Multi-level optimization. Space mapping and manifold mapping
Echeverria, D.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 5

Practical Applications I: Optimization with Inexpensive Constraints

5.1 Introduction

In this chapter we apply some of the two-level optimization methods described and analyzed in Chapters 2, 3 and 4, to problems of practical relevance. Two different fields are considered: magnetostatics and steady state photon transport equations (Sections 5.2 and 5.3, respectively). In both cases we consider exact gradient information as unavailable. This circumstance, together with the intrinsic multi-level nature of a modeling procedure, makes this type of problems well suited for the space-mapping and the manifold-mapping approaches. The optimization constraints for the problems in this chapter are always of the box type, i.e., they are expressed in terms of constant upper and lower bounds for each component of the design variable \( x \). Therefore, the constraint manifolds \( k_r(\bar{X}) \) and \( k_r(\check{X}) \) coincide and thus, there is no need to introduce a mapping between them (Section 3.3.2).

This chapter is structured as follows. In Section 5.2.1 we present a reachable design in magnetostatics and with that we verify that space mapping does perform efficiently and accurately. In the example in Section 5.2.2 the perfect mapping condition holds in such an approximate way that, within the tolerance required, the space-mapping solution coincides with the true optimum. In these two cases we expect space mapping and manifold mapping to perform equally well. The design problem in Section 5.2.3 combines the difficulties of those in the previous subsections (material nonlinearities and force computation) and it
represents a case where the space-mapping and manifold-mapping solutions differ significantly. In Section 5.3 we treat a simulation-based optimization in the field of photon transport. Besides rechecking the efficiency of the two-level approach, there we see that a proper choice of the coarse model can also be used to improve design problems that are badly scaled.

The examples in Section 5.2 have been taken from publications in the field of magnetics. They are a first step towards the designs proposed in Chapter 6 (Sections 6.2.1 and 6.2.2). These problems and other applications from electronics described in that chapter have in common that the constraints are given by functions that require a significant computational cost.

5.2 Finite-Element based applications

Optimization problems with objective functions based on finite-element simulations are a suited field for applying two-level approaches [7, 30, 50, 51, 56, 74, 86]. Accurate finite-element solutions are in many cases computationally expensive and it is not complicated to find cheaper and less accurate approximations of them (for example by increasing the size of the elements). In this section we apply space mapping and manifold mapping to three optimal design problems from magnetostatics proposed earlier in [30, 77] and we compare our results with other efficient and commonly used optimization techniques. The software package FemLab [31] is used for solving the discrete finite-element formulations.

5.2.1 A C-shaped circuit (EPE1)

This design problem (denoted by EPE1) was introduced in [30] to check the performance of the SM technique. EPE1 is a two-dimensional magnetostatics problem for a C-shaped circuit. We use it as a first comparison of SM with other minimization schemes. We show that the success of the SM technique depends on the similarity between the models considered and their relative computational cost. For this problem it is possible to improve the accuracy of the coarse model without significantly increasing its execution time. This apparently results in a more efficient SM based algorithm.

A permanent magnet with residual flux density $B_r = 1.0$ T is placed between two ferromagnetic cores as shown in Figure 5.1. The sizes of the air gap, $d$ and $g$, are taken as 10 and 1 mm, respectively. In the first experiment with EPE1 the core is assumed to be a linear material with magnetic relative permeability equal to 5000. The magnet is always considered linear with unit magnetic relative permeability. The design specifications are the flux densities at the air gap center and in the core$^1$, $B_g$ and $B_c$, respectively, and the magnet permeability.

\[^1\text{Taking the upper left corner of the C-shaped circuit as the origin, the core flux density } B_c \text{ is measured at } (x_1 + d, -x_3/2).\]
Figure 5.1: EPE1: a two-dimensional magnetostatics problem.

The permeance coefficient $P_m$ (equivalently the magnet working point). The particular design objective is $y = [B_g, B_c, P_m] = [0.5, 1.0, 14]$. The design parameters $x = [x_1, x_2, x_3]$ are the dimensions of the magnet, $x_1$ and $x_2$, and the main width of the core $x_3$. The control parameter spaces $X$ and $Z$ coincide and they are equal to the positive octant of $\mathbb{R}^3$.

The fine model $f(x)$ is a finite-element resolution of a vector magnetic potential formulation of the magnetostatic equations [79]. The basis functions are quadratic Lagrangian interpolants and adaptive mesh refinement is applied, leading to discretizations of around 40000 degrees of freedom. This fine model takes into account phenomena like the fringing effect at the air gap and different flux leakages.

