Optimization and Engineering, 2, 369–384, 2001

An Introduction to the Space Mapping Technique

MOHAMED H. BAKR, JOHN W. BANDLER
Simulation Optimization Systems Research Laboratory and the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada, L8S 4K1

KAJ MADSEN, JACOB SØNDERGAARD
Informatics and Mathematical Modelling, Technical University of Denmark, DK 2800 Lyngby, Denmark

Received March 30, 2001; Revised January 16, 2002

Abstract. The space mapping technique is intended for optimization of engineering models which involve very expensive function evaluations. It is assumed that two different models of the same physical system are available: Besides the expensive model of primary interest (denoted the fine model), access to a cheaper (coarse) model is assumed which may be less accurate.

The main idea of the space mapping technique is to use the coarse model to gain information about the fine model, and to apply this in the search for an optimal solution of the latter. Thus the technique iteratively establishes a mapping between the parameters of the two models which relate similar model responses. Having this mapping, most of the model evaluations can be directed to the fast coarse model.

In many cases this technique quickly provides an approximate optimal solution to the fine model that is sufficiently accurate for engineering purposes. Thus the space mapping technique may be considered a preprocessing technique that perhaps must be succeeded by use of classical optimization techniques. We present an automatic scheme which integrates the space mapping and classical techniques.

Keywords: non-linear optimization, space mapping, surrogate modelling

1. Introduction

When engineers encounter a mathematical problem which they cannot solve, it is common practice to consider another formulation which is solvable and intends to contribute to the original problem solution.

The space mapping technique, which was introduced by Bandler et al. (1994), is based on this principle. It is an optimization technique for engineering design in the following situation: Assume the performance of some physical object depends on a number of parameters. We search for an optimal parameter setting and during the search procedure we need to find model responses corresponding to some intermediate sets of parameters. This may for instance be based on function evaluations requested by a mathematical optimization algorithm. These evaluations are assumed to be so expensive that traditional optimization becomes unrealistic in practice. Even cases where function evaluations involve physical experiments may occur. Therefore, the aim is to make a shortcut using a cheaper, but presumably less accurate, model of the same physical system, in order to gain information about the optimal parameter setting of the original model.

Thus we assume two different models are available:

1. An accurate but expensive model, represented by a residual function $f : \Omega^{(f)} \to \mathbb{R}^m$, which must be minimized as indicated below. Here $\Omega^{(f)} \subseteq \mathbb{R}^n$, and $m \geq n$. This model is denoted the *fine model*. Gradients of f are assumed not to be available.
2. A cheap (i.e., fast) model, represented by a residual function $c : \Omega^{(c)} \to \mathbb{R}^m$, which must be minimizable in the same sense as f. Here $\Omega^{(c)} \subseteq \mathbb{R}^n$, and $m \geq n$. This model is denoted the *coarse model*. Gradients of c are assumed to be available.

In this context a residual function is the difference between a response function, originating from a model, and some predefined specifications. A response function may for instance be model responses at a specific set of sample points $t^{(j)}$, $j = 1, \ldots, m$, hence $f(x)$, $c(z)$ are vector functions with elements $f^{(j)}(x) = \varphi(t^{(j)}; x)$, $c^{(j)}(z) = \sigma(t^{(j)}; z)$ being the difference between the model response and the specification at a given sample point $t^{(j)}$. We wish to find an optimal set of parameters $x^* \in \Omega^{(f)}$ which makes the fine model response meet the specifications as well as possible, hence minimizing the fine model residual function f

$$x^* \in \arg\min_{x \in \Omega^{(f)}} H(f(x)) \tag{1}$$

with respect to some merit function H, e.g., a norm in $\mathbb{R}^m$. Since the fine model is considered too expensive for direct optimization, we want to use the coarse model to gain information about the fine model.

The general idea of how this is achieved can be illustrated by the following simple example:

Consider an archery contest, and assume for simplicity that the archer has a steady hand: he always shoots in exactly that direction he has planned. The goal of course is to hit the bull's-eye y^*, hence y^* represents the given set of specifications. The shooting situation is simulated with a coarse model which hits the spot the archer is pointing at, not taking forces like wind and gravity into account.

We represent the points y in the target plane as vectors in $\mathbb{R}^2$. The coarse objective function is a vector function $c : \Omega^{(c)} \to \mathbb{R}^2$, where $\Omega^{(c)} \subseteq \mathbb{R}^2$ is the set of possible directions from the archer to the target. Let $z \in \Omega^{(c)}$ be a direction pointing to the spot $y^{(c)}$ at the target. Then the objective $c(z)$ is the difference between $y^{(c)}$ and the target, i.e., $c(z) = y^{(c)} - y^*$. The fine model is a representation of the actual shot towards the target, i.e., in this case the fine model represents physical experiments. The fine objective function is a vector function $f : \Omega^{(f)} \to \mathbb{R}^2$, $\Omega^{(f)} \subseteq \mathbb{R}^2$. For a direction $x \in \Omega^{(f)}$ the objective $f(x)$ is the difference between the spot $y^{(f)}$ at the target which is hit and the target, i.e., $f(x) = y^{(f)} - y^*$. We wish to find a direction $x^* \in \Omega^{(f)}$ such that $\|f(x^*)\| = 0$.

At first the archer aims at y^*, i.e., he optimizes the coarse model by finding the direction $z^* \in \Omega^{(c)}$ which points at y^*. This can be formulated as follows,

$$z^* = \arg\min_{z \in \Omega^{(c)}} H(c(z)) \tag{2}$$

for some norm H. In this case $\|c(z^*)\| = 0$.

