

Titre: Amélioration du rendement énergétique des amplicateurs à post-compensation (feedforward) pour signaux à facteur de crête élevé
Title: Compensation (feedforward) for signals with high crest factor to improve the energy efficiency of amplifiers

Auteur: Colin Larose
Author:

Date: 2002

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Larose, C. (2002). Amélioration du rendement énergétique des amplicateurs à post-compensation (feedforward) pour signaux à facteur de crête élevé [Ph.D. thesis, École Polytechnique de Montréal]. PolyPublie.
Citation: <https://publications.polymtl.ca/7082/>

Document en libre accès dans PolyPublie Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/7082/>
PolyPublie URL:

Directeurs de recherche: Fadhel M. Ghannouchi
Advisors:

Programme: Unspecified
Program:

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

**ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600**



UNIVERSITÉ DE MONTRÉAL

**AMÉLIORATION DU RENDEMENT ÉNERGÉTIQUE
DES AMPLIFICATEURS À POST-COMPENSATION (FEEDFORWARD)
POUR SIGNAUX À FACTEUR DE CRÊTE ÉLEVÉ**

**COLIN LAROSE
DÉPARTEMENT DE GÉNIE ÉLECTRIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL**

**THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIAE DOCTOR (Ph.D.)
(GÉNIE ÉLECTRIQUE)
NOVEMBRE 2002**



National Library
of Canada

Acquisitions and
Bibliographic Services

385 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

385, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file: [View reference](#)

Our file: [View reference](#)

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-75938-5

Canada

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée:

AMÉLIORATION DU RENDEMENT ÉNERGÉTIQUE
DES AMPLIFICATEURS À POST-COMPENSATION (FEEDFORWARD)
POUR SIGNAUX À FACTEUR DE CRÊTE ÉLEVÉ

présentée par: LAROSE Colin
en vue de l'obtention du diplôme de: Philosophiae Doctor
a été dûment acceptée par le jury d'examen constitué de:

M. BOSISIO Renato, M.Sc., président
M. GHANNOUCHI Fadhel, Ph.D., membre et directeur de recherche
M. SAWAN Mohamad, Ph.D., membre
M. GAGNON François, Ph.D., membre

DÉDICACE

Je rends hommage à mon Créateur, Sauveur et Conseiller, qui donne un sens à tous mes efforts, ainsi qu'à mes parents terrestres, dont l'origine modeste n'a pas empêché de semer en moi les germes de la réussite. Je fais de même pour ces géants qui m'ont précédé et reçu sur leurs épaules, en particulier le Dr Jim Cavers de Simon Fraser University, dont les travaux sur l'amplificateur à post-compensation n'ont pas encore fini de m'inspirer. Je dédie particulièrement cette thèse aux quatre personnes qui ont à tout jamais transformé ma vie pendant la période de mon doctorat, soit mon épouse et mes trois enfants. Ils m'ont donné bien des façons de ne pas perdre la tête et bien des raisons de persévérer dans les moments difficiles. Je la dédie également à mes proches et amis qui ont dû s'étonner de mon isolement temporaire et qui se réjouiront avec moi de la conclusion heureuse de cette étape de ma vie.

REMERCIEMENTS

Je remercie le Dr Stanley Kubina de l'Université Concordia de m'avoir recommandé à celui qui allait être mon patron de 1990 à 1995 à l'Agence spatiale canadienne, le Dr Shantnu Mishra. Cette étape m'a fourni l'expérience de recherche et la base financière me permettant d'envisager un programme d'études doctorales. Une coopération étroite avec le Dr Christopher Trueman, également de Concordia, a joué un rôle clé dans mon bilan de publication et dans l'obtention d'une bourse d'études supérieures. Je remercie le CRSNG pour cette bourse, l'École Polytechnique de Montréal pour les suppléments de bourse qu'elle a rajoutés, ainsi que mon beau-père Jacques de Billy pour la générosité qu'il n'a cessé de démontrer envers ma famille.

Je désire remercier mon directeur de recherche, M. Fadhel Ghannouchi, pour sa patience pendant les phases moins spectaculaires mais non moins productives de mes travaux, son ouverture par rapport à mon changement de sujet de recherche, et ses interventions et conseils judicieux lorsque rien n'allait plus. Les responsabilités qu'il m'a confiées et les conférences auxquelles il m'a permis d'assister ont été une source d'enrichissement, et ses encouragements répétés vis-à-vis mon concept d'amplificateur à post-compensation pré-écrêteur se sont avérés par la suite être déterminants pour l'issue de mes travaux. Je remercie également ceux qui ont servi sur mes différents jurys, en particulier M. Gilles Brassard de l'Agence spatiale canadienne et M. François Gagnon de l'École de technologie supérieure.

Parmi les étudiants et le personnel amicaux qui m'ont entouré, je me dois de citer Mme G. Zhao de la compagnie AmpliX pour m'avoir gracieusement accordé plusieurs entretiens techniques bien compliqués, et l'administrateur de réseau René Archambault pour avoir promptement éteint plusieurs de mes « feux ». Enfin, je rends hommage à la grande dame déjà cachée derrière le grand homme que j'espère devenir, mon épouse

Josée. Je ne saurais la remercier suffisamment d'avoir marié un étudiant et d'avoir eu assez de confiance en moi pour se redéfinir en tant que mère à temps plein, favorisant ainsi le développement de ma carrière aux dépens de la sienne. Je remercie également mes parents, qui n'ont cessé de me supporter par leurs diverses ressources et compétences, ainsi que tous ceux de la communauté anglicane St. Stephen's qui m'ont accompagné par leur amitié et leurs prières.

RÉSUMÉ

Les réseaux de communication sans fil actuels requièrent des stations de base pouvant amplifier fidèlement des signaux dont l'enveloppe manifeste d'importantes modulations. Ces modulations surviennent lorsqu'on combine plusieurs signaux à enveloppe constante (MF, par exemple), d'où la nécessité d'utiliser le soi-disant *amplificateur de puissance multi-porteuse* (APMP). Les signaux, combinés ou non, peuvent aussi avoir des modulations d'enveloppe inhérentes, comme dans le système cellulaire numérique AMRC (accès multiple par répartition de code). Quoi qu'il en soit, le signal résultant est nécessairement affecté par la non-linéarité inhérente de l'amplificateur de puissance, et voit alors son spectre déborder sur les canaux adjacents. Les standards de non-interférence très sévères demandent de la part de l'amplificateur une linéarité qui s'obtient malheureusement au détriment de son rendement énergétique. Aux frais de consommation électrique accrus s'ajoutent les frais d'installation et d'entretien d'un système de refroidissement adapté. L'alternative communément adoptée est de linéariser un amplificateur à bon rendement énergétique au moyen de circuits externes à faible consommation.

L'amplificateur à *post-compensation* (feedforward), obtenu au moyen de la technique de linéarisation du même nom, est devenu une réalité commerciale notamment dans le système AMRC en raison de sa haute linéarité sur une large bande passante. Le linéarisateur forme une première boucle en soustrayant le signal d'origine d'un échantillon du signal amplifié pour générer un signal d'erreur. Dans une seconde boucle, le signal d'erreur est amplifié par un amplificateur auxiliaire (ou d'erreur) et soustrait du signal principal pour annuler la distorsion. La qualité de cette annulation dépend de la linéarité de l'amplificateur d'erreur. Les précisions des deux soustractions de signaux sont aussi déterminantes et rendent le circuit vulnérable aux changements dans les conditions de fonctionnement. Ces changements que sont, en ordre croissant de rapidité,

le vieillissement des composantes, les fluctuations de température et les altérations des propriétés statistiques du signal d'entrée, demeurent beaucoup plus lents que la modulation de l'enveloppe même et peuvent être neutralisés par des systèmes adaptatifs. Ces systèmes maintiennent l'équilibre des deux boucles en contrôlant deux modulateurs vectoriels; ces gains complexes ajustables sont adjacents aux deux amplificateurs afin de ne pas modifier le gain composite du circuit, et les précèdent afin de minimiser les pertes d'énergie. De même, les délais de propagation dans chacune des boucles doivent être harmonisés en équilibrant les délais substantiels des deux amplificateurs par des lignes à délai dans les branches parallèles.

Au plan énergétique, la linéarisation par post-compensation est désavantageée par rapport à d'autres méthodes. En effet, le rendement énergétique de l'amplificateur principal est dégradé de façon significative par l'atténuation de sa sortie dans une des lignes à délai et par la consommation en puissance de l'amplificateur d'erreur. Les quelques études déjà publiées sur ce problème simplifient l'analyse par des suppositions plus ou moins réalistes, notamment l'emploi du signal classique à deux tonalités dont la modulation d'enveloppe est bien plus faible que celle employée dans les signaux de communication actuels.

Le présent ouvrage débute donc par le développement d'une procédure numérique adéquate pour optimiser l'amplificateur à post-compensation. Cette procédure est basée sur un modèle réaliste où les modulateurs vectoriels précèdent les amplificateurs principal et d'erreur (ce qui affecte les performances des méthodes d'adaptation) et les relations régissant les diverses pertes, délais et tolérances dans le circuit sont prises en considération. Par l'entremise de sept paramètres de conception, la procédure optimise le rendement énergétique en fonction des exigences de gain et de linéarité de l'usager, et non l'inverse. L'analyse complexe en bande de base assure un meilleur rendement de calcul et permet de tenir compte de la conversion amplitude-phase (MA-MP) qui s'avérera centrale aux phénomènes étudiés. Bien qu'elle soit à bande étroite, cette analyse

s'est avérée bonne à prédire le débordement spectral d'un signal AMRC à la sortie d'un amplificateur principal typique. Le traitement statistique accélère encore les calculs et s'accommode d'une description statistique du signal d'entrée dérivée de la fonction de distribution cumulée complémentaire (FDCC), dont l'usage est courant. Finalement, les relations non analytiques sont effectivement résolues au moyen d'une représentation graphique particulièrement évocatrice.

La procédure d'optimisation de l'amplificateur à post-compensation ne peut optimiser directement le type d'amplificateur et la méthode d'adaptation employés dans chaque boucle. La procédure requiert donc au départ quatre fonctions normalisées traduisant les caractéristiques de compression du gain et de consommation de puissance des deux amplificateurs. Ces caractéristiques dépendent du type de transistor et de la classe d'amplification déterminée par la polarisation du transistor. Les classes A, AB, B et C sont en ordre décroissant de linéarité, mais le choix de la classe A n'est pas automatique pour autant puisque ces mêmes classes sont en ordre croissant de rendement énergétique, offrant ainsi différents compromis linéarité-rendement. Par exemple, l'industrie préfère actuellement l'amplificateur principal de classe AB. La procédure requiert aussi la méthode d'adaptation à employer pour chaque boucle. L'adaptation de la deuxième boucle se fait toujours en fonction de l'objectif ultime qui est de minimiser la distorsion restante dans le signal de sortie. On pense généralement servir le même objectif en adaptant la première boucle par la *méthode de puissance*, laquelle minimise simplement la puissance moyenne du signal d'erreur. Cette méthode n'est cependant pas optimale puisque la distorsion restante à la sortie est reliée à la distorsion instantanée dans l'amplificateur d'erreur, distorsion qui est elle-même une fonction très *non linéaire* de la puissance instantanée du signal d'erreur. En adaptant la première boucle comme la deuxième en fonction de la distorsion restante à la sortie, deux nouvelles méthodes optimales ont été mises en évidence, selon le point de départ de l'adaptation: la méthode optimale de droite et la méthode optimale de gauche.

L'ouvrage conclut par des efforts en vue d'optimiser les types d'amplificateurs et la méthode d'adaptation à employer dans l'amplificateur à post-compensation. Sous des contraintes de conception typiques, le rendement énergétique de l'amplificateur à post-compensation a été optimisé pour une série d'amplificateurs principaux variant de la classe AB à la classe C (l'amplificateur d'erreur demeurant en classe A) et pour les trois méthodes d'adaptation de la première boucle mentionnées ci-haut. Les influences des deux facteurs investigués s'avèrent faibles pour un signal d'entrée à deux tonalités mais relativement importantes pour un signal AMRC. Ces influences étant interdépendantes, les deux nouvelles méthodes d'adaptation promettent des rendements énergétiques accrus avec des amplificateurs principaux plus près de la classe B ou carrément de classe C, selon le type de transistor employé et la linéarité requise à la sortie. En supposant deux amplificateurs à MESFET typiques et une linéarité de sortie couramment utilisée dans la conception de systèmes mobiles, les méthodes optimales de droite et de gauche ont permis des améliorations de rendement relatives de 21 et de 36 pour cent avec des amplificateurs principaux près de la classe B et de classe B, respectivement. Ces découvertes contredisent la perception voulant que l'industrie ait déjà réalisé tout le potentiel de rendement énergétique de l'amplificateur à post-compensation.

Avec des MESFET moins typiques à transconductance plus constante, les méthodes optimales de droite et de gauche offrent les meilleurs rendements avec des amplificateurs principaux de classe C. Le fonctionnement du circuit avec adaptation optimale de gauche rejoint alors le concept de l'amplificateur à post-compensation *pré-écrêteur*, anticipé au début des présents travaux. Dans ce concept, l'amplificateur principal de classe C amplifie efficacement les crêtes du signal d'entrée au-delà d'un seuil donné. Il crée ainsi un signal d'erreur équivalent au signal d'entrée écrêté, donc peu agressant envers l'amplificateur d'erreur. Ce dernier amplificateur est effectivement un nouvel amplificateur principal qui bénéficie d'un écrêtage du signal d'entrée effectué au préalable dans la première boucle dite *d'écrêtage* à l'aide de l'amplificateur de crête. Les crêtes sont restituées à la sortie dans la deuxième boucle dite *de restitution des crêtes*.

ABSTRACT

Today's wireless communication networks require base stations that can amplify signals with strongly modulated envelopes. These modulations occur when signals of constant envelope (for example FM) are combined, justifying the need for the so-called *multi-carrier power amplifier* (MCPA). Whether combined or not, signals may also have inherent envelope modulation, as in the CDMA (Code Division Multiple Access) digital cellular system. In either case, the resulting signal is necessarily affected by the inherent non-linearity of the power amplifier, and its power spectrum spills over into the adjacent channels. Stringent standards of non-interference require that the amplifier be highly linear, but unfortunately this characteristic is obtained at the expense of power efficiency. Added to the increased electrical consumption are the installation and maintenance costs of a suitable cooling system. A commonly adopted alternative consists in linearizing a power-efficient amplifier with external circuitry of low power consumption.

The feedforward amplifier, obtained via the feedforward linearization technique, has become a commercial reality (notably in the CDMA system) by virtue of its high linearity over a wide bandwidth. The feedforward linearizer forms a first loop by subtracting the original signal from a sample of the amplified signal to generate an error signal. In a second loop, the error signal is amplified by an auxiliary (or error) amplifier and subtracted from the main signal to cancel the distortion. The effectiveness of this cancellation depends on the linearity of the error amplifier. The accuracies of the two signal subtractions are also critical and make the circuit vulnerable to changes in the operating conditions. These changes which are, in order of speed, the aging of the components, the fluctuations in temperature, and the alterations in the statistical properties of the input signal, remain much slower than the modulation of the envelope itself and can be neutralized by adaptive systems. These systems maintain the balance in each loop by controlling two vector modulators; these adjustable complex gains are

adjacent to the two amplifiers in order not to modify the overall gain of the circuit, and precede them in order to minimize power losses. Likewise, propagation delays in each loop must be matched by compensating for the longer delays of the amplifiers with delay lines in the parallel branches.

In terms of power efficiency, feedforward linearization is at a disadvantage compared with other methods. Indeed, the power efficiency of the main amplifier is significantly degraded by the attenuation of its output in one of the delay lines and by the power consumption of the error amplifier. The few studies already published on this subject simplify the analysis with more or less realistic assumptions, notably with the use of the classic two-tone signal, whose envelope modulation is much weaker than that used in present communication signals.

The present work therefore begins with the development of an adequate numerical method of optimizing the feedforward amplifier. The procedure is based on a realistic model in which the vector modulators precede the main and error amplifiers (which influences how the adaptation methods perform) and the relationships governing the various losses, delays, and tolerances in the circuit are taken into account. By way of seven design parameters, the procedure optimizes power efficiency in accordance with the customer's gain and linearity requirements, and not the other way around. The complex baseband analysis makes for better computational efficiency and allows taking into account the AM-PM conversion, which will prove to be central to the phenomena under study. Although narrowband, this analysis proved adequate in predicting the spectral regrowth of a CDMA signal exiting a typical main amplifier. The statistical approach further accelerates the computations and merely requires of the input signal a statistical description derived from the complementary cumulative distribution function (CCDF), whose use is widespread. Finally, the non-analytical relationships are effectively solved by means of a particularly intuitive graphical representation.

The feedforward optimization procedure cannot directly optimize the type of amplifier and the adaptation method employed in each loop. Hence, the procedure requires from the outset four normalized functions to convey the gain compression and dc power characteristics of the two amplifiers. These characteristics depend on the type of transistor and the class of amplification as determined by the biasing of the transistor. Although classes A, AB, B, and C are in descending order of linearity, the choice of Class A is not automatic because these same classes are in ascending order of power efficiency and therefore offer different linearity-efficiency tradeoffs. For example, the Class AB main amplifier is now preferred by the industry. The procedure also requires the adaptation method employed in each loop. Adaptation of the second loop is always done toward the bottom-line objective of minimizing the remaining distortion in the output signal. It is generally thought that the same objective can be served by adapting the first loop according to the *power method*, which simply minimizes the average power in the error signal. This method is not optimal, however, because the remaining distortion at the output is related to the instantaneous distortion in the error amplifier, and the latter distortion is itself a very *non-linear* function of the instantaneous power in the error signal. By adapting the first loop, like the second, in terms of the remaining distortion at the output, two new optimal methods were brought to light, depending on the starting point of the adaptation: the optimal-R (Right) method and the optimal-L (Left) method.

The work concludes with efforts to optimize the types of amplifiers and the adaptation method to be used in the feedforward amplifier. Under typical design constraints, the power efficiency of the feedforward amplifier was optimized for a series of main amplifiers ranging from Class AB to Class C (the error amplifier remaining in Class A) and for the three above-mentioned methods of adapting the first loop. The influences of the two factors investigated are found to be weak for a two-tone input signal but relatively important for a CDMA signal. These influences being interdependent, the two new adaptation methods promise better power efficiencies with main amplifiers biased closer to Class B or even in Class C, depending on the type of transistor employed and

the required output linearity. Assuming two typical MESFET amplifiers and an output linearity commonly used in mobile systems design, the optimal-R and optimal-L methods allowed relative efficiency improvements of 21 and 36 percent with main amplifiers biased near Class B and at Class B, respectively. These findings contradict the perception that the efficiency potential of the feedforward amplifier has already been fully realized by the industry.

With less typical MESFETs of more constant transconductance, the optimal-R and optimal-L methods yield the best efficiency figures with Class C main amplifiers. Circuit operation with optimal-L adaptation then approaches the concept of the *pre-clipping* feedforward amplifier, anticipated from the outset of this work. In this concept, the efficient Class C main amplifier amplifies the input signal peaks beyond a given threshold. The resulting error signal is equivalent to the clipped input signal and is therefore less stressful to the error amplifier. The latter amplifier is in effect a new main amplifier that benefits from a clipping of the input signal done beforehand in the first *clipping* loop with the help of the *peaking* amplifier. The peaks are reinserted at the output in the second *peak reinsertion* loop.

TABLE DES MATIÈRES

LISTE DES ANNEXES	xviii
LISTE DES TABLEAUX	xix
LISTE DES FIGURES	xx
LISTE DES SIGLES ET ABRÉVIATIONS	xxii
AVANT-PROPOS	xxiii
CHAPITRE I: INTRODUCTION	1
1.1 Signaux à enveloppe constante	1
1.2 Amplificateur de classe A en recul de puissance	3
1.3 Linéarisation	4
1.4 Contre-réaction	4
1.5 Prédistorsion adaptative en bande de base	6
1.6 Post-compensation	6
1.6.1 Travaux de H. S. Black	6
1.6.2 Contributions de H. Seidel	8
1.6.3 Amplificateur à post-compensation moderne	9
1.7 Prédistorsion RF	11
1.8 Rendement de l'amplificateur à post-compensation	11
1.8.1 Stratégies « de matériel »	12
1.8.2 Stratégies « de logiciel »	12
CHAPITRE II: REVUE DE LA LITTÉRATURE	13
2.1 Kenington et autres	13
2.1.1 Formule de Kenington	13
2.1.2 Formule de Kenington avec perte de délai	14
2.1.3 Réduction de la ligne à délai	14

2.1.4	Limites de la formule de Kenington	15
2.2	Horiguchi et autres	16
2.2.1	Amplificateurs à tailles fixes	16
2.2.2	Amplificateurs à rendements fixes	16
2.3	Au-delà de Kenington	17
2.3.1	Hickson et autres	17
2.3.2	Larose et Ghannouchi	18
CHAPITRE III: ORGANISATION GÉNÉRALE		19
3.1	Articles en annexe	19
3.2	Détails théoriques et expérimentaux	19
3.3	Rédaction en français	20
CHAPITRE IV: DÉTAILS TECHNIQUES ET EXPÉRIMENTAUX		21
4.1	Premier article (annexe I)	21
4.1.1	Modèle de l'amplificateur de puissance	21
4.1.2	Modèle de l'amplificateur à post-compensation	25
4.1.3	Méthode d'optimisation pentagonale	27
4.1.4	Effets de température sur l'adaptation par suppression de signal	32
4.1.5	Effets d'une suppression imparfaite de signal	35
4.1.6	Vérification des paramètres optimaux triviaux	39
4.1.7	Calcul du couplage principal	41
4.1.8	Méthode d'optimisation polaire	42
4.1.9	Localisation des points de solution	44
4.1.10	Spectre de densité de puissance	48
4.2	Deuxième article (annexe II)	50
4.2.1	Schéma des stratégies d'adaptation de la première boucle	50
4.2.2	Tolérance du gain du second modulateur vectoriel	52
4.2.3	Discontinuités dans la solution graphique	54

4.2.4 Amplificateur à post-compensation pré-écrêteur	55
CHAPITRE V: CONCLUSION	57
5.1 Introduction	57
5.2 Logiciel d'optimisation énergétique	58
5.3 Nouvelles méthodes d'adaptation	62
BIBLIOGRAPHIE	68

LISTE DES ANNEXES

Annexe I:	Premier article: « Optimization of feedforward amplifier power efficiency on the basis of drive statistics »	75
Annexe II:	Deuxième article: « Optimal amplifier classes and adaptation methods for feedforward amplifiers driven by high-stress signals »	102

LISTE DES TABLEAUX

Tableau 4.1: Possibilités mathématiques sur (E, G, K)	41
Tableau 4.2: Comparaisons se rapportant à la tolérance δ_2 , pour trois méthodes d'adaptation et $IMSR_{req} = -60$ dB	53

LISTE DES FIGURES

Figure 1.1:	Technique traditionnelle d'amplification de canaux multiples pour transmission par antenne unique	2
Figure 1.2:	Technique d'amplification de canaux multiples au moyen d'un amplificateur de puissance multi-porteuse (APMP)	3
Figure 1.3:	Schéma fonctionnel de l'amplificateur à post-compensation tel qu'envisagé par H. S. Black	7
Figure 1.4:	Modèle complexe en bande de base de l'amplificateur à post-compensation	9
Figure 4.1:	Écart typique entre les puissances de référence d'entrée $P_{in,1dB}$ et $P_{in,ref}$ pour un amplificateur à semi-conducteurs de classe A	22
Figure 4.2:	Effets du gain de puissance ou de la puissance de sortie de saturation sur la puissance de sortie, la phase de sortie relative, et la puissance consommée	24
Figure 4.3:	Première étape de la construction du modèle normalisé de l'amplificateur à post-compensation	25
Figure 4.4:	Exemple de recherche en deux dimensions par la méthode du simplexe: (a) description classique, (b) description alternative	28
Figure 4.5:	Exemple d'optimisation pentagonale avec tolérance de 0,085 sur les coordonnées	31
Figure 4.6:	Deuxième boucle de l'amplificateur à post-compensation adaptée par la méthode CGS (Ghannouchi et autres. 2001)	33
Figure 4.7:	Diagramme des délais pour le circuit de la figure 4.6 à une température de référence	33
Figure 4.8:	Diagramme modifié des délais pour le circuit de la figure 4.6 à une température de référence	35

Figure 4.9:	Signaux du modèle de l'amplificateur à post-compensation dans une analyse inspirée de Cavers (1995)	36
Figure 4.10:	Minimisation de la fonction banane (fonction de Rosenbrock) par la méthode polaire avec tolérance de 0,15	43
Figure 4.11:	(a) Contours de rendement énergétique (η) et de linéarité de sortie ($SIMR$). (b) Fonction d'erreur pour $SIMR = 60$ dB	45
Figure 4.12:	Données permettant l'interpolation bicubique de la fonction z au point (s,t)	47
Figure 4.13:	Relation entre la transformée de Fourier de l'enveloppe complexe d'un signal et la transformée de Fourier du signal réel	49
Figure 4.14:	Adaptation de la première boucle: (a) conventionnelle. (b) optimale avec contrôles séparés. (c) optimale avec contrôles intégrés	50
Figure 4.15:	Contours de rendement énergétique (η) et de linéarité de sortie ($SIMR$) avec points de solution. Solutions locales et globales pour $SIMR = 62$ à 65 dB	54
Figure 4.16:	Représentation conservatrice du concept d'amplificateur à post-compensation pré-écrêteur	56
Figure 4.17:	Représentation alternative du concept d'amplificateur à post-compensation pré-écrêteur	56

LISTE DES SIGLES ET ABRÉVIATIONS

(Anglais entre parenthèses)

ADC	Attribution dynamique de canaux (DCA)
AMRC	Accès multiple par répartition de code (CDMA)
AMRT	Accès multiple par répartition dans le temps (TDMA)
APMP	Amplificateur de puissance multi-porteuse (MCPA)
ATOP	Amplificateur à tube à ondes progressives (TWTA)
CC	Courant continu (DC)
CGS	(Complex Gain Stabilization) Stabilisation du gain complexe
FDCC	Fonction de distribution cumulée complémentaire (CCDF)
IEEE	(Institute of Electrical and Electronics Engineers)
IMSR	(IM-to-Signal Ratio) Rapport intermodulation à signal
MA	Modulation d'amplitude (AM)
MA-MA	Réponse amplitude-amplitude (AM-AM)
MA-MP	Réponse amplitude-phase (AM-PM)
MESFET	(Metal Semiconductor Field Effect Transistor) Transistor à effet de champ métal-semiconducteur
MF	Modulation de fréquence (FM)
MIC	Modulation par impulsions codées (PCM)
MP	Modulation de phase (PM)
RF	Radiofréquence
SIMR	(Signal-to-IM Ratio) Rapport signal à intermodulation
VHF	(Very High Frequency) Très haute fréquence
WCDMA	(Wideband CDMA) AMRC large bande

AVANT-PROPOS

Une chronologie des événements ayant abouti à cette thèse illustre bien les chemins tortueux qu'emprunte parfois le processus de découverte scientifique. À mon arrivée à l'automne 1995, je me vois proposer comme sujet de recherche l'adaptation de l'amplificateur à post-compensation au moyen d'un signal pilote à spectre étalé. Mon étude préliminaire du circuit met aussitôt en évidence une lacune dans la littérature. Si l'objectif de l'amplificateur à post-compensation, avec ses amplificateurs principal et auxiliaire, est d'amplifier linéairement un signal avec un rendement meilleur qu'il ne l'est possible avec un amplificateur unique, où sont les analyses rigoureuses assurant la réalisation optimale de cet objectif? Je prends quelques initiatives de recherche en ce sens, et en 1997 la pertinence du sujet s'impose au sein du laboratoire de recherche du groupe Ampli. Le prototype d'amplificateur à post-compensation recevant un signal AMRC « stressant » ne rencontre pas les critères de linéarité, dû à des crêtes de puissance inattendues qui saturent l'amplificateur auxiliaire. Comment y remédier au moindre coût en rendement énergétique? Doit-on augmenter la capacité de l'ampli principal ou celle de l'ampli auxiliaire (ou les deux)? Faut-il ajuster le couplage de leur signaux de sortie, et de combien? Au milieu du doute général, il me vient l'intuition d'une solution optimale faisant appel à un ampli principal en classe C, que j'appellerai plus tard *amplificateur à post-compensation pré-écrêteur*. Il ne m'en faut pas plus pour changer officiellement mon sujet de recherche.

En 1998, je batis peu à peu un logiciel d'optimisation de l'ampli à post-compensation qui promet de tenir compte de tous les paramètres essentiels de conception: ce sera ma « traversée du désert ». Je démarre ensuite en 1999 une longue série de simulations visant à valider le concept de l'ampli pré-écrêteur. Je découvre ainsi les effets surprenants d'un signal à facteur de crête élevé sur les paramètres de conception optimaux, comparativement au signal à deux tonalités classique, mais rien pour justifier l'emploi

d'un ampli principal en classe C. Résigné, j'entreprends en 2000 d'améliorer mes programmes et de préparer des validations expérimentales en vue d'un dépôt de thèse. En 2001, non loin du but, une série d'événements fortuits change pour une deuxième fois le cours de mes travaux. Dans un premier temps, une vérification de routine que j'aurais très bien pu négliger révèle un détail hallucinant: une « erreur » d'ajustement de la première boucle de l'ampli à post-compensation améliore la linéarité du circuit! Le phénomène ayant été confirmé, je commence à me demander si ce pourrait être la clé manquante pour valider le concept de l'ampli pré-écrêteur, que mon directeur de thèse ne m'avait pas laissé oublier. Ne voulant pas passer le reste de ma vie à me demander si j'avais passé à côté de quelque chose, j'exécute quelques simulations difficiles avec peu d'enthousiasme, et trouve à ma grande surprise une variante du concept recherché, avantageuse au point de vue énergétique. Le concept original, encore plus avantageux, est à son tour découvert quelques mois plus tard de façon bien peu glorieuse. À une occasion, en effet, une minimisation en deux dimensions est démarrée par erreur à partir d'un point très éloigné de la solution, et tout juste assez loin pour tomber dans une deuxième cuvette dont j'ignorais l'existence.

Intuitions folles, coups de chance, tâtonnements désespérés et erreurs d'exécution: autant de réalités que cet ouvrage s'est efforcé de dissimuler sous le couvert de raisonnements rectilignes. Ne soyez pas dupes, et bonne lecture!

CHAPITRE I

INTRODUCTION

1.1 Signaux à enveloppe constante

Les amplificateurs de puissance ont tous des fonctions de transfert d'enveloppe plus ou moins non linéaires et par conséquent introduisent de la distorsion d'intermodulation dans les signaux de communication modulés en amplitude. Cette distorsion est particulièrement indésirable puisqu'elle déborde sur les canaux adjacents et qu'elle ne peut être filtrée. Afin d'éviter ce type de distorsion, les systèmes de communications doivent utiliser des techniques de modulation à enveloppe constante, telles que MF, MP et MIC. Ces systèmes peuvent du même coup utiliser des amplificateurs de classe C fonctionnant au point de saturation pour un rendement énergétique optimal.

À cette solution apparemment idéale correspondent tout de même de sérieux inconvénients. D'une part, la modulation à enveloppe constante entraîne une consommation spectrale relativement élevée, ce qui peut être coûteux à une époque où le trafic de communication est sans cesse augmentant. Remplacer les techniques à enveloppe constante par les techniques de modulation *linéaire* permet d'améliorer le rendement spectral et par conséquent la capacité des systèmes. Un exemple est la proposition de Seidel (1971) d'utiliser la transmission MA à bande latérale unique à la place de la transmission MF dans les stations-relais à micro-ondes.

D'autre part, les systèmes multi-porteuse tels que les stations cellulaires et les répéteurs de satellite requièrent la transmission simultanée de plusieurs canaux d'information (des centaines dans beaucoup de systèmes à satellites), et même si la modulation à enveloppe constante est utilisée de façon exclusive, le signal multi-porteuse aura, lui, une enveloppe variable. Pour éviter la distorsion d'intermodulation, les porteuses doivent donc être

amplifiées individuellement par des amplificateurs monovoie avant d'être rassemblées dans un combinateur à grande puissance. Malheureusement, cette utilisation du combinateur comporte des inconvénients aux points de vue énergétique et spectral. Pour N porteuses, la puissance de sortie d'un combinateur hybride est réduite à $1/N$ de la puissance totale à l'entrée, ce qui entraîne des pertes de puissance considérables (Bolan et Eisenberg, 1995). Le combinateur à cavité, lui, gaspille moins de puissance (la moitié), mais même avec les nouveaux modèles à syntonisation automatique, on ne peut pas utiliser l'attribution dynamique de canaux (ADC) parce que les flancs de courbe des filtres à cavité exigent une séparation minimale de 630 kHz entre les canaux. On évite ces inconvénients en combinant les signaux à des niveaux relativement faibles, là où les pertes de puissance sont tolérables, avec un combinateur hybride insensible à la fréquence. La sortie est par la suite amplifiée à l'aide d'un amplificateur de puissance multi-porteuse (APMP) dont la nature sera précisée. Les configurations de transmission de sites cellulaires avec combinateur à cavité et avec APMP sont illustrées aux figures 1.1 et 1.2, respectivement (Drucker, 1994).

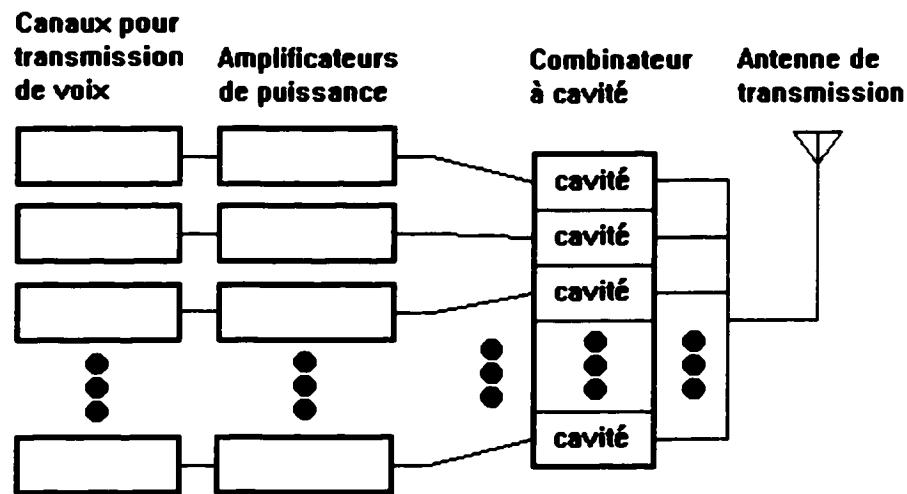


Figure 1.1 Technique traditionnelle d'amplification de canaux multiples pour transmission par antenne unique

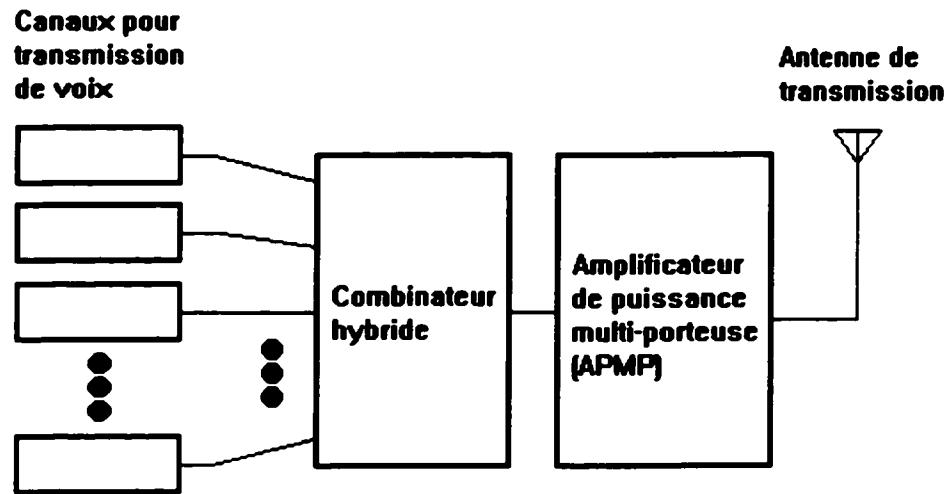


Figure 1.2 Technique d'amplification de canaux multiples au moyen d'un amplificateur de puissance multi-porteuse (APMP)

1.2 Amplificateur de classe A en recul de puissance

Les signaux de communication à enveloppe variable requièrent l'usage d'amplificateurs à haut degré de linéarité pour minimiser la distorsion d'intermodulation. Dans les satellites de télécommunications, l'usage d'un amplificateur conventionnel de classe A ou même AB est souvent acceptable puisque les normes d'intermodulation sont moins sévères que dans les systèmes terrestres. Traditionnellement, l'APMP est un amplificateur de classe A auquel on demande de produire une puissance totale de sortie de 6 à 10 dB inférieure à sa puissance de sortie en régime de saturation. À ce recul de puissance correspond une réduction du rendement énergétique puisque le niveau de fonctionnement en CC doit être maintenu au niveau original. Ceci entraîne à son tour une augmentation considérable de poids et donc de coût pour une charge utile de satellite.

Dans les émetteurs de stations cellulaires, les produits d'intermodulation doivent être inférieurs aux signaux désirés d'au moins 60 dB (et plus dans certains systèmes) si l'on veut surmonter les problèmes « near-far » qui prédominent (Kenington, 1992). Il est

possible d'atteindre ce niveau de performance avec un amplificateur de classe A, mais seulement avec un grand recul de puissance (typiquement 25 à 30 dB), donc avec un faible rendement énergétique (typiquement à 1 ou 2 %) et une forte dissipation de chaleur. Cela augmente les coûts au niveau de l'achat, de l'utilisation et de l'entretien du système de refroidissement ainsi qu'au niveau de la consommation électrique. De plus, le coût d'un amplificateur pouvant saturer aussi haut par rapport au niveau de puissance requis peut être très élevé.

1.3 Linéarisation

Au point de vue du rendement énergétique, une solution préférable à un amplificateur de classe A en recul de puissance est un amplificateur à meilleur rendement qu'on linéarise au moyen d'un montage externe à faible consommation appelé *linéarisateur*. Dans les télécommunications par satellite, l'amplificateur à linéariser peut fonctionner à saturation ou en classe C. La linéarisation d'un amplificateur de classe C à haut rendement lui permet de rivaliser en performance avec l'amplificateur de classe A et de convenir pour la première fois aux applications multi-porteuse. Les stations cellulaires, qui font face à des normes d'intermodulation plus sévères, exigent que l'amplificateur à linéariser soit déjà « linéaire », mais celui-ci peut fonctionner en classe A ou AB avec un recul de puissance raisonnable, tout comme les amplificateurs utilisés dans les systèmes à satellites en l'absence de linéarisation (voir à la section 1.2). Les trois méthodes de linéarisation couramment employées sont la contre-réaction, la post-compensation et la prédistorsion.

1.4 Contre-réaction

L'ingénieur en électronique Harold S. Black s'est rendu célèbre par son invention de la contre-réaction en 1927. Cette technique, avec l'attrait de la simplicité, retourne à l'entrée de l'amplificateur un échantillon du signal de sortie pour y être comparé. La contre-réaction ignore donc le fait que l'entrée et la sortie ne sont pas simultanées et ne peuvent être

comparées directement. En pratique, elles sont à peu près simultanées si la vitesse de l'amplificateur est grande comparée au débit d'information. Dans le cas contraire, des considérations de stabilité excluent l'usage de la technique. La largeur de bande de la méthode est donc limitée par le délai de transmission de l'amplificateur. Dans le cas de l'amplificateur à tube à ondes progressives (ATOP), qui dépend précisément du temps de transit pour l'amplification, la contre-réaction devient tout à fait inapplicable. Il ne suffit pas de diminuer le délai de l'amplificateur en diminuant son nombre d'étages d'amplification et par conséquent son gain. Ce gain a effectivement peu d'influence sur le gain moyen du circuit (qui est surtout déterminé par le facteur de la boucle de réaction) mais il influe sur la qualité de la correction apportée par la méthode. Il y a donc un rapport inverse entre la largeur de bande du circuit et sa performance de linéarisation.

Dans le cas des amplificateurs de puissance à semi-conducteurs, la contre-réaction convient aux applications à bande étroite telles que les combinés cellulaires, qui sont des systèmes à porteuse unique, mais pas aux applications à large bande telles que les communications par satellite. Dans les stations cellulaires, la contre-réaction permet la modulation linéaire utilisée par exemple dans le système cellulaire numérique AMRT (accès multiple par répartition dans le temps), mais les porteuses doivent être amplifiées individuellement au sein de la configuration traditionnelle illustrée à la figure 1.1. avec les inconvénients que cela implique. Cette configuration ne convient pas au système AMRC (accès multiple par répartition de code) parce que la bande passante relativement large des canaux individuels empêche la contre-réaction de rencontrer les hauts degrés de linéarité requis (Drucker, 1994).

