Adaptive solution of
PDE-constrained optimal control problems

Rolf Rannacher∗

Version of March 30, 2006

Abstract

The efficient numerical solution of optimal control problems governed by PDEs requires work reduction by adaptive discretization and, in the nonstationary case, storage reduction by checkpointing techniques. This article explains the principles underlying these approaches and illustrates them by examples.

1 Adaptive discretization in optimal control

We consider optimal control problems

\[ J(u, q) \rightarrow \min \quad A(u, q) = 0, \]

with cost functionals \( J(\cdot, \cdot) \) and partial differential operators \( A(\cdot, \cdot) \) for states \( u \) and controls \( q \). The application of mesh adaptation in the discretization of such problems raises several fundamental questions:

- **What is the right notion of ‘admissibility’ of states \( u = u(q) \)?** Discretization inevitably introduces perturbation of the state equation. Achieving high accuracy in the discretization of PDEs is expensive. Hence, the extent to which ‘admissibility’ is relevant for the optimization process becomes a critical question and is a modeling issue.

- **How should admissibility be ‘measured’?** In solving ODEs, one may require the error to be uniformly ‘small’, but in the context of PDEs the choice of an appropriate error measure is less clear.

The efficient numerical solution of optimal control problems governed by PDEs requires work reduction by adaptive discretization and, in the nonstationary case, storage reduction by checkpointing techniques.

In the following, we consider finite element discretization using continuous piecewise polynomial trial and test functions for all unknowns on meshes \( T_h = \{ K \} \) which consist of non-degenerate quadrilaterals (‘cells’) \( K \) of width \( h_K := \text{diam}(K) \). The ‘global mesh size’ is \( h := \max_{K \in T_h} h_K \). These meshes are allowed to have ‘hanging nodes’ for simplifying local mesh refinement and coarsening.

∗This article is based on joint work with Roland Becker (University of Pau, France), Boris Vexler (RADON Institute, Linz, Austria), and Dominik Meidner (University of Heidelberg, Germany). We gratefully acknowledge the support by the German Research Foundation (DFG) through SFB 359 at Heidelberg University.
The question of how to organize systematic mesh adaptation in the approximation of optimal control problems will be discussed within the general framework of ‘goal-oriented’ error analysis. Let the goal of a numerical simulation be the computation of a quantity \( J(u) \) from the solution of a continuous model \( A(u) = 0 \) with accuracy \( TOL \) by using a discrete model \( A_h(u_h) = 0 \) of dimension \( N \). Accordingly, the goal of adaptivity is the optimal use of computing resources, i.e.,

\[
\{ TOL \text{ given: } N \to \min! \} \quad \text{or} \quad \{ N \text{ given: } TOL \to \min! \}.
\]

To achieve this goal, a posteriori information is used in terms of cell-wise refinement indicators \( \eta_K := \rho_K \omega_K \) based on ‘smoothness’ or ‘residual’ information, where \( \omega_K \) are certain weighting factors. These weights are obtained from the approximate solution of an associated linear ‘dual problem’, which is driven by the target functional \( J(\cdot) \) as right-hand side. This approach, called the 'Dual Weighted Residual (DWR) Method', has been developed in Becker/Rannacher [14, 15]. For a comprehensive discussion of the DWR method and related references, we refer to the survey article Becker/Rannacher [16] and the book Bangerth/Rannacher [2]. Related strategies of duality-based error control and mesh adaptation are described in Eriksson/Estep/Hansbo/Johnson [24] and Giles/Süli [27]. The application of the DWR method for mesh adaptation in optimal control problems has been developed in Becker/Kapp/Rannacher [11], Rannacher [40], Kapp [32], Becker [5], Bangerth [1], and Vexler [43]. For publications of these results and other references, we refer to the survey articles Becker et al. [10] and Carraro et al. [20].

2 An abstract a posteriori error estimate

Let \( V \) be a Hilbert space of states and \( Q \) a space of controls. Then, with a differentiable cost functional \( J(\cdot, \cdot) : V \times Q \to \mathbb{R} \) and semi-linear form \( a(\cdot, \cdot)(\cdot) \) on \( V \times Q \times V \), we consider the following optimization problem:

\[
J(u, q) \to \min! \quad a(u, q)(\psi) = 0 \quad \forall \psi \in V.
\]  

(2)

We assume that there is a locally unique minimum \( \{ u, q \} \in V \times Q \) which is characterized as corresponding to a saddle-point \( \{ u, q, \lambda \} \in V \times Q \times V \) of the Lagrangian functional

\[
\mathcal{L}(u, q, \lambda) := J(u, q) - a(u, q)(\lambda),
\]

where \( \lambda \) denotes the associated co-state (‘adjoint’ or ‘dual’ variable). Conditions under which this is true can be found in the standard literature on optimal control theory, e.g., Lions [35] and Fursikov [26]. The triplet \( \{ u, q, \lambda \} \in V \times Q \times V \) is determined by the saddle-point problem (so-called ‘optimality system’ or ‘Karush-Kuhn-Tucker (KKT) system’)

\[
a'_u(u, q)(\varphi, \lambda) = J'_u(u, q)(\varphi) \quad \forall \varphi \in V;
\]

\[
a'_q(u, q)(\chi, \lambda) = J'_q(u, q)(\chi) \quad \forall \chi \in Q,
\]

\[
a(u, q)(\psi) = 0 \quad \forall \psi \in V.
\]

(3)

Here, \( a'_u(u, q)(\varphi, \cdot), \quad a'_q(u, q)(\chi, \cdot), \quad J'_u(u, q)(\varphi), \) and \( J'_q(u, q)(\chi) \) denote the directional derivatives of the semi-linear form \( a(u, q)(\cdot) \) and the functional \( J(u, q) \) in the directions \( \varphi \) and \( \chi \), respectively.
The optimization problem is discretized by a standard Galerkin method using finite dimensional subspaces $V_h \times Q_h \subset V \times Q$:

$$J(u_h, q_h) \rightarrow \min! \quad a(u_h, q_h)(\psi_h) = 0 \quad \forall \psi_h \in V_h.$$  

