ROBUSTIZING CIRCUIT OPTIMIZATION

USING HUBER FUNCTIONS

OSA-92-MT-15-R

November 24, 1992

(Revised February 19, 1993)

1993 IEEE International Microwave Symposium Digest, Atlanta, GA, June 1993

ROBUSTIZING CIRCUIT OPTIMIZATION USING HUBER FUNCTIONS

J.W. Bandler*, S.H. Chen*, R.M. Biernacki*, K. Madsen**, L. Gao*** and H. Yu*

Optimization Systems Associates Inc.
P.O. Box 8083, Dundas, Ontario, Canada L9H 5E7

ABSTRACT

We introduce a novel approach to "robustizing" circuit optimization using Huber functions: both two-sided and one-sided. We compare Huber optimization with ℓ_1, ℓ_2 and minimax methods in the presence of faults, large and small measurement errors, bad starting points and statistical uncertainties. We demonstrate FET statistical modeling, multiplexer optimization, analog fault location and data fitting.

INTRODUCTION

Engineering designers are often concerned with the robustness of numerical optimization techniques, and rightly so, knowing that engineering data is, with few exceptions, contaminated by model/measurement/statistical errors.

The classical least-squares method is well known for its vulnerability to gross errors: a few wild data points can alter the least-squares solution significantly. The ℓ_1 method is robust against gross errors [1,2]. We will show, however, that when the data contains many small variations, the ℓ_1 solution can become undesirably biased.

We introduce, to microwave circuit CAD, the Huber concept [3,4]. The Huber optimization is more robust than ℓ_2 w.r.t. large errors, and smoother, less biased than ℓ_1. We demonstrate the benefits of this novel approach in FET statistical modeling, analog fault location and data fitting.

We extend the Huber concept by introducing a "one-sided" Huber function for large-scale optimization. For large-scale problems, the designer often attempts, by intuition, a "preliminary" optimization by selecting a small number of dominant variables. We demonstrate, through multiplexer optimization, that the one-sided Huber function can be more effective and efficient than minimax in overcoming a bad starting point.

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 OGP0007239, OGP0042444 and STR0117819.

* J.W. Bandler, S.H. Chen and R.M. Biernacki are also with and H. Yu is with the Simulation Optimization Systems Research Laboratory, Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada L8S 4L7.

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

*** L. Gao is with the Department of Mathematics and Institute of Mathematics, Peking University, Beijing 100871, China.

THEORY

The Huber optimization problem is defined as [3,4]

$$\underset{x}{\text{minimize}} \quad F(x) \equiv \sum_{j=1}^{m} \rho_k(f_j(x)) \tag{1}$$

where

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

x is the set of optimization variables, k is a positive constant and f_j, $j = 1, 2, \ldots, m$, are error functions.

The Huber function ρ_k is a hybrid of the ℓ_2 (when $|f| \le k$) and the ℓ_1 (when $|f| > k$) functions. By varying k, we can alter the fraction of error functions to be treated in the least-squares sense. The choice of k defines the threshold between "large" and "small" errors. If k is set to a sufficiently large value, the optimization problem (1) becomes least squares (ℓ_2). On the other hand, as k approaches zero, ρ_k will approach the ℓ_1 function.

We extend the Huber concept for design optimization with upper and lower specifications by introducing a "one-sided" Huber function. Negative errors are truncated, i.e., $\rho_k = 0$ for $f \le 0$, because the corresponding design specification is satisfied.

Our exposition utilizes a dedicated and efficient Huber optimizer [4] already implemented in the CAD system OSA90/hope™ [5] as a standard feature.

COMPARISON OF ℓ_1, ℓ_2 AND HUBER METHODS IN DATA FITTING

We consider the approximation of $\sqrt{t}$ by a rational function for $0 \le t \le 1$ [2]. Large errors are deliberately introduced at 5 of the sample points and small variations to the remaining data. The ℓ_1, ℓ_2 and Huber approximations are shown in Fig. 1. Fig. 2 shows an enlarged portion for a clearer view of the details.

