Adaptive Space-Time Finite Element Methods for Parabolic Optimization Problems

Dominik Meidner and Boris Vexler

Abstract. In this paper we summerize recent results on a posteriori error estimation and adaptivity for space-time finite element discretizations of parabolic optimization problems. The provided error estimates assess the discretization error with respect to a given quantity of interest and separate the influences of different parts of the discretization (time, space, and control discretization). This allows us to set up an efficient adaptive strategy producing economical (locally) refined meshes for each time step and an adapted time discretization. The space and time discretization errors are equilibrated, leading to an efficient method.

Mathematics Subject Classification (2000). 65N30, 49K20, 65M50, 35K55, 65N50.

Keywords. parabolic equations, optimal control, parameter identification, a posteriori error estimation, mesh refinement, space-time finite elements, dynamic meshes.

1. Introduction

In this paper, we discuss adaptive algorithms for the efficient solution of time-dependent optimization problems governed by parabolic partial differential equations.

The use of adaptive techniques based on a posteriori error estimation is well accepted in the context of finite element discretization of partial differential equations, see, e.g., [4, 12, 39]. The application of these techniques is also investigated for optimization problems governed by elliptic partial differential equations. Energy-type error estimators for the error in the state, control, and adjoint variable are developed, e.g., in [18, 21, 26, 27]—see also [29] and the references therein—in the context of distributed and boundary control problems subject to a elliptic state equation and pointwise control constraints.

For optimal control problems governed by linear parabolic equations a posteriori error estimates with respect to some norms are presented in [25, 28]. In [25],

the state equation is discretized by space-time finite elements, and the reliability of error estimates is shown under the H^2-regularity assumption on the domain. In [34], an anisotropic error estimate is derived for the error due to the space discretization of an optimal control problem governed by the linear heat equation. In [37], an adaptive algorithm for optimal control of time-dependent differential algebraic equations is presented. Space-time adaptivity within an SQP method is investigated in [10]. In [23], adaptive techniques are applied for boundary control of time-dependent equations describing radiative heat transfer in glass.

However, in many applications, one is interested in the precise computation of certain physical quantities ("quantities of interest") rather than in reducing the discretization error with respect to global norms. In [2, 4], a general concept for a posteriori estimation of the discretization error with respect to the cost functional in the context of stationary optimal control problems is presented. In the papers [5, 6], this approach was extended to the estimation of the discretization error with respect to an arbitrary functional depending on both, the control and the state variable, i.e., with respect to a quantity of interest. This allows—among other things—an efficient treatment of parameter identification and model calibration problems. Moreover, a posteriori error estimates for optimal control problems with pointwise inequality constraints were derived, see, e.g., [7, 17, 19, 41].

Here, we discuss the extension of this "quantity of interest" driven approach of error estimation to optimization problems governed by parabolic partial differential equations, see [30, 31]. The results and the numerical examples of [31] are extended to the case of dynamically changing meshes, see [36], which is essential for an efficient adaptive algorithm.

We consider optimization problems under constraints of (nonlinear) parabolic differential equations:

$$\begin{aligned} \partial_t u + A(q,u) &= f \\ u(0) &= u_0(q). \end{aligned} \tag{1.1}$$

Here, the state variable is denoted by u and the control variable by q. Both, the differential operator A and the initial condition u_0 may depend on q. This allows a simultaneous treatment of optimal control and parameter identification problems. For optimal control problems, the operator A is typically given by

$$A(q,u) = \bar{A}(u) - B(q),$$

with a (nonlinear) operator $\bar{A}$ and a (usually linear) control operator B. In parameter identification problems, the variable q denotes the unknown parameters to be determined and may enter the operator A in a nonlinear way. The case of initial control is included via the q-dependent initial condition $u_0(q)$. The target of the optimization is to minimize a given cost functional $J(q,u)$ subject to the state equation (1.1).

For the numerical solution of this optimization problem, the state variable has to be discretized in space and in time. Moreover, if the control (parameter) space is infinite dimensional, it is usually discretized too, cf. [17] for the approach

without discretizing the control in the context of elliptic problems. Both, time and space discretization of the state equation are based on the finite element method as proposed, e.g., in [13, 14]. In [3] we have shown that this type of discretization allows for a natural translation of the optimality conditions from the continuous to the discrete level. This gives rise to exact computation of the derivatives required in the optimization algorithms on the discrete level.

Our main contribution is the development of a posteriori error estimates which assess the error between the solutions of the continuous and the discrete optimization problem with respect to a given quantity of interest. This quantity of interest may coincide with the cost functional or expresses another goal of the computation. In order to set up an efficient adaptive algorithm, we separate the influences of the time and space discretizations on the error in the quantity of interest. This allows to balance different types of errors and to successively improve the accuracy by constructing locally refined meshes for time and space discretizations.

We introduce σ as a general discretization parameter including the space, time, and control discretization and denote the solution of the discrete problem by (q_σ, u_σ). For this discrete solution we derive an a posteriori error estimate with respect to the cost functional J of the following form:

$$J(q,u) - J(q_\sigma, u_\sigma) \approx \eta_k^J + \eta_h^J + \eta_d^J. \tag{1.2}$$

Here, η_k^J, η_h^J, and η_d^J denote the error estimators which can be evaluated from the computed discrete solution: η_k^J assess the error due to the time discretization, η_h^J due to the space discretization, and η_d^J due to the discretization of the control space. In the case of not discretizing the control space—cf. [17] for elliptic problems—the contribution of η_d^J vanishes. However, in many situations it is reasonable, to use coarser time grids for the discretization of the control variable than for the state variable, see, e. g., [30]. The structure of the error estimate (1.2) allows for equilibration of different discretization errors within an adaptive refinement algorithm to be described in the sequel.

The error estimator for the error with respect to the cost functional does in general not provide a bound for the norm of the error $\|q - q_\sigma\|$. This allows for very economical meshes—in terms of degrees of freedom—for approximating the optimal value $J(q,u)$. If an estimate for $\|q - q_\sigma\|$ is requested, our technique can be adapted using the (local) coercivity of the reduced cost functional, see [1] for this approach in the context of elliptic problems.

For many optimization problems the quantity of physical interest coincides with the cost functional which explains the choice of the error measure (1.2). However, in the case of parameter identification or model calibration problems, the cost functional is only an instrument for the estimation of the unknown parameters. Therefore, the value of the cost functional in the optimum and the corresponding discretization error are of secondary importance. This motivates error estimation with respect to a given functional I depending on the state and the control (parameter) variable. In this paper we discuss the extension of the corresponding results

from [5, 6, 40] to parabolic problems, see [31], and present an a posteriori error estimator of the form

$$I(q,u) - I(q_\sigma, u_\sigma) \approx \eta_k^I + \eta_h^I + \eta_d^I,$$

where again η_k^I and η_h^I estimate the temporal and spatial discretization errors and η_d^I estimates the discretization error due to the discretization of the control space.

In Section 5.2 we describe an adaptive algorithm based on these error estimators. Within this algorithm the time, space, and control discretizations are refined separately to achieve an efficient reduction of the total error by equilibrating different types of the error. This local refinement relies on the computable representation of the error estimators as a sum of local contributions (error indicators), see the discussion in Section 5.1.

The outline of the paper is as follows: In the next section we introduce the functional analytic setting of the considered optimization problem and describe necessary optimality conditions for the problem under consideration. In Section 3 we present the space time finite element discretization of the optimization problem. Section 4 is devoted to the derivation of the error estimators in a general setting. In Section 5 we discuss numerical evaluation of these error estimators and the adaptive algorithm in details. In the last section we present two numerical examples illustrating the behavior of the proposed methods. The first example deals with the optimal control of surface hardening of steel, whereas the second one is concerned with the identification of Arrhenius parameters in a simplified gaseous combustion model.

2. Optimization problems

The optimization problems considered in this paper are formulated in the following abstract setting: Let Q be a Hilbert space for the controls (parameters) with inner product $(\cdot,\cdot)_Q$. Moreover, let V and H be Hilbert spaces, which build together with the dual space V^* of V a Gelfand triple $V \hookrightarrow H \hookrightarrow V^*$. The duality pairing between the Hilbert spaces V and its dual V^* is denoted by $\langle\cdot,\cdot\rangle_{V^*\times V}$ and the inner product in H by $(\cdot,\cdot)$. A typical choice for these space could be

$$V = \left\{ v \in H^1(\Omega) \;\middle|\; v\big|_{\partial\Omega_{\mathrm{D}}} = 0 \right\} \text{ and } H = L^2(\Omega), \tag{2.1}$$

where $\partial\Omega_{\mathrm{D}}$ denotes the part of the boundary of the computational domain Ω with prescribed Dirichlet boundary conditions.

For a time interval $(0,T)$ we introduce the Hilbert space $X := W(0,T)$ defined as

$$W(0,T) = \left\{ v \;\middle|\; v \in L^2((0,T),V) \text{ and } \partial_t v \in L^2((0,T),V^*) \right\}. \tag{2.2}$$

It is well known that the space X is continuously embedded in $C([0,T],H)$, see, e.g., [11]. Furthermore, we use the inner product of $L^2((0,T),H)$ given by

$$(u,v)_I := (u,v)_{L^2((0,T),H)} = \int_0^T (u(t),v(t))\,dt \tag{2.3}$$

for setting up the weak formulation of the state equation.

By means of the spatial semi-linear form $\bar{a}\colon Q \times V \times V \to \mathbb{R}$ defined for a differential operator $A\colon Q \times V \to V^*$ by

$$\bar{a}(q,\bar{u})(\bar{\varphi}) := \langle A(q,\bar{u}),\bar{\varphi}\rangle_{V^*\times V},$$

we can define the semi-linear form $a(\cdot,\cdot)(\cdot)$ on $Q \times L^2(I,V) \times L^2(I,V)$ as

$$a(q,u)(\varphi) := \int_0^T \bar{a}(q,u(t))(\varphi(t))\,dt$$

which is assumed to be linear in the third argument and three times directional differentiable with derivatives which are linear in the direction. This would be implied if the semi-linear form is Gâteaux differentiable. However, in our analysis, we do not require the boundedness of the derivatives.

Remark 2.1. If the control variable q depends on time, this has to be incorporated by an obvious modification of the definitions of the semi-linear forms.

After these preliminaries, we pose the *state equation* in a weak form: Find for given control $q \in Q$ the *state* $u \in X$ such that

$$\begin{aligned} (\partial_t u,\varphi)_I + a(q,u)(\varphi) &= (f,\varphi)_I \quad \forall \varphi \in X,\\ u(0) &= u_0(q), \end{aligned} \tag{2.4}$$

where $f \in L^2((0,T),V^*)$ represents the right hand side of the state equation and $u_0\colon Q \to H$ denotes a three times directional differentiable mapping with derivatives which are linear in the direction describing parameter-dependent initial conditions.

Remark 2.2. There are several sets of assumptions on the nonlinearity in $\bar{a}(\cdot,\cdot)(\cdot)$ and its dependence on the control variable q allowing the state equation (2.4) to be well-posed. Typical examples are different semilinear equations, where the form $\bar{a}(\cdot,\cdot)(\cdot)$ consists of a linear elliptic part and a nonlinear term depending on u and ∇u. Due to the fact that the development of the proposed adaptive algorithm does not depend on the particular structure of the nonlinearity in $\bar{a}$, we do not specify a set of assumptions on it, but assume that the state equation (2.4) possess a unique solution $u \in X$ for each $q \in Q$.