De façon générale, la contre-réaction n'est pas adaptée aux fréquences micro-ondes. En premier lieu, l'ATOP dont il était question plus haut est le dispositif traditionnellement utilisé à ces fréquences. En deuxième lieu, les micro-ondes sous-entendent habituellement un système à large bande. En effet, les fréquences élevées rendent disponible une plus

grande largeur de bande par canal, permettant ainsi de satisfaire à la demande croissante pour les communications à haute vitesse (Golio, 2001).

1.5 Prédistorsion adaptative en bande de base

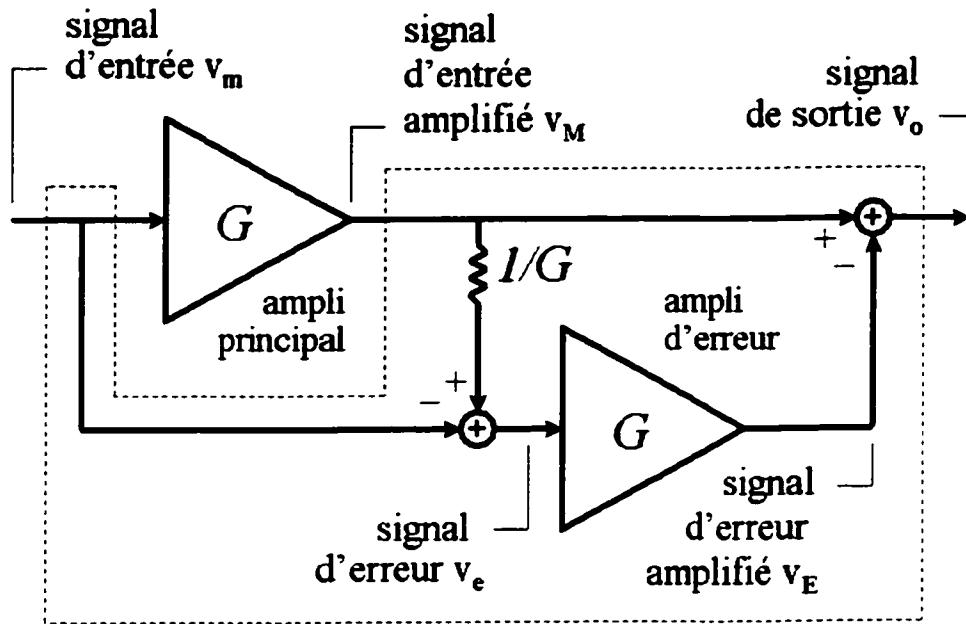
La prédistorsion adaptative en bande de base est une autre méthode de linéarisation qui s'apparente à la contre-réaction en raison d'une bande passante limitée. Sa réalisation est plus complexe que celle de la contre-réaction, mais les progrès réalisés ces dernières années au niveau de la technologie numérique ont permis d'élargir sa bande passante au point où elle pourrait éventuellement concurrencer la méthode de post-compensation (section 1.6). Par contre, bien que la prédistorsion n'affecte pas directement le rendement de l'amplificateur, ce dernier doit fonctionner avec suffisamment de recul pour éviter l'écrêtage de l'enveloppe, ce que la prédistorsion est incapable de corriger.

1.6 Post-compensation

1.6.1 Travaux de H. S. Black

La popularité que connaît la post-compensation depuis une décennie fait souvent oublier que Black l'avait déjà inventée trois ans avant la contre-réaction (Black, 1928). La figure 1.3 donne le schéma conceptuel d'un amplificateur linéarisé par post-compensation, tel que Black l'avait envisagé à l'époque. Comme le linéarisateur comporte lui-même un amplificateur, ce dernier et l'amplificateur à linéariser sont respectivement qualifiés d'*auxiliaire* et de *principal*, et l'ensemble est désigné *amplificateur à post-compensation*. Dans la première boucle, le linéarisateur prélève un échantillon du signal d'entrée amplifié et l'atténue au niveau d'origine pour en soustraire le signal d'entrée. L'annulation du signal isole le signal d'erreur qui est ensuite amplifié par l'amplificateur auxiliaire (ou *d'erreur*) pour le ramener à son propre niveau d'origine. Dans la deuxième

boucle. le signal d'erreur amplifié est finalement soustrait du signal d'entrée amplifié pour annuler l'erreur et produire un signal de sortie théoriquement sans distorsion.



Linéarisateur à post-compensation

Figure 1.3 Schéma fonctionnel de l'amplificateur à post-compensation tel qu'envisagé par H. S. Black

Dans ce processus, contrairement à la contre-réaction, la sortie de l'amplificateur ne retourne pas à l'entrée puisque les comparaisons se font toujours vers l'avant. Sans boucle de réaction, le fonctionnement du circuit devient inconditionnellement stable sur toutes les bandes passantes, mais la comparaison de signaux non simultanés reste un problème puisqu'elle affecte la précision de la correction obtenue. Le processus est donc lui aussi limité en largeur de bande, cette fois par les délais de transmission des deux amplificateurs. Pour cette raison, la post-compensation était au début presque exclusivement utilisée comme technique de redondance. En effet, si les deux amplificateurs sont identiques, une défaillance de l'un ou de l'autre produit un retour

gracieux au fonctionnement sans correction plutôt qu'une interruption complète. Cette possibilité de réserve est hautement désirable pour les applications de communications. En ce qui concerne la linéarisation, les chercheurs ont reconnu le besoin de maintenir l'équilibre des phases, mais sans doute parce que la pression était moindre à l'époque, ils n'ont pas su voir le problème en termes de synchronisation des signaux et produire les techniques de circuit correspondant à ces besoins.

1.6.2 Contributions de H. Seidel

Seidel et ses collègues ont commencé dès 1966 à réaliser le potentiel linéarisateur de l'amplificateur à post-compensation en ajoutant des lignes à délai pour compenser les délais des amplificateurs principal et d'erreur et rendre les signaux à comparer compatibles dans le temps. Leurs travaux ont ainsi produit des résultats impressionnantes sur des bandes passantes de 10 à 30 MHz, premièrement avec des amplificateurs VHF et finalement aux fréquences micro-ondes avec des ATOP à plus de 13 ns de délai (Seidel, 1971). Les amplificateurs à temps de transit zéro n'étaient donc pas une exigence fondamentale mais bien une façon de rendre acceptable le fonctionnement de la contre-réaction. Cette barrière maintenant enlevée, et la post-compensation disponible sans exception, le problème de la fabrication de dispositifs à temps de transit zéro disparaissait, et les efforts de conception pouvaient être reconscérés à des usages plus productifs.

Une autre contribution importante de Seidel a été de sacrifier la fonction de redondance du circuit à des considérations énergétiques en remplaçant le combinateur hybride de sortie par un coupleur directif favorisant le signal principal, et en réduisant la taille de l'amplificateur d'erreur relativement à celle de l'amplificateur principal. Ainsi, l'amplificateur à post-compensation se consacrait exclusivement à sa fonction de correction, mais Seidel n'a pas mis l'accent sur ce développement parce que le rendement énergétique n'était pas sa préoccupation première. Seidel a par ailleurs reconnu l'influence des changements dans les conditions de fonctionnement en ajoutant un

contrôle adaptatif pour maintenir l'équilibre de gain et de phase dans la deuxième boucle. En fait, les précisions des soustractions de signaux sont essentielles dans les deux boucles, et sont affectées entre autres par le vieillissement des composantes, les changements de température et les altérations des propriétés statistiques du signal d'entrée (en ordre croissant de rapidité). Ces changements, tous beaucoup plus lents que la modulation de l'enveloppe même, sont neutralisés aujourd'hui par des systèmes adaptatifs qui contrôlent selon une variété de méthodes deux gains complexes sous la forme de modulateurs vectoriels précédant les amplificateurs.

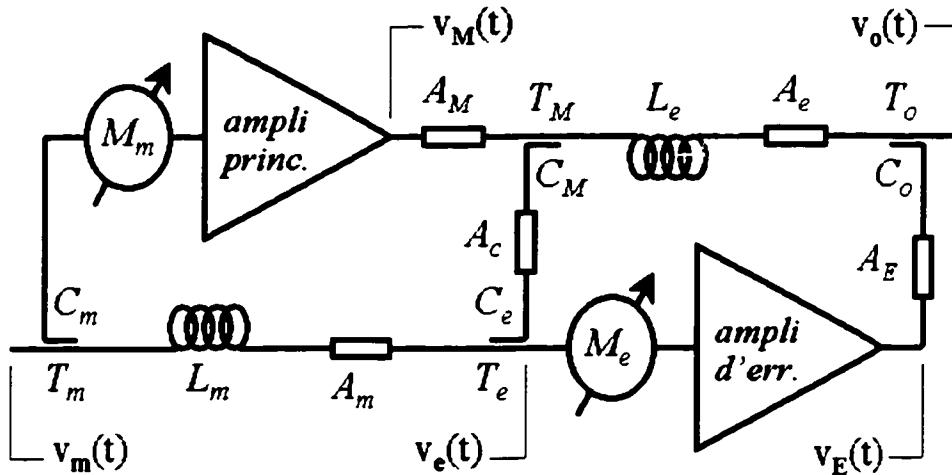


Figure 1.4 Modèle complexe en bande de base de l'amplificateur à post-compensation

1.6.3 Amplificateur à post-compensation moderne

Les améliorations apportées par Seidel complètent essentiellement l'amplificateur à post-compensation moderne tel que représenté par le modèle de la figure 1.4. Les signaux sont représentés par des enveloppes complexes variant dans le temps, et les composantes le sont par des gains complexes. En plus des éléments décrits ci-haut, le modèle divise et combine les signaux à l'aide de coupleurs directifs, et représente les inévitables pertes par

la série « A » d'atténuateurs. Le coupleur d'entrée favorise le signal de référence afin de préserver au maximum l'intégrité de ce signal (Jurgen, 1972).

Le degré d'annulation de la distorsion se situe typiquement entre 25 et 30 dB avec la technologie actuelle, selon la linéarité de l'amplificateur d'erreur ainsi que l'équilibre en gain et en phase obtenu dans la deuxième boucle sur la bande passante requise. Ces facteurs ne dépendant aucunement du niveau original de la distorsion, le niveau absolu de distorsion à la sortie peut être aussi bas que lui permet ce niveau original de distorsion. En fait, il est théoriquement possible de supprimer la distorsion à un niveau arbitrairement bas en superposant le nombre requis d'étages de post-compensation, chaque étage utilisant le même signal d'entrée comme référence. Ceci distingue la post-compensation des autres techniques de linéarisation où le niveau de distorsion à la sortie ne peut descendre sous un certain plancher.

La version non adaptative de l'amplificateur à post-compensation a rapidement trouvé un usage dans les grands systèmes à large bande tels que les systèmes à satellite et la câblodistribution, où la complexité du circuit était largement justifiée par les économies d'énergie réalisables. Dans les stations cellulaires, dû aux normes d'intermodulation plus sévères, l'usage de la post-compensation a dû attendre le développement de systèmes d'adaptation dont la taille et le coût permettaient l'exploitation commerciale (Narahashi et Nojima, 1991). En 1995, l'avenir de la post-compensation était assuré avec le lancement des premiers réseaux AMRC, dont les stations de base requièrent un amplificateur ultra-linéaire à large bande passante. Depuis lors, l'AMRC est devenu la technologie sans fil à la croissance la plus rapide, avec plus de 100 millions d'abonnés mondialement (Qualcomm Inc., 2002).

1.7 Prédistorsion RF

La prédistorsion RF, qu'on ne doit confondre avec sa version adaptative en bande de base (section 1.5), produit des niveaux de linéarité modestes mais fonctionne sans égard à la bande passante, ce qui lui permet de concurrencer la post-compensation dans les applications à linéarité moyenne comme les systèmes à satellite. Comme dans la prédistorsion adaptative, l'amplificateur doit fonctionner avec suffisamment de recul, mais selon Parsons et Kenington (avr. 1994), le rendement qu'on peut réaliser est supérieur à ce qui est possible avec la post-compensation. La prédistorsion RF pourrait aussi être le moyen d'atteindre des temps de conversation raisonnables avec les futurs combinés AMRC large bande (WCDMA) (Bateman. 1999). Les stations de base demeurent donc le seul terrain où l'exclusivité de l'amplificateur à post-compensation n'est pas contestée.

1.8 Rendement de l'amplificateur à post-compensation

Pour donner une idée de la puissance consommée par les stations de base, certaines prévisions placent le nombre requis de stations à 50 000 pour un réseau complet de communications de future génération dans un pays européen majeur. Avec une demande d'environ 0.625 kWh par cellule, basée sur huit porteuses et un amplificateur à post-compensation, cela représente une consommation annuelle de 830 millions de kWh, soit environ l'équivalent de ce que consomme une ville de taille moyenne en deux mois (Kenington, 1999). On voit donc l'intérêt qu'il y aurait à améliorer le rendement de l'amplificateur à post-compensation. Comme on le constate à la figure 1.4, ce rendement est affecté par la puissance consommée dans l'amplificateur d'erreur ainsi que par les pertes subies par le signal principal, notamment dans le coupleur de sortie et la ligne à délai. Les stratégies proposées pour améliorer le rendement du circuit peuvent être dites « de matériel » ou « de logiciel », selon qu'elles se rapportent à des technologies plus performantes ou à des façons plus intelligentes d'utiliser la technologie existante.

1.8.1 Stratégies « de matériel »

La majorité des articles publiés proposent des stratégies de matériel, et se penchent surtout sur l'amplificateur principal. Ceci est dû au fait que la consommation du linéarisateur à post-compensation diminue selon la linéarité initiale du signal à corriger. Par conséquent, pour un rendement donné de l'amplificateur principal, il est avantageux que celui-ci soit aussi linéaire que possible. Ceci confère une nouvelle pertinence à l'optimisation des amplificateurs. Par exemple, c'est dans le contexte de l'amplificateur à post-compensation que Vennema (1996) fait valoir son amplificateur de classe AB à linéarité supérieure. De même, cela confère une importance renouvelée aux autres techniques de linéarisation qui procurent un niveau de linéarité plus modeste mais ne réduisent pas de façon significative le rendement de l'amplificateur. Ainsi, Parsons et autres proposaient en 1994 d'utiliser l'autre méthode à large bande, soit la prédistorsion RF, pour corriger une première fois l'amplificateur principal avant de compléter le travail par post-compensation. Cette approche est souvent utilisée aujourd'hui. D'autre part, Faulkner proposait en 1998 de combiner la post-compensation avec la contre-réaction, mais comme on pouvait s'y attendre, cela diminue la bande passante. D'autres stratégies de matériel récentes concernaient l'amplificateur d'erreur (Kouki et Outaleb, 2000) et la ligne à délai de sortie (Wixon, 2001; Jeckeln et autres, 2001).

1.8.2 Stratégies « de logiciel »

L'interaction complexe de divers dispositifs micro-ondes au sein de l'amplificateur à post-compensation multiplie les différentes configurations de circuit susceptibles de délivrer la performance requise à la sortie. Les configurations qui maximisent le rendement énergétique global sont les stratégies « de logiciel » qui constituent le sujet de la présente thèse.

CHAPITRE II

REVUE DE LA LITTÉRATURE

Au titre de stratégies « de logiciel » visant à améliorer le rendement de l'amplificateur à post-compensation, la présente thèse comble une lacune dans la littérature puisque les articles rigoureux publiés dans cette catégorie, à l'exclusion des articles de l'auteur, se chiffrent à cinq sur une période de dix ans (plus un sixième récemment soumis pour publication et révisé par l'auteur). La difficulté du sujet réside dans les relations non linéaires régissant la puissance consommée et la linéarité des amplificateurs principal et d'erreur.

2.1 Kenington et autres

2.1.1 Formule de Kenington

Afin d'éviter les non-linéarités dans l'analyse, Kenington (1992) a fixé au départ le type et le rendement de chacun des amplificateurs, de même que l'erreur d'ajustement dans chacune des boucles (dans le cas présent, des erreurs nulles). Le type et le rendement fixes de l'amplificateur principal lui confèrent une linéarité fixe (Raab, 1986). Ceci, étant donné l'annulation parfaite du signal dans la première boucle, fixe à son tour le format du signal d'erreur, de sorte que le type et le rendement fixes de l'amplificateur d'erreur lui confèrent à lui aussi une linéarité fixe. Finalement, l'ajustement optimal de la deuxième boucle fixe la linéarité de sortie du circuit. Avec ces suppositions, le rendement global du circuit a comme seule variable le couplage de sortie, et tend vers zéro lorsque le couplage approche l'une ou l'autre de ses valeurs extrêmes. Certains auteurs (Lubell et autres, 1974; Bennett et Clements, 1974) avaient pressenti le dilemme entourant le coupleur de sortie. Diminuer le couplage réduit la perte de couplage que doit subir le signal principal et diminue la taille que doit avoir l'amplificateur principal pour produire la puissance

requise à la sortie du circuit. En contrepartie, la taille de l'amplificateur d'erreur doit augmenter pour compenser le couplage réduit. La formule de Kenington donne le couplage de sortie optimal en fonction des rendements des amplificateurs et de la linéarité de l'amplificateur principal.

2.1.2 Formule de Kenington avec perte de délai

La formule de Kenington a été plus tard améliorée en tenant compte des pertes dans la ligne à délai de la branche principale (Parsons et Kenington, mai 1994). Le délai dans l'amplificateur d'erreur étant relié au gain de ce dernier, réduire ce gain permet de réduire la longueur de la ligne à délai et les pertes qu'elle occasionne. En contrepartie, il faut rétablir l'équilibre des gains dans la seconde boucle en augmentant le couplage du coupleur de sortie ou du coupleur principal (qui suit l'amplificateur principal), ce qui augmente les pertes de couplage dans la branche principale. Il semble donc exister un délai optimal que la formule de Kenington améliorée ne peut nous révéler.

2.1.3 Réduction de la ligne à délai

En avril 1994, Parsons et Kenington soulignaient que dans les systèmes à bande étroite, la ligne à délai de l'amplificateur à post-compensation pouvait être réduite ou même supprimée afin d'améliorer le rendement énergétique tout en maintenant en grande partie la linéarité. L'idée n'a apparemment pas été retenue par l'industrie. Le problème n'a pas été étudié dans le contexte de cette thèse puisqu'un déséquilibre dans les délais est équivalent à une réponse fréquentielle, ce dont l'analyse statistique employée ne peut tenir compte.

2.1.4 Limites de la formule de Kenington

On peut reprocher à la formule de Kenington de tenir compte uniquement des produits d'intermodulation de troisième ordre dans un signal classique à deux tonalités, dont la modulation d'enveloppe est bien plus faible que celle employée dans les signaux de communication actuels. Encore plus inquiétant, la formule suppose des annulations parfaites dans les deux boucles, ce qui n'est pas réaliste. Par exemple, il suffit d'une légère imprécision dans la première boucle pour augmenter de façon notable la puissance du signal d'erreur. Eid et autres (1995) utilisent quand même la formule de Kenington, mais ils se contredisent en supposant des annulations imparfaites dans le reste de leur analyse.

Par ailleurs, ces auteurs illustrent bien certaines difficultés rencontrées à l'usage de la formule de Kenington. En premier lieu, la formule ne révèle pas la combinaison optimale des rendements d'amplificateurs qui procurera au circuit la linéarité requise par le cahier des charges. Il faut donc fixer arbitrairement le rendement de l'amplificateur principal et déterminer expérimentalement le rendement de l'amplificateur d'erreur qui procurera la linéarité demandée. Eid et autres, qui n'ont pas fait le lien entre la linéarité de l'amplificateur d'erreur et son rendement, n'imposent à ce rendement que la limite supérieure de 50 % associée aux amplificateurs de puissance de classe A (limite qui, par surcroît, suppose un impossible signal d'erreur à enveloppe constante). En deuxième lieu, le couplage de sortie optimal impose aux deux amplificateurs les puissances de sortie qui procureront au circuit le gain requis par le cahier des charges, et ces puissances à leur tour imposent aux amplificateurs les tailles qui confirmeront les rendements et les linéarités supposées au préalable dans le calcul du couplage optimal. Dans le cas où les amplificateurs étaient déjà choisis, cela veut dire que leurs rendements et leurs linéarités doivent être préservés en changeant dans la même proportion le gain et la taille de chaque amplificateur. Ces changements peuvent s'avérer difficiles en pratique. Eid et autres, qui choisissent au départ leur amplificateur principal et n'ont pas l'intention de le changer,

n'ont pas le choix que d'ignorer leur promesse de procurer au circuit un gain désiré. L'alternative était d'ignorer la formule de Kenington et de calculer le couplage de sortie uniquement en fonction de ce gain désiré.

2.2 Horiguchi et autres

2.2.1 Amplificateurs à tailles fixes

Andrenko, Horiguchi et autres (1999) ont essayé de surmonter les limites de la formule de Kenington avec leur propre analyse d'optimisation. Ils ont pris le point de départ plus réaliste où l'on dispose de deux amplificateurs donnés dont on ne peut changer que les gains, ce qui se fait facilement en pratique. Par contre, en figeant les tailles des amplificateurs et en permettant à leurs rendements de changer, ils ont sacrifié le mécanisme par lequel Kenington préservait la linéarité du circuit. Cette erreur étant déjà fatale, on ne mentionnera pas ici les autres erreurs conceptuelles remarquées à la lecture de l'article.

2.2.2 Amplificateurs à rendements fixes

Dans un article subséquent (Horiguchi et autres, 2001), la même équipe japonaise est revenue au concept de rendements fixes et a repris la formule de Kenington, ayant corrigé ses deux faiblesses majeures au moyen d'éléments empruntés de leur article précédent. En premier lieu, ils ont redéfini la linéarité du signal sortant de l'amplificateur principal en tant que la fraction de sa puissance totale attribuée à la distorsion, ce qui permet d'accorder tous les types de signaux. Ils ont ensuite tenu compte correctement de l'imprécision dans l'ajustement de la première boucle, mais ils oublient de mentionner que la phase de l'erreur d'ajustement influe sur le rendement requis de l'amplificateur d'erreur. En effet, les présents travaux ont révélé une influence marquée de cette phase sur le format du signal d'erreur. Idéalement, on voudrait supposer des erreurs

d'ajustement réalistes dans les deux boucles, ainsi que des phases à ces erreurs qui exigent de l'amplificateur d'erreur le pire rendement pour une linéarité donnée du circuit, tel que déterminé expérimentalement. Une autre amélioration apportée est l'inclusion de la perte de couplage due au coupleur principal. Les auteurs, tout en se concentrant exclusivement sur le couplage de sortie, prétendent que leurs équations leur donnent aussi la possibilité d'optimiser le couplage principal ainsi que la perte de délai. Cependant, on constate par inspection qu'on obtiendrait les solutions triviales impossibles correspondant à un couplage principal nul et un délai nul. Ceci est dû au fait que la relation entre le gain et le délai de l'amplificateur d'erreur, expliquée à la section 2.1.2, n'a pas été prise en considération.

2.3 Au-delà de Kenington

2.3.1 Hickson et autres

Même exploitée à sa limite, l'approche de Kenington laisse une question sans réponse. En effet, parmi les combinaisons de rendements d'amplificateurs qui procurent au circuit la linéarité requise, quelle est celle qui maximise le rendement global du circuit? Hickson et autres (1994) ont obtenu par simulation la linéarité de sortie d'un amplificateur à post-compensation en fonction de la taille de l'amplificateur d'erreur et de la puissance de sortie du circuit. Ce processus, équivalent à calculer la linéarité de sortie en fonction des rendements des deux amplificateurs, a illustré le fait qu'une linéarité donnée puisse s'obtenir par différents rapports de rendements des amplificateurs. Par contre, n'ayant pas utilisé la formule de Kenington pour calculer les rendements globaux réalisables avec ces différents rapports, les auteurs ne pouvaient juger du rapport de rendement optimal. Ils pouvaient seulement éliminer des possibilités en constatant que passé une certaine taille de l'amplificateur d'erreur, la linéarité du circuit plafonnait. Cette observation, expliquée par Wilkinson et Kenington (1992), indique qu'ils ont dû simuler des déviations de gain et de phase dans l'ajustement de la deuxième boucle. Parmi les possibilités éliminées

figurait la configuration classique de Black dans laquelle la taille de l'amplificateur d'erreur rejoint celle de l'amplificateur principal.

2.3.2 Larose et Ghannouchi

L'approche d'optimisation du présent auteur a été présentée en conférence pour la première fois il y a deux ans (Larose et Ghannouchi, 2000). Tous les éléments étaient déjà en place, sauf que l'atténuation de la ligne à délai de sortie, ainsi que les phases des erreurs d'ajustement des boucles, devaient encore être fixées au préalable. L'approche devait encore démontrer tout son potentiel en permettant l'étude des méthodes d'adaptation ou des types d'amplificateurs employés. Ceux-ci ne pouvaient être optimisés directement mais pouvaient néanmoins être investigués par essais et erreurs. C'est ainsi qu'une nouvelle méthode d'adaptation avantageuse au plan énergétique, devant plus tard être désignée *optimale de droite*, a été découverte et présentée l'année suivante (Larose et Ghannouchi, juin 2001). Un autre article de conférence a suivi pour préciser la meilleure réalisation pratique de la méthode (Larose et Ghannouchi, août 2001). L'année suivante, une équipe coréenne présentait un prototype d'amplificateur à post-compensation utilisant avec succès l'adaptation optimale de droite (Woo et autres, 2002).

L'adaptation est sans doute le dernier terrain où on aurait pensé trouver de nouvelles options énergétiques pour l'amplificateur à post-compensation. Personne n'avait auparavant fait un tel lien parce qu'avant la méthode optimale de droite, les méthodes disponibles se valaient toutes dans la mesure où elles étaient correctement appliquées. De plus, toute découverte antérieure de la méthode optimale de droite aurait laissé indifférent dans le contexte d'un signal d'entrée à deux tonalités, où la méthode procure peu de bénéfices.

CHAPITRE III

ORGANISATION GÉNÉRALE

3.1 Articles en annexe

En 1992 (résolution CRES 64-9). l'École Polytechnique a introduit la possibilité de présenter la thèse par un regroupement d'articles. L'auteur s'est prévalu ici de cette possibilité avec deux articles soumis pour publication dans des revues avec comité de lecture. Ces revues, ayant publié de nombreux articles traitant de l'amplificateur à post-compensation, sont reconnues dans la discipline. Le premier article, « Optimization of feedforward amplifier power efficiency on the basis of drive statistics », a été accepté le 11 mars 2002 par le *IEEE Transactions on Microwave Theory and Techniques* et paraîtra dans le numéro de janvier 2003. Le deuxième, « Optimal amplifier classes and optimization methods for feedforward amplifiers driven by high-stress signals », a été soumis le 10 mai 2002 au *IEEE Transactions on Vehicular Technologies*. Conformément à la formule de présentation des articles *en annexe*, les deux articles sont regroupés en référence dans les annexes I et II.

3.2 Détails théoriques et expérimentaux

Les deux articles couvrent toute la matière destinée à la présente thèse. Ils représentent les deux phases de la recherche effectuée, soit en premier lieu la mise au point d'un logiciel d'optimisation du rendement énergétique de l'amplificateur à post-compensation, et en deuxième lieu l'exploitation systématique de cet outil pour découvrir les classes d'amplificateur et les méthodes d'adaptation optimales pour l'amplificateur à post-compensation. La rédaction a fait l'objet d'un soin qui convient pour des articles destinés en même temps à des revues de prestige et à une thèse de doctorat. Par contre, les contraintes d'espace imposées par les revues ont demandé des efforts considérables de condensation de l'écriture. Les détails non essentiels sacrifiés dans cet exercice, malheureusement, sont

souvent parmi les plus évocateurs et contribuent à une compréhension intuitive de la matière. Dans d'autres cas, au contraire, ils alourdissent la lecture des articles mais aident néanmoins à vérifier la matière et à l'exploiter de façon optimale. Ces nombreux détails trouvent ici leur place au sein du Chapitre IV: Détails techniques et expérimentaux. Ce chapitre est suivi d'une conclusion comprenant une discussion générale et intégrée des résultats et les recommandations pertinentes.

3.3 Rédaction en français

La rédaction française de cette thèse a été grandement aidée par le site Internet *Termium* (<http://www.termiumplus.bureauodelatraduction.gc.ca>) du Bureau de la traduction du Canada, ainsi que le site *Le Grand dictionnaire terminologique* de l'Office de la langue française (<http://www.granddictionnaire.com>). Pour ce qui est du mot anglais *feedforward*, cependant, l'abondance des traductions proposées (régénération, correction aval, réaction précurseive, action directe secondaire, etc.) et le souci de continuité par rapport à une rédaction antérieure à l'École Polytechnique (Eid, 1995) ont convaincu l'auteur de maintenir le terme *post-compensation* adopté par Eid et Beauregard après consultation auprès de l'Office de la langue française.

CHAPITRE IV

DÉTAILS TECHNIQUES ET EXPÉIMENTAUX

4.1 Premier article (annexe I)

Cette section accompagne l'article en annexe I, « Optimization of feedforward amplifier power efficiency on the basis of drive statistics », accepté le 11 mars 2002 par le *IEEE Transactions on Microwave Theory and Techniques* et devant paraître dans le numéro de janvier 2003.

4.1.1 Modèle de l'amplificateur de puissance

Un modèle complexe en bande de base d'un amplificateur de puissance, utilisé pour les amplificateurs principal et d'erreur, est décrit à la section II-B du premier article par le biais des fonctions de compression de gain et de puissance consommée, et ensuite décrit plus directement à la section II-A du second article. Ce modèle mérite des explications supplémentaires puisqu'il constitue la pièce centrale du modèle de l'amplificateur à post-compensation. L'*excitation* d'entrée d (de l'anglais *drive*) de l'amplificateur est définie comme étant le rapport de la puissance de sortie non compressée à la puissance de sortie de saturation,

$$d = \frac{P_{in}|G|^2}{P_{sat}} = \frac{P_{in}}{P_{sat}/|G|^2}. \quad \text{Équation 4.1}$$

On voit dans l'équation que le ratio des puissances de sortie produit une excitation d'entrée parce que la puissance de sortie non compressée est reliée à la puissance d'entrée par le gain de puissance $|G|^2$. P_{sat} est utilisée à la place de la puissance de sortie de compression à 1 dB, P_{1dB} , parce qu'elle s'applique à toutes les classes d'amplification.

Pour les amplificateurs à l'état solide de classe A, où G serait normalement le gain complexe en petits signaux, la puissance de référence à l'entrée $P_{\text{sat}} / |G|^2$ n'est pas très différente de la référence conventionnelle qui est la puissance d'entrée de compression à 1 dB, $P_{\text{in},1\text{dB}}$. La relation entre ces deux puissances est illustrée à la figure 4.1. Comme P_{sat} dépasse $P_{1\text{dB}}$ de 1 à 2 dB, en pratique, $P_{\text{sat}} / |G|^2$ dépasse $P_{\text{in},1\text{dB}}$ de moins de 1 dB.

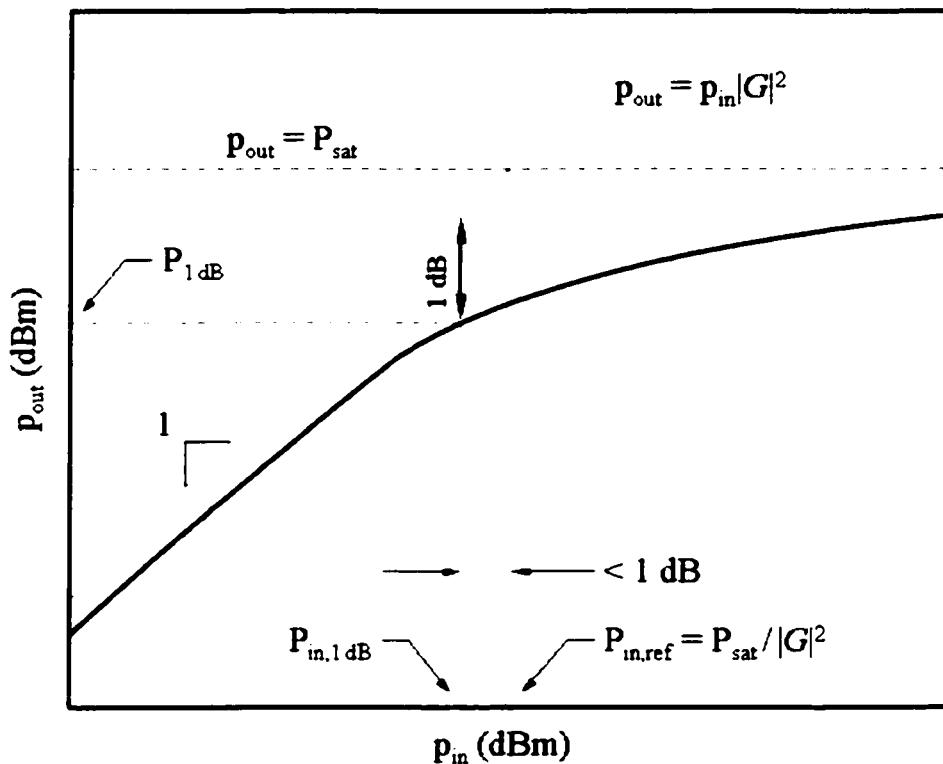


Figure 4.1 Écart typique entre les puissances de référence d'entrée $P_{\text{in},1\text{dB}}$ et $P_{\text{in},\text{ref}}$ pour un amplificateur à semi-conducteurs de classe A

Les fonctions de compression de gain et de puissance consommée sont obtenues respectivement par

$$f(d) = \frac{v_{\text{out}}(d)}{v_{\text{in}}(d)G}, \quad h(d) = \frac{p_{\text{dc}}(d)}{P_{\text{sat}}}.$$
Équation 4.2

Dans le régime A en général, lorsque G est le gain complexe en petits signaux et que l'excitation d augmente à partir de zéro, la réponse MA-MA $|f(d)|$ décroît à partir d'une valeur initiale unitaire, et la réponse MA-MP $\arg[f(d)]$ change à partir d'une valeur initiale nulle. La figure 4.2 montre que les deux fonctions définissent un type d'amplificateur indépendamment de son gain de puissance et de sa puissance de sortie de saturation. Dans cet exemple, f et h se rapportent à un amplificateur de classe C typique, et on observe l'effet d'un doublage de $|G|^2$ ou de P_{sat} sur la puissance de sortie p_{out} , la phase de sortie relative $\Delta\theta_{out}$ et la puissance consommée p_{dc} . L'expression

$$\frac{p_{out}}{P_{sat}} = \frac{p_{in}|G|^2}{P_{sat}} \left| f\left(\frac{p_{in}|G|^2}{P_{sat}}\right) \right|^2. \quad \text{Équation 4.3}$$

obtenue à partir des équations 4.1 et 4.2, démontre qu'augmenter $|G|^2$ comprime la courbe p_{out} vs. p_{in} vers l'axe de p_{out} , préservant ainsi la puissance de sortie de saturation, alors qu'augmenter P_{sat} étire la courbe le long des deux axes, préservant ainsi les proportions de la courbe et le gain qui lui est associé. D'autre part, l'expression

$$\Delta\theta_{out} = \arg(v_{out}) - \arg(v_{in}G) = \arg\left[f\left(\frac{p_{in}|G|^2}{P_{sat}}\right)\right] \quad \text{Équation 4.4}$$

démontre qu'augmenter $|G|^2$ comprime la courbe $\Delta\theta_{out}$ vs. p_{in} vers l'axe de $\Delta\theta_{out}$, alors qu'augmenter P_{sat} éloigne la courbe de cet axe, ce qui est compatible avec le comportement de la courbe de puissance de sortie. Finalement, l'expression

$$\frac{p_{dc}}{P_{sat}} = h\left(\frac{p_{in}|G|^2}{P_{sat}}\right) \quad \text{Équation 4.5}$$

démontre qu'augmenter $|G|^2$ ou P_{sat} affecte la courbe p_{dc} vs. p_{in} de la même façon que cela affectait la courbe de puissance de sortie, comme on s'y attendrait.

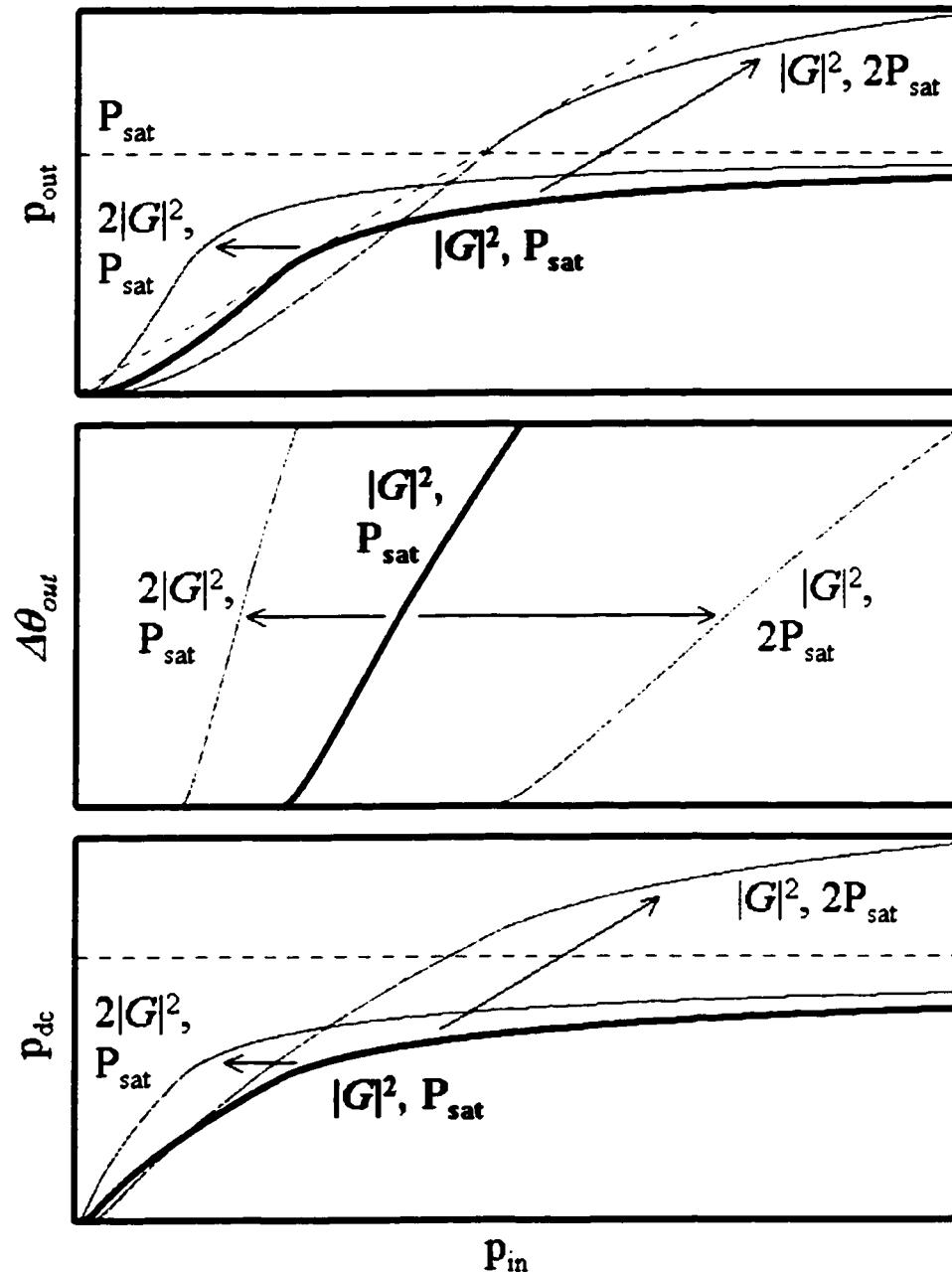


Figure 4.2 Effets du gain de puissance ou de la puissance de sortie de saturation sur la puissance de sortie, la phase de sortie relative, et la puissance consommée

La figure montre aussi que la définition du gain de puissance supposée ici et pour tout autre amplificateur ne fonctionnant pas en classe A est arbitraire, puisqu'il demeurera toujours possible de générer les mêmes infinités de courbes au moyen de $|G|^2$ et de P_{sat} .

4.1.2 Modèle de l'amplificateur à post-compensation

La figure 2 du premier article, pour des motifs de concision, omet d'illustrer la première étape de la construction du modèle normalisé de l'amplificateur à post-compensation, où le modèle d'amplificateur de puissance est substitué aux amplificateurs principal et d'erreur et le gain A_M est réparti sur deux branches. Cette étape est illustrée ici à la figure 4.3. Une répartition subséquente des gains $A_M C_M A_c C_e$ et $A_E C_o$ mènera à l'étape suivante illustrée à la figure 2(b) de l'article. Une ambiguïté sera créée par le fait que le symbole d'amplification a été retenu pour chaque combinaison d'amplificateur et de son modulateur vectoriel, mais que dans le modèle normalisé ces combinaisons seront identifiées comme étant des gains normalisés de modulateurs vectoriels. Cette ambiguïté n'a pas encore trouvé de résolution satisfaisante.

