Deutsche Forschungsgemeinschaft

Priority Program 1253
Optimization with Partial Differential Equations

D. Clever, J. Lang and D. Schröder

Model Hierarchy Based Optimal Control of Radiative Heat Transfer

October 2012

Preprint-Number SPP1253-133

http://www.am.uni-erlangen.de/home/spp1253
Model Hierarchy Based Optimal Control of Radiative Heat Transfer

Debora Clever  
*Department of Mathematics, Graduate School of Computational Engineering, Technische Universität Darmstadt, Dolivostr. 15, D-64293 Darmstadt, Germany*  
clever@mathematik.tu-darmstadt.de

Jens Lang  
*Department of Mathematics, Center of Smart Interfaces, Graduate School of Computational Engineering, Technische Universität Darmstadt, Dolivostr. 15, D-64293 Darmstadt, Germany*  
lang@mathematik.tu-darmstadt.de

Dirk Schröder  
*Department of Mathematics, Technische Universität Darmstadt, Dolivostr. 15, D-64293 Darmstadt, Germany*  
schroeder@mathematik.tu-darmstadt.de
Abstract

We present a model hierarchy multilevel optimisation approach to solve an optimal boundary control problem in glass manufacturing. The process is modelled by radiative heat transfer and formulated as an optimal control problem restricted by partial differential algebraic equations (PDAE) and additional control constraints. We consider a sequence of model approximations given by space-time dependent non-linear PDAEs of ascending accuracy. The different models allow for a model hierarchy based optimisation approach, where the models are shifted automatically as the optimisation proceeds. We present a realisation of a multilevel generalised SQP method within the fully space-time adaptive optimisation environment KARDOS using linearly implicit methods of Rosenbrock type and multilevel finite elements. We apply the optimal control algorithm to a glass cooling problem and present numerical experiments for the model hierarchy based approach in two spatial dimensions and for a fully space-time adaptive optimisation in three spatial dimensions.

Keywords: optimal boundary control; multilevel; SQP-methods; control constraints; adaptivity; Rosenbrock methods; 3D finite elements; model hierarchy; partial differential algebraic equations; PDAE; radiative heat transfer
1 Introduction

During the last decade it has been becoming of growing interest not only to simulate the behaviour of engineering, medical or financial applications but also to optimise their input such that the resulting output follows a desired profile. Usually, the considered process is described by a system of space-time dependent partial differential equations, possibly coupled with algebraic constraints (PDAEs). Mathematically, this results in a so called PDAE-constrained optimal control problem, in general with additional constraints on control and state. For real-world applications, the bottleneck of solving such problems is the high complexity of the involved PDAEs, which have to be solved several times within each optimisation iteration. Therefore, an efficient optimisation environment has to combine the following two aspects: (i) an optimisation technique with a high order of convergence, such that the number of PDAE solves is brought to a moderate level, and (ii) a fully space-time adaptive PDAE solver of high order, such that the involved PDAEs are solved as efficiently as possible. To even gain more efficiency without loss of accuracy for the optimal control, multilevel techniques are an attractive instrument.

In this work we present an optimal boundary control problem for the cooling process of hot and already formed glass, which is an important task in glass manufacturing. The cooling process is steered within a furnace, such that the furnace temperature acts on the glass surface only. The task is now to choose a furnace temperature profile that enforces a certain glass temperature evolution, to promote chemical reactions and minimise internal stresses, both at moderate cost. Furthermore, the furnace temperature has to be restricted to a feasible set due to the operation interval of the oven. Because of the high temperatures that occur especially at the beginning of the cooling process, the direction- and frequency-dependent thermal radiation field and the spectral radiative properties of semi-transparent glass play a dominant role. In Section 2 we introduce a seven-dimensional cooling model that couples the radiative heat transfer equations with the evolution of the glass temperature in dependence on the furnace temperature. Since this model is quite expensive, especially for optimisation purposes, we derive suitable approximations, using simplified spherical harmonics and a practically relevant frequency bands model (Farina et al., 2010; Golbarm, 1961; Klar et al., 2005; Larsen et al., 2002; Pinnau and Thömmes, 2004). The different models allow for a model hierarchy based optimisation approach, where the optimisation is started on the cheapest model. While the optimisation proceeds, the models are shifted automatically, such that the optimal control is carried out on the most accurate model at the end.

In Section 3 we formulate the cooling process as a PDAE-constrained optimal control problem with control constraints and present a realisation of a multilevel generalised SQP method (Ziems, 2010; Ziems and Ulbrich, 2011) within the fully space-time adaptive optimisation environment KARDO (Clever et al., 2010, 2012; Erdmann et al., 2002). Control constraints are handled by an appropriate projection. Reduced gradients and actings of the reduced Hessian are computed with the continuous adjoint approach. We follow Rothe’s method with adaptive Rosenbrock methods in time and adaptive multilevel finite elements in space. To be able to choose the discretisation scheme in accordance to the structure of the considered PDAE, we explicitly allow for an independent discretisation of state and adjoint systems. The resulting inexactness is controlled by refining grids adaptively in space and time as the optimisation proceeds.

In Section 4 we apply the presented algorithm to the glass cooling problem. Note that the presented environment is not restricted to the solution of glass cooling problems. It is a suitable optimisation tool for a great variety of boundary control problems restricted by space-time dependent PDAEs of similar type and constraints on the control. For more details on the class of PDAEs that can be handled we refer to Erdmann et al. (2002). The Section is divided into two subsections. In Subsection 4.1 we solve the optimal control problem for the highest model using the model hierarchy based approach and compare its performance to a similar optimisation run, carried out on the highest model only. Due to the high complexity of the considered model, and the high computing time in three spatial dimensions, we first approximate the three dimensional computational domain by a two dimensional cross section. In Subsection 4.2, we then consider the entire domain and solve the optimal control problem in three space dimensions, considering the less complex grey scale model.
2 Glass Cooling

One important step in glass manufacturing is the cooling of the hot and already formed glass down to room temperature. Because the quality of the final product depends highly on the temperature evolution within the glass during the cooling process, there is the need to control the behaviour of the glass temperature. To this end, the hot glass is cooled within a furnace, which is preheated in the beginning. Choosing an optimal course for the temperature reduction within the oven, the temperature evolution within the glass can be influenced in such a way that the resulting product is of high quality.

