Centrum voor Wiskunde en Informatica

**Modelling, Analysis and Simulation**

MAS Space mapping and defect correction

D. Echeverría, P.W. Hemker

CWI is the National Research Institute for Mathematics and Computer Science. It is sponsored by the Netherlands Organization for Scientific Research (NWO).
CWI is a founding member of ERCIM, the European Research Consortium for Informatics and Mathematics.

CWI's research has a theme-oriented structure and is grouped into four clusters. Listed below are the names of the clusters and in parentheses their acronyms.

Probability, Networks and Algorithms (PNA)

Software Engineering (SEN)

Modelling, Analysis and Simulation (MAS)

Information Systems (INS)

Space mapping and defect correction

ABSTRACT

In this paper we show that space-mapping optimization can be understood in the framework of defect correction. Then, space-mapping algorithms can be seen as special cases of defect correction iteration. In order to analyze properties of space mapping and the space-mapping function we introduce the new concept of flexibility of the underlying models. The best space-mapping results are obtained for so-called equally flexible models. By introducing an affine operator as a left preconditioner, two models can be made equally flexible, at least in the neighborhood of a solution. This motivates an improved space-mapping (or manifold-mapping) algorithm. The left preconditioner complements traditional space mapping where only a right preconditioner is used. In the last section a few simple examples illustrate some of the phenomena analyzed in this paper.

Space Mapping and Defect Correction

D. Echeverría and P.W. Hemker

CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

ABSTRACT

In this paper we show that space-mapping optimization can be understood in the framework of defect correction. Then, space-mapping algorithms can be seen as special cases of defect correction iteration. In order to analyze properties of space mapping and the space-mapping function we introduce the new concept of flexibility of the underlying models. The best space-mapping results are obtained for so-called equally flexible models.

By introducing an affine operator as a left preconditioner, two models can be made equally flexible, at least in the neighborhood of a solution. This motivates an improved space-mapping (or manifold-mapping) algorithm. The left preconditioner complements traditional space mapping where only a right preconditioner is used. In the last section a few simple examples illustrate some of the phenomena analyzed in this paper.

2000 Mathematics Subject Classification: 65K05, 65B05
Keywords and Phrases: Space mapping, defect correction, two-level optimization

1. Introduction

The space-mapping idea was conceived by Bandler [4] in the field of microwave filter design. It aims at reducing the cost of accurate optimization computations by iteratively correcting a sequence of rougher approximations. In technological applications this allows us to couple simple rules that represent expert knowledge accumulated over the years with the accuracy of expensive simulation techniques based on the numerical solution of partial differential equations. This combination may yield an efficient method with good accuracy of the final solution. As an example, in Section 5.2 we significantly accelerate the solution process for an optimization problem from magnetostatics by combining finite elements for the precise computations of the magnetic field with rather simple magnetic circuit calculations.

The space-mapping technique has been mainly applied in electromagnetics [6, 8] but, since the underlying principles are quite general, it could also be used in other areas [23, 28]. Space mapping has developed enormously during the last decade. As can be seen in the rather complete survey [6], the original idea has gone through a large number of changes and improvements. Further, it can be applied in model reduction for the construction of cheap surrogates [10, 29]. But this modeling purpose is still of secondary importance.

In the present paper we consider space mapping from the perspective of defect correction. Defect correction is a well-know principle in computational mathematics [13] to solve accurate but complex operator equations in an iterative manner with the help of approximate but easier

operators. The idea of solving a problem with the help of a sequence of simpler versions is quite old and well established. Famous examples are, e.g., Newton iteration and multigrid methods [20].

Also the space-mapping technique can be considered as defect correction iteration. This viewpoint allows some mathematical analysis and the construction of a family of more general iteration procedures, generalizing space mapping.

In this paper we also introduce the concept of model *flexibility*, that gives additional insight and understanding of space-mapping performance. The most significant conclusion derived from that concept is that the best space-mapping techniques are those in which both models involved have equal flexibility, while the less accurate model can be solved much faster.

Traditional space-mapping techniques cannot change the model flexibility. In this paper we present a new space-mapping approach that, by additional left preconditioning [33] with an affine operator, aims at locally endowing the two given models with equal flexibility. In fact, this is rather a manifold-mapping than a space-mapping technique. There appears to be a similarity with the recently published output space mapping (OSM) [7] and space-mapping based interpolating surrogate (SMIS) scheme [9]. However, our flexibility analysis makes clear that properly adapted left-preconditioning is essentially more effective than the traditional space mapping, which corresponds with right-preconditioning.

The present paper is structured as follows. Section 2 briefly presents the most fundamental concepts and definitions of space mapping. Section 3 first introduces the two types of space-mapping solutions and presents typical algorithms found in literature for their computation. Then the notions of perfect mapping and flexibility of a model are presented and some relations between these are explained. In Section 4 the defect correction principle is introduced and its relation with space mapping is shown. Some defect correction iteration schemes can be recognized as space mapping methods but also generalizations can be found. These processes are analyzed and a procedure is constructed so that locally the model flexibility for the different models coincides. This yields our improved space mapping method. By means of a few simple examples, in Section 5 we illustrate some essential concepts and lemmas from the space-mapping theory presented in the previous sections. We conclude the paper with a more realistic magnetic shape design problem which corroborates the efficiency of the space-mapping technique for optimization.

2. Aims and first definitions

Two aims. Although space mapping is mostly employed as a means for optimization, it is useful to know that it can be applied with two different objectives in mind –optimization and model reduction– and that, with the different use, different emphasis is put on particular aspects of space mapping.

We distinguish between: (1) The space-mapping technique aiming at efficient optimization of complex technical problems. The idea is to combine the advantages of simple (*coarse*, less accurate, easy to evaluate) models and more complex (*fine*, more accurate, expensive) models of a phenomenon, to accelerate the numerical computation of a fine-model optimum. Typical examples of such an optimization problem are, e.g. fitting a model to a set of measured data or determining the shape of an electromagnet in order to achieve a magnetic field with certain properties. More generally, (2) space mapping deals with constructing a simpler or *surrogate model* instead of a more expensive and accurate one, and with calibrating this surrogate.

The main difference between the two purposes of space mapping (*optimization* or *model reduction*) is the required validity region in which the surrogate approximation should correspond with the fine one. In the first case it is sufficient if the two models correspond in the neighborhood of the optimum where the optimization procedure dwells. In the latter case the surrogate model should be sufficiently accurate over a larger range of its parameters.

Though most of the concepts can also be applied in model reduction, this paper is primarily concerned with the use of space mapping in optimization

The optimization problem.   Let the specifications for the data of an optimization problem be denoted by $(\mathbf{t}, \mathbf{y}) \equiv (\{t_i\}, \{y_i\})_{i=1,\dots,m}$. The independent variable $\mathbf{t} \in \mathbb{R}^m$ could be, e.g., time, frequency, space, etc. The dependent variable $\mathbf{y} \in Y \subset \mathbb{R}^m$ represents the quantities that describe the behavior of the phenomena under study or design. The set $Y \subset \mathbb{R}^m$ is called the *set of possible aims.* In addition to each value y_i a range of its probability (or accuracy) can be given.

The behavior of the variables $\mathbf{y}$ depends not only on the independent variable $\mathbf{t}$ but also on an additional set of control/design variables. What are relevant control variables depends on the situation. With $\mathbf{x}$ the vector of relevant control variables, we may write $y_i \approx y(t_i, \mathbf{x})$. The behavior of the phenomenon is described by the function $y(t, \mathbf{x})$ and the difference between the measured data and the physical reality may be the result of, e.g., measurement errors or of the imperfection of the mathematical description.

Since in reality the behavior of the phenomenon may be extremely complex or even impossible to describe in all its details, we usually describe it by some mathematical model. Such models, however appear in several degrees of sophistication. It is the purpose of this research to exploit the combination of the simplicity of the less sophisticated methods with the accuracy of the more complex ones. Therefore, we first distinguish two types of model: fine and coarse.

The fine model.   The *fine model* response is denoted by $\mathbf{f}(\mathbf{x}) \in \mathbb{R}^m$, where $\mathbf{x} \in X \subset \mathbb{R}^{n_{\mathbf{f}}}$ is the *control variable*[1]. The set X of possible control variables is usually a closed and bounded subset of $\mathbb{R}^{n_{\mathbf{f}}}$. The set $\mathbf{f}(X) \subset \mathbb{R}^m$ of all possible fine model responses is called the set of *fine model reachable aims*. The fine model is assumed to be *accurate* but *expensive* to evaluate. We assume that $\mathbf{f}(\mathbf{x})$ is continuous. Possibly it is differentiable but its Jacobian matrix $J_{\mathbf{f}} = \mathbf{df}/\mathbf{dx}$ is generally supposed to be either unavailable or very expensive to compute.

For the optimization problem a *fine model cost function*, $F(\mathbf{x}) = \|\,\mathbf{f}(\mathbf{x}) - \mathbf{y}\,\|$ is defined, which is a measure for the discrepancy between the data and a particular response of the mathematical model. This cost function should be minimized. So we look for

$$\mathbf{x}^* = \operatorname*{argmin}_{\mathbf{x} \in X} \|\,\mathbf{f}(\mathbf{x}) - \mathbf{y}\,\| \,. \tag{2.1}$$

The cost function can be quite complex, e.g., it may take into account requirements for responses (such as positivity) or expectations for the errors (possibility of large outliers). Convexity and differentiability are generally not required. However, the cost function should be continuous. For simplicity, in this paper we take for $\|\cdot\|$ a norm on $\mathbb{R}^m$.

---

[1]In fact $\mathbf{f}(\mathbf{x}) = (f(t_1, \mathbf{x}), f(t_2, \mathbf{x}), \cdots, f(t_m, \mathbf{x}))$, but in this paper we distinguish the elements in $\mathbb{R}^m$ by their index.

Definition 1. A design problem, characterized by the model $\mathbf{f}(\mathbf{x})$, the aim $\mathbf{y} \in Y$, and the space of possible controls $X \subset \mathbb{R}^{n_{\mathbf{f}}}$ is called a *reachable design* iff the equality $\mathbf{f}(\mathbf{x}^*) = \mathbf{y}$ can be achieved for some $\mathbf{x}^* \in X$. $\hfill\square$

The coarse model.   The *coarse model* is denoted by $\mathbf{c}(\mathbf{z}) \in \mathbb{R}^m$ with $\mathbf{z} \in Z \subset \mathbb{R}^{n_{\mathbf{c}}}$ the coarse model control variable. This model is assumed to be *cheap* to evaluate but *less accurate* than the fine model. The set $\mathbf{c}(Z) \subset \mathbb{R}^m$ is the set of *coarse model reachable aims* . For the coarse model we have the *coarse model cost function*, $C(\mathbf{z}) = \|\,\mathbf{c}(\mathbf{z}) - \mathbf{y}\,\|$. We denote its minimizer by $\mathbf{z}^*$,

$$\mathbf{z}^* = \underset{\mathbf{z} \in Z}{\operatorname{argmin}} \, \|\,\mathbf{c}(\mathbf{z}) - \mathbf{y}\,\| \,. \tag{2.2}$$

Here the norm $\|\cdot\|$ is similar to the one in (2.1), but for the coarse model we assume that there exists an efficient method for its optimization, whereas this is not the case for (2.1). In principle, the spaces X and Z can have different dimensions.

The assumption of a well defined optimization problem.   We assume that the fine and coarse optimization problems, characterized by $\mathbf{y}$, $\mathbf{f}(\mathbf{x})$ and X, respectively $\mathbf{y}$, $\mathbf{c}(\mathbf{z})$ and Z, are uniquely solvable and *well defined*, i.e.,

$$\forall \mathbf{y} \in Y \;\; \exists! \; \mathbf{x}^* \in X \qquad \mathbf{x}^* = \underset{\mathbf{x} \in X}{\operatorname{argmin}} \, \|\,\mathbf{f}(\mathbf{x}) - \mathbf{y}\,\| \,, \tag{2.3}$$

and

$$\forall \mathbf{y} \in Y \;\; \exists! \; \mathbf{z}^* \in Z \qquad \mathbf{z}^* = \underset{\mathbf{z} \in Z}{\operatorname{argmin}} \, \|\,\mathbf{c}(\mathbf{z}) - \mathbf{y}\,\| \,. \tag{2.4}$$

In many practical cases this assumption is reasonable. If X and Z are closed and bounded non-empty sets in $\mathbb{R}^m$ and $\mathbf{f}$ and $\mathbf{c}$ continuous functions, the existence of the solutions is guaranteed. Generally, uniqueness can be achieved by properly reducing the sets X or Z. If the models are non-injective (or extremely ill-conditioned) in a small neighborhood of a solution, essential difficulties may arise. We assume that such situations do not exist for the fine model. If it occurs for the coarse model, possibly regularisation can remove the difficulty [6, 2].

Space-mapping function.   It is supposed to be much easier to compute $\mathbf{z}^*$ than $\mathbf{x}^*$. With this knowledge and with the information regarding the similarity between the two models $\mathbf{f}$ and $\mathbf{c}$, we aim at an *efficient* algorithm to compute an accurate approximation of $\mathbf{x}^*$. Obviously, the similarity or discrepancy between the responses of two models used for the same phenomenon is here an important property. It is expressed by the *misalignment function* (see Figure 1),

$$r(\mathbf{z}, \mathbf{x}) = \|\,\mathbf{c}(\mathbf{z}) - \mathbf{f}(\mathbf{x})\,\| \,.$$

In particular, for a given $\mathbf{x} \in X$ it is useful to know which $\mathbf{z} \in Z$ yields the smallest discrepancy. This information would allow us to improve the coarse model. Therefore, the space-mapping function is introduced.

Figure 1: Misalignment and space-mapping function.
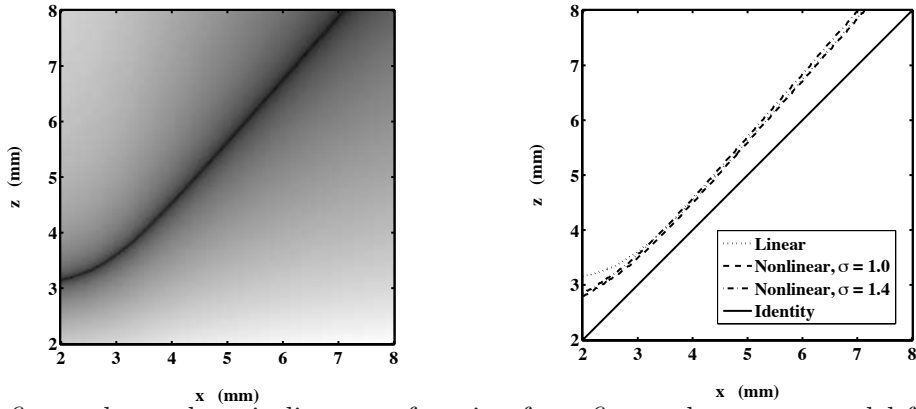
The left figure shows the misalignment function for a fine and a coarse model for the problem in Section 5.2. Darker shading shows a smaller misalignment. The right figure shows the identity function and a few space-mapping functions for different coarse models used in Section 5.2. (Only the variable x_3 is shown; $x_1 = 5$ mm, $x_2 = 7.5$ mm are kept fixed.)

Definition 2. The *space-mapping function* $\mathbf{p} : X \subset \mathbb{R}^{n_\mathbf{f}} \to Z \subset \mathbb{R}^{n_\mathbf{c}}$ is defined by

$$\mathbf{p}(\mathbf{x}) = \operatorname*{argmin}_{\mathbf{z} \in Z} r(\mathbf{z}, \mathbf{x}) = \operatorname*{argmin}_{\mathbf{z} \in Z} \| \mathbf{c}(\mathbf{z}) - \mathbf{f}(\mathbf{x}) \| . \tag{2.5}$$

If (2.5) does not allow for a unique minimum, no space mapping $\mathbf{p} : X \to Z$ exists. $\square$

The definition implies that for a given control $\mathbf{x}$ the function $\mathbf{p}$ delivers the best coarse-model control $\mathbf{z}$ that yields a similar (or the same) response as $\mathbf{x}$ for the fine-model. The assumption of the existence of a unique minimizer is implied by (2.4) if $\mathbf{f}(X) \subset Y$, which can be expected in practise.