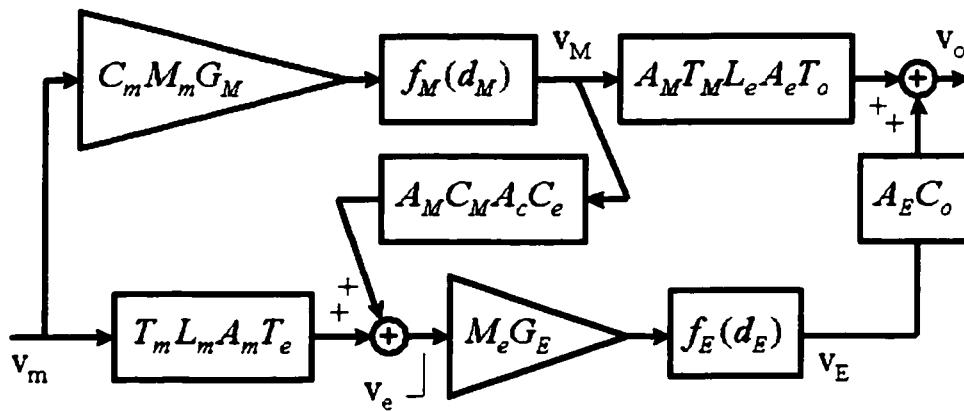


Figure 4.3 Première étape de la construction du modèle normalisé de l'amplificateur à post-compensation

La figure 4.3 procure l'occasion de souligner à nouveau la constance du gain du circuit, démontrée mathématiquement à la section III-B du premier article mais avec un impact diminué par la normalisation du modèle. Cette propriété, qui constitue la base de l'adaptation par suppression *non adaptive* du signal d'entrée (section III-D de l'article), continue d'être mal comprise par de nombreux auteurs même si d'autres l'ont comprise dès le début (par exemple Lubell et autres, 1974). La démonstration en est pourtant fort simple. Si on néglige la compression de l'amplificateur d'erreur en fixant $f_E(d_E) = 1$, le signal de sortie devient

$$v_o = v_M A_M (T_M L_e A_e T_o + C_M A_c C_e M_e G_e A_e C_o) + v_m T_m L_m A_m T_e M_e G_e A_e C_o. \quad \text{Équation 4.6}$$

M_e compense toute variation dans G_e afin de maintenir l'équilibre des gains dans la seconde boucle. Cet équilibre étant exprimé par $T_M L_e A_e T_o + C_M A_c C_e M_e G_e A_e C_o = 0$, le signal de sortie se réduit à

$$v_o = v_m T_m L_m A_m T_e M_e G_e A_e C_o = -v_m \frac{T_m L_m A_m T_e T_M L_e A_e T_o}{C_M A_c C_e}. \quad \text{Équation 4.7}$$

La sortie est donc indépendante du signal amplifié v_M , et est égale à l'entrée multipliée par un gain stable exprimable en termes des gains des trois branches passives. Par conséquent, parmi tous les types de distorsion que le circuit a la réputation d'annuler indifféremment (intermodulation, effets de mémoire, bruit, etc.), il faut inclure les fluctuations dans le gain du signal amplifié. Ces fluctuations sous-entendent évidemment une compensation imparfaite de la part du premier modulateur vectoriel. Ce résultat ne nie pas l'importance du modulateur vectoriel; on sait déjà qu'une erreur d'ajustement peut augmenter le signal d'erreur de manière à dégrader la linéarité de l'amplificateur d'erreur. Le gain du circuit, par contre, ne sera pas affecté de façon significative. La confusion qui règne encore à ce sujet est l'héritage d'une pratique qui consistait à transférer le premier modulateur vectoriel à la branche passive du bas, d'où il influençait

le gain du circuit (Bauman, 1983). On note que Cavers (1995), qui simplifie son analyse par cette approche, conclut avec Bauman à la nécessité d'une suppression *adaptative* du signal d'entrée. La confusion provient également d'un faux raisonnement persistant concernant le résidu de signal d'entrée qui infiltre le signal d'erreur à la suite d'un changement dans le gain du signal amplifié. Comme l'intermodulation réinjectée à la sortie interfère toujours de façon destructive avec l'intermodulation présente dans le signal amplifié, on raisonne que le résidu de signal d'entrée réinjecté à la sortie doit lui aussi interférer de façon destructive avec le signal amplifié, quelle qu'ait été la phase du changement de gain (Eid et autres, 1995). Par contre, lorsqu'on reconnaît le résidu de signal d'entrée comme étant une partie intégrante de la distorsion réinjectée, on comprend que ce résidu interfère de façon destructive non pas avec le signal amplifié mais plutôt avec le *changement* dans le gain du signal amplifié. Dans le cas où ce gain aurait diminué en module, par exemple, le résidu de signal d'entrée doit interférer de façon constructive avec le signal amplifié.

4.1.3 Méthode d'optimisation pentagonale

Les minimisations en deux dimensions dont il est question dans les sections III-C à III-E du premier article sont répétées des centaines sinon des milliers de fois dans une simulation typique et constituent le facteur limitant au plan de la vitesse d'exécution. Les dérivées des surfaces n'étant pas disponibles, une méthode d'optimisation par recherche plutôt que par gradient est indiquée, comme par exemple la méthode du simplexe avec modifications de Nelder et Mead, disponible dans MATLAB. Cependant, la lenteur constatée de la méthode, conjuguée au fait que les vallées explorées sont presque circulaires et ne sauraient présenter de difficultés, ont contraint l'auteur à développer un algorithme fait sur mesure. La nouvelle méthode, dite *pentagonale*, combine des éléments de la méthode du simplexe et de l'algorithme de Davies, Swann et Campey, et comporte les avantages de ces deux méthodes.

La figure 4.4 présente la description classique de la méthode du simplexe en deux dimensions (Adby et Dempster, 1974) ainsi qu'une description alternative due à l'auteur. Dans la description classique en (a), le triangle équilatéral 1-2-3 subit des réflexions successives sur le côté joignant les deux points les plus bas, jusqu'à ce qu'il oscille autour du point 16. La description alternative en (b) reprend le même exemple en commençant par les points 1 et 2.

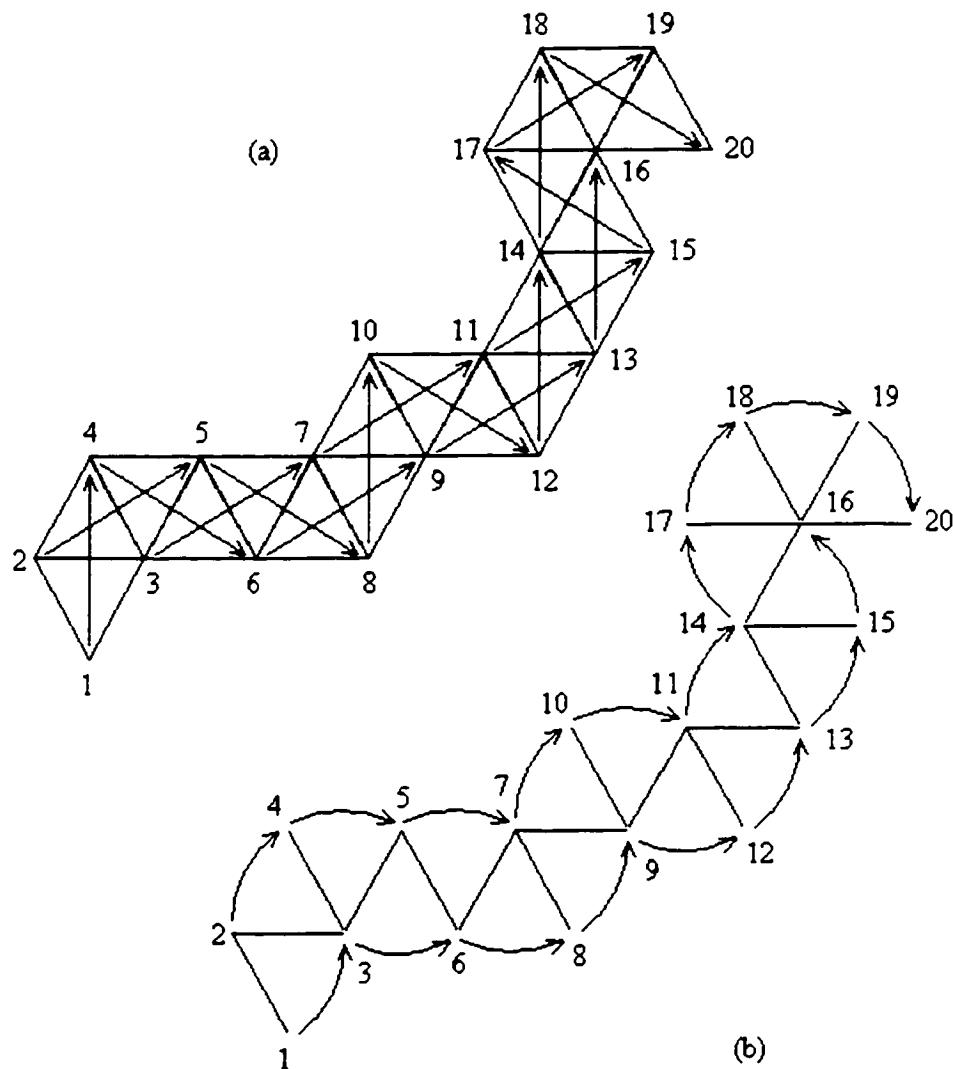


Figure 4.4 Exemple de recherche en deux dimensions par la méthode du simplexe: (a) description classique, (b) description alternative

Dans ce dernier cas. la surface étant plus basse au point 2, on démarre une rotation par sauts de 60 degrés autour de ce point en espérant trouver cinq nouveaux points plus élevés que le point 2 (on a déjà le point 1). Sitôt arrivé au point 3, qui est plus bas que le point 2, on interrompt la rotation et on en démarre une nouvelle en sens inverse autour du point 3, espérant trouver quatre nouveaux points plus élevés que le point 3 (on a déjà les points 1 et 2). Le point 4 passe le test. mais pas le point 5. autour duquel on démarre une nouvelle rotation en inversant encore le sens. espérant trouver quatre nouveaux points plus élevés que le point 5 (on a déjà les points 3 et 4). On continue ainsi jusqu'au point 16 où, finalement. les quatre nouveaux points 17 à 20 qui l'entourent sont tous plus élevés (on avait déjà les points 14 et 15). On peut alors contracter le simplexe et recommencer.

Cette autre façon de comprendre la méthode du simplexe en deux dimensions a permis d'apporter à l'algorithme les améliorations suivantes. Pour commencer, on peut au début de chaque nouvelle rotation doubler le rayon utilisé afin d'accélérer la recherche. comme on le fait en une dimension dans l'algorithme de Davies, Swann et Campey. Cela signifie qu'une nouvelle rotation ne peut se compléter qu'après l'évaluation de six nouveaux points (si on fait des sauts de 60 degrés), les points calculés précédemment ne pouvant pas être réutilisés. On constate alors qu'il n'y a plus d'avantage particulier à faire des rotations en six points, et qu'on pourrait utiliser un nombre inférieur ou supérieur de points. Par exemple. on peut effectuer les rotations par sauts de 72 degrés. de sorte que seuls cinq nouveaux points sont requis pour compléter une rotation. Une rotation complétée (ou oscillation) produit ainsi six points constitués d'un pentagone entourant un point central. Ce complexe de points se prête idéalement à une évaluation quadratique du minimum, comme on le fait en une dimension dans l'algorithme de Davies, Swann et Campey. En supposant la surface quadratique $z = a_1x^2 + a_2y^2 + a_3xy + a_4x + a_5y + a_6$ et en substituant les six points, on obtient les six équations nécessaires au calcul des six coefficients a_1 à a_6 . Les coordonnées du minimum peuvent ensuite être calculées en résolvant le système d'équations

$$\frac{dz}{dx} = 2a_1x + a_3y + a_4 = 0, \quad \frac{dz}{dy} = a_3x + 2a_2y + a_5 = 0. \quad \text{Équation 4.8}$$

Si on s'en tenait aux sauts de 60 degrés et qu'on utilisait les six points de l'hexagone final pour tenter une approximation quadratique, on aboutirait invariablement à un système d'équations surdéterminé. Par exemple, si le point central est à l'origine, le rayon est unitaire, et les six angles progressent de zéro à 300 degrés, la matrice à inverser devient

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 1/4 & 3/4 & \sqrt{3}/4 & 1/2 & \sqrt{3}/2 & 1 \\ 1/4 & 3/4 & -\sqrt{3}/4 & -1/2 & \sqrt{3}/2 & 1 \\ 1 & 0 & 0 & -1 & 0 & 1 \\ 1/4 & 3/4 & \sqrt{3}/4 & -1/2 & -\sqrt{3}/2 & 1 \\ 1/4 & 3/4 & -\sqrt{3}/4 & 1/2 & -\sqrt{3}/2 & 1 \end{bmatrix}. \quad \text{Équation 4.9}$$

mais cette matrice a une déterminante de zéro, ce qui rend l'inversion impossible.

L'étape suivante est de comparer le point obtenu par approximation quadratique avec le point central du pentagone et de démarrer une nouvelle recherche du simplexe à partir du moins élevé des deux points (probablement l'approximation quadratique). Chaque recherche du simplexe, incluant l'originale, doit commencer avec un rayon de rotation égal à la tolérance requise des deux coordonnées. L'optimisation se termine lorsqu'une recherche du simplexe oscille autour de son point de départ, ce qui veut dire que le rayon de rotation n'a jamais doublé mais est demeuré égal à la tolérance des coordonnées. À ce stade il est peu onéreux de faire une dernière approximation quadratique et de choisir le point résultant, s'il est meilleur que le point central du pentagone. Lorsqu'on est situé sur une vallée étroite, il arrive que ce dernier point approximé, tout en étant meilleur que le point central, se retrouve à l'extérieur du pentagone. On doit dans ce cas démarrer une nouvelle recherche du simplexe à partir du point approximé.

La méthode d'optimisation pentagonale cherche donc à accomplir en deux dimensions ce que l'algorithme de Davies, Swann et Campey réussit à faire en une dimension: cerner le minimum (ici avec un pentagone) avant d'en calculer une approximation quadratique. Dans les tests effectués, la méthode s'est avérée robuste, rapide et précise. La figure 4.5 montre les résultats d'une optimisation pentagonale avec une tolérance de 0,085 sur les coordonnées. Une première recherche du simplexe nécessitant 13 évaluations de fonction a été suivie d'une deuxième et d'une dernière recherche de 7 et 6 évaluations respectivement. Sur cette figure, la dernière approximation quadratique se confond avec le point central du dernier pentagone, ce qui suggère une précision du résultat bien supérieure à la tolérance demandée.

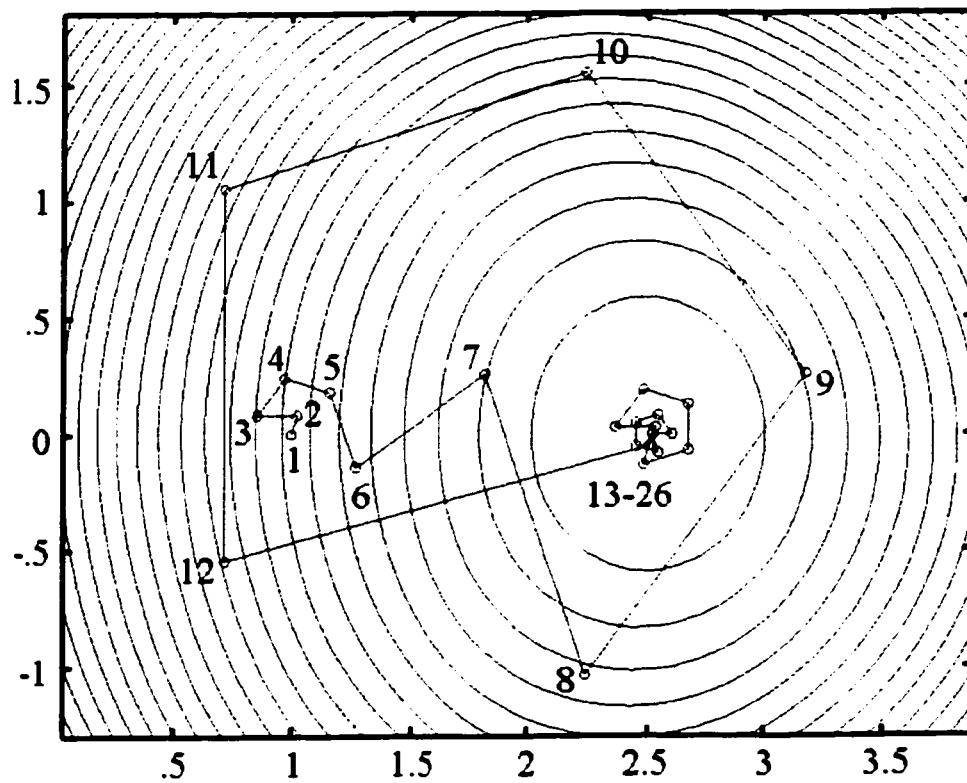


Figure 4.5 Exemple d'optimisation pentagonale avec tolérance de 0.085
sur les coordonnées

4.1.4 Effets de température sur l'adaptation par suppression de signal

La méthode d'analyse de l'amplificateur à post-compensation traitée dans le premier article restreint l'adaptation du circuit aux méthodes sans signal pilote ni filtrage. À ce titre, l'adaptation de la deuxième boucle par suppression non adaptative du signal d'entrée ou par suppression du signal d'erreur a l'avantage de la simplicité par rapport à la méthode adaptative préconisée par Bauman (1983) et Cavers (1995). La suppression non adaptative du signal d'entrée doit cette simplicité au fait qu'elle exploite la stabilité des composantes passives qui fixent le gain et le délai du circuit, et de même, la suppression du signal d'erreur exploite la stabilité des composantes passives de la seconde boucle. Étant donné le traitement exclusif accordé à ces deux méthodes dans le premier article, il importe de prendre au sérieux le brevet de Ghannouchi et autres (2001). Ce brevet, qui introduit la méthode de suppression du signal d'erreur sous le nom de *Complex Gain Stabilization* (CGS), s'inquiète de l'effet des changements de température sur le délai de la seconde boucle, et propose une stratégie de compensation. Le schéma qu'il utilise pour la deuxième boucle ainsi adaptée est reproduit à la figure 4.6, et les délais qu'il suppose à une température de référence sont résumés dans le diagramme de la figure 4.7.

En vertu de la symétrie des coupleurs directifs, les délais de transmission et de couplage dans un coupleur donné sont considérés comme étant égaux. Ainsi, les deux délais respectivement associés aux coupleurs 17 et 29 peuvent être sortis de la boucle, le premier précédant la division des signaux et le second suivant la jonction des signaux. Par rapport aux coupleurs 37 et 38, le brevet est ambigu puisque les symboles D_1 et D_2 semblent admettre deux délais différents alors que les deux délais de couplage sont traités comme s'ils étaient équivalents. Par souci de cohérence, la figure 4.7 traite ces deux coupleurs comme elle a traité le coupleur 17, ce qui a l'effet de faire sortir D_1 de la boucle CGS plutôt que D_2 , comme dans le brevet. Cette différence ne change évidemment en rien la substance de l'analyse.

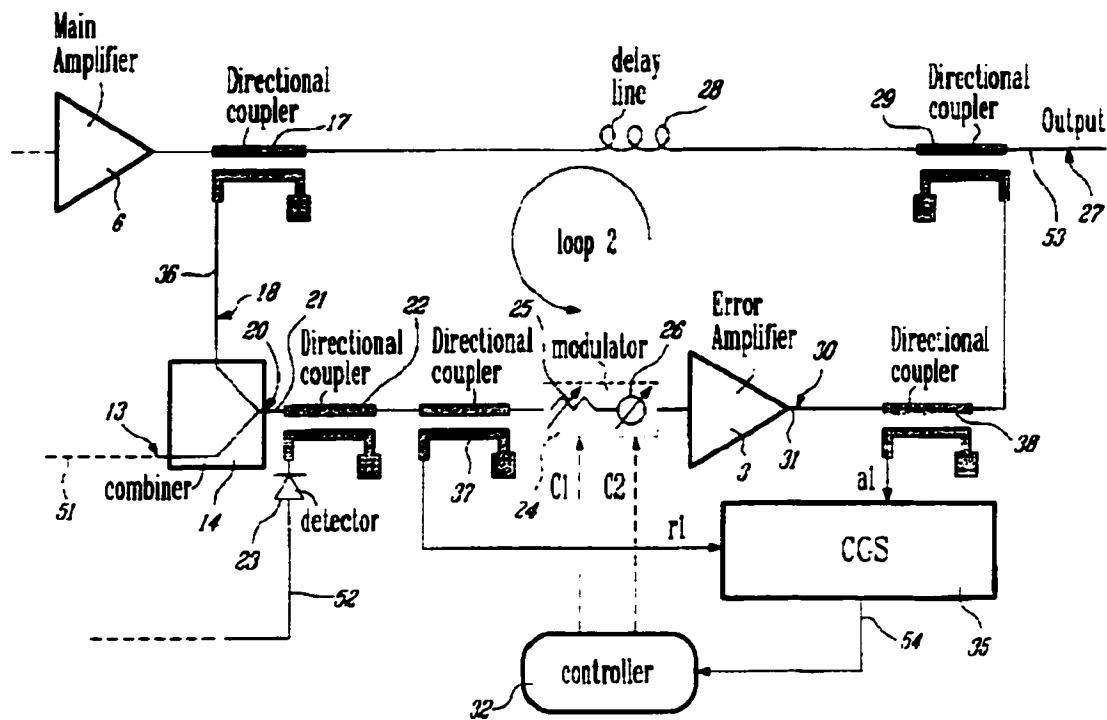


Figure 4.6 Deuxième boucle de l'amplificateur à post-compensation
adaptée par la méthode CGS (Ghannouchi et autres. 2001)

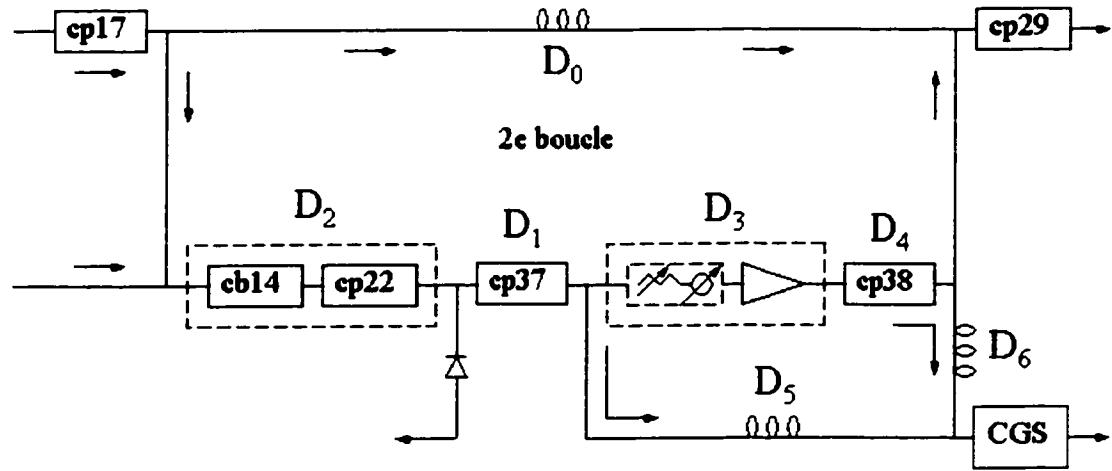


Figure 4.7 Diagramme des délais pour le circuit de la figure 4.6 à une température de référence

On a négligé les délais de tous les branchements entre composantes sauf ceux dénotés a_1 et r_1 à la figure 4.6, lesquels nécessitent les lignes à délai D_5 et D_6 . En effet, la boucle CGS minimise la différence entre deux signaux pour imposer à la branche supérieure de la boucle CGS la fonction de transfert qui équilibre la deuxième boucle. Au point de vue des délais, cette condition est assurée lorsque

$$D_5 - D_6 = D_0 - (D_2 + D_1). \quad \text{Équation 4.10}$$

Il est entendu que le modulateur vectoriel ne peut pas se faire imposer un délai différent de son délai propre, mais il peut quand même répondre à une « suggestion » en ajustant sa phase de telle sorte qu'un écart de délais dans la boucle CGS soit compensé le mieux possible sur la totalité de la bande passante.

La ligne à délai D_0 pourra en pratique subir des changements de température par rapport à la température de référence, bien que ces changements ne soient pas de l'ordre des variations environnementales. En particulier, une augmentation de température rallongera quelque peu la ligne à délai et augmentera son délai. Il peut être rassurant de penser que le même réchauffement s'étendra à l'ensemble du circuit pour augmenter tous les délais dans la même proportion. Le brevet suppose effectivement des changements de température uniformes, mais il suppose aussi que les délais du combinateur et des coupleurs sont relativement insensibles à ces changements. Dans la mesure où ces suppositions correspondent à la réalité, la boucle CGS devra hausser le délai suggéré à

$$D_0(1+\varepsilon) - (D_2 + D_1) = D_0 - (D_2 + D_1) + D_0\varepsilon. \quad \text{Équation 4.11}$$

mais en réalité, d'après l'équation 4.10, elle le haussera à

$$(D_5 - D_6)(1+\varepsilon) = D_0 - (D_2 + D_1) + [D_0 - (D_2 + D_1)]\varepsilon. \quad \text{Équation 4.12}$$

Il ne manquera donc à l'augmentation du délai suggéré qu'une petite fraction égale à $(D_2 + D_1) / D_0$. La stratégie du brevet pour compenser ce problème était d'augmenter le délai D_5 pour le rendre égal à D_0 , ce qui demande $D_6 = D_2 + D_1$, mais ayant constaté que D_5 n'est pas relié au problème, on voudra simplifier en éliminant le délai D_6 et en utilisant $D_5 = D_0 - (D_2 + D_1)$. Une stratégie qui permet à tout le moins d'atténuer le problème est de réduire $D_2 + D_1$ en transférant le coupleur 22 à l'intérieur de la boucle CGS, ce qui n'affecte pas le coupleur dans sa fonction de détection de la puissance relative du signal d'erreur. Ce changement, illustré à la figure 4.8, a pour effet de retrancher la contribution du coupleur 22 au délai D_2 .

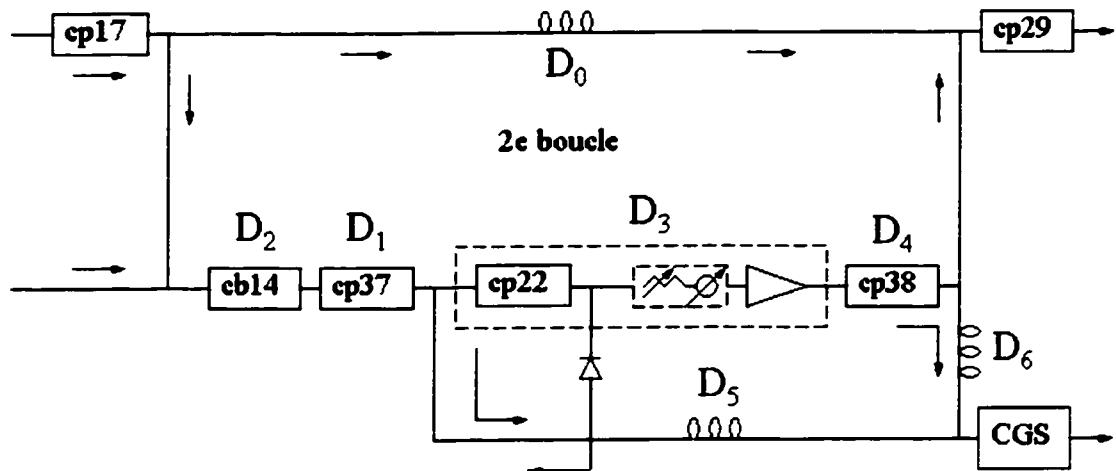


Figure 4.8 Diagramme modifié des délais pour le circuit de la figure 4.6 à une température de référence

4.1.5 Effets d'une suppression imparfaite de signal

La section III-D du premier article tentait d'expliquer en peu de mots les conséquences d'une suppression imparfaite du signal d'entrée ou d'erreur sur l'adaptation de la seconde boucle et sur le rapport intermodulation à signal à la sortie du circuit (IMSR). Il aurait

été plus utile de reprendre le type d'analyse employé par Cavers (1995) et d'utiliser les résultats pour vérifier certaines valeurs simulées d'IMSR apparaissant au tableau 1 de l'article, mais la contrainte d'espace a repoussé cet exercice aux paragraphes qui suivent.

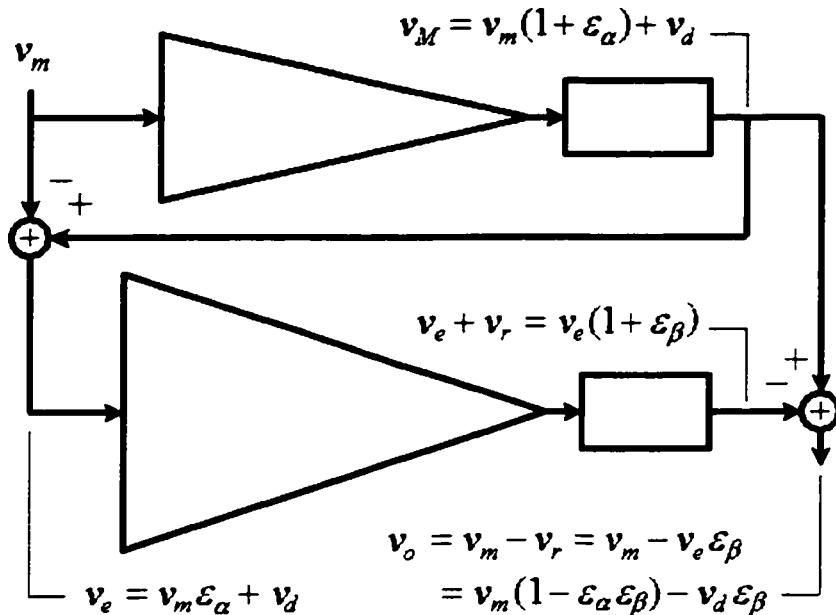


Figure 4.9 Signaux du modèle de l'amplificateur à post-compensation dans une analyse inspirée de Cavers (1995)

La figure 2(c) du premier article, reproduite ici à la figure 4.9, montre les différents signaux du circuit lorsque le signal d'entrée amplifié est exprimé par $v_m(1 + \varepsilon_\alpha) + v_d$ où la distorsion v_d est sans corrélation avec v_m , ou $X_{dm} = E[v_d v_m^*] = 0$. L'amplificateur d'erreur, supposé parfaitement linéaire, produit le signal résiduel $v_r = v_e \varepsilon_\beta$ où ε_β est l'erreur d'ajustement du second modulateur vectoriel. Le signal d'entrée amplifié et le signal de sortie ont respectivement les puissances (normalisées) et les IMSR suivants:

$$E[|v_M|^2] = |1 + \varepsilon_\alpha|^2 + X_d \Rightarrow IMSR_M = X_d / |1 + \varepsilon_\alpha|^2, \quad \text{Équation 4.13}$$

$$E[|v_o|^2] = |1 - \varepsilon_\alpha \varepsilon_\beta|^2 + X_d |\varepsilon_\beta|^2 \Rightarrow IMSR = X_d |\varepsilon_\beta|^2 / |1 - \varepsilon_\alpha \varepsilon_\beta|^2. \quad \text{Équation 4.14}$$

En combinant ces deux équations, on relie les deux IMSR par l'équation

$$IMSR = \frac{|1+\varepsilon_\alpha|^2}{|1-\varepsilon_\alpha\varepsilon_\beta|^2} |\varepsilon_\beta|^2 IMSR_M. \quad \text{Équation 4.15}$$

Avec une erreur ε dans la suppression du signal *d'erreur*, les expressions en (26) et (27) du premier article, respectivement minimisées et annulées par ε_β dans les méthodes de puissance et du gradient, sont maintenant $E[|v_e\varepsilon_\beta - v_e\varepsilon|^2]$ et $E[(v_e\varepsilon_\beta - v_e\varepsilon)v_e^*]$. Dans les deux cas on obtient la solution $\varepsilon_\beta = \varepsilon$ et le résultat $v_r = v_e\varepsilon$ qui avait été anticipé dans l'article. Le point important à retenir est que la convergence de ε_β n'est pas couplée à celle de ε_α . En substituant la solution dans l'équation 4.15 et en tenant compte du fait que $IMSR_M$ est une fonction de ε_α , on obtient l'IMSR de sortie et son approximation pour un petit ε_α .

$$IMSR = \frac{|1+\varepsilon_\alpha|^2}{|1-\varepsilon_\alpha\varepsilon|^2} |\varepsilon|^2 IMSR_M(\varepsilon_\alpha) \approx |\varepsilon|^2 IMSR_M(0). \quad \text{Équation 4.16}$$

Avec une erreur ε dans la suppression du signal *d'entrée*, les expressions en (29) et (30) du premier article, respectivement minimisées et annulées, sont maintenant $E[|v_e\varepsilon_\beta + v_m\varepsilon|^2]$ et $E[(v_e\varepsilon_\beta + v_m\varepsilon)v_e^*]$. Encore une fois, la théorie d'estimation linéaire (Haykin, 1991) garantit deux solutions identiques pour ε_β . En annulant la deuxième expression, on obtient facilement $\varepsilon_\beta = -X_{em}^*\varepsilon / X_e$ où $X_{em} = E[(v_m\varepsilon_\alpha + v_d)\varepsilon_m^*] = \varepsilon_\alpha$ et $X_e = |\varepsilon_\alpha|^2 + X_d$, donc $\varepsilon_\beta = -\varepsilon_\alpha^*\varepsilon / (|\varepsilon_\alpha|^2 + X_d)$. Cette fois, la convergence de ε_β est couplée à celle de ε_α . En substituant la solution de ε_β et ensuite l'équation 4.13 dans l'équation 4.15, on obtient l'IMSR de sortie et son approximation pour petits ε_α et ε .

$$IMSR = \frac{|\varepsilon|^2 |\varepsilon_a|^2 |1 + \varepsilon_a|^2 IMSR_M(\varepsilon_a)}{|(1 + \varepsilon)| |\varepsilon_a|^2 + |1 + \varepsilon_a|^2 IMSR_M(\varepsilon_a)|^2} \approx \frac{|\varepsilon_a|^2}{(|\varepsilon_a|^2 + IMSR_M(0))^2} |\varepsilon|^2 IMSR_M(0).$$

Équation 4.17

En comparaison avec la suppression du signal d'erreur à l'équation 4.16, on constate que le couplage des convergences de ε_β et ε_a associé à la suppression du signal d'entrée introduit dans l'IMSR approximatif le facteur additionnel $|\varepsilon_a|^2/(|\varepsilon_a|^2 + IMSR_M(0))^2$. Lorsque $|\varepsilon_a| < IMSR_M(0)$, ce facteur additionnel est inférieur à l'unité et donne l'avantage à la suppression du signal d'entrée. Dans le cas contraire, et particulièrement lorsque la première boucle est adaptée par la méthode optimale *de droite* présentée dans le deuxième article, le facteur additionnel dépasse l'unité et la suppression du signal d'erreur est indiquée. Ce point a déjà été démontré au moyen de simulations dans un article de Larose et Ghannouchi (août 2001).

Le tableau 1 du premier article donne des valeurs simulées de SIMR (la réciproque de l'IMSR) lorsque l'adaptation de la seconde boucle utilise la suppression du signal d'entrée. Le cas $\varepsilon = 0$ devrait donner en théorie des SIMR infinis mais la non-linéarité de l'amplificateur d'erreur a fait plafonner les valeurs à environ 60 dB. Le cas $\varepsilon = -1$, qui correspond à une absence totale de suppression du signal d'entrée, viole la supposition d'un petit ε , et devrait donner selon la partie exacte de l'équation 4.17

$$SIMR = \frac{|1 + \varepsilon_a|^2}{|\varepsilon_a|^2 SIMR_M(\varepsilon_a)}.$$

Équation 4.18

Dans ce cas, l'adaptation de la première boucle par la méthode du gradient donne en théorie $\varepsilon_a = 0$ et $SIMR \rightarrow \infty$, ce qui fait encore plafonner les valeurs à environ 60 dB. L'adaptation de la première boucle par la méthode de puissance, par contre, a produit $\varepsilon_a \approx 0.039 \angle -168^\circ$ et $SIMR_M(\varepsilon_a) = 19,79$ dB, de sorte qu'on prévoit $SIMR \approx 8.05$ dB.

Les valeurs simulées de 8.9115 dB et 7.7773 dB, obtenues respectivement en adaptant la seconde boucle par les méthodes du gradient et de puissance, sont proches du SIMR prévu, et diffèrent entre elles à cause de la non-linéarité de l'amplificateur d'erreur.

4.1.6 Vérification des paramètres optimaux triviaux

Dans la section III-F du premier article, on a déterminé que pour un couple (D_{0M}, D_{0E}) donné et les valeurs de H_M , H_E , S , T , V et W qui lui sont associées, les gains de puissance A à R (excluant les gains d'amplificateur C et I qui ont été éliminés) et les constantes a et b du modèle de l'amplificateur à post-compensation de la figure 4 doivent maximiser le rendement

$$\eta = \frac{1}{X_{satM} H_M + X_{satE} H_E} \quad \text{Équation 4.19}$$

où les puissances de sortie de saturation normalisées X_{satM} et X_{satE} sont données par

$$X_{satM} = \frac{1}{D_{0M} D} \left[\left(\frac{T}{EFGHJK} \right)^b \frac{1}{(P-E)Q(R-K)} \right]^{\frac{1}{1+b}} \frac{1}{V}. \quad \text{Équation 4.20}$$

$$X_{satE} = \frac{1}{D_{0E} JK V}. \quad \text{Équation 4.21}$$

sujet à la contrainte de l'équation (61).

$$W = \left[\left(\frac{ABD}{S} \right)^a (L-A) M (N-G) \right]^{\frac{1}{1+a}} \left(\frac{1}{EFG} \right)^{\frac{1-ab}{(1+a)(1+b)}} \left[\left(\frac{HJK}{T} \right)^b (P-E) Q (R-K) \right]^{\frac{1}{1+b}}. \quad \text{Équation 4.22}$$

On s'est penché uniquement sur les solutions des couplages A , E , G et K parce qu'il s'agit des seules solutions non triviales, mais la vérification des solutions triviales apportée ci-dessous confirmera le bien-fondé de l'approche.

En prenant pour point de départ toute solution qui satisfait la contrainte de l'équation 4.22, on peut prouver qu'il est avantageux, si possible, d'augmenter B , D , H , J , L , M , N , P , Q , ou R . Ces changements, qui laissent inchangées ou diminuent les puissances de saturation X_{satA} et X_{satE} et par conséquent laissent inchangé ou augmentent η , augmentent en même temps le côté droit de l'équation 4.22. L'égalité brisée dans cette contrainte peut alors être rétablie en augmentant G aussi près de N qu'il est nécessaire, ce qui diminue X_{satA} encore davantage pour confirmer le gain en rendement. De même, il est avantageux d'augmenter F . L'égalité brisée dans la contrainte doit cette fois être rétablie en diminuant G , mais ayant ainsi augmenté le terme $N - G$ au numérateur de la contrainte, on a forcément augmenté la combinaison FG au dénominateur, de sorte que X_{satA} se trouve une fois de plus diminué. Ces raisonnements confirment le fait évident que B , D , F , H , J , L , M , N , P , Q , et R doivent atteindre leurs plus hautes valeurs réalisables si on veut minimiser les pertes dans le circuit.

On peut aussi prouver l'avantage de diminuer la constante $a = p_1 / q_1$ en décomposant cette opération en une première étape où on diminue simultanément a à a' et M à M' de telle sorte que $C^{-a'}M' = C^{-a}M$, et une deuxième étape où on retourne M' à M . La première étape modifie la solution de départ sans affecter le rendement du circuit, et la deuxième augmente ce rendement de la façon décrite plus haut. Le même raisonnement s'appliquant à la constante $b = p_2 / q_2$, il est évident que les rapports délai-gain des amplificateurs, p_1 et p_2 [ns/dB], et les qualités des lignes à délai, q_1 et q_2 [ns/dB], doivent atteindre leurs valeurs réalisables les plus basses et les plus hautes, respectivement.