To be able to compute such an optimal boundary control, it is necessary to derive a suitable model of the cooling process. Because of the high temperatures that occur especially at the beginning of the cooling process, the direction- and frequency-dependent thermal radiation field and the spectral radiative properties of semi-transparent glass play a dominant role. In the following we describe radiation by its intensity $I(x,t,\nu,s)$, which depends on the spatial variable $x \in \Omega \subset \mathbb{R}^d$, $d=2,3$, time $t \in [0,t_c)$, frequency $\nu \in [0,\infty)$, and direction $s \in S^2$ of the unit sphere. Since the prolongation of energy caused by radiation is significantly faster than the one caused by diffusion, it is reasonable to model radiation in a quasi-static manner. The furnace temperature, denoted by $u(t)$, is assumed to be constant in space. Note that, based on the physical background, it occurs only in the boundary conditions. The prolongation of the glass temperature $T(x,t)$ is modelled by the heat equation and the non-linear exchange of energy between radiation field and glass temperature $T(x,t)$. We write the model in a dimensionless form, such that the involved quantities are arranged more clearly. Multiplication with appropriate reference values brings the dimensionless quantities back to its physical counterparts (see e.g. Klar et al. (2005)). The equations read as follows:

$$
e^2 \partial_t T - e^2 \nabla \cdot (k_e \nabla T) = - \int_{\Omega} \kappa_B (B(T,\nu) - I(x,t,\nu,s)) \, ds \, d\nu,$$

(1)

$$\epsilon \cdot \nabla I(x,t,\nu,s) + (\sigma_\nu + \kappa_\nu) I(x,t,\nu,s) = \frac{\sigma_\nu}{4\pi} \int_{\Omega} I(x,t,\nu,s) \, ds + \kappa_B B(T,\nu), \quad \forall \nu > \nu_0,$$

(2)

with boundary and initial conditions

$$\epsilon k_e n \cdot \nabla T = h_e (u - T) + \alpha \pi \left( \frac{n_g}{n_0} \right)^2 \int_0^{\nu_0} (B(u,\nu) - B(T,\nu)) \, d\nu,$$

(3)

$$I(x,t,\nu,s) = r(n \cdot s) I(x,t,\nu,s') + (1 - r(n \cdot s)) B(u,\nu), \quad \forall (x,s) \in \partial \Omega \times S^2 : n(x) \cdot s < 0,$$

(4)

$$T(x,0) = T_0(x).$$

(5)

The non-linear exchange of energy between radiative field and the glass itself is described by the scaled Planck function

$$B(T,\nu) = \frac{n_g^2}{\epsilon_0^2 \lambda_{ref}} \frac{2h_e \nu^3}{c^4 \mu (c^4 \nu/(k_B T_{ref}) - 1)},$$

(6)

with Planck constant $h_p = 6.626 \times 10^{-34} \text{Js}$, Boltzmann constant $k_B = 1.381 \times 10^{-23} \text{J}^\circ\text{K}$, the speed of light in vacuum $c_0 = 2.998 \times 10^8 \text{m/s}$, and the reference quantities for the glass temperature $T_{ref}$ and the intensity $I_{ref}$, respectively. Generally, we set $T_{ref} = 1 \text{K}$ and $I_{ref} = 1 \text{W/m}^2\text{sr}$, such that the values of the dimensionless quantities coincide with their dimension-assigned counterpart. For more details concerning the dimensionless model, we e.g. refer to Farina et al. (2010); Larsen et al. (2002).

The definition of the remaining dimensionless parameters is given in Table 1. The concrete values that are used in the numerical experiments are listed in Table 3.

The challenge in solving system (1)-(5) lies in

- the differential algebraic structure, caused by the different time scales of radiation and diffusion,
- the highly non-linear coupling of the glass temperature with the radiative field,
- the six or even seven dimensional phase space.

The high dimension of the phase space makes the numerical solution of the full radiative heat transfer equation (1)-(5) very expensive, which is especially demanding for optimisation purposes, where the system has to be solved several times. Various approximate models that are less time consuming, yet sufficiently accurate, have been developed (Klar et al., 2005; Larsen et al., 2002; Turpault et al., 2004; Turpault, 2005).
Table 1: Dimensionless parameters of radiative heat transfer model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\epsilon$</td>
<td>optical thickness coefficient, with $0 &lt; \epsilon \leq 1$ for an optically thick, diffusive regime</td>
</tr>
<tr>
<td>$k_c$</td>
<td>heat conduction coefficient</td>
</tr>
<tr>
<td>$h_c$</td>
<td>convective heat transfer coefficient</td>
</tr>
<tr>
<td>$\kappa_\nu$</td>
<td>frequency-dependent absorption coefficient</td>
</tr>
<tr>
<td>$\sigma_\nu$</td>
<td>frequency-dependent scattering coefficient</td>
</tr>
<tr>
<td>$n_g$</td>
<td>refractive index: ratio of the speed of light in vacuum and in glass</td>
</tr>
<tr>
<td>$n_a$</td>
<td>refractive index: ratio of the speed of light in vacuum and in air</td>
</tr>
<tr>
<td>$\nu_0$</td>
<td>upper bound of the opaque spectral region $[0, \nu_0]$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>mean hemispheric surface emissivity in the opaque spectral region</td>
</tr>
<tr>
<td>$r$</td>
<td>reflectivity coefficient, with $r \in [0, 1]$</td>
</tr>
<tr>
<td>$n(x)$</td>
<td>outwards normal on glass surface</td>
</tr>
</tbody>
</table>

2.1 The Models

In the following, we use a first order approximation of simplified spherical harmonics (SP$_1$), including a practically relevant frequency bands model. The main idea of this approach is to substitute the direction dependent radiative intensity by an integral mean using asymptotic and variational analysis. The idea of simplified spherical harmonics (SP$_n$) goes back to Gelbard (1961). In comparison to the spherical harmonics approach (P$_n$), which results in a system of $(n + 1)^2$ coupled equations, the SP$_n$ approximation only results in a system of $(n + 1)/2$ equations (Zeng and Han, 2011).

In a second step, the continuous frequency spectrum is discretised into $N$ disjoint groups (also referred to as bands), assuming constant coefficients on each of these bands. By varying the number $N$ of frequency bands, we derive three different models of rising complexity and accuracy.

The considered $N$-band SP$_1$-approximations have been tested fairly extensively for various radiation transfer problems in glass and have proven to be an efficient way to improve the classical diffusion approximations (Klar et al., 2005; Larsen et al., 2002).

Note, that there are many other approaches to efficiently reduce the complexity of the full radiative heat transfer system, including grey and multigroup, half and full space moment models (Turpault, 2005; Turpault et al., 2004) and that there is still active research in improving these approximations (Turpault, 2010; Zeng and Han, 2011).

The Augmented Rosseland Approximation. In its classical form, the Rosseland approximation, which can be also interpreted as an SP$_0$ approximation, lacks in a proper description of boundary layer effects. Considering the SP$_1$ approximation of (1)-(5) and replacing the equation of transfer by a pure algebraic equation lead to a significant improvement of the Rosseland approximation, where the boundary conditions are augmented by a radiative term.

