

**A NOVEL APPROACH TO STATISTICAL
MODELING USING CUMULATIVE PROBABILITY
DISTRIBUTION FITTING**

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

SOS-93-25-R

November 1993

© J.W. Bandler, R.M. Biernacki, Q. Cai and S.H. Chen 1993

No part of this document may be copied, translated, transcribed or entered in any form into any machine without written permission. Address enquiries in this regard to Dr. J.W. Bandler. Excerpts may be quoted for scholarly purposes with full acknowledgement of source. This document may not be lent or circulated without this title page and its original cover.

**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

Tel 905 525 9140 Ext. 24818
Fax 905 523 4407

Abstract

A novel approach to statistical modeling is presented. The statistical model is directly extracted by fitting the cumulative probability distributions 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 physics-oriented MESFET modeling.

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.

* J.W. Bandler, R.M. Biernacki and S.H. Chen are also with and Q. Cai is with Optimization Systems Associates Inc., P.O. Box 8083, Dundas, Ontario, Canada L9H 5E7, Tel 905 628 8228.

SUMMARY

Introduction

Yield analysis and optimization which take into account the manufacturing tolerances and model uncertainties have been recently addressed in microwave CAD, e.g., [1-5]. Accurate and reliable statistical modeling is a prerequisite for accurate yield analysis and optimization [1].

In our previous work [6], 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) [7] of the model responses to those of the measured data. Efficient ℓ_1 optimization [8] 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 Ladbroke model (KTL) [9].

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 two different designs and several different specifications.

HarPE [10] and OSA90/hope [11] were used to implement the new technique and to carry out the calculations presented in this paper.

Definition of CPD and Matching Error

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

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

where n_x is the number of data points in \mathcal{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 $\mathcal{S}_a = [X_{a1} X_{a2} \dots X_{an_a}]^T$ and $\mathcal{S}_b = [X_{b1} X_{b2} \dots X_{bn_b}]^T$. 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

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

The matching error between the two CPDs can be defined as

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

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

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

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

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

$$\mathcal{S}_i = [X_{i1} X_{i2} \dots X_{in_{mo}}]^T, \quad i = 1, 2, \dots, n_r \quad (5)$$

The statistical model is simulated by Monte Carlo analysis with n_{so} outcomes and for each simulated

response corresponding to \mathcal{S}_i we have the sample

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

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 \mathcal{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

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

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

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

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

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

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 [9]. 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 [9] was selected for statistical modeling. The attractive statistical characteristics of the KTL model have been presented by Bandler *et al.* [2,9] 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 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 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.

In order to compare these results with the PEP technique 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 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 S_{21} versus frequency are shown in Figs. 6 and 7, respectively. From Fig. 5 we can observe that the CPD matching of the real part of S_{21} of the KTL model obtained by CPD fitting is better than that obtained by the PEP technique, while the CPD matching of the imaginary part of S_{21} of both models are about the same. 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.

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]. Two

designs were carried out using OSA90/hope [11]. The passband of the amplifier is 8 GHz - 12 GHz in Design I and 10 GHz - 14 GHz in Design II. Three different specifications were considered in both cases. 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 in Designs I and II are listed in Tables III and IV, respectively. We can see that the yields predicted by both models are in good agreement in both designs for every specification.

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. Also the applicability and robustness of the Huber norm to CPD modeling should be investigated in addition to the ℓ_1 optimization used in the experiments reported in this paper. 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] E.M. Bastida, G.P. Donzelli and M. Pagani, "Efficient development of mass producible MMIC circuits," *IEEE Trans. Microwave Theory Tech.*, vol. 40, 1992, pp. 1364-1373.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] J.W. Bandler, R.M. Biernacki, Q. Cai, S.H. Chen and S. Ye, "A robust physics-oriented statistical GaAs MESFET model," Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, Report SOS-93-14-R, 1993.
- [10] *HarPE™*, Optimization Systems Associates Inc., P.O. Box 8083, Dundas, Ontario, Canada L9H 5E7, 1993.
- [11] *OSA90/hope™*, Optimization Systems Associates Inc., P.O. Box 8083, Dundas, Ontario, Canada L9H 5E7, 1993.

