A NOVEL APPROACH TO STATISTICAL MODELING USING CUMULATIVE PROBABILITY DISTRIBUTION FITTING

J.W. Bandler, R.M. Biernacki, Q. Cai and S.H. Chen

Simulation Optimization Systems Research Laboratory
and Department of Electrical and Computer Engineering
McMaster University, Hamilton, Ontario, Canada L8S 4L7

ABSTRACT

A novel approach to statistical modeling is presented. The statistical model is directly extracted by fitting the cumulative probability distributions (CPDs) of the model responses to those of the measured data. This new technique is based on a solid mathematical foundation and, therefore, should prove more reliable and robust than the existing methods. The approach is illustrated by statistical MESFET modeling based on a physics-oriented model which combines the modified Khatibzadeh and Trew model and the Ladbrooke model (KTL). The approach is compared with the established parameter extraction/postprocessing approach (PEP) in the context of yield verification.

INTRODUCTION

Yield analysis and optimization which take into account the manufacturing tolerances and model uncertainties have been recently addressed in microwave CAD, e.g., [1-4]. Accurate and reliable statistical modeling is a prerequisite for accurate yield analysis and optimization [1].

In our previous work [5], we established the parameter extraction/postprocessing approach (PEP) to statistical modeling. Optimization is applied to extract parameters of individual devices by fitting the simulated responses to the corresponding measured data. The parameter statistics, i.e., the mean values, standard deviations, discrete distribution functions and the correlation matrix, are then obtained by postprocessing the resulting models. That approach strongly relies on the uniqueness of the parameter extraction process and, therefore, the resulting statistical models may not reflect the actual distribution of measurement data, even if the fit of the simulated responses to the corresponding measurements for individual device models is excellent.

In this paper we propose a novel approach to statistical modeling. The statistical model is determined by fitting the cumulative probability distributions (CPDs) [6] of the model responses to those of the measured data. Efficient ℓ_1 optimization [7] is used for CPD fitting. The optimization variables include the mean values and standard deviations of the statistical parameters. Thus, the model parameter statistics are obtained directly instead of from postprocessing a set of individually extracted models.

The new technique was applied to statistical modeling of a MESFET. The statistical model used is based on a physics-oriented model which combines the modified Khatibzadeh and Trew model and the Ladbrooke model (KTL) [8].

The resulting model was tested using the yield verification technique presented in [2]. Monte Carlo simulation of a broadband small-signal amplifier was performed and compared using the new model and the data for several different specifications.

HarPE [9] and OSA90/hope [10] were used to implement the new technique and to carry out the calculations presented in this paper.

This work was supported in part by Optimization Systems Associates Inc. and in part by the Natural Sciences and Engineering Research Council of Canada under Grants OGP0042444, OGP0007239 and STR0117819.

The authors are also with Optimization Systems Associates Inc., P.O. Box 8083, Dundas, Ontario, Canada L9H 5E7.

DEFINITION OF CPD AND MATCHING ERROR

Given a sample of data $S = [X_1\ X_2\ \dots\ X_n]^T$, the measured CPD of S, denoted by $C(x)$, is defined as

$$C(x) = \frac{n_x}{n} \tag{1}$$

where n_x is the number of data points in S which are smaller than or equal to x. When n is adequately large $C(x)$ provides a very good approximation to the theoretical probability distribution from which the sample was drawn. Therefore, we can test whether two samples of data come from the same probability distribution by comparing their measured CPDs.

Consider two samples of data $S_a = [X_{a1}\ X_{a2}\ \dots\ X_{an_a}]^T$ and $S_b = [X_{b1}\ X_{b2}\ \dots\ X_{bn_b}]^T$. We can calculate their corresponding CPDs $C_a(x)$ and $C_b(x)$ using (1). The distance between the two CPDs at the point x is

$$D_{ab}(x) = |\, C_a(x) - C_b(x)\, | \tag{2}$$

The matching error between the two CPDs can be defined as

$$e_{ab} = \int_{-\infty}^{\infty} D_{ab}(x)\,dx \tag{3}$$

If we merge S_a and S_b to form $S_c = [X_{c1}\ X_{c2}\ \dots\ X_{cn_c}]^T$, $n_c = n_a + n_b$, with all the points sorted in ascending order, the calculation of e_{ab} becomes

TU 3F

CH3389-4/94/0000-0385$01.00 © 1994 IEEE

$$e_{ab} = \sum_{i=1}^{n_c - 1} D_{ab}(X_{ci})\,(X_{c(i+1)} - X_{ci}) \tag{4}$$

which is the absolute value of the area between the two CPDs, as shown in Fig. 1.

