

Optimal Control of the Solid Fuel Ignition Model

Andreas Kauffmann* Karl Kunisch†

Abstract

Optimal control problems for the solid fuel ignition model are analysed. For a class of cost functionals optimality conditions are derived. Their solution by means of SQP-methods with special emphasis on appropriate preconditioners is discussed. Some results on feedback stabilization are included as well.

1 The optimal control problem

The aim of this contribution is the study of optimal control problems related to the solid fuel ignition problem [2] with distributed control

$$y_t - \Delta y - \delta e^y = u, \quad y|_{t=0} = y_0, \quad y|_{\Sigma} = 0, \quad (1)$$

or with Neumann boundary control

$$y_t - \Delta y - \delta e^y = 0, \quad y|_{t=0} = y_0, \quad \frac{\partial}{\partial n} y|_{\Sigma} = u, \quad (2)$$

where Σ is the lateral boundary of a time-space cylinder $Q = (0, T) \times \Omega$ with Ω a bounded domain in \mathbb{R}^n with smooth boundary. The existence problem for optimal control problems governed by (1) or (2) is complicated due to the lack of coercivity of the state equation. A-priori bounds for both y and u must - in an appropriate sense - be guaranteed through the cost functional. Two natural choices are to either involve ∇y or e^y in the cost functional. The first possibility is considered in [3], the second one is considered here, where we choose as cost the tracking type functional

$$J(y, u) = \frac{1}{2} \|y - y_d\|_{L^2(Q)}^2 + \frac{1}{2} \|e^y - e^{y_d}\|_{L^2(Q)}^2 + \frac{\nu}{2} \|u\|_X^2, \quad (3)$$

where $\nu > 0$, $y_d \in L^\infty(Q)$ and $X = L^2(Q)$ in the case of distributed control and $X = H^1(\Sigma)$ for Neumann boundary control. The set of admissible control-state pairs is defined to be

$$\mathcal{P} := \{(y, u) \in H^{2,1}(Q) \times U_{ad} : (y, u) \text{ satisfy (1), respectively (2)}\}, \quad (4)$$

*Fachbereich Mathematik, Technische Universität Berlin, D-10623 Berlin, Straße des 17. Juni 136, Germany

†Institut für Mathematik, Karl-Franzens-Universität Graz, A-8010 Graz, Austria, supported by the Fonds zur Förderung der wissenschaftlichen Forschung under SFB 03 „Optimierung und Kontrolle“.

with $U_{ad} \subset X$ a nonempty, closed, convex set. The problems under investigation are

$$\min_{(y,u) \in \mathcal{P}} J(y, u), \tag{5}$$

with J given by (3). Henceforth, in the case of boundary control we assume that $\mathcal{P} \neq \emptyset$. For the distributed control case with $U_{ad} = L^2(Q)$ this is satisfied without further assumption.

Proving existence of an optimal pair (y^*, u^*) is then straightforward: $\mathcal{P} \neq \emptyset$ implies existence of a minimizing sequence $\{(y_n, u_n)\}$. Utilizing the state equation one argues existence of a strongly in $L^2(Q)$ convergent subsequence of $\{y_n\}$. Using lower semicontinuity of J with respect to (y, u) it can be shown that subsequential limits of $\{(y_n, u_n)\}$ are a solution to J . Let us note that, provided that $\mathcal{P} \neq \emptyset$, solutions to (5) exist also for values of δ for which the uncontrolled state equation may have no solution.

Turning to optimality conditions it will be convenient at first to introduce some notation and to consider differentiability properties of the exponential function. We shall utilize the anisotropic Sobolev space

$$H^{2,1}(Q) = \{\varphi \in L^2(0, T; H^2(\Omega)) : \varphi_t \in L^2(0, T; L^2(\Omega))\},$$

and the space $H_{0,\cdot}^{2,1}(Q) = \{\varphi \in H^{2,1}(Q) : \varphi|_{\Sigma} = 0\}$. Applying the theory of interpolation spaces and traces one obtains that $H^{2,1}(Q) \subset L^\infty(0, T; H^1(\Omega))$ [7].

Proposition 1 *Let $n \leq 2$ and $\Omega \subset \mathbb{R}^n$ be bounded. Then the operator $F : H_{0,\cdot}^{2,1}(Q) \rightarrow L^\infty(0, T; L^p(\Omega))$ with $F(\varphi) = e^\varphi$ is Fréchet differentiable with Fréchet derivative $F'(\varphi) = e^\varphi$ (for $p < \infty$). Moreover there exist constants c_1, c_2 depending only on Q such that*

$$\|e^\varphi\|_{L^\infty(0, T; L^p(\Omega))}^p \leq c_2 \exp\left(c_1 p^2 \|\varphi\|_{H^{2,1}(Q)}^2\right).$$

An analogous result holds with $H_{0,\cdot}^{2,1}(Q)$ replaced by $H^{2,1}(Q)$, if Ω additionally is convex.

The following example demonstrates that it is not possible to extend these results to $n \geq 3$.