4.1.7 Calcul du couplage principal

La section III-F enchaîne avec le système d'équations en E , G et K ci-dessous,

$$cEG = N - G, \quad dEGK = (P - E)(R - K), \quad eEGK = 1. \quad \text{Équation 4.23}$$

et la solution quadratique recherchée du couplage principal,

$$E = \frac{-a_1 - \sqrt{a_1^2 - 4a_2a_0}}{2a_2}. \quad a_2 = eNR - c, \quad a_1 = dN - a_2P - 1, \quad a_0 = P. \quad \text{Équation 4.24}$$

On n'a pas expliqué la raison pratique pour laquelle on calcule E avant G et K , et on n'a pas prouvé que la solution quadratique de E obtenue avec le signe positif ne peut pas être la solution recherchée. Pour le premier point, étant donné que c , d et e sont positifs selon l'équation (66) du premier article, un examen systématique du système d'équations dégage les cinq possibilités mathématiques énumérées au tableau 4.1. Les couplages réels situés entre zéro et les gains d'insertion sont réalisables (✓), mais non les autres (✗).

Tableau 4.1 Possibilités mathématiques sur (E, G, K)

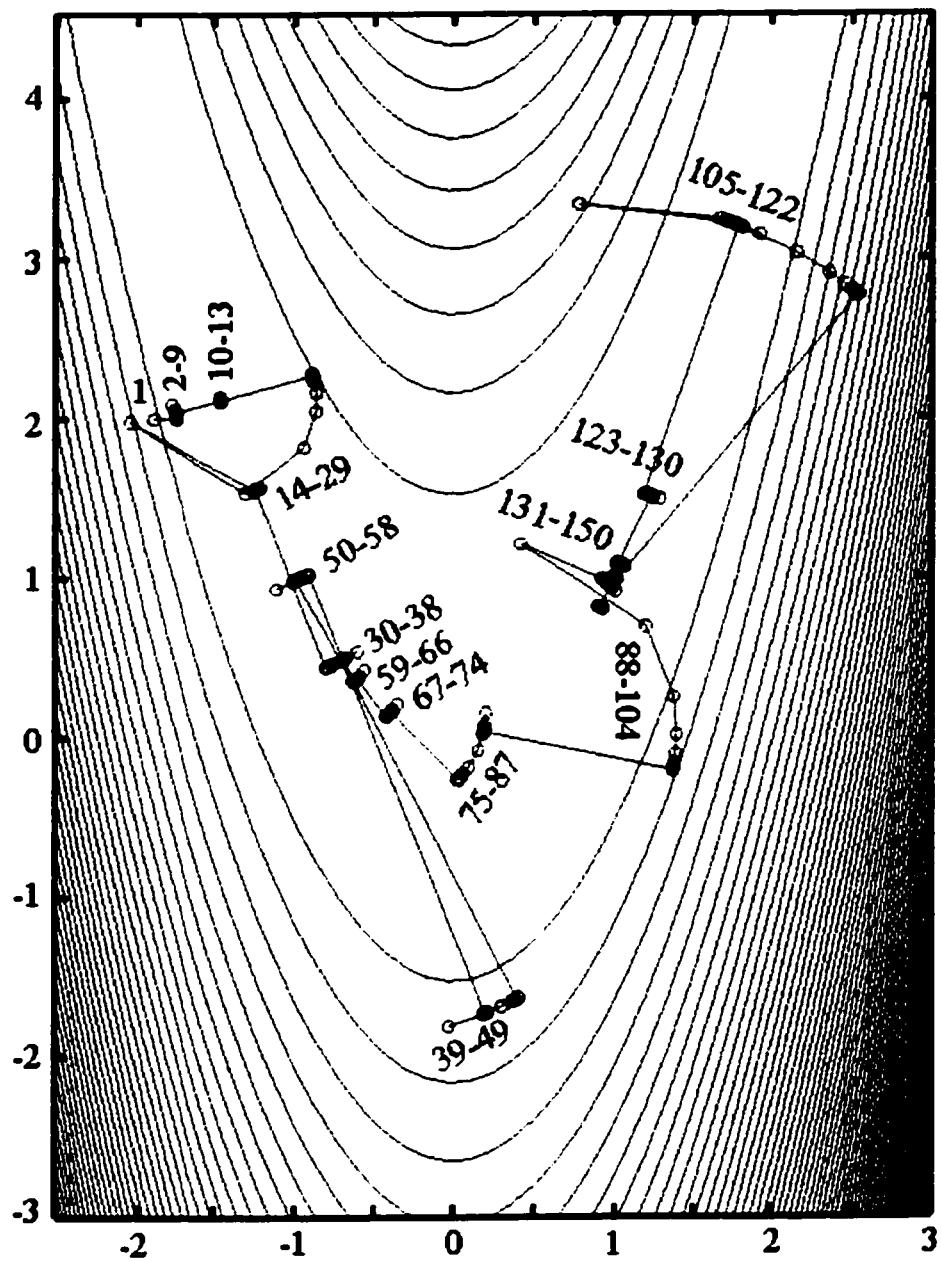
Possibilité	E	G	K
1 (✓)	$0 < E < P$ (✓)	$0 < G < N$ (✓)	$0 < K < R$ (✓)
2 (✗)	$E < 0$ (✗)	$G < 0$ (✗)	$0 < K < R$ (✓)
3 (✗)	$E < 0$ (✗)	$G > N$ (✗)	$K < 0$ (✗)
4 (✗)	$E > P$ (✗)	$0 < G < N$ (✓)	$K > R$ (✗)
5 (✗)	E complexe (✗)	G complexe (✗)	K complexe (✗)

On constate que dans toute solution (E, G, K) non réalisable, E est non réalisable, mais pas nécessairement G ou K . Par conséquent, évaluer E en premier permet d'anticiper toute solution (E, G, K) non réalisable avant d'en calculer les deux autres couplages.

Pour le deuxième point, supposons qu'on ait obtenu au moyen du signe *positif* une première solution quadratique de E réalisable, donc réelle, positive et inférieure à P . Selon le tableau 4.1, la solution (E,G,K) correspondante est réalisable. Dans l'expression $a_2 = [1 + N(R - K)/GK]/E$, obtenue à partir des équations 4.23 et 4.24, tous les termes sont positifs et par conséquent a_2 lui-même est positif, tout comme $a_0 = P$. Le discriminant positif $a_1^2 - 4a_2a_0$ est donc inférieur à a_1^2 , de sorte que E a deux solutions du même signe. La deuxième solution obtenue au moyen du signe négatif est donc positive comme la première, et puisque plus petite, également réalisable. Elle est aussi préférable à la première puisque dans l'expression $K = (c + 1/E)/eN$, obtenue en éliminant G de l'équation 4.23, un E plus petit donne un K plus grand, donc une puissance de saturation X_{sat} plus petite à l'équation (69) du premier article et un rendement énergétique accru. On en conclut effectivement qu'il n'est jamais nécessaire de vérifier la solution E obtenue par le signe positif, puisque même si cette solution s'avérait réalisable, la solution alternative lui serait supérieure.

4.1.8 Méthode d'optimisation polaire

La *minimisation de puissance de crête*, étudiée dans le deuxième article comme méthode d'adaptation pour la première boucle de l'amplificateur à post-compensation, demande de minimiser une surface dont le gradient est discontinu. Ni la méthode d'optimisation pentagonale (section 4.1.3), ni la méthode du simplexe de MATLAB ne réussissent à localiser le minimum situé au fond d'une vallée en V. En raison du succès de l'initiative d'optimisation pentagonale, un autre algorithme fait sur mesure a été conçu par l'auteur pour traiter les cas problématiques. Dans le contexte du premier article, la méthode s'est avérée essentielle pour la maximisation en deux dimensions expliquée à la section III-F et illustrée à la figure 9, ainsi que pour la localisation automatique des points de solution illustrés à la figure 10 (section 4.1.9). Dans les deux cas, en effet, les vallées ou les sommets explorés peuvent être très étroits.



**Figure 4.10 Minimisation de la fonction banane (fonction de Rosenbrock)
par la méthode polaire avec tolérance de 0,15**

La nouvelle méthode de recherche, dite *polaire*, adapte l'algorithme de Davies, Swann et Campey au cas bi-dimensionnel en orientant chacun des déplacements de façon à

minimiser (ou maximiser) la surface. Cette orientation, *qui n'est généralement pas celle de plus grande pente*, est déterminée par un deuxième niveau de recherche qui utilise à son tour l'algorithme de Davies, Swann et Campey. La figure 4.10 montre comment la méthode s'est accommodée de la fonction banane (aussi appelée fonction de Rosenbrock), qui est notoire par la difficulté qu'elle présente à la plupart des méthodes¹.

Avec une tolérance de 0.15 dans les coordonnées, le minimum [1 1] est atteint à partir du point [-1.9 2] après 150 itérations. L'orientation du premier déplacement de longueur 0.15 est optimisée par une recherche à huit itérations (points 2 à 9) débutant à zéro degrés. La longueur du prochain déplacement est doublée à 0.3, et son orientation est optimisée par une seconde recherche débutant à l'orientation optimale précédente (points 10 à 13). L'orientation change abruptement au prochain déplacement pour ensuite se maintenir jusqu'au groupe des points 39 à 49. La fonction ayant remonté à cet endroit, le prochain déplacement est négatif et d'une longueur obtenue par estimation quadratique à partir des hauteurs de points optimales des groupes 14-29, 30-38 et 39-49. Puisque la hauteur optimale du nouveau groupe 50-58 n'est pas inférieure à celle du groupe central 30-38, la recherche recommence à partir du point optimal du groupe 30-38 avec un déplacement de longueur 0.15. Un scénario semblable fait recommencer la recherche à partir du point optimal du groupe 88-104, après quoi 20 autres itérations suffisent.

4.1.9 Localisation des points de solution

Les valeurs du rendement énergétique η et de la linéarité de sortie $SIMR$, calculées sur une grille de points $(D_{0,M}, D_{0,E})$, permettent de tracer les contours approximatifs de ces deux fonctions sur un même plan et de repérer visuellement, pour chacun des contours de linéarité, le point de solution qui maximise le rendement. Pour des raisons de convenance et de précision, cette procédure de solution graphique a été automatisée de la façon suivante.

¹ On peut obtenir des détails sur cette fonction par la commande « optdemo » de MATLAB.

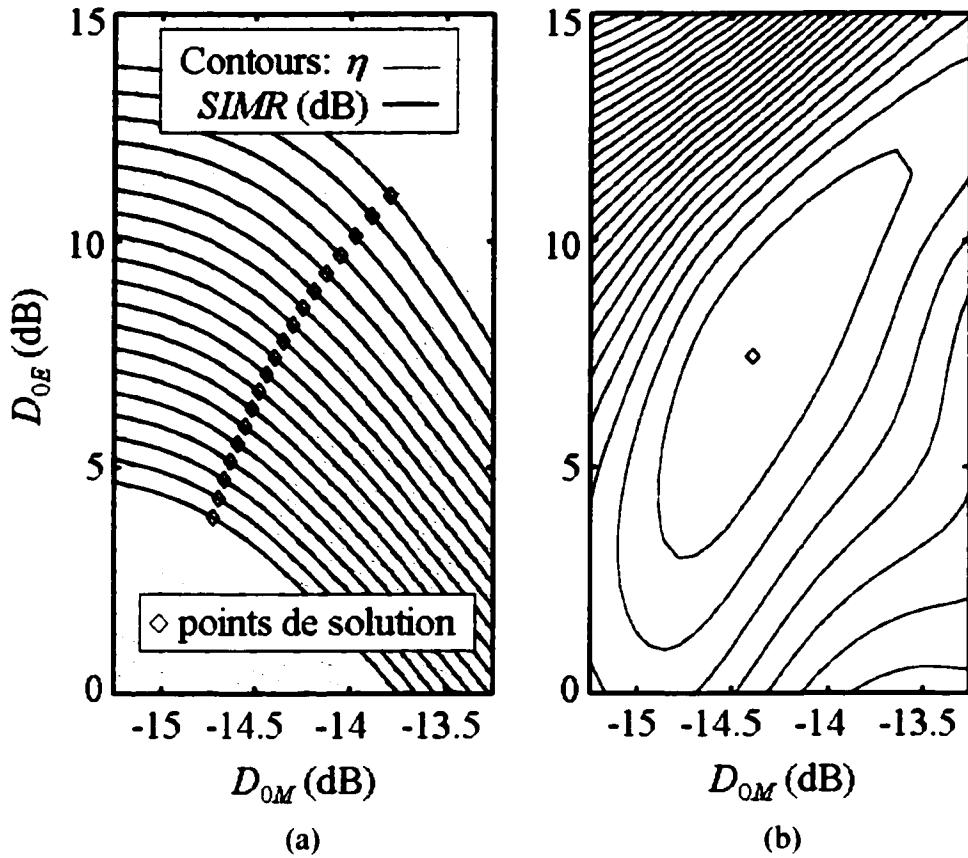


Figure 4.11 (a) Contours de rendement énergétique (η) et de linéarité de sortie ($SIMR$). (b) Fonction d'erreur pour $SIMR = 60$ dB.

L'exemple de contours donné à la figure 10(a) du premier article est reproduit ici à la figure 4.11(a). Chaque point de solution se situe à l'endroit où le contour du SIMR pertinent devient parallèle aux contours de rendement. Il est donc possible de localiser un point de solution au moyen d'une recherche du minimum de zéro de la fonction d'erreur

$$E = (SIMR - SIMR_0)^2 + m \sin^2 \theta \quad \text{Équation 4.25}$$

où $SIMR$ et $SIMR_0$ sont les SIMR réel et désiré en dB, θ est l'angle entre les gradients des deux surfaces, et m est une constante positive arbitraire. Avec $SIMR_0 = 60$ dB et

$m = 5000$, les contours de la fonction d'erreur pour la figure 4.11(a) sont tels que rapportés à la figure 4.11(b). Une valeur aussi élevée de m étire les contours dans le sens où se succèdent les points de solution, de sorte qu'en localisant les points en séquence, on peut exploiter l'habileté de la méthode de recherche polaire (section 4.1.8) à suivre les vallées étroites.

L'interpolation bic cubique du SIMR à partir des valeurs de la grille se fait aisément à l'aide de la fonction « `interp2` » de MATLAB, mais l'évaluation des deux gradients a demandé l'écriture d'un nouveau programme basé sur une compréhension intime de l'interpolation bic cubique². Celle-ci opère à partir des seize points entourant le point demandé, tel qu'illustré à la figure 4.12 (on fait des extrapolations quadratiques au besoin si le point demandé est près des bords de la grille). Si on suppose une fonction z définie sur la grille, et les paramètres s et t définissant la position du point demandé, une première interpolation dans le sens vertical donne le point

$$\begin{aligned} z_1 &= [t(t(2-t)-1)z_{11} + (t^2(3t-5)+2)z_{12} + t(t(4-3t)+1)z_{13} + t^2(t-1)z_{14}] / 2 \\ &= [t^3 \quad t^2 \quad t \quad 1] \begin{bmatrix} -.5 & 1.5 & -1.5 & .5 \\ 1 & -2.5 & 2 & -.5 \\ -.5 & 0 & .5 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} z_{11} \\ z_{12} \\ z_{13} \\ z_{14} \end{bmatrix} = \mathbf{tC} \begin{bmatrix} z_{11} \\ z_{12} \\ z_{13} \\ z_{14} \end{bmatrix}. \end{aligned} \quad \text{Équation 4.26}$$

Par extension, les quatre valeurs interpolées dans le sens vertical sont données par

$$[z_1 \quad z_2 \quad z_3 \quad z_4] = \mathbf{tC} \begin{bmatrix} z_{11} & z_{21} & z_{31} & z_{41} \\ z_{12} & z_{22} & z_{32} & z_{42} \\ z_{13} & z_{23} & z_{33} & z_{43} \\ z_{14} & z_{24} & z_{34} & z_{44} \end{bmatrix} = \mathbf{tCZ}^T. \quad \text{Équation 4.27}$$

² On peut examiner la fonction « `cubic` » de `interp2` avec la commande « `edit interp2` ». On y trouve une référence (Keys, 1981).

et en utilisant ce résultat, la valeur finale interpolée dans le sens horizontal devient

$$z = [s^3 \ s^2 \ s \ 1] \mathbf{C} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = \mathbf{s} \mathbf{C} (\mathbf{t} \mathbf{C} \mathbf{Z}^T)^T = \mathbf{s} \mathbf{C} \mathbf{Z} \mathbf{C}^T \mathbf{t}^T. \quad \text{Équation 4.28}$$

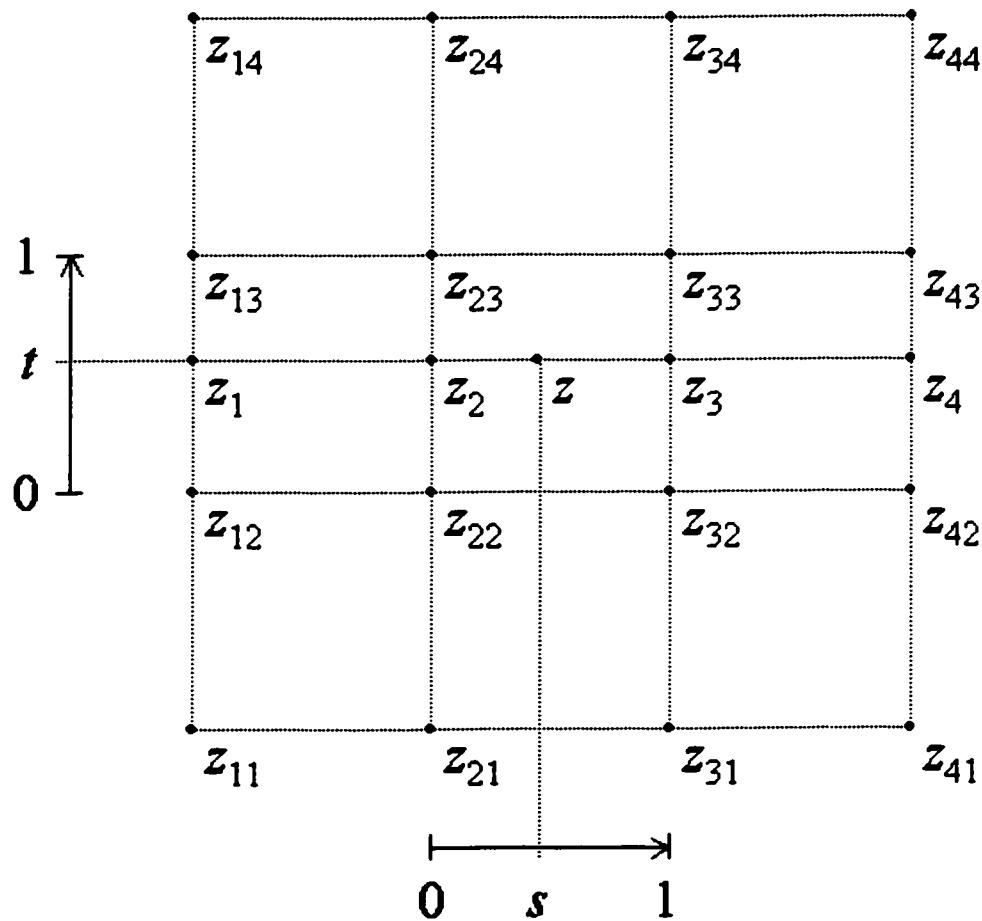


Figure 4.12 Données permettant l'interpolation bic cubique de la fonction z au point (s,t)

En remplaçant maintenant z par $SIMR$ et en définissant $\mathbf{C} \cdot \mathbf{SIMR} \cdot \mathbf{C}^T = \mathbf{A}$, la valeur interpolée est $SIMR = \mathbf{sAt}^T$. De même, on peut remplacer z par η , définir $\mathbf{C}\eta\mathbf{C}^T = \mathbf{B}$, et interpoler $\eta = \mathbf{sBt}^T$, mais cette dernière fonction ne nous concerne que par son gradient. Le sinus de l'angle entre les gradients est donné par la formule élémentaire

$$\sin \theta = \frac{|\nabla SIMR \times \nabla \eta|}{|\nabla SIMR| |\nabla \eta|} \quad \text{Équation 4.29}$$

où les gradients sont donnés par

$$\nabla SIMR = [\partial SIMR / \partial s \quad \partial SIMR / \partial t] = [\mathbf{s}' \mathbf{At}^T \quad \mathbf{sAt'}^T] = [u_1 \quad u_2]. \quad \text{Équation 4.30}$$

$$\nabla \eta = [\partial \eta / \partial s \quad \partial \eta / \partial t] = [\mathbf{s}' \mathbf{Bt}^T \quad \mathbf{sBt'}^T] = [v_1 \quad v_2]. \quad \text{Équation 4.31}$$

et les dérivées sont $\mathbf{s}' = [3s^2 \quad 2s \quad 1 \quad 0]$ et $\mathbf{t}' = [3t^2 \quad 2t \quad 1 \quad 0]$. En termes des composantes vectorielles, le carré du sinus s'exprime par

$$\sin^2 \theta = \frac{(u_1 v_2 - u_2 v_1)^2}{(u_1^2 + u_2^2)(v_1^2 + v_2^2)}. \quad \text{Équation 4.32}$$

En général les quadrilatères de la figure 4.12 seront rectangulaires plutôt que carrés, de sorte que l'angle calculé de chaque gradient sera une transformation de l'angle réel, mais cela ne changera pas les points de solution où les angles des deux gradients, réels ou calculés, seront identiques.

4.1.10 Spectre de densité de puissance

Bien que la méthode d'optimisation conçue dans cet ouvrage utilise la distribution statistique de l'enveloppe de puissance des signaux, l'enveloppe échantillonnée des

signaux a été utilisée au besoin pour produire des spectres de densité de puissance semblables à celui de la figure 7(b). Il a fallu auparavant vérifier comme suit le rapport précis entre la transformée de Fourier de l'enveloppe complexe $v(t)$ et celle du signal réel $\tilde{v}(t)$. La relation entre les deux signaux étant

$$\tilde{v}(t) = \operatorname{Re}\{v(t)e^{j\omega_c t}\} = \frac{v(t)e^{j\omega_c t} + [v(t)e^{j\omega_c t}]^*}{2}. \quad \text{Équation 4.33}$$

on peut utiliser les propriétés $\Im\{v(t)e^{j\omega_c t}\} = V(\omega - \omega_c)$ et $\Im\{v^*(t)\} = V^*(-\omega)$ pour obtenir la relation entre les deux transformées de Fourier.

$$\tilde{V}(\omega) = \frac{1}{2}V(\omega - \omega_c) + \frac{1}{2}V^*(-\omega - \omega_c). \quad \text{Équation 4.34}$$

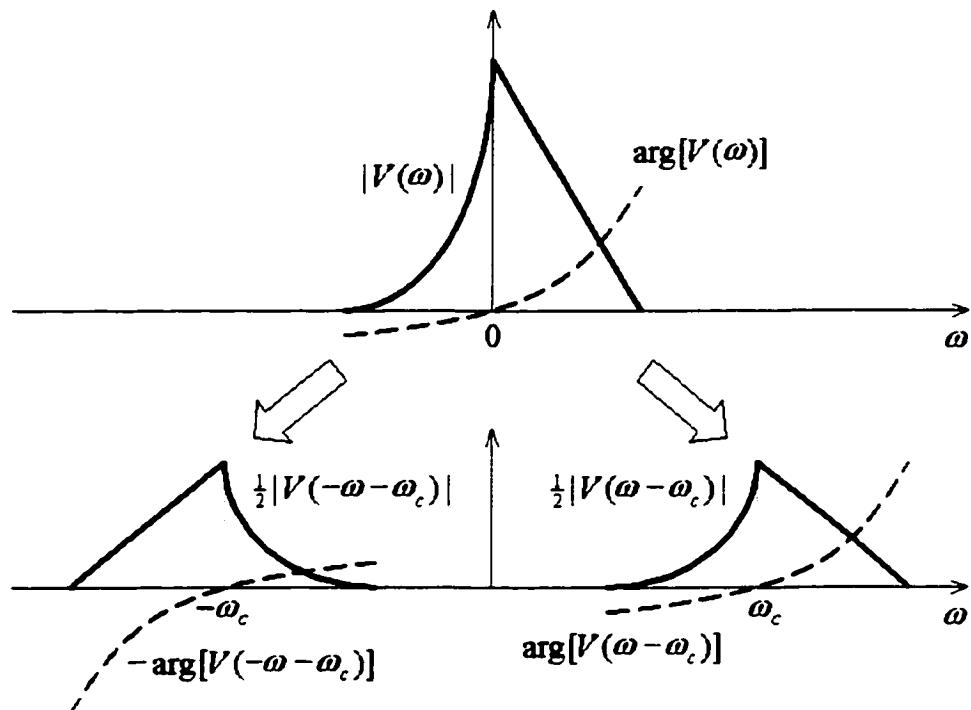


Figure 4.13 Relation entre la transformée de Fourier de l'enveloppe complexe d'un signal et la transformée de Fourier du signal réel

Cette relation est illustrée dans le diagramme de la figure 4.13. En résumé, la transformée de Fourier de l'enveloppe complexe d'un signal est la moitié de droite de la transformée de Fourier du signal réel, doublée en amplitude et abaissée à la fréquence zéro.

4.2 Deuxième article (annexe II)

Cette section accompagne l'article en annexe II. « Optimal amplifier classes and adaptation methods for feedforward amplifiers driven by high-stress signal ». soumis le 10 mai 2002 au *IEEE Transactions on Vehicular Technologies*.

4.2.1 Schéma des stratégies d'adaptation de la première boucle

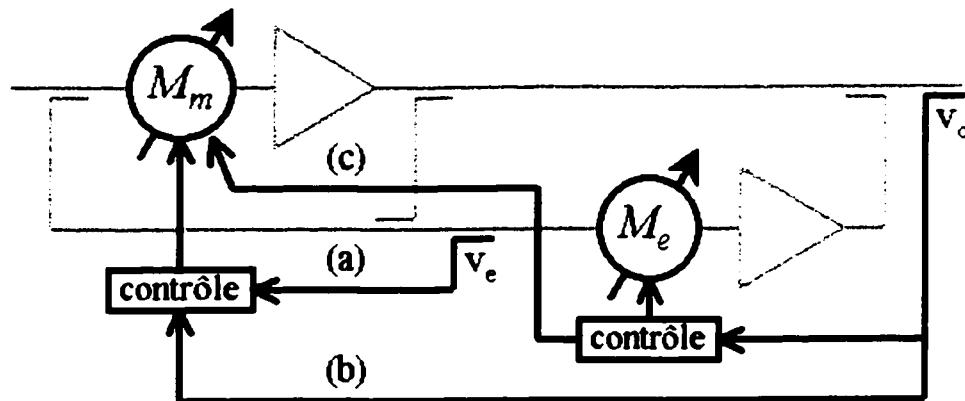


Figure 4.14 Adaptation de la première boucle: (a) conventionnelle. (b) optimale avec contrôles séparés, (c) optimale avec contrôles intégrés

La figure 4.14, qui n'apparaît nulle part dans le deuxième article, donne un bref aperçu des différences fondamentales entre les méthodes d'adaptation traditionnelles de l'amplificateur à post-compensation et les méthodes optimales de gauche ou de droite. Dans tous les cas, l'adaptation de la deuxième boucle utilise le critère décisif qui est de minimiser la distorsion à l'intérieur du signal de sortie du circuit. Ceci est symbolisé à la figure 4.14 par un bloc de contrôle adaptatif qui reçoit le signal de sortie. Les différences

se situent donc au niveau de l'adaptation de la première boucle. Dans l'adaptation conventionnelle en (a), un deuxième contrôle adaptatif dédié à la première boucle reçoit le signal d'erreur dont il cherche à minimiser la puissance moyenne. Ce critère reflète mal le précédent critère décisif dans le cas des signaux à facteur de crête élevé, mais il a l'avantage que l'adaptation de la première boucle ne dépend pas de l'adaptation de la seconde. Il suffit donc d'une séquence de deux optimisations bi-dimensionnelles, une pour la première boucle suivie d'une autre pour la seconde, pour adapter le circuit en entier.

Dans l'adaptation optimale de gauche ou de droite, chaque boucle est contrôlée selon le même critère décisif basé sur le signal de sortie. L'adaptation du circuit devient donc effectivement une optimisation unique en quatre dimensions. Le contrôle adaptatif dédié à la première boucle, s'il est conservé, reçoit maintenant le signal de sortie tel qu'illustré en (b). Dans ce cas, les adaptations des deux boucles sont interdépendantes, et il faut alterner entre deux optimisations bi-dimensionnelles jusqu'à la convergence globale du circuit (Woo et autres, 2002). En deux dimensions, cela est équivalent à minimiser une surface en explorant une dimension à la fois, ce qui est d'une lenteur inacceptable. Une méthode d'adaptation doit pouvoir compenser des changements rapides, par exemple.

“...effects due to power ramping: the loop is kept aligned even when the output power is changed. For example, WCDMA uses power control on the downlink (BS to mobile) with power control periods of 0.625 ms and a power stepsize of 1 dB.” (Potecary, 1999)

Le problème provient du fait que chacun des contrôles adaptatifs ne sait pas ce que fait l'autre. La solution consiste à intégrer les deux contrôles en un seul, tel qu'illustré en (c). Dans ce dernier cas, il devient possible d'effectuer une seule optimisation en quatre dimensions pour adapter le circuit en entier. Cette opération a été simulée sur MATLAB au moyen de la fonction pré définie de recherche du simplexe de Nelder et Mead. Un examen ultérieur des résultats des simulations a démontré que l'algorithme avait eu de la difficulté à faire converger la partie imaginaire du gain du second modulateur vectoriel à

partir d'une valeur initiale de zéro. La cause de cette difficulté, heureusement sans conséquences ultimes, s'est avérée être une « amélioration » de l'algorithme de MATLAB documentée au sein du programme³. À la place du simplexe initial quasi-régulier utilisé par Nelder et Mead, qui est de largeur égale dans chacune des dimensions (Wolfe, 1978), quelqu'un a suggéré des largeurs égales à 5 % des coordonnées initiales, et égales à la valeur arbitraire de 0,00025 dans les dimensions correspondant à des coordonnées initiales nulles. Des simulations récemment répétées avec l'usage du simplexe initial quasi-régulier n'ont plus manifesté de problème de convergence. Qui plus est, la vitesse de convergence a pu être améliorée par le choix judicieux de la taille du simplexe initial.

4.2.2 Tolérance du gain du second modulateur vectoriel

On doit corriger une erreur à la section V-B du second article concernant la tolérance δ_2 , qui est la valeur admissible de l'erreur $|\varepsilon_\beta|$ dans le gain du second modulateur vectoriel.

On lit que l'IMSR de sortie dû uniquement à l'erreur $|\varepsilon_\beta| = \delta_2$ est approximativement égal à $\delta_2^2 IMSR_M$, et que si cet IMSR est bien meilleur que la valeur requise $IMSR_{req}$, c'est-à-dire $\delta_2 < \sqrt{IMSR_{req} / IMSR_M}$, on peut omettre δ_2 de la liste des paramètres de simulation. En réalité, l'approximation $\delta_2^2 IMSR_M$ est trop mauvaise pour les valeurs élevées de déséquilibre ε_α dans la première boucle qui sont associées à l'adaptation optimale de gauche. On tient compte de ε_α en utilisant l'équation exacte 4.15 avec $|\varepsilon_\beta| = \delta_2$, soit $IMSR \approx |1 + \varepsilon_\alpha|^2 \delta_2^2 IMSR_M$. La condition sur δ_2 devient alors

$$\delta_2 < \frac{\sqrt{IMSR_{req} / IMSR_M}}{|1 + \varepsilon_\alpha|}. \quad \text{Équation 4.35}$$

³ On peut examiner l'algorithme de recherche du simplexe de MATLAB et son « amélioration » avec la commande « edit fminsearch ».

Pour illustrer les implications de cette correction, le tableau 4.2 calcule plusieurs expressions reliées aux points optimaux de la colonne centrale de la figure 15 du second article pour $IMSR_{req} = -60$ dB. Dans la séquence des trois méthodes d'adaptation au tableau, $IMSR_M$ augmente graduellement puisque la classe optimale d'amplificateur principal passe de AB ($V_{dc} = 0.18$) à B ($V_{dc} = 0$). Les valeurs de $\sqrt{IMSR_{req} / IMSR_M}$ donnent à penser que la tolérance δ_2 doit être dix fois plus petite pour la méthode optimale de gauche que pour la méthode de minimisation de puissance. Lorsque ces chiffres sont corrigés par $|1 + \varepsilon_a|$, qui passe de 1 à 0.41 dans la séquence du tableau, la tolérance est seulement quatre fois plus petite pour la méthode optimale de gauche, et n'est plus très éloignée de la tolérance pour la méthode optimale de droite, laquelle est trois fois plus petite.

Tableau 4.2 Comparaisons se rapportant à la tolérance δ_2 pour trois méthodes d'adaptation et $IMSR_{req} = -60$ dB

Expression	Adaptation	Minim. de puissance	Optimale de droite	Optimale de gauche
$IMSR_M$		-31.592 dB	-20.200 dB	-11.411 dB
$\sqrt{IMSR_{req} / IMSR_M}$		0.0380	0.0102	0.0037
$ 1 + \varepsilon_a $		1	0.87	0.41
$\sqrt{IMSR_{req} / IMSR_M} / 1 + \varepsilon_a $		0.0380	0.0118	0.0090

Ces résultats plus avantageux sont dus au fait que le rapport signal à intermodulation à la sortie, après avoir été amélioré par l'annulation de l'intermodulation, se voit encore amélioré de façon substantielle par la réinjection de la portion du signal d'entrée ayant infiltré le signal d'erreur. En effet, comme expliqué à la section 4.1.2, cette réinjection a un effet destructif non pas sur le signal amplifié mais sur l'*erreur* ε_a dans le gain de ce

signal. Dans les cas d'adaptation optimale, où le gain du signal amplifié est diminué en module, la réinjection du signal d'entrée a un effet constructif.

4.2.3 Discontinuités dans la solution graphique

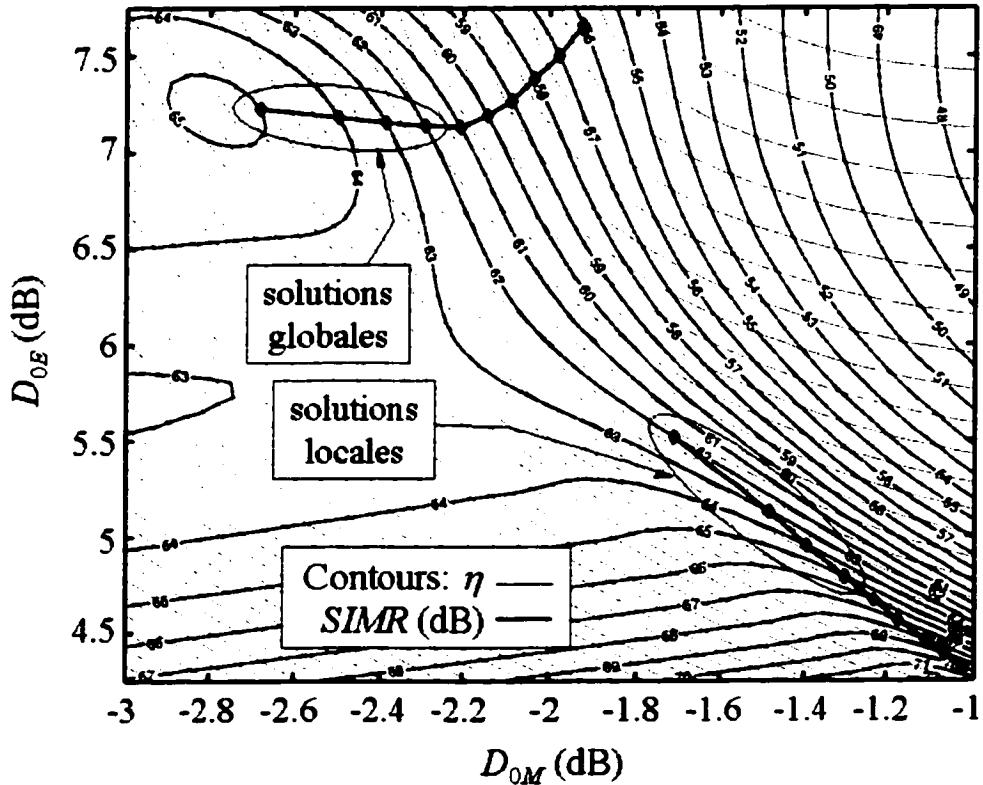


Figure 4.15 Contours de rendement énergétique (η) et de linéarité de sortie (SIMR) avec points de solution. Solutions locales et globales pour $SIMR = 62$ à 65 dB.

Les discontinuités observées dans les courbes de paramètres optimaux de la figure 16 du second article ont été reliées aux changements nécessaires de minimums locaux dans la portion graphique de la procédure d'optimisation, où des points de solution sont localisés à partir du tracé des contours de linéarité et de rendement. L'exemple de tracé offert à la

figure 4.15 est basé sur le signal d'entrée à deux tonalités. étant donné que le signal AMRC n'a pas tendance à produire ce genre de situation. Le trajet des points de solution qui vaut pour les valeurs inférieures de $SIMR$ s'arrête ici à $SIMR = 65$ dB, et de même, le trajet qui vaut pour les valeurs supérieures ne va pas en dessous de $SIMR = 62$ dB. On peut vérifier sur le tracé que les solutions globales pour $SIMR = 62$ à 65 dB appartiennent au premier groupe de points. La possibilité de solutions locales est une des raisons pour laquelle la localisation d'une série de points de solution ne peut pas être entièrement automatisée (section 4.1.9) et requiert un minimum de supervision humaine.

4.2.4 Amplificateur à post-compensation pré-écrêteur

La section V-D du second article intitulée « Input Signal Splitting » propose un mode de fonctionnement de l'amplificateur à post-compensation qui devait par la suite justifier la nouvelle désignation d'amplificateur à post-compensation *pré-écrêteur*. Le diagramme de la figure 18 de l'article, redessiné ici à la figure 4.16, montre que le signal d'entrée subit effectivement une division (*splitting*) en deux parties inférieure et supérieure. Il est plus significatif cependant de noter que l'amplificateur de classe A peut maintenant être perçu comme un nouvel amplificateur principal qui bénéficie d'un écrêtage préalable du signal d'entrée, d'où le terme *pré-écrêteur*. En effet, la distinction entre les deux amplificateurs est devenue brouillée par le fait qu'ils se partagent presque également le travail d'amplification du signal d'entrée, et la consommation en CC de l'amplificateur d'erreur peut atteindre et même dépasser celle de l'amplificateur principal. La figure 4.16 représente donc le concept de l'amplificateur à post-compensation pré-écrêteur de façon conservatrice en conservant les similarités qui le rattachent au concept original de post-compensation. La première boucle est cependant déjà renommée *boucle d'écrêtage*, et la seconde, *boucle de restitution des crêtes*. La figure 4.17 qui suit se libère de toute contrainte en élevant l'amplificateur de classe A au rang principal et en renommant l'amplificateur de classe C *amplificateur de crête*.

