Performance and Limitations of Decomposition-Based Parameter-Extraction Procedures for FET Small-Signal Models

Cornell van Niekerk, *Student Member, IEEE,* and Petrie Meyer, *Member, IEEE*

Abstract— A recently proposed optimizer-based parameter-extraction technique using adaptive decomposition is subjected to a systematic and rigorous evaluation. The technique is shown to be robust and accurate under varying starting conditions. A study of convergence performance based on decomposition theory and test results is presented. Robustness tests are used to show that commonly used statistical descriptions such as mean and standard deviation are inadequate for presenting these types of test data.

Index Terms—Decomposition, MESFET, parameter extraction, optimizer.

I. INTRODUCTION

THE design and analysis of millimeter-wave nonlinear circuits increasingly require accurate nonlinear models. Most of the currently available models in computer-aided design (CAD) packages are of the equivalent-circuit type and, despite efforts to develop black-box and fast physical models, it is expected that these models will still be prominent for some time. This is mainly due to their computational efficiency, availability in commercial programs, and the ease with which they can be integrated into existing design techniques.

A key step in the construction of lumped-element nonlinear models is the extraction of the small-signal equivalent circuit from *s*-parameters at different bias settings. Fig. 1 shows the 13-element small-signal model that is most often used to describe the GaAs MESFET. Until recently, the extraction of this model was performed using standard gradient and random optimizers [1] or with the aid of analytical techniques [2]. The first approach can lead to nonphysical and nonunique solutions, while the second relies on additional measurement steps or special structures. Analytical methods are faster than optimizer-based methods, but they are susceptible to measurement errors and their implementation is device specific.

In the last few years, new parameter-extraction methods have been published, which strive to overcome these limitations. Lin and Kompa [3] demonstrated an extraction technique which, by using a bidirectional multiplane search, reduces the dimensions of the optimization problem to improve robustness and efficiency. Shirakawa *et al.* [4] published an extraction technique with some features similar to the method of Lin
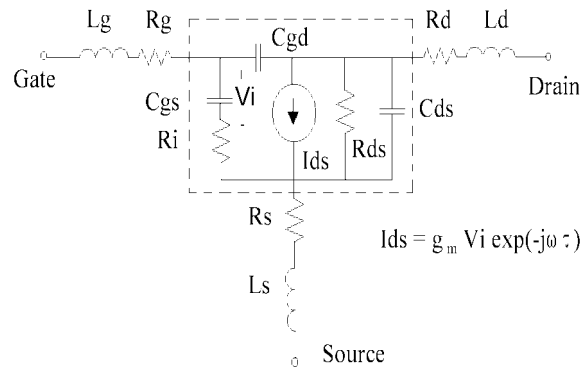
Fig. 1.   The standard 13-element MESFET model.

and Kompa. Both these extraction methods only optimize the extrinsic components of the model (shown in Fig. 1) while using analytical methods to determine the intrinsic components. Both achieve good results.

Van Niekerk and Meyer [5] recently demonstrated a very robust and powerful parameter-extractor based on the method proposed by Kondoh [6]. Kondoh showed that good results can be achieved by breaking the optimization problem into eight subfunctions and optimizing the different model elements only with respect to specific subfunctions. The eight subproblems are iteratively repeated in a specific order until the model elements have converged to their final values. Van Niekerk and Meyer extended this method by using the maximum number of subproblems, and optimizing them in a sequence in which the order is calculated with a principle components sensitivity analysis. The decomposition process used in the new procedure is, therefore, adaptive, and not based on experimentation, such as in [6].

No rigorous test results of decomposition-based parameter-extraction algorithms have been presented in the literature. In particular, the effects of starting values and the final distribution of results have been largely neglected. In this paper, the results of such tests are presented for the algorithm proposed in [5], and for the original algorithm proposed by Kondoh [6]. Robustness tests are performed using a large number of randomly chosen element starting values for each extraction. Histograms of the extracted-element values reveal them not to conform to often used standard statistical distributions, and the reasons for this are discussed.

In addition, a detailed look at the fundamental convergence mechanisms is presented, as the convergence behavior of decomposition-based optimization has only been touched on very lightly in the past. It is shown why these algorithms display superior performance with respect to problems like local minima.

Finally, a general discussion of the principle components' sensitivity analysis—the key to the basic algorithm—is also presented.

II. THE THEORY OF DECOMPOSITION-BASED PARAMETER EXTRACTION

Decomposition is defined as a process by which a function that is to be optimized is broken up into several subfunctions. The independent variables of the function are divided into groups according to their influence on a particular subfunction. Should the ith variable have its largest influence on the jth subfunction, it is assigned to that function. The subfunctions are optimized in a specific order, and only with respect to the variables assigned to them. This order is repeated until the whole problem has converged to its final value.