Example 1 *Let $Q = (0, 1) \times \Omega$ where $\Omega \subset \mathbb{R}^3$ is the unit ball. Define*

$$\varphi(t, x) = (t^\alpha + |x|^2)^{-\rho}$$

and choose $0 < \rho \leq \frac{1}{8}$ and $\frac{4}{5} < \alpha < 1$. Then $\varphi \in H^{2,1}(Q)$ and $e^\varphi \notin L^p(Q)$ for every $1 \leq p \leq \infty$.

For proving the existence of Lagrange multipliers and for deriving first order optimality systems an adapted penalty functional technique, see [1], is utilized. For a given optimal pair (y^*, u^*) the adapted penalty functional $J_\varepsilon^\alpha(y, u)$ is defined by

$$\begin{aligned} J_\varepsilon^\alpha(y, u) &= \frac{1}{2} \|y - y_d\|_{L^2(Q)}^2 + \frac{1}{2} \|e^y - e^{y_d}\|_{L^2(Q)}^2 + \frac{\nu}{2} \|u\|_{L^2(Q)}^2 \\ &+ \frac{1}{2\varepsilon} \|y_t - \Delta y - \delta e^y - u\|_{L^2(Q)}^2 + \frac{1}{2} \|y - y^*\|_{L^2(Q)}^2 \\ &+ \frac{1}{2} \|u - u^*\|_{L^2(Q)}^2, \end{aligned} \tag{6}$$

respectively by

$$\begin{aligned}
J_\varepsilon^a(y, u) &= \frac{1}{2} \|y - y_d\|_{L^2(Q)}^2 + \frac{1}{2} \|e^y - e^{y_d}\|_{L^2(Q)}^2 + \frac{\nu}{2} \|u\|_{H^1(\Sigma)}^2 \\
&+ \frac{1}{2\varepsilon_1} \|y_t - \Delta y - \delta e^y\|_{L^2(Q)}^2 + \frac{1}{2\varepsilon_2} \left\| \frac{\partial}{\partial n} y|_\Sigma - u \right\|_{H^{\frac{1}{2}, \frac{1}{4}}(\Sigma)}^2 \\
&+ \frac{1}{2} \|y - y^*\|_{L^2(Q)}^2 + \frac{1}{2} \|u - u^*\|_{L^2(\Sigma)}^2.
\end{aligned} \tag{7}$$

The unconstrained optimization problem $(P_\varepsilon^a) : \min_{(y,u)} J_\varepsilon^a(y, u)$ has an optimal solution $(y^\varepsilon, u^\varepsilon)$ for every $\varepsilon > 0$, and $(y^\varepsilon, u^\varepsilon)$ converges weakly in $H^{2,1}(Q) \times X$ to (y^*, u^*) . Thus there exists an approximation procedure for every optimal pair (y^*, u^*) .

An optimality system for (5) is derived by establishing the optimality system for (P_ε^a) and passing to the limit $\varepsilon \rightarrow 0$. For this purpose it is necessary to assure Fréchet differentiability of $J_\varepsilon^a(y, u)$, $J(y, u)$ and of the equality constraint (1) or (2). This requires the restriction $n \leq 2$. We now focus on the distributed control case, the boundary control being treated in [5]. The optimality system for $J_\varepsilon^a(y, u)$ is found to be

$$\begin{aligned}
(-\partial_t - \Delta - \delta e^{y^\varepsilon}) \lambda^\varepsilon &= -(y^\varepsilon - y_d + e^{y^\varepsilon} (e^{y^\varepsilon} - e^{y_d}) + y^\varepsilon - y^*), \\
\lambda^\varepsilon|_{t=T} &= 0, \quad \lambda^\varepsilon|_\Sigma = 0,
\end{aligned} \tag{8}$$

$$\int_Q (\nu u^\varepsilon - \lambda^\varepsilon + u^\varepsilon - u^*) (v - u^\varepsilon) dQ \geq 0 \quad \text{for all } v \in U_{ad}, \tag{9}$$

where

$$\lambda^\varepsilon := \frac{1}{\varepsilon} (y^\varepsilon_t - \Delta y^\varepsilon - \delta e^{y^\varepsilon} - u^\varepsilon) \tag{10}$$

is an Ansatz for the Lagrange multiplier. To prove boundedness of λ^ε for $\varepsilon \rightarrow 0$ the solvability of the adjoint equation (8) has to be investigated. It has a (unique) solution λ^ε , if the right hand side is in $L^2(0, T; H^{-1}(\Omega))$ and $e^{y^\varepsilon} \in L^2(0, T; L^3(\Omega))$ (see [7]). The boundedness of the cost functional for $(y^\varepsilon, u^\varepsilon)$ just guarantees $e^{y^\varepsilon} \in L^2(Q)$, which is not enough. But $y^\varepsilon_t - \Delta y^\varepsilon - \delta e^{y^\varepsilon} - u^\varepsilon \in L^2(Q)$ implies $y^\varepsilon \in H^{2,1}(Q)$ [8]. Applying Proposition 1 it is possible to pass to the limit as $\varepsilon \rightarrow 0^+$ in (8) and (9) to obtain the optimality system for (5):

