measurement time.

V. TOWARDS OPTIMUM FILTERING AND SAMPLING

Although the ML estimation procedure is the same for all filtering and sampling methods, the statistics of the estimator function and of the estimates do depend on the filtering and

the sampling. In this experiment, the utilizable bandwidth was restricted by the maximum possible sampling rate to much less than the receiver bandwidth. Thus the filters could be chosen essentially arbitrarily. For this case, it is shown in Reference 3 that a filter which integrates over the time between samples (a sliding window integrator) is nearly ideal in the sense of maximizing both the minimum and the average signal-to-noise ratios of the estimator function. In conjunction with this filter, the sampling times in the various channels should be staggered as shown in Figure 3. Both the cosine and sine channels at both receivers are sampled with a uniform interval of T = 4 μsec between samples, but the sine channel is sampled T/2 later than the cosine channel at one receiver, and T/4 later at the other receiver.

The optimization problem is considerably different when the utilized bandwidth is limited by the receiver RF bandwidths. In this case the receiver transfer function may be the principal factor determining the effective filter characteristics, and the primary design parameters to optimize are the sampling rate and phase relationships.

VI. PROPERTIES AND EXAMPLES OF THE ESTIMATOR FUNCTION

The statistics of the estimator function have been evaluated both analytically and by simulation.[3] We summarize here some of the key statistics, and then examine graphically some typical sample functions which were observed in the experiment.

The estimation procedure is considered to be reliable when the probability is high that the estimates are in the general vicinity of the correct values of the parameters, rather than being completely extraneous. This depends on the probability distributions of G for the correct and widely erroneous values of the parameters. Once the form of the distributions are known, the performance can be well predicted by a figure of merit which we call the signal-to-noise ratio of the estimator. It is defined as the square of the difference in the means of G for the correct and incorrect values of the parameters, to the variance of G at the correct values. When G is normalized in the natural manner, its mean is unity for widely incorrect assumed values of the parameters, and is unity also when $\rho = 0$, so

$$R = \frac{[E\{G(\tau, f)\} - 1]^2}{\text{Var } G(\tau, f)} \tag{8}$$

The estimator signal-to-noise ratio varies approximately as ρ^2, i.e., as the product of the input signal-to-noise ratios, or alternatively as the square of the source flux density. For the particular filtering and sampling method used, it is given by

$$R = \frac{1}{2} \frac{r}{1 + \left(\frac{1}{2r}\right)} \approx \frac{r}{2} \tag{9}$$

where

$$r = 0.267 \rho^2 N$$

and N is the number of samples in each channel at each receiver. Since the system bandwidth is the inverse of the time between samples in one channel, N is also the system time-bandwidth product.

Estimation will be reliable whenever R exceeds about 10, because the maximum value of G will almost always occur in the vicinity of the correct values of τ and f unless the initial uncertainty in these parameters is large. For example, when the initial uncertainty in f is negligibly small, the number of independent values of G which must be calculated is approximately equal to the time uncertainty times twice the system bandwidth. For the 250 kHz bandwidth of this experiment, time uncertainties of ±10 to ±100 μsec would require calculation of only 10 to 100 independent values of G. It can be seen from the curves of Reference 3 that, for these uncertainties, the results would be reliable about 98 to 99 percent of the time with R = 10.

The resolution of the estimates depends on the peakedness of G more than on R. An approximation to the rms error is estimation of τ is presented in Reference 3, and is

$$\sigma_\tau = \frac{0.79T}{\rho N^{1/2}} \tag{10}$$

where T is the time between samples in one channel, or the inverse system bandwidth. In terms of R,

$$\sigma_\tau \approx \frac{0.289T}{R^{1/2}}$$

so that R = 10 is sufficient to reduce the rms error to less than 0.1 T as well as to result in reliable estimation.

Insight into the capabilities of the estimator function to resolve time and frequency can be gained by studying the function at high signal-to-noise ratios. Figure 4 shows a plot of an actual sample function of $G(\tau_k, f_j)$ observed for a fairly high intensity source, 3C279, with R estimated to be 24.8. The maximum of G is 52.844 and occurs for $f_j = -0.20$, $\tau_k = 40.97$, so that these are the estimates $\hat{f}$ and $\hat{\tau}$ of f and τ. In the time domain, G is nominally symmetrical, and decreases to half its maximum in under ± 2 μsec, and approximately to zero in ±4 μsec. In the frequency domain, G is also nominally symmetrical about the actual value of f, although this is not apparent from the sample function because the maximum did not occur at $f_j = 0$. The measurement time of the experiment was NT ≈ 0.64 sec., and the effective bandwidth of G is slightly less than the inverse of this time. It is observed that for different f_j, the maximum of G occurs at very close to the same value of τ_k. This implies that it may be unnecessary to maximize over f_j when only estimates of τ are required, provided that the initial uncertainty in f is small compared to 1/NT, say less than ±0.1/NT.

