A SUGGESTED ALGORITHM FOR MULTI-POINT PARAMETER EXTRACTION

M.H. Bakr and J.W. Bandler

SOS-98-25-V

August 1998

A SUGGESTED ALGORITHM 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

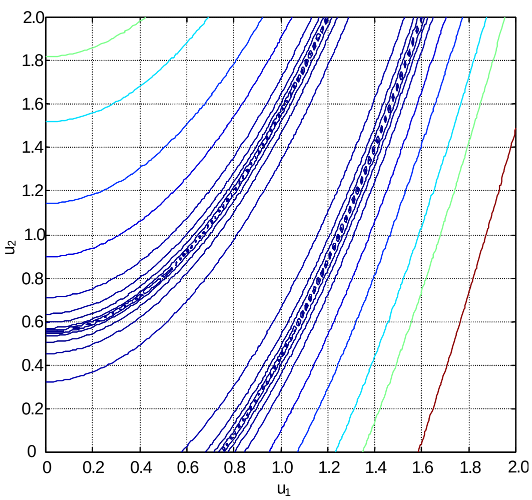
presented at

SOS Research Laboratory Meeting, Hamilton, August 24, 1998.

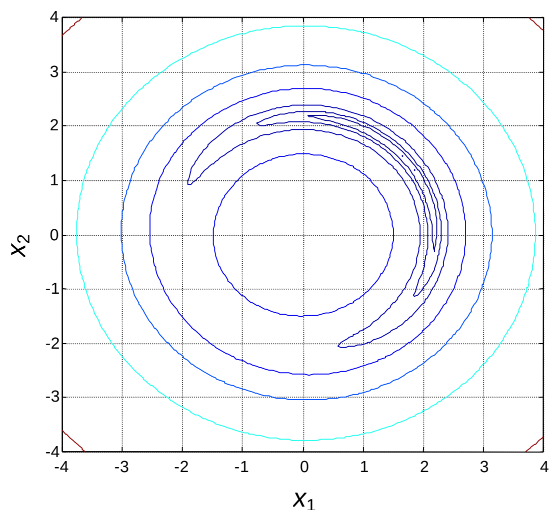
Classification of the Solution of Parameter Extraction

a solution $\boldsymbol{x}_{os}^{e}$ of the parameter extraction problem is labeled locally unique if there exists an open neighborhood of $\boldsymbol{x}_{os}^{e}$ containing no other point $\boldsymbol{x}_{os}$ such that $\boldsymbol{R}(\boldsymbol{x}_{os}) = \boldsymbol{R}(\boldsymbol{x}_{os}^{e})$, where $\boldsymbol{R}$ is the vector of matched coarse model responses

it can be shown that the local uniqueness condition is equivalent to the condition that the Jacobian of the vector of matched coarse model responses $\boldsymbol{R}$ has a rank n, where n is the number of optimizable parameters.



(a) non locally unique



b) locally unique

Problem Definition

assume that multi-point parameter extraction is carried out at the fine model point $\boldsymbol{x}_{em}$ using N_e fine model points

it follows that the vector of matched coarse model responses $\boldsymbol{R}$ is given by

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

where $\Delta\boldsymbol{x}^{(i)} \in V_p$; the set of utilized perturbations

it is required to find the perturbation $\Delta\boldsymbol{x}$ that can be added to the set V_P and maximizes the improvement of the uniqueness of the extraction step

Problem Assumptions

the coarse model is assumed to be much faster than the fine model

few extra coarse model simulations add negligible overhead to the computational time of the problem

it is also assumed that first and second order derivatives of the coarse model responses can be obtained

in the absence of information about the mapping between the two spaces we take $\boldsymbol{B}=\boldsymbol{I}$

the mapping $\boldsymbol{B}$ can be easily integrated with the suggested algorithm if it is available

A Suggested Method: the Non Locally Unique Case

assume that the rank of the Jacobian of matched coarse model responses at the point $\boldsymbol{x}_{os}^{e}$ is $k < n$

we impose the condition that the gradients of $(n–k)$ of the 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 $\boldsymbol{R}$

define the set of linearly independent gradients by

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

the gradient of each of the $(n–k)$ selected responses can be approximated by

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

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 perturbation $\Delta\boldsymbol{x}$ that satisfies the orthogonality condition is obtained by solving

Suggested Method: the Non Locally Unique Case

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

where

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

and

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

this system of linear equations may be under-determined, over-determined or well-determined

the pseudoinverse of the matrix $\boldsymbol{A}^T$ is used to find the solution of minimum length in all cases

the perturbation $\Delta \boldsymbol{x}$ is rescaled to satisfy a certain trust region condition

rescaling implies that the gradients might not be orthogonal but the independency property is retained