The cost functional $J\colon Q \times \{\, v \in L^2(I,V) \mid v(T) \in H \,\} \to \mathbb{R}$ is defined using two three times direction differentiable functionals $J_1\colon V \to \mathbb{R}$ and $J_2\colon H \to \mathbb{R}$ with derivatives which are linear in the direction by

$$J(q,u) = \int_0^T J_1(u)\,dt + J_2(u(T)) + \frac{\alpha}{2}\|q-\bar{q}\|_Q^2, \tag{2.5}$$

where the regularization (or cost) term is added which involves $\alpha \geq 0$ and a reference control (parameter) $\bar{q} \in Q$. We note, that due to the embedding $X \hookrightarrow C(\bar{I}, H)$ the functional J is well-defined for arguments $u \in X$.

The corresponding optimization problem is formulated as follows:

$$\text{Minimize } J(q,u) \text{ subject to the state equation (2.4), } (q,u) \in Q \times X. \tag{2.6}$$

The question of existence and uniqueness of solutions to such optimization problems is discussed, e.g., in [24, 16, 38]. Throughout the paper, we assume problem (2.6) to admit a (locally) unique solution. Moreover, we assume the existence of a neighborhood $W \subset Q \times X$ of the optimal solution, such that the linearized form $\bar{a}'_u(q, u(t))(\cdot,\cdot)$ considered as a linear operator

$$\bar{a}'_u(q, u(t))\colon V \to V^*$$

is an isomorphism for all $(q,u) \in W$ and almost all $t \in (0,T)$. This assumption will allow all considered linearized and adjoint problems to be well posed.

To set up the optimality system we introduce the Lagrangian $\mathcal{L}\colon Q \times X \times X \to \mathbb{R}$, defined as

$$\mathcal{L}(q,u,z) = J(q,u) + (f - \partial_t u, z)_I - a(q,u)(z) - (u(0) - u_0(q), z(0)). \tag{2.7}$$

By definition, $\mathcal{L}$ inherits the differentiability properties from those of a, u_0, and J.

The optimality system of the considered optimization problem (2.6) is given by the derivatives of the Lagrangian:

$$\begin{aligned}
&\mathcal{L}'_z(q,u,z)(\varphi) = 0, \quad \forall \varphi \in X \qquad \text{(State equation)},\\
&\mathcal{L}'_u(q,u,z)(\varphi) = 0, \quad \forall \varphi \in X \qquad \text{(Adjoint state equation)},\\
&\mathcal{L}'_q(q,u,z)(\psi) = 0, \quad \forall \psi \in Q \qquad \text{(Gradient equation)}.
\end{aligned} \tag{2.8}$$

For the explicit formulation of the dual equation in this setting see, e.g., [3].

3. Discretization

In this section, we discuss the discretization of the optimization problem (2.6). To this end, we use Galerkin finite element methods in space and time to discretize the state equation. Our systematic approach to a posteriori error estimation relies on using such Galerkin-type discretizations.

The first of the following subsections is devoted to the semi-discretization in time by *discontinuous Galerkin (dG)* methods. Subsection 3.2 deals with the space discretization of the semi-discrete problems arising from time discretization. For the numerical analysis of these schemes we refer to [13]. In the context of optimal control problems, a priori error estimates for this type of discretization are derived in [32, 33].

The discretization of the control space Q is kept rather abstract by choosing a finite dimensional subspace $Q_d \subset Q$. A possible concretion of this choice is shown in the numerical examples in Section 6.

3.1. Time discretization of the states

To define a semi-discretization in time, let us partition the time interval $[0,T]$ as

$$[0,T] = \{0\} \cup I_1 \cup I_2 \cup \cdots \cup I_M$$

with subintervals $I_m = (t_{m-1}, t_m]$ of size k_m and time points

$$0 = t_0 < t_1 < \cdots < t_{M-1} < t_M = T.$$

We define the discretization parameter k as a piecewise constant function by setting $k\big|_{I_m} = k_m$ for $m = 1, \ldots, M$.

By means of the subintervals I_m, we define for $r \in \mathbb{N}_0$ the semi-discrete space X_k^r by

$$X_k^r = \left\{ v_k \in L^2((0,T),V) \;\middle|\; v_k\big|_{I_m} \in \mathcal{P}^r(I_m,V) \text{ and } v_k(0) \in H \right\}$$

Here, $\mathcal{P}^r(I_m,V)$ denotes the space of polynomials up to order r defined on I_m with values in V. Thus, the functions in X_k^r may have discontinuities at the edges of the subintervals I_m. This space is used in the sequel as trial and test space in the discontinuous Galerkin method. For functions in X_k^r we use the notation

$$(v,w)_{I_m} \coloneqq (v,w)_{L^2(I_m,L^2(\Omega))}.$$

To define the dG(r) discretization we employ the following definition for functions $v_k \in X_k^r$:

$$v_{k,m}^+ := \lim_{t\to 0^+} v_k(t_m+t), \quad v_{k,m}^- := \lim_{t\to 0^+} v_k(t_m-t) = v_k(t_m), \quad [v_k]_m := v_{k,m}^+ - v_{k,m}^-.$$

Then, the dG(r) semi-discretization of the state equation (2.4) reads: Find for given control $q_k \in Q$ a state $u_k \in X_k^r$ such that

$$\begin{aligned} \sum_{m=1}^{M} (\partial_t u_k, \varphi)_{I_m}\, dt + a(q_k,u_k)(\varphi) + \sum_{m=0}^{M-1} ([u_k]_m, \varphi_m^+) = (f,\varphi)_I,\ \forall \varphi \in X_k^r, \\ u_{k,0}^- = u_0(q_k). \end{aligned} \tag{3.1}$$

Remark 3.1. This equation is assumed to posses a unique solution for each $q_k \in Q$, cf. Remark 2.2. In special cases, the existence and uniqueness can be shown by separation of variables and using the fact that X_k^r is finite dimensional with respect to time.

The semi-discrete optimization problem for the dG(r) time discretization has the form:

$$\text{Minimize } J(q_k,u_k) \text{ subject to the state equation (3.1)},\ (q_k,u_k) \in Q \times X_k^r. \tag{3.2}$$

Then, we pose the Lagrangian $\widetilde{\mathcal{L}}\colon Q \times X_k^r \times X_k^r \to \mathbb{R}$ associated with the dG(r) time discretization for the state equation as

$$\begin{aligned}\widetilde{\mathcal{L}}(q_k, u_k, z_k) = J(q_k, u_k) + (f, z_k)_I - \sum_{m=1}^{M} (\partial_t u_k, z_k)_{I_m}\, dt \\ - a(q_k, u_k)(z_k) - \sum_{m=0}^{M-1} ([u_k]_m, z_{k,m}^+) - (u_{k,0}^- - u_0(q_k), z_{k,0}^-).\end{aligned}$$

As on the continuous level for $\mathcal{L}$, $\widetilde{\mathcal{L}}$ inherits here the differentiability properties from those of a, u_0, and J.

3.2. Space discretization of the states

In this subsection, we first describe the finite element discretization in space. To this end, we consider two or three dimensional shape-regular meshes, see, e.g., [9]. A mesh consists of quadrilateral or hexahedral cells K which constitute a non-overlapping cover of the computational domain $\Omega \subset \mathbb{R}^n$, $n \in \{2, 3\}$. The corresponding mesh is denoted by $\mathcal{T}_h = \{K\}$, where we define the discretization parameter h as a cellwise constant function by setting $h\big|_K = h_K$ with the diameter h_K of the cell K.

We allow dynamic mesh changes in time whereas the time steps k_m are kept constant in space. Therefore, we associate with each time point t_m a triangulation $\mathcal{T}_h^m$ and a corresponding finite element space $V_h^{s,m} \subseteq V$ which is used as spatial trial and test space in the adjacent time interval I_m.

The finite element space $V_h^{s,m} \subset V$ is constructed in a standard way by

$$V_h^{s,m} = \left\{ v \in V \;\middle|\; v\big|_K \in \mathcal{Q}^s(K) \text{ for } K \in \mathcal{T}_h^m \right\},$$

where $\mathcal{Q}^s(K)$ consists of shape functions obtained via bi- or tri-linear transformations of polynomials in $\widehat{\mathcal{Q}^s}(\widehat{K})$ defined on the reference cell $\widehat{K} = (0,1)^n$, see, e.g., [9] for details.

To obtain the fully discretized versions of the time discretized state equation (3.1), we utilize the space-time finite element space

$$X_{k,h}^{r,s} = \left\{ v_{kh} \in L^2((0.T), H) \;\middle|\; v_{kh}\big|_{I_m} \in \mathcal{P}^r(I_m, V_h^{s,m}) \text{ and } v_{kh}(0) \in V_h^{s,0} \right\}.$$

Then, the cG(s)dG(r) discretization has the form: Find for given control $q_{kh} \in Q$ a state $u_{kh} \in X_{k,h}^{r,s}$ such that

$$\begin{aligned}\sum_{m=1}^{M} (\partial_t u_{kh}, \varphi)_{I_m}\, dt + a(q_{kh}, u_{kh})(\varphi) + \sum_{m=0}^{M-1} ([u_{kh}]_m, \varphi_m^+) + (u_{kh,0}^-, \varphi_0^-) \\ = (f, \varphi)_I + (u_0(q_{kh}), \varphi_0^-) \quad \forall \varphi \in X_{k,h}^{r,s}. \quad (3.3)\end{aligned}$$

These fully discretized state equations are assumed to posses unique solutions for each $q_{kh} \in Q$, see Remark 3.1.

Thus, the optimization problem with fully discretized state is given by

$$\text{Minimize } J(q_{kh}, u_{kh}) \text{ subject to the state equation (3.3)} \quad (q_{kh}, u_{kh}) \in Q \times X_{k,h}^{r,s}. \quad (3.4)$$

The definition of the Lagrangian $\widetilde{\mathcal{L}}$ for fully discretized states can directly be transferred from the formulations for semi-discretization in time just by restriction of the state space X_k^r to the subspaces $X_{k,h}^{r,s}$, respectively.

3.3. Discretization of the controls

As proposed in the beginning of the current section, the discretization of the control space Q is kept rather abstract. It is done by choosing a finite dimensional subspace $Q_d \subset Q$. Then, the formulation of the state equation, the optimization problems and the Lagrangians defined on the fully discretized state space can directly be transferred to the level with fully discretized control and state spaces by replacing Q by Q_d. The full discrete solutions will be indicated by the subscript σ which collects the discretization indices k, h, and d.

4. A posteriori error estimators

In this section, we establish a posteriori error estimators for the error arising due to the discretization of the control and state spaces in terms of the cost functional J and in terms of an arbitrary quantity of interest I.

4.1. Error estimator for the cost functional

In the sequel, we derive error estimators for the discretization error with respect to the cost functional J:

$$J(q, u) - J(q_\sigma, u_\sigma)$$

Here, $(q, u) \in Q \times X$ denotes the continuous optimal solution of (2.6) and $(q_\sigma, u_\sigma) = (q_{khd}, u_{khd}) \in Q_d \times X_{k,h}^{r,s}$ is the optimal solution of the fully discretized problem.

To separate the influences of the different discretizations on the discretization error, we split

$$\begin{aligned} J(q, u) - J(q_\sigma, u_\sigma) = \quad & J(q, u) - J(q_k, u_k) \\ & + J(q_k, u_k) - J(q_{kh}, u_{kh}) \\ & + J(q_{kh}, u_{kh}) - J(q_\sigma, u_\sigma), \end{aligned}$$

where $(q_k, u_k) \in Q \times X_k^r$ is the solution of the time discretized problem (3.2) and $(q_{kh}, u_{kh}) \in Q \times X_{k,h}^{r,s}$ is the solution of the time and space discretized problem (3.4) with still undiscretized control space Q.

