A NOVEL TECHNIQUE FOR MULTI-POINT PARAMETER EXTRACTION

M.H. Bakr and J.W. Bandler

SOS-98-23-R

August 1998

A NOVEL TECHNIQUE FOR MULTI-POINT PARAMETER EXTRACTION

M.H. Bakr and J.W. Bandler

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

Abstract

In this paper we present a novel technique for multi-point parameter extraction. This technique addresses the optimal selection of the perturbations used to improve the uniqueness and robustness of the parameter extraction step. The uniqueness of the extracted parameters is of great importance. We establish a criterion for the selection of the best possible perturbations that can be used in the multi-point parameter extraction. The selection of the perturbation is dependent on the characteristics of the minimum reached in the previous iteration. Two different approaches are suggested for two possible classifications of the extracted parameters. A novel theorem motivates the selection of the perturbation for a certain type of minimum. A robust algorithm that integrates the two approaches discussed is introduced. The algorithm terminates if the extracted coarse model parameters can be trusted. This algorithm was applied to a number of examples. The results obtained are encouraging.

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Grants OGP0007239, OGP0042444, STP0201832 and through the Micronet Network of Centres of Excellence.

J.W. Bandler is also with Bandler Corporation, P.O. Box 8083, Dundas, Ontario, Canada L9H 5E7.

I. INTRODUCTION

In this work we present a new algorithm for multi-point parameter extraction. Parameter extraction plays a crucial role in the space mapping algorithm [1,2,3]. It utilizes an optimization problem that may yield a nonunique solution. This nonuniqueness may lead to divergence or oscillatory behavior of the algorithm.

It should be noted that the new technique for the multi-point parameter extraction is discussed in the context of space mapping. We assume the presence of a fine model that generates the target responses and a coarse model whose parameters are to be extracted. However, there is no restriction on the application of this method to other problems that require parameter extraction.

Several authors have addressed the nonuniqueness problem. For example, Daijavad [4] proposed the idea of making unknown perturbations to a certain system whose parameters are to be extracted. This approach was only concerned about reaching a locally unique minimum of the problem without much concern about whether it is the only existing minimum. Later Bandler et al. [5] suggested that a multi-point parameter extraction be used to match the first order derivatives of the two models and thus to have only a global minimum. The perturbations used in this approach are known. However, these perturbations were arbitrary and the optimality of the selection of these perturbations was not addressed. Recently, a recursive multi-point parameter extraction technique was suggested by Bakr et al. [3]. This new approach utilizes the information available about the mapping between the two spaces.

The new algorithm introduced in this work aims at minimizing the number of perturbations used in the multi-point parameter extraction by utilizing the best possible perturbation in each iteration. Each perturbation requires a circuit simulation which would be very CPU intensive. We classify the different solutions returned by the multi-point extraction process and based on this classification a new perturbation that is likely to enhance the uniqueness is suggested. Some of the concepts discussed in this work are borrowed from the context of analog fault diagnosis [6].

We start by defining the multi-point parameter extraction problem in Section II. Different approaches presented in the literature are reviewed. In Section III we characterize the different minima that may be returned by the parameter extraction process. Two approaches are suggested to enhance the uniqueness of the extraction based on the character of the returned parameters. In Section IV we introduce a new algorithm for multi-point parameter extraction which is based on the discussion of Section III. The new algorithm was applied successfully to a number of examples which are discussed in Section V. Finally, the conclusions are given in Section VI.

II. PARAMETER EXTRACTION

The object of parameter extraction is to find a set of parameters whose responses match a given set of measurements. It can be formulated as

$$\underset{\boldsymbol{x}_{os}}{minimize} \left\| \boldsymbol{R} - \boldsymbol{R}_{os}(\boldsymbol{x}_{os}) \right\| \tag{1}$$

where $\boldsymbol{R}$ is the vector of given measurements, $\boldsymbol{R}_{os}$ is the vector of system responses and $\boldsymbol{x}_{os}$ is the vector of extracted parameters. In the context of space mapping the fine model responses $\boldsymbol{R}_{em}$ at a certain fine model point $\boldsymbol{x}_{em}$ supplies the target response $\boldsymbol{R}$. It is required to find a set of coarse model parameters $\boldsymbol{x}_{os}$ whose response matches the given fine model response. It can be formulated as

$$\underset{\boldsymbol{x}_{os}}{minimize} \left\| \boldsymbol{R}_{em}(\boldsymbol{x}_{em}) - \boldsymbol{R}_{os}(\boldsymbol{x}_{os}) \right\| \tag{2}$$

The solution of this optimization problem may be nonunique. This nonuniqueness may cause the space mapping algorithm to diverge or to oscillate [5].

A multi-point parameter extraction procedure [5] was suggested to improve the uniqueness of the step. The two models are matched at a number of points. The vector of extracted coarse model parameters $\boldsymbol{x}_{os}$ should satisfy

$$\boldsymbol{R}_{os}(\boldsymbol{x}_{os} + \Delta\boldsymbol{x}) = \boldsymbol{R}_{em}(\boldsymbol{x}_{em} + \Delta\boldsymbol{x}) \tag{3}$$

for a discrete set of perturbations $\Delta \boldsymbol{x} \in V_P$; the set of perturbations where $\mathbf{0} \in V_P$. It follows that the solution to this multi-point parameter extraction not only matches the responses at the given fine model point $\boldsymbol{x}_{em}$ but also matches the responses of a set of corresponding points in both spaces. The number of fine model points needed for this step is arbitrary. There was also no clear way of how to select the set of perturbations. Also, the available information about the mapping between the two spaces was not utilized.

In a later work Bakr et al. [3] suggested a better extraction procedure in which the vector of extracted parameters should satisfy

$$\boldsymbol{R}_{os}(\boldsymbol{x}_{os} + \boldsymbol{B}\Delta\boldsymbol{x}) = \boldsymbol{R}_{em}(\boldsymbol{x}_{em} + \Delta\boldsymbol{x}) \tag{4}$$

for a set of perturbations $\Delta \boldsymbol{x} \in V_P$ where $\mathbf{0} \in V_P$. In the context of space mapping this step is superior to (3) as it integrates the available mapping $\boldsymbol{B}$ into the multi-point extraction step. It is also suggested in [3] that the parameter extraction step terminates if the vector of extracted coarse model parameters approaches a limit. The algorithm generated the perturbations used for the multi-point parameter extraction process.

The perturbations used in [3] are not guaranteed to result in the best possible improvement in the uniqueness of the extracted parameters. A large number of additional fine model simulations might be needed to ensure the uniqueness of the step.

The main aim of this work is to find the optimal set of perturbations in (4) that results in the maximum possible improvement in the uniqueness of the step with every new perturbation. The perturbations are determined in the coarse model space using coarse model simulations only. Once a perturbation is predicted in the coarse model space it is mapped back to the fine model space.

In the next section we introduce some of the terms used to characterize the solutions of the parameter extraction problem. Also, we introduce a novel theorem that helps in determining the best possible perturbation according to the information available about the problem under consideration.

III. THE SELECTION OF PERTURBATIONS

Assume that the set of perturbations V_p in (4) has a cardinality N_e. It follows that additional N_e–1 fine model points are used in the multi-point parameter extraction process in addition to the point $\boldsymbol{x}_{em}$. The vector of responses $\boldsymbol{R}$ used to match the two models is thus given by

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{R}_{em}(\boldsymbol{x}_{em}) \\ \boldsymbol{R}_{em}(\boldsymbol{x}_{em} + \Delta \boldsymbol{x}^{(1)}) \\ . \\ . \\ \boldsymbol{R}_{em}(\boldsymbol{x}_{em} + \Delta \boldsymbol{x}^{(N_e-1)}) \end{bmatrix} \tag{5}$$

The dimensionality of the vector $\boldsymbol{R}$ is m_e where $m_e = N_e\, m$, where m is the dimensionality of $\boldsymbol{R}_{em}$. Applying multi-point parameter extraction using an optimization algorithm results in a vector of extracted coarse model parameters $\boldsymbol{x}_{os}^{e}$. This vector is labeled locally unique [6] if there exists an open neighborhood of $\boldsymbol{x}_{os}^{e}$ containing no other point $\boldsymbol{x}_{os}$ such that $\boldsymbol{R}(\boldsymbol{x}_{os}) = \boldsymbol{R}(\boldsymbol{x}_{os}^{e})$. It was shown in [6] that the local uniqueness condition is equivalent to the condition that the Jacobian of the vector of matched responses $\boldsymbol{R}$ has rank n.