A Suggested Method: the Locally Unique Case

assume that a solution for the multi-point parameter extraction $\boldsymbol{x}_{os,1}^{e}$ was obtained and that the rank of the Jacobian of matched coarse model responses is k=n

a perturbation $\Delta\boldsymbol{x}$ is sought that distinguishes this locally unique minimum from other minima that may exist

assume the presence of another locally unique minimum $\boldsymbol{x}_{os,2}^{e}$

the quadratic models for the objective function in the neighborhood of these two minima are given by, respectively

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

and

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

where f_1 and f_2 are the values of the objective function at the points $\boldsymbol{x}_{os,1}^{e}$ and $\boldsymbol{x}_{os,2}^{e}$, respectively and $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ are the corresponding Hessian matrices

it can be shown that the direction that maximizes the difference between the two quadratic models is an eigenvector for the matrix $(\boldsymbol{H}_1 - \boldsymbol{H}_2)$

Suggested Method: the Locally Unique Case

the matrix $\boldsymbol{H}_2$ is assumed to be the identity as there is no available information about the other minima that may exist

it follows that the direction of $\Delta\boldsymbol{x}$ is an eigenvector for the matrix $\boldsymbol{H}_1$

the obtained direction of $\Delta\boldsymbol{x}$ supplies a direction with a high rate of change of the coarse model responses and is a characteristic of the minimum $\boldsymbol{x}_{os,1}^{e}$

proper scaling is applied to the obtained perturbation $\Delta\boldsymbol{x}$ to satisfy a certain trust region condition

a scheme was developed to select one of the eigenvectors of the matrix $\boldsymbol{H}_1$ on a sequential basis

if all the perturbations with certain trust region size are exhausted the size of the trust region is increased

An Alternative Perturbation for the Locally Unique Case without Using Second Order Derivatives

a perturbation of $\Delta\boldsymbol{x}$ results in a perturbation of the coarse model responses at the two minima by

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

and

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

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

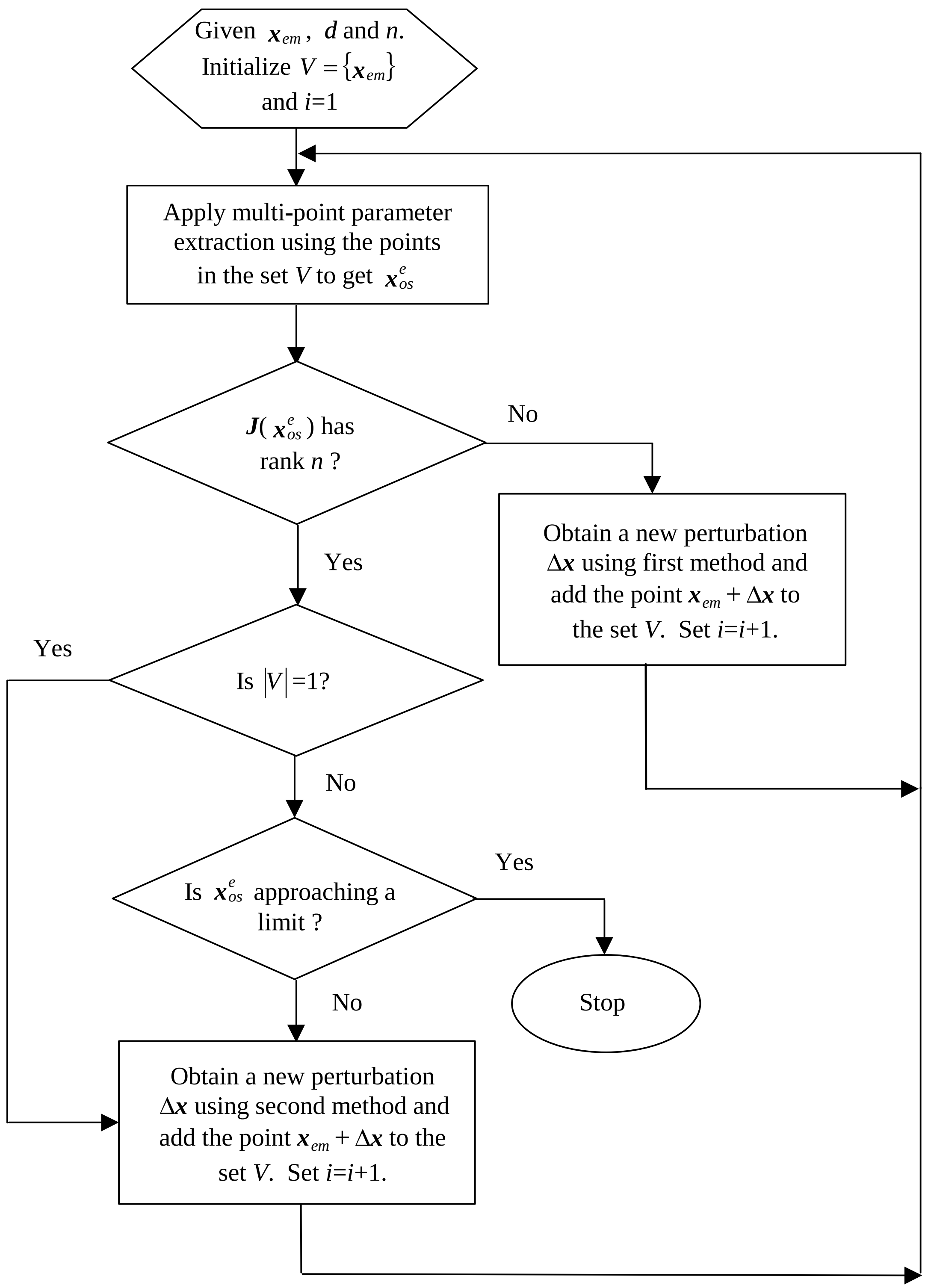
$$L(\Delta\boldsymbol{x},\ l) = \Delta\boldsymbol{x}^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os,1}^{e})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os,1}^{e})\Delta\boldsymbol{x} - \Delta\boldsymbol{x}^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os,2}^{e})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os,2}^{e})\Delta\boldsymbol{x}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