Theorem 4.1. *Let $\xi = (q, u, z)$, $\xi_k = (q_k, u_k, z_k)$, $\xi_{kh} = (q_{kh}, u_{kh}, z_{kh})$, and $\xi_\sigma = (q_\sigma, u_\sigma, z_\sigma)$ be stationary points of $\mathcal{L}$ resp. $\widetilde{\mathcal{L}}$ on the different levels of discretization,*

i.e.,

$$\begin{aligned}
\mathcal{L}'(\xi)(\hat{q},\hat{u},\hat{z}) = \widetilde{\mathcal{L}}'(\xi)(\hat{q},\hat{u},\hat{z}) = 0, &\quad \forall (\hat{q},\hat{u},\hat{z}) \in Q \times X \times X,\\
\widetilde{\mathcal{L}}'(\xi_k)(\hat{q}_k,\hat{u}_k,\hat{z}_k) = 0, &\quad \forall (\hat{q}_k,\hat{u}_k,\hat{z}_k) \in Q \times X_k^r \times X_k^r,\\
\widetilde{\mathcal{L}}'(\xi_{kh})(\hat{q}_{kh},\hat{u}_{kh},\hat{z}_{kh}) = 0, &\quad \forall (\hat{q}_{kh},\hat{u}_{kh},\hat{z}_{kh}) \in Q \times X_{k,h}^{r,s}, \times X_{k,h}^{r,s}\\
\widetilde{\mathcal{L}}'(\xi_\sigma)(\hat{q}_\sigma,\hat{u}_\sigma,\hat{z}_\sigma) = 0, &\quad \forall (\hat{q}_\sigma,\hat{u}_\sigma,\hat{z}_\sigma) \in Q_d \times X_{k,h}^{r,s} \times X_{k,h}^{r,s}.
\end{aligned}$$

Then, there holds for the errors with respect to the cost functional due to the time, space, and control discretizations:

$$\begin{aligned}
J(q,u) - J(q_k,u_k) &= \frac{1}{2}\widetilde{\mathcal{L}}'(\xi_k)(q-\hat{q}_k, u-\hat{u}_k, z-\hat{z}_k) + \mathcal{R}_k\\
J(q_k,u_k) - J(q_{kh},u_{kh}) &= \frac{1}{2}\widetilde{\mathcal{L}}'(\xi_{kh})(q_k-\hat{q}_{kh}, u_k-\hat{u}_{kh}, z_k-\hat{z}_{kh}) + \mathcal{R}_h\\
J(q_{kh},u_{kh}) - J(q_\sigma,u_\sigma) &= \frac{1}{2}\widetilde{\mathcal{L}}'(\xi_\sigma)(q_{kh}-\hat{q}_\sigma, u_{kh}-\hat{u}_\sigma, z_{kh}-\hat{z}_\sigma) + \mathcal{R}_d.
\end{aligned}$$

Here, $(\hat{q}_k,\hat{u}_k,\hat{z}_k) \in Q\times X_k^r\times X_k^r$, $(\hat{q}_{kh},\hat{u}_{kh},\hat{z}_{kh}) \in Q\times X_{k,h}^{r,s}\times X_{k,h}^{r,s}$, *and* $(\hat{q}_\sigma,\hat{u}_\sigma,\hat{z}_\sigma) \in Q_d \times X_{k,h}^{r,s} \times X_{k,h}^{r,s}$ *can be chosen arbitrary. The remainder terms* $\mathcal{R}_k$, $\mathcal{R}_h$, *and* $\mathcal{R}_d$ *are of third order in the error, see* [31]. $\mathcal{R}_k$, *for instance, is given for* $e_k := \xi - \xi_k$ *by*

$$\frac{1}{2}\int_0^1 \widetilde{\mathcal{L}}'''(\xi_k + se)(e,e,e)\cdot s\cdot(1-s)\,ds.$$

Remark 4.2. Due to the structure of $\widetilde{\mathcal{L}}$, the remainder terms $\mathcal{R}_k$, $\mathcal{R}_h$, and $\mathcal{R}_d$ depend only on third derivatives of the semi-linear form a and the cost functional J. If the cost functional is of tracking type, then J''' vanishes and the remainder term depend only on a'' and a'''. For given applications, one can explicitly calculate the concrete form of the remainder terms.

By means of the residuals of the three equations building the optimality system (2.8)

$$\begin{aligned}
\rho^u(q,u)(\varphi) := \widetilde{\mathcal{L}}'_z(\xi)(\varphi), \quad \rho^z(q,u,z)(\varphi) := \widetilde{\mathcal{L}}'_u(\xi)(\varphi),\\
\rho^q(q,u,z)(\varphi) := \widetilde{\mathcal{L}}'_q(\xi)(\varphi),
\end{aligned}$$

the statement of Theorem 4.1 can be rewritten as

$$J(q,u) - J(q_k,u_k) \approx \frac{1}{2}\Big(\rho^u(q_k,u_k)(z-\hat{z}_k) + \rho^z(q_k,u_k,z_k)(u-\hat{u}_k)\Big) \tag{4.1a}$$

$$\begin{aligned}
J(q_k,u_k) - J(q_{kh},u_{kh}) \approx\\
\frac{1}{2}\Big(\rho^u(q_{kh},u_{kh})(z_k-\hat{z}_{kh}) + \rho^z(q_{kh},u_{kh},z_{kh})(u_k-\hat{u}_{kh})\Big)
\end{aligned} \tag{4.1b}$$

$$J(q_{kh},u_{kh}) - J(q_\sigma,u_\sigma) \approx \frac{1}{2}\rho^q(q_\sigma,u_\sigma,z_\sigma)(q_{kh}-\hat{q}_\sigma). \tag{4.1c}$$

Here, we employed the fact, that the terms

$$\rho^q(q_k,u_k,z_k)(q-\hat{q}_k), \qquad \rho^q(q_{kh},u_{kh},z_{kh})(q_k-\hat{q}_{kh}),$$
$$\rho^u(q_\sigma,u_\sigma)(z_{kh}-\hat{z}_\sigma), \qquad \rho^z(q_\sigma,u_\sigma,z_\sigma)(u_{kh}-\hat{u}_\sigma)$$

are zero for the choice

$$\hat{q}_k = q \in Q, \qquad \hat{q}_{kh} = q_k \in Q,$$
$$\hat{z}_\sigma = z_{kh} \in X_{k,h}^{r,s}, \qquad \hat{u}_\sigma = u_{kh} \in X_{k,h}^{r,s}.$$

This is possible since for the errors $J(q,u)-J(q_k,u_k)$ and $J(q_k,u_k)-J(q_{kh},u_{kh})$ only the state space is discretized and for $J(q_{kh},u_{kh})-J(q_\sigma,u_\sigma)$ we keep the discrete state space while discretizing the control space Q.

4.2. Error estimator for an arbitrary functional

We now tend toward an error estimation of the different types of discretization errors in terms of a given functional $I\colon Q\times X\to\mathbb{R}$ describing the quantity of interest. This will be done using solutions of some auxiliary problems. In order to ensure the solvability of these problems, we assume that the semi-discrete and the fully discrete optimal solutions (q_k,u_k), (q_{kh},u_{kh}), and (q_σ,u_σ) are in the neighborhood $W\subset Q\times X$ of the optimal solution (q,u) introduced in Section 2.

We define exterior Lagrangians $\mathcal{M}\colon [Q\times X\times X]^2\to\mathbb{R}$ and $\widetilde{\mathcal{M}}\colon [Q\times X_k^r\times X_k^r]^2\to\mathbb{R}$ as

$$\mathcal{M}(\xi,\chi) = I(q,u)+\mathcal{L}'(\xi)(\chi)$$

with $\xi=(q,u,z),\chi=(p,v,y)$ and

$$\widetilde{\mathcal{M}}(\xi_k,\chi_k) = I(q_k,u_k)+\widetilde{\mathcal{L}}'(\xi_k)(\chi_k)$$

with $\xi_k=(q_k,u_k,z_k),\chi_k=(p_k,v_k,y_k)$.

Now we are in a similar setting as in the subsection before: We split the total discretization error with respect to I as

$$\begin{aligned} I(q,u)-I(q_\sigma,u_\sigma) = \quad & I(q,u)-I(q_k,u_k)\\ &+I(q_k,u_k)-I(q_{kh},u_{kh})\\ &+I(q_{kh},u_{kh})-I(q_\sigma,u_\sigma)\end{aligned}$$

and obtain the following theorem:

Theorem 4.3. *Let (ξ,χ), (ξ_k,χ_k), (ξ_{kh},χ_{kh}), and (ξ_σ,χ_σ) be stationary points of $\mathcal{M}$ resp. $\widetilde{\mathcal{M}}$ on the different levels of discretization, i.e.*

$$\begin{aligned}\mathcal{M}'(\xi,\chi)(\hat{\xi},\hat{\chi}) = \widetilde{\mathcal{M}}'(\xi,\chi)(\hat{\xi},\hat{\chi}) = 0, &\quad \forall(\hat{\xi},\hat{\chi})\in[Q\times X\times X]^2,\\ \widetilde{\mathcal{M}}'(\xi_k,\chi_k)(\hat{\xi}_k,\hat{\chi}_k) = 0, &\quad \forall(\hat{\xi}_k,\hat{\chi}_k)\in[Q\times X_k^r\times X_k^r]^2,\\ \widetilde{\mathcal{M}}'(\xi_{kh},\chi_{kh})(\hat{\xi}_{kh},\hat{\chi}_{kh}) = 0, &\quad \forall(\hat{\xi}_{kh},\hat{\chi}_{kh})\in[Q\times X_{k,h}^{r,s}\times X_{k,h}^{r,s}]^2,\\ \widetilde{\mathcal{M}}'(\xi_\sigma,\chi_\sigma)(\hat{\xi}_\sigma,\hat{\chi}_\sigma) = 0, &\quad \forall(\hat{\xi}_\sigma,\hat{\chi}_\sigma)\in[Q_d\times X_{k,h}^{r,s}\times X_{k,h}^{r,s}]^2.\end{aligned}$$

Then, there holds for the errors with respect to the quantity of interest due to the time, space, and control discretizations:

$$\begin{aligned} I(q,u) - I(q_k,u_k) &= \frac{1}{2}\widetilde{\mathcal{M}}'(\xi_k,\chi_k)(\xi-\hat{\xi}_k,\chi-\hat{\chi}_k) + \mathcal{R}_k, \\ I(q_k,u_k) - I(q_{kh},u_{kh}) &= \frac{1}{2}\widetilde{\mathcal{M}}'(\xi_{kh},\chi_{kh})(\xi_k-\hat{\xi}_{kh},\chi_k-\hat{\chi}_{kh}) + \mathcal{R}_h, \\ I(q_{kh},u_{kh}) - I(q_\sigma,u_\sigma) &= \frac{1}{2}\widetilde{\mathcal{M}}'(\xi_\sigma,\chi_\sigma)(\xi_{kh}-\hat{\xi}_\sigma,\chi_{kh}-\hat{\chi}_\sigma) + \mathcal{R}_d. \end{aligned}$$

Here, $(\hat{\xi}_k,\hat{\chi}_k) \in [Q \times X_k^r \times X_k^r]^2$, $(\hat{\xi}_{kh},\hat{\chi}_{kh}) \in [Q \times X_{k,h}^{r,s} \times X_{k,h}^{r,s}]^2$, *and* $(\hat{\xi}_\sigma,\hat{\chi}_\sigma) \in [Q_d \times X_{k,h}^{r,s} \times X_{k,h}^{r,s}]^2$ *can be chosen arbitrary and the remainder terms* $\mathcal{R}_k$, $\mathcal{R}_h$, *and* $\mathcal{R}_d$ *are of third order in the error, see* [31].