To achieve this condition of local uniqueness, it was suggested in the context of system identification [4] that increasing the number of perturbations used increases the dimensionality of the vector $\boldsymbol{R}$ and thus enhances the possibility that the Jacobian matrix $\boldsymbol{J}$ of these responses becomes a full rank matrix. The perturbations suggested by Daijavad [4] were unidentified perturbations and thus result in an increase in the number of the optimizable parameters. However, it was pointed out that the improvement in rank of the matrix $\boldsymbol{J}$ outweighs the increase in the number of designable parameters.

In a later work, the idea of using known perturbations to achieve global uniqueness of the parameter extraction step was introduced [5]. By global uniqueness we mean that there exists only one minimum $\boldsymbol{x}_{os}^{e}$ for the multi-point parameter extraction problem. The existence of more than one locally unique solution for the parameter extraction step may result in the divergence or oscillation of the space mapping algorithm. It was also pointed out in [5] that using multi-point parameter extraction is equivalent to matching the first-order derivatives of the coarse and fine models and thus enhances the uniqueness of the parameter extraction step.

We suggest two different types of perturbation depending on whether the solution of the multi-point extraction step is locally unique or not. If the solution obtained is not locally unique we choose a perturbation that is more likely to make the new extracted parameters locally unique. First we define the vector of matched coarse model responses

$$\boldsymbol{R}_s = \begin{bmatrix} \boldsymbol{R}_{os}(\boldsymbol{x}_{os}) \\ \boldsymbol{R}_{os}(\boldsymbol{x}_{os} + \Delta\boldsymbol{x}^{(1)}) \\ . \\ . \\ \boldsymbol{R}_{os}(\boldsymbol{x}_{os} + \Delta\boldsymbol{x}^{(N_e-1)}) \end{bmatrix} \tag{6}$$

Assume that a locally nonunique minimum was obtained and that the rank of the Jacobian $\boldsymbol{J}$ of $\boldsymbol{R}_s$ is k where $k < n$. We suggest a perturbation $\Delta\boldsymbol{x}$ that attempts to increase the rank of the Jacobian of the matched responses at this minimum by at least one. This is achieved by imposing the condition that the gradients of $(n\text{–}k)$ of the coarse model responses generated by the new coarse model point $\boldsymbol{x}_{os}^e + \Delta\boldsymbol{x}$ be normal to the gradients of a linearly independent set of gradients of cardinality k of the responses in the vector $\boldsymbol{R}_s$ at the point $\boldsymbol{x}_{os} = \boldsymbol{x}_{os}^e$. We denote the set of linearly independent vectors by S where

$$S = \left\{ \boldsymbol{g}^{(1)}, \ldots, \boldsymbol{g}^{(k)} \right\} \tag{7}$$

We denote the set of the gradients of the newly selected $(n\text{–}k)$ responses in $\boldsymbol{R}_{os}(\boldsymbol{x}_{os}^e + \Delta\boldsymbol{x})$ by S_a where

$$S_a = \left\{ \boldsymbol{g}_a^{(k+1)}, \ldots, \boldsymbol{g}_a^{n} \right\} \tag{8}$$

where each of these gradients is approximated by

$$\boldsymbol{g}_a^{(i)} = \boldsymbol{g}^{(i)} + \boldsymbol{G}^{(i)}\Delta\boldsymbol{x}, \qquad i=k+1, \ldots, n \tag{9}$$

where $\boldsymbol{g}^{(i)}$ is the gradient of the ith response at the point $\boldsymbol{x}_{os}^e$ and $\boldsymbol{G}^{(i)}$ is the corresponding Hessian. The imposed condition on the perturbation is that

$$\boldsymbol{g}_a^{(i)T}\boldsymbol{g}^{(j)} = 0 \quad \forall\, \boldsymbol{g}^{(j)} \in S \;\text{ and }\; \forall\, \boldsymbol{g}_a^{(i)} \in S_n \tag{10}$$

From (9) and (10) it is clear that the perturbation $\Delta \boldsymbol{x}$ is obtained by solving the system of linear equations

$$\boldsymbol{A}^T \Delta \boldsymbol{x} = -\boldsymbol{c} \tag{11}$$

where the matrix $\boldsymbol{A}$ is given by

$$\boldsymbol{A} = \left[\boldsymbol{G}^{(k+1)} \boldsymbol{g}^{(1)} \,.\;.\, \boldsymbol{G}^{(n)} \boldsymbol{g}^{(1)} \ldots.. \boldsymbol{G}^{(n)} \boldsymbol{g}^{(k)}\right] \tag{12}$$

and the vector $\boldsymbol{c}$ is given by

$$\boldsymbol{c} = \begin{bmatrix} \boldsymbol{g}^{(k+1)T} \boldsymbol{g}^{(1)} \\ . \\ \boldsymbol{g}^{(n)T} \boldsymbol{g}^{(1)} \\ . \\ \boldsymbol{g}^{(n)T} \boldsymbol{g}^{(k)} \end{bmatrix} \tag{13}$$

A complete derivation of (11) is given in Appendix A. It should be noted that the system of linear equations (11) may be an over-determined, under-determined or well-determined system of equations depending on k and n. The pseudoinverse of the matrix $\boldsymbol{A}^T$ obtains the solution with minimum L_2 norm for all cases. The fact that this solution is a minimum length solution is of importance since (11) is based on a linear approximation to the gradients of the responses which can only be trusted within a certain trust region. If the perturbation $\Delta \boldsymbol{x}$ is outside this trust region it is rescaled to satisfy this condition.

The perturbation predicted by (11) attempts to increase the rank of the Jacobian of the matched responses in the multi-point parameter extraction by at least one. It should be noted that no condition is imposed for the dependency of the gradients of the selected (n-k) responses at the new point. It follows that the improvement in the rank may be higher than one especially if $k << n$.

If the minimum obtained by the multi-point parameter extraction is a locally unique minimum we still have to ensure that this is the true solution to the extraction problem. The following theorem leads us to the best possible way to eliminate any other existing locally unique minimum.

Theorem 1

Assume that there exist two locally unique minima $\boldsymbol{x}_{os}^{e,1}$ and $\boldsymbol{x}_{os}^{e,2}$ for the multi-point parameter extraction problem obtained using a set of perturbations V_p. A possible perturbation $\Delta \boldsymbol{x}$ that can be added to the set V_p and be can be used to eliminate one of these minima as a solution for the multi-point parameter extraction is in the direction of an eigenvector for the matrix $\boldsymbol{H}_1 - \boldsymbol{H}_2$ where $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ are the Hessian matrices for the objective function used by the optimizer at the points $\boldsymbol{x}_{os,1}^{e}$ and $\boldsymbol{x}_{os,2}^{e}$, respectively.

Proof

It is assumed that a Hessian matrix can be defined for the objective function utilized. The quadratic approximations of the objective function in a neighborhood centered at the two locally unique minima $\boldsymbol{x}_{os}^{e,1}$ and $\boldsymbol{x}_{os}^{e,2}$, respectively are given by

$$q_1(\Delta\boldsymbol{x}) = f_1 + 0.5\,\Delta\boldsymbol{x}^T \boldsymbol{H}_1\,\Delta\boldsymbol{x} \tag{14}$$

$$q_2(\Delta\boldsymbol{x}) = f_2 + 0.5\,\Delta\boldsymbol{x}^T \boldsymbol{H}_2\,\Delta\boldsymbol{x} \tag{15}$$

where f_1 and f_2 are the values of the objective function at the two respective points. The perturbation $\Delta\boldsymbol{x}$ that results in the maximum difference between the two quadratic models (14) and (15) for a specific trust region $\boldsymbol{d}$ is obtained by formulating the Lagrangian

$$L(\boldsymbol{x},\ l) = (f_2 - f_1) + 0.5\,\Delta\boldsymbol{x}^T(\boldsymbol{H}_2 - \boldsymbol{H}_1)\Delta\boldsymbol{x} + \ l(\Delta\boldsymbol{x}^T\Delta\boldsymbol{x} \ - d^2) \tag{16}$$

Taking the derivative with respect to $\Delta\boldsymbol{x}$ gives

$$(\boldsymbol{H}_1 - \boldsymbol{H}_2)\Delta\boldsymbol{x} = 2\ l\Delta\boldsymbol{x} \tag{17}$$

It follows that $\Delta\boldsymbol{x}$ is an eigenvector of the matrix $(\boldsymbol{H}_1 - \boldsymbol{H}_2)$. The perturbation $\Delta\boldsymbol{x}$ provides the direction that maximizes the difference between the quadratic models. In other words, it provides a perturbation that maximizes the contrast between the changes of the coarse model responses at these two minima. It follows that the true minimum is the one whose response changes better match the changes of the fine model responses obtained using the perturbation $\Delta\boldsymbol{x}$.