it can be shown that the perturbation $\Delta\boldsymbol{x}$ is an eigenvector for the matrix $\boldsymbol{J}_{os}(\boldsymbol{x}_{os,1}^{e})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os,1}^{e}) - \boldsymbol{J}_{os}(\boldsymbol{x}_{os,2}^{e})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os,2}^{e})$

the perturbation is then scaled to satisfy the length condition

again we make the assumption that $\boldsymbol{J}_{os}(\boldsymbol{x}_{os,2}^{e})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os,2}^{e}) = \boldsymbol{I}$ because of the lack of information about other minima

The Algorithm Flowchart



Example 1: the Rosenbrock Function

the coarse model is

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

the fine model is

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

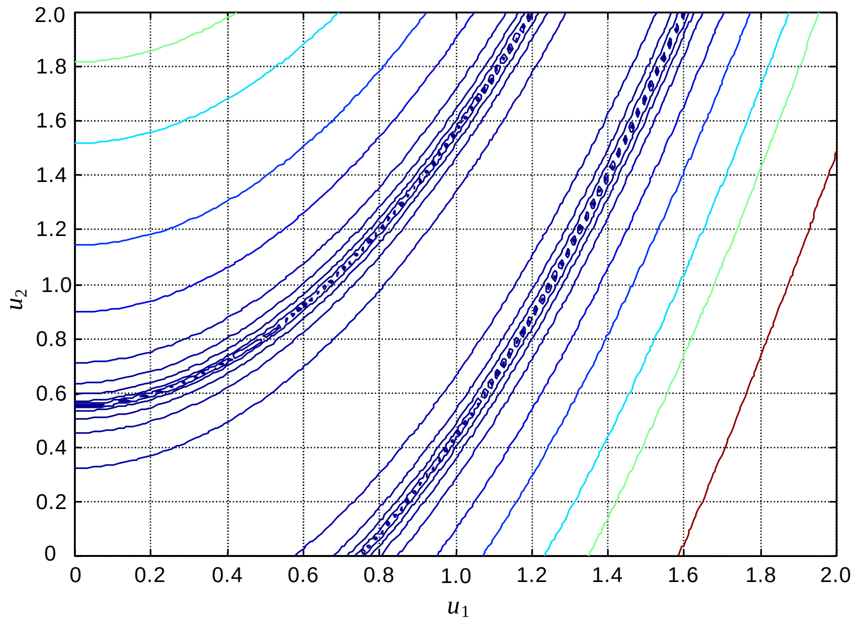
it is required to extract the vector of coarse model parameters at the point $[1.0 \quad 1.0]^T$