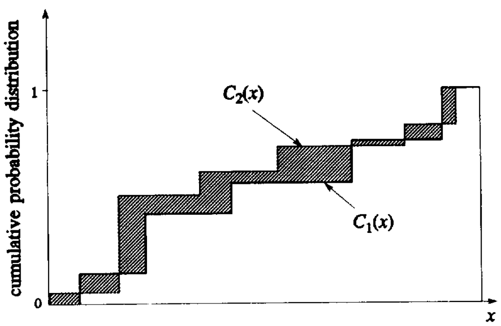

Fig. 1 Two cumulative probability distributions and their matching error (shaded area).

FORMULATION OF CPD FITTING FOR STATISTICAL MODELING

Suppose that the measurement data contains n_r measured responses for n_{mo} manufactured outcomes. For each measured response we thus have the sample

$$S_i = [X_{i1}\ X_{i2}\ \dots\ X_{in_{mo}}]^T, \quad i = 1, 2, \dots, n_r \tag{5}$$

The statistical model is simulated by Monte Carlo analysis with n_{so} outcomes and for each simulated response corresponding to S_i we have the sample

$$R_i(\phi) = [R_{i1}(\phi)\ R_{i2}(\phi)\ \dots\ R_{in_{so}}(\phi)]^T \tag{6}$$

where $\phi = [\phi_1\ \phi_2\ \dots\ \phi_{n_\phi}]^T$ is the set of optimization variables such as the mean values and standard deviations of a normal distribution, the nominal values and tolerances of a uniform distribution. For each pair S_i and R_i we calculate their CPDs using (1), the difference between these two CPDs using (2), and finally the matching error $e_i(\phi)$ using (4). Let

$$e(\phi) = [e_1(\phi)\ e_2(\phi)\ \dots\ e_{n_r}(\phi)]^T \tag{7}$$

then the optimization problem of CPD fitting for statistical modeling can be defined as

$$\underset{\phi}{\text{minimize}}\quad U(\phi) \triangleq H[e(\phi)] \tag{8}$$

where $H[e(\phi)]$ represents a norm of $e(\phi)$ such as the ℓ_1, ℓ_2 or the Huber norm. In our CPD fitting we have used the ℓ_1 norm, which can be written as

$$H[e(\phi)] = \sum_{i=1}^{n_r} |\ e_i(\phi)\ | \tag{9}$$

STATISTICAL MODEL EXTRACTION

The proposed statistical modeling technique of CPD fitting was applied to a sample of GaAs MESFET data which was obtained by aligning the Plessey wafer measurements to a consistent bias condition [8]. There were 35 data sets (devices) containing the small-signal S parameters measured at the frequencies from 1 to 21 GHz with 2 GHz step under the bias condition of $V_{gs} = -0.7$ V and $V_{ds} = 5$ V.

The KTL model [8] was selected for statistical modeling. The attractive statistical characteristics of the KTL model have been presented by Bandler *et al.* [2,8] using the method of multi-device parameter extraction and statistical postprocessing. That method was also used here for finding the starting point for optimization.

We considered 16 statistical parameters assuming normal distributions. This resulted in 32 optimization variables, namely all mean values and standard deviations. The initial values for the means and standard deviations were estimated from multi-device parameter extraction and statistical postprocessing based on 15 devices. The resulting correlation matrix was used to represent the correlations between the statistical parameters. By applying CPD fitting we obtained the KTL model parameter values listed in Table I. The CPDs of the real part of S_{21}, Re{S_{21}}, at 11 GHz from the data and from the statistical KTL model before and after optimization are shown in Fig. 2. The mean values and the standard deviations of Re{S_{21}} versus frequency from the data and from the model before and after optimization are depicted in Figs. 3 and 4, respectively. From Fig. 2 we can see that after optimization the CPD matching between the data and the KTL model is significantly improved. The mean values and standard deviations of model responses after optimization are also closer to those of the data, as indicated in Figs. 3 and 4.