Its local solutions $\{u_h, q_h\} \in V_h \times Q_h$ correspond to saddle points $x_h := \{u_h, q_h, \lambda_h\} \in X_h := V_h \times Q_h \times V_h$ of the Lagrangian $L(\cdot, \cdot, \cdot)$, i.e., to solutions of the discrete saddle-point problems

$$a_u'(u_h, q_h)(\varphi_h, \lambda_h) = J_u'(u_h, q_h)(\varphi_h) \quad \forall \varphi_h \in V_h,$$

$$a_q'(u_h, q_h)(\chi_h, \lambda_h) = J_q'(u_h, q_h)(\chi_h) \quad \forall \chi_h \in Q_h,$$

$$a(u_h, q_h)(\psi_h) = 0 \quad \forall \psi_h \in V_h.$$  

The idea is now to seek control of the discretization error with respect to the cost functional $J(\cdot, \cdot)$ of the optimal control problem in terms of the ‘primal’, ‘dual’, and ‘control’ residuals defined by

$$\rho^*(\cdot) := J_u'(u_h, q_h)(\cdot) - a_u'(u_h, q_h)(\cdot, \lambda_h) \quad \text{(dual residual)},$$

$$\rho^q(\cdot) := J_q'(u_h, q_h)(\cdot) - a_q'(u_h, q_h)(\cdot, \lambda_h) \quad \text{(control residual)},$$

$$\rho(\cdot) := -a(u_h, q_h)(\cdot) \quad \text{(primal residual)}.$$  

From Becker/Rannacher [16], we recall the following fundamental result.

**Theorem 1.** There holds the a posteriori error representation

$$J(u, q) - J(u_h, q_h) = \frac{1}{2} \rho^*(u - i_h u) + \frac{1}{2} \rho^q(q - i_h q) + \frac{1}{2} \rho(\lambda - i_h \lambda) + \mathcal{R}_h,$$  

for arbitrary approximations $i_h u, i_h q \in V_h, \text{ and } i_h \lambda \in Q_h$. The remainder $\mathcal{R}_h$ is cubic in the errors $e_u := u - u_h$, $e_q := q - q_h$, and $e_\lambda := \lambda - \lambda_h$.

**Proof.** On the product spaces $X := V \times Q \times V$ and $X_h := V_h \times Q_h \times V_h$ of triplets $x := \{u, q, \lambda\}$ and $x_h := \{u_h, q_h, \lambda_h\}$, respectively, we define the functional $L(x) := L(u, q, \lambda)$. Then, solving the saddle-point problems (3) and (5) is equivalent to determining stationary points $x$ and $x_h$ of $L(\cdot)$ on $X$ and $X_h$, respectively,

$$L'(x)(y) = 0 \quad \forall y \in X, \quad L'(x_h)(y_h) = 0 \quad \forall y_h \in X_h.$$  

By elementary calculus setting $e := x - x_h$, we have

$$L(x) - L(x_h) = \int_0^1 L'(x_h + se)(e) \, ds,$$

and, approximating the integral by the trapezoidal rule,

$$L(x) - L(x_h) = \frac{1}{2} \{L'(x_h)(e) + L'(x)(e)\} + \frac{1}{2} \int_0^1 L''(x_h + se)(e, e, e) \, s(s - 1) \, ds.$$  

Next, using the equations (7), we conclude that

$$L(x) - L(x_h) = \frac{1}{2} L'(x_h)(x - i_h x) + \mathcal{R}_h,$$  

3
for arbitrary \( i_h x \in X_h \), and with the cubic remainder term
\[
R_h := \frac{1}{2} \int_0^1 L''(x_h + se)(e, e, e) s(s - 1) ds.
\]

Finally, observing that
\[
L(x) - L(x_h) = \mathcal{L}(u, q, \lambda) - \mathcal{L}(u_h, q_h, \lambda_h)
= J(u, q) - a(u, q)(\lambda) - J(u_h, q_h) + a(u_h, q_h)(\lambda_h) = J(u, q) - J(u_h, q_h),
\]
the identity (6) follows.

**Remark 1.** The statement of Theorem 1 needs some explanations.

- The derivation of the error representation (6) does not require the uniqueness of solutions. The a priori assumption \( \{u_h, q_h, \lambda_h\} \rightarrow \{u, q, \lambda\}, as \ h \rightarrow 0 \), makes the result meaningful for non-unique solutions.
- The generally nonlinear saddle-point problem (5) can be solved by a Newton or SQP iteration with mesh adaptation in each step. Hence, the combined process may be viewed as a successive ‘model enrichment’.
- The evaluation of second-order sufficient optimality conditions requires the solution of auxiliary linear boundary value problems.
- In concrete situations the error identity (6) can be converted into an a posteriori error estimates of the form
\[
J(u, q) - J(u_h, q_h) \approx \eta(u_h) := \sum_{K \in \mathcal{T}_h} \rho_K \omega_K,
\]
with cell residuals \( \rho_K \) and cell weights \( \omega_K \) (influence factors). The cubic remainder term \( R_h \) is usually neglected. The evaluation of these estimates requires guesses for the continuous solution \( \{u, q, \lambda\} \) which are obtained from \( \{u_h, q_h, \lambda_h\} \) by post-processing. The cell weights \( \omega_K \) are approximated by local post-processing of the approximate solutions \( x_h \), e.g., by patch-wise higher-order interpolation,
\[
(x - i_h x)|_K \approx (i_{2h}^{(2)} x_h - x_h)|_K,
\]
where the information from the piecewise bilinear function \( x_h = \{u_h, q_h, \lambda_h\} \) is used to define a piecewise biquadratic interpolant \( i_{2h}^{(2)} x_h \) on a coarser mesh.
- On the basis of the resulting approximate error indicators \( \tilde{\eta}_K := \rho_K \tilde{\omega}_K \) the mesh adaptation is oriented towards ‘error balancing’,
\[
\eta := \sum_{K \in \mathcal{T}_h} \eta_K, \quad \eta_K \approx \frac{TOL}{N}, \quad N := \# \{K \in \mathcal{T}_h\} \quad \Rightarrow \quad \eta \approx TOL.
\]
In practice this process may be organized as follows: Determine the 30% cells with largest and the 10% cells with smallest values of \( \eta_K \). The cells of the first group are refined and those of the second group coarsened. In two dimensions this strategy leads to about a doubling of the number of cells in each refinement cycle. By a similar strategy it can be achieved that the number of cells stays about constant during the adaptation process within a time stepping procedure.
The systematic mesh adaptation based on the a posteriori error representation (6) allows the computation of approximate optimal controls \( q_{h}^{opt} \) on rather coarse meshes, while the corresponding discrete states \( u_{h}^{opt} \) may lack admissibility, though the target value of the cost functional is well achieved. In order to recover more admissible states, one may compute a better approximation \( u_{h}^{opt} \) to \( u^{opt} \) from the optimal control \( q_{h} \) by solving the state equation again in a larger space \( V_{H} \supset V_{h} \),

\[
a(u_{h}^{opt}, q_{h}^{opt})(\psi_{H}) = 0 \quad \forall \psi_{H} \in V_{H}.
\]

The natural choice of the cost functional \( J(\cdot, \cdot) \) for error control may not be appropriate in parameter identification problems, in which the error in the control itself is more interesting. This requires a more refined approach as will be discussed below.