Theorem 1 *Suppose $n \leq 2$ and (y^*, u^*) is an optimal solution of (5) with distributed control. Then there exists a unique Lagrange multiplier $\lambda^* \in H^{2,1}(Q)$ such that*

$$\begin{aligned}
(-\partial_t - \Delta - \delta e^{y^*}) \lambda^* + y^* - y_d + e^{y^*} (e^{y^*} - e^{y_d}) &= 0, \\
\lambda^*|_{t=T} &= 0, \quad \lambda^*|_\Sigma = 0,
\end{aligned}$$

$$\int_Q (\nu u^* - \lambda^*) (v - u^*) dQ \geq 0 \quad \text{for all } v \in U_{ad},$$

$$\begin{aligned}
y^*_t - \Delta y^* - \delta e^{y^*} - u^* &= 0, \\
y^*|_{t=0} &= y_0, \quad y^*|_\Sigma = 0.
\end{aligned}$$

Moreover one obtains sufficiency of the first order conditions, if $\|y^* - y_d\|_{H^{2,1}(Q)}$ is sufficiently small.

Theorem 2 *If $\|y^* - y_d\|_{H^{2,1}(Q)}$ is sufficiently small then $\mathcal{L}''(y^*, u^*; \lambda^*)$ is positive definite on the nullspace of the linearized constraint (1) in (y^*, u^*) .*

An analogous result for the boundary control case can be found in [5].

For numerical purposes it is of interest to substitute the nonlinearity e^y by introduction of an additional variable $w = e^y$. Then the constraint is split into a linear PDE and an algebraic constraint. For these problems one obtains the following optimality systems:

$$\begin{aligned} (-\partial_t - \Delta) \lambda^* + e^{y^*} \mu^* + y^* - y_d &= 0, \\ \lambda^*|_{t=T} &= 0, \quad \lambda^*|_{\Sigma} = 0, \\ w^* - e^{y^*} - \delta \lambda^* - \mu^* &= 0, \\ \int_Q (\nu u^* - \lambda^*) (v - u^\varepsilon) dQ &\geq 0 \quad \text{for all } v \in U_{ad}, \\ y^*_t - \Delta y^* - \delta e^{y^*} - u^* &= 0, \\ y^*|_{t=T} = y_0, \quad y^*|_{\Sigma} &= 0, \\ e^{y^*} - w^* &= 0, \end{aligned}$$

in the case of distributed control and

$$\begin{aligned} (-\partial_t - \Delta) \lambda^* + e^{y^*} \mu^* + y^* - y_d &= 0, \\ \lambda^*|_{t=T} = 0, \quad \frac{\partial}{\partial n} \lambda^*|_{\Sigma} &= 0, \\ w^* - e^{y^*} - \delta \lambda^* - \mu^* &= 0, \\ \int_{\Sigma} (\nu (I - \mathcal{D}_{tt} - \Delta_{\Gamma}) u^* - \lambda^*|_{\Sigma}) (v - u^\varepsilon) d\Sigma &\geq 0 \quad \text{for all } v \in U_{ad}, \\ y^*_t - \Delta y^* - \delta e^{y^*} - u^* &= 0, \\ y^*|_{t=T} = y_0, \quad \frac{\partial}{\partial n} y^*|_{\Sigma} &= u^*, \\ e^{y^*} - w^* &= 0, \end{aligned}$$

in the case of Neumann boundary control. Here $\Gamma = \partial\Omega$ and Δ_{Γ} denotes the Laplace-Beltrami operator and $\mathcal{D}_{tt} := -\partial_t^* \partial_t$, where $\partial_t^* : L^2(\Sigma) \rightarrow H^1(\Sigma)^*$ is the adjoint operator of $\partial_t : H^1(\Sigma) \rightarrow L^2(\Sigma)$. In both cases existence and boundedness of Lagrange multipliers $\lambda^* \in H^{2,1}(Q)$ arises from bijectivity of the operator $(\partial_t - \Delta - \delta e^y) : H^{2,1}(Q) \rightarrow L^2(Q)$ for $y \in H^{2,1}(Q)$, if $n \leq 2$.

We define $\mathcal{L} : H^{2,1}(Q) \times L^2(Q) \times H^{2,1}(Q) \rightarrow \mathbb{R}$ by

$$\mathcal{L}(y, u; \lambda) = J(y, u) + (y_t, \lambda)_Q + (\nabla y, \nabla \lambda)_Q - \delta(e^y, \lambda)_Q - (u, \lambda)_Q$$

and let denote ' the Fréchet derivative w. r. t. (y, u) .

2 Numerical solution using SQP methods

For numerical solution of the optimal control problem (5) an augmented Lagrangian SQP method was used. This will be described for the following general problem: Given some real Hilbert spaces X, Y a cost functional $J : X \rightarrow \mathbb{R}$ and a constraint operator $e : X \rightarrow Y$

$$\min J(x) \quad \text{subject to} \quad e(x) = 0. \quad (11)$$

The Lagrange and the augmented Lagrange functionals are defined by

$$\begin{aligned} \mathcal{L}(x; \lambda) &:= J(x) + \langle e(x), \lambda \rangle_Y, \\ \mathcal{L}_c(x; \lambda) &:= J(x) + \langle e(x), \lambda \rangle_Y + \frac{c}{2} \|e(x)\|_Y^2. \end{aligned}$$

To prove convergence of augmented Lagrangian SQP methods the following assumptions are necessary:

