Author(s): Wolfmayr, Monika

Title: Guaranteed lower bounds for cost functionals of time-periodic parabolic optimization problems

Year: 2020

Version: Accepted version (Final draft)

GUARANTEED LOWER BOUNDS FOR COST FUNCTIONALS OF TIME-PERIODIC PARABOLIC OPTIMIZATION PROBLEMS

MONIKA WOLFMAYR

Abstract. In this paper, a new technique is shown for deriving computable, guaranteed lower bounds of functional type (minorants) for two different cost functionals subject to a parabolic time-periodic boundary value problem. Together with previous results on upper bounds (majorants) for one of the cost functionals, both minorants and majorants lead to two-sided estimates of functional type for the optimal control problem. Both upper and lower bounds are derived for the second new cost functional subject to the same parabolic PDE-constraints, but where the target is a desired gradient. The time-periodic optimal control problems are discretized by the multiharmonic finite element method leading to large systems of linear equations having a saddle point structure. The derivation of preconditioners for the minimal residual method for the new optimization problem is discussed in more detail. Finally, several numerical experiments for both optimal control problems are presented confirming the theoretical results obtained. This work provides the basis for an adaptive scheme for time-periodic optimization problems.

1. Introduction

In this work, we derive fully computable, guaranteed lower bounds (minorants) for cost functionals of parabolic optimization problems with given time-periodic conditions. The fully computable upper bounds (majorants) for one of the cost functionals were derived in [26]. The second cost functional is new and for this one both upper and lower bounds are presented. The motivation for the second problem lies in applications, where the target is the gradient or flux instead of the state function. Lower bounds for cost functionals of time-periodic parabolic optimal control problems have not been discussed yet. However, optimal control problems are highly important for different applications (see e.g. [36] as well as the original work [29]). These applications also include time-periodic problems, see for instance [1] and [15] considering problems in electromagnetics and biochemistry, respectively. For these types of problems the multiharmonic (or harmonic-balanced) finite element method (short MhFEM) is a natural choice. The functions are approximated by truncated Fourier series in time and by the finite element method (FEM) in space – more precisely, the Fourier coefficients. We refer to the application of this discretization technique already in [45] as well as later in [3, 4, 5, 9] for non-linear time-harmonic eddy current problems. Moreover, time-periodic optimal control problems and the MhFEM were discussed in e.g. [16, 17, 27] and [18, 19]. Recent works on robust preconditioners for time-periodic parabolic and eddy-current optimal control problems are discussed in [28] and [2], respectively. In this work, a standard finite element discretization is used with continuous, piecewise linear finite elements and a regular grid as discussed, e.g., in [6, 7, 43]. However, the method is wider applicable using also other finite elements or also finite differences (for instances, if the given domain is geometrically rather non-complex).

Functional a posteriori estimation provides a useful machinery to derive computable and guaranteed quantities for the desired unknown solution, see, e.g., [39, 12] on parabolic problems. Recent works on new estimates for parabolic problems and parabolic optimal control problems can be found in [32] and [42], respectively. A posteriori estimates of functional type for elliptic optimal control problems can be found in [10, 11, 40, 31]. First functional type estimates for inverse problems, which are related to optimal control problems, can be found in [41, 8]. Moreover, recent results on guaranteed computable estimates for convection-dominated diffusion problems

This research was supported by the Academy of Finland, grant 295897.

are presented in [33]. In [26], majorants for one cost functional of a time-periodic parabolic optimal control and for the corresponding optimality system were presented. This work presents the corresponding minorants for this cost functional using the new technique presented in [44], which makes use of ideas derived by Mikhlin [34] but generalized for the class of optimal control problems. We mention here that [30] presents a different approach for the derivation of a lower bound for a class of elliptic optimal control problems.

We extend the analysis in this paper and consider a second cost functional with respect to the same parabolic time-periodic boundary value problem. In the second optimal control problem, the target is a given desired gradient. Problems of that type have been earlier discussed in [11]. The results on computable lower bounds together with the upper bounds lead to two-sided estimates which can be used to derive majorants for the discretization error in state and control. These majorants and minorants provide a new formulation of the optimization problems since they can, in principle, be used as objects of direct minimization on their difference. The majorants and minorants can be used in order to derive an adaptive scheme in time and in space. The linearity of the optimal control problems leads to decoupled problems in the Fourier modes including decoupling in the majorants and minorants, which is shown in this work. The overall estimators provide the modes/mode numbers which are necessary for computations. The problems for the different modes can then be computed on different grids, for which estimators in space can be used with respect to the finite element discretized Fourier coefficients. Altogether we derive a space-time adaptive method. Its idea has been for the first time introduced in [26] and has been called the adaptive multiharmonic finite element method (AMhFEM).

In this work, robust preconditioners for the preconditioned minimal residual (MinRes) method (see [37]) are discussed for the second optimization problem, which are new for this case. Also the practical performance of the AMLI preconditioner MinRes solver is presented in various numerical experiments for both optimization problems. For additional numerical tests regarding the AMLI preconditititioned MinRes solver used in this work and its performance for different cases of given data in time-periodic parabolic optimal control problems, we refer the reader to [27].

The work is arranged in the following sections: In Section 2, the two types of cost functionals are presented including some preliminary results. We denote the problems by optimization problem I and II. Moreover, the former result on the majorant for problem I is summarized there. The new minorant for optimization problem I is derived in Section 3 followed by the discussion of the majorant and minorant for optimization problem II in Section 4. In Section 5, the MhFEM for both optimization problems is presented. Robust preconditioners for applying the preconditioned MinRes method on the problems discretized by the MhFEM are presented in Section 6. Section 7 discusses detailed a set of various numerical experiments for both optimization problems I and II. A few final remarks are drawn in Section 8.

2. Time-periodic parabolic PDE, the two cost functionals and preliminary results

We denote by $\Omega \subset \mathbb{R}^d$ with possible dimensions $d = \{1, 2, 3\}$ the spatial bounded Lipschitz domain with boundary $\Gamma := \partial\Omega$. Also we denote by $Q := \Omega \times (0, T)$ the space-time domain and $\Sigma := \Gamma \times (0, T)$ its lateral surface, where $(0, T)$ is the given time interval. The optimization problems are both subject to the following parabolic PDE with given homogeneous Dirichlet boundary conditions and time-periodical condition:

$$\sigma\, \partial_t y - \nabla \cdot (\nu \nabla y) = u \qquad \text{in } Q, \tag{1}$$

$$y = 0 \qquad \text{on } \Sigma, \tag{2}$$

$$y(0) = y(T) \qquad \text{in } \overline{\Omega}. \tag{3}$$

The function y is the state and u will be the control function. We assume that the coefficient functions σ and ν are positive and bounded satisfying $0 < \underline{\sigma} \leq \sigma(\boldsymbol{x}) \leq \overline{\sigma}$ and $0 < \underline{\nu} \leq \nu(\boldsymbol{x}) \leq \overline{\nu}$ for $\boldsymbol{x} \in \Omega$ with constants $\underline{\sigma}$, $\overline{\sigma}$, $\underline{\nu}$ and $\overline{\nu}$. The time-periodic problems in this paper are motived by real-life applications such as computational electromagnetics, where these parameters correspond to the reluctivity and conductivity being usually piecewise constant because of the various materials of the electrical devices.

2.1. **Preliminaries.** In the following, we present a proper functional space setting for time-periodic problems which starts by defining the Hilbert spaces

$$
\begin{aligned}
&H^{1,0}(Q) = \{u \in L^2(Q) : \nabla u \in [L^2(Q)]^d\},\ H_0^{1,0}(Q) = \{u \in H^{1,0}(Q) : u = 0 \text{ on } \Sigma\},\\
&H^{0,1}(Q) = \{u \in L^2(Q) : \partial_t u \in L^2(Q)\},\ H^{0,1}_{per}(Q) = \{u \in H^{0,1}(Q) : u(0) = u(T) \text{ in } \overline{\Omega}\},\\
&H^{1,1}(Q) = \{u \in L^2(Q) : \nabla u \in [L^2(Q)]^d, \partial_t u \in L^2(Q)\},
\end{aligned}
$$

(see also [23, 24]). For instance, the norm in $H^{1,1}$ is given by

$$
\|u\|_{1,1} := \left(\int_Q \left(u(\boldsymbol{x},t)^2 + |\nabla u(\boldsymbol{x},t)|^2 + |\partial_t u(\boldsymbol{x},t)|^2 \right) d\boldsymbol{x}\, dt \right)^{1/2}.
$$

In the following, we skip the subindex for the norms and inner products in $L^2(Q)$ and denote them by $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$. For $L^2(\Omega)$ and $H^1(\Omega)$, we denote them by $\|\cdot\|_\Omega$ and $\langle\cdot,\cdot\rangle_\Omega$ and $\|\cdot\|_{1,\Omega}$ and $\langle\cdot,\cdot\rangle_{1,\Omega}$, respectively.

Time-periodic functions which are at least from the space L^2 can be naturally represented by Fourier series as

$$
u(\boldsymbol{x},t) = u_0^c(\boldsymbol{x}) + \sum_{k=1}^{\infty} \left(u_k^c(\boldsymbol{x}) \cos(k\omega t) + u_k^s(\boldsymbol{x}) \sin(k\omega t) \right)
$$

for $\omega = 2\pi/T$ being the frequency, T the period and with the Fourier cofficients

$$
u_0^c(\boldsymbol{x}) = \frac{1}{T}\int_0^T u(\boldsymbol{x},t)\, dt,\ u_k^c(\boldsymbol{x}) = \frac{2}{T}\int_0^T u(\boldsymbol{x},t)\cos(k\omega t)\, dt,\ u_k^s(\boldsymbol{x}) = \frac{2}{T}\int_0^T u(\boldsymbol{x},t)\sin(k\omega t)\, dt.
$$

We define the norm in Fourier space as follows

$$
\left\|\partial_t^{1/2} u\right\|^2 := |u|^2_{0,\frac{1}{2}} := \frac{T}{2}\sum_{k=1}^{\infty} k\omega \|\boldsymbol{u}_k\|^2_\Omega \tag{4}
$$

as well as the spaces $H^{0,\frac{1}{2}}_{per}(Q) = \{u \in L^2(Q) : \|\partial_t^{1/2} u\| < \infty\}$, $H^{1,\frac{1}{2}}_{per}(Q) = \{u \in H^{1,0}(Q) : \|\partial_t^{1/2} u\| < \infty\}$, $H^{1,\frac{1}{2}}_{0,per}(Q) = \{u \in H^{1,\frac{1}{2}}_{per}(Q) : u = 0 \text{ on } \Sigma\}$, where $\boldsymbol{u}_k = (u_k^c, u_k^s)^T$, $k \in \mathbb{N}$. We also introduce the orthogonal vector $\boldsymbol{u}_k^\perp = (-u_k^s, u_k^c)^T$. The inner products (including also a σ-weighted version) in these spaces are defined by $\langle \partial_t^{1/2} u, \partial_t^{1/2} v\rangle := \frac{T}{2}\sum_{k=1}^\infty k\omega \langle \boldsymbol{u}_k, \boldsymbol{v}_k\rangle_\Omega$, and $\langle \sigma\partial_t^{1/2} u, \partial_t^{1/2} v\rangle := \frac{T}{2}\sum_{k=1}^\infty k\omega \langle \sigma\boldsymbol{u}_k, \boldsymbol{v}_k\rangle_\Omega$. The $H^{1,\frac{1}{2}}_{per}(Q)$-seminorm is defined by

$$
|u|^2_{1,\frac{1}{2}} = T\,\|\nabla u_0^c\|^2_\Omega + \frac{T}{2}\sum_{k=1}^{\infty}\left(k\omega\|\boldsymbol{u}_k\|^2_\Omega + \|\nabla \boldsymbol{u}_k\|^2_\Omega \right) = \|\nabla u\|^2 + \|\partial_t^{1/2} u\|^2
$$

and the corresponding norm is $\|u\|^2_{1,\frac{1}{2}} := \|u\|^2 + |u|^2_{1,\frac{1}{2}}$. Finally, we also define the product $\langle \kappa, \partial_t^{1/2} u\rangle := \frac{T}{2}\sum_{k=1}^\infty (k\omega)^{1/2}\langle \boldsymbol{\kappa}_k, \boldsymbol{u}_k\rangle_\Omega$ as well as the orthogonal Fourier series

$$
u^\perp(\boldsymbol{x},t) := \sum_{k=1}^{\infty} (-\boldsymbol{u}_k^\perp)^T \cdot (\cos(k\omega t), \sin(k\omega t))^T.
$$

Using this notation we can prove that $(u^\perp)^\perp =$ -u, $\|u^\perp\| = \|u\|$ and $\|\partial_t^{1/2} u^\perp\| = \|\partial_t^{1/2} u\|$ for all $u \in H^{0,\frac{1}{2}}_{per}(Q)$ and also that $\|\boldsymbol{u}_k^\perp\|_\Omega = \|\boldsymbol{u}_k\|_\Omega$. Briefly, we recall from [25], the following identities

$$
\langle \sigma\partial_t^{1/2} u, \partial_t^{1/2} v\rangle = \langle \sigma\partial_t u, v^\perp\rangle,\ \langle \sigma\partial_t^{1/2} u, \partial_t^{1/2} v^\perp\rangle = \langle \sigma\partial_t u, v\rangle, \quad \forall u \in H^{0,1}_{per}(Q)\,\forall v \in H^{0,\frac{1}{2}}_{per}(Q), \tag{5}
$$

as well as orthogonality relations

$$
\langle \sigma\partial_t u, u\rangle = 0 \text{ and } \langle \sigma u^\perp, u\rangle = 0 \qquad \forall\, u \in H^{0,1}_{per}(Q), \tag{6}
$$

$$
\langle \sigma\partial_t^{1/2} u, \partial_t^{1/2} u^\perp\rangle = 0 \text{ and } \langle \nu\nabla u, \nabla u^\perp\rangle = 0 \qquad \forall\, u \in H^{1,\frac{1}{2}}_{per}(Q). \tag{7}
$$

Note that the following identity is valid (in Fourier series sense)

$$\int_Q \kappa\, \partial_t^{1/2} u^\perp \, d\boldsymbol{x}\, dt = -\int_Q \partial_t^{1/2} \kappa^\perp \, u \, d\boldsymbol{x}\, dt \qquad \forall\, \kappa, u \in H_{per}^{0,\frac{1}{2}}(Q). \tag{8}$$

Friedrichs' inequality in Q can be proved by using standard Friedrichs' inequality on the Fourier coefficients with respect to the spatial domain Ω. We have that $\|\nabla u\|^2 \geq \frac{1}{C_F^2}\|u\|^2$, where $C_F > 0$ is Friedrichs' constant.