The process of finding $\mathbf{p}(\mathbf{x})$ for a given $\mathbf{x}$ is called *parameter extraction* or *single point extraction* because it finds the best coarse-model parameter that corresponds with a given fine-model control $\mathbf{x}$. It should be noted that this evaluation of the space-mapping function $\mathbf{p}(\mathbf{x})$ requires both an evaluation of $\mathbf{f}(\mathbf{x})$ and a minimization process with respect to $\mathbf{z}$ in $\| \mathbf{c}(\mathbf{z}) - \mathbf{f}(\mathbf{x}) \|$. Hence, in algorithms we should make economic use of space-mapping function evaluations.

The success of the space-mapping technique relies on the existence of possible, desirable properties for the space-mapping function. One is its injectivity. When $n_\mathbf{f} > n_\mathbf{c}$, that is, when the model $\mathbf{f}$ has more control parameters than $\mathbf{c}$, the space-mapping function cannot be injective. For that reason, together with the fact that both fine and coarse control variables mostly have the same interpretation, in this paper we restrict ourselves to spaces X and Z having the same dimension, $n_\mathbf{c} = n_\mathbf{f} = n$. But the case $n_\mathbf{f} < n_\mathbf{c}$ has still some practical interest: at the expense of losing uniqueness in the (coarse) optimization, the extra parameters in $\mathbf{c}$ could be used for a better model alignment.

In circumstances where the coarse and the fine model reachable aims are each other's subset, i.e. if $\mathbf{c}(Z) \subset \mathbf{f}(X)$ or $\mathbf{f}(X) \subset \mathbf{c}(Z)$, or coincide, the notion of *flexibility of a model* can be introduced (Section 3.3) to derive properties of the corresponding space-mapping

function. A priori there is no reason to assume that these sets have such a relation or even a non-empty intersection. However, given that both models describe the same phenomenon, in some situations it may occur.

Perfect mapping.   For reasons that will become immediately clear in the section below, we introduce the following definition.

Definition 3. A space-mapping function $\mathbf{p}$ is called a *perfect mapping* iff

$$\mathbf{z}^* = \mathbf{p}(\mathbf{x}^*) \,. \tag{2.6}$$

$\square$

Using the definition of space mapping we see that (2.6) can be written as

$$\operatorname*{argmin}_{\mathbf{z} \in Z} \| \, \mathbf{c}(\mathbf{z}) - \mathbf{y} \| \; = \operatorname*{argmin}_{\mathbf{z} \in Z} \| \, \mathbf{c}(\mathbf{z}) - \mathbf{f}(\mathbf{x}^*) \| \,, \tag{2.7}$$

i.e., a perfect space-mapping function maps $\mathbf{x}^*$, the solution of the fine model optimization, exactly onto $\mathbf{z}^*$, the minimizer of the coarse model design. However, we notice that *perfection* is not a property of the space-mapping function alone but it also depends on the data $\mathbf{y}$ considered. A space-mapping function can be perfect for one set of data but imperfect for a different data set.

3. SPACE-MAPPING OPTIMIZATION

In literature many space mapping based algorithms can be found [6], but they all have the same basis. In this section we first describe the original space-mapping idea and two principal approaches (primal and dual). In Section 3.2 we show a few typical space-mapping algorithms used in practise, and in Section 3.3 we analyze what approximations to a solution are obtained under certain conditions.

3.1 Space-mapping approaches: original, primal and dual

The basic idea behind space-mapping optimization is the following: if either the fine model allows for an almost reachable design (i.e., $\mathbf{f}(\mathbf{x}^*) \approx \mathbf{y}$) or if both models are similar near their respective optima (i.e., $\mathbf{f}(\mathbf{x}^*) \approx \mathbf{c}(\mathbf{z}^*)$)) we expect

$$\mathbf{p}(\mathbf{x}^*) = \operatorname*{argmin}_{\mathbf{z} \in Z} \| \, \mathbf{c}(\mathbf{z}) - \mathbf{f}(\mathbf{x}^*) \| \; \approx \operatorname*{argmin}_{\mathbf{z} \in Z} \| \, \mathbf{c}(\mathbf{z}) - \mathbf{y} \| \; = \mathbf{z}^* \,.$$

Based on this relation, the *original* space-mapping approach [4] assumes $\mathbf{p}(\mathbf{x}^*) \approx \mathbf{z}^*$. It first determines $\mathbf{z}^*$ and then solves

$$\mathbf{p}(\mathbf{x}_{sm}^*) = \mathbf{z}^* \,. \tag{3.1}$$

The algorithm for its computation is shown in Figure 2. However, in general $\mathbf{p}(\mathbf{x}^*) \neq \mathbf{z}^*$ and even $\mathbf{z}^* \in \mathbf{p}(X)$ is not guaranteed, so that the existence of $\mathbf{x}_{sm}^*$ cannot be assumed. Therefore the *primal* space-mapping approach seeks for a solution of the minimization problem

$$\mathbf{x}_p^* = \operatorname*{argmin}_{\mathbf{x} \in X} \| \mathbf{p}(\mathbf{x}) - \mathbf{z}^* \| \,. \tag{3.2}$$

Figure 2: The original space-mapping algorithm.

$\mathbf{x}_0 = \mathbf{z}^* = \operatorname{argmin}_{\mathbf{z} \in Z} \| \mathbf{c}(\mathbf{z}) - \mathbf{y} \|$;
$\mathbf{p}_0 = I$;
for $k = 0, 1, \cdots$, **while** $\| \mathbf{p}(\mathbf{x}_k) - \mathbf{z}^* \| >$ tolerance
do $\mathbf{z}_k = \mathbf{p}(\mathbf{x}_k) = \operatorname{argmin}_{\mathbf{z} \in Z} \| \mathbf{c}(\mathbf{z}) - \mathbf{f}(\mathbf{x}_k) \|$;
   from $\{(\mathbf{x}_j, \mathbf{z}_j)\}_{j=0}^{k}$ determine an updated approximation[a] $\mathbf{p}_{k+1}$ for $\mathbf{p}$;
   $\mathbf{x}_{k+1} = \mathbf{p}_{k+1}^{-1}(\mathbf{z}^*)$;
enddo

---

[a]the approximation is assumed to be invertible

Figure 3: The iteration in the primal space-mapping algorithm.

for $k = 0, 1, \ldots$, **while** $\ldots$
do $\mathbf{z}_k = \mathbf{p}(\mathbf{x}_k)$;
   compute an updated approximation $\mathbf{p}_{k+1}$ by means of $\{(\mathbf{x}_j, \mathbf{z}_j)\}_{j=0}^{k}$;
   $\mathbf{x}_{k+1} = \operatorname{argmin}_{\mathbf{x} \in X} \| \mathbf{p}_{k+1}(\mathbf{x}) - \mathbf{z}^* \|$;
enddo

Figure 4: The iteration in the dual space-mapping algorithm.

for $k = 0, 1, \ldots$, **while** $\ldots$
do $\mathbf{z}_k = \mathbf{p}(\mathbf{x}_k)$;
   compute updated approximation $\mathbf{p}_{k+1}$ by means of $\{(\mathbf{x}_j, \mathbf{z}_j)\}_{j=0}^{k}$;
   $\mathbf{x}_{k+1} = \operatorname{argmin}_{\mathbf{x} \in X} \| \mathbf{c}(\mathbf{p}_{k+1}(\mathbf{x})) - \mathbf{y} \|$;
enddo

This leads to the *primal space-mapping algorithm* as shown in Figure 3.

An alternative approach can be taken. The idea behind space-mapping optimization is the replacement of the expensive fine model optimization by a surrogate model. For the surrogate model we can take the coarse model $\mathbf{c}(\mathbf{z})$, and improve its accuracy by the space mapping $\mathbf{p} : X \to Z$. Now the improved or *mapped coarse model* $\mathbf{c}(\mathbf{p}(\mathbf{x}))$ may serve as the better *surrogate model*. Because of (2.5) we expect that $\mathbf{c}(\mathbf{p}(\mathbf{x})) \approx \mathbf{f}(\mathbf{x})$ and hence

$$\| \mathbf{f}(\mathbf{x}) - \mathbf{y} \| \approx \| \mathbf{c}(\mathbf{p}(\mathbf{x})) - \mathbf{y} \| .$$

Then the minimization of $\| \mathbf{c}(\mathbf{p}(\mathbf{x})) - \mathbf{y} \|$ will usually give us a value, $\mathbf{x}_d^*$, close to the desired optimum $\mathbf{x}^*$:

$$\mathbf{x}_d^* = \operatorname*{argmin}_{\mathbf{x} \in X} \| \mathbf{c}(\mathbf{p}(\mathbf{x})) - \mathbf{y} \| . \tag{3.3}$$

This leads to the *dual space-mapping algorithm* as shown in Figure 4.

We will see in Section 3.3 that all three approaches coincide when $\mathbf{z}^* \in \mathbf{p}(X)$ and $\mathbf{p}$ is injective. In Section 5.1 the case where the space-mapping solutions are different will be illustrated. A notable difficulty of all the approaches is the expensive evaluation of $\mathbf{p}(\mathbf{x})$.

Figure 5: The ASM algorithm.

$$
\begin{array}{l}
\mathbf{x}_0 = \mathbf{z}^* = \text{argmin}_{\mathbf{z} \in Z} \, \| \, \mathbf{c}(\mathbf{z}) - \mathbf{y} \| \\
\mathbf{B}_0 = I \\
\textbf{for} \quad k = 0, 1, \ldots \\
\textbf{while} \quad \| \, \mathbf{p}(\mathbf{x}_k) - \mathbf{z}^* \| \, > \text{tolerance} \\
\textbf{do}^a \; \mathbf{h}_k = -\mathbf{B}_k^\dagger (\mathbf{p}(\mathbf{x}_k) - \mathbf{z}^*) \\
\qquad \mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{h}_k \\
\qquad \mathbf{B}_{k+1} = \mathbf{B}_k + \frac{(\mathbf{p}(\mathbf{x}_{k+1}) - \mathbf{z}^*)\mathbf{h}^T}{\mathbf{h}^T \mathbf{h}} \\
\textbf{enddo}
\end{array}
$$

$^a\mathbf{B}_k^\dagger$ denotes the pseudo-inverse of $\mathbf{B}_k$.

One way to deal with this problem is via convenient choices for the approximations to $\mathbf{p}(\mathbf{x})$ indicated in each of the generic schemes given in Figures 2–4. The selection of a suitable approximation is made in the algorithms below.

3.2 Space-mapping algorithms

The three algorithms given in the Figures 2–4 are vague in the sense that no choice has been made yet for the approximation of the space-mapping function $\mathbf{p}$. Using a linear approximation gives rise to the more popular space-mapping optimization algorithms. The two most representative examples are shown below. An extensive survey of available algorithms can be found in [6].

The ASM algorithm [5]. The space-mapping function is approximated by linearisation to obtain

$$
\mathbf{p}_k(\mathbf{x}) = \mathbf{p}(\mathbf{x}_k) + \mathbf{B}_k (\mathbf{x} - \mathbf{x}_k).
$$

In each space-mapping iteration step the matrix $\mathbf{B}_k$ is adapted by a rank-one update. For that purpose a Broyden-type approximation [14] for the Jacobian of the space-mapping function $\mathbf{p}(\mathbf{x})$ is used,

$$
\mathbf{B}_{k+1} = \mathbf{B}_k + \frac{\mathbf{p}(\mathbf{x}_{k+1}) - \mathbf{p}(\mathbf{x}_k) - \mathbf{B}_k \mathbf{h}}{\mathbf{h}^T \mathbf{h}} \, \mathbf{h}^T ,
$$

where $\mathbf{h} = \mathbf{x}_{k+1} - \mathbf{x}_k$. This is combined with original space mapping, so that $\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{B}_k^{-1}(\mathbf{p}(\mathbf{x}_k) - \mathbf{z}^*)$ and the *aggressive space mapping* (*ASM*) is obtained (Figure 5). Note that Broyden's method cannot assure that every $\mathbf{B}_k$ is nonsingular, so, the inverse in the original space-mapping scheme should be read as a pseudo-inverse. ASM just solves equation (3.2) by a quasi-Newton iteration with an approximate Jacobian. We notice that only one evaluation of the fine model is needed per iteration.

The TRASM algorithm [3]. Like ASM, the *trust region aggressive space mapping* (TRASM) algorithm solves (3.2), now taking into account a trust region. I.e., to stabilize the algorithm, specially in its initial stage, every iterant $\mathbf{x}_{k+1}$ is found in a region not too far away from

Figure 6: The TRASM algorithm.

$$\mathbf{x}_0 = \mathbf{z}^* = \text{argmin}_{\mathbf{z} \in Z} \, \| \mathbf{c}(\mathbf{z}) - \mathbf{y} \|$$
$$\mathbf{B}_0 = \mathbf{I}, \ \lambda_0 = 0$$
$$\mathbf{for} \quad k = 0, 1, \dots$$
$$\mathbf{while} \quad \| \mathbf{p}(\mathbf{x}_k) - \mathbf{z}^* \| > \text{tolerance}$$
$$\mathbf{do} \ \mathbf{h}_k = -(\mathbf{B}_k^T \mathbf{B}_k + \lambda_k \mathbf{I})^{-1} \mathbf{B}_k^T (\mathbf{p}(\mathbf{x}_k) - \mathbf{z}^*) \ ;$$
$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{h}_k \ ;$$
$$\mathbf{B}_{k+1} = \mathbf{B}_k + \frac{(\mathbf{p}(\mathbf{x}_{k+1}) - \mathbf{z}^*)\mathbf{h}^T}{\mathbf{h}^T \mathbf{h}} \ ;$$
update λ_k depending on the local linearity assumption
and the trust region required
$$\mathbf{enddo}$$

the earlier one. Further, making use of the Levenberg-Marquardt algorithm [26], TRASM restricts itself to minimizing the residue in the Euclidean norm. By Levenberg-Marquardt the inner loop of TRASM uses the update formula

$$(\mathbf{B}_k^T \mathbf{B}_k + \lambda_k \mathbf{I}) \, \mathbf{h}_k = -\mathbf{B}_k^T (\mathbf{p}(\mathbf{x}_k) - \mathbf{z}^\star) \ ,$$
$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{h}_k \ ,$$

where $\mathbf{B}_k$ is the kth approximation of the Jacobian of $\mathbf{p}(\mathbf{x})$ at $\mathbf{x}_k$. Notice that the adaptation of the trust-region is controlled by adapting λ_k. For $\lambda_k = 0$ the iteration step reduces to a quasi-Newton one, whereas for larger λ_k the step-length reduces and the method tends to steepest descent for minimizing $\| \mathbf{p}(\mathbf{x}) - \mathbf{x}^* \|$. Generally $\mathbf{B}_k$ is constructed by rank-one updates as in ASM. The TRASM algorithm is shown in Figure 6.

The HASM algorithm [1]. In general the space-mapping function $\mathbf{p}$ will not be perfect, and hence, a space mapping based algorithm will *not* yield the solution of the fine model optimization. Therefore space mapping is sometimes combined with a classical optimization method. This can be done in several ways ways. One obvious way is using a space-mapping solution as an initial guess for an arbitrary other (classical) optimization algorithm. Another approach is by constructing a combination of a space-mapping and a classical optimization method. Such an algorithm is called *hybrid aggressive space mapping* (HASM). It exploits space mapping when effective, otherwise it defaults to the linearizion of the fine model response $\mathbf{f}(\mathbf{x})$

$$\mathbf{l}_k(\mathbf{x}) = \mathbf{f}(\mathbf{x}_k) + \widehat{\mathbf{B}}_k \, (\mathbf{x} - \mathbf{x}_k) \ ,$$

where $\widehat{\mathbf{B}}_k$ is a Broyden rank one approximation to the Jacobian of $\mathbf{f}(\mathbf{x})$ in $\mathbf{x}_k$. It is proved in [24] that (under mild conditions) the iterative process

$$\mathbf{x}_{k+1} = \underset{\mathbf{x} \in X}{\text{argmin}} \, \| \mathbf{l}_k(\mathbf{x}) - \mathbf{y} \|$$

converges to $\mathbf{x}^*$.

The *combined model* $\mathbf{v}_k(\mathbf{x})$ is introduced as a convex combination of the mapped coarse and the linearized model:

$$\mathbf{v}_k(\mathbf{x}) = \omega_k \, \mathbf{c}(\mathbf{p}_k(\mathbf{x})) + (1 - \omega_k) \, \mathbf{l}_k(\mathbf{x}) \tag{3.4}$$

Figure 7: A hybrid aggressive space-mapping (HASM) algorithm.