three fine model points were needed for the algorithm to terminate

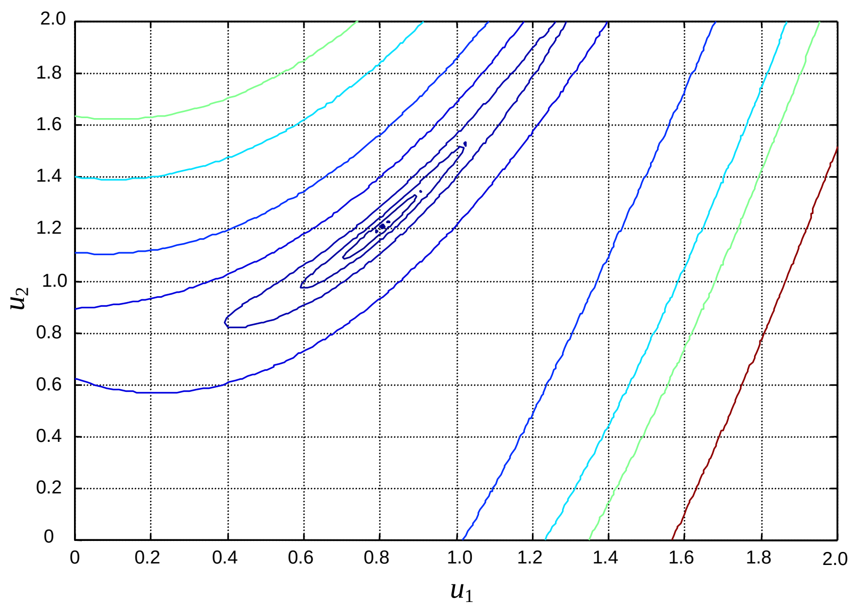
the variation of the extracted parameters with the number of fine model points is shown in the following table

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

The Contours of the L_2 Objective Function for the Rosenbrock Function

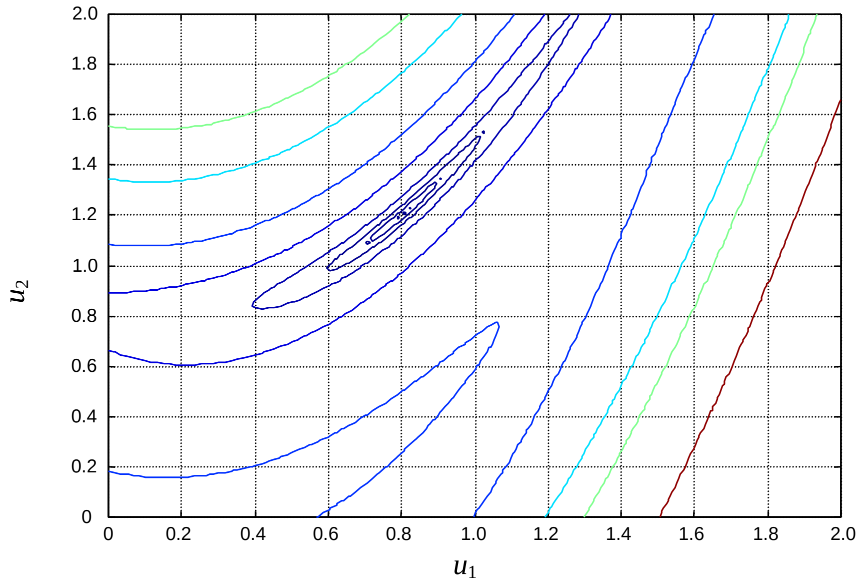


(a) single-point extraction



(b) two-point extraction

The Contours of the L_2 Objective Function for the Rosenbrock Function



(c) three-point extraction

Example 2: a Quadratic Function

the coarse model for this problem is given by

$$R_{os} = x_1^2 + x_2^2$$

the fine model for this problem is given by

$$R_{em} = (0.9x_1 + 0.1x_2)^2 + (0.1x_1 + 0.9x_2)^2$$

it is required in this problem to extract the coarse model parameters corresponding to the fine model parameters $\boldsymbol{x}_{em}=[2.0 \quad 1.0]^T$

four fine model points were needed to ensure the uniqueness of the extracted parameters

The Variation of the Extracted Parameters with the Number of Fine Model Points for the Quadratic Function

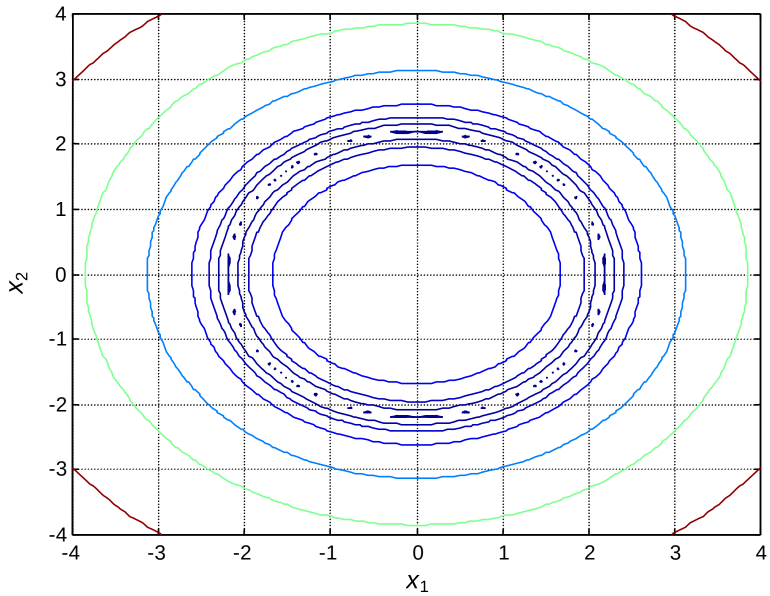
the following table shows the variation of the extracted coarse model parameters with the number of fine model points used in the multi-point parameter extraction