The coarse model $c(z)$ is a magnetic equivalent circuit (MEC) [29], as used in [30]. Here, the flux density is assumed to be confined to the magnet, the core and the air gap, and to be constant in each of these three regions. The magnetic permeability in the core is taken as infinity. No fringing effect around the air gap or flux leakage is considered. Just by applying the divergence and Stokes' theorems to the integral form of the magnetostatic equations we obtain the coarse model

$$c(z) = c([z_1, z_2, z_3]) = \begin{bmatrix} B_r, z_1 z_2 \\ \frac{d}{g z_1 + d z_2} \\ \frac{d B_r, z_1 z_2}{g z_1 z_3 + d z_2 z_3}, \frac{d z_2}{g z_1} \end{bmatrix}, \quad (5.1)$$

where the first component approximates the flux density at the air gap, the second one the flux density in the core and the third one the magnet permeance coefficient. Computing the coarse optimum $z^*$ is straightforward. It should be noted that the coarse model minimization time can be neglected with respect to that for the finite-element computation. The resulting value for $z^*$ is found in Table 5.1.
<table>
<thead>
<tr>
<th></th>
<th># evals.</th>
<th>Final design (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x^*$</td>
<td></td>
<td>[5.3571, 7.5000, 5.0000]</td>
</tr>
<tr>
<td>SM I</td>
<td>4</td>
<td>[7.9897, 7.5821, 6.5396]</td>
</tr>
<tr>
<td>SM II</td>
<td>6</td>
<td>[7.9797, 7.5806, 6.5381]</td>
</tr>
<tr>
<td>qN</td>
<td>6</td>
<td>[7.9891, 7.5821, 6.5394]</td>
</tr>
<tr>
<td>NMS</td>
<td>62</td>
<td>[8.0000, 7.5872, 6.5372]</td>
</tr>
<tr>
<td>DIRECT</td>
<td>186</td>
<td>[7.9806, 7.5823, 6.5432]</td>
</tr>
</tbody>
</table>

Table 5.1: Efficiency comparison between different optimization methods applied to EPE1. SM I: the ASM algorithm; SM II: the simplified ASM algorithm with $B_k = I$; qN: a quasi-Newton scheme; NMS: Nelder-Mead simplex and DIRECT: a direct search method. The first column indicates the number of fine model evaluations required to solve the EPE1 problem for a tolerance $\tau = 0.001$ (approximately proportional to the total computing time). The second column represents the resulting vector of design parameters. For comparison, in the first row also the coarse optimum $x^*$ is included.

**Results**

Two versions of the SM algorithm are compared with three standard optimization schemes. Because the number of design parameters is equal to the number of specifications, we may expect that the design is reachable (i.e., $f(x^*) = y$) and therefore, that the condition for a perfect mapping holds. Thus, it is justified to use the stopping criterion adopted in [30]

$$\frac{\|p(x_k) - z^*\|_2}{\|z^*\|_2} \leq \tau. \quad (5.2)$$

The advantage of (5.2) over, for example, the residual criterion $\|f(x_k) - y\|_2/\|y\|_2$ is that in the former the three vector components are well scaled (the optimal dimensions are comparable in size).

The optimization problem is solved by means of five different algorithms. ASM (see Figure 2.5) is implemented together with a variant in which $B_k$ is taken as the identity in every iteration. (This is the special case (2.36) considered in Section 2.5.3.) A quasi-Newton method was also tried for solving $f(x^*) = y$. The Jacobian was approximated by Broyden’s method and for the initial estimates for the solution and for the Jacobian we took $z^*$ and $J_c(z^*)$, respectively. As this method uses coarse model information as well, it cannot be considered completely different from two-level optimization. Since derivative information for the fine model is not easily obtained, two direct optimization schemes are used: the Nelder-Mead simplex method (NMS) [55] and the global

---

2 The MatLab function fminssearch is used.
search algorithm DIRECT [52]. The former takes the coarse optimum as the initial guess. The latter does not require a starting value but the space where the solution has to be found (Z or a subset) should be specified. (We took the cube centered in $z^*$ and with the point $z^*/2$ at a corner.) Table 5.1 shows the number of fine model evaluations for each method until the stopping criterion with $\tau = 0.001$ is satisfied. It should be emphasized that the number of evaluations is approximately proportional to the total computing time. ASM turns out to perform best, and the quasi-Newton scheme still seems to be efficient. Nelder-Mead shows a slow convergence. DIRECT's global optimization nature justifies its poor results.

In the second experiment with EPEI, the core is assumed to be made of steel and, hence, it shows a nonlinear $\mathbf{B-H}$ characteristic. The fine model is refined by taking this fact into account. We want to show how important the selection of the coarse model is for the SM approach. Therefore, we improve the coarse model above in two different ways. First, in the part of the core where its width is $x_3$, the magnetic permeability is now finite and derived from the $\mathbf{B-H}$ curve. (The treatment of the rest of the core seems to be of no relevance for the results.) Second, the fringing effect is taken into account in the coarse model by increasing the width of the air gap $d$ to $\sigma d$, with $\sigma > 1$ a correction factor. The experiments show that taking into account the fringing effect is more important than adding the nonlinearity. In Figure 5.2 the convergence history is shown for ASM applied to the nonlinear fine model with four different coarse models. The first one is

---

Footnote: 

3MatLab implementation by D.E. Finkel [42] is employed.
the same linear model as above in (5.1), the second takes the core nonlinearity into account and the other two incorporate the fringing effect at the air gap by increasing its cross section.