$\mathbf{x}_0 = \mathbf{z}^* = \operatorname{argmin}_{\mathbf{z} \in Z} \| \mathbf{c}(\mathbf{z}) - \mathbf{y} \|$;
$\mathbf{p}_0 = I$;
$\mathbf{B}_0 = \mathbf{dc}/\mathbf{dz}\,(\mathbf{z}^*)$;
for $k = 0, 1, \ldots$
do select ω_k
$\qquad \mathbf{x}_{k+1} = \operatorname{argmin}_{\mathbf{x} \in X} \| \omega_k \, \mathbf{c}(\mathbf{p}_k(\mathbf{x})) + (1 - \omega_k)(\mathbf{f}(\mathbf{x}_k) + \widehat{\mathbf{B}}_k(\mathbf{x} - \mathbf{x}_k)) - \mathbf{y} \|$
$\qquad \mathbf{z}_{k+1} = \mathbf{p}(\mathbf{x}_{k+1})$
$\qquad$ compute updated $\mathbf{p}_{k+1}$ by means of $\{(\mathbf{x}_j, \mathbf{z}_j)\}_{j=0}^k$
$\qquad$ compute updated $\widehat{\mathbf{B}}_{k+1}$ by means of $\{(\mathbf{x}_j, \mathbf{z}_j)\}_{j=0}^k$

enddo

with $\omega_k \in [0, 1]$. Often ω_k is used as a simple switch: either $\omega_k = 1$ or $\omega_k = 0$. Now $\| \mathbf{v}_k(\mathbf{x}) - \mathbf{y} \|$ is the functional to be minimized in the HASM scheme (Figure 7). Conditions for convergence are given in [25]. It should be noted that the combined model $\mathbf{v}_k(\mathbf{x})$ is only slightly more expensive than the mapped coarse one $\mathbf{c}(\mathbf{p}_k(\mathbf{x}))$ since for both the same number of fine model evaluations is used.

3.3 Perfect mapping, flexibility and reachability
By its definition the perfect mapping relates the similarity of the models and the specifications. If the fine model allows for a *reachable design*, then it is immediate that, independent of the coarse model used, the mapping is always perfect. Also if the coarse and the fine model optimal responses are identical, the space-mapping function is perfect. These two facts are summarized in the following lemma.

Lemma 1. If either a design is reachable for the fine model, or if the fine and the coarse model give the same optimal response, then the corresponding space-mapping function is perfect. In formula: (i) If $\mathbf{f}(\mathbf{x}^*) = \mathbf{y}$ then $\mathbf{p}(\mathbf{x}^*) = \mathbf{z}^*$. (ii) If $\mathbf{f}(\mathbf{x}^*) = \mathbf{c}(\mathbf{z}^*)$ then $\mathbf{p}(\mathbf{x}^*) = \mathbf{z}^*$.
$\square$

The fine and the coarse model in space-mapping should, by their nature, show important similarities. Normally the coarse model needs only a small correction in order to be aligned with the fine one. Then the space-mapping function is a just a small perturbation of the identity. Therefore we can expect that, with the tolerances handled in practice, there exists a point $\widetilde{\mathbf{x}} \in X$ such that $\mathbf{f}(\widetilde{\mathbf{x}}) = \mathbf{c}(\mathbf{z}^*)$ and thus, that $\mathbf{p}(\widetilde{\mathbf{x}}) = \mathbf{z}^*$. This is a most interesting situation in which many of the different situations discussed above coincide and space mapping may lead us to the desired optimum. This is shown in the following lemma.

Lemma 2. (i) If $\mathbf{z}^* \in \mathbf{p}(X)$, then $\mathbf{p}(\mathbf{x}_{sm}^*) = \mathbf{p}(\mathbf{x}_p^*) = \mathbf{p}(\mathbf{x}_d^*) = \mathbf{z}^*$;
(ii) If, in addition, $\mathbf{p}$ is an injective perfect mapping then $\mathbf{x}^* = \mathbf{x}_{sm}^* = \mathbf{x}_p^* = \mathbf{x}_d^*$.

Proof. In case $\mathbf{z}^* \in \mathbf{p}(X)$ by definition (3.1) $\mathbf{x}_{sm}^*$ exists and $\mathbf{p}(\mathbf{x}_{sm}^*) = \mathbf{z}^*$. Now by definition (3.2) $\mathbf{p}(\mathbf{x}_p^*) = \mathbf{p}(\mathbf{x}_{sm}^*)$. Further, $\mathbf{c}(\mathbf{p}(\mathbf{x}_{sm}^*)) - \mathbf{y} = \mathbf{c}(\mathbf{z}^*) - \mathbf{y}$, so that by definition (3.3) $\mathbf{p}(\mathbf{x}_d^*) = \mathbf{p}(\mathbf{x}_{sm}^*)$. Summarizing, $\mathbf{p}(\mathbf{x}_{sm}^*) = \mathbf{p}(\mathbf{x}_p^*) = \mathbf{p}(\mathbf{x}_d^*) = \mathbf{z}^*$. If, further, $\mathbf{p}$ is injective,

the original, primal and dual approach yield the same result. If, in addition, $\mathbf{p}$ is a perfect mapping then $\mathbf{x}_{sm}^* = \mathbf{x}_p^* = \mathbf{x}_d^* = \mathbf{x}^*$. $\qquad\square$

In some cases we can expect that the sets of fine and coarse reachable aims may overlap in a region of $\mathbb{R}^m$ close to their respective optima. The concept of model flexibility is now introduced and from that notion some results concerning properties of the space-mapping functions can be derived.

Definition 4. A model is called more *flexible* than another if the set of its reachable aims contains the set of reachable aims of the other. Two models are *equally flexible* if their respective sets of reachable aims coincide. $\qquad\square$

Thus, a coarse model $\mathbf{c}$ is more flexible than the fine one $\mathbf{f}$ if $\mathbf{c}(Z) \supset \mathbf{f}(X)$, i.e., if the coarse model response can reproduce all the fine model reachable aims. E.g., if $n \geq m$, linearisation makes a flexible coarse model. Similarly the fine model is more flexible if $\mathbf{f}(X) \supset \mathbf{c}(Z)$. Model flexibility is closely related to properties of the space-mapping function. This is shown in the following lemmas, where $\mathbf{p}$ denotes the space-mapping function.

Lemma 3. If $\mathbf{c}$ is more flexible than $\mathbf{f}$ then
(i) $\mathbf{c}(\mathbf{p}(\mathbf{x})) = \mathbf{f}(\mathbf{x}) \quad \forall \mathbf{x} \in X$;
(ii) $\mathbf{p} : X \to Z$ is a perfect mapping $\Leftrightarrow \mathbf{c}(\mathbf{z}^*) = \mathbf{f}(\mathbf{x}^*)$;
(iii) if $\mathbf{f} : X \to Y$ is injective then $\mathbf{p} : X \to Z$ is injective;
(iv) if $\mathbf{c}(Z) \setminus \mathbf{f}(X) \neq \emptyset$, then $\mathbf{p} : X \to Z$ cannot be surjective.

Proof. (i) $\forall \mathbf{f}(\mathbf{x}) \exists \mathbf{c}(\mathbf{z}) : \mathbf{c}(\mathbf{z}) = \mathbf{f}(\mathbf{x}) \Rightarrow \mathbf{c}(\mathbf{p}(\mathbf{x})) = \mathbf{f}(\mathbf{x}) \quad \forall \mathbf{x} \in X$.
(ii$\Rightarrow$) $\mathbf{f}(\mathbf{x}^*) = \mathbf{c}(\mathbf{p}(\mathbf{x}^*)) = \mathbf{c}(\mathbf{z}^*)$ because of (i) and perfect mapping.
(ii$\Leftarrow$) $\mathbf{p}(\mathbf{x}^*) = \mathrm{argmin}_{\mathbf{z} \in Z} \ \|\mathbf{c}(\mathbf{z}) - \mathbf{f}(\mathbf{x}^*)\| = \mathrm{argmin}_{\mathbf{z} \in Z} \ \|\mathbf{c}(\mathbf{z}) - \mathbf{c}(\mathbf{z}^*)\| = \mathbf{z}^*$ because the optimum is unique.
(iii) Let $\mathbf{p}(\mathbf{x}_1) = \mathbf{p}(\mathbf{x}_2) \Rightarrow \mathbf{c}(\mathbf{p}(\mathbf{x}_1)) = \mathbf{c}(\mathbf{p}(\mathbf{x}_2)) \Rightarrow \mathbf{f}(\mathbf{x}_1) = \mathbf{f}(\mathbf{x}_2)$ because of (i) $\Rightarrow \mathbf{x}_1 = \mathbf{x}_2$ because $\mathbf{f}$ is injective.
(iv) Let be $\widetilde{\mathbf{z}} : \mathbf{c}(\widetilde{\mathbf{z}}) \in \mathbf{c}(Z) \setminus \mathbf{f}(X)$. But then, because of (i), we see that it is not possible to find $\widetilde{\mathbf{x}} \in X : \mathbf{p}(\widetilde{\mathbf{x}}) = \widetilde{\mathbf{z}}$. $\qquad\square$

Remark. Because of (ii) generally we cannot expect space-mapping functions to be perfect for flexible coarse models unless the two models are equally flexible near the optimum. This fact is illustrated in Figure 8. However, we remind that if the design is reachable, the perfect mapping property holds, even if $\mathbf{c}(Z) \setminus \mathbf{f}(X) \neq \emptyset$.

Lemma 4. If $\mathbf{f}$ is more flexible than $\mathbf{c}$ then
(i) $\mathbf{p} : X \to Z$ is surjective;
(ii) if $\mathbf{f}(X) \setminus \mathbf{c}(Z) \neq \emptyset$, then $\mathbf{p}$ cannot be injective.

Proof. (i) $\mathbf{f}(X) \supset \mathbf{c}(Z) \Rightarrow \forall \mathbf{c}(\mathbf{z}) \exists \mathbf{f}(\mathbf{x}) : \mathbf{c}(\mathbf{z}) = \mathbf{f}(\mathbf{x})$
$\quad \Rightarrow \forall \mathbf{z} \exists \mathbf{x} : \mathbf{z} = \mathrm{argmin}_{\zeta \in Z} \ \|\mathbf{c}(\zeta) - \mathbf{f}(\mathbf{x})\| \Rightarrow \forall \mathbf{z} \exists \mathbf{x} : \mathbf{z} = \mathbf{p}(\mathbf{x})$.
(ii) All elements $\mathbf{f}(\mathbf{x}) \in \mathbf{f}(X)$ can be divided over equivalence classes for which $\mathbf{c}(\mathbf{z})$ is the closest element in $\mathbf{c}(Z)$, i.e., classes
$\mathbf{f_z} = \{\mathbf{f}(\mathbf{x}) \mid \mathrm{argmin}_{\zeta \in Z} \ \|\mathbf{c}(\zeta) - \mathbf{f}(\mathbf{x})\| = \mathbf{z}\}$ with $\mathbf{z} \in Z$. All $\mathbf{x}$ with $\mathbf{f}(\mathbf{x}) \in \mathcal{F}_\mathbf{z}$ now share the property that $\mathbf{p}(\mathbf{x}) = \mathbf{z}$. If $\mathbf{f}(\mathbf{x}) \in \mathbf{f}(X) \setminus \mathbf{c}(Z)$ it follows that $\|\mathbf{c}(\mathbf{p}(\mathbf{x})) - \mathbf{f}(\mathbf{x})\| \neq 0$ and hence $\mathbf{f_{p(x)}}$ contains at least two elements. $\qquad\square$

Figure 8: Perfect mapping cannot be expected in general for more flexible coarse models.
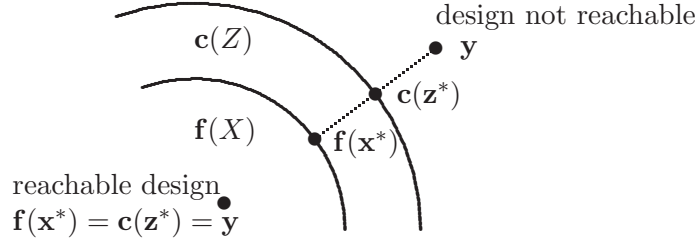
We combine the previous two lemmas in the following.

Lemma 5. If $\mathbf{f}$ and $\mathbf{c}$ are equally flexible and $\mathbf{f} : X \to Y$ is injective, then (i) $\mathbf{p}$ is a bijection, and (ii) $\mathbf{p}$ is a perfect mapping.

Proof. (i) Combine lemmas 3 and 4.
(ii) Because $\mathbf{f}(X) \equiv \mathbf{c}(Z)$, the optimum verifies $\mathbf{f}(\mathbf{x}^*) = \mathbf{c}(\mathbf{z}^*)$. Since the coarse model is more flexible than the fine one we conclude that $\mathbf{p}$ is a perfect mapping. □

The conclusions s in Lemma 2 can now be derived from assumptions about model flexibility.

Lemma 6. (i) If $\mathbf{f}$ is more flexible than $\mathbf{c}$, then $\mathbf{p}(\mathbf{x}^*_{sm}) = \mathbf{p}(\mathbf{x}^*_p) = \mathbf{p}(\mathbf{x}^*_d) = \mathbf{z}^*$. (ii) If $\mathbf{f}$ and $\mathbf{c}$ are equally flexible and $\mathbf{f}$ is injective, then $\mathbf{x}^* = \mathbf{x}^*_{sm} = \mathbf{x}^*_p = \mathbf{x}^*_d$. □

Remark. It is not really needed for the space-mapping function to be a bijection over the whole domain in which it is defined. In fact, perfect mapping is a property that concerns only a point, and it is enough if the function is injective in a (small) neighborhood. Thus the assumptions for the former lemmas can be relaxed and stated just locally.

4. DEFECT CORRECTION AND SPACE MAPPING
The main idea underlying the space-mapping technique, i.e. the efficient solution of a complex problem by the iterative use of a simpler one, is known since long in computational mathematics. Generally it is known as defect correction iteration (of which, e.g., Newton's method, relaxation procedures [13, 20], iterative refinement [13] and multigrid methods [20] are examples). The general principles of the idea are well known and studied in a number of papers [12, 13, 31]. In this section we first briefly summarize the defect correction principle for the solution of operator equations and extend the idea to optimization problems. We compare defect correction with space mapping and show how space mapping can be derived from defect correction. Then, in the general framework of defect correction new space-mapping algorithms are constructed and analyzed.

4.1 The defect correction principle
Let us first consider the problem of solving a nonlinear operator equation

$$\mathcal{F}\mathbf{x} = \mathbf{y}, \tag{4.1}$$

where $\mathcal{F} : D \subset E \to \widehat{D} \subset \widehat{E}$ is a continuous, generally nonlinear operator and E and $\widehat{E}$ are Banach spaces. In general, neither injectivity nor surjectivity of the mapping are assumed, but in many cases these properties can be achieved by a proper choice of the subsets D and $\widehat{D}$.

The classical *defect correction* iteration for the solution of the equation (4.1) with $\mathbf{y} \in \widehat{D}$ is based on a sequence of operators $\widetilde{\mathcal{F}}_k : D \to \widehat{D}$ approximating $\mathcal{F}$. We assume that each $\widetilde{\mathcal{F}}_k$ has an easy-to-calculate inverse $\widetilde{\mathcal{G}}_k : \widehat{D} \to D$. Actually, it is the existence of the easy-to-evaluate operator $\widetilde{\mathcal{G}}_k$, rather than the existence of $\widetilde{\mathcal{F}}_k$, that is needed for defect correction, and we do not need to assume neither $\widetilde{\mathcal{F}}_k$ nor $\widetilde{\mathcal{G}}_k$ to be invertible.