In the following, the parameter $\lambda > 0$ denotes the regularization or cost parameter.

2.2. **Optimization problem I.** In the first case, we want to minimize the following cost functional with respect to the unknown state y and control u:

$$\mathcal{J}(y,u) := \frac{1}{2}\|y - y_d\|^2 + \frac{\lambda}{2}\|u\|^2 \tag{9}$$

subject to the time-periodic boundary value problem (1)–(3). The given desired state $y_d \in L^2(Q)$ does not have to be time-periodic. It only has to be from the space $L^2(Q)$. The cost functional $\mathcal{J}$ defined in (9) can be written as

$$\mathcal{J}(y,u) = T\mathcal{J}_0(y_0^c, u_0^c) + \frac{T}{2}\sum_{k=1}^{\infty} \mathcal{J}_k(\boldsymbol{y}_k, \boldsymbol{u}_k),$$

where $\mathcal{J}_0(y_0^c, u_0^c) = \frac{1}{2}\|y_0^c - y_{d0}^c\|_\Omega^2 + \frac{\lambda}{2}\|u_0^c\|_\Omega^2$ and $\mathcal{J}_k(\boldsymbol{y}_k, \boldsymbol{u}_k) = \frac{1}{2}\|\boldsymbol{y}_k - \boldsymbol{y}_{dk}\|_\Omega^2 + \frac{\lambda}{2}\|\boldsymbol{u}_k\|_\Omega^2$. In [26], the corresponding optimality system is derived, which is given in weak formulation as follows: Given $y_d \in L^2(Q)$, find $y, p \in H_{0,per}^{1,\frac{1}{2}}(Q)$ such that

$$\int_Q \left(y\, z - \nu \nabla p \cdot \nabla z + \sigma \partial_t^{1/2} p\, \partial_t^{1/2} z^\perp \right) d\boldsymbol{x}\, dt = \int_Q y_d\, z \, d\boldsymbol{x}\, dt, \qquad \forall z \in H_{0,per}^{1,\frac{1}{2}}(Q), \tag{10}$$

$$\int_Q \left(\nu \nabla y \cdot \nabla q + \sigma \partial_t^{1/2} y\, \partial_t^{1/2} q^\perp + \lambda^{-1} p\, q \right) d\boldsymbol{x}\, dt = 0, \qquad \forall q \in H_{0,per}^{1,\frac{1}{2}}(Q). \tag{11}$$

The reduced optimality system (10)–(11) i.a. results from using the condition $u = -\lambda^{-1}p$, since no box constraints are imposed on the control function u in this paper. This also leads to the space of admissible controls being given by $H_{0,per}^{1,\frac{1}{2}}(Q)$.

Remark 1. *Since the optimality system is derived first, and then the discretization is performed later, we can say that the* first optimize, then discretize *approach is applied here (as discussed, e.g., earlier in* [14]*).*

Let us define $V := H_0^1(\Omega)$ and $\mathbb{V} := V \times V$. Expanding all functions into Fourier series in (10)–(11) together with using the orthogonality of the cosine and sine functions yields the following problems corresponding to the kth and 0th Fourier coefficients and which are decoupled due to the linearity of the optimal control problem: Find $\boldsymbol{y}_k, \boldsymbol{p}_k \in \mathbb{V}$ so that

$$\int_\Omega \left(\boldsymbol{y}_k \cdot \boldsymbol{z}_k - \nu \nabla \boldsymbol{p}_k \cdot \nabla \boldsymbol{z}_k + k\omega\, \sigma \boldsymbol{p}_k \cdot \boldsymbol{z}_k^\perp \right) d\boldsymbol{x} = \int_\Omega \boldsymbol{y_d}_k \cdot \boldsymbol{z}_k \, d\boldsymbol{x}, \quad \forall \boldsymbol{z}_k \in \mathbb{V}, \tag{12}$$

$$\int_\Omega \left(\nu \nabla \boldsymbol{y}_k \cdot \nabla \boldsymbol{q}_k + k\omega\, \sigma \boldsymbol{y}_k \cdot \boldsymbol{q}_k^\perp + \lambda^{-1} \boldsymbol{p}_k \cdot \boldsymbol{q}_k \right) d\boldsymbol{x} = 0, \quad \forall \boldsymbol{q}_k \in \mathbb{V}. \tag{13}$$

and $y_0^c, p_0^c \in V$ so that

$$\int_\Omega \left(y_0^c \cdot z_0^c - \nu \nabla p_0^c \cdot \nabla z_0^c \right) d\boldsymbol{x} = \int_\Omega y_{d0}^c \cdot z_0^c \, d\boldsymbol{x}, \quad \forall z_0^c \in V, \tag{14}$$

$$\int_\Omega \left(\nu \nabla y_0^c \cdot \nabla q_0^c + \lambda^{-1} p_0^c \cdot q_0^c \right) d\boldsymbol{x} = 0, \quad \forall q_0^c \in V. \tag{15}$$

Both problems (12)–(13) and (14)–(15) are uniquely solvable (see [27]).

2.3. **Majorant for cost functional** (9)**.** Here, the results of [26] on upper bounds for optimization problem I are summerized, since they are needed later to derive the two-sided estimate for the cost functional (9), which deepens and extends the a posteriori error analysis for optimization problem I. Let $y = y(v)$ be the corresponding state to an arbitrary control v. The following upper bound can be proved:

$$\mathcal{J}(y(v), v) \leq \mathcal{J}^{\oplus}(\alpha, \beta; \eta, \boldsymbol{\tau}, v) \qquad \forall\, v \in L^2(Q),$$

for arbitrary $\alpha, \beta > 0$, $\eta \in H^{1,1}_{0,per}(Q)$ and

$$\boldsymbol{\tau} \in H(\mathrm{div}, Q) := \{\boldsymbol{\tau} \in [L^2(Q)]^d : \nabla \cdot \boldsymbol{\tau}(\cdot, t) \in L^2(\Omega) \text{ for a.e. } t \in (0, T)\},$$

where, for any $\boldsymbol{\tau} \in H(\mathrm{div}, Q)$, the identity

$$\int_\Omega \nabla \cdot \boldsymbol{\tau}\, w \, d\boldsymbol{x} = -\int_\Omega \boldsymbol{\tau} \cdot \nabla w \, d\boldsymbol{x} \qquad \forall\, w \in V$$

is valid. The guaranteed and fully computable majorant is given by

$$\mathcal{J}^{\oplus}(\alpha, \beta; \eta, \boldsymbol{\tau}, v) := \frac{1+\alpha}{2}\|\eta - y_d\|^2 + \gamma(\|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|^2 + \frac{C_F^2}{\beta}\|\mathcal{R}_1(\eta, \boldsymbol{\tau}, v)\|^2) + \frac{\lambda}{2}\|v\|^2, \tag{16}$$

where $\underline{\mu_1} = \frac{1}{\sqrt{2}} \min\{\underline{\nu}, \underline{\sigma}\}$ and we have set $\gamma := \frac{(1+\alpha)(1+\beta)C_F^2}{2\alpha\underline{\mu_1}^2}$. The parameters $\alpha, \beta > 0$ have been introduced in order to obtain a quadratic functional by applying Young's inequality.

Remark 2. *The arbitrary functions $\eta \in H^{1,1}_{0,per}(Q)$ and $v \in L^2(Q)$ can be taken later as the approximate solutions of the optimal control problem* (9) *subject to* (1)–(3) *and $\boldsymbol{\tau} \in H(\mathrm{div}, Q)$ represents the image of the flux $\nu\nabla\eta$. Note again that* first optimize, then discretize *is applied in this paper, see also Remark 1.*

For the derivation of (16), the following estimate for the approximation error has been used:

$$|y(v) - \eta|_{1,\frac{1}{2}} \leq \frac{1}{\underline{\mu_1}}\left(C_F\,\|\mathcal{R}_1(\eta, \boldsymbol{\tau}, v)\| + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|\right), \tag{17}$$

where

$$\mathcal{R}_1(\eta, \boldsymbol{\tau}, v) := \sigma\partial_t\eta - \nabla\cdot\boldsymbol{\tau} - v \qquad \text{and} \qquad \mathcal{R}_2(\eta, \boldsymbol{\tau}) := \boldsymbol{\tau} - \nu\nabla\eta. \tag{18}$$

The derivation of estimate (17) can be found in [25]. The function $\mathcal{J}^{\oplus}(\alpha, \beta; \eta, \boldsymbol{\tau}, v)$ is a sharp upper bound on $\mathcal{J}(y(v), v)$ for arbitrary but fixed v as well as on the optimal value $\mathcal{J}(y(u), u)$

$$\inf_{\substack{\eta \in H^{1,1}_{0,per}(Q), \boldsymbol{\tau} \in H(\mathrm{div}, Q)\\ v \in L^2(Q), \alpha, \beta > 0}} \mathcal{J}^{\oplus}(\alpha, \beta; \eta, \boldsymbol{\tau}, v) = \mathcal{J}(y(u), u), \tag{19}$$

since the infimum is attained for the optimal control u, its corresponding state $y(u)$ and its exact flux $\nu\nabla y(u)$, and for α going to zero. Therefore, we have the estimate

$$\mathcal{J}(y(u), u) \leq \mathcal{J}^{\oplus}(\alpha, \beta; \eta, \boldsymbol{\tau}, v) \quad \forall\, \eta \in H^{1,1}_{0,per}(Q),\ \boldsymbol{\tau} \in H(\mathrm{div}, Q),\ v \in L^2(Q),\ \alpha, \beta > 0. \tag{20}$$

2.4. **Optimization problem II.** In the second case, we want to minimize the following cost functional with respect to the unknown state y and control u:

$$\tilde{\mathcal{J}}(y, u) := \frac{1}{2}\|\nabla y - \boldsymbol{g}_d\|^2 + \frac{\lambda}{2}\|u\|^2 \tag{21}$$

subject to the time-periodic boundary value problem (1)–(3), where $\boldsymbol{g}_d \in [L^2(Q)]^d$ is the given desired gradient. The optimality system can analogously be derived as for optimization problem I using the Lagrange functional

$$\tilde{\mathcal{L}}(y, u, p) = \tilde{\mathcal{J}}(y, u) - \int_Q (\sigma\,\partial_t y - \nabla\cdot(\nu\nabla y) - u)\, p \, d\boldsymbol{x}\, dt.$$

Optimality equations $\partial_u \tilde{\mathcal{L}}(y,u,p) = 0$ and $\partial_p \tilde{\mathcal{L}}(y,u,p) = 0$ are similar. Equation $\partial_y \tilde{\mathcal{L}}(y,u,p) = 0$ is different. The optimality conditions are given in weak form as follows: Given $\boldsymbol{g}_d \in [L^2(Q)]^d$, find $y, p \in H^{1,\frac{1}{2}}_{0,per}(Q)$ such that

$$\int_Q \left(\nabla y \cdot \nabla z - \nu \nabla p \cdot \nabla z + \sigma \partial_t^{1/2} p \, \partial_t^{1/2} z^\perp \right) d\boldsymbol{x}\, dt = \int_Q \boldsymbol{g}_d \cdot \nabla z \, d\boldsymbol{x}\, dt, \quad \forall z \in H^{1,\frac{1}{2}}_{0,per}(Q), \tag{22}$$

$$\int_Q \left(\nu \nabla y \cdot \nabla q + \sigma \partial_t^{1/2} y \, \partial_t^{1/2} q^\perp + \lambda^{-1} p \, q \right) d\boldsymbol{x}\, dt = 0, \quad \forall q \in H^{1,\frac{1}{2}}_{0,per}(Q). \tag{23}$$

The optimality systems corresponding to every mode k are analogously derived as for optimization problem I (similar to (12)–(13) and (14)–(15)). In Section 4, we will derive new two-sided bounds for optimization problem II.

3. Guaranteed lower bounds leading to two-sided bounds for optimization problem I

In this work, we complement the guaranteed upper bounds for the discretization error in state and control of minimizing cost functional $\mathcal{J}$ defined in (9) subject to (1)–(3). This is done by obtaining fully computable lower bounds for $\mathcal{J}$ following the technique from [44] (derived for elliptic problems) leading to two-sided bounds for the cost functional (9).