A similar result to that obtained in (17) can be obtained using a different approach. Assume that a perturbation of $\Delta \boldsymbol{x}$ is sought. This perturbation results in a perturbation of the coarse model responses at the two minima by

$$\Delta \boldsymbol{R}_1 = \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,1})\Delta \boldsymbol{x} \tag{18}$$

and

$$\Delta \boldsymbol{R}_2 = \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,2})\Delta \boldsymbol{x} \tag{19}$$

We impose the condition that the difference between the ℓ_2 norms of these two response perturbations be maximized subject to certain trust region size. It follows that the following Lagrangian can be formed

$$L(\Delta \boldsymbol{x},\ l) = \Delta \boldsymbol{x}^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,1})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,1})\Delta \boldsymbol{x} - \Delta \boldsymbol{x}^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,2})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,2})\Delta \boldsymbol{x} + l(\Delta \boldsymbol{x}^T \Delta \boldsymbol{x} - d^2) \tag{20}$$

Using a similar approach to that used in deriving (17) it can be shown that the perturbation $\Delta \boldsymbol{x}$ is an eigenvector for the matrix

$$\boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,1})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,1}) - \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,2})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,2}) \tag{21}$$

The perturbation predicted by (17) is almost identical to that predicted by (21) if the L_2 objective function is utilized in (17). This follows directly when the Hessian matrix of the L_2 objective function is derived [7].

The perturbations given by (17) or alternatively by (21) are hard to evaluate. This is simply attributed to the fact that once a locally unique minimum is reached the Hessian of the objective function at this point can be obtained while no information is available about the Hessian of the objective function at the other locally unique minima that may exist. In such a case, the only reasonable assumption that can be made about the other minima is that $\boldsymbol{H}_2$=$\boldsymbol{I}$, the identity matrix in (17) or alternatively that the matrix $\boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,2})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,2})$ is the identity in (21). It follows that $\Delta \boldsymbol{x}$ should be an eigenvector of the matrix $\boldsymbol{H}_1$ in (17) and that $\Delta \boldsymbol{x}$ should be an eigenvector of the matrix $\boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,1})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,1})$ in (21). For all the examples that was tried the l_2 objective function was utilized and thus (17) and (21) are practically identical.

The perturbation determined by (11) or (17) is a suggested perturbation in the coarse model space. The new fine model point that should be added to the set of fine model points used for the multi-point parameter extraction is

$$\boldsymbol{x} = \boldsymbol{x}_{em} + \boldsymbol{B}^{-1}\Delta\boldsymbol{x} \tag{22}$$

The relation (22) is used if some available information is available about the mapping between the two spaces. However, in most cases we make the assumption that $\boldsymbol{B}=\boldsymbol{I}$.

The scheme that we utilized for the selection of points in (17) is as follows. First, the eigenvalue problem is solved. The eigenvector $\boldsymbol{v}^{(1)}$ with the largest eigenvalue in modulus is initially selected as the candidate eigenvector. The suggested perturbation in this case is

$$\Delta\boldsymbol{x} = \frac{d}{\left\|\boldsymbol{v}^{(1)}\right\|}\boldsymbol{v}^{(1)} \tag{23}$$

This perturbation is tested to see whether it belongs to the already used set of perturbations. It follows that the perturbation is rejected if the condition

$$\frac{\Delta\boldsymbol{x}^T \Delta\,\boldsymbol{x}^{(i)}}{\left\|\Delta\boldsymbol{x}\right\|^2} > (1-e) \tag{24}$$

is satisfied for a perturbation $\Delta\boldsymbol{x}^{(i)} \in V_p$. In this case the alternative perturbation

$$\Delta\boldsymbol{x} = \frac{-\,d}{\left\|\boldsymbol{v}^{(1)}\right\|}\boldsymbol{v}^{(1)} \tag{25}$$

is tested against the condition (24). If it also fails we switch to the eigenvector with second largest eigenvalue in modulus and repeat steps (23) – (25). This is repeated until either a perturbation is found such that condition (24) is not satisfied or all the eigenvectors are exhausted for perturbations of length d In this case the trust region size d is scaled by a where a >1.0 and the perturbation is taken in the direction of eigenvector with largest eigenvalue in modulus.

In the next section we introduce an automated multi-point parameter extraction algorithm. The approaches developed in this section are utilized in this algorithm.

IV. THE MULTI-POINT PARAMTER EXTRACTION ALGORITHM

In this section we present an algorithm for the multi-point parameter extraction process. This algorithm is based on the two methods discussed in the previous section. These two methods tackle the cases of a locally unique minimum and the non locally unique minimum. The algorithm is given by the following steps

Step 0 Given $\boldsymbol{x}_{em}$, d and n. Initialize $V = \{\boldsymbol{x}_{em}\}$ and i=1.

Comment The set V contains the points used for the multi-point parameter extraction. The index i is an iteration count and also indicates the number of fine model points used in the multi-point parameter extraction.

Step 1 Apply multi-point parameter extraction using the points in the set V to get $\boldsymbol{x}_{os}^{e}$.

Step 2 If the Jacobian of the matched responses at $\boldsymbol{x}_{os}^{e}$ has a full rank go to Step 4.

Step 3 Obtain a new perturbation $\Delta\boldsymbol{x}$ using (11) and add the point $\boldsymbol{x}_{em} + \Delta\boldsymbol{x}$ to the set V. Set i=i+1 and go to Step 1.

Comment The perturbation $\Delta\boldsymbol{x}$ is rescaled to satisfy the trust region condition $\|\Delta\boldsymbol{x}\| = d$.

Step 4 If $|V|$ is equal to one go to Step 6.

Comment $|V|$ denotes the cardinality of the set V.

Step 5 If $\boldsymbol{x}_{os}^{e}$ is approaching a limit stop.

Step 6 Obtain a new perturbation $\Delta\boldsymbol{x}$ using (17) with $\boldsymbol{H}_2$=$\boldsymbol{I}$, update d and add the point $\boldsymbol{x}_{em} + \Delta\boldsymbol{x}$ to the set V, set i=i+1 and go to Step 1.

Comment In Step (6) the eigenvalue problem is solved and the perturbation $\Delta\boldsymbol{x}$ is selected according to scheme discussed in the previous section. This scheme may result in updating the trust region size. The algorithm terminates if the vector of extracted coarse model parameters obtained using i fine model points is close enough in terms of some norm to the vector of extracted coarse model parameters obtained using ($i-1$) fine model points.

A flowchart of this algorithm is shown in Fig. 1.

V. EXAMPLES

In this section we review a number of examples that are used to illustrate the algorithm. These examples are discussed in the context of space mapping. The parameters to be extracted are the coarse model parameters. The L_2 optimizer available in OSA90/hope [8] was utilized in the parameter extraction process for all the examples.

The Rosenbrock Function

The first example utilizes the Rosenbrock function [7]. The coarse model for this problem is given by

$$R_{os} = 100\,(u_2 - u_1^2)^2 + (1 - u_1)^2 \tag{26}$$

The fine model is another Rosenbrock function but with a shift applied to the parameters

$$R_f = 100\,((u_2 + 0.2) - (u_1 - 0.2)^2)^2 + (1 - (u_1 - 0.2))^2 \tag{27}$$

It is required to extract the coarse model parameters corresponding to the fine model point $[1.0 \quad 1.0]^T$. In the first iteration we apply single point parameter extraction at this point. The vector $\boldsymbol{x}_{os}^e = [1.21541 \quad 0.91728\,]^T$ is obtained. The contours corresponding to this single point parameter extraction process are shown in Fig. 2. It is clear from the contour plot that the minimum obtained is not a locally unique minimum. The algorithm detects this and generates a step that attempts to improve the rank of the Jacobian of the matched responses in the two-point parameter extraction using (11). Utilizing a trust region size of 0.25 the set of points used for the multi-point parameter extraction is

$$V = \left\{ [1.0 \quad 1.0]^T, [0.7643 \quad 1.0833]^T \right\} \tag{28}$$

The contours of the L_2 objective function for the two-point parameter extraction step based on the set V in (28) is shown in Fig. 3. It is clear that by using only one other fine model simulation the uniqueness of the problem has improved dramatically. Actually, the only existing minimum is a global unique minimum. To ensure the uniqueness of this step a third point is generated by the algorithm using the eigenvectors of the Hessian of the objective function. The set of points used for the three-point parameter extraction are

$$V = \left\{ \begin{bmatrix} 1.0 & 1.0 \end{bmatrix}^T, \begin{bmatrix} 0.7643 & 1.0833 \end{bmatrix}^T, \begin{bmatrix} 0.8185 & 0.8281 \end{bmatrix}^T \right\} \tag{29}$$

The contours of the L_2 objective function are shown in Fig. 4. The algorithm then terminates as it detects that extracted parameters are approaching a limit and returns this last set of extracted parameters as the solution for the multi-point parameter extraction. The change of the extracted parameters obtained using the L_2 optimizer with the number of fine model points used are shown in Table I.