To apply Theorem 4.3 for instance to $I(q_{kh},u_{kh}) - I(q_\sigma,u_\sigma)$, we have to require that

$$\widetilde{\mathcal{M}}'(\xi_\sigma,\chi_\sigma)(\hat{\xi}_\sigma,\hat{\chi}_\sigma) = 0, \quad \forall (\hat{\xi}_\sigma,\hat{\chi}_\sigma) \in [X_{k,h}^{r,s} \times X_{k,h}^{r,s} \times Q_d]^2.$$

For solving this system, we have to consider the concrete form of $\widetilde{\mathcal{M}}'$:

$$\begin{aligned} &\widetilde{\mathcal{M}}'(\xi_\sigma,\chi_\sigma)(\delta\xi_\sigma,\delta\chi_\sigma) = \\ &\qquad I_q'(q_\sigma,u_\sigma)(\delta q_\sigma) + I_u'(q_\sigma,u_\sigma)(\delta u_\sigma) + \widetilde{\mathcal{L}}'(\xi_\sigma)(\delta\chi_\sigma) + \widetilde{\mathcal{L}}''(\xi_\sigma)(\chi_\sigma,\delta\xi_\sigma) \end{aligned}$$

Since (q_σ,u_σ) is the solution of the discrete optimization problem, $\xi_\sigma = (q_\sigma,u_\sigma,z_\sigma)$ is a stationary point of the Lagrangian $\widetilde{\mathcal{L}}$ and consequently, it fulfills $\widetilde{\mathcal{L}}'(\xi_\sigma)(\delta\chi_\sigma) = 0$. Thus, the solution triple $\chi_\sigma = (p_\sigma,v_\sigma,y_\sigma) \in Q_d \times X_{k,h}^{r,s} \times X_{k,h}^{r,s}$ has to fulfill

$$\begin{aligned} &\widetilde{\mathcal{L}}''(\xi_\sigma)(\chi_\sigma,\delta\xi_\sigma) = \\ &\qquad - I_q'(q_\sigma,u_\sigma)(\delta q_\sigma) - I_u'(q_\sigma,u_\sigma)(\delta u_\sigma), \quad \forall \delta\xi_\sigma \in Q_d \times X_{k,h}^{r,s} \times X_{k,h}^{r,s}. \end{aligned} \tag{4.2}$$

Solving this system of equations is apart from a different right-hand side equivalent to the execution of one step of a (reduced) SQP-type method, see [31] for details.

By means of the residuals of the equations for p, v and y, i.e.,

$$\begin{aligned} \rho^v(\xi,p,v)(\varphi) &:= \widetilde{\mathcal{L}}_{uz}''(\xi)(v,\varphi) + \widetilde{\mathcal{L}}_{qz}''(\xi)(p,\varphi) \\ \rho^y(\xi,p,v,y)(\varphi) &:= \widetilde{\mathcal{L}}_{zu}''(\xi)(y,\varphi) + \widetilde{\mathcal{L}}_{qu}''(\xi)(p,\varphi) + \widetilde{\mathcal{L}}_{uu}''(\xi)(v,\varphi) + I_u'(q,u)(\varphi) \\ \rho^p(\xi,p,v,y)(\varphi) &:= \widetilde{\mathcal{L}}_{qq}''(\xi)(p,\varphi) + \widetilde{\mathcal{L}}_{uq}''(\xi)(v,\varphi) + \widetilde{\mathcal{L}}_{zq}''(\xi)(y,\varphi) + I_q'(q,u)(\varphi), \end{aligned}$$

and the already defined residuals ρ^u, ρ^z, and ρ^q the result of Theorem 4.3 can be expressed as

$$\begin{aligned} I(q,u) - I(q_k,u_k) \approx \frac{1}{2}\Big(&\rho^u(q_k,u_k)(y-\hat{y}_k) + \rho^z(q_k,u_k,z_k)(v-\hat{v}_k) \\ &+ \rho^v(\xi_k,p_k,v_k)(z-\hat{z}_k) + \rho^y(\xi_k,p_k,v_k,y_k)(u-\hat{u}_k)\Big) \end{aligned}$$

$$I(q_k,u_k)-I(q_{kh},u_{kh}) \approx \frac{1}{2}\Big(\rho^u(q_{kh},u_{kh})(y_k-\hat{y}_{kh})+\rho^z(q_{kh},u_{kh},z_{kh})(v_k-\hat{v}_{kh}) + \rho^v(\xi_{kh},p_{kh},v_{kh})(z_k-\hat{z}_{kh})+\rho^y(\xi_{kh},p_{kh},v_{kh},y_{kh})(u_k-\hat{u}_{kh})\Big)$$

$$I(q_{kh},u_{kh})-I(q_\sigma,u_\sigma) \approx \frac{1}{2}\Big(\rho^q(q_\sigma,u_\sigma,z_\sigma)(p_{kh}-\hat{p}_\sigma)+\rho^p(\xi_\sigma,p_\sigma,v_\sigma,y_\sigma)(q_{kh}-\hat{q}_\sigma)\Big).$$

As for the estimator for the error in the cost functional, we employed here the fact, that the terms

$$\begin{aligned}
\rho^q(q_k,u_k,z_k)(p-\hat{p}_k), &\qquad \rho^p(\xi_k,p_k,v_k,y_k)(q-\hat{q}_k),\\
\rho^q(q_{kh},u_{kh},z_{kh})(p_k-\hat{p}_{kh}), &\qquad \rho^p(\xi_{kh},p_{kh},v_{kh},y_{kh})(q_k-\hat{q}_{kh}),\\
\rho^u(q_\sigma,u_\sigma)(y_{kh}-\hat{y}_\sigma), &\qquad \rho^z(q_\sigma,u_\sigma,z_\sigma)(v_{kh}-\hat{v}_\sigma),\\
\rho^v(\xi_\sigma,p_\sigma,v_\sigma)(z_{kh}-\hat{z}_\sigma), &\qquad \rho^y(\xi_\sigma,p_\sigma,v_\sigma,y_\sigma)(u_{kh}-\hat{u}_\sigma)
\end{aligned}$$

vanish if $\hat{p}_k$, $\hat{q}_k$, $\hat{p}_{kh}$, $\hat{q}_{kh}$, $\hat{y}_\sigma$, $\hat{v}_\sigma$, $\hat{z}_\sigma$, $\hat{u}_\sigma$ are chosen appropriately.

Remark 4.4. For the error estimation with respect to the cost functional no additional equations have to be solved. The error estimation with respect to a given quantity of interest requires the computation of the auxiliary variables p_σ, v_σ, y_σ. The additional numerical effort is similar to the execution of one step of the SQP or Newton's method.

5. Numerical realization

5.1. Evaluation of the error estimators

In this subsection, we concretize the a posteriori error estimator developed in the previous section for the cG(1)dG(0) space-time discretization on quadrilateral meshes in two space dimensions. That is, we consider the combination of dG(0) time discretization with piecewise bi-linear finite elements for the space discretization.

The error estimates presented in the previous section involve interpolation errors of the time, space, and the control discretizations. We approximate these errors using interpolations in higher order finite element spaces. To this end, we introduce linear operators Π_h, Π_k, and Π_d, which will map the computed solutions

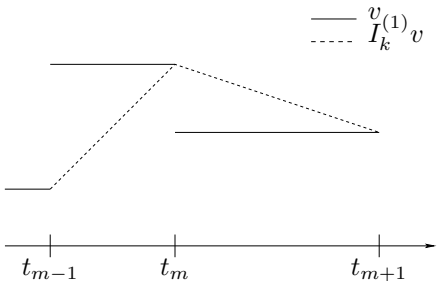


FIGURE 1. Temporal Interpolation

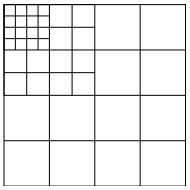


FIGURE 2. Patched Mesh

to the approximations of the interpolation errors:

$$\begin{aligned} z - \hat{z}_k &\approx \Pi_k z_k & u - \hat{u}_k &\approx \Pi_k u_k \\ z_k - \hat{z}_{kh} &\approx \Pi_h z_{kh} & u_k - \hat{u}_{kh} &\approx \Pi_h u_{kh} \\ q_{kh} - \hat{q}_\sigma &\approx \Pi_d q_\sigma \end{aligned}$$

$$\begin{aligned} y - \hat{y}_k &\approx \Pi_k y_k & v - \hat{v}_k &\approx \Pi_k v_k \\ y_k - \hat{y}_{kh} &\approx \Pi_h y_{kh} & v_k - \hat{v}_{kh} &\approx \Pi_h v_{kh} \\ p_{kh} - \hat{p}_\sigma &\approx \Pi_d p_\sigma \end{aligned}$$

For the here considered case of cG(1)dG(0) discretization of the state space, the operators are chosen as

$$\begin{aligned} \Pi_k &= I_k^{(1)} - \mathrm{id} \qquad \text{with} \quad I_k^{(1)} \colon X_k^0 \to \widetilde{X}_k^1, \\ \Pi_h &= I_{2h}^{(2)} - \mathrm{id} \qquad \text{with} \quad I_{2h}^{(2)} \colon X_{k,h}^{0,1} \to \widetilde{X}_{k,2h}^{0,2}. \end{aligned}$$

Here, $\widetilde{X}_k^1$ denotes the space of continuous piecewise linear polynomials with values in V. The action of the piecewise linear interpolation operator $I_k^{(1)}$ in time is depicted in Figure 1.

The piecewise bi-quadratic spatial interpolation $I_{2h}^{(2)}$ can be easily computed if the underlying mesh provides a patch structure. That is, one can always combine four adjacent cells to a macro cell on which the bi-quadratic interpolation can be defined. An example of such an patched mesh is shown in Figure 2.

The choice of Π_d depends on the discretization of the control space Q. If the finite dimensional subspaces Q_d are constructed similar to the discrete state spaces, one can directly choose for Π_d a modification of the operators Π_k and Π_h defined above. If for example the controls q only depend on time and the discretization is done with piecewise constant polynomials, we can choose $\Pi_d = I_d^{(1)} - \mathrm{id}$. If the control space Q is already finite dimensional and Q_d is chosen equal to Q, which is usually the case in the context of parameter estimation, it is possible to choose $\Pi_d = 0$ and thus, the estimator for the error $J(q_{kh}, u_{kh}) - J(q_\sigma, u_\sigma)$ is zero—as well as this discretization error itself.

In order to make the error representations from the previous section computable, we replace the residuals linearized on the solution of semi-discretized problems by the linearization at full discrete solutions.

We finally obtain the following computable a posteriori error estimator for the cost functional J

$$J(q,u) - J(q_\sigma, u_\sigma) \approx \eta_k^J + \eta_h^J + \eta_d^J$$

with

$$\begin{aligned}
\eta_k^J &:= \frac{1}{2}\Big(\rho^u(q_\sigma, u_\sigma)(\Pi_k z_\sigma) + \rho^z(q_\sigma, u_\sigma, z_\sigma)(\Pi_k u_\sigma)\Big)\\
\eta_h^J &:= \frac{1}{2}\Big(\rho^u(q_\sigma, u_\sigma)(\Pi_h z_\sigma) + \rho^z(q_\sigma, u_\sigma, z_\sigma)(\Pi_h u_\sigma)\Big)\\
\eta_d^J &:= \frac{1}{2}\rho^q(q_\sigma, u_\sigma, z_\sigma)(\Pi_d q_\sigma).
\end{aligned}$$

For the quantity of interest I the error estimator is given by

$$I(q,u) - I(q_\sigma, u_\sigma) \approx \eta_k^I + \eta_h^I + \eta_d^I$$

with

$$\begin{aligned}
\eta_k^I &:= \frac{1}{2}\Big(\rho^u(q_\sigma, u_\sigma)(\Pi_k y_\sigma) + \rho^z(q_\sigma, u_\sigma, z_\sigma)(\Pi_k v_\sigma)\\
&\qquad + \rho^v(\xi_\sigma, v_\sigma, p_\sigma)(\Pi_k z_\sigma) + \rho^y(\xi_\sigma, v_\sigma, y_\sigma, p_\sigma)(\Pi_k u_\sigma)\Big)\\
\eta_h^I &:= \frac{1}{2}\Big(\rho^u(q_\sigma, u_\sigma)(\Pi_h y_\sigma) + \rho^z(q_\sigma, u_\sigma, z_\sigma)(\Pi_h v_\sigma)\\
&\qquad + \rho^v(\xi_\sigma, v_\sigma, p_\sigma)(\Pi_h z_\sigma) + \rho^y(\xi_\sigma, v_\sigma, y_\sigma, p_\sigma)(\Pi_h u_\sigma)\Big)\\
\eta_d^I &:= \frac{1}{2}\Big(\rho^q(q_\sigma, u_\sigma, z_\sigma)(\Pi_d p_\sigma) + \rho^p(\xi_\sigma, v_\sigma, y_\sigma, p_\sigma)(\Pi_d q_\sigma)\Big).
\end{aligned}$$

The presented a posteriori error estimators are directed towards two aims: assessment of the discretization error and improvement of the accuracy by local refinement. For the second aim, the information provided by the error estimators have to be localized to cellwise or nodewise contributions (local error indicators).

