Optimal Boundary Control of Glass Cooling Processes *

René Pinnau[†] Guido Thömmes[‡]

Abstract

An optimal control problem for glass cooling processes is considered. We model glass cooling using the SP_1 approximations to the radiative heat transfer equations. The control variable is the temperature at the boundary of the domain. This results in a boundary control problem for a parabolic/elliptic system which is treated by a constrained optimization approach. We consider several cost functionals of tracking type, define the corresponding Lagrange functionals and derive the first-order optimality system. We investigate several numerical methods based on the adjoint variables and present results of numerical simulations illustrating the feasibility and performance of the different approaches.

Key words. Radiative heat transfer, glass manufacturing, thermal stresses, SP_N approximation, optimal control, first–order optimality system, Lagrange multipliers, descent algorithm, numerics.

AMS(MOS) subject classification. 35K55, 49K20, 80A20.

1 Introduction

In glass manufacturing, a hot melt of glass is cooled down to room temperature. During cooling, large temperature differences i.e. large gradients have to be avoided since they lead to thermal stress in the material. This may cause cracks or, in the case of high quality glass, affect the quality of the resulting product or

*This work was supported by DFG grant KL 1105/7 and by the Fraunhofer ITWM, Kaiserslautern.

†Department of Mathematics, Technical University of Darmstadt, 64289 Darmstadt, Germany.

‡Department of Mathematics, Technical University of Darmstadt, 64289 Darmstadt, Germany.

device. Hence, the process has to be managed in such a way that temperature gradients are sufficiently small [1]. The main goal of the optimal control problem that we investigate in this paper is to minimize these gradients. Besides, further criteria may be of interest in practice. In order to reduce energy consumption the surrounding temperature should be small and it should be reduced as fast as possible to room temperature [11]. It may be noted that, in particular, fast cooling and small gradients are conflicting goals where automated control can support or even improve procedures based on experience and heuristics. These subordinate criteria can be included in our model as well.

In order to model glass cooling we consider for notational simplicity a frequency independent, grey model without scattering. Stated on a bounded spatial domain $\Omega \subset \mathbb{R}^d$, $d = 1, 2$ or 3, the scaled conductive radiative heat transfer equations read [9]

$$\varepsilon^2 \frac{\partial T}{\partial t} = \varepsilon^2 \nabla \cdot k \nabla T - \int_{S^2} \sigma(B - I)\, d\omega \tag{1.1a}$$

$$\forall \omega \in S^2 : \ \varepsilon \omega \cdot \nabla I = \sigma(B - I). \tag{1.1b}$$

Ingoing radiation is prescribed by transparent boundary conditions

$$I(x, t, \omega) = I_b(x, t, \omega), \qquad n \cdot \omega < 0, \tag{1.1c}$$

and temperature is assumed to obey Robin-type boundary conditions

$$\frac{h}{\varepsilon k} T + n \cdot \nabla T = \frac{h}{\varepsilon k} T_b, \tag{1.1d}$$

At initial time $t = 0$, the temperature is $T(x, 0) = T_0(x)$. In these equations, $I(x, t, \omega)$ denotes the specific radiation intensity at point $x \in \Omega$ traveling in direction $\omega \in S^2$ at time $t \geq 0$. The outside radiation I_b is assumed to be known for the ingoing directions (i.e. $n \cdot \omega < 0$) on the boundary. We denote the outward normal on $\partial\Omega$ by n. Furthermore, $T(x, t)$ denotes the material temperature and T_b is the exterior temperature on the boundary. The equations contain the parameters opacity $\kappa(\nu)$, heat conductivity k and convective heat transfer coefficient h, which are assumed to be constant. Moreover, B denotes Planck's function

$$B(\nu, T) = \frac{2 h_P \nu^3}{c^2} \left(e^{\frac{h_P \nu}{k_B T}} - 1 \right)^{-1}$$

for black body radiation in glass, which involves Planck's constant h_P, Boltzmann's constant k_B and the speed of light in vacuum c.

This form of the model is much too complex for optimization purposes due to the dependence on the direction $\omega \in S^2$. A simple solve for the state system i.e. a coupled system of a heat equation and a radiative transfer equation would already need a large amount of computing time. Instead, we use the diffusion–type SP_N approximations [5, 9] to the radiative heat transfer equations. These approximations have recently been investigated by one of the authors. They perform well in the diffusive, optically thick regime [7] where ε is small. But also for $\varepsilon \approx 1$ they perform better than the widely–used, standard Rosseland approximation [6]. We therefore chose the simplest approximation of this class for simulating the temperature evolution during the annealing of a glass slab. The SP_1 approximation to the radiative heat transfer equations is the system

$$\frac{\partial T}{\partial t} = k\Delta T + \frac{1}{3\sigma}\Delta\rho, \tag{1.2a}$$

$$-\varepsilon^2\frac{1}{3\sigma}\Delta\rho + \sigma\rho = \sigma(4\pi a T^4), \tag{1.2b}$$

with Robin–type boundary conditions at the boundary

$$\frac{h}{\varepsilon k}T + n\cdot\nabla T = \frac{h}{\varepsilon k}u, \tag{1.2c}$$

$$\frac{3\sigma}{2\varepsilon}\rho + n\cdot\nabla\rho = \frac{3\sigma}{2\varepsilon}(4\pi a u^4), \tag{1.2d}$$

and an initial condition $T(0,x) = T_0(x)$ for the temperature. Here, ρ is the radiative flux and the prescribed temperature at the boundary is from now on denoted by u, in accordance with the control theory literature. In fact, the evolution of the temperature T can only be controlled via u. There are several physical parameters and constants, namely the heat conductivity k, the convective heat transfer coefficient h, the opacity σ and Stefan–Boltzmann's constant a. The above system is a heat equation for T with a source term depending on the radiation. Glass at temperature T emits frequency-dependent radiation that is proportional to black body radiation. Integrating wrto. frequency we obtain the total thermal radiation $B(T) = aT^4$ according to Stefan's law. It appears on the right side of the flux equation (1.2b) and in the boundary condition (1.2d) in terms of T and u, respectively.

We intend to minimize cost functionals of tracking type having the form

$$J = J(T,\rho,u) = \frac{1}{p}\int_0^1 \|\nabla T\|^p_{L^p(\Omega)}\ dt + \frac{1}{p}\int_0^1 \|T - T_d\|^p_{L^p(\Omega)}\ dt + \frac{\delta}{2}\int_0^1 \|u - u_d\|^2_{L^2(\partial\Omega)}\ dt,$$

where (T,ρ) solves (1.2) and the time is also scaled to the unit interval $(0,1)$. Here, $T_d = T_d(t,x)$ is a specified temperature profile and $u_d = u_d(t,x)$ is a given control of the ambient temperature which shall be optimized. The control variable u appears in the cost functional as a term penalizing large deviations from u_d.

Typically, such a profile is given by engineers. It follows a certain path in time which is essential in order to achieve the desired material properties of the glass. Since these are not yet taken into account in our model, we enforce them via this tracking type penalization. Furthermore, the positive constant δ allows to adjust the weight of the penalty term.