As expected, the ℓ_2 solution suffers significantly from the presence of gross errors. The ℓ_1 solution, according to the optimality condition, is dictated by a subset of residual functions

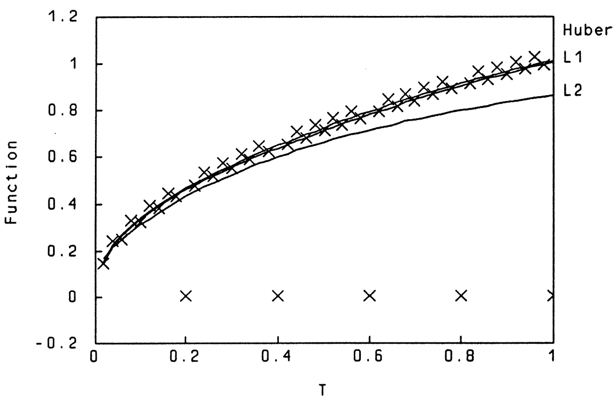

Fig. 1. ℓ_1, ℓ_2 and Huber solutions for data fitting in the presence of errors.

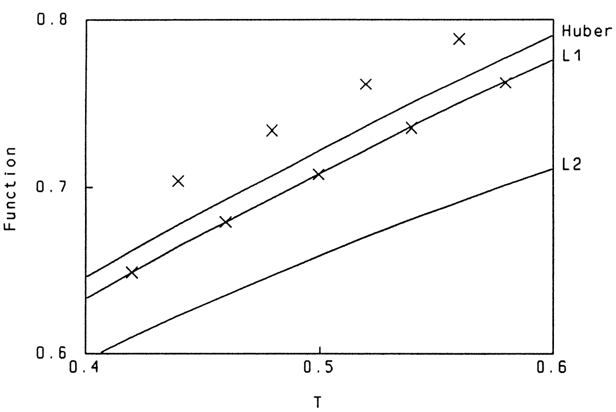

Fig. 2. An enlarged portion of Fig. 1.

which have zero values at the solution. In a sense, all the nonzero residuals are treated as large errors. Such a biased ℓ_1 solution, as dramatized in our example, is undesirable if we wish to model the small variations in the data.

The Huber solution provides a flexible combination of the robustness of the ℓ_1 and the unbiasedness of the ℓ_2. In fact, the Huber solution is equivalent to an ℓ_2 solution with the gross errors reduced to the threshold value k.

HUBER ESTIMATOR FOR STATISTICAL MODELING OF DEVICES

We use the Huber function as an automated robust estimator for FET statistical modeling. Model parameters are extracted from the measurements of 80 FETs using HarPE™ [6], and then postprocessed to estimate the parameter statistics.

Fig. 3 shows the run chart of the extracted values of the time-delay τ. Most of the values are between 2ps - 2.5ps, but there are a few abnormal values due to faulty devices and/or gross measurement errors. In our earlier work [7] using the ℓ_2 estimator, the abnormal data sets were manually excluded from the statistical modeling process.

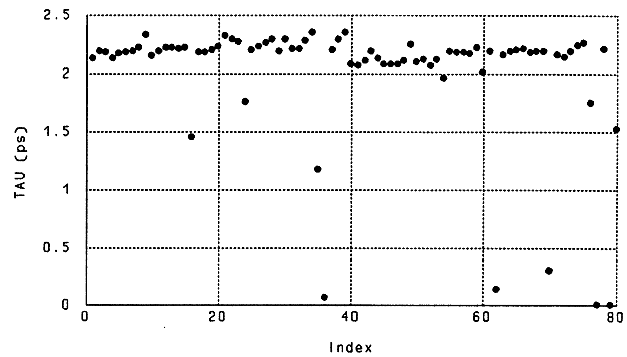

Fig. 3. Run chart of the extracted model parameter τ.

