THE HUBER CONCEPT IN DEVICE MODELING, CIRCUIT DIAGNOSIS AND DESIGN CENTERING

J.W. Bandler, S.H. Chen, R.M. Biernacki and K. Madsen

OSA-93-OS-17-R

October 18, 1993

© Optimization Systems Associates Inc. 1993

THE HUBER CONCEPT IN DEVICE MODELING, CIRCUIT DIAGNOSIS

AND DESIGN CENTERING

J.W. Bandler, S.H. Chen, R.M. Biernacki and K. Madsen

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

> Tel +1 905 525 9140 Ext. 24818 Fax +1 905 523 4407

E-mail: biernack@mcmail.cis.mcmaster.ca or E-mail: bandler@mcmail.cis.mcmaster.ca

Technical Track: Computer-Aided Design

Preferred Presentation Format: Lecture

J.W. Bandler, S.H. Chen and R.M. Biernacki are also with Optimization Systems Associates Inc., P.O. Box 8083, Dundas, Ontario, Canada L9H 5E7.

K. Madsen is with the Institute for Numerical Analysis, The Technical University of Denmark, DK-2800 Lyngby, Denmark.

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 HUBER CONCEPT IN DEVICE MODELING, CIRCUIT DIAGNOSIS AND DESIGN CENTERING

Abstract

We present exciting applications of the Huber concept in circuit modeling and optimization. By combining the desirable properties of the ℓ_1 and ℓ_2 norms, the Huber function is robust against gross errors and smooth w.r.t. small variations in the data. We extend the Huber concept by introducing a one-sided Huber function tailored to design optimization with upper and lower specifications. We demonstrate the advantages of Huber optimization in the presence of faults, large and small measurement errors, bad starting points and statistical uncertainties through meaningful circuit applications, including multicavity filter parameter identification and design optimization, FET statistical modeling and analog fault location. Furthermore, we present a new one-sided Huber approach to yield optimization of linear and nonlinear circuits. Comparisons are made, where appropriate, with optimization using ℓ_1 , one-sided ℓ_1 , ℓ_2 and minimax objective functions.

Extended Summary

INTRODUCTION

Realistic circuit optimization must take into account model/measurement/statistical errors, variations and uncertainties. Least-squares (ℓ_2) solutions are notoriously susceptible to the influence of gross errors: just a few "wild" data points can alter the results significantly. The ℓ_1 method is robust against gross errors [1-3]. However, there may also be many small variations in the data, in addition to and distinct from gross errors, which should be included in statistical models and/or accommodated as design tolerances. The ℓ_1 method inappropriately treats such small variations in the same category as gross errors. In other words, neither the ℓ_1 nor ℓ_2 alone is capable of providing solutions which are robust against large (catastrophic) errors and flexible w.r.t. small (deterministic or statistical) variations in the data.

The Huber function [4-7] is a hybrid of the ℓ_1 and ℓ_2 norms. It separates large and small errors w.r.t. an appropriately chosen threshold. The large errors are treated in the ℓ_1 sense and the small errors are measured in terms of least squares. Consequently, the Huber solution can provide a smooth model from data which contains many small variations and such a model is also robust against gross errors. We demonstrate the benefits of this approach through applications to multicavity filter parameter identification, FET statistical modeling and analog fault location.

We extend the Huber concept by introducing a "one-sided" Huber function for design optimization where the specifications are one-sided (i.e., upper and/or lower specifications). The minimax method is often chosen to achieve an equal-ripple design. However, the success of minimax optimization may depend on the starting point. Given a "bad" starting point, a minimax optimizer can be trapped by the initial large errors. To overcome a bad starting point, the onesided Huber function can be employed in a "preprocessing" optimization. We have demonstrated this approach to large-scale multiplexer optimization problems [7]. In this paper, we expand our investigation by comparing minimax optimization of multicavity filters with and without one-sided Huber preprocessing from randomly generated starting points. Yield optimization takes into account assumed random variations in the manufacturing process to maximize design yield in order to reduce the manufacturing cost. We present, for the first time, a one-sided Huber approach to yield optimization of linear and nonlinear circuits.

Our approach is implemented in the CAD system OSA90/hope[™] [8] which is used to produce the examples in this presentation.

THEORY

The Huber function is defined as [3,4,7]

$$\rho_k(f) = \begin{cases} f^2/2 & if |f| \le k \\ k|f| - k^2/2 & if |f| > k \end{cases}$$
(1)

where k is a positive constant threshold value and f represents an error function.