TABLE I

CPD OPTIMIZED KTL MODEL PARAMETERS

| Parameter | Mean | Std. Dev. (%) |
|--------------------------------|-------------------------|---------------|
| $L(\mu\text{m})$ | 0.4685 | 3.57 |
| $a(\mu\text{m})$ | 0.1308 | 5.19 |
| $N_d(\text{m}^{-3})$ | 2.3272×10^{23} | 3.25 |
| $v_{sat}(\text{m/s})$ | 10.4849×10^4 | 2.27 |
| $\mu_0(\text{m}^2/\text{Vns})$ | 6.50×10^{-10} | 2.16 |
| $L_{G0}(\text{nH})$ | 0.0396 | 10.9 |
| $R_d(\Omega)$ | 1.2867 | 4.32 |
| $R_s(\Omega)$ | 3.9119 | 1.91 |
| $R_g(\Omega)$ | 8.1718 | 0.77 |
| $L_d(\text{nH})$ | 0.0659 | 5.74 |
| $L_s(\text{nH})$ | 0.0409 | 5.49 |
| $G_{ds}(1/\Omega)$ | 3.92×10^{-3} | 1.78 |
| $C_{ds}(\text{pF})$ | 0.0547 | 1.58 |
| $C_{ge}(\text{pF})$ | 0.0807 | 5.92 |
| $C_{de}(\text{pF})$ | 0.0098 | 6.22 |
| $C_x(\text{pF})$ | 2.4231 | 4.03 |
| $Z(\mu\text{m})$ | 300 | * |
| ϵ | 12.9 | * |
| $V_{b0}(\text{V})$ | 0.6 | * |
| $r_{01}(\Omega/\text{V}^2)$ | 0.35 | * |
| $r_{02}(\Omega)$ | 2003 | * |
| $r_{03}(\Omega)$ | 7.0 | * |
| a_0 | 1.0 | * |

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.

* Assumed fixed (non-statistical) parameters.

TABLE II

PEP OPTIMIZED KTL MODEL PARAMETERS

| Parameter | Mean | Std. Dev. (%) |
|--------------------------------|--------------------------|---------------|
| $L(\mu\text{m})$ | 0.5190 | 4.72 |
| $a(\mu\text{m})$ | 0.1584 | 8.20 |
| $N_d(\text{m}^{-3})$ | 2.2313×10^{23} | 4.68 |
| $v_{sat}(\text{m/s})$ | 10.724×10^4 | 2.24 |
| $\mu_0(\text{m}^2/\text{Vns})$ | 5.9379×10^{-10} | 1.89 |
| $L_{G0}(\text{nH})$ | 0.0331 | 12.2 |
| $R_d(\Omega)$ | 1.1190 | 9.43 |
| $R_s(\Omega)$ | 3.3226 | 2.69 |
| $R_g(\Omega)$ | 6.6209 | 1.59 |
| $L_d(\text{nH})$ | 0.0533 | 10.4 |
| $L_s(\text{nH})$ | 0.0407 | 9.75 |
| $G_{ds}(1/\Omega)$ | 3.8367×10^{-3} | 2.51 |
| $C_{ds}(\text{pF})$ | 0.0486 | 2.84 |
| $C_{ge}(\text{pF})$ | 0.0698 | 9.72 |
| $C_{de}(\text{pF})$ | 0.0109 | 10.5 |
| $C_x(\text{pF})$ | 3.3046 | 3.69 |
| $Z(\mu\text{m})$ | 300 | * |
| ϵ | 12.9 | * |
| $V_{b0}(\text{V})$ | 0.6 | * |
| $r_{01}(\Omega/\text{V}^2)$ | 0.35 | * |
| $r_{02}(\Omega)$ | 2003 | * |
| $r_{03}(\Omega)$ | 7.0 | * |
| a_0 | 1.0 | * |

The parameter definitions are the same as listed in Table I.

* Assumed fixed (non-statistical) parameters.

TABLE III

YIELD PREDICTED BY THE KTL MODELS AND VERIFIED BY DATA IN DESIGN I
(PASSBAND: 8 GHz - 12 GHz)

| Specification | Before Yield Optimization | | | After Yield Optimization | | |
|---------------|--------------------------------|---------------------------------|-------------------------------|--------------------------------|---------------------------------|-------------------------------|
| | CPD* Predicted Yield (%) | PEP** Predicted Yield (%) | Data Verified Yield (%) | CPD* Predicted Yield (%) | PEP** Predicted Yield (%) | Data Verified Yield (%) |
| 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: $7.5 \text{ dB} < |S_{21}| < 8.5 \text{ dB}$, $|S_{11}| < 0.5$, $|S_{22}| < 0.5$.

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

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

* Obtained by cumulative probability distribution fitting.

** Obtained by multi-device parameter extraction and postprocessing.

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

TABLE IV

YIELD PREDICTED BY THE KTL MODELS AND VERIFIED BY DATA IN DESIGN II
(PASSBAND: 10 GHz - 14 GHz)

| Specification | Before Yield Optimization | | | After Yield Optimization | | |
|---------------|--------------------------------|---------------------------------|-------------------------------|--------------------------------|---------------------------------|-------------------------------|
| | CPD* Predicted Yield (%) | PEP** Predicted Yield (%) | Data Verified Yield (%) | CPD* Predicted Yield (%) | PEP** Predicted Yield (%) | Data Verified Yield (%) |
| Spec. 1 | 26 | 26.5 | 30.9 | 71 | 63 | 70.5 |
| Spec. 2 | 33 | 38.5 | 40.5 | 80 | 75.5 | 87.6 |
| Spec. 3 | 72.5 | 68 | 78.5 | 97.5 | 93.5 | 94.3 |

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

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

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

* Obtained by cumulative probability distribution fitting.