- There exists a solution x^* of problem (5).
- The functionals J and e are twice continuously Fréchet differentiable in a neighbourhood $V(x^*)$ of x^* with J'', e'' continuous on $V(x^*)$.
- The linearization $e'(x^*) : X \rightarrow Y$ is surjective.
- There exists $\kappa > 0$ such that

$$\mathcal{L}''(x^*, \lambda^*)(\xi, \xi) \geq \kappa \|\xi\|_X^2 \quad \text{for all} \quad \xi \in \ker e'(x^*) \subset X.$$

From the first three conditions one can conclude existence of a Lagrange multiplier λ^* such that $\mathcal{L}'(x^*, \lambda^*) = 0$ and we know $\mathcal{L}''(x^*, \lambda^*)$ to be positive semidefinite on $\ker e'(x^*) \subset X$. Together these assumptions guarantee local convergence of the following Algorithm 1 for appropriate choices of c . (see e.g. [4], or [6]).

Algorithm 1 *Let $\lambda_0 \in Y$, $x_0 \in X$ and $c \geq 0$ be given.*

- *Update the Lagrange multiplier and determine x_{k+1} by solving*

$$\begin{pmatrix} \mathcal{L}''_c(x_k, \lambda_k) & e'(x_k)^* \\ e'(x_k) & 0 \end{pmatrix} \begin{pmatrix} x_{k+1} - x_k \\ \lambda_{k+1} - \lambda_k \end{pmatrix} + \begin{pmatrix} \mathcal{L}'_c(x_k, \lambda_k) \\ e(x_k) \end{pmatrix} = 0.$$

The reason for utilizing the penalty term involving the parameter c is that it can have a globalizing effect. We stress the fact that c is not considered as a parameter which needs to be increased to obtain the solution to (11), rather a small $c \neq 0$ can be computationally advantageous. For the application of Algorithm 1 to problem (5) containing the severe exponential nonlinearity we also utilized at times a cheap linesearch based on retraction from the full step by bisecting the steplength. Moreover the computation of $\mathcal{L}'_c, \mathcal{L}''_c$ is reduced to the computation of $\mathcal{L}'_0, \mathcal{L}''_0$ by the well known technique of using an intermediate first order Lagrange multiplier update. This results in:

Algorithm 2 Let $\lambda_0 \in Y$, $x_0 \in X$ and $c \geq 0$ be given.

- Set $\tilde{\lambda} = \lambda_k + ce(x_k)$.
- Solve

$$\begin{pmatrix} \mathcal{L}_0''(x_k, \tilde{\lambda}) & e'(x_k)^* \\ e'(x_k) & 0 \end{pmatrix} \begin{pmatrix} \partial x \\ \partial \lambda \end{pmatrix} + \begin{pmatrix} \mathcal{L}_0'(x_k, \tilde{\lambda}) \\ e(x_k) \end{pmatrix} = 0 \quad (12)$$

- Determine $\vartheta \in (0, 1]$ by means of a cheap line search and set $x_{k+1} = x_k + \vartheta \partial x$ and $\lambda_{k+1} = \tilde{\lambda} + \vartheta \partial \lambda$.

In [5] it is verified that the abstract algorithm described above is applicable to (5) with $x = (y, u)$ or $x = (y, u, w)$ depending on whether the additional splitting is utilized or not. - Many choices have to be made before an efficient implementation of Algorithm 2 is accomplished. We address some of them.

1. There are two forms of problem formulations : The original problem with variables $(y, u; \lambda)$, called **3x3** in the following, and the problem formulation with substitution $e^y = w$ with variables $(y, u, w; \lambda, \mu)$, called **5x5** in the following. Which of these problems is easier to solve by augmented Lagrangian SQP ?
2. Are line searches essential for convergence ?
3. What is the role of c ?
4. Which solver should be used for (12) ?

The experience with many numerical tests for distributed as well as boundary optimal control for the solid fuel ignition model are summarized in the following discussion.

For solving the 3x3 problem by augmented Lagrangian SQP methods the choice of the augmentation parameter c is crucial. In some examples it is impossible to obtain convergence using $c = 0$. Admittedly, in some tests the range of c for which convergence is achieved is small. The introduction of the linesearch rules does not significantly increase convergence speed. But it enlarges the range of usable parameters c . This is quite helpful if bad initial values are given. So in the 3x3 case linesearch rules and use of $c > 0$ seem to be helpful, if one does not start near the optimum.

When solving the 5x5 problem by augmented Lagrangian SQP methods without linesearch rule the choice of $c > 0$ did not influence significantly convergence (or lack thereof). But the use (of moderate values!) of $c > 0$ together with a linesearch rule increases convergence speed compared to the unaugmented SQP method.

The tests also indicated, that substitution of the exponential term is useful numerically and the (5x5) problem is more amenable computationally than the (3x3) problem. It is much easier to obtain convergence when solving problem (5x5) than when solving problem (3x3). The choice of c for (5x5) is not so crucial for convergence as the choice of c for (3x3). Also (at least together with a linesearch rule) convergence is obtained faster when solving problem (5x5).

We turn now to the most crucial and costly step of the SQP method applied to (5): the realization of the linear systems (12).