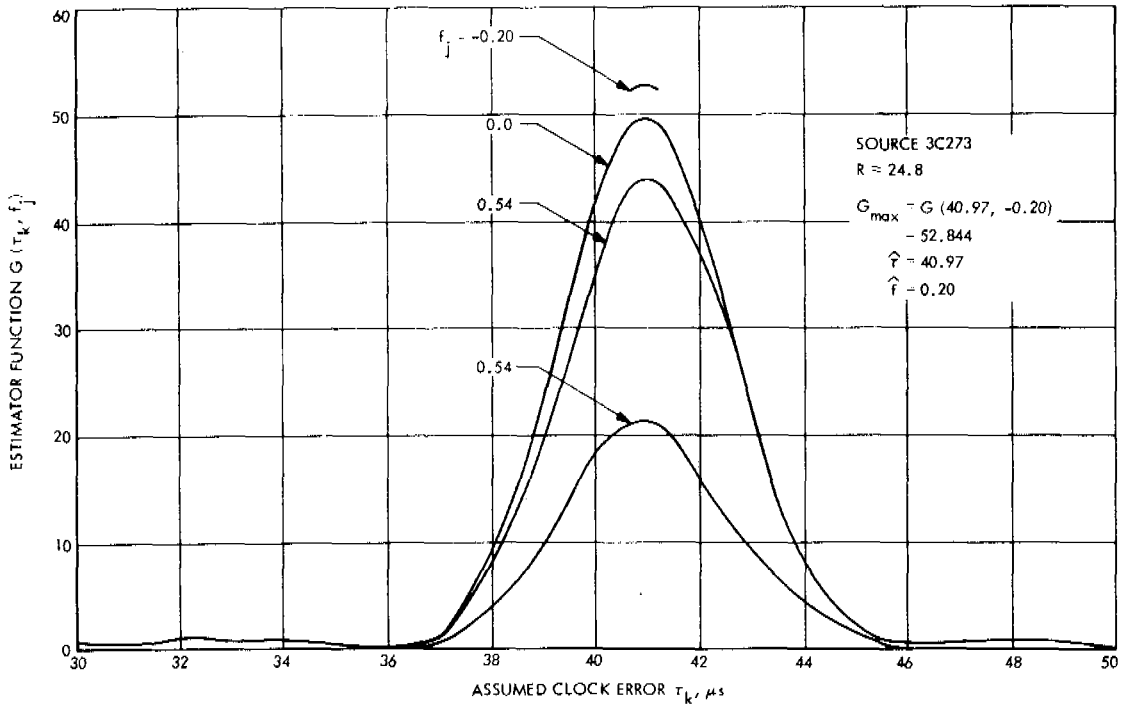

Figure 4. An estimator sample function at a high signal-to-noise ratio.

The performance of the estimator when the noise is significant is illustrated in Figures 5 and 6. Each presents three sample functions from different realizations of the experiment, with the time dependence shown only for the frequency variable fixed at the nominal value, $f_j = 0$. Figure 5 is for a weaker source, 4C39.25, with R estimated to be 3.81, which is significantly below the suggested design value of 10. In one of the three cases, the maximum of G occurs near $\tau_k = 21$ μsec, far removed from the true value which is near 41 μsec. Extraneous results like this occur frequently at these low signal-to-noise ratios. Figure 6 is for source P1127-14, with R estimated to be 8.20, which is only marginally below the design point of 10. Fairly wide variations in the maximum value of G occur at this signal-to-noise ratio, but no extraneous maxima were observed in the 72 sets of data taken for this source.

VII. DETAILED RESULTS

A total of 504 sets of data were taken using five different radio sources, and independent estimates of the time and frequency differences at the two receivers were made for each set of data. The most important results are the means and standard deviations of the estimates of τ and f as a function of the estimator signal-to-noise ratio, R. In order to present these results, it was necessary to estimate R from the data. The method for estimating R is presented later.