In the FET parameter-extraction problem, the global-error function that is to be optimized is defined as

$$F(\underline{x}) = \sum_{i=1}^{N} e_i(\underline{x})^2 \qquad (1)$$

where

$$e_i(\underline{x}) = \frac{1}{\sigma_{11}}\left|R_i^{11}(\underline{x}) - S_i^{11}\right| + \frac{1}{\sigma_{21}}\left|R_i^{21}(\underline{x}) - S_i^{21}\right|$$
$$+ \frac{1}{\sigma_{12}}\left|R_i^{12}(\underline{x}) - S_i^{12}\right| + \frac{1}{\sigma_{22}}\left|R_i^{22}(\underline{x}) - S_i^{22}\right|. \qquad (2)$$

In (1), $e_i(\underline{x})$ is the difference between the measured and modeled s-parameters at frequency point i, $\underline{x}$ is the parameter vector containing the element values of the circuit shown in Fig. 1, and N is the number of frequency points. Equation (2) shows the definition of $e_i(\underline{x})$, where R_i is the modeled s-parameter, S_i is the measured s-parameter at frequency point i, and σ is a normalization constant equal to the magnitude of the largest measured s-parameter value.

The difference between each of the four measured and modeled s-parameters is defined as a subfunction. Bandler and Zhang [7] developed an automated approach for assigning model elements to the different subfunctions using a Monte Carlo analysis, which confirmed the assignment of model elements to subfunctions proposed by Kondoh. This assignment has been retained in the current extraction method. It is important to note that Kondoh did not optimize all the model parameters across the complete set of frequency points. This subdivision in frequency is not used in the new method of Van Niekerk and Meyer [5] since it was found to not provide any accuracy improvements. Kondoh subdivided his optimization problem to obtain eight subproblems. The level of decomposition was extended in the new optimization procedure to allow for maximum decomposition. Every model element is optimized on its own, leading to a number of one-dimensional suboptimization problems equal to the number

TABLE I
ELEMENT VALUES FOR TWO EXAMPLES

Parameter		Sub Function	Transistor 1		Transistor 2	
			Element Value	Opt. Order	Element Value	Opt Order
1.	Lg (nH)	S_{11}	0.184	Cgd	0.1	Cgd
2.	Rg (Ω)	S_{11}	0.15	Cgs	1.3	Cgs
3.	Cgs (pF)	S_{11}	0.28	Cds	0.2	Cds
4.	Ri (Ω)	S_{11}	5.23	τ	2.5	τ
5.	Cgd (pF)	S_{12}	0.0095	Ls	0.025	Ls
6.	gm (mS)	S_{21}	24.5	Lg	50	Lg
7.	τ (ps)	S_{21}	5	Rs	0.85	Ld
8.	Rds (Ω)	S_{22}	552	Ld	188	Rs
9.	Cds (pF)	S_{22}	0.0988	Rd	0.049	Rd
10.	Rd (Ω)	S_{22}	4	gm	1.3	gm
11.	Ld (nH)	S_{22}	0.151	Rds	0.1	Rds
12.	Rs (Ω)	S_{12}	1.5	Ri	1.3	Rg
13.	Ls (nH)	S_{12}	0.0054	Rg	0.08	Ri

of model elements. Table I shows the assignment of model elements to the different suboptimization problems.

The different suboptimization problems are not linearly independent [7]. In order to ensure convergence, they have to be solved in a specific sequence, which is repeated until all the model elements have converged to their final values. Kondoh determined his optimization sequence through experimentation [6], while Bandler and Zhang demonstrated an adaptive algorithm that determines the optimization sequence as a function of the subfunction error and dimensions.

In the new decomposition algorithm, the order of optimization is based on the sensitivity of the global-error function to the model elements. A principle components sensitivity analysis [5], [8] is used to order the model elements in descending order of their influence on the global-error function. The suboptimization problems are then optimized in this order. This allows the dominant model elements to get close to their correct values faster, providing the other elements with a better chance of converging.

The sensitivity analysis used to calculate the order of optimization was presented by Patterson et al. [8] to improve the conditioning of a conventional multidimensional gradient search as performed on a global-error function. In order to distinguish between maxima, minima, and saddle points, a second-order analysis was used. The error function $F(\underline{x})$ is expanded in a Taylor series around the optimal point. Equation (3) shows this series, truncated at the second term, where $\underline{x}^*$ is the optimal set of element values at which $F(\underline{x})$ is a minimum, and $\delta\underline{x}^t$ is a small difference:

$$F(\underline{x}^* + \delta\underline{x}^*) = F(\underline{x}^*) + \delta\underline{x}^t \nabla F(\underline{x}^*) + \frac{1}{2}\delta\underline{x}^t \nabla^2 F(\underline{x}^*)\delta\underline{x} \qquad (3)$$

The Hessian matrix $\nabla^2 F(\underline{x}^*)$ can be approximated with

$$H = 2J^t J \qquad (4)$$

where J is the Jacobian matrix [8]. The Jacobian matrix is defined as

$$J(x) = \begin{bmatrix} \dfrac{\partial e_1}{\partial x_1} & \dfrac{\partial e_1}{\partial x_2} & \cdots & \dfrac{\partial e_1}{\partial x_j} \\ \dfrac{\partial e_2}{\partial x_1} & \dfrac{\partial e_2}{\partial x_2} & \cdots & \dfrac{\partial e_2}{\partial x_j} \\ \vdots & \vdots & \cdots & \vdots \\ \dfrac{\partial e_N}{\partial x_1} & \dfrac{\partial e_N}{\partial x_2} & \cdots & \dfrac{\partial e_N}{\partial x_j} \end{bmatrix} \quad (5)$$

where e_i is defined in (2) and x_1 to x_j represent the model elements shown in Fig. 1.