 ρ_k is a hybrid of the ℓ_2 (when $|f| \le k$) and the ℓ_1 (when |f| > k) norms, as illustrated in Fig. 1. The definition of ρ_k ensures a smooth transition between ℓ_2 and ℓ_1 at |f| = k. The threshold k separates "large" and "small" errors. With a sufficiently large k, ρ_k becomes least squares. As k approaches zero, ρ_k approaches the ℓ_1 norm. By changing k, we can alter the proportion of error functions to be treated in the ℓ_1 or ℓ_2 sense.

We define the "one-sided" Huber function as

$$\rho_{k}^{+}(f) = \begin{cases} 0 & if \ f \leq 0 \\ f^{2}/2 & if \ 0 < f \leq k \\ kf - k^{2}/2 & if \ f > k \end{cases}$$
(2)

This definition is tailored to one-sided (upper/lower) specifications. A negative value of f indicates that the corresponding specification is satisfied and is therefore truncated.

We have implemented dedicated, efficient algorithms for minimizing the one- and two-sided Huber objective functions, as described in [4,5,7]. The basic algorithm is derived from the framework of Madsen [9] and consists of two stages. The first stage is a Gauss-Newton method which solves a linearized subproblem within a trust region. The algorithm may switch to a quasi-Newton stage when close to a solution to accelerate the convergence.

MULTICAVITY FILTER PARAMETER IDENTIFICATION

We consider a 6th-order multicavity filter represented by the equivalent circuit shown in Fig. 2 [2,10]. The filter has a 40 MHz bandwidth centered at 4000 MHz.

The input reflection coefficient of the filter is used as simulated measurement. Two large errors are deliberately introduced at two of the frequency points. The task is to identify the parameters from the contaminated data [2].

Our investigation is organized into six different cases. In Case A, the two large errors are the only errors contained in the data. The results in Table I shows that the ℓ_2 solution is hopelessly corrupted by the gross errors, whereas the ℓ_1 and Huber solutions are equally robust. This is also illustrated in Fig. 3.

In Cases B to D, the data is truncated to the first two significant digits to emulate the limited accuracy of measurement equipment. In these three cases, the numbers of frequencies considered are 26, 51 and 101, respectively. The truncation errors are small relative to the two gross errors. We choose a threshold value commensurate with the magnitude of the truncation errors so that they are treated in the ℓ_2 sense by the Huber method. Consequently, we expect the Huber solution to be less affected than the corresponding ℓ_1 solution. Our expectation is confirmed by the results in Table I.

In Case E, we introduced into the data small errors randomly generated from the uniform distribution [-0.01 0.01]. Again, the Huber solution is better in comparison with the ℓ_1 solution, as shown in Table I.

In Case F, the data is perturbed by +0.005 and -0.005 at alternate frequency points. As depicted in Fig. 4, the ℓ_1 solution is dictated by a subset of the data points which correspond to a subset of zero residual functions at the solution, in accordance with the ℓ_1 optimality condition.

The Huber solution, on the other hand, provides a smooth interpolation of all the data points except the two large errors, as shown in Fig. 4. This case dramatizes the advantage of the Huber function over the ℓ_1 norm in accommodating small data variations.

FET STATISTICAL MODELING

One approach to statistical modeling of devices [11-13] is to extract the model parameters from a sample of device measurements and then postprocessing the sample of model parameters to estimate their statistics (means, standard deviations and correlations).

To estimate the mean of a parameter by optimization, we define the error functions as

$$f_j(\overline{\phi}) = \overline{\phi} - \phi^j, \quad j = 1, 2, ..., N$$
(3)

where ϕ^{j} is the extracted parameter value for the *j*th device and N is the total number of devices. Similarly, to estimate the variances, we define

$$f_j(V_{\phi}) = V_{\phi} - (\phi^j - \overline{\phi})^2, \quad j = 1, 2, ..., N$$
 (4)

where V_{ϕ} denotes the estimated variance from which we can calculate the standard deviation σ_{ϕ} . The model parameters we use are extracted from the measurements of 80 FETs [14].

Fig. 5 shows the run chart of an extracted model parameter, namely the FET gate inductance L_G . Clearly, the sample contains a few wild points (likely due to faulty devices) which will severely degrade a least-squares estimate. In our earlier work [11,12] using the ℓ_2 estimator, the wild points were manually excluded.