A 10:1 *Impedance Transformer*

The second example is a 10:1 impedance transformer [9]. The parameters for this problem are the characteristic impedance of the two transmission lines while the two lengths of the transmission lines are kept fixed at their optimal values (quarter wave length). The coarse model utilizes non scaled parameters while the fine model scales each of the two impedances by a factor of 1.6.

It is required in this problem to extract the coarse model parameters which matches the fine model responses at the point $[2.2628 \quad 4.5259]^T$. This point is the optimal coarse model design according to the specifications in [9].

First we assume that the two models are matched using the single response

$$r = \max_i r_i, \qquad 1 \le i \le 11 \tag{30}$$

where r_i, $1 \le i \le 11$ are the reflection coefficient calculated at 11 equally spaced frequencies in the frequency range 0.5 GHz $\le f \le$ 1.5 GHz. Initially, the algorithm carries out a single point parameter extraction. The contours of the objective function are shown in Fig. 5. It is clear from the figure that the extracted solution is not a locally unique minimum. The algorithm detects that and generates a perturbation that enhances the uniqueness for the two-point parameter extraction using (11). The set of points used for the two-point parameter extraction are

$$V = \left\{ [2.2628 \quad 4.5259]^T, [2.7092 \quad 4.7510]^T \right\} \tag{31}$$

The contours of the objective function are shown in Fig. 6. It is clear that two locally unique minima have been formed. For the next three steps of the extraction algorithm, perturbations were generated based on the eigenvectors of the Hessian of objective function at the reached minima. The contours of the L_2 objective function are shown in Figs. 7, 8 and 9. It is clear from Fig. 9 that only a global unique minimum exists. The last perturbation added to the set of points used for parameter extraction enhanced the uniqueness significantly. This is due to the fact the eigenvectors of the two existent locally unique minima were quite different. The variation of the extracted parameters with the number of points used for the multi-point parameter extraction are shown in Table II.

The same example was tried using the reflection coefficients at the 11 equally spaced frequencies directly as matching responses. First, we applied single point parameter extraction. The response of the fine model point and the corresponding extracted coarse model point are shown in Fig. 10. The contours of the L_2 objective function for single point parameter extraction are shown in Fig. 11. It is clear from this figure that there exist three locally unique minima for the extraction problem. The algorithm generates a suggested perturbation using (17). The set of fine model points utilized in the two-point parameter extraction is

$$V = \left\{ [2.26277 \quad 4.52592]^T, [1.49975 \quad 4.76634]^T \right\} \tag{32}$$

The fine model response for each point in the set V in (32) and the corresponding extracted coarse model point are shown in Fig. 12. The corresponding L_2 contours of the two-point parameter extraction are shown in Fig. 13. It is clear that there still exist two locally unique minima. The algorithm generates a third perturbation using (17). The set of fine model points utilized in the three-point parameter extraction is

$$V = \left\{ \begin{bmatrix} 2.26277 & 4.52592 \end{bmatrix}^T, \begin{bmatrix} 1.49975 & 4.76634 \end{bmatrix}^T, \begin{bmatrix} 3.02024 & 4.26855 \end{bmatrix}^T \right\} \qquad (33)$$

The fine model response for each point in the set V in (32) and the coarse model response at the corresponding extracted coarse model point are shown in Fig. 14. The corresponding L_2 contours of the three-point parameter extraction are shown in Fig. 15. The algorithm terminates as the termination condition is satisfied. The change of the extracted coarse model point with the number of fine model points is given in Table III.

VI. CONCLUSIONS

In this work we presented a novel algorithm for the multi-point parameter extraction. The algorithm enhances the uniqueness of the problem based on the nature of the local minimum reached using parameter extraction. The algorithm systematically improves the uniqueness of the problem until only one minimum of the extraction problem exists. The suggested algorithm was successfully tried on a number of examples.

REFERENCES

[1] J.W. Bandler, R.M. Biernacki, S.H. Chen, P.A. Grobelny and R.H. Hemmers, "Space mapping technique for electromagnetic optimiztion," *IEEE Trans. Microwave Theory Tech.*, vol. 42, 1994, pp. 2536-2544.

[2] J.W. Bandler, R.M. Biernacki, S.H. Chen, R.H. Hemmers and K. Madsen, "Electromagnetic optimization exploiting aggressive space mapping," *IEEE Trans. Microwave Theory Tech.*, vol. 43, 1995, pp. 2874-2882.

[3] M.H. Bakr, J.W. Bandler, R.M. Biernacki, S.H. Chen and K. Madsen, " A trust region aggressive space mapping algorithm for EM optimization," *IEEE MTT-S Int. Microwave Symp. Dig.* (Baltimore, MD.), 1998, pp. 1759-1762.

[4] S. Daijavad, Design and Modelling of Microwave Circuits Using Optimization Methods, Ph.D. Thesis, McMaster University, 1986.

[5] J.W. Bandler, R.M. Biernacki and S.H. Chen, "Fully automated space mapping optimization of 3D structures," *IEEE MTT-S Int. Microwave Symp. Dig.* (San Francisco, CA), 1996, pp. 753-756.

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

[7] R. Fletcher, Practical Methods of Optimization, Second Edition, John Wiley & Sons Ltd., 1987.

[8] OSA90/hope, Hewlett-Packard Co., 1400 Fountaingrove Parkway, Santa Rosa, CA 95403, 1997.

[9] J.W. Bandler, "Optimization methods for computer-aided design," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, 1969, pp. 533-552.

TABLE I
THE VARIATIONS OF THE EXTRACTED PARAMETERS FOR THE ROSENBROCK FUNCTION WITH THE NUMBER OF POINTS USED FOR EXTRACTION

Number of Points	$x_{os,1}^{e}$	$x_{os,2}^{e}$
1	1.21541	0.91728
2	0.80008	1.20012
3	0.80008	1.20014

TABLE II
THE VARIATIONS OF THE EXTRACTED PARAMETERS FOR THE 10:1 IMPEDANCE TRANSFORMER WITH THE NUMBER OF POINTS USED FOR EXTRACTION USING A SINGLE MATCHED RESPONSE

Number of Points	$x_{os,1}^{e}$	$x_{os,2}^{e}$
1	1.81215	5.73398
2	1.38297	2.62400
3	1.63656	3.08426
4	4.31022	7.17462
5	3.98739	7.26278

TABLE III
THE VARIATIONS OF THE EXTRACTED PARAMETERS FOR THE 10:1 IMPEDANCE TRANSFORMER WITH THE NUMBER OF POINTS USED FOR EXTRACTION USING ELEVEN MATCHED RESPONSE

Number of Points	$x_{os,1}^{e}$	$x_{os,2}^{e}$
1	3.62043	7.24147
2	3.47160	7.43214
3	3.60357	7.35052

APPENDIX A

Let the two sets S and S_a be defined as in (7) and (8), respectively. We impose the condition that every gradient on the set S_a should be orthogonal to all gradients in the set S. It follows that

$$\boldsymbol{g}_a^{(i)T}\boldsymbol{g}^{(j)} = 0 \quad \forall\, \boldsymbol{g}^{(j)} \in S \text{ and } \forall\, \boldsymbol{g}_a^{(i)} \in S_n \tag{34}$$

But each gradient in the set S_a can be approximated by

$$\boldsymbol{g}_a^{(i)} = \boldsymbol{g}^{(i)} + \boldsymbol{G}^{(i)}\Delta\boldsymbol{x}, \qquad i=k+1, . ., n \tag{35}$$

where $\boldsymbol{g}^{(i)}$ is the gradient of the ith response at the point $\boldsymbol{x}_{os}^{e}$ and $\boldsymbol{G}^{(i)}$ is the corresponding Hessian. It follows that the condition (10) can be restated as

$$\boldsymbol{g}^{(j)\boldsymbol{T}}\boldsymbol{g}^{(i)} = -\boldsymbol{g}^{(j)\boldsymbol{T}}\boldsymbol{G}^{(i)}\Delta\boldsymbol{x}, \quad j=1, . ., k \text{ and } i=k+1, . ., n \tag{36}$$

Equation (36) is a linear equation in n unknowns (the components of $\Delta\boldsymbol{x}$). There are $k(n-k)$ such linear equations. Putting these equations into a matrix form we have

$$\boldsymbol{B}\Delta\boldsymbol{x} = -\boldsymbol{c}. \tag{37}$$

where the mth row of the matrix $\boldsymbol{B}$ is

$$\boldsymbol{B}^{(m)T} = \boldsymbol{g}^{(j)T}\boldsymbol{G}^{(i)}, \qquad i=k+1, . ., n \text{ and } j=1, . ., k \tag{38}$$

where $m=(i-1)(n-k)+j$. Similarly the mth component of the vector $\boldsymbol{c}$ is

$$\boldsymbol{c}_{m,1} = \boldsymbol{g}^{(j)T}\boldsymbol{g}^{(i)}, \qquad i=k+1, . ., n \text{ and } j=1, . ., k \tag{39}$$

Thus (11) follows.