We concretize this procedure here for the error estimators η_k^J and η_h^J assessing the error with respect to the cost functional. For concrete choices of discretizations

for the control space one can proceed with η_d^J in the same manner. Of course, the error indicators η_k^I, η_h^I, and η_d^I can be treated similarly, too.

For localizing the error estimators, we split up the error estimates η_k^J and η_h^J into their contributions on each subinterval I_m by

$$\eta_k^J = \sum_{m=1}^{M} \eta_k^{J,m} \qquad \text{and} \qquad \eta_h^J = \sum_{m=0}^{M} \eta_h^{J,m},$$

where the contributions $\eta_k^{J,m}$ and $\eta_h^{J,m}$ are given in terms of the time stepping residuals ρ_m^u and ρ_m^z as

$$\eta_k^{J,m} = \frac{1}{2}\Big\{\rho_m^u(q_\sigma, u_\sigma)(P_k z_\sigma) + \rho_m^z(q_\sigma, u_\sigma, z_\sigma)(P_k u_\sigma)\Big\},$$
$$\eta_h^{J,m} = \frac{1}{2}\Big\{\rho_m^u(q_\sigma, u_\sigma)(P_h z_\sigma) + \rho_m^z(q_\sigma, u_\sigma, z_\sigma)(P_h u_\sigma)\Big\}.$$

Thereby, the time stepping residuals ρ_m^u and ρ_m^z are those parts of the global residuals ρ^u and ρ^z belonging to the time interval I_m or to the initial time $t = 0$ for $m = 0$.

Whereas the temporal indicators $\eta_k^{J,m}$ can be used directly for determining the time intervals to be refined, the indicators $\eta_h^{J,m}$ for the spatial discretization error have to be further localized to indicators on each spatial mesh. For details of this further localization procedure we refer, e.g., to [8, 36].

5.2. Adaptive algorithm

Goal of the adaption of the different types of discretizations has to be the equilibrated reduction of the corresponding discretization errors. If a given tolerance TOL has to be reached, this can be done by refining each discretization as long as the value of this part of the error estimator is greater than $1/3$ TOL. We want to present here a strategy which will equilibrate the different discretization errors even if no tolerance is given.

Aim of the equilibration algorithm presented in the sequel is to obtain discretization such that

$$|\eta_k| \approx |\eta_h| \approx |\eta_d|$$

and to keep this property during the further refinement. Here, the estimators η_i denote the estimators η_i^J for the cost functional J or η_i^I for the quantity of interest I.

For doing this equilibration, we choose an "equilibration factor" $e \approx 2 - 5$ and propose the following strategy: We compute a permutation (a, b, c) of the discretization indices (k, h, d) such that

$$|\eta_a| \geq |\eta_b| \geq |\eta_c|,$$

and define the relations

$$\gamma_{ab} := \left|\frac{\eta_a}{\eta_b}\right| \geq 1, \qquad \gamma_{bc} := \left|\frac{\eta_b}{\eta_c}\right| \geq 1.$$

TABLE 1. Equilibration Strategy

Relation between the estimators	Discretizations to be refined
$\gamma_{ab} \le e$ and $\gamma_{bc} \le e$	a, b, and c
$\gamma_{bc} > e$	a and b
else ($\gamma_{ab} > e$ and $\gamma_{bc} \le e$)	a

Then we decide by means of Table 1 in every repetition of the adaptive refinement algorithm which discretization shall be refined. For every discretization to be adapted we select by means of the local error indicators the cells for refinement. For this purpose there are several strategies available, see, e.g., [4].

6. Numerical examples

6.1. Surface hardening of steel

We consider the optimal control of laser surface hardening of steel. In this process, a laser beam moves along the surface of a workpiece. The heating induced by the laser is accompanied by a phase transition, in which austenite, the high temperature phase in steel, is produced. Due to further phase transitions (which are not contained in the considered model) the desired hardening effect develops.

The goal is to control this hardening process such that a desired hardening profile is produced. Since in practical applications, the moving velocity of the laser beam is kept constant, the most important control parameter is the energy of the laser beam. Especially when there are large variations in the thickness of the workpiece or in regions near the boundaries of the workpiece, the proper adjustment of the laser energy is crucial to meet the given hardening profile.

The configuration of the control problem to be investigated in this section is taken from [15, 20]:

$$\begin{aligned}
\partial_t a &= \frac{1}{\tau(\theta)}(a_{\mathrm{eq}}(\theta) - a)\mathcal{H}_\delta(a_{\mathrm{eq}}(\theta) - a) && \text{in } \Omega \times (0,T),\\
\rho c_{\mathrm{p}} \partial_t \theta - \varepsilon \Delta \theta &= -\rho L \partial_t a + q\Lambda && \text{in } \Omega \times (0,T),\\
\partial_n \theta &= 0 && \text{on } \partial\Omega \times (0,T)\\
a &= 0 && \text{on } \Omega \times \{\,0\,\},\\
\theta &= \theta_0 && \text{on } \Omega \times \{\,0\,\},
\end{aligned} \tag{6.1}$$

Here a is the volume fraction of austenite, a_{eq} is the equilibrium volume fraction of austenite, and τ is a time constant. Both a_{eq} and τ depend on the temperature θ. Furthermore, as in [20], $\mathcal{H}_\delta$ is a monotone regularization of the Heaviside function and the density ρ, the heat capacity c_{p}, the heat conductivity ε, and the latent heat L are assumed to be positive constants. The term $q(t)\Lambda(x,t)$ describes the

volumetric heat source due to laser radiation, where q acts as time-dependent control variable. Thus, the optimal control is searched for in $Q = L^2(0,T)$.

Here, because of the Neumann boundary conditions for θ, the space X is defined using $V = H^1(\Omega)$ and $H = L^2(\Omega)$.

As cost functional to be minimized, we choose

$$J(q,u) = \frac{\beta}{2}\int_0^T \|a(t) - \hat{a}(t)\|^2_{L^2(\Omega)}\,dt + \frac{\alpha}{2}\int_0^T q(t)^2\,dt,$$

where $\hat{a}$ is a given desired volume fraction of austenite. For the computations, we choose the physical parameters for the heat equation accordingly to [20] as

$$\rho c_{\mathrm{p}} = 1.17, \qquad \varepsilon = 0.153, \qquad \text{and} \qquad \rho L = 150.$$

The equilibrium volume fraction a_{eq} and the time constant τ are constructed by cubic spline interpolation from values taken also from [20]. The initial condition for the temperature is set to $\theta_0 = 20$. The laser source Λ is modeled by

$$\Lambda(x,t) = \frac{4\kappa A}{\pi D^2}\exp\left(-\frac{2(x_1 - vt)^2}{D^2}\right)\exp(\kappa x_2), \quad x = (x_1,x_2)^T,$$

where the values of the parameters are taken from [20] as $D = 0.47$, $\kappa = 60$, $A = 0.3$, and $v = 1.15$.

For the numerical computations, we choose the domain Ω to be $(0,5)\times(-1,0)$ and determine the final time T such that the laser, which moves from $(0,0)$ to $(5,0)$, reaches the boundary at $(5,0)$ at time T. Thus, we set $T = 5/v \approx 4.34782$. The desired volume fraction $\hat{a}$ (cf. Figure 3(c)) is chosen as

$$\hat{a}(x) := \begin{cases} 1 & \text{for } 0 \geq x_1 \geq -\frac{1}{8} \\ 0 & \text{for } -\frac{1}{8} > x_1 \geq -1 \end{cases}, \quad x = (x_1,x_2)^T,$$

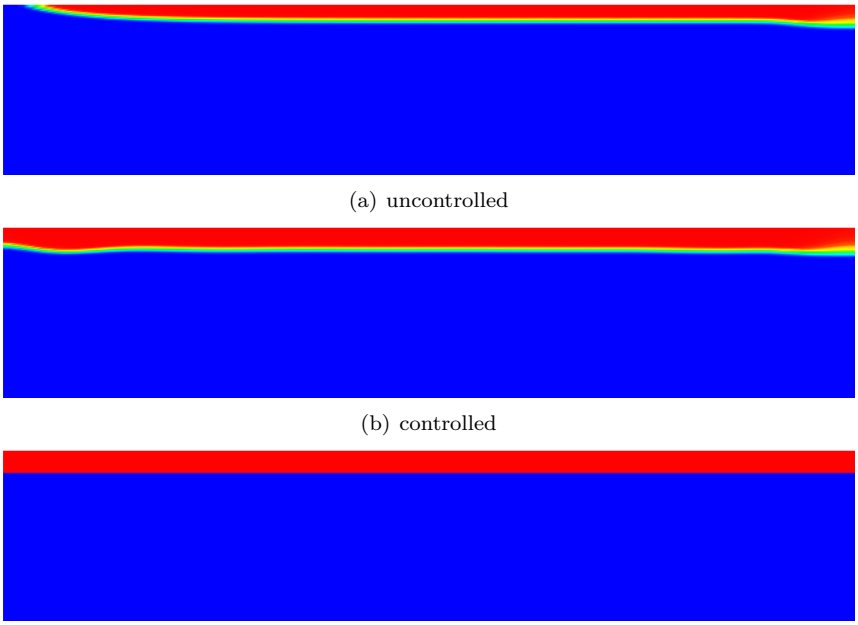
and for the parameters α and β from the definition of the objective functional J we take $\alpha = 10^{-4}$ and $\beta = 3500$. The parameter δ for the regularized Heaviside function is chosen as $\delta = 0.15$.

For discretizing the state space, we employ the cG(1)dG(0) discretization Since the control space is given by $Q = L^2((0,T),\mathbb{R})$, we have to discretize the controls only in time. Correspondingly to the state space, we choose a dG(0) discretization based on a possibly coarser step size than the step size used for discretizing the state space.

At first, we investigate the qualitative behavior of the optimization algorithm. Figure 3 presents the distribution of austenite at final time T before (a) and after (b) the optimization on a fine discretization of the state space. To compare with, the desired state is depicted in Figure 3(c).