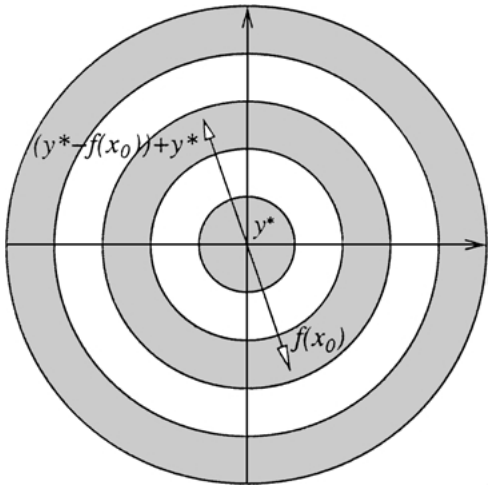

Figure 1. "Calculation" of the first shot $f(x_0)$, $x_0 = z^*$. In the next shot the archer will aim at $(y^* - f(x_0)) + y^*$.

After taking aim the archer fires the shot in the direction chosen, hence we "calculate" $f(x_0)$ for $x_0 = z^*$, as illustrated in figure 1.

Since the coarse model does not take the influence of wind and gravity into account, the arrow may fail to hit y^*, which in mathematical terms means that $x^* \neq z^*$.

After failing a shot any good archer would adjust the sight in order to obtain a better result with the next shot. The natural adjustment would be to "mirror the error". If, for instance, the first shot has hit too low on the right side of y^*, then the next aim should be directly opposite: upwards on the left side of y^*. In our notation the second shot would aim at $(y^* - f(x_0)) + y^*$. Thus, if we let z_0 be the direction which points at $f(x_0)$ (i.e., $c(z_0) = f(x_0)$), then $c(z^*) = y^*$ implies that the direction of the second shot becomes $x_1 = (z^* - z_0) + z^*$. Since $x_0 = z^*$ this is the same as the tentative iterate $\tilde{x}_1$ suggested by the first space mapping iteration (see (7) *and* (8) *below where $B_0 = I$).*

Essentially this way of a coarse model interacting with a fine model (or as here: the physical reality) has been used in engineering practice for decades.

The idea of the space mapping technique is to establish a connection between the coarse and the fine models, through a parameter mapping, and to utilize this mapping for finding an optimal set of parameters for the fine model. In other words we are interested in establishing a parameter mapping $p : \Omega^{(f)} \to \Omega^{(c)}$ which yields an approximation of the form

$$f(x) \simeq c(p(x)), \tag{3}$$

where the mapping function p relates similar responses in the following sense: For $x \in \Omega^{(f)}$ we obtain $z = p(x) \in \Omega^{(c)}$ as a solution to the subproblem

$$z \in \arg \min_{\hat{z} \in \Omega^{(c)}} \| f(x) - c(\hat{z}) \|, \tag{4}$$

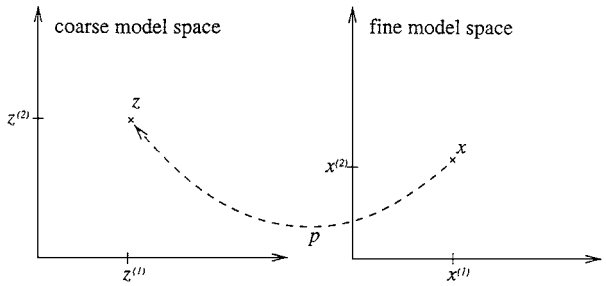

Figure 2. The mapping function relating the fine and the coarse model spaces, shown here for the two-dimensional case, $[z^{(1)}\ z^{(2)}]^T = p([x^{(1)}\ x^{(2)}]^T)$.

for some specific norm. In the present paper we assume that this optimal solution is unique. For the problem of multiple solutions we refer to Bakr et al. (2000b). The concept of the mapping function is illustrated in figure 2 for the two-dimensional case.

If the approximation (3) is close then the composite function $c \circ p$ is applicable as a surrogate for f. Hence the optimal solution of $c \circ p$ can be expected to be close to the optimal solution of f. In other words we might optimize $c \circ p$ rather than f which is expected to be easier under the condition that c and f have similar structures: Then we expect p to be a well behaved function, and since c is cheap to calculate, the composite function $c \circ p$ may be easier to optimize than f. This way of replacing f by $c \circ p$ is the basis of the space mapping technique.

Note for the subproblem (4) that for a given x, a calculation of $p(x)$ involves one evaluation of f succeeded by an optimization in the coarse model space $\Omega^{(c)}$. Hence an evaluation of the mapping function is at least as expensive as an evaluation of the fine model.

The space mapping technique assumes the two models are related in such a way that (3) is a close approximation. Hence $c \circ p$ is optimized in the effort of finding a solution to (1) and for this we apply classical optimization techniques. The problem formulation is

$$\bar{x} \in \arg \min_{x \in \Omega^{(f)}} H(c(p(x))), \tag{5}$$

where $\bar{x}$ may be close to x^* if $c \circ p$ is close to f. Observe that if the optimal solution z^* of $H \circ c$ is unique then the solution of (5) is equivalent of solving the system of n non-linear equations

$$p(x) = z^* \tag{6}$$

for x. In other words $\bar{x} = p^{-1}(z^*)$.

In the first space mapping paper Bandler et al. (1994) estimate the mapping p on the basis of some predefined weighted fundamental functions and evaluations of f at a selected set of base points in $\Omega^{(f)}$. Bandler et al. (1995) formulated the problem as solving (6) for x using Broyden's method for non-linear equations (Broyden, 1965). Bakr et al. (1998) introduced a trust-region methodology to enhance the global convergence properties. The details of

these different approaches are described in the review paper by Bakr et al. (2000b). Recent results of combining space mapping and direct optimization in the field of microwave circuit design are described in Bakr et al. (2000a)

2. Space mapping details

In our formulation the space mapping intends to solve (5) by iteration. At the kth iteration the mapping function p as defined in (4) is replaced by a local estimate p_k, and then the optimal solution of $H \circ c \circ p_k$ is the next iterate. The question is how to find a good approximation p_k. In this presentation we choose to iteratively approximate p by a first order approximation, with the Jacobian matrix approximated using Broyden's rank one update formula.