In many model situations the stability properties of the state equation allow for coercivity estimates for the KKT system of the form

\[
\sup_{y \in X} \frac{L'(x)(y)}{\|y\|_{X}} \geq \beta \|y\|_{X}, \quad y \in X.
\]

However, usually the stability constant \( \beta \) depends on the regularization parameter like \( \beta \sim \alpha \) and is worst-case oriented. This coercivity allows for the derivation of likewise \( \alpha \)-dependent a posteriori error estimates

\[
\|x - x_{h}\|_{X} \leq \gamma(\alpha) \sup_{y \in X} \frac{L'(x_{h})(y)}{\|y\|_{X}}, \quad \gamma(\alpha) \sim \alpha^{-1} \gg 1,
\]

which will be referred to as ‘energy-norm-type’ error estimates. They may lead to crude overestimation of the error in the cost functional and to less economical meshes compared to those obtained by the ‘weighted’ a posteriori error estimates of the DWR approach. Energy-norm-type a priori and a posteriori error estimates have been derived for various types of optimal control problems, e.g., in Falk [25], Neittaanmäki/Tai [39], Kärkkäinen [33], Liu/Yan [36, 37, 38], Rösch [42], Dahmen/Kunoth [22], Gaevskaia/Hoppe/Repin [28], Kunoth [34], and Rannacher/Vexler [41].

### 2.1 Case of parameter estimation

We consider the model problem

\[
-\Delta u + qu = f \quad \text{in } \Omega, \quad u_{|\partial\Omega} = 0.
\]

The goal is to determine the coefficient \( q \) by comparing the resulting state \( u(q) \) to given measurements \( \bar{C} \),

\[
J(u, q) := \frac{1}{2} \|C(u) - \bar{C}\|_{Z}^{2} + \frac{1}{2} \alpha \|q\|_{Q}^{2} \rightarrow \min! \quad (0 < \alpha \ll 1).
\]

Here, \( \|\cdot\|_{Z} \) and \( \|\cdot\|_{Q} \) denote appropriate norms for observables and controls, respectively. The first-order optimality system reads

\[
-\langle \varphi, C(u) - \bar{C} \rangle_{Z} + \langle \nabla \varphi, \nabla \lambda \rangle + \langle q \varphi, \lambda \rangle = 0 \quad \forall \varphi,
\]

\[
\alpha \langle \chi, q \rangle_{Q} + \langle \chi u, \lambda \rangle = 0 \quad \forall \chi,
\]

\[
(\nabla u, \nabla \psi) + \langle qu, \psi \rangle - \langle f, \psi \rangle = 0 \quad \forall \psi.
\]
If the parameter \( q > 0 \) is identifiable, there holds
\[
-\Delta \lambda + q \lambda = C(u) - \bar{C} = 0, \tag{10}
\]
and consequently \( \lambda \equiv 0 \). This would imply that in an a posteriori error estimate for the least-squares functional \( J(\cdot, \cdot) \) the weights vanish and no mesh refinement would be induced. This demonstrates that the (artificial) least-squares cost functional in parameter estimation is not appropriate for error-based mesh design.

An energy-norm-type a posteriori error estimate for \( |q - q_h|^2 \) can be derived based on a coercivity estimates for the saddle-point problem, see Liu/Yan [36]. However, the stability constant in this estimate blows up with shrinking stabilization, \( \alpha \to 0 \). Within the framework of the DWR methods a more useful a posteriori error estimate for \( q - q_h \) can be derived by an ‘outer’ duality argument applied to the approximation of the full KKT system. From Becker/Vexler [17], we recall the following result.

**Theorem 2.** Let \( \{u, q\} \) be a local solution of the optimization problem and \( \{u_h, q_h\} \) its finite element approximation. Further, let \( G \) be the derivative of the solution operator \( u = S(q) \) with respect to \( q \). For a prescribed control error functional \( E(\cdot) \), let \( z \in V \) be the corresponding dual solution determined by the problem
\[
a'_u(u, q)(\varphi, z) = -\langle G(G^* G)^{-1} \nabla E(q), C'(u)(\varphi) \rangle_Z \quad \forall \varphi \in V. \tag{11}
\]
Then, there holds the a posteriori error representation
\[
E(q) - E(q_h) = \frac{1}{2} \rho(z - i_h z) + \frac{1}{2} \rho^*(u - i_h u) + \mathcal{P}_h + \mathcal{R}_h, \tag{12}
\]
for arbitrary \( i_h u, i_h \lambda \in V_h \), with the discretization residuals
\[
\rho(\psi) := -a(u_h, q_h)(\psi), \quad \rho^*(\varphi) := \langle G_h (G_h^* G_h)^{-1} \nabla E(q_h), C'(u_h)(\varphi) \rangle_Z - a'_u(u_h, q_h)(\varphi, z_h).
\]
The remainder \( \mathcal{P}_h \) is bounded by
\[
|\mathcal{P}_h| \leq c \| e \|_V \| C(u) - \bar{C} \|_Z, \tag{13}
\]
and the linearization remainder \( \mathcal{R}_h \) is again cubic in the error \( e := \{e^u, e^q, e^z\} \).

Notice that the choice \( E(\chi) := \| q - \chi \|_Q \) results in an a posteriori estimate for the error norm \( \| q - q_h \|_Q \).