We consider the optimal control problem as a constrained optimization problem [3, 2, 4] and derive the corresponding first–order optimality system via the Lagrange functional. For the computation of the optimal control u we present two algorithms relying on the adjoint variables. These techniques were successfully used in many control problems in fluid flow, see e.g. [2, 4] and the references therein. The advantage of this approach is threefold:

- Compared to finite differences for the computation of directional derivatives of the reduced cost functional the numerical effort is kept constant also for an increasing number of discrete design variables.
- It can be easily adopted to different cost functionals.
- Numerical methods can be formulated on the continuous level and the subsequent discretization remains free.

Nevertheless, the optimal boundary control problem in this paper poses new challenging difficulties, mathematically and numerically. First, there is a fourth–order algebraic nonlinearity in T. Secondly, the SP_1 system is indefinite and, thirdly, the control enters the equations in a genuinely nonlinear way as u and u^4.

Hence, the first method we propose is a stable variable step–length descent algorithm, where the gradient is computed via the adjoint variables, while the second directly solves the first–order optimality system via a nonlinear iteration scheme. In particular, for the descent algorithm we intensively investigate the choice of an appropriate step–length. This step is crucial for the performance of the algorithm and nonstandard owing to the nonlinearity in the control.

The paper is organized as follows. In Section 2, we define the cost functionals and in Section 3 the first–order optimality system is derived. The gradient-descent algorithm is discussed in Section 4 and a nonlinear iteration scheme is outlined in Section 5. Finally, various numerical experiments showing the performance of our approach are presented in Section 6 and concluding remarks can be found in Section 7.

2 Definition of the Functionals

We want to control the temperature profile such that the local temperature gradients are minimized. As we can only influence the outside temperature u we use this parameter as control variable. We consider various cost functionals of tracking–type which have the general form

$$J = J(T, u) = \int_0^1 \int_\Omega F(T, \nabla T)\, dx\, dt + \frac{\delta}{2} \int_0^1 \|u - u_d\|^2_{L^2(\partial\Omega)}\, dt, \tag{2.1}$$

where $F : \mathbb{R}^2 \to \mathbb{R}_0^+$ is a nonnegative, differentiable function, e.g.

$$F_1(T, \nabla T) = |T - T_d|^p\,, \tag{2.2a}$$
$$F_2(T, \nabla T) = |\nabla T|^p\,, \tag{2.2b}$$

or

$$F_3(T, \nabla T) = F_1 + F_2. \tag{2.2c}$$

Here, T_d is a given desired temperature profile and u is the control parameter, i.e. the outside temperature. Further, u_d describes a given initial outer temperature profile at the boundary, which represents the status quo, since the outside temperature has to follow a certain path to ensure certain quality constraints of the glass. The constant δ is a positive parameter which allows to adjust the weight of the cost and the observation. We choose $p \in (1, \infty)$ according to engineering specifications. Clearly, in order to incorporate more sophisticated stress models different functions F could be constructed, which might depend in a highly nonlinear manner on T and ∇T.

Naturally, considering the function F_2 will minimize the gradients in the temperature profile T, but also F_1 can yield this effect as T_d can be chosen constant in space. The larger p the smaller will be the difference $T - T_d$, uniformly in space, too. Finally, F_3 combines these two effects. Note, that we want to follow the Lagrange formalism such that the choice $p = \infty$ is not possible. Moreover, it is not clear from an analytical point of view wether boundedness of ∇T in $(0, 1) \times \Omega$ can be expect for complex domains Ω.

We consider the optimal boundary control problem as a constrained optimization problem, where the cost functional (2.1) is minimized with respect to the constraint given by system (1.2), i.e.

$$\begin{aligned} &\text{minimize } J(T, u) \text{ wrto. } u, \\ &\text{subject to system (1.2).} \end{aligned} \tag{2.3}$$

There exist various ways for the computation of a minimizer of (2.3). In this paper, we address the problem via the adjoint approach. To this purpose we derive in the next section the first–order optimality system in a systematic manner.

3 The First–order Optimality System

To derive the first–order optimality system we embed the minimization problem (2.3) in a precise analytical setting.

Given a Hilbert space H, let $L^2(H) = L^2(0,1,H)$ consist of all measurable functions $v : (0,1) \to H$ such that $\int_0^1 \|v(t)\|^2_{L^2(\Omega)} \; dt$ is bounded. For notational convenience we define

$$
\begin{aligned}
Q &\stackrel{\text{def}}{=} (0,1) \times \Omega, \\
\Sigma &\stackrel{\text{def}}{=} (0,1) \times \partial\Omega \\
X &\stackrel{\text{def}}{=} [L^2(0,1;H^1(\Omega))]^2 \\
U &\stackrel{\text{def}}{=} L^2(0,1;L^2(\partial\Omega)),
\end{aligned}
$$

where X is the space of states $x \stackrel{\text{def}}{=} (T,\rho)$ and U the space of controls and set

$$
\alpha = \frac{h}{\varepsilon\, k}, \quad \gamma = \frac{3\sigma}{2\varepsilon}.
$$

Then, the weak formulation of (1.2) reads: Find $(T,\rho) \in [L^2(H^1)]^2$ with $T_t \in L^2(H^{-1})$ and $T(0,x) = T_0(x)$ in $L^2(\Omega)$ such that

$$
\begin{aligned}
\langle e_1(T,\rho,u), \xi_T\rangle \stackrel{\text{def}}{=} & \int_0^1 \langle T_t, \xi_T\rangle \; dt + \int_Q k\, \nabla T\, \nabla \xi_T \; dx\, dt + \int_Q \frac{1}{3\sigma} \nabla\rho\, \nabla\xi_T \; dx\, dt \\
& + \int_\Sigma k\, \alpha(T-u)\xi_T \; ds\, dt + \int_\Sigma \frac{1}{3\sigma}\gamma(\rho - 4\pi a u^4)\xi_T \; ds\, dt = 0
\end{aligned}
$$

and

$$
\begin{aligned}
\langle e_2(T,\rho,u), \xi_\rho\rangle \stackrel{\text{def}}{=} & \int_Q \frac{\varepsilon^2}{3\sigma} \nabla\rho\, \nabla\xi_\rho \; dx\, dt + \int_Q (\sigma\rho - 4\pi\sigma a T^4)\xi_\rho \; dx\, dt \\
& + \int_\Sigma \frac{\varepsilon^2}{3\sigma}\gamma(\rho - 4\pi a u^4)\xi_\rho \; ds\, dt = 0
\end{aligned}
$$

for all $(\xi_T, \xi_\rho) \in [L^2(H^1)]^2$.

Defining the operator $e \stackrel{\text{def}}{=} (e_1, e_2)$ we can write this shortly as:

Find $(T,\rho) \in [L^2(H^1)]^2$ with $T_t \in L^2(H^{-1})$ and $T(0,x) = T_0(x)$ in $L^2(\Omega)$ such that

$$
\langle e(T,\rho), (\xi_T, \xi_\rho)\rangle = 0
$$

for all $(\xi_T, \xi_\rho) \in [L^2(H^1)]^2$, where $\langle \cdot, \cdot \rangle$ denotes the canonical dual pairing.

Since we only consider controls $u \in L^2(\Sigma)$ we will restrict u to the set of admissible controls

$$U_{ad} \stackrel{\text{def}}{=} \{u \in U : \quad \underline{u} \le u \le \overline{u}\},$$

where $0 < \underline{u} < \overline{u}$. On the one hand, this is physically reasonable since the outside temperature is controlled via a furnace which can be operated only in a limited temperature range. On the other hand, this restriction ensures the existence of an optimal control $u \in U$ (see [10]).