Let the kth iterate be x_k and assume $z_k = p(x_k)$ has been found by (4). Letting the kth Jacobian approximation be B_k, the corresponding linearization is

$$p_k(x) = B_k(x - x_k) + z_k. \tag{7}$$

The $(k + 1)$th tentative iterate is:

$$\tilde{x}_{k+1} \in \arg\min_{x \in \Omega^{(f)}} H(c(p_k(x)). \tag{8}$$

In case of multiple optimal solutions we choose the one having the shortest distance to the previous iterate x_k. If $H(f(\tilde{x}_{k+1})) < H(f(x_k))$ then the next iterate x_{k+1} is chosen as $\tilde{x}_{k+1}$, otherwise $x_{k+1} = x_k$.

Now $\tilde{z}_{k+1} = p(\tilde{x}_{k+1})$ is found by (4) and finally the Jacobian approximation is updated by Broyden's formula:

$$B_{k+1} = B_k + \frac{\tilde{z}_{k+1} - z_k - B_k h_k}{h_k^T h_k} h_k^T, \tag{9}$$

where $h_k = \tilde{x}_{k+1} - x_k$. Notice that the update is always performed, independently of the acceptance of the tentative point $\tilde{x}_{k+1}$.

Initially the optimal solution z^* of $H \circ c$ is found and used as the first iterate: $x_0 = z^*$. This can be interpreted as an assumption that p is close to the identity mapping:

$$f(x) \simeq c(p(x)) \simeq c(Ix) \tag{10}$$

where $I = I(n)$. It corresponds to the initial aim at the bull's eye in the archery example of the previous section.

The motivation for the initial choice of the Jacobian approximation is another intuition used in the archery example: To mirror the error. This intuition is based on the assumption that the difference between the two model functions is close to a parameter translation:

$$f(x) \simeq c(p(x)) \simeq c(Ix + C_0) \tag{11}$$

where C_0 is a constant, i.e., $p_1(x) = Ix + C_0$. Since $p(x_0) = z_0$ we obtain $C_0 = z_0 - Ix_0$, and thus (11) suggests that $p_0(x)$ is given by (7) with $k = 0$ and $B_0 = I(n)$. Hence the traditional choice of B_0 in Broyden's method is motivated by the archer's simplification (11).

The validity of the mapping approximation p_k is confined to a trust region of size δ_k, hence the feasible set at iteration k is

$$x \in \Omega^{(p_k)} \equiv \{\tilde{x} \mid \|\tilde{x} - x_k\| \leq \delta_k\} \cap \Omega^{(f)}, \tag{12}$$

for some specific norm, thus (8) is replaced by

$$\tilde{x}_{k+1} \in \arg \min_{x \in \Omega^{(p_k)}} H(c(p_k(x)). \tag{13}$$

The update of the trust region size δ_k follows the classical scheme: Significant improvement in the objective compared to the predicted improvement by the approximation is rewarded by enlarging the trust region, whereas insufficient improvement leads to decreasing the trust region size, see Moré (1982) for a thorough treatment of this subject.

For many engineering purposes this formulation yields sufficiently accurate results. However, the convergence of the approach depends on the similarity between the two models. Now, assume the sequence $\{x_k\}$ generated using (13) converges to the solution $\bar{x}$ of (6), then $z^* = p(\bar{x})$. If $\bar{x} = x^*$ then $z^* = p(x^*)$; if, however, the response of the coarse model is less accurate than that of the fine model then we cannot expect z^* and x^* to correspond. Hence in general we must expect $\bar{x} \neq x^*$.

In case of convergence the typical performance we have noticed is a decrease of $\|x_k - x^*\|$ as long as this distance is of a larger order of magnitude than $\|\bar{x} - x^*\|$. Finally, as x_k approaches $\bar{x}$, $\|x_k - x^*\|$ starts to increase.

This observation indicates that the space mapping technique may be considered a good preprocessing process, but not a method for obtaining an accurate solution. If the latter is required then another (i.e., locally convergent) method of optimization will be necessary in the final stages. A switch of method should ideally take place when the distance $\|x_k - x^*\|$ has reached the same order of magnitude as $\|\bar{x} - x^*\|$. The combined strategy of the following section represents some early attempts to reach this ideal goal.

3. Combining with classical methods

This section demonstrates how the space mapping technique can be combined with classical methods of optimization, based on local Taylor type approximations.

Assume the space mapping technique has been used for a number of iterations. Hence a number of fine model evaluations $f(x_k)$ have been calculated. On the basis of these we build an approximation of the Jacobian of f using, for instance, Broyden's formula:

$$D_{k+1} = D_k + \frac{f(x_{k+1}) - f(x_k) - D_k h_k}{h_k^T h_k} h_k^T, \tag{14}$$

where $h_k = x_{k+1} - x_k$. The initial Jacobian approximation is related to the Jacobian of the mapped coarse model at x_0:

$$\begin{aligned} D_0 &= \nabla_{x=x_0}[c(p(x))] = \nabla_{z=z_0}[c(z)] \cdot \nabla_{x=x_0} p(x) \\ &\approx \nabla_{z=z_0}[c(z)] \cdot \nabla_{x=x_0} p_1(x) = \nabla_{z=z_0}[c(z)] \end{aligned} \tag{15}$$

where the "$\approx$" is probably not very precise but in accordance with the intuition (11) used when we start the space mapping. This yields a local linearization of the fine model

$$l_k(x) = D_k(x - x_k) + f(x_k). \tag{16}$$

Traditionally we would minimize $H \circ f$ iteratively using (16) as a basis for finding the $(k+1)$th tentative iterate:

$$\tilde{x}_{k+1} \in \arg \min_{x \in \Omega^{(l_k)}} H(l_k(x)), \tag{17}$$

where $\Omega^{(l_k)}$ is some trust region to be updated during the iteration. The next iterate is $x_{k+1} = \tilde{x}_{k+1}$ if the objective $H \circ f$ is improved, otherwise $x_{k+1} = x_k$. Under mild conditions this iteration yields convergence to a stationary point x^* (see (1)) of f, see e.g., Madsen (1986).