### 3 Special optimization problems

#### 3.1 A stationary control problem

We consider ‘Neumann boundary control’ of the stationary semi-linear diffusion problem (Kapp [32] and Becker/Kapp/Rannacher [11])
\[
-\Delta u + s(u) = f \quad \text{in} \quad \Omega \subset \mathbb{R}^2, \quad \partial_n u|_{\Gamma_N} = 0, \quad \partial_n u|_{\Gamma_C} = q, \tag{14}
\]
with \( s(u) = u^3 \). The control \( q \) on the control boundary \( \Gamma_C \) is to be determined, such that the resulting observation \( u|_{\Gamma_O} \) in the least-squares sense matches a prescribed profile \( \bar{u} \) on the observation boundary \( \Gamma_O \),

\[
J(u, q) = \frac{1}{2} \| u - \bar{u} \|_{\Gamma_O}^2 + \alpha \frac{1}{2} \| q \|_{\Gamma_C}^2 \rightarrow \min \quad (\bar{u} \equiv 1, \; \alpha \geq 0)
\]

We consider two different configurations depicted in Figure 1. Configuration 1 represents the extreme case that control and observation boundary coincide, while in Configuration 2 they lie opposite to each other, such that the information has to be passed through the entire domain and the presence of the reentrant corners will be felt. Hence, we expect rather different structures of the optimally adapted meshes in the two cases.

\[
\text{Figure 1: Configuration 1 (left) and Configuration 2 (right)}
\]

The state equation in variational form reads

\[
(\nabla u, \nabla \psi) + (s(u), \psi) - (q, \psi|_{\Gamma_C}) = (f, \psi) \quad \forall \psi \in V,
\]

(15)

where \((\cdot, \cdot)\) denotes the \(L^2\) scalar product on \(\Omega\). Applying the Lagrangian formalism, we obtain the corresponding KKT system

\[
(\varphi, u - \bar{u})|_{\Gamma_O} + (\nabla \varphi, \nabla \lambda) + (s'(u) \varphi, \lambda) = 0 \quad \forall \varphi \in V,
\]

\[
\alpha(q, \chi|_{\Gamma_C}) - (\lambda, \chi|_{\Gamma_C}) = 0 \quad \forall \chi \in Q,
\]

(16)

or in its strong form

\[
-\Delta \lambda + s'(u) \lambda = 0 \quad \text{in} \; \Omega, \quad \partial_n \lambda|_{\Gamma_N \cup \Gamma_C \setminus \Gamma_O} = 0, \quad \partial_n \lambda|_{\Gamma_O} = u - \bar{u},
\]

\[
aq - \lambda|_{\Gamma_C} = 0,
\]

\[
-\Delta u + s(u) = 0 \quad \text{in} \; \Omega, \quad \partial_n u|_{\Gamma_N} = 0, \quad \partial_n u|_{\Gamma_C} = q.
\]

(17)

For the Galerkin approximation, we use ‘bilinear’ finite elements for the state and co-state and piecewise constant elements for the controls. The corresponding spaces are denoted by \(V_h\) and \(Q_h\), respectively, and the discretized KKT system reads

\[
(\varphi_h, u_h - \bar{u})|_{\Gamma_O} + (\nabla \varphi_h, \nabla \lambda_h) + (s'(u_h) \varphi_h, \lambda_h) = 0 \quad \forall \varphi_h \in V_h,
\]

\[
\alpha(q_h, \chi_h|_{\Gamma_C}) - (\lambda_h, \chi_h|_{\Gamma_C}) = 0 \quad \forall \chi_h \in Q_h,
\]

(18)

\[
(\nabla u_h, \nabla \psi_h) + (s(u_h), \psi_h) - (f, \psi_h) - (q_h, \psi_h|_{\Gamma_C}) = 0 \quad \forall \psi_h \in V_h.
\]
Then, the general result of Theorem 1 yields the a posteriori error estimate
\[ |J(u, q) - J(u_h, q_h)| \approx \eta_\omega := \sum_{K \in \mathcal{T}_h} \{ \rho_K^\omega \omega_K^\lambda + \rho_K^q \omega_K^q + \rho_K^\lambda \omega_K^u \}, \tag{19} \]
with the cell-residuals and weights defined by
\[
\begin{align*}
\rho_K^\lambda := & \| R_K^\lambda \|_K + h_K^{-1/2} \| \lambda \|_{\partial K}, \\
\rho_K^q := & h_K^{-1/2} \| q_h \|_{\partial K}, \\
\rho_K^u := & \| R_K^u \|_K + h_K^{-1/2} \| u_h \|_{\partial K}, \\
\omega_K^u := & \| u - i_h u \|_K + h_K^{1/2} \| u - i_h u \|_{\partial K}, \\
\omega_K^q := & h_K^{1/2} \| q - i_h q \|_{\partial K}, \\
\omega_K^\lambda := & \| \lambda - i_h \lambda \|_K + h_K^{1/2} \| \lambda - i_h \lambda \|_{\partial K},
\end{align*}
\]
with arbitrary approximations \( \{ i_h u, i_h q, i_h \lambda \} \in V_h \times Q_h \times V_h \), where \( R_{h|K}^u := f + \Delta u_h - s(u_h) \), \( R_{h|K}^\lambda := \Delta \lambda_h - s'(u_h) \lambda_h \), and
\[
\begin{align*}
r_{h|\Gamma}^\lambda := & \begin{cases} 
\frac{1}{2} [\partial_n u_h], & \text{if } \Gamma \not\subset \partial \Omega, \\
\partial_n u_h, & \text{if } \Gamma \subset \partial \Omega \setminus \Gamma_C,
\end{cases} \\
r_{h|\Gamma} = & \begin{cases} 
\frac{1}{2} [\partial_n \lambda_h], & \text{if } \Gamma \not\subset \partial \Omega, \\
\partial_n \lambda_h, & \text{if } \Gamma \subset \partial \Omega \setminus \Gamma_O,
\end{cases} \\
r_{h|\Gamma}^q := & \begin{cases} 
\lambda_h - \alpha q_h, & \text{if } \Gamma \subset \Gamma_C, \\
0, & \text{if } \Gamma \not\subset \Gamma_C.
\end{cases}
\end{align*}
\]
The mesh adaptation is driven by the cell-error indicators \( \eta_{\omega} := \rho_K^\omega \omega_K^\lambda + \rho_K^q \omega_K^q + \rho_K^\lambda \omega_K^u \). We will compare the performance of this ‘weighted’ error estimator with the following two heuristic ‘energy-norm-type’ error estimators:
\[
\begin{align*}
\eta_{E}^u := c_I \left( \sum_{K \in \mathcal{T}_h} h_K^2 (\rho_K^u)^2 \right)^{1/2}, \\
\eta_{E}^\lambda := c_I \left( \sum_{K \in \mathcal{T}_h} h_K^2 \{(\rho_K^\lambda)^2 + (\rho_K^\lambda)^2\} \right)^{1/2}.
\end{align*}
\]