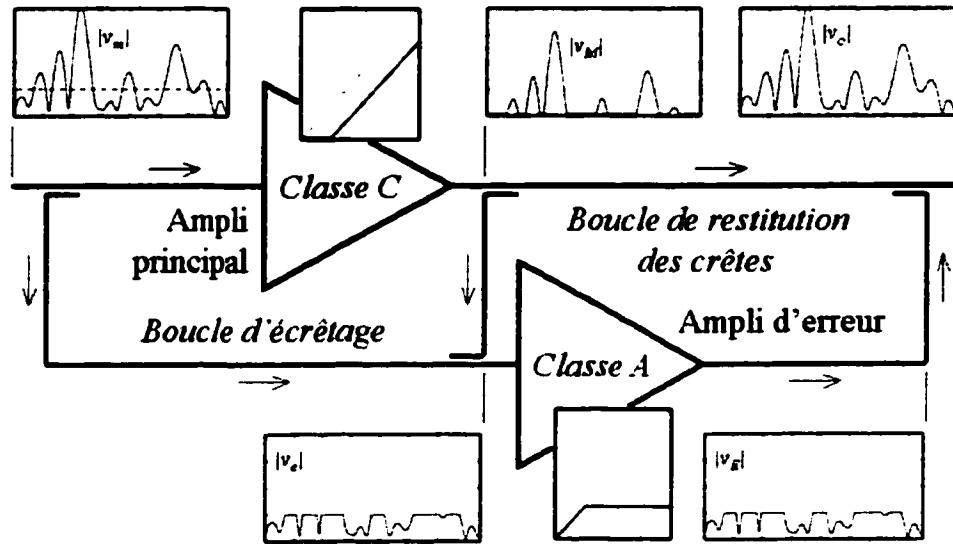


Figure 4.16 Représentation conservatrice du concept d'amplificateur à post-compensation pré-écrêteur

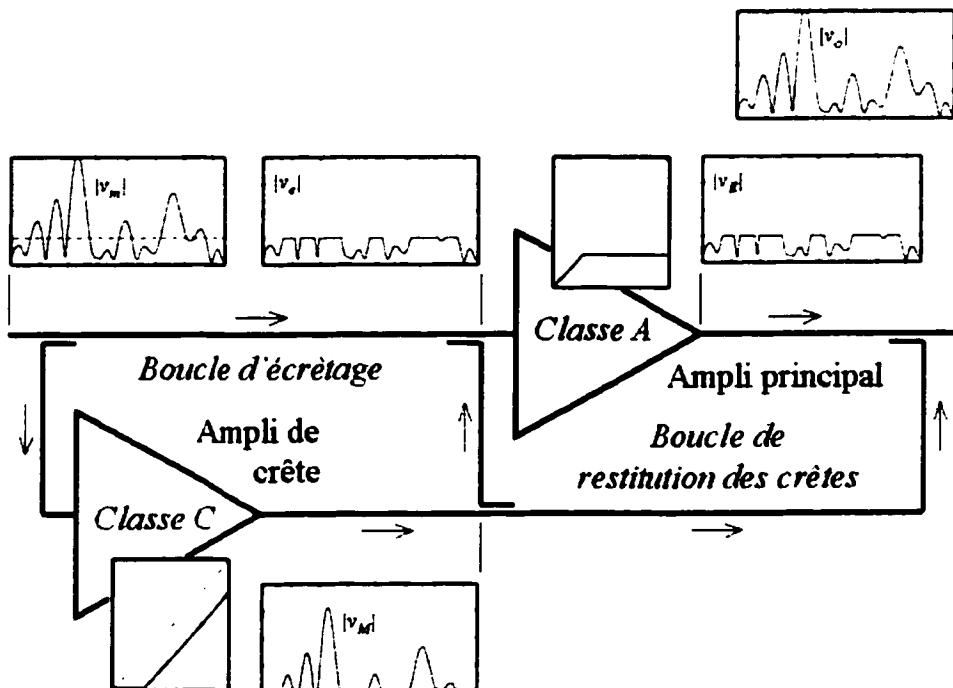


Figure 4.17 Représentation alternative du concept d'amplificateur à post-compensation pré-écrêteur

CHAPITRE V

CONCLUSION

Ce chapitre comprend une discussion générale et intégrée des résultats de la recherche et les recommandations pertinentes. Les résultats peuvent se résumer à l'élaboration d'un logiciel versatile d'optimisation énergétique de l'amplificateur à post-compensation et à la démonstration, au moyen de ce même logiciel, du potentiel énergétique supérieur de deux nouvelles méthodes d'adaptation du circuit lorsque le signal d'entrée a un facteur de crête élevé.

5.1 Introduction

Sans idées radicalement nouvelles, il aurait été risqué d'investir dans le sujet du rendement énergétique de l'amplificateur à post-compensation, étant donné le degré de maturité de cette technologie. La démarche était motivée en grande partie par l'intuition de l'amplificateur à post-compensation *pré-écrèteur*, même si la présentation de ce concept en toute fin du second article donne l'impression qu'il a été une réflexion après-coup. Ce concept, sans référence aucune à de nouvelles méthodes d'adaptation, s'appuyait sur un amplificateur principal en classe C idéalisé. Il promettait une amplification efficace du signal d'entrée en échange d'un signal d'erreur plus substantiel mais au facteur de crête grandement diminué. La validation du nouveau concept a dû attendre la découverte d'une nouvelle approche de contrôle adaptatif. En effet, l'adaptation traditionnelle de la première boucle, en cherchant à minimiser la puissance moyenne du signal d'erreur, empêchait la condition qui aurait réduit comme prévu le facteur de crête de ce signal. De la nouvelle approche adaptative est née l'adaptation optimale *de droite*, avantageuse au niveau du rendement mais au comportement différant des attentes, et ensuite l'adaptation optimale *de gauche*, encore plus avantageuse et au comportement correspondant aux attentes. Plus encourageant encore, les bénéfices des nouvelles méthodes d'adaptation ont

été obtenus tout aussi bien avec des amplificateurs non idéalisés. Dans ce dernier cas, le régime de fonctionnement optimal de l'amplificateur principal se situait le plus souvent en classe B ou AB, ce qui est plus accessible au plan technique que le fonctionnement en classe C demandé par le concept initial de pré-écrêtage.

Ces résultats laissent souvent incrédule parce que les méthodes d'adaptation connues se valent toutes dans la mesure où elles sont correctement implantées. On demandera par exemple pourquoi les nouvelles méthodes d'adaptation n'ont pas été découvertes plus tôt. En fait, il est possible que ces méthodes aient déjà été découvertes puis oubliées dans le contexte du signal classique à deux tonalités, où elles auraient procuré peu de bénéfices. On demandera aussi s'il y a quelque part une « attrape ». À ce titre, il faut admettre que la non-linéarité accrue de l'amplificateur principal nécessitera une précision accrue dans l'annulation de l'erreur effectuée par la deuxième boucle, mais plusieurs points donnés à la section 5.3 permettront d'atténuer cette préoccupation. Finalement, on demandera à voir des réalisations expérimentales. Même si celles-ci n'en sont qu'au stade de la planification, le fait que la théorie repose sur des représentations bien établies du comportement des composantes micro-ondes passives ou actives procure une certaine mesure de confiance. Une comparaison convaincante du débordement spectral d'un signal AMRC a confirmé l'utilité de la modélisation en quadrature d'un amplificateur sur la bande passante concernée. L'adaptation optimale de droite a été réalisée manuellement et de façon informelle au laboratoire de la compagnie AmpliX⁴, et a fait l'objet d'une réalisation automatique par une équipe coréenne (Woo et autres, 2002).

5.2 Logiciel d'optimisation énergétique

L'article de Cavers (1995), sans se préoccuper du rendement de l'amplificateur à post-compensation, a tout de même fourni plusieurs des fondements du logiciel d'optimisation énergétique. On y trouve en effet une normalisation de la seconde boucle du circuit, un

⁴ <http://www.mitctelcom.com>

modèle complexe en bande de base de l'amplificateur principal qui résume à la fois ses réponses MA-MA et MA-MP, et une description statistique du signal d'entrée rendue suffisante par l'usage de l'IMSR comme seul critère de linéarité. La normalisation sert à simplifier l'analyse sans sacrifier la généralité. La considération de la nature distinctive du signal d'entrée est nécessaire étant donné la prépondérance du signal à deux tonalités dans les analyses publiées, et l'importance actuelle des signaux ayant des facteurs de crête bien supérieurs. L'analyse complexe en bande de base offre un rendement de calcul amélioré qui devait s'avérer salutaire étant donné le grand nombre d'options devant être explorées par le logiciel, et la réponse MA-MP de l'amplificateur de puissance, qui a tendance à être négligée par les chercheurs, devait s'avérer centrale aux phénomènes étudiés. L'approche statistique augmente encore et de façon dramatique le rendement de calcul, et exempte d'un long échantillonnage du signal d'entrée en permettant l'utilisation de la fonction de distribution cumulée complémentaire (FDCC), dont l'usage est devenu standard. L'approche statistique n'admet que l'IMSR comme critère de linéarité parce qu'elle empêche la représentation spectrale des signaux, mais comme le prévoyait Cavers, l'IMSR s'est avéré simplement relié au rapport de densité spectrale normalement utilisé par les organismes de réglementation. Mis à part ces emprunts, le logiciel comporte de nombreuses innovations exclusives à l'auteur au point de vue du modèle utilisé et du type d'analyse effectuée.

Au point de vue du modèle utilisé, l'innovation centrale est une version améliorée du modèle de l'amplificateur de puissance proposé par Cavers. Le nouveau modèle spécifie indépendamment le gain, la puissance de sortie de saturation, la caractéristique de compression du gain, et pour introduire le volet énergétique, la caractéristique de puissance consommée. Cette versatilité permet d'optimiser le gain et la taille d'un amplificateur utilisant un type de transistor donné et fonctionnant en une classe donnée. En étendant l'usage du modèle à l'amplificateur d'erreur, on tenait compte de la contribution de ce dernier à la consommation totale du circuit, et non moins significatif, on tenait compte simultanément des effets de la non-linéarité de l'amplificateur d'erreur et d'une adaptation

imparfaite de la seconde boucle sur la non-linéarité résiduelle du circuit. Entre autres innovations, le premier modulateur vectoriel précède immédiatement l'amplificateur principal comme le veut la pratique; cette condition qui est critique au niveau du fonctionnement des techniques d'adaptation est évitée par les chercheurs pour cause de complexité. On prend aussi en considération les rapports qui doivent exister entre les pertes dues aux lignes à délai et les gains des amplificateurs dont ils compensent les délais. Ces rapports évitent d'avoir à spécifier au préalable des gains ou des délais arbitraires. Finalement, on emploie un modèle réaliste du coupleur avec perte d'insertion, et on inclut dans toutes les branches du circuit des atténuateurs pour simuler les pertes possibles dues entre autres aux échantillonnages requis par les contrôles adaptatifs.

Au point de vue du type d'analyse effectuée, on a étendu la normalisation à l'ensemble du circuit. La position pratique du premier modulateur vectoriel, ainsi que l'usage d'un modèle non linéaire pour l'amplificateur d'erreur, amènent deux changements dans l'analyse. En premier lieu, ils forcent le recours à des recherches en deux dimensions pour déterminer les gains des deux modulateurs vectoriels qui satisfont les critères d'adaptation employés. L'adaptation de la deuxième boucle à l'aide d'un signal pilote n'est donc plus possible dans le nouveau contexte statistique puisque son critère est évalué à partir de la description temporelle ou spectrale du signal de sortie. En deuxième lieu, les modules des erreurs de gain des modulateurs vectoriels ne déterminent plus à eux seuls l'effet sur l'IMSR de sortie, les phases de ces erreurs ayant elles aussi un effet déterminant. Le scénario de la pire éventualité est donc identifié en supposant les modules des erreurs relatives de gain égales aux tolérances des modulateurs vectoriels et en recherchant les phases de ces erreurs qui dégradent au maximum l'IMSR de sortie. Le logiciel emploie l'approche pratique d'optimiser le rendement énergétique du circuit en fonction de la performance demandée en termes d'IMSR et de gain à la sortie, et non pas l'inverse. L'optimisation nécessite une solution graphique qui a pu être partiellement automatisée et qui a l'avantage d'être particulièrement évocatrice. Elle produit pour une séquence prédéterminée d'IMSR les valeurs optimales des gains et des puissances de

sortie de saturation des amplificateurs principal et d'erreur, ainsi que les valeurs optimales de trois couplages. Le quatrième couplage, soit celui d'entrée, est exclu parce que sa valeur optimale a été identifiée en termes de deux autres valeurs de paramètres d'entrée.

La restriction aux techniques d'adaptation sans signal pilote a fourni l'occasion de réexaminer cette famille de techniques, laquelle a été associée à une plus grande flexibilité et capacité du circuit (Kenington et autres, 1991). On note que l'approche statistique, en rendant impossible la représentation spectrale des signaux, n'a pas permis de modifier l'analyse complexe en bande de base pour qu'elle tienne compte des réponses fréquentielles en bande (Loyka, 2001), par exemple les effets de mémoire dans les amplificateurs et autres composantes, les déséquilibres dans les délais, et le filtrage par lequel on peut assister l'adaptation (Cavers, 1995). La restriction sur ce filtrage, ajoutée à la restriction sur le signal pilote, a restreint l'adaptation aux méthodes à suppression de signal par troisième boucle. Un travail de synthèse a permis de réunir sous ce même concept la suppression adaptative du signal d'entrée (Bauman, 1983), la suppression du signal d'erreur (Ghannouchi et autres, 2001), et la suppression *non adaptative* du signal d'entrée (Larose et Ghannouchi, juin 2001). L'introduction du déséquilibre ε dans la troisième boucle, en plus de tenir compte d'inévitables erreurs de réalisation, démarque conceptuellement les deuxième et troisième méthodes et permet d'en inclure une quatrième, soit l'usage du signal de sortie sans suppression aucune du signal d'entrée (Cavers, 1995).

L'approche statistique « sans mémoire » est plus difficile à soutenir pour les largeurs de bande employées par exemple dans le système cellulaire AMRC. On a vérifié expérimentalement, et avec grand succès, le débordement spectral d'un signal AMRC amplifié par l'entremise du modèle complexe en bande de base d'un amplificateur de puissance. Cependant, la détermination précise des valeurs absolues des paramètres de conception n'était pas l'objectif premier de l'élaboration du logiciel. Celui-ci était plutôt

destiné à l'exploration des types d'amplificateurs et des méthodes d'adaptation permettant un rendement optimal du circuit. Ces deux facteurs, de par leur nature complexe, ne peuvent être optimisés directement et requièrent une recherche par tâtonnements qui nécessite un grand nombre de simulations.

5.3 Nouvelles méthodes d'adaptation

Pour la deuxième boucle du circuit, les méthodes d'adaptation sans signal pilote ni filtrage demandent l'ajout d'une troisième boucle de circuit qui extrait la distorsion résiduelle du circuit à partir du signal de sortie ou du signal d'erreur amplifié en soustrayant respectivement une copie du signal d'entrée ou d'erreur. La méthode de suppression du signal d'erreur a déjà fait l'objet d'un dépôt de brevet (Ghannouchi et autres, 2001), et la méthode de suppression du signal d'entrée a été suggérée par certains auteurs (Bauman, 1983; Cavers, 1995) mais seulement dans sa version *adaptive*, c'est-à-dire que la troisième boucle devait être équipée de son propre modulateur vectoriel. Il a été démontré ici que l'amplificateur à post-compensation, lorsque ses deux modulateurs vectoriels sont positionnés comme le veut la pratique, délivre un gain constant à la sortie en dépit de toute fluctuation dans le gain du signal amplifié, et que par conséquent la suppression du signal d'entrée peut fort bien être effectuée en mode non adaptatif. En fait, les méthodes de suppression du signal d'erreur et de suppression non adaptative du signal d'entrée sont conceptuellement équivalentes lorsqu'elles sont parfaitement exécutées.

Traditionnellement, la première boucle du circuit n'est pas adaptée en fonction de la distorsion résiduelle du circuit, qui constitue pourtant le critère décisif, mais plutôt en fonction du signal d'erreur qu'on cherche à minimiser en puissance. Étant donné la relation très non linéaire entre le signal d'erreur instantané et la distorsion résiduelle instantanée, il est étonnant avec le recul que cette approche n'ait pas été remise en question plus tôt. Les deux nouvelles méthodes d'adaptation proposées ici ont en

commun le fait que la première boucle est adaptée comme la deuxième en fonction du critère de distorsion résiduelle. Avec un seul critère à évaluer, le matériel est quelque peu simplifié, et la paire d'optimisations à deux dimensions est remplacée par une seule optimisation à quatre dimensions. Le succès constaté de la méthode de recherche du simplexe dans cette dernière optimisation, après correction de l'algorithme de MATLAB, est d'autant plus significatif que cette méthode convient à la réalisation pratique des contrôles adaptatifs. En effet, par opposition aux méthodes d'optimisation par gradient, les méthodes de recherche directe telles que la méthode du simplexe sont indiquées lorsque la fonction, dans le cas présent la distorsion résiduelle du circuit, dépend d'une mesure physique et par conséquent est sujette à des erreurs aléatoires (Swann, 1972).

La nouvelle méthode d'adaptation dite *optimale de droite* est obtenue lorsque la distorsion résiduelle est extraite par suppression du signal d'erreur ou par suppression *non adaptative* du signal d'entrée. Par rapport à l'adaptation traditionnelle, cette méthode optimale produit un signal d'erreur à puissance moyenne supérieure mais à facteur de crête bien inférieur, ce qui le rend dans l'ensemble moins stressant pour l'amplificateur d'erreur et diminue la distorsion résiduelle générée par ce dernier.

Lorsque la distorsion résiduelle est extraite par suppression *adaptative* du signal d'entrée, un deuxième minimum local de distorsion apparaît beaucoup plus à gauche dans le plan complexe du gain du premier modulateur vectoriel, et peut être exploité si le point de départ de l'optimisation est choisi en conséquence. On obtient ainsi la nouvelle méthode d'adaptation dite *optimale de gauche*. Cette méthode produit une distorsion résiduelle substantielle au point de vue de l'amplificateur d'erreur, mais à très forte corrélation avec le signal d'entrée, donc négligeable *au point de vue du signal de sortie*. Dans la suppression du signal d'erreur ou la suppression non adaptative du signal d'entrée, la distorsion résiduelle était jugée selon le point de vue de l'amplificateur d'erreur, ce qui explique que le deuxième minimum local de distorsion était absent. La méthode optimale de gauche crée une relation particulière entre les fonctions de transfert non linéaires des

deux amplificateurs qui a pour résultat, dans le contexte de la post-compensation, de compenser la fonction de transfert non linéaire de l'amplificateur principal. Paradoxalement, elle crée un certain déséquilibre des gains dans la seconde boucle qui dégrade l'annulation des autres types de distorsion, comme par exemple la distorsion linéaire (effets de mémoire) et le bruit. La méthode s'apparente à la technique de linéarisation LINC (Cox, 1974) en ce qu'elle utilise deux amplificateurs fonctionnant en mode non linéaire, mais elle a l'avantage que les distorsions autres que la distorsion d'intermodulation sont elles aussi annulées, bien qu'imparfaitement.

Pour un amplificateur à post-compensation donné, l'amélioration de la linéarité de sortie due aux nouvelles méthodes d'adaptation optimales s'est avérée peu significative avec un signal d'entrée à deux tonalités, mais parfois impressionnante avec un signal d'entrée AMRC à facteur de crête élevé. En corollaire, le logiciel d'optimisation énergétique a prouvé qu'avec le signal d'entrée AMRC, des contraintes de conception typiques, et une linéarité de sortie donnée, le rendement du système pouvait être grandement amélioré par l'application des méthodes d'adaptation optimales. Cette amélioration s'est avérée très dépendante de la classe de l'amplificateur principal. Afin d'étudier systématiquement le phénomène, des caractéristiques d'amplificateurs allant de la classe A à la classe C ont été générés au moyen du modèle en grands signaux d'un MESFET typique. L'approche de Turlington (2000) a été utilisée, mais en remplaçant son modèle comportemental par le modèle de Statz, lequel convient particulièrement aux classes autres que A (Pitzalis, 1993), et en modifiant la méthode de génération de la réponse MA-MP pour la rendre applicable aux classes B et C. En balayant cette fois la classe de l'amplificateur principal, le logiciel a permis d'estimer les gains possibles en rendement pour différentes valeurs de linéarité de sortie, avec les résultats qu'on retrouve au second article. Pour un signal AMRC et un SIMR de sortie de 60 dB, qui est typique pour les systèmes mobiles, la méthode optimale de droite a permis avec un amplificateur principal fonctionnant près de la classe B un rendement de 21 % supérieur à ce que l'adaptation conventionnelle

permettait avec un amplificateur de classe AB, et la méthode optimale de gauche a permis un rendement 36 % supérieur avec un amplificateur de classe B.

Les méthodes d'adaptation optimales demandent donc que l'amplificateur principal fonctionne plus près de la classe B, et parfois même en classe C dans l'adaptation optimale de gauche. Par conséquent, l'avantage énergétique réel dépassera les chiffres précédents si le niveau du signal d'entrée est appelé à varier, par exemple si le nombre de porteuses change dans un signal multi-porteuse. Dans ce cas, les systèmes optimaux auront été conçus en fonction du plus haut niveau de signal d'entrée anticipé, et pendant la majorité du temps où le signal d'entrée descendra au-dessous de ce niveau, le système dont l'amplificateur principal est proche de la classe B maintiendra mieux son rendement optimal que le système dont l'amplificateur principal est plus près de la classe A.

En contrepartie, la linéarité de l'amplificateur fonctionnant plus près de la classe B sera moins bonne, et pour amener cette linéarité au niveau requis, il faudra davantage de précision dans la deuxième boucle, tant au niveau du contrôle adaptatif qu'au niveau des réponses fréquentielles sur la bande passante. Cette tendance n'est pas forcément un problème puisque historiquement, l'industrie a déjà fait un mouvement en ce sens, préférant les amplificateurs principaux de classe AB à ceux de classe A pour des motifs de rendement énergétique, et relevant les défis de précision par les progrès de la technologie (Zhu et autres, 1999). Parmi les progrès récents, on mentionnera la compensation adaptative des délais et des réponses fréquentielles, par laquelle a été obtenue une réduction de 40 dB des produits d'intermodulation sur une bande passante de 40 MHz (Smith et Cavers, 1998). Il peut être utile de souligner, cependant, que la sensibilité du rendement à la classe d'amplificateur principal est considérablement diminuée par rapport à l'adaptation conventionnelle, comme on peut le constater à la figure 15 du second article. On peut donc au besoin bouger en direction de la classe A avec une perte de rendement relativement petite. Finalement, on peut recourir à la solution « des grands moyens » qui consiste à distribuer le degré de linéarisation requis sur deux étages de post-compensation. En échange d'une complexité accrue, l'emploi

d'un premier étage à objectif de linéarité moins ambitieux, suivi d'un deuxième étage à consommation en CC réduite, comporte un avantage énergétique certain. Cette option intrigante mériterait d'être explorée au moyen d'une mise à jour appropriée du logiciel d'optimisation.

Dans la méthode optimale de droite, la diminution de la tolérance dans l'ajustement de la seconde boucle a l'avantage d'être accompagnée par une augmentation plus que proportionnelle de la tolérance dans l'ajustement de la première boucle. L'inévitable déséquilibre des gains dans la troisième boucle contribue à l'imprécision dans le contrôle adaptatif de la seconde boucle, mais on réduit cette contribution à un niveau égal au déséquilibre de la troisième boucle en choisissant la suppression du signal d'erreur. L'alternative de la suppression non adaptative du signal d'entrée n'est pas viable dans ce cas parce qu'elle sensibilise l'adaptation de la seconde boucle au déséquilibre qui existe dans la première (Larose et Ghannouchi, août 2001). On sait en effet que la méthode optimale de droite, tout comme la méthode optimale de gauche, laisse passer une portion importante du signal d'entrée dans le signal d'erreur, d'où l'expression *annulation imparfaite du signal* (Woo et autres, 2002).

La méthode optimale de gauche, quant à elle, ne bénéficie pas d'une augmentation de la tolérance dans l'ajustement de la première boucle. Par rapport à la méthode optimale de droite, on s'attendrait à ce que la diminution de la tolérance dans l'ajustement de la seconde boucle soit d'autant plus forte que la non-linéarité de l'amplificateur principal est plus prononcée, mais heureusement, cette tendance est en grande partie annulée par l'effet linéarisant d'un résidu plus important de signal d'entrée réinjecté à la sortie. La nécessité d'une suppression *adaptative* du signal d'entrée est un désavantage au plan de la complexité, mais lorsque le signal d'entrée occupe en tout temps la même bande de fréquence, comme dans les systèmes AMRC, on peut utiliser un filtre pour séparer le signal désiré de la distorsion, comme le recommande Cavers (1995) et comme l'on fait Woo et autres (2002) en rapport avec la méthode optimale de droite. Cette option « avec

mémoire » aurait pu être simulée en abandonnant l'approche statistique, mais les temps de calcul seraient devenus prohibitifs compte tenu des ressources disponibles. L'autre option n'ayant pas été simulée, soit l'usage d'un signal pilote, est fondamentalement incompatible avec les méthodes d'adaptation optimales parce que l'objectif de la suppression du signal pilote à la sortie du circuit peut être atteint indépendamment de l'ajustement de la première boucle, qui requiert alors son propre critère d'adaptation.

Un dernier point digne de mention est que l'adaptation optimale de gauche sacrifie l'aptitude distinctive et légendaire de l'amplificateur à post-compensation à annuler aussi bien tous les types de distorsion. Le léger déséquilibre dans la seconde boucle que requiert la méthode pour annuler la distorsion d'intermodulation diminue en même temps la précision de l'annulation des effets de mémoire et du bruit. En idéalisant quelque peu les amplificateurs principal et d'erreur au moyen de transistors à transconductance plus constante, on réduit de beaucoup ce phénomène, et du même coup on met clairement en évidence le concept de l'amplificateur à post-compensation *pré-écrêteur* qui avait motivé les présents travaux. Quelle que soit sa viabilité en pratique, ce concept conserve une valeur éducative en faisant ressortir les différences fondamentales entre les systèmes optimaux utilisant l'adaptation optimale de gauche et l'adaptation conventionnelle.

BIBLIOGRAPHIE

ADBY, P.R. et DEMPSTER, M.A.H. (1974). *Introduction to Optimization Methods*. Chapman and Hall.

AGILENT TECHNOLOGIES. (2000). Performing cdma2000 measurements today. Application Note 1325.

AGILENT TECHNOLOGIES. (2001). Three carrier IS-95A CDMA waveform for the Agilent ESG-D series. [En ligne]. Disponible: <http://www.agilent.com>. Quick Search avec « DAB ».

ANDRENKO, A.S., HORIGUCHI, K. et autres. (1999). Optimization analysis of feedforward power amplifier. *1999 Asia Pacific Microwave Conf. Proc.*, 3, 626-629.

BATEMAN, A. (1999). Extending talk time through improved transmitter efficiency. *1999 IEEE Wireless Communications and Networking Conf.*, 2, 727-731.

BATEMAN, A. et HAINES, D. (1989). Direct conversion transceiver design for compact low-cost portable mobile radio terminals. *Proc. IEEE Veh. Technol. Conf.*, 1, 57-62.

BAUMAN, R.M. (1983). Adaptive feed-forward system. U.S. Patent 4 389 618.

BENNETT, T.J. et CLEMENTS, R.F. (1974). Feedforward—an alternative approach to amplifier linearization. *Radio and Electronic Engineer*, 44, 257-262.

BLACK, H.S. (1928). Translating system. US Patent 1 686 792.

BOLAN, D. et EISENBERG, J. (juil. 1995). MCPA strategy for emerging technologies. *Cellular Business*, 56-63.

CAVERS, J.K. (1995). Adaptation behavior of a feedforward amplifier linearizer. *IEEE Trans. Veh. Technol.*, 44, 31-40.

COX, D.C. (1974). Linear amplification with nonlinear components. *IEEE Trans. Commun.*, 22, 1942-1945.

DRUCKER, E.H. (nov. 1994). The advantages of broadband LPAs. *Cellular Business*, 46-52.

EID, E.E. (1995). *Linéarisation d'amplificateurs micro-ondes par une méthode «feedforward» adaptative*. Thèse de doctorat. École Polytechnique de Montréal. Canada.

EID, E.E., GHANNOUCHI, F.M. et BEAUREGARD, F. (nov. 1995). Optimal feedforward linearization system design. *Microwave Journal*, 38, 78-86.

FAULKNER, M. (1998). Amplifier linearization using RF feedback and feedforward techniques. *IEEE Trans. Veh. Technol.*, 47, 209-215.

GHANNOUCHI, F.M., BEAUREGARD, F. et autres. (2001). Adaptive linearization of a feedforward amplifier by complex gain stabilization of the error amplifier. U.S. Patent 6 275 105 B1.

GOLIO, M. (2001). *The RF and Microwave Handbook*, CRC Press.

HAYKIN, S.S. (1991). *Adaptive Filter Theory*, Prentice Hall.

HICKSON, M.T., PAUL, D.K. et autres. (1994). High efficiency feedforward linearizers. *24th European Microwave Conf. Proc.*, 1, 819-824.

HORIGUCHI, K., NAKAYAMA, M. et autres. (2001). A high efficiency feedforward amplifier with a series diode linearizer for cellular base stations. *IEEE MTT-S Int. Microwave Symp. Dig.*, 2, 797-800.

JECKELN, E.G., GHANNOUCHI, F.M. et autres. (2001). Efficient baseband/RF feedforward linearizer through a mirror power amplifier using software-defined radio and quadrature digital up-conversion. *2001 IEEE MTT-S Int. Microwave Symp. Dig.*, 2, 789-792.

JURGEN, R.K. (avr. 1972). Feedforward correction: a late-blooming design. *IEEE Spectrum*, 9, 41-43.

KENINGTON, P.B., WILKINSON, R.J.. et MARVILL, J.D. (1991). Broadband linear amplifier design for a PCN base-station. *Proc. IEEE Veh. Technol. Conf.*, 155-160.

KENINGTON, P.B. (1992). Efficiency of feedforward amplifiers. *IEE Proc. G. Circuits Devices Syst.*, 139, 591-593.

KENINGTON, P. (sep. 1999). Achieving high efficiency in multi-carrier base-station power amplifiers. *Microwave Engineering Europe*, 83-90.

KEYS, R.G. (1981). Cubic convolution interpolation for digital image processing. *IEEE Trans. Acoust., Speech, Signal Processing*, 29, 1153-1160.

KOUKI, A.B. et OUTALEB, N. (2000). Performance improvement of feedforward power amplifiers by using a novel injection combining technique. *2000 IEEE MTT-S Int. Microwave Symp. Dig.*, 1, 481-484.

KUMAR, M., WHARTENBY, J. et WOLKSTEIN, H. (1985). Predistortion linearizer using GaAs dual-gate MESFET for TWTA and SSPA used in satellite transponders. *IEEE Trans. Microwave Theory Tech.*, 33, 1479-1488.

LAROSE, C.L. et GHANNOUCHI, F.M. (2000). Optimization of feedforward amplifier power efficiency on the basis of input power statistics. *IEEE MTT-S Int. Microwave Symp. Dig.*, 3, 1491-1494.

LAROSE, C.L. et GHANNOUCHI, F.M. (juin 2001). First-Loop control in feedforward amplifiers for high-stress communications signals. *Proc. Int. Symp. on Microwave and Optical Technol.*, 195-198.

LAROSE, C.L. et GHANNOUCHI, F.M. (août 2001). Pilotless adaptation of feedforward amplifiers driven by high-stress signals. *Proc. Radio Wireless Conf.*, 81-84.

LAROSE, C.L. et GHANNOUCHI, F.M. (2003). Optimization of feedforward amplifier power efficiency on the basis of drive statistics. *IEEE Trans. Microwave Theory Tech.*, 51.

LAROSE, C.L. et GHANNOUCHI, F.M. (soumis pour publication). Optimal amplifier classes and adaptation methods for feedforward amplifiers driven by high-stress signals. *IEEE Trans. Veh. Technol.*

LOYKA, S.L. et MOSIG, J.R. (2000). New behavioral-level simulation technique for RF/microwave applications. Part I: basic concepts. *Int. J. RF and Microwave CAE.* 10, 221-237.

LUBELL, P.D., DENNISTON, W.B., et HERTZ, R.F. (avr. 1974). Linearizing amplifiers for multi-signal use. *Microwaves*, 13, 46-50.

MAAS, S.A. (1988). *Nonlinear Microwave Circuits*, Artech House.

NARAHASHI, S. et NOJIMA, T. (1991). Extremely low-distortion multi-carrier amplifier—self-adjusting feed-forward (SAFF) amplifier. *Int. Conf. on Communications Conf. Rec.*, 3, 1485-1490.

PARSONS, K.J. et KENINGTON, P.B. (avr. 1994). Effect of delay mismatch on a feedforward amplifier. *IEE Proc., Circuits Devices Syst.*, 141, 140-144.

PARSONS, K.J. et KENINGTON, P.B. (mai 1994). The efficiency of a feedforward amplifier with delay loss. *IEEE Trans. Veh. Technol.*, 43, 407-412.

PARSONS, K.J., KENINGTON, P.B. et MCGEEHAN, J.P. (1994). Efficient linearisation of RF power amplifiers for wideband applications. *IEE Colloquium on "Linear RF Amplifiers and Transmitters"*, 7/1-7.

PITZALIS, O. (1993). Computer-aided design of GaAs FET power amplifiers. *High-Power GaAs FET Amplifiers*. J. L. B. Walker, Ed., Artech House, 147-188.

POTHECARY, N. (1999). *Feedforward Linear Power Amplifiers*, Artech House.

PULSAR MICROWAVE CORP., NJ. Directional coupler specifications. [En ligne]. Disponible: <http://www.pulsarmicrowave.com/products/couplers/couplers.htm>

QUALCOMM INCORPORATED. (2002). About CDMA: history of CDMA. [En ligne]. Disponible: <http://www.qualcomm.com/cdma/history.html>

RAAB, F.H. (1986). Average efficiency of power amplifiers. *Proc. of RF Technology Expo 86*, 473-486.

SEIDEL, H. (1971). A microwave feed-forward experiment. *Bell Syst. Tech. J.*, 50, 2879-2916.

SMITH, A.M. et CAVERS, J.K. (1998). A wideband architecture for adaptive feedforward linearization. *Proc. IEEE Veh. Technol. Conf.*, 3, 2488-2492.

SWANN, W.H. (1972). Direct search methods. *Numerical Methods for Unconstrained Optimization*, W. Murray, Ed., Academic Press, 13-28.

SYNERGY MICROWAVE CORP. Couplers: parameter definitions. [En ligne]. Disponible: <http://www.synergymwave.com/products/couplers/CouplerTutorial.pdf>

TURLINGTON, T.R. (2000). *Behavioral Modeling of Nonlinear RF and Microwave Devices*, Artech House.

VENNEMA, K. (fév. 1996). Two-tone linearity in a 900 MHz silicon bipolar class AB amplifier. *Microwave Journal*, 39, 88-93.

WALKER, J.L.B. (1993). Introduction and basic theory. *High-Power GaAs FET Amplifiers*, J.L.B. Walker, Ed., Artech House, 1-41.

WILKINSON, R.J. et KENINGTON, P.B. (1992). Specification of error amplifiers for use in feedforward transmitters. *IEE Proc. G, Circuits Devices Syst.*, 139, 477-480.

WIXON, B. (mai 2001). Delay filters—size, integration, and cost advantages. *Microwave Product Digest*, 56-87.

WOLFE, M.A. (1978). *Numerical Methods for Unconstrained Optimization: An Introduction*. International Student Edition, Van Nostrand Reinhold.

WOO, Y.Y., YANG, Y. et autres (2002). Feedforward amplifier for WCDMA base stations with a new adaptive control method. *IEEE MTT-S Int. Microwave Symp. Dig.*, 2, 769-772.

WRIGHT, A. et DURTIER, W. (1992). Experimental performance of an adaptive digital linearized power amplifier. *IEEE Trans. Veh. Technol.*, 41, 395-400.

ZHU, X., ZHOU, J., FENG, L. et autres. (1999). A wide-band high-linearity feedforward power amplifier for digital communication systems. *1999 Asia Pacific Microwave Conf. Proc.*, 1, 29-32.

ANNEXE I

PREMIER ARTICLE:
« OPTIMIZATION OF FEEDFORWARD AMPLIFIER POWER EFFICIENCY
ON THE BASIS OF DRIVE STATISTICS »

Cet article a été accepté le 11 mars 2002 par le *IEEE Transactions on Microwave Theory and Techniques* et paraîtra dans le numéro de janvier 2003.

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— microwave power amplifiers, linearization, feedforward amplifiers, power efficiency, design methodology

I. INTRODUCTION

The power efficiency of a microwave amplifier is best when it is operated near saturation. Amplification of a communications signal in this non-linear range generates intermodulation (IM) distortion that interferes with neighbouring 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 broadband techniques, such as predistortion [2]-[3] and feedforward, has been made necessary by the emergence of the new CDMA (Code Division Multiple Access) 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 multi-channel 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

This work was supported in part by a Postgraduate Scholarship from the Natural Sciences and Engineering Research Council of Canada (NSERC).

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

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 modelling [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 modelling 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 IM power (SIMR) [8]. Under relatively broadband 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 the 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 $v_m(t)$ to the main amplifier, which outputs an amplified input signal $v_M(t)$ 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 $v_m(t)$ to produce an error signal $v_e(t)$. This closes the circuit's first loop, also known as the signal cancellation loop because it cancels the signal component in $v_M(t)$ in order to isolate its IM distortion. For signals to enter the coupler simultaneously, a delay line of gain L_m adds delay to $v_m(t)$ as necessary to compensate for the longer delay introduced in $v_M(t)$. The error signal $v_e(t)$ exits the error amplifier as the amplified error signal $v_E(t)$ after suffering its share of IM distortion. This signal is sent to an output coupler, with subtracts it from the main portion of $v_M(t)$ to yield a largely corrected output signal $v_o(t)$. This closes the circuit's second loop, also known as the error cancellation loop because it cancels the error signal present in $v_M(t)$. A delay line of gain L_e adds delay to $v_M(t)$ as necessary to compensate for the longer delay introduced in $v_E(t)$.

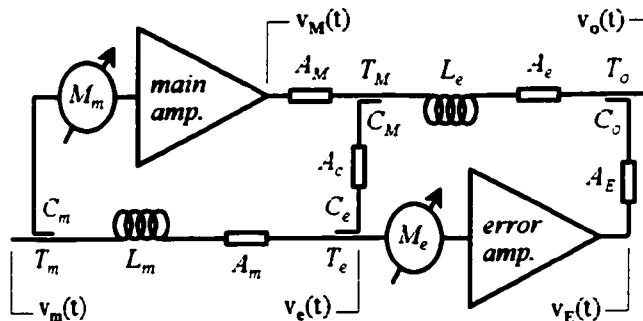


Fig. 1. Complex baseband feedforward amplifier model, with complex envelopes of signals and complex gains of components.

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 here no value of the main path power gain $|T_M L_e|^2$ is assumed.

Input parameters:

$\{x_m\}$	Set of instantaneous-to-average input powers
$P(\{x_m\})$	Probabilities of $\{x_m\}$
$\{S/IMR_{req}\}$	Required signal-to-IM power ratios 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
adp_1	Adaptation method for first loop
adp_2	Adaptation method for second loop (pilotless)
p_1	Delay-to-gain ratio of main amplifier [ns/dB]
p_2	Delay-to-gain ratio of error amplifier [ns/dB]
τ_1	Residual delay in first loop [ns]
τ_2	Residual delay in second loop [ns]
q_1	Delay-to-loss ratio of first delay line [ns/dB]
q_2	Delay-to-loss ratio of second delay line [ns/dB]
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)
$ C_m ^2$	Coupling of input coupler (<1)
$ A_c ^2$	Power gain in central branch (<1)
$ A_m ^2$	Power gain opposite main amplifier (<1)
$ A_e ^2$	Power gain opposite error amplifier (<1)
$ A_M ^2$	Power gain after main amplifier (<1)
$ A_E ^2$	Power gain after error amplifier (<1)
$ M_m ^2$	Power gain before main amplifier (<1)

$ M_e ^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 $\{S/MR_{req}\}$:

$\{ G_M ^2\}$	Uncompressed power gains of main amplifier
$\{ G_E ^2\}$	Uncompressed power gains of error amplifier
$\{ C_M ^2\}$	Couplings of main coupler (<1)
$\{ C_e ^2\}$	Couplings of error coupler (<1)
$\{ C_o ^2\}$	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 non-negotiable. 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. The optimization is therefore 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.

i) Input Parameter Descriptions:

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

$$E[x_m] = E[p_m / P_m] = E[p_m] / P_m = 1. \quad (1)$$

the statistical data set should satisfy the two conditions

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

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 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 $P_o = P_s + P_{IM}$, and then to describe the output signal linearity using the signal-to-IM power ratio, $SIMR = P_s / P_{IM}$. 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 $\{SIMR_{req}\}$ is specified.

In conservative manner, definitions involving the system's output power will use the linear output power P_s instead of the total output power 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 P_s / P_m rather than the total power gain P_o / 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

$$d = \frac{P_{in}|G|^2}{P_{sat}} = \frac{P_{in}}{P_{sat}/|G|^2}. \quad (3)$$

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 $P_{sat}/|G|^2$ 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,

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

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

Parameters adp_1 and 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, adp_2 like adp_1 must refer to a pilotless adaptation method, which uses only samples of existing signals.