In the present context we use a combination of (16) and the space mapping model $c \circ p_k$ of f: At the kth iteration the combined surrogate for f is

$$s_k(x) = \omega_k \cdot c(p_k(x)) + (1 - \omega_k) \cdot l_k(x), \tag{18}$$

where $\omega_k \in [0; 1]$. Thus the $(k+1)$th tentative iterate is:

$$\tilde{x}_{k+1} \in \arg \min_{x \in \Omega^{(s_k)}} H(s_k(x)). \tag{19}$$

where $\Omega^{(s_k)}$ is a trust region to be updated during the iteration. In case of multiple solutions we choose the one closest to x_k. The next iterate is $x_{k+1} = \tilde{x}_{k+1}$ if the objective $H \circ f$ is improved, otherwise $x_{k+1} = x_k$.

The intention is to use the space mapping surrogate initially (i.e., $\omega_k = 1$) and the local approximation (i.e., $\omega_k = 0$) in the final stages of the iteration. Hence the weighting factor ω_k can be used in a transition from the space mapping surrogate $c \circ p_k$ to a local linearization l_k.

We expect the usefulness of the linear model to increase as the iteration approaches the optimal solution of f. On the other hand, we expect $c \circ p_k$ to be insubstantial in describing f accurately in the vicinity of the optimal solution. Hence we would like to use the information given in the coarse model at the initial stages of the iterations, and as we approach the optimal solution we would like to do a direct optimization, by having the linear model l_k dominate s_k.

In general, we do not wish to change the value of ω if the steps produced by the space mapping algorithm yield a sufficient reduction in the objective function $H \circ f$.

A very simple method of updating ω_k which fulfills these conditions is to define $\omega_{k+1} = \omega_k$ if the objective has been improved, and $\omega_{k+1} = \omega_k/2$ otherwise. More sophisticated

updating strategies are currently being investigated. Some suggestions are found in Bakr (2000), Bakr et al. (2000a), and Søndergaard (1999). The challenge is to find a good combination of the trust region radius update and the ω_k update.

4. Examples

Example 1a. To illustrate the space mapping method we consider the design of a two-section capacitively-loaded 10 : 1 impedance transformer. The coarse and the fine models are shown in figure 3. Assume that the fine model is very expensive and is not recommended for direct optimization. The values of the fine model capacitances are given in Table 1. The characteristic impedances are kept fixed at the optimal values given in Table 1. The physical lengths L_1 and L_2 of the two transmission lines are selected as designable parameters. Eleven frequency points are simulated per sweep. We consider the input reflection coefficient response $f^{(j)}(x) = |S_{11}(t^{(j)}; x)|$ (notice that $S_{11}(t^{(j)}; x) > 0$ for all x) of both models which is a function of the real frequency t and the designable parameters $x = [L_1\ L_2]^T$.

The design specifications are $|S_{11}(t^{(j)}; x)| \leq 0.50$ for the frequency interval $t \in [0.5; 1.5]$ GHz. Hence we wish to find a design $x = x^*$ of the fine model yielding

$$H(f(x)) \equiv \max_j \left\{ f^{(j)}(x) \right\} \leq 0.50. \tag{20}$$

In the following we review some results of applying the combined method (18) on this problem.

Table 1. The fine model capacitances, and the characteristic impedances for the two-section capacitively-loaded impedance transformer.

Capacitance	Value (pF)	Impedance	Value (ohm)
C_1	10	Z_1	4.47214
C_2	10	Z_2	2.23607
C_3	10		

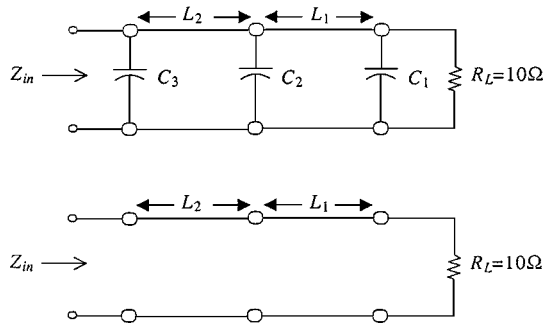

Figure 3. Fine and coarse model, two-section capacitively-loaded impedance transformer.

Table 2. The optimal coarse and fine model parameters z^* and x^* (physical lengths of the transmission lines) for the two-section capacitively-loaded impedance transformer.

z^* (m)	x^* (m)
0.01724138	0.06186103
0.01724138	0.06605482

Given the optimal coarse model parameters z^* (in Table 2), initially we let $x_0 = z^*$, figure 4 shows the fine model response $f(x_0)$. The figure illustrates how the initial fine model design at x_0 violates the specifications (20). Solving the subproblem (4) we find $z_0 = p(x_0)$, such that $c(z_0)$ (also shown in figure 4) is close to $f(x_0)$.

After the first iteration x_1 is found using (19) and from figure 5 we note how the fine model response $f(x_1)$ meets the specifications. For the engineering purpose of finding a design satisfying the specifications (20) a result like this is sufficient. Until this stage the algorithm has used two fine model evaluations.