Considering the (5x5) problem with distributed control the linear operator

$$M(x; \lambda) := \begin{pmatrix} \mathcal{L}_0''(x_k, \tilde{\lambda}) & e'(x_k)^* \\ e'(x_k) & 0 \end{pmatrix}$$

is given by

$$M(x; \lambda) = \left(\begin{array}{ccc|cc} I + \mu e^y & & & -\partial_t - \Delta & e^y \\ & I & & -\delta I & -I \\ & & \nu I & -I & \\ \hline \partial_t - \Delta & -\delta I & -I & & \\ e^y & & -I & & 0 \end{array} \right).$$

Every discretization of the problem results in a large and sparse matrix representation of $M(x; \lambda)$. So to solve (12) iterative methods should be used. One of the most powerful iterative solvers for linear equations with a symmetric structure is the preconditioned conjugate gradient (CG) method. For convergence speed of preconditioned CG the choice of a preconditioning operator B^{-1} is crucial. One has to balance between good approximation quality of $B \approx M(x; \lambda)$ and cheap computation of $V := B^{-1}$ applied to a vector.

The fact that parabolic operators occur in the operator M suggests the use of parabolic equation solvers as part of preconditioning, utilizing their smoothing properties. The following abbreviations will be used for parabolic operators:

$$\begin{aligned} P &:= \partial_t - \Delta, & R &:= (\partial_t - \Delta)^{-1}, \\ P^* &:= -\partial_t - \Delta, & R^* &:= (-\partial_t - \Delta)^{-1}. \end{aligned}$$

The inverse parabolic operators are understood with zero initial and zero boundary conditions. Then two possible preconditioners basing on a blockwise construction have the following form:

$$\begin{aligned} V_1 &:= \left(\begin{array}{ccc|cc} \cdot & \cdot & \cdot & R & \cdot \\ \cdot & I & \cdot & \cdot & \cdot \\ \cdot & \cdot & \frac{1}{\nu} I & \cdot & \cdot \\ \hline R^* & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & I \end{array} \right) = \left(\begin{array}{ccc|cc} \cdot & \cdot & \cdot & P^* & \cdot \\ \cdot & I & \cdot & \cdot & \cdot \\ \cdot & \cdot & \nu I & \cdot & \cdot \\ \hline P & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & I \end{array} \right)^{-1}, \\ V_2 &:= \left(\begin{array}{ccc|cc} \cdot & \cdot & \cdot & R & \cdot \\ \cdot & I & \cdot & \cdot & \cdot \\ \cdot & \cdot & \frac{1}{\nu} I & \cdot & \cdot \\ \hline R^* & \cdot & \cdot & Q & \cdot \\ \cdot & \cdot & \cdot & \cdot & I \end{array} \right) = \left(\begin{array}{ccc|cc} I + \mu e^y & \cdot & \cdot & P^* & \cdot \\ \cdot & I & \cdot & \cdot & \cdot \\ \cdot & \cdot & \frac{1}{\nu} I & \cdot & \cdot \\ \hline P & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & I \end{array} \right)^{-1}, \end{aligned}$$

where $Q = -R^*(I + \mu e^y)R$. Using V_1 or V_2 there are 2 parabolic equations to solve in each preconditioning step.

Numerical tests with a quite coarse discretization, enabling numerical evaluation of eigenvalues, demonstrated that V_1 as well as V_2 reduced the condition of the

linear system drastically, with V_2 slightly preferable compared to V_1 . Other, more sophisticated preconditioners, also based on blockwise construction, turned out to be less successfully. For details see [5].)

Although the matrix $M(x; \lambda)$ is not positive definite, the preconditioned CG method was quite efficient for solving (12), [5]. But the numerical solution can still be speeded up considerably by utilizing a preconditioned CG method with suitable preconditioners applied to the Schur complement formulation. Here the following system is to be solved:

$$\mathcal{O}_{SC}(x; \lambda) \partial \lambda = e(x) - (e'(x) \mathcal{L}''(x; \lambda)^{-1}) \mathcal{L}'(x; \lambda), \quad (13)$$

where the operator \mathcal{O}_{SC} is given by

$$\mathcal{O}_{SC}(x; \lambda) = e'(x) \mathcal{L}''(x; \lambda)^{-1} e'(x)^*.$$

Setting $\vartheta = 1 + \mu e^y$ and $\omega = \vartheta^{-1}$, then in the (5x5) case of distributed control we obtain

$$\mathcal{O}_{SC}(y, w, u; \lambda, \mu) = \begin{pmatrix} P\omega P^* + (\delta^2 + \frac{1}{\nu}) I & P\omega e^y + \delta I \\ e^y \omega P^* + \delta I & e^y \omega e^y + I \end{pmatrix}.$$

A preconditioner, which turned out to be very effective is given by

$$V_{SC} = \begin{pmatrix} R^*(e^y e^y + \vartheta) R & -R^* e^y \\ -e^y R & 1 \end{pmatrix} = \begin{pmatrix} P\omega P^* & P\omega e^y \\ e^y \omega P^* & e^y \omega e^y + 1 \end{pmatrix}^{-1}.$$

Using the CG method with this preconditioner solving the Schur complement system (13) was at least five times faster than solving the original system (12) with the CG method preconditioned by V_2 .