In Figure 4, we present a comparison of different refinement strategies for the spatial discretization. We depict the development of the error $e_h^J := J(q_{kd},u_{kd}) - J(q_\sigma,u_\sigma)$ caused by the spatial discretization of the state space. Thereby, we consider the following three types of refinement:

- Uniform refinement



(a) uncontrolled

(b) controlled

(c) desired

FIGURE 3. Distribution of austenite at final time T

- Local refinement based on the error indicator η_h^J with one fixed mesh for all time steps
- Local refinement based on η_h^J but allowing separate spatial meshes for each time interval by using dynamically changing meshes

We observe that by the usage of local refinement the number of grid points can be reduced from $N = 16{,}641$ to $N = 5{,}271$ to achieve an error of 10^{-5}. Moreover, if we allow dynamically changing meshes, we only need $N_{\max} = 3{,}873$ grid points. The total number of degrees of freedom in the space discretization ($\dim X_{k,h}^{0,1}$) is reduced even by a factor of 5.7 when employing local refinement on dynamic meshes.

In Figure 5, a selection from the sequence of locally refined meshes is given. Thereby, we detect a strong refinement at the position where the laser currently acts and at the region around the transition from hardened to not hardened steel. In this region, the optimal distribution of austenite as well as the desired hardening profile exhibits spatial discontinuities.

We now couple the temporal and spatial estimators by the equilibration strategy described in Section 5.2. Since in this example, we do not benefit from local refinement in time, we allow only uniform refinement of the time steps. In space, we allow the adaptation procedure to use dynamically changing meshes. Results

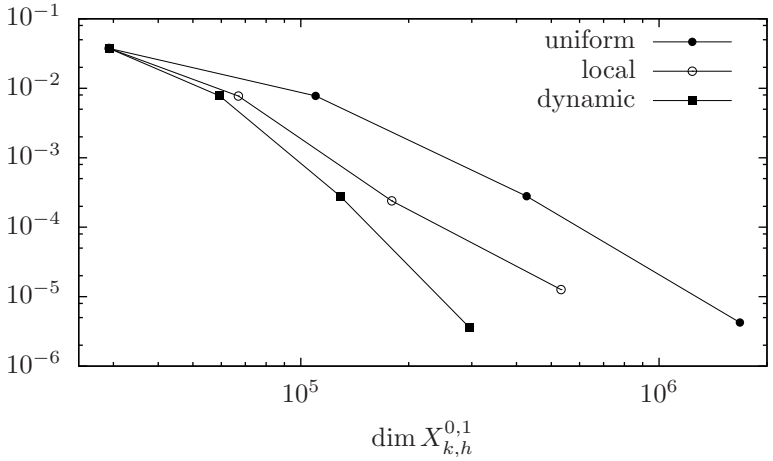


FIGURE 4. Comparison of the relative error $|e_h^J|/J$ for uniform and local refinement of the triangulation using dynamically changing meshes

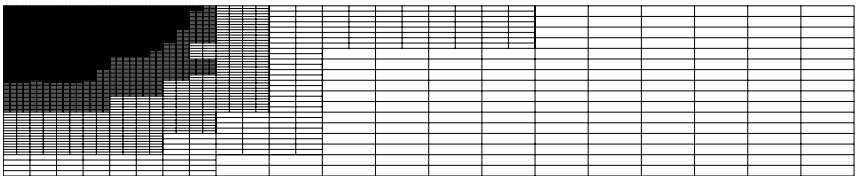
TABLE 2. Local refinement on dynamic meshes with equilibration

N_{tot}	$N_{\max}$	M	η_h^J/J	η_k^J/J	$\eta_h^J/J + \eta_k^J/J$	e_{kh}^J/J	I_{eff}
14739	289	50	$-7.4\cdot10^{-03}$	$2.3\cdot10^{-02}$	$1.622\cdot10^{-02}$	$-4.916\cdot10^{-03}$	-0.30
59325	675	100	$-2.8\cdot10^{-03}$	$1.3\cdot10^{-02}$	$1.049\cdot10^{-02}$	$7.828\cdot10^{-03}$	0.74
257867	1659	200	$-3.9\cdot10^{-04}$	$6.3\cdot10^{-03}$	$6.040\cdot10^{-03}$	$7.445\cdot10^{-03}$	1.23
515115	1659	400	$-3.9\cdot10^{-04}$	$3.2\cdot10^{-03}$	$2.827\cdot10^{-03}$	$3.454\cdot10^{-03}$	1.22
1029611	1659	800	$-4.3\cdot10^{-04}$	$1.6\cdot10^{-03}$	$1.193\cdot10^{-03}$	$1.424\cdot10^{-03}$	1.19
4721397	3911	1600	$9.8\cdot10^{-06}$	$7.8\cdot10^{-04}$	$8.143\cdot10^{-04}$	$9.375\cdot10^{-04}$	1.15

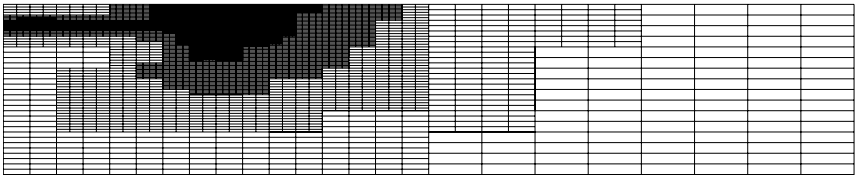
of this computation are given in Table 2. Therein, we observe that for the chosen initial levels of discretizations, the contribution from the spatial discretization error to the overall error is much smaller than the contribution from the temporal discretization error. Consequently, the equilibration procedure decides for example to keep the spatial meshes fixed while increasing the number of time steps from 200 over 400 to 800 time steps. The effectivity index given in the last column of this table is defined as usual by

$$I_{\text{eff}} := \frac{e_{kh}^J}{\eta_h^J + \eta_k^J}.$$

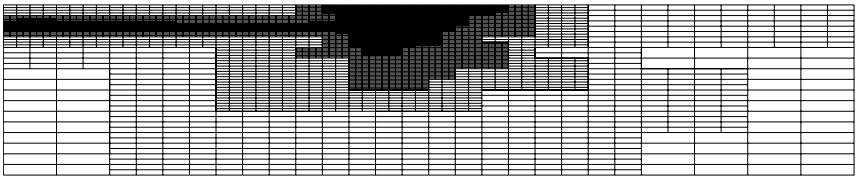
This implies that uniform refinement of the time and space discretizations without the knowledge of the size of the different error contributions can not be competitive. For the efficient equilibration—and thus the efficient reduction of the error—estimations of the size of each involved discretization errors are essential. Furthermore, the table demonstrates that the estimator $\eta_h^J/J + \eta_k^J/J$ is in very



(a) $t = 0.43$

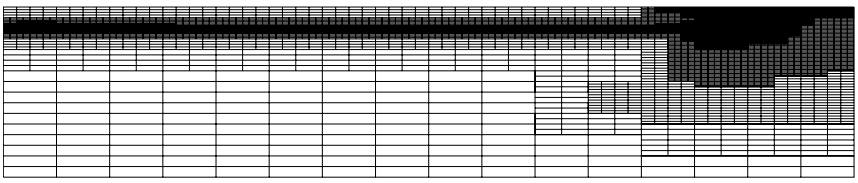


(b) $t = 1.30$



(c) $t = 2.17$

(d) $t = 3.04$



(e) $t = 3.91$

FIGURE 5. Locally refined meshes for $t \in \{\, 0.43, 1.30, 2.17, 3.04, 3.91 \,\}$

good agreement with the relative error $|e^J_{kh}|/J$ Thereby, the error e^J_{kh} is defined as $e^J_{kh} := J(q_d, u_d) - J(q_\sigma, u_\sigma)$ with an approximation of $J(q_d, u_d)$ obtained by extrapolation. For normalizing the errors and the estimators we use J, which denotes an approximation of the exact value of the cost functional $J(q, u)$.

7. Propagation of laminar flames

In this section, we consider a parameter estimation problem arising in chemical combustion processes. We aim at the identification of an unknown parameter in a reaction mechanism governed by an Arrhenius law. This formulation is employed to model the propagation of laminar flames through a channel. The channel is narrowed by two heat absorbing obstacles influencing the traveling of the flame.

The identification of the unknown parameter is done employing measurements of the solution components at four spatial points at final time. Using these values, the cost functional is constructed by means of a least-squares formalism.

The governing equations for the considered problem are taken from an example given in [22]. They describe the major part of gaseous combustion under the low Mach number hypothesis. In this approach, the dependency of the fluid density on the pressure is eliminated while the temperature dependence remains. If additionally the dependence on the temperature is neglected, the motion of the fluid becomes independent on the temperature and the species concentration. Hence, one can solve the temperature and the species equation alone specifying any solenoidal velocity field v. In particular, $v = 0$ is an interesting case.

Introducing the dimensionless temperature θ, denoting by Y the species concentration, and assuming constant diffusion coefficients yields the system of equations

$$\begin{aligned} \partial_t \theta - \Delta\theta &= \omega(Y,\theta) && \text{in } \Omega \times (0,T), \\ \partial_t Y - \frac{1}{\mathrm{Le}}\Delta Y &= -\omega(Y,\theta) && \text{in } \Omega \times (0,T), \\ \theta &= \theta_0 && \text{on } \Omega \times \{0\}, \\ Y &= Y_0 && \text{on } \Omega \times \{0\}, \end{aligned} \tag{7.1}$$

where the Lewis number Le is the ratio of diffusivity of heat and diffusivity of mass. We use a simple one-species reaction mechanism governed by an Arrhenius law given by

$$\omega(Y,\theta) = \frac{\beta^2}{2\mathrm{Le}} Y \mathrm{e}^{\frac{\beta(\theta-1)}{1+\alpha(\theta-1)}}, \tag{7.2}$$

in which an approximation for large activation energy has been employed.

Here, we consider a freely propagating laminar flame described by (7.1) and its response to a heat absorbing obstacle, a set of cooled parallel rods with rectangular cross section (cf. Figure 6). The computational domain has width 16 and length 60. The obstacle covers half of the width and has length $L/4$. The boundary

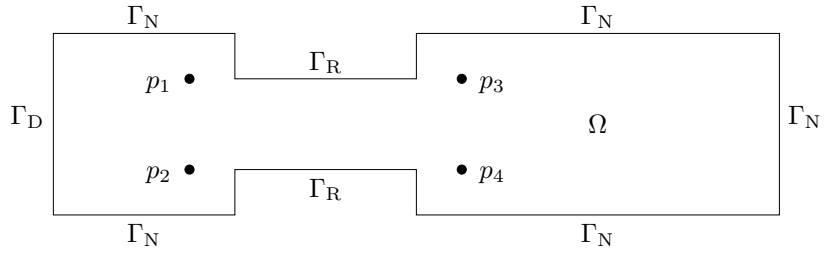


FIGURE 6. Computational domain Ω and measurement points p_i

conditions are chosen as

$$\begin{aligned} \theta &= 1 \quad \text{on } \Gamma_D \times (0,T), & \partial_n \theta &= 0 \quad \text{on } \Gamma_N \times (0,T), \\ Y &= 0 \quad \text{on } \Gamma_D \times (0,T), & \partial_n Y &= 0 \quad \text{on } \Gamma_N \times (0,T), \end{aligned}$$

$$\begin{aligned} \partial_n \theta &= -\kappa\theta \quad \text{on } \Gamma_R \times (0,T), \\ \partial_n Y &= 0 \quad \text{on } \Gamma_R \times (0,T), \end{aligned}$$

where the heat absorption is modeled by boundary conditions of Robin type on Γ_R.

The initial condition is the analytical solution of a one-dimensional right-traveling flame in the limit $\beta \to \infty$ located left of the obstacle:

$$\theta_0(x) = \begin{cases} 1 & \text{for } x_1 \le \tilde{x}_1 \\ \mathrm{e}^{\tilde{x}_1 - x_1} & \text{for } x_1 > \tilde{x}_1 \end{cases},$$

$$Y_0(x) = \begin{cases} 0 & \text{for } x_1 \le \tilde{x}_1 \\ 1 - \mathrm{e}^{\mathrm{Le}(\tilde{x}_1 - x_1)} & \text{for } x_1 > \tilde{x}_1 \end{cases}.$$

For the computations, the occurring parameters are set as in [22] to

$$\mathrm{Le} = 1, \qquad \beta = 10, \qquad \kappa = 0.1, \qquad \tilde{x}_1 = 9,$$

whereas the temperature ratio α, which determines the gas expansion in non-constant density flows, is the objective of the parameter estimation.