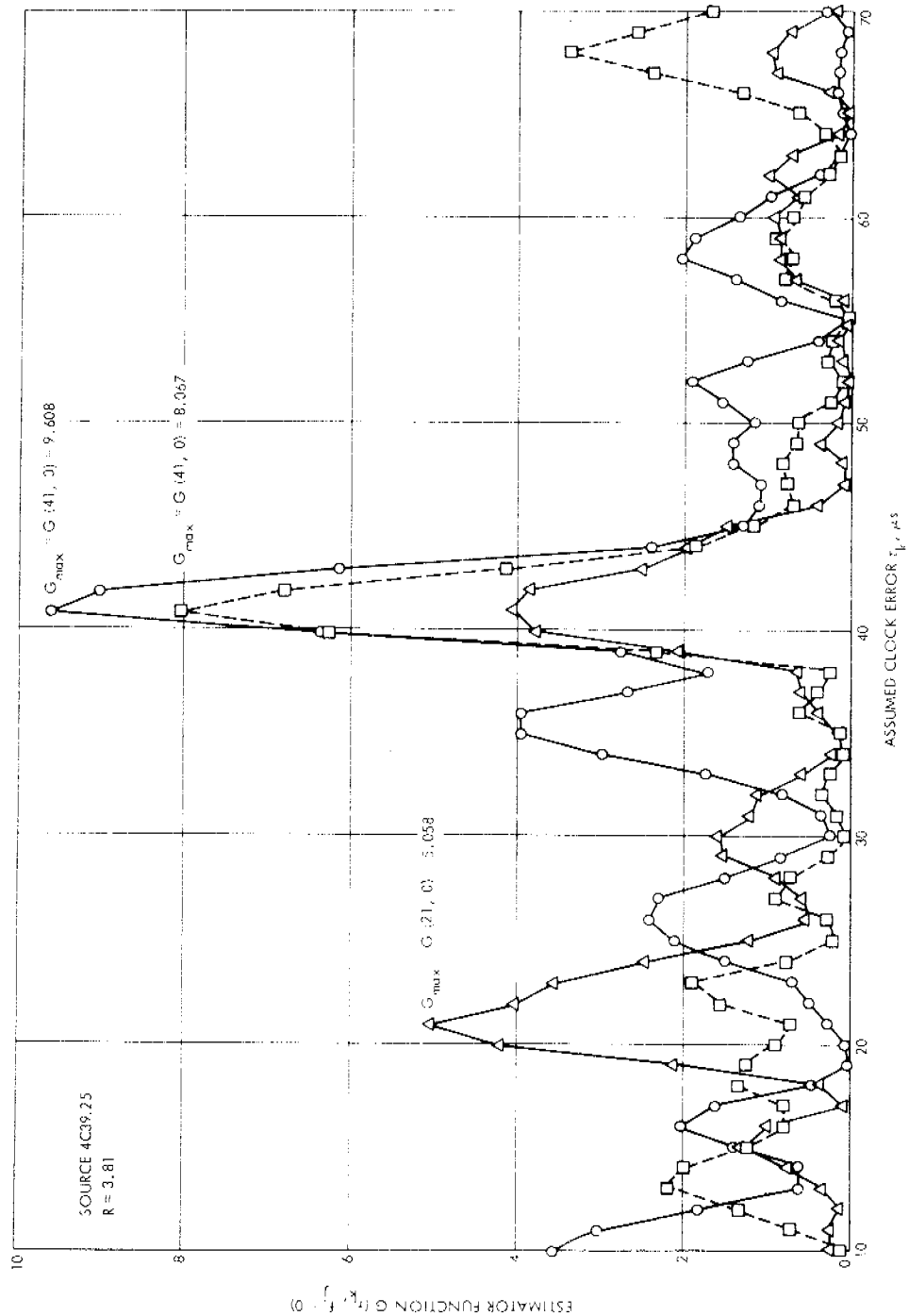

Figure 5. Three estimator sample functions at a low signal-to-noise ratio.