The concepts described here for the case of distributed control are easily transferred to the case of Neumann boundary control.

Note that the advantages of the Schur complement approach are due to the fact that it is possible to give a closed form expression for $L''(x; \lambda)^{-1}$.

The Schur complement approach reduces the dimension of the system (12). Another possibility is to utilize a reduced SQP method. If one investigates the linear operators arising from the reduced SQP method for the distributed control problem it turns out that for the (3x3) case the systems corresponding to the reduced SQP methods and the Schur complement method can be understood as differently preconditioned versions of the same equation and that the reduced SQP method for (5x5) leads to equations quite similar to the equations that are obtained from the reduced SQP method for the (3x3) formulation. (For the case of Neumann boundary control this does not hold.)

3 A numerical example

For a test was chosen $Q = (0, T) \times \Omega = (0, 1) \times (-1, +1)$. The parameters occurring in state equation and cost functional were $\delta = 1$, $\nu = 1$ and the desired state is $y_d = (e^{2.0t} - 1) \cos \frac{\pi}{2} x$. Starting data was

$$y_0 = 4t \cos \frac{\pi}{2} x, \quad u_0 = 0.9 \cdot (y_t - \Delta y - \delta e^y), \quad \lambda_0 = 0$$

and in the (5x5) case additionally

$$w_0 = e^{y_0}, \quad \mu_0 = 0.$$

In the (5x5) case only the linear constraint $y_t - \Delta y - \delta w - u = 0$ is penalized in Algorithm 2.

The partial differential equations were discretized using finite differences with $m = 60$ gridpoints in time and $n = 91$ gridpoints in space.

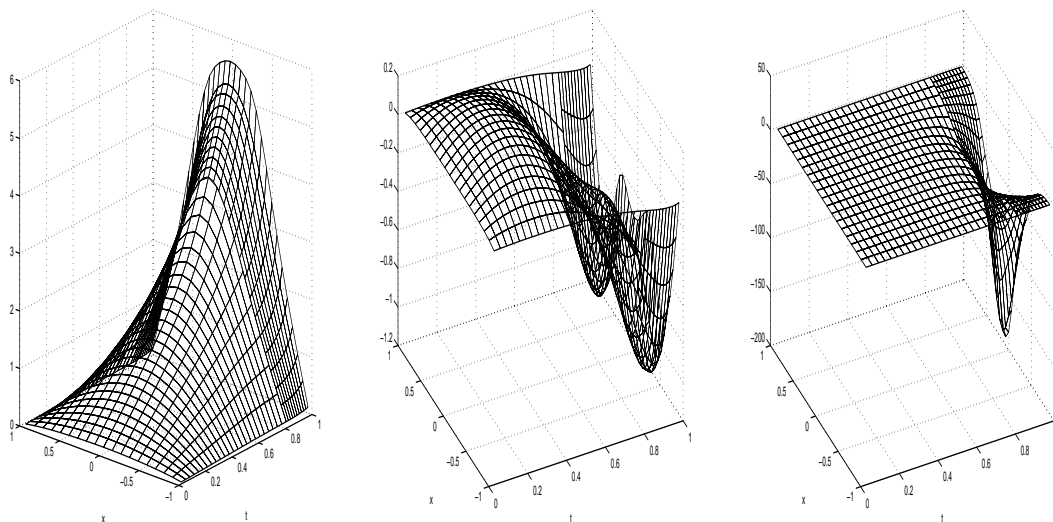


Figure 1: y^* , $y^* - y_d$ and $e^{y^*} - e^{y_d}$

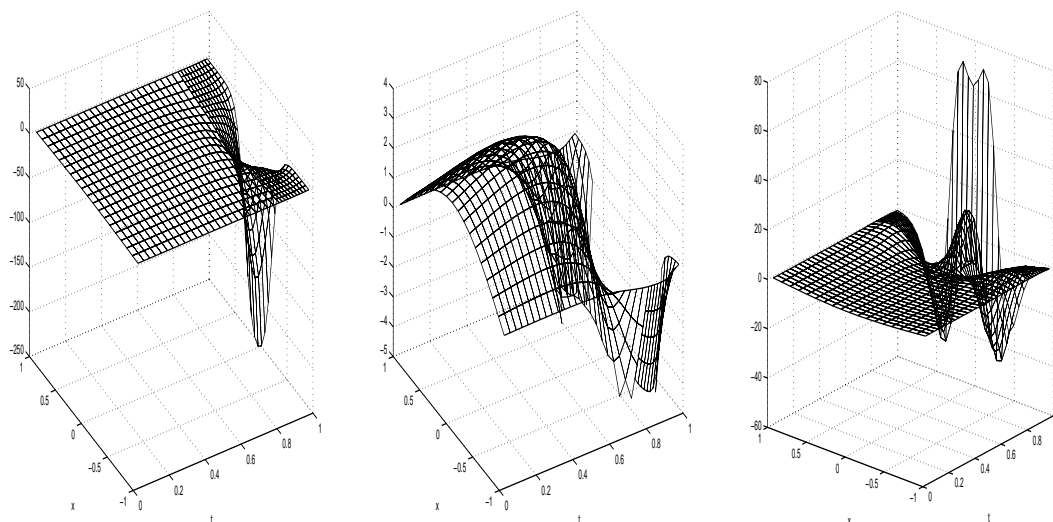


Figure 2: u^* , u_+ and μ^*

To document the test plots of the optimal solutions (y^*, u^*) are given in Figures 1 and 2. Also the Lagrange multiplier μ^* for the substitution constraint $w = e^y$ is given. There is no extra plot of the Lagrange multiplier λ^* , since $\lambda^* = u^*$ in the given setting. We observe that the absolute value of the control u^* increases very

strongly in time. To depict details of the control in the first time phase also the graph $u_+^* = \{(x, u^*(x)) : u^*(x) \geq -5\}$ is plotted for this test.