### 3.1.1 Numerical example

Figure 2 shows the solution and corresponding meshes obtained for Configuration 1 by using the different error estimators. As predicted the ‘weighted’ estimator \( \eta_\omega \) in contrast to the ‘energy-norm-type’ error estimators does not enforce any mesh refinement at the reentrant corners. From the corresponding results in Table 1, we see that in this case the ‘weighted’ error estimator yields by far more economical meshes than the ‘energy-norm-type’ error estimators. Further, the overestimation factor \( I_{\text{eff}} := |J(u - u_h, q - q_h)|/\eta_\omega \) is rather moderate considering the many crude estimations necessary to derive the error estimate (19) from the error representation (6). The corresponding results for Configuration 2 are different. Again the efficiency of the ‘weighted’ error estimator is acceptable, but its superiority over the ‘energy-norm-type’ error estimator is less distinct. This is due to the fact that both estimators must allow the information to pass through the entire domain and consequently have to cope with the effect of the reentrant corners by mesh refinement.
Figure 2: Solutions and meshes obtained by $\eta_{E}^{u,\lambda}$ (left) and $\eta_{\omega}$ (right), for Configuration 1

<table>
<thead>
<tr>
<th>N</th>
<th>$E_{\text{rel}}$</th>
<th>$I_{\text{eff}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>596</td>
<td>2.5e-4</td>
<td>0.34</td>
</tr>
<tr>
<td>1616</td>
<td>2.3e-4</td>
<td>0.81</td>
</tr>
<tr>
<td>5084</td>
<td>8.2e-5</td>
<td>0.46</td>
</tr>
<tr>
<td>8648</td>
<td>4.2e-5</td>
<td>0.29</td>
</tr>
<tr>
<td>15512</td>
<td>3.9e-5</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Table 1: Results for $\eta_{\omega}$ (left), and efficiencies for $\eta_{E}^{u}$, $\eta_{E}^{u,\lambda}$, and $\eta_{\omega}$ (right), for Configuration 1

Figure 3: Solutions and meshes obtained by $\eta_{E}^{u,\lambda}$ (left) and $\eta_{\omega}$ (right), for Configuration 2
<table>
<thead>
<tr>
<th>( N )</th>
<th>( E_{\text{rel}} )</th>
<th>( L_{\text{eff}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>512</td>
<td>9.3e-5</td>
<td>1.32</td>
</tr>
<tr>
<td>15368</td>
<td>8.1e-7</td>
<td>0.56</td>
</tr>
<tr>
<td>27800</td>
<td>4.9e-7</td>
<td>0.35</td>
</tr>
<tr>
<td>57632</td>
<td>2.3e-7</td>
<td>0.42</td>
</tr>
<tr>
<td>197408</td>
<td>4.6e-8</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Table 2: Results for \( \eta_\omega \) (left), and efficiencies for \( \eta_E^\omega \), \( \eta_{E^\omega}^u \), and \( \eta_\omega \) (right), for Configuration 2

### 3.2 A nonstationary model control problem

We consider ‘initial control’ in the nonstationary semi-linear diffusion problem (Becker/Meidner/Vexler [13])

\[
\partial_t u - \Delta u + s(u) = f \quad \text{in} \quad Q_T = \Omega \times (0,T], \quad \partial_n u|_{\partial \Omega} = 0, \quad u|_{t=0} = q. \tag{20}
\]

The control \( q \) is to be determined in \( \Omega \), such that a prescribed profile \( \bar{u} \) is matched in the least-squares sense by the solution at the final time \( T \),

\[
J(u,q) := \frac{1}{2} \|u(\cdot,T) - \bar{u}\|_\Omega^2 + \frac{1}{2} \alpha \|q\|_\Omega^2 \rightarrow \min! \quad (0 < \alpha \ll 1).
\]

The variational formulation of the state equation in the space-time region \( Q_T = \Omega \times (0,T] \) reads

\[
(\partial_t u, \psi) + (\nabla u, \nabla \psi) + (s(u), \psi) = (f, \psi) + (q - u(0), \psi(0))_\Omega \quad \forall \psi \in V, \tag{21}
\]

where \((\cdot,\cdot)\) denotes the \( L^2 \) scalar product on \( Q_T \). Applying the Lagrangian formalism results in the following KKT system in space-time for the triplet \( \{u, q, \lambda\} \):

\[
-(\varphi, \partial_t \lambda) + (\nabla \varphi, \nabla \lambda) + (\varphi, s'(u)\lambda) = (\varphi(T), u(T) - \bar{u} - \lambda(T))_\Omega,
\]

\[
\lambda(0) + \alpha q, \lambda)_\Omega = 0, \tag{22}
\]

\[
(\partial_t u, \psi) + (\nabla u, \nabla \psi) + (s(u), \psi) = (q - u(0), \psi(0))_\Omega,
\]

for all admissible test triplets \( \{\psi, \chi, \varphi\} \). Here, the initial condition is incorporated into the variational formulation. This system reads in strong form:

\[
-\partial_t \lambda - \Delta \lambda + s'(u)\lambda = 0 \quad \text{in} \quad Q_T, \quad \lambda|_{t=T} = u(T) - \bar{u}, \quad \partial_n \lambda|_{\partial \Omega} = 0,
\]

\[
\alpha q - \lambda|_{t=0} = 0, \quad \partial_n \lambda|_{\partial \Omega} = 0, \tag{23}
\]

\[
\partial_t u - \Delta u + s(u) = 0 \quad \text{in} \quad Q_T, \quad u|_{t=0} = q, \quad \partial_n u|_{\partial \Omega} = 0.
\]
For solving this saddle-point system, we consider a space-time finite element Galerkin discretization using the so-called ‘dG(r)’ (‘discontinuous Galerkin’) or ‘cG(r)’ (‘continuous Galerkin’) time discretization method (see, e.g., Eriksson/Johnson/Thomee [23]). For a time grid

\[ 0 = t_0 < ... < t_m < ... < t_M = T, \quad k_m := t_m - t_{m-1}, \]

and a corresponding sequence of spatial meshes \( T^n_h \), we introduce finite element spaces \( V^m_h \) consisting of spatially continuous functions which are cell-wise (isoparametric) bilinear in space and polynomial (constant or linear) in time. Depending on the time discretization used the trial functions may be discontinuous or continuous in time, while the test functions are always discontinuous in time. Then, the two coupled subproblems for the primal state \( u_h \) and the adjoint state \( \lambda_h \) can be written in form of successive time-stepping schemes running from \( t_0 = 0 \) forward to \( t_M = T \) and from \( t_M = T \) backward to \( t_0 = 0 \), respectively. This space-time Galerkin finite element method is very similar to the backward Euler (cG(0) method) and the Crank-Nicolson (cG(1) method) scheme in time together with a time-varying spatial finite element discretization. For more details, we refer to Becker [5] and Bangerth/Rannacher [2].