The augmented Rosseland approximation results in a space-time dependent partial differential equation of mixed parabolic-elliptic type in one component with state $y := T$, given by

$$\partial_T - \nabla \cdot ((k_c + k_r(T))\nabla T) = 0,$$

equipped with boundary and initial conditions

$$n \cdot (k_c + k_r(T))\nabla T = \frac{h_c}{\epsilon} (u - T) + \frac{\alpha \pi}{\epsilon} \left( \frac{n_a}{n_g} \right)^2 \int_0^{\nu_0} (B(u, \nu) - B(T, \nu))d\nu$$

$$+ \frac{4\pi a_1}{\epsilon} \int_{\nu_0}^\infty (B(u, \nu) - B(T, \nu))d\nu,$$

$$T(x, 0) = T_0(x),$$

the boundary condition coefficient $a_1 = 1.149e - 1$ and the thermal conductivity

$$k_r(T) = \frac{4\pi}{3} \int_{\nu_0}^\infty \frac{1}{\sigma + \kappa} \partial_T B(T, \nu)d\nu,$$

which is caused by radiation. To evaluate $k_r$, we divide the continuous frequency spectrum into eight bands with piecewise constant absorption coefficients $\kappa_i$, $i = 1, \ldots, 8$, (see Table 2).
The Grey Scale Model. Considering the SP$_1$ approximation of (1)-(5) and discretising the continuous frequency spectrum into eight bands $[\nu_{i-1}, \nu_i], i = 0, \ldots, 8$, where we formally set $\nu_8 := \infty$ and $\nu_{-1} := 0$, see Figure 1. On each of the bands we interpret frequency dependent quantities as constants. The values are given in Table 2. Defining the frequency-independent mean of the Planck function and the radiated energy coefficient $a_2 = 1.8 e^{-8}$. Note, that the absorption coefficient $\kappa$ and the scattering coefficient $\sigma$ are assumed to be constant.

The Eight Band Model. We again consider the SP$_1$ approximation of (1)-(5), but now we discretise the continuous frequency spectrum into eight bands $[\nu_{i-1}, \nu_i], i = 0, \ldots, 8$, where we formally set $\nu_8 := \infty$ and $\nu_{-1} := 0$, see Figure 1. On each of the bands we interpret frequency dependent quantities as constants. The values are given in Table 2. Defining the frequency-independent mean of the Planck function

$$B^{(i)}(\nu) := \int_{\nu_{i-1}}^{\nu_i} B(\nu, \nu) d\nu, \quad i = 1, \ldots, 8,$$

the eight band model is given by the following system of space-time dependent partial differential algebraic equations of mixed parabolic-elliptic type in nine components with state $y := (T, \phi_1, \ldots, \phi_8)^T$:

$$\partial_t T - k_c \Delta T = \frac{1}{3(\kappa + \sigma)} \Delta \phi = 0$$

with boundary and initial conditions

$$k_c n \cdot \nabla T + \frac{1}{3(\kappa + \sigma)} n \cdot \nabla \phi = \frac{h_c}{\epsilon} (u - T) + \frac{1}{2\epsilon} (4\pi a_2 u^4 - \phi)$$

$$-\frac{\epsilon^2}{3(\kappa + \sigma)} n \cdot \nabla \phi = \frac{\epsilon}{2} (4\pi a_2 u^4 - \phi)$$

$$T(x, 0) = T_0(x),$$

and the radiated energy coefficient $a_2 = 1.8 e^{-8}$. Note, that the absorption coefficient $\kappa$ and the scattering coefficient $\sigma$ are assumed to be constant.
Figure 1: The continuous frequency spectrum is approximated by a discrete eight-band model with constant coefficients on each interval.

Table 2: Absorption coefficients of eight band model kindly provided by (Fraunhofer-Institut für Techno- und Wirtschaftsmathematik, Kaiserslautern, Germany)

<table>
<thead>
<tr>
<th>Band i</th>
<th>(\nu_{i-1}(10^{13}\text{s}^{-1}))</th>
<th>(\nu_{i}(10^{13}\text{s}^{-1}))</th>
<th>(\kappa_{i}(\text{m}^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>2.9334638</td>
<td>opaque</td>
</tr>
<tr>
<td>1</td>
<td>2.9334638</td>
<td>3.4223744</td>
<td>7136.00</td>
</tr>
<tr>
<td>2</td>
<td>3.4223744</td>
<td>3.7334994</td>
<td>576.32</td>
</tr>
<tr>
<td>3</td>
<td>3.7334994</td>
<td>4.5631659</td>
<td>276.98</td>
</tr>
<tr>
<td>4</td>
<td>4.5631659</td>
<td>5.1356161</td>
<td>27.98</td>
</tr>
<tr>
<td>5</td>
<td>5.1335616</td>
<td>5.8669276</td>
<td>15.45</td>
</tr>
<tr>
<td>6</td>
<td>5.8669276</td>
<td>6.8447489</td>
<td>7.70</td>
</tr>
<tr>
<td>7</td>
<td>6.8447488</td>
<td>102.6712329</td>
<td>0.50</td>
</tr>
<tr>
<td>8</td>
<td>102.6712329</td>
<td>(\infty)</td>
<td>0.40</td>
</tr>
</tbody>
</table>

2.2 Model Hierarchy

With the three models described above, we have a sequence of approximations to the full radiative heat transfer model with ascending accuracy and complexity. To determine an optimal furnace temperature evolution for a real cooling process, it is clearly desirable to consider the most accurate model. However, from an efficiency point of view the lowest model is the natural candidate of choice, especially for optimisation purposes. To serve both accuracy and efficiency, we consider the optimal control problem on a model hierarchy, where most of the optimisation iterations are carried out on cheap models but where at least the last iteration, and hence the optimal control, is computed using the best model available.

Let a sequence of models \(M_{i}\), \(i = 1, \ldots, m\), with ascending accuracy be given. We take for \(M_1\) the Rosseland approximation, for \(M_2\) the grey scale, and for \(M_3\) the eight band model. Furthermore, we define some criticality measure \(C_{M}\), which decreases while the control iterates approach the optimal control. We now shift from the current model \(M_i\) to the next higher order model \(M_{i+1}\) if an appropriate error estimate \(E(M_i, M_{i+1})\) exceeds the scaled criticality measure, i.e.,

\[
\text{shift from model } M_i \text{ to model } M_{i+1}, \text{ if } E(M_i, M_{i+1}) > k_i C_{M}. \tag{21}
\]

For the glass cooling problem, it is reasonable to estimate the error between the different models by considering the difference in the glass temperature over space and time. Note that neither the mean intensities nor the entire state are a meaningful candidate, since they differ in the three models. For \(T: \Omega \times [0, t_e) \rightarrow \mathbb{R}\), we define

\[
E(M_i, M_{i+1}) := \|T_{M_i} - T_{M_{i+1}}\|_{L^2(\Omega \times [0, t_e)])} = \left( \int_0^{t_e} \int_{\Omega} (T_{M_i}(x,t) - T_{M_{i+1}}(x,t))^2 dx dt \right)^{\frac{1}{2}}, \tag{22}
\]

which can be evaluated with nearly no extra effort, see Subsection 3.3. An implementation of a model hierarchy based SQP method and a comparison to the corresponding realisation on the highest order model is presented and discussed in Section 4.
To determine a furnace temperature profile that enforces an efficient cooling while maintaining a high quality of the manufactured glass we interpret the setting as an optimal boundary control problem. To this end we define an objective functional that measures the quality of the furnace temperature profile and the resulting temperature evolution within the glass. The optimal control is computed by minimising this objective with respect to the considered glass cooling model and with upper and lower bounds for the furnace temperature. To this end, we use the fully-space time adaptive optimisation environment developed in Clever et al. (2010, 2012). It includes several derivative based optimisation algorithms based on the so called continuous adjoint approach, which contains the following two aspects. The first aspect is, that the derivatives are evaluated based on backwards sensitivities rather than forward sensitivities. Hence, the evaluation of the reduced gradient for example, requires only two PDAE solves, that of the state system and that of the adjoint system, independently of the dimension of the control space discretisation. In contrast, in a forward sensitivity based approach the effort rises with the dimension. The second aspect is, that the optimality systems are derived in the context of infinite dimensional vector spaces. This has the advantage that any discretisation of one of the up to four resulting PDAEs coincides naturally with its underlying PDAE, which therefore allows the use of independent space-time grids and even independent integration and discretisation schemes.