Delay-to-gain ratios p_1 and p_2 specify linear relationships between the group delays of the two amplifiers and their dB 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 q_1 and q_2 are quality figures for the delay lines. The delay and dB 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

$$L = |T_m|^2 + |C_m|^2, \quad P = |T_M|^2 + |C_M|^2, \quad N = |T_e|^2 + |C_e|^2, \quad R = |T_o|^2 + |C_o|^2. \quad (5)$$

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 $L/(1+q_1/p_1)$, 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 v_E or the output signal 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 $P_{S,req} = P_m U_{req}$. This normalization is expected since neither $P_{S,req}$ nor P_m are listed among the input parameters. For a given $P_{S,req}$, the absolute saturated output powers are obtained by using

$$\{P_{satM}\} = P_{S,req} \{X_{satM}\}, \quad \{P_{satE}\} = P_{S,req} \{X_{satE}\}. \quad (6)$$

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

$$\eta = \frac{P_s}{P_{dcM} + P_{dcE}}, \quad (7)$$

where P_{dcM} and 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 Subsection 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 (Subsection C) and the second (Subsection D). Hence, the output SIMR is found to be a non-analytic function of only the two reference input drives D_{0M} and D_{0E} . The gain and power efficiency are found to be non-analytic functions of the same, but also analytic functions of other optimization parameters (Subsection E). A partial optimization reduces these parameters and the power efficiency to non-analytic 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 (Subsection 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 , P_{satM} , $f_M(d_M)$, and $h_M(d_M)$ for the main amplifier and with G_E , P_{satE} , $f_E(d_E)$, and $h_E(d_E)$ for the error amplifier. The input drives therefore become

$$d_M = \frac{P_m |C_m M_m|^2 |G_M|^2}{P_{satM}}, \quad d_E = \frac{P_e |M_e|^2 |G_E|^2}{P_{satE}}. \quad (8)$$

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

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

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

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

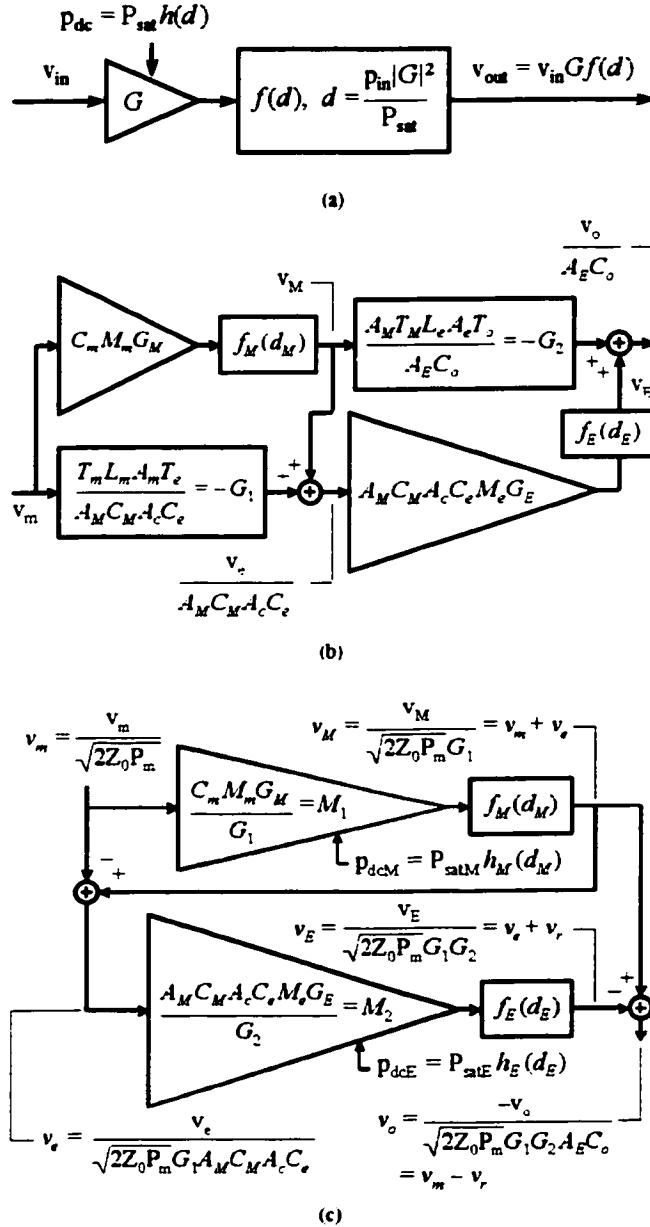


Fig. 2. Complex baseband feedforward amplifier models, with complex signal envelopes and complex gains. (a) Power amplifier model only. (b) Complete system, partially normalized. (c) Fully normalized.

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

$$v_m = \frac{v_m}{\sqrt{2Z_0 P_m}}, \quad v_e = \frac{v_e}{\sqrt{2Z_0 P_m} G_1 A_M C_M A_c C_e}, \quad v_o = \frac{-v_o}{\sqrt{2Z_0 P_m} G_1 G_2 A_E C_o}. \quad (11)$$

so that the squared magnitude

$$|v_m|^2 = \frac{|v_m|^2}{2Z_0 P_m} = \frac{P_m}{P_m} = x_m \quad (12)$$

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

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

featuring the main amplifier's reference input drive

$$D_{0M} = \frac{P_m |G_1|^2}{P_{\text{sat}}}. \quad (14)$$

Given that $v_M = v_m M_1 f_M(d_M(x_m, D_{0M}, M_1))$, the instantaneous fractional imbalance in the first loop and the instantaneous error power (normalized) may be obtained as

$$\epsilon_1(x_m, D_{0M}, M_1) = \frac{v_e}{v_m} = \frac{v_M - v_m}{v_m} = M_1 f_M(d_M(x_m, D_{0M}, M_1)) - 1. \quad (15)$$

$$x_e(x_m, D_{0M}, M_1) = |v_e|^2 = |v_m \epsilon_1|^2 = x_m |\epsilon_1(x_m, D_{0M}, M_1)|^2. \quad (16)$$

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

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

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.

$$E[v_e v_m^*] = E[v_m \epsilon_1 v_m^*] = E[x_m \epsilon_1(x_m, D_{0M}, M_1)] = X_{em}(D_{0M}, M_1). \quad (18)$$

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

$$M_1(D_{0M}, \phi_1) = M_{10}(D_{0M}) + |M_{10}(D_{0M})| \delta_1 e^{j\theta_1}, \quad (19)$$

from which follow the functions

$$d_M(x_m, D_{0M}, \phi_1), \quad \varepsilon_1(x_m, D_{0M}, \phi_1), \quad x_e(x_m, D_{0M}, \phi_1), \quad X_e(D_{0M}, \phi_1), \quad X_{em}(D_{0M}, \phi_1). \quad (20)$$

D. Balancing the Second Loop

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

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

featuring the error amplifier's reference input drive

$$D_{0E} = \frac{P_m |G_1 G_2|^2}{P_{satE}}. \quad (22)$$

Given that $v_E = v_e M_2 f_E(d_E)$, the instantaneous fractional imbalance in the second loop may be obtained as

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

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

$$E[|v_r|^2] = E[|v_e \varepsilon_2|^2] = E[x_e(x_m, D_{0M}, \phi_1) |\varepsilon_2(x_m, D_{0M}, \phi_1, D_{0E}, M_2)|^2] = X_r(D_{0M}, \phi_1, D_{0E}, M_2), \quad (24)$$

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

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

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. Refs. [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.

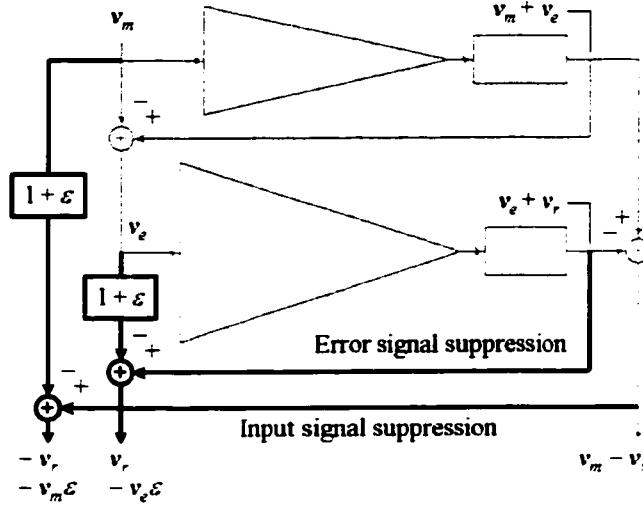


Fig. 3. Third circuit loop added to Fig. 2(c) for memoryless and pilotless adaptation of second loop with error or input signal suppression.

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 non-zero ϵ and error signal suppression, the expressions respectively minimized and nulled by M_2 in the power and gradient methods are now

$$E[|v_r - v_e \epsilon|^2] = X_r - 2 \operatorname{Re}\{X_{re} \epsilon^*\} + X_e |\epsilon|^2. \quad (26)$$

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

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 $v_E = v_e(1 + \epsilon)$ with either adaptation method, leaving an incompletely cancelled error signal $v_r = v_e \epsilon$ at the output. Now with a non-zero ϵ and input signal suppression, and using (1) and

$$\begin{aligned} E[v_r v_m^*] &= E[v_e \epsilon_2 v_m^*] = E[v_m \epsilon_1 \epsilon_2 v_m^*] \\ &= E[x_m \epsilon_1(x_m, D_{0M}, \phi_1) \epsilon_2(x_m, D_{0M}, \phi_1, D_{0E}, M_2)] = X_{rm}(D_{0M}, \phi_1, D_{0E}, M_2), \end{aligned} \quad (28)$$

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

$$E[|v_r + v_m \epsilon|^2] = X_r + 2 \operatorname{Re}\{X_{rm} \epsilon^*\} + |\epsilon|^2. \quad (29)$$

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

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 $X_{cm} = 0$. Then the output signal $v_o = v_m - v_r$ itself can 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_{cm} that is not entirely realistic. It actually prohibits the use of the power minimization method in the first loop, which does not make X_{cm} 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

$$M_2(D_{0M}, \phi_1, D_{0E}, \phi_2) = M_{20}(D_{0M}, \phi_1, D_{0E}) + |M_{20}(D_{0M}, \phi_1, D_{0E})| \delta_2 e^{j\theta_2}. \quad (31)$$

from which follow the functions

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

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

$$v_o = v_m - v_r = v_m G_S + v_{IM} \quad (33)$$

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

$$X_o = E[|v_m G_S + v_{IM}|^2] = |G_S|^2 + X_{IM}. \quad (34)$$

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

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

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

$$SIMR = \frac{|G_S|^2}{X_{IM}} = \frac{|X_{rm}|^2}{X_o - |X_{rm}|^2}. \quad (36)$$

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

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

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

Eqns. (33) and (37)-(38) agree that if v_r and v_m are already uncorrelated, $G_s = 1$ and $X_{IM} = X_r$. *SIMR* follows as

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

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-dimensional 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

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

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 $U = |v_o/v_m|^2$ yields

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

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

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

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

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

featuring the average dc powers normalized to the saturated output powers.

$$H_M(D_{0M}, D_{0E}) = E[h_M(d_M(x_m, D_{0M}, D_{0E}))], \quad H_E(D_{0M}, D_{0E}) = E[h_E(d_E(x_m, D_{0M}, D_{0E}))] \quad (44)$$

and the normalized saturated output powers

$$X_{satM} = \frac{P_{satM}}{P_s}, \quad X_{satE} = \frac{P_{satE}}{P_s}. \quad (45)$$

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

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

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 P_{satM} and 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 through K denote the power gains along the S-shaped path containing the two amplifiers.

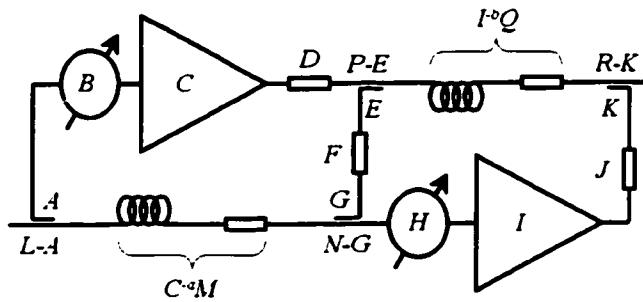


Fig. 4. Feedforward amplifier model with real power gains of components.

Comparing with Fig. 1,

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

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

$$|T_m|^2 = L - A, \quad |T_M|^2 = P - E, \quad |T_e|^2 = N - G, \quad |T_o|^2 = R - K. \quad (48)$$

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

$$q_1(-10\log |L_m|^2) = p_1(10\log C) + \tau_1. \quad (49)$$

Therefore, the composite power gain in $L_m A_m$, and likewise for $L_e A_e$, can be expressed as

$$|L_m A_m|^2 = C^{-a} M, \quad |L_e A_e|^2 = I^{-b} Q, \quad (50)$$

where

$$a = p_1/q_1, \quad b = p_2/q_2, \quad (51)$$

$$M = 10^{-\tau_1/10q_1} |A_m|^2, \quad Q = 10^{-\tau_2/10q_2} |A_c|^2 \quad (52)$$

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

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

into (46) yields the normalized saturated output powers

$$X_{satM} = \frac{1}{D_{0M} D(P-E) I^{-b} Q(R-K) V}, \quad X_{satE} = \frac{1}{D_{0E} J K V}. \quad (54)$$

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}) , $SIMR = \{SIMR_{req}\}$ is not yet an optimization constraint. The three constraints at this point are the required power gain balance in the two loops, as well as $U = U_{req}$. Eliminating $|G_1|^2$ and $|G_2|^2$ respectively from the power equivalents of (9) and (10), and expanding $U = U_{req}$ using (41) and (10), the three constraints become

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

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

$$ABCDEFGHIJK = STW. \quad (57)$$

where

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

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

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

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)–(56) into (57), yielding the single constraint

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

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

$$X_{satM} = \frac{1}{D_{0M} D} \left[\left(\frac{T}{EFGHJK} \right)^b \frac{1}{(P-E)Q(R-K)} \right]^{\frac{1}{1-b}} \frac{1}{V}, \quad X_{satE} = \frac{1}{D_{0E} J K V}. \quad (62)$$

Any solution that satisfies (61) can be updated by bringing A closer to $L/(1+1/a) = L/(1+q_1/p_1)$ and then by raising G as near N as necessary to restore the broken equality in (61). The second change decreases X_{out} in (62) so as to increase the efficiency in (43). It follows that A has the non-trivial optimal value of $L/(1+q_1/p_1)$, 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

$$cEG = N - G, \quad (63)$$

$$dEGK = (P - E)(R - K), \quad (64)$$

$$eEGK = 1 \quad (65)$$

where

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

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

$$E = \frac{-a_1 - \sqrt{a_1^2 - 4a_2a_0}}{2a_2}, \quad (67)$$

$$a_2 = eNR - c, \quad a_1 = dN - a_2P - 1, \quad a_0 = P. \quad (68)$$

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

$$X_{\text{out}M} = \frac{AB}{D_{0M}SU_{\text{req}}} C, \quad X_{\text{out}E} = \frac{1}{D_{0E}JV} \frac{1}{K}. \quad (69)$$

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 non-analytic 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. 10(a). Then the maximum of efficiency for a given $SIMR$ can be located by tracking η along the relevant $SIMR$ contour. The values of the seven design parameters $C = |G_M|^2$, $I = |G_E|^2$, $E = |C_M|^2$, $G = |C_e|^2$, $K = |C_o|^2$, $X_{\text{out}M}$, and $X_{\text{out}E}$ 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 $SIMR = \{SIMR_{\text{req}}\}$, as in Fig. 10(a), the behaviors of the optimized variables as a function of $SIMR$ are best conveyed in the type of plot shown in Fig. 10(b), which pictures the seven design parameters in dB and η in percent versus $SIMR$ in dB. By convention, couplings are represented by positive dB 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.

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. 5(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 Figs. 5(b) and 5(c), where the x_m and ccdf axes are linear and have been swapped. In Fig. 5(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 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. 5(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.

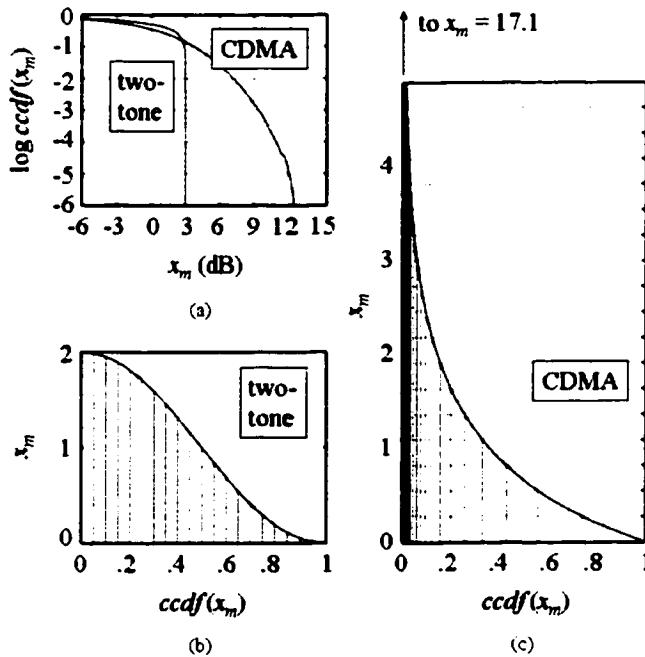


Fig. 5. Ccdfs of a two-tone signal and a three-carrier IS-95A CDMA signal.
(a) On the same standard logarithmic plot. (b)-(c) On separate linear plots
and sampled horizontally and vertically, respectively.

B. Amplifier Models

Fig. 6 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.

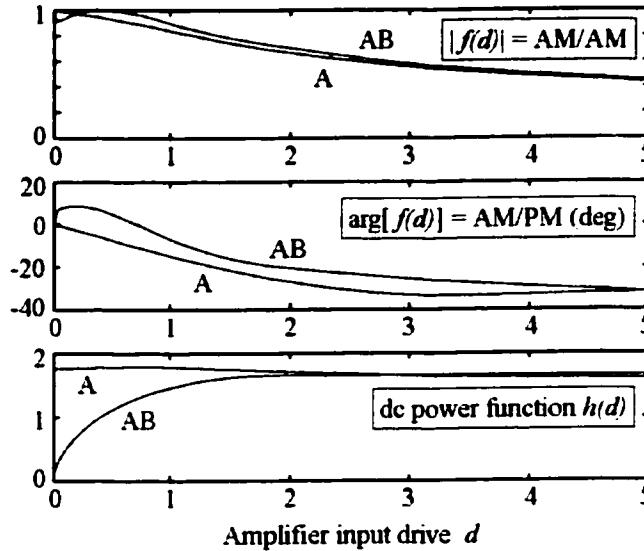


Fig. 6. Gain compression function $f(d)$ and dc power function $h(d)$ for typical Class AB main amplifier (measured) and Class A error amplifier (simulated).

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 ACPR (Adjacent Channel Power Ratio) 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. 7(a). Due to some residual frequency response over the measurement bandwidth, the ACPR values for the lower sideband are about 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 $P_{\text{satM}} = 54.33$ dBm and the same output power. Using $P_M = E[|v_M|^2 / 2Z_0]$, $v_M = v_m / (\sqrt{2Z_0 P_m} G_1) = v_m M_1 f_M(d_M)$, and Eqns. (13)-(14), the output power of the amplifier model is given by

$$P_M = D_{0M} |M_1|^2 E[x_m |f_M(x_m D_{0M} |M_1|^2)|^2] P_{\text{satM}} \quad (70)$$

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

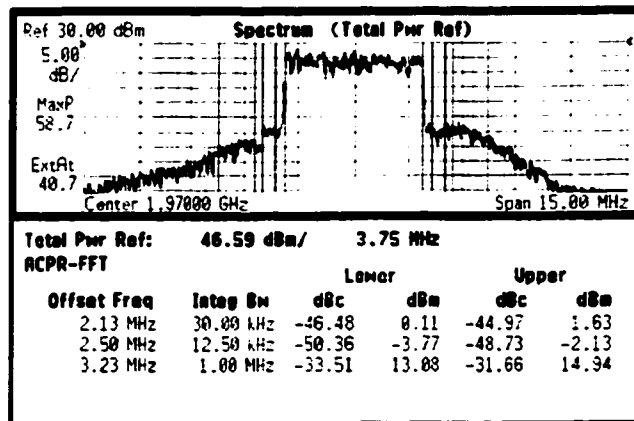
$$v_M / M_1 = v_m f_M(d_M) = v_m f_M(x_m D_{0M} |M_1|^2) \quad (71)$$

was computed and converted to the power spectrum of Fig. 7(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.

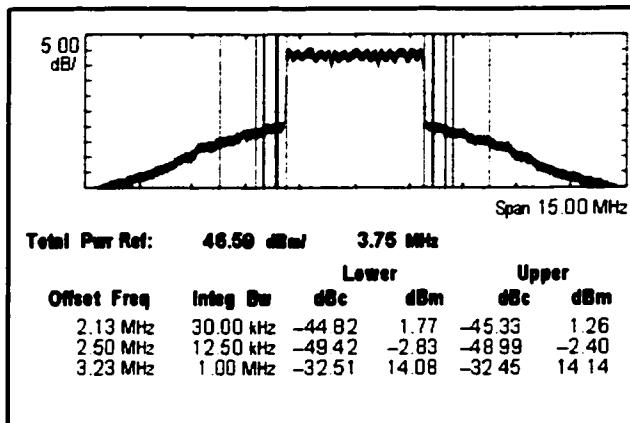
The SIMR of v_M is denoted $SIMR_M$, and using (36) with the proper subscripts, is obtained from

$$\frac{1}{SIMR_M} = \frac{|X_M - |X_{AM}|^2|}{|X_{AM}|^2} = \frac{E[x_m | f_M(d_M)]|^2}{|E[x_m f_M(d_M)]|^2} - 1. \quad (72)$$

A value of $SIMR_M = 19.79$ 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 about 24 dB found in Fig. 7(b).



(a)



(b)

Fig. 7. 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.

C. Adaptation Methods

With 402 x_n 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 $M_{10} = 1.0434 \angle -4.68^\circ$. The power minimization method brings a relative change of 3.9% $\angle -168^\circ$ to this value, causing the error power X_e to drop from 0.0126 to its minimum of 0.011 but raising $|X_{cm}|^2$ from zero to 0.00125. With power minimization, the average input drive $D_{0A}/|M_1|^2$ works out to the earlier value of -7.419 dB and hence $SIMR_M = 19.79$ 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. 8 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 about 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.

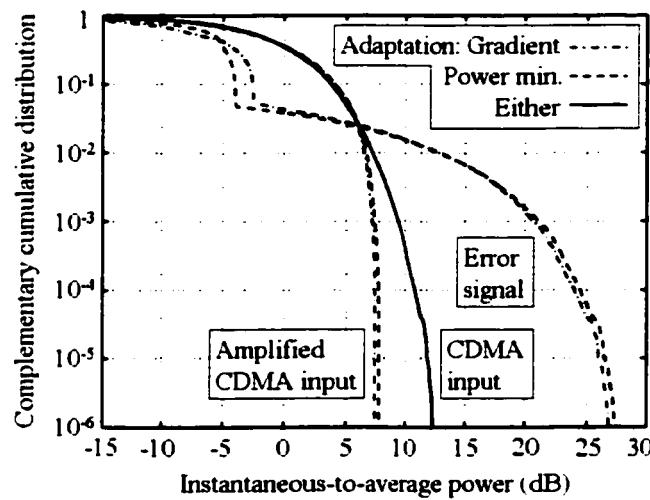


Fig. 8. Ccdfs of the signals in the first loop, assuming the gradient and power minimization adaptation methods with $D_{0M} = -7.4481$ dB

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, $SIMR = 59.63$ dB is obtained. However, with the power method in the first loop and the consequent 3.9% change in M_1 , $SIMR = 8.9115$ 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 non-linear 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.

1st loop	Gradient	Power min.
2nd loop		
$\varepsilon = -1$, 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

Table 1. 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.

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 the paper. Then the present example has values of M_1 and M_2 such that

$$|M_1|^2 = 1.0067, \quad |M_2|^2 = 1.0069, \quad |G_S|^2 = 1.0004, \quad H_M = 0.5883, \quad H_E = 1.7564. \quad (73)$$

Assuming now the following realistic specifications for a feedforward circuit,

$$\begin{aligned} U_{req} &= 35 \text{ dB}, \quad p_1 = p_2 = 0.15 \text{ ns/dB}, \quad \tau_1 = \tau_2 = 1.5 \text{ ns}, \quad q_1 = q_2 = 15 \text{ ns/dB}, \\ L &= P = N = R = -0.2 \text{ dB}, \quad |C_m|^2 = L/(1 + q_1/p_1) = -20.243 \text{ dB}, \quad |A_c|^2 = 0 \text{ dB}, \\ |A_m|^2 &= |A_e|^2 = -0.2 \text{ dB}, \quad |A_M|^2 = |A_E|^2 = -0.3 \text{ dB}, \quad |M_m|^2 = |M_e|^2 = -16 \text{ dB}. \end{aligned} \quad (74)$$

Fig. 9 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 $X_{outM} = 15.53$ dB and $X_{outE} = 16.22$ 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 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 $X_{outM} = 15.65$ dB and $X_{outE} = 16.24$ 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 by 26 is pictured in Fig. 10(a), on which are plotted contours of SIMR from 51 to 69 dB and contours of η .

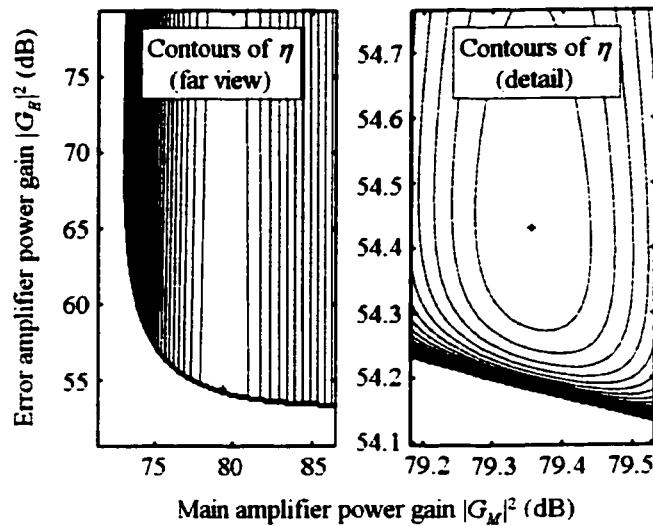


Fig. 9. Contours of power efficiency η as a function of the main and error amplifier gains for $D_{0M} = -7.4481$ dB and $D_{0E} = -14.6295$ dB. Far view (left), and detail around the point of maximum efficiency (right).

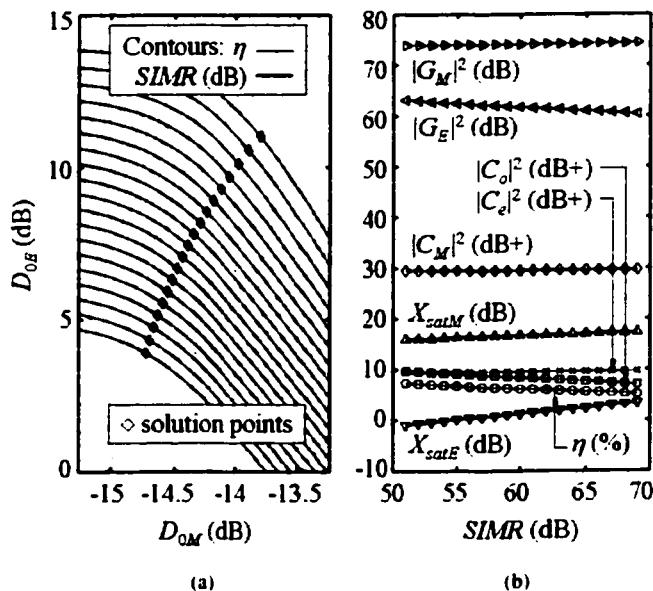


Fig. 10. (a) Contours of output SIMR and power efficiency in the D_{0M} - D_{0E} plane, and points of optimal efficiency. (b) Corresponding optimal values of amplifier gains, couplings, normalized saturated output powers, and efficiency.

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. 10(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, $X_{out} = 16.73$ dB, and $X_{outE} = 1.25$ 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 six-fold improvement in efficiency. Inserting the updated value of $D_{0M}|M_1|^2$ in (72) yields $SIMR_M = 29.44$ dB instead of 19.79 dB, and the resulting error signal is found to have a crest factor of about 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. 10(a) now yields a worst-case $SIMR$ of 56.64 dB with $\phi_1 = -38.1^\circ$ and $\phi_2 = 114^\circ$. 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 .

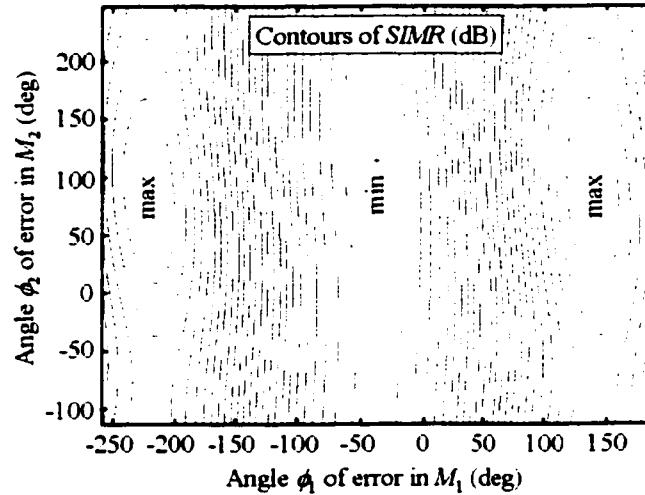


Fig. 11. Contours of output SIMR as a function of the angles of the 1 % errors in each of the vector modulator gains, assuming $D_{0M} = -14.396$ dB and $D_{0E} = 7.4326$ dB. Minimum SIMR is the worst-case scenario.

The complete simulation of Fig. 10 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 coupling, and the efficiency. 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.

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 times higher, due to required saturated output powers about 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.

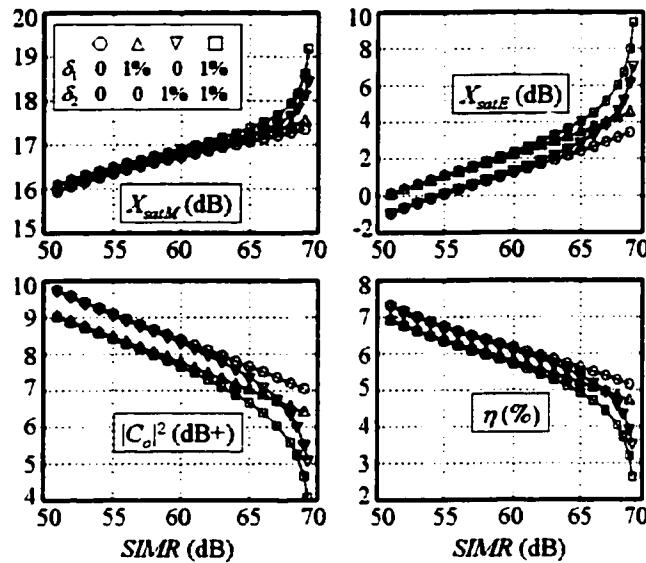


Fig. 12. Optimal values of normalized saturated output powers, output coupling, and power efficiency vs. output SIMR for none, either, or both tolerances of 1 % in the vector modulator gains.

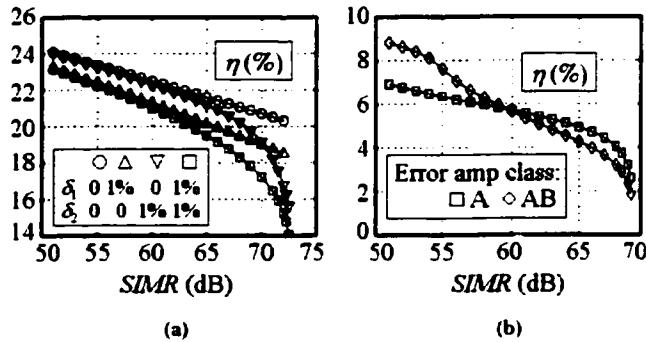


Fig. 13. (a) Equivalent of efficiency curves in Fig. 12 for a two-tone test signal instead of a three-carrier IS-95A CDMA signal. (b) Optimal power efficiency versus output SIMR with Class A and Class AB error amplifier. Tolerances of 1 % in both vector modulator gains are assumed.

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. 6 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 was 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 narrowband by nature. However, under relatively broadband 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 re-examination 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 six-fold 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 Class A 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.

ACKNOWLEDGEMENT

The authors wish to thank Mrs. G. Zhao of AmpliX Inc. 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 Vehicular Technol. Conf.*, 1989, vol. 1, pp. 57-62.
- [2] A. Wright and W. Durtler, "Experimental performance of an adaptive digital linearized power amplifier," *IEEE Trans. Veh. Technol.*, vol. 41, no. 4, 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. 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 Journal*, 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, no. 2, 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 and Microwave CAD*, 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, no. 1, 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.*, 2000, vol. 3, pp. 1491- 1494.
- [10] C. L. Larose and F. M. Ghannouchi, "First-Loop control in feedforward amplifiers for high-stress communications signals", in *Proc. Int. Symp. on Microwave and Optical Technol.*, 2001, pp. 195-198.
- [11] C. L. Larose and F. M. Ghannouchi, "Pilotless adaptation of feedforward amplifiers driven by high-stress signals," in *Proc. IEEE Radio and Wireless Conf.*, 2001, pp. 81-84.
- [12] C. L. Larose and F. M. Ghannouchi, "Optimal amplifier classes and adaptation methods for feedforward amplifiers driven by high-stress signals," *IEEE Trans. Veh. Technol.*, to be published.
- [13] PULSAR MICROWAVE Corp., NJ. Directional coupler specifications. [Online]. Available: <http://www.pulsarmicrowave.com/products/couplers/couplers.htm>
- [14] SYNERGY MICROWAVE Corp., NJ. Couplers: Parameter definitions. [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] AGILENT TECHNOLOGIES, CA. Three carrier IS-95A CDMA waveform for the Agilent ESG-D series. [Online]. Available: <http://www.agilent.com>. Quick Search on "DAB".
- [20] AGILENT TECHNOLOGIES, CA. "Performing cdma2000 measurements today," Application Note 1325, May 2000.

Colin L. Larose received the B.Eng. and M.Eng. degrees in electrical engineering from Concordia University, Montréal, QC, Canada in 1983 and 1986 respectively. After receiving a Bachelor of Arts in 1989 in California, he worked for five years in the area of radar cross-section measurement and simulation at the David Florida Laboratory of the Canadian Space Agency in Ottawa, ON, Canada. He has been the recipient of a research grant in 1985 and a Postgraduate Scholarship in 1995 from the Natural Sciences and Engineering Research Council of Canada (NSERC).

He is currently completing his Ph.D. degree at the École Polytechnique de Montréal, Montréal, QC, Canada. His fields of interest include feedforward and other linearization techniques, wireless communications, and measurement and simulation techniques for antennas and radar.

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 et de génie informatique, 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, non- linear 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.

ANNEXE II

DEUXIÈME ARTICLE:
« OPTIMAL AMPLIFIER CLASSES AND ADAPTATION METHODS
FOR FEEDFORWARD AMPLIFIERS DRIVEN BY HIGH-STRESS SIGNAL »

Cet article a été soumis le 10 mai 2002 au *IEEE Transactions on Vehicular Technologies*.

Optimal Amplifier Classes and Adaptation Methods for Feedforward Amplifiers Driven by High-Stress Signals

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

Abstract—This paper introduces two new adaptation methods for the feedforward amplifier, and systematically studies the influence of the adaptation and of the main amplifier class of operation on the achievable system power efficiency under typical design constraints. These two influences are weak for a two-tone driving signal but interact powerfully for a real-life high-stress signal, leading to optimal designs involving Class B or even Class C main amplifiers, depending on the transistors employed and the required output linearity. Assuming a Code Division Multiple Access (CDMA) signal, typical MESFETs, and an output linearity commonly used in mobile systems design, the new adaptation methods allowed relative efficiency improvements of 21 and 36 percent using main amplifiers biased near Class B and at Class B, respectively.

Index terms—microwave power amplifiers, linearization, power efficiency, adaptive control, code division multiaccess

I. INTRODUCTION

There is considerable industrial interest in producing microwave amplifiers with good linearity and power efficiency. These two contradictory requirements can be reconciled by using external circuitry to linearize an efficient amplifier. Relatively simple linearization methods have been used [1]-[2], but the use of more broadband techniques, such as analog (RF) predistortion [3] and feedforward, has been made necessary by the emergence of the new CDMA digital cellular system. The feedforward linearizer is sensitive to changes in operating conditions, but the recent development of adaptation methods to compensate for such changes has renewed interest in the technique. Feedforward is more effective than predistortion, but it has relatively modest efficiency due to the use of an error amplifier and compensating delay line.

Early studies on the power efficiency of feedforward amplifiers [4]-[6] made various simplifying assumptions, notably that of a two-tone test signal, which has much smaller envelope fluctuations than today's real-life CDMA signals. Ref. [7] describes a method of optimizing the dc-RF conversion efficiency of a feedforward amplifier for any given driving signal, subject to a required linearity and gain at the linearizer's output. The gains and saturated output powers of the main and error amplifiers and the various couplings are optimized, but the remaining amplifier characteristics and the adaptation method must be specified at the outset. The optimization is therefore really a partial one, and many such optimizations using different amplifier characteristics and adaptation schemes are required to identify a final optimal design. The scope of [7] could only afford a glimpse into this type of activity, and Refs. [8]-[9] only proposed a new adaptation method. The present paper undertakes a systematic search of many available options. A typical large-signal transistor model is used to compute amplifier characteristics ranging from Class A to Class C. The error amplifier is specified as usual as Class A, but the main amplifier is swept in class of operation, each class offering a unique trade-off between linearity and efficiency. Three adaptation methods are also investigated, including two new methods of which one has never been reported. All optimizations

This work was supported in part by a Postgraduate Scholarship from the Natural Sciences and Engineering Research Council of Canada (NSERC).

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

are performed with the assumptions of two-tone and CDMA input signals, a wide range of required linearities at the output, and typical design constraints. For a CDMA signal, optimizations are also repeated with a transistor device of more constant transconductance.

II. BACKGROUND

This Section summarizes prerequisite material from [7].

A. Normalized Feedforward Amplifier Model

The procedure in [7] uses the behavioral-level technique of quadrature modeling [10], which is computationally efficient as it uses the complex envelope of a signal, or the complex baseband signal, instead of the real signal. The complex baseband model of the feedforward amplifier shown in Fig. 1(a) 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 $v_m(t)$ to the main amplifier, which outputs an amplified input signal $v_M(t)$ containing IM distortion. A small portion of this signal is sent to be compared to the remaining portion of $v_m(t)$ to produce an error signal $v_e(t)$. This closes the circuit's first loop, or signal cancellation loop. The error signal $v_e(t)$ exits the error amplifier as the amplified error signal $v_E(t)$ after suffering its share of IM distortion. This signal is sent to an output coupler, with subtracts it from the main portion of $v_M(t)$ to yield a largely corrected output signal $v_o(t)$. This closes the circuit's second loop, or error cancellation loop. To balance propagation delays within each loop, delay lines of gains L_m and L_e are added to compensate for the delays of the amplifiers. Gain balance must also be maintained in each loop to ensure precise signal cancellations. For this purpose, any drifting in the amplifiers is automatically compensated for by suitable adaptation of vector modulator gains M_m and M_e . To simplify the analysis, authors sometimes place the vector modulators 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,

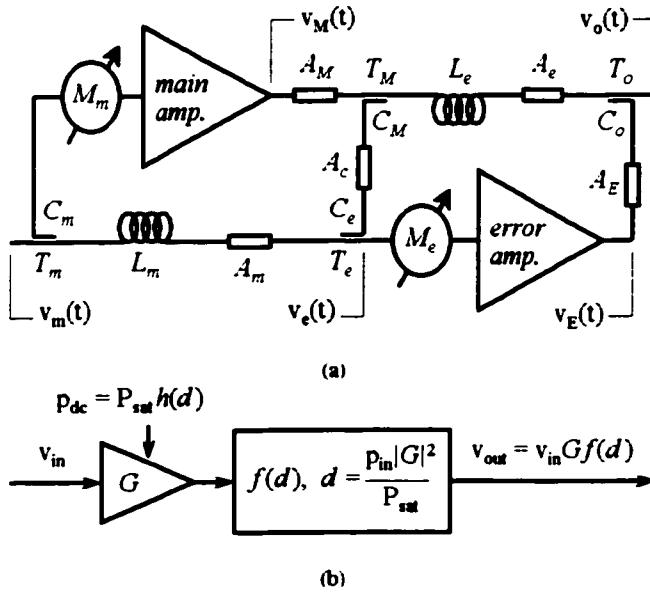


Fig. 1. (a) Complex baseband feedforward amplifier model, with complex envelopes of signals and complex gains of components. (b) Complex baseband power amplifier model.

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 to account for other possible attenuations in the circuit.

In Fig. 1(b), a power amplifier is modeled as a complex gain G followed by a gain compression function $f(d)$ where d is the instantaneous input drive. A complex $f(d)$ allows one to specify both the amplifier's AM/AM characteristic $|f(d)|$ and its AM/PM characteristic $\arg[f(d)]$. The amplifier's dc power consumption is its saturated output power P_{sat} multiplied by a normalized dc power function $h(d)$. Functions f and h describe an amplifier (e.g. transistor type, class of operation) independently of its gain and saturated output power, which are to be optimized. Input drive d is the reciprocal of input backoff and is defined here as the ratio of the uncompressed output power to the saturated output power, or

$$d = \frac{P_{in} |G|^2}{P_{sat}} = \frac{P_{in}}{P_{sat} / |G|^2}. \quad (1)$$

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, but in Class C operation, 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 specifying $f(d)$ and $h(d)$.