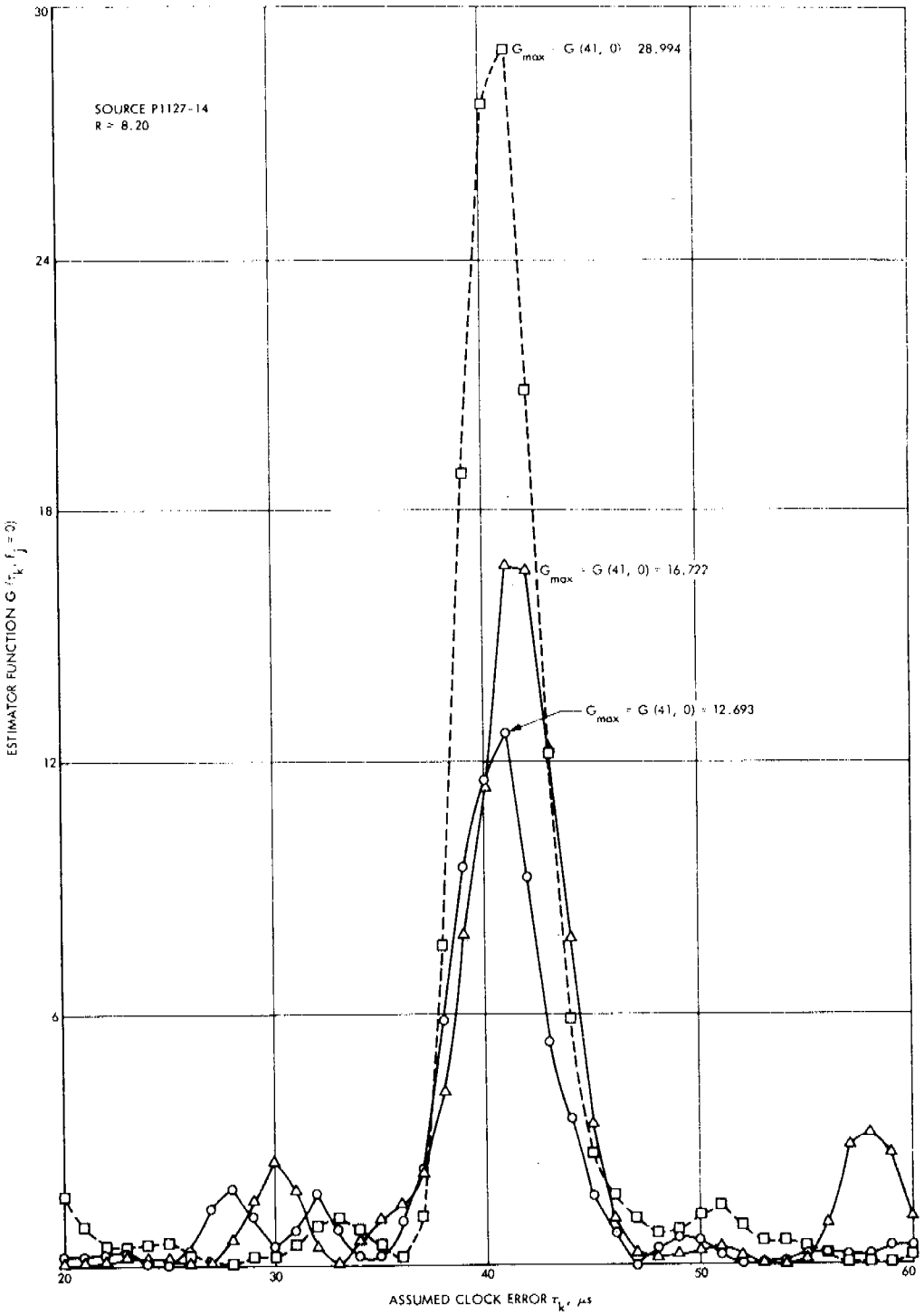

Figure 6. Three estimator sample functions at a marginal signal-to-noise ratio.

Joint Estimate of τ and f

Table 2 presents the results of the joint estimation of τ and f for the five sources. The statistics are based on 144 independent estimates for each of the two strongest sources, 3C273 and 3C279, and on 72 cases for the other sources. Since the true values of τ and f were not known, it was not possible to compute the actual rms errors, therefore, the standard deviations were estimated from the data using the estimates of the means. The standard deviation of the mean estimate for one source is equal to the standard deviation of one estimate for that source, divided by the square root of the number of cases. None of the mean estimates of τ differ from the value 40.97 by more than two standard deviations. All variations in the mean estimates can thus be attributed to noise. There is no evidence to suggest any effects due to errors in source or station positions, changes in the clock synchronization during the experiment, or errors in data processing.

The statistics of the estimates of f cannot be attributed entirely to noise, because of local oscillator instabilities. A hydrogen maser was used for the S-band reference at DSS 12, and rubidium was used at DSS 11, so the rubidium standard contribution dominated. Both the long- and short-term stabilities are on the order of one part in 10^{11}. Errors in the nominal value of f of up to 0.1 Hz were anticipated, as were short term variations with a standard deviation on the order of 0.01 to 0.1 Hz.

Due to noise alone, the standard deviation of f should vary as $R^{1/2}$, provided R is ~ 10 or greater. This relationship was nominally satisfied for the second and third strongest sources, with R = 24.8 and 8.20, and standard deviations of 0.0946 and 0.159 Hz. For the strongest source, the frequency instability was not negligible compared to the noise. Therefore, its effect was estimated from the results for two strongest sources, assuming the noise and instability errors to add in the mean square. The rms error due to frequency instability was estimated at 0.036 Hz, which is well within the range of uncertainty of this effect. The rms frequency estimation error due to noise is then approximately

$$\frac{0.468}{R^{1/2}} \quad \text{Hz}$$

This relationship was also closely satisfied for the next weakest source. An R of 10 thus results in an rms error in frequency estimate of less than 0.1 divided by the measurement time of 0.64 sec, just as it results in a timing error of less than 0.1 divided by the bandwidth.

Estimation of τ For Fixed f

When the a priori uncertainty in frequency is small, τ can be estimated by maximizing G over τ only, assuming no frequency error, that is, f = 0. This results in better estimates of τ than does joint estimation of τ and f when the long- and short-term frequency instabilities are very small. Before this experiment was actually performed, it was felt that the frequency stabilities would be sufficiently good to omit maximization over f, and this was confirmed in the experiment. However, the amount of long-term drift is random, and the

Table 2
Joint Estimation of τ and f.

Radio Source	Estimated SNR of Estimator (R)	ESTIMATION OF τ		ESTIMATION OF f	
		Mean (μsec)	Standard Deviation (μsec)	Mean (Hz)	Standard Deviation (Hz)
3C273	148.0	40.955	0.0956	-0.0807	0.0503
3C279	24.8	41.00	0.228	-0.0799	0.0946
P1127-14	8.20	40.95	0.403	-0.0810	0.159
DW0742+10	4.24	40.95	0.719	-0.0772	0.239
4C39.25	3.81	40.78	1.26	-0.0355	0.249

frequency offsets in the local oscillators might have been too large on another day. It was, therefore, necessary to process the data in both manners, in order to be able to predict future performance.