The main difficulty in solving the space-time KKT problem (23) is due to the enormous storage requirements for evaluating the \( u \)-dependent coefficients of the adjoint equation over the whole time interval. These storage requirements can be reduced by so-called ‘checkpointing’ techniques (Griewank [29]), which will be described in the following:

**Checkpointing algorithm:** First, for numbers \( P, Q \in \mathbb{N}_+ \), with \( M = PQ \), the discrete time values \( \{u^m\} \) of the primal state are computed over the whole time interval \([t_0, t_M]\) and the \( P + 1 \) samples \( \{u^0, u^Q, \ldots, u^{QP}\} \) are stored. Additionally, the \( Q - 1 \) values of \( u^m \) in the last section are stored such that now the adjoint state \( \{\lambda^m\} \) can be computed backward in time in the last section. In the following process the \( Q - 1 \) stored values of the primal state in the last section are no longer needed and their storage can be used for storing the primal state over other time sections. Next, starting from the stored value \( u^{(P-2)Q} \) the primal state is recomputed in the \((P - 1)\)-st section and the resulting \( Q - 1 \) of its values are stored. Then, in turn, the corresponding adjoint state can be calculated backward in time in the \((P - 1)\)-st section. This process is repeated until primal and adjoint state are computed on the first section. Denoting by \( S_1 \) the storage (per level) and by \( W_1 \) the work (number of time steps), we have

\[
S_1(P, Q) = (P + 1) + (Q - 1) = P + Q, \\
W_1(P, Q) = M + (P - 1)(Q - 1) = 2M - P - Q + 1.
\]

(24)

This ‘simple’ checkpointing can be recursively refined to ‘multi-level’ checkpointing. Assuming that \( M = M_0 \cdot \ldots \cdot M_L \), with certain \( M_l \in \mathbb{N}_+ \), the ‘one-level checkpointing’ can be applied for the factorization \( M = PQ \) with \( P = M_0 \) and \( Q = M_1 \cdot \ldots \cdot M_L \), and then recursively to each of the \( P \) sections. For the corresponding memory requirement and total number of forward time steps one finds by induction:

\[
S_L = \sum_{l=0}^{L} (M_l - 1) + 2, \quad W_L = (L + 1)M - \sum_{l=0}^{L} \frac{M}{M_l} + 1.
\]

Hence, for \( L \sim \log_2(M) \) minimal storage \( S_{\text{opt}} \) and work count \( W_{\text{opt}} \) are achieved for \( M_l \sim M^{1/(L+1)}, l = 0, \ldots, L \):

\[
S_{\text{min}} = O\left(\log_2(M)\right), \quad W_{\text{min}} = O\left(M \log_2(M)\right).
\]

(25)
Remark 2. Multi-level checkpointing makes the simultaneous computation of primal and dual states feasible even for fine space-time meshes. However, in order to judge the actual value of this approach, we have to compare the time for performing the additional work with the time needed for loading the primal solution from external memory when needed. Depending on the available hardware employing the checkpointing strategy may be less efficient than simply using fast external hard disk memory.

3.2.1 Numerical example

The model problem (20) is considered with nonlinearity \( s(u) = u^2 \), posed on the space-time region \( Q_T := \Omega \times [0, T] \) where \( \Omega = (0, 1)^3 \subset \mathbb{R}^3 \) and \( [0, T] = [0, 1] \). The initial control is parametrized in the form

\[
u|_{t=0} = g_0 + \sum_{i=1}^{8} q_i g_i,
\]

where the shape functions \( g_i \) are given by \( g_0 = (1 - 2\|x - \tilde{x}\|^3) \), with \( \tilde{x} = (0.5, 0.5, 0.5)^T \), and \( g_i = (1 - 0.5\|x - \tilde{x}_i\|^3) \), with \( \tilde{x}_i \in \{0.2, 0.8\}^3 \). Hence, the control space is \( Q = \mathbb{R}^8 \), while the space for states and co-states is \( V = H^1(\Omega) \) as before. The reference solution is \( \bar{u}(x) = (3 + x_1 + x_2 + x_3)/6 \), and the parameter in the regularization term is set to \( \alpha = 10^{-4} \).

The discretization of this problem uses the cG(1) method in space (trilinear shape functions) and the cG(1) method in time (Crank-Nicolson scheme) with uniform mesh size \( h = 0.0625 \) in space and uniform time step \( k = 0.01 \), resulting in \( N = 4096 \) hexahedral cells and 100 time steps. Table 3 shows the effect on storage requirements compared to work of different forms of multi-checkpointing. By checkpointing the storage requirements can be substantially reduced.

<table>
<thead>
<tr>
<th>Factorization</th>
<th>Memory in MB</th>
<th>Time Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>274</td>
<td>35,000</td>
</tr>
<tr>
<td>5 \cdot 100</td>
<td>58</td>
<td>87,948</td>
</tr>
<tr>
<td>10 \cdot 50</td>
<td>2</td>
<td>90,783</td>
</tr>
<tr>
<td>5 \cdot 10 \cdot 10</td>
<td>13</td>
<td>113,463</td>
</tr>
<tr>
<td>4 \cdot 5 \cdot 5 \cdot 5</td>
<td>9</td>
<td>130,788</td>
</tr>
</tbody>
</table>

Table 3: Reduction of the storage requirement due to multi-level checkpointing with cG(1) discretization and 32768 cells in each time step

3.3 Stationary flow control

We consider a boundary control problem governed by the stationary incompressible Navier-Stokes system (Becker [3, 4], and Becker/Heuveline/Rannacher [9]),

\[-\nu \Delta v + v \cdot \nabla v + \nabla p = 0, \quad \nabla \cdot v = 0, \quad \text{in } \Omega,\]

where the physical unknowns are the velocity and the pressure \( u = \{v, p\} \) for prescribed viscosity \( \nu \), and constant density \( \rho = 1 \). The flow configuration is shown in Figure 4.