Patterson shows that the eigenvalues of the Hessian matrix determines the sensitivity of the error function to the model elements, with those to which the error function is the most sensitive, corresponding to the largest eigenvalues. By looking at the largest components of the eigenvectors of the different eigenvalues, the model elements can be ordered from the most to least sensitive. The principle components' sensitivity analysis can, therefore, be viewed as the study of the eigenvalues and eigenvectors of the Hessian matrix. An example illustrating how this is done is presented in [5].

III. EVALUATION PROCEDURE

The aim of this paper is to provide a rigorous evaluation of the performance of the method proposed by Van Niekerk and Meyer [5], together with a comparison of the results with earlier algorithms. In order to achieve this, the algorithm was tested with simulated data to provide an absolute measure for determining accuracy. Simulated data was generated using the model in Fig. 1 and the element values shown in Table I. The element values listed in Table I describes the FLR016XV (transistor 1) and the FHR02X (transistor 2) devices from Fujitsu.[1] The s-parameters were generated at 40 frequency points, from 1 to 40 GHz.

Two tests were performed to determine the accuracy and robustness of the extraction procedure. In the first test, 100 random starting values were chosen in the search space using a uniform distribution. The search space ranged from 0.1 times to five times the nominal parameter values given in Table I. An extraction was performed using each set of starting values. A similar procedure was used for the second test, but at the start of each extraction, the simulated s-parameters were contaminated with noise. The noise was defined to have a Gaussian magnitude distribution and a uniform phase distribution between 0 and 2π radians. The magnitude distribution has a standard deviation of 2% of magnitude of the measured s-parameter. The second test used 500 random starting values.

The results of the study can be divided between those related to the optimization sequence, convergence of the algorithm, and the accuracy of the final values.

IV. THE OPTIMIZATION SEQUENCE

When the optimization sequences of different examples are compared with one another, a basic pattern emerges.

[1] *Fujitsu Microwave Semiconductors Databook*, Fujitsu Compound Semiconductors Inc., San Jose, CA, 1994.
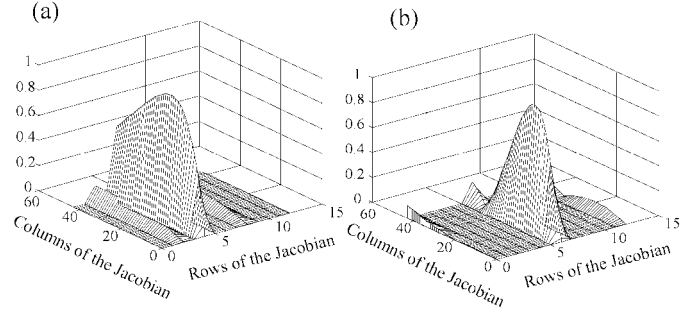

Fig. 2. Normalized amplitude of the Jacobian matrix calculated for the first example given in Table I. The rows correspond with the sequence of the model elements in Table I. (a) The whole matrix, illustrating the dominant rows corresponding to the intrinsic capacitors. (b) The matrix with these columns set to zero to show the inductors and the channel delay τ as the next level of dominant elements.

Capacitors are always dominant, followed by τ and the extrinsic inductances. The intrinsic and extrinsic resistors and the transconductance are always the least sensitive elements in the model. Table I presents the optimization sequences for two MESFET's as an example.

This behavior can be explained by a detailed look at the principle components' sensitivity analysis, which is normally viewed as a study of the Hessian matrix, but which is very similar to the Karhunen–Loéve (KL) transformation used in digital signal processing to identify the less important dimensions of a problem. The KL transform is applied to a data matrix consisting of D-dimensional vectors, which describe a system. This matrix does not need to be symmetrical. A Hessian matrix is generated by premultiplying the data matrix with its transpose, followed by an eigenvector/eigenvalue analysis. As this process is, in effect, used to analyze the data matrix, the principle components' sensitivity analysis can, in a similar vein, be said to analyze the Jacobian matrix.

Relating the sensitivities to the Jacobian matrix has the advantage of allowing a one-to-one correspondence between the rows of the matrix and the model elements, something which is lost in the calculation of the Hessian matrix. Fig. 2 shows the normalized amplitude of a representative Jacobian matrix. It is clear that the elements with the largest frequency dependence corresponds to the most prominent columns of the matrix, which will account for their prominence in the sensitivity analysis. In contrast, the importance of the frequency-independent elements, such as the resistors and the transconductance, is related more to their position in the model and their size. The position of a model element in the order of optimization is, therefore, determined by the following three factors (in order of importance):

1) frequency dependence of the element value;
2) magnitude of the element value;
3) position of the element in the model.

These three criteria provides us with rough indicators as to how the sensitivity analysis works and what really influences the order of optimization. The effect of the starting point of the search is not included in this discussion, but
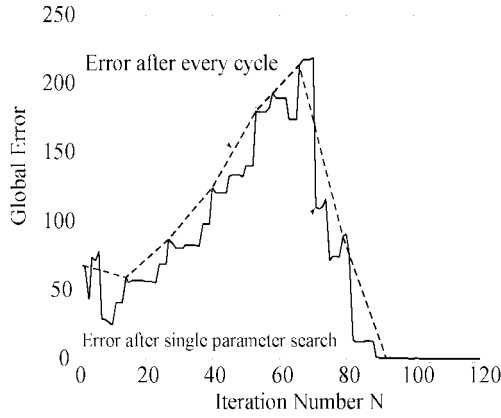
Fig. 3. The change in the global error (complete objective function) as the search progresses.

