EM-Based Optimization Exploiting Partial Space Mapping and Exact Sensitivities

J.W. Bandler, A.S. Mohamed, M.H. Bakr, K. Madsen and J. Søndergaard

Simulation Optimization Systems Research Laboratory
McMaster University

Bandler Corporation, www.bandler.com
john@bandler.com

presented at

2002 IEEE MTT-S International Microwave Symposium, Seattle, WA, June 02, 2002

Outline

ASM for microwave circuit design

Gradient Parameter Extraction (GPE)

Partial Space Mapping (PSM)

mapping update

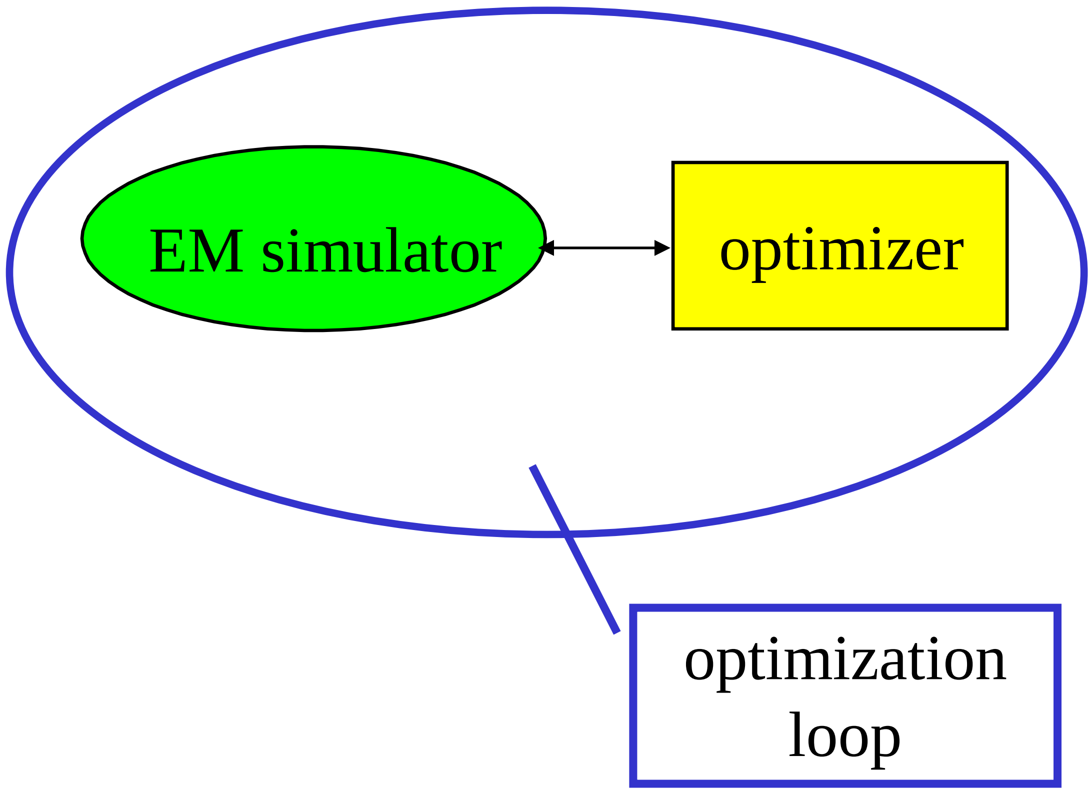
proposed algorithm

examples

conclusions

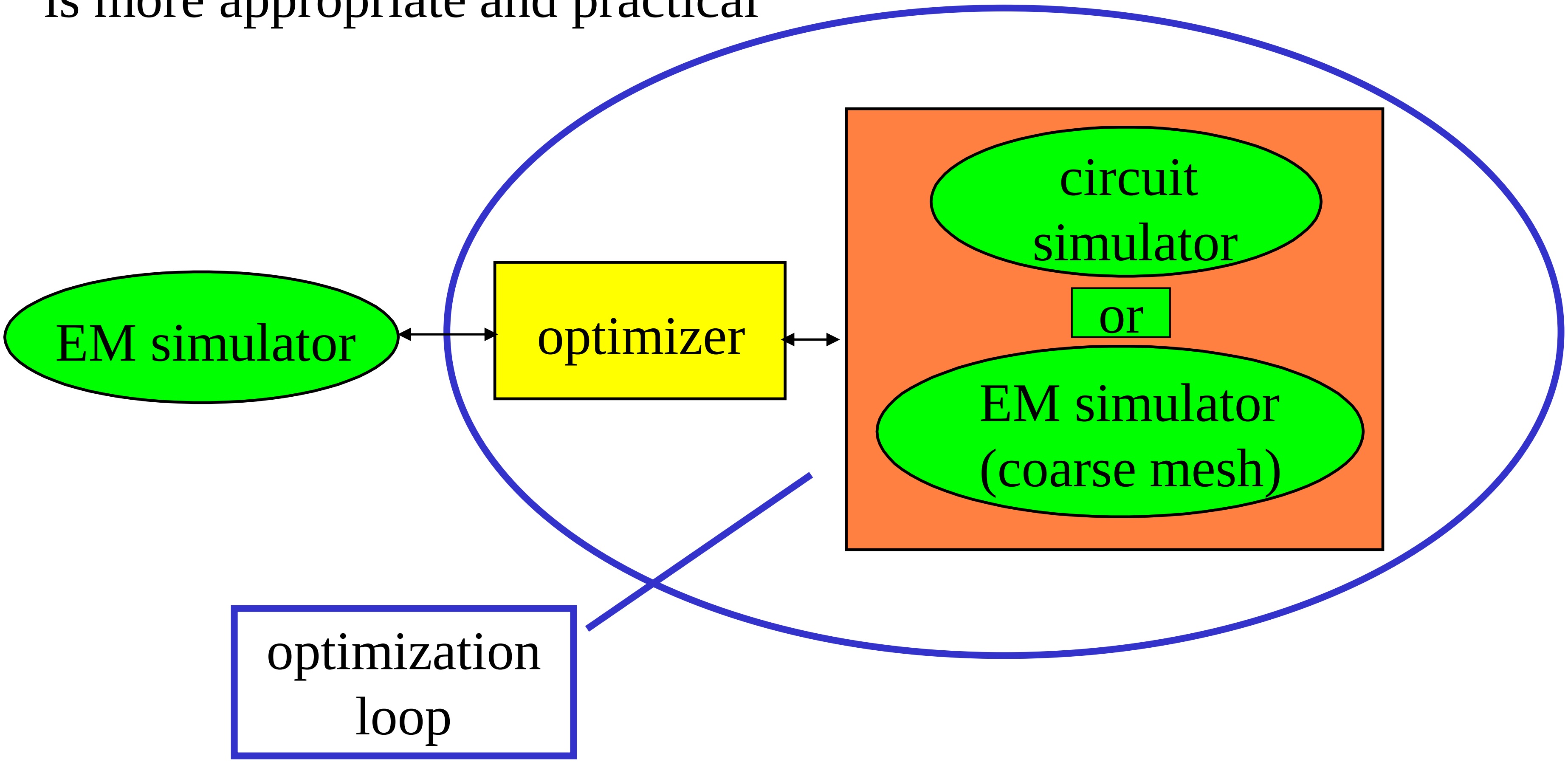
Introduction

using full wave EM simulator (fine model) inside the optimization loop is prohibitive