12
Figure 4: Configuration of the flow control problem

The boundary conditions are

\[ v|_{\Gamma_{\text{rigid}}} = 0, \quad v|_{\Gamma_{\text{in}}} = v^\text{in}, \quad \nu\partial_n v - np|_{\Gamma_{\text{out}}} = 0 \]

with the ‘rigid’ boundary \( \Gamma_{\text{rigid}} \), the ‘inflow’ boundary \( \Gamma_{\text{in}} \), the ‘outflow’ boundary \( \Gamma_{\text{out}} \), and the ‘control’ boundary \( \Gamma_Q \). The Reynolds number is \( \text{Re} = \frac{U^2 D}{\nu} = 40 \), defined in terms of viscosity \( \nu \), characteristic length \( D = \text{diam(cylinder)} \), and maximal inflow velocity \( \bar{U} \), corresponding to stationary (uncontrolled) flow.

The goal of the optimization is the minimization of the drag coefficient

\[ J_{\text{drag}}(u) := \frac{2}{U D} \int_S \nabla^T \sigma(v, p) e_1 \, ds, \]

taken over the surface \( S \) of the cylinder, with the stress tensor \( \sigma(v, p) = -pI + \nu(\nabla v + \nabla^T v) \), and the unit vector \( e_1 \) of the main flow direction. The piecewise constant control \( q = \{q_1, q_2\} \in \mathbb{R}^2 \) acts at the two components of the control boundary \( \Gamma_Q \) as a Neumann control. The corresponding state equation posed in variational form seeks \( u = \{v, p\} \in \{v^\text{in} + H\} \times L^2(\Omega) \) and \( q \in \mathbb{R}^2 \), such that

\[ \nu(\nabla \psi, \nabla \varphi^v) + (v \cdot \nabla \psi, \varphi^v) - (p, \nabla \cdot \varphi^v) - (\varphi^p, \nabla \cdot v) - (q, n \cdot \varphi^v)_{\Gamma_Q} = 0, \quad (27) \]

for test functions \( \varphi = \{\varphi^v, \varphi^p\} \in V \), with the spaces \( V = H \times L^2(\Omega) \) and \( H = H^1_0(\Omega; \Gamma_{\text{rigid}} \cup \Gamma_{\text{in}})^2 \). In this case the optimality system for the triplet \( \{u, q, \lambda\} \), with \( u = \{v, p\} \) and \( \lambda = \{z, \pi\} \), in variational form reads as follows:

\[ \nu(\nabla \psi^v, \nabla z) + (v \cdot \nabla \psi^v, z) + (\psi^v \cdot \nabla v, z) - (\pi, \nabla \cdot \psi^v) - (\psi^p, \nabla \cdot z) = J_{\text{drag}}(\psi), \]

\[ (r, z \cdot n)_{\Gamma_Q} = 0, \quad (28) \]

for all test triplets \( \{\psi, r, \varphi\} \in V \times \mathbb{R}^2 \times V \). This nonlinear system is discretized by the finite element Galerkin method using (isoparametric) bilinear trial functions (equal-order ‘\( Q_1 \)-Stokes element’) on quadrilateral meshes with consistent least-squares stabilization of pressure-velocity coupling and transport. We omit the explicit statement of the discrete Galerkin equations and the corresponding a posteriori error estimates. The resulting discrete equations are solved by a simplified Newton method.

Table 4 shows the advantages of using mesh adaptation in this optimal control model problem. The same minimal drag value can be reached on adapted meshes which have only about
6% of the cells of corresponding uniform meshes. Figure 5 shows the streamline plots for the uncontrolled flow \((q = 0)\) and the optimally controlled flow together with the corresponding adapted ‘control mesh’. The controlled flow has been computed by post-processing as described above on a globally refined mesh. The structure of this (stationary) flow pattern indicates that it may not be dynamically stable and hence not physically relevant. In fact, an accompanying linear stability analysis confirms this expectation (Heuveline/Rannacher [31]). The results of this test case have been obtained using the finite element toolkit GASCOIGNE documented in Becker/Braack [6].

<table>
<thead>
<tr>
<th>Uniform refinement</th>
<th>Adaptive refinement</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N)</td>
<td>(J_{\text{drag}})</td>
</tr>
<tr>
<td>10512</td>
<td>3.31321</td>
</tr>
<tr>
<td>41504</td>
<td>3.21096</td>
</tr>
<tr>
<td>164928</td>
<td>3.11800</td>
</tr>
</tbody>
</table>

Table 4: Uniform refinement versus adaptive refinement

Figure 5: Velocity of the uncontrolled flow (top), controlled flow (middle), corresponding adapted mesh (bottom)
4 Parameter estimation

4.1 Example: Fitting of reaction parameters

We consider the scalar reaction-diffusion model problem (Becker/Vexler [17], Vexler [43]),

\[ \beta \cdot \nabla u - \mu \Delta u + f(u) = 0 \quad \text{in } \Omega, \]
\[ u|_{\Gamma_{in}} = \hat{u}, \quad \partial_n u|_{\partial \Omega \setminus \Gamma_{in}} = 0, \]

(29)
on the domain depicted in Figure 6. The reaction term has the form of an Arrhenius law,

\[ f(u) = A \exp \left( - \frac{E}{1-u} \right) u(1-u). \]

(30)
The goal is to determine the parameters \( A \) and \( E \) from 'measured' line averages

\[ \int_{\Gamma_i} u \, ds, \quad i = 1, \ldots, 10. \]

The results obtained by using the DWR approach for mesh adaptation are shown in Figure 6 and Figure 7. The clear superiority of systematic local mesh adaptation over uniform mesh refinement is seen in Figure 8. The results in this section have been obtained using the C++ library RODOBO documented in Becker/Meidner/Vexler [12].

Figure 6: Configuration of the reaction-diffusion problem, measurements are vertical line-integrals of concentration (left), initial solution for \( A = 54.6, E = 0.15 \) (right).

Figure 7: Estimated solution for \( A = 992.3, E = 0.07 \) (left) and refined mesh (right).
4.2 Model calibration

We consider a reactive flow problem governed by the full set of stationary conservation equations for mass, momentum, energy and species concentrations, the ‘compressible’ Navier-Stokes equations (Becker/Vexler [17] and Vexler [43])

\[
\nabla \cdot (\rho u) = 0, \\
(\rho u \cdot \nabla) u + \nabla \cdot \tau + \nabla p = 0, \\
\rho u \cdot \nabla T - c_p^{-1} \nabla \cdot Q = f_T, \\
\rho u \cdot \nabla y_k + \nabla \cdot F_k = f_k, \quad k = 1, \ldots, 9,
\]

with an accurate diffusion model of ‘multi-component diffusion’ type,

\[
F_k = q_k D_k^* \nabla y_k, \quad D_k^* = (1 - y_k) \left( \sum_{l \neq k} \frac{x_l}{D_{kl}^\text{bin}} \right)^{-1}.
\]

The set-up of this hydrogen diffusion flame is taken from Braack/Ern [19] and is shown schematically in Figure 9. At the inflow boundary of the center tube, 10% mass fraction of hydrogen \( y_{O_2} \) and 90% of nitrogen \( y_{N_2} \) is prescribed. The inflow temperature is \( T = 273 \) K. Along the upper and lower boundary the constant values \( y_{O_2} = 0.22 \) and \( y_{N_2} = 0.78 \) are given. The peak velocity of the three parabolic velocity profiles is 1 m/s.