Fig. 4. The change in the element values for three typical parameters of the 13-element model as the search progresses.

differences in starting values and measured data do have small influences on the order of optimization. This is because e_i is also a function of both the model element values and the measured data [see (2)].

V. CONVERGENCE CONSIDERATIONS

No conclusive theoretical work is available to explain why the decomposed optimization process converges to the correct solution. Kondoh [6] found that the order in which the functions are to be optimized is crucial for convergence, and Bandler and Zhang [7] showed that this was due to the fact that the defined subfunctions are not linearly independent.

In this section, new results concerning the convergence of the decomposition-based algorithm are presented. The convergence of the algorithm can be divided into two regions. In the first region, there occurs what can best be described as the preconvergence maneuvering of the model elements. During this phase, the model-element values change quickly, covering the whole search space in no apparent pattern. This is accompanied by large increases and decreases in the global error, giving the impression that the search routine is hill climbing. During this phase, the model elements will frequently run into the boundaries of the optimization space, and may stay there for more than one iteration of the search. This behavior is illustrated in Figs. 3 and 4. Fig. 5 shows how the global objective function differs from the subfunction that is being optimized during phase one, and how rapidly the shape of the objective function changes from iteration to iteration. All the element values, except one, were fixed at their value after the shown number of iterations, and the free variable was varied around its current value. Fig. 5 contains both the normalized global and subfunction errors as a function of the free variable and, thus, the optimization landscape seen by the algorithm at that iteration.

The behavior during phase one can be understood by looking at the relationships between the elements. Since the dominant model elements are initially far from their correct values, they will have a large influence on one another and on the other less dominant model parameters. This will increase the dependance of the decomposed functions on one another, explaining the violent changes in element values and
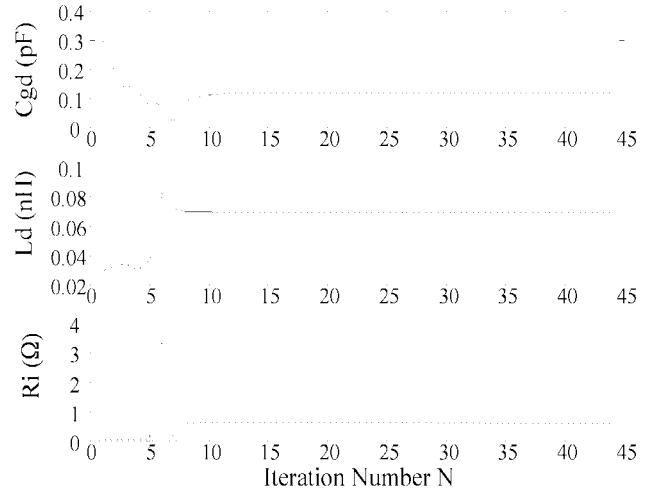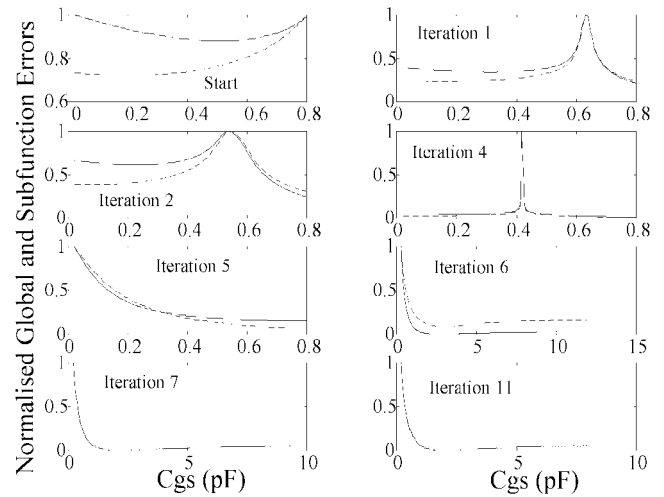

Fig. 5. The change in the landscape seen by the search routine for the element C_{gs}. The y-axis represents both the subfunction- and global-error function values, both normalized to one (– – –: subfunction error, —: global error).

the global-error function that is evident during the first few iterations of the search.

The second phase of the algorithm's convergence is also evident from Figs. 3 and 4. During this part, the algorithm converges smoothly to the final solution. A comparison of the normalized subfunction and the normalized global-error function in Fig. 5 shows that they approximate each other in the region of the solution, indicating that by minimizing the subfunction, the algorithm also minimizes the global-error function in the second phase of convergence.

The decomposed optimization algorithms are by no means globally convergent and can still be caught in local minima if the search is started too far from the global minimum. The routines are, however, far more robust than most conventional search algorithms, and are capable of covering a far larger search area. This robustness is due to the following reasons.