TABLE I
ESTIMATED STATISTICS OF SELECTED FET PARAMETERS

Parameter	$\bar{\phi}(\ell_2)$	$\bar{\phi}(\mathrm{H})$	$\bar{\phi}(\ell_2^*)$	$\sigma_\phi(\ell_2)$	$\sigma_\phi(\mathrm{H})$	$\sigma_\phi(\ell_2^*)$
L_G(nH)	0.04387	0.03464	0.03429	94.6%	21.8%	17.4%
G_{DS}(1/KΩ)	1.840	1.820	1.839	28.6%	6.3%	4.9%
I_{DSS}(mA)	47.36	47.53	47.85	14.0%	12.7%	11.3%
τ(ps)	2.018	2.154	2.187	26.3%	5.8%	3.4%
C_{10}(pF)	0.3618	0.3658	0.3696	8.2%	4.6%	3.5%
K_1	1.2328	1.231	1.233	15.5%	10.8%	8.7%

$\bar{\phi}$ denotes the mean and σ_ϕ the standard deviation.
H denotes Huber estimates.
ℓ_2^* denotes ℓ_2 estimates after 11 abnormal data sets are manually excluded [7].

To estimate the mean of a parameter, we define

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

where ϕ^j is the extracted parameter value for the jth device and N is the total number of devices. The threshold value k for the Huber function is chosen according to the normal spread of the parameter values (e.g., we chose $k = 0.25$ for τ).

We also define

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

where V_ϕ denotes the estimated variance from which we can calculate the standard deviation σ_ϕ.

Table I lists the ℓ_2 and Huber estimates of the statistics of a selected number of model parameters. We also list the results obtained using ℓ_2 *after the abnormal data sets are manually excluded*. In comparison, the Huber estimator does not require manual manipulation of the data and is clearly more appropriate when there are data points which cannot be clearly classified as normal or abnormal.

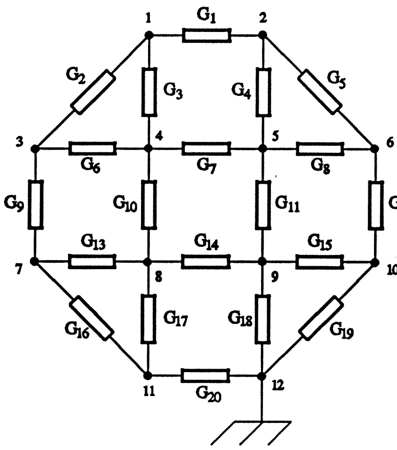

Fig. 4. The resistive mesh circuit.

APPLICATION TO ANALOG FAULT LOCATION

Analog fault location [1,8,9] can be formulated as the Huber optimization

$$\underset{x}{minimize} \sum_{j=1}^{n+K} \rho_k(f_j(x)) \tag{5}$$

where

$$f_i(x) \equiv \Delta x_i / x_i^0, \quad i = 1, 2, ..., n$$
$$f_{n+i}(x) \equiv \beta_i(V_i^c - V_i^m), \quad i = 1, 2, ..., K \tag{6}$$

where $x \equiv [x_1\ x_2\ ...\ x_n]^T$ is a vector of circuit parameters, x^0 represents the nominal values, and $\Delta x \equiv x - x^0$ represents the deviations from the nominal. V_1^m, ..., V_K^m are K measurements (such as time-domain voltages and split real and imaginary parts of complex frequency-domain voltages). V_1^c, ..., V_K^c are the calculated circuit responses. β_i, i = 1, 2, ..., K, are appropriate multipliers.

Consider the resistive mesh network shown in Fig. 4 [1,8]. The nominal parameter values are G_i = 1.0 with tolerances ϵ_i = ±0.05, i = 1, 2, ..., 20. Nodes 4, 5, 8 and 9 are assumed to be internal and inaccessible for measurement.

Two faults are assumed, namely G_2 and G_{18}. Table II compares the results from the ℓ_1 and Huber optimizations utilizing voltage measurements under a single excitation applied to node 1.

We tested this example for 4 other different starting points. The Huber method correctly located the faults in all the cases. The ℓ_1 method was successful in 3 of the cases, but failed in one of the cases (trapped in a different local minimum).

ONE-SIDED HUBER OPTIMIZATION FOR CIRCUIT DESIGN

In a large-scale design problem, we often wish to optimize a small number of dominant variables in order to obtain a good starting point for the full-scale optimization.