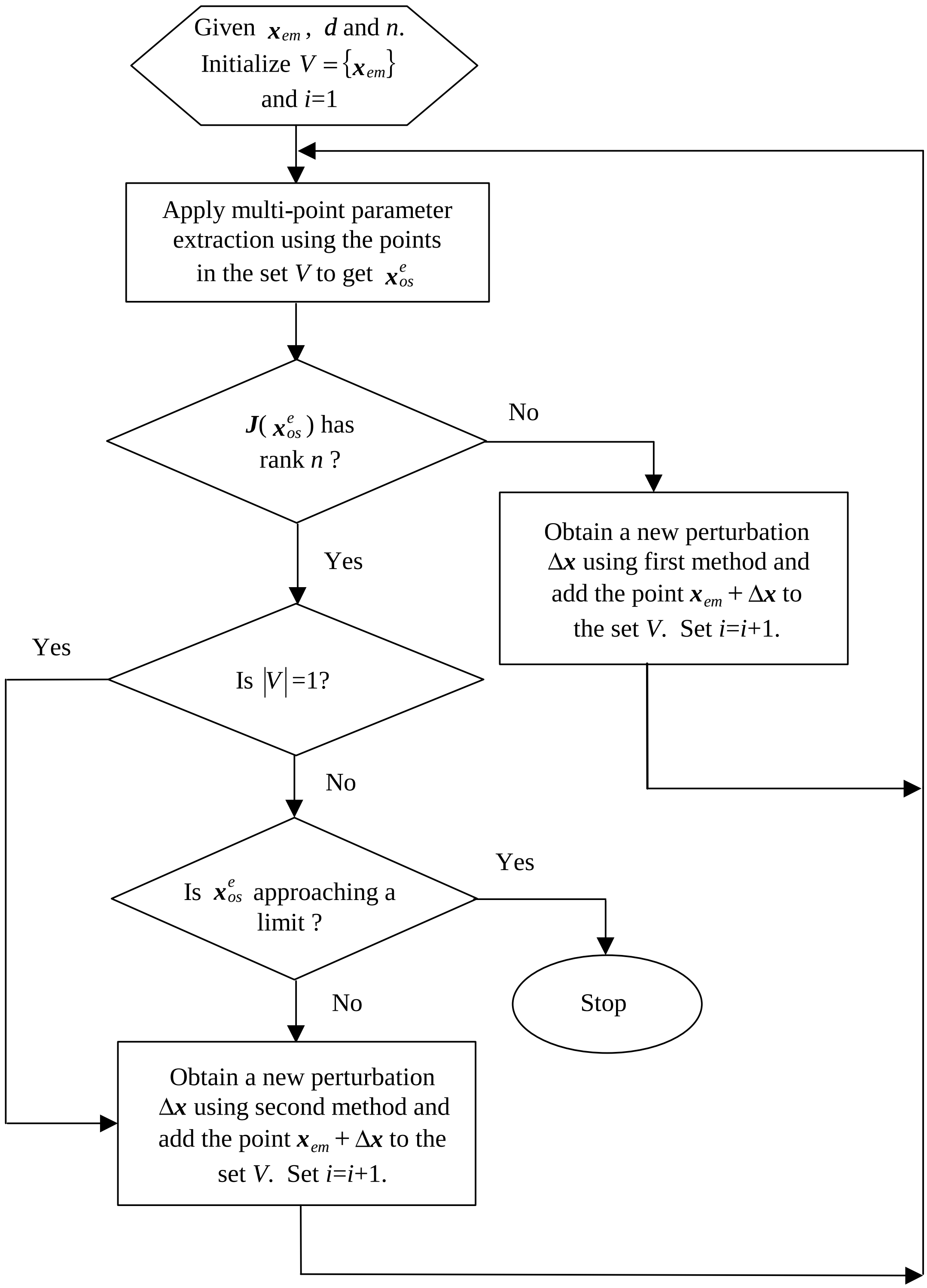


Fig. 1. The FlowChart of the Algorithm.

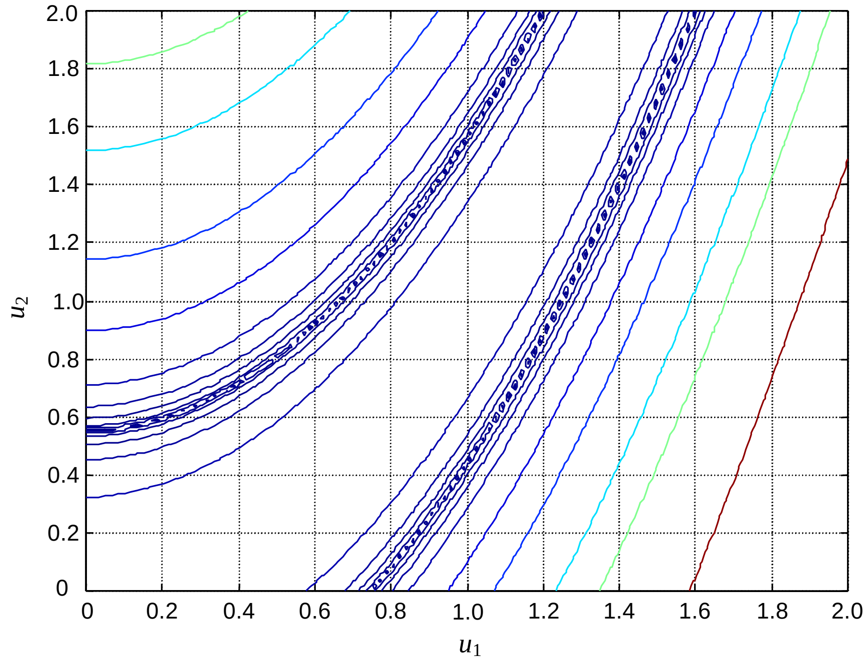


Fig. 2. The contours of the L_2 objective function of the single point parameter extraction for the Rosenbrock function.

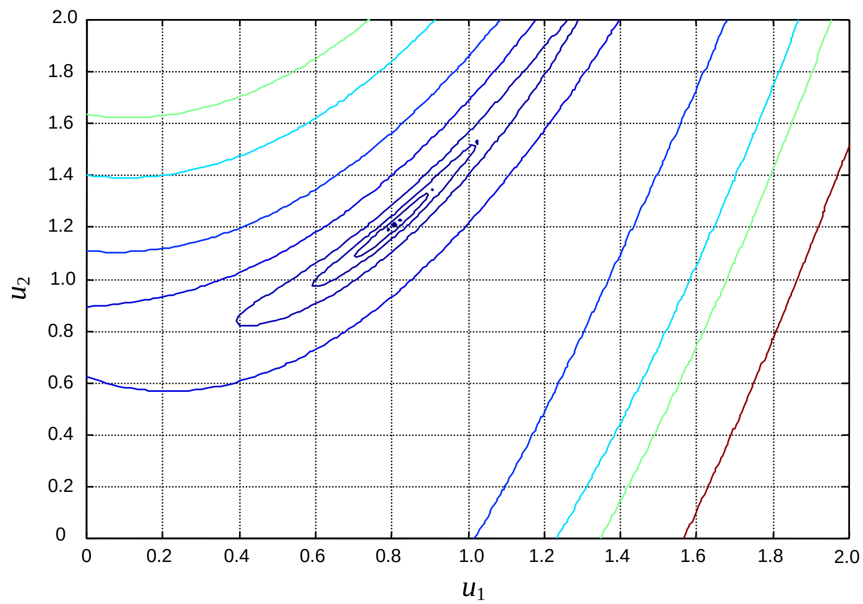


Fig. 3. The contours of the L_2 objective function of the two-point parameter extraction for the Rosenbrock function.