The results of estimation of τ for f fixed at zero are presented in Table 3. The theoretical rms errors in estimation of τ are also presented, as calculated for the estimated values of R. For the three highest signal-to-noise ratio cases, the observed and calculated rms errors were very close. For the lowest two signal-to-noise ratios, the observed errors were significantly higher than the calculated values. This is because the theory breaks down when R is low enough so that extraneous results occur.

The observed rms errors at low signal-to-noise ratios would have been still higher if the assumed region of uncertainty in τ had been greater, because there would have been more extraneous results due to noise. Throughout the experiment, the uncertainty region was assumed to be from 30 to 50 μsec.

Table 3
Estimation of τ for f = 0.

Radio Source	Estimated R	Estimation of τ		Theoretical rms Error in τ (μsec)
		Mean (μsec)	Standard Deviation (μsec)	
3C273	148.0	40.955	0.0978	0.095
3C279	24.8	41.00	0.224	0.232
P1127-14	8.20	40.95	0.408	0.403
DW0742+10	4.24	40.98	0.641	0.560
4C39.25	3.81	40.92	0.952	0.591

Comparison of Estimation Methods

In comparing the results of estimating τ jointly with f and with f fixed at zero, it is seen that there is negligible difference in the standard deviations of the estimates for the three highest signal-to-noise ratio cases, and that all are close to theory. For the two lower signal-to-noise ratios, the errors are significantly higher when f is estimated instead of assumed to be zero. There are two reasons for this. First, the estimates of f are poor enough to degrade the estimate of τ. Second, more extraneous estimates occurred, because there were effectively more independent calculations of G for noise only.

Estimation of R, ρ, and Flux Density

For each independent case, the approximate maximum likelihood estimate for ρ is the square root of the maximum value of G, divided by the proper normalization factor. This is the best estimate of ρ only because the maximum value of G occurs at the best estimates of τ and f. A better estimate of ρ would be obtained from value of G at the correct values of τ and f. Therefore, since it was desired to have the overall best estimates of ρ, and hence of R, the values of ρ were estimated using the best overall estimates of τ and f. These best estimates were taken as τ = 40.955 and f = 0.0807 Hz, the values obtained from the strongest source. The overall estimates of ρ for each source were taken as the average of the estimates of ρ for all of the cases J or that source.

The estimates of R were obtained from the estimates for ρ according to equations 8 and 9. To estimate the correlated fluxes, it was assumed that the system temperatures at DSS 11 and 12 were the cold sky temperatures of 37K and 16.3K, respectively, raised by the source total flux at the rate of 0.11K per flux unit. Then the correlated fluxes are given by Equation 1.

Table 4 presents the estimated flux densities, input signal-to-noise ratios, and estimator signal-to-noise ratios for the five sources.

Table 4
Estimated Flux Densities and Estimator SNR's.

Radio Source	Number of Cases	Total Flux [Ref. 4] (fu)	Estimated Correlated Flux (fu)	Estimated Geo. Mean Input SNR (ρ)	Estimated Estimator SNR (R)
3C273	144	39.0	22.0	0.0834	148.0
3C279	144	12.2	8.1	0.0344	24.8
P1127-14	72	6.2	4.6	0.0199	8.20
DW0742+10	72	3.7	3.3	0.0145	4.24
4C39.25	72	3.8	3.1	0.0138	3.81

APPENDIX

This appendix presents a precise formulation of the problem and the notation, and the derivation of the approximate maximum likelihood estimation procedure. The optimization of the filtering and sampling, an analysis of the statistics of the estimator function, and an approximation to the rms error of the time estimate are presented in Reference 3.

Problem Formulation and Data Sampling

Figure 3 illustrates the demodulation, filtering, and sampling of the radio-source signal and receiver noise at the two ground stations. The radio energy emitted by the radio point source is essentially white and gaussian. However, because we can only observe the energy in the bandwidth of our receivers, we can consider the signal to be a narrowband gaussian process. The signal plus noise at the outputs of the two receivers can be represented as

$$X(t) = [n(t)+s(t)] \cos(\omega_1 t+\phi_1) + [m(t)+r(t)] \sin(\omega_1 t+\phi_1) \qquad (A1)$$

and

$$Z(t) = [p(t)+s(t-\delta)] \cos(\omega_2 t+\phi_2)+[q(t)+r(t-\delta)] \sin(\omega_2 t+\phi_2) \qquad (A2)$$

where

t = time

$\delta = \delta(t)$ = time lag from receiver 1 to receiver 2

$\omega_1 - \omega_2$ = difference in doppler shift, or actual fringe frequency

ϕ_1, ϕ_2 = random phase angles

$s(t), r(t)$ = noise processes representing signal

$n(t), m(t), p(t), q(t)$ = receiver noise

All of the noise processes are assumed independent and bandlimited only by the receivers. The difference frequency $\omega_2 - \omega_2$ and difference phase $\phi_1 - \phi_2$ are assumed to be constant over the observation time, however, the time delay $\delta(t)$ varies due to the rotation of the Earth. We can assume this to be linear and known, $\delta(t) = \delta_0 + \delta t$. The difference frequency and phase are essentially constant only because the change in δ is small compared to the reciprocal of the difference frequency.

