Optimization of Feedforward Amplifier Power Efficiency on the Basis of Drive Statistics

Colin L. Larose, *Member, IEEE*, and Fadhel M. Ghannouchi, *Senior Member, IEEE*

***Abstract*—Among power amplifier linearization techniques, feedforward delivers the best performance, but at the cost of significant degradation in the amplifier's power efficiency. This paper details a procedure to design feedforward amplifiers for optimal dc–RF conversion efficiency. The procedure has the convenience of requiring only the power statistics of the driving signal, of being computationally efficient, and of lending itself to a highly intuitive graphical representation. The gains and normalized saturated output powers of the main and error amplifiers, as well as the various couplings, are optimized for a specified linearity and gain at the output of the system. The amplifier types and adaptation methods employed must be specified at the outset, but results are presented that begin to reveal the impact of these factors on efficiency and, thus, to demonstrate the investigative potential of the procedure.**

***Index Terms*—Design methodology, feedforward amplifiers, linearization, microwave power amplifiers, power efficiency.**

I. Introduction

THE power efficiency of a microwave amplifier is best when it is operated near saturation. Amplification of a communications signal in this nonlinear range generates intermodulation (IM) distortion that interferes with neighboring channels, unless the carrier is modulated with techniques that ensure a constant envelope. However, the cost benefits of increasing the channel capacity of microwave radio systems have forced the development of linear modulation schemes where the transmitted signal has a fluctuating envelope. A class-A amplifier is capable of linear operation if it saturates at an RF level much superior to the required level, but such an inefficient amplifier is costly to purchase, operate, and to put into orbit for satellite applications. For these reasons, there is considerable industrial interest in producing linear amplifiers with good power efficiency. Both requirements can be reconciled by using external circuitry to linearize an efficient amplifier. Cartesian loop feedback is a relatively simple linearization method that has been used [1], but the development of more broad-band techniques, such as predistortion [2], [3] and feedforward, has been made necessary by the emergence of the new code-division multiple-access (CDMA) digital cellular system. The feedforward linearizer is by nature sensitive to changes in operating conditions, but the development of adaptation methods to compensate for such changes has renewed interest in the technique. Feedforward linearization is more effective than predistortion, but its use of an auxiliary amplifier and compensating delay line significantly degrades the efficiency of the linearized amplifier.

Published studies on the power efficiency of feedforward amplifiers [4]–[6] have assumed a two-tone test signal, which has much smaller envelope fluctuations than today's real-life CDMA or multichannel signals, and have made various other simplifying assumptions. Design rules-of-thumb now commonly used can, therefore, be viewed with suspicion. This paper describes a method of optimizing the dc–RF conversion efficiency of a feedforward amplifier for a given driving signal, subject to a required linearity and gain at the linearizer's output. The procedure uses the behavioral-level technique of quadrature modeling [7], which has been helpful in analyzing the adaptive behavior of feedforward amplifiers [8]. This technique is computationally efficient, as it uses the complex envelope of a signal, or the complex baseband signal, instead of the real signal. Since quadrature modeling assumes frequency responses to be flat over the simulation bandwidth, the procedure cannot account for frequency responses in the amplifiers, delay mismatches in the circuit, or adaptation schemes that use filtering. On the other hand, the procedure has the convenience of requiring knowledge of only the power statistics of the driving signal, provided that: 1) the adaptation uses no pilot signal and 2) the required output linearity is expressed as the ratio of total signal power to intermodulation power (SIMR) [8]. Under relatively broad-band CDMA excitation, the properly measured quadrature model of a typical amplifier will be shown to still predict spectral regrowth well enough for the purposes of this paper.

II. Overview

An overview of the optimization procedure will be provided by introducing a complex baseband model of the feedforward amplifier, followed by the lists and descriptions of the input and output parameters of the procedure.

A. Complex Baseband Feedforward Amplifier Model

The feedforward model of Fig. 1 allows a rigorous analysis involving all design parameters. Signal symbols denote complex voltage envelopes, and component symbols denote complex voltage gains. An input coupler with coupling and transmission gains of C_m and T_m sends a portion of input signal $\mathrm{v_m(t)}$ to the main amplifier, which outputs an amplified input signal $\mathrm{v_M(t)}$ containing IM distortion. The "main" coupler sends a small portion of this signal to the "error" coupler, which subtracts from it the remaining portion of $\mathrm{v_m(t)}$ to produce an error signal $\mathrm{v_e(t)}$. This closes the circuit's first loop also

Manuscript received October 11, 2001; revised March 9, 2002. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under a Postgraduate Scholarship.

The authors are with the Groupe de Recherches Avancées en Microondes et en Électronique Spatiale (Poly-Grames), Département de génie électrique, École Polytechnique de Montréal, Montréal, QC, Canada H3C 3A7.

Digital Object Identifier 10.1109/TMTT.2002.806946

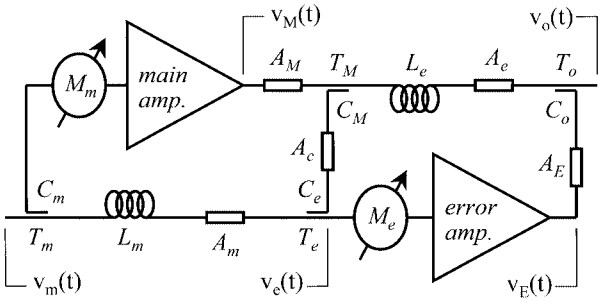


Fig. 1. Complex baseband feedforward amplifier model, with complex envelopes of signals and complex gains of components.

known as the signal cancellation loop because it cancels the signal component in $\mathrm{v_M(t)}$ in order to isolate its IM distortion. For signals to enter the coupler simultaneously, a delay line of gain L_m adds delay to $\mathrm{v_m(t)}$ as necessary to compensate for the longer delay introduced in $\mathrm{v_M(t)}$. The error signal $\mathrm{v_e(t)}$ exits the error amplifier as the amplified error signal $\mathrm{v_E(t)}$ after suffering its share of IM distortion. This signal is sent to an output coupler, with subtracts it from the main portion of $\mathrm{v_M(t)}$ to yield a largely corrected output signal $\mathrm{v_o(t)}$. This closes the circuit's second loop also known as the error cancellation loop because it cancels the error signal present in $\mathrm{v_M(t)}$. A delay line of gain L_e adds delay to $\mathrm{v_M(t)}$ as necessary to compensate for the longer delay introduced in $\mathrm{v_E(t)}$.

Gain balance must be maintained in each loop to ensure precise signal cancellations. For this purpose, any drifting in the amplifier gains is automatically compensated for by suitable adaptation of vector modulator gains M_m and M_e. In the literature, the vector modulators are sometimes seen in the alternate paths of each loop [4], [5], [8], but this allows the drifting in the amplifiers to influence the overall system gain. To avoid this, the vector modulators must be placed in the paths of the amplifiers, as shown, and are placed before the amplifiers in order to reduce power losses. Gains A_m, A_M, A_c, A_e, and A_E are included in this model in order to account for other possible attenuations in the branches of a practical circuit.

B. Input and Output Parameters

The input and output parameter lists shown below are longer than those in a preliminary version of this work [9] because no value of the main path power gain $|T_M L_e|^2$ is assumed here.

Input Parameters:

$\{x_m\}$	Set of instantaneous-to-average input powers.
$P(\{x_m\})$	Probabilities of $\{x_m\}$.
$\{\mathrm{SIMR_{req}}\}$	Required SIMRs at output.
U_{req}	Required linear power gain of system.
$f_M(d_M)$	Gain compression function of main amplifier.
$f_E(d_E)$	Gain compression function of error amplifier.
$h_M(d_M)$	DC power function of main amplifier.
$h_E(d_E)$	DC power function of error amplifier.
$\mathrm{adp_1}$	Adaptation method for first loop.
$\mathrm{adp_2}$	Adaptation method for second loop (pilotless).
$\mathrm{p_1}$	Delay-to-gain ratio of main amplifier (nanoseconds per decibel).
$\mathrm{p_2}$	Delay-to-gain ratio of error amplifier (nanoseconds per decibel).
τ_1	Residual delay in first loop (nanoseconds).
τ_2	Residual delay in second loop (nanoseconds).
$\mathrm{q_1}$	Delay-to-loss ratio of first delay line (nanoseconds per decibel).
$\mathrm{q_2}$	Delay-to-loss ratio of second delay line (nanoseconds per decibel).
L	Insertion gain of input coupler (<1).
P	Insertion gain of main coupler (<1).
N	Insertion gain of error coupler (<1).
R	Insertion gain of output coupler (<1).
$\lvert C_m\rvert^2$	Coupling of input coupler (<1).
$\lvert A_c\rvert^2$	Power gain in central branch (<1).
$\lvert A_m\rvert^2$	Power gain opposite main amplifier (<1).
$\lvert A_e\rvert^2$	Power gain opposite error amplifier (<1).
$\lvert A_M\rvert^2$	Power gain after main amplifier (<1).
$\lvert A_E\rvert^2$	Power gain after error amplifier (<1).
$\lvert M_m\rvert^2$	Power gain before main amplifier (<1).
$\lvert M_e\rvert^2$	Power gain before error amplifier (<1).
δ_1	Tolerance of first vector modulator gain.
δ_2	Tolerance of second vector modulator gain.
ε	Relative gain error in third loop.

Optimized Output Sets Corresponding to $\{\mathrm{SIMR_{req}}\}$*:*

$\{\lvert G_M\rvert^2\}$	Uncompressed power gains of main amplifier.
$\{\lvert G_E\rvert^2\}$	Uncompressed power gains of error amplifier.
$\{\lvert C_M\rvert^2\}$	Couplings of main coupler (<1).
$\{\lvert C_e\rvert^2\}$	Couplings of error coupler (<1).
$\{\lvert C_o\rvert^2\}$	Couplings of output coupler (<1).
$\{X_{satM}\}$	Normalized saturated powers of main amplifier.
$\{X_{satE}\}$	Normalized saturated powers of error amplifier.
$\{\eta\}$	DC–RF conversion efficiencies of system.

The input parameters above are listed in three groups according to their reasons for not being subject to optimization. The first group of parameters, which specify the input signal format and the required performance of the system, are by their very nature nonnegotiable. The second group of "parameters," consisting of the amplifier characteristics and the adaptation methods used, cannot be optimized by the current procedure in part because they cannot be reduced to single variables. Therefore, the optimization is, strictly speaking, a partial one, and many such optimizations using different amplifier types and adaptation schemes would be required to identify a final optimal design. The scope of this paper can only afford a glimpse into this type of activity, which has already produced significant findings [10]–[12]. The third group consists of those 21 parameters for which optimal solutions are already known. For example, the various delays, losses, and tolerances that reduce performance in a practical circuit have trivial optimal solutions of zero delays, zero losses, and zero tolerances. Values specified are the achievable values closest to optimal, which depends on a designer's access to the best technology.