Remark 3.1. Alternatively, we can seek the control in a Hilbert space with more regularity which would require to add some technical terms to the cost functional such that compactness arguments hold for the minimizing sequence [10].

Now the precise statement of the constrained optimization problem reads

$$\begin{aligned} &\text{minimize} \quad J(x,u) \quad \text{over} \quad (x,u) \in X \times U_{ad} \\ &\text{subject to} \quad e(x,u) = 0. \end{aligned} \tag{3.1}$$

Let X^* be the dual space of X. The Lagrangian $\mathcal{L} : X \times U \times X^* \to \mathbb{R}$ associated to (3.1) is defined by

$$\mathcal{L}(x,u,\xi) \stackrel{\text{def}}{=} J(x,u) + \langle e(x,u), \xi \rangle ,$$

or, more explicitly,

$$\begin{aligned} L(x,u,\xi) = J(x,u) &+ \int_0^1 < \partial_t T, \xi_T >_{H^{-1},H^1} \; dt + \int_Q k \,\nabla T \nabla \xi_T \; dx\, dt \\ &+ \int_\Sigma k\alpha \,(T-u)\xi_T \; ds\, dt + \int_Q \frac{1}{3\sigma} \nabla\rho\nabla\xi_T \; dx\, dt \\ &+ \int_\Sigma \frac{1}{3\sigma}\gamma(\rho - 4\pi a u^4)\xi_T \; ds\, dt + \int_Q \frac{\varepsilon^2}{3\sigma}\nabla\rho\nabla\xi_\rho \; dx\, dt \\ &+ \int_Q \sigma(\rho - 4\pi a T^4)\xi_\rho \; dx\, dt + \int_\Sigma \frac{\varepsilon^2}{3\sigma}\gamma(\rho - 4\pi a u^4)\xi_\varphi \; ds\, dt. \end{aligned}$$

where $\xi \stackrel{\text{def}}{=} (\xi_T, \xi_T) \in X^*$ denotes the Lagrange–multiplier or the adjoint variable. Then the first–order optimality system corresponding to (2.3) formally reads

$$\nabla_{(x,u,\xi)} \mathcal{L}(x,u,\xi) = 0. \tag{3.2}$$

Remark 3.2. Corresponding to (3.1) the equation $\nabla_u L = 0$ has to be replaced by the inequality $\nabla_u L \,\xi \ge 0$ for all $\xi \in X^*$.

In the following, we want to rewrite this equation in a more concise form. Taking variations of $\mathcal{L}$ with respect to the adjoint variable $\xi \in X^*$ yields again the *state system*

$$\text{given } u \in U: \quad e(x,u) = 0, \text{ in } X.$$

Secondly, taking variations of $\mathcal{L}$ with respect to the state variable $x \in X$ we derive the *adjoint system*

$$\text{given } (x,u) \in U \times X : e_x^*(x,u)\xi = -J_x(x,u), \quad \text{in } X^*$$

or, equivalently,

$$\forall \tilde{x} \in X : \quad \langle e_x(x,u)\tilde{x}, \xi \rangle = - \langle J_x(x,u), \tilde{x} \rangle .$$

Here, $e_x^*(x,u)\xi$ denotes the adjoint of the linearization of e at (x,u) in the direction ξ. We have the specific equations

$$\begin{aligned}
&\int_0^1 < \partial_t \tilde{T}, \xi_T >_{H^{-1},H^1} \, dt + \int_Q k \, \nabla \tilde{T} \, \nabla \xi_T \, dx \, dt \\
&+ \int_\Sigma k \, \alpha \, \tilde{T} \, \xi_T \, ds dt - \int_Q 16 \pi \sigma a \, T^3 \, \tilde{T} \xi_\rho \, dx dt = - \int_Q DF(T, \nabla T) \cdot (\tilde{T}, \nabla \tilde{T}) \, dx dt,
\end{aligned}$$

and

$$\begin{aligned}
&\int_Q \frac{1}{3\sigma} \nabla \tilde{\rho} \, \nabla \xi_T \, dx \, dt + \int_Q \frac{\varepsilon^2}{3\sigma} \nabla \tilde{\rho} \, \nabla \xi_\rho \, dx \, dt + \int_Q \sigma \tilde{\rho} \xi_\rho \, dx \, dt \\
&\qquad + \int_\Sigma \frac{\varepsilon^2}{3\sigma} \gamma \, \tilde{\rho} \, \xi_\rho \, ds \, dt + \int_\Sigma \frac{1}{3\sigma} \gamma \, \tilde{\rho} \, \xi_T \, ds \, dt = 0.
\end{aligned}$$

They hold for all $(\tilde{T}, \tilde{\rho}) \in [L^2(H^1)]^2$ and, hence, they are nothing else but the weak formulation of

$$-\partial_t \xi_T = k \, \Delta \xi_T + 16 \pi a \sigma \, T^3 \, \xi_\rho - \partial_1 F(T, \nabla T) + \nabla \cdot \partial_2 F(T, \nabla T), \tag{3.3a}$$

$$-\frac{\varepsilon^2}{3\sigma} \, \Delta \xi_\rho + \sigma \, \rho = \frac{1}{3\sigma} \, \Delta \xi_T, \tag{3.3b}$$

with boundary conditions

$$n \cdot \nabla \xi_T + \alpha \, \xi_T = n \cdot \partial_2 F(T, \nabla T), \tag{3.3c}$$

$$n \cdot \nabla \xi_T + \gamma \, \xi_T + \varepsilon^2 (n \cdot \nabla \xi_\rho + \gamma \, \xi_\rho) = 0, \tag{3.3d}$$

and terminal condition

$$\xi_T(1) = 0. \tag{3.3e}$$

Remark 3.3. In the case of the function F_1, we have $\partial_2 F_1 = 0$, such that the above equations simplify significantly.

Finally, variations of $\mathcal{L}$ with respect to u in a direction $\tilde{u}$ yield

$$\langle e_u(x,u)\tilde{u}, \xi\rangle = \langle J_u(x,u), \tilde{u}\rangle\,,$$

or

$$\int_\Sigma -(k\alpha\,\tilde{u} - \frac{4}{3\sigma}u^3\,\tilde{u})\xi_T - \frac{4\varepsilon^2}{3\sigma}u^3\,\tilde{u}\,\xi_\rho\,ds dt = -\delta\int_\Sigma (u-u_d)\tilde{u}\,ds dt.$$

Since this is true for all $\tilde{u} \in L^2(0,1;L^2(\partial\Omega))$ we end up with a third order algebraic equation for u

$$-k\alpha\,\xi_T - \frac{4}{3\sigma}(\xi_T + \varepsilon^2\,\xi_\rho)u^3 = -\delta\,(u-u_d), \qquad \text{on } \Sigma. \tag{3.4}$$

4 The Gradient Algorithm

In this section, we present a robust descent algorithm, where the descent direction is computed via the adjoint variables. We formulate this method in terms of the continuous variables such that it is independent of a specific discretization which can be chosen subsequently.