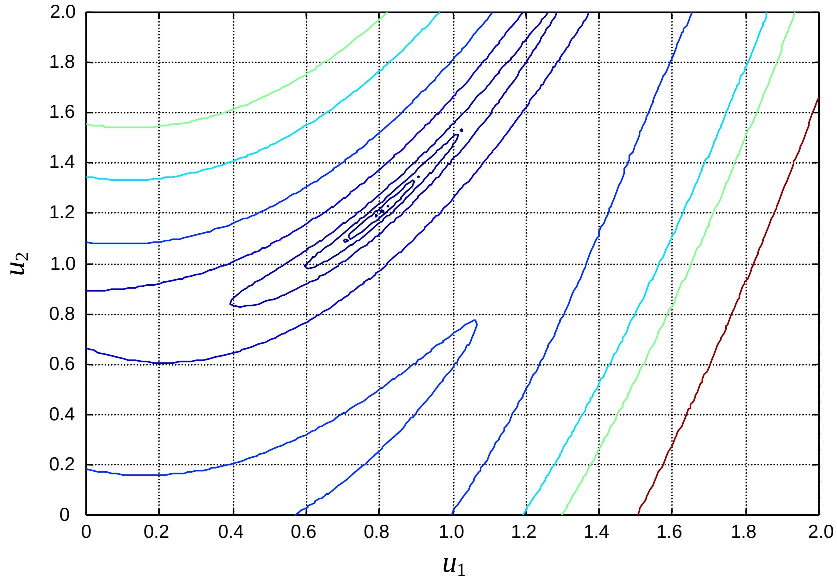


Fig. 4. The contours of the L_2 objective function of the three-point parameter extraction for the Rosenbrock function.

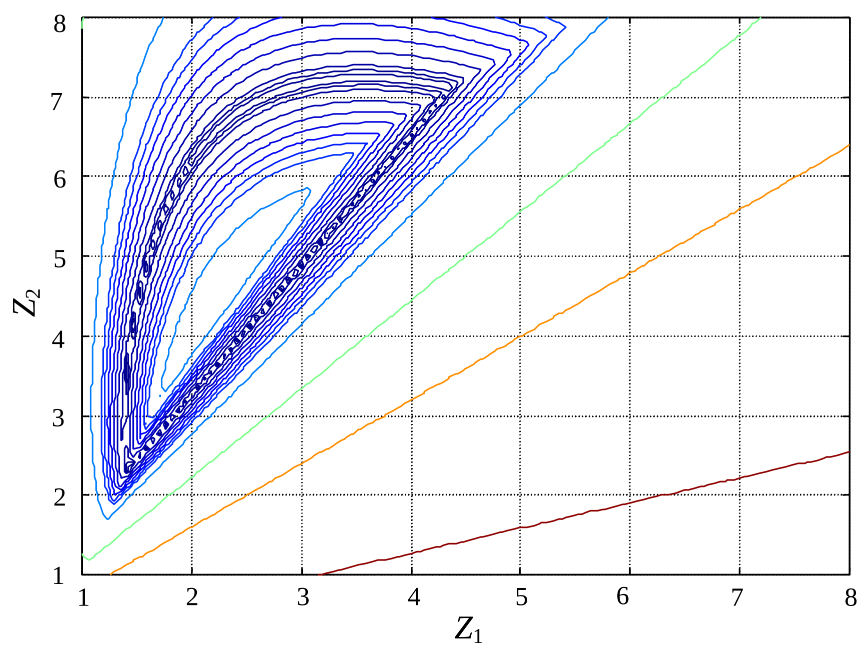


Fig. 5. The contours of the L_2 objective function of the single-point parameter extraction for the 10:1 transformer using a single matched response.

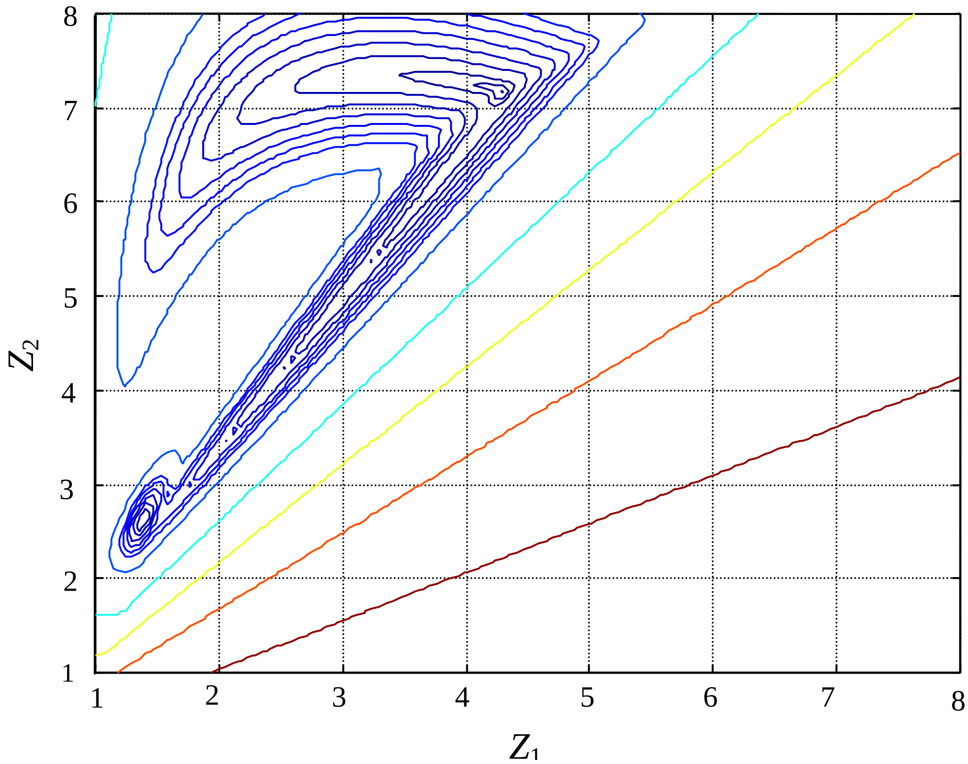


Fig. 6. The contours of the L_2 objective function of the two-point parameter extraction for the 10:1 transformer using a single matched response.

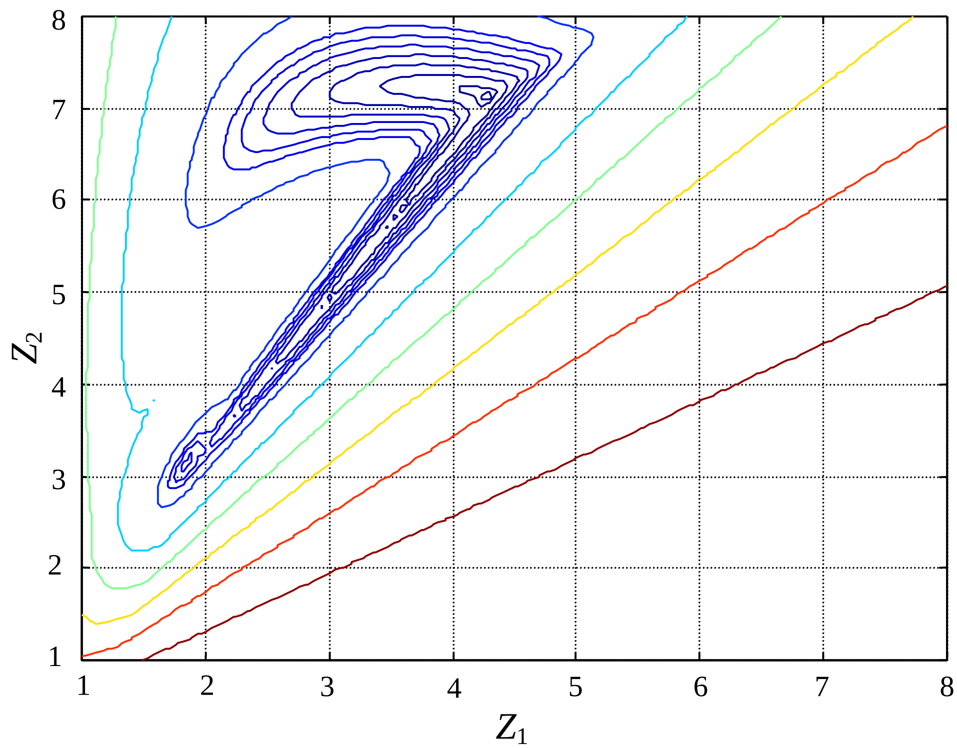


Fig. 7. The contours of the L_2 objective function of the three-point parameter extraction for the 10:1 transformer using a single matched response.