Suppose now that we observe $X(t)$ beginning at $t = 0$, and $Z(t)$ beginning at $t = \tau$. This time offset τ is not precisely known, because the clocks at the two stations are not precisely synchronized. We desire to form an estimate $\hat{\tau}$ of τ from the received signals, and to use this estimate to synchronize the clocks.

In order to extract the maximum information from the received signals, both the sine and cosine components of the random processes must be processed. The received signals are thus demodulated to baseband in two channels, using quadrature phase reference signals

derived from rubidium frequency standards which we require to be frequency and phase stable over the observation interval. The signals are then filtered and sampled, with the filtering assuring that all samples in each channel are independent of one another. The demodulated and filtered signals, with * denoting convolution, are

$$x(t) = [X(t)\cos(\omega_3 t+\phi_3)] * h_x(t) \tag{A3}$$

and

$$y(t) = [X(t)\sin(\omega_3 t+\phi_3)] * h_y(t) \tag{A4}$$

at the X receiver, and

$$z(t) = [Z(t)\cos(\omega_4 t+\phi_4)] * h_z(t) \tag{A5}$$

and

$$w(t) = [Z(t)\sin(\omega_4 t+\phi_4)] * h_w(t) \tag{A6}$$

at the Z receiver. We have represented the filtering by convolutions with h_x, h_y, h_z, and h_w, the filter weighting functions.

Since the frequency and phase reference for a narrowband process can be chosen arbitrarily, we can choose the frequency and phase reference of either X or Z arbitrarily. For convenience, we chose $\omega_1 = \omega_3$ and $\phi_1 = \phi_3$, and we define $\omega = \omega_2 - \omega_4$ and $\phi = \phi_2 - \phi_4$. The difference frequency ω, also called the stopped fringe rate, is determined by the relative doppler between X and Z, as reflected by ω_2, and by the reference ω_4. The difference or fringe phase ϕ is random, and uniformly distributed. With this simplification, the observed processes are

$$x(t) = [n(t)+s(t)] * h_x(t) \tag{A7}$$

$$y(t) = [m(t)+r(t)] * h_y(t) \tag{A8}$$

$$z(t) = \{[p(t)+s(t-\delta)]\cos(\omega t+\phi) + [q(t)+r(t-\delta)]\sin(\omega t+\phi)\} * h_w(t) \tag{A9}$$

and

$$w(t) = \{[q(t)+r(t-\delta)]\cos(\omega t+\phi) - [p(t)+s(t-\delta)]\sin(\omega t+\phi)\} * h_w(t) \tag{A10}$$

The four observables are now sampled, all at a uniform and identical rate, with a sampling interval T. Independence of the samples in each channel is assured by having the weighting functions be zero outside of the interval (O, T), and by the whiteness of the noise processes. A remaining parameter which can be varied is the relative times of the samples in the sine and cosine channels, so we leave this arbitrary. As references, we assume that the sampling

of x(t) begins at t = 0, and the sampling of z(t) begins at t = τ, that is, at the delay we wish to estimate. The samples of y and w occur Δ_1 and Δ_2 after the samples of x and z. Thus the samples are

$$X_j = x(jT)$$

$$Y_j = y(jT+\Delta_1)$$

$$Z_j = z(jT+\tau)$$

and

$$W_j = w(jT+\tau+\Delta_2)$$

At this point we make the further assumption that ω is a very low frequency compared to the sampling rate, so that the factors $\cos(\omega t + \phi)$ are constant over T and can be brought outside of the convolution integrals. This assumption is reasonable, since ω can be chosen by the experimenter.

We now normalize the observables to unit variance, and express the observable covariances as

$$E(X_i Z_j) = A_{ij} = \rho a_{ij} \cos(jT\omega + \phi) \quad \text{(A11)}$$

$$E(X_i W_j) = B_{ij} = \rho b_{ij} \sin(jT\omega + \phi) \quad \text{(A12)}$$

$$E(Y_i Z_j) = C_{ij} = \rho c_{ij} \sin(jT\omega + \phi) \quad \text{(A13)}$$

$$E(Y_i W_j) = D_{ij} = \rho d_{ij} \cos(jT\omega + \phi) \quad \text{(A14)}$$

The a_{ij}, b_{ij}, c_{ij}, d_{ij} reflect the dependence on τ-δ (t), and are constant for fixed i-j when τ-δ is constant. In any case, they vary slowly in i-j. Also, the sinusoidal variation in the covariances is slow in j, because $\omega T \ll 1$. Thus for each i-j there is a range of j for which the covariances are essentially constant.