Owing to the fact that the system (1.2) is uniquely solvable [10], we may reformulate the minimization problem (3.1) introducing the *reduced* cost functional

$$\begin{array}{ll} \text{minimize} \quad \hat{J}(u) \stackrel{\text{def}}{=} J(x(u),u) \quad \text{over} \quad u \in U_{ad} \\ \text{where} \quad x(u) \in X \quad \text{satisfies} \quad e(x(u),u) = 0. \end{array} \tag{4.1}$$

Then, a gradient algorithm for the computation of a minimizer of $\hat{J}$ is given by

Algorithm 1.

1. *Set $k = 0$ and choose initial control $u_0 \in U$.*
2. *Given u_k, compute the gradient $d_k \stackrel{\text{def}}{=} \nabla\hat{J}(u_k)$.*
3. *Given $\beta > 0$, set $u_{k+1} = u_k - \beta\,d_k$.*
4. *Set $k \to k+1$ and goto 2.*

For the computation of the gradient $\nabla \hat{J}(u)$ we employ the adjoint variables, since from variational calculus we have the identity

$$\nabla \hat{J}(u) = J_u(x(u), u) + J_x(x(u), u)x_u(u) = J_u(x(u), u) + e_u^* \xi.$$

Hence, for one evaluation of the gradient at $u \in U$ we have to solve first the nonlinear state system (1.2) for $x(u) \in X$ forward in time and secondly the linear adjoint system (3.3) for ξ backward in time. Eventually, we evaluate

$$\nabla \hat{J}(u) = -k\alpha\, \xi_T + \delta\,(u - u_d) - \frac{4}{3\sigma}(\xi_T + \varepsilon^2\, \xi_\rho)u^3 \tag{4.2}$$

Crucial for the convergence of Algorithm 1 is the choice of the step size β in the third step. Clearly, the best choice would be the result of a line search

$$\beta^* = \mathrm{argmin}_{\beta > 0}\, \hat{J}(u - \beta\, d)$$

which, unfortunately, is numerically much too expensive although it is a one-dimensional minimization problem. This is due to the fact that each evaluation of the cost functional $\hat{J}$ requires the solution of the nonlinear state system.

To avoid this computational drawback we propose two methods to derive an appropriate approximation for β^*. The first, purely heuristic approach is

$$\beta_1 \stackrel{\text{def}}{=} \min\left(1, \|d\|_\infty^{-1}\right), \tag{4.3}$$

and it actually means that we make small steps as long as the gradient is large, while we switch to the step–length 1 when the gradient is small, i.e. near the optimal solution. Note, that especially during the first steps we have to expect very large gradients due to the cubic power of u in (4.2).

The second approximation is based on the linearization of $x(u - \beta\, d)$ at u, which yields

$$x(u - \beta\, d) \approx x(u) - x_u(u)(\beta\, d).$$

Then, we can solve

$$\beta_2 = \mathrm{argmin}_{\beta > 0}\ J\Big(\ x(u) - x_u(u)(\beta\, d)\ ,\ u - \beta\, d\ \Big)$$

exactly. In fact, this results in an algebraic equation for β.

Remark 4.1. In case of the cost functional J_1 with $p = 2$ we get explictely

$$\beta_2 = \frac{\|d\|_\Sigma^2}{\varepsilon\, \|d\|_\Sigma^2 + \|v\|_Q^2}, \tag{4.4}$$

where $v = (v_T, v_\rho) \stackrel{\text{def}}{=} x_u(u)d$ solves the linear backward system

$$-\partial_t v_T = k\,\Delta v_T + 16\pi\sigma a\,T^3\,v_\rho - (T - T_d) \tag{4.5a}$$

$$-\frac{\varepsilon^2}{3\sigma}\,\Delta v_\rho + \sigma\,v_\rho = -\frac{1}{3\sigma}\,\Delta v_T \tag{4.5b}$$

with boundary conditions

$$n \cdot \nabla v_T + \alpha\,v_T = \alpha\,d, \tag{4.5c}$$

$$n \cdot \nabla v_T + \gamma\,v_T + \varepsilon^2(n \cdot \nabla v_\rho + \gamma\,v_\rho) = \gamma(16\pi a u^3)d \tag{4.5d}$$

and terminal condition

$$v_T(1) = 0. \tag{4.5e}$$

In fact, independent of the choice of F the computation of β_2 requires the solution of an additional linear parabolic problem.

Clearly, one could also consider higher order terms in the Taylor expansion. Nevertheless, this would amount in solving additional linear problems, such that it is numerically cheaper to perform a few more gradient steps.

5 A Nonlinear Iteration Scheme

Another possibility to compute a minimizer of (3.1) is to solve the first–order optimality system (3.2) directly. To achieve this we propose a nonlinear fixed point iteration which decouples the equations in an appropriate way. Starting from an initial guess for u we solve the nonlinear state system for a new state $x = x(u)$ and then the adjoint system at x and u for $\xi = \xi(x, u)$. Using the state and the adjoint variable we solve the algebraic equation (3.4) for u_{new}. The detailed procedure reads

Algorithm 2.

1. *Set $k = 0$ and choose $u_0 \in U$.*
2. *Given u_k, solve*

$$\begin{aligned}
\partial_t T &= k\Delta T + \frac{1}{3\sigma}\Delta\varphi, \\
-\varepsilon^2\frac{1}{3\sigma}\Delta\varphi + \sigma\varphi &= \sigma(4\pi a T^4) \qquad \text{in } Q, \\
\frac{h}{\varepsilon k}T + n \cdot \nabla T &= \frac{h}{\varepsilon k}\,u_k, \\
\frac{3\sigma}{2\varepsilon}\varphi + n \cdot \nabla\varphi &= \frac{3\sigma}{2\varepsilon}(4\pi a\,u_k^4) \qquad \text{on } \Sigma, \\
T(0, x) &= T_0(x) \qquad \text{in } \Omega,
\end{aligned}$$

for $(T, \rho) \in X$.

3. *Given* (T,ρ) *corresponding to* u_k*, solve*

$$
\begin{gathered}
-\partial_t \xi_T - k\,\Delta\xi_T - 16\pi\sigma a\,T^3 \;-\xi_\rho = -\partial_1 F(T,\nabla T) + \nabla\cdot\partial_2 F(T,\nabla T),\\
-\frac{\varepsilon^2}{3\sigma}\,\Delta\xi_\rho + \sigma\,\rho = -\frac{1}{3\sigma}\,\Delta\xi_T \qquad in\ Q,\\
n\cdot\nabla\xi_T + \alpha\,\xi_T = -n\cdot\partial_2 F(T,\nabla T),\\
n\cdot\nabla\xi_T + \gamma\,\xi_T + \varepsilon^2(n\cdot\nabla\xi_\rho + \gamma\,\xi_\rho) = 0 \quad ,\ on\ \Sigma,\\
\xi_T(1) = 0 \quad ,\ in\ \Omega,
\end{gathered}
$$

for $\xi = (\xi_T, \xi_\rho) \in X^*$.

4. *Given* (ξ_T, ξ_ρ)*, solve*

$$
-k\alpha\,\xi_T + \delta\,(u - u_d) - \frac{4}{3\sigma}(\xi_T + \varepsilon^2\,\xi_\rho)u^3 = 0 \qquad on\ \Sigma,
$$

for u_{k+1}.