** Obtained by multi-device parameter extraction and postprocessing.

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

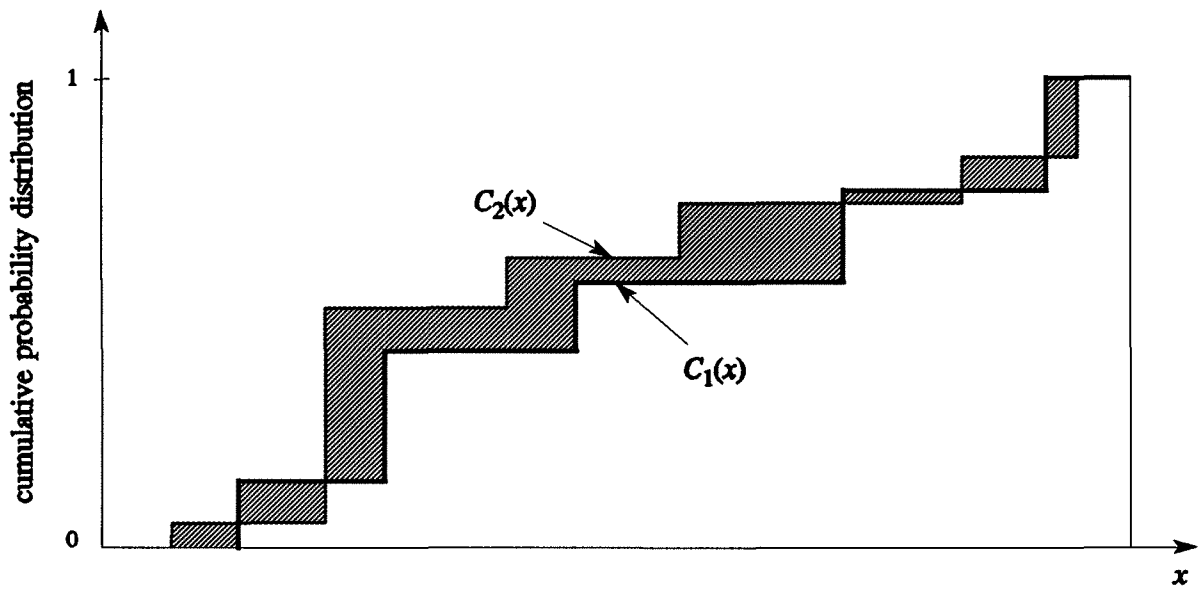
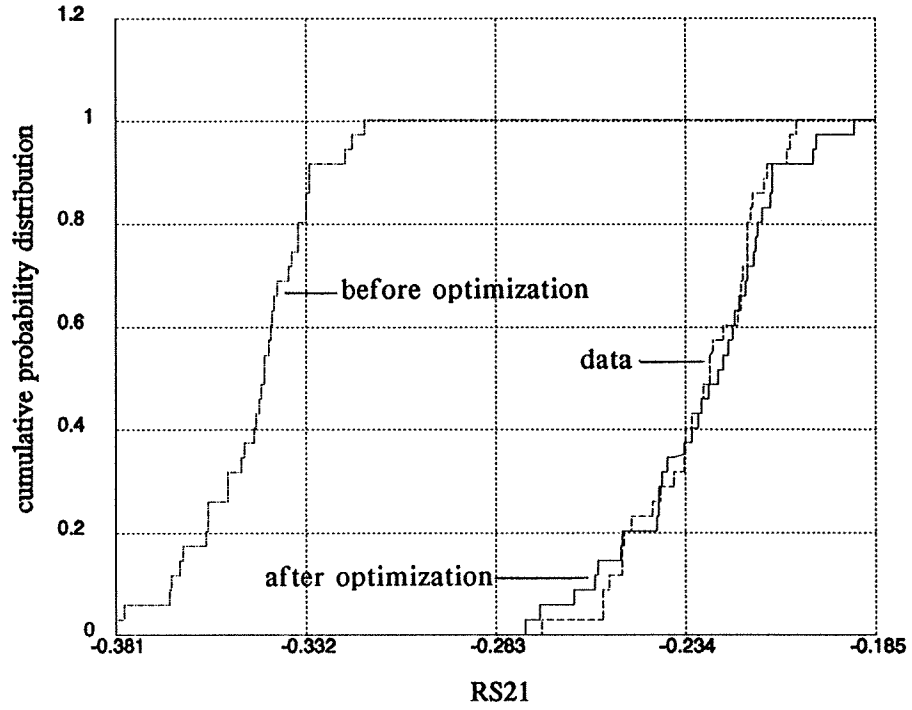
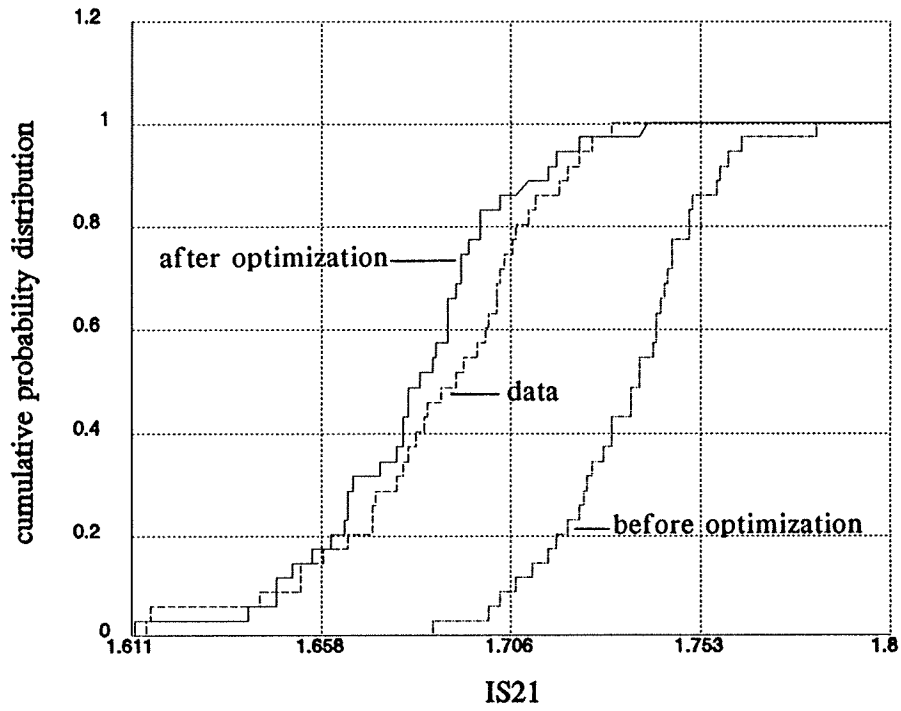