The Huber function can be used as an automatic robust statistical estimator. Table II lists the means and standard deviations of a selected number of model parameters we have obtained using the ℓ_2 and the Huber estimators. For comparison, we also list the results obtained using the ℓ_2 estimator after the abnormal data sets are manually excluded.

The threshold value k is chosen to reflect the normal spread of the parameter values (e.g., we chose k = 0.015 for L_G).

ANALOG FAULT LOCATION

Consider the resistive mesh network shown in Fig. 6, which has been used to demonstrate the l_1 approach to analog fault location [2,15,16]. We have reported successful application of the Huber function to this problem [7].

In this paper, we present new results which take into account data truncation errors to represent the limited accuracy of measurement equipment.

The parameter values of the mesh network are listed in Table III. Two faults are assumed, namely G_2 and G_{18} . A single excitation (a DC current source) is applied to node 1. Nodes 4, 5, 8 and 9 are assumed to be inaccessible for measurement. The voltages at the other nodes are calculated, truncated to the first two significant digits and used as simulated measurement.

The nominal parameter values are used as the starting point for optimization. The results listed in Table III show that the ℓ_1 optimization failed to isolate the faults.

The data truncation may translate into deviations in the network parameter values, in addition to the faults and tolerances already considered. The ℓ_1 optimization attempts to suppress as many parameter deviations as possible to exactly zero, which may lead to an incorrect solution, as demonstrated in this case. By allowing for many small deviations, the Huber solution is more robust w.r.t. data truncation and measurement errors.

ONE-SIDED HUBER FORMULATION FOR YIELD OPTIMIZATION

In Monte Carlo analysis, we consider a number of statistical outcomes of circuit parameters denoted by ϕ^i . The design yield can be estimated as the percentage of acceptable outcomes out of the total number of outcomes considered.

Following the generalized ℓ_p centering approach of Bandler and Chen [1], for each outcome we create a generalized ℓ_p function $\nu(\phi^i)$ which has a positive value if the outcome violates the design specifications or a zero or negative value if the specifications are satisfied.

In our earlier work [1,17,18], we have formulated yield optimization as a one-sided ℓ_1 problem in which the objective function is defined as

6

$$U(\phi^0) = \sum_{i \in I} \alpha_i \ v(\phi^i) \tag{5}$$

where ϕ^0 represents the nominal circuit parameters to be centered, α_i is a positive multiplier associated with the *i*th outcome, and

$$I = \{ i \mid v(\phi^{i}) > 0 \}$$
(6)

is an index set defined over all the outcomes.

In this paper, we formulate yield optimization as one-sided Huber problem in which the objective function is defined as

$$U(\boldsymbol{\phi}^{0}) = \sum_{i=1}^{N} \rho_{k}^{+}(\alpha_{i} v(\boldsymbol{\phi}^{i}))$$
(7)

where N is the total number of outcomes and ρ_k^+ is the one-sided Huber function defined in (2).

We consider the linear LC filter [19] depicted in Fig. 7. During the yield optimization a multidimensional Q-model is utilized to reduce the statistical circuit simulation time [17]. The yield optimization results using our new one-sided Huber approach are summarized in Table IV. The one-sided Huber approach proved to be a competitive alternative to the one-sided ℓ_1 centering approach.

Also, consider the nonlinear frequency doubler depicted in Fig. 8 [20]. Uniform tolerances are assumed for the linear matching circuits and multidimensional Gaussian distributions with correlations are assumed for the intrinsic FET parameters. The yield at the nominal design is 28% (all the yield figures reported in this paper are estimated by Monte Carlo simulation with 500 outcomes). The yield at the centered design obtained using one-sided ℓ_1 centering is 76%, after 17 iterations and 337 CPU seconds on a SPARCstation 10. The yield after one-sided Huber optimization is 77%, after 29 iterations and \$74 CPU seconds.

ONE-SIDED HUBER PREPROCESSING OF ARBITRARY STARTING POINTS

We have exploited the potential of using one-sided Huber preprocessing to overcome bad starting points in large-scale multiplexer optimization [7].

In this paper, we expand our investigation by testing randomly generated starting points for multicavity filter optimization. The same 6th-order multicavity filter shown in Fig. 2 is used.

Two sets of random starting points were generated. Each set consists of 30 starting points of uniform distribution centered at a "good" starting point. The spread of the parameter values is $\pm 30\%$ for the first set and $\pm 40\%$ for the second set. The input return loss of the filter at the first set of starting points is shown in Fig. 9. We can see that some of the starting points are indeed very bad (the return loss response is always symmetrical w.r.t. the center frequency due to the filter configuration considered, hence only one half of the frequency band needs to be simulated during optimization).