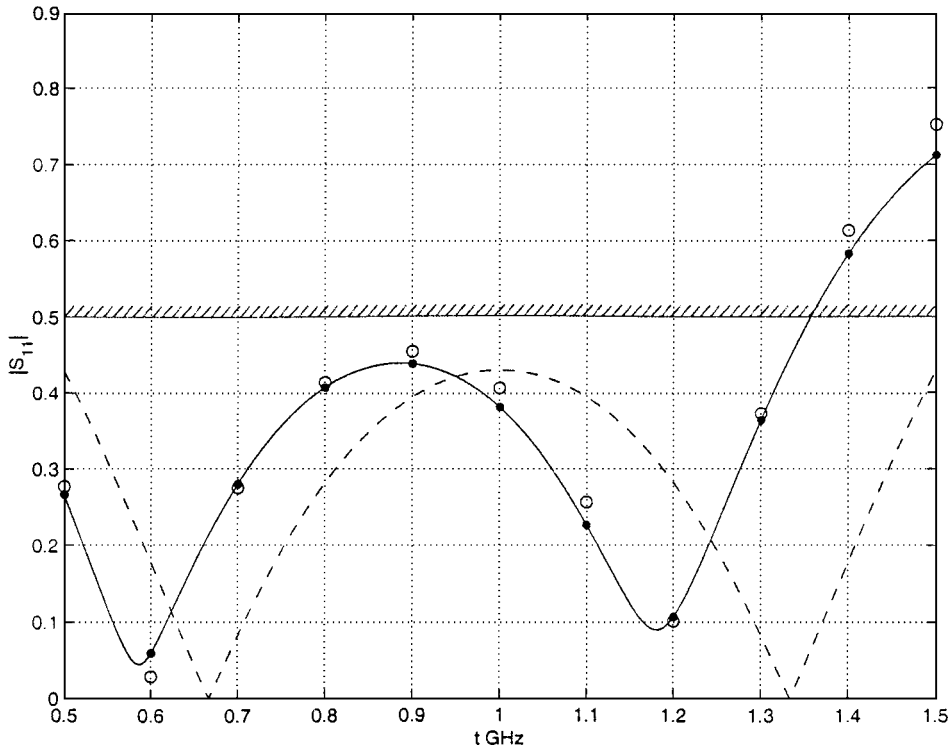

Figure 4. Two-section capacitively-loaded impedance transformer: The fine model response $f(x_0)$ ($\circ$) at the coarse model optimal solution $x_0 = z^*$ and the coarse model response $c(z_0)$ (—•—), $z_0 = p(x_0)$. The dashed curve is the optimal coarse model response $c(z^*)$ which the mapped coarse model $c \circ p$ is aiming for, see (5).

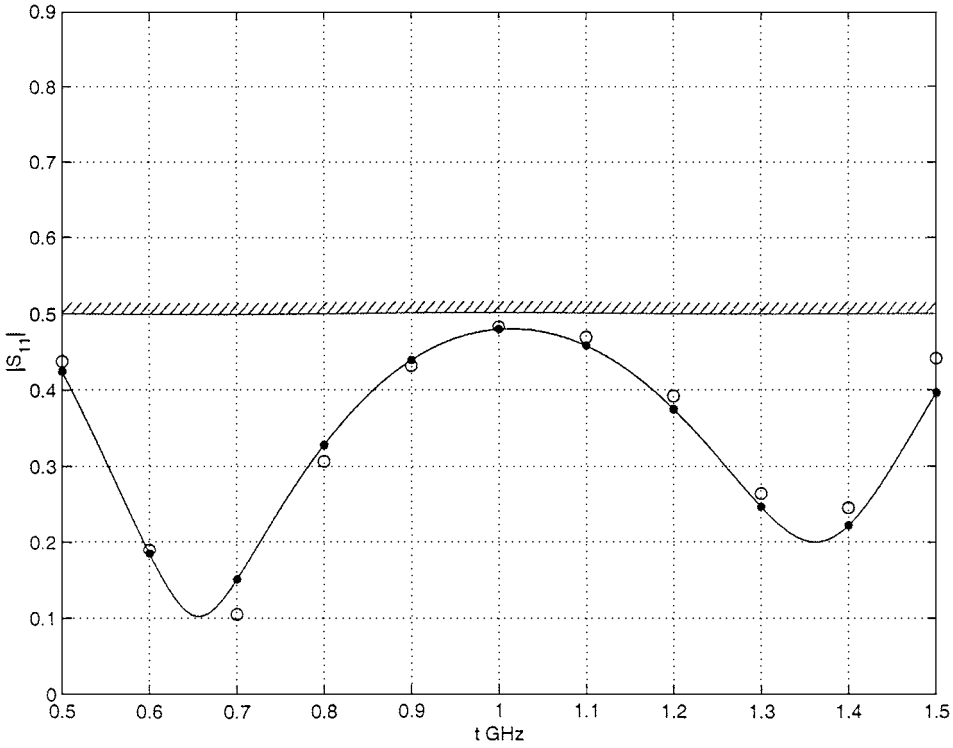

Figure 5. Two-section capacitively-loaded impedance transformer: The fine model response $f(x_1)$ ($\circ$) and the coarse model response $c(z_1)$ (—•—), $z_1 = p(x_1)$.

The visual difference from the fine model design at x_1 to the optimal design x^* (given in Table 2) is rather small: figures 5 and 6 show that from the first iteration to the solution the objective is decreased only from $H(f(x_1)) = 0.481$ to $H(f(x^*)) = 0.455$. It turns out that the distance between x_1 and the solution x^* is so small that the coarse model is unable to provide sufficient improvements after x_1 (in accordance with the argument at the end of Section 2). Hence the algorithm switches rapidly to the local linear model which—in the near neighbourhood of an iterate x_k—is more accurate than the mapped coarse model. The fact that the local linear model is preferable when only small steps are needed is illustrated in Example 1b.

Example 1b. Using the same problem, we here give a graphical illustration of how the mapped coarse model approximation $c \circ p_k$ is a valid approximation to f in a larger region than a linearization l_k of f. The following point is to be made: When large steps are needed then the mapped coarse model approximation is the better, and when small steps are needed (e.g., when we are close to x^*) then the linearization l_k is the better. In order to make the argument more clear we insert accurate Jacobian approximations, B_k to $p'(x_k)$ in (7), and D_k to $f'(x_k)$ in (16) (these approximations being found using finite differences).