Fig. 1 Two cumulative probability distributions and their matching error (shaded area).

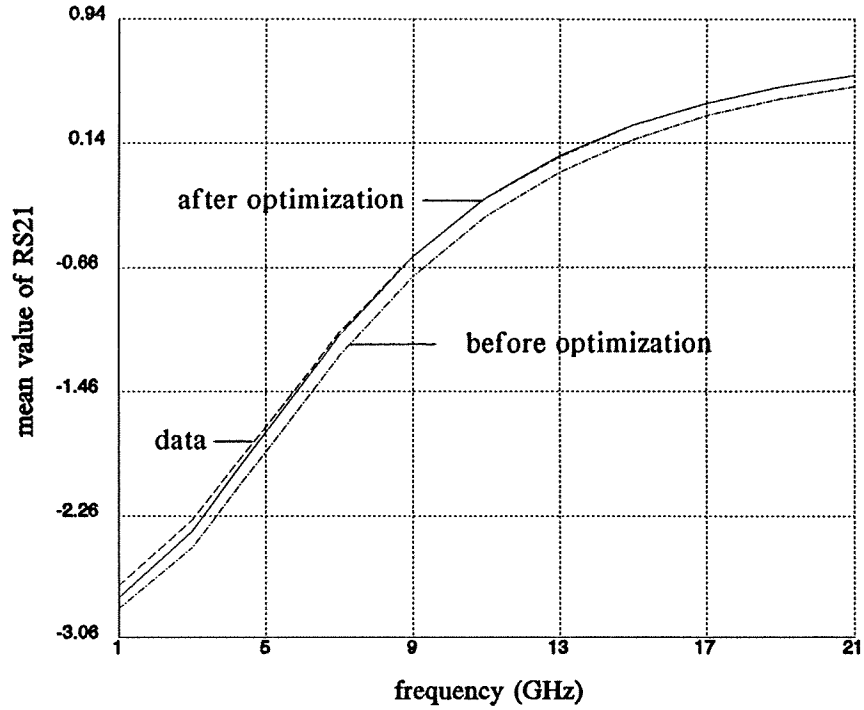


(a)