Defect correction comes in two brands [12], depending on the space, E or $\widehat{E}$, in which linear combinations for extrapolation are made. The two basic iterative defect correction procedures to generate a (hopefully convergent) sequence of approximations to the solution of (4.1) are

$$
\begin{cases}
\mathbf{x}_0 & = \quad \widetilde{\mathcal{G}}_0 \, \mathbf{y} \,, \\
\mathbf{x}_{k+1} & = \quad (I - \widetilde{\mathcal{G}}_{k+1} \, \mathcal{F}) \, \mathbf{x}_k + \widetilde{\mathcal{G}}_{k+1} \, \mathbf{y},
\end{cases}
\tag{4.2}
$$

and

$$
\begin{cases}
\mathbf{l}_0 & = \quad \mathbf{y} \,, \\
\mathbf{l}_{k+1} & = \quad (I - \mathcal{F} \, \widetilde{\mathcal{G}}_k) \, \mathbf{l}_k + \mathbf{y}.
\end{cases}
\tag{4.3}
$$

In the latter we identify the approximate solution as $\mathbf{x}_k \equiv \widetilde{\mathcal{G}}_k \mathbf{l}_k$. We see that the two iteration processes are dual in the sense that the extrapolation in (4.2) is in the space E, whereas the additions in (4.3) are in $\widehat{E}$. If $\widetilde{\mathcal{G}}_k$ is injective, then an operator $\widetilde{\mathcal{F}}_k$ exists such that $\widetilde{\mathcal{F}}_k \widetilde{\mathcal{G}}_k = \mathcal{I}_{\widehat{D}}$, i.e., $\widetilde{\mathcal{F}}_k$ is the left-inverse of $\widetilde{\mathcal{G}}_k$. Then $\widetilde{\mathcal{F}}_k \mathbf{x}_k = \mathbf{l}_k$ and (4.3) is equivalent with the iterative procedure

$$
\begin{cases}
\widetilde{\mathcal{F}}_0 \mathbf{x}_0 & = \quad \mathbf{y} \,, \\
\widetilde{\mathcal{F}}_{k+1} \mathbf{x}_{k+1} & = \quad \widetilde{\mathcal{F}}_k \, \mathbf{x}_k - \mathcal{F} \widetilde{\mathcal{G}}_k \widetilde{\mathcal{F}}_k \, \mathbf{x}_k + \mathbf{y} \,.
\end{cases}
\tag{4.4}
$$

In order to apply (4.4), the injectivity of $\widetilde{\mathcal{G}}_k$ is not really needed and it is immediately seen that neither (4.3) nor (4.4) converge if $\mathbf{y} \notin \widehat{D}$. However, (4.4) can be modified so that it can be used for $\mathbf{y} \notin \widehat{D}$. Therefore, we require injectivity for $\widetilde{\mathcal{F}}_k$ and take $\widetilde{\mathcal{G}}_k$ its left-inverse, i.e., $\widetilde{\mathcal{G}}_k \widetilde{\mathcal{F}}_k = \mathcal{I}_D$. Then (4.4) leads to

$$
\begin{cases}
\mathbf{x}_0 & = \quad \widetilde{\mathcal{G}}_0 \, \mathbf{y} \,, \\
\mathbf{x}_{k+1} & = \quad \widetilde{\mathcal{G}}_{k+1} \left(\widetilde{\mathcal{F}}_k \, \mathbf{x}_k - \mathcal{F} \, \mathbf{x}_k + \mathbf{y} \right) \,.
\end{cases}
\tag{4.5}
$$

Because (4.5) allows for a non-injective $\widetilde{\mathcal{G}}_k$ this procedure can be used with optimization purposes. In case of an invertible $\widetilde{\mathcal{G}}$ both (4.4) and (4.5) are equivalent with (4.3).

Assuming that for large enough k the approximate operators $\widetilde{\mathcal{G}}_k = \widetilde{\mathcal{G}}$ and $\widetilde{\mathcal{F}}_k = \widetilde{\mathcal{F}}$ do not change anymore, upon convergence we see that with $\lim_{k\to\infty} \mathbf{x}_k = \overline{\mathbf{x}}$ and $\lim_{k\to\infty} \mathbf{l}_k = \overline{\mathbf{l}}$ we find for (4.2), (4.3), (4.4) and (4.5) respectively

$$
\widetilde{\mathcal{G}} \, \mathcal{F} \, \overline{\mathbf{x}} = \widetilde{\mathcal{G}} \, \mathbf{y} \,,
\tag{4.6}
$$

$$\mathcal{F} \widetilde{\mathcal{G}} \bar{\mathbf{1}} = \mathcal{F} \bar{\mathbf{x}} = \mathbf{y} \,, \tag{4.7}$$

$$\mathcal{F} \widetilde{\mathcal{G}} \widetilde{\mathcal{F}} \bar{\mathbf{x}} = \mathbf{y} \,, \qquad \text{and} \tag{4.8}$$

$$\bar{\mathbf{x}} = \widetilde{\mathcal{G}} \left(\widetilde{\mathcal{F}} \bar{\mathbf{x}} - \mathcal{F} \bar{\mathbf{x}} + \mathbf{y} \right) . \tag{4.9}$$

The process (4.2) is convergent if the operator $\mathcal{I} - \widetilde{\mathcal{G}}\mathcal{F} : D \to D$ is a contraction, i.e., if the Lipschitz constant $\|\mathcal{I} - \widetilde{\mathcal{G}}\mathcal{F}\|_{D \subset E}$ is less than one. This implies that $\mathcal{F}$ should be injective. Similarly (4.3) converges if $\mathcal{I} - \mathcal{F}\widetilde{\mathcal{G}} : \widehat{D} \to \widehat{D}$ is a contraction, which implies that $\mathcal{F}$ should be surjective. When $\mathcal{F}$ is not surjective, (4.7) shows that (4.3) and (4.4) do not always allow for a fixed point. However, the processes (4.2) and (4.5) may allow for one, provided $\widetilde{\mathcal{G}}$ is not injective. Further we see that, for constant, non-singular affine $\widetilde{\mathcal{G}} = \widetilde{\mathcal{G}}_k$, all processes (4.2)–(4.5) yield identical sequences of approximants. If a regular $\widetilde{\mathcal{G}}$ is not affine, the processes (4.2) and (4.3) give different results.

For our optimization problems, where the design may be not reachable, $\mathbf{y} \in \widehat{D}$, but $\mathbf{y} \notin \mathcal{F}(D)$, i.e., $\mathcal{F}$ is no surjection so that no solution for (4.1) exists and (4.3)-(4.4) cannot converge. Therefore, we drop the idea of finding an $x \in D$ satisfying (4.1) and we replace the aim by looking for a solution $\mathbf{x}^* \in D$ so that the distance between $\mathcal{F}\mathbf{x}$ and $\mathbf{y}$ is minimal, i.e., we want to find

$$\mathbf{x}^* = \operatorname{argmin}_{\mathbf{x} \in D} \| \mathcal{F}\mathbf{x} - \mathbf{y} \|_{\widehat{E}} \,.$$

For a compact non-empty D and a continuous $\mathcal{F}$, at least a solution exists. and if the operators $\widetilde{\mathcal{G}}_k$ are such that (4.2) or (4.5) converges, the stationary point $\bar{\mathbf{x}}$ satisfies (4.6) or (4.9) respectively. For a linear operator $\widetilde{\mathcal{G}}$ this means that $\mathcal{F}\bar{\mathbf{x}} - \mathbf{y} \in \operatorname{Ker}(\widetilde{\mathcal{G}})$.

4.2 Defect correction and space mapping
Because the theory of defect correction is formulated in terms of Banach spaces, to compare it with the space-mapping theory we restrict ourselves to cost functions $\| \cdot \|$ that take the form of a norm. In fact, for our space-mapping algorithms we simply take $E = \mathbb{R}^n$, $\widehat{E} = \mathbb{R}^m$ and $D = X \subset \mathbb{R}^n$ a compact (closed and bounded) subset. It is tempting to identify the operator $\mathcal{F}$ with the fine model function $\mathbf{f}$, i.e.,

$$\mathcal{F}\mathbf{x} = \mathbf{f}(\mathbf{x}) \,, \qquad \text{with } \mathbf{x} \in D = X \subset \mathbb{R}^n \,, \tag{4.10}$$

but, to be more precise, we denote by $\mathcal{F}^\dagger$ the operator for the minimization problem

$$\mathcal{F}^\dagger \mathbf{y} = \operatorname*{argmin}_{\mathbf{x} \in X} \| \mathbf{f}(\mathbf{x}) - \mathbf{y} \| \,. \tag{4.11}$$

Now the operator $\mathcal{F} : D \to \widehat{D}$ in (4.10) is the right-inverse of $\mathcal{F}^\dagger$. We stress that $\mathcal{F}$ generally is not surjective. Defect correction for operator equations is extended to optimization by approximating the operator $\mathcal{F}^\dagger$.

The space mapping idea comes back by identifying the role of the approximate pseudo-inverse. Since the mapped coarse model acts as a surrogate for the fine, one clear choice of the approximate operator can be

$$\widetilde{\mathcal{F}}_k \mathbf{x} = \mathbf{c}(\mathbf{p}_k(\mathbf{x})) \quad \forall k \,. \tag{4.12}$$

Then $\widetilde{\mathcal{G}}_k$, the approximate inverse of $\mathcal{F}$ is the the mapped coarse model optimization

$$\widetilde{\mathcal{G}}_k \, \mathbf{y} = \operatorname*{argmin}_{\mathbf{x} \in X} \| \, \mathbf{c}(\mathbf{p}_k(\mathbf{x})) - \mathbf{y} \| \, . \tag{4.13}$$

At this stage we make no particular choice for the functions $\mathbf{p}_k$, but we only assume that (4.13) is easily evaluated. For example, if the spaces X and Z can be identified, we simply can take $\mathbf{p}_k = I$, the identity operator. We can try and find better choices for $\mathbf{p}_k$, but –as we assumed that both X and Z are in $\mathbb{R}^n$– only bijections for the functions $\mathbf{p}_k$ are considered. As for $\widetilde{\mathcal{G}}_k$, we also assume for $\mathbf{p}_k$ that they do not to change anymore for k large enough, i.e., $\mathbf{p}_k = \overline{\mathbf{p}}$ for $k \geq k_0$.

4.3 Defect correction algorithms for optimization

Construction of new iterative optimization algorithms   Using the basic defect correction processes and the relations (4.10), (4.12) and (4.13) that are associated with space mapping, we derive two new defect-correction iteration schemes that can be used for optimization. Substitution of the relations into the processes (4.2) and (4.5) respectively yields the following initial estimate and iteration processes for $k = 0, 1, 2, \cdots$

$$\mathbf{x}_0 \;=\; \operatorname{argmin}_{\mathbf{x} \in X} \| \, \mathbf{c}(\mathbf{p}_0(\mathbf{x})) - \mathbf{y} \| \, , \tag{4.14}$$

$$(\text{DeC-X}) \quad \mathbf{x}_{k+1} \;=\; \mathbf{x}_k - \operatorname{argmin}_{\mathbf{x} \in X} \| \, \mathbf{c}(\mathbf{p}_{k+1}(\mathbf{x})) - \mathbf{f}(\mathbf{x}_k) \|$$
$$+ \operatorname{argmin}_{\mathbf{x} \in X} \| \, \mathbf{c}(\mathbf{p}_{k+1}(\mathbf{x})) - \mathbf{y} \| \, , \tag{4.15}$$

$$(\text{DeC-Y}) \quad \mathbf{x}_{k+1} \;=\; \operatorname{argmin}_{\mathbf{x} \in X} \| \, \mathbf{c}(\mathbf{p}_{k+1}(\mathbf{x})) - \mathbf{c}(\mathbf{p}_k(\mathbf{x}_k)) + \mathbf{f}(\mathbf{x}_k) - \mathbf{y} \| \, . \tag{4.16}$$

Again, the two processes (4.15) and (4.16) are dual in the sense that extrapolation is applied in the space X for process (4.15) and extrapolation in Y for process (4.16). In the case when the spaces X and Z can be identified, for the initial estimate (4.14) we take $\mathbf{p}_0 = I$, the identity, like in the space-mapping algorithms.

In the above iterations all iterated minimizations are over functions involving the surrogate model, $\mathbf{c} \circ \mathbf{p}_k$. However, it is the coarse model that was assumed to be cheaply optimized. Therefore, it is convenient to derive procedures, similar to the ones given, such that direct optimization over the coarse model becomes central. By taking $\mathcal{F} \, \mathbf{z} = \mathbf{f}(\mathbf{q}(\mathbf{z}))$, $\widetilde{\mathcal{F}}_k \, \mathbf{z} = \mathbf{c}(\mathbf{z})$ and $\widetilde{\mathcal{G}}_k \, \mathbf{y} = \operatorname{argmin}_{\mathbf{z} \in Z} \| \, \mathbf{c}(\mathbf{z}) - \mathbf{y} \|$, with $\mathbf{q}$ and $\mathbf{q}_k$ bijections from Z to X fulfilling in every iteration $\mathbf{q} \, \mathbf{z}_k = \mathbf{q}_k \, \mathbf{z}_k$ we obtain for $k = 0, 1, 2, \cdots$

$$\mathbf{z}_0 \;=\; \mathbf{z}^* = \operatorname{argmin}_{\mathbf{z} \in Z} \| \, \mathbf{c}(\mathbf{z}) - \mathbf{y} \| \, , \tag{4.17}$$

$$(\text{DeC-Z}) \quad \mathbf{z}_{k+1} \;=\; \mathbf{z}_k - \operatorname{argmin}_{\mathbf{z} \in Z} \| \, \mathbf{c}(\mathbf{z}) - \mathbf{f}(\mathbf{q}_k(\mathbf{z}_k)) \| + \mathbf{z}^* \, , \tag{4.18}$$

$$(\text{DeC-Y}) \quad \mathbf{z}_{k+1} \;=\; \operatorname{argmin}_{\mathbf{z} \in Z} \| \, \mathbf{c}(\mathbf{z}) - \mathbf{c}(\mathbf{z}_k) + \mathbf{f}(\mathbf{q}_k(\mathbf{z}_k)) - \mathbf{y} \| \, . \tag{4.19}$$

As the solution is wanted in terms of fine-model control variables, the procedures are complemented with $\mathbf{x}_k = \mathbf{q}_k(\mathbf{z}_k)$. The bijections can be interpreted as $\mathbf{q}_k = \mathbf{p}_k^{-1}$. Again, for $k > k_0$, we assume the iteration process to be stationary: $\mathbf{q}_k = \overline{\mathbf{q}}$.

Lemma 7. (1) The initial estimates (4.14) and (4.17) are equivalent, as well as the processes (4.16) and (4.19). (2) If $\mathbf{p}_k$ is linear and the same for all k, $\mathbf{p}_k = \overline{\mathbf{p}}$, then also (4.15) and (4.18) are equivalent.

Proof. (1) We substitute in (4.14) $\mathbf{p}_0(\mathbf{x}_0) = \mathbf{z}_0$ to obtain (4.19). In (4.16) we substitute $\mathbf{p}_k(\mathbf{x}_k) = \mathbf{z}_k$, $\mathbf{p}_{k+1}(\mathbf{x}_{k+1}) = \mathbf{z}_{k+1}$ and $\mathbf{q}(\mathbf{z}_k) = \mathbf{x}_k$ to obtain (4.17). (2) The same substitutions in (4.15) make

$$
\begin{aligned}
\mathbf{q}_{k+1}\,\mathbf{z}_{k+1} \;=\; &\mathbf{q}_k\,\mathbf{z}_k \;\; - \; \mathbf{q}_{k+1}\operatorname{argmin}_{\mathbf{z}\in X}\;\|\,\mathbf{c}(\mathbf{x}) - \mathbf{f}(\mathbf{q}_k(\mathbf{z}_k))\,\| \\
&+ \; \mathbf{q}_{k+1}\operatorname{argmin}_{\mathbf{x}\in X}\;\|\,\mathbf{c}(\mathbf{x}) - \mathbf{y}\,\| \; .
\end{aligned}
$$

Now linearity of $\mathbf{q}_{k+1}$ and $\mathbf{p}_{k+1}\,\mathbf{q}_k\,\mathbf{z}_k = \mathbf{z}_k$ complete the proof. $\qquad\square$

Recovering space mapping from defect correction   With a proper choice of $\mathbf{p}_k$ we can recover the space-mapping algorithms presented in Section 3.1. The primal space-mapping generic scheme is obtained just by applying defect correction to the problem

$$
\mathcal{F}^\dagger\,\mathbf{z}^* = \operatorname{argmin}_{x\in X}\;\|\mathbf{p}(\mathbf{x}) - \mathbf{z}^*\| \, ,
$$

where the easy to calculate pseudo-inverses are

$$
\widetilde{\mathcal{G}}_k\,\mathbf{z}^* = \operatorname{argmin}_{x\in X}\;\|\mathbf{p}_k(\mathbf{x}) - \mathbf{z}^*\| \, ,
$$

where $\mathbf{p}_k$ is an arbitrary approximation of $\mathbf{p}$.

The dual space-mapping approach can be derived as follows. Let $\mathbf{p}_k$ be an approximation of $\mathbf{p}$ satisfying $\mathbf{p}_k(\mathbf{x}_k) = \mathbf{p}(\mathbf{x}_k)$, then (4.15) reduces to

$$
\mathbf{x}_{k+1} = \operatorname{argmin}_{\mathbf{x}\in X}\;\|\,\mathbf{c}(\mathbf{p}_k(\mathbf{x})) - \mathbf{y}\,\| \, , \tag{4.20}
$$

which coincides with the dual space-mapping algorithm.