More details on the optimal control problem and the determination of the derivatives are given in Subsection 3.1. In Subsection 3.2 we introduce a multilevel generalised SQP method (Clever et al., 2010, 2012; Ziems, 2010; Ziems and Ulbrich, 2011), which allows for point-wise constraints on the control. Its coupling with the state-of-the-art PDAE solver KARDOS is presented in Subsection 3.3. Numerical experiments in Clever et al. (2010, 2012) show a significant reduction of computing time when using this method in comparison to a gradient method or an SQP method without multilevel strategy.

### 3.1 Optimal Control Problem

To determine an appropriate optimal control, it is essential to formulate a sound objective. In the context of glass cooling we deal with at least two contrary criteria. An important aim is to force the glass temperature function as close as possible to a desired temperature profile. Such a profile, for which good performance of the involved chemical processes is known, is generally given by engineers. A common approach is to choose the tracking function for the glass temperature spatially constant in order to enforce a homogeneous cooling with small temperature gradients. This is necessary to reduce internal stresses and avoid cracks within the glass. Note that, because the cooling is controlled at the boundary only, such a guiding function can only be approached but not necessarily reached within the entire domain. Furthermore, it is desirable to pay certain attention to the glass temperature at the final time, which is realised by an additional term in the objective. Especially in the context of the continuous adjoint calculus, such a term is of great importance, since it affects the initial values of the adjoint systems (Clever and Lang, 2012). Finally, the objective has to include a regularisation of the control itself. This term can be used to search an optimal control close to a preferable profile or to minimise the manufacturing costs. An objective functional that meets all the requirements stated above can be defined by

\[
J(T, u) := \frac{1}{2} \int_0^t \|T - T_d\|_{L^2(\Omega)}^2 \, dt + \frac{\delta_a}{2} \|(T - T_d)(t_e)\|_{L^2(\Omega)}^2 + \frac{\delta_u}{2} \int_0^t (u - u_d)^2 \, dt, \tag{23}
\]

with the desired glass temperature distribution \(T_d(x, t)\), the guideline for the control \(u_d(t)\) and the Tikhonov parameters \(\delta_a\) and \(\delta_u\). For more complex objective functionals within the context of glass cooling, including time dependent weights and a term forcing the minimisation of internal stresses explicitly, we refer to Clever and Lang (2012). Additionally, it is important to restrict the control \(u\) to a convex, feasible set \(U_{ad}\), which represents the operation interval of the furnace.

In the following, let the state system of the optimal control problem be denoted by

\[
ed(y, u) = 0. \tag{24}
\]

Depending on the model chosen, the state system is defined by one of the approximations (7)-(9), (10)-(14) or (16)-(20), respectively. Assuming \(J(y, u)\) and \(e(y, u)\) to be twice continuously Fréchet differentiable, which is obvious for the objective (23) and was shown for the grey scale model in Pinnau (2007), it is a common approach to reduce the optimisation problem to its control component, such that the reduced optimal control problem reads

\[
\min_{u \in U_{ad}} \tilde{J}(u) := J(y(u), u), \quad \text{where } y = y(u) \text{ satisfies } e(y, u) = 0. \tag{25}
\]
The feasible subset \( U_{ad} \) of the control space \( U \) is defined by

\[
U_{ad} := \{ u \in U : u_{\text{low}} \leq u(t) \leq u_{\text{up}}, \forall t \in [0, t_e] \},
\]  
(26)

with upper bound \( u_{\text{up}} \) and lower bound \( u_{\text{low}} \). Denoting the adjoint operator of the state operator by \( e^*(y, u) \), the reduced gradient is given by

\[
\nabla \hat{J}(u) = \nabla_u J(y, u) + \nabla_u e^*(y, u) \xi,
\]  
(27)

where \( \xi \) is the adjoint state, which can be computed by solving the adjoint system

\[
\partial_y J(y, u) + \partial_y e^*(y, u) \xi = 0.
\]  
(28)

A formal description of this process is shown in Figure 2.

Figure 2: Computation of reduced gradient by adjoint calculus

Instead of the whole reduced Hessian \( \hat{J}''(u) \), we only consider actings of it, that can efficiently be evaluated by computing

\[
\hat{J}''(u) s_u = \partial_{uu} J(y, u) s_u + \partial_{uu} e^*(y, u) \xi s_u + \partial_u e^*(y, u) w + \partial_{uy} J(y, u) s_y + \partial_{uy} e^*(y, u) \xi s_y,
\]  
(29)

with some direction \( s_u \), the linearised state \( s_y \), and the second adjoint state \( w \). The linearised state \( s_y \) is the solution of the linearised state system

\[
\partial_y e(y, u) s_y = -\partial_u e(y, u) s_u,
\]  
(30)

with input \( s_u \), and the second adjoint state is the solution of the second adjoint system

\[
\partial_y e^*(y, u) w = -\partial_{yy} J(y, u) s_y - \partial_{yy} e^*(y, u) \xi s_y - \partial_{yu} J(y, u) s_u - \partial_{yu} e^*(y, u) \xi s_u,
\]  
(31)

which depends on state \( y \), linearised state \( s_y \) and adjoint state \( \xi \). The complete data flow is illustrated in Figure 3. Note that it is inefficient to compute all components of the reduced Hessian, since this would require at least \( 2m \) additional PDAE solves, where \( m \) is the number of discretisation points of the control space. Contrarily, the adjoint based approach presented above requires only two additional PDAE solves, independent of the discrete control space.