From each starting point, we performed two experiments: (1) direct minimax optimization and (2) one-sided Huber optimization (preprocessing) followed by minimax optimization. The optimized responses are shown in Fig. 10 for the first set of starting points and in Fig. 11 for the second set. Although the one-sided Huber preprocessing did not guarantee convergence to the optimal solution from all the starting points, it produced more focused results by eliminating many spurious local minima.

CONCLUSIONS

We have presented exciting breakthroughs in applying a novel Huber approach to modeling, parameter identification, fault diagnosis and design centering of linear and nonlinear circuits. Compared with ℓ_1 , ℓ_2 and minimax methods, the Huber approach has demonstrated robustness and consistency in the presence of large and small errors, deterministic and statistical variations, which are critical considerations for practical CAD in an engineering environment.

The numerical results reported in this paper were obtained using a research-oriented implementation of the algorithms. We are confident that the computational efficiency can be further improved in a more polished implementation.

REFERENCES

- [1] J.W. Bandler and S.H. Chen, "Circuit optimization: the state of the art," *IEEE Trans. Microwave Theory Tech.*, vol. 36, 1988, pp. 424-443.
- [2] J.W. Bandler, W. Kellermann and K. Madsen, "A nonlinear l₁ optimization algorithm for design, modeling and diagnosis of networks", *IEEE Trans. Circuits and Systems*, vol. CAS-34, 1987, pp. 174-181.
- [3] J.W. Bandler, S.H. Chen and S. Daijavad, "Microwave device modeling using efficient l₁ optimization: a novel approach", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-34, 1986, pp. 1282-1293.
- [4] P. Huber, Robust Statistics. New York: Wiley, 1981.
- [5] G. Li and K. Madsen, "Robust nonlinear data fitting", in *Numerical Analysis*, D.F. Griffiths and G.A. Watson, Eds., Pitman Research Notes in Mathematics Series 170. UK: Longman, 1988, pp. 176-191.
- [6] H. Ekblom and K. Madsen, "Algorithms for nonlinear Huber estimation", *BIT* 29, 1989, pp. 60-76.
- [7] J.W. Bandler, S.H. Chen, R.M. Biernacki, L. Gao, K. Madsen and H. Yu, "Huber optimization of circuits: a robust approach", *IEEE Trans. Microwave Theory Tech.*, vol. 41, December 1993.
- [8] OSA90/hope[™], Optimization Systems Associates Inc., P.O. Box 8083, Dundas, Ontario, Canada L9H 5E7, 1993.
- [9] K. Madsen, Minimization of Non-Linear Approximation Functions, Dr. Techn. Thesis, Institute of Numerical Analysis, Tech. Univ. of Denmark, DK2800 Lyngby, Denmark, 1985.
- [10] A.E. Atia and A.E. Williams, "New types of waveguide bandpass filters for satellite transponders", COMSAT Technical Review, vol. 1, 1971, pp. 21-43.
- [11] J.W. Bandler, R.M. Biernacki, S.H. Chen, J. Song, S. Ye and Q.J. Zhang, "Statistical modeling of GaAs MESFETs", *IEEE MTT-S Int. Microwave Symp. Dig.* (Boston, MA), 1991, pp. 87-90.
- [12] J.W. Bandler, R.M. Biernacki, Q. Cai, S.H. Chen, S. Ye and Q.J. Zhang, "Integrated physicsoriented statistical modeling, simulation and optimization", *IEEE Trans. Microwave Theory Tech.*, vol 40, 1992, pp.1374-1400.
- [13] HarPE[™], Optimization Systems Associates Inc., P.O. Box 8083, Dundas, Ontario, Canada L9H 5E7, 1993.
- [14] Measurement data provided by Plessey Research Caswell Ltd., Caswell, Towcester, Northamptonshire, England NN12 8EQ, 1990.
- [15] J.W. Bandler, R.M. Biernacki, A.E. Salama and J.A. Starzyk, "Fault isolation in linear analog circuits using the l_1 norm", *Proc. IEEE Int. Symp. Circuits and Systems* (Rome, Italy),

1982, pp. 1140-1143.