In the following table the number of augmented Lagrangian SQP steps necessary to achieve convergence is given. As stopping criterion was used equality of the first eight digits of $\mathcal{L}_c(x_{k+1}, \lambda_{k+1})$ and $\mathcal{L}_c(x_k, \lambda_k)$. Then ususally $\|\mathcal{L}'_c(x_k, \lambda_k)\|$ was somewhere between 10^{-10} and 10^{-7} .

If the line search rule slows down the convergence speed such that no more progress was visible the algorithm was aborted. These tests are marked by x_n .

c		0	1	2	4	8	12	100	400	1200
3x3	$\vartheta = 1$	∞	∞	13	14	∞				
	line search	∞	x_{15}	16	16	13	11	11	14	19
5x5	$\vartheta = 1$	14	14	13	14	14	14	14		14
	line search	x_{10}	10	8	8	6	6	5	5	9

x_n : Not really near the optimums after n SQP-steps

∞ : Failed to converge – numerical values out of range

Figure 3 results from the first 3 steps of Algorithm 2 applied to the test problem (5x5).

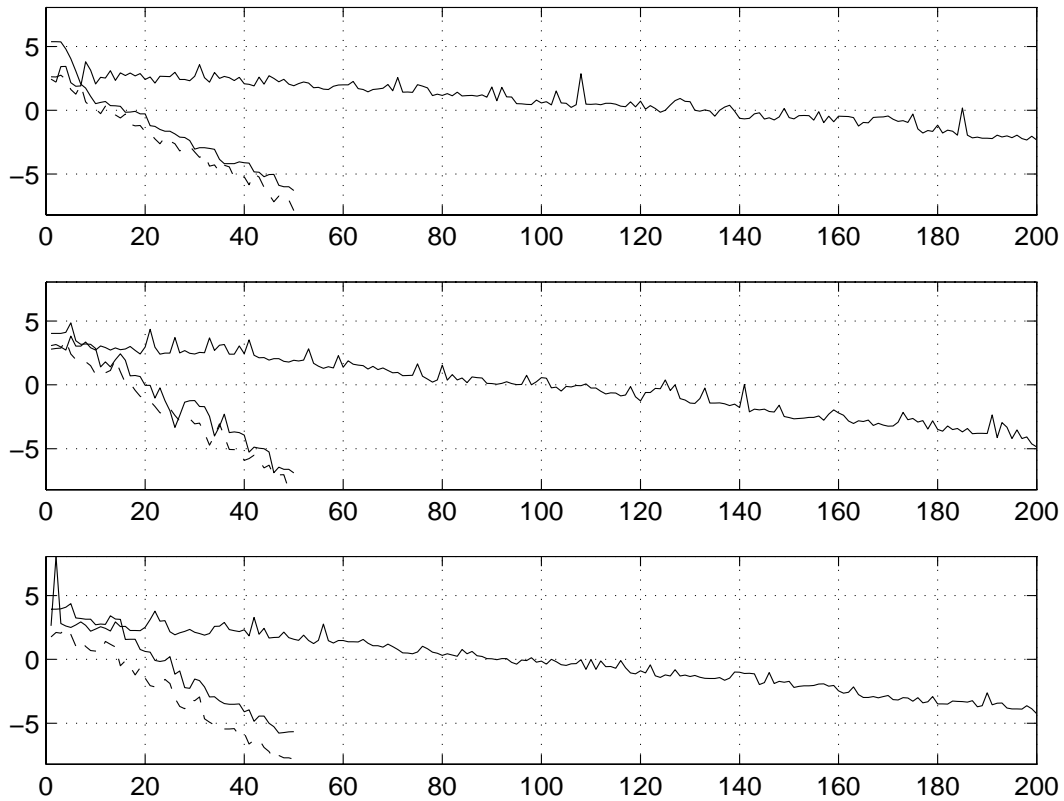


Figure 3: \log_{10} -norm of residua

In the plots we see:

- 200 iteration steps of preconditioned CG for solving the 5x5 system described in the previous subsection. CG was preconditioned with V_2 . The plot shows $\log_{10} |M\partial z^{(k)} + r|$ where $\partial z = (\partial x, \partial \lambda)^T$.
One 5x5 step of preconditioned CG costs about 725 000 flops.
- 50 iteration steps of preconditioned CG solving the Schur complement system. CG was preconditioned with V_{SC} . The plot shows $\log_{10} |S\partial \lambda^{(k)} - r_{SC}|$.
One Schur complement step of preconditioned CG costs about 650 000 flops.
- The dashed line shows $\log_{10} |M\partial z^{(k)} + r|$ where $\partial z^{(k)}$ was computed from $\partial \lambda^{(k)}$ using $\partial z = (-\mathcal{L}''(x; \lambda)^{-1}(\mathcal{L}'(x; \lambda) + e'(x)^* \partial \lambda), \partial \lambda)^T$.