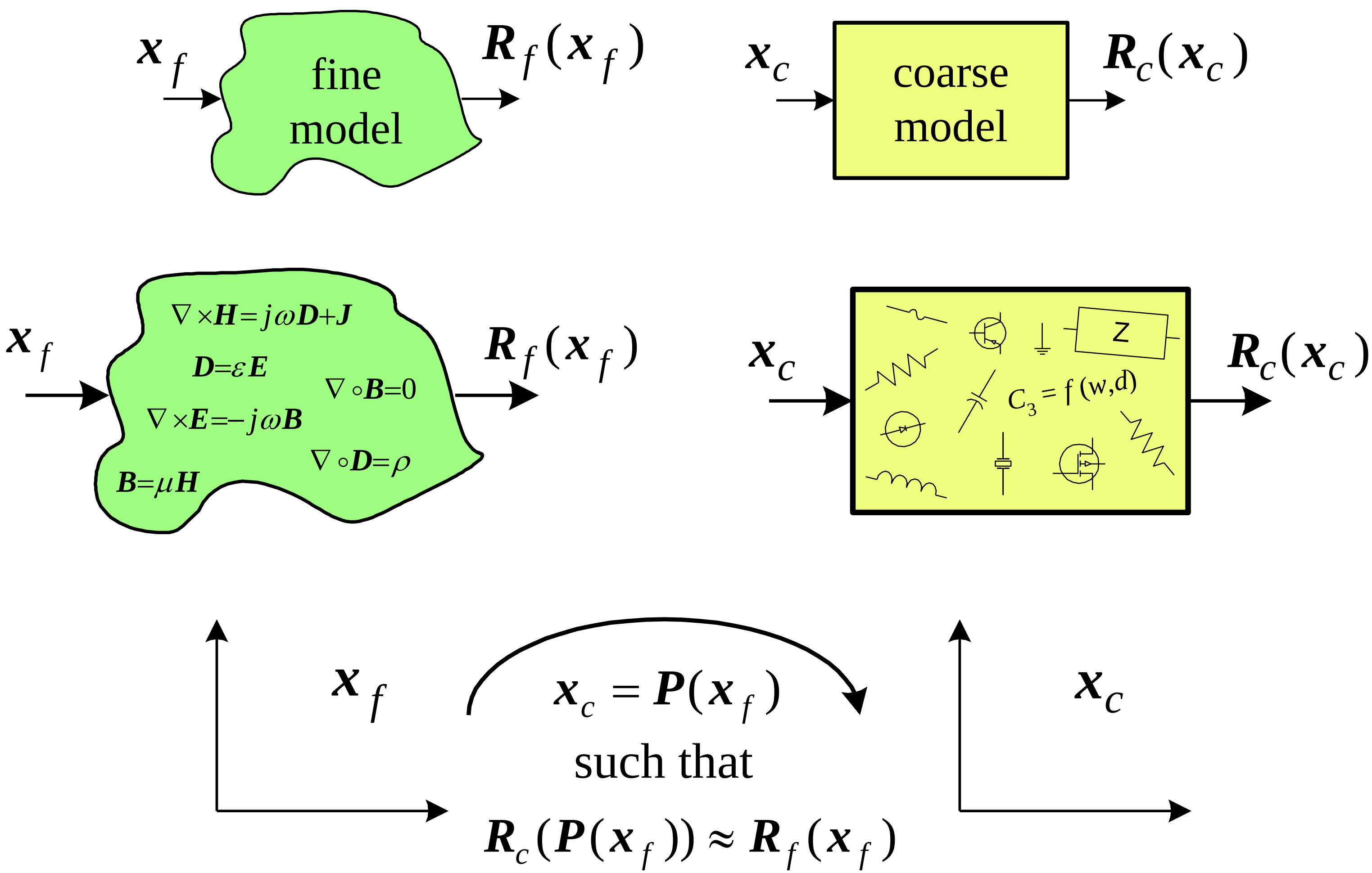
Introduction

using simpler (less accurate) model inside the optimization loop is more appropriate and practical



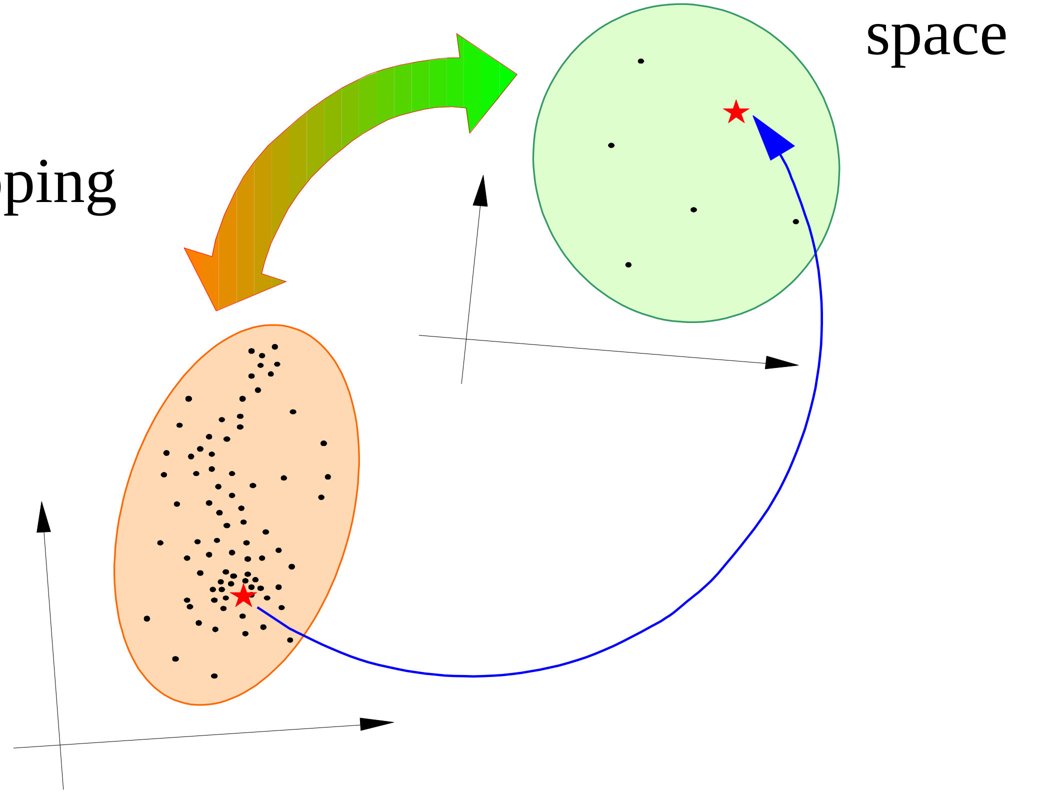
The Space Mapping Concept

(*Bandler et al.*, 1994-)



The Space Mapping Concept

(*Bandler et al., 1994-*)



Jacobian-Space Mapping Relationship

(*Bakr et al., 1999*)

through PE we match the responses

$$\boldsymbol{R}_f(\boldsymbol{x}_f) \approx \boldsymbol{R}_c(\boldsymbol{P}(\boldsymbol{x}_f))$$

by differentiation

$$\left(\frac{\partial \boldsymbol{R}_f^T}{\partial \boldsymbol{x}_f}\right)^T \approx \left(\frac{\partial \boldsymbol{R}_c^T}{\partial \boldsymbol{x}_c}\right)^T \cdot \left(\frac{\partial \boldsymbol{x}_c^T}{\partial \boldsymbol{x}_f}\right)^T$$