TABLE I
CPD OPTIMIZED KTL MODEL PARAMETERS

Parameter	Mean	σ(%)	Parameter	Mean	σ(%)
$L(\mu m)$	0.4685	3.57	C_{ds}(pF)	0.0547	1.58
$a(\mu m)$	0.1308	5.19	C_{ge}(pF)	0.0807	5.92
$N_d(m^{-3})$	2.3×10^{23}	3.25	C_{de}(pF)	0.0098	6.22
v_{sat}(m/s)	10.5×10^4	2.27	C_x(pF)	2.4231	4.03
$\mu_0(m^2/Vns)$	6.5×10^{-10}	2.16	$Z(\mu m)$	300	*
L_{G0}(nH)	0.0396	10.9	ε	12.9	*
$R_d(\Omega)$	1.2867	4.32	V_{b0}(V)	0.6	*
$R_s(\Omega)$	3.9119	1.91	$r_{01}(\Omega/V^2)$	0.35	*
$R_g(\Omega)$	8.1718	0.77	$r_{02}(\Omega)$	2003	*
L_d(nH)	0.0659	5.74	$r_{03}(\Omega)$	7.0	*
L_s(nH)	0.0409	5.49	a_0	1.0	*
$G_{ds}(1/\Omega)$	3.9×10^{-3}	1.78			

L is the gate length, a the channel thickness, N_d the doping density, v_{sat} the saturation electron drift velocity, μ_0 the low-field mobility of GaAs, L_{g0} the inductance from gate bond wires and pads, Z the gate width, ε the dielectric constant and V_{b0} the zero-bias barrier potential. r_{01}, r_{02}, r_{03} and a_0 are fitting coefficients. R_d, R_s, R_g, L_d, L_s, G_{ds}, C_{ds}, C_{ge}, C_{de} and C_x are extrinsic parameters.

σ denotes standard deviation.

* Assumed fixed (non-statistical) parameters.

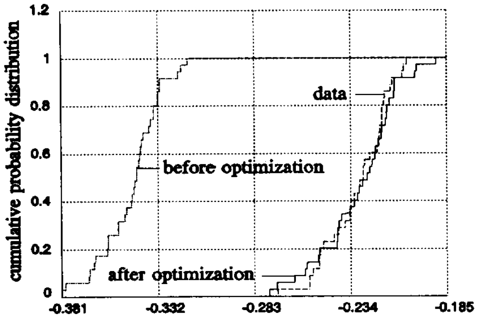

Fig. 2 CPDs of Re{S_{21}} at 11 GHz from data (---) and from the statistical KTL model before (-•-) and after (—) CPD matching.

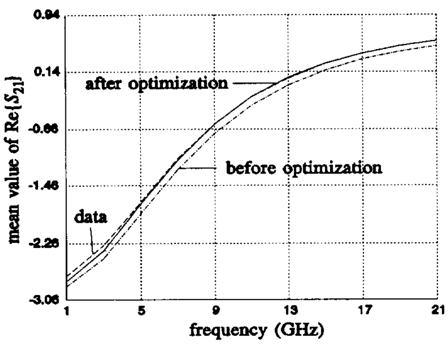

Fig. 3 Mean values of Re{S_{21}} versus frequency from data (---) and from the statistical KTL model before (-•-) and after (—) CPD matching.

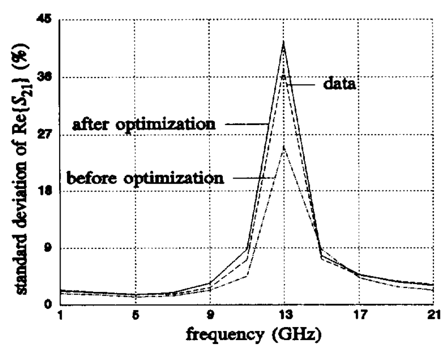

Fig. 4 Standard deviations of Re{S_{21}} versus frequency from data (---) and from the statistical KTL model before (-•-) and after (—) CPD matching.