The optimized outputs listed above come in sets that correspond to the set of required output SIMR values specified as input parameters. The list comprises the seven optimization parameters and the resulting system efficiency. How well the optimal solution associated with a required output SIMR can be approximated depends once again on the designer's access to the best technology.

1) Input Parameter Descriptions: With pilotless adaptation and the SIMR as the linearity criterion, the required knowledge

of the input signal $\mathrm{v_m}$ will later prove to be the averages of various functions of the instantaneous-to-average input power $x_m = \mathrm{p_m}/\mathrm{P_m}$. These averages can be estimated by using a set of powers x_{mk} and probabilities $P(x_{mk})$, $k = 1, \ldots, n$ that offer enough resolution to ensure solution convergence. Given a total probability of unity and the fact that

$$\mathrm{E}[x_m] = \mathrm{E}[\mathrm{p_m}/\mathrm{P_m}] = \mathrm{E}[\mathrm{p_m}]/\mathrm{P_m} = 1 \tag{1}$$

the statistical data set should satisfy the two conditions

$$\sum_{k=1}^{n} P(x_{mk}) = 1 \quad \sum_{k=1}^{n} P(x_{mk}) x_{mk} = 1. \tag{2}$$

Such a data set can be derived from the widely used complementary cumulative distribution function (ccdf) of x_m, which gives the probability of x_m exceeding any given value.

Working from a statistical description of the input signal means that the power spectral density of the output signal $\mathrm{v_o}$ is not available. However, it remains possible to break up the output signal power into the linear signal power and the IM power, or $\mathrm{P_o} = \mathrm{P_S} + \mathrm{P_{IM}}$, and then to describe the output signal linearity using $\mathrm{SIMR} = \mathrm{P_S}/\mathrm{P_{IM}}$. Furthermore, as explained in [8], the SIMR is related to the power spectral density ratio that regulatory authorities normally deal with, and is very roughly $10\log 3 = 4.8$ dB less favorable. In general, a set of required values $\{\mathrm{SIMR_{req}}\}$ is specified.

In a conservative manner, definitions involving the system's output power will use the linear output power $\mathrm{P_S}$ instead of the total output power $\mathrm{P_o}$, although the difference is small for reasonably large values of SIMR. Accordingly, the power gain U of the system is defined as the linear power gain $\mathrm{P_S}/\mathrm{P_m}$ rather than the total power gain $\mathrm{P_o}/\mathrm{P_m}$. A single required value U_{req} is specified.

Function pairs $f_M(d_M)$ and $h_M(d_M)$ for the main amplifier, and $f_E(d_E)$ and $h_E(d_E)$ for the error amplifier, allow to describe an amplifier (e.g., transistor type, class of operation) independently of its gain and saturated output power, which are to be optimized. The instantaneous input drive d_M or d_E is the reciprocal of the input backoff and is defined here as the ratio of the uncompressed output power to the saturated output power or

$$d = \frac{\mathrm{p_{in}}|G|^2}{\mathrm{P_{sat}}} = \frac{\mathrm{p_{in}}}{\mathrm{P_{sat}}/|G|^2}. \tag{3}$$

The uncompressed output power depends on what is understood to be the uncompressed gain G. In class-A operation, G is normally the small-signal complex gain, and for a solid-state amplifier, the reference input power $\mathrm{P_{sat}}/|G|^2$ typically exceeds the more usual 1-dB compression reference by less than 1 dB. In class-C operation, however, the small-signal definition of G is clearly unsuitable. In general, G can be defined as deemed convenient for a given class of operation, as long as the same definition is used in the specification of $f_M(d_M)$, $f_E(d_E)$, $h_M(d_M)$, and $h_E(d_E)$. The complex gain compression function $f_M(d_M)$ or $f_E(d_E)$ is the amplifier's output voltage relative to its uncompressed output voltage, while the normalized dc power function $h_M(d_M)$ or $h_E(d_E)$ is the amplifier's dc power consumption relative to its saturated output power. In equation form

$$f(d) = \frac{\mathrm{v_{out}}(d)}{\mathrm{v_{in}}(d)G} \quad h(d) = \frac{\mathrm{p_{dc}}(d)}{\mathrm{P_{sat}}}. \tag{4}$$

The magnitude and phase of $f(d)$ are particular expressions of the amplifier's AM/AM and AM/PM characteristics.

Parameters $\mathrm{adp_1}$ and $\mathrm{adp_2}$ are identifiers of the methods used to adapt the two vector modulators. Adaptation of the second vector modulator may use a pilot signal injected into the main amplified signal as a type of distortion that is easily monitored at the system's output by coherent detection or by filtering. Since such detection cannot be simulated within a statistical solution procedure, $\mathrm{adp_2}$, like $\mathrm{adp_1}$, must refer to a pilotless adaptation method, which uses only samples of existing signals.

Delay-to-gain ratios $\mathrm{p_1}$ and $\mathrm{p_2}$ specify linear relationships between the group delays of the two amplifiers and their decibel gains since both quantities are roughly proportional to the number of amplification stages. Residual delays τ_1 and τ_2 are the fixed delays required in lines L_m and L_e when the delays of the amplifiers are neglected. Delay-to-loss ratios $\mathrm{q_1}$ and $\mathrm{q_2}$ are quality figures for the delay lines. The delay and decibel loss of a line are linearly related because both are proportional to the length of the line.

The insertion loss of a directional coupler being the fraction of the power injected at one port that is dissipated within the unit [13], [14], the "insertion gain" of the coupler is defined here as that remaining fraction of power that is recovered at the transmit, coupled, or isolated ports. Accordingly, insertion gains L, P, N, and R are computed as

$$\begin{aligned} L &= |T_m|^2 + |C_m|^2 \\ P &= |T_M|^2 + |C_M|^2 \\ N &= |T_e|^2 + |C_e|^2 \\ R &= |T_o|^2 + |C_o|^2 \end{aligned} \tag{5}$$

neglecting losses at the isolated ports. Coupling $|C_m|^2$ is also specified in the input list because its optimal value is already known to be $L/(1 + \mathrm{q_1}/\mathrm{p_1})$, as will be shown. Although the other three couplings are yet to be determined, specifying insertion gains for these couplers is realistic because manufacturers typically specify a unique insertion loss value for a given type and frequency range of coupler, independently of the coupling.

The seven power gains that follow represent various power losses in the circuit in part due to the signal samplings needed for the adaptation. A maximum value of unity indicates zero loss. $|M_m|^2$ and $|M_e|^2$ include the attenuations of the vector modulators at the centers of their intended operating regions. These attenuations must be sufficient to allow the vector modulators to adapt to the anticipated ranges of drifting in the amplifiers. Hence, reducing drifting in the amplifiers may allow for better values of $|M_m|^2$ and $|M_e|^2$.

Tolerances in the gains of the two vector modulators are the results of inaccuracies in the responses of the devices to their control signals, as well as inaccuracies in the control signals themselves. As an example of the latter, power minimization methods require occasional perturbations in the vector modulator gains to determine whether they are still at their minimizing points [8]. A specified tolerance δ_1 in the first vector modulator gain M_m signifies

that this gain may depart from its intended value M_{m0} by a complex error of magnitude as high as $|M_{m0}|\delta_1$. Tolerance δ_2 has the same meaning in reference to M_e.

Memoryless and pilotless adaptation of the second loop requires a third circuit loop that compares a sample of the amplified error signal $\mathrm{v_E}$ or the output signal $\mathrm{v_o}$ to a reference signal. To model an error in practical implementation, complex number ε specifies a relative gain error in the branch that supplies the reference signal.

2) Output Parameter Descriptions: $\{|G_M|^2\}$ and $\{|G_E|^2\}$ are the power gains of the main and error amplifiers according to the definitions of gain that were used in specifying functions $f_M(d_M)$, $h_M(d_M)$, $f_E(d_E)$, and $h_E(d_E)$. $\{|C_M|^2\}$, $\{|C_e|^2\}$, and $\{|C_o|^2\}$ are the couplings of the main, error, and output couplers, respectively. Values superior to 0.5 are numerically possible and signify that the coupled branch of the coupler has become its transmit branch and vice versa. A value of 0.5 refers to a hybrid junction. $\{X_{satM}\}$ and $\{X_{satE}\}$ are the saturated output powers of the amplifiers relative to the required linear output power $\mathrm{P_{S,req}} = \mathrm{P_m}U_{\mathrm{req}}$. This normalization is expected since neither $\mathrm{P_{S,req}}$, nor $\mathrm{P_m}$ are listed among the input parameters. For a given $\mathrm{P_{S,req}}$, the absolute saturated output powers are obtained by using

$$\{\mathrm{P_{satM}}\} = \mathrm{P_{S,req}}\{X_{satM}\} \quad \{\mathrm{P_{satE}}\} = \mathrm{P_{S,req}}\{X_{satE}\}. \tag{6}$$

Finally, $\{\eta\}$ is the resultant set of dc–RF conversion efficiencies for the system. This efficiency is defined as

$$\eta = \frac{\mathrm{P_S}}{\mathrm{P_{dcM}} + \mathrm{P_{dcE}}} \tag{7}$$

where $\mathrm{P_{dcM}}$ and $\mathrm{P_{dcE}}$ are the average dc power consumptions of the main and error amplifiers. The more complicated power-added efficiency (PAE) is not used because for a fixed system gain, it is related to η by a constant multiplier [15]. The dc consumptions of the vector modulators and control circuitry are not included because they are considered constant and, hence, irrelevant for optimization purposes.

III. Optimization of Power Efficiency

A. Summary

This section derives the procedure to optimize the power efficiency of the feedforward amplifier model in Fig. 1. The model is normalized beforehand, as detailed in Section III-B, thus simplifying the derivations considerably. For a given input signal, the normalized model is used to derive the output SIMR, gain, and power efficiency of the feedforward amplifier as a function of its optimization parameters so that, given a required output SIMR and gain, a constrained optimization of the power efficiency can be performed with respect to the optimization parameters. A prerequisite task is to obtain the vector modulator gains that balance the first loop (Section III-C) and the second (Section III-D). Hence, the output SIMR is found to be a nonanalytic function of only the two reference input drives D_{0M} and D_{0E}. The gain and power efficiency are found to be nonanalytic functions of the same, but also analytic functions of other optimization parameters (Section III-E). A partial optimization reduces these parameters and the power efficiency to nonanalytic functions of D_{0M} and D_{0E}, like the output SIMR. A graphical approach identifies the values of D_{0M} and D_{0E} that optimize the efficiency for given values of the output SIMR and, thus, the fully optimized efficiencies and optimization parameters emerge (Section III-F).