Jacobian-Space Mapping Relationship

(*Bakr et al., 1999*)

given coarse model Jacobian $\boldsymbol{J}_c$ and space mapping matrix $\boldsymbol{B}$ we estimate

$$\boldsymbol{J}_f(\boldsymbol{x}_f) \approx \boldsymbol{J}_c(\boldsymbol{x}_c)\boldsymbol{B}$$

given $\boldsymbol{J}_c$ and $\boldsymbol{J}_f$ we estimate (least squares)

$$\boldsymbol{B} \approx (\boldsymbol{J}_c^T \boldsymbol{J}_c)^{-1} \boldsymbol{J}_c^T \boldsymbol{J}_f$$

Gradient Parameter Extraction (GPE)

at the jth iteration

$$\boldsymbol{x}_c^{(j)} = \arg\min_{\boldsymbol{x}_c} \left\| [\boldsymbol{e}_0^T \quad \lambda\boldsymbol{e}_1^T \quad \cdots \quad \lambda\boldsymbol{e}_n^T]^T \right\|, \ \lambda \geq 0$$

where λ is a weighting factor and $\boldsymbol{E} = [\boldsymbol{e}_1\ \boldsymbol{e}_2 \ldots \boldsymbol{e}_n]$

$$\boldsymbol{e}_0 = \boldsymbol{R}_f(\boldsymbol{x}_f^{(j)}) - \boldsymbol{R}_c(\boldsymbol{x}_c)$$

$$\boldsymbol{E} = \boldsymbol{J}_f(\boldsymbol{x}_f^{(j)}) - \boldsymbol{J}_c(\boldsymbol{x}_c)\boldsymbol{B}$$

Partial Space Mapping (PSM)

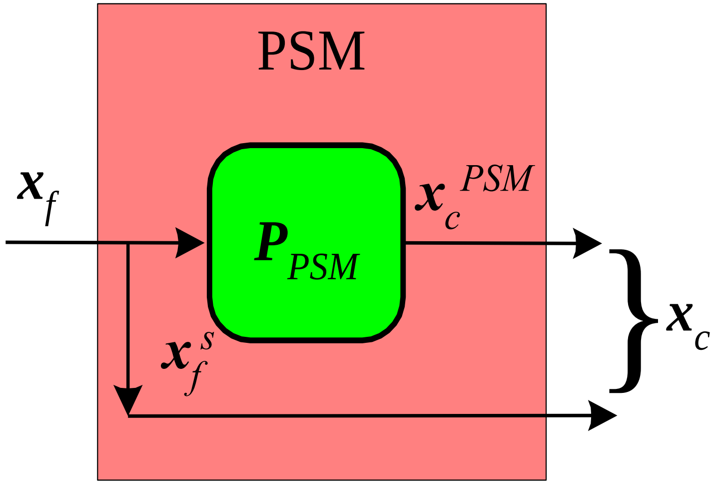
a few coarse parameters may be sufficient

elapsed optimization time is reduced

reflects the idea of postproduction tuning

Partial Space Mapping (PSM)

$$\boldsymbol{x}_c = \begin{bmatrix} \boldsymbol{x}_c^{PSM} \\ \boldsymbol{x}_f^s \end{bmatrix} = \begin{bmatrix} \boldsymbol{P}_{PSM}(\boldsymbol{x}_f) \\ \boldsymbol{x}_f^s \end{bmatrix}$$



the Jacobian-PSM relationship

$$\boldsymbol{J}_f \approx \boldsymbol{J}_c^{PSM} \boldsymbol{B}^{PSM}$$

the minimum norm solution for a quasi-Newton step $\boldsymbol{h}$

$$\boldsymbol{h}_{\text{min norm}}^{(j)} = \boldsymbol{B}^{PSM(j)T} (\boldsymbol{B}^{PSM(j)} \boldsymbol{B}^{PSM(j)T})^{-1} (-\boldsymbol{f}^{(j)})$$

Mapping Update Using Exact Derivatives

$$\boldsymbol{B}^{PSM(j)} = (\boldsymbol{J}_c^{PSM(j)T}\boldsymbol{J}_c^{PSM(j)})^{-1}\boldsymbol{J}_c^{PSM(j)T}\boldsymbol{J}_f^{(j)}$$

Mapping Update Using Hybrid Approach

finite difference initialization used

$$\boldsymbol{B}^{PSM(0)} = (\boldsymbol{J}_c^{PSM(0)T}\boldsymbol{J}_c^{PSM(0)})^{-1}\boldsymbol{J}_c^{PSM(0)T}\boldsymbol{J}_f^{(0)}$$

then update using Broyden formula

Mapping Update By Constraining *B*

(*Bakr et al., 2000*)

$$\boldsymbol{B} = (\boldsymbol{J}_c^T\boldsymbol{J}_c + \eta^2\boldsymbol{I})^{-1}(\boldsymbol{J}_c^T\boldsymbol{J}_f + \eta^2\boldsymbol{I})$$

Proposed PSM/GPE Algorithm

Step 1 set $j = 1$, $\boldsymbol{B} = \boldsymbol{I}$ for the PE process

Step 2 obtain the optimal coarse model design $\boldsymbol{x}_c^*$

Step 3 set $\boldsymbol{x}_f^{(1)} = \boldsymbol{x}_c^*$

Step 4 if derivatives exist execute GPE
otherwise, execute the traditional PE with $\lambda = 0$

Step 5 initialize the mapping matrix $\boldsymbol{B}^{PSM}$

Step 6 stop if

$$\left\| \boldsymbol{f}^{(j)} \right\| < \varepsilon_1 \text{ or } \left\| \boldsymbol{R}_f^{(j)} - \boldsymbol{R}_c^* \right\| < \varepsilon_2$$

Proposed PSM/GPE Algorithm (continued)

Step 7 evaluate $\boldsymbol{h}^{(j)}$ using

$$\boldsymbol{h}_{\text{min norm}}^{(j)} = \boldsymbol{B}^{PSM(j)T}(\boldsymbol{B}^{PSM(j)}\boldsymbol{B}^{PSM(j)T})^{-1}(-\boldsymbol{f}^{(j)})$$

Step 8 find the next $\boldsymbol{x}_f^{(j+1)}$

Step 9 perform GPE or PE as in Step 4

Step 10 if derivatives exist obtain $\boldsymbol{B}^{PSM(j)}$ using

$$\boldsymbol{B}^{PSM(j)} = (\boldsymbol{J}_c^{PSM(j)T}\boldsymbol{J}_c^{PSM(j)})^{-1}\boldsymbol{J}_c^{PSM(j)T}\boldsymbol{J}_f^{(j)}$$

otherwise update $\boldsymbol{B}^{PSM(j)}$ using a Broyden formula

Proposed PSM/GPE Algorithm (continued)