Number of Points	$x_{os,1}$	$x_{os,2}$
1	1.95724	0.99458
2	2.10283	0.63094
3	1.92787	1.05337
4	1.89571	1.10868

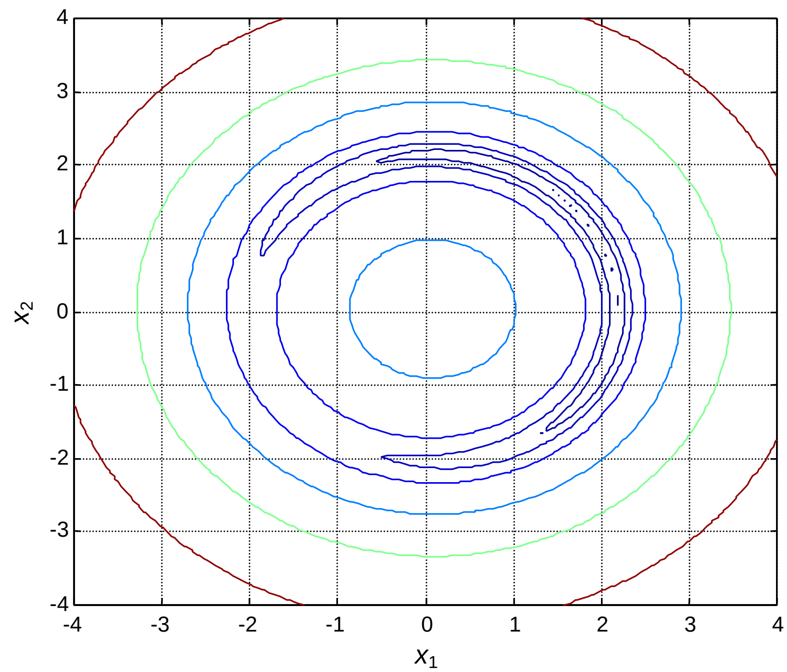
the exact solution for the parameter extraction problem is