The complex baseband power amplifier model is narrowband by nature, but it was found in the example of Fig. 2 to predict spectral regrowth well enough under relatively broadband CDMA excitation. Fig. 2(a) shows an ACPR (Adjacent Channel Power Ratio) measurement for a three-carrier CDMA signal exiting a typical Class AB main amplifier. Fig. 2(b) simulates the same measurement using the amplifier's AM/AM and AM/PM characteristics, measured in pulse mode to better reflect actual operating conditions. The simulated ACPR values listed are typically within 1 dB of the measured ones.

The model of Fig. 1(a) is completed by substituting Fig. 1(b) with G_M , P_{satM} , $f_M(d_M)$, and $h_M(d_M)$ for the main amplifier and with G_E , P_{satE} , $f_E(d_E)$, and $h_E(d_E)$ for the error amplifier. The model is then reduced to four components and normalized with the two normalization gains

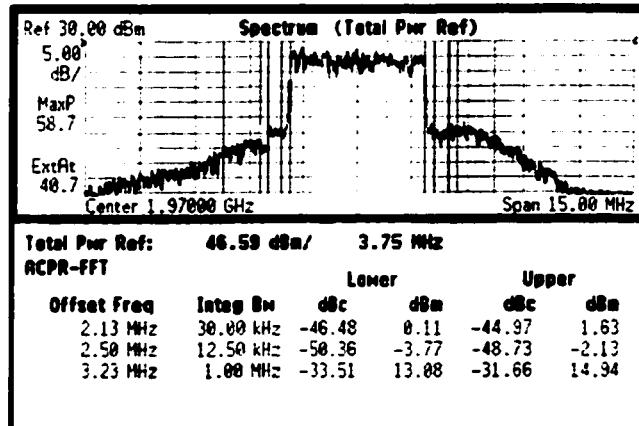
$$G_1 = -\frac{T_m L_m A_m T_c}{A_M C_M A_c C_e}, \quad G_2 = -\frac{A_M T_M L_c A_e T_n}{A_E C_o}. \quad (2)$$

as shown in Fig. 3. The original vector modulator gains are now normalized as M_1 and M_2 . Signals have been further normalized by the constant $\sqrt{2Z_0 P_m}$ where Z_0 is the reference impedance, and are now denoted using italics to distinguish them from their absolute counterparts. In this way, the squared magnitude

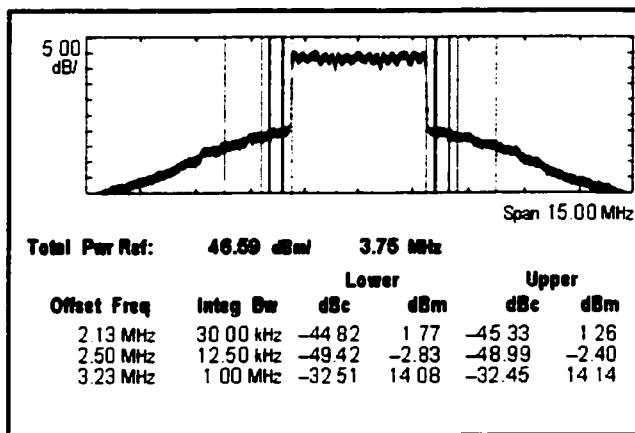
$$|v_m|^2 = \frac{|v_m|^2}{2Z_0 P_m} = \frac{p_m}{P_m} = x_m \quad (3)$$

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. The instantaneous input drives of the two amplifiers are

$$d_M = x_m D_{0M} |M_1|^2, \quad d_E = x_e D_{0E} |M_2|^2. \quad (4)$$



(a)



(b)

Fig. 2. ACPR measurement for a three-carrier CDMA signal exiting a typical Class AB amplifier. (a) Measured. (b) Predicted using measured AM/AM and AM/PM characteristics.

featuring $x_e = |v_e|^2$ and the amplifiers' reference input drives

$$D_{0M} = \frac{P_m |G_1|^2}{P_{\text{satM}}}, \quad D_{0E} = \frac{P_m |G_1 G_2|^2}{P_{\text{satE}}}. \quad (5)$$

With three circuit branches now having unity gain, 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. Also, the amplified input signal is 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 delivers an invariant gain normalized here to unity. The gain is unaffected by drifting in the amplifier gains, or 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.

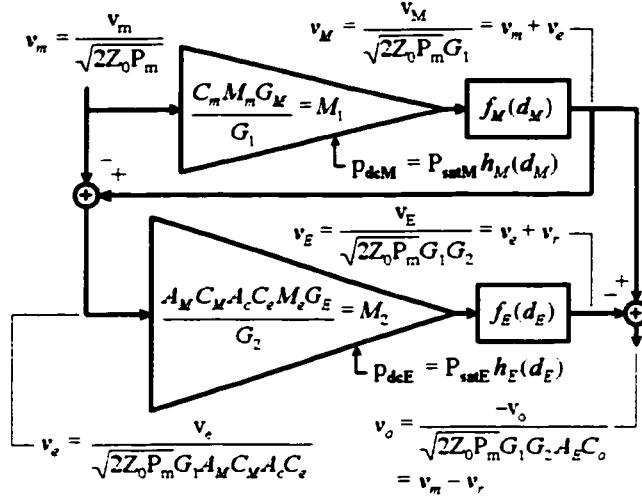


Fig. 3 Normalized complex baseband feedforward amplifier model, with complex signal envelopes and complex gains.

B. Output IMSR

Ref. [7] uses the IMSR, or ratio of IM distortion power to linear signal power, to summarize the output linearity of the feedforward amplifier and thus benefit from the convenience of a statistical analysis. As explained in [11], the IMSR 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. For a given input signal, the model of Fig. 3 will now be used to derive the output IMSR of the feedforward amplifier as a function of its design parameters.

Given $v_M = v_m M_1 f_M(d_M)$ and (4), the instantaneous fractional imbalance in the first loop and the instantaneous error power (normalized) may be obtained as

$$\varepsilon_1(x_m, D_{0M}, M_1) = \frac{v_e}{v_m} = \frac{v_M - v_m}{v_m} = M_1 f_M(x_m D_{0M} |M_1|^2) - 1, \quad (6)$$

$$x_e(x_m, D_{0M}, M_1) = |v_e|^2 = |v_m \varepsilon_1|^2 = x_m |\varepsilon_1(x_m, D_{0M}, M_1)|^2. \quad (7)$$

Next, given $v_E = v_e M_2 f_E(d_E)$ and (4), the instantaneous fractional imbalance in the second loop may be obtained as

$$\varepsilon_2(x_m, D_{0M}, M_1, D_{0E}, M_2) = \frac{v_r}{v_e} = \frac{v_E - v_e}{v_e} = M_2 f_E(x_e(x_m, D_{0M}, M_1) D_{0E} |M_2|^2) - 1. \quad (8)$$

The output IMSR is found by rewriting the output signal as

$$v_o = v_m - v_r = v_m G_S + v_{IM} \quad (9)$$

where the linear signal gain G_S minimizes the average power in the IM distortion v_{IM} . According to linear estimation theory [12], this also uncorrelates v_{IM} and v_m . Therefore, the total output power is the sum of the linear signal power and the IM power, or

$$X_o = E[|v_o|^2] = E[|v_m G_S|^2] + E[|v_{IM}|^2] = |G_S|^2 + X_{IM}. \quad (10)$$

given from (3) that $E[|v_m|^2] = E[p_m]/P_m = 1$. The IMSR is the ratio of the second power to the first, and is obtained in terms of v_o by using (9) in the correlation of v_{IM} and v_m .

$$X_{IM} = E[v_{IM} v_m^*] = E[(v_o - v_m G_S) v_m^*] = X_{rm} - G_S. \quad (11)$$

Since this expression must equal zero, $G_S = X_{rm}$, and combining this with (10), the IMSR is

$$IMSR = \frac{X_{IM}}{|G_S|^2} = \frac{X_o - |X_{rm}|^2}{|X_{rm}|^2} = \frac{X_o}{|X_{rm}|^2} - 1. \quad (12)$$

The IMSR is then easily obtained in terms of v_r . Replacing v_o with $v_m - v_r$ in $G_S = X_{rm}$.

$$G_S = E[(v_m - v_r) v_m^*] = 1 - X_{rm}. \quad (13)$$

and since from (9) $v_r = v_m(1 - G_S) - v_{IM}$, X_r is also the sum of two uncorrelated powers.

$$X_r = |1 - G_S|^2 + X_{IM}. \quad (14)$$

Using (13) and (14), the output IMSR is obtained as

$$IMSR(D_{0M}, M_1, D_{0E}, M_2) = \frac{X_{IM}}{|G_S|^2} = \frac{X_r - |X_{rm}|^2}{|1 - X_{rm}|^2} \quad (15)$$

where

$$\begin{aligned} X_r(D_{0M}, M_1, D_{0E}, M_2) &= \\ E[|v_r|^2] &= E[|v_r \epsilon_2|^2] = E[x_r(x_m, D_{0M}, M_1) |\epsilon_2(x_m, D_{0M}, M_1, D_{0E}, M_2)|^2], \end{aligned} \quad (16)$$

$$X_{rm}(D_{0M}, M_1, D_{0E}, M_2) = E[v_r v_m^*] = E[v_m \epsilon_1 \epsilon_2 v_m^*] = E[x_m \epsilon_1(x_m, D_{0M}, M_1) \epsilon_2(x_m, D_{0M}, M_1, D_{0E}, M_2)]. \quad (17)$$

Eqns. (9), (13), and (15) agree that if v_r and v_m are already uncorrelated, $G_S = 1$ and $IMSR = X_{IM} = X_r$.

The benefit of the IMSR is realized here in that the statistics of x_m are all that must be known of the input signal to evaluate the required expectancies. In terms of single variables, the output IMSR and its contributing terms are functions of only $(D_{0M}, M_1, D_{0E}, M_2)$.

III. AMPLIFIER CHARACTERISTICS

Two large-signal transistor models will now be used to simulate amplifiers ranging from Class A to Class C and hence compute their gain compression and dc power functions. Ref. [13] is used as a guide, except that the proposed MESFET behavioral model is replaced by the Statz MESFET model, and the method of generating AM/PM characteristics is modified to extend its usefulness beyond Classes A and AB.

A. Transistor Models

The Statz MESFET model produces very good fitting of the transconductance nonlinearities near pinch-off, making it a good choice for simulating amplifiers other than Class A [14]. The transistor is viewed as the gate-controlled drain current source

$$I_{ds}(V_{gs}) = \frac{\beta(V_{gs} - V_{po})^2}{1 + b(V_{gs} - V_{po})} \quad (18)$$

with zero output current below pinch-off and maximum deliverable current I_{ds} at zero gate voltage. By imposing the latter condition on (18), coefficient β is evaluated in terms of I_{ds} . The gate voltage and drain current can then be normalized because the amplifier characteristics sought are independent of gain and saturated output power. Hence, the normalized version of the model is

$$I(V) = \frac{(1+c)V^2}{1+cV} \quad (19)$$

where $V = 1 - V_{gs}/V_{po}$ is the fraction of the distance from pinch-off to saturation. $I = I_{ds}/I_{ds}$ is the fraction of the maximum drain current, and $c = -bV_{po}$ uniquely determines the aspect of the I - V characteristic, with $c = 0$ corresponding to square-law behavior. The example of a Statz model for the Fujitsu FLK202XV MESFET chip [14] had $b = 0.536$ and $V_{po} = -2.72$ V, hence $c = 1.4579$. Therefore $c = 1.5$ is used here as a typical value, and $c = 15$ is used to simulate a transistor of more constant transconductance. The I - V characteristics for both values of c are shown in the middle plot of Fig. 4 and saturate at unity as indicated by the symbol \circlearrowright .

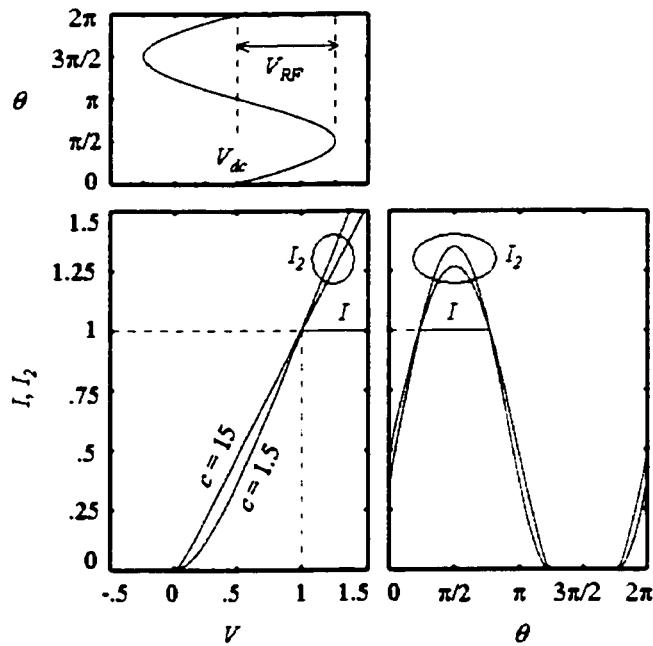


Fig. 4. Normalized Statz I_{ds} - V_{gs} characteristics for $c = 1.5$ and 15 , with and without saturation. Normalized voltage and current cycles are shown in an overdriven case.

B. Amplifier Simulations

For a given gate bias voltage V_{dc} and excitation voltage V_{RF} , the gate voltage and drain current waveforms $V(\theta)$ and $I(\theta)$ may be expressed as

$$V(\theta) = V_{dc} + V_{RF} \sin \theta, \quad I(\theta) = I(V(\theta)). \quad (20)$$

These are shown in Fig. 4 for one cycle in an overdriven case. Fourier analysis obtains the dc and fundamental components of the drain current as

$$I_{dc} = \frac{1}{2\pi} \int_0^{2\pi} I(\theta) d\theta = E[I(\theta)], \quad (21)$$

$$I_{RF} = \frac{1}{\pi} \int_0^{2\pi} I(\theta) \sin \theta d\theta = 2E[I(\theta) \sin \theta]. \quad (22)$$

Using 512 samples from $\theta = 2\pi/512$ to 2π , I_{dc} and I_{RF} were estimated over a wide range of V_{RF} for 52 distinct V_{dc} values. Class A operation was simulated with $V_{dc} = 0.5$, and Class AB through Class C was simulated with $V_{dc} = 0.25$ to -0.25 in steps of 0.01. Fig. 5 shows the results for both values of c and with steps of 0.05 in V_{dc} . Class B curves at $V_{dc} = 0$ are highlighted in bold, and the fundamental component of drain current saturates as indicated by the symbol $I_{RF,2}$.

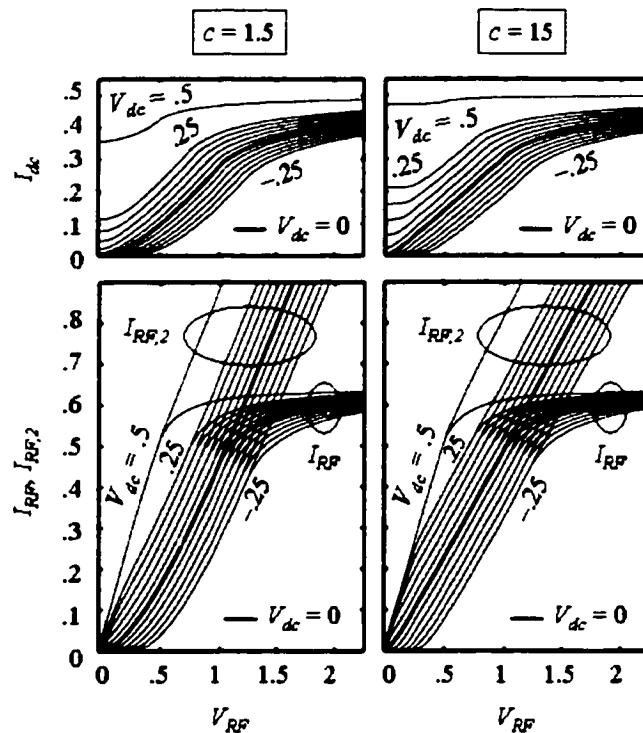


Fig. 5. Dc and fundamental components of drain current versus excitation voltage for normalized Statz models with $c = 1.5$ and 15 and with V_{dc} in steps of 0.05. Class B curves at $V_{dc} = 0$ are highlighted in bold.

For the purposes of the simulation procedure to be used, V_{RF} , I_{dc} , and I_{RF} must be converted to the input drive d , the gain compression function f , and the dc power function h . This exercise begins with the uncompressed I_{RF} , defined for convenience as V_{RF} times the transconductance $g_m = dI/dV$ at $V = 0.5$, independently of the gate bias voltage V_{dc} , or

$$I_{RF,0} = g_m|_{V=0.5} V_{RF} = \begin{cases} 1.1224V_{RF}, & c = 1.5 \\ 1.0519V_{RF}, & c = 15 \end{cases} \quad (23)$$

Saturation occurs when the drain current is a square wave [15], in which case its fundamental component is $I_{RF,sat} = 2/\pi$. The input drive is the uncompressed output power relative to the saturated output power, or

$$d = \frac{I_{RF,0}^2 R_L / 2}{I_{RF,sat}^2 R_L / 2} = \begin{cases} 3.1087V_{RF}^2, & c = 1.5 \\ 2.7302V_{RF}^2, & c = 15 \end{cases} \quad (24)$$

The magnitude of function f is I_{RF} relative to $I_{RF,0}$, or

$$|f| = \frac{I_{RF}}{I_{RF,0}} = \begin{cases} 0.8909I_{RF}/V_{RF}, & c = 1.5 \\ 0.9507I_{RF}/V_{RF}, & c = 15 \end{cases} \quad (25)$$

Ref. [13] proposes that phase shift be proportional to the dB gain compression attributed to saturation, up to a moderate compression depth of 8 dB. This compression is estimated here by re-computing the fundamental component of drain current with the saturation effect removed. For this purpose, the I - V characteristic in the saturation region is replaced by a straight line tangent at $V = 1$. For both values of c , the updated I - V characteristic $I_2(V)$ is added in Fig. 4 along with the corresponding current cycle in the overdriven case, and the curves of the updated fundamental component of drain current $I_{RF,2}$ are added in Fig. 5. As the latter curves become straight lines, the corresponding gains converge as required by [13]. Assuming a typical sensitivity of five degrees per dB of compression for all compression depths, the phase shift in degrees is

$$\arg[f] = 5 \cdot 20 \log \frac{I_{RF,2}}{I_{RF}} = 100 \log \frac{I_{RF,2}}{I_{RF}} \quad [\text{deg}]. \quad (26)$$

A load selected for power output has nominal impedance $R_L = 2(V_S - V_K)/I_{ds}$ where V_S is the drain bias or supply voltage and V_K is the knee voltage. Since in absolute terms the dc component of drain current is $I_{dc} = I_{dc} I_{ds}$ and its saturated fundamental component is $I_{RF,sat} = I_{RF,sat} I_{ds}$, the dc power function is

$$h = \frac{P_{dc}}{P_{sat}} = \frac{V_S I_{dc}}{I_{RF,sat}^2 R_L / 2} = \left(\frac{V_S}{V_S - V_K} \right) \frac{\pi^2}{4} I_{dc}. \quad (27)$$

Using $V_S = 5$ V throughout, as in [13], and the typical $V_K = 1$ V, the dc power function for both values of c becomes

$$h = \frac{5\pi^2}{16} I_{dc} = 3.0843 I_{dc}. \quad (28)$$

Using (24)-(26) and (28), the curves of Fig. 5 translate into the curves of $|f(d)|$, $\arg[f(d)]$, and $h(d)$ shown in Fig. 6.

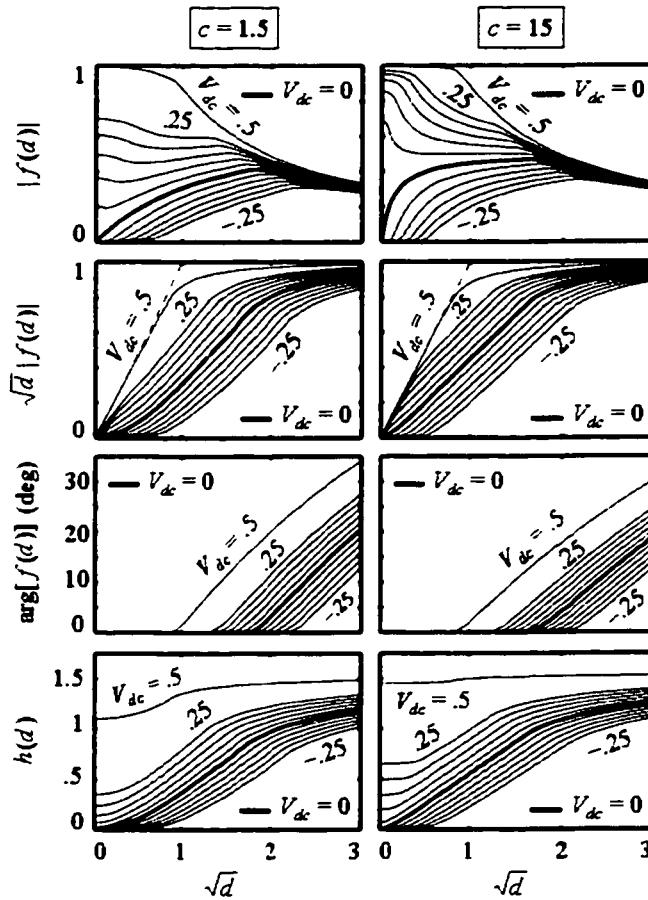


Fig. 6. Simulated amplifier characteristics for normalized Statz models with $c = 1.5$ and 15 and with V_{dc} in steps of 0.05 . Class B curves at $V_{dc} = 0$ are highlighted in bold.

The curves are plotted with respect to \sqrt{d} to show more detail in the small-signal region but also to allow the inclusion of the curves of $\sqrt{d}|f(d)|$ versus \sqrt{d} . These are the curves of amplified input signal magnitude $|v_M|$ versus input signal magnitude $|v_m|$ when $D_{0M} = |M_1| = 1$, or the curves of $|v_E|$ versus $|v_e|$ when $D_{0E} = |M_2| = 1$. To prove the former, the relation $v_M = v_m M_1 f_M(d_M)$ combined with (4) produces

$$|v_M \sqrt{D_{0M}}| = |v_m \sqrt{D_{0M}} M_1| |f_M(|v_m \sqrt{D_{0M}} M_1|^2)|, \quad (29)$$

and therefore $|v_m| = \sqrt{d_M}$ and $|v_M| = \sqrt{d_M} |f_M(d_M)|$ if $D_{0M} = |M_1| = 1$. If $|M_1|$ is not unity, inspection of (29) reveals that the $|v_m| - |v_M|$ curves are still the $\sqrt{d}|f(d)|$ curves of Fig. 6 but compressed toward the $|v_M|$ axis by the factor $|M_1|$, thus changing the gains of the curves but not their saturation. If on the other hand D_{0M} is not unity, the $|v_m| - |v_M|$ curves are the $\sqrt{d}|f(d)|$ curves compressed toward both axes by the factor $\sqrt{D_{0M}}$, thus making the curves saturate to $1/\sqrt{D_{0M}}$ without changing their gains.

IV. ADAPTATION METHODS

This section is devoted to adaptation of the feedforward amplifier. The quadrature modeling technique of [7] cannot account for adaptation methods that use filtering because it assumes frequency responses to be flat over the simulation bandwidth. Furthermore, in order to benefit from the convenience of a statistical analysis, adaptation methods must use samples of existing signals in the circuit and may not be assisted by a pilot signal. This will provide an opportunity for a fresh re-examination of pilotless adaptation methods.

A. Overview

The various adaptation schemes that can be devised differ in the criteria they employ to adjust M_1 and M_2 . The ideal adaptation scheme would employ the bottom-line output IMSR of (15) as its only criterion and therefore resort to a four-dimensional search on the two complex vector modulator gains. However, traditional schemes have relied on separate criteria for adjusting M_1 , thus resorting to a pair of well-behaved two-dimensional searches. Criteria for M_1 minimize in different ways the error signal, as this is thought to best serve the objective of minimizing the output IMSR. In the power minimization method, M_1 is searched in order to minimize the average error power

$$E[x_e(x_m, D_{0M}, M_1)] = X_e(D_{0M}, M_1). \quad (30)$$

The circuitry to do so is simple, but the method has some drawbacks. In the gradient method [11], a gradient signal for the adaptation of M_1 is obtained from the correlation of the error and input signals,

$$E[v_e v_m^*] = E[v_m \epsilon_1 v_m^*] = E[x_m \epsilon_1(x_m, D_{0M}, M_1)] = X_{em}(D_{0M}, M_1). \quad (31)$$

When the loop is balanced, M_1 nulls the above correlation. For simulation purposes, M_1 may be searched in order to minimize to zero the squared magnitude $|X_{em}(D_{0M}, M_1)|^2$. Finally, in the peak power minimization method, M_1 is searched in order to minimize the peak error power

$$\max[x_e(x_m, D_{0M}, M_1)] = x_{e,max}(D_{0M}, M_1). \quad (32)$$

The statistics of x_m are sufficient knowledge of the input signal to simulate the three adaptation methods. Each adaptation criterion is a function of only D_{0M} and M_1 .

The gradient method is widely thought to be equivalent to power minimization in that it minimizes the average error power. According to linear estimation theory [12], this would be exactly true if the vector modulator were placed in the linear path of the input signal, which is opposite the main amplifier, but in the present case the two schemes are not equivalent. As for the peak power method, it can be impractical in that accurate peak power readings can be difficult to obtain with practical signals, and the peak power surface may be difficult to minimize because it has discontinuous derivatives. The method remains highly valuable, however, from a tutorial standpoint.

Fig. 7 compares the above three adaptation methods for M_1 , when the input is a two-tone signal, the main and error amplifiers are $c = 1.5$ transistors with $V_{ds} = 0.06$ (Class AB) and 0.5 (Class A) respectively, and $D_{0M} = -2.5386$ dB and $D_{0E} = 8.7088$ dB. Contours are plotted over the M_1 complex plane. Contours of X_e , $|X_{em}|^2$, and $x_{e,max}$ represent the surfaces minimized in the power, gradient, and peak power methods, respectively, while those of $IMSR$ give the output $IMSR$ as minimized by M_2 . Because M_1 seeks to balance a loop whose reference branch has unity gain, as shown in Fig. 3, and because the relevant compression function $f_M(d_M)$ in Fig. 6 has a magnitude never exceeding 0.5 and a positive phase, the optimal M_1 values marked in the first three plots have magnitudes exceeding two and negative phases. The surprising discrepancy between the power and gradient methods is one that subsides as the main amplifier is biased more toward Class A, but worsens toward class B. The plot of $IMSR$ reveals a pattern of two local minima that is consistently observed when the main amplifier is biased far enough from Class A. The minima on the left and right sides are marked with triangles pointing in these respective directions and will be designated as optimal-L and optimal-R. These correspond here to $IMSR = -60$ dB and -59 dB respectively. Since these values are less favorable than the ratios of power spectra, as noted earlier, they more than meet the -60 dB criterion commonly used in mobile systems design [11]. The three optimal values of M_1 previously identified are revealed to be sub-optimal when superimposed on the plot of $IMSR$. Although they tend to cluster around the optimal-R value, they differ from it by complex relative errors of magnitude 6.1 %, 31.2 %, and 4.5 %, resulting in $IMSR = -55.6$ dB, -43.2 dB, and -58.4 dB respectively.

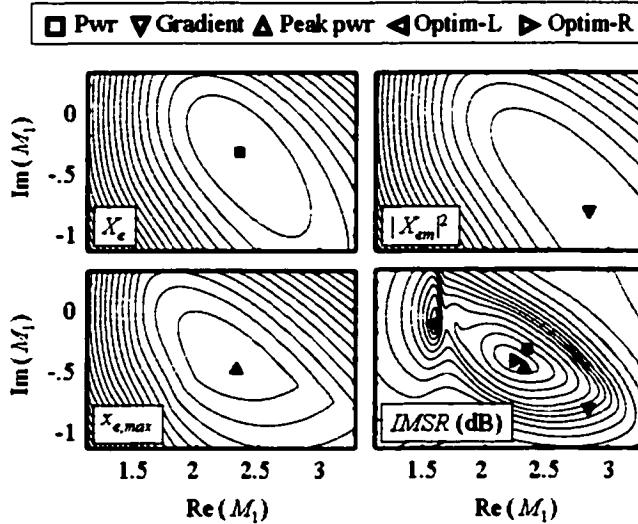


Fig. 7. For two-tone excitation, contours of three adaptation criteria for the first loop and of the achievable output $IMSR$ versus vector modulator tuning in the first loop. Minima are marked with symbols.

The above experiment was repeated in Fig. 8 with a three-carrier IS-95A CDMA waveform downloaded from Agilent Technologies' Web site [16]. The same amplifiers were used with lower reference input drives of $D_{0M} = -11.387$ dB and $D_{0E} = 2.1938$ dB because the higher crest factor requires more backoff for comparable $IMSR$ values. In the $IMSR$ plot, the optimal-L and optimal-R points now correspond to $IMSR = -60$ dB and -62.2 dB respectively. The optimal M_1 values from the power, gradient, and peak power methods are also superimposed. In comparison to the two-tone case, these three points are now spread out considerably further from the optimal-R point, and with far greater consequence on their

IMSR values. They now depart from the optimal-R location by 21.2 %, 20.2 %, and 13.2 %, resulting in $IMSR = -46.176$ dB, -45.427 dB, and -49.403 dB respectively. The earlier discrepancy between the power and gradient methods is all but removed.

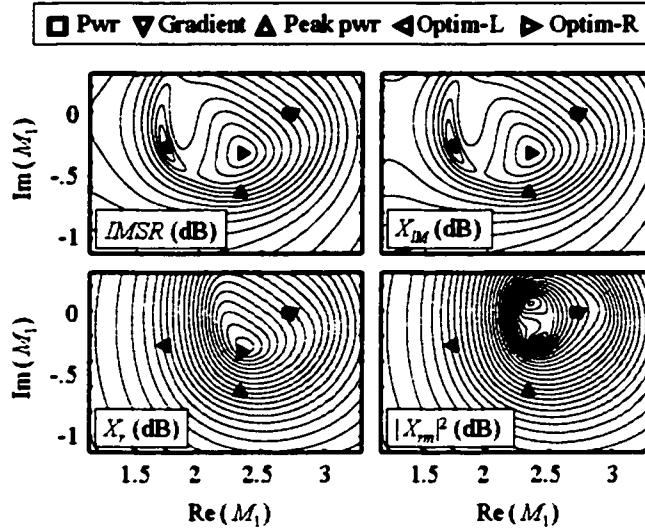


Fig. 8. For CDMA excitation, contours of the achievable output *IMSR* and of three contributing terms versus vector modulator tuning in the first loop. Symbols have same significance as in Fig. 7.

In conclusion, the need to conform M_1 to either of the optimal-L or optimal-R values is not justified for a two-tone signal, but it cannot be ignored for signals of higher crest factor such as CDMA. This need can be met by performing a four-dimensional search on M_1 and M_2 in order to minimize the output *IMSR*, as mentioned earlier. The existence of two local minima only requires initial search values appropriate for the minimum sought.

B. Implementation Issues

In the five adaptation schemes portrayed in Figs. 7 and 8, the practical difficulty resides in monitoring the output *IMSR*. Doing so via the input and output signals v_m and v_n as specified in (12) is not accurate because it involves the subtraction of two nearly equal numbers to obtain a result many orders of magnitude smaller. The alternative is to use the input and residual signals v_m and v_r as specified in (15). The residual signal is not directly available for measurement 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 [11]. These methods will therefore be referred to as error signal suppression and input signal suppression. Fig. 9 shows the third circuit loop that must be added to Fig. 3 for each of these purposes. In either case, the diagram allows for the gain of the 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 . The implications of imperfect extraction were examined in [9], but this paper assumes perfect extraction for simplicity.

The numerator $X_{IM} = X_r - |X_m|^2$ in (15) is measurable provided that its two terms are sufficiently different. To help verify this, Fig. 8 provides the contours of the X_r and $|X_m|^2$ associated with the optimal *IMSR* contours in the same Figure. The two sets of contours appear different around the optimal-R location

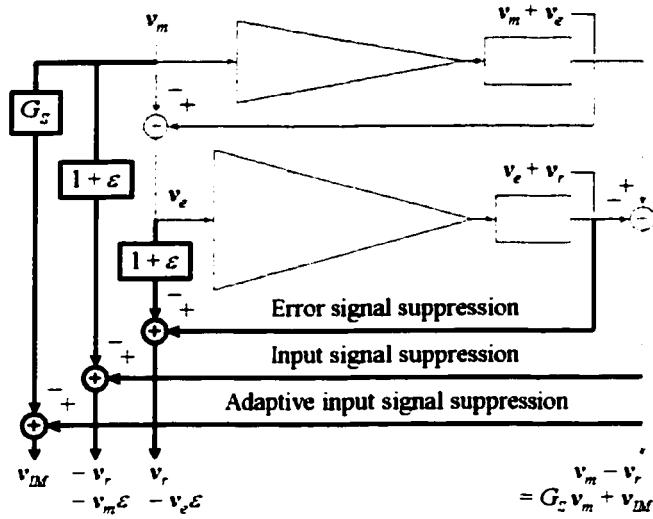


Fig. 9. Third circuit loop added to Fig. 3 for pilotless adaptation of second loop with error signal suppression, input signal suppression, or adaptive input signal suppression.

but remarkably similar near the optimal-L location. Indeed, at the optimal-R location, $X_r = -59.7$ dB and $|X_{rm}|^2 = -63.3$ dB differ by a safe 3.6 dB, but at the optimal-L location, $X_r = -23.2197$ dB and $|X_{rm}|^2 = -23.2205$ dB differ by only 0.0008 dB. Therefore, while the $IMSR$ minimum at the optimal-R location is attributable to a very low value of X_r , the minimum at the optimal-L location is attributable to a convergence between much higher values of X_r and $|X_{rm}|^2$. This makes (15) unsuitable for use in optimal-L adaptation. In that case, the IM distortion v_{IM} must be extracted from the feedforward circuit in order to directly measure X_{IM} . Then v_r is no longer needed to compute the denominator $|1 - X_{rm}|^2$ in (15) because minimizing X_{IM} is virtually equivalent to minimizing $IMSR$. To demonstrate this, Fig. 8 provides the contours of X_{IM} associated with the optimal $IMSR$ contours in the same Figure. The two sets of contours are very similar even near the optimal-L location because the magnitude of X_{rm} is still too small to have much of an effect on $|1 - X_{rm}|^2$. Extraction of v_{IM} is done by input signal suppression with an adaptive third loop, as shown in Fig. 9 [18]. Given current values of M_1 and M_2 , a third vector modulator in the reference branch of the third loop seeks to minimize the extracted signal in power, or equivalently, to uncorrelate it with v_m . The objective is met when the vector modulator gain is G_s and the extracted signal is v_{IM} . When the third loop is adapted by power minimization, all three vector modulator gains are searched toward minimizing the extracted signal power, and therefore the system could in principle be adapted via a single six-dimensional search rather than separate four-dimensional and two-dimensional searches. In computer simulations, a single four-dimensional search is conducted because the numerical precision allows computing X_{IM} as $X_r - |X_{rm}|^2$.

In the remaining four adaptation schemes in Fig. 8, the similarity of the $IMSR$ and X_r contours at the relevant M_1 locations suggests that minimizing $IMSR$ might be achieved well enough by minimizing X_r instead, thus dispensing from the need to measure X_{rm} . To confirm this, Fig. 10 shows the contours of X_r and $IMSR$ in the M_1 plane when the tuning of M_2 minimizes X_r instead of $IMSR$. The near-perfect correspondence between the two sets of contours indicates that minimizing X_r practically removes the

correlation between v_r and v_m , making $X_{rm} \approx 0$. The original M_1 locations from the power, gradient, and peak power methods are superimposed on both plots and now correspond to $\text{IMSR} = -46.174 \text{ dB}$, -45.427 dB , and -44.681 dB respectively. Compared to the former values, IMSR is virtually unchanged in the power and gradient methods, but suffers close to 5 dB degradation in the peak power method. A new optimal M_1 at the minimum of the X_r plot is marked with a circle and will be designated as optimal-r. As before, this value is obtained through a four-dimensional search on M_1 and M_2 , but now the search is insensitive to the initial search value because X_r has a single minimum. The resulting IMSR is -61.9 dB , or only 0.3 dB more than the optimal-R value. This small degradation is typical and is certainly justified by the above advantages of optimal-r adaptation.

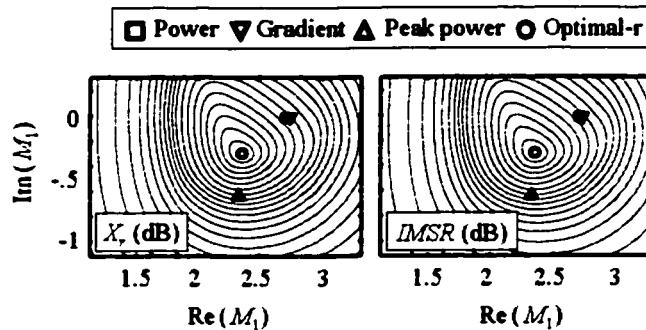


Fig. 10 For CDMA excitation, contours of minimized average residual power and corresponding output IMSR versus vector modulator tuning in the first loop.

Outside the context of optimal-L adaptation, this paper will henceforth assume that M_2 minimizes X_r . Ref. [7] noted that this is analogous to the power minimization method where M_1 minimizes X_e . By further analogy, a gradient signal may be used to some advantage in adapting M_2 . This signal is the correlation of the residual and error signals,

$$\mathbb{E}[v_r v_e^*] = \mathbb{E}[v_e \varepsilon_2 v_e^*] = \mathbb{E}[x_e(x_m, D_{0M}, M_1) \varepsilon_2(x_m, D_{0M}, M_1, D_{0E}, M_2)] = X_{re}(D_{0M}, M_1, D_{0E}, M_2). \quad (33)$$

Like X_e in the gradient method, X_r is not exactly minimized when M_2 nulls the above correlation, but the difference in the output IMSR is negligible. Unfortunately, since $X_{re} = 0$ can be satisfied for any given value of M_1 , it cannot serve as a criterion for both M_1 and M_2 , and is of no utility in optimal-r adaptation.

C. Additional Insights

Several signal representations such as statistical, vectorial, time-domain, and frequency-domain are required to gain full insight into the results of Figs. 7, 8, and 10. Fig. 11 is a statistical and vectorial representation of signals in the feedforward amplifier for the case of Fig. 7 with the power minimization method. Since the two-tone signal has a crest factor of $3 \text{ dB} = 2$, the input signal magnitude $|v_m|$ extends from zero to $\sqrt{2} = 1.414$ on the horizontal scale. The curves labeled $|v_M|$ and $|v_m|$ compare the magnitudes of the signals arriving from the upper (nonlinear) and lower (linear) paths of the first loop to be subtracted, and the curve labeled $|v_e|$ is the resulting error signal magnitude. The first curve is the $\sqrt{d}|f(d)|$ curve with $V_u = 0.06$ that may be deduced from Fig. 6, but stretched along both axes by the

factor $1/\sqrt{D_{0M}} = 1.34$ to make it saturate at this value, and then compressed toward the vertical axis by the factor $|M_1| = 2.36$ to approximate unity gain. The error magnitude $|v_e| = |v_M - v_m|$ is indeed a function of the magnitude of v_m and not of its phase, as (7) shows, but it must be evaluated vectorially in the complex plane. This is done in Fig. 11 for regularly spaced real values of v_m . The AM/PM conversion is manifested in the final bending of the v_M trajectory. The results are transferred in the original curve plot, on which is then superimposed the probability density function (pdf) of $|v_m|$, scaled as needed for clarity. The pdf is quite flat up to $|v_m| = 1$ but then shoots rapidly upward due to the zero slope of the signal envelope at its peaks. Hence, the minimization of the average error power put more weight on keeping $|v_e|$ small in the compression region of the main amplifier, where AM/PM occurs. The vector diagram shows that M_1 tried to compensate for the positive AM/PM by taking on a negative angle of -8.1° .