3.1. **Minorant for cost functional** (9).

Let $y = y(u)$ be the optimal state corresponding to the optimal control function $u \in L^2(Q)$, which is connected with the adjoint state $p = p(u)$ by the identity $u = -\lambda^{-1} p(u)$. Then $y = y(u)$ is the solution of the variational formulation

$$\int_Q \left(\nu \nabla y \cdot \nabla q + \sigma \partial_t^{1/2} y \, \partial_t^{1/2} q^\perp + \lambda^{-1} p \, q \right) d\boldsymbol{x}\, dt = 0 \qquad \forall\, q \in H^{1,\frac{1}{2}}_{0,per}(Q) \tag{24}$$

of the boundary value problem (1)–(3) (see also (10)–(11)). For any arbitray function $\eta \in H^{1,1}_{0,per}(Q)$, one can obtain that

$$\mathcal{J}(y(u), u) = \frac{1}{2} \|y - \eta\|^2 + \int_Q (y - \eta)\,(\eta - y_d)\, d\boldsymbol{x}\, dt + \frac{1}{2} \|\eta - y_d\|^2 + \frac{\lambda}{2} \|u\|^2.$$

Since $\frac{1}{2}\|y - \eta\|^2 \geq 0$ and using the identity $u = -\lambda^{-1} p(u)$, $\mathcal{J}$ can be estimated from below by

$$\mathcal{J}(y(u), u) = \mathcal{J}(y(u), -\lambda^{-1} p(u)) \geq \frac{1}{2} \|\eta - y_d\|^2 + \frac{1}{2\lambda} \|p\|^2 + \int_Q (y - \eta)\,(\eta - y_d)\, d\boldsymbol{x}\, dt. \tag{25}$$

For $\eta \in H^{1,1}_{0,per}(Q)$, let $p_\eta, \tilde{p}_\eta \in H^{1,\frac{1}{2}}_{0,per}(Q)$ be the solutions to the equations

$$\int_Q \left(\nu \nabla p_\eta \cdot \nabla z - \sigma \partial_t^{1/2} p_\eta \, \partial_t^{1/2} z^\perp \right) d\boldsymbol{x}\, dt = \int_Q (\eta - y_d) z \, d\boldsymbol{x}\, dt, \qquad \forall z \in H^{1,\frac{1}{2}}_{0,per}(Q), \tag{26}$$

$$\int_Q \left(\nu \nabla \eta \cdot \nabla q + \sigma \partial_t^{1/2} \eta \, \partial_t^{1/2} q^\perp + \lambda^{-1} \tilde{p}_\eta \, q \right) d\boldsymbol{x}\, dt = 0, \qquad \forall q \in H^{1,\frac{1}{2}}_{0,per}(Q). \tag{27}$$

Remark 3. *Note that we assumed that $\eta \in H^{1,1}_{0,per}(Q)$ according to the derivation of the majorant, but so far the assumption $\eta \in H^{1,\frac{1}{2}}_{0,per}(Q)$ would be enough.*

Adding and subtracting p_η in (25) together with $\frac{1}{2\lambda}\|p - p_\eta\|^2 \geq 0$ yields the estimate

$$\mathcal{J}(y(u), u) \geq \frac{1}{2} \|\eta - y_d\|^2 + \frac{1}{2\lambda} \|p_\eta\|^2 + \int_Q (y - \eta)\,(\eta - y_d)\, d\boldsymbol{x}\, dt + \int_Q \lambda^{-1}\,(p - p_\eta)\, p_\eta \, d\boldsymbol{x}\, dt.$$

Using equation (26) and identity (8) leads to the estimate

$$
(28)\quad \begin{cases}
\mathcal{J}(y(u),u) \geq \frac{1}{2}\|\eta - y_d\|^2 + \frac{1}{2\lambda}\|p_\eta\|^2 + \int_Q \lambda^{-1}\,(p - p_\eta)\,p_\eta\,d\boldsymbol{x}\,dt \\
\qquad + \int_Q \Big(\nu\nabla p_\eta\cdot\nabla\,(y-\eta) - \sigma\partial_t^{1/2} p_\eta\,\partial_t^{1/2}\,(y-\eta)^{\perp}\Big)d\boldsymbol{x}\,dt \\
\quad = \frac{1}{2}\|\eta - y_d\|^2 + \frac{1}{2\lambda}\|p_\eta\|^2 + \int_Q \lambda^{-1}\,(p - p_\eta)\,p_\eta\,d\boldsymbol{x}\,dt \\
\qquad + \int_Q \Big(\left(\nu\nabla y - \nu\nabla\eta\right)\cdot\nabla p_\eta + \Big(\sigma\partial_t^{1/2} y - \sigma\partial_t^{1/2}\eta\Big)\,\partial_t^{1/2} p_\eta^{\perp}\Big)d\boldsymbol{x}\,dt \\
\quad = \frac{1}{2}\|\eta - y_d\|^2 + \frac{1}{2\lambda}\|p_\eta\|^2 + \int_Q \Big(\nu\nabla y\cdot\nabla p_\eta + \sigma\partial_t^{1/2} y\,\partial_t^{1/2} p_\eta^{\perp} + \lambda^{-1} p\,p_\eta\Big)d\boldsymbol{x}\,dt \\
\qquad - \int_Q \Big(\nu\nabla\eta\cdot\nabla p_\eta + \sigma\partial_t^{1/2}\eta\,\partial_t^{1/2} p_\eta^{\perp} + \lambda^{-1} p_\eta\,p_\eta\Big)d\boldsymbol{x}\,dt.
\end{cases}
$$

By applying equations (24) and (27), it follows that

$$
\mathcal{J}(y(u),u) \geq \frac{1}{2}\|\eta - y_d\|^2 + \frac{1}{2\lambda}\|p_\eta\|^2 + \int_Q \lambda^{-1}(\tilde{p}_\eta - p_\eta)p_\eta\,d\boldsymbol{x}\,dt.
$$

We introduce now the arbitrary function $\zeta \in H^{1,1}_{0,per}(Q)$. Note that at the moment $\zeta \in H^{1,\frac{1}{2}}_{0,per}(Q)$ would be enough, but the higher regularity in time will be needed in another step. This goes along with the higher regularity assumption on η (see Remark 3). Since $\frac{1}{2\lambda}\|p_\eta - \zeta\|^2 \geq 0$, we have that

$$
\mathcal{J}(y(u),u) \geq \frac{1}{2}\|\eta - y_d\|^2 + \frac{1}{2\lambda}\|\zeta\|^2 + \int_Q \lambda^{-1}\left(p_\eta\zeta - \zeta^2 + \tilde{p}_\eta p_\eta - p_\eta^2\right)\,d\boldsymbol{x}\,dt.
$$

Now we add and subtract $\lambda^{-1}\tilde{p}_\eta\zeta$ in the last integral as well as use equation (27) again. Moreover, we exploit the fact that we have assumed that $\eta \in H^{1,1}_{0,per}(Q)$, hence, we can apply the identities (5). Altogether these steps yield the estimate

$$
(29)\quad \begin{aligned}
\mathcal{J}(y(u),u) \geq{}& \frac{1}{2}\|\eta - y_d\|^2 + \frac{1}{2\lambda}\|\zeta\|^2 - \int_Q \left(\nu\nabla\eta\cdot\nabla\zeta + \sigma\partial_t\eta\,\zeta + \lambda^{-1}\zeta^2\right)\,d\boldsymbol{x}\,dt \\
&+ \int_Q \lambda^{-1}(\zeta - p_\eta)(p_\eta - \tilde{p}_\eta)\,d\boldsymbol{x}\,dt.
\end{aligned}
$$

In the following, we need to estimate the last integral of this expression in order to formulate a computable lower bound for the cost functional. For that let us first prove a computable upper bound for the error in the adjoint state, which is presented in the following theorem. Note that here we will need the higher regularity assumption (in time) on ζ.

Theorem 1. *Let $y_d \in L^2(Q)$ be given. Let $p_\eta \in H^{1,\frac{1}{2}}_{0,per}(Q)$ solve (26) for an arbitrary $\eta \in H^{1,1}_{0,per}(Q)$. The following estimate holds:*

$$
(30)\quad \|\nabla(p_\eta - \zeta)\| \leq \frac{1}{\underline{\mu_1}}\left(C_F\|\mathcal{R}_3(\zeta,\boldsymbol{\rho},\eta)\| + \|\mathcal{R}_4(\zeta,\boldsymbol{\rho})\|\right)
$$

for any $\zeta \in H^{1,1}_{0,per}(Q)$, where $\underline{\mu_1} = \frac{1}{\sqrt{2}}\min\{\underline{\nu},\underline{\sigma}\}$, $\mathcal{R}_3(\zeta,\boldsymbol{\rho},\eta) = \eta - y_d + \nabla\cdot\boldsymbol{\rho} + \sigma\partial_t\zeta$ and $\mathcal{R}_4(\zeta,\boldsymbol{\rho}) = \boldsymbol{\rho} - \nu\nabla\zeta$ with $\boldsymbol{\rho} \in H(\mathrm{div}, Q)$ and $C_F > 0$.

Proof. Let us consider the adjoint equation (26). Adding and subtracting $\zeta \in H^{1,1}_{0,per}(Q)$ in the left-hand side of the equation leads to

$$
(31)\quad \begin{aligned}
\int_Q \Big(\nu\nabla(p_\eta - \zeta)\cdot\nabla z - \sigma\partial_t^{1/2}(p_\eta - \zeta)\,\partial_t^{1/2} z^{\perp}\Big)d\boldsymbol{x}\,dt ={}& \int_Q (\eta - y_d)z\,d\boldsymbol{x}\,dt \\
&- \int_Q \nu\nabla\zeta\cdot\nabla z\,d\boldsymbol{x}\,dt + \int_Q \sigma\partial_t^{1/2}\zeta\,\partial_t^{1/2} z^{\perp} d\boldsymbol{x}\,dt.
\end{aligned}
$$

Next we introduce the auxiliary variable $\boldsymbol{\rho} \in H(\mathrm{div}, Q)$. Together with using that $\zeta \in H^{1,1}_{0,per}(Q)$ as well as applying Cauchy–Schwarz' and Friedrichs' inequalities, the following estimate for the right-hand side of (31) can be obtained:

$$\begin{aligned}
&\sup_{0\neq z\in H^{1,\frac{1}{2}}_{0,per}(Q)} \frac{\int_Q (\eta - y_d + \nabla\cdot\boldsymbol{\rho} + \sigma\partial_t\zeta) z \, d\boldsymbol{x}\, dt + \int_Q (\boldsymbol{\rho} - \nu\nabla\zeta)\cdot\nabla z \, d\boldsymbol{x}\, dt}{|z|_{1,\frac{1}{2}}} \\
&\leq \sup_{0\neq z\in H^{1,\frac{1}{2}}_{0,per}(Q)} \frac{\|\eta - y_d + \nabla\cdot\boldsymbol{\rho} + \sigma\partial_t\zeta\|\|z\| + \|\boldsymbol{\rho} - \nu\nabla\zeta\|\|\nabla z\|}{|z|_{1,\frac{1}{2}}} \\
&\leq \sup_{0\neq z\in H^{1,\frac{1}{2}}_{0,per}(Q)} \frac{C_F\|\eta - y_d + \nabla\cdot\boldsymbol{\rho} + \sigma\partial_t\zeta\|\|\nabla z\| + \|\boldsymbol{\rho} - \nu\nabla\zeta\|\|\nabla z\|}{\|\nabla z\|} \\
&\leq C_F\|\eta - y_d + \nabla\cdot\boldsymbol{\rho} + \sigma\partial_t\zeta\| + \|\boldsymbol{\rho} - \nu\nabla\zeta\|.
\end{aligned}$$

Using the boundedness of the coefficients σ and ν, the orthogonality relations (7) and applying that $\|z^\perp\| = \|z\|$, we can prove the estimate from below for the left-hand side of (31), which is

$$\sup_{0\neq z\in H^{1,\frac{1}{2}}_{0,per}(Q)} \frac{\int_Q \left(\nu\nabla(p_\eta - \zeta)\cdot\nabla z - \sigma\partial_t^{1/2}(p_\eta - \zeta)\,\partial_t^{1/2} z^\perp\right) d\boldsymbol{x}\, dt}{|z|_{1,\frac{1}{2}}}.$$