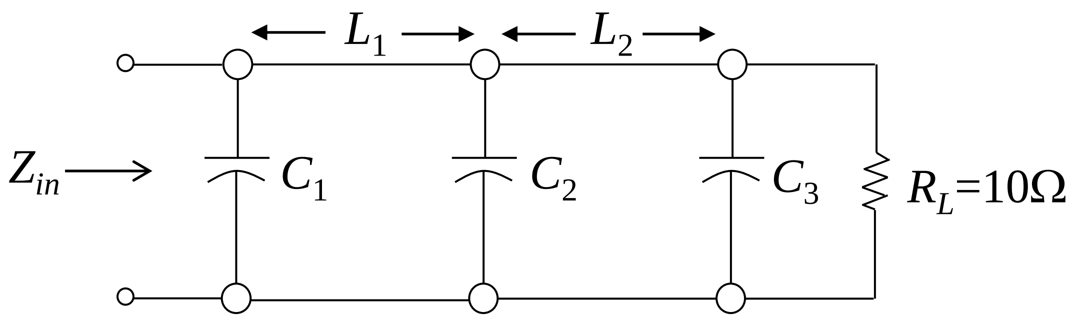
Step 11 set $j = j+1$ and go to *Step* 6

the result is the solution $\bar{\boldsymbol{x}}_f$ and mapping matrix $\boldsymbol{B}^{PSM}$

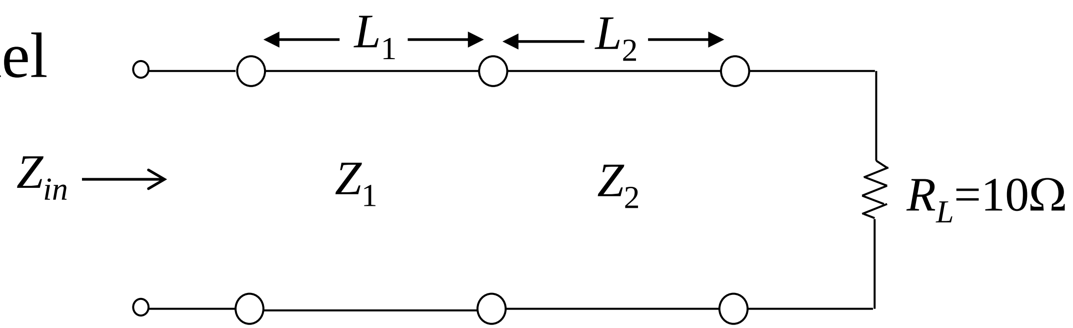
A Two-section 10:1 Capacitively-loaded Impedance Transformer

(*Bakr et al. 2000*)

fine model



coarse model



specifications

$$|S_{11}| \leq 0.50 \text{ for } 0.5 \text{ GHz} \leq \omega \leq 1.5 \text{ GHz}$$

Optimization of the Impedance Transformer

consider $\boldsymbol{x}_c^{PSM} = [L_1 \; L_2]^T$ while $\boldsymbol{x}_f^s = [Z_1 \; Z_2]^T$ kept fixed at the optimal solution during the PE

exact adjoint sensitivity analysis gives $\boldsymbol{J}_\text{c}$ and $\boldsymbol{J}_f$

exact derivatives to update mapping

the final mapping is

$$\boldsymbol{B}^{PSM} = \begin{bmatrix} 1.044 & -0.017 & 0.009 & 0.002 \\ -0.011 & 1.079 & -0.011 & 0.006 \end{bmatrix}$$

Optimization of the Impedance Transformer (continued)

initial and final designs

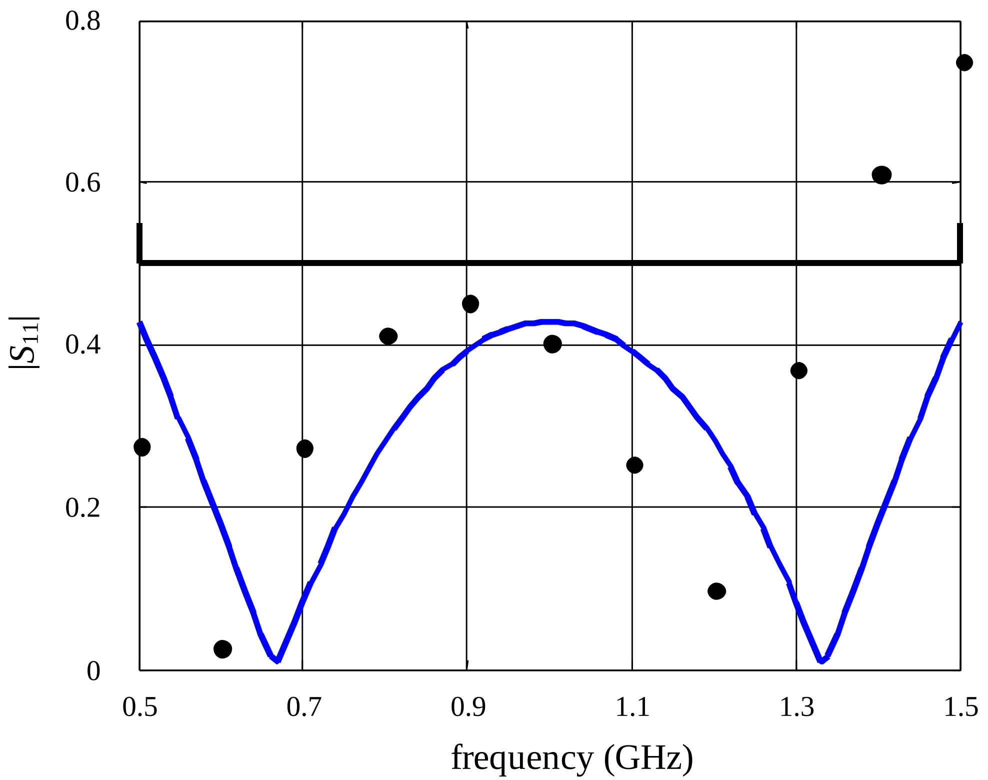
parameter	$x_f^{(0)}$	$x_f^{(1)}$
L_1	1.0	0.9111
L_2	1.0	0.8082
Z_1	2.2362	2.2371
Z_2	4.4723	4.4708