### 3.2 Multilevel Generalised SQP Method

In the following, we consider a multilevel generalised SQP method that explicitly allows the use of independent discretisation schemes and independent spatial meshes for state and adjoint equations, see Clever et al. (2010, 2012); Ziems (2010); Ziems and Ulbrich (2011). Enriched with a fully space time
The main idea of the SQP method is to approximate the reduced optimisation problem (25) by a sequence of reduced SQP-subproblems:

\[
\min_{s_u \in U_{ad} - \mathbf{u}_k} \hat{q}_k(s_u) := J(y_k, \mathbf{u}_k) + \left\langle \hat{J}(\mathbf{u}_k), s_u \right\rangle + \frac{1}{2} \left\langle s_u, \hat{J}''(\mathbf{u}_k) s_u \right\rangle \quad \text{s.t. } \|s_u\| \leq \Delta_k,
\]

with reduced gradient \( \hat{J}'(\mathbf{u}_k) \), reduced Hessian \( \hat{J}''(\mathbf{u}_k) \) and trust region radius \( \Delta_k \). When solving (32) we have to account for four characteristics. The first two points are, that the search direction \( s_u \) is restricted to the set \( U_{ad} - \mathbf{u}_k \) and its length to the trust region radius \( \Delta_k \). The third point is, that we only need actions of the reduced Hessian \( \hat{J}''(\mathbf{u}_k) s_u \) and the last, that the reduced Hessian is not necessarily symmetric due to the different discretisation schemes in state and adjoint systems and the independent spatial meshes. To handle the control constraints, we define the \( \varepsilon \)-active set

\[
\mathcal{A}^\varepsilon(u_k^h) = \{ i | (u_{up}^h)_i - (u_k^h)_i \leq \varepsilon \text{ or } ((u_k^h)_i - (u_{low}^h)_i) \leq \varepsilon \},
\]

where \((u^h)_i\) denotes the \( i \)-th component of the discretised control \( u^h \). The \( \varepsilon \)-inactive index set \( \mathcal{I}^\varepsilon(u_k^h) \) is defined by the complement \( \mathcal{A}^\varepsilon(u_k^h)^c \). Here and in the following we set \( \varepsilon = \min\{ C_M, (u_{up} - u_{low})/2 \} \), with criticality measure \( C_M \), upper bound \( u_{up} \) and lower bound \( u_{low} \), to avoid an overlapping of active regions. Because we only have actions of the slightly unsymmetrical Hessian, we consider a restricted version of BiCGstab, which solves the linear equation

\[
\hat{J}''(u_k) s_u = -\nabla \hat{J}(u_k)
\]

only on the \( \varepsilon \)-inactive part, delivering the solution \( s_u^{\mathcal{A}^\varepsilon(u_k)} \). In the \( \varepsilon \)-active region, a projected gradient step \( s_u^{\mathcal{A}^\varepsilon(u_k)} := -\nabla \hat{J}(u_k)|_{\mathcal{A}^\varepsilon(u_k)} \) is considered. Using Armijo’s line search, a proper scaling \( \sigma_k \) of the trial step

\[
s_{u,\text{proj}} := P_{U_{ad} - \mathbf{u}_k}(\sigma_k s_u), \quad \text{with } s_u := s_u^{\mathcal{A}^\varepsilon(u_k)} + s_u^{\mathcal{I}^\varepsilon(u_k)},
\]

is determined. The projection \( P_{U_{ad} - \mathbf{u}_k}(d) \) of a quantity \( d \) is defined, such that \( u_k + P_{U_{ad} - \mathbf{u}_k}(d) \) is within the feasible set \( U_{ad} \). A new control \( u_{k+1} := u_k + \sigma s_{u,\text{proj}} \) is accepted, if the actual reduction

\[
\text{ared}_k = J(y_k, u_k) - J(y_{k+1}, u_{k+1})
\]

is at least a fraction of the model based predicted reduction

\[
\text{pred}_k = \hat{q}_k(0) - \hat{q}_k(s_{u,\text{proj}}).
\]

If a step is accepted, the ratio of actual and predicted reduction is used in a standard fashion to adjust the trust region radius \( \Delta_k \). If a step is rejected, it has to be verified, if either the SQP-subproblem (32) does not approximate the optimisation problem (25) well enough or if full and reduced model do not coincide well enough on the current discretisation level. In the first case, the trust region gets reduced and in the second, the accuracy level gets increased.

Independently of acceptance or rejection of a new control, for a sufficient quality of space and time grids, we require

\[
\eta_y \leq c_1 C_M,
\]

\[
\eta_\xi \leq c_2 C_M,
\]

with criticality measure \( C_M \) and global discretisation error estimators \( \eta_y \) and \( \eta_\xi \), see Subsection 3.3. If the criticality measure descents below a predefined tolerance, the algorithm is stopped. For more details on the algorithm we refer to Clever et al. (2010, 2012); Ziems (2010); Ziems and Ulbrich (2011).

In later computations we define the criticality measure \( C_M \) by the projection of the reduced gradient

\[
C_M = \|P_{U_{ad} - \mathbf{u}_k}(\hat{J}'(u_k))\|.
\]

Note that we use the same criticality measure \( C_M \) to control the model hierarchy in (21). A flow chart that illustrates the performance of the considered multilevel hierarchy based SQP-method is presented in Figure 4.
Compute quotient of predicted reduction (37) and actual reduction (36) to decide if $u_{k+1}$ is accepted.

Set $k := k+1$.

Solve state system (24) and adjoint system (28) to evaluate objective functional (23) and reduced gradient (27).

Provide global error estimators $\eta_y$ and $\eta_\xi$.

Check whether the step is rejected because the trust region is too large or because the grid is too coarse.

Determine proper scaling $\sigma_k$ using Armijo line search.

Set $u_{k+1} := u_k + s_k \cdot \text{proj}$.

If model hierarchy is enabled, determine model error (22) and check if switching criteria (21) holds.

Solve $J(y_k, u_k) = -\nabla J(u_k)$ on $e$-inactive set $T^e(\alpha)$ using restricted BiCGstab.

Evaluate $s_k := s_k^T(\alpha) + s_k^A(\alpha)$.

Increase accuracy yes

Stop, if criticality measure (40) is small enough, else.

Increase accuracy no

Check if accuracy level is sufficient yes

Weight to higher model no

Increase accuracy yes

Determine proper scaling $\sigma_k$ using Armijo line search.

Reduce trust region yes

Figure 4: Schematic representation of multilevel hierarchy based SQP-method

In the flow chart decisions are visualised by octagons, other tasks by rectangles. Rounded corners mark tasks, that include at least one PDAE solve. Note that the BiCGstab box includes the evaluation of actions of the reduced Hessian (29) in every iteration.

3.3 Realisation

To solve the involved systems of PDAEs, to evaluate functionals like objective and reduced derivatives and to estimate discretisation errors, we couple the optimisation algorithm with the fully space-time adaptive software package Kardos. It is based on Rothe’s method with adaptive linearly implicit one-step method of Rosenbrock type for the time integration and adaptive multilevel finite elements for the spatial discretisation. The linearly implicit structure of the Rosenbrock methods is advantageous to handle non-linearities, like the Planck function in the glass cooling problem. The one-step character allows rapid change of step sizes, which are adjusted with respect to local error estimates determined by an embedded scheme of inferior order. To control the adaptive grid refinement in space, the spatial discretisation error is estimated locally by the hierarchical basis concept (Lang, 2001).