First, we estimate the supremum from below with choosing $z = (p_\eta - \zeta) + (p_\eta - \zeta)^\perp$, for which

$$|z|_{1,\frac{1}{2}} = |(p_\eta - \zeta) + (p_\eta - \zeta)^\perp|_{1,\frac{1}{2}} = \sqrt{2}\,|p_\eta - \zeta|_{1,\frac{1}{2}},$$

using the orthogonality relations (7). Next, applying the second equation in (7) gives

$$\begin{aligned}
\langle\nu\nabla(p_\eta - \zeta), \nabla z\rangle &= \langle\nu\nabla(p_\eta - \zeta), \nabla((p_\eta - \zeta) + (p_\eta - \zeta)^\perp)\rangle \\
&= \langle\nu\nabla(p_\eta - \zeta), \nabla(p_\eta - \zeta)\rangle + \langle\nu\nabla(p_\eta - \zeta), \nabla(p_\eta - \zeta)^\perp\rangle \\
&= \langle\nu\nabla(p_\eta - \zeta), \nabla(p_\eta - \zeta)\rangle,
\end{aligned}$$

and applying the first equation in (7) as well as using the identity $((p_\eta - \zeta)^\perp)^\perp = -(p_\eta - \zeta)$ gives

$$\begin{aligned}
\langle\sigma\partial_t^{1/2}(p_\eta - \zeta), \partial_t^{1/2} z^\perp\rangle &= \langle\sigma\partial_t^{1/2}(p_\eta - \zeta), \partial_t^{1/2}((p_\eta - \zeta) + (p_\eta - \zeta)^\perp)^\perp\rangle \\
&= \langle\sigma\partial_t^{1/2}(p_\eta - \zeta), \partial_t^{1/2}(p_\eta - \zeta)^\perp\rangle + \langle\sigma\partial_t^{1/2}(p_\eta - \zeta), \partial_t^{1/2}((p_\eta - \zeta)^\perp)^\perp\rangle \\
&= -\langle\sigma\partial_t^{1/2}(p_\eta - \zeta), \partial_t^{1/2}(p_\eta - \zeta)\rangle
\end{aligned}$$

leading to the estimate

$$\begin{aligned}
&\sup_{0\neq z\in H^{1,\frac{1}{2}}_{0,per}(Q)} \frac{\int_Q \left(\nu\nabla(p_\eta - \zeta)\cdot\nabla z - \sigma\partial_t^{1/2}(p_\eta - \zeta)\,\partial_t^{1/2} z^\perp\right) d\boldsymbol{x}\, dt}{|z|_{1,\frac{1}{2}}} \\
&\geq \frac{\langle\nu\nabla(p_\eta - \zeta), \nabla(p_\eta - \zeta)\rangle + \langle\sigma\partial_t^{1/2}(p_\eta - \zeta), \partial_t^{1/2}(p_\eta - \zeta)\rangle}{|(p_\eta - \zeta) + (p_\eta - \zeta)^\perp|_{1,\frac{1}{2}}} \\
&= \frac{\langle\nu\nabla(p_\eta - \zeta), \nabla(p_\eta - \zeta)\rangle + \langle\sigma\partial_t^{1/2}(p_\eta - \zeta), \partial_t^{1/2}(p_\eta - \zeta)\rangle}{\sqrt{2}\,|p_\eta - \zeta|_{1,\frac{1}{2}}} \\
&\geq \frac{\underline{\nu}\|\nabla(p_\eta - \zeta)\|^2 + \underline{\sigma}\|\partial_t^{1/2}(p_\eta - \zeta)\|^2}{\sqrt{2}\,|p_\eta - \zeta|_{1,\frac{1}{2}}} \\
&\geq \underline{\mu_1}\,|p_\eta - \zeta|_{1,\frac{1}{2}},
\end{aligned}$$

where $\underline{\mu_1} = \frac{1}{\sqrt{2}}\min\{\underline{\nu}, \underline{\sigma}\}$. Combining now both estimates together with $\|\nabla(p_\eta - \zeta)\| \leq |p_\eta - \zeta|_{1,\frac{1}{2}}$ we finally derive (30). □

Now we have all the tools in order to estimate the last term of (29). We obtain as follows

$$\int_Q \lambda^{-1}(\zeta-p_\eta)(p_\eta-\zeta+\zeta-\tilde{p}_\eta)\,d\boldsymbol{x}\,dt=\int_Q(\lambda^{-1}(\zeta-p_\eta)(p_\eta-\zeta)+\lambda^{-1}(\zeta-p_\eta)(\zeta-\tilde{p}_\eta))\,d\boldsymbol{x}\,dt$$
$$=-\lambda^{-1}\|\zeta-p_\eta\|^2+\int_Q(\lambda^{-1}(\zeta-p_\eta)(\zeta-\tilde{p}_\eta))\,d\boldsymbol{x}\,dt$$
$$=-\lambda^{-1}\|\zeta-p_\eta\|^2+\int_Q\left(\nu\nabla\eta\cdot\nabla(\zeta-p_\eta)+\sigma\partial_t^{1/2}\eta\,\partial_t^{1/2}(\zeta-p_\eta)^{\perp}\right)d\boldsymbol{x}\,dt+\int_Q\lambda^{-1}\zeta(\zeta-p_\eta)\,d\boldsymbol{x}\,dt$$
$$=-\lambda^{-1}\|\zeta-p_\eta\|^2+\int_Q\left(\sigma\partial_t\eta-\nabla\cdot\boldsymbol{\tau}+\lambda^{-1}\zeta\right)(\zeta-p_\eta)\,d\boldsymbol{x}\,dt+\int_Q(\nu\nabla\eta-\boldsymbol{\tau})\cdot\nabla(\zeta-p_\eta)\,d\boldsymbol{x}\,dt$$
$$\geq-\lambda^{-1}C_F^2\|\nabla(\zeta-p_\eta)\|^2-(C_F\|\mathcal{R}_1(\eta,\boldsymbol{\tau},-\lambda^{-1}\zeta)\|+\|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|)\|\nabla(\zeta-p_\eta)\|$$

leading to

$$\int_Q\lambda^{-1}(\zeta-p_\eta)(p_\eta-\zeta+\zeta-\tilde{p}_\eta)\,d\boldsymbol{x}\,dt\geq-\frac{C_F^2}{\underline{\mu_1}^2\lambda}\left(C_F\|\mathcal{R}_3(\zeta,\boldsymbol{\rho},\eta)\|+\|\mathcal{R}_4(\zeta,\boldsymbol{\rho})\|\right)^2$$
$$-\frac{1}{\underline{\mu_1}}(C_F\|\mathcal{R}_1(\eta,\boldsymbol{\tau},-\lambda^{-1}\zeta)\|+\|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|)\left(C_F\|\mathcal{R}_3(\zeta,\boldsymbol{\rho},\eta)\|+\|\mathcal{R}_4(\zeta,\boldsymbol{\rho})\|\right),\quad(32)$$

where $\boldsymbol{\tau},\boldsymbol{\rho}\in H(\mathrm{div},Q)$ and we have used equation (27), relations (6)–(7), Cauchy–Schwarz' and Friedrichs' inequalities, estimate (30) and that $\eta\in H^{1,1}_{0,per}(Q)$.

Finally, we obtain the following estimate from below:

$$\mathcal{J}(y(u),u)\geq\mathcal{J}^{\ominus}(\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho})\qquad\forall\,\eta,\zeta\in H^{1,1}_{0,per}(Q),\boldsymbol{\tau},\boldsymbol{\rho}\in H(\mathrm{div},Q),\quad(33)$$

where the (fully computable) minorant is given by

$$\mathcal{J}^{\ominus}(\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho})=\frac{1}{2}\|\eta-y_d\|^2+\frac{1}{2\lambda}\|\zeta\|^2-\int_Q\left(\nu\nabla\eta\cdot\nabla\zeta+\sigma\partial_t\eta\,\zeta+\lambda^{-1}\zeta^2\right)\,d\boldsymbol{x}\,dt$$
$$-\frac{C_F^2}{\underline{\mu_1}^2\lambda}\left(C_F\|\mathcal{R}_3(\zeta,\boldsymbol{\rho},\eta)\|+\|\mathcal{R}_4(\zeta,\boldsymbol{\rho})\|\right)^2\quad(34)$$
$$-\frac{1}{\underline{\mu_1}}(C_F\|\mathcal{R}_1(\eta,\boldsymbol{\tau},-\lambda^{-1}\zeta)\|+\|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|)\left(C_F\|\mathcal{R}_3(\zeta,\boldsymbol{\rho},\eta)\|+\|\mathcal{R}_4(\zeta,\boldsymbol{\rho})\|\right).$$

Theorem 2. *The supremum of the minorant $\mathcal{J}^{\ominus}$ as defined in* (34) *is attained for the minimum of the cost functional* (9) *subject to* (1)–(3)*, which is equivalent to the optimal value of the optimality system* (10)–(11) *as follows*

$$\sup_{\eta,\zeta\in H^{1,1}_{0,per}(Q),\boldsymbol{\tau},\boldsymbol{\rho}\in H(\mathrm{div},Q)}\mathcal{J}^{\ominus}(\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho})=\mathcal{J}(y(u),u).\quad(35)$$

Proof. The estimate is sharp for the exact control u, $\eta=y(u)$, $\zeta=p(u)$, $\boldsymbol{\tau}=\nu\nabla y(u)$ and $\boldsymbol{\rho}=\nu\nabla p(u)$. Hence, $\mathcal{J}^{\ominus}(y(u),p(u),\nu\nabla y(u)),\nu\nabla p(u))=\frac{1}{2}\|y-y_d\|^2+\frac{1}{2\lambda}\|p\|^2=\frac{1}{2}\|y-y_d\|^2+\frac{\lambda}{2}\|u\|^2=\mathcal{J}(y(u),u)$. □

Remark 4. *It has been shown in* [10] *that if we choose finite dimensional subspaces that are limit dense in the spaces of the exact solution $y(u)$, of its adjoint $p(u)$ and of their fluxes, $\nu\nabla y(u)$ and $\nu\nabla p(u)$, which are $H^{1,1}_{0,per}(Q)$ and $H(\mathrm{div},Q)$, and we choose sequences of the functions $(\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho})$ in these finite dimensional subspaces, so for instance $(\eta_i,\zeta_i,\boldsymbol{\tau}_i,\boldsymbol{\rho}_i)$, then they converge (let $i\to\infty$) to the exact solution, its adjoint and their fluxes. The corresponding majorants $\mathcal{J}_i^{\oplus}$ and minorants $\mathcal{J}_i^{\ominus}$ converge to the exact value of the cost functional $\mathcal{J}$. As shown in* [10]*, majorants can be used in order to produce sequences of state and controls with values of the cost functional being as close to the exact cost functional value as one needs it.*

3.2. **Guaranteed upper bounds for the discretization errors of the control and the state.** Here, we present a posteriori error estimates for control and state measured by the norm $|||u-v|||^2 := \frac{1}{2}\|y(u)-y(v)\|^2 + \frac{\lambda}{2}\|u-v\|^2$. The next theorem was proved for the elliptic case (together with control constraints) in [44]. The norm $|||\cdot|||$ can be represented in terms of the state and the adjoint state (instead of the control), since there are no control constraints imposed. Hence, $u = -\lambda^{-1}p(u)$, $v = -\lambda^{-1}p(v)$, and $|||u-v|||^2 = \frac{1}{2}\|y(u)-y(v)\|^2 + \frac{1}{2\lambda}\|p(u)-p(v)\|^2$.

Theorem 3. *We obtain the following identity for an arbitrary control $v \in L^2(Q)$:*

$$|||u-v|||^2 = \mathcal{J}(y(v),v) - \mathcal{J}(y(u),u). \tag{36}$$

Proof. Together with $u = -\lambda^{-1}p(u)$ and $v = -\lambda^{-1}p(v)$ the difference can be computed as

$$\begin{aligned}
\mathcal{J}(y(v),v) - \mathcal{J}(y(u),u) &= \frac{1}{2}\|y(v)-y_d\|^2 - \frac{1}{2}\|y(u)-y_d\|^2 + \frac{\lambda}{2}\|v\|^2 - \frac{\lambda}{2}\|u\|^2 \\
&= \frac{1}{2}\int_Q (y(v)-y(u)+2y(u)-2y_d)(y(v)-y(u))\,d\boldsymbol{x}\,dt + \frac{\lambda}{2}\int_Q (v-u+2u)(v-u)\,d\boldsymbol{x}\,dt \\
&= \frac{1}{2}\|y(u)-y(v)\|^2 + \int_Q (y(u)-y_d)(y(v)-y(u))\,d\boldsymbol{x}\,dt \\
&\quad + \frac{\lambda}{2}\|u-v\|^2 + \lambda^{-1}\int_Q p(u)(p(v)-p(u))\,d\boldsymbol{x}\,dt.
\end{aligned}$$

Since the adjoint states $p(u), p(v) \in H^{1,1}_{0,per}(Q)$ fulfill (10)–(11) for the corresponding states $y(u)$, $y(v) \in H^{1,1}_{0,per}(Q)$, we obtain

$$\begin{aligned}
\mathcal{J}(y(v),v) - \mathcal{J}(y(u),u) &= \frac{1}{2}\|y(u)-y(v)\|^2 + \frac{\lambda}{2}\|u-v\|^2 + \lambda^{-1}\int_Q p(u)(p(v)-p(u))\,d\boldsymbol{x}\,dt \\
&+ \int_Q (\nu\nabla p(u)(\nabla y(v)-\nabla y(u)) - \sigma\partial_t^{1/2}p(u)\,\partial_t^{1/2}(y(v)-y(u))^{\perp})\,d\boldsymbol{x}\,dt = \frac{1}{2}\|y(u)-y(v)\|^2 \\
&+ \frac{\lambda}{2}\|u-v\|^2 + \lambda^{-1}\int_Q p(u)p(v)\,d\boldsymbol{x}\,dt + \int_Q (\nu\nabla y(v)\cdot\nabla p(u) + \sigma\partial_t^{1/2}y(v)\,\partial_t^{1/2}p(u)^{\perp})\,d\boldsymbol{x}\,dt \\
&= \frac{1}{2}\|y(u)-y(v)\|^2 + \frac{\lambda}{2}\|u-v\|^2.
\end{aligned}$$

This proves now the equality relation (36) by applying the equations (10)–(11) for $(u, y(u), p(u))$ as well as $(v, y(v), p(v))$. □

Using the result of Theorem 3, we can derive the majorant for the discretization errors of control and state in the norm $|||\cdot|||$.