L_1 and L_2 are normalized lengths

Z_1 and Z_2 are in ohm

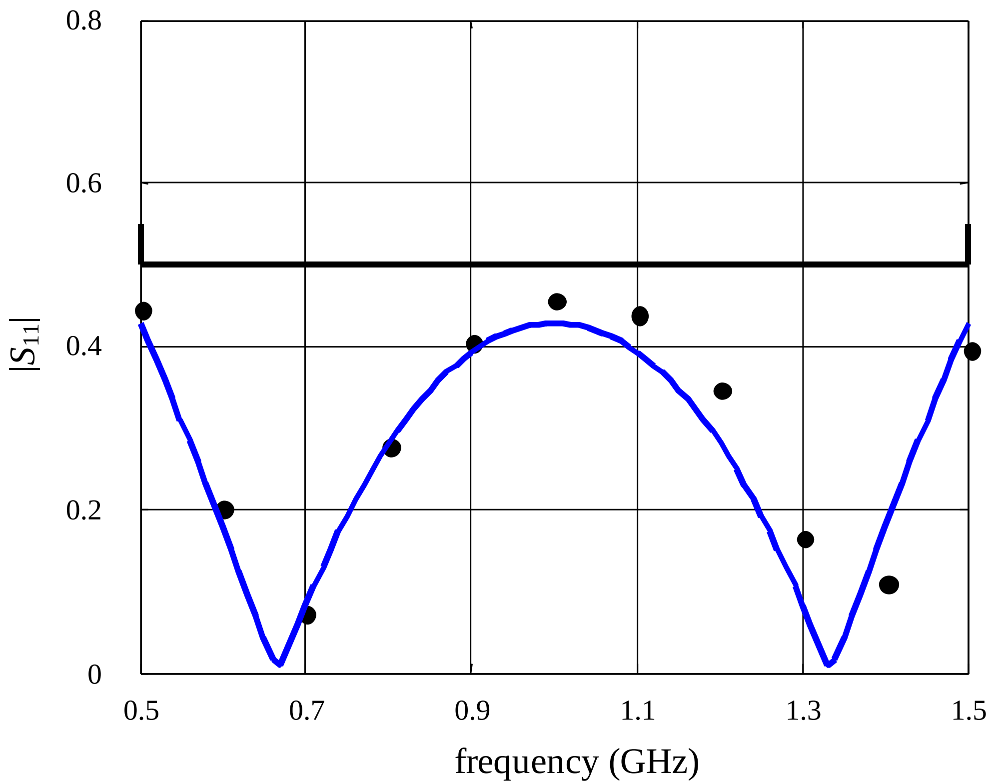
Optimization of the Impedance Transformer (continued)

initial coarse model (target) response(–)
initial fine model response (•)



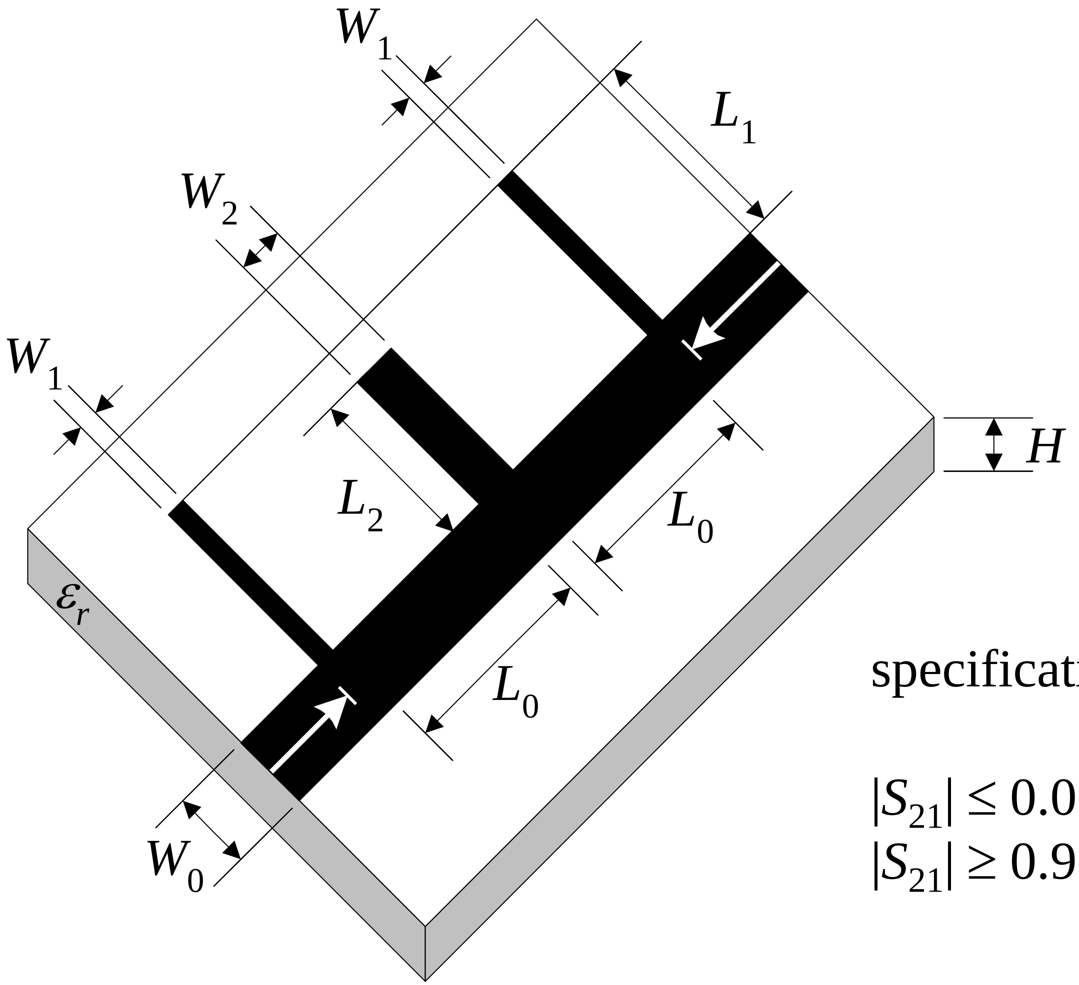
Optimization of the Impedance Transformer (continued)

initial coarse model (target) response (–)
final fine model response (•)



Bandstop Microstrip Filter with Quarter-Wave Open Stubs

(*Bakr et al., 2000*)



H = 25 mil, W_0 = 25 mil,
ε_r = 9.4 (alumina)

the design parameters are
$\boldsymbol{x}_f = [W_1\ W_2\ L_0\ L_1\ L_2]^T$

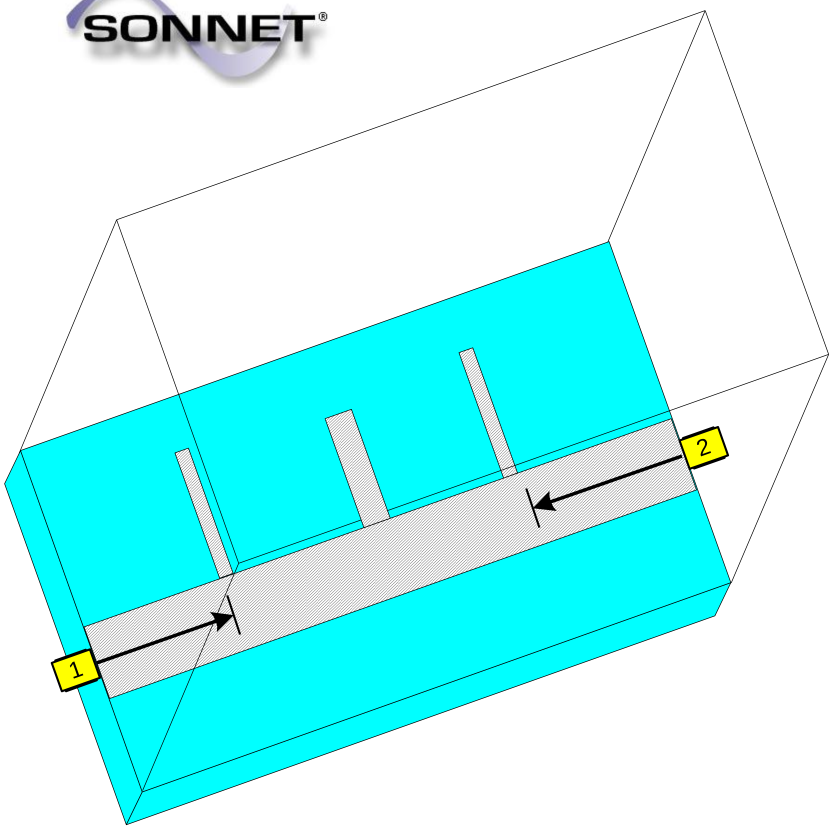
specifications

$|S_{21}| \leq 0.05$ for 9.3 GHz $\leq \omega \leq$ 10.7 GHz
$|S_{21}| \geq 0.9$ for $\omega \leq$ 8 GHz and $\omega \geq$ 12 GHz