B. Normalized Feedforward Amplifier Model

To analyze the complex baseband feedforward amplifier model, the power amplifier model described in (4) is first expressed in block form as in Fig. 2(a) and substituted into Fig. 1 with G_M, $\mathrm{P_{satM}}$, $f_M(d_M)$, and $h_M(d_M)$ for the main amplifier and with G_E, $\mathrm{P_{satE}}$, $f_E(d_E)$, and $h_E(d_E)$ for the error amplifier. The input drives, therefore, become

$$d_M = \frac{\mathrm{p_m}|C_m M_m|^2 |G_M|^2}{\mathrm{P_{satM}}} \quad d_E = \frac{\mathrm{p_e}|M_e|^2 |G_E|^2}{\mathrm{P_{satE}}}. \tag{8}$$

Gain A_M, followed by gains $A_M C_M A_c C_e$ and $A_E C_o$ are then distributed in the circuit to produce the diagram of Fig. 2(b) with the normalization gains

$$G_1 = -\frac{T_m L_m A_m T_e}{A_M C_M A_c C_e} \quad G_2 = -\frac{A_M T_M L_e A_e T_o}{A_E C_o}. \tag{9}$$

In Fig. 2(c), the model is finally reduced to four components by normalizing it using G_1 and G_2. The original vector modulator gains are now normalized as

$$M_1 = \frac{C_m M_m G_M}{G_1} \quad M_2 = \frac{A_M C_M A_c C_e M_e G_E}{G_2}. \tag{10}$$

All signals have been further normalized by the constant $\sqrt{2\mathrm{Z_0 P_m}}$, where $\mathrm{Z_0}$ is the reference impedance. The resulting normalized signals are denoted using italics to distinguish them from their absolute counterparts. In particular

$$\begin{aligned} v_m &= \frac{\mathrm{v_m}}{\sqrt{2\mathrm{Z_0 P_m}}} \\ v_e &= \frac{\mathrm{v_e}}{\sqrt{2\mathrm{Z_0 P_m}} G_1 A_M C_M A_c C_e} \\ v_o &= \frac{-\mathrm{v_o}}{\sqrt{2\mathrm{Z_0 P_m}} G_1 G_2 A_E C_o} \end{aligned} \tag{11}$$

so that the squared magnitude

$$|v_m|^2 = \frac{|\mathrm{v_m}|^2}{2\mathrm{Z_0 P_m}} = \frac{\mathrm{p_m}}{\mathrm{P_m}} = x_m \tag{12}$$

is the instantaneous-to-average input power. The highest possible value of x_m, if finite, is the input signal's peak-to-average power ratio, otherwise known as the crest factor.

With three circuit branches now having unity gain, the values of the normalized vector modulator gains will differ from the reference solution $M_1 = M_2 = 1$ only to the extent that the gain compression functions deviate from unity. The amplified input signal is now the input signal itself plus the error signal, or $v_M = v_m + v_e$ and, similarly, the amplified error signal is the error signal itself plus an unwanted residual signal or $v_E = v_e + v_r$. The output signal is the input signal itself minus this residual signal or $v_o = v_M - v_E = v_m - v_r$. Thus, as long as a linear error amplifier in a balanced second loop ensures a small v_r, the present feedforward configuration takes on the added role of automatic gain and phase controller, delivering an invariant gain normalized here to unity. The gain is unaffected by drifting in the amplifier gains, or

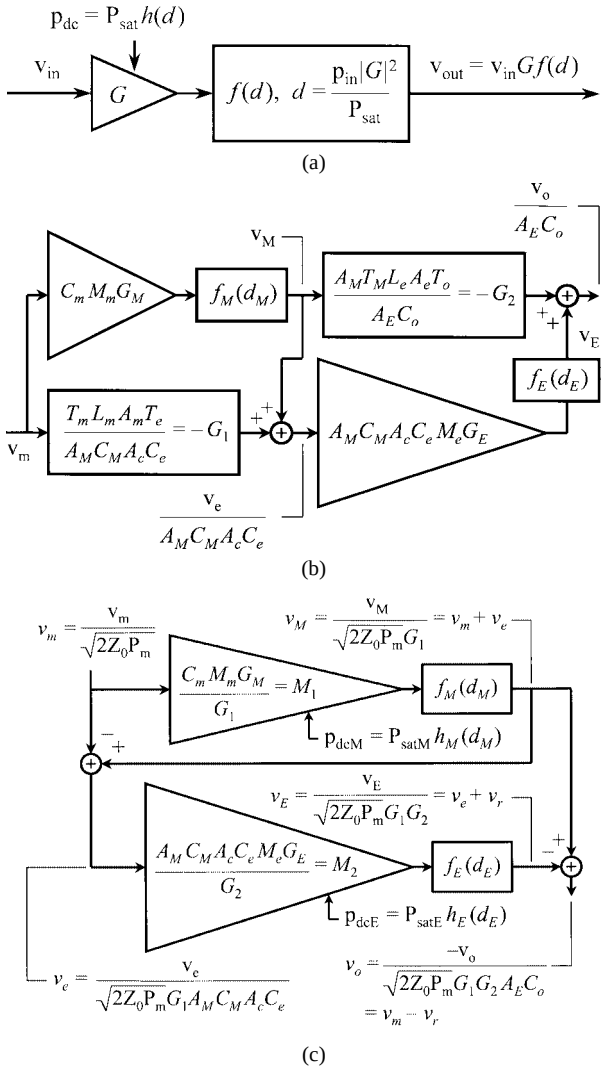


Fig. 2. Complex baseband feedforward amplifier models, with complex signal envelopes and complex gains. (a) Power amplifier model only. (b) Complete system, partially normalized. (c) Fully normalized.

even by an imbalance in the first loop. A balanced first loop, however, minimizes the error signal power and, thus, minimizes the required error amplifier power rating.

C. Balancing the First Loop

Combining (8), (10), and (12), the main amplifier's instantaneous input drive can be expressed as

$$d_M(x_m, D_{0M}, M_1) = x_m D_{0M} |M_1|^2 \tag{13}$$

featuring the main amplifier's reference input drive

$$D_{0M} = \frac{\mathrm{P_m}|G_1|^2}{\mathrm{P_{satM}}}. \tag{14}$$

Given that $v_M = v_m M_1 f_M(d_M)$, the instantaneous fractional imbalance in the first loop and the instantaneous error power (normalized) may be obtained as

$$\varepsilon_1(x_m, D_{0M}, M_1) = \frac{v_e}{v_m} = \frac{v_M - v_m}{v_m} = M_1 f_M\big(d_M(x_m, D_{0M}, M_1)\big) - 1 \tag{15}$$

$$x_e(x_m, D_{0M}, M_1) = |v_e|^2 = \big|v_m \varepsilon_1\big|^2 = x_m \big|\varepsilon_1(x_m, D_{0M}, M_1)\big|^2. \tag{16}$$

The value of M_1 that balances the first loop will depend on the adaptation method used. In the power minimization method, M_1 minimizes the average error power

$$\mathrm{E}\big[x_e(x_m, D_{0M}, M_1)\big] = X_e(D_{0M}, M_1). \tag{17}$$

The circuitry to do power minimization is simple, but the method has drawbacks alluded to earlier. In the alternative known as the gradient method [8], a gradient signal for the adaptation of M_1 is obtained from the correlation of the error and input signals

$$\begin{aligned}\mathrm{E}[v_e v_m^*] &= \mathrm{E}[v_m \varepsilon_1 v_m^*] \\ &= \mathrm{E}\big[x_m \varepsilon_1(x_m, D_{0M}, M_1)\big] \\ &= X_{em}(D_{0M}, M_1).\end{aligned} \tag{18}$$

When the loop is balanced, M_1 nulls the above correlation or, equivalently, minimizes to zero its squared magnitude $|X_{em}(D_{0M}, M_1)|^2$. Thus, in either method, the solution can be found by a two-dimensional search on a real-valued function over the M_1 complex plane. The gradient method is widely thought to be equivalent to power minimization in that it minimizes the average error power. However, according to linear estimation theory [16], this would be exactly true only if the vector modulator were placed in the linear path of the input signal, which is opposite the main amplifier. In the present case, the two schemes are not equivalent, but regardless of the method, the optimal solution M_{10} is reduced to a function of only D_{0M}. The worst case error, which it is appropriate to assume, has M_1 somewhere on the circle of radius $|M_{10}|\delta_1$ centered on M_{10} or

$$M_1(D_{0M}, \phi_1) = M_{10}(D_{0M}) + \big|M_{10}(D_{0M})\big|\delta_1 e^{j\phi_1} \tag{19}$$

from which follow the functions

$$\begin{aligned}&d_M(x_m, D_{0M}, \phi_1) \\ &\varepsilon_1(x_m, D_{0M}, \phi_1) \\ &x_e(x_m, D_{0M}, \phi_1) \\ &X_e(D_{0M}, \phi_1) \\ &X_{em}(D_{0M}\phi_1).\end{aligned} \tag{20}$$

D. Balancing the Second Loop

Next, combining (8), (10), and (11), the error amplifier's instantaneous input drive can be expressed as

$$d_E(x_m, D_{0M}, \phi_1, D_{0E}, M_2) = x_e(x_m, D_{0M}, \phi_1) D_{0E} |M_2|^2 \tag{21}$$

featuring the error amplifier's reference input drive

$$D_{0E} = \frac{\mathrm{P_m}|G_1 G_2|^2}{\mathrm{P_{satE}}}. \tag{22}$$

Given that $v_E = v_e M_2 f_E(d_E)$, the instantaneous fractional

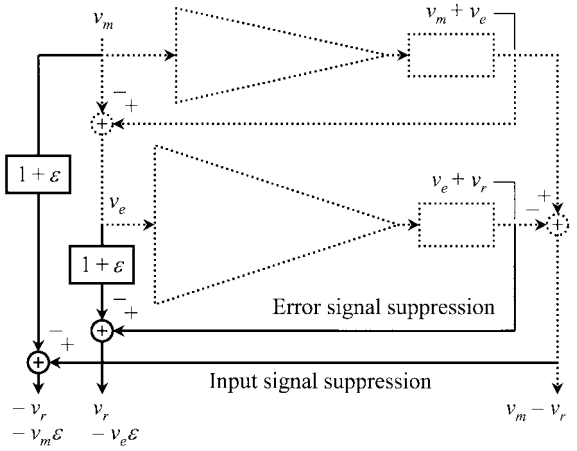


Fig. 3. Third circuit loop added to Fig. 2(c) for memoryless and pilotless adaptation of second loop with error or input signal suppression.

imbalance in the second loop may be obtained as

$$\begin{aligned}\varepsilon_2(x_m, D_{0M}, \phi_1, D_{0E}, M_2) &= \frac{v_r}{v_e} \\ &= \frac{v_E - v_e}{v_e} \\ &= M_2 f_E(d_E(x_m, D_{0M}, \phi_1, D_{0E}, M_2)) - 1. \end{aligned} \quad (23)$$

By analogy with M_1, M_2 in the power minimization method minimizes the average residual power

$$\begin{aligned}\mathrm{E}\left[|v_r|^2\right] &= \mathrm{E}\left[|v_e\varepsilon_2|^2\right] \\ &= \mathrm{E}\left[x_e(x_m, D_{0M}, \phi_1)|\varepsilon_2(x_m, D_{0M}, \phi_1, D_{0E}, M_2)|^2\right] \\ &= X_r(D_{0M}, \phi_1, D_{0E}, M_2) \end{aligned} \quad (24)$$

and, in the gradient method, which does not exactly minimize the residual power, M_2 nulls the correlation between the residual and error signals

$$\begin{aligned}\mathrm{E}[v_r v_e^*] &= \mathrm{E}[v_e\varepsilon_2 v_e^*] \\ &= \mathrm{E}\left[x_e(x_m, D_{0M}, \phi_1)\varepsilon_2(x_m, D_{0M}, \phi_1, D_{0E}, M_2)\right] \\ &= X_{re}(D_{0M}, \phi_1, D_{0E}, M_2). \end{aligned} \quad (25)$$

In practice, the residual signal is not available on its own and must be extracted either from the amplified error signal $v_e + v_r$ by suppressing v_e [17] or from the output signal $v_m - v_r$ by suppressing v_m [8]. These methods will, therefore, be referred to as error signal suppression and input signal suppression. Fig. 3 shows the third circuit loop that must be added to Fig. 2(c) for each of these purposes. References [8] and [18] insist with good reason that the reference branch for input signal suppression be adaptive because they place the first vector modulator opposite the main amplifier, where it influences the gain of the output signal. Such adaptation is unnecessary with the present feedforward configuration.