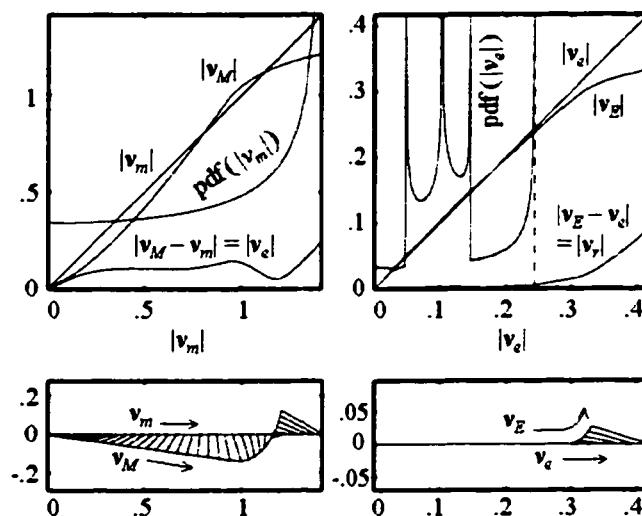


Fig. 11. Statistical and vectorial representation of signals in the feedforward amplifier for the case of Fig. 7 with the power minimization method.

Fig. 11 also conveys in analogous fashion the signals found in the second loop. The pdf of $|v_e|$ is relatively complex, and extends to $|v_e| = 0.24$ as indicated by the dashed line. With the Class A error amplifier saturating as high as $|v_E| = 1/\sqrt{D_{0E}} = 0.367$, M_2 was able to balance the second loop very well over the entire relevant range, and hence to produce the output IMSR of -55.6 dB.

Signal representations like those of Fig. 11, supplemented by time-domain samples of the waveforms, are provided in Fig. 12 to help analyze the results obtained in Figs. 8 and 10 with a CDMA signal. The power minimization method is featured in the first column for comparison with Fig. 11. The CDMA signal having a crest factor of 12.33 dB = 17.1 , $|v_m|$ now extends to $\sqrt{17.1} = 4.135$ on the horizontal scale. The $|v_m|$ - $|v_M|$ curve is the same as before, but now saturates at $1/\sqrt{D_{0E}} = 3.71$. The pdf of the CDMA signal is qualitatively opposite to that of the two-tone signal, peaking in the small-signal region and tapering off in the compression region. Consequently, a higher value of $|M_1| = 2.71$ compresses the $|v_m|$ - $|v_M|$ curve further toward the vertical axis in an effort to reduce the error magnitude in the small-signal region. With an angle of only -0.275° , M_1 makes little effort to compensate for the AM/PM, so that $|v_e|$ peaks at $1.6 = 4.1$ dB at the end of the range. The next plots in the same column show samples of the various signal envelopes in dB over a time interval that features several strong peaks, including the maximum of $|v_m| =$

12.33 dB. The error signal has a minimized average power of -17.2 dB, but with its maximum at 4.1 dB, has a crest factor of 21.3 dB, or 9 dB higher than that of the input signal. The residual signal exaggerates these characteristics, dropping to below the -60 dB level but featuring occasional strong peaks that may approach the 0 dB mark. The Figure's first column ends with a statistical and vectorial representation of signals in the second loop. With $|v_E|$ saturating at $1/\sqrt{D_{0E}} = 0.777$, M_2 was not able to balance the second loop in the low-probability range extending to $|v_e| = 1.6$, resulting as we know in $IMSR = -46.174$ dB.

The other three columns of Fig. 12 repeat the first column for the peak power, optimal-r, and optimal-L adaptation methods. These correspond to $|M_1|$ values of 2.42 , 2.38 , and 1.75 respectively, resulting in a progressive stretching of the $|v_m|$ - $|v_e|$ curve away from the vertical axis. In the peak power method, the maximum $|v_e|$ is reduced to $0.74 = -2.6$ dB, but as shown in the complex plane, M_1 achieves this by overcompensating for the AM/PM with an angle of -15.34° . The result, as shown in the time-domain traces, is a relatively high average error power of -10.1 dB but a crest factor of 7.5 dB, or 4.8 dB lower than that of the input signal. These two trends are likewise reflected in the residual signal. The first trend undoes the benefits of the second, with $IMSR$ now at -44.681 dB.

Optimal-r adaptation, illustrated in the third column of Fig. 12, strikes a spectacular compromise between the two former methods. M_1 compensates for the AM/PM with an angle of -7.25° that is reminiscent of the -8.1° angle obtained in the two-tone case with the power minimization method. In the time-domain plots, the resulting residual signal clearly stands out as having the lowest occurrence of strong peaks and the smallest average power overall. As mentioned before, the resulting $IMSR$ is -61.9 dB.

In the fourth column of Fig. 12, optimal-L adaptation stands in stark contrast to the previous methods. The favorable $IMSR$ of -60 dB that is obtained appears contradicted by the much-increased residual power observed in the time-domain plot, until one notices that the residual signal envelope almost perfectly mirrors that of the input signal. It is therefore possible for M_1 and M_2 to "match" the main and error amplifiers in such a way that v_r approaches αv_m where α is in general complex. The ideal case of $v_r = \alpha v_m$ would not corrupt v_m in the output signal $v_o = v_m - v_r = (1 - \alpha)v_m$ but would only modify the gain of the system. In terms previously discussed, it would also correlate v_r fully with v_m , creating the condition $X_r = |X_{rm}|^2 = |\alpha|^2$ and hence the convergence between X_r and $|X_{rm}|^2$ that brings $IMSR$ to zero in (15). A modified gain for a system that was said to behave as an automatic gain and phase controller can be explained by the fact that optimal-L adaptation does not employ a balanced second loop, as seen in Fig. 12. The $|v_e|$ - $|v_E|$ curve for this method has a small-signal gain of $0.9\angle -2.46^\circ$, which differs significantly from unity. As a consequence, any distortion in v_m that is not the direct result of the main amplifier's AM/AM or AM/PM conversion, such as a frequency response or noise, will not be cancelled as efficiently in the second loop. The resulting process differs from the classic feedforward process in which any and all types of distortion are given equal treatment.

For additional insight into optimal-L adaptation, Fig. 13 supplies a complete vectorial representation of signals for the current case. As in Fig. 12, v_m takes on regularly spaced real values, but for clarity these extend only up to 3.4 to avoid the compression region. Also for clarity, the real and imaginary axes are unequally scaled, and the vectors that apply to $v_m = 1.7$ are highlighted in bold. The AM/AM in both amplifiers and the AM/PM in the error amplifier interact in such a way that the output signal v_o takes on almost equally spaced values along an almost straight line. The output signal approximates v_m scaled by the factor $1 - \alpha = 0.94\angle -1.8^\circ$, from which $\alpha = 0.07\angle 25^\circ$.

A final useful comparison of the four adaptation methods of Fig. 12 is provided in the frequency-domain plot of Fig. 14. The signal symbols have been capitalized to indicate Fourier transformation. The input signal V_m has an emission bandwidth of 3.75 MHz, and has a low enough noise floor to allow to accurately

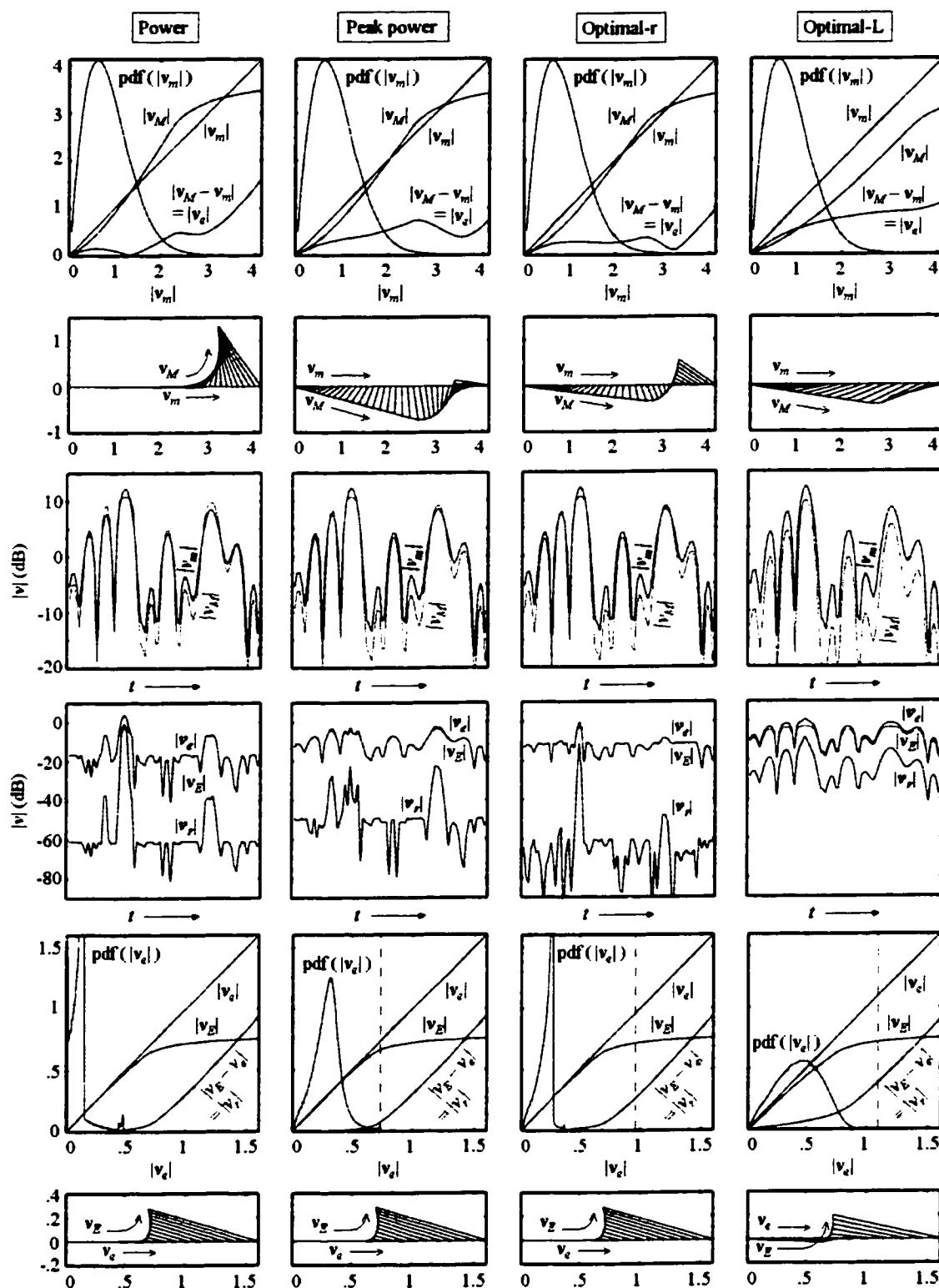


Fig. 12. Statistical, vectorial, and time-domain representations of signals in the feedforward amplifier for the case of Figs. 8 and 9 with four adaptation methods.

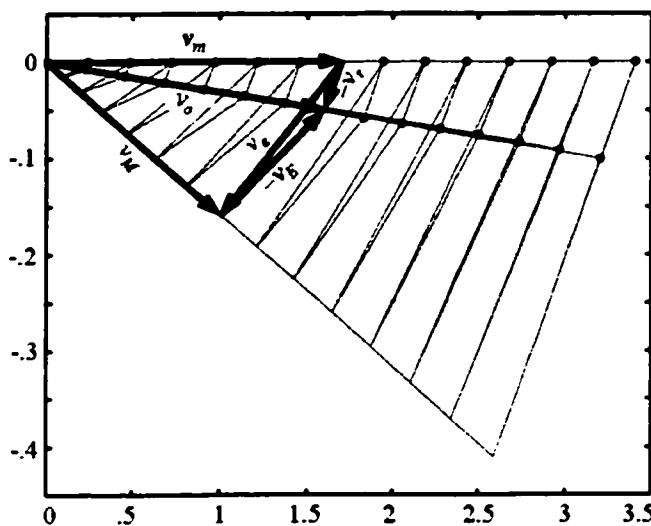


Fig. 13. Complete vectorial representation of signals in the feedforward amplifier for the case of Fig. 8 with optimal-L adaptation.

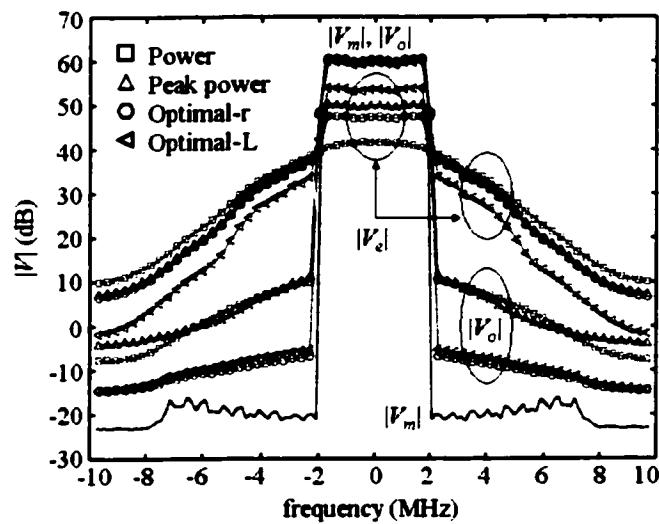


Fig. 14. Frequency-domain representations of signals in the feedforward amplifier for the case of Figs. 8 and 10 with four adaptation methods.

assess the spectral regrowth of the other signals. The spectra of the error signals show different amounts of input signal being leaked into the error signal, from virtually none in the power minimization method to a large amount in the optimal-L method. Clearly, the extent of this leaking is not a reliable predictor of the residual spectral regrowth in the output signal V_o . The highest spectral density ratios for the four output signals, in the usual order, can be read as -50 dB, -49 dB, -67 dB, and -65 dB. As expected [11], these figures are between 4 and 5 dB more favorable than the corresponding output IMSRs of -46.174 dB, -44.681 dB, -61.9 dB, and -60 dB. Barely visible in this plot is the fact that the output signal in optimal-L adaptation lies $-20\log(|1-\alpha|) = 0.54$ dB below the other output signals within the emission bandwidth.

V. OPTIMIZATION OF POWER EFFICIENCY

A. Summary

The above analysis has established that in a given feedforward amplifier driven by a high-stress input signal such as CDMA, a change in the adaptation method alone can have a major impact on the system's linearization performance. Therefore it can be hoped that in practical feedforward design, where a given output linearity would need to be achieved, a better choice of adaptation scheme would allow smaller amplifier power ratings and thus would translate into a more power-efficient system. The same hopes can also be entertained in relation to the proper choice of main amplifier characteristics. The present section seeks to verify and quantify such efficiency gains.

For a wide range of required output SIMRs, the design parameters of a feedforward amplifier were optimized for maximum power efficiency using the statistical technique detailed in [7]. This was done for main amplifiers ranging from Class AB to Class C, and for the power minimization, optimal-r, and optimal-L methods of adaptation. For maximum relevance, this exercise initially assumed a CDMA input signal and $c = 1.5$ transistors, but was later repeated with only the input signal changed to two-tone, and then with only the transistors changed to $c = 15$.

B. Input and Output Parameters

All optimization results presented here assume the same feedforward system specifications as in [7] in regard to the linear power gain, delay-to-gain ratios of amplifiers, residual delays in loops, delay-to-loss ratios of delay lines, insertion losses of couplers, input coupling, miscellaneous losses, and central attenuations of vector modulators. These input values are typical attainable ones and are not subject to optimization because they either specify a required performance, in the case of the system's linear power gain, or their optimal solutions are already known. The remaining specifications assumed here are

$$\{IMSR_{req}\} = \{-40, -41, -42, \dots, -80 \text{ dB}\}, \quad \delta_1 = \delta_2 = \varepsilon = 0. \quad (34)$$

$\{IMSR_{req}\}$ specifies a generous set of required output IMSRs, equally spaced and centered at -60 dB. Tolerances δ_1 and δ_2 are the worst-case magnitudes of the complex relative errors in M_1 and M_2 . Typical tolerances were simulated in [7], but the extra computational effort would have impossibly burdened the present work. The ideal situation $\delta_2 = 0$ in the second loop is a valid approximation if the actual δ_2 is small enough to be overshadowed by the nonlinearity of the error amplifier. This condition is met if the output IMSR due to δ_2 alone, approximately equal to $\delta_2^2 / IMSR_{M1}$, where $IMSR_{M1}$ is the IMSR of the amplified input signal [11], is much better than $IMSR_{req}$, or

$$\delta_2 \ll \sqrt{IMSR_{req} / IMSR_{M1}}. \quad (35)$$

An expression for δ_1 can also be obtained if the gradient method of adapting M_1 is assumed. With $\delta_1 = 0$, v_{M1} is $v_m + v_d$ where $v_d = v_e$ is uncorrelated with v_m , and the average error power X_e is $X_d = IMSR_{M1}$. Assuming now a finite δ_1 with M_1 after the main amplifier, the extra signal $\delta_1 e^{j\theta} (v_m + v_d)$ is at worst added to v_e . Counting only the dominant contribution of $\delta_1 e^{j\theta} v_m$, the updated X_e is $\delta_1^2 + IMSR_{M1}$, for a relative change in X_e of $\delta_1^2 / IMSR_{M1}$. Given therefore a reference case involving $IMSR_{M10}$ and δ_{10} , and a new case involving $IMSR_{M1}$, the new δ_1 that would cause the same relative change in X_e , and presumably the same degradation in $IMSR$, is

$$\delta_1 = \delta_{10} \sqrt{IMSR_{M1} / IMSR_{M10}}. \quad (36)$$

This dependency on IMSR_{M} is the inverse of that in (35), and verifies by simulation for each adaptation method. The absolute requirement on δ_1 , however, depends on the said method as shown in Figs. 8-9. Optimal-L adaptation has clearly a more stringent requirement than optimal-r adaptation. The power method is also at a disadvantage in that M_1 is located on a nearby slope of the IMSR surface. In that case, however, an error in M_1 is as likely to improve IMSR as it is to worsen it, and therefore the approximation $\delta_1 = 0$ may be used to obtain an average IMSR in place of the worst-case IMSR sought in [7].

Error ε is that of the gain of the reference branch in a non-adaptive third loop doing error or input signal suppression, and therefore is relevant to the power and optimal-r adaptation methods. With such an error, [9] advises against using input signal suppression with the present adaptation methods, and reports that with error signal suppression, an error of approximately ε is introduced in the targeted M_2 regardless of M_1 . Assuming error signal suppression, therefore, the assumption (34) places the same demands on $|\varepsilon|$ as on the tolerance δ_2 .

The outputs of the optimization procedure come in sets that correspond to $\{\text{IMSR}_{\text{req}}\}$. They consist of the seven optimization parameters

$$\{|G_M|^2\}, \{|G_E|^2\}, \{|C_M|^2\}, \{|C_e|^2\}, \{|C_u|^2\}, \{X_{\text{satM}}\} = \frac{\{P_{\text{satM}}\}}{P_{S,\text{req}}}, \{X_{\text{satE}}\} = \frac{\{P_{\text{satE}}\}}{P_{S,\text{req}}}. \quad (37)$$

and the resulting dc-RF conversion efficiencies $\{\eta\}$ for the system. The seven parameters are the two amplifier gains, the three couplings besides the input coupling, and the amplifiers' saturated output powers normalized with the system's required output power $P_{S,\text{req}}$. The dc-RF conversion efficiency is defined as

$$\eta = \frac{P_S}{P_{\text{dcM}} + P_{\text{dcE}}}. \quad (38)$$

where P_S is the system's output power and $P_{\text{dcM}}, P_{\text{dcE}}$ are the amplifiers' average dc power consumptions. Any other dc consumptions are disregarded because they are considered constant and hence irrelevant for optimization purposes.

C. Results

Fig. 15 summarizes the efficiencies computed in over 200 optimizations of the type described above, involving different combinations of input signal, transistor devices, main amplifier gate biasing, and adaptation method. Efficiencies are plotted on logarithmic scales to emphasize relative rather than absolute changes in efficiency. The main amplifier's normalized gate bias voltage V_{d} in the horizontal axes of the plots was varied in increments of 0.01, and the efficiency curves were smoothed as necessary. The SIMR employed in the plots is analogous to the signal-to-noise ratio (SNR), and is the reciprocal of IMSR , or the absolute value of IMSR in dB. Each column in the Figure represents a combination of input signal format and transistor devices. The curves resulting from the power minimization and optimal-r adaptation methods are plotted together because of their similarities. A plotting symbol identifies the adaptation method in each curve and is placed where V_{d} maximizes efficiency.

The first column of Fig. 15 refers to a two-tone input and $c = 1.5$ transistors, and reveals little advantage in using optimal-r adaptation. This reflects the fact that in Section IV A, optimal-R adaptation had little IMSR advantage over the power method when the input was a two-tone signal. The first column also puts optimal-L adaptation at a slight disadvantage. Whatever the method, the optimal V_{d} decreases in tandem

with $SIMR_{req}$. In the first two methods, the optimal V_{ds} values decrease to around zero, which is Class B operation. In the third method they are lower and decrease to -0.145 , well into Class C operation.

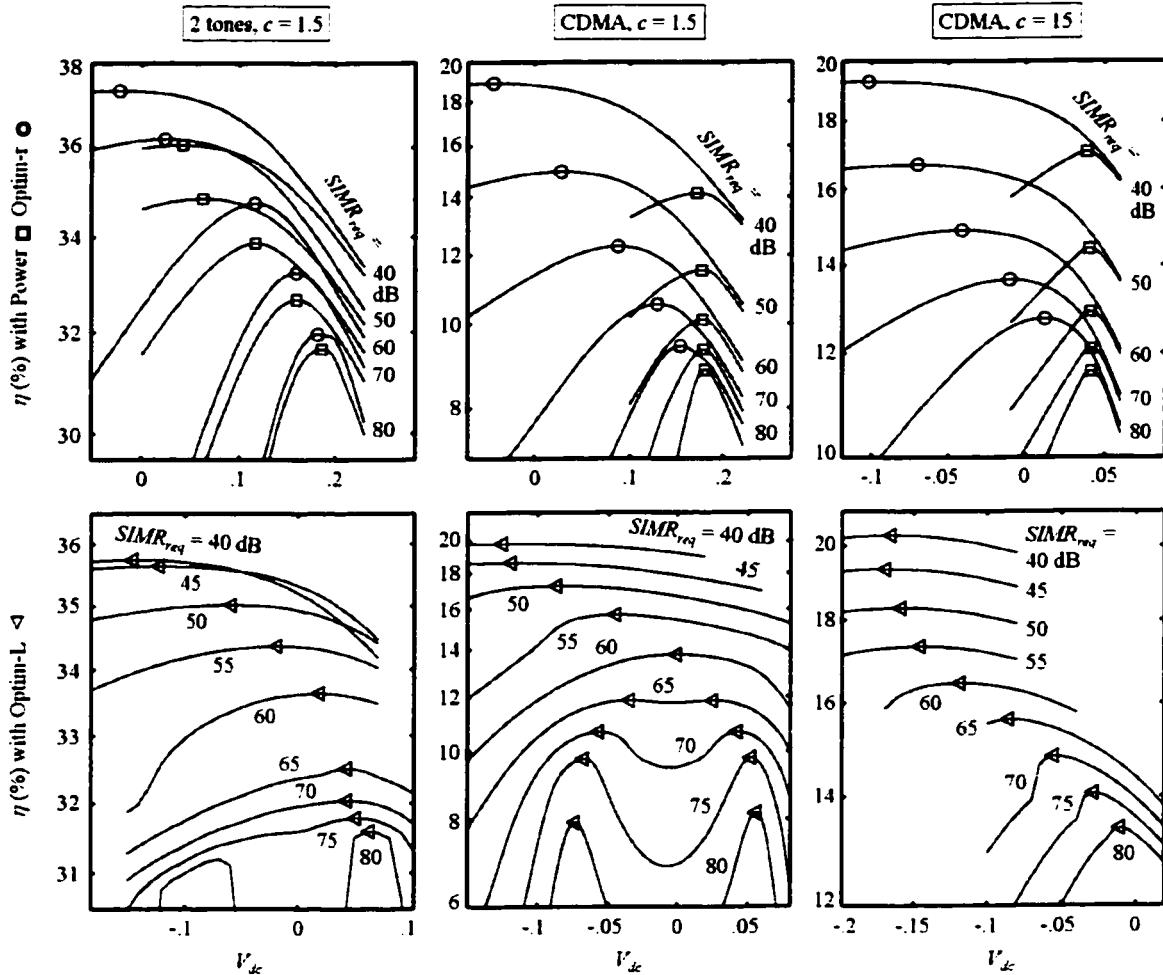


Fig. 15. Optimal feedforward amplifier efficiencies computed under typical design constraints versus main amplifier gate bias voltage. The plots invoke subsets of the required output SIMRs, three adaptation methods, and three combinations of input signal and transistor devices.

The second column of Fig. 15 repeats the first column with a CDMA input signal instead of two-tone. Aside from the fact that all efficiency figures are much lower, the optimal V_{dg} in power minimization is now a steady 0.18 regardless of the required $SIMR$, hence the industry's preference for Class AB main amplifiers in high-power applications [20]. To the right of $V_{dg} = 0.18$, the optimal-r curves differ little from those for the power method, but moving left toward Class B, the former curves continue to rise while the latter dive. The wide separations already evident at $V_{dg} = 0.06$ reflect the fact that in Section IV A, optimal-r adaptation had an important $IMSR$ advantage over the power method when the input was a CDMA signal. The resulting optimal-r curves restore the trend of decreasing optimal V_{dg} as observed in the two-tone case. The new efficiency peaks are considerably higher than those with the power method, and they are also broader, which reduces the sensitivity to V_{dg} . In optimal-L adaptation, the efficiency peaks are the highest of all three methods for up to $SIMR_{req} = 70$ dB. Since this method is based on matching amplifier

compression characteristics, its optimal V_{de} values may be expected to be similar to those found with the same $c = 1.5$ transistors and two-tone excitation. This is indeed the case except that a new series of local maxima, of which the two-tone case only gave a hint, now trace an equally good second path beyond $SIMR_{req} = 60$ dB.

The third column of Fig. 15 repeats the second column with $c = 15$ transistors in both amplifiers. As may be expected, this change causes a systematic increase in all efficiency figures. The optimal V_{de} values are also found to have systematically decreased. In the power minimization method, they are now steady at 0.04, and in the other two methods, extend from about Class B to well into Class C. Among other changes, the optimal-r method has an even more pronounced advantage over the power method in terms of sensitivity to V_{de} , and optimal-L adaptation now has the efficiency advantage over the full range of $SIMR_{req}$.

All the efficiencies optimized with respect to V_{de} , and the corresponding optimal values of V_{de} and of the last three optimization parameters in (37), are plotted in Fig. 16 with respect to $SIMR_{req}$. This is the type of presentation used in [7] but with the additional parameter V_{de} , since the original optimization procedure has in effect been upgraded, at considerable effort, with this eighth optimization parameter. In the plots not shown in Fig. 16, $|G_M|^2$ falls almost exclusively within a 3 dB bracket. $|G_E|^2$ is such that $|G_E|^2|C_o|^2$ remains about constant, and $|C_M|^2$ and $|C_e|^2$ remain at around -30 dB and -10 dB respectively.

In the second column of Fig. 16, at the typical $SIMR_{req}$ of 60 dB, the optimal η of 10.1 % in the power method rises to 12.3 % in optimal-r adaptation and 13.8 % in optimal-L adaptation, for relative improvements of 21 % and 36 %. In accordance with Fig. 15, optimal-L curves in this column, excluding the efficiency curve, split in two branches after $SIMR_{req} = 60$ dB. The discontinuities also observed in these curves near 40 dB, and in the same curves of the first column at 65 dB, are due to necessary changes of local minima in the graphical portion of the optimization procedure [7]. For any given $SIMR_{req}$ in any of the three columns, moving from power minimization to optimal-r adaptation, and from there unto optimal-L adaptation, successively decreases the optimal V_{de} . The main amplifier's normalized saturated output power X_{outM} hardly changes, but the transitions toward Class C bring large decreases in the main amplifier's dc power consumption. The accompanying increases in error signal power require increases in the error amplifier's normalized saturated output power X_{outE} and proportional increases in its dc power consumption. The optimal output coupling $|C_o|^2$ meanwhile increases; this appears as a decrease in Fig. 16 due to the practice of using positive dB values. This last trend reflects the fact that the power consumption of the error amplifier is catching up with that of the main amplifier. The two power consumptions actually meet in the center column, where the output coupling in optimal-L adaptation reaches about -3 dB and the coupler becomes a hybrid junction. Going even further, the alternative path of this coupling curve shows values between -2.3 dB and -1.4 dB, at which point the coupled path of the output coupler receives the main signal instead of the error signal.

D. Input Signal Splitting

A closer look at optimal-L adaptation with $c = 15$ transistors is warranted, as this situation produces by far the best efficiency figures. Fig. 17 shows the statistical and vectorial representation of signals in the optimized feedforward amplifier for $SIMR_{req} = 60$ dB, where $\eta = 16.5$ %. Comparing with Fig. 12, in which $c = 1.5$, the $|v_m| - |v_M|$ curve is parallel to the $|v_m|$ line over a longer length, thanks to a much sharper elbow. Concerns over an unbalanced second loop are also less important, the $|v_e| - |v_E|$ curve having a small-signal gain of $0.98 \angle -0.22^\circ$ and being more linear. These trends point toward a mode of feedforward operation depicted in Fig. 18. The system in effect splits the input signal in two: a top portion

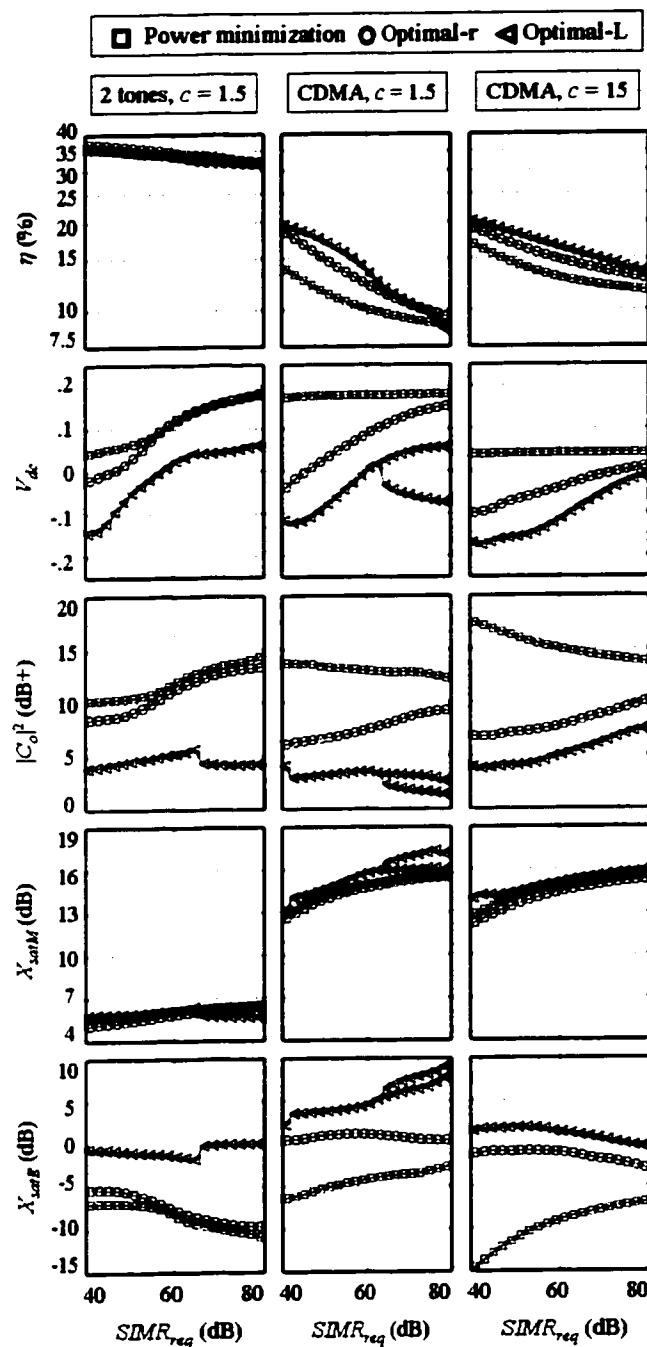


Fig. 16. Optimal values of efficiency and of four design parameters versus required output SIMR of feedforward amplifier for three adaptation methods and three combinations of input signal and transistor devices.

extracted by the efficient Class C main amplifier, and a bottom portion remaining after subtraction of the top portion. The low-crest-factor error signal is amplified by a Class A amplifier operated just below saturation, and the signal peaks are added back at the output of a balanced second loop. At first glance the feedforward architecture appears to have become redundant, since the input signal could be truncated by

feeding it directly into the error amplifier. However, this scheme would fail for two reasons: 1) the error amplifier would overheat from the near-constant saturation, and 2) the saturating error amplifier would introduce AM/PM distortion into the amplified error signal.

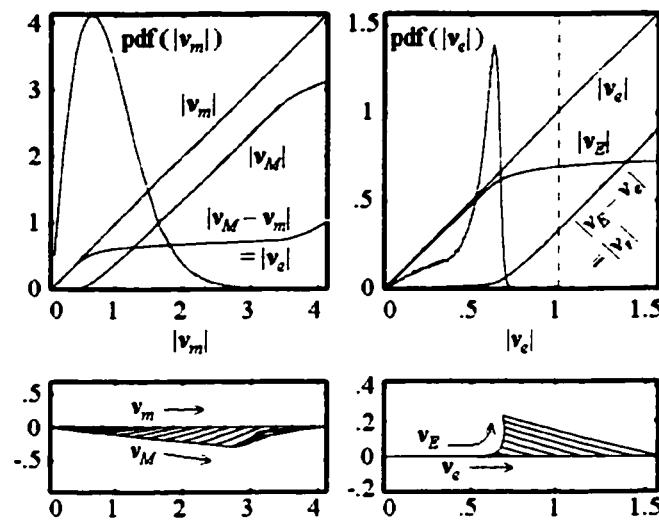


Fig. 17. Statistical and vectorial representation of signals in the fully optimized feedforward amplifier with CDMA excitation. $C = 15$ transistor devices, optimal-L adaptation, and a required output SIMR of 60 dB.

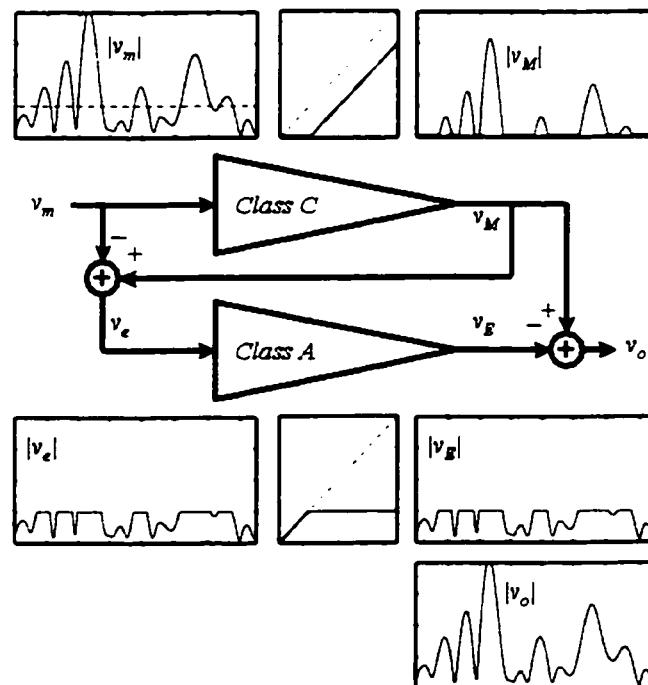


Fig. 18. Time-domain representation of signals in a mode of feedforward operation where the input signal is effectively split into top and bottom portions.

VI. CONCLUSIONS

The well-established feedforward technology is commonly thought to have reached its potential in terms of power efficiency, but two new pilotless adaptation methods now promise significant gains in this area. These findings were made possible by the combination of two factors: 1) the assumption of a modern high-stress communication signal instead of the classic two-tone signal, within a simulation tool that can account efficiently for the input signal format, and 2) the simultaneous investigation into the main amplifier class of operation, using appropriately derived families of curves describing both AM/AM and AM/PM characteristics and dc power consumption. The new adaptation methods reject the notion that the optimal tuning of the first loop is that which minimizes the average error power, and instead tune the first loop using the same bottom-line criterion already used in tuning the second loop, which is the minimization in power of a signal extracted via a third circuit loop. A four-dimensional search may now be chosen to replace the former two-dimensional searches, leading to a more complex search algorithm but more importantly to reduced hardware requirements.

When a non-adaptive third loop extracts the so-called residual signal, a type of adaptation designated optimal-r results. For a typical required linearization performance, traditional adaptation allowed 10.1 % system efficiency within a narrow window of class AB operation in the main amplifier, whereas optimal-r adaptation produced 12.3 % efficiency within a more comfortable range of class AB operation closer to class B. Thus, efficiency is improved by 21 % and system design is less sensitive to the main amplifier's gate biasing. The main amplifier being less linear, the tuning tolerance must be tightened in the second loop, which is the only apparent drawback. In the first loop, the tolerance can be relaxed in the same proportion, and then further relaxed since the targeted tuning has moved very near the minimizing point where the output IMSR surface has zero slope.

With some added complexity in making the third loop adaptive, the actual IM distortion can be extracted from the output signal, and so-called optimal-L adaptation becomes possible with adequate initial loop tunings. For the same linearization performance as before, 13.8 % efficiency was obtained within a comfortable range of class of operation in the main amplifier, now centered on class B, for an improvement of 36 % over traditional adaptation. The main amplifier being even less linear, the tuning tolerance must be tightened yet further in the second loop. The parallel relaxation of tolerance expected in the first loop is undone by the inherently more stringent requirement of optimal-L adaptation. Another possible drawback of the method is that distortion other than AM/AM and AM/PM in the main amplifier is not cancelled as efficiently, but this concern is lessened by the use of transistors of more constant transconductance.

REFERENCES

- [1] A. Bateman and D. Haines, "Direct conversion transceiver design for compact low-cost portable mobile radio terminals," in *Proc. IEEE Vehicular Technol. Conf.*, 1989, vol. 1, pp. 57-62.
- [2] A. Wright and W. Durtler, "Experimental performance of an adaptive digital linearized power amplifier," *IEEE Trans. Veh. Technol.*, vol. 41, no. 4, 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. 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 Journal*, 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, no. 2, pp. 407-412, May 1994.
- [7] C. L. Larose and F. M. Ghannouchi, "Optimization of feedforward amplifier power efficiency on the basis of drive statistics," *IEEE Trans. Microwave Theory Tech.*, vol. 51, no. 1, Jan. 2003.
- [8] C. L. Larose and F. M. Ghannouchi, "First-Loop control in feedforward amplifiers for high-stress communications signals," in *Proc. Int. Symp. on Microwave and Optical Technol.*, 2001, pp. 195-198.
- [9] C. L. Larose and F. M. Ghannouchi, "Pilotless adaptation of feedforward amplifiers driven by high-stress signals," in *Proc. IEEE Radio and Wireless Conf.*, 2001, pp. 81-84.
- [10] S. L. Loyka and J. R. Mosig, "New behavioral-level simulation technique for RF/microwave applications. Part I: basic concepts," *Int. J. RF and Microwave CAD*, vol. 10, no. 4, pp. 221-237, July 2000.
- [11] J. K. Cavers, "Adaptation behavior of a feedforward amplifier linearizer," *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 31-40, Feb. 1995.
- [12] S. S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice Hall, 1991.
- [13] T. R. Turlington, *Behavioral Modeling of Nonlinear RF and Microwave Devices*. Norwood, MA: Artech House, 2000.

- [14] O. Pitzalis, "Computer-aided design of GaAs FET power amplifiers," in *High-Power GaAs FET Amplifiers*, J. L. B. Walker, Ed. Norwood, MA: Artech House, 1993.
- [15] J. L. B. Walker, "Introduction and basic theory," in *High-Power GaAs FET Amplifiers*, J. L. B. Walker, Ed. Norwood, MA: Artech House, 1993.
- [16] AGILENT TECHNOLOGIES, CA. Three carrier IS-95A CDMA waveform for the Agilent ESG-D series. [Online]. Available: <http://www.agilent.com>. Quick Search on "DAB".
- [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] S. A. Maas, *Nonlinear Microwave Circuits*. Norwood, MA: Artech House, 1988, p. 370.
- [20] N. Potocary, *Feedforward Linear Power Amplifiers*. Norwood, MA: Artech House, 1999.

Colin L. Larose received the B.Eng. and M.Eng. degrees in electrical engineering from Concordia University, Montréal, QC, Canada in 1983 and 1986 respectively. After receiving a Bachelor of Arts in 1989 in California, he worked for five years in the area of radar cross-section measurement and simulation at the David Florida Laboratory of the Canadian Space Agency in Ottawa, ON, Canada. He has been the recipient of a research grant in 1985 and a Postgraduate Scholarship in 1995 from the Natural Sciences and Engineering Research Council of Canada (NSERC).

He is currently completing his Ph.D. degree at the École Polytechnique de Montréal, Montréal, QC, Canada. His fields of interest include feedforward and other linearization techniques, wireless communications, and measurement and simulation techniques for antennas and radar.

Fadel 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 et de génie informatique, 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.