Derivation of Approximate Maximum Likelihood Estimator

The general procedure of maximum-likelihood estimation is to maximize the a posteriori probability density function (PDF) of the observables, conditioned on the unknown parameters. The values of the parameters which maximize the PDF for the given set of observables are chosen as the maximum-likelihood (ML) estimates. The parameters to be estimated here are ρ, τ, ϕ, and ω. In this section, we derive approximate maximizations of the PDF with respect to ρ and ϕ. The resulting function must then be maximized numerically with respect to τ and ω in order to obtain estimates of all the parameters.

The first step in our problem is to find the joint PDF of the observables X_i, Y_i, Z_i, and W_i, conditioned on the unknown parameters ρ, ϕ, τ, and ω. This PDF depends only on the conditional covariance matrix, since the observables are jointly gaussian and zero mean.

Suppose we define a row vector U having as its components all of the observables:

$$U = (X_1, X_2, \ldots, X_N, Y_1, Y_2, \ldots, Y_N, Z_1, Z_2, \ldots, Z_N, W_1, W_2, \ldots, W_N) \quad (A15)$$

where N is the number of samples of each variable.

Then the covariance matrix of U is

$$\Lambda = \begin{pmatrix} I & O & A & B \\ O & I & C & D \\ A^t & C^t & I & O \\ B^t & D^t & O & I \end{pmatrix} \quad (A16)$$

where A, B, C, and D are the covariance matrices with elements A_{ij}, B_{ij}, and so on, given by equations (A11) through (A14), and the conditional PDF of the observables is

$$P(U\rho, \phi\tau, \omega) = \frac{C}{|\Lambda|^{1/2}} \exp\left[-\frac{1}{2} U\Lambda^{-1} U^t\right]$$

The covariance matrix Λ depends on the parameters ρ, ϕ, τ, and ω, and c is a constant.

The major problem at this point is to invert the covariance matrix. We can do this only in series form, and it is the truncation of this series in the maximization procedure which causes our estimator to be only approximately maximum likelihood.

To proceed we define a matrix P such that

$$\Lambda = I + P \quad (A18)$$

The matrix P has at most four non-zero elements in each row and column, because A, B, C, and D have at most two non-zero elements in each row and column. Furthermore, the non-zero elements of P are proportional to ρ and do not exceed ρ in absolute value. Since ρ is small ($<10^{-2}$), we can expand Λ^{-1} in a power series, and bound the terms:

$$\Lambda^{-1} = I - P + P^2 - P^3 + \cdots \quad (A19)$$

Since the two principal quadrants of P are zero, the principal diagonal elements of P^n are zero for odd n. The other elements are bounded by

$$\begin{aligned} \max_{i,j} \left| (P^n)_{ij} \right| &\leq 4\rho \max_{i,j} \left| (P^{n-1})_{i,j} \right| \\ &\leq 4^{n-1} \rho^n \end{aligned} \quad (A20)$$

where $(P^n)_{ij}$ denotes the ij elements of P^n.

Closer bounds can be obtained utilizing properties of the cross covariances for particular cases.

The conditional pdf can now be written as

$$P(U|\rho,\phi,\tau,\omega) = c \exp -\left[\tfrac{1}{2}\, U(I+P)^{-1}\, U^t - \tfrac{1}{2} \log \det (I+P)\right] \tag{A21}$$

Using a well-known matrix identity,

$$\log \det (I+P) \equiv \mathrm{Tr} \log (I+P) = \mathrm{Tr}\left(P - \frac{P^2}{2} + \frac{P^3}{3} - \frac{P^4}{4} + \cdots\right) \tag{A22}$$

The odd power terms can be deleted, since the principal diagonal of P^n is zero for odd n. Thus

$$\log \det (I+P) = -\mathrm{Tr}\left(\frac{P^2}{2} + \frac{P^4}{4} + \cdots\right) \tag{A23}$$

We now define a likelihood function $L_1(U|\rho,\phi,\tau,\omega)$ as the exponent of the conditional pdf, and maximization of L_1 is equivalent to maximization of the pdf.

$$L_1(U|\rho,\phi,\tau,\omega) = -\tfrac{1}{2}\, U(I - P + P^2 - + \cdots)U^t \tag{A24}$$

$$+ \tfrac{1}{2}\, \mathrm{Tr}\left(\frac{P^2}{2} + \frac{P^4}{4} + \cdots\right)$$

It is not feasible to maximize L_1 analytically with respect to any of the parameters without neglecting terms in P of higher order than P^2. With this approximation, we can maximize with respect to ρ and ϕ. Since normally τ and ω are the parameters of primary interest, the approximate solutions for ρ and ϕ usually suffice, but greater accuracy can be obtained numerically if required.