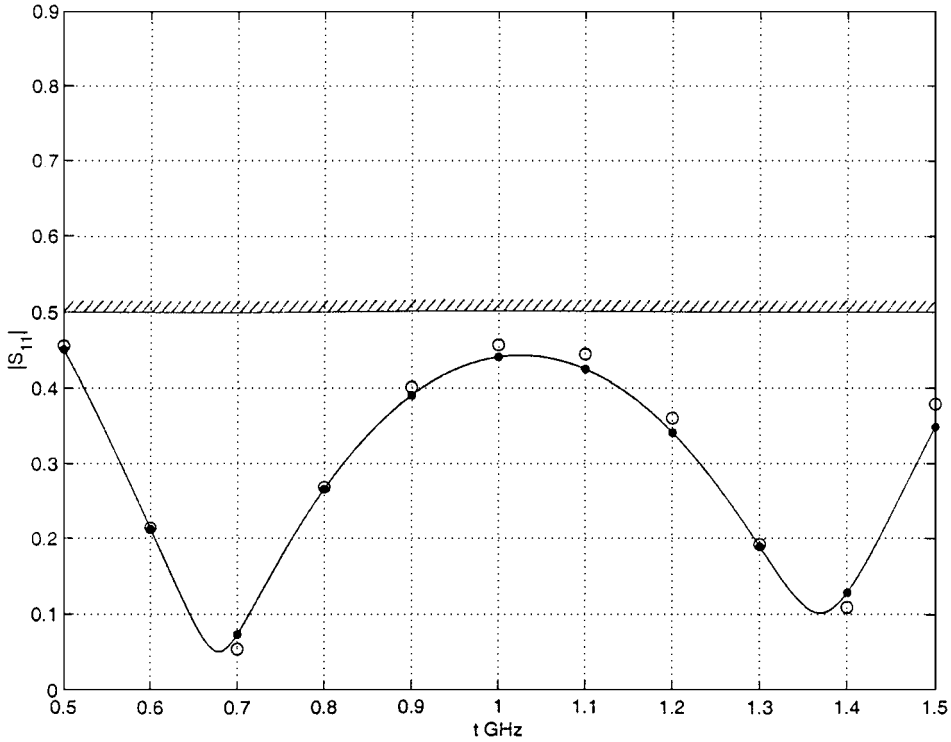

Figure 6. Two-section capacitively-loaded impedance transformer: The fine model response $f(x^*)$ ($\circ$) and the coarse model response $c(p(x^*))$ (—●—).

In figure 7 the mapped coarse model approximation error $\|c(p_k(x)) - f(x)\|_2$ is plotted for points on a mesh in a square region centered at x_k. The linearized fine model approximation error $\|l_k(x) - f(x)\|_2$ is plotted for points at the same mesh. The figure illustrates, as expected, how the approximation error of the linear approximation l_k (which is zero at x_k) grows with the square of the distance from x_k. The approximation error of the mapped coarse model $c \circ p_k$, however, does not grow systematically with the distance from x_k, in fact it is almost constant in the region considered. Furthermore we note that $c \circ p_k$ does not interpolate f, i.e., $\|c(p_k(x)) - f(x)\|_2$ is non-zero at $x = x_k$.

From these observations we conclude that close to x_k the better approximation to f is l_k, whereas $c \circ p_k$ is the better away from x_k. In fact $c \circ p_k$ is a valid approximation to f in the whole region shown in figure 7.

Example 2. In this example we consider the design of a seven-section capacitively-loaded impedance transformer. The load impedance is 100 Ω and the line impedance is 50 Ω. The coarse and the fine models are shown in figure 8. The values of the fine model capacitances are given in Table 3. The characteristic impedances are synthesized using an equi-ripple approximate design procedure (Pozar, 1998) and are kept fixed at these values given in Table 3. The physical lengths L_i, $i = 1, \ldots, 7$, of the seven transmission lines are selected

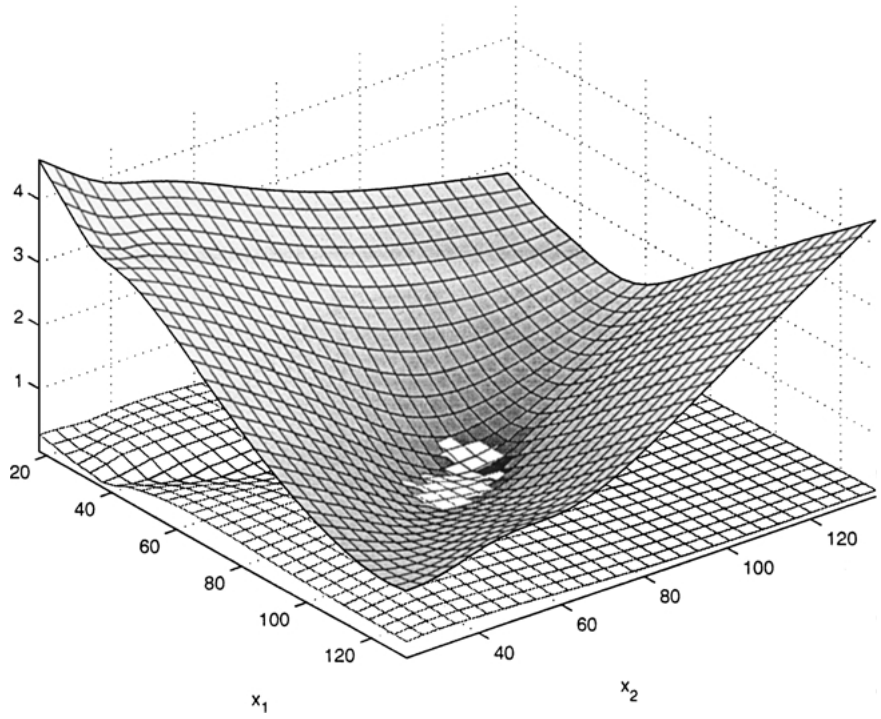

Figure 7. Two-section capacitively-loaded impedance transformer: Mapped coarse model approximation error $\|c(B_k(x-x_k)+p(x_k))-f(x)\|_2$ (white mesh), linearized fine model approximation error $\|D_k(x-x_k)+f(x_k)-f(x)\|_2$ (gray scale mesh). For both meshes: x_k the point of linearization is in the center of the plot.