Kardos offers the possibility to make PDAE solutions from previous runs available to the current PDAE solve. This mechanism can be used to solve state and adjoint systems sequentially, to evaluate the reduced gradient as illustrated in Figure 2 and to evaluate the reduced Hessian, see Figure 3. Figure 5 gives a closer look at the underlying grid and data management in Kardos. Because state and linearised state equation have initial conditions, those two systems are solved forward in time, whereas adjoint and second adjoint system with their terminal conditions are solved backwards in time. Each of the four systems can be solved with a different integration method that matches the structure of the underlying PDAE. On each accuracy level, an adaptive time discretisation $t_0, t_1, \ldots, t_n$ is determined during the state system solve. This discretisation is reused for all other PDAE solves as long as the resulting global error estimates $\eta_y$ and $\eta_\xi$ fulfil (38) and (39). In each point of time $t_k$, the Rosenbrock discretisation results in a sequence of elliptic PDEs, which are solved on their own suitable grid. Hence, when setting up adjoint and second adjoint system the environment has to manage the transfer from up to three PDAE solutions, which have been computed on their own independent spatial mesh. The computations are done.
partly forwards and partly backwards in time. Note that the same mechanism can be used to evaluate the

\[
y, s, y, s, y, \xi, y, s, \xi
\]

Linearized state system solve, forward in time

Second adjoint system solve, backwards in time

Figure 5: Grid and data management

On each accuracy level, an adaptive time discretisation \(t_0, t_1, \ldots, t_n\) is determined during the state system solve, which is then used for all other PDAE solves on this level. In each point of time \(t_k\), the Rosenbrock discretisation results in a sequence of elliptic PDEs, which are solved on their own suitable spatial grid by finite elements. Solving the adjoint system backwards in time on the predefined time grid, the spatial meshes are individually adapted to reach the desired accuracy. To provide a current PDAE solve with solutions from previous runs, the solution is imported on its computational grid and transferred to the current grid.

model error (22). In this case a lower model solution is read in during a higher model solve. Having both solutions on the same grid, the \(L^2\)-norm of their difference can be efficiently evaluated by a weighted sum. On each accuracy level the procedure is the same as presented in the flow chart 5, but with increasing number of time nodes. This results in an increasing number of elliptic differential equations, each of them solved on an individually refined grid ensuring the new accuracy. To provide global error estimates \(\eta_y\) and \(\eta_\xi\) for the multilevel strategy in (38)-(39), we make use of the already computed local error estimates and rely on tolerance proportionality. With respect to the semi-discretisation, we distinguish between the local error in time \(le_t = (0, le^t_1, \ldots, le^t_n)\) and the local error in space \(le_x = (0, le^x_1, \ldots, le^x_n)\) and define the global estimate

\[
\eta = (t_e - t_0)^{-1/2} \left( \sum_{k=0}^{n-1} \frac{\tau_k}{2} \left( (le^t_k)^2 + (le^t_{k+1})^2 \right) \right)^{1/2} + \left( \sum_{k=0}^{n-1} \frac{\tau_k}{2} \left( (le^x_k)^2 + (le^x_{k+1})^2 \right) \right)^{1/2},
\]

(41)

with adaptive time step sizes \(\tau_k, k = 0, \ldots, n - 1\).

Considering the local error estimate (41) within the state system (24) evaluates \(\eta_y\). Using the local error estimate (41) within the adjoint system (28) evaluates \(\eta_\xi\). The unknown proportionality factors can formally be handled by the constants \(c_1\) and \(c_2\).

We want to point out, that the KARDOS based multilevel SQP algorithm is not restricted to the solution of the glass cooling problem. It is a suitable optimisation tool for a great variety of boundary control problems restricted by space-time dependent PDAEs of similar type. For more details on the class of PDAEs that can be handled we refer to Erdmann et al. (2002).
4 Numerical Experiments

The following Section is divided into two subsections. In Subsection 4.1 we solve the optimal control problem for the eight band model using the model hierarchy based approach, explained in Subsection 2.2. We compare its performance with respect to quality and effort to a similar optimisation run, carried out on the eight band model only. Due to the high complexity of the considered model, and the high computing time in three spatial dimensions, in this subsection we approximate the three dimensional computational domain by a two dimensional cross section. In Subsection 4.2, we then consider the entire domain and solve the optimal control problem in three space dimensions, considering the less complex grey scale model. The computational domain $\Omega^{3d}$ is given by the convex hull of the eight points

$$p_1 = (-1, -1, -1), \quad p_2 = (1, -1, -1), \quad p_3 = (1, 1, -1), \quad p_4 = (-1, 1, -1),$$

$$p_5 = (0.5, 0.5, 1), \quad p_6 = (1, 0.5, 1), \quad p_7 = (1, 1, 1), \quad p_8 = (0.5, 1, 1).$$

To define a two dimensional cross section we set $z = \frac{1}{3}$, which results in $\Omega^{2d} = [0, 1] \times [0, 1]$. Both geometries, together with their initial grids, are shown in Figure 6.

![Figure 6: Three dimensional computational domain and two dimensional cross section $z = \frac{1}{3}$, with initial grids](image)

Further parameters concerning the different models and the optimisation algorithm are given in Table 3. In the following numerical experiments we use linear finite elements and the third order Rosenbrock method ROS3PL (Lang and Verwer, 2001; Lang and Teleaga, 2008), which is an L-stable time integration scheme and does not suffer from order reduction.

### 4.1 Model Hierarchy Approach

In this section we present the model hierarchy based SQP method presented in Section 3. For the model hierarchy, we consider all three models (7)-(9), (10)-(14) and (16)-(20), which means that we finish the algorithm with an optimal control for the eight band model. As an approximative weighted mean of the absorption rates of the different bands, given in Table 2, the mean absorption coefficient $\kappa$ in the
The glass temperature for different times and models is shown in Figure 7. The glass temperature at final time $t=0.1$ is shown in Figure 8. For a better visualisation of the temperature differences, the colour is scaled between 300 and 450, which is the lower quarter of the feasible set $U_{\text{ad}}$. The computations were carried out on an AMD Athlon(tm) 64 X2 Dual Core Processor 6000+, with a CPU cache size of 512 KB and a clock rate of 1000 MHz. On the predefined grid the state system solve all PDAE solves on a predefined space-time grid, which is refined in the boundary region, with 76 time steps and 1053 spatial nodes. The mesh is chosen in such a way, that there is no further refinement necessary during the entire optimisation. Note that this simplification is only made to be able to study the benefit of the model hierarchy. Full space-time adaptivity and a grid refining multilevel strategy as explained in Section 3 are considered in Subsection 4.2 for three space dimensions. For an application of the grid refining multilevel strategy to the glass cooling problem in two space dimensions, we refer to Clever et al. (2010, 2012).

The constants $k_1$ and $k_2$ that steer the shifting from one model to another, see (21), are set to $k_1 = k_2 = 1.0$. To get an impression about the similarity of the three models, Figure 7 shows the glass temperature on the cut section through $y = 0.5$ at different points of time for all three models, resulting from the initial control $u_0$. The corresponding temperature distribution over the entire domain $[0, 1] \times [0, 1] \in \mathbb{R}^2$ for the final time is shown in Figure 8. For a better visualisation of the temperature
This optimal control together with the initial control is presented in Figure 9(a). The evolution of the glass temperature over time that results from the optimal control is shown in Figure 9(b) for the boundary point $Q_1 = (0, 0.5)$, the corner point $Q_2 = (1, 0)$, and the interior point $Q_3 = (0.5, 0.25)$. Note that the glass temperature within the computational domain is mainly located between the temperature of the boundary point $Q_1$ (green dashed line) and the interior point $Q_3$ (red dot-dashed line), see 9(b). Hence, the optimal glass temperature is evenly distributed around the desired profile $T_d(t)$. Contrarily, for the initial control, the resulting glass temperature distribution lies above $T_d(t)$ in any space-time point of the computational domain. The glass temperature distribution over the entire domain at final time, resulting from the optimal control, is shown in Figure 9(c). Comparing it to the distribution resulting from the initial control, see Figure 8(c), we can observe a significant improvement. More details about the optimisation performance are given in Table 4.