Fig. 3 allows the gain of each reference branch to depart from unity by a fixed complex error ε. When ε is not zero, error and input signal suppression cease to be equivalent, as they leak, respectively, some amount of v_e and v_m into the extracted v_r. With a nonzero ε and *error* signal suppression, the expressions, minimized and nulled by M_2 in the power and gradient methods, respectively, are now

$$\mathrm{E}\left[|v_r - v_e\varepsilon|^2\right] = X_r - 2\mathrm{Re}\{X_{re}\varepsilon^*\} + X_e|\varepsilon|^2 \quad (26)$$

$$\mathrm{E}[(v_r - v_e\varepsilon)v_e^*] = X_{re} - X_e\varepsilon. \quad (27)$$

The last term in (26) is optional because $X_e(D_{0M}, \phi_1)$ is independent of M_2. The effect of the new terms in the power and gradient expressions is obvious in the case of a perfectly linear error amplifier. The amplified error signal then converges toward $v_E = v_e(1+\varepsilon)$ with either adaptation method, leaving an incompletely canceled error signal $v_r = v_e\varepsilon$ at the output. Now, with a nonzero ε and *input* signal suppression, and using (1) and

$$\begin{aligned}&\mathrm{E}[v_r v_m^*] \\ &\quad = \mathrm{E}[v_e\varepsilon_2 v_m^*] \\ &\quad = \mathrm{E}[v_m\varepsilon_1\varepsilon_2 v_m^*] \\ &\quad = \mathrm{E}\left[x_m\varepsilon_1(x_m, D_{0M}, \phi_1)\varepsilon_2(x_m, D_{0M}, \phi_1, D_{0E}, M_2)\right] \\ &\quad = X_{rm}(D_{0M}, \phi_1, D_{0E}, M_2) \end{aligned} \quad (28)$$

the expressions, minimized and nulled by M_2 in the power and gradient methods, respectively, are

$$\mathrm{E}\left[|v_r + v_m\varepsilon|^2\right] = X_r + 2\mathrm{Re}\{X_{rm}\varepsilon^*\} + |\varepsilon|^2 \quad (29)$$

$$\mathrm{E}[(v_r + v_m\varepsilon)v_e^*] = X_{re} + X_{em}^*\varepsilon. \quad (30)$$

The last term in (29) is optional. In the gradient method, ε has no effect on the convergence if the gradient method was also used in the first loop to make $X_{em} = 0$. The output signal $v_o = v_m - v_r$ itself can then be used with no signal suppression, which corresponds to an error of $\varepsilon = -1$. However, as reported in [8], using the output signal in this way requires a tolerance in X_{em} that is not entirely realistic. It actually prohibits the use of the power minimization method in the first loop, which does not make X_{em} converge to zero. Moreover, the dominance of v_m over v_r requires extremely long averaging times and leads to slow convergence.

Depending on the adaptation method used, M_2 minimizes either a power or the squared magnitude of a correlation. Thus, in all cases, the solution can be found by a two-dimensional search on a real-valued function over the M_2 complex plane. Since the expression minimized is invariably a function of $(D_{0M}, \phi_1, D_{0E}, M_2)$, the optimal solution M_{20} is reduced to a function of only (D_{0M}, ϕ_1, D_{0E}). The worst case error, which it is appropriate to assume, has M_2 somewhere on the circle of radius $|M_{20}|\delta_2$ centered on M_{20} or

$$\begin{aligned}&M_2(D_{0M}, \phi_1, D_{0E}, \phi_2) \\ &\quad = M_{20}(D_{0M}, \phi_1, D_{0E}) + \left|M_{20}(D_{0M}, \phi_1, D_{0E})\right|\delta_2 e^{j\phi_2} \end{aligned} \quad (31)$$

from which follow the functions

$$\begin{aligned}&d_E(x_m, D_{0M}, \phi_1, D_{0E}, \phi_2) \\ &X_r(D_{0M}, \phi_1, D_{0E}, \phi_2) \\ &X_{rm}(D_{0M}, \phi_1, D_{0E}, \phi_2). \end{aligned} \quad (32)$$

E. Output SIMR, Gain, and Power Efficiency

Having solved for M_1 and M_2 allows the derivation of the circuit's output linearity, defined here as the SIMR. The output SIMR is found by breaking up the output signal as

$$v_o = v_m - v_r = v_m G_S + v_{\rm IM} \tag{33}$$

where the linear signal gain G_S minimizes the average power in the IM distortion $v_{\rm IM}$. Since this also uncorrelates $v_{\rm IM}$ with v_m, the total output power is the sum of the powers in the linear and IM signals

$$X_o = \mathrm{E}\left[|v_m G_S + v_{\rm IM}|^2\right] = |G_S|^2 + X_{\rm IM}. \tag{34}$$

The output SIMR is $|G_S|^2/X_{\rm IM}$ and is obtained in terms of v_o by using (33) in the correlation of $v_{\rm IM}$ and v_m

$$X_{{\rm IM}m} = \mathrm{E}[(v_o - v_m G_S)v_m^*] = X_{om} - G_S. \tag{35}$$

Since $X_{{\rm IM}m} = 0$, $G_S = X_{om}$, and combining this with (34)

$$\mathrm{SIMR} = \frac{|G_S|^2}{X_{\rm IM}} = \frac{|X_{om}|^2}{X_o - |X_{om}|^2}. \tag{36}$$

Inserting now $v_o = v_m - v_r$ into (36) yields

$$G_S(D_{0M}, \phi_1, D_{0E}, \phi_2) = 1 - X_{rm}(D_{0M}, \phi_1, D_{0E}, \phi_2) \tag{37}$$

$$X_{\rm IM}(D_{0M}, \phi_1, D_{0E}, \phi_2) = X_r(D_{0M}, \phi_1, D_{0E}, \phi_2) - \left|X_{rm}(D_{0M}, \phi_1, D_{0E}, \phi_2)\right|^2. \tag{38}$$

Equations (33), (37), and (38) agree that if v_r and v_m are already uncorrelated, $G_S = 1$ and $X_{\rm IM} = X_r$. SIMR follows as

$$\mathrm{SIMR}(D_{0M}, \phi_1, D_{0E}, \phi_2) = \frac{|G_S(D_{0M}, \phi_1, D_{0E}, \phi_2)|^2}{X_{\rm IM}(D_{0M}, \phi_1, D_{0E}, \phi_2)}. \tag{39}$$

One or both angles may be omitted as parameters if their associated tolerances δ_1 and δ_2 are zero. The remaining angles, if any, are searched for the values that minimize SIMR, which is the worst case scenario. The one- or two-dimensional search must be done with care, as local minima may exist. Thus, the angles become functions of (D_{0M}, D_{0E}), from which follow the functions

$$\begin{aligned} &M_1(D_{0M}, D_{0E}) \\ &d_M(x_m, D_{0M}, D_{0E}) \\ &M_2(D_{0M}, D_{0E}) \\ &d_E(x_m, D_{0M}, D_{0E}) \\ &G_S(D_{0M}, D_{0E}) \\ &\mathrm{SIMR}(D_{0M}, D_{0E}). \end{aligned} \tag{40}$$

SIMR has, therefore, become a function of only the two reference input drives D_{0M} and D_{0E}. Also made possible is the derivation of the linear power gain U. Substituting $v_o = v_m G_S$ in (11) and then using $U = |\mathrm{v_o}/\mathrm{v_m}|^2$ yields

$$U = |G_1 G_2|^2 |A_E C_o|^2 \left|G_S(D_{0M}, D_{0E})\right|^2 \tag{41}$$

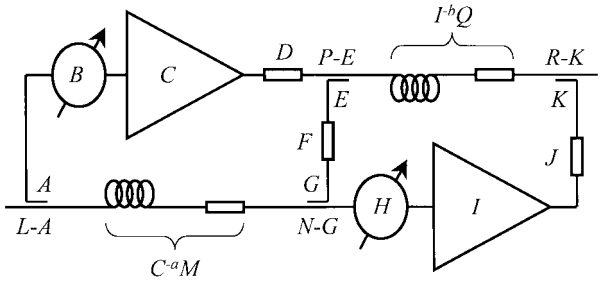


Fig. 4. Feedforward amplifier model with real power gains of components.

or if (9) is used to replace $|G_2|^2$

$$U = |G_1|^2 |A_M T_M L_e A_e T_o|^2 |G_S(D_{0M}, D_{0E})|^2. \tag{42}$$

Finally, the circuit's dc–RF conversion efficiency η can be derived. Inserting the amplifiers' instantaneous dc power consumptions $\mathrm{p_{dcM}} = \mathrm{P_{satM}} h_M(d_M)$ and $\mathrm{p_{dcE}} = \mathrm{P_{satE}} h_E(d_E)$ into (7) yields

$$\eta = \frac{1}{X_{satM} H_M(D_{0M}, D_{0E}) + X_{satE} H_E(D_{0M}, D_{0E})} \tag{43}$$

featuring the average dc powers normalized to the saturated output powers

$$\begin{aligned} H_M(D_{0M}, D_{0E}) &= \mathrm{E}\left[h_M\Big(d_M\big(x_m, D_{0M}, D_{0E}\big)\Big)\right] \\ H_E(D_{0M}, D_{0E}) &= \mathrm{E}\left[h_E\Big(d_E\big(x_m, D_{0M}, D_{0E}\big)\Big)\right] \end{aligned} \tag{44}$$

and the normalized saturated output powers

$$X_{satM} = \frac{\mathrm{P_{satM}}}{\mathrm{P_S}} \quad X_{satE} = \frac{\mathrm{P_{satE}}}{\mathrm{P_S}}. \tag{45}$$

These two saturated powers can also be expressed in terms of D_{0M} and D_{0E} by substituting $\mathrm{P_S} = \mathrm{P_m} U$, then (42) and (41) for U, respectively, and then by using (14) and (22). This yields

$$\begin{aligned} X_{satM} &= \frac{1}{D_{0M} |A_M T_M L_e A_e T_o|^2 |G_S(D_{0M}, D_{0E})|^2} \\ X_{satE} &= \frac{1}{D_{0E} |A_E C_o|^2 |G_S(D_{0M}, D_{0E})|^2}. \end{aligned} \tag{46}$$

It follows that η, like U, is a function of (D_{0M}, D_{0E}) plus the attenuations of several circuit components.