Bandstop Microstrip Filter: Fine and Coarse Models

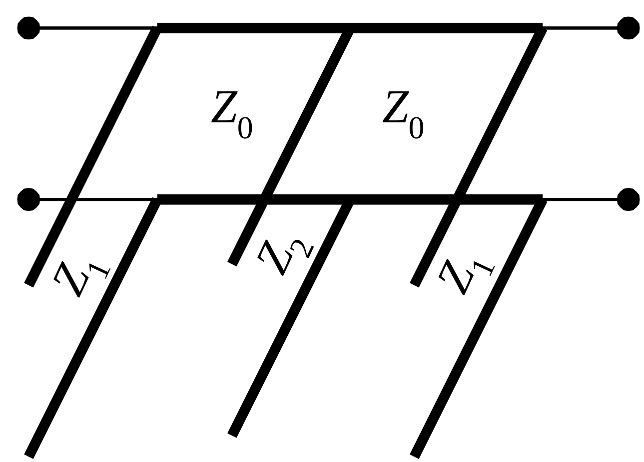
fine model:

Sonnet's ***em***™ with high resolution grid



coarse model:

OSA90/hope™ ideal transmission line sections and empirical formulas



Optimization of the Bandstop Filter

during PE we consider $\boldsymbol{x}_c^{PSM} = [L_1\ L_2]^T$ while $\boldsymbol{x}_f^s = [W_1\ W_2\ L_0]^T$ are held fixed

finite differences estimate the fine and coarse Jacobians

use hybrid approach to update mapping

the final mapping is

$$\boldsymbol{B}^{PSM} = \begin{bmatrix} 0.570 & 0.168 & 0.209 & 0.911 & 0.214 \\ -0.029 & 0.154 & 0.126 & -0.024 & 0.470 \end{bmatrix}$$

Optimization of the Bandstop Filter (continued)

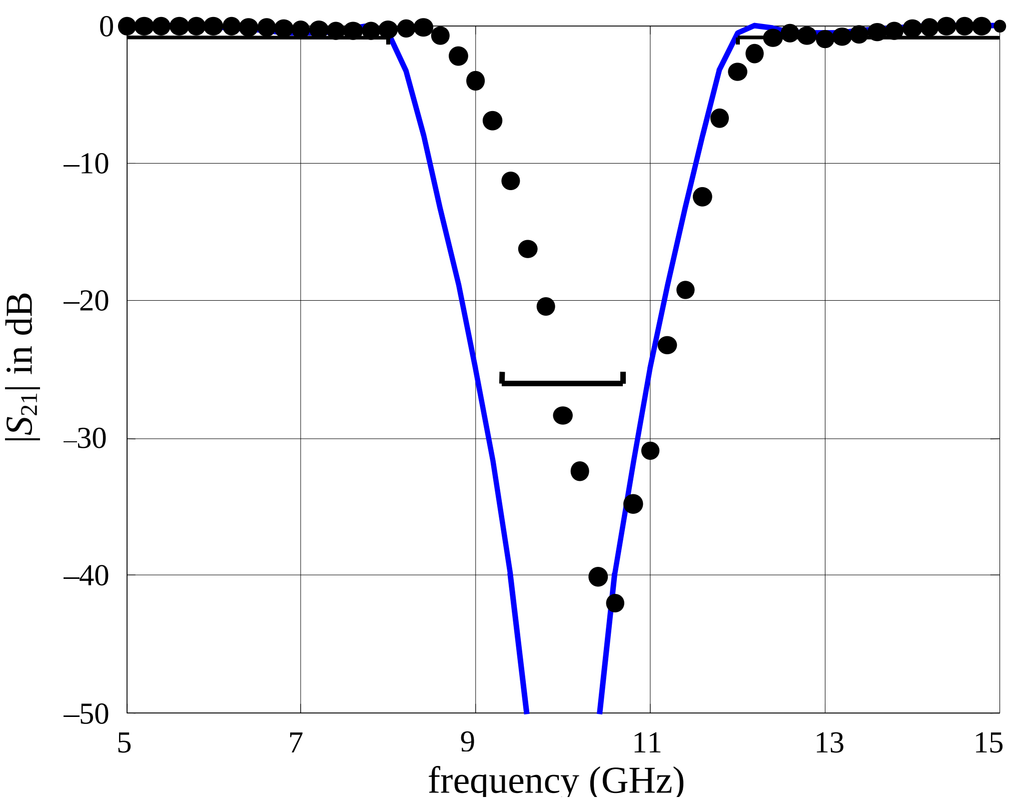
initial and final designs

parameter	$\boldsymbol{x}_f^{(0)}$	$\boldsymbol{x}_f^{(5)}$
W_1	4.560	7.329
W_2	9.351	10.672
L_0	107.80	109.24
L_1	111.03	115.53
L_2	108.75	111.28

all values are in mils

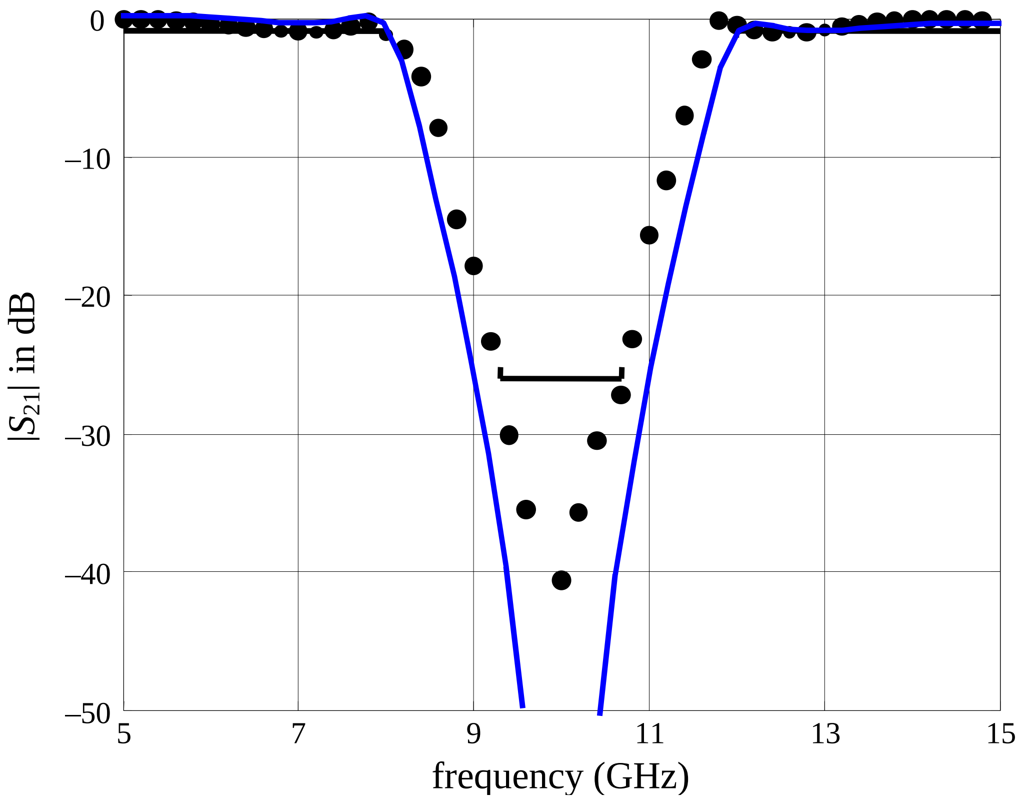
Optimization of the Bandstop Filter (continued)

initial coarse model OSA90™ response (–)
initial fine response ***em***™ (•)