$$\boldsymbol{x}_{os} = [1.9 \quad 1.1]^T$$

The Contours of the L_2 Objective Function for the Quadratic Function

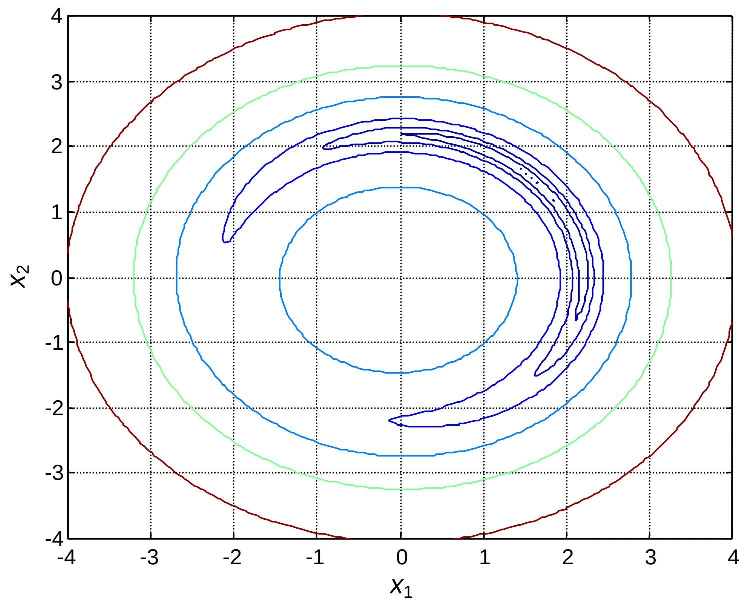


(a) single-point extraction

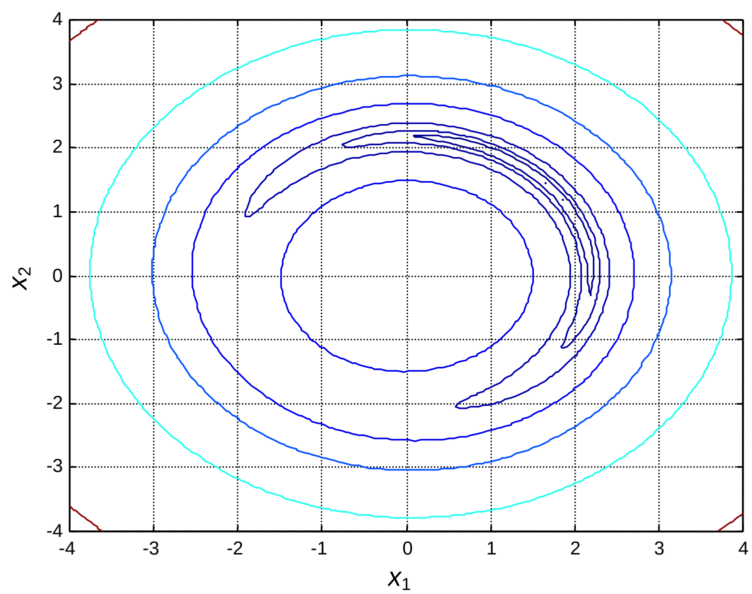


(b) two-point extraction

The Contours of the L_2 Objective Function for the Quadratic Function



(c) three-point extraction



(d) four-point extraction

Example 3: the 10:1 Impedance Transformer

the parameters for this problem are the characteristic impedances of the two transmission lines

the lengths of both lines are kept fixed at their optimal values (quarter wave length)

the coarse model is an ideal 10:1 impedance transformer

the fine model scales each of the two impedances by 1.6

the responses of both models at 11 different frequencies in the frequency band $0.5 \text{ GHz} \leq f \leq 1.5 \text{ GHz}$ were used to match the two models

it is required to extract the coarse model parameters corresponding to the fine model point $[2.2628 \quad 4.5259]^T$

three fine model points were needed to improve the uniqueness of the problem

The Matched Responses for the 10:1 Transformer (Single-Point Extraction)

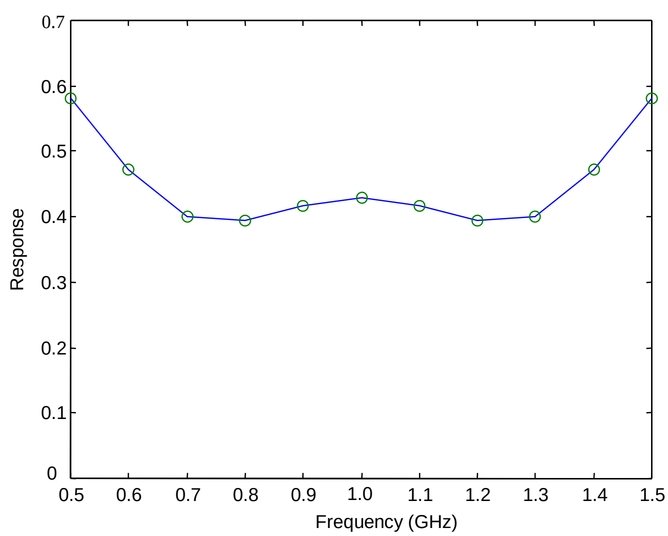
the set of fine model points utilized in parameter extraction is

$$V = \left\{ \begin{bmatrix} 2.26277 \\ 4.52592 \end{bmatrix} \right\}$$

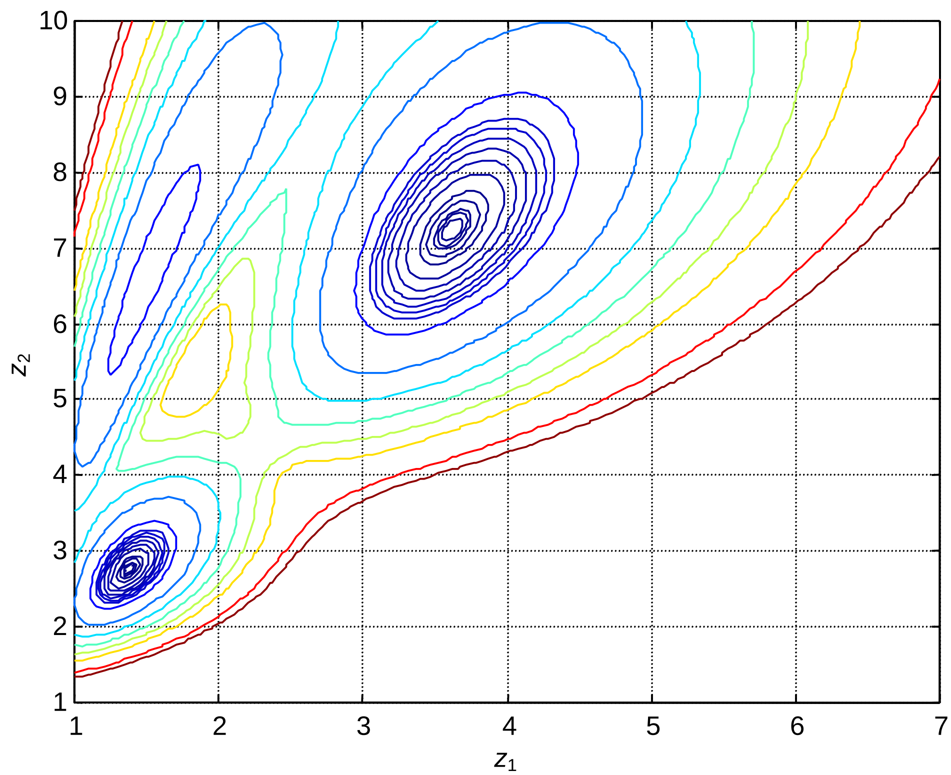
the extracted coarse model parameters are