1) The functions that are optimized are less complex than the global-error function. This reduces the amount of local minima that pose a danger to the extraction process

TABLE II
THE AVERAGE AND MAXIMUM % ERROR IN 100 EXTRACTIONS OF
THE 13-ELEMENT FET MODEL FROM IDEAL AND NOISY DATA

| Parameter | Test 1 - Ideal Data | | | | Test 2 - Noisy Data | | | |
| | Transistor 1 | | Transistor 2 | | Transistor 1 | | Transistor 2 | |
	Avg.	Max.	Avg.	Max.	Avg.	Max.	Avg.	Max.
Cgs	0.021	0.16	0.51	4.37	0.053	1.11	0.16	7.11
Cgd	0.041	0.082	0.60	3.35	0.047	1.63	0.19	6.43
Cds	0.0042	0.028	0.74	5.61	0.02	1.44	1.50	8.24
gm	0.0019	0.14	0.055	5.30	0.017	1.10	0.70	7.55
τ	0.013	0.12	1.50	7.48	0.039	1.76	1.61	15.62
Ri	1.64	2.92	8.00	89.91	1.82	15.86	4.46	96.10
Rds	0.011	0.057	0.41	4.06	0.042	5.50	0.22	7.77
Lg	0.0072	0.015	0.007	0.19	0.032	1.09	0.089	1.68
Ld	0.0012	0.01	1.00	4.35	0.032	1.75	1.53	5.90
Ls	0.025	0.062	0.055	0.25	0.05	3.58	0.037	2.14
Rg	50.62	398	10.25	104.8	56.36	400	3.30	118.6
Rd	0.16	0.35	23.25	95.51	0.25	9.55	29.49	142.3
Rs	0.50	3.93	4.73	61.62	0.53	2.19	4.54	85.46

since the subfunctions are not evaluated simultaneously and the global-error function is not used for any decisions during the search.

2) The functions being optimized display a high degree of sensitivity to the model elements associated with them.

3) The order in which the subfunctions are repeatedly optimized leads to highly consistent extraction results. There are a large group of optimization sequences that will allow the decomposition-based optimization procedure to converge, but they are not equivalent in terms of the accuracy that can be obtained and the amount of iterations needed for the routine to converge.

The first robustness experiment was also repeated using a larger search space. Model-element starting values were chosen in the ranges of 0.05–10 times the nominal element values and 0.033–20 times the nominal element values. Both ranges caused some of the extractions in the experiments not to converge to the correct answer. These extractions were easy to identify since most of the model elements would take on values close to the optimization boundaries and, in some cases, the final global error was larger than the global error at the start of the search. However, better than 90% of the extractions converged to the correct element values.

VI. ACCURACY OF FINAL VALUES

A. Results of the Principle Components' Sensitivity-Analysis-Based Extraction Procedure

Table II contains the percentage error made in the final values of the model elements for both tests for the two transistor examples defined in Table I. Both the maximum and mean error are provided. The accuracy of the method is clear.

A novel representation of the performance of the algorithm is shown in Figs. 6 and 7. The x-axes represent 100 extractions performed during the tests, while the y-axes represent the element values. Each circle shows the starting value for that particular element for one search, while the solid line shows
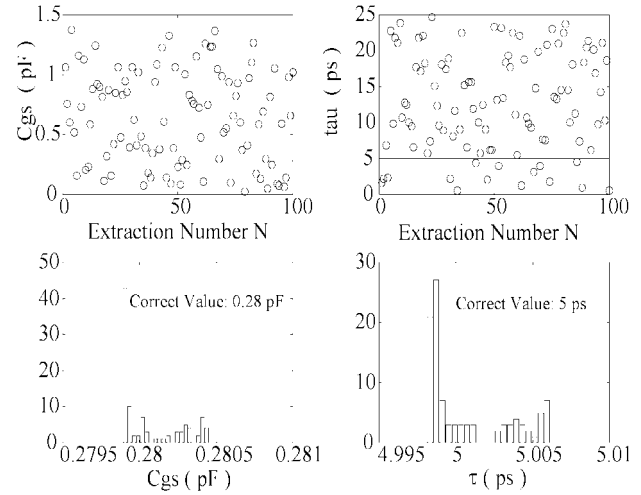

Fig. 6. Converge diagram and histogram for model elements C_{gs} and τ for test one. Circles indicate the starting values for each extraction and the solid line the final values.
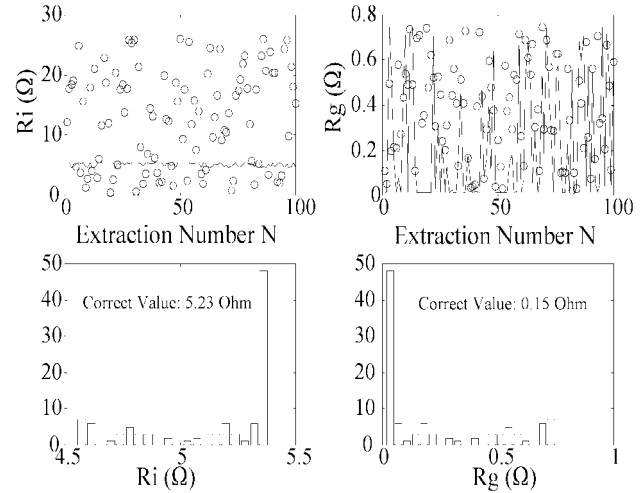

Fig. 7. Converge diagram and histogram for model elements R_i and R_g for test one. Circles indicate the starting values for each extraction and the solid line the final values.

the final value. Ideally, one should see a perfect horizontal line for the final values, representing no variation in the extracted value. The variation of the solid line from the perfect straight line gives a clear graphical view of the sensitivity of the final values to the starting values.