All these coarse models have a negligible computational cost compared with the fine model. Hence, the number of SM iterations within ASM is still a good measure for the final computing time. The proper selection of the coarse model is an important key for the SM technique since convergence can be greatly improved by an adequate choice. However, a better coarse model requires additional knowledge and insight about the system modeled.

5.2.2 A coreless actuator (EPE2)

This optimization problem (here denoted by EPE2) was proposed in [77] as a test for nonlinear optimization applied to electro-magnetic actuator design. EPE2 is a three-dimensional magnetostatic problem for an axisymmetric coreless actuator. An electro-magnetic actuator [23] is a device that converts electro-magnetic energy into mechanical force and motion. This one is called coreless because actuators usually have ferromagnetic cores (see Sections 5.2.3, 6.2.1 and 6.2.2).

Here we study EPE2 with a double aim. Apart from seeing the effect of SM compared with other efficient minimization methods, we also want to find a proper strategy for the construction of an appropriate coarse model. In view of our experience with the EPE1 experiment, the latter is probably a delicate task. We will apply a whole hierarchy of coarse models and this will provide us with some insight into the question how coarse a model can be selected in a problem like EPE2. We will verify that in this problem, the solution found by space mapping has a similar cost function value to those for the MM and GMM algorithms.

The geometry of the actuator is shown in Figure 5.3. It consists of a moving cylindrical magnet with length $l_m = 36$ mm and radius $r_m = 14.7$ mm and two fixed toroidal coils with inner radius $r_c = 15$ mm. The geometry is further specified by the design parameters $x_1$ (internal thickness), $x_2$ (half distance), $x_3$ (length) and $x_4$ (external thickness). We combine these four parameters in the control variable vector $x = [x_1, x_2, x_3, x_4]$. The design in [77] is bounded by

\[
\begin{align*}
2.5 \text{ mm} & \leq x_1 \leq 20 \text{ mm} \\
5.0 \text{ mm} & \leq x_2 \leq 25 \text{ mm} \\
2.5 \text{ mm} & \leq x_3 \leq 50 \text{ mm} \\
2.5 \text{ mm} & \leq x_4 \leq 20 \text{ mm}
\end{align*}
\] (5.3)

Both the sets $X$ and $Z$ coincide with this region.

The permanent magnet in [77] has a constant residual flux density $B_r = 0.35$ T and it is assumed to be linear with a magnetic relative permeability equal to one. The current density in the coils is taken constant, with a
value $J_c = 2.68 \text{ A/mm}^2$. The magnet can be moved from the center, along its axis over some distance $d$.

The specification is the force exerted on the rod at different locations $d$, which should be constant and equal to 5 N. By force response we mean the force (statically) computed when the magnet is positioned over a distance $d$ along the $z$-axis. For this purpose the force is determined at 20 regularly distributed displacements $d_i$ between 0 and 15 mm from the center. Thus, both the specification and the computed force are vectors of length 20.

Again, finite elements are employed in the fine model. The axisymmetric vector potential formulation of the magnetostatic equations is solved using second order Lagrangian interpolants as basis functions. The mesh is adaptively refined and the number of degrees of freedom is around 7000 – 18000. The force is computed via Lorentz’ formula [29]. This is done by post-processing the finite-element computation.

A hierarchy of coarse models

Although we know that the coarse model should be an easy-to-calculate approximation of the fine one, it is difficult to establish beforehand how (in)accurate it may be, still providing an acceleration of the optimization process. In order to investigate this question, here we build a coarse model that can be tuned as (im)precise as the user desires by means of two parameters.
Figure 5.4: A hierarchy of models for EPE2. The left figure shows the cells \((N_c = 5)\) in the coil for which the flux density is considered constant, with its value computed at the barycenter (indicated by a cross). These points also indicate the position of the loops that approximate the coil. The right figure shows the force response for different coarse models, compared with the fine model force response. In the right figure legend, FE refers to the finite-element simulation (fine model).

The first parameter, \(N_m\), refers to the treatment of the magnet. In EPE2 the effect of the magnetization is equivalent to a surface current density \(J_m = B_r/\mu_0\) [29]. In our coarse model, this current density is approximated by \(N_m\) current loops. They are equally distributed over the vertical surface of the magnet, each with radius \(r_m\) and carrying a current \(J_m l_m/N_m\). The magnetic flux density is computed by the Biot-Savart law.