### Table 4: Optimisation protocol for the model hierarchy based approach

<table>
<thead>
<tr>
<th>opt. iter.</th>
<th>target value</th>
<th>criticality measure $E(\mathcal{M}<em>i, \mathcal{M}</em>{i+1})$</th>
<th>time to compute $E$ [s]</th>
<th>#BiCGstab iterations</th>
<th>accumulated CPU time [s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start with Rosseland approximation</td>
<td>2.4213889e+02</td>
<td>1.320e+01</td>
<td>2.804e–01</td>
<td>45</td>
<td>139</td>
</tr>
<tr>
<td>1</td>
<td>8.9556646e+01</td>
<td>6.420e–01</td>
<td>3.143e–01</td>
<td>45</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>8.9018226e+01</td>
<td>6.828e–02</td>
<td>3.143e–01</td>
<td>44</td>
<td>4</td>
</tr>
<tr>
<td>Shift to grey scale model</td>
<td>1.5892804e+02</td>
<td>6.010e+00</td>
<td>1.504e–01</td>
<td>392</td>
<td>2 570</td>
</tr>
<tr>
<td>3</td>
<td>1.3618194e+02</td>
<td>3.657e–01</td>
<td>1.575e–01</td>
<td>391</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>1.3605275e+02</td>
<td>7.212e–03</td>
<td>1.579e–01</td>
<td>391</td>
<td>3</td>
</tr>
<tr>
<td>Shift to eight band model</td>
<td>1.9479777e+02</td>
<td>2.230e+00</td>
<td>-</td>
<td>-</td>
<td>5 575</td>
</tr>
<tr>
<td>1</td>
<td>1.9113854e+02</td>
<td>1.126e–01</td>
<td>-</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>1.9112199e+02</td>
<td>2.54e–03</td>
<td>-</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>1.9112197e+02</td>
<td>2.367e–04</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
</tbody>
</table>

Comparing the model hierarchy based approach to a similar optimisation that is already started with the eight band model, see Table 5, it can be observed, that now the algorithm determines an optimal control of comparable accuracy in $k = 5$ instead of $k = 7$ optimisation iterations. However, the model hierarchy based approach needs only 3 optimisation iterations on the time consuming eight band model, see Table 4. It is necessary to additionally compute the model error for the two lower models. However the effort for the optimisation iterations on these models is almost negligible in comparison to the complexity of the eight band model. Having a look at the computing time (Table 4 and Table 5), it can be seen that the model hierarchy based approach saves about 20% of the computational time required by the one-model approach.

Studying the results of the two algorithms, we clearly observe a significant different development of
Table 5: Optimisation protocol for the eight band model

<table>
<thead>
<tr>
<th>optimisation iteration</th>
<th>target value</th>
<th>criticality measure</th>
<th>#BiCGstab iterations</th>
<th>accumulated CPU time [s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.5934272e+02</td>
<td>1.835e+01</td>
<td>1</td>
<td>1268</td>
</tr>
<tr>
<td>2</td>
<td>1.9478452e+02</td>
<td>1.763e+00</td>
<td>4</td>
<td>10387</td>
</tr>
<tr>
<td>3</td>
<td>1.9118843e+02</td>
<td>2.163e–01</td>
<td>3</td>
<td>17716</td>
</tr>
<tr>
<td>4</td>
<td>1.9112234e+02</td>
<td>7.690e–03</td>
<td>3</td>
<td>25043</td>
</tr>
<tr>
<td>5</td>
<td>1.9112200e+02</td>
<td>7.283e–04</td>
<td>1</td>
<td>28831</td>
</tr>
<tr>
<td>6</td>
<td>1.9112198e+02</td>
<td>1.576e–04</td>
<td>1</td>
<td>32625</td>
</tr>
</tbody>
</table>

The control iterates in each optimisation step, see Figure 10. But, as expected, the optimal controls that result from the two approaches are nearly identical. Their relative $L^2$-difference is $1.578e–06$. Due to the great conformance of the optimal control from both approaches, we abstain from showing further results for the one-model approach, since they coincide with those presented for the model hierarchy based approach in Figure 9.

Summarising, it can be said that for a reasonable hierarchy of different models with ascending accuracy, the presented model hierarchy based approach is a suitable tool to decrease the computational effort while maintaining the desired accuracy. As long as the optimisation is entirely carried out on the same level of accuracy we observe savings in the computing time of around 20%. However, when the grid refining multilevel strategy, presented in Section 3, is enabled as well, the last optimisation iteration might require more than 90% of the entire computing time. To serve the desired accuracy, this expensive step has to be carried out in the model hierarchy based approach as well. Hence in such a case, a benefit of 20% from the remaining 10% does only result in a benefit of 2% of the computing time.

4.2 Numerical Experiments in 3D

In the previous 2d example, we did not apply grid adaption and the grid refining multilevel strategy for a better visualisation of the benefit achieved by the model hierarchy based approach. However, to solve the optimal boundary problem for glass cooling on a complex 3d geometry, the space-time adaptivity together with the grid refining multilevel strategy are essential. Since the combination of model hierarchy and multilevel strategy does not promise so much additional benefit, in this subsection we concentrate on the one-model approach, using the grey scale model (10)-(14).

The involved PDAEs are solved by *Kardos3d* (Erdmann et al., 2002). Different radiation models on a cube were already solved successfully with this software (Klar et al., 2005), although no optimisation was considered.
As in the 2d case, we consider the problem and model parameters given in Table 3 and set the mean absorption coefficient $\kappa = 10$. The computations were done on an AMD Opteron DualCore 8218 with 3.0 GHz and 128 GB RAM with 533 MHz FSB.

Since the arising PDAEs are solved adaptively in time and space, error tolerances have to be set. At the beginning of the optimisation process the local time error of state and adjoint state are forced below the tolerance $1.0e-02$, the local space error of the state below $7.5e-02$, and that of the adjoint state below $5.0e-1$. During the optimisation these tolerances are reduced as described in Figure 4. Relying on tolerance proportionality a suitable refinement is found within just one single refinement iteration.

In Table 6 the optimisation protocol of the glass cooling process modelled by the grey scale approximation can be seen. In column 2 and 3 we present the evolution of the objective value and the criticality measure. Note, that after a refinement step, the objective value and the criticality measure are not necessarily monotonically decreasing. Jumps within the objective value are caused by the different results of the numerical integration of (23) on the different meshes. Jumps within the criticality measure are additionally caused by the fact, that a good control iterate on a coarse grid is usually not as good on a finer grid.