Results are presented for the first test for two dominant model elements (C_{gs} and τ) and for the two least sensitive elements (R_i and R_g) for transistor one. The channel resistance R_i and the parasitic gate resistance R_g are of particular interest because of their positions in the model and their electrical equivalence. This makes the determination of their individual values very difficult [8].

A study of the information obtained from various extraction examples revealed the following.

1) Not all parameter-extraction problems can be solved with equal accuracy. This is similar to conclusions reached by Curtise [1].

2) All the parameters, with the exception of R_g, could be extracted with high accuracy.

TABLE III
THE AVERAGE AND MAXIMUM % ERROR IN 100 EXTRACTIONS OF
THE 13-ELEMENT FET MODEL FROM PERFECT AND NOISY
DATA WHEN R_g IS FIXED AT ITS CORRECT VALUE

| | Test 1 - Perfect Data | | | | Test 2 - Noisy Data | | | |
| | Transistor 1 | | Transistor 2 | | Transistor 1 | | Transistor 2 | |
Parameter	Avg.	Max.	Avg.	Max.	Avg.	Max.	Avg.	Max.
Cgs	0.0014	0.0095	0.022	0.37	0.0322	1.11	0.11	2.66
Cgd	0.0008	0.0094	0.21	0.42	0.0035	1.75	0.22	3.34
Cds	0.002	0.0121	0.91	4.89	0.0219	1.43	1.51	7.72
gm	0.0009	0.0059	0.37	2.22	0.0012	1.07	0.66	4.45
τ	0.0028	0.0117	1.01	5.27	0.0234	1.78	1.73	12.51
Ri	0.0002	0.0028	2.16	8.78	0.0023	5.60	2.75	21.49
Rds	0.0071	0.0372	0.11	1.18	0.0605	5.70	0.27	2.97
Lg	0.0005	0.0058	0.034	0.14	0.0257	1.10	0.12	1.72
Ld	0.0023	0.0155	0.98	4.60	0.0305	1.72	1.53	6.18
Ls	0.0023	0.0169	0.0047	0.36	0.0310	3.65	0.089	2.18
Rg	0	0	0	0	0	0	0	0
Rd	0.0134	0.0851	19	93.15	0.0542	9.35	30.27	130.86
Rs	0.0019	0.0120	2.75	17.27	0.0231	1.98	5.01	42.43

3) The accuracy with which the parasitic resistances R_d and R_s and the channel resistance R_i could be found varied from example to example.

4) Most of the extracted element values occur in a very narrow range around the correct value. However, from the histograms, it can be observed that the lowest concentration of extracted-parameter values occurs in the middle (near the correct element value) of the distribution, which is somewhat contrary to the behavior of normal statistical processes.

The unexpected distribution is also observed in the results of the second robustness test where noise is added to the simulated measurements. Although the added noise has a Gaussian magnitude distribution, the distribution of the parameter values is decidedly non-Gaussian. The following two mechanisms are mainly responsible for this phenomenon.

1) The Distribution of Local Minima Close to the Global Minimum: This is a phenomena that is dependent on the shape of the multidimensional error function, and it cannot be described using statistical methods [10]. As these minima are very close to the global minimum, one can expect the algorithm to converge to these minima more often than to the correct value.

2) The Low Level of Added Noise: These perturbations can be described with statistical distributions, but in order for the histograms of the extracted element values to reflect these distributions, the noise has to be the dominant factor that causes extraction errors.

The second test also illustrates that the uncertainty in the determination of the less dominant parameters greatly increases in the presence of imperfections in the data. This is especially true for the parasitic resistances R_g, R_d, R_s, and the channel resistance R_i.

R_g is the model element that is the most difficult to determine because of the small influence it has on the value of the global-error function. This causes the value of R_g to
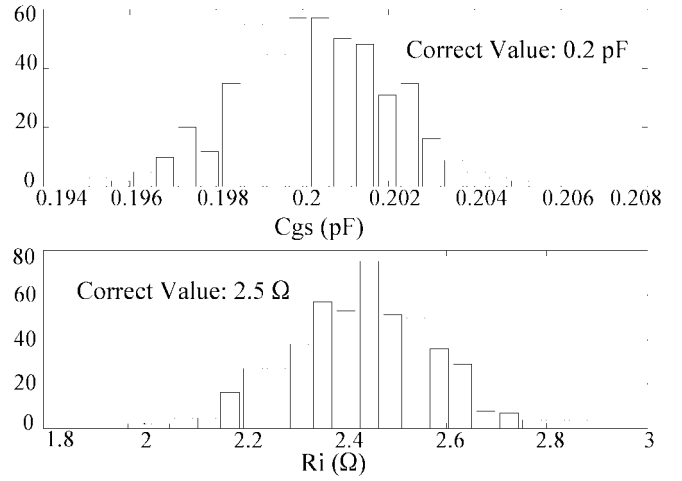

Fig. 8. The distribution of the model elements C_{gs} and R_i when the second robustness test is done with the parameter R_g fixed at its correct value. The distribution of the element values is now Gaussian.