The second parameter, \(N_c\), relates to the representation of the coil. Since the coil volume-current density is constant and because of the axial symmetry of the problem, the force computation is reduced to the integration of the radial component of the flux-density caused by the magnet over each coil cross-section. This cross-section is divided into \((N_c - 1) \times (N_c - 1)\) pieces. The regular partition is made by \(N_c - 1\) layers of \(N_c - 1\) cells, as shown in Figure 5.4 for \(N_c = 5\). The coarse model considers the flux density to be piecewise constant over this partitioning. The value for the radial flux density in every cell is computed at its barycenter. The integral over the cross-section is approximated by the sum of the calculated values, weighted by their cell area. This is equivalent to substituting the coil by \((N_c - 1)^2\) loops situated in the barycenter of each region and with a current equal to the one in that small area.

Thus, a coarse model is identified by the pair of integer parameters \((N_m, N_c)\). As long as the values for \(N_m\) and \(N_c\) are increased, the coarse model will be more accurate but also computationally more expensive. This behavior is reflected in
Figure 5.4 where $x_1 = x_4 = 16$ mm, $x_2 = 18.50$ mm and $x_3 = 15$ mm, and the coarse models $(N_m, N_c) = (4, 2)$, $(8, 4)$ and $(16, 8)$ are used. In that figure, FE refers to the finite-element simulation (fine model).

At first sight, the response associated with model $(4, 2)$ seems very different from the one obtained with the finite-element simulation. But apparently this does not influence its use as a coarse model. As will be seen in the next subsection, this very coarse model performs more efficiently than $(8, 4)$. The SM function compensates for the strong model misalignment and guarantees the final accuracy.

Results

In this section we compare the efficiencies for several optimization techniques and different SM strategies. The SM technique performs best. Surprisingly, extremely inaccurate coarse models lead to very efficient optimization schemes.

Unlike EPE1, the design is not reachable and hence, we cannot expect a perfect mapping. Therefore, the criterion (5.2) is no longer adequate for comparing SM with some other schemes. As in [77], we take $F(x) = \|f(x) - y\|_2/\|y\|_2$ for the cost function. Except for the methods in [77], in the experiments the solution process is stopped when this measure is below 0.08.

It is interesting to notice that there is a whole region in the parameter space $X$ where the fine model satisfies the specifications with the desired accuracy. We see in Table 5.2 how the final design parameters obtained by the different techniques are not unique. In fact it appears that there exists a whole manifold of possible solutions in $X$, so additional design constraints can be imposed. No regularization is needed in this problem to cope with difficulties arising from such manifold of solutions. In Section 6.2.1 we will see a situation where this is not the case.

The techniques analyzed in [77] are the penalty method (PM) and the method of moving asymptotes (MMA). PM reduces the cost function to 0.074 in 14 fine model evaluations and MMA to 0.083 in only 9. Both algorithms need an initial guess. In the paper [77] two of them are considered. The given results start from the point $x_1 = x_4 = 16$ mm, $x_2 = 18.50$ mm and $x_3 = 15$ mm and for both PM and MMA they are better (less fine model evaluations for an equal or smaller cost function value) than those corresponding to the other starting point.

The NMS is modified (by penalizing the cost function) in order to deal with the constraints. The initial guesses are again the same as in [77]. Now the best results are obtained when the starting point is $x_1 = x_3 = x_4 = 18$ mm and $x_2 = 10$ mm. The cost function decreases to 0.070 in 12 fine model evaluations. The performance of DIRECT is much better than in EPE1. It appears that the existence of multiple solutions makes the global search easier. The cost function drops to 0.068 in 11 fine model evaluations. A method based on sequential quadratic programming (SQP) [65] is also tested. It yields to 0.047 in 7 fine

---

The Matlab function fmincon is used.
Table 5.2: Efficiency comparison between different optimization methods applied to EPE2. PM: penalty method; MMA: method of moving asymptotes; DIRECT: direct search method; NMS: Nelder-Mead simplex; SQP: sequential quadratic programming (MatLab’s fmincon) and SM-(N_m, N_c): space mapping with coarse model (N_m, N_c). The first column indicates the total amount of computational work expressed in the equivalent number of fine model evaluations (approximately proportional to the total computing time). The second column shows the minimum cost function obtained. The design tolerance is 0.08, except for PM and MMA for which comparable results were taken from [77]. The third column represents the final design parameters.