A special case.   If $X = Z$ and $\mathbf{p}_k = I$, then (4.15) reduces to

$$
\left\{
\begin{aligned}
\mathbf{x}_0 \;\; &= \;\; \operatorname{argmin}_{\mathbf{z}\in Z}\;\|\,\mathbf{c}(\mathbf{z}) - \mathbf{y}\,\| = \mathbf{z}^* \, , \\
\mathbf{x}_{k+1} \;\; &= \;\; \mathbf{x}_k - \operatorname{argmin}_{\mathbf{z}\in Z}\;\|\,\mathbf{c}(\mathbf{z}) - \mathbf{f}(\mathbf{x}_k)\,\| + \mathbf{z}^* \, , \\
&= \;\; \mathbf{x}_k - \mathbf{p}(\mathbf{x}_k) + \mathbf{z}^* \, , \qquad k = 0, 1, \cdots .
\end{aligned}
\right. \tag{4.21}
$$

This is the iterative method to solve the original space-mapping problem (2.6) which has been published in [11, 27]. Clearly, the iteration can only converge if $\mathbf{p}(\mathbf{x}_{sm}^*) = \mathbf{z}^*$ allows for a solution and the convergence rate is determined by the Lipschitz constant $\|\,I - \mathbf{p}(\cdot)\,\|$. Hence, a necessary and sufficient condition for convergence of (4.21) is $\|\,I - \mathbf{p}(\cdot)\,\| < 1$.

Properties of the new algorithms

Convergence of DeC-X.   In the case of convergence of (4.15) we obtain, with fixed point $\lim_{k\to\infty}\mathbf{x}_k = \overline{\mathbf{x}}$, the equality

$$
\widehat{\mathbf{x}} := \operatorname{argmin}_{\mathbf{x}\in X}\;\|\,\mathbf{c}(\overline{\mathbf{p}}(\mathbf{x})) - \mathbf{f}(\overline{\mathbf{x}})\,\| = \operatorname{argmin}_{\mathbf{x}\in X}\;\|\,\mathbf{c}(\overline{\mathbf{p}}(\mathbf{x})) - \mathbf{y}\,\| \, , \tag{4.22}
$$

which shows that the design $\overline{\mathbf{x}} \in X$ is associated with the design result $\widehat{\mathbf{y}} = \mathbf{c}(\overline{\mathbf{p}}(\widehat{\mathbf{x}})) \in \mathbf{c}(\overline{\mathbf{p}}(X))$ that is the best approximation of $\mathbf{y}$ reachable in $\mathbf{c}(\overline{\mathbf{p}}(X)) = \mathbf{c}(Z)$, and at the same time such that $\mathbf{f}(\overline{\mathbf{x}})$ is best approximated in $\mathbf{c}(Z)$ by the same $\widehat{\mathbf{y}}$. Hence, both $\mathbf{f}(\overline{\mathbf{x}})$ and $\mathbf{y}$ belong to the

Figure 9: The relative location of $\mathbf{c}(\mathbf{z}^*)$, $\mathbf{f}(\mathbf{x}^*)$ and $\mathbf{f}(\overline{\mathbf{x}})$.
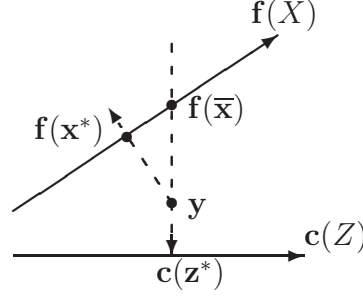
class $W_{\widehat{\mathbf{y}}} \subset \mathbb{R}^m$ of points in $\mathbb{R}^m$ for which $\widehat{\mathbf{y}}$ is the closest point in $\mathbf{c}(Z) \subset \mathbb{R}^m$. Noting that $\overline{\mathbf{p}}$ is a bijection, the equality (4.22) can be written briefly as

$$\widehat{\mathbf{x}} := \overline{\mathbf{p}}^{-1}(\mathbf{p}(\overline{\mathbf{x}})) = \overline{\mathbf{p}}^{-1}(\mathbf{z}^*) \tag{4.23}$$

where $\mathbf{p}$ denotes the space-mapping function (2.5).

If we assume $\widehat{E}$ to be a Hilbert space, from (4.22) we see that both $\mathbf{c}(\overline{\mathbf{p}}(\widehat{\mathbf{x}})) - \mathbf{f}(\overline{\mathbf{x}})$ and $\mathbf{c}(\overline{\mathbf{p}}(\widehat{\mathbf{x}})) - \mathbf{y}$ are perpendicular to the tangent plane in $\mathbf{c}(\overline{\mathbf{p}}(X))$ at $\mathbf{c}(\overline{\mathbf{p}}(\widehat{\mathbf{x}}))$. We denote the linear manifold of points perpendicular to the tangent plane in $\mathbf{c}(\overline{\mathbf{p}}(X))$ at $\mathbf{c}(\overline{\mathbf{p}}(\widehat{\mathbf{x}}))$ by $\mathbf{c}(\overline{\mathbf{p}}(X))^\perp(\widehat{\mathbf{x}})$. Thus

$$\begin{aligned}
\mathbf{c}(\overline{\mathbf{p}}(\widehat{\mathbf{x}})) - \mathbf{f}(\overline{\mathbf{x}}) &\in \mathbf{c}(\overline{\mathbf{p}}(X))^\perp(\widehat{\mathbf{x}}), \\
\mathbf{c}(\overline{\mathbf{p}}(\widehat{\mathbf{x}})) - \mathbf{y} &\in \mathbf{c}(\overline{\mathbf{p}}(X))^\perp(\widehat{\mathbf{x}}),
\end{aligned}$$

and hence

$$\mathbf{f}(\overline{\mathbf{x}}) - \mathbf{y} \in \mathbf{c}(\overline{\mathbf{p}}(X))^\perp(\widehat{\mathbf{x}}) = \mathbf{c}(\overline{\mathbf{p}}(X))^\perp(\overline{\mathbf{p}}^{-1}\mathbf{p}(\overline{\mathbf{x}})) = \mathbf{c}(Z)^\perp(\mathbf{z}^*).$$

This situation is shown in Figure 9. We summarize this result in the following lemma.

Lemma 8. Let $\widehat{E}$ be a Hilbert space. In case of convergence of (4.15) with a fixed point $\overline{\mathbf{x}}$ we obtain

$$\mathbf{f}(\overline{\mathbf{x}}) - \mathbf{y} \in \mathbf{c}(Z)^\perp(\mathbf{z}^*) = \mathbf{c}(\overline{\mathbf{p}}(X))^\perp(\overline{\mathbf{p}}^{-1}\mathbf{p}(\overline{\mathbf{x}})). \qquad \square$$

Because $\overline{\mathbf{p}}$ is a bijection, from (4.23) we directly see the following result.

Lemma 9. Let $\mathbf{p}(\mathbf{x})$ be the space-mapping function as defined in (2.5). If (4.15) converges to $\overline{\mathbf{x}}$ then $\mathbf{p}(\overline{\mathbf{x}}) = \mathbf{z}^*$. $\qquad \square$

Corollary. It follows from the definition of a *perfect mapping* that, with $\overline{\mathbf{x}}$ the convergence result of (4.15), $\mathbf{p}(\overline{\mathbf{x}}) = \mathbf{p}(\mathbf{x}^*)$ if and only if $\mathbf{p}$ is a perfect mapping

Remark. We note that convergence of (4.15) implies $\mathbf{z}^* \in \mathbf{p}(X)$. Hence in the case $\mathbf{z}^* \notin \mathbf{p}(X)$ the algorithm cannot reach a stationary point.

In case we have knowledge about the flexibility of the models we can go a bit further with the analysis and state the next lemma.

Lemma 10. If the coarse model is more flexible than the fine one, i.e., $\mathbf{c}(Z) \supset \mathbf{f}(X)$, and $\mathbf{y} \in \mathbf{c}(Z) \setminus \mathbf{f}(X)$, then (4.15) cannot converge.

Proof. Because $\mathbf{y} \in \mathbf{c}(\overline{\mathbf{p}}(X))$ there exists a $\widehat{\mathbf{x}}$ such that $\mathbf{y} = \mathbf{c}(\overline{\mathbf{p}}(\widehat{\mathbf{x}}))$. If (4.15) converges, then (4.22) holds and $\widehat{\mathbf{x}} = \mathrm{argmin}_{\mathbf{x}} \, \| \mathbf{c}(\overline{\mathbf{p}}(\mathbf{x})) - \mathbf{f}(\overline{\mathbf{x}}) \|$.

Since the surrogate model is more flexible than the fine model $\mathbf{c}(\overline{\mathbf{p}}(\widehat{\mathbf{x}})) = \mathbf{f}(\overline{\mathbf{x}})$. This yields a contradiction. Hence (4.15) cannot converge. $\qquad\square$

Remark. So in the end we have what one would have expected. If $\mathbf{c}(Z) = \mathbf{f}(X)$ and the process (4.15) converges, the space mapping is perfect and the fixed point of the iteration is the solution of our optimization problem.

Convergence of DeC-Y. In Section 4.3 Lemma 7, it was shown that the iterations (4.16) and (4.19) are equivalent. Hence, their convergence results coincide. For (4.16) we derive

Lemma 11. Let the space $\widehat{E}$ be a Hilbert space. In the case of convergence of (4.16), with fixed point $\lim_{k\to\infty} \mathbf{x}_k = \overline{\mathbf{x}}$ we obtain

$$\mathbf{f}(\overline{\mathbf{x}}) - \mathbf{y} \in \mathbf{c}(\overline{\mathbf{p}}(X))^{\perp}(\overline{\mathbf{x}}) \,.$$

Proof. In the case of convergence of (4.16), with fixed point $\lim_{k\to\infty} \mathbf{x}_k = \overline{\mathbf{x}}$ we obtain

$$\overline{\mathbf{x}} := \mathop{\mathrm{argmin}}_{\mathbf{x} \in X} \, \| \mathbf{c}(\overline{\mathbf{p}}(\mathbf{x})) - \mathbf{c}(\overline{\mathbf{p}}(\overline{\mathbf{x}})) + \mathbf{f}(\overline{\mathbf{x}}) - \mathbf{y} \| \,. \tag{4.24}$$

So, it follows that for all $\mathbf{h} \in \mathbb{R}^n$ with $\|\mathbf{h}\| \geq 0$

$$\| \mathbf{c}(\overline{\mathbf{p}}(\overline{\mathbf{x}} + \mathbf{h})) - \mathbf{c}(\overline{\mathbf{p}}(\overline{\mathbf{x}})) + \mathbf{f}(\overline{\mathbf{x}}) - \mathbf{y} \| \geq \| \mathbf{f}(\overline{\mathbf{x}}) - \mathbf{y} \| \,.$$

With $\widehat{E}$ a Hilbert space, this implies that

$$\mathbf{f}(\overline{\mathbf{x}}) - \mathbf{y} \perp \mathbf{c}(\overline{\mathbf{p}}(\overline{\mathbf{x}} + \mathbf{h})) - \mathbf{c}(\overline{\mathbf{p}}(\overline{\mathbf{x}})) \qquad \forall \mathbf{h} \in \mathbb{R}^n \,,$$

so that $\mathbf{f}(\overline{\mathbf{x}}) - \mathbf{y}$ is perpendicular to the linear manifold through $\mathbf{c}(\overline{\mathbf{p}}(\overline{\mathbf{x}}))$ that is tangential to $\mathbf{c}(\overline{\mathbf{p}}(X))$. $\qquad\square$

Remark. Notice the slight difference with the result for DeC-X.

Lemma 12. With a flexible coarse model and $\mathbf{p}_k(\mathbf{x}_k) = \mathbf{p}(\mathbf{x}_k)$ the iteration processes (4.15) and (4.16) yield identical sequences of approximations. Hence, under these conditions, the lemmas 9 and 10 for (4.15) also hold for (4.16).

Proof. The main iteration in DeC-Y is

$$\mathbf{x}_{k+1} = \mathrm{argmin}_{\mathbf{x} \in X} \, \| \mathbf{c}(\mathbf{p}_{k+1}(\mathbf{x}))) - \mathbf{c}(\mathbf{p}_k(\mathbf{x}_k))) + \mathbf{f}(\mathbf{x}_k) - \mathbf{y} \| \,.$$

From Lemma 3 for a more flexible coarse model

$$\mathbf{c}(\mathbf{p}(\mathbf{x}_k)) = \mathbf{f}(\mathbf{x}_k) \,.$$

Since $\mathbf{p}_k(\mathbf{x}_k) = \mathbf{p}(\mathbf{x}_k)$ we have

$$\mathbf{x}_{k+1} = \mathrm{argmin}_{\mathbf{x} \in X} \, \| \mathbf{c}(\mathbf{p}_{k+1}(\mathbf{x}))) - \mathbf{y} \| \,.$$

which coincides with DeC-X (4.15) because we satisfy the hypothesis of recovering space mapping from defect correction (Section 4.3). $\qquad\square$

Convergence of DeC-Z. Because of the similarity of the lemmas and the proofs for the scheme (4.18) and those for (4.15), here we omit the proofs.

Lemma 13. Let $\widehat{E}$ be a Hilbert space. In case of convergence of (4.18) with a fixed point $\overline{\mathbf{x}} = \overline{\mathbf{q}}(\overline{\mathbf{z}})$ we obtain

$$\mathbf{f}(\overline{\mathbf{x}}) - \mathbf{y} \ \in \ \mathbf{c}(Z)^\perp(\mathbf{z}^*) \equiv \mathbf{c}(\overline{\mathbf{p}}(X))^\perp(\overline{\mathbf{p}}^{-1}\mathbf{p}(\overline{\mathbf{x}})) \ . \square$$

Lemma 14. Let $\mathbf{p}(\mathbf{x})$ be the space-mapping function as defined in (2.5). If (4.18) converges to $\overline{\mathbf{x}} = \overline{\mathbf{q}}(\overline{\mathbf{z}})$ then $\mathbf{p}(\overline{\mathbf{x}}) = \mathbf{z}^*$. $\hspace{2cm}\square$

Again it follows that, with $\overline{\mathbf{x}}$ the convergence result of (4.18), $\mathbf{p}(\overline{\mathbf{x}}) = \mathbf{p}(\mathbf{x}^*)$ if and only if $\mathbf{p}$ is a perfect mapping. If $\mathbf{c}(Z) = \mathbf{f}(X)$ and the process converges, the fixed point of the iteration is the solution of our optimization problem.

Lemma 15. If the coarse model is more flexible than the fine one, i.e., $\mathbf{c}(Z) \supset \mathbf{f}(X)$, and $\mathbf{y} \in \mathbf{c}(Z) \setminus \mathbf{f}(X)$, then (4.18) cannot converge. $\hspace{2cm}\square$

The processes with linear coarse model and mappings A coarse model linear in its control variables together with linear mappings for either $\mathbf{p}_k$ and $\overline{\mathbf{p}}$ or $\mathbf{q}_k$ and $\overline{\mathbf{q}}$, can be attractive in practise and the analysis can give some additional insight. Under these conditions of linearity, all defect correction processes above coincide. For the analysis we take the iteration (4.18) with the Euclidean norm as a starting point.

Denoting by C the $m \times n$ matrix that describes the linear coarse model, we obtain

$$\begin{cases} \mathbf{z}_0 &= \ \mathbf{z}^* = \operatorname{argmin}_{\mathbf{z} \in Z} \|C\mathbf{z} - \mathbf{y}\|_2 \ , \\ \mathbf{z}_{k+1} &= \ \mathbf{z}_k - \operatorname{argmin}_{\mathbf{z} \in Z} \|C\mathbf{z} - \mathbf{f}(\mathbf{q}_k(\mathbf{z}_k))\|_2 + \mathbf{z}^* \ , \end{cases} \tag{4.25}$$

or

$$\begin{cases} \mathbf{z}_0 &\in \ C^\dagger \mathbf{y} + \operatorname{Span}(V_0) \ , \\ \mathbf{z}_{k+1} &\in \ \mathbf{z}_k - C^\dagger(\mathbf{f}(\mathbf{q}_k(\mathbf{z}_k)) - \mathbf{y}) + \operatorname{Span}(V_0) \ , \end{cases} \tag{4.26}$$

where $C^\dagger$ is the pseudo-inverse of C and $\operatorname{Span}(V_0) \subset \mathbb{R}^m$ is the kernel of $C^T C$. The fixed point $\mathbf{f}(\overline{\mathbf{x}}) = \mathbf{f}(\overline{\mathbf{q}}(\overline{\mathbf{z}}))$ of iteration (4.26) is characterized by

$$C^\dagger(\mathbf{y} - \mathbf{f}(\overline{\mathbf{x}})) \in \operatorname{Span}(V_0) \ .$$

In the case $m \geq n$ and with a full rank C operator, the term $\operatorname{Span}(V_0)$ vanishes, so that $C^T(\mathbf{y} - \mathbf{f}(\overline{\mathbf{x}})) = 0$, i.e. the coarse model cannot recognize data components in the kernel of C^T.