5. *Set* $k \to k+1$ *and goto 2.*

Remark 5.1. If the control u appeared in the system only linearly this iteration could be easily interpreted as a gradient algorithm with fixed step length $1/\delta$.

Hence, in each iteration step we only have to solve the nonlinear state equations, the linear adjoint system and an algebraic equation for the update of u. The advantage of this method lies in the fact that it can be easily implemented in an existing finite difference or finite element code and that it incorporates more of the nonlinearity of the problem that arises ¿from the generic nonlinearity wrto. the control variable u.

Remark 5.2. Another possibility is to use the solution $\bar{u}$ of step 4 to compute a descent direction d for the classical descent algorithm via $d = u - \bar{u}$. This can also be interpreted as a damping in Algorithm 2.

6 Numerical Results

This section is devoted to the presentation of numerical experiments assessing the performance of the various algorithms and to show results which demonstrate the capability of our approach to minimize local temperature gradients in glass cooling processes.

For the numerical simulations we first have to discretize the continuous algorithms.

6.1 Discretization

Standard finite differences were used for the space discretization of the Laplacian both for the state and the adjoint system. For simplicity we used a uniform grid with $\Delta x = 0.02$. The time derivative was treated in an implicit way using the forward Euler method with 50 uniform time steps. In the case of our generic example based on the functional J_1 with $p = 2$, we have the following discretized form of the SP_1 state system (1.2)

$$\frac{T_i^{n+1} - T_i^n}{\Delta t} = k\frac{T_{i+1}^{n+1} - 2T_i^{n+1} + T_{i-1}^{n+1}}{\Delta x^2} + \frac{1}{3\sigma}\frac{\rho_{i+1}^{n+1} - 2\rho_i^{n+1} + \rho_{i-1}^{n+1}}{\Delta x^2}, \tag{6.1a}$$

$$-\varepsilon^2\frac{1}{3\sigma}\frac{\rho_{i+1}^{n+1} - 2\rho_i^{n+1} + \rho_{i-1}^{n+1}}{\Delta x^2} + \sigma\rho_i^{n+1} = \sigma\ 4\pi a\big(T_i^n\big)^4, \tag{6.1b}$$

and the corresponding discretization of the adjoint equations (3.3) reads

$$-\frac{\xi_{T,i}^{n+1} - \xi_{T,i}^n}{\Delta t} = k\frac{\xi_{T,i+1}^{n+1} - 2\xi_{T,i}^{n+1} + \xi_{T,i-1}^{n+1}}{\Delta x^2} + 16\pi a\sigma\left(T_i^{n+1}\right)^3\xi_{\rho,i}^{n+1} - (T_i^{n+1} - T_{d,i}^{n+1}), \tag{6.2a}$$

$$-\frac{\varepsilon^2}{3\sigma}\frac{\xi_{\rho,i+1}^{n+1} - 2\xi_{\rho,i}^{n+1} + \xi_{\rho,i-1}^{n+1}}{\Delta x^2} + \sigma\,\xi_{\rho,i}^{n+1} = \frac{1}{3\sigma}\frac{\xi_{T,i+1}^{n+1} - 2\xi_{T,i}^{n+1} + \xi_{T,i-1}^{n+1}}{\Delta x^2}. \tag{6.2b}$$

Furthermore, note that the third order algebraic equation (3.4) for the control u, which occurs in the first order optimality system in the form

$$p_3u^3 + p_1u + p_0 = 0,$$

in general has multiple real roots. Cardano's formula was used to determine their numerical values and in the case of non-uniqueness the ambiguity was resolved by choosing the root that lay in the acceptable range of ambient temperatures between 300 K and 1000 K. Numerical experiments indicated that this heuristics was appropriate in our context and could make the solution unique.

The physical parameters appearing in the equations above had the following values in our numerical experiments. Heat conductivity k and opacity σ were

ε	σ	k	a
1.0	1.0	1.0	$1.806 \cdot 10^{-8}$

assumed to be constant wrto. space and time.

6.2 Comparison of the Functionals

In order to assess which functional is the best model for the optimization criteria described in Section 1 we evaluated the results of the gradient-descent algorithm.

Besides the reduction of the functional, we focus in particular on the value of the sup–norm of the temperature gradients, which is the crucial quantity in applications.

$F_1 = 0.5\,\lvert T - T_d\rvert^2$					
δ	J	$\lVert u-u_d\rVert_2$	$\lVert T-T_d\rVert_2$	$\lVert\nabla T\rVert_2$	$\lVert\nabla T\rVert_\infty$
–	249.3	–	22.33	8.173	153.7
5.00	236.4	1.251	21.56	8.196	151.7
1.00	200.7	5.054	19.39	8.337	145.6
0.50	174.5	8.236	17.75	8.556	140.2
0.10	110.2	18.77	13.61	10.09	126.8
0.05	88.97	24.17	12.20	11.31	143.8

$F_2 = 0.5\,\lvert\nabla T\rvert^2$					
δ	J	$\lVert u-u_d\rVert_2$	$\lVert T-T_d\rVert_2$	$\lVert\nabla T\rVert_2$	$\lVert\nabla T\rVert_\infty$
–	33.40	–	22.33	8.173	153.7
5.00	33.16	0.172	22.32	8.134	153.7
1.00	32.21	0.833	22.28	7.983	154.0
0.50	31.17	1.594	22.26	7.815	154.3
0.10	25.43	6.108	22.32	6.865	157.0
0.05	21.66	9.593	22.66	6.222	160.3

$F_3 = F_1 + F_2$					
δ	J	$\lVert u-u_d\rVert_2$	$\lVert T-T_d\rVert_2$	$\lVert\nabla T\rVert_2$	$\lVert\nabla T\rVert_\infty$
–	282.7	–	22.33	8.173	153.7
5.00	269.8	1.276	21.56	8.158	151.8
1.00	234.9	5.075	19.43	8.164	145.9
0.50	210.6	8.118	17.90	8.238	141.0
0.10	159.6	16.71	14.58	8.874	126.5
0.05	147.1	19.84	13.73	9.277	121.1

Table 1: Reduction of the value of the functionals J_1, J_2, J_3

¿From Table 1 it can be seen that J_1 and the difference $T - T_d$ are significantly diminished when δ is decreased. At the beginning this is true for the L^∞–norm of ∇T but it tends to increase again when $\delta < 0.1$. We explain this by the fact that the control temperature u had values below room temperature in which is not sensible from an engineering point view. As a result, gradients at the boundary become very large due to the large difference between the temperature in the exterior and the interior. Although functional J_2 effectively reduces the L^2-norm of the gradient we observe that, in contrast, the corresponding maximum norm increases. And, lastly, the third functional, which incorporates both terms, is seen