<table>
<thead>
<tr>
<th>Method</th>
<th># f evals.</th>
<th>Cost function</th>
<th>Final design (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM</td>
<td>14</td>
<td>0.074</td>
<td>[17.20, 24.10, 23.90, 15.87]</td>
</tr>
<tr>
<td>MMA</td>
<td>9</td>
<td>0.083</td>
<td>[16.77, 23.50, 24.27, 16.48]</td>
</tr>
<tr>
<td>NMS</td>
<td>12</td>
<td>0.070</td>
<td>[16.13, 26.53, 16.46, 16.51]</td>
</tr>
<tr>
<td>DIRECT</td>
<td>11</td>
<td>0.068</td>
<td>[17.08, 21.67, 26.25, 11.25]</td>
</tr>
<tr>
<td>SQP</td>
<td>7</td>
<td>0.047</td>
<td>[10.25, 17.50, 34.00, 10.25]</td>
</tr>
<tr>
<td>SM-(8,4)</td>
<td>7.9</td>
<td>0.050</td>
<td>[11.25, 21.75, 24.49, 17.08]</td>
</tr>
<tr>
<td>SM-(4,2)</td>
<td>2.5</td>
<td>0.041</td>
<td>[14.33, 20.02, 27.16, 10.75]</td>
</tr>
<tr>
<td>SM-(6,2)</td>
<td>1.5</td>
<td>0.039</td>
<td>[17.73, 21.67, 26.97, 11.25]</td>
</tr>
</tbody>
</table>

model evaluations. The initial guess is the same as for the NMS.

Applying the SM technique entails solving two types of optimization sub-problems. The first one is the computation of the coarse optimum z'. Because the coarse model is very cheap to evaluate we can afford a global minimization method such as DIRECT for that task. The second subproblem concerns the SM function p. Since usually this function is close to the identity in a great part of the domain (see Figure 2.1), the computation of p(x) brings no significant problems. The NMS is used there and the coarse optimum z' is always taken as the initial guess.

Testing several coarse models in the hierarchy shows that the inaccurate (6,2) performs best in optimizing EPE2, within the tolerance required: only two iterations (about 2.5 fine model evaluations) are needed to reach a cost function value of 0.026, a value clearly acceptable in practice. We notice that the computational cost associated to all coarse model evaluations is taken into account in Table 5.2 (a fine model evaluation is approximately equivalent in computational time to 20, 200 or 300 of the coarse models (8,4), (6,2) or (4,2), respectively).

The manifold mapping algorithms MM and GMM are also applied to EPE2. In GMM we estimate the Jacobians of both the fine and the coarse model by means of Broyden’s method. The results are shown in Table 5.3. Both the MM and GMM scheme yield a design solution with a small cost function (0.022) in
Table 5.3: Efficiency comparison between space mapping and manifold mapping applied to EPE2.

<table>
<thead>
<tr>
<th></th>
<th># (f, c) evals.</th>
<th># f evals.</th>
<th>Cost function</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM</td>
<td>(2, 132)</td>
<td>2.4</td>
<td>0.022</td>
</tr>
<tr>
<td>GMM</td>
<td>(2, 148)</td>
<td>2.5</td>
<td>0.022</td>
</tr>
<tr>
<td>SM</td>
<td>(3, 150)</td>
<td>3.5</td>
<td>0.023</td>
</tr>
</tbody>
</table>

MM: manifold mapping; GMM: generalized manifold mapping; SM: space mapping. The second column shows the total amount of equivalent fine model evaluations needed in the optimization (approximately proportional to the total computing time). The coarse model used is \((N_m, N_c) = (4, 2)\) and it is approximately 300 times faster than the fine one.

less than three equivalent fine model evaluations. The number of coarse model evaluations is larger for the second scheme because in each iteration step the coarse model Jacobian is approximated by finite differences. In order for SM to obtain a comparable value for the cost function as for the two manifold-mapping algorithms, all coarse models from Table 5.2 are considered. The best option found is the coarse model with \((N_m, N_c) = (4, 2)\) which requires more than three equivalent fine model evaluations.

The results obtained with SM are equally cheap and as accurate as those with MM. From this fact we can conclude that (for the accuracy required) the perfect mapping condition practically holds for EPE2. However, this is not the case for the actuator design in the next section.

5.2.3 An automotive actuator (EPE25)

We denote this design problem by EPE25. Automotive actuators typically generate high levels of force and they can be used in devices such as electromagnetic switches, relays, valves, etc. Figure 5.5 is a schematic view of the cylindrical plunger electromagnet. It was originally introduced in [77] and it consists of a core and a plunger, both made of iron, and a copper coil. The sizes \(x = [x_1, x_2]\) have to be optimized in order to assure a maximum magnetic thrust force versus displacement in the \(z\)-axis. The volume of the device is kept constant. The specifications \(y\) are a constant force response of 100 N, for six vertical plunger displacements. The design space \(X\) proposed in [77] is the rectangle \([2.5, 4.5] \times [12, 18]\), where all the bounds have been specified in mm. The set \(Z\) coincides with \(X\). More details on the problem are found in [76, 77].