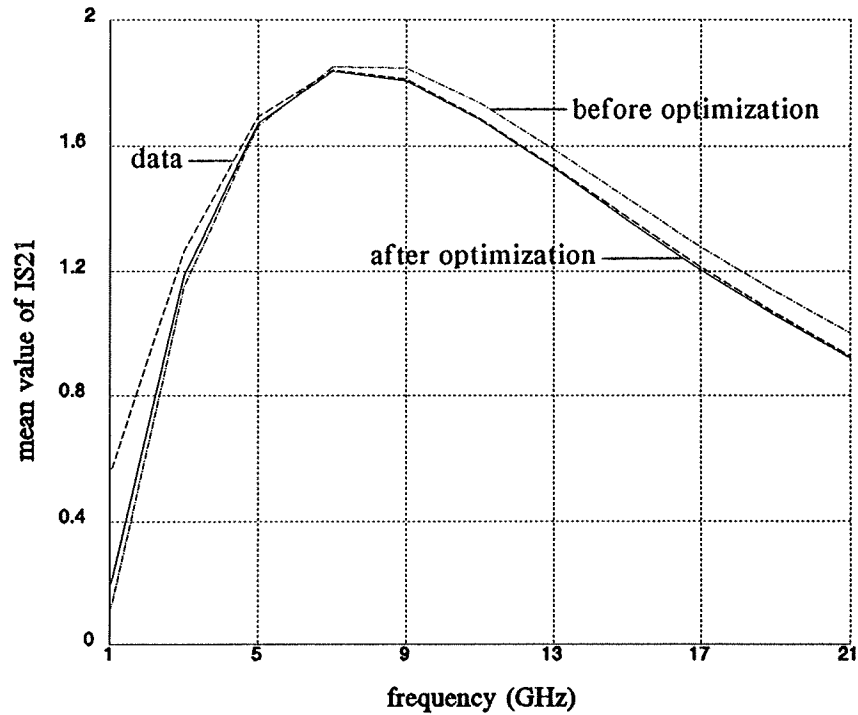


(b)

Fig. 2 CPDs of S_{21} at 11 GHz from data (---) and from statistical KTL model before (---) and after (—) CPD matching, (a) the real part and (b) the imaginary part.

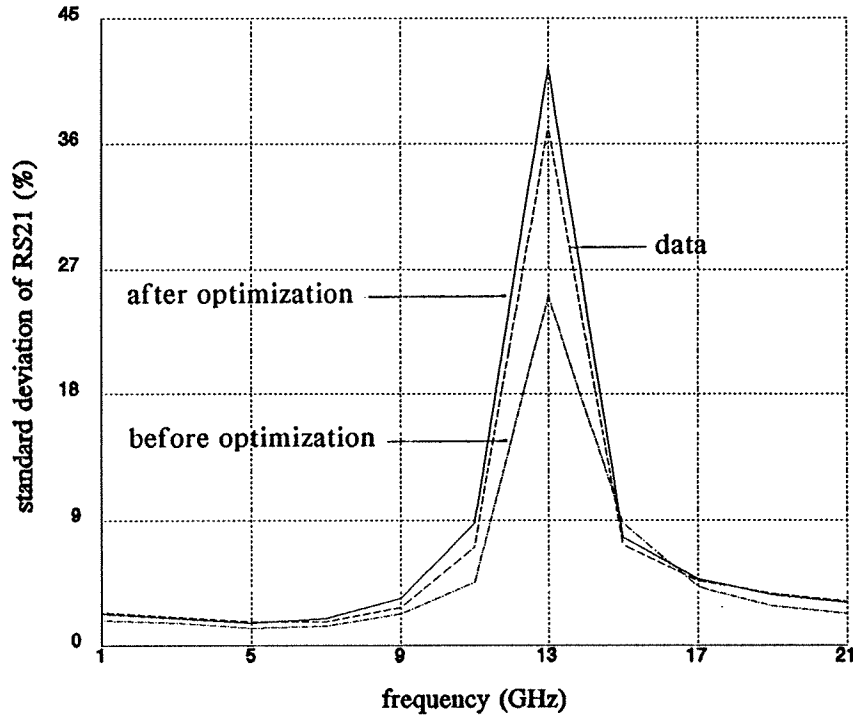


(a)