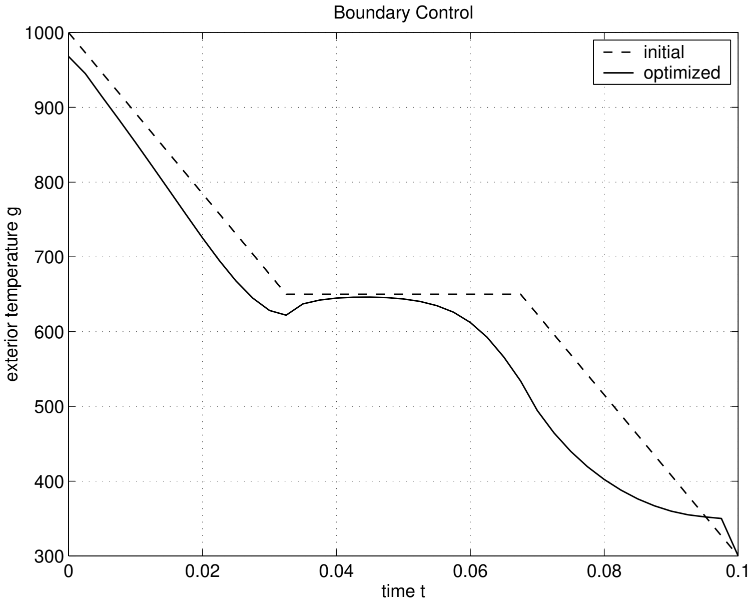

Figure 1: Optimal boundary control u for functional J_3. The weight of the control was $\delta = 0.05$.

to result in decreasing norms of the difference while the 2-norms of the gradient increase when δ becomes smaller. The overall value of the functional, however, is continuously decreased as expected. It should be highlighted that although the 2-norm of the gradient becomes larger the maximum norm shows the opposite behaviour. We, therefore, believe that J_3 is the best choice among the three functionals under consideration here. The experiments below were done using this functional. The computed cooling profile can be found in Figure 1 and the corresponding distribution of $|\nabla T|$ is given in Figure 4, where one also finds the uncontrolled state for reference. Note that at a first glance the region with large gradients seems broader, but it has to be observed that the controlled profile has already reached a smaller end temperature (see Figure 3).

6.3 Effect of Higher Order L^p–Norms

As already mentioned in Section 2, the maximum norm of the gradient would be the most appropriate norm for measuring local temperature gradients. But this is mathematically much more difficult since the standard Lagrange formalism is no more applicable. Norms of order $p > 2$, however, could be a compromise as they might better emulate the maximum norm while avoiding these technical problems. Therefore, it is interesting to ascertain whether they can give better results.

$F = \frac{1}{4}\|T - T_d\|^4 + \frac{1}{4}\|\nabla T\|^4$					
δ [$\cdot 10^3$]	J [$\cdot 10^6$]	$\|\|u - u_d\|\|_2$	$\|\|T - T_d\|\|_4$	$\|\|\nabla T\|\|_4$	$\|\|\nabla T\|\|_\infty$
–	2.422	–	55.09	26.32	153.7
100	2.307	0.815	54.18	26.26	152.0
10	1.756	5.406	49.44	26.08	143.1
1	0.879	17.70	39.59	25.62	121.6
0.1	0.565	27.08	35.40	27.18	108.5

$F = \frac{1}{8}\|T - T_d\|^8 + \frac{1}{8}\|\nabla T\|^8$					
δ [$\cdot 10^{12}$]	J [$\cdot 10^{12}$]	$\|\|u - u_d\|\|_2$	$\|\|T - T_d\|\|_8$	$\|\|\nabla T\|\|_8$	$\|\|\nabla T\|\|_\infty$
–	812.3	–	94.56	56.74	153.7
10	605.9	2.661	90.41	56.16	148.3
1	288.6	9.336	81.11	55.30	135.7
0.1	92.6	19.63	69.32	51.92	120.0
0.01	40.4	26.45	63.08	50.81	111.4
0.001	34.9	27.40	62.35	51.22	110.7

$F = \frac{1}{16}\|T - T_d\|^{16} + \frac{1}{16}\|\nabla T\|^{16}$					
δ [$\cdot 10^{30}$]	J [$\cdot 10^{30}$]	$\|\|u - u_d\|\|_2$	$\|\|T - T_d\|\|_{16}$	$\|\|\nabla T\|\|_{16}$	$\|\|\nabla T\|\|_\infty$
–	366.7	–	129.0	89.44	153.7
10	269.5	1.450	126.2	88.45	150.8
1	127.1	4.896	119.9	86.56	144.2
0.1	45.5	9.419	112.4	83.77	136.5
0.01	20.0	12.86	107.2	81.32	132.1
0.001	15.6	13.83	105.7	80.67	130.8

Table 2: Reduction of J_3 for higher order L^p-norms with $p = 4$, 8 and 16.

Table 2 reveals in the cases $p = 4$ and $p = 8$ that the sup–norm of the temperature gradient could be reduced even more than in the case $p = 2$ (see Table 1). The reduction obtained here is approximately 30% of the value without optimization compared with 20% in the case $p = 2$. Figure 2 shows the proposed profile of u for $p = 8$ after optimization with $\delta = 10^9$ and the distribution of $|\nabla T|$ is given in Figure 5. Increasing the order of p did not seem to result in further improvements because for $p = 16$ the gain was lower than in the previous cases. When it is necessary to use higher order norms it seems to be a good choice to use $p = 4$ or 8 for the L^p-norms in the functional.

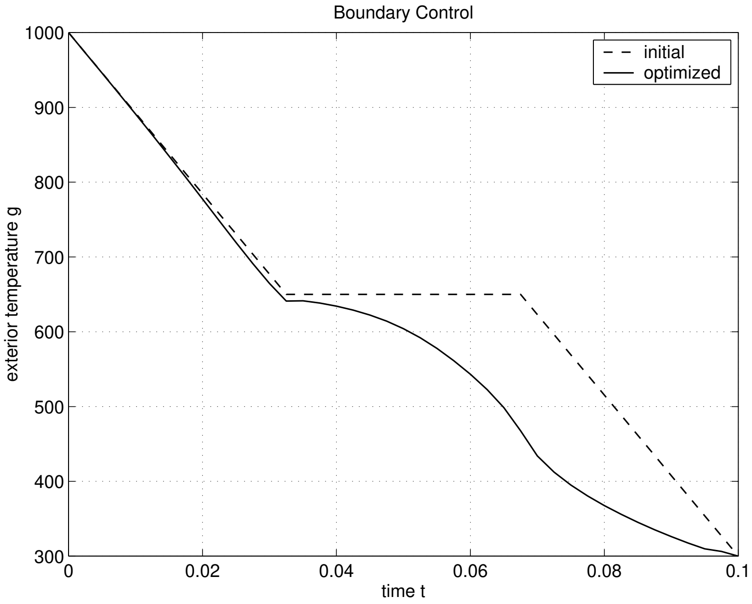

Figure 2: Optimal boundary control u for functional J_3 with $p = 8$. Weight of the control: $\delta = 10^9$.

6.4 Performance of Different Algorithms

We consider the gradient–descent Algorithm 1 from Section 4 with the two step size choices (4.3) and (4.4)

$$\beta_1 = \min\left(1, \|d\|_\infty^{-1}\right), \qquad \beta_2 = \frac{\|d\|_\Sigma^2}{\varepsilon\,\|d\|_\Sigma^2 + \|v\|_Q^2},$$

where $d = \nabla \hat{J}$ and v is the solution of (4.5). Furthermore, the nonlinear iteration of Algorithm 2 is taken into account. And finally, we investigate the behaviour of a hybrid method which uses the direction $d = u - \bar{u}$ as descent direction instead of the gradient, where $\bar{u}$ is the control coming from the third order equation

$$-k\alpha\,\xi_T + \delta\,(\bar{u} - u_d) - \frac{4}{3\sigma}(\xi_T + \varepsilon^2\,\xi_\rho)\bar{u}^3 = 0 \quad \text{ on } \Sigma$$

in step 4 of Algorithm 2 of the adjoint method. In our numerical experiments we used the functional J_1 with the step profile u_d as before. The iteration was terminated when the update for u was < 0.01 uniformly for all discrete times.