Consider a 5-channel 12 GHz waveguide manifold multiplexer [10].

TABLE II
FAULT LOCATION OF THE RESISTIVE MESH CIRCUIT

Element	Nominal Value	Actual Value	Percentage Deviation		
			Actual	ℓ_1	Huber
G_1	1.0	0.98	-2.0	0.00	-0.11
G_2	1.0	0.50	-50.0*	-48.89	-47.28
G_3	1.0	1.04	4.0	0.00	-2.46
G_4	1.0	0.97	-3.0	0.00	-1.18
G_5	1.0	0.95	-5.0	-2.70	-3.16
G_6	1.0	0.99	-1.0	0.00	-0.06
G_7	1.0	1.02	2.0	0.00	-0.19
G_8	1.0	1.05	5.0	0.00	-0.41
G_9	1.0	1.02	2.0	2.41	3.75
G_{10}	1.0	0.98	-2.0	0.00	0.39
G_{11}	1.0	1.04	4.0	0.00	-0.37
G_{12}	1.0	1.01	1.0	2.73	1.32
G_{13}	1.0	0.99	-1.0	0.00	-0.26
G_{14}	1.0	0.98	-2.0	0.00	-0.50
G_{15}	1.0	1.02	2.0	0.00	-0.05
G_{16}	1.0	0.96	-4.0	-3.36	-2.67
G_{17}	1.0	1.02	2.0	0.00	-0.61
G_{18}	1.0	0.50	-50.0*	-50.09	-47.33
G_{19}	1.0	0.98	-2.0	-1.41	-3.81
G_{20}	1.0	0.96	-4.0	-4.40	-4.72

* Faults

The responses before optimization are shown in Fig. 5. From a total of 75 optimizable variables, we first select 10 dominant variables including spacings and the channel input transformer ratios. The minimax solution with these variables is shown in Fig. 6 and the one-sided Huber solution is shown in Fig. 7. The worst-case errors in these two figures are similar.

Since the worst-case errors cannot be further reduced with only 10 variables, the minimax optimizer sees no point in

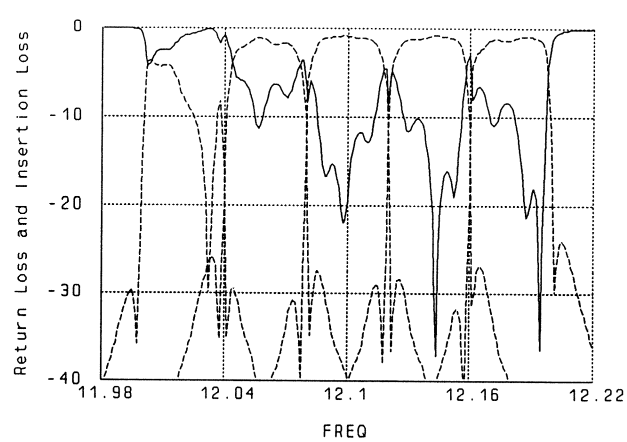

Fig. 5. Multiplexer responses at the starting point, showing the common port return loss (———) and the individual channel insertion losses (------).

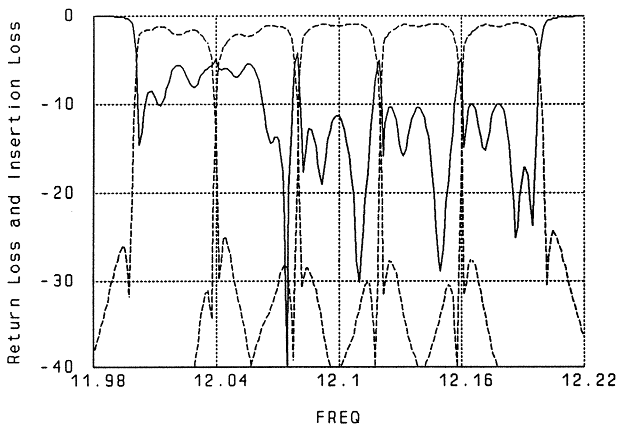

Fig. 6. Multiplexer responses after minimax optimization with 10 variables: spacings and channel input transformer ratios; the common port return loss (———) and the individual channel insertion losses (------). This result hardly improved upon the starting point shown in Fig. 5.