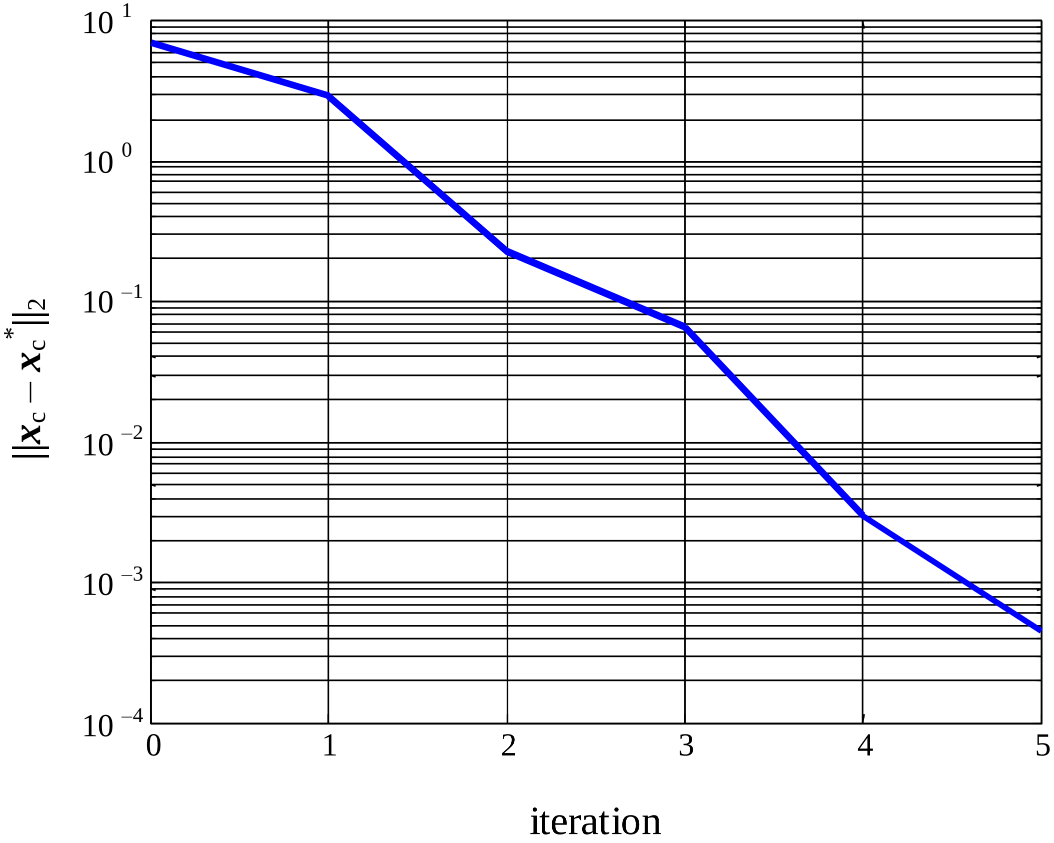
Optimization of the Bandstop Filter (continued)

initial coarse model OSA90™ response (–)
final fine response ***em***™ (•)



Optimization of the Bandstop Filter (continued)

$||\boldsymbol{x}_c - \boldsymbol{x}_c^*||_2$ versus iteration for the bandstop microstrip filter



Original Rosenbrock Function (Coarse Model)

$$R_c(\boldsymbol{x}_c) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

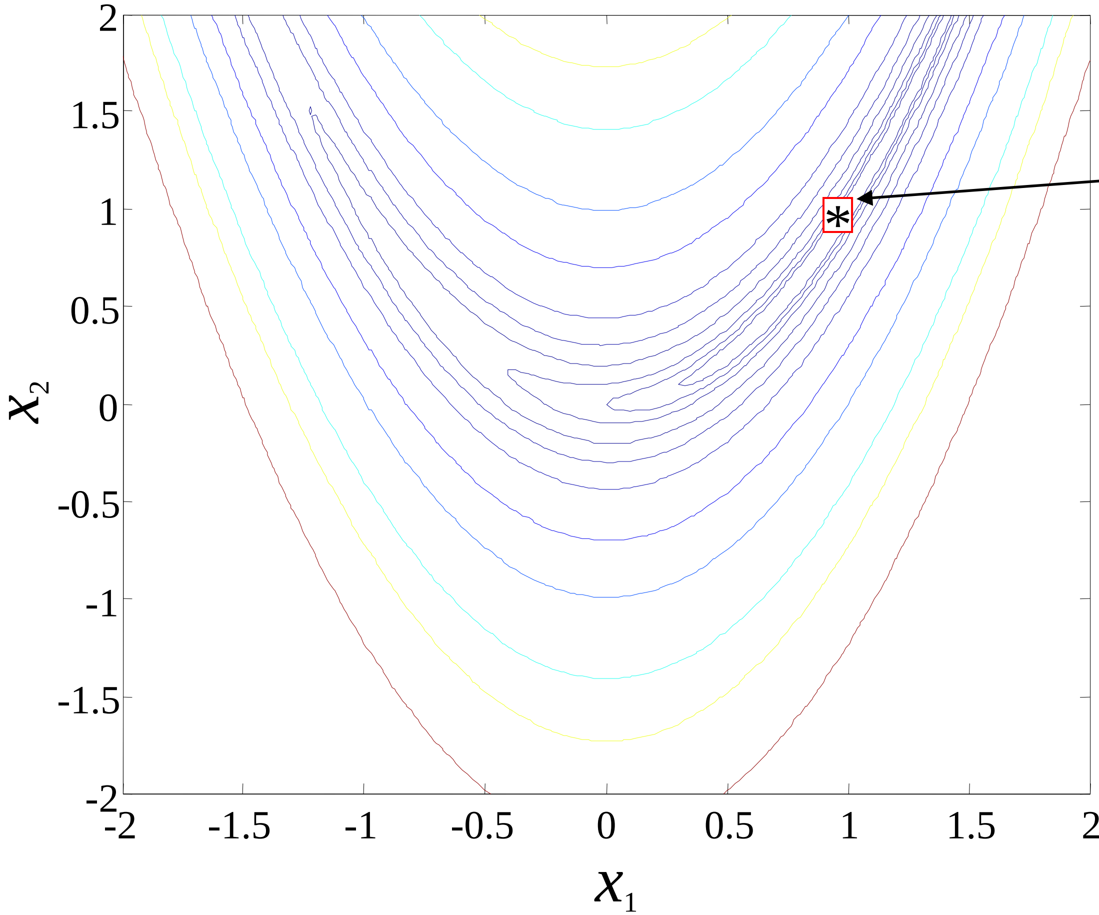
$$\text{where } \boldsymbol{x}_c = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ and } \boldsymbol{x}_c^* = \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix}$$