F. Optimal Power Efficiency

The power-efficiency function η may now be optimized with respect to its variables, subject to the constraints imposed by a required SIMR and U. The function's dependency on (D_{0M}, D_{0E}) is not analytic, but since these two variables are independent ones, as they alone involve $\mathrm{P_{satM}}$ and $\mathrm{P_{satE}}$, a logical first step is to optimize η with respect to its other variables for any given (D_{0M}, D_{0E}). This partial optimization is made much easier to manage by a temporary change of notation whereby power gains are denoted with single uppercase letters instead of squared magnitudes of voltage gains. In the power diagram of Fig. 4, symbols A–K denote the power gains along

the S-shaped path containing the two amplifiers. Comparing with Fig. 1

$$\begin{aligned}|C_m|^2 &= A\\ |M_m|^2 &= B\\ |G_M|^2 &= C\\ |A_M|^2 &= D\\ |C_M|^2 &= E\\ |A_c|^2 &= F\\ |C_e|^2 &= G\\ |M_e|^2 &= H\\ |G_E|^2 &= I\\ |A_E|^2 &= J\\ |C_o|^2 &= K. \end{aligned} \tag{47}$$

If power gains B, D, F, H, and J are already known, the amplifier gains C and I and the couplings A, E, G, and K are the only six unknowns. The remaining power gains in Fig. 4 are expressed in terms of these six unknowns. Using the known insertion gains in (5), the four transmission gains in the couplers can be expressed as

$$\begin{aligned}|T_m|^2 &= L - A\\ |T_M|^2 &= P - E\\ |T_e|^2 &= N - G\\ |T_o|^2 &= R - K. \end{aligned} \tag{48}$$

Also, in the first loop, the delay in the delay line must equal the delay in the main amplifier plus the residual delay or

$$\mathrm{q}_1\left(-10\log|L_m|^2\right) = \mathrm{p}_1(10\log C) + \tau_1. \tag{49}$$

Therefore, the composite power gain in L_mA_m, and likewise for L_eA_e, can be expressed as

$$|L_mA_m|^2 = C^{-a}M \quad |L_eA_e|^2 = I^{-b}Q \tag{50}$$

where

$$\begin{aligned} a &= \mathrm{p}_1/\mathrm{q}_1\\ b &= \mathrm{p}_2/\mathrm{q}_2 \end{aligned} \tag{51}$$

$$\begin{aligned} M &= 10^{-\tau_1/10\mathrm{q}_1}|A_m|^2\\ Q &= 10^{-\tau_2/10\mathrm{q}_2}|A_e|^2 \end{aligned} \tag{52}$$

are known, and exponents a and b are typically much less than unity. Substituting the new symbols and

$$V(D_{0M}, D_{0E}) = \left|G_S(D_{0M}, D_{0E})\right|^2 \tag{53}$$

into (46) yields the normalized saturated output powers

$$\begin{aligned} X_{satM} &= \frac{1}{D_{0M}D(P-E)I^{-b}Q(R-K)V}\\ X_{satE} &= \frac{1}{D_{0E}JKV}. \end{aligned} \tag{54}$$

These must be made to maximize efficiency in (43) for a given (D_{0M}, D_{0E}). Since SIMR is precisely a function of (D_{0M}, D_{0E}), $\mathrm{SIMR} = \{\mathrm{SIMR}_{\mathrm{req}}\}$ is not yet an optimization constraint. The three constraints at this point are the required power gain balances in the two loops, as well as $U = U_{\mathrm{req}}$. Eliminating $|G_1|^2$ and $|G_2|^2$, respectively, from the power equivalents of (9) and (10), and expanding $U = U_{\mathrm{req}}$ using (41) and (10), the three constraints become

$$\frac{ABCDEFG}{(L-A)C^{-a}M(N-G)} = S \tag{55}$$

$$\frac{EFGHIJK}{(P-E)I^{-b}Q(R-K)} = T \tag{56}$$

$$ABCDEFGHIJK = STW \tag{57}$$

where

$$S(D_{0M}, D_{0E}) = |M_1(D_{0M}, D_{0E})|^2 \tag{58}$$

$$T(D_{0M}, D_{0E}) = |M_2(D_{0M}, D_{0E})|^2 \tag{59}$$

$$W(D_{0M}, D_{0E}) = \frac{U_{\mathrm{req}}}{V(D_{0M}, D_{0E})}. \tag{60}$$

The three constraints above can be reduced to one if two unknowns are removed from the discussion. A useful choice is to eliminate C and I by inserting (55) and (56) into (57), yielding the single constraint

$$\begin{aligned} W = &\left[\left(\frac{ABD}{S}\right)^a (L-A)M(N-G)\right]^{1/(1+a)}\\ &\cdot\left(\frac{1}{EFG}\right)^{(1-ab)/((1+a)(1+b))}\\ &\cdot\left[\left(\frac{HJK}{T}\right)^b (P-E)Q(R-K)\right]^{1/(1+b)} \end{aligned} \tag{61}$$

and by inserting (56) into (54) to eliminate I, yielding the normalized saturated output powers

$$\begin{aligned} X_{satM} &= \frac{1}{D_{0M}D}\left[\left(\frac{T}{EFGHJK}\right)^b \cdot\frac{1}{(P-E)Q(R-K)}\right]^{1/(1+b)}\frac{1}{V}\\ X_{satE} &= \frac{1}{D_{0E}JKV}. \end{aligned} \tag{62}$$

Any solution that satisfies (61) can be updated by bringing A closer to $L/(1+1/a) = L/(1+\mathrm{q}_1/\mathrm{p}_1)$ and then by raising G as near N as necessary to restore the broken equality in (61). The second change decreases X_{satM} in (62) so as to increase the efficiency in (43). It follows that A has the nontrivial optimal value of $L/(1+\mathrm{q}_1/\mathrm{p}_1)$, independently of (D_{0M}, D_{0E}). Similar reasoning confirms the obvious fact that the optimal B, D, F, H, J, L, M, N, P, Q, and R are their highest achievable values, thus minimizing losses in the circuit. The optimal values of the other couplings E, G, and K cannot be verified analytically, but a numerical solution procedure can be obtained as described below. The three circuit equations (55)–(57) are first rewritten

as

$$cEG = N - G \tag{63}$$

$$dEGK = (P - E)(R - K) \tag{64}$$

$$eEGK = 1 \tag{65}$$

where

$$c = \frac{ABC^{1+a}DF}{S(L-A)M} \quad d = \frac{FHI^{1+b}J}{TQ} \quad e = \frac{ABCDFHIJ}{STW} \tag{66}$$

and eliminating G and K instead of C and I leads to the quadratic solution

$$E = \frac{-a_1 - \sqrt{a_1^2 - 4a_2a_0}}{2a_2} \tag{67}$$

$$a_2 = eNR - c$$

$$a_1 = dN - a_2P - 1$$

$$a_0 = P. \tag{68}$$

It can be shown that the alternative quadratic solution with a positive square root in (67) can never return the desired E. For any (C, I) pair, the solution obtained for E allows to compute G through (63), then K through (64) or (65), and then the normalized saturated output powers using

$$X_{satM} = \frac{AB}{D_{0M}SU_{\text{req}}}C \quad X_{satE} = \frac{1}{D_{0E}JV}\frac{1}{K} \tag{69}$$

as obtained by combining (54), (56), (57), and (60) in this order. The efficiency η is finally computed via (43). It follows that η can be maximized by a two-dimensional search in (C, I). This procedure reduces η to a nonanalytic function of (D_{0M}, D_{0E}), just like the remaining SIMR constraint. This suggests a simple graphical method of optimizing η with respect to (D_{0M}, D_{0E}). By computing the values of SIMR and η on a sufficiently fine rectangular grid in the $D_{0M} - D_{0E}$ plane, accurate contours of the two functions can be plotted on the same plane, as shown in the example of Fig. 5(a). The maximum of efficiency for a given SIMR can then be located by tracking η along the relevant SIMR contour. The values of the seven design parameters $C = |G_M|^2$, $I = |G_E|^2$, $E = |C_M|^2$, $G = |C_e|^2$, $K = |C_o|^2$, X_{satM}, and X_{satE} computed as intermediate steps must have been stored to enable their interpolation at the same location in the $D_{0M} - D_{0E}$ plane. For a sequence of contours $\text{SIMR} = \{\text{SIMR}_{\text{req}}\}$, as in Fig. 5(a), the behaviors of the optimized variables as a function of SIMR are best conveyed in the type of plot shown in Fig. 5(b), which pictures the seven design parameters in decibels and η in percent versus SIMR in decibels. By convention, couplings are represented by positive decibel values.

IV. Results

This section illustrates with a representative example the steps involved in optimizing power efficiency. To further illustrate the usefulness of the method, additional results are obtained in response to changes in vector modulator tolerances, input signal format, and error amplifier class.

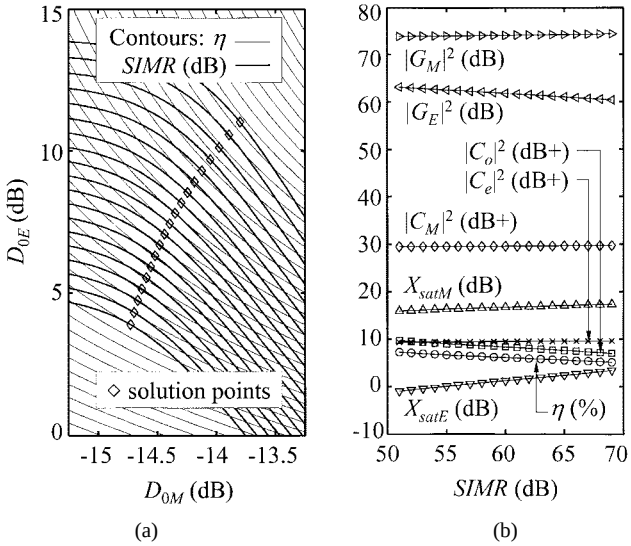


Fig. 5. (a) Contours of output SIMR and power efficiency in the D_{0M}–D_{0E} plane, and points of optimal efficiency. (b) Corresponding optimal values of amplifier gains, couplings, normalized saturated output powers, and efficiency.