TABLE IV
THE AVERAGE AND MAXIMUM % ERROR IN 100 EXTRACTIONS OF THE
13 ELEMENT FET MODEL FROM PERFECT DATA USING [6],
FOR R_g BOTH FREE AND FIXED AT ITS CORRECT VALUE

| | Rg Free | | | | Rg Fixed | | | |
| | Transistor 1 | | Transistor 2 | | Transistor 1 | | Transistor 2 | |
Parameter	Avg.	Max.	Avg.	Max.	Avg.	Max.	Avg.	Max.
Cgs	0.11	1.91	3.72	63.53	3.37	71.36	0.13	0.44
Cgd	0.33	4.19	6.88	159.93	1.97	41.47	0.0027	0.42
Cds	0.24	3.24	7.03	90	0.40	6.31	1.036	4.54
gm	4.66	71.66	8.30	52.68	7.35	84.89	0.50	2.11
tau	26.13	400	10.78	393.98	29.37	400	0.16	4.80
Ri	9.19	90	124.4	400	4.44	90	1.60	7.20
Rds	0.86	13.49	5.61	138.27	1.41	15.27	0.28	1.33
Lg	0.085	4.60	2.76	51.33	11.63	269.78	0.032	0.22
Ld	0.39	5.57	12.87	90	0.95	15.57	0.93	4.17
Ls	0.36	6.73	2.34	45.33	13.48	315.86	0.11	0.45
Rg	163.84	400	134.34	400	0	0	0	0
Rd	0.71	20.55	221.48	400	0.63	25.02	19.50	84.11
Rs	2.47	19.85	35.47	400	1.85	38.43	3.99	17.47

be obscured by measurement uncertainty and the small errors made in the extraction of the other model elements. Due to this, there is no correlation between the value of the global-error function and the error made in the extraction of R_g. Despite this, Novotny and Kompa [11] found that when R_g is fixed at its correct value, and there is a large increase in the overall extraction accuracy. Table III shows the results of the robustness tests performed with R_g fixed at its correct value, while Fig. 8 shows the distributions of the extracted values for the elements C_{gs} and R_i for transistor two. It should be noted that the distributions are now Gaussian, indicating that the influence of the local minima has been greatly reduced.

B. Experimental Results for the Algorithm of Kondoh [6]

The first robustness test was also performed on the original algorithm proposed by Kondoh. The test was done with R_g as part of the optimization problem and repeated with R_g fixed at its correct value. Table IV contains the results. Figs. 9 and 10
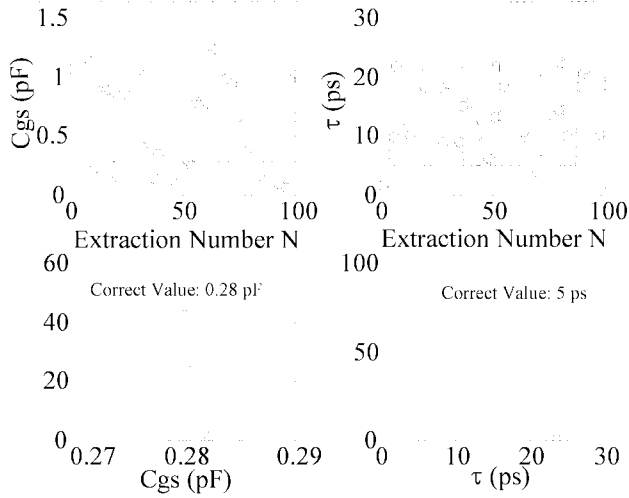
Fig. 9. Convergence diagram and histograms for the first robustness test using the Kondoh algorithm for C_{gs} and τ.
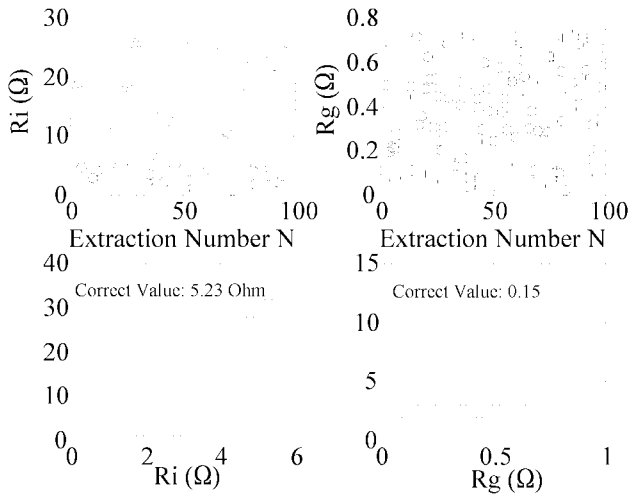

Fig. 10. Convergence diagram and histograms for the first robustness test using the Kondoh algorithm for R_i and R_g.

provides a graphical representation of the extraction results for the dominant elements C_{gs} and τ, and for the two insensitive elements R_i and R_g for transistor one. It can clearly be seen that τ gets caught in local minima for quite a number of starting values. Although not being the least sensitive element in the model, τ is optimized last in the sequence proposed by Kondoh. This has the added effect of causing R_i to be determined incorrectly.