Shifted Rosenbrock Function (Fine Model)

$$R_f(\boldsymbol{x}_f) = 100\left((x_2 + \alpha_2) - (x_1 + \alpha_1)^2\right)^2 + \left(1 - (x_1 + \alpha_1)\right)^2$$

$$\text{where } \boldsymbol{x}_f = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} -0.2 \\ 0.2 \end{bmatrix} \text{ hence } \boldsymbol{x}_f^* = \begin{bmatrix} 1.2 \\ 0.8 \end{bmatrix}$$

Original Rosenbrock Function (Coarse Model Contour Plot)



$$\boldsymbol{x}_c^* = \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix}$$

$$R_c(\boldsymbol{x}_c^*) = 0$$

Shifted Rosenbrock Function Results

iteration	$\boldsymbol{x}_c^{(j)}$	$\boldsymbol{f}^{(j)}$	$\boldsymbol{B}^{(j)}$	$\boldsymbol{h}^{(j)}$	$\boldsymbol{x}_f^{(j)}$	R_f
0	$\begin{bmatrix}1.0\\1.0\end{bmatrix}$	---	---	---	$\begin{bmatrix}1.0\\1.0\end{bmatrix}$	31.4
1	$\begin{bmatrix}0.8\\1.2\end{bmatrix}$	$\begin{bmatrix}-0.2\\0.2\end{bmatrix}$	$\begin{bmatrix}1.0 & 0.0\\0.0 & 1.0\end{bmatrix}$	$\begin{bmatrix}0.2\\-0.2\end{bmatrix}$	$\begin{bmatrix}1.2\\0.8\end{bmatrix}$	0
	$\begin{bmatrix}1.0\\1.0\end{bmatrix}$	$\begin{bmatrix}0\\0\end{bmatrix}$				

Transformed Rosenbrock Function (Fine Model)

linear transformation of the original Rosenbrock function

$$R_f(\boldsymbol{x}_f) = 100(u_2 - u_1^2)^2 + (1 - u_1)^2$$

$$\text{where } \boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 1.1 & -0.2 \\ 0.2 & 0.9 \end{bmatrix} \boldsymbol{x}_f + \begin{bmatrix} -0.3 \\ 0.3 \end{bmatrix}$$

$$\boldsymbol{x}_f^* = \begin{bmatrix} 1.2718447 \\ 0.4951456 \end{bmatrix}$$

Transformed Rosenbrock Function Final Results

iteration	$x_c^{(j)}$	$f^{(j)}$	$B^{(j)}$	$h^{(j)}$	$x_f^{(j)}$	R_f
0	$\begin{bmatrix}1.0\\1.0\end{bmatrix}$	---	---	---	$\begin{bmatrix}1.0\\1.0\end{bmatrix}$	108.3
1	$\begin{bmatrix}0.526\\1.384\end{bmatrix}$	$\begin{bmatrix}-0.474\\0.384\end{bmatrix}$	$\begin{bmatrix}1.01 & -0.05\\0.01 & 1.01\end{bmatrix}$	$\begin{bmatrix}0.447\\-0.385\end{bmatrix}$	$\begin{bmatrix}1.447\\0.615\end{bmatrix}$	5.119
2	$\begin{bmatrix}1.185\\1.178\end{bmatrix}$	$\begin{bmatrix}0.185\\0.178\end{bmatrix}$	$\begin{bmatrix}0.96 & -0.12\\-0.096 & 1.06\end{bmatrix}$	$\begin{bmatrix}-0.218\\-0.187\end{bmatrix}$	$\begin{bmatrix}1.23\\0.427\end{bmatrix}$	4.4E–3
3	$\begin{bmatrix}0.967\\0.929\end{bmatrix}$	$\begin{bmatrix}-0.033\\-0.071\end{bmatrix}$	$\begin{bmatrix}1.09 & -0.19\\0.168 & 0.92\end{bmatrix}$	$\begin{bmatrix}0.0429\\0.0697\end{bmatrix}$	$\begin{bmatrix}1.273\\0.4970\end{bmatrix}$	1.8E–6
4	$\begin{bmatrix}1.001\\1.001\end{bmatrix}$	$\begin{bmatrix}0.001\\0.001\end{bmatrix}$	$\begin{bmatrix}1.10001 & -0.1999\\0.1999 & 0.9001\end{bmatrix}$	$\begin{bmatrix}-0.001\\-0.002\end{bmatrix}$	$\begin{bmatrix}1.2719\\0.4952\end{bmatrix}$	5E–10

Transformed Rosenbrock Function Final Results (continued)

iteration	$\boldsymbol{x}_c^{(j)}$	$\boldsymbol{f}^{(j)}$	$\boldsymbol{B}^{(j)}$	$\boldsymbol{h}^{(j)}$	$\boldsymbol{x}_f^{(j)}$	R_f
5	$\begin{bmatrix}1.00002\\1.00004\end{bmatrix}$	$1E-4*\begin{bmatrix}0.2\\0.4\end{bmatrix}$	$\begin{bmatrix}1.1 & -0.2\\0.2 & 0.9\end{bmatrix}$	$1E-4*\begin{bmatrix}0.3\\0.5\end{bmatrix}$	$\begin{bmatrix}1.2718\\0.4951\end{bmatrix}$	3E–17
6	$\begin{bmatrix}1.0\\1.0\end{bmatrix}$	$1E-8*\begin{bmatrix}0.1\\0.3\end{bmatrix}$	$\begin{bmatrix}1.1 & -0.2\\0.2 & 0.9\end{bmatrix}$	$1E-8*\begin{bmatrix}0.2\\0.3\end{bmatrix}$	$\boldsymbol{x}_f^*$	9E–29

$$\boldsymbol{x}_f^* = \begin{bmatrix}1.27184466\\0.49514563\end{bmatrix}$$

Conclusions

new Aggressive Space Mapping techniques

Gradient Parameter Extraction (GPE) exploiting available Jacobian (exact or approximate)

Partial Space Mapping (PSM) with reduced set of optimization variables in the PE phase

consideration of mapping updates

available Jacobians can be used to build the mapping

Mapping Update By Constraining *B*

(Bakr et al., 2000)

to constrain the mapping matrix to be close to ***I***

$$\boldsymbol{B} = \arg\min_{\boldsymbol{B}} \left\| [\boldsymbol{e}_1^T \cdots \boldsymbol{e}_n^T \; \eta\Delta\boldsymbol{b}_1^T \cdots \eta\Delta\boldsymbol{b}_n^T]^T \right\|_2^2$$

where η is a weighting factor, $\boldsymbol{e}_i$ and $\Delta\boldsymbol{b}_i$ are the ith columns of $\boldsymbol{E}$ and $\Delta\boldsymbol{B}$

$$\boldsymbol{E} = \boldsymbol{J}_f - \boldsymbol{J}_c\boldsymbol{B}$$

$$\Delta\boldsymbol{B} = \boldsymbol{B} - \boldsymbol{I}$$

analytical solution is

$$\boldsymbol{B} = (\boldsymbol{J}_c^T\boldsymbol{J}_c + \eta^2\boldsymbol{I})^{-1}(\boldsymbol{J}_c^T\boldsymbol{J}_f + \eta^2\boldsymbol{I})$$