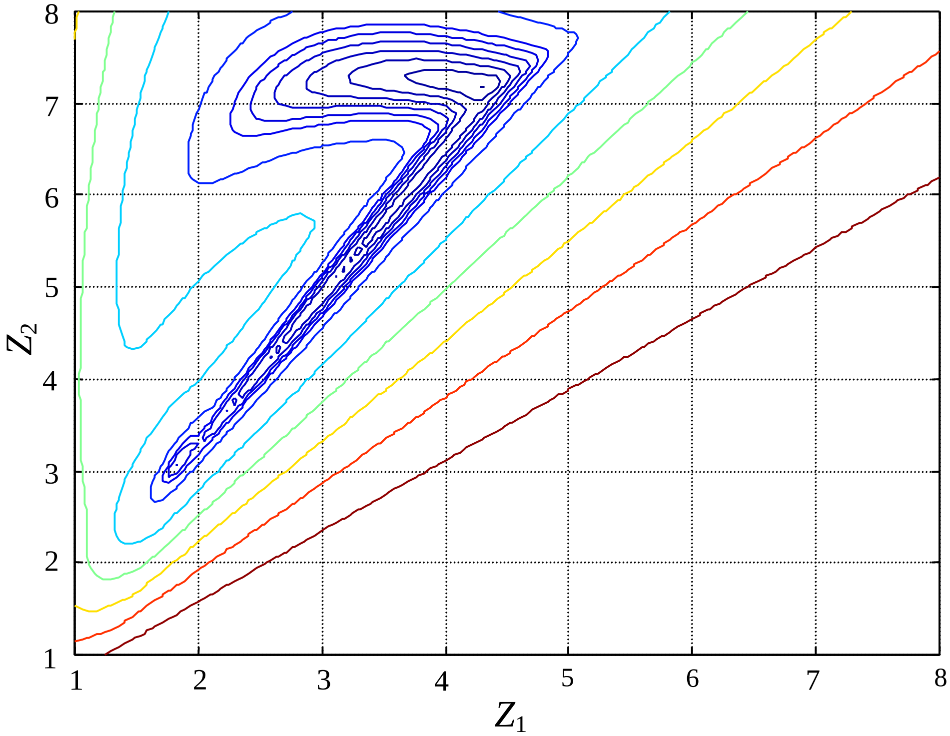


Fig. 8. The contours of the L_2 objective function of the four-point parameter extraction for the 10:1 transformer using a single matched response.

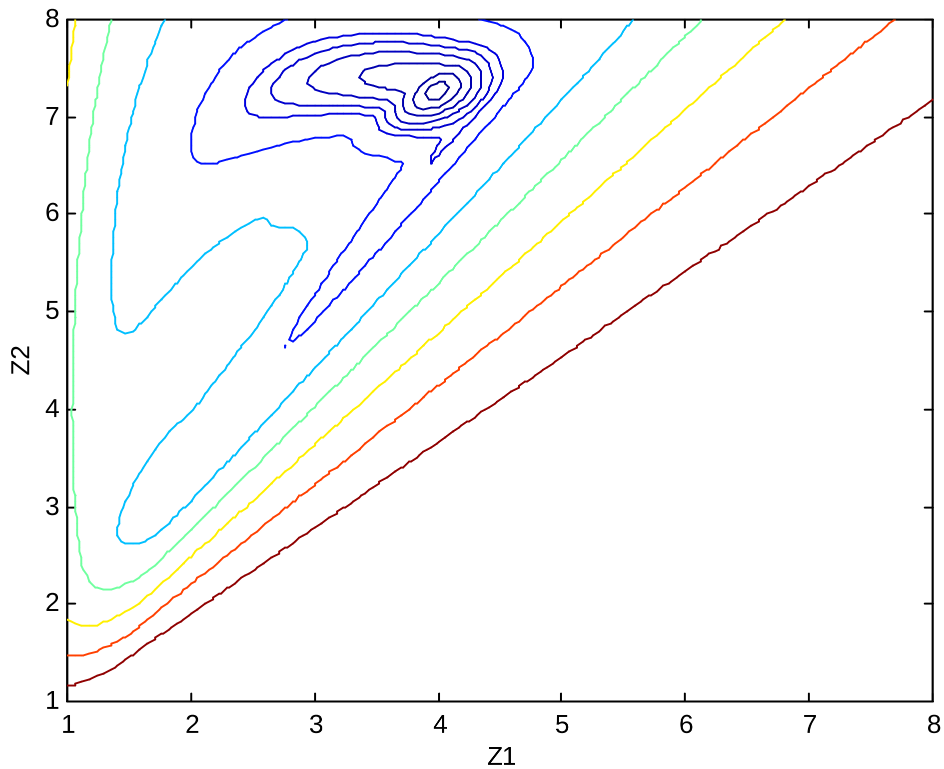


Fig. 9. The contours of the L_2 objective function of the five-point parameter extraction for the 10:1 transformer using a single matched response.

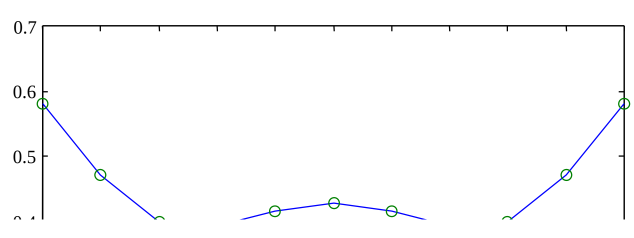


Fig. 10. The responses of the given fine model point (o) and the coarse model response (−) at the corresponding coarse model point obtained using single point parameter extraction.

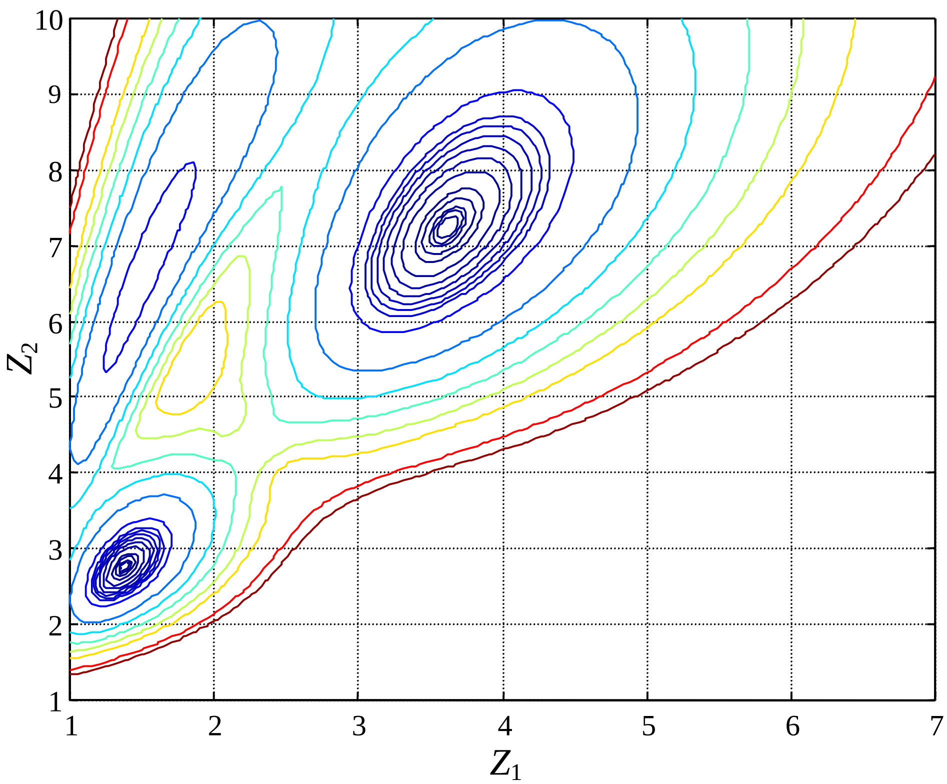
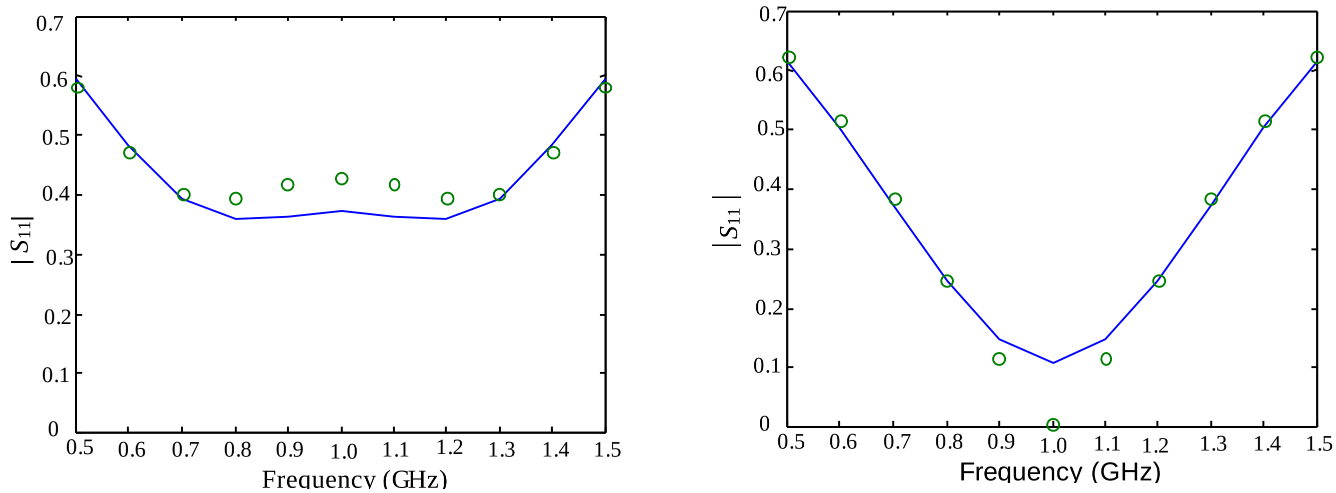


Fig. 11. The contours of the L_2 objective function of the single-point parameter extraction for the 10:1 transformer.