To proceed, we define a new matrix Q by

$$Q = \frac{1}{\rho} P = \frac{1}{\rho} \begin{pmatrix} 0 & 0 & A & B \\ 0 & 0 & C & D \\ A^t & C^t & 0 & 0 \\ B^t & D^t & 0 & 0 \end{pmatrix} \tag{A25}$$

Next we drop the $U\, I\, U^t$ term in L_1, which is independent of the parameters, to obtain

$$L_2(U|\rho,\phi,\tau,\omega) \approx \tfrac{1}{2}\, U(\rho Q - \rho^2 Q^2)U^t + \tfrac{1}{4}\, \rho^2\, \mathrm{Tr}(Q^2) \tag{A26}$$

By differentiating with respect to ρ, we see that L_2 is maximized for the conditional estimate of ρ

$$\hat{\rho} = \frac{UQU^t}{2UQ^2U^t - \mathrm{Tr}(Q^2)} \tag{A27}$$

The denominator of this expression can be approximated by its mean, which is $\mathrm{Tr}(Q^2)$, so

$$\hat{\rho} \approx \frac{UQU^t}{\mathrm{Tr}(Q^2)} \tag{A28}$$

The variance of the denominator of equation (A27) is also on the order of $\mathrm{Tr}(Q^2)$. Therefore, since $\mathrm{Tr}(Q^2) \approx 4\,N$, the approximation is good when N is large, say 10^4 or greater, which will always be true in VLBI problems.

A new likelihood function is now obtained by substituting the value of $\hat{\rho}$ into equation (A26), and again approximating uQ^2u^t by $\mathrm{Tr}(Q^2)$:

$$L_3(U\,|\rho,\phi,\tau,\omega) \approx \frac{[UQU^t]^2}{\mathrm{Tr}(Q^2)} \tag{A29}$$

Since the elements of Q vary slowly except for the sinusoidal variation, $\mathrm{Tr}(Q^2)$ is essentially independent of Q and ω so long as $N\Delta\omega \gg \pi$. This can be assured by controlling ω by selecting the local oscillator frequencies. Neglecting any slight variation of $\mathrm{Tr}(Q^2)$, L_2 can be maximized over φ. To do this, Q is expressed

$$Q = R\cos\phi + S\sin\phi \tag{A30}$$

where R and S do not depend on ϕ and are given by

$$R = \begin{pmatrix} 0 & R_o \\ R_o^t & 0 \end{pmatrix} \tag{A31}$$

$$S = \begin{pmatrix} 0 & S_o \\ S_o^t & 0 \end{pmatrix} \tag{A32}$$

where

$$R_o = \begin{pmatrix} (a_{ij}\cos j\Delta\omega & (-b_{ij}\sin j\Delta\omega) \\ (c_{ij}\sin j\Delta\omega) & (-d_{ij}\cos j\Delta\omega) \end{pmatrix} \tag{A33}$$

$$S_o \quad \begin{pmatrix} (-a_{ij}\sin j\Delta\omega) & (-b_{ij}\cos j\Delta\omega) \\ (c_{ij}\cos j\,\Delta\omega) & (-d_{ij}\sin j\Delta\omega) \end{pmatrix} \tag{A34}$$

The derivative of the likelihood ratio with respect to ϕ is then

$$\frac{d}{d\phi} L_3 = \frac{2(UQU^t)\, U(S\cos\phi - R\sin\phi)\, U^t}{Tr(Q^2)} \tag{A35}$$

and the value of ϕ which maximizes L_3 is

$$\hat{\phi} = \arctan \frac{USU^t}{URU^t} \tag{A36}$$

The new likelihood ratio is the maximum of L_3, that is, $L_3(\hat{\phi})$, which we renormalize to obtain the final estimator function G:

$$G(\tau,\omega) = \frac{(URU^t)^2 + (USU^t)^2}{4Tr(Q^2)} \tag{A37}$$

This is as far as we can proceed analytically. To find the final approximate ML estimates of all the parameters, G is maximized numerically over τ and ω. When only $\hat{\tau}$ is required, ω is usually known a priori, so that the numerical maximization is only over one parameter, τ.

ACKNOWLEDGEMENTS

The author acknowledges the assistance of P.F. MacDoran, J.G. Williams, and D.S. Spitzmesser, who shared their knowledge and experience in VLBI experiments, and of S.S. Brokl and the station personnel at DSS 11 and 12, who assisted in implementing the experiment.

REFERENCES

1. C.E. Hildebrand, V.J. Ondrasik, and G.A. Ransford, "Earth-Based Navigation Capabilities for Outer Planet Missions," AIAA/AAS Astrodymanics Conference, Palo Alto, California, September 11-12, 1972. AIAA Paper No. 72-925.
2. V.J. Ondrasik, C.E. Hildebrand, and G.A. Ransford, "Preliminary Evaluation Radio Data Orbit Determination Capabilities for the Saturn Portion of a Jupiter-Saturn-Pluto 1977 Mission," DSN Progress Report, JPL TR 32-1526, Vol. X, August 15, 1972, pp. 59-75.
3. W.J. Hurd, "DSN Station Clock Synchronization by Maximum Likelihood VLBI," DSN Progress Report, JPL TR 32-1526, Vol. X, August 15, 1972, pp. 82-95.
4. K.I. Kellermann, et.al., "High Resolution Observations of Compact Radio Sources at 13 Centimeters," Astrophysical J., Vol. 151, September 1970, pp. 803-809.