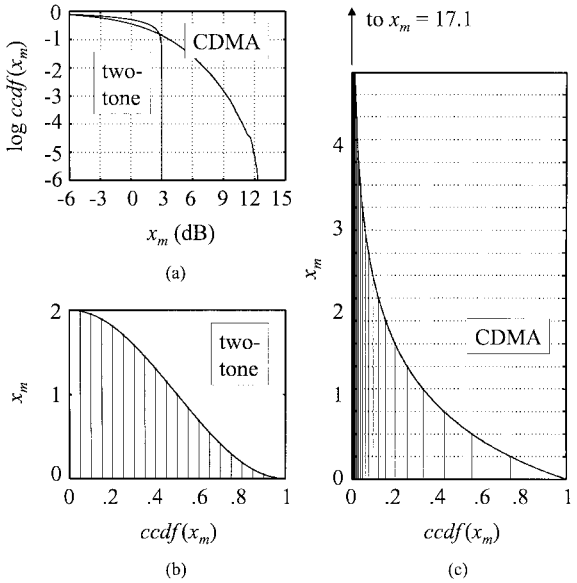


Fig. 6. Ccdfs of a two-tone signal and a three-carrier IS-95A CDMA signal. (a) On the same standard logarithmic plot. (b) and (c) On separate linear plots and sampled horizontally and vertically, respectively.

A. Input Signals

In what will be called the reference simulation, a three-carrier IS-95A CDMA waveform [19] was used as the input test signal. The waveform consists of 524 288 complex envelope samples that were used to approximate the ccdf of the normalized power in Fig. 6(a), in contrast to that of a two-tone signal. The crest factor exceeds 12 dB, compared to 3 dB for a two-tone signal. For the purposes of discretization, the ccdfs of both signals are replotted as shown in Fig. 6(b) and (c), where the x_m and ccdf axes are linear and have been swapped. In Fig. 6(b), this corresponds to a descending half-cycle of the two-tone signal's power envelope, which is sinusoidal. There, the horizontal scale is subdivided into 20 equal intervals of probability 1/20 for which average values of x_m can be computed. In trials, this small

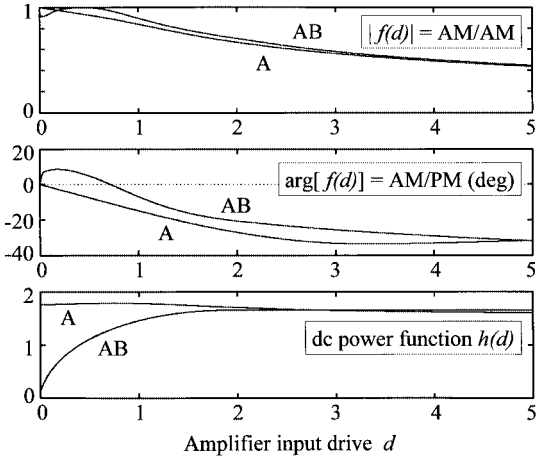
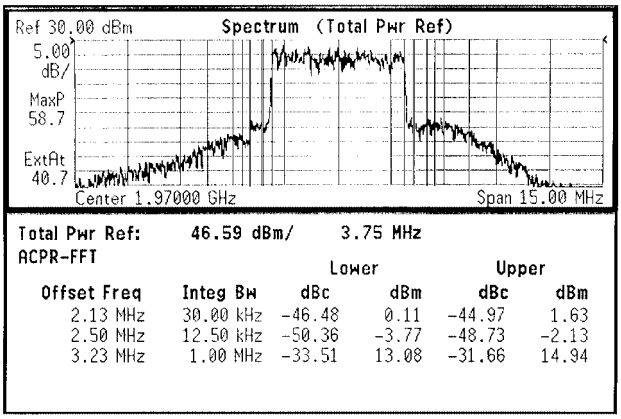


Fig. 7. Gain compression function $f(d)$ and dc power function $h(d)$ for typical class-AB main amplifier (measured) and class-A error amplifier (simulated).

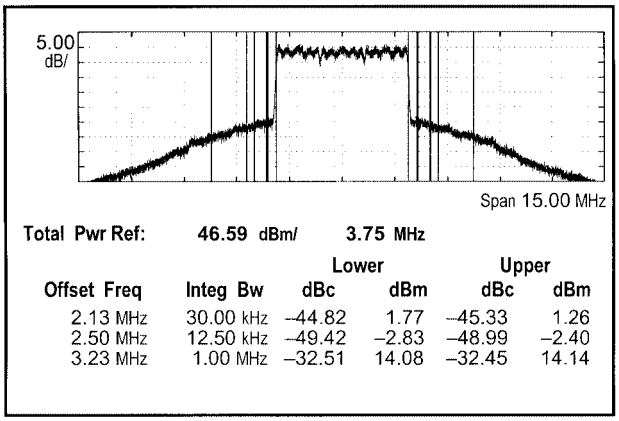
number has been found sufficient for convergence of the optimization procedure. For the CDMA signal, however, too many equal-probability intervals are required for convergence because they provide poor definition of the critical high-power portion of the curve in Fig. 6(c). Instead, 64 equal intervals are taken on the vertical scale from $x_m = 0$ to its maximum of 17.1, resulting in 64 unequal intervals on the horizontal scale. The x_m samples of the waveform are averaged within these x_m intervals, and probabilities assigned to these intervals according to their numbers of samples. Although 64 intervals has been found sufficient for convergence, this paper uses 512 intervals at little cost in computing time, which effectively leaves 402 x_m values once the no-sample intervals have been discarded.

B. Amplifier Models

Fig. 7 shows the compression and dc power functions for the typical class-AB and class-A amplifiers to be used as the main and error amplifiers, respectively. The curves for class A were simulated using HP's Microwave Design System, while the curves for class AB were obtained from measurements on a power amplifier destined for use in a feedforward circuit. Pulse mode measurements with a 0.5% duty cycle were performed, as they reflected actual operating conditions with a CDMA signal better than continuous-wave measurements, and produced a quite different saturated output power. The validity of this choice, and with it the validity of the complex baseband amplifier model, were confirmed by an adjacent channel power ratio (ACPR) measurement for a CDMA2000 SR3 MC signal [20] having a ccdf similar to that of the downloaded waveform. The class-AB amplifier was driven with such a signal to an output power of 46.62 dBm, or a reference of 46.59 dBm within the 3.75-MHz emission bandwidth, resulting in the output power density spectrum of Fig. 8(a). Due to some residual frequency response over the measurement bandwidth, the ACPR values for the lower sideband are approximately 1.5 dB inferior to those for the upper sideband. ACPR values were also simulated using the complete downloaded CDMA signal and the class-AB amplifier model with the measured value $\mathrm{P_{satM}} = 54.33$ dBm and the same output power. Using $\mathrm{P_M} = \mathrm{E}[|\mathrm{v_M}|^2/2\mathrm{Z_0}]$,



(a)



(b)

Fig. 8. Output power spectrum and ACPR values for a class-AB amplifier driven by a three-carrier CDMA signal. (a) Measured. (b) Predicted using measured AM/AM and AM/PM characteristics.

$v_M = \mathrm{v_M}/(\sqrt{2\mathrm{Z_o}\mathrm{P_m}}G_1) = v_m M_1 f_M(d_M)$, and (13) and (14), the output power of the amplifier model is given by

$$\mathrm{P_M} = D_{0M}|M_1|^2\mathrm{E}\left[x_m\left|f_M\left(x_m D_{0M}|M_1|^2\right)\right|^2\right]\mathrm{P_{satM}} \tag{70}$$

where $D_{0M}|M_1|^2$ is the amplifier's average input drive. An average input drive of -7.419 dB was required in (70) to produce 46.62 dBm of output power. With this input drive, the normalized signal

$$v_M/M_1 = v_m f_M(d_M) = v_m f_M\left(x_m D_{0M}|M_1|^2\right) \tag{71}$$

was computed and converted to the power spectrum of Fig. 8(b). An averaging factor of 32, properly applied to linear values, was used to improve the appearance of the trace. A Hanning window was also applied to compensate for the time truncation, but was not required since it did not affect the relative power levels. The computed ACPR values listed are typically within 1 dB of the measured ones.

The SIMR of v_M is denoted SIMR_M, and using (36) with the proper subscripts, is obtained from

$$\frac{1}{\mathrm{SIMR}_M} = \frac{X_M - |X_{Mm}|^2}{|X_{Mm}|^2} = \frac{\mathrm{E}\left[x_m|f_M(d_M)|^2\right]}{|E[x_m f_M(d_M)]|^2} - 1. \tag{72}$$

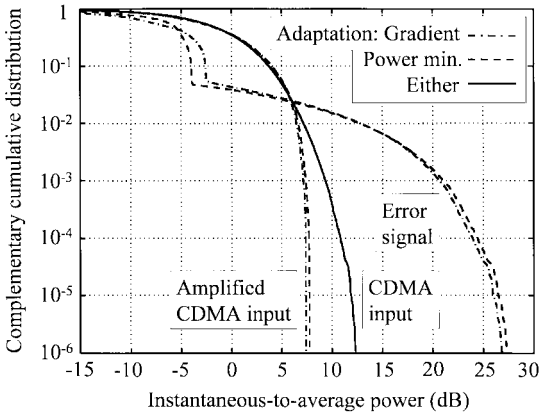


Fig. 9. Ccdfs of the signals in the first loop, assuming the gradient and power minimization adaptation methods with $D_{0M} = -7.4481$ dB.

TABLE I
OUTPUT SIMR VERSUS ADAPTATION SCHEMES FOR $D_{0M} = -7.4481$ dB AND $D_{0E} = -14.6295$ dB. ASSUMES INPUT SIGNAL SUPPRESSION IN THE THIRD LOOP AND ZERO TOLERANCES IN THE VECTOR MODULATOR GAINS

2nd loop \ 1st loop	Gradient	Power min.
$\varepsilon = -1$, Gradient	59.63 dB	8.9115 dB
$\varepsilon = -1$, Power min.	50.65 dB	7.7773 dB
$\varepsilon = 0$, Gradient	59.63 dB	60.013 dB
$\varepsilon = 0$, Power min.	59.634 dB	60 dB

A value of $\mathrm{SIMR}_M = 19.79$ dB was obtained from (72) for the current example. As expected, this is almost 5 dB less favorable than the worst power density ratio of approximately 24 dB found in Fig. 8(b).