Convergence of the linearized process. From (4.26) it is clear that for a non-full rank C matrix or if $m < n$, the solution is not uniquely determined and the iteration cannot be a contraction whereas for a full-rank C with $m \geq n$, the convergence rate is determined by the Lipschitz constant

$$\| I - C^\dagger \mathbf{f}(\overline{\mathbf{q}}(\cdot)) \| \ .$$

If we consider the Euclidean norm and the operators $\mathbf{f} : X \to Y$ and $\overline{\mathbf{q}} : Z \to X$ have Jacobians $F \in \mathbb{R}^{m \times n}$ and $\overline{Q} \in \mathbb{R}^{n \times n}$ respectively, then the convergence rate is bounded by

$$\|I - C^{\dagger} F \overline{Q}\|_2 = \|(C^T C)^{-1} C^T \left(C - F \overline{Q} \right)\|_2 \leq \|(C^T C)^{-1}\|_2 \, \|C^T \left(C - F \overline{Q} \right)\|_2 \, .$$

This shows how the operators $\mathbf{c}$ and $\mathbf{f} \circ \overline{\mathbf{q}}$ should be related for the iteration process to converge: the coarse-model $\mathbf{c}$ should be full rank and be sufficiently stable (well-conditioned) and the fine-model $\mathbf{f}$, preconditioned by $\overline{\mathbf{q}}$, should sufficiently correspond with the coarse-model $\mathbf{c}$ (at least with respect to the components in its range).

Left- and right-preconditioning.   The efficiency of the space-mapping algorithms depends essentially on the similarity between the fine and the coarse model used and their relative costs for solution. The space-mapping procedure is then determined by the choice how the space-mapping function is approximated, with ASM and TRASM as obvious choices.

In the defect correction algorithms we also have to deal with the similarity between the fine and the coarse model but, instead of in the selection of the space-mapping approximation, there the freedom is in the preconditioning mapping $\overline{\mathbf{p}}$ or $\overline{\mathbf{q}} = \overline{\mathbf{p}}^{-1}$. This choice can be used to improve the convergence of the procedure. From the linear case we see that better convergence can be expected if the approximation $\mathbf{c}(\overline{\mathbf{p}}(\mathbf{x})) \approx \mathbf{f}(\mathbf{x})$ is more accurate. From $\mathbf{c}(\mathbf{z}) \approx \mathbf{f}(\mathbf{q}(\mathbf{z}))$ we see that $\mathbf{q}$ acts as a right-preconditioner [33].

On the other hand, by the orthogonality relations in the lemmas above, we see that it is advantageous, both for the conditioning of the problem and for the minimization of the residual, if the manifolds $\mathbf{f}(X)$ and $\mathbf{c}(Z)$ are found parallel in the neighborhood of the solution. However, by space mapping or by right-preconditioning the relation between the manifolds $\mathbf{f}(X)$ and $\mathbf{c}(Z)$ remains unchanged. This relation can be improved by the effect of an additional left-preconditioner. Therefore we have to introduce such a preconditioner $\mathbf{r}$ so that near $\mathbf{f}(\mathbf{x}^*) \in Y$ the manifold $\mathbf{f}(X) \subset Y$ is mapped onto $\mathbf{c}(Z) \subset Y$:

$$\mathbf{r}(\mathbf{f}(\mathbf{q}(\mathbf{z}))) \approx \mathbf{c}(\mathbf{z}) \, .$$

Hence, in the next section we propose an affine operator $\mathbf{S} = \mathbf{r}^{-1}$, which maps $\mathbf{c}(Z)$ onto $\mathbf{f}(X)$, at least in the neighborhood of the solution. This improves significantly the traditional approach. The recently introduced output space mapping and space-mapping-based interpolating surrogate scheme [9] also take such left preconditioning into account, but in a slightly different manner.

4.4 Improved space mapping
Except for the hybrid method HASM, all methods thus far have the clear disadvantage that, in general, the fixed point of the iteration does not coincide with the solution of the fine model minimization problem. This is due to the fact that the approximate solution $\overline{\mathbf{x}}$ satisfies

$$
\begin{aligned}
\mathbf{f}(\overline{\mathbf{x}}) - \mathbf{y} &\in \;\; \mathbf{c}(\overline{\mathbf{p}}(X))^{\perp}(\overline{\mathbf{p}}^{-1}\mathbf{p}(\overline{\mathbf{x}})) &&\text{for DeC-X and DeC-Z,} \\
\mathbf{f}(\overline{\mathbf{x}}) - \mathbf{y} &\in \;\;\;\;\; \mathbf{c}(\overline{\mathbf{p}}(X))^{\perp}(\overline{\mathbf{x}}) &&\text{for DeC-Y,}
\end{aligned}
$$

whereas a (local) minimum for $\|\mathbf{f}(\mathbf{x}) - \mathbf{y}\|$ satisfies

$$\mathbf{f}(\mathbf{x}^*) - \mathbf{y} \in \mathbf{f}(X)^{\perp}(\mathbf{x}^*) \, .$$

Hence, differences between $\overline{\mathbf{x}}$ and $\mathbf{x}^*$ will be larger for larger distances between $\mathbf{y}$ and the sets $\mathbf{f}(X)$ and $\mathbf{c}(Z)$, and for larger angles between the linear manifolds tangential at $\mathbf{c}(Z)$ and $\mathbf{f}(X)$ near the optima. This disadvantage, however, can be easily removed by introducing an affine mapping $\mathbf{S} : Y \rightarrow Y$ such that $\mathbf{S}\, \mathbf{c}(\overline{\mathbf{z}}) = \mathbf{f}(\mathbf{x}^*)$ for a proper $\overline{\mathbf{z}} \in Z$. This means that $\mathbf{S}$ maps $\mathbf{c}(\overline{\mathbf{z}})$ onto $\mathbf{f}(\mathbf{x}^*)$ and the linear manifold tangential to $\mathbf{c}(Z)$ in $\mathbf{c}(\overline{\mathbf{z}})$ onto the one tangential to $\mathbf{f}(X)$ in $\mathbf{f}(\mathbf{x}^*)$. Because, in the non-degenerate case when $m \geq n$, both $\mathbf{f}(X)$ and $\mathbf{c}(Z)$ are n-dimensional sets in $\mathbb{R}^m$, the mapping $\mathbf{S}$ can be described by

$$\mathbf{S}\, \mathbf{v} = \mathbf{f}(\mathbf{x}^*) + S\ (\mathbf{v} - \mathbf{c}(\overline{\mathbf{z}}))\ ,$$

where S is an $m \times m$-matrix S of rank n. This mapping $\mathbf{S}$ is not a priori available, but an approximation to it can be computed iteratively during the optimization. A full rank $m \times m$-matrix S can be constructed, which has a well-determined part of rank n, while a remaining part of rank $m - n$ is free to choose. Because of the supposed similarity between the models $\mathbf{f}$ and $\mathbf{c}$ we keep the latter part close to the identity.

We propose the following algorithm ($\overline{\mathbf{p}} : X \rightarrow Z$ being a proper but yet arbitrary mapping)

1. Set $k = 0$, set $S_0 = I$ the $m \times m$ identity matrix, and compute

$$\mathbf{x}_0 = \operatorname{argmin}_{\mathbf{x} \in X}\ \| \mathbf{c}(\overline{\mathbf{p}}(\mathbf{x})) - \mathbf{y} \|\ .$$

2. Compute $\mathbf{f}(\mathbf{x}_k)$ and $\mathbf{c}(\overline{\mathbf{p}}(\mathbf{x}_k))$.

3. If $k > 0$, with $\Delta \mathbf{c}_i = \mathbf{c}(\overline{\mathbf{p}}(\mathbf{x}_{k-i})) - \mathbf{c}(\overline{\mathbf{p}}(\mathbf{x}_k))$ and $\Delta \mathbf{f}_i = \mathbf{f}(\mathbf{x}_{k-i}) - \mathbf{f}(\mathbf{x}_k)$, $i = 1, \cdots, \min(n, k)$, we define ΔC and ΔF to be the rectangular $m \times \min(n, k)$-matrices with respectively $\Delta \mathbf{c}_i$ and $\Delta \mathbf{f}_i$ as columns. Their singular value decompositions are respectively $\Delta C = U_c \Sigma_c V_c^T$ and $\Delta F = U_f \Sigma_f V_f^T$.
   Now, $S_k = \Delta F\, \Delta C^\dagger + (I - U_{\mathbf{f}}\, U_{\mathbf{f}}^T)\, (I - U_{\mathbf{c}}\, U_{\mathbf{c}}^T)$ and the approximate affine mapping is

$$\mathbf{S}_k\, \mathbf{v} = \mathbf{f}(\mathbf{x}_k) + S_k(\mathbf{v} - \mathbf{c}(\overline{\mathbf{p}}(\mathbf{x}_k))), \quad \forall \mathbf{v} \in Y,$$

   which, for $l > 0$ and $l = k - 1, \cdots, \max(0, k - n)$, satisfies

$$S_k\, (\mathbf{c}(\overline{\mathbf{p}}(\mathbf{x}_l)) - \mathbf{c}(\overline{\mathbf{p}}(\mathbf{x}_k))) = \mathbf{f}(\mathbf{x}_l) - \mathbf{f}(\mathbf{x}_k)\ .$$

4. The next iterand is computed as

$$\begin{aligned}
\mathbf{x}_{k+1} = \\
\operatorname{argmin}_{\mathbf{x} \in X}\ \| \mathbf{S}_k(\mathbf{c}(\overline{\mathbf{p}}(\mathbf{x}))) - \mathbf{y} \| = \\
\operatorname{argmin}_{\mathbf{x} \in X}\ \| S_k(\mathbf{c}(\overline{\mathbf{p}}(\mathbf{x})) - \mathbf{c}(\overline{\mathbf{p}}(\mathbf{x}_k))) + \mathbf{f}(\mathbf{x}_k) - \mathbf{y} \| = \\
\operatorname{argmin}_{\mathbf{x} \in X}\ \| \mathbf{c}(\overline{\mathbf{p}}(\mathbf{x})) - \mathbf{c}(\overline{\mathbf{p}}(\mathbf{x}_k)) + \left[\Delta C\, \Delta F^\dagger + I - U_c U_c^T \right] (\mathbf{f}(\mathbf{x}_k) - \mathbf{y}) \|.
\end{aligned} \tag{4.27}$$

5. Set $k := k + 1$ and goto 2.

Here, † again denotes the pseudo-inverse, $\Delta F^\dagger = V_f \Sigma_f^{-1} U_f^T$. In the last line in (4.27), the freedom in making S_k full-rank is used, replacing $\Delta C\ \Delta F^\dagger + (I - U_c U_c^T)(I - U_f U_f^T)$ by $\Delta C\ \Delta F^\dagger + I - U_c U_c^T$, in order to stabilize the algorithm. This does not change the solution.

If the above iteration converges with fixed point $\overline{\mathbf{x}}$ and mappings $\overline{\mathbf{S}}$ and $\overline{\mathbf{p}}$, we have

$$\mathbf{f}(\overline{\mathbf{x}}) - \mathbf{y} \in \overline{\mathbf{S}}(\mathbf{c}(\overline{\mathbf{p}}(X)))^\perp(\overline{\mathbf{x}}) = \mathbf{f}(X)^\perp(\overline{\mathbf{x}})\ .$$

From this relation and the fact that $\mathbf{S}_k(\mathbf{c}(\overline{\mathbf{p}}(\mathbf{x}_k))) = \mathbf{f}(\mathbf{x}_k)$, it can be concluded that, under convergence to $\overline{\mathbf{x}}$, the fixed point is a (local) optimum of the fine model minimization.

The improved space-mapping scheme

$$\mathbf{x}_{k+1} = \underset{\mathbf{x}}{\operatorname{argmin}}\ \|\,\mathbf{S}_k(\mathbf{c}(\overline{\mathbf{p}}_k(\mathbf{x})))) - \mathbf{y}\|$$

can also be recognized as defect correction iteration with either $\widetilde{\mathcal{F}}_k = \mathbf{S}_k \circ \mathbf{c} \circ \overline{\mathbf{p}}_k$ and $\mathcal{F} = \mathbf{f}$ in (4.15) or (4.16), or with $\widetilde{\mathcal{F}}_k = \mathbf{S}_k \circ \mathbf{c}$ and $\mathcal{F} = \mathbf{f} \circ \overline{\mathbf{q}}$ in (4.18) or (4.19).

5. Examples

5.1 Simple examples and counterexamples

Coinciding space-mapping optima and deviation from perfect mapping   Each of the four points $\mathbf{z}^*$, $\mathbf{x}_p^*$, $\mathbf{x}_d^*$ and $\mathbf{x}^*$ introduced in the previous sections is the solution of a minimization problem, the first three being approximations to the last. In this section by a few very simple examples we illustrate that in some cases these approximations coincide, but in other cases they can be all different.

As a first example we consider a simple least squares data fitting problem: the data are $\mathbf{y} = [y_{-1}, y_0, y_1]$ for $\mathbf{t} = [-1, 0, 1]$, so we have $m = 3$. The fine model is the family of polynomials over $X = \mathbb{R}^2$,

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(x_1, x_2) = \left[x_1^2\, t_i^2 + x_1\, t_i + x_2\right]_{i=1,2,3}\ . \tag{5.1}$$

The coarse model is the linear $\mathbf{c}(\mathbf{z}) = \mathbf{c}(z_1, z_2) = z_1\, t_i + z_2$. The number of control parameters is $n = 2$ and the two models are clearly similar for small x_1. The search is made over $X = Z = \mathbb{R}^2$. It is immediate that the coarse model minimum is

$$\mathbf{z}^* = \left[\frac{y_1 - y_{-1}}{2}, \frac{y_{-1} + y_0 + y_1}{3}\right]\ . \tag{5.2}$$

Using the definition of the space-mapping function, it is easily determined as

$$\mathbf{p}(\mathbf{x}) = \left[x_1, x_2 + \frac{2}{3}x_1^2\right]\ ,$$

which for small x_1 clearly resembles the identity operator. It is a bijection between X and Z and by applying Lemma 2 in Section 3.3 we see that for this problem all the space-mapping solutions exist and coincide

$$\mathbf{x}_{sm}^* = \mathbf{x}_p^* = \mathbf{x}_d^* = \left[\frac{y_1 - y_{-1}}{2}, \frac{y_{-1} + y_0 + y_1 - (y_1 - y_{-1})^2/2}{3}\right]\ . \tag{5.3}$$

Notice that this is not the true solution for fitting with (5.1). Only in the case of perfect mapping the solution $\mathbf{x}_{sm}^*$ corresponds with the true solution $\mathbf{x}^*$.

Figure 10: Quality of the space-mapping optimum depending on the deviation from perfect mapping.
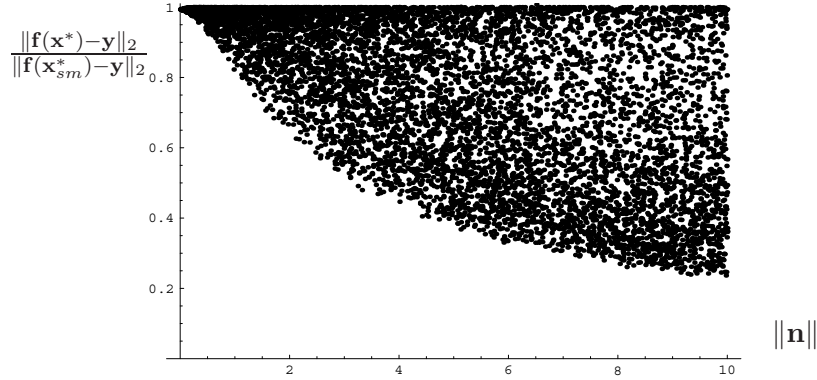
Results for 8000 experiments with random perturbations of size $\|\mathbf{n}\|_{\mathbf{2}} < 10$.
In the example $\|\mathbf{f}(\widehat{\mathbf{x}})\|_2 = 0.2328$

Deviation from perfect mapping.   In order to check the value of the space-mapping approach and see the effect of a non-perfect mapping, we perform the following experiment. We create the data $\mathbf{y} = \mathbf{f}(\widehat{\mathbf{x}}) + \mathbf{n}$, with $\widehat{\mathbf{x}} = [0.1, 0.1]$ and $\mathbf{n}$ a random vector, uniformly distributed on a sphere with radius $\|\mathbf{n}\| = r$. Figure 10 shows the quotient $\|\mathbf{f}(\mathbf{x}^*) - \mathbf{y}\|_2/\|\mathbf{f}(\mathbf{x}^*_{sm}) - \mathbf{y}\|_2$. When $\mathbf{n}$ is not significant compared to the fine model response, then the mapping is almost perfect and the space-mapping solution yields a response with a cost function close to the true optimum. With this test we corroborate that a coarse model that can accurately reproduce the response of the fine is a good candidate for use with space mapping, and second that fine models that meet the specifications well improve the quality of the space-mapping result.