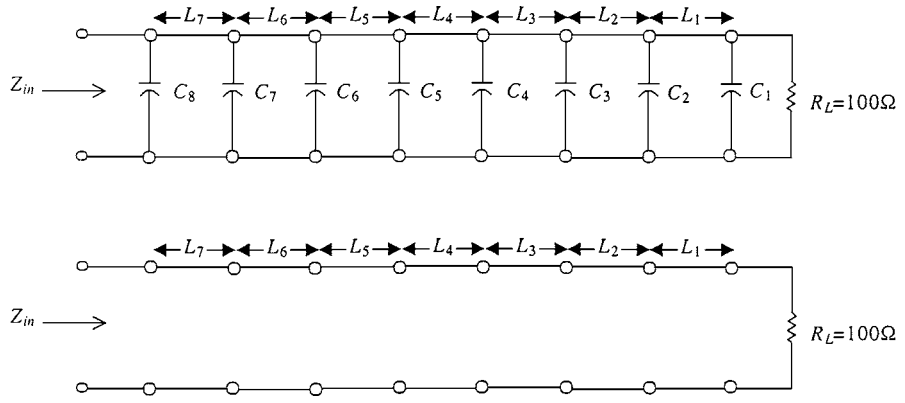

Figure 8. Fine and coarse model, seven-section capacitively-loaded impedance transformer.

Table 3. The fine model capacitances, and the characteristic impedances for the seven-section capacitively-loaded impedance transformer.

Capacitance	Value (pF)	Impedance	Value (ohm)
C_1	0.025	Z_1	91.9445
C_2	0.025	Z_2	85.5239
C_3	0.025	Z_3	78.1526
C_4	0.025	Z_4	70.7107
C_5	0.025	Z_5	63.9774
C_6	0.025	Z_6	58.4632
C_7	0.025	Z_7	54.3806
C_8	0.025		

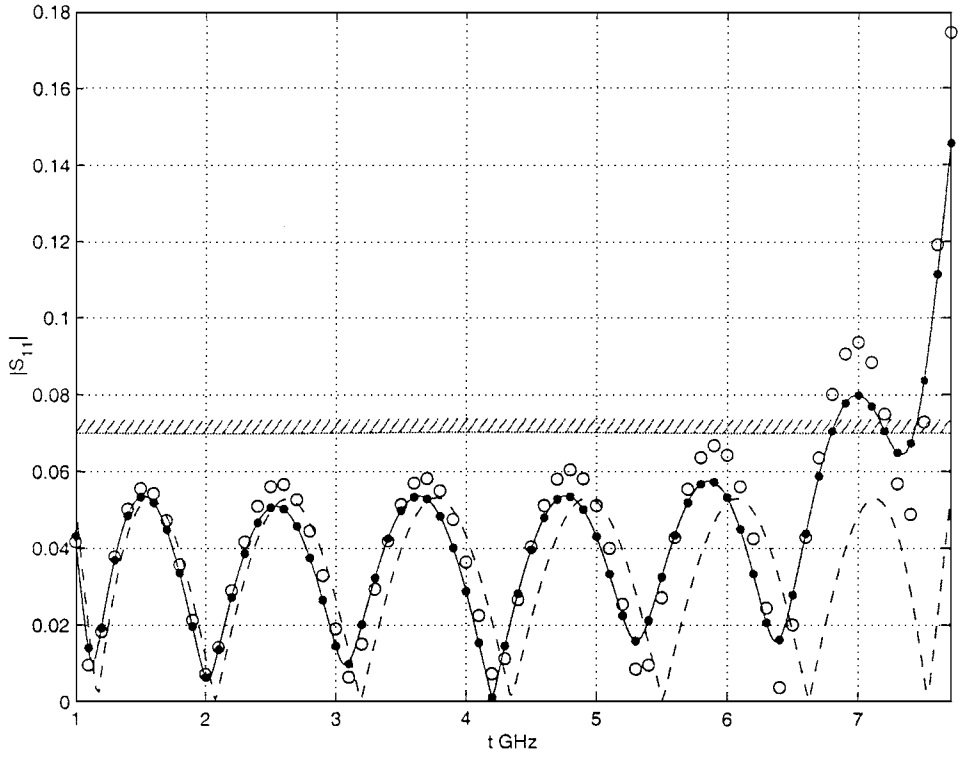

Figure 9. Seven-section capacitively-loaded impedance transformer: The fine model response $f(x_0)$ ($\circ$) at the coarse model optimal solution $x_0 = z^*$ and the coarse model response $c(z_0)$ (—•—), $z_0 = p(x_0)$. The dashed curve is the optimal coarse model response $c(z^*)$ which the mapped coarse model $c \circ p$ is aiming for, see (5).

as designable parameters. We consider the input reflection coefficient response $f^{(j)}(x) = |S_{11}(t^{(j)}; x)|$, and the design specifications are $|S_{11}(t^{(j)}; x)| \leq 0.07$ for the frequency interval $t \in [1; 7.7]$ GHz.

In figure 9 the fine model response is plotted at the optimal design of the coarse model, $x_0 = z^*$. The coarse model response at the design z_0 (being the design at which the coarse model response is closest to the fine model response $f(x_0)$) is also plotted in figure 9. It is seen that this coarse model response is not very accurate in describing the fine model response indicating that the correspondence between the two models is less obvious in this case.

In figure 10 the optimal fine model response is plotted together with the closest coarse model response. We see how the coarse model poorly describes the fine model at this design, in this case the space mapping algorithm is depending heavily on the classical method to be able to converge to the optimal solution (not another local minimum). The optimal coarse and fine model parameters are given in Table 4.