C. Adaptation Methods

With 402 x_m values, the class-AB main amplifier, a reference input drive of $D_{0M} = -7.4481$ dB, and zero tolerance in the first loop, the normalized vector modulator gain M_1 with the gradient method is equal to the optimal value $M_{10} = 1.0434\angle -4.68°$. The power minimization method brings a relative change of $3.9\%\angle -168°$ to this value, causing the error power X_e to drop from 0.0126 to its minimum of 0.011, but raising $|X_{em}|^2$ from zero to 0.00125. With power minimization, the average input drive $D_{0M}|M_1|^2$ works out to the earlier value of -7.419 dB and, hence, $\mathrm{SIMR}_M = 19.79$ dB, as obtained before. A criterion commonly used in mobile systems design will be assumed, whereby this SIMR needs to be raised to 60 dB at the linearizer's output [8].

Fig. 9 compares the ccdf of the CDMA input with those of the amplified CDMA input and the error signal. The crest factor of the CDMA input is reduced by the amplifier to approximately 7.5 dB in both adaptation methods, indicating some clipping of the signal. The striking result is that the error signal has a 27-dB crest factor, or a full 15 dB above that of the CDMA input. Consequently, a much larger backoff in the error amplifier, hence, a much smaller value of D_{0E}, is required to meet the target linearity than would be suggested by merely the average error power. Assuming a value of $D_{0E} = -14.6295$ dB, Table I shows the output SIMRs obtained when the second loop is adapted using input signal suppression and with zero tolerance or $M_2 = M_{20}$. As explained earlier, $\varepsilon = -1$ dispenses with the signal suppression and requires accurate application of the gradient method in the first loop. With the gradient method applied in both loops, $\mathrm{SIMR} = 59.63$ dB is obtained. However, with the power method in the first loop and the consequent 3.9% change in M_1, $\mathrm{SIMR} = 8.9115$ dB is actually much worse than SIMR_M itself. In the second loop, the power method would be equivalent to the gradient method if the error amplifier were perfectly linear. In the present nonlinear case, however, the power method yields significantly worse results than the gradient method. Finally, assuming perfect signal suppression with $\varepsilon = 0$, all four combinations of loop adaptations yield output SIMRs around 60 dB. As expected, the result with both gradient adaptations is identical to the equivalent result without signal suppression.

D. Optimization of Efficiency

Clearly, $D_{0E} = -14.6295$ dB was chosen so that an output SIMR of precisely 60 dB would result when power minimization was used in both loops with $\varepsilon = 0$. This adaptation scheme will be assumed in the remainder of this paper. The present example then has values of M_1 and M_2 such that

$$\begin{aligned} |M_1|^2 &= 1.0067 \\ |M_2|^2 &= 1.0069 \\ |G_S|^2 &= 1.0004 \\ H_M &= 0.5883 \\ H_E &= 1.7564. \end{aligned} \tag{73}$$

Assuming now the following realistic specifications for a feedforward circuit:

$$\begin{aligned} U_{\mathrm{req}} &= 35 \text{ dB} \\ \mathrm{p}_1 &= \mathrm{p}_2 = 0.15 \text{ ns/dB} \\ \tau_1 &= \tau_2 = 1.5 \text{ ns} \\ \mathrm{q}_1 &= \mathrm{q}_2 = 15 \text{ ns/dB} \\ L &= P = N = R = -0.2 \text{ dB} \\ |C_m|^2 &= L/(1 + \mathrm{q}_1/\mathrm{p}_1) = -20.243 \text{ dB} \\ |A_c|^2 &= 0 \text{ dB} \\ |A_m|^2 &= |A_e|^2 = -0.2 \text{ dB} \\ |A_M|^2 &= |A_E|^2 = -0.3 \text{ dB} \\ |M_m|^2 &= |M_e|^2 = -16 \text{ dB} \end{aligned} \tag{74}$$

Fig. 10 shows the contours of power efficiency η as a function of the amplifier power gains $|G_M|^2$ and $|G_E|^2$. Gain pairs in the lower left blank areas require unfeasible values of the main coupling $|C_M|^2$, and have been assigned $\eta = 0$. A poor best value of $\eta = 1.06\%$ is obtained with $|G_M|^2 = 79.36$ dB and $|G_E|^2 = 54.43$ dB, corresponding to couplings of $|C_M|^2 = -32.3$ dB, $|C_e|^2 = -12.3$ dB, and $|C_o|^2 = -1.29$ dB, and normalized saturated output powers of $X_{satM} = 15.53$ dB and $X_{satE} = 16.22$ dB. The two amplifiers have similar saturated output powers, but the class-A error amplifier consumes more dc power, which explains why the output coupling is higher than -3 dB. In effect, given the insertion gain, the output coupler has become a 6.74-dB coupler with the transmit branch in the path of

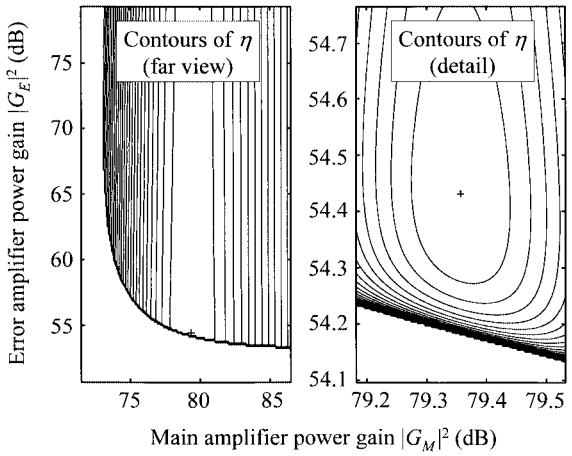


Fig. 10. Contours of power efficiency η as a function of the main and error amplifier gains for $D_{0M} = -7.4481$ dB and $D_{0E} = -14.6295$ dB. (*left-hand side*) Far view and (*right-hand side*) detail around the point of maximum efficiency.

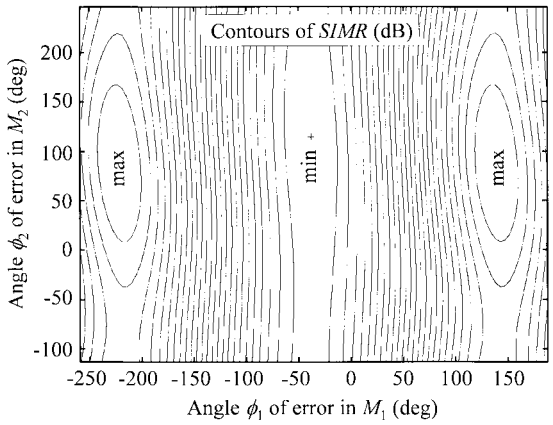


Fig. 11. Contours of output SIMR as a function of the angles of the 1% errors in each of the vector modulator gains, assuming $D_{0M} = -14.396$ dB and $D_{0E} = 7.4326$ dB. Minimum SIMR is the worst case scenario.

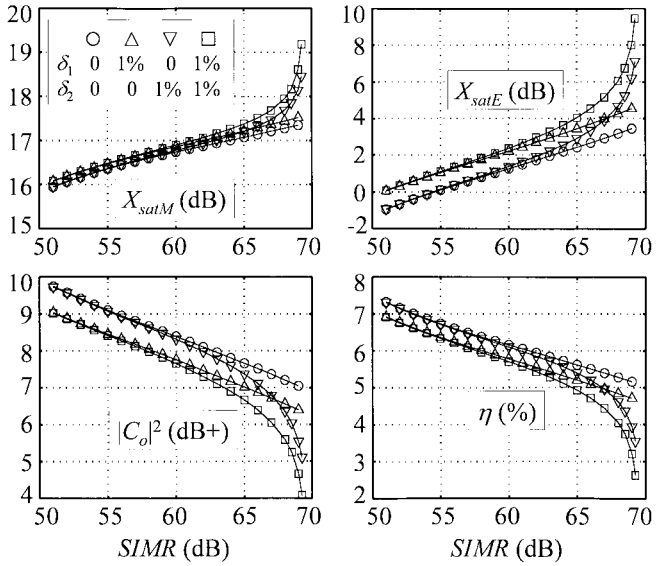


Fig. 12. Optimal values of normalized saturated output powers, output coupling, and power efficiency versus output SIMR for none, either, or both tolerances of 1% in the vector modulator gains.

the error amplifier. Another result contrary to expectation is that the main amplifier gain exceeds the error amplifier gain considerably. As it turns out, the relative gains of the two amplifiers can be adjusted at little cost in efficiency by merely changing the input coupling $|C_m|^2$. For example, having swapped the coupled and transmit branches of the original 20.243-dB input coupler, the new $|C_m|^2 = -0.2432$ dB yields a best value of $\eta = 1.05\%$ with $|G_M|^2 = 59.47$ dB, $|G_E|^2 = 74.01$ dB, $|C_M|^2 = -42.1$ dB, $|C_e|^2 = -22.1$ dB, $|C_o|^2 = -1.31$ dB, and the almost unchanged values of $X_{satM} = 15.65$ dB and $X_{satE} = 16.24$ dB.

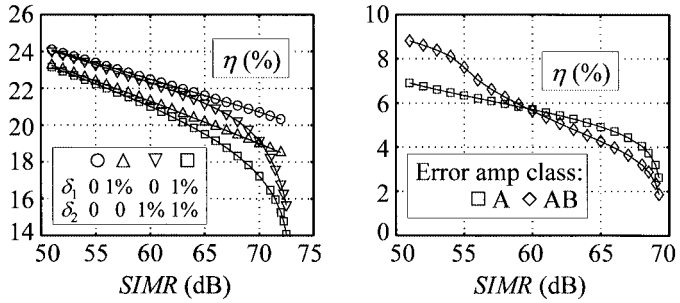
Finding SIMR and the best η at a single (D_{0M}, D_{0E}) point, as done above, must be repeated for a grid of such points in order to optimize η for given values of SIMR, including 60 dB. An adequate grid of 26×26 is pictured in Fig. 5(a), on which are plotted contours of SIMR from 51 to 69 dB and contours of η. The locations of optimal η on the SIMR contours are used to interpolate the corresponding values of amplifier gains, couplings, and normalized saturated output powers shown in Fig. 5(b) along with η. As expected, increasing the linearity requirement increases the two required saturated output powers and decreases efficiency. The error amplifier rises faster than the main amplifier, causing the output coupling to dip significantly below the classic 10-dB value. At SIMR $= 60$ dB, an optimal value of $\eta = 6.16\%$ is found at $D_{0M} = -14.396$ dB and $D_{0E} = 7.4326$ dB, corresponding to $|G_M|^2 = 74.13$ dB, $|G_E|^2 = 61.75$ dB, $|C_M|^2 = -29.5$ dB, $|C_e|^2 = -9.53$ dB, $|C_o|^2 = -8.39$ dB, $X_{satM} = 16.73$ dB, and $X_{satE} = 1.25$ dB. Compared to the previous situation, an increase of 1.2 dB in the main amplifier power rating has allowed a decrease of 14.97 dB in that of the error amplifier, resulting in an almost sixfold improvement in efficiency. Inserting the updated value of $D_{0M}|M_1|^2$ in (72) yields $\text{SIMR}_M = 29.44$ dB instead of 19.79 dB, and the resulting error signal is found to have a crest factor of approximately 24 dB instead of 27 dB. These two results give some insight into the difference.