As can be seen in Table 3, the heuristic approach performs very well while the linear method using the step size β_2 needs significantly more iterations. This is due to the fact that the proposed step sizes were in general rather small which indicates that this approach underestimates the true step size. Nevertheless,

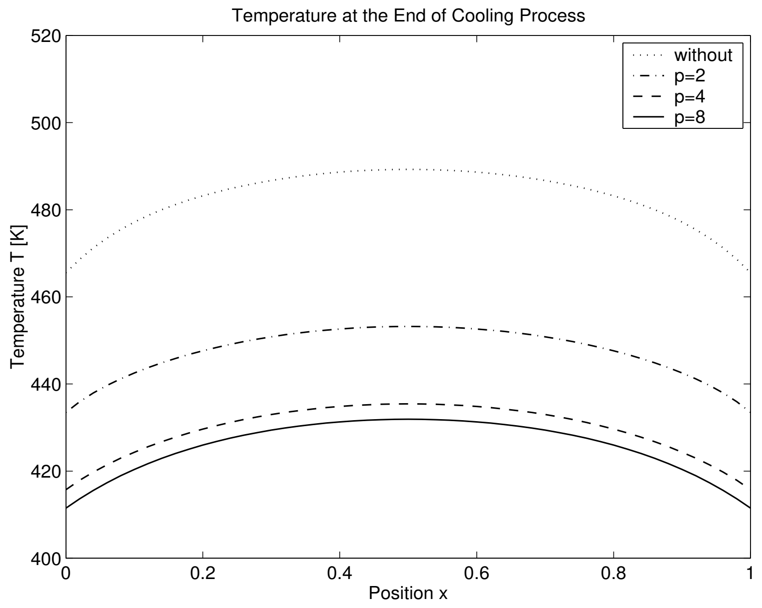

Figure 3: End temperatures at $t = \tau$ without optimization and with optimization using J_3 for $p = 2(\delta = 0.05)$, $p = 4(\delta = 100)$ and $p = 8(\delta = 10^9)$, respectively

δ	heuristic	linear	nonlinear	hybrid
1.00	8	37	7	34
0.75	4	44	9	50
0.50	7	55	17	63
0.25	13	68	–	90
0.10	26	73	–	134

Table 3: Number of iterations needed by different algorithms.

underestimation at least does not affect the overall global convergence of the algorithm. The nonlinear iteration scheme gives similar performance, when the number of iterations is concerned, as the heuristic approach but it fails to work when δ becomes smaller. This can be understood by the fact that the contractivity of the fixed point mapping defined by Algorithm 2 is violated for small δ. To overcome this problem we finally considered a hybrid method, which is a damped version of Algorithm 2 where the damping parameter is computed via (4.4). This method converges slowly and does not seem to be competitive. Numerical experiments indicate that in some cases the scalar product $(u - \bar{u}, \nabla \hat{J})$ can oscillate between negative and positive values such that $u - \bar{u}$ is in general not guaranteed to be a descent direction.

7 Conclusions

We studied an optimal boundary control problem for glass cooling processes with the aim to minimize thermal stresses which strongly correlate with local temperature gradients. Allowing for general cost functionals we embedded this problem into the setting of general constrained optimization problems and derived the first order optimality system. For the numerical solution we proposed two algorithms based on adjoint variables, i.e. a gradient–descent method and a nonlinear iteration scheme. The performance of the algorithms was tested for the optimal cooling of a glass slab and, in particular, different step size rules were considered.

In fact, the gradient algorithm with a heuristic step size rule proved to be very robust, performed most successfully and yielded a significant reduction of the local temperature gradients of approximately 30%. We emphasize that the adjoint approach can be easily adopted for higher dimensional problems and also allows for the direct computation of the sensitivities of this problem.

Future work will focus on better approximations of the radiative heat transfer equation, e.g. frequency dependence [8] or the SP_3 system [9], and more sophisticated second order optimization algorithms, which may better resolve the high nonlinearity in this problem. This will be essential to keep computing times low in the three dimensional case.

References

[1] M. K. Choudhary and N. T. Huff. Mathematical modeling in the glass industry: An overview of status and needs. *Glastech. Ber. Glass Sci. Technol.*, 70:363–370, 1997.

[2] L.S. Hou and S.S. Ravindran. Computations of boundary optimal control problems for an electrically conducting fluid. *J. Comp. Phys.*, 128:319–330, 1996.

[3] K. Ito and K. Kunisch. Augmented Lagrangian-SQP-methods for nonlinear optimal control problems of tracking type. *SIAM J. Control and Optimization*, 34:874–891, 1996.

[4] K. Ito and S.S. Ravindran. Optimal control of thermally convected fluid flows. *SIAM J. Sci. Comput.*, 19(6):1847–1869, 1998.

[5] E. Larsen, G. Pomraning, and V. C. Badham. Asymptotic Analysis of Radiative Transfer Problems. *J. Quant. Spectr. and Radiative Transfer*, 29:285–310, 1983.

[6] E. W. Larsen. Diffusion Theory as an Asymptotic Limit of Transport Theory for Nearly Critical Systems with Small Mean Free Paths. *Annals of Nuclear Energy*, 7:249, 1980.

[7] E. W. Larsen, J. E. Morel, and W. F. Miller. Asymptotic Solution of Numerical Transport Problems in Optically Thick, Diffusive Regimes. *J. Comp. Phys.*, 69:283–324, 1987.

[8] E. W. Larsen, G. Thömmes, and A. Klar. Frequency Averaged Approximations to the Equations of Radiative Heat Transfer in Glass. *submitted*, 2001.

[9] E. W. Larsen, G. Thömmes, and A. Klar. Simplified P_N Approximations to the Equations of Radiative Heat Transfer in Glass I: Modelling. *submitted*, 2001.

[10] R. Pinnau and G. Thömmes. Analysis of an optimal boundary control problem for the SP_1 system. *In preparation*, 2001.

[11] G. Thömmes, R. Pinnau, M. Seaid, T. Götz, and A. Klar. Numerical methods and optimal control for glass gooling processes. *Submitted for Publication*, 2001.