Different optima $\mathbf{x}^*$, $\mathbf{x}^*_p$ *and* $\mathbf{x}^*_d$ However, in general, if $\mathbf{z}^* \notin \mathbf{p}(X)$ the solutions $\mathbf{x}^*_p$ and $\mathbf{x}^*_d$ are expected to be different. We expect that by increasing the misalignment between the models, the condition $\mathbf{z}^* \in \mathbf{p}(X)$ fails to be fulfilled. This is shown by taking another simple fine model:

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(x_1, x_2) = [x_1 \, (x_2 \, t_i + 1)^2]_{i=1,2,3} \,. \tag{5.4}$$

We keep the linear coarse model $\mathbf{c}(\mathbf{z}) = \mathbf{c}(z_1, z_2) = z_1\mathbf{t} + z_2$. Proceeding as above, the space-mapping function can be determined as

$$\mathbf{p}(\mathbf{x}) = \left[2\, x_1 \, x_2 \,, x_1 \left(1 + \frac{2}{3}\, x_2^2 \right) \right] \,. \tag{5.5}$$

This mapping clearly is not injective and is substantially different from the identity. Furthermore, $\mathbf{p}(X) \neq \mathbb{R}^2$ and it is easily verified that $\mathbf{p}(X) = \{\mathbf{z} \in Z; \, 3\, z_2^2 \geq 2\, z_1^2\}$.

For this example we can distinguish four different situations: (1) the design is reachable, (2) the mapping is perfect but the design is not reachable, (3) the primal and the dual solution correspond but the mapping is not perfect, and (4) the four points $\mathbf{z}^*$, $\mathbf{p}(\mathbf{x}^*_p)$, $\mathbf{p}(\mathbf{x}^*_d)$ and $\mathbf{p}(\mathbf{x}^*)$ are all different.

Figure 11: Different space-mapping solutions for problem (5.4).
Four situations: (1) and (2): coinciding $\mathbf{z}^* = \mathbf{p}(\mathbf{x}_p^*) = \mathbf{p}(\mathbf{x}_d^*) = \mathbf{p}(\mathbf{x}^*)$;
(3): coinciding $\mathbf{p}(\mathbf{x}_p^*) = \mathbf{p}(\mathbf{x}_d^*)$ separate $\mathbf{z}^*$ and $\mathbf{p}(\mathbf{x}^*)$;
(4): separate $\mathbf{z}^*$, $\mathbf{p}(\mathbf{x}_p^*)$, $\mathbf{p}(\mathbf{x}_d^*)$ and $\mathbf{p}(\mathbf{x}^*)$.
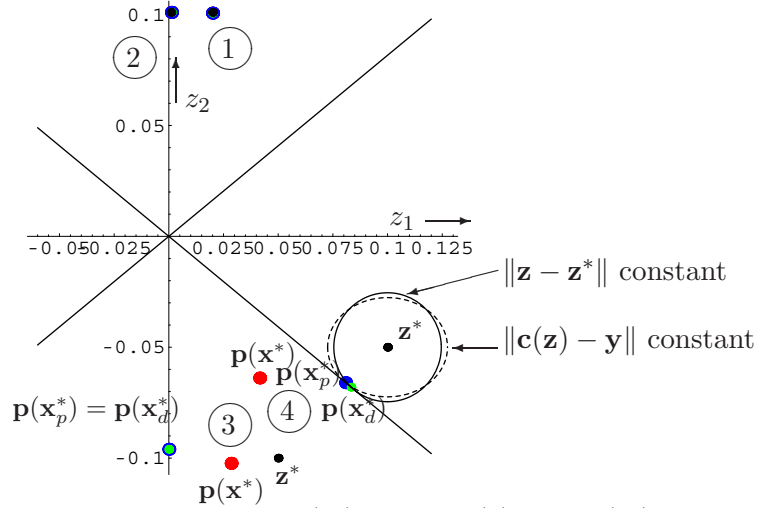
The top and bottom quadrant are $\mathbf{p}(X)$. In case (4) $\mathbf{z}^* \notin \mathbf{p}(X)$.

These four cases are shown in Table 1 and in Figure 11. Case (1) is characterized by the fact that the cost function for the fine model vanishes at $\mathbf{x}_p^* = \mathbf{x}_d^* = \mathbf{x}^*$. This occurs if the data $\mathbf{y}$ fit the model (5.4) so that a reachable design is obtained. In case (2) we also see $\mathbf{x}_p^* = \mathbf{x}_d^* = \mathbf{x}^*$, but the cost function for the fine model does not vanish. In case (3) still $\mathbf{z}^* \in \mathbf{p}(X)$, but the data are such that $\mathbf{p}$ is no perfect mapping so that $\mathbf{x}_p^* = \mathbf{x}_d^*$, but they do not coincide with $\mathbf{x}^*$. For the case (4) the data $\mathbf{y}$ are selected such that $\mathbf{z}^* \notin \mathbf{p}(X)$. Now $\mathbf{x}_p^* \neq \mathbf{x}_d^*$ and, in fact, all values differ.

Because of the quadratic form in $\mathbf{p}(\mathbf{x})$, each of the minimization procedures for the primal and the dual approach yields two solutions. In Table 1 we omit the one that has the larger fine cost function.

The optimization problem that defines the primal approach finds the point in $\mathbf{p}(X)$ that is closest in the Euclidean norm to $\mathbf{z}^*$. It lies on a circle centered in $\mathbf{z}^*$, with minimum radius so that it has one point of contact with $\mathbf{p}(X)$. For the dual approach the contour lines for the cost function are ellipses in $\mathbf{p}(X)$ centered in $\mathbf{z}^*$. Therefore, $\mathbf{p}(\mathbf{x}_p^*)$ and $\mathbf{p}(\mathbf{x}_d^*)$, and consequently $\mathbf{x}_p^*$ and $\mathbf{x}_d^*$, do not coincide.

Remark. Case (4) has two local optima. The behavior of the fine cost function in the neighborhood of one of them is similar to that of the coarse cost function close to the coarse optimum. But for the other one, the global fine optimum, the difference is substantial. Space-mapping finds the local optimum with the similar behavior (see Figure 12). This observation emphasizes the importance of the similarity between the models used.

Table 1: Different space-mapping solutions for the least squares data fitting problem associated with (5.4).

	y_{-1}	y_0	y_1	$\mathbf{z}^*$	$C(\mathbf{z}^*)$	$F(\mathbf{z}^*)$
1	0.08100	0.10000	0.12100	[0.020,0.101]	8.16e-04	1.41e-01
2	0.10011	0.10125	0.10241	[0.001,0.101]	8.16e-06	1.73e-01
3	0.00000	-0.40000	0.10000	[0.050,-0.100]	3.67e-01	4.58e-01
4	0.00000	-0.35000	0.20000	[0.100,-0.050]	3.67e-01	4.76e-01

	$\mathbf{x}^*$	$\mathbf{x}_p^*$	$\mathbf{x}_d^*$	$F(\mathbf{x}^*)$	$F(\mathbf{x}_p^*)$	$F(\mathbf{x}_d^*)$
1	[0.100,0.100]	[0.100,0.100]	[0.100,0.100]	1.96e-17	3.10e-17	1.96e-17
2	[0.101,0.006]	[0.101,0.006]	[0.101,0.006]	5.50e-06	5.50e-06	5.50e-06
3	[-0.101,-0.141]	[-0.096,-0.261]	[-0.096,-0.261]	3.70e-01	3.73e-01	3.73e-01
4 a	[-0.059,-0.352]	[-0.034,-1.225]	[-0.033,-1.225]	3.83e-01	4.12e-01	4.10e-01
b	[0.007,4.007]			3.63e-01		

$$C(\mathbf{z}) = \|\mathbf{c}(\mathbf{z}) - \mathbf{y}\|_2 \text{ and } F(\mathbf{x}) = \|\mathbf{f}(\mathbf{x}) - \mathbf{y}\|_2.$$

The importance of the use of similar models. Now follows an example showing how strong dissimilarities between the models make both space-mapping approaches fail. For the fine model we take the family of exponentials

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(x_1, x_2) = [x_1 \exp(x_2 t_i)]_{i=1,2,3} ,$$

and the coarse is linear as before. In both cases $X = Z = \mathbb{R}^2$. The space-mapping function is again uniquely determined

$$\mathbf{p}(\mathbf{x}) = \left[x_1 \sinh(x_2), \frac{x_1}{3}(1 + 2\cosh(x_2)) \right].$$

Now $\mathbf{p}(X) = \{\mathbf{z} \in Z; 9z_2^2 > 4z_1^2\}$ is an open set and thus, in the case $\mathbf{z}^* \notin \mathbf{p}(X)$, the primal and dual space-mapping solutions do not exist because no minimum exists for the minimization problems that define these approximations. The fine and the coarse model have little in common, to the extent that for certain specifications the space-mapping technique cannot be applied. The conclusion is that in order to have a proper space-mapping scheme, it is crucial to use models that are sufficiently similar.

Preference for the dual approach The dual space-mapping solution for case (4) above has a smaller cost function associated than the primal one. Additional experiments with the same problem were carried out in order to check if this better performance can generally be expected, but the results were inconclusive. In [30, section 4.1.2] the following lemma in favor of the dual approach is given and checked experimentally.

Lemma 16. If $\exists \epsilon, \delta > 0$, $\epsilon \leq \delta$ such that
(i) $\|\mathbf{c}(\mathbf{p}(\mathbf{x}_p^*)) - \mathbf{f}(\mathbf{x}_p^*)\| \leq \epsilon/2$, $\|\mathbf{c}(\mathbf{p}(\mathbf{x}_d^*)) - \mathbf{f}(\mathbf{x}_d^*)\| \leq \epsilon/2$, and
(ii) $\|\mathbf{c}(\mathbf{p}(\mathbf{x}_p^*)) - \mathbf{y}\| - \|\mathbf{c}(\mathbf{p}(\mathbf{x}_d^*)) - \mathbf{y}\| \geq \delta$
then $\|\mathbf{f}(\mathbf{x}_d^*) - \mathbf{y}\| \leq \|\mathbf{f}(\mathbf{x}_p^*) - \mathbf{y}\|$. □

Figure 12: Level curves and minima of the fine and coarse cost functions for the case (4) in Table 1. The solution found by the improved algorithm, $\overline{\mathbf{x}}$, corresponds with the local optimum for the fine model, which shows similarity with the coarse model.
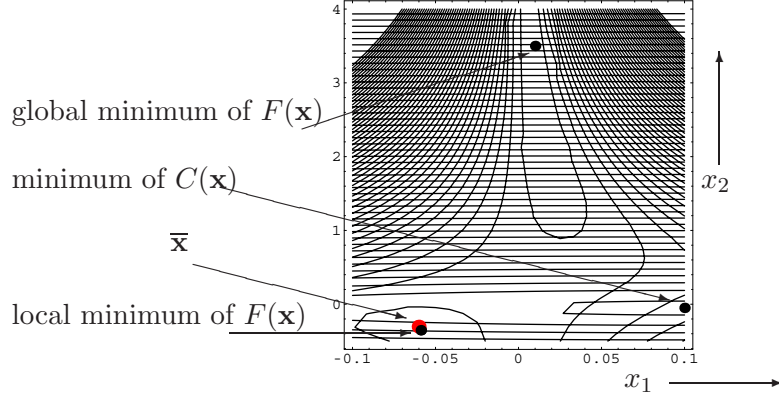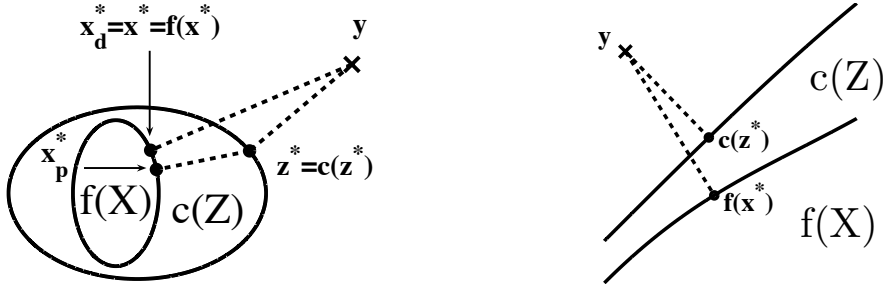

Figure 13: Left : A simple counterexample for $\mathbf{x}_p^* \neq \mathbf{x}^* = \mathbf{x}_d^*$. The Euclidean norm is taken as cost function. Right: Usually the fine and coarse models are similar enough near the optima that $\mathbf{c}(\mathbf{z}^*) \notin \mathbf{f}(X)$ implies a behavior as shown.

We give two theoretical examples supporting a choice in favor of the dual approach, specially in the case where the flexability concept can be used.

Example I.   In this example X is a non-trivial subset of Z. On X the fine and coarse model responses are taken equal. Then the space-mapping function is the identity and the mapped coarse model matches the fine one, i.e., $\mathbf{c}(\mathbf{p}(\mathbf{x})) = \mathbf{f}(\mathbf{x})\ \forall \mathbf{x} \in X$ (see Lemma 3 (i) Section 3.3). Under these circumstances the dual solution $\mathbf{x}_d^*$ coincides with the fine optimum $\mathbf{x}^*$. But, in general, this is not the case for the primal approach, as can be seen for the particular sets X and Z shown in Figure 13 (left), where $\mathbf{c}$ is the identity, the Euclidean norm is used, and with $\mathbf{y} \in Y \setminus \mathbf{c}(Z)$ we see that $\mathbf{x}_p^* \neq \mathbf{x}^*$. This argument, based on projection onto subsets, is more general and can be used in less obvious situations.

Example II.   In Figure 13 (right) we analyze the possibility of $\mathbf{x}_p^* = \mathbf{x}^* \neq \mathbf{x}_d^*$ with $\mathbf{c}$ (locally) more flexible than $\mathbf{f}$ in the region near the optimal responses. This means that we assume the existence of a set $R \subset \mathbb{R}^n$ such that $R \supset R \cap \mathbf{c}(Z) \supset R \cap \mathbf{f}(X)$ and that $\mathbf{c}(\mathbf{p}(\mathbf{x}_d^*)), \mathbf{f}(\mathbf{x}^*), \mathbf{c}(\mathbf{z}^*) \in R$. We define the auxiliary set $Q = \{\mathbf{x} \in X \,|\, \mathbf{f}(\mathbf{x}) \in R\}$. By Lemma 3

Table 2: Results for the improved algorithm

	$\overline{\mathbf{x}}$	$F(\overline{\mathbf{x}})$
1	[0.100,0.100]	6.30e-06
2	[0.101,0.006]	5.53e-06
3	[-0.101,-0.141]	3.70e-01
4	[-0.060,-0.304]	3.83e-01

$$\tau = 0.00001, \ \lambda = 0.01$$

(i) Section 3.3, applied locally, we know $\mathbf{c}(\mathbf{p}(\mathbf{x})) = \mathbf{f}(\mathbf{x})$ for all $\mathbf{x} \in Q$, consequently

$$
\begin{aligned}
\mathbf{x}^* &= \mathrm{argmin}_{\mathbf{x} \in X} \ \|\mathbf{f}(\mathbf{x}) - \mathbf{y}\| = \mathrm{argmin}_{\mathbf{x} \in Q} \ \|\mathbf{f}(\mathbf{x}) - \mathbf{y}\| \\
&= \mathrm{argmin}_{\mathbf{x} \in Q} \ \|\mathbf{c}(\mathbf{p}(\mathbf{x})) - \mathbf{y}\| = \mathrm{argmin}_{\mathbf{x} \in X} \ \|\mathbf{c}(\mathbf{p}(\mathbf{x})) - \mathbf{y}\| = \mathbf{x}_d^*,
\end{aligned}
$$

i.e., the dual space-mapping approach also yields the fine optimum. We conclude that in this case the equality $\mathbf{x}_p^* = \mathbf{x}^*$ implies $\mathbf{x}_d^* = \mathbf{x}^*$.