The plots indicate, that the Schur complement systems will be solved about at least 5 times faster by of preconditioned CG than the 5x5 systems. In addition it requires less flops.

4 Feedback heuristics

In the case of distributed control the optimal control can be expressed in feedback form as

$$u = -\frac{1}{\nu} \mathcal{E}(y)^{-1} \mathcal{F}(y) \quad (14)$$

where

$$\mathcal{F}(y) = y - y_d + e^y(e^y - e^{y_d})$$

and $\mathcal{E}(y) : \{\varphi \in H^{2,1}(Q) : \varphi|_{\Sigma} = 0, \varphi|_{t=T} = 0\} \rightarrow L^2(Q)$ is given by

$$\mathcal{E}(y) = -\partial_t - \Delta - \delta e^y.$$

Of course, this representation is not numerically feasible but it can be used to conjecture stabilizing and implementable feedback laws. We discuss some of them. First, since y_d is not assumed to be a steady state of (1) we define

$$u_d = y_{d_t} - \Delta y_d - \delta e^{y_d}.$$

In the following result we address the feedback law

$$u = u_d - \frac{1}{\nu} (\delta e^y)^{-1} (y - y_d + e^y(e^y - e^{y_d})).$$

Theorem 3 *Suppose that $\delta > 0$, $0 < \nu \leq \delta^{-2}$, $2 \leq p, n < p < \infty$, and that y_d satisfies $y_d \in C^\infty([0, \infty) \times \bar{\Omega})$ with $y_d|_{\Sigma} = 0$. Then there exists a solution $y \in W_{loc}^{2,1;p}((0, \infty); \Omega)$ of*

$$y_t - \Delta y = \delta e^y + u_d - \frac{1}{\nu} (\delta e^y)^{-1} (y - y_d + e^y(e^y - e^{y_d})),$$

$$y|_{t=0} = y_0 \in W_0^{2;p}(\Omega), \quad y|_{\Sigma} = 0.$$

Moreover there exists a constant $C > 0$ such that

$$\|y(t) - y_d(t)\|_{L^2(\Omega)}^2 \leq e^{-Ct} \|y(0) - y_d(0)\|_{L^2(\Omega)}^2 \quad \text{for all } t \geq 0.$$

For boundary control the feedback solution analogous to (14) is given by

$$u = \frac{1}{\nu} \mathcal{G} \mathcal{E}(y)^{-1} \mathcal{F}(y), \quad (15)$$

where $\mathcal{G} = (I - \mathcal{D}_{tt} - \Delta_{\Gamma})^{-1} \circ \tau_{\Sigma}$, with $\tau_{\Sigma} : Q \rightarrow \Sigma$ the trace onto the lateral boundary Σ of Q . Attempts to replace (15) by feedback laws requiring only data of y along Σ did not result in control systems for which it was possible to avoid blow up in numerical experiments. We used a feedback law of the form

$$\frac{\partial y}{\partial n}|_{\Sigma} = u_d + u_c,$$

where $u_c = -\frac{1}{\nu} \tau_{\Gamma} \circ (-\Delta_M)^{-1} (y - y_d + e^y (e^y - e^{y_d}))$, and $-\Delta_M^{-1}$ denotes the solution operator to

$$\begin{cases} -\Delta \varphi = \psi \\ (\frac{\partial}{\partial n} \varphi + \kappa \varphi)|_{\Sigma} = 0, \kappa \neq 0. \end{cases}$$

For this choice of feedback control it is possible, numerically, to track the solution of (2) into the neighborhood of an unstable steady state y_d , see [5].

References

- [1] V. Barbu. Necessary conditions for distributed control problems governed by parabolic variational inequalities. *SIAM J. Control and Optimization*, 19(1):64–86, 1981.
- [2] J. Bebernes and D. Eberly. *Mathematical Problems from Combustion Theory*. Springer–Verlag, 1989.
- [3] K. Ito and K. Kunisch. Optimal control of the solid fuel ignition model with h^1 -cost. preprint.
- [4] K. Ito and K. Kunisch. Augmented Lagrangian–SQP–methods in hilbert spaces and application to control in the coefficients problem. *SIAM J. on Optimization*, 32(3):96–125, 1996.
- [5] A. Kauffmann. *Optimal Control of the Solid Fuel Ignition Model*. PhD thesis, Technische Universität Berlin, 1998.
- [6] D. Kleis and E.W. Sachs. Convergence rate of the Augmented Lagrangian SQP method. Forschungsbericht 95-19, Universität Trier, 1995.
- [7] J.L. Lions. *Control of Distributed Singular Systems*. Gauthier–Villars, 1985.
- [8] J.L. Lions and E. Magenes. *Problemes aux limites non homogenes et applications*, volume 2. Dunod, Paris 1968.