The results summarized in Table IV can be misleading. The algorithm described in [6] can provide accurate extractions, but due to the fact that it is more susceptible to local minima, it gets caught far from the correct solution more often. This leads to a higher average error.

A comparison of the results in Table IV and Figs. 9 and 10 with those presented in Table II and Figs. 6 and 7 reveals the following: the sensitive elements of the model that are placed low down in the optimization sequence in [6] is more likely to get caught in local minima than in the new algorithm. This will also cause the insensitive elements to be determined incorrectly. Fixing R_g at its correct value improves

the extraction accuracy obtained for both methods, but it does not prevent the sensitive elements from getting stuck in local minima for the Kondoh algorithm.

Although the extraction accuracy obtained with [6] is still very high for most of the extractions in test one, the procedure is not as robust or accurate as the new adaptive optimization sequence.

VII. CONCLUSION

Rigorous test results have been presented for the new adaptive decomposition-based optimization procedure proposed in [5]. The results show the accuracy and insensitivity of the parameter extractor to optimization starting values. Robustness tests were performed on the procedure using user-defined data, allowing an absolute level of accuracy to be determined. The tests show that care has to be taken when presenting the extraction results since normal statistical descriptions such as mean and standard deviation may not be appropriate due to the shape of the histograms. The tests also confirm the large influence that the element R_g has on the extraction accuracy that can be obtained and the difficulty of determining this element from only one set of measured s-parameters. The convergence of the new procedure and the role this plays in the improved robustness of this method have been discussed.

The results presented in this paper has been generated with in-house software written in FORTRAN 77 and compiled on a variety of systems. Executing the program on an SGI Power Indigo workstation with a 195-MHz R10000 processor typically took about 16 s for one extraction.

REFERENCES

[1] W. R. Curtise and R. L. Camisa, "Self-consistent GaAs FET models for amplifier design and device diagnostics," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-32, pp. 1573–1578, Dec. 1984.
[2] G. Dambrine, A. Cappy, F. Heliodore, and E. Playez, "A new method for determining the FET small-signal equivalent circuit," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 1151–1159, July 1988.
[3] F. Lin and G. Kompa, "FET model parameter extraction based on optimization with multiplane data-fitting and bidirectional search—A new concept," *IEEE Trans. Microwave Theory Tech.*, vol. 42, pp. 1114–1121, July 1994.
[4] K. Shirakawa, H. Oikawa, T. Shimura, Y. Kawasaki, Y. Ohashi, T. Saito, and Y. Daido, "An approach to determining an equivalent circuit for HEMT's," *IEEE Trans. Microwave Theory Tech.*, vol. 43, pp. 499–503, Mar. 1995.
[5] C. van Niekerk and P. Meyer, "A new approach for the extraction of an FET equivalent circuit from measured S-parameters," *Microwave Opt. Technol. Lett.*, vol. 11, no. 5, pp. 281–284, Apr. 1996.
[6] H. Kondoh, "An accurate FET modeling from measured S-parameters," in *IEEE MTT-S Symp. Dig.*, Baltimore, MD, June 1986, pp. 377–380.
[7] J. W. Bandler and Q.-J. Zhang, "An automatic decomposition approach to optimization of large microwave systems," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-35, pp. 1231–1239, Dec. 1987.
[8] A. D. Patterson, V. F. Fusco, J. J. McKeown, and J. A. C. Stewart, "A systematic optimization strategy for microwave device modeling," *IEEE Trans. Microwave Theory Tech.*, vol. 41, pp. 395–405, Mar. 1993.
[9] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
[10] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C—The Art of Scientific Computing*. Cambridge, U.K.: Cambridge Univ. Press, 1988.
[11] G. Kompa and M. Novotny, "Highly consistent FET model parameter extraction based on broadband S-parameter measurements," in *IEEE MTT-S Symp. Dig.*, Albuquerque, NM, 1992, pp. 293–296.

Cornell van Niekerk (S'96) was born in Swellendam, South Africa, in 1971. He received the B.Ing. and M.Eng. degrees from the University of Stellenbosch, Stellenbosch, South Africa, in 1993 and 1996, and is currently working toward the Ph.D. degree, involved with the problem of experimental MESFET model extraction.

In 1998, he joined the staff at the Department of Electrical and Electronic Engineering, University of Stellenbosch, where he is currently a Lecturer. His current research interests are the extraction of transistor CAD models, cryogenic circuits, and hybrid high-temperature superconductor circuits.

Petrie Meyer (S'87–M'88) received the M.Eng. and Ph.D. (Eng.) degrees from the University of Stellenbosch, Stellenbosch, South Africa, in 1986 and 1995, respectively.

In 1988, he joined the Department of Electrical and Electronic Engineering, University of Stellenbosch, where he is currently a Senior Lecturer. His main interests include hybrid numerical electromagnetic (EM) techniques, with specific attention to the method-of-lines and mode-matching techniques, passive microwave circuits, and design algorithms.

Dr. Meyer has served as chair for the IEEE South African Section in 1997, and as technical chair for the 1999 IEEE Region 8 AFRICON Conference.