The result does not hold for a more flexible fine model in R. Since $\mathbf{x}_p^* \neq \mathbf{x}_d^*$, by Lemma 2 (Section 3.3) the coarse optimum $\mathbf{z}^*$ cannot be in $\mathbf{p}(X)$. From $\mathbf{z}^* \notin \mathbf{p}(X)$ it follows that $\mathbf{c}(\mathbf{z}^*) \notin \mathbf{f}(X)$, otherwise $\mathbf{c}(\mathbf{z}^*) = \mathbf{f}(\widehat{\mathbf{x}})$ for some $\widehat{\mathbf{x}} \in X$, would yield $\mathbf{z}^* = \mathbf{p}(\widehat{\mathbf{x}})$. This, because $\mathbf{c}(\mathbf{z}^*) \in R$ clearly contradicts a fine model more flexible than the coarse in R.

Results for the improved algorithm   We apply the improved space-mapping (manifold-mapping) algorithm from Section 4.4 to the same four examples from Section 5.1. The actual implementation of the algorithm contains two parameters, τ and λ. In order to determine the rank of the matrices ΔC and ΔF we use a threshold value τ for the singular values. The same τ is used to stop the iteration if either $\|\mathbf{x}_k - \mathbf{x}_{k+1}\|_2 < \tau$ or if $\|\mathbf{f}(\mathbf{x}_k) - \mathbf{f}(\mathbf{x}_{k+1})\|_2 / \|\mathbf{f}(\mathbf{x}_k)\|_2 < \tau$. Further we regularise the pseudo-inverse by replacing $\Delta F^\dagger = U_f \Sigma_f^\dagger V_f^T$ by $U_f(\Sigma_f^\dagger + \lambda \sigma_1 I)V_f^T$, where σ_1 is the largest singular value of ΔF and $0 \leq \lambda \ll 1$ is a stabilization (trust region) parameter.

The results of the optimization procedure are given in Table 2. As expected, the results compare favorable with those obtained by the traditional space mapping in Table 1. We see that $\overline{\mathbf{x}} \neq \mathbf{x}^*$, but this is a consequence of the tolerance used in the algorithm.

5.2 A more realistic case: magnetic size design

This is an example [16, 18] how space mapping can be applied for efficiently solving practical optimization problems. The performance is compared with several other minimization schemes. The problem concerns a magnetic device with an air-gap.

A permanent magnet is placed between two ferromagnetic cores, as shown in Figure 14. The lengths d and g are prescribed. The design specifications $\mathbf{y}$ are the magnetic flux densities at three given points in space $(m = 3)$. Three control variables are proposed: the dimensions of the magnet and the main width of the core. Thus, $n = m = 3$.

The fine model considered is a finite element discretisation of a vector magnetic potential formulation of the two-dimensional magnetostatic equations [32]. The approximating functions are Lagrangian interpolants of second order and adaptive mesh refinement is applied,

Figure 14: A two-dimensional magnetostatics problem.
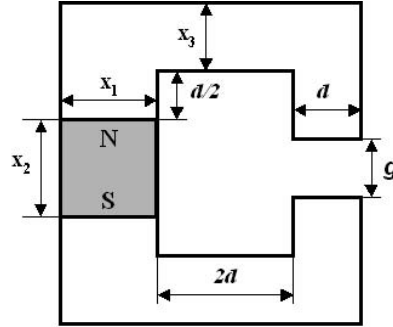
leading to a discretisation with around 40000 degrees of freedom. The software package FEMLAB [17] is used for this purpose. This fine model takes into account important magnetic phenomena like the fringing effect at the air-gap and the flux leakage through the air central window but it still ignores some effects (e.g., nonlinearities in the core). The control parameter space X is the positive octant of $\mathbb{R}^3$.

A simple magnetic circuit description [15] (analogous to electrical circuits) can also be used for the computation of the quantities of interest in the magnetic circuit. The amount of work for such magnetic circuit computations is much less than for finite elements, but at the expense of accuracy. Such magnetic circuit descriptions are ideal candidates for coarse models in magnetic design problems. In our case the coarse model reads [16]

$$\mathbf{c}(\mathbf{z}) = \mathbf{c}(z_1, z_2, z_3) = \left[\frac{\mathbf{B}_r \, z_1 \, z_2}{g \, z_1 + d \, z_2} \, , \ \frac{d \, \mathbf{B}_r \, z_1 \, z_2}{g \, z_1 \, z_3 + d \, z_2 \, z_3} \, , \ \frac{d \, z_2}{g \, z_1} \right] , \tag{5.6}$$

where the $\mathbf{B}_r$ is a constant that describes the material of the permanent magnet and g and d are the fixed dimensions given. The meaning of the control parameters in Z is the same as in X, thus $X = Z$. In the coarse model neither the fringing effect around the air-gap, nor the flux leakage through the window, nor core nonlinearities are considered. However, computing the coarse optimum $\mathbf{z}^*$ is very simple. The computing time for optimization of the coarse model can be neglected when compared with a fine model simulation. The resulting $\mathbf{z}^*$ is given in Table 3.

Results. Two versions of the space-mapping algorithm were compared with three standard optimization schemes. Because $n = m$ we may expect the design to be reachable and, hence, the space mapping to be perfect. So, the stopping criterion from [16],

$$\frac{\|\mathbf{p}(\mathbf{x}_k) - \mathbf{z}^*\|_2}{\|\mathbf{z}^*\|_2} \leq \tau , \tag{5.7}$$

makes sense for finding $\mathbf{x}^*$, even for algorithms not based on space mapping. We use (5.7) with the tolerance $\tau = 0.001$. The advantage of (5.7) over, e.g., $\| \mathbf{f}(\mathbf{x}_k) - \mathbf{y} \|_2 / \| \mathbf{y} \|_2$ is that in the former choice the three vector components are well scaled (because the optimal dimensions are comparable in size). The reachability condition is eventually checked.

The optimization problem is solved by means of five different algorithms. The ASM is implemented together with a variant in which the $\mathbf{B}_k$ is kept the identity matrix during iteration. Also a quasi-Newton method was used for solving $\mathbf{f}(\mathbf{x}^*) = \mathbf{y}$, where the Jacobian was approximated by Broyden's method and the initial estimates for the solution and for the Jacobian were respectively $\mathbf{z}^*$ and $\frac{\delta \mathbf{c}}{\delta \mathbf{z}}(\mathbf{z}^*)$. As this method uses coarse model information as well, it cannot be considered completely independent from principles used in space mapping.

Since derivative information for the fine model is not easily obtained, two direct optimization methods are considered: the Nelder-Mead simplex method (NMS) [21][2] and the global search algorithm DIRECT [22][3]. The first takes the coarse optimum as the initial guess. The second needs no starting value but Z or a proper subset should be specified. We take for the search region the cube centered in $\mathbf{z}^*$ and with the point $\mathbf{z}^*/2$ at a corner. Table 3 shows the number of fine model evaluations for each method until the stopping criterion is satisfied as well as the final value of the three design parameters. The number of evaluations is approximately proportional to the total computing time. ASM turns out to perform best and the quasi-Newton scheme is still efficient. Nelder-Mead shows slow convergence. DIRECT's global optimization nature justifies its poor results.

Table 3: Efficiency comparison between different optimization methods applied to the first (linear) fine model. ASM : the ASM algorithm; SMI : the simplified ASM algorithm with $\mathbf{B}_k = I$; qN : a quasi-Newton scheme; NMS : Nelder-Mead simplex and DIRECT: a direct search method.

	# Evaluations	Final design (mm)
$\mathbf{z}^*$		[5.3571, 7.5000, 5.0000]
ASM	4	[7.9897, 7.5821, 6.5396]
SMI	6	[7.9797, 7.5808, 6.5381]
qN	6	[7.9891, 7.5821, 6.5394]
NMS	62	[8.0000, 7.5872, 6.5372]
DIRECT	186	[7.9806, 7.5823, 6.5432]

Including nonlinearities. The above problem is also solved with a similar finite element fine model, but including nonlinearities of the core. In Figure 15 the convergence history is shown for ASM applied to the nonlinear fine model with four different coarse models. The first one is the same linear model as above (5.6). The second takes the core nonlinearities into account and the other two incorporate the fringing effect at the air-gap by increasing its cross section (by the σ parameter). All these coarse models have a negligible computational cost compared with that for the fine model. Hence, the number of space-mapping iterations within ASM is still a good measure for the relative efficiency. Again, the proper selection of the coarse model appears to be an important key for the space-mapping technique since convergence can be greatly improved by an adequate choice.

---

[2] The MATLAB function *fminsearch* is used.
[3] MatLab implementation by D.E. Finkel [19] is employed.

Figure 15: The convergence history for the fine cost function in the ASM iteration. Design with the nonlinear fine model and different coarse models.
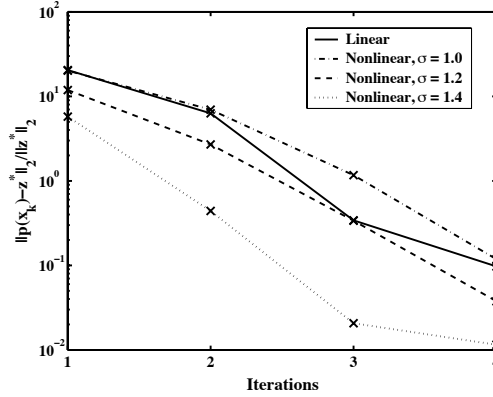
6. Conclusions

The space-mapping technique aims at accelerating expensive optimization procedures by combining problem descriptions with different degrees of accuracy. In numerical analysis, for the solution of operator equations, the same principle is known as defect correction iteration. Taking the defect correction principle as a starting point, space mapping can be interpreted, extended and analyzed.

Introducing the concept of flexibility of a model, important properties of space mapping can be formulated in these terms, and better understood. Equally flexible models are most desirable in space mapping, but this property does not generally hold for every pair of problem formulations used in practice. As a consequence, the solution found by traditional space mapping differs from the solution for the fine model. By introducing proper left-preconditioning we can improve the flexibility of the models. Thus we derive a new improved space-mapping algorithm which yields a more accurate solution. The phenomena analyzed in the paper are illustrated by a few examples, and finally a simple design problem from magnetostatics confirms once more the efficiency of the space-mapping/defect correction approach in optimization.

References

1. M.H. Bakr, J.W. Bandler, N.K. Georgieva, and K. Madsen. A hybrid aggressive space mapping algorithm for em optimization. *IEEE Trans. Microwave Theory Tech*, 47:2440–2449, 1999.

2. 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*, 1(3):241–276, 2000.

3. Mohamed H. Bakr, John W. Bandler, Radoslaw M. Biernacki, Shao Hua (Steve) Chen, and Kaj Madsen. A trust region aggressive space mapping algorithm for em optimization. *IEEE Transactions on Microwave Theory and Techniques*, 46(12):2412 – 2425, 1998.

4. 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*, 42:2536–2544, 1994.

5. J.W. Bandler, R.M. Biernacki, S.H. Chen, R.H. Hemmers, and K. Madsen. Electromagnetic optimization exploiting aggressive spacemapping. *IEEE Trans. Microwave Theory Tech*, 43:2874–2882, 1995.

6. J.W. Bandler, Q.S. Cheng, A.S. Dakroury, A.S. Mohamed, M.H. Bakr, K. Madsen, and J. Søndergaard. Space mapping: The state of the art. *IEEE Transactions on Microwave Theory and Techniques*, 52:337–360, 2004.

7. J.W. Bandler, Q.S. Cheng, D.H. Gebre-Mariam, K. Madsen, F. Pedersen, and J. Søndergaard. Em-based surrogate modeling and design exploiting implicit, frequency and output space mappings. *IEEE MTT-S International Microwave Symposium Digest*, pages 1003–1006, 2003.

8. J.W. Bandler, Q.S. Cheng, D.M. Hailu, and N.K. Nikolova. A space-mapping design framework. *IEEE Trans. Microwave Theory Tech.*, 52(11):2601–2610, 2004.

9. J.W. Bandler, D.M. Hailu, Kaj Madsen, and F. Pedersen. A space-mapping interpolating surrogate algorithm for highly optimized em-based design of microwave devices. *IEEE*

Trans. Microwave Theory Tech., 52(11):2593–2600, 2004.

10. J.W. Bandler, M.A. Ismail, J.E. Rayas-Sanchez, and Q.J. Zhang. Neuromodeling of microwave circuits exploiting space mapping technology. *IEEE Trans. Microwave Theory Tech.*, 47:2417 – 2427, 1999.

11. S. Bila, D. Baillargeat, S. Verdeyme, and P. Guillon. Automated design of microwave devices using full EM optimization method. *in IEEE MTT-S Int. Microwave Symp. Dig., Baltimore, MD*, pages 1771–1774, 1998.

12. K. Böhmer, P. W. Hemker, and H. J. Stetter. The defect correction approach. In K. Böhmer and H. J. Stetter, editors, *Defect Correction Methods: Theory and Applications*, Computing Suppl. 5, pages 1–32. Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1984.

13. K. Böhmer and H.J. Stetter. *Defect Correction Methods: Theory and Applications*. Springer, Berlin, 1984.

14. C. G. Broyden. A class of methods for solving nonlinear simultaneous equations. *Math.Comp.*, 19:577–593, 1965.

15. D.K. Cheng. *Field and Wave Electromagnetics*. Addison-Wesley, 1989.

16. H. Choi, D. Kim, I. Park, and S. Hahn. A new design technique of magnetic systems using space mapping algorithm. *IEEE Transactions on Magnetics*, 37(5):3627–3630, 2001.

17. COMSOL. FEMLAB 3: User's Guide, 2004.

18. D. Echeverria, D. Lahaye, L. Encica, and P.W. Hemker. Optimisation in electromagnetics with the space-mapping technique. *COMPEL The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, 24(3), 2005. Accepted for publication in a special issue.

19. D.E. Finkel. Direct optimization algorithm user guide. Technical Report CRSC-TR03-11, Center for Research in Scientific Computation, North Carolina State University, 2003.

20. W. Hackbush. *Multigrid methods and Applications*, volume 4 of *Series in Computational Mathematics*. Springer-Verlag, Berlin, 1985.

21. M.H. Wright J.C. Lagarias, J.A. Reeds and P.E. Wright. Convergence properties of the nelder-mead simplex method in low dimensions. *SIAM Journal of Optimization*, 9(1):112–147, 1998.

22. D.R. Jones, C.D. Perttunen, and B.E. Stuckman. Lipschitzian optimization without the lipschitz constant. *79*, (1):157–181, 1993.

23. S.J. Leary, A. Bhaskar, and A.J. Keane. A constraint mapping approach to the structural optimization of an expensive model using surrogates. *Optimization and Engineering*, 2(4), 2001.

24. K. Madsen. Minimization of non-linear approximation functions. Technical report, Technical University of Denmark, 1986. dr. techn. thesis, Institute for Numerical Analysis.

25. K. Madsen and J. Søndergaard. Convergence of hybrid space mapping algorithms. *submitted, Optimization and Engineering*, 2003.

26. D. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM*

Journal on Applied Mathematics, 11:431–441, 1963.

27. A. M. Pavio. The electromagnetic optimization of microwave circuits using companion models, 1999. presented at the IEEE MTT-S Int. Microwave Symp. Workshop, Anaheim, CA.

28. M. Redhe and L. Nilsson. Using space mapping and surrogates models to optimize vehicle crashworthiness design. *9th AIAA/ISSMO Multidisciplinary Analysis and Optimization Symp.*, (AIAA-2002-5536), 2002.

29. J. Snel. Space mapping models for rf components. *IEEE MTT-S Int. Microwave Symp. Workshop*, 2001.

30. J. Søndergaard. *Optimization Using Surrogate Models - by the Space Mapping Technique.* PhD thesis, IMM, DTU, DK-2800 Lyngby, 2003.

31. H.J. Stetter. The defect correction principle and discretization methods. *Numer. Math*, 29:425–443, 1978.

32. P.P Sylvster and R.L. Ferrari. *Finite Element for Electrical Engineers.* Cambridge University Press, New York, 1996. Third Edition.

33. H. A. van der Vorst. *Iterative Krylov Methods for Large Linear Systems*, volume 13 of *Cambridge Monographs on Applied and Computational Mathematics.* Cambridge University Press, 2003.