Theorem 4. *The functional*

$$\begin{aligned}
\mathcal{M}^{\oplus}(\alpha,\beta;\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho},v) &= \frac{\alpha}{2}\|\eta-y_d\|^2 + \gamma(\|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|^2 + \frac{C_F^2}{\beta}\|\mathcal{R}_1(\eta,\boldsymbol{\tau},v)\|^2) + \frac{\lambda}{2}\|v\|^2 - \frac{1}{2\lambda}\|\zeta\|^2 \\
&+ \int_Q \left(\nu\nabla\eta\cdot\nabla\zeta + \sigma\partial_t\eta\,\zeta + \lambda^{-1}\zeta^2\right)\,d\boldsymbol{x}\,dt + \frac{C_F^2}{\underline{\mu_1}^2\lambda}\left(C_F\|\mathcal{R}_3(\zeta,\boldsymbol{\rho},\eta)\| + \|\mathcal{R}_4(\zeta,\boldsymbol{\rho})\|\right)^2 \\
&+ \frac{1}{\underline{\mu_1}}\left(C_F\|\mathcal{R}_1(\eta,\boldsymbol{\tau},-\lambda^{-1}\zeta)\| + \|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|\right)\left(C_F\|\mathcal{R}_3(\zeta,\boldsymbol{\rho},\eta)\| + \|\mathcal{R}_4(\zeta,\boldsymbol{\rho})\|\right)
\end{aligned}$$

for an arbitrary $v \in L^2(Q)$, $\eta,\zeta \in H^{1,1}_{0,per}(Q)$, $\boldsymbol{\tau},\boldsymbol{\rho} \in H(\mathrm{div}, Q)$ and $\alpha,\beta > 0$, and where $\underline{\mu_1} = \frac{1}{\sqrt{2}}\min\{\underline{\nu},\underline{\sigma}\}$, is a majorant for the discretization error

$$|||u-v|||^2 \le \mathcal{M}^{\oplus}(\alpha,\beta;\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho},v) = \mathcal{J}^{\oplus}(\alpha,\beta;\eta,\boldsymbol{\tau},v) - \mathcal{J}^{\ominus}(\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho}). \tag{37}$$

Proof. Applying (36) together with (20) and (33) yields estimate (37). □

Proposition 1. *The infimum of the majorant $\mathcal{M}^{\oplus}(\alpha,\beta;\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho},v)$ in (37) is attained for the minimum of the optimization problem I, which is equivalent to the solution of the optimality system (10)–(11) ($v=u$, $\eta=y(u)$, $\zeta=p(u)=-\lambda u$, $\boldsymbol{\tau}=\nu\nabla y(u)$, $\boldsymbol{\rho}=\nu\nabla p(u)$) as follows*

$$\inf_{\substack{\eta,\zeta\in H^{1,1}_{0,per}(Q),\boldsymbol{\tau},\boldsymbol{\rho}\in H(\mathrm{div},Q),\\ v\in L^2(Q),\alpha,\beta>0}} \mathcal{M}^{\oplus}(\alpha,\beta;\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho},v)=0.$$

Proof. The majorant $\mathcal{M}^{\oplus}(\alpha,\beta;y(u),p(u),\nu\nabla y(u),\nu\nabla p(u),u)=\frac{\alpha}{2}\|y(u)-y_d\|^2$ equals zero if α goes to zero. □

The majorant $\mathcal{M}^{\oplus}$ is a sharp, guaranteed and fully computable upper bound for the control-state error in $|||\cdot|||$. However, it overestimates the L^2-norm $|||\cdot|||$ which is of order h^2, since the majorant only decreases with order h. Following the idea from [44] a weighted H^1-norm is introduced decreasing with the same order as the majorant. For that, we define the norm $|||u-v|||_1^2 := \frac{1}{2}\|y(u)-y(v)\|^2+\frac{\lambda{\mu_1}^2}{2C_F^2}|y(u)-y(v)|^2_{1,\frac{1}{2}}$.

Theorem 5. *The following estimate:*

$$|||u-v|||_1^2 \le \mathcal{J}(y(v),v)-\mathcal{J}(y(u),u)+\frac{3\lambda}{4C_F^2}\left(C_F\|\mathcal{R}_1(\eta,\boldsymbol{\tau},v)\|+\|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|\right)^2 \tag{38}$$

is valid for an arbitrary control function $v\in L^2(Q)$ and with $\mathcal{R}_1(\eta,\boldsymbol{\tau},v)$ and $\mathcal{R}_2(\eta,\boldsymbol{\tau})$ defined as in (18).

Proof. Let the parameter $\delta>0$ be arbitrary but fixed. We add and subtract η, apply triangle inequality on $\frac{{\mu_1}^2}{C_F^2\delta}|y(u)-y(v)|^2_{1,\frac{1}{2}}$ and obtain

$$\frac{{\mu_1}^2}{C_F^2\delta}|y(u)-y(v)|^2_{1,\frac{1}{2}} \le \frac{{\mu_1}^2}{2C_F^2\delta}\left(|y(u)-\eta|^2_{1,\frac{1}{2}}+|y(v)-\eta|^2_{1,\frac{1}{2}}\right),$$

where we further add and subtract v. Then we apply triangle inequality two times leading to

$$\begin{aligned}\frac{{\mu_1}^2}{C_F^2\delta}|y(u)-y(v)|^2_{1,\frac{1}{2}} &\le \frac{1}{2C_F^2\delta}\Bigg(\left(\|\boldsymbol{\tau}-\nu\nabla\eta\|+C_F\|\sigma\partial_t\eta-\nabla\cdot\boldsymbol{\tau}-v\|+C_F\|u-v\|\right)^2\\ &\qquad+\left(\|\boldsymbol{\tau}-\nu\nabla\eta\|+C_F\|\sigma\partial_t\eta-\nabla\cdot\boldsymbol{\tau}-v\|\right)^2\Bigg)\\ &\le \frac{3}{2C_F^2\delta}(\|\boldsymbol{\tau}-\nu\nabla\eta\|+C_F\|\sigma\partial_t\eta-\nabla\cdot\boldsymbol{\tau}-v\|)^2+\frac{1}{\delta}\|u-v\|^2.\end{aligned}$$

Together with (36) this yields

$$\begin{aligned}|||u-v|||^2+\frac{{\mu_1}^2}{C_F^2\delta}|y(u)-y(v)|^2_{1,\frac{1}{2}}-\frac{1}{\delta}\|u-v\|^2 = \frac{1}{2}\|y(u)-y(v)\|^2+\frac{{\mu_1}^2}{C_F^2\delta}|y(u)-y(v)|^2_{1,\frac{1}{2}}\\ +\left(\frac{\lambda}{2}-\frac{1}{\delta}\right)\|u-v\|^2 \le \mathcal{J}(y(v),v)-\mathcal{J}(y(u),u)+\frac{3}{2C_F^2\delta}\left(C_F\|\mathcal{R}_1(\eta,\boldsymbol{\tau},v)\|+\|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|\right)^2.\end{aligned}$$

We see that the choice $\delta=2/\lambda$ finally provides estimate (38). □

This theorem directly leads to the following two results presented in Propositions 2 and 3.

Proposition 2. *The following error majorant for any control $v\in L^2(Q)$ is obtained:*

$$\begin{aligned}|||u-v|||_1^2 \le \mathcal{M}_1^{\oplus}(\alpha,\beta;\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho},v) := \mathcal{J}^{\oplus}(\alpha,\beta;\eta,\boldsymbol{\tau},v)-\mathcal{J}^{\ominus}(\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho})\\ +\frac{3\lambda}{4C_F^2}\left(C_F\|\mathcal{R}_1(\eta,\boldsymbol{\tau},v)\|+\|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|\right)^2\end{aligned} \tag{39}$$

with

$$\mathcal{M}_1^{\oplus}(\alpha,\beta;\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho},v) = \frac{\alpha}{2}\|\eta - y_d\|^2 + \gamma(\|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|^2 + \frac{C_F^2}{\beta}\|\mathcal{R}_1(\eta,\boldsymbol{\tau},v)\|^2) + \frac{\lambda}{2}\|v\|^2 - \frac{1}{2\lambda}\|\zeta\|^2$$
$$+ \int_Q \left(\nu\nabla\eta\cdot\nabla\zeta + \sigma\partial_t\eta\,\zeta + \lambda^{-1}\zeta^2\right)\,d\boldsymbol{x}\,dt + \frac{C_F^2}{\underline{\mu_1}^2\lambda}\left(C_F\|\mathcal{R}_3(\zeta,\boldsymbol{\rho},\eta)\| + \|\mathcal{R}_4(\zeta,\boldsymbol{\rho})\|\right)^2$$
$$+ \frac{1}{\underline{\mu_1}}(C_F\|\mathcal{R}_1(\eta,\boldsymbol{\tau},-\lambda^{-1}\zeta)\| + \|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|)\,(C_F\|\mathcal{R}_3(\zeta,\boldsymbol{\rho},\eta)\| + \|\mathcal{R}_4(\zeta,\boldsymbol{\rho})\|)$$
$$+ \frac{3\lambda}{4C_F^2}\left(C_F\|\mathcal{R}_1(\eta,\boldsymbol{\tau},v)\| + \|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|\right)^2$$

where $\underline{\mu_1} = \frac{1}{\sqrt{2}}\min\{\underline{\nu},\underline{\sigma}\}$, $\alpha,\beta > 0$ *as well as arbitrary* $\eta,\zeta \in H^{1,1}_{0,per}(Q)$ *and* $\boldsymbol{\tau},\boldsymbol{\rho} \in H(\mathrm{div},Q)$.

Proposition 3. *The infimum of the majorant* $\mathcal{M}_1^{\oplus}(\alpha,\beta;\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho},v)$ *in* (39) *is attained for the minimum of the optimization problem I being equivalent to solving the optimality system* (10)–(11) ($v = u$, $\eta = y(u)$, $\zeta = p(u) = -\lambda u$, $\boldsymbol{\tau} = \nu\nabla y(u)$, $\boldsymbol{\rho} = \nu\nabla p(u)$) *as follows*

$$\inf_{\substack{\eta,\zeta\in H^{1,1}_{0,per}(Q),\boldsymbol{\tau},\boldsymbol{\rho}\in H(\mathrm{div},Q),\\ v\in L^2(Q),\alpha,\beta>0}} \mathcal{M}_1^{\oplus}(\alpha,\beta;\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho},v) = 0,$$

and α *going to zero.*

4. Two-sided bounds for optimization problem II

In this section, we analogously derive the majorants and minorants for the second cost functional, however, skipping details which are similar to the derivation in the case of optimization problem I.

4.1. **Majorant for cost functional** (21)**.** Adding and subtracting $\nabla\eta$ in the cost functional $\tilde{\mathcal{J}}(y(v),v)$, applying the triangle inequality as well as using the estimate

$$\|\nabla y(v) - \nabla\eta\|^2 \le |y(v) - \eta|^2_{1,\frac{1}{2}} = \|\nabla y(v) - \nabla\eta\|^2 + \|\partial_t^{1/2}y(v) - \partial_t^{1/2}\eta\|^2,$$

we conclude that

$$\tilde{\mathcal{J}}(y(v),v) \le \frac{1}{2}\left(\|\nabla\eta - \boldsymbol{g}_d\| + |y(v)-\eta|_{1,\frac{1}{2}}\right)^2 + \frac{\lambda}{2}\|v\|^2.$$

Together with (17) this leads to the estimate

$$\tilde{\mathcal{J}}(y(v),v) \le \frac{1}{2}\left(\|\nabla\eta - \boldsymbol{g}_d\| + \frac{1}{\underline{\mu_1}}\|\mathcal{R}_2(\eta,\boldsymbol{\tau})\| + \frac{C_F}{\underline{\mu_1}}\|\mathcal{R}_1(\eta,\boldsymbol{\tau},v)\|\right)^2 + \frac{\lambda}{2}\|v\|^2,$$

where again $\underline{\mu_1} = \frac{1}{\sqrt{2}}\min\{\underline{\nu},\underline{\sigma}\}$ as well as $\mathcal{R}_1(\eta,\boldsymbol{\tau},v)$ and $\mathcal{R}_2(\eta,\boldsymbol{\tau})$ are defined as in (18). Finally, introducing parameters $\alpha,\beta > 0$ and applying Young's inequality, we can reformulate the estimate such that the right-hand side is given by a quadratic functional as follows

$$\tilde{\mathcal{J}}(y(v),v) \le \tilde{\mathcal{J}}^{\oplus}(\alpha,\beta;\eta,\boldsymbol{\tau},v) \qquad \forall\, v \in L^2(Q),$$

where

$$\tilde{\mathcal{J}}^{\oplus}(\alpha,\beta;\eta,\boldsymbol{\tau},v) := \frac{1+\alpha}{2}\|\nabla\eta - \boldsymbol{g}_d\|^2 + \gamma(\|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|^2 + \frac{C_F^2}{\beta}\|\mathcal{R}_1(\eta,\boldsymbol{\tau},v)\|^2) + \frac{\lambda}{2}\|v\|^2. \tag{40}$$