E. Additional Results

If tolerances in M_1 and M_2 of $\delta_1 = \delta_2 = 1\%$ are introduced, the optimal (D_{0M}, D_{0E}) point that yielded SIMR $= 60$ dB in Fig. 5(a) now yields a worst case SIMR of 56.64 dB with $\phi_1 = -38.1°$ and $\phi_2 = 114°$. These angles were found by a two-dimensional search in the plane of ϕ_1 versus ϕ_2 shown in Fig. 11. The mostly vertical contours of SIMR indicate that, in this case, the tolerance in M_1 was more consequential than the tolerance in M_2.

The complete simulation of Fig. 5 was performed using none, either, or both of these tolerances, and Fig. 12 shows the resulting optimal traces for the two saturated output powers, the output couplingy. It is clear that δ_1 alone has an effect on the saturated output powers that is moderate and independent of the desired output linearity, whereas the effect of δ_2 is negligible at first, but becomes dominant as the required SIMR exceeds 67 dB. The error amplifier is also more affected by tolerances than the main amplifier, and correlates closely in behavior with the output coupling and the efficiency.

Fig. 13(a) repeats the efficiency plot of Fig. 12 with a two-tone input signal instead of the CDMA signal. The most obvious difference is that the efficiencies are more than three



(a) (b)

Fig. 13. (a) Equivalent of efficiency curves in Fig. 12 for a two-tone test signal instead of a three-carrier IS-95A CDMA signal. (b) Optimal power efficiency versus output SIMR with class-A and class-B error amplifier. Tolerance of 1% in both vector modulator gains are assumed.

times higher, due to required saturated output powers approximately 9 dB lower. Not shown here is the fact that the error amplifier now rises no faster than the main amplifier. This causes the optimal output coupling to persist at values close to 10.5 dB.

The final test presented here is to confront class-A operation for the error amplifier with class-AB operation similar to that of the main amplifier. This is done by using the class-AB curves of Fig. 7 for both amplifiers. The efficiency results with a CDMA input signal and $\delta_1 = \delta_2 = 1\%$ are shown in Fig. 13(b). The curve that applies to class A has been carried over from Fig. 12. Class-AB operation for the error amplifier appears advantageous for required SIMR levels below 60 dB.

V. Conclusions

A method of designing feedforward amplifiers for optimal power efficiency has been presented. The method uses a detailed complex baseband model of an adaptive feedforward amplifier with vector modulators in their practical positions. The model accounts for the unique compression and dc power characteristics of each amplifier, the adaptation schemes and their related tolerances, delay-to-gain relationships in the amplifiers and delay lines, and power losses in each branch and in directional couplers. The procedure takes into account the input signal format and the required linearity and gain at the system's output to optimize the gains and normalized power ratings of the amplifiers, as well as the various couplings.

The complex baseband analysis is computationally efficient, but is narrow-band by nature. However, under relatively broad-band CDMA excitation, the properly obtained complex baseband model of a typical amplifier was found to predict spectral regrowth well enough for present purposes. For added efficiency and convenience, the method uses only the power statistics of the input signal. This requires pilotless adaptation in the second loop, and provided the impetus for a fresh reexamination of pilotless adaptation methods. It also requires as the sole linearity criterion the SIMR, which was confirmed to be simply related to the regulated power spectral density ratio.

The optimization procedure centers on a judicious normalization of the feedforward model and the so-called reference input drives D_{0M} and D_{0E}, in terms of which can be found the output SIMR and also the power efficiency after it is partially optimized with respect to other variables. This leads to a highly intuitive graphical representation in which contours of SIMR and efficiency are superimposed and, thus, efficiency can be maximized for given SIMR values. A significant find is that the optimal input coupling is readily derived from a small number of component specifications.

For illustration, the reader was led through the details of an optimization involving a CDMA signal, a class-AB main amplifier, a class-A error amplifier, a required system gain of 35 dB, typical component specifications, and zero tolerance in the vector modulators. Before settling with an adaptation scheme, differences were shown in the performances of different schemes under certain conditions, and before settling for the optimal input coupling, it was shown that this coupling can be adjusted to change the balance of the amplifier gains with little impact on efficiency. At an SIMR of 60 dB, the optimal design produced a remarkable 24-dB crest factor in the error signal, severely curtailing the efficiency and making the balance of amplifier power ratings critical. At the same SIMR, a decrease of only 1.2 dB in the main amplifier power rating required an increase of 15 dB in the error amplifier power rating, for an almost sixfold decrease in efficiency. The reference optimization above was repeated with a tolerance in one or both vector modulators, revealing marked differences in the relative impacts of these tolerances as a function of the SIMR. These tests, then repeated for a two-tone input signal, produced important differences in the optimal design parameters and showed to what extent the optimization of efficiency is a signal-dependent affair. Finally, substituting a class-AB amplifier for the error amplifier showed the promise that lies in the judicious use of different amplifier types.

Although the proposed procedure cannot directly optimize amplifier types and adaptation methods, it is being used to investigate such factors [10]–[12], and therein may lie its greatest value.

Acknowledgment

The authors wish to thank G. Zhao, AmpliX Inc., Montréal, QC, Canada, for providing the measured data for validation purposes and for her willingness to discuss related technical issues.

References

[1] A. Bateman and D. Haines, "Direct conversion transceiver design for compact low-cost portable mobile radio terminals," in *Proc. IEEE Veh. Technol. Conf.*, vol. 1, 1989, pp. 57–62.

[2] A. Wright and W. Durtler, "Experimental performance of an adaptive digital linearized power amplifier," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 395–400, Nov. 1992.

[3] M. Kumar, J. Whartenby, and H. Wolkstein, "Predistortion linearizer using GaAs dual-gate MESFET for TWTA and SSPA used in satellite transponders," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-33, pp. 1479–1488, Dec. 1985.

[4] K. Konstantinou and D. K. Paul, "Analysis and design of broadband, high efficiency feedforward amplifiers," in *IEEE MTT-S Int. Microwave Symp. Dig.*, 1996, pp. 867–870.

[5] E. E. Eid, F. M. Ghannouchi, and F. Beauregard, "Optimal feedforward linearization system design," *Microwave J.*, vol. 38, no. 11, pp. 78–84, Nov. 1995.

[6] K. J. Parsons and P. B. Kenington, "The efficiency of a feedforward amplifier with delay loss," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 407–412, May 1994.

[7] S. L. Loyka and J. R. Mosig, "New behavioral-level simulation technique for RF/microwave applications. Part I: Basic concepts," *Int. J. RF Microwave Computer-Aided Eng.*, vol. 10, no. 4, pp. 221–237, July 2000.
[8] J. K. Cavers, "Adaptation behavior of a feedforward amplifier linearizer," *IEEE Trans. Veh. Technol.*, vol. 44, pp. 31–40, Feb. 1995.
[9] C. L. Larose and F. M. Ghannouchi, "Optimization of feedforward amplifier power efficiency on the basis of input power statistics," in *IEEE MTT-S Int. Microwave Symp. Dig.*, vol. 3, 2000, pp. 1491–1494.
[10] ——, "First-Loop control in feedforward amplifiers for high-stress communications signals," in *Proc. Int. Microwave and Opt. Technol. Symp.*, 2001, pp. 195–198.
[11] ——, "Pilotless adaptation of feedforward amplifiers driven by high-stress signals," in *Proc. IEEE Radio and Wireless Conf.*, 2001, pp. 81–84.
[12] ——, "Optimal amplifier classes and adaptation methods for feedforward amplifiers driven by high-stress signals," *IEEE Trans. Veh. Technol.*, to be published.
[13] Directional coupler specifications. Pulsar Microwave Corporation, Clifton, NJ. [Online]. Available: http://www.pulsarmicrowave.com/products/couplers/couplers.htm.
[14] Couplers: Parameter definitions. Synergy Microwave Corporation, Paterson, NJ. [Online]. Available: http://www.synergymwave.com/products/couplers/CouplerTutorial.pdf.
[15] S. A. Maas, *Nonlinear Microwave Circuits*. Norwood, MA: Artech House, 1988, p. 370.
[16] S. S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1991.
[17] F. M. Ghannouchi, F. Beauregard, A. B. Kouki, and M. Bouchard, "Adaptive linearization of a feedforward amplifier by complex gain stabilization of the error amplifier," U.S. Patent 6 275 105 B1, Aug. 14, 2001.
[18] R. M. Bauman, "Adaptive feed-forward system," U.S. Patent 4 389 618, June 21, 1983.
[19] Three carrier IS-95A CDMA waveform for the Agilent ESG-D series. Agilent Technol., Palo Alto, CA. [Online]. Available: http://www.agilent.com, Quick Search on "DAB."
[20] "Performing cdma2000 measurements today," Agilent Technol., Palo Alto, CA, Application Note 1325, May 2000.

Colin L. Larose (M'91) received the B.Eng. and M.Eng. degrees in electrical engineering from Concordia University, Montréal, QC, Canada, in 1983 and 1986, respectively, the B.A. degree from Ambassador University, Pasadena, CA, in 1989, and the Ph.D. degree in electrical engineering from the École Polytechnique de Montréal, Montréal, QC, Canada.

For five years, he was involved in the area of radar cross-sectional measurement and simulation with the David Florida Laboratory, Canadian Space Agency, Ottawa, ON, Canada. His fields of interest include feedforward and other linearization techniques, wireless communications, and measurement and simulation techniques for antennas and radar.

Mr. Larose was the recipient of a 1985 research grant and a 1995 Postgraduate Scholarship presented by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Fadhel M. Ghannouchi (S'84–M'88–SM'93) received the B.Eng. degree in engineering physics and the M.S. and Ph.D. degrees in electrical engineering from the École Polytechnique de Montréal, Montréal, QC, Canada, in 1983, 1984, and 1987, respectively.

He is currently a Professor with the Département de génie électrique, École Polytechnique de Montréal, where he has taught electromagnetics and microwave theory and techniques since 1984. He has provided consulting services to a number of microwave companies. He is also the founder of AmpliX Inc., Montréal, QC, Canada, a company that offer linearization products and services to wireless and satcom equipment manufacturers. His research interests are in the areas of microwave/millimeter-wave instrumentation and measurements, nonlinear modeling of microwave active devices, and design of power and spectrum efficient microwave amplification systems.

Dr. Ghannouchi is a Registered Professional Engineer in the Province of Quebec, Canada. He is on the Editorial Board of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES and has served on the Technical Committees of several international conferences and symposiums.