$$\boldsymbol{x}_{os}^{e} = \begin{bmatrix} 3.62043 \\ 7.24147 \end{bmatrix}$$

the matched responses are shown in the following figure



The Contours of the L_2 Objective Function for the 10:1 Transformer (Single-Point Extraction)



The Matched Responses for the 10:1 Transformer (Two-Point Extraction)

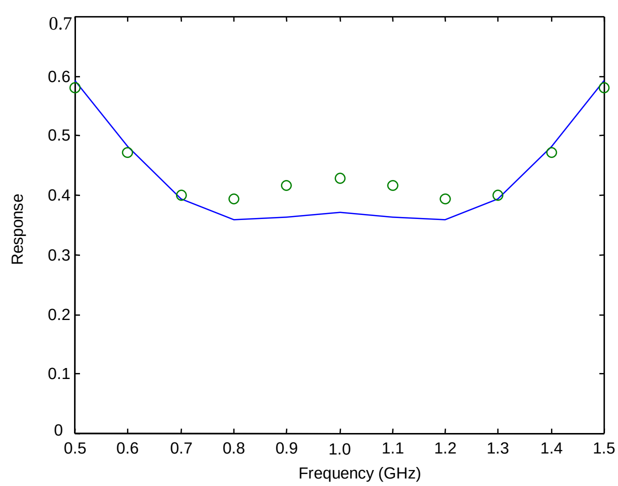
the set of fine model points utilized in parameter extraction is

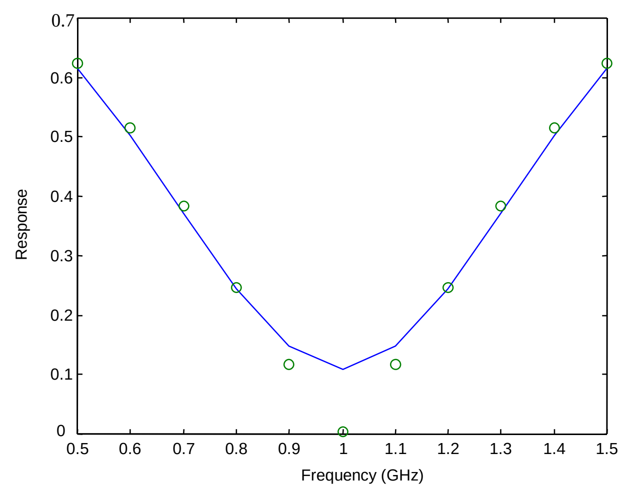
$$V = \left\{ \begin{bmatrix} 2.26277 \\ 4.52592 \end{bmatrix}, \begin{bmatrix} 1.49975 \\ 4.76634 \end{bmatrix} \right\}$$

the extracted coarse model parameters are

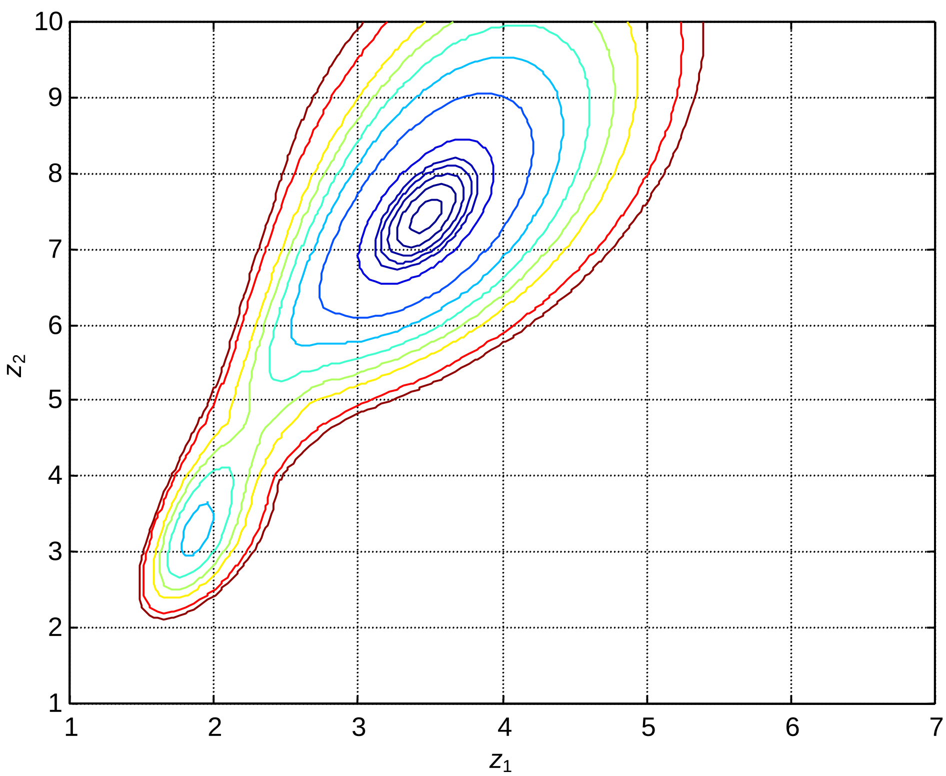
$$\boldsymbol{x}_{os}^{e} = \begin{bmatrix} 3.4716 \\ 7.43214 \end{bmatrix}$$