The infimum of the majorant (40) is attained for the minimum of the cost functional (21) subject to (1)–(3), which is equivalent to the optimal value of the optimality system (22)–(23). Analogously to (19), we can show that

$$\inf_{\substack{\eta\in H^{1,1}_{0,per}(Q),\boldsymbol{\tau}\in H(\mathrm{div},Q)\\ v\in L^2(Q),\alpha,\beta>0}} \tilde{\mathcal{J}}^{\oplus}(\alpha,\beta;\eta,\boldsymbol{\tau},v) = \tilde{\mathcal{J}}(y(u),u), \tag{41}$$

and that

$$\tilde{\mathcal{J}}(y(u),u) \le \tilde{\mathcal{J}}^{\oplus}(\alpha,\beta;\eta,\boldsymbol{\tau},v) \quad \forall\,\eta \in H^{1,1}_{0,per}(Q),\ \boldsymbol{\tau} \in H(\mathrm{div},Q),\ v \in L^2(Q),\ \alpha,\beta > 0. \tag{42}$$

4.2. **Minorant for cost functional** (21). Let us derive now the minorant. For any $\eta \in H^{1,1}_{0,per}(Q)$, we have that

$$\tilde{\mathcal{J}}(y(v), v) = \frac{1}{2}\|\nabla y - \nabla \eta\|^2 + \int_Q (\nabla y - \nabla \eta) \cdot (\nabla \eta - \boldsymbol{g}_d)\, d\boldsymbol{x}\, dt + \frac{1}{2}\|\nabla \eta - \boldsymbol{g}_d\|^2 + \frac{\lambda}{2}\|v\|^2$$

for all $v \in L^2(Q)$. The first norm is again greater or equal to zero, together with the identity $v = -\lambda^{-1} p(v)$, we can estimate $\tilde{\mathcal{J}}$ from below by

$$\tilde{\mathcal{J}}(y(v), v) = \tilde{\mathcal{J}}(y(v), -\lambda^{-1} p(v)) \geq \frac{1}{2}\|\nabla \eta - \boldsymbol{g}_d\|^2 + \frac{1}{2\lambda}\|p\|^2 + \int_Q (\nabla y - \nabla \eta) \cdot (\nabla \eta - \boldsymbol{g}_d)\, d\boldsymbol{x}\, dt.$$

Note that Remark 3 applies here as well. For $\eta \in H^{1,1}_{0,per}(Q)$, let $p_\eta, \tilde{p}_\eta \in H^{1,\frac{1}{2}}_{0,per}(Q)$ be the solutions to the equations

$$\int_Q \left(\nu \nabla p_\eta \cdot \nabla z - \sigma \partial_t^{1/2} p_\eta\, \partial_t^{1/2} z^\perp\right) d\boldsymbol{x}\, dt = \int_Q (\nabla \eta - \boldsymbol{g}_d) \cdot \nabla z\, d\boldsymbol{x}\, dt, \qquad \forall z \in H^{1,\frac{1}{2}}_{0,per}(Q), \tag{43}$$

$$\int_Q \left(\nu \nabla \eta \cdot \nabla q + \sigma \partial_t^{1/2} \eta\, \partial_t^{1/2} q^\perp + \lambda^{-1} \tilde{p}_\eta\, q\right) d\boldsymbol{x}\, dt = 0, \qquad \forall q \in H^{1,\frac{1}{2}}_{0,per}(Q). \tag{44}$$

Deriving the minorant for the second minimization functional uses similar ideas now as presented for problem I (see Subsection 3.1). However, in the following we present the main steps which are important for problem II. Adding and subtracting p_η together with $\frac{1}{2\lambda}\|p - p_\eta\|^2 \geq 0$ yields

$$\begin{aligned}
&\tilde{\mathcal{J}}(y(v), v) = \tilde{\mathcal{J}}(y(v), -\lambda^{-1} p(v)) \\
&\quad \geq \frac{1}{2}\|\nabla \eta - \boldsymbol{g}_d\|^2 + \frac{1}{2\lambda}\|p_\eta\|^2 + \int_Q (\nabla y - \nabla \eta) \cdot (\nabla \eta - \boldsymbol{g}_d)\, d\boldsymbol{x}\, dt + \int_Q \lambda^{-1} (p - p_\eta)\, p_\eta\, d\boldsymbol{x}\, dt.
\end{aligned}$$

Appying equation (43), identity (8) (analogously to (28)) and then using equations (24) and (44) provides the estimate

$$\tilde{\mathcal{J}}(y(u), u) \geq \frac{1}{2}\|\nabla \eta - \boldsymbol{g}_d\|^2 + \frac{1}{2\lambda}\|p_\eta\|^2 + \int_Q \lambda^{-1} (\tilde{p}_\eta - p_\eta) p_\eta\, d\boldsymbol{x}\, dt.$$

Together with introducing an arbitrary function $\zeta \in H^{1,1}_{0,per}(Q)$, following (29), applying Theorem 1 and using (32), we finally derive the estimate

$$\tilde{\mathcal{J}}(y(u), u) \geq \tilde{\mathcal{J}}^\ominus(\eta, \zeta, \boldsymbol{\tau}, \boldsymbol{\rho}) \qquad \forall\, \eta, \zeta \in H^{1,1}_{0,per}(Q), \boldsymbol{\tau}, \boldsymbol{\rho} \in H(\mathrm{div}, Q) \tag{45}$$

with the fully computable minorant

$$\begin{aligned}
\tilde{\mathcal{J}}^\ominus(\eta, \zeta, \boldsymbol{\tau}, \boldsymbol{\rho}) &= \frac{1}{2}\|\nabla \eta - \boldsymbol{g}_d\|^2 + \frac{1}{2\lambda}\|\zeta\|^2 - \int_Q \left(\nu \nabla \eta \cdot \nabla \zeta + \sigma \partial_t \eta\, \zeta + \lambda^{-1} \zeta^2\right) d\boldsymbol{x}\, dt \\
&\quad - \frac{C_F^2}{\underline{\mu_1}^2 \lambda} \left(C_F \|\mathcal{R}_3(\zeta, \boldsymbol{\rho}, \eta)\| + \|\mathcal{R}_4(\zeta, \boldsymbol{\rho})\|\right)^2 \\
&\quad - \frac{1}{\underline{\mu_1}} (C_F \|\mathcal{R}_1(\eta, \boldsymbol{\tau}, -\lambda^{-1}\zeta)\| + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|)\left(C_F \|\mathcal{R}_3(\zeta, \boldsymbol{\rho}, \eta)\| + \|\mathcal{R}_4(\zeta, \boldsymbol{\rho})\|\right).
\end{aligned} \tag{46}$$

The following theorem is analogous to Theorem 2 and so is its proof. Also Remark 4 can be applied for optimization problem II.

Theorem 6. *The supremum of the minorant $\tilde{\mathcal{J}}^\ominus$ as defined in* (46) *is attained for the minimum of the cost functional* (21) *subject to* (1)–(3)*, which is equivalent to the optimal value of the optimality system* (22)–(23) *as follows*

$$\sup_{\eta, \zeta \in H^{1,1}_{0,per}(Q), \boldsymbol{\tau}, \boldsymbol{\rho} \in H(\mathrm{div}, Q)} \tilde{\mathcal{J}}^\ominus(\eta, \zeta, \boldsymbol{\tau}, \boldsymbol{\rho}) = \tilde{\mathcal{J}}(y(u), u). \tag{47}$$

4.3. **Guaranteed upper bounds for the discretization errors of the control and the state.** We present the majorants for the control-state errors measured in the norm $|||u-v|||^2 := \frac{1}{2}\|\nabla y(u) - \nabla y(v)\|^2 + \frac{\lambda}{2}\|u-v\|^2 = \frac{1}{2}\|\nabla y(u) - \nabla y(v)\|^2 + \frac{1}{2\lambda}\|p(u)-p(v)\|^2$. We obtain the identity $|||u-v|||^2 = \tilde{\mathcal{J}}(y(v),v) - \tilde{\mathcal{J}}(y(u),u)$ for an arbitrary control $v \in L^2(Q)$ yielding the error majorant

$$|||u-v|||^2 \leq \tilde{\mathcal{M}}^{\oplus}(\alpha,\beta;\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho},v) := \tilde{\mathcal{J}}^{\oplus}(\alpha,\beta;\eta,\boldsymbol{\tau},v) - \tilde{\mathcal{J}}^{\ominus}(\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho}) \tag{48}$$

for $\alpha,\beta > 0$, arbitrary $\eta,\zeta \in H^{1,1}_{0,per}(Q)$, $\boldsymbol{\tau},\boldsymbol{\rho} \in H(\mathrm{div},Q)$ with

$$\begin{aligned}\tilde{\mathcal{M}}^{\oplus}(\alpha,\beta;\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho},v) = {} & \frac{\alpha}{2}\|\nabla\eta - \boldsymbol{g}_d\|^2 + \gamma(\|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|^2 + \frac{C_F^2}{\beta}\|\mathcal{R}_1(\eta,\boldsymbol{\tau},v)\|^2) + \frac{\lambda}{2}\|v\|^2 - \frac{1}{2\lambda}\|\zeta\|^2 \\ & + \int_Q \left(\nu\nabla\eta\cdot\nabla\zeta + \sigma\partial_t\eta\,\zeta + \lambda^{-1}\zeta^2\right)\,d\boldsymbol{x}\,dt + \frac{C_F^2}{\underline{\mu_1}^2\lambda}\left(C_F\|\mathcal{R}_3(\zeta,\boldsymbol{\rho},\eta)\| + \|\mathcal{R}_4(\zeta,\boldsymbol{\rho})\|\right)^2 \\ & + \frac{1}{\underline{\mu_1}}(C_F\|\mathcal{R}_1(\eta,\boldsymbol{\tau},-\lambda^{-1}\zeta)\| + \|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|)\,(C_F\|\mathcal{R}_3(\zeta,\boldsymbol{\rho},\eta)\| + \|\mathcal{R}_4(\zeta,\boldsymbol{\rho})\|)\,,\end{aligned}$$

where $\underline{\mu_1} = \frac{1}{\sqrt{2}}\min\{\underline{\nu},\underline{\sigma}\}$. The infimum of this majorant is attained for the minimum of the optimization problem II being equivalent to solving the optimality system (22)–(23) ($v = u$, $\eta = y(u)$, $\zeta = p(u) = -\lambda u$, $\boldsymbol{\tau} = \nu\nabla y(u)$ and $\boldsymbol{\rho} = \nu\nabla p(u)$) as follows

$$\inf_{\substack{\eta,\zeta\in H^{1,1}_{0,per}(Q),\boldsymbol{\tau},\boldsymbol{\rho}\in H(\mathrm{div},Q),\\ v\in L^2(Q),\alpha,\beta>0}} \tilde{\mathcal{M}}^{\oplus}(\alpha,\beta;\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho},v) = 0.$$

Analogously, defining $|||u-v|||_1^2 := \frac{1}{2}\|\nabla y(u) - \nabla y(v)\|^2 + \frac{\lambda\underline{\mu_1}^2}{2C_F^2}|y(u)-y(v)|^2_{1,\frac{1}{2}}$ we derive

$$\begin{aligned}|||u-v|||_1^2 \leq \tilde{\mathcal{M}}_1^{\oplus}(\alpha,\beta;\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho},v) := {} & \tilde{\mathcal{J}}^{\oplus}(\alpha,\beta;\eta,\boldsymbol{\tau},v) - \tilde{\mathcal{J}}^{\ominus}(\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho}) \\ & + \frac{3\lambda}{4C_F^2}\left(C_F\|\mathcal{R}_1(\eta,\boldsymbol{\tau},v)\| + \|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|\right)^2,\end{aligned}$$

where now

$$\begin{aligned}\tilde{\mathcal{M}}_1^{\oplus}(\alpha,\beta;\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho},v) = {} & \frac{\alpha}{2}\|\nabla\eta - \boldsymbol{g}_d\|^2 + \gamma(\|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|^2 + \frac{C_F^2}{\beta}\|\mathcal{R}_1(\eta,\boldsymbol{\tau},v)\|^2) + \frac{\lambda}{2}\|v\|^2 - \frac{1}{2\lambda}\|\zeta\|^2 \\ & + \int_Q \left(\nu\nabla\eta\cdot\nabla\zeta + \sigma\partial_t\eta\,\zeta + \lambda^{-1}\zeta^2\right)\,d\boldsymbol{x}\,dt + \frac{C_F^2}{\underline{\mu_1}^2\lambda}\left(C_F\|\mathcal{R}_3(\zeta,\boldsymbol{\rho},\eta)\| + \|\mathcal{R}_4(\zeta,\boldsymbol{\rho})\|\right)^2 \\ & + \frac{1}{\underline{\mu_1}}(C_F\|\mathcal{R}_1(\eta,\boldsymbol{\tau},-\lambda^{-1}\zeta)\| + \|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|)\,(C_F\|\mathcal{R}_3(\zeta,\boldsymbol{\rho},\eta)\| + \|\mathcal{R}_4(\zeta,\boldsymbol{\rho})\|) \\ & + \frac{3\lambda}{4C_F^2}\left(C_F\|\mathcal{R}_1(\eta,\boldsymbol{\tau},v)\| + \|\mathcal{R}_2(\eta,\boldsymbol{\tau})\|\right)^2.\end{aligned}$$

All other results similar to Propositions 2 and 3 follow completely.