The fine model \(f(x)\) is based on a static force computation for the six plunger displacements. This force is evaluated numerically by quadratic Lagrangian finite elements with adaptive refinement. The number of degrees of freedom in every
Figure 5.5: EPE25: Axisymmetrical geometry and design variables of the automotive actuator.

Figure 5.6: Force responses for the fine and coarse model for EPE25 in two points of $X$.

...case is around 5000 – 7000 yielding three digits of accuracy in the force. The objective function here is $F(x) = \|f(x) - y\|_2/\|y\|_2 \times 100$, i.e., a relative Euclidean norm of the model discrepancy.

The coarse model $c(x)$ is based on a magnetic equivalent circuit (MEC) [29]...
Table 5.4: Force optimization of the automotive actuator EPE25 in the original design space.

<table>
<thead>
<tr>
<th>Method</th>
<th># fevals</th>
<th>Cost function</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMA</td>
<td>7</td>
<td>49.0</td>
</tr>
<tr>
<td>PM</td>
<td>9</td>
<td>48.8</td>
</tr>
<tr>
<td>EM</td>
<td>13</td>
<td>48.8</td>
</tr>
<tr>
<td>RQP</td>
<td>7</td>
<td>48.8</td>
</tr>
<tr>
<td>SQP</td>
<td>30</td>
<td>48.5</td>
</tr>
<tr>
<td>SM</td>
<td>1.0</td>
<td>48.5</td>
</tr>
<tr>
<td>MM</td>
<td>1.0</td>
<td>48.5</td>
</tr>
</tbody>
</table>

MMA: method of moving asymptotes; PM: penalty method; EM: ellipsoid method; RQP: recursive quadratic programming; SQP: sequential quadratic programming; SM: space mapping; MM: manifold mapping. The first column indicates the total amount of computational work expressed in the equivalent number of fine model evaluations (approximately proportional to the total computing time). The second column shows the minimum cost function obtained.

derived from the actuator. The MEC is nonlinear, in the sense that a \( B-H \) curve is used for the plunger. An infinite core magnetic permeability is taken and neither fringing nor leakage effects are included. A linear interpolation is applied for the force between the points for maximum and minimum displacement. This coarse model is around 400 times faster than the fine one. In Figure 5.6 the fine and coarse model responses are shown for two points in the design space \( X \): the geometrical center \( x_0 \) and the optimum obtained by MM. Each optimization involving the coarse model is solved by sequential quadratic programming (SQP).

Table 5.4 compares the performance of seven minimization methods. The figures in the first column (amount of work expressed in number of equivalent fine model evaluations) are approximately proportional to the total computing time. All seven techniques yield solutions with almost the same objective function value. The first two (MMA: method of moving asymptotes, PM: penalty method) correspond with the methods from [77] commented in Section 5.2.2, and the next two (EM: ellipsoid method, RQP: recursive quadratic programming) to other schemes studied by the same author in a different work [76]. The coarse model optimum \( z^* = [4.5 \text{ mm}, 18 \text{ mm}] \) has a lower fine cost function than any of the solutions given by the methods above. SM and MM converge to \( z^* \) (i.e., the coarse model optimum \( z^* \) is a minimizer of \( F(x) \)) after just one fine model evaluation (the number of coarse model evaluations is also taken into account in the amount of work shown in Table 5.4). SQP applied to the fine model and starting from the coarse model optimum performs also very efficiently.
Table 5.5: Force optimization of the automotive actuator EPE25 in the modified design space.

<table>
<thead>
<tr>
<th></th>
<th># fevals.</th>
<th>F(x)</th>
<th>Final design (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SQP</td>
<td>12.1</td>
<td>45.0</td>
<td>[5.00, 16.45]</td>
</tr>
<tr>
<td>MM</td>
<td>3.2</td>
<td>45.0</td>
<td>[5.00, 16.61]</td>
</tr>
<tr>
<td>SM</td>
<td>6.9</td>
<td>45.3</td>
<td>[5.00, 15.00]</td>
</tr>
</tbody>
</table>

SQP: sequential quadratic programming; MM: manifold mapping; SM: space mapping. The first column indicates the total amount of computational work expressed in the equivalent number of fine model evaluations. The second and third columns show the minimum cost function obtained and the resulting vector of design parameters, respectively. MM and SM stop when the reduction in the cost function value is smaller than 0.01. SQP takes the coarse model optimum z* as its initial guess and iterates, for comparison purposes, until the cost function value decreases below 45.0.