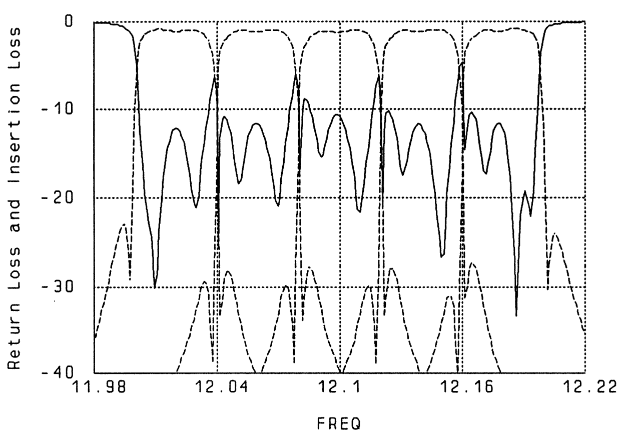

Fig. 7. Multiplexer responses after one-sided Huber optimization with 10 variables: spacings and channel input transformer ratios; the common port return loss (———) and the individual channel insertion losses (------). This result is significantly better than the minimax solution of Fig. 6.

spending effort elsewhere. Using the one-sided Huber function, however, we were able to obtain a good starting point for the subsequent full-scale minimax optimization, which results in the multiplexer responses shown in Fig. 8.

CONCLUSIONS

We have introduced the concept and some applications of the Huber method to microwave circuit CAD. This novel concept is consistent with practical engineering intuition and will have a far-reaching and profound impact on modeling, design, design validation, fault diagnosis and statistical processing of circuits and devices. We have presented strong evidence in a number of application areas, and without doubt we will find the Huber optimization of significant benefit in other areas as well.

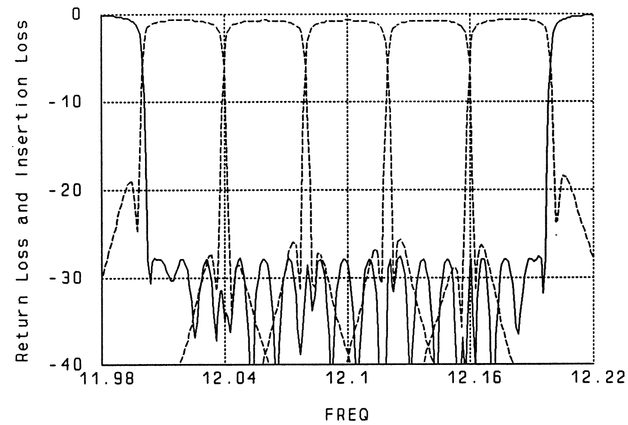

Fig. 8. Multiplexer responses after minimax optimization with the full set of 75 variables, showing the common port return loss (———) and the individual channel insertion losses (------).

REFERENCES

[1] J.W. Bandler, W. Kellermann and K. Madsen, "A nonlinear ℓ_1 optimization algorithm for design, modeling and diagnosis of networks", *IEEE Trans. Circuits and Systems*, vol. CAS-34, 1987, pp. 174-181.

[2] 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.

[3] P. Huber, *Robust Statistics*. New York: Wiley, 1981.

[4] 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.

[5] *OSA90/hope*™, Optimization Systems Associates Inc., P.O. Box 8083, Dundas, Ontario, Canada L9H 5E7, 1992.

[6] *HarPE*™ *Version 1.6*, Optimization Systems Associates Inc., P.O. Box 8083, Dundas, Ontario, Canada L9H 5E7, 1992.

[7] 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.

[8] 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.

[9] J.W. Bandler and A.E. Salama, "Fault diagnosis of analog circuits", *Proc. IEEE*, vol. 73, 1985, pp. 1279-1325.

[10] 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.