- [16] J.W. Bandler and A.E. Salama, "Fault diagnosis of analog circuits", Proc. IEEE, vol. 73, 1985, pp. 1279-1325.
- [17] R.M. Biernacki, J.W. Bandler, J. Song and Q.J. Zhang, "Efficient quadratic approximation for statistical design", *IEEE Trans. Circuits and Systems*, vol. 36, 1989, pp. 1449-1454.
- [18] J.W. Bandler, S.H. Chen and K. Madsen, "An algorithm for one-sided l₁ optimization with application to circuit design centering," Proc. IEEE Int. Symp. Circuits and Systems (Espoo, Finland), 1988, pp. 1795-1798.
- [19] K. Singhal and J.F. Pinel, "Statistical design centering and tolerancing using parametric sampling", *IEEE Trans. Circuits and Systems*, vol. CAS-28, 1981, pp. 692-701.
- [20] J.W. Bandler, Q.J. Zhang, J. Song and R.M. Biernacki, "FAST gradient based yield optimization of nonlinear circuits", *IEEE Trans. Microwave Theory Tech.*, vol. 38, 1990, pp. 1701-1710.

 TABLE I

 MULTICAVITY FILTER PARAMETER IDENTIFICATION RESULTS

Couplings	<i>M</i> ₁₂	M ₂₃	M ₃₄	M ₄₅	M ₅₆	M ₁₆	M ₂₅
Actual Values	0.859956	0.526602	0.791894	0.526602	0.859956	0.087293	-0.393685
Starting Point	0.819006	0.511264	0.824890	0.511264	0.819006	0.093863	-0.357895
Case A: ℓ_2	-11%	6.5%	-0.76%	7.3%	-10%	278%	31%
ℓ_1	0.05%	0.06%	0.00%	-0.06%	-0.05%	-0.01%	0.00%
Huber	0.02%	0.01%	-0.02%	0.01%	0.02%	-1.2%	-0.16%
Case B: ℓ_1	0.47%	0.07%	-0.69%	0.36%	0.59%	-33%	-5.1%
Huber	0.34%	0.3%	-0.5%	0.3%	0.34%	-19%	-3.3%
Case C: ℓ_1	0.86%	1.8%	0.78%	-3.1%	-0.42%	-12%	0.47%
Huber	0.19%	-0.1%	0.00%	-0.1%	0.19%	-8.9%	-1%
Case D: ℓ_1	0.51%	1.8%	0.51%	-2.9%	-0.22%	-14%	-0.33%
Huber	0.15%	0.00%	-0.11%	-0.01%	0.15%	-8.3%	-1.2%
Case E: ℓ_1	1.8%	2.4%	0.6%	-4.1%	-0.53%	-43%	-3.1%
Hube	r 0.41%	0.04%	-0.52%	0.04%	0.41%	-27%	-4%
Case F: ℓ_1	1.5%	3.8%	0.23%	-4.2%	-1.3%	-25%	-2.5%
Hube	r 0.05%	0.15%	-0.24%	0.15%	0.05%	-4.9%	-1.1%

The percentage entries represent the relative differences between the identified parameter values and the actual parameter values.

The data contains two large errors in all cases.

Case A: the data is not truncated.

Case B: the data is truncated to two significant digits; 26 frequencies. Case C: the data is truncated to two significant digits; 51 frequencies. Case D: the data is truncated to two significant digits; 101 frequencies. Case E: the data contains random deviates from uniform distribution [-0.01 +0.01]. Case F: the data is perturbed by +/-0.005 at alternate frequencies.

Parameter	$\overline{\phi}$ (ℓ_2)	$\overline{\phi}$ (Huber)	$\overline{\phi} (\ell_2^*)$	$\sigma_{\phi}(\ell_2)$	$\sigma_{\phi}(\mathrm{Huber})$	$\sigma_{\phi}(\ell_2^*)$
$L_{c}(nH)$	0.04387	0.03464	0.03429	94.6%	21.8%	17.4%
$G_{\rm DS}(1/{\rm K}\Omega)$	1.840	1.820	1.839	28.6%	6.3%	4.9%
$I_{\rm DSS}({\rm mA})$	47.36	47.53	47.85	14.0%	12.7%	11.3%
$\tau(\text{ps})$ $C_{10}(\text{pF})$	2.018 0.3618	2.154 0.3658	2.187 0.3696	26.3% 8.2%	5.8% 4.6%	3.4% 3.5%
K_1^{100}	1.2328	1.231	1.233	15.5%	10.8%	8.7%
•						

 TABLE II

 ESTIMATED STATISTICS OF SELECTED FET PARAMETERS

 L_G represents the FET gate lead inductance, G_{DS} the drain-source conductance, I_{DSS} the drain saturation current, τ the time-delay, C_{10} and K_1 are parameters in the definition of the gate nonlinear capacitor.