In column 4-6 (Table 6) the number of grid points on the different levels is given. We want to point out, that the high number of more than 10 000 spatial nodes on the first level only occurs at one point of time. The entire number of nodes (which is the sum of the number of spatial nodes over all time steps) increases from one level to the next.

After four optimisation iterations, discretisation errors and criticality measure are small enough and the computed optimal control can be trusted to approximate well the infinite dimensional optimal control. We have adaptive grids with a minimum number of 4570 and a maximum number of 69474 spatial nodes at 42 adaptive points in time. This leads to a global state error in time of $1.43e-03$ and in space of $2.84e-03$ and a global adjoint error in time of $1.31e-03$ and in space of $2.93e-02$. One can see that the last optimisation step, that is computed on the finest level, consumes about 95% of the total computing time.

![Figure 11: Control iterates and glass temperature profile at different points for the optimal control](image)

In 11(a) the control iterates are displayed. A fast convergence to the optimal control can be observed. In 11(b) the temperature profile at four different points for the optimal control is compared with the desired temperature profile.

<table>
<thead>
<tr>
<th>opt. iter.</th>
<th>target value</th>
<th>criticality measure</th>
<th>time nodes</th>
<th>max. space nodes</th>
<th>entire no. nodes</th>
<th>accumulated CPU time [s]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>State</td>
<td>Adjoint</td>
<td>State</td>
<td>Adjoint</td>
</tr>
<tr>
<td>0</td>
<td>1.566e+03</td>
<td>4.048e+01</td>
<td>16</td>
<td>11 037</td>
<td>4 531</td>
<td>73 756 64 118</td>
</tr>
<tr>
<td>1</td>
<td>5.761e+02</td>
<td>3.237e+00</td>
<td>16</td>
<td>11 037</td>
<td>9 767</td>
<td>79 167 69 529</td>
</tr>
<tr>
<td>2</td>
<td>5.690e+02</td>
<td>5.022e-01</td>
<td>16</td>
<td>11 037</td>
<td>10 180</td>
<td>84 941 75 303</td>
</tr>
<tr>
<td>ref</td>
<td>5.544e+02</td>
<td>7.280e-01</td>
<td>26</td>
<td>4 520</td>
<td>4 520</td>
<td>100 630 103 132</td>
</tr>
<tr>
<td>3</td>
<td>5.539e+02</td>
<td>5.501e-02</td>
<td>26</td>
<td>4 520</td>
<td>4 520</td>
<td>103 524 106 026</td>
</tr>
<tr>
<td>ref</td>
<td>5.743e+02</td>
<td>1.249e+00</td>
<td>42</td>
<td>67 068</td>
<td>67 068</td>
<td>988 469 973 240</td>
</tr>
<tr>
<td>4</td>
<td>5.741e+02</td>
<td>5.533e-02</td>
<td>42</td>
<td>69 474</td>
<td>69 474</td>
<td>1 059 143 1 043 914</td>
</tr>
</tbody>
</table>

![Table 6: Optimisation protocol for grey scale model in three spatial dimensions](image)
The control iterates of the optimisation are shown in Figure 11(a). The convergence to the optimal control is quite fast. Most work has to be done around the transition from inactive to active control constraints at around $t = 0.064$, where the number of spatial nodes reaches the maximum of about 70,000 nodes per time step.

In Figure 11(b) the temperature during the cooling process in four different points of the geometry is compared to the desired temperature $T_d$. These points are $Q_1 = (0, 0, 0)$ in the interior, $Q_2 = (1, 1, -1)$ in a corner, $Q_3 = (0, -1, -1)$ on a boundary edge and $Q_4 = (0.5, 0.25, 0.33)$ on the cross section $\Omega_{2d}$. The points on the edge and in the corner ($Q_2$ and $Q_3$) cool down uniformly and quicker than the points in the interior ($Q_1$ and $Q_4$). Comparing Figure 9(a) and 9(b) to Figure 11(a) and 11(b) one can observe a great analogy between the results computed on the two dimensional cross section and the 3d case.

![Figure 12: Terminal glass temperature distributions resulting from initial and optimal control](image)

The temperature distribution at the end of the cooling process at different slices through the geometry resulting from the initial control is presented in the first row, the one resulting from the optimal control in the second row. On the left $z = -1$, which is essentially the bottom of the geometry (12(a) and 12(d)), in the middle $z = 1/3$ for the cross section $\Omega_{2d}$ (12(b) and 12(e)) and finally on the right the slice is parallel to the $y$-axis with $y = 0.5$ (12(c) and 12(f)).

In Figure 12 the terminal temperature distribution in the geometry resulting from the computed optimal control is compared to the one resulting from the initial control. While there are still great temperature differences between the corner and the interior for the initial control, the temperature distribution after an optimal cooling process is quite uniform and significantly lower. Having a look at the cross section (Figure 12(b) and Figure 12(e)), we again observe great similarity to the 2d case (Figure 8 and Figure 9(c)).

Summarising, it can be said, that the presented optimisation algorithm is capable of solving complex three dimensional problems. Further, we have presented a fully adaptive optimisation scenario for the glass cooling problem on a three dimensional computational domain. It can be seen, that here the grid refining multilevel strategy is of great advantage and makes the difference whether one can solve the optimisation problem in an acceptable time or not. From Clever et al. (2010, 2012) we deduce that an optimisation using only the highest accuracy level is expected to take five times longer than the fully adaptive multilevel optimisation strategy, which for the presented case would mean a computing time of more than a month.
5 Conclusions

We have presented an adaptive multilevel and model hierarchy strategy to solve optimal control problems for glass cooling processes in two and three-dimensional geometries. Both strategies are useful to drastically reduce the computing time necessary to reach practically relevant accuracies. The model hierarchy based optimisation approach is especially attractive if no grid refinement options are available to solve the PDAE constraints. In this case, most of the work can be done on the basis of lower order and less expensive models. However, if adaptive discretisation schemes are used and therefore the last optimisation iteration is in general carried out on significantly fine meshes, then this advantage becomes less important.

Adaptive multilevel optimisation strategies based on successive improvement of the approximation property of the space-time discretisations perform remarkably robust and have the potential to solve even complex three-dimensional glass cooling problems within moderate time. Their great advantage is, that only a few optimisation steps have to be evaluated on the highest discretisation level. Needless to say, the two approaches presented can be applied to general PDAE-constrained optimal control problems. An implementation of a second boundary control problem within the presented environment, namely a thermistor problem in steel hardening, is work in progress. Furthermore, the environment is meant to be augmented to optimal control problems with distributed control such as the redistribution of dopant into silicon.
References


Acknowledgement

The authors gratefully acknowledge the support of the German Research Foundation (DFG) within the Priority Program 1253 "Optimization with Partial Differential Equations" under grants LA1372/6-1,2 and UL158/7-1,2, the 'Excellence Initiative' of the German Federal and State Governments and the Graduate School of Computational Engineering at Technische Universität Darmstadt. Furthermore, we thank the referees for the constructive comments.