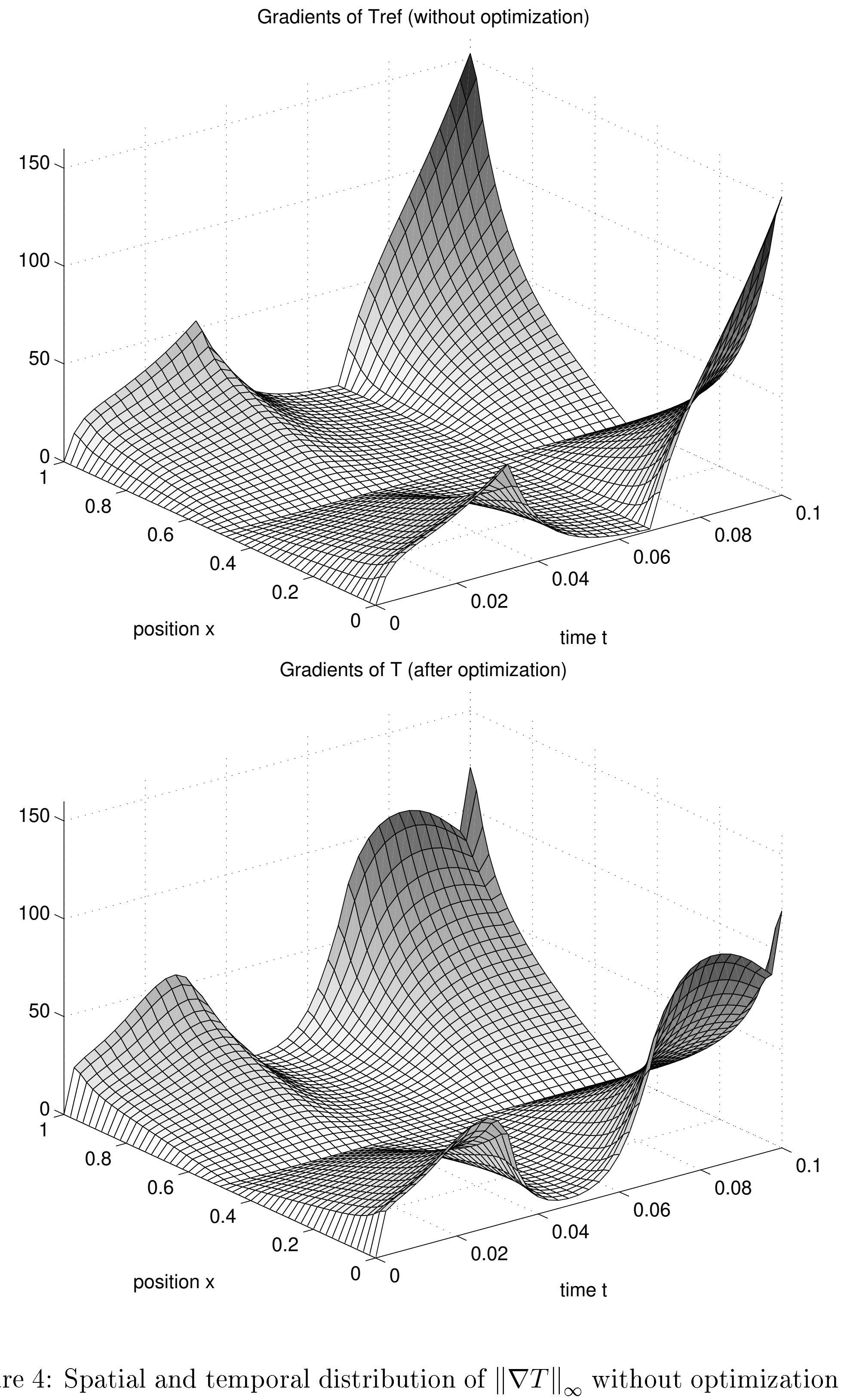

Figure 4: Spatial and temporal distribution of $\|\nabla T\|_\infty$ without optimization and with optimization using J_3 for $p = 2(\delta = 0.05)$

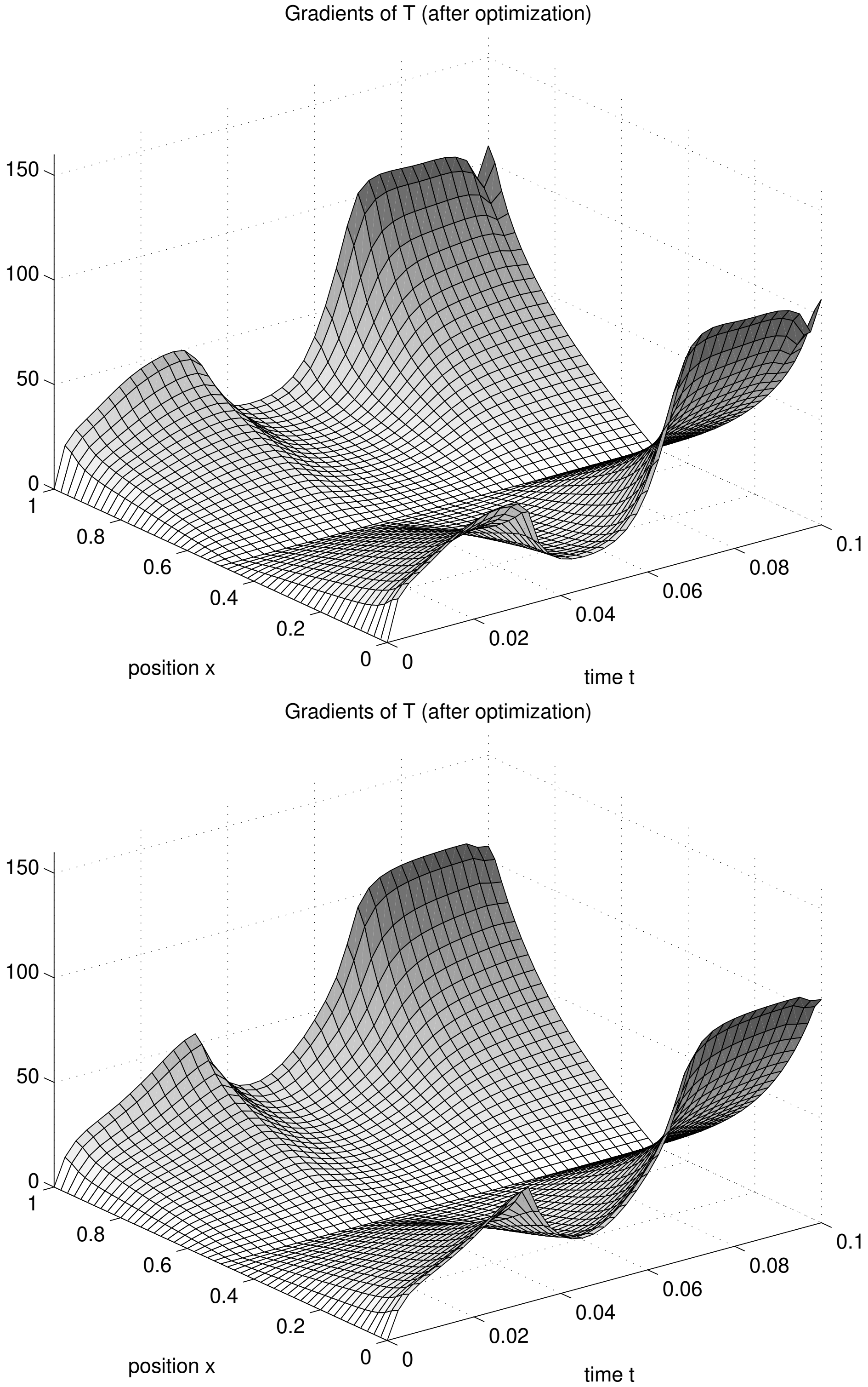

Figure 5: Spatial and temporal distribution of $\|\nabla T\|_\infty$ with optimization using J_3 for $p = 4(\delta = 100)$ and $p = 8(\delta = 10^9)$