5. Multiharmonic finite element (MhFE) discretization

The desired state y_d and desired gradient $\boldsymbol{g}_d$ belong to $L^2(Q)$ and $[L^2(Q)]^d$, respectively. So they can be represented as Fourier series having Fourier coefficients from $L^2(\Omega)$. Moreover, we assume that η and ζ approximating the exact state y and adjoint state p, respectively, as well as the vector-valued functions $\boldsymbol{\tau}$ and $\boldsymbol{\rho}$ are given by truncated Fourier series. We also have the multiharmonic time derivative defined by $\partial_t\eta(\boldsymbol{x},t) = \sum_{k=1}^N (k\omega\,\eta_k^s(\boldsymbol{x})\cos(k\omega t) - k\omega\,\eta_k^c(\boldsymbol{x})\sin(k\omega t))$ as well as the gradient and divergence by

$$\nabla\eta(\boldsymbol{x},t) = \nabla\eta_0^c(\boldsymbol{x}) + \sum_{k=1}^N \left(\nabla\eta_k^c(\boldsymbol{x})\,\cos(k\omega t) + \nabla\eta_k^s(\boldsymbol{x})\,\sin(k\omega t)\right),$$

$$\nabla\cdot\boldsymbol{\tau}(\boldsymbol{x},t) = \nabla\cdot\boldsymbol{\tau}_0^c(\boldsymbol{x}) + \sum_{k=1}^N \left(\nabla\cdot\boldsymbol{\tau}_k^c(\boldsymbol{x})\,\cos(k\omega t) + \nabla\cdot\boldsymbol{\tau}_k^s(\boldsymbol{x})\,\sin(k\omega t)\right).$$

For the numerical treatment, we truncate the Fourier series expansions of all appearing functions at an index $N \in \mathbb{N}$ creating Fourier series approximations of the functions. Next, we approximate the Fourier coefficients $\boldsymbol{y}_k = (y_k^c, y_k^s)^T \in \mathbb{V}$ and $\boldsymbol{p}_k = (p_k^c, p_k^s)^T \in \mathbb{V}$ of the unknown state and adjoint state functions by finite element functions $\boldsymbol{y}_{kh} = (y_{kh}^c, y_{kh}^s)^T \in \mathbb{V}_h$ and $\boldsymbol{p}_{kh} = (p_{kh}^c, p_{kh}^s)^T \in \mathbb{V}_h$. The finite element space $\mathbb{V}_h = V_h \times V_h \subset \mathbb{V}$ with $V_h = \text{span}\{\phi_1, \ldots, \phi_n\}$ and $\{\phi_i(\boldsymbol{x}) : i = 1, 2, \ldots, n_h\}$ is conforming. We have defined $n = n_h = \dim V_h = O(h^{-d})$ and h is the discretization parameter. The basis of the finite element space V_h on the triangulation $\mathcal{T}_h$, which is regular, consists of piecewise linear and continuous elements.

5.1. **Optimization problem I.** The MhFE discretization (see also [27] for more details on the multiharmonic finite element analysis for such an optimal control problem) yields the system of linear equations having a saddle point structure corresponding to the decoupled problems (12)–(13) as follows

$$\begin{pmatrix} M_h & 0 & -K_{h,\nu} & k\omega M_{h,\sigma} \\ 0 & M_h & -k\omega M_{h,\sigma} & -K_{h,\nu} \\ -K_{h,\nu} & -k\omega M_{h,\sigma} & -\lambda^{-1} M_h & 0 \\ k\omega M_{h,\sigma} & -K_{h,\nu} & 0 & -\lambda^{-1} M_h \end{pmatrix} \begin{pmatrix} \underline{y}_k^c \\ \underline{y}_k^s \\ \underline{p}_k^c \\ \underline{p}_k^s \end{pmatrix} = \begin{pmatrix} M_h \underline{y}_{dk}^c \\ M_h \underline{y}_{dk}^s \\ 0 \\ 0 \end{pmatrix} \forall k = 1, \ldots, N. \tag{49}$$

The functions $y_{kh}^c(\boldsymbol{x}) = \sum_{j=1}^n y_{k,j}^c \phi_j(\boldsymbol{x})$, $y_{kh}^s(\boldsymbol{x}) = \sum_{j=1}^n y_{k,j}^s \phi_j(\boldsymbol{x})$, $p_{kh}^c(\boldsymbol{x}) = \sum_{j=1}^n p_{k,j}^c \phi_j(\boldsymbol{x})$ and $p_{kh}^s(\boldsymbol{x}) = \sum_{j=1}^n p_{k,j}^s \phi_j(\boldsymbol{x})$ are defined by the corresponding nodal function values $\underline{y}_k^c = (y_{k,j}^c)_{j=1,\ldots,n}$, $\underline{y}_k^s = (y_{k,j}^s)_{j=1,\ldots,n}$, $\underline{p}_k^c = (p_{k,j}^c)_{j=1,\ldots,n}$, $\underline{p}_k^s = (p_{k,j}^s)_{j=1,\ldots,n} \in \mathbb{R}^n$. We have defined the stiffness matrix $K_{h,\nu}$ as well as the mass matrices M_h and $M_{h,\sigma}$ by their entries

$$K_{h,\nu}^{ij} = \int_\Omega \nu \, \nabla\phi_i \cdot \nabla\phi_j \, d\boldsymbol{x}, \quad M_h^{ij} = \int_\Omega \phi_i \phi_j \, d\boldsymbol{x}, \quad M_{h,\sigma}^{ij} = \int_\Omega \sigma \, \phi_i \phi_j \, d\boldsymbol{x}.$$

The right-hand side can be obtained by computing the vectors

$$M_h \underline{y}_{dk}^c = \left(\int_\Omega y_{dk}^c \phi_i \, d\boldsymbol{x} \right)_{i=1,\ldots,n} \quad \text{and} \quad M_h \underline{y}_{dk}^s = \left(\int_\Omega y_{dk}^s \phi_i \, d\boldsymbol{x} \right)_{i=1,\ldots,n}.$$

Note that for the case $k = 0$, hence for (14)–(15), we obtain

$$\begin{pmatrix} M_h & -K_{h,\nu} \\ -K_{h,\nu} & -\lambda^{-1} M_h \end{pmatrix} \begin{pmatrix} \underline{y}_0^c \\ \underline{p}_0^c \end{pmatrix} = \begin{pmatrix} M_h \underline{y}_{d0}^c \\ 0 \end{pmatrix}. \tag{50}$$

Solving all the systems of linear equations finally lead to the contributions for computing the MhFE approximations $y_{Nh}(\boldsymbol{x}, t)$ and $p_{Nh}(\boldsymbol{x}, t)$ given by $y_{Nh}(\boldsymbol{x}, t) = y_{0h}^c(\boldsymbol{x}) + \sum_{k=1}^N (y_{kh}^c(\boldsymbol{x}) \cos(k\omega t) + y_{kh}^s(\boldsymbol{x}) \sin(k\omega t))$, $p_{Nh}(\boldsymbol{x}, t) = p_{0h}^c(\boldsymbol{x}) + \sum_{k=1}^N (p_{kh}^c(\boldsymbol{x}) \cos(k\omega t) + p_{kh}^s(\boldsymbol{x}) \sin(k\omega t))$. For some proper fast solvers for these systems we refer to [17, 21, 27]. Both, majorant (16) and minorant (34) of the cost functional $\mathcal{J}$ can be computed by choosing the MhFE approximations y_{Nh}, p_{Nh} and $u_{Nh} = -\lambda^{-1} p_{Nh}$ as η, ζ and v, respectively. The arbitrary functions $\boldsymbol{\tau}$ and $\boldsymbol{\rho}$ can also be represented in form of multiharmonic functions $\boldsymbol{\tau}_{Nh} = \boldsymbol{\tau}_{0h}^c(\boldsymbol{x}) + \sum_{k=1}^N (\boldsymbol{\tau}_{kh}^c(\boldsymbol{x}) \cos(k\omega t) + \boldsymbol{\tau}_{kh}^s(\boldsymbol{x}) \sin(k\omega t))$ and $\boldsymbol{\rho}_{Nh} = \boldsymbol{\rho}_{0h}^c(\boldsymbol{x}) + \sum_{k=1}^N (\boldsymbol{\rho}_{kh}^c(\boldsymbol{x}) \cos(k\omega t) + \boldsymbol{\rho}_{kh}^s(\boldsymbol{x}) \sin(k\omega t))$. Hence, majorant (16) and minorant (34) have a multiharmonic structure too. The linearity of the problem again yields the decoupling of the problems introducing $\alpha_k, \beta_k > 0$ and resulting into majorants $\mathcal{J}_k^\oplus$ and minorants $\mathcal{J}_k^\ominus$ corresponding to every Fourier mode. We start with the majorants $\mathcal{J}_0^\oplus = \mathcal{J}_0^\oplus(\alpha_0, \beta_0; y_{0h}^c, p_{0h}^c, \boldsymbol{\tau}_{0h}^c)$ and $\mathcal{J}_k^\oplus = \mathcal{J}_k^\oplus(\alpha_k, \beta_k; \boldsymbol{y}_{kh}, \boldsymbol{p}_{kh}, \boldsymbol{\tau}_{kh})$ together with defining the parameter $\gamma_k := ((1 + \alpha_k)(1 + \beta_k) C_F^2)/(2\alpha_k \underline{\mu_1}^2)$. We have that

$$\mathcal{J}_0^\oplus = \frac{1 + \alpha_0}{2} \|y_{0h}^c - y_{d0}^c\|_\Omega^2 + \frac{1}{2\lambda} \|p_{0h}^c\|_\Omega^2 + \gamma_0 (\|\mathcal{R}_{20}^c\|_\Omega^2 + \frac{C_F^2}{\beta_0} \|\mathcal{R}_{10}^c\|_\Omega^2) \tag{51}$$

and

$$\mathcal{J}_k^\oplus = \frac{1 + \alpha_k}{2} \|\boldsymbol{y}_{kh} - \boldsymbol{y}_{dk}\|_\Omega^2 + \frac{1}{2\lambda} \|\boldsymbol{p}_{kh}\|_\Omega^2 + \gamma_k (\|\mathcal{R}_{2k}\|_\Omega^2 + \frac{C_F^2}{\beta_k} \|\mathcal{R}_{1k}\|_\Omega^2). \tag{52}$$

Defining $\boldsymbol{\alpha}_{N+1} = (\alpha_0, \ldots, \alpha_{N+1})^T$ and $\boldsymbol{\beta}_N = (\beta_0, \ldots, \beta_N)^T$, we can write the overall majorant as

$$\mathcal{J}^{\oplus}(\boldsymbol{\alpha}_{N+1}, \boldsymbol{\beta}_N; y_{Nh}, p_{Nh}, \boldsymbol{\tau}_{Nh}) = T\, \mathcal{J}_0^{\oplus} + \frac{T}{2} \sum_{k=1}^{N} \mathcal{J}_k^{\oplus} + \frac{1+\alpha_{N+1}}{2}\, \mathcal{E}_N. \tag{53}$$

Here, the terms are ${\mathcal{R}_1}_0^c = \nabla \cdot \boldsymbol{\tau}_{0h}^c - \lambda^{-1} p_{0h}^c$, ${\mathcal{R}_2}_0^c = \boldsymbol{\tau}_{0h}^c - \nu \nabla y_{0h}^c$,

$$\begin{aligned} \mathcal{R}_{1k} &= k\omega\, \sigma \boldsymbol{y}_{kh}^{\perp} + \mathbf{div}\, \boldsymbol{\tau}_{kh} - \lambda^{-1} \boldsymbol{p}_{kh} \\ &= ({\mathcal{R}_1}_k^c, {\mathcal{R}_1}_k^s)^T = (-k\omega\, \sigma y_{kh}^s + \nabla \cdot \boldsymbol{\tau}_{kh}^c - \lambda^{-1} p_{kh}^c, k\omega\, \sigma y_{kh}^c + \nabla \cdot \boldsymbol{\tau}_{kh}^s - \lambda^{-1} p_{kh}^s)^T \end{aligned}$$

and $\mathcal{R}_{2k} = \boldsymbol{\tau}_{kh} - \nu \nabla \boldsymbol{y}_{kh} = ({\mathcal{R}_2}_k^c, {\mathcal{R}_2}_k^s)^T = (\boldsymbol{\tau}_{kh}^c - \nu \nabla y_{kh}^c, \boldsymbol{\tau}_{kh}^s - \nu \nabla y_{kh}^s)^T$. The truncation's remainder term $\mathcal{E}_N := \|y_d - y_{dN}\|^2 = \frac{T}{2} \sum_{k=N+1}^{\infty} \|\boldsymbol{y_d}_k\|_\Omega^2 = \frac{T}{2} \sum_{k=N+1}^{\infty} \left(\|y_{dk}^c\|_\Omega^2 + \|y_{dk}^s\|_\Omega^2 \right)$ can always be computed with any desired accuracy, since y_d is known (see also [25]). The minorant (34) can be written as

$$\mathcal{J}^{\ominus}(y_{Nh}, p_{Nh}, \boldsymbol{\tau}_{Nh}, \boldsymbol{\rho}_{Nh}) = T\, \mathcal{J}_0^{\ominus} + \frac{T}{2} \sum_{k=1}^{N} \mathcal{J}_k^{\ominus} + \frac{\mathcal{E}_N}{2}, \tag{54}$$