We define the pair of solution components $u := (\theta, Y) \in \tilde{u} + X^2$ and denote the parameter α to be estimated by $q \in Q := \mathbb{R}$. For the definition of the state space X, we use here the spaces V and H given as

$$V = \left\{ v \in H^1(\Omega) \;\middle|\; v\big|_{\Gamma_D} = 0 \right\} \qquad \text{and} \qquad H = L^2(\Omega).$$

The function $\tilde{u}$ is defined to fulfill the prescribed Dirichlet data as $\tilde{u}\big|_{\Gamma_D} = (1, 0)$.

The unknown parameter α is estimated here using information from pointwise measurements of θ and Y at four points $p_i \in \Omega$, $i = 1, 2, 3, 4$, at final time $T = 60$. This parameter identification problem can be formulated by means of a cost

functional of least-squares type, that is

$$J(q,u) = \frac{1}{2}\sum_{i=1}^{4}\big(\theta(p_i,T) - \hat{\theta}_i\big)^2 + \frac{1}{2}\sum_{i=1}^{4}\big(Y(p_i,T) - \hat{Y}_i\big)^2.$$

The values of the artificial measurements $\hat{\theta}_i$ and $\hat{Y}_i$, $i = 1,2,3,4$, are obtained from a reference solution computed on fine space and time discretizations.

The consideration of point measurements does not fulfill the assumption on the cost functional in (2.5), since the point evaluation is not bounded as a functional on H. Therefore, the point functionals here have to be understood as regularized functionals defined on H. An a priori analysis of parameter estimation problems governed by elliptic equations and using such types of point functionals can be found in [35].

For the considered type of parameter estimation problems, one is usually not interested in reducing the discretization error measured in terms of the cost functional J. The focus is rather on the error in the parameter q itself. Hence, we define the quantity of interest I as

$$I(q,u) = q,$$

and apply the techniques presented in Section 4 for estimating the discretization error with respect to I. Since the control space in this application is given by $Q = \mathbb{R}$, we do not discretize Q. Thus, there is no discretization error due to the control discretization and the a posteriori error estimator η^I consists only of η_k^I and η_h^I.

For the computations, the state is discretized using the cG(1)dG(0) approach. We define the temporal and spatial discretization errors e_k^I and e_h^I as

$$e_k^I := I(q_h,u_h) - I(q_{kh},u_{kh}) \qquad \text{and} \qquad e_h^I := I(q_k,u_k) - I(q_{kh},u_{kh}).$$

The values of $I(q_h,u_h)$ and $I(q_k,u_k)$ are extrapolated from computations on a sequence of fine time and space discretizations, respectively. Since we have $I(q,u) = q = \alpha \approx 0.8$, there is no difference between the consideration of relative or absolute errors.

We apply the adaptive algorithm described above to the considered problem using simultaneous refinement of the space and time discretizations. Thereby, the refinements are coupled by the equilibration strategy introduced in Section 5.2. The Tables 3 and 4 demonstrate the effectivity of the error estimator $\eta_h^I + \eta_k^I$ on locally refined discretizations using fixed and dynamically changing spatial triangulations. The effectivity index given in the last column of these tables is defined as usual by

$$I_{\text{eff}} := \frac{e^I}{\eta_h^I + \eta_k^I}.$$

In Figure 7, we compare uniform refinement of the space and time discretizations with local refinement of both discretizations on a fixed spatial triangulation and on dynamically changing triangulations. We gain an remarkable reduction of

TABLE 3. Local refinement on a *fixed mesh* with equilibration

N	M	η_h^I	η_k^I	$\eta_h^I + \eta_k^I$	e^I	I_{eff}
269	512	$4.3 \cdot 10^{-02}$	$-8.4 \cdot 10^{-03}$	$3.551 \cdot 10^{-02}$	$-2.889 \cdot 10^{-02}$	-0.81
635	512	$5.5 \cdot 10^{-03}$	$-9.1 \cdot 10^{-03}$	$-3.533 \cdot 10^{-03}$	$-4.851 \cdot 10^{-02}$	13.72
1847	722	$-1.5 \cdot 10^{-02}$	$-3.6 \cdot 10^{-03}$	$-1.889 \cdot 10^{-02}$	$-3.024 \cdot 10^{-02}$	1.60
5549	1048	$-6.5 \cdot 10^{-03}$	$-2.5 \cdot 10^{-03}$	$-9.074 \cdot 10^{-03}$	$-1.097 \cdot 10^{-02}$	1.20
14419	1088	$-2.4 \cdot 10^{-03}$	$-2.5 \cdot 10^{-03}$	$-5.064 \cdot 10^{-03}$	$-5.571 \cdot 10^{-03}$	1.10
43343	1102	$-8.5 \cdot 10^{-04}$	$-2.5 \cdot 10^{-03}$	$-3.453 \cdot 10^{-03}$	$-3.693 \cdot 10^{-03}$	1.06

TABLE 4. Local refinement on *dynamic meshes* with equilibration

N_{tot}	$N_{\max}$	M	η_h^I	η_k^I	$\eta_h^I + \eta_k^I$	e^I	I_{eff}
137997	269	512	$4.3 \cdot 10^{-02}$	$-8.4 \cdot 10^{-03}$	$3.551 \cdot 10^{-02}$	$-2.889 \cdot 10^{-02}$	-0.81
238187	663	512	$3.5 \cdot 10^{-03}$	$-8.6 \cdot 10^{-03}$	$-5.192 \cdot 10^{-03}$	$-5.109 \cdot 10^{-02}$	9.84
633941	1677	724	$-1.6 \cdot 10^{-02}$	$-3.5 \cdot 10^{-03}$	$-2.015 \cdot 10^{-02}$	$-3.227 \cdot 10^{-02}$	1.60
1741185	2909	1048	$-7.3 \cdot 10^{-03}$	$-2.5 \cdot 10^{-03}$	$-9.869 \cdot 10^{-03}$	$-1.214 \cdot 10^{-02}$	1.23
3875029	4785	1098	$-2.2 \cdot 10^{-03}$	$-2.5 \cdot 10^{-03}$	$-4.792 \cdot 10^{-03}$	$-5.432 \cdot 10^{-03}$	1.13
9382027	10587	1140	$-7.9 \cdot 10^{-04}$	$-2.5 \cdot 10^{-03}$	$-3.301 \cdot 10^{-03}$	$-3.588 \cdot 10^{-03}$	1.08
23702227	25571	1160	$-2.8 \cdot 10^{-04}$	$-2.4 \cdot 10^{-03}$	$-2.756 \cdot 10^{-03}$	$-2.944 \cdot 10^{-03}$	1.06

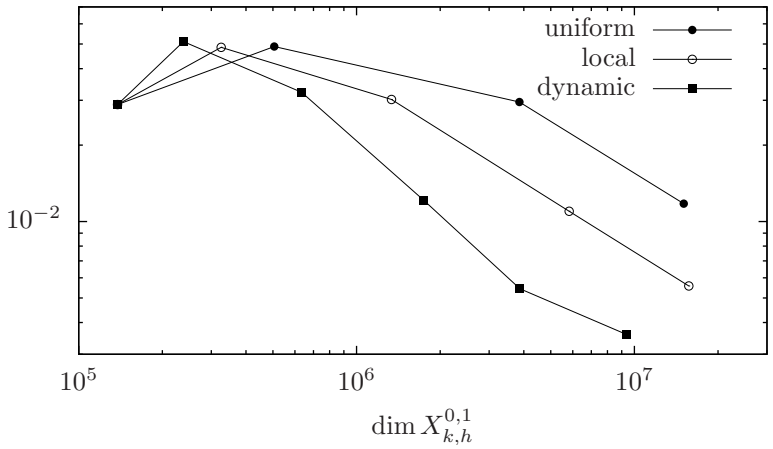


FIGURE 7. Comparison of the error $|e^I|$ for different refinement strategies

the required degrees of freedom for reaching a given tolerance. To meet for instance an error of $|e^I| \approx 10^{-2}$, the uniform refinement requires in total 15,056,225 degrees of freedom, the local refinement needs 5,820,901 degrees of freedom, and the dynamical refinement necessitates only 1,741,185 degrees of freedom. Thus, we gain a reduction of about 8.6.

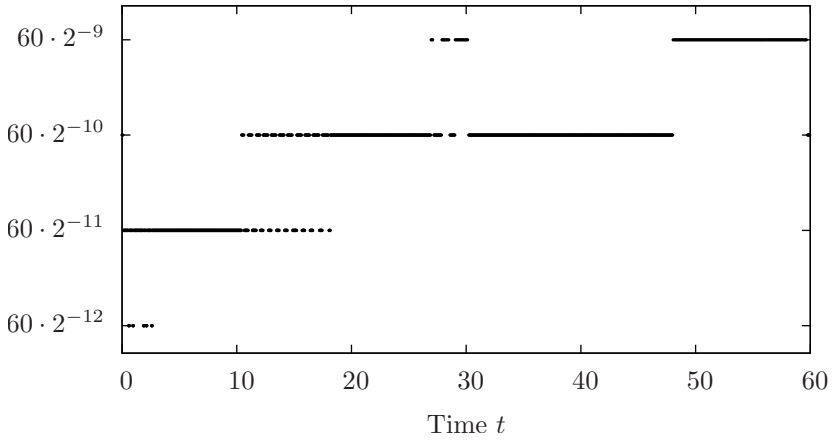


FIGURE 8. Visualization of the adaptively determined time step size k

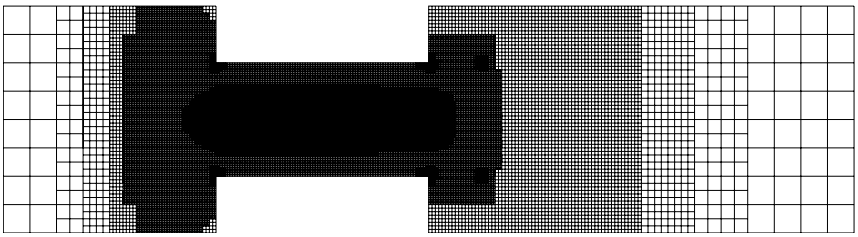


FIGURE 9. Locally refined fixed mesh

Figure 8 depicts the distribution of the temporal step size k resulting from a fully adaptive computation on dynamic meshes. We observe a strong refinement of the time steps at the beginning of the time interval, whereas the time steps at the end are determined by the adaptation to be eight times larger.

Before presenting a sequence of dynamically changing meshes, we show in Figure 9 a typical locally refined mesh obtained by computations on a fixed spatial triangulation. We note, that the refinement is especially concentrated at the four reentrant corners and the two measurement points behind the obstacle. The interior of the region with restricted cross section is also strongly refined.

Finally, Figure 10 shows the spatial triangulation and the reaction rate ω for certain selected time points. Thereby, ω is computed from the numerical solution by means of formula (7.2). We observe, that the refinement traces the front of the reaction rate ω until $t \approx 56$ (cf. Figure 10(d)). Afterwards, the mesh around the front becomes coarser and the refinement is concentrated at the four measurement

points p_i. Compared to the usage of one fixed triangulation, the usage of dynamically changing meshes enables us here to reduce the discretization error in terms of the quantity of interest at lower computational costs; cf. Figure 7.