(a) (b)

Fig. 12. The fine model responses (o) and the corresponding coarse model responses (–) at the first point (a) and the second point (b) utilized in the two-point parameter extraction.

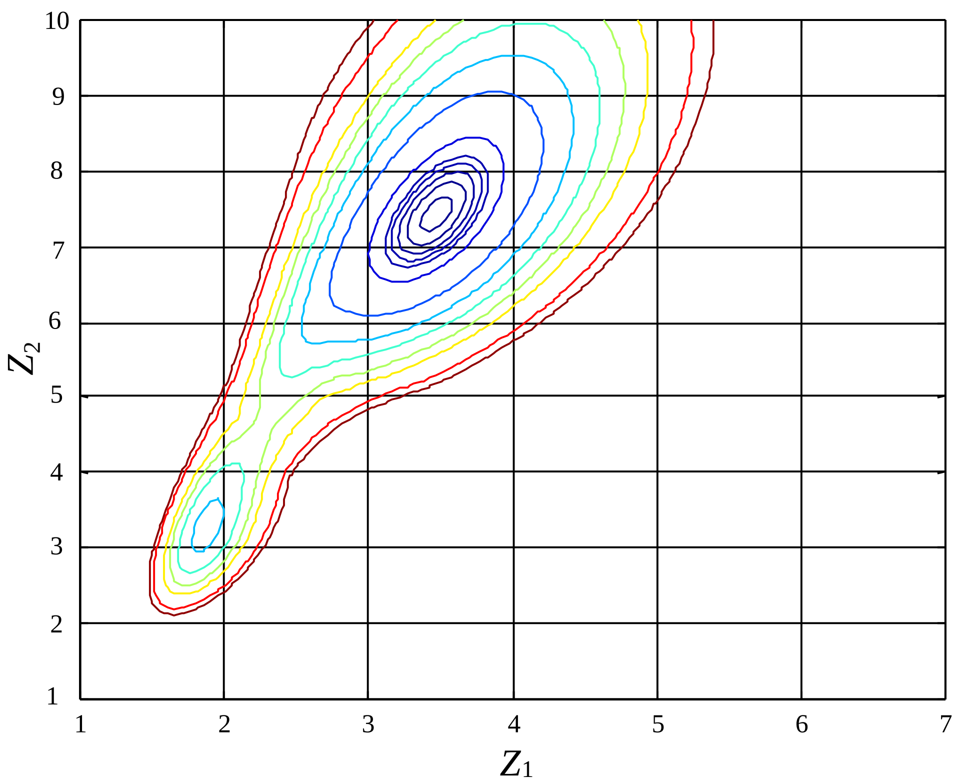


Fig. 13. The contours of the L_2 objective function of the two-point parameter extraction for the 10:1 transformer.

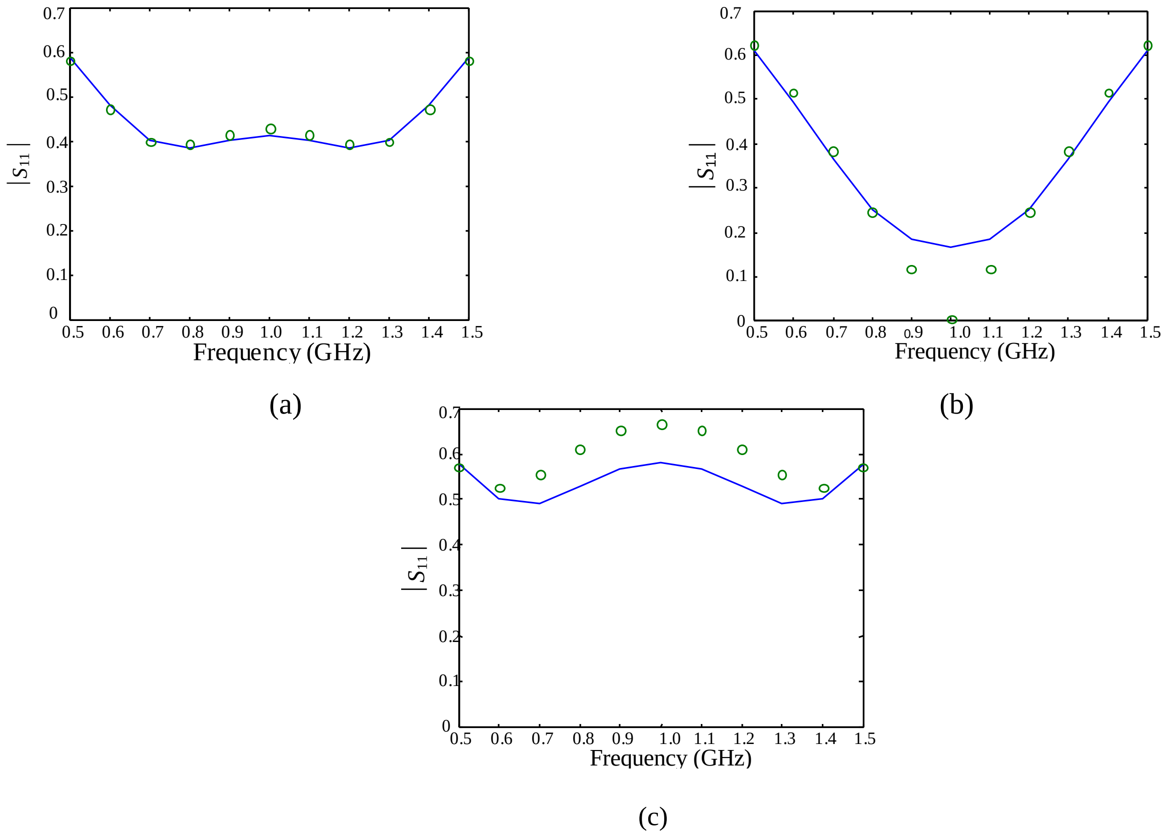


Fig. 14. The fine model responses (o) and the corresponding coarse model responses (–) at the first point (a), the second point (b) and the third point (c) utilized in the three-point parameter extraction.

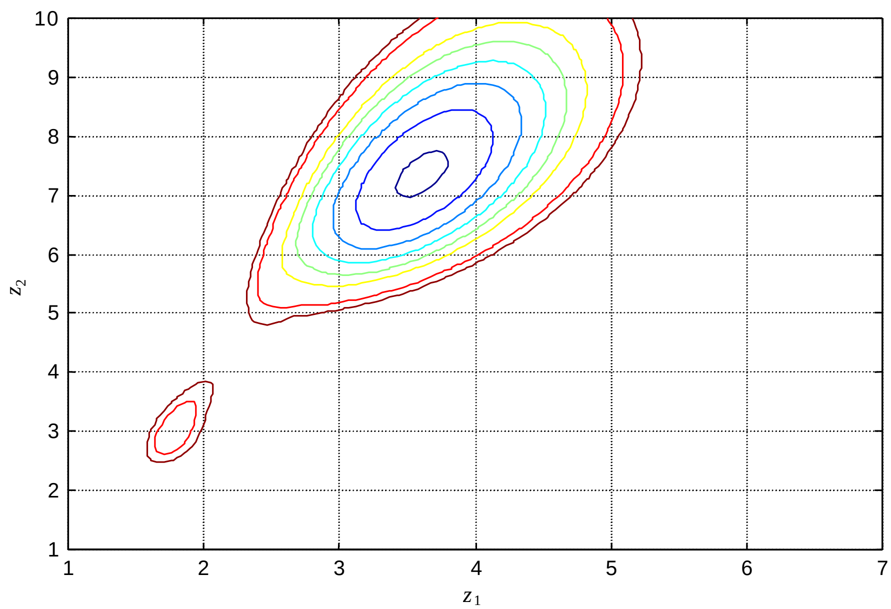


Fig. 15. The contours of the L_2 objective function of the three-point parameter extraction for the 10:1 transformer.