Next, the space X is slightly enlarged to [2.5, 5] × [15, 30] (again in mm) in order to obtain a design in which the coarse optimum does not solve the optimization process\(^5\). Thus, a better comparison can be made between the three methods that perform similarly above. Now, the coarse optimum is the point [4.68 mm, 17.72 mm] which has an associated cost function of 46.5. The optimization results are given in Table 5.5. MM and SQP yield both a solution with the same quality, but MM is almost four times faster. SM improves also SQP in efficiency but the result obtained is not the optimal design. This experiment corroborates MM as an efficient method with the additional property of convergence towards the fine optimum. GMM, with the fine model Jacobian estimated via Broyden’s method, performs identically to MM.

5.3 A design problem in photon transmission

This optimization problem, introduced in [38], has practical relevance in photonics. Unlike the previous designs, its associated fine model is not based on finite elements. The phenomenon under analysis is the photon transmission and absorption in random media [70]. For example, the propagation of light in biological tissues can be studied within this framework. We will verify again that both the MM and SM approaches solve the design problem efficiently and that the former yields a more accurate solution than the latter. Additionally, we will

---

\(^5\)This modification of the design space is introduced for testing purposes. As observed in [76], large values for \(x_1\) could lead to devices with significant mass and a poor dynamic regime performance.
Figure 5.7: The photon transmission takes places through a cylinder which contains random media. The two-dimensional design variable $\mathbf{x} = [x_1, x_2]$ represents the absorption coefficients of the random media considered.

see that coarse models can be used as an efficient means for improving situations where a bad model scaling complicates the whole optimization process.

The setup is a cylinder of length $L$ which contains two different materials distributed randomly (see Figure 5.7). Photons are emitted at one end of the cylinder (source) and collected at the other end (target). During their travel, some of the photons will be scattered and some others will be absorbed. The one-dimensional steady state transport equation [70] is solved over the cylinder. Due to the stochastic nature of the media, Monte Carlo simulations are performed.

The model output is two-dimensional: the first component is the average transmission coefficient and the second one is its variance. We perform $N$ direct numerical simulations of the transport equation in order to estimate these expected values. The two-dimensional design variable $\mathbf{x} = [x_1, x_2]$ represents the absorption coefficients of the random media considered. The sets $X$ and $Z$ coincide and they are equal to $[1, 10]^2$.

The model is considered to be fine enough for practical purposes when $L = 1$ (complete cylinder) and when the number of simulations is $N = 100000$. Thus, we characterize the fine model $f(\mathbf{x})$ by the pair $(L, N) = (1, 100000)$. We can obtain cheaper (and less accurate) models either by considering a smaller domain for solving the transport equation (i.e., $L < 1$) or by performing much less transport simulations (i.e., $N < 100000$). This is shown in Figure 5.8 and Figure 5.9 where the transmission coefficient and its variance are computed for different
Figure 5.8: Different models for the transmission coefficient. The second component of the design variable $x_2$ is kept constant and equal to five. The length $L$ is taken as one in the left figure. The number of realizations $N$ is 100000 for the right plot.

Figure 5.9: Different models for the transmission variance. The second component of the design variable $x_2$ is kept constant and equal to five. The length $L$ is taken as one in the left figure. The number of realizations $N$ is 100000 for the right plot.

values of $L$ and $N$. There, the second component of the design variable $x_2$ is kept constant and equal to five. (When $x_1 = x_2$ the cylinder contains just one medium and then the transmission variance is zero.) We observe that for the transmission coefficient, $L$ and $N$ can be reduced to a very low value without
Table 5.6: Results for the reachable photon design problem.

<table>
<thead>
<tr>
<th>#f evals</th>
<th>$F(\cdot)$</th>
<th>Final design</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMS</td>
<td>74.2</td>
<td>0.000</td>
</tr>
<tr>
<td>SM</td>
<td>55.8</td>
<td>0.000</td>
</tr>
<tr>
<td>MM</td>
<td>16.9</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#f evals</th>
<th>$F(\cdot)$</th>
<th>Final design</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMS</td>
<td>66.3</td>
<td>0.020</td>
</tr>
<tr>
<td>SM</td>
<td>15.7</td>
<td>0.020</td>
</tr>
<tr>
<td>MM</td>
<td>13.6</td>
<td>0.020</td>
</tr>
</tbody>
</table>

NMS: Nelder-Mead simplex; SM: space mapping; MM: manifold mapping. The first column indicates the equivalent number of fine model evaluations required to solve the optimization problem. $F(\cdot)$ represents the cost function. The stopping criterion is in every case $h_k = \|x_{k+1} - x_k\|_2 < 10^{-4}$. The coarse model employed is around forty times faster than the fine one.

(apparently) losing much accuracy in the model. The transmission variance is very sensitive to the varying of the length $L$, but concerning the number of realizations $N$, the behavior is analog to the one observed in Figure 5.8. Thus, the same pair $(L, N)$ can be acceptable for computing the transmission coefficient but not for its variance.