References

[1] R. Becker: *Estimating the control error in discretized PDE-constraint optimization.* J. Numer. Math. **14** (2006), 163–185.

[2] R. Becker, H. Kapp, and R. Rannacher: *Adaptive finite element methods for optimal control of partial differential equations: Basic concepts.* SIAM J. Control Optimization **39** (2000), 113–132.

[3] R. Becker, D. Meidner, and B. Vexler: *Efficient numerical solution of parabolic optimization problems by finite element methods.* submitted to Optim. Methods Softw. (2005). In revision.

[4] R. Becker and R. Rannacher: *An optimal control approach to a-posteriori error estimation.* In A. Iserles, editor, *Acta Numerica 2001*, pages 1–102. Cambridge University Press, 2001.

[5] R. Becker and B. Vexler: *A posteriori error estimation for finite element discretizations of parameter identification problems.* Numer. Math. **96** (2004), 435–459.

[6] R. Becker and B. Vexler: *Mesh refinement and numerical sensitivity analysis for parameter calibration of partial differential equations.* J. Comp. Physics **206** (2005), 95–110.

[7] O. Benedix and B. Vexler: *A posteriori error estimation and adaptivity for elliptic optimal control problems with state constraints.* Comput. Optim. Appl. **44** (2009), 3–25.

[8] M. Braack and A. Ern: *A posteriori control of modeling errors and discretization errors.* Multiscale Model. Simul. **1** (2003), 221–238.

[9] P. G. Ciarlet: *The Finite Element Method for Elliptic Problems*, volume 40 of *Classics Appl. Math.* SIAM, Philadelphia, 2002.

[10] D. Clever, J. Lang, S. Ulbrich, and J. C. Ziems: *Combination of an adaptive multilevel SQP method and a space-time adaptive PDAE solver for optimal control problems.* Preprint SPP1253-094, SPP 1253, 2010.

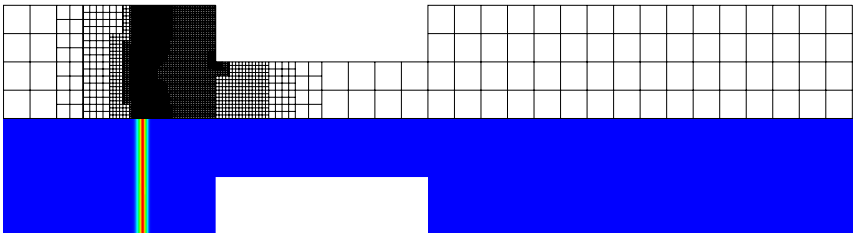
[11] R. Dautray and J.-L. Lions: *Mathematical Analysis and Numerical Methods for Science and Technology: Evolution Problems I*, volume 5. Springer-Verlag, Berlin, 1992.

[12] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson: *Introduction to adaptive methods for differential equations.* In A. Iserles, editor, *Acta Numerica 1995*, pages 105–158. Cambridge University Press, 1995.

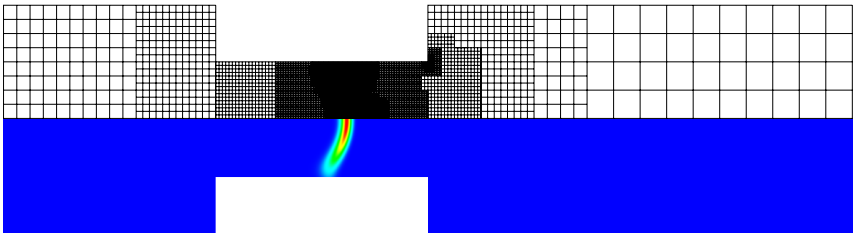
[13] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson: *Computational differential equations.* Cambridge University Press, Cambridge, 1996.

[14] K. Eriksson, C. Johnson, and V. Thomée: *Time discretization of parabolic problems by the discontinuous Galerkin method.* RAIRO Modelisation Math. Anal. Numer. **19** (1985), 611–643.

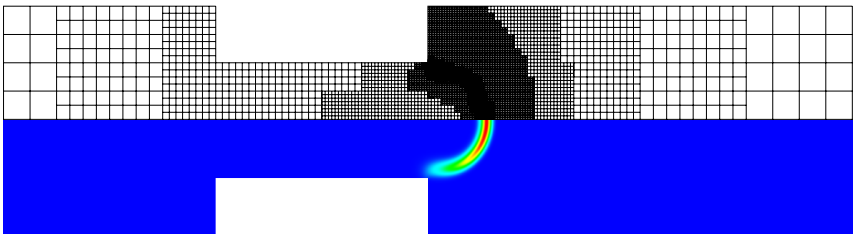
[15] J. Fuhrmann and D. Hömberg: *Numerical simulation of the surface hardening of steel.* Internat. J. Numer. Methods Heat Fluid Flow **9** (1999), 705–724.



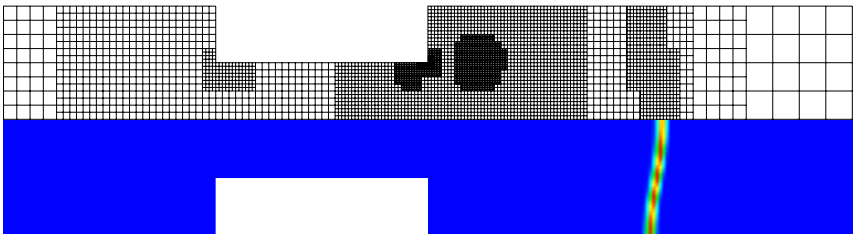
(a) $t = 1$



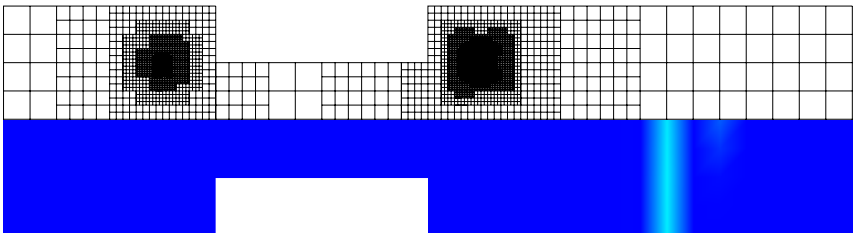
(b) $t = 20$



(c) $t = 40$



(d) $t = 56$



(e) $t = 60$

FIGURE 10. Locally refined meshes and reaction rate ω for $t \in \{ 1, 20, 40, 56, 60 \}$

[16] A. V. Fursikov: *Optimal Control of Distributed Systems: Theory and Applications*, volume 187 of *Transl. Math. Monogr.* AMS, Providence, 1999.

[17] A. Günther and M. Hinze: *A-posteriori error control of a state constrained elliptic control problem.* J. Numer. Math. **16** (2008), 307–322.

[18] M. Hintermüller, R. Hoppe, Y. Iliash, and M. Kieweg: *An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints.* ESIAM Control Optim. Calc. Var. **14** (2008), 540–560.

[19] M. Hintermüller and R. H. W. Hoppe: *Goal-oriented adaptivity in control constrained optimal control of partial differential equations.* SIAM J. Control Optim. **47** (2008), 1721–1743.

[20] D. Hömberg and S. Volkwein: *Suboptimal control of laser surface hardening using proper orthogonal decomposition.* Preprint 639, WIAS Berlin, 2001.

[21] R. Hoppe, Y. Iliash, C. Iyyunni, and N. Sweilam: *A posteriori error estimates for adaptive finite element discretizations of boundary control problems.* J. Numer. Math. **14** (2006), 57–82.

[22] J. Lang: *Adaptive Multilevel Solution of Nonlinear Parabolic PDE Systems. Theory, Algorithm, and Applications*, volume 16 of *Lecture Notes in Earth Sci.* Springer-Verlag, Berlin, 1999.

[23] J. Lang: *Adaptive computation for boundary control of radiative heat transfer in glass.* J. Comput. Appl. Math. **183** (2005), 312–326.

[24] J.-L. Lions: *Optimal Control of Systems Governed by Partial Differential Equations*, volume 170 of *Grundlehren Math. Wiss.* Springer-Verlag, Berlin, 1971.

[25] W. Liu, H. Ma, T. Tang, and N. Yan: *A posteriori error estimates for discontinuous galerkin time-stepping method for optimal control problems governed by parabolic equations.* SIAM J. Numer. Anal. **42** (2004), 1032–1061.

[26] W. Liu and N. Yan: *A posteriori error estimates for distributed convex optimal control problems.* Adv. Comput. Math **15** (2001), 285–309.

[27] W. Liu and N. Yan: *A posteriori error estimates for control problems governed by nonlinear elliptic equations.* Appl. Num. Math. **47** (2003), 173–187.

[28] W. Liu and N. Yan: *A posteriori error estimates for optimal control problems governed by parabolic equations.* Numer. Math. **93** (2003), 497–521.

[29] W. Liu and N. Yan: *Adaptive finite element methods for optimal control governed by PDEs*, volume 41 of *Series in Information and Computational Science.* Science Press, Beijing, 2008.

[30] D. Meidner: *Adaptive Space-Time Finite Element Methods for Optimization Problems Governed by Nonlinear Parabolic Systems.* PhD Thesis, Institut für Angewandte Mathematik, Universität Heidelberg, 2007.

[31] D. Meidner and B. Vexler: *Adaptive space-time finite element methods for parabolic optimization problems.* SIAM J. Control Optim. **46** (2007), 116–142.

[32] D. Meidner and B. Vexler: *A priori error estimates for space-time finite element approximation of parabolic optimal control problems. Part I: Problems without control constraints.* SIAM J. Control Optim. **47** (2008), 1150–1177.

[33] D. Meidner and B. Vexler: *A priori error estimates for space-time finite element approximation of parabolic optimal control problems. Part II: Problems with control constraints.* SIAM J. Control Optim. **47** (2008), 1301–1329.

[34] M. Picasso: *Anisotropic a posteriori error estimates for an optimal control problem governed by the heat equation.* Int. J. Numer. Methods PDE **22** (2006), 1314–1336.

[35] R. Rannacher and B. Vexler: *A priori error estimates for the finite element discretization of elliptic parameter identification problems with pointwise measurements.* SIAM J. Control Optim. **44** (2005), 1844–1863.

[36] M. Schmich and B. Vexler: *Adaptivity with dynamic meshes for space-time finite element discretizations of parabolic equations.* SIAM J. Sci. Comput. **30** (2008), 369–393.

[37] R. Serban, S. L. 2, and L. R. Petzold: *Adaptive algorithms for optimal control of time-dependent partial differential-algebraic equation systems.* Int. J. Numer. Meth. Engng. **57** (2003), 1457–1469.

[38] F. Tröltzsch: *Optimale Steuerung partieller Differentialgleichungen.* Friedr. Vieweg & Sohn Verlag, Wiesbaden, 2005.

[39] R. Verfürth: *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques.* Wiley/Teubner, New York-Stuttgart, 1996.

[40] B. Vexler: *Adaptive Finite Elements for Parameter Identification Problems.* PhD Thesis, Institut für Angewandte Mathematik, Universität Heidelberg, 2004.

[41] B. Vexler and W. Wollner: *Adaptive finite elements for elliptic optimization problems with control constraints.* SIAM J. Control Optim. **47** (2008), 509–534.

Dominik Meidner
Lehrstuhl für Mathematische Optimierung
Technische Universität München
Fakultät für Mathematik
Boltzmannstraße 3
85748 Garching b. München
Germany
e-mail: `meidner@ma.tum.de`

Boris Vexler
Lehrstuhl für Mathematische Optimierung
Technische Universität München
Fakultät für Mathematik
Boltzmannstraße 3
85748 Garching b. München
Germany
e-mail: `vexler@ma.tum.de`