A feasible solution is found after 18 fine model evaluations. At this stage the combination parameter ω_k of (18) has been downdated to 0.016, so the space mapping is almost abandoned

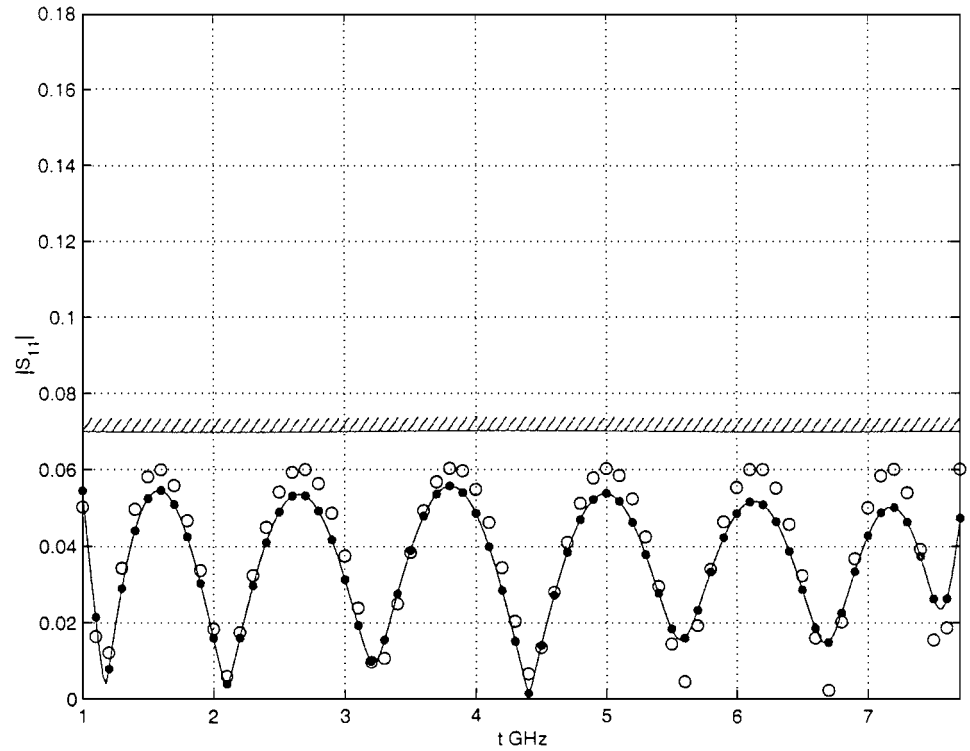

Figure 10. Seven-section capacitively-loaded impedance transformer: The fine model response $f(x^*)$ (○) and the coarse model response $c(p(x^*))$ (—●—).

Table 4. The optimal coarse and fine model parameters z^* and x^* (physical lengths of the transmission lines) for the seven-section capacitively-loaded impedance transformer.

z^* (m)	x^* (m)
0.01724138	0.01564205
0.01724138	0.01638347
0.01724138	0.01677145
0.01724138	0.01697807
0.01724138	0.01709879
0.01724138	0.01723238
0.01724138	0.01625988

from this stage, i.e., the rest of the iterations are practically speaking based on the local linear model l_k.

For comparison we have solved this problem directly using an implementation of the minimax optimization method of Hald and Madsen (1981) with finite differences to approximate the fine model Jacobians. As initial iterate we use the coarse model optimal solution, i.e., $x_0 = z^*$. This way we find a feasible solution after 25 fine model evaluations.

5. Conclusions

The basic principles of the space mapping technique have been presented. It is shown how the space mapping technique can be combined with classical optimization strategies. The combined method is illustrated by a simple two-dimesional example and a more complicated seven-dimensional example. The space mapping surrogate is shown by example to be a valid approximation to the fine model in a larger region than a corresponding fine model linearization using the same number of fine model evaluations.

The space mapping has proved to be an efficient preprocessing technique in many difficult engineering optimization problems. The solution accuracy is often sufficient for practical purposes. Otherwise the technique can be combined with other methods of optimization.

Acknowledgments

The authors would like to thank José Rayes-Sanchez and A. S. Mohamed of McMaster University. The work in Canada was supported by NSERC under Grants OGP0007239 and STR234854-00, and by Bandler Corporation.

References

M. H. Bakr, "Advances in space mapping optimization of microwave circuits," Ph.D. Thesis, McMaster University, 2000. Available at `http://www.sos.mcmaster.ca/theses.htm`

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 Trans. Microwave Theory Tech*. vol. 46, no. 12, pp. 2412–2425, 1998.

M. H. Bakr, J. W. Bandler, K. Madsen, J. E. Rayas-Sánchez, and J. Søndergaard, "Space mapping optimization of microwave circuits exploiting surrogate models," *IEEE Trans. Microwave Theory Tech*. vol. 48, pp. 2297–2306, 2000a.

M. H. Bakr, J. W. Bandler, K. Madsen, and J. Søndergaard, "Review of the space mapping approach to engineering optimization and modeling," *Optimization and Engineering* vol. 1, no. 3, pp. 241–276, 2000b.

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

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, no. 12, pp. 2874–2882, 1995.

C. G. Broyden, "A class of methods for solving non-linear simultaneous equations," *Math. Comp*. vol. 19, pp. 577–593, 1965.

J. Hald and K. Madsen, "Combined LP and quasi-Newton methods for minimax optimization," *Mathematical Programming* vol. 20, pp. 49–62, 1981.

K. Madsen, "Minimax solution of non-linear equations without calculating derivatives," *Mathematical Programming Study* vol. 3. pp. 110–126, 1975.

K. Madsen, "Minimization of approximation functions," dr. techn. thesis, Technical University of Denmark, 1986.

J. J. Moré, "Recent developments in algorithms and software for trust region methods," *Mathematical Programming: The State of The Art*, pp. 258–287, 1982.

D. M. Pozar, *Microwave Engineering*, 2nd edn., John Wiley: New York, 1998.

J. Søndergaard, "Non-linear optimization using space mapping," Master Thesis IMM-EKS-1999-23 at IMM, DTU, 1999. Available at `http://www.imm.dtu.dk/~km/jsmaster.ps.gz`