 ℓ_2^* denotes ℓ_2 estimates after 11 abnormal data sets are manually excluded [11].

			Percentage Deviation		
Element	Nominal Value	Actual Value	Actual	ℓ_1	Huber
$\begin{array}{c} G_{1} \\ G_{2} \\ G_{3} \\ G_{4} \\ G_{5} \\ G_{6} \\ G_{7} \\ G_{8} \\ G_{9} \\ G_{10} \\ G_{11} \\ G_{12} \\ G_{13} \\ G_{14} \\ G_{15} \\ G_{15} \end{array}$	$ \begin{array}{c} 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\$	0.95 0.50 1.05 0.95 0.95 1.05 1.05 1.05 1.05 1.05 1.05 1.05 0.95 0.95 0.95	-5.0 -50.0* 5.0 -5.0 -5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 -5.0 5.0 -5.0 5.0 -5.0 5.0 -5.0 -	-6.98 -47.55 -25.45 0.00 0.00 0.00 0.00 0.00 0.00 0.00	-15.14 -54.40 -3.68 -0.63 -4.07 -0.05 -1.06 -0.94 3.69 0.48 0.67 6.00 -0.32 -0.68 0.18 -3.53
$G_{16} G_{17} G_{18}$	1.0 1.0	1.05 0.50	5.0 -50.0 *	0.00 -8.90	-0.81 -49.97
$G_{19} \\ G_{20}$	1.0 1.0	0.95 0.95	-5.0 -5.0	-25.32 -20.73	-4.74 -5.98

TABLE III FAULT LOCATION OF THE RESISTIVE MESH CIRCUIT

* Faults.

The simulated voltage measurement was truncated to two significant digits.

TABLE IV DESIGN CENTERING OF THE LC FILTER

	Yield	Number of Iterations	CPU Time (Seconds)
Nominal Design	52%		
One-Sided ℓ_1	75%	11	160
One-Sided Huber, k=0.05	76%	22	336
One-Sided Huber, k=0.10	75%	19	265
One-Sided Huber, k=0.20	75%	9	123
One-Sided Huber, $k=0.25$	74%	21	260
One-Sided Huber, $k=0.30$	74%	12	160
One-Sided Huber, $k=0.40$	72%	12	149

The yield figures were estimated by Monte Carlo simulation using 500 outcomes.

The CPU time were measured on a SPARCstation 10.

k is the threshold value for the one-sided Huber function.



Fig. 1 The Huber, ℓ_1 and ℓ_2 objective functions in the one-dimensional case. The strikes and dots represent points on the ℓ_1 and ℓ_2 curves, respectively. The solid line represents the Huber function.



Fig. 2 Equivalent circuit for an unterminated multicavity filter.



Fig. 3 Comparison of the results from parameter identification using ℓ_1 , ℓ_2 and Huber objective functions. The circles represent the data with two large errors. The solid line represents the responses after ℓ_1 and Huber optimization (they are indistinguishable). The dotted line represents the ℓ_2 solution.



Fig. 4 Comparison of parameter identification from truncated data using ℓ_1 and Huber objective functions. A portion of the passband is shown, enlarged for detail. The circles represent the data. The solid line represents the response after optimization. The ℓ_1 solution (top) is dictated by a subset of data which corresponds to a subset of zero residual functions at the solution. The Huber solution (bottom) interpolates the data variations smoothly (in the ℓ_2 sense).



Fig. 5 Run chart of the FET statistical model parameter L_G (gate lead inductance).



Fig. 6 The resistive mesh circuit.







Fig. 8 The FET frequency doubler.



Fig. 9 The input return loss of the multicavity filter at the 30 starting points randomly generated from uniform distributions with a $\pm 30\%$ parameter spread.



Fig. 10 The return loss of the multicavity filter at the solutions from the 30 randomly generated starting points with a ±30% parameter spread. The top figure shows the results after direct minimax optimization. The bottom figure shows the results obtained from one-sided Huber preprocessing followed by minimax optimization.



Fig. 11 The return loss of the multicavity filter at the solutions from the 30 randomly generated starting points with a ±40% parameter spread. The top figure shows the results after direct minimax optimization. The bottom figure shows the results obtained from one-sided Huber preprocessing followed by minimax optimization.