In order to compare these results with the PEP method we also performed statistical modeling using multi-device parameter extraction and postprocessing based on the same 35 data sets. The KTL model parameter values obtained by the PEP method are listed in Table II. The CPDs of Re{S_{21}} at 11 GHz for both models are plotted in Fig. 5 together with the corresponding CPDs from the data. The mean values and the standard deviations of Re{S_{21}} versus frequency are shown in Figs. 6 and 7, respectively. From Fig. 5 we can observe that the CPD matching of Re{S_{21}} of the KTL model obtained by CPD fitting is better than that obtained by the PEP method. From Figs. 6 and 7 we see that our new approach gives better standard deviation match though the mean value matches of both models are similar.

TABLE II
PEP OPTIMIZED KTL MODEL PARAMETERS

Parameter	Mean	σ(%)	Parameter	Mean	σ(%)
L(μm)	0.5190	4.72	C_{ds}(pF)	0.0486	2.84
a(μm)	0.1584	8.20	C_{ge}(pF)	0.0698	9.72
N_d(m^{-3})	2.2x10^{23}	4.68	C_{de}(pF)	0.0109	10.5
v_{sat}(m/s)	10.7x10^4	2.24	C_x(pF)	3.3046	3.69
μ_0(m^2/Vns)	5.9x10^{-10}	1.89	Z(μm)	300	*
L_{G0}(nH)	0.0331	12.2	ε	12.9	*
R_d(Ω)	1.1190	9.43	V_{b0}(V)	0.6	*
R_s(Ω)	3.3226	2.69	r_{01}(Ω/V^2)	0.35	*
R_g(Ω)	6.6209	1.59	r_{02}(Ω)	2003	*
L_d(nH)	0.0533	10.4	r_{0S}(Ω)	7.0	*
L_s(nH)	0.0407	9.75	a_0	1.0	*
G_{ds}(1/Ω)	3.8x10^{-3}	2.51			

The parameter definitions are the same as in Table I.

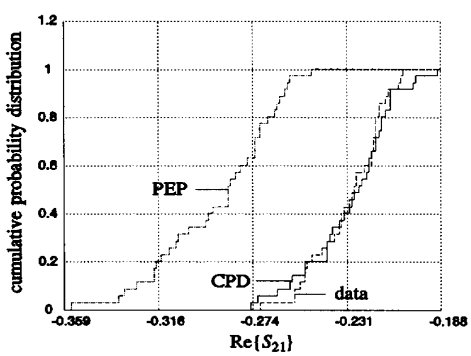

Fig. 5 CPDs of Re{S_{21}} at 11 GHz from data (---) and from the CPD (—) and PEP (-•-) statistical KTL models.

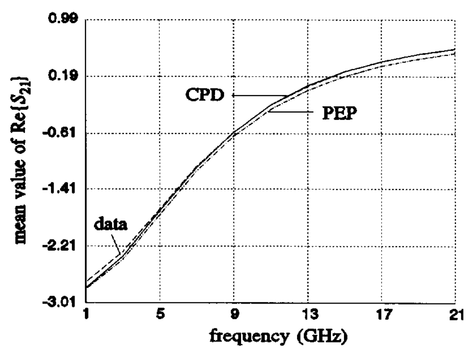

Fig. 6 Mean values of Re{S_{21}} versus frequency from data (---) and from the CPD (—) and PEP (-•-) statistical KTL models.

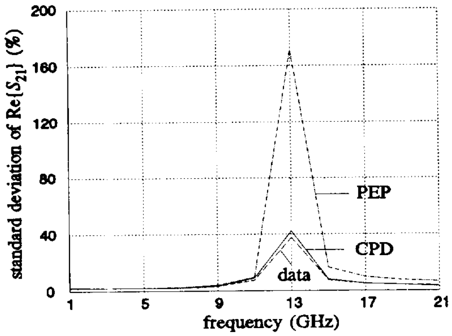

Fig. 7 Standard deviations of Re$\{S_{21}\}$ versus frequency from data (---) and from the CPD (——) and PEP (-•-) statistical KTL models.

YIELD VERIFICATION

The ultimate goal of statistical modeling is to provide accurate statistical models for yield optimization. Therefore, the statistical model can be further verified by comparing the yield estimations by the model and data [2]. To this end we performed yield verification using Monte Carlo simulation.