the matched responses of every fine model point are shown





The Contours of the L_2 Objective Function for the 10:1 Transformer (Two-Point Extraction)



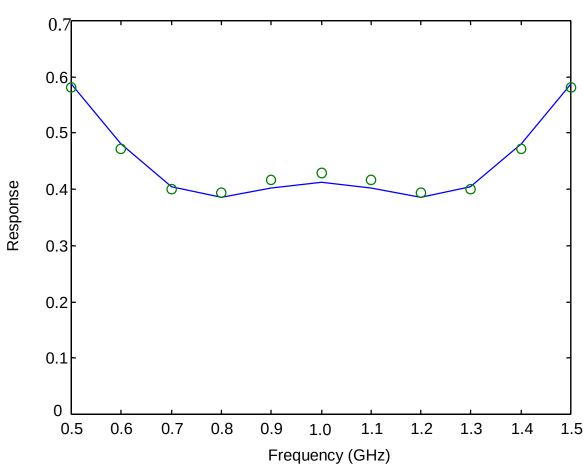
The Matched Response for the 10:1 Transformer (Three-Point Extraction)

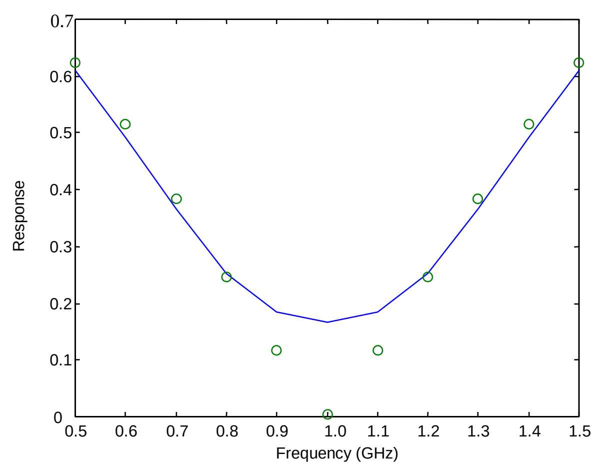
the set of used fine model points is

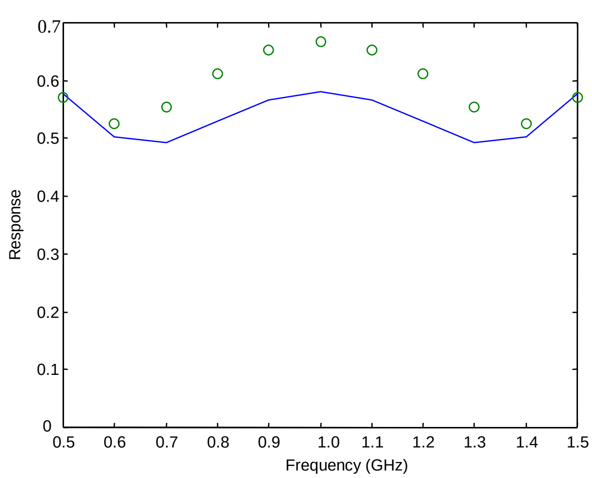
$$V = \left\{ \begin{bmatrix} 2.26277 \\ 4.52592 \end{bmatrix}, \begin{bmatrix} 1.49975 \\ 4.76634 \end{bmatrix}, \begin{bmatrix} 3.02024 \\ 4.26855 \end{bmatrix} \right\}$$

the extracted coarse model parameters are $\boldsymbol{x}_{os}^{e} = \begin{bmatrix} 3.60357 \\ 7.35052 \end{bmatrix}$

the matched responses of every fine model point are shown







The Contours of the L_2 Objective Function for the 10:1 Transformer (Three-Point Extraction)