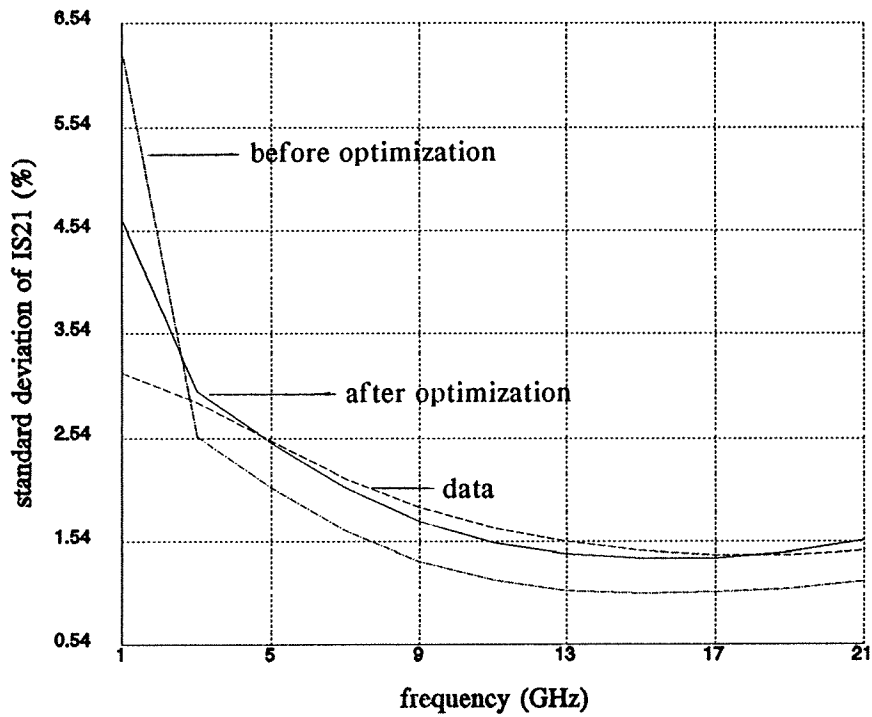


(b)

Fig. 3 Mean values of S_{21} versus frequency from data (---) and from statistical KTL model before (---) and after (—) CPD matching, (a) the real part and (b) the imaginary part.

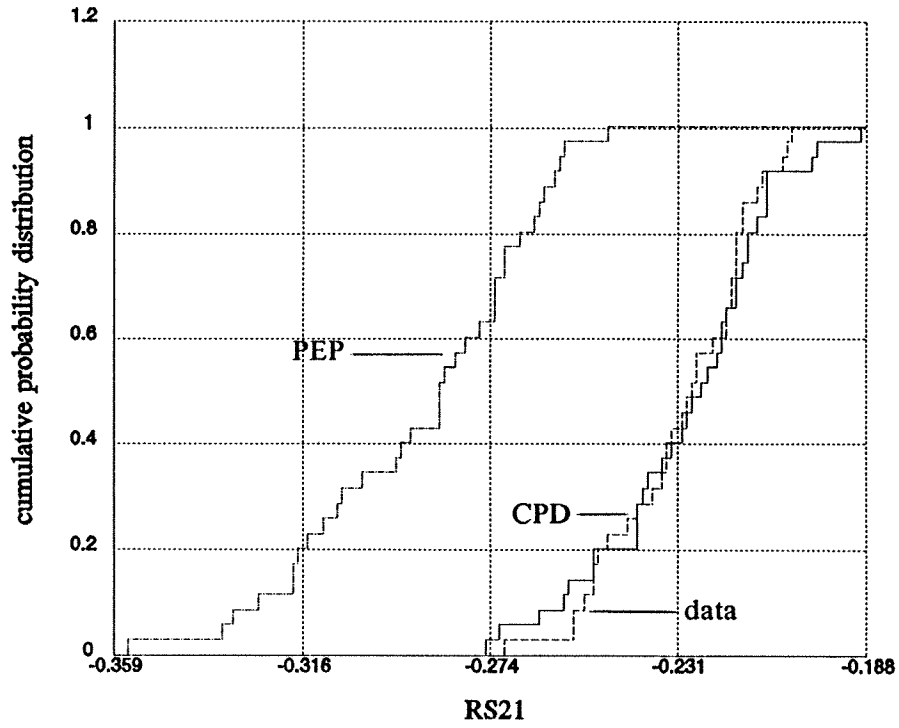


(a)

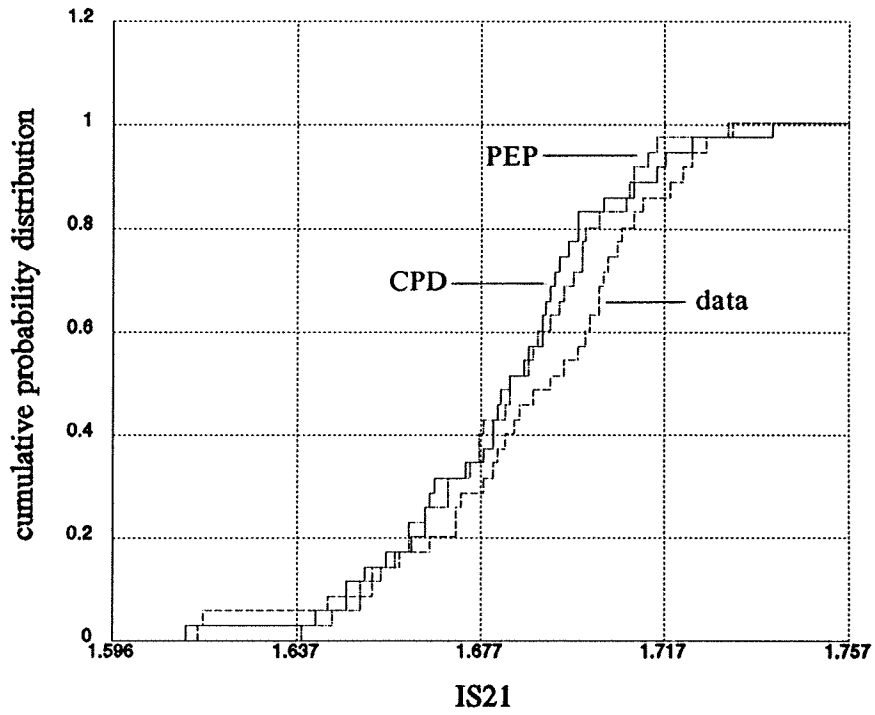


(b)

Fig. 4 Standard deviations of S_{21} versus frequency from data (---) and from statistical KTL model before (-.-) and after (—) CPD matching, (a) the real part and (b) the imaginary part.