We will take a coarse model $g(\bar{z})$ given by the pair $(L, N) = (1, 100)$. This model is around forty times faster than the fine one. The cost function considered is again $F(\bar{z}) = \|f(\bar{z}) - c(\bar{z})\|_2/\|\bar{z}\|_2 \times 100$, a relative measurement of the model discrepancy. Every optimization process stops when the increase of the Euclidean norm of the design parameters is smaller than $10^{-4}$, i.e., $h_k = \|x_{k+1} - x_k\|_2 < 10^{-4}$.

In this design problem we first try a reachable design. The specifications are $\bar{y} = f([3.4567, 7.6543]) = [7.0348, 0.0002]$. We observe from Figure 5.8 and Figure 5.9 (and also from $\bar{y}$) that the transmission variance takes much smaller values than the transmission coefficient. This difference in size suggests that a scaling could be necessary since the cost function considered gives more emphasis to the first component of every model than to the second one. Consequently, values far away from the fine model optimum could yield, for practical purposes, the same cost function value.

We can use inexpensive coarse model evaluations in order to approximate a scaling for the fine model. We proceed as follows. An average value of the transmission variance is first computed based on the coarse model. Then, the output of both fine and coarse models is weighted according to that value. (From
another perspective, a scaled cost function is used.

The results of this reachable two-dimensional optimization problem are shown in Table 5.6. Space mapping (with NMS for the coarse model optimizations) leads to the fine model optimum (again perfect mapping holds). However, SM takes almost the same computational effort as NMS. The MM algorithm (with NMS for all the coarse model optimizations) yields a speed-up of around a factor four. This acceleration value is similar to those obtained in the previous section.

If no scaling is introduced, the problem is ill-conditioned and the solutions obtained differ clearly from the fine optimum ([3.2784, 7.7062] for NMS and [3.3043, 7.6975] for MM).

If manifold mapping is applied with the coarse model characterized by \((L, N) = (0.01, 100000)\) (which is around thirty times faster than the fine one), the solution [3.4516, 7.6557] is obtained after 85.6 equivalent fine model evaluations. The associated computational cost is even higher than that for the one-level NMS scheme. If we use the model given by \((L, N) = (1, 100)\) just for computing the coarse model optimum \(z^*\) and then we continue with \((L, N) = (0.01, 100000)\), the resulting procedure accelerates NMS by a factor of three ([3.4526, 7.6554] is obtained after 26.9 equivalent fine model evaluations). We can infer that intermediate coarse models can be used for computing better initial guesses and with that improve a previously observed speed-up.

We also try a non-reachable design, given by the specifications \(y = [5, 0.004]\). The second component of \(y\) is about twenty times larger than in the reachable case and this seems to be enough for not needing the additional weighting. The results without scaling are shown in Table 5.7. The speed-up factors for SM and MM are five and four, respectively. With the use of scaling, the convergence behavior is about the same and the limit point is [1.0000, 5.6710] for SM and [1.0007, 5.6967] for MM.

Though in this design problem a hierarchy of models is characterized by the pair \((L, N)\), the difference in computational cost between the fine and coarse models considered is not big enough to fully exploit the advantages of the multi-level approach. Moreover, a really simple model as \((L, N) = (1, 100)\) presents no important convergence difficulties in the two-level iteration. An intermediate model would presumably be of not much help. If higher accuracy is desired in solving the transport equation, the distance between the fine and coarse models will increase, and then, we believe that a multi-level scheme could offer some noticeable improvement over the two-level strategy.

5.4 Conclusions

In this chapter we have verified that the two-level optimization approach is an efficient strategy for solving various design problems of practical relevance. The constraints in these problems do not require special care because they are inex-
pensive to evaluate. The first three examples (EPE1, EPE2 and EPE25) have a cost function based on solving a finite-element discretization of the magneto-statics equations. EPE1 is a reachable design that can be solved by classical space mapping. It also shows how improving a coarse model can lead to an improvement in performance. EPE2 and EPE25 are examples of electro-magnetic actuator design. EPE2 represents a case where the perfect mapping condition holds approximately and thus, the solution found by space mapping coincides (within the tolerance needed) with the true optimum. In EPE2 a hierarchy of models is available and there even a very simple model can accelerate the optimization process significantly. In EPE25 there is no perfect mapping and classical space mapping yields a solution with a cost function noticeably larger than the fine model optimum. The last optimization problem is taken from the field of photon transport in random media. The fine model requires the solution of the discretized steady state transport equation. The results are consistent with those from the previous three examples, in the sense that space mapping and manifold mapping are also efficient minimization techniques. Moreover, we see with the photon transport example that coarse models can be beneficial when used with model scaling purposes. The design problems in the next chapter are comparable in complexity to those studied here. However, they carry the additional difficulty of having optimization constraints that are described by functions of considerable computational cost.