We considered yield optimization of a small-signal broadband amplifier used in [2]. The design was carried out using OSA90/hope [10]. The passband of the amplifier is 8 GHz - 12 GHz. Three different specifications were considered. Yield optimization was performed using the two statistical KTL models (CPD and PEP). The yields predicted by Monte Carlo simulation from the data and from both models are listed in Tables III. We can see that the yields predicted by both models are in good agreement for every specification.

TABLE III
YIELD PREDICTED BY THE KTL MODELS AND VERIFIED BY DATA

	Yield Before Optimization			Yield After Optimization		
Spec.	CPD (%)	PEP (%)	Data (%)	CPD (%)	PEP (%)	Data (%)
Spec. 1	22	26	28.6	71	69.5	77.6
Spec. 2	30	38.5	37.1	76.5	78.5	90.9
Spec. 3	64.5	67.5	76.7	98.5	93.5	99.5

Spec. 1: 7.5 dB < $|S_{21}|$ < 8.5 dB, $|S_{11}|$ < 0.5, $|S_{22}|$ < 0.5.

Spec. 2: 6.5 dB < $|S_{21}|$ < 7.5 dB, $|S_{11}|$ < 0.5, $|S_{22}|$ < 0.5.

Spec. 3: 6.0 dB < $|S_{21}|$ < 8.0 dB, $|S_{11}|$ < 0.5, $|S_{22}|$ < 0.5.

200 outcomes are used for yield prediction by the statistical KTL model, 210 for yield verification using the device data.

CONCLUSIONS

We have presented a novel approach to statistical modeling. The parameter mean values and standard deviations are directly optimized to match the cumulative probability distributions of the model responses to those of the data. This approach avoids parameter extraction of individual devices and, therefore, is not affected by possible pitfalls of the parameter extraction process. Our investigations set the stage for further research, which could include determining parameter correlations in addition to mean values and standard deviations, as well as possible extensions to other than normal distributions. In principle, the proposed method is not limited to normal distributions. Finally, we point out that although the PEP technique normally provides adequate statistical models, the new CPD technique is based on a solid mathematical foundation and, therefore, should prove more reliable and robust.

REFERENCES

[1] J.W. Bandler, R.M. Biernacki, Q. Cai, S.H. Chen, S. Ye and Q.J. Zhang, "Integrated physics-oriented statistical modeling, simulation and optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 40, 1992, pp. 1374-1400.

[2] J.W. Bandler, S. Ye, Q. Cai, R.M. Biernacki and S.H. Chen, "Predictable yield-driven circuit optimization," *IEEE Int. Microwave Symp. Digest* (Albuquerque, NM), 1992, pp. 837-840.

[3] F. Filicori, G. Ghione and C.U. Naldi, "Physics-based electron device modelling and computer-aided MMIC design," *IEEE Trans. Microwave Theory Tech.*, vol. 40, 1992, pp. 1333-1352.

[4] D.E. Stoneking, G.L. Bilbro, P.A. Gilmore, R.J. Trew and C.T. Kelley, "Yield optimization using a GaAs process simulator coupled to a physical device model," *IEEE Trans. Microwave Theory Tech.*, vol. 40, 1992, pp. 1353-1363.

[5] J.W. Bandler, R.M. Biernacki, S.H. Chen, J.F. Loman, M.L. Renault and Q.J. Zhang, "Combined discrete/normal statistical modeling of microwave devices," *Proc. 19th European Microwave Conf.* (London, England), 1989, pp. 205-210.

[6] W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, *Numerical Recipes. The Art of Scientific Computing*. Cambridge University Press: Cambridge, 1986.

[7] J.W. Bandler, S.H. Chen and S. Daijavad, "Microwave device modeling using efficient ℓ_1 optimization: a novel approach," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-34, 1986, pp. 1282-1293.

[8] J.W. Bandler, R.M. Biernacki, Q. Cai and S.H. Chen, "A robust physics-oriented statistical GaAs MESFET model," *GAAS94* (Turin, Italy), 1994.

[9] *HarPE*™, Optimization Systems Associates Inc., P.O. Box 8083, Dundas, Ontario, Canada L9H 5E7, 1993.

[10] *OSA90/hope*™, Optimization Systems Associates Inc., P.O. Box 8083, Dundas, Ontario, Canada L9H 5E7, 1993.