The numerical evaluation of the ‘multi-component diffusion’ model (32) is rather expensive since it requires the inversion of linear systems in all nodal points. Therefore, it is desirably to replace this complicated model by the much simpler ‘Fick’s model’,

\[
F_k = q_k D_k \nabla y_k,
\]

with diffusion coefficients independent of the other species. The goal is now to choose these coefficients \( D_k \) in such a way that the resulting solution is close to that obtained by using the full
multi-component diffusion. This parameter identification is done on the basis of given measured point-values of species concentrations. Since the case of pointwise data is not directly covered by our Hilbert-space-based theory, we may replace point-values by certain local averages. An a priori error analysis of the discretization of point-value-based parameter estimation problems is given in Rannacher/Vexler [41].

Figure 10 shows the result of a successful adaptation measured in terms of the position of the combustion front which is usually a very sensitive parameter. Figure 11 shows zooms into automatically adapted meshes.

Figure 9: Configuration of the hydrogen diffusion flame

Figure 10: Multicomponent diffusion reference solution (left), Fick’s law as initial parameters (middle), Fick’s law with estimated parameters (right)

Figure 11: Locally refined meshes (zooms)
4.3 Flow stabilization

We consider a nonstationary optimal control problem in a high-temperature flow reactor (Carraro [21]). The goal is the determination of the reaction rate

\[ k(T) = A \left( \frac{T}{300K} \right)^\beta \exp \left( - \frac{E_a}{RT} \right) \]

of the elementary chemical reaction \( \text{O}^{(1D)} + \text{H}_2 \rightarrow \text{OH} + \text{H} \) at temperature \( T = 780 \text{ K} \). This reaction is important for the understanding of atmospheric chemistry. Until recently its rate has been known only for room temperature \( T \sim 300 \text{ K} \). The reaction model involves 7 species and 6 elementary reactions, with argon Ar as inert gas. By laser-induced fluorescence (LIF) techniques the time-dependent decay of local average values of H-concentration can be measured in the reactor. The result of the parameter identification process is the significantly increased value of \( k = 1.5 \cdot 10^{-10} \text{[cm}^3\text{mol}^{-1}\text{s}^{-1}] \), for \( T = 780 \text{K} \), over the known value \( k = 1.0 \cdot 10^{-10} \text{[cm}^3\text{mol}^{-1}\text{s}^{-1}] \), for \( T = 300 \text{K} \).

The first step in preparing for the simulation of the flow in the reactor is the determination of a complete set of initial and boundary conditions for which the model can be expected to be well-posed. For the initial concentration of the species, their partial pressures are calculated from the values of their fluxes and the total pressure, and from the partial pressures the initial number of molecules per \( m^3 \) is obtained. The inflow profiles are calculated from the geometry of the inflow and the fluxes measured experimentally by the flux controllers. The outflow boundary conditions, the no-slip condition at rigid walls and the symmetry condition along the cylinder axis are imposed as usual. Only the temperature boundary condition at the outer wall is not so clearly defined. This is a typical difficulty for the simulation of real-life experiments. The reactor is heated at the exterior wall in order to achieve the desired high temperature in the interior measurement area. However, due to constructional constraints the exact temperature distribution along the outer wall cannot be directly measured and is actually nonuniform due to the cold inflow at the upper inlets. The only information available is the temperature profile along the inner reactor axis which is obtained from thermo-sensor measurements. This is taken as input data for the implicit determination of the unknown temperature boundary data by a parameter estimation process. This model calibration has been performed prior to the actual parameter identification.

Figure 12 shows the twodimensional geometry of the flow reactor used for the model calibration and afterwards for the parameter identification together with the computed temperature distribution and an adapted mesh. Figure 13 shows the result of the parameter identification versus the set of measurements. The results in this section have been obtained using the Finite Element Package HiFlow documented in Heuveline [30].
Figure 12: Geometry sketch of the experimental setup (left), computed temperature distribution with $780\,K$ in the measurement area (middle), adapted mesh for the flow simulation (right)

Figure 13: Computed $H$ concentrations using estimated parameters versus experimental data at room temperature 300$K$ (left) and at higher temperature 780$K$ (right)

The design of the reactor is largely based on a priori numerical simulation. Especially the question of the achievable temperature in the measurement area is crucial and has been systematically investigated. These computations have shown that at high temperature and ‘high’ pressure (atmospheric pressure) the inclusion of gravity may result in persisting nonstationary flow behavior; see Figure 14. This has enforced the restriction of the experiment to low pres-
sure conditions (15 – 50 Torr). In order to do controlled measurements in the reactor at high temperature and atmospheric pressure the flow needs to be stabilized. For suppressing this nonstationary behavior, one may try to minimize the cost functional

\[
J(u) = \frac{K}{2} \int_0^T \left\{ ||\partial_t v(t)||_{\Omega_{\text{meas}}}^2 + ||\partial_t T(t)||_{\Omega_{\text{meas}}}^2 \right\} dt + \text{‘Regularization’},
\]

where \( \Omega_{\text{meas}} \) is the measurement zone in the reactor and \( u = \{v, p, T, y\} \) is determined by the flow model. It is expected that in this way the LIF measurements can be extended to atmospheric pressure.

Figure 14: A time series of temperature plots in the case of non-zero gravity indicating nonstationary behavior

5 Conclusion and outlook

We have demonstrated that systematic mesh adaptation in solving optimal control problems by the DWR approach has great potential for significantly reducing the computational work. Current and future developments are on the following topics:

- Theoretical convergence proof for the mesh adaptation process.
- Combination of spatial mesh adaptation with the checkpointing technique in solving the space-time KKT system.
- Incorporation of control and state constraints.
- Identification of distributed parameters.
- Parameter identification with uncertainties.
- Shape optimization with fluid-structure interaction.
References