(a)



(b)

Fig. 5 CPDs of S_{21} at 11 GHz from data (---) and from the CPD (—) and PEP (---) statistical KTL models, (a) the real part and (b) the imaginary part.

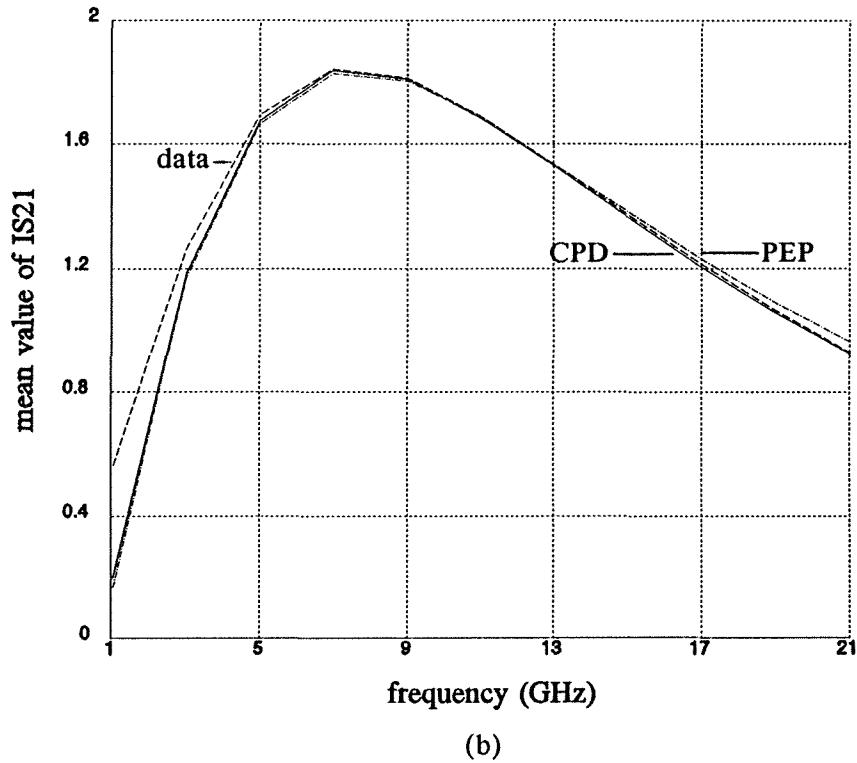
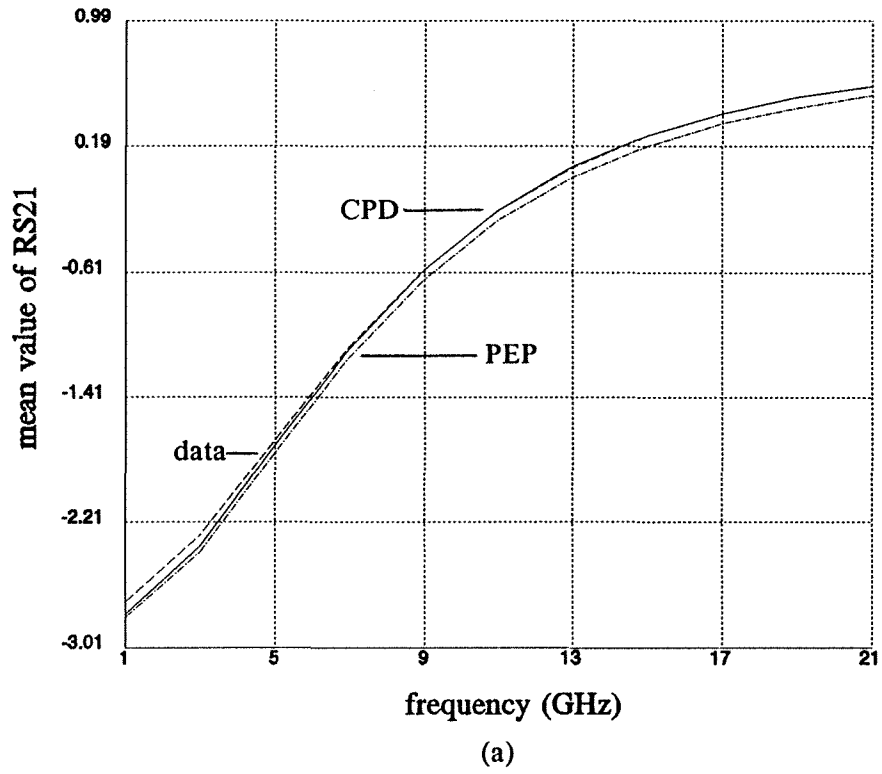
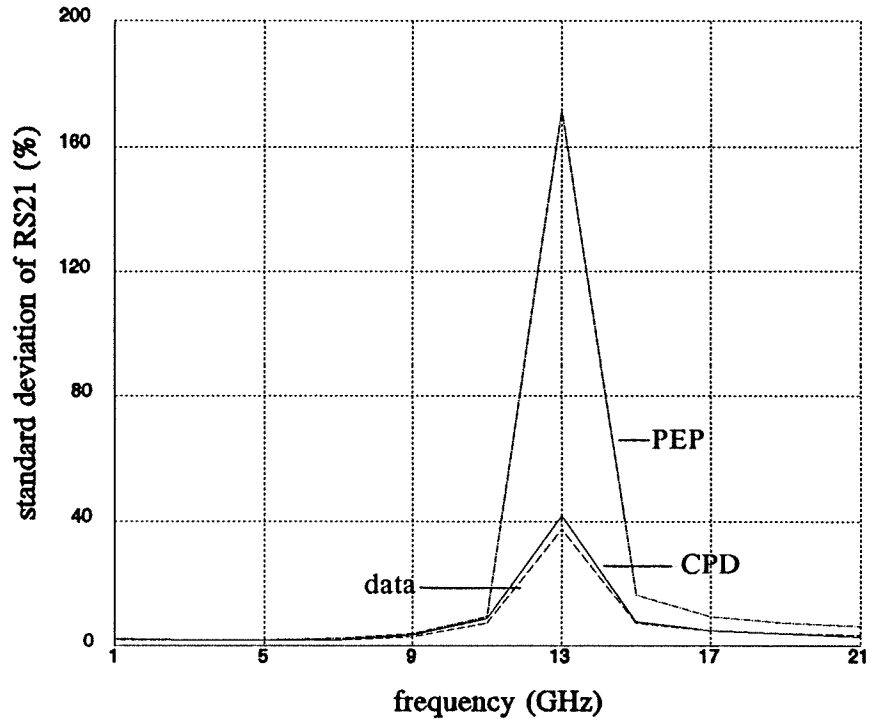
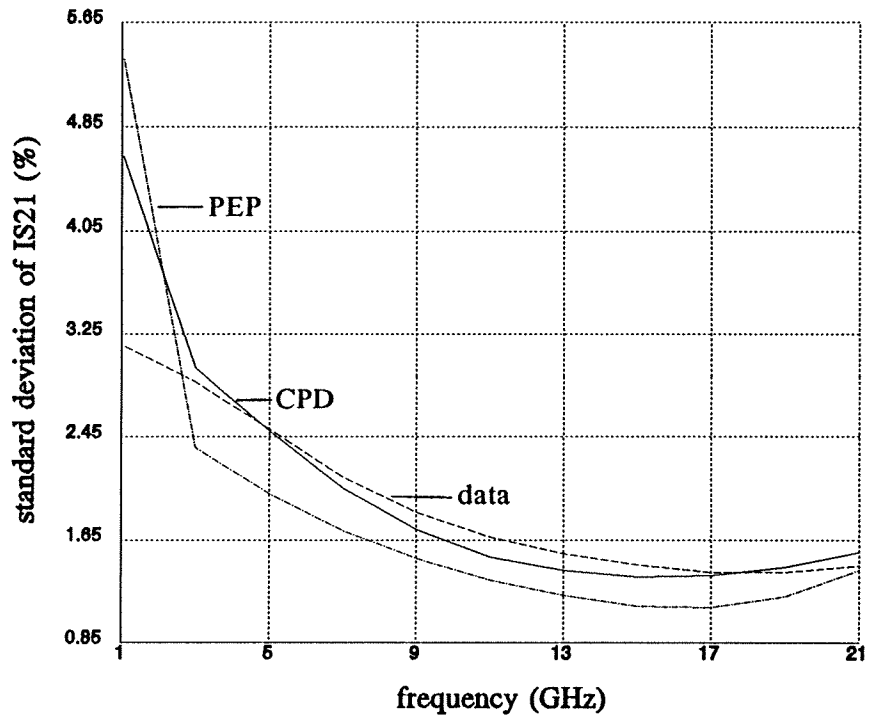


Fig. 6 Mean values of S_{21} versus frequency from data (---) and from the CPD (—) and PEP (-·-) statistical KTL models, (a) the real part and (b) the imaginary part.



(a)



(b)

Fig. 7 Standard deviations of S_{21} versus frequency from data (---) and from the CPD (—) and PEP (-.-) statistical KTL models, (a) the real part and (b) the imaginary part.