where $\mathcal{J}_0^{\ominus} = \mathcal{J}_0^{\ominus}(y_{0h}^c, p_{0h}^c, \boldsymbol{\tau}_{0h}^c, \boldsymbol{\rho}_{0h}^c)$ and $\mathcal{J}_k^{\ominus} = \mathcal{J}_k^{\ominus}(\boldsymbol{y}_{kh}, \boldsymbol{p}_{kh}, \boldsymbol{\tau}_{kh}, \boldsymbol{\rho}_{kh})$ are given by

$$\begin{aligned} \mathcal{J}_0^{\ominus} &= \frac{1}{2} \|y_{0h}^c - {y_d}_0^c\|_\Omega^2 + \frac{1}{2\lambda} \|p_{0h}^c\|_\Omega^2 - \int_\Omega \left(\nu \nabla y_{0h}^c \cdot \nabla p_{0h}^c + \lambda^{-1} (p_{0h}^c)^2 \right) d\boldsymbol{x} \\ &- \frac{C_F^2}{\underline{\mu_1}^2 \lambda} \left(C_F \|{\mathcal{R}_3}_0^c\|_\Omega + \|{\mathcal{R}_4}_0^c\|_\Omega \right)^2 - \frac{1}{\underline{\mu_1}} \left(C_F \|{\mathcal{R}_1}_0^c\|_\Omega + \|{\mathcal{R}_2}_0^c\|_\Omega \right) \left(C_F \|{\mathcal{R}_3}_0^c\|_\Omega + \|{\mathcal{R}_4}_0^c\|_\Omega \right) \end{aligned} \tag{55}$$

with ${\mathcal{R}_3}_0^c = \nabla \cdot \boldsymbol{\rho}_{0h}^c + y_{0h}^c - {y_d}_0^c$, ${\mathcal{R}_4}_0^c = \boldsymbol{\rho}_{0h}^c - \nu \nabla p_{0h}^c$, and

$$\begin{aligned} \mathcal{J}_k^{\ominus} &= \frac{1}{2} \|\boldsymbol{y}_{kh} - \boldsymbol{y}_{dk}\|_\Omega^2 + \frac{1}{2\lambda} \|\boldsymbol{p}_{kh}\|_\Omega^2 - \int_\Omega \left(\nu \nabla \boldsymbol{y}_{kh} \cdot \nabla \boldsymbol{p}_{kh} - k\omega\, \sigma \boldsymbol{y}_{kh}^{\perp} \cdot \boldsymbol{p}_{kh} + \lambda^{-1} \boldsymbol{p}_{kh}^2 \right) d\boldsymbol{x} \\ &- \frac{C_F^2}{\underline{\mu_1}^2 \lambda} \left(C_F \|\mathcal{R}_{3k}\|_\Omega + \|\mathcal{R}_{4k}\|_\Omega \right)^2 - \frac{1}{\underline{\mu_1}} \left(C_F \|\mathcal{R}_{1k}\|_\Omega + \|\mathcal{R}_{2k}\|_\Omega \right) \left(C_F \|\mathcal{R}_{3k}\|_\Omega + \|\mathcal{R}_{4k}\|_\Omega \right) \end{aligned} \tag{56}$$

with $\mathcal{R}_{3k} = k\omega\, \sigma \boldsymbol{p}_{kh}^{\perp} + \nabla \cdot \boldsymbol{\rho}_{kh} + \boldsymbol{y}_{kh} - \boldsymbol{y}_{dk} = ({\mathcal{R}_3}_k^c, {\mathcal{R}_3}_k^s)^T = (-k\omega\, \sigma p_{kh}^s + \nabla \cdot \boldsymbol{\rho}_{kh}^c + y_{kh}^c - {y_d}_k^c, k\omega\, \sigma p_{kh}^c + \nabla \cdot \boldsymbol{\rho}_{kh}^s + y_{kh}^s - {y_d}_k^s)^T$ and $\mathcal{R}_{4k} = \boldsymbol{\rho}_{kh} - \nu \nabla \boldsymbol{p}_{kh} = ({\mathcal{R}_4}_k^c, {\mathcal{R}_4}_k^s)^T = (\boldsymbol{\rho}_{kh}^c - \nu \nabla p_{kh}^c, \boldsymbol{\rho}_{kh}^s - \nu \nabla p_{kh}^s)^T$.

Remark 5. *For any index $\bar{N} \in \mathbb{N}$, $\bar{N} > N$, the truncated remainder term*

$$\mathcal{E}_{N,\bar{N}} := \frac{T}{2} \sum_{k=N+1}^{\bar{N}} \|\boldsymbol{y_d}_k\|_\Omega^2$$

is a fully computable lower bound for the remainder term $\mathcal{E}_N$. For any given $y_d \in L^2(Q)$, this provides an arbitrarily tight lower bound for the minorant (54), which is in return optimal.

The fluxes $\boldsymbol{\tau}_{0h}^c$, $\boldsymbol{\rho}_{0h}^c$ and $\boldsymbol{\tau}_{kh}$, $\boldsymbol{\rho}_{kh}$ for all $k = 1, \ldots, N$, denoted by $\boldsymbol{\tau}_{kh} = R_h^{\text{flux}}(\nu \nabla \boldsymbol{y}_{kh})$ and $\boldsymbol{\rho}_{kh} = R_h^{\text{flux}}(\nu \nabla \boldsymbol{p}_{kh})$ are reconstructed by lowest-order Raviart-Thomas elements mapping L^2-functions to $H(\text{div}, \Omega)$, see [38] as well as [25, 26], leading to $\boldsymbol{\tau}_{Nh} = R_h^{\text{flux}}(\nu \nabla y_{Nh})$ and $\boldsymbol{\rho}_{Nh} = R_h^{\text{flux}}(\nu \nabla p_{Nh})$. We minimize $\mathcal{J}^{\oplus}$ with respect to the positive parameters α_k and β_k leading to the optimized $\boldsymbol{\alpha}_{N+1}$ and $\boldsymbol{\beta}_N$. Finally, the multiharmonic majorant (53) and minorant (54) lead to upper and lower bounds for $\mathcal{J}$ which are guaranteed and computable. The computation of the infimum of $\mathcal{J}^{\oplus}$ and the supremum of $\mathcal{J}^{\ominus}$ provide the minimum of $\mathcal{J}$, see also [26].

5.2. **Optimization problem II.** For the second problem, we only summarize the main results and changes. The MhFE discretization leads to the following discrete problem:

$$\begin{pmatrix} K_h & 0 & -K_{h,\nu} & k\omega M_{h,\sigma} \\ 0 & K_h & -k\omega M_{h,\sigma} & -K_{h,\nu} \\ -K_{h,\nu} & -k\omega M_{h,\sigma} & -\lambda^{-1} M_h & 0 \\ k\omega M_{h,\sigma} & -K_{h,\nu} & 0 & -\lambda^{-1} M_h \end{pmatrix} \begin{pmatrix} \underline{y}_k^c \\ \underline{y}_k^s \\ \underline{p}_k^c \\ \underline{p}_k^s \end{pmatrix} = \begin{pmatrix} K_h \underline{g}_{dk}^c \\ K_h \underline{g}_{dk}^s \\ 0 \\ 0 \end{pmatrix} \tag{57}$$

and

$$\begin{pmatrix} K_h & -K_{h,\nu} \\ -K_{h,\nu} & -\lambda^{-1}M_h \end{pmatrix}\begin{pmatrix} \underline{y}_0^c \\ \underline{p}_0^c \end{pmatrix} = \begin{pmatrix} K_h\underline{g}_{d0}^c \\ 0 \end{pmatrix}. \tag{58}$$

Now, the right-hand side vectors are computed by

$$K_h\underline{g}_{dk}^c = \left(\int_\Omega \boldsymbol{g}_{dk}^{\,c}\cdot\nabla\phi_i\,d\boldsymbol{x}\right)_{i=1,\dots,n} \quad\text{and}\quad K_h\underline{g}_{dk}^s = \left(\int_\Omega \boldsymbol{g}_{dk}^{\,s}\cdot\nabla\phi_i\,d\boldsymbol{x}\right)_{i=1,\dots,n}.$$

We summarize the discrete majorant (40) and minorant (46) of the cost functional $\tilde{\mathcal{J}}$ computed by choosing the MhFE approximations for all the (arbitrary) functions. Defining now $\tilde{\mathcal{E}}_N := \|\boldsymbol{g}_d - \boldsymbol{g}_{dN}\|^2 = \frac{T}{2}\sum_{k=N+1}^{\infty}\|\boldsymbol{g}_{dk}\|_\Omega^2 = \frac{T}{2}\sum_{k=N+1}^{\infty}\left(\|\boldsymbol{g}_{dk}^c\|_\Omega^2 + \|\boldsymbol{g}_{dk}^s\|_\Omega^2\right)$ as truncation's remainder term, we write the majorant (40) as

$$\tilde{\mathcal{J}}^{\oplus}(\boldsymbol{\alpha}_{N+1},\boldsymbol{\beta}_N;y_{Nh},p_{Nh},\boldsymbol{\tau}_{Nh}) = T\,\tilde{\mathcal{J}}_0^{\oplus} + \frac{T}{2}\sum_{k=1}^{N}\tilde{\mathcal{J}}_k^{\oplus} + \frac{1+\alpha_{N+1}}{2}\,\tilde{\mathcal{E}}_N, \tag{59}$$

where $\tilde{\mathcal{J}}_0^{\oplus} = \tilde{\mathcal{J}}_0^{\oplus}(\alpha_0,\beta_0;y_{0h}^c,p_{0h}^c,\boldsymbol{\tau}_{0h}^c)$ and $\tilde{\mathcal{J}}_k^{\oplus} = \tilde{\mathcal{J}}_k^{\oplus}(\alpha_k,\beta_k s;\boldsymbol{y}_{kh},\boldsymbol{p}_{kh},\boldsymbol{\tau}_{kh})$ are given by

$$\tilde{\mathcal{J}}_0^{\oplus} = \frac{1+\alpha_0}{2}\|\nabla y_{0h}^c - \boldsymbol{g}_{d0}^{\,c}\|_\Omega^2 + \frac{1}{2\lambda}\|p_{0h}^c\|_\Omega^2 + \gamma_0(\|\mathcal{R}_{2}{}_0^c\|_\Omega^2 + \frac{C_F^2}{\beta_0}\|\mathcal{R}_{1}{}_0^c\|_\Omega^2) \tag{60}$$

and

$$\tilde{\mathcal{J}}_k^{\oplus} = \frac{1+\alpha_k}{2}\|\nabla\boldsymbol{y}_{kh} - \boldsymbol{g}_{dk}\|_\Omega^2 + \frac{1}{2\lambda}\|\boldsymbol{p}_{kh}\|_\Omega^2 + \gamma_k(\|\mathcal{R}_{2k}\|_\Omega^2 + \frac{C_F^2}{\beta_k}\|\mathcal{R}_{1k}\|_\Omega^2). \tag{61}$$

The minorant (46) can be written as

$$\tilde{\mathcal{J}}^{\ominus}(y_{Nh},p_{Nh},\boldsymbol{\tau}_{Nh},\boldsymbol{\rho}_{Nh}) = T\,\tilde{\mathcal{J}}_0^{\ominus} + \frac{T}{2}\sum_{k=1}^{N}\tilde{\mathcal{J}}_k^{\ominus} + \frac{\tilde{\mathcal{E}}_N}{2}, \tag{62}$$

where $\tilde{\mathcal{J}}_0^{\ominus} = \tilde{\mathcal{J}}_0^{\ominus}(y_{0h}^c,p_{0h}^c,\boldsymbol{\tau}_{0h}^c,\boldsymbol{\rho}_{0h}^c)$ and $\tilde{\mathcal{J}}_k^{\ominus} = \tilde{\mathcal{J}}_k^{\ominus}(\boldsymbol{y}_{kh},\boldsymbol{p}_{kh},\boldsymbol{\tau}_{kh},\boldsymbol{\rho}_{kh})$ are given by

$$\begin{aligned}\tilde{\mathcal{J}}_0^{\ominus} &= \frac{1}{2}\|\nabla y_{0h}^c - \boldsymbol{g}_{d0}^{\,c}\|_\Omega^2 + \frac{1}{2\lambda}\|p_{0h}^c\|_\Omega^2 - \int_\Omega\left(\nu\nabla y_{0h}^c\cdot\nabla p_{0h}^c + \lambda^{-1}(p_{0h}^c)^2\right)d\boldsymbol{x}\\ &\quad - \frac{C_F^2}{\underline{\mu_1}^2\lambda}\left(C_F\|\mathcal{R}_{3}{}_0^c\|_\Omega + \|\mathcal{R}_{4}{}_0^c\|_\Omega\right)^2 - \frac{1}{\underline{\mu_1}}\left(C_F\|\mathcal{R}_{1}{}_0^c\|_\Omega + \|\mathcal{R}_{2}{}_0^c\|_\Omega\right)\left(C_F\|\mathcal{R}_{3}{}_0^c\|_\Omega + \|\mathcal{R}_{4}{}_0^c\|_\Omega\right)\end{aligned} \tag{63}$$

and

$$\begin{aligned}\tilde{\mathcal{J}}_k^{\ominus} &= \frac{1}{2}\|\nabla\boldsymbol{y}_{kh} - \boldsymbol{g}_{dk}\|_\Omega^2 + \frac{1}{2\lambda}\|\boldsymbol{p}_{kh}\|_\Omega^2 - \int_\Omega\left(\nu\nabla\boldsymbol{y}_{kh}\cdot\nabla\boldsymbol{p}_{kh} - k\omega\,\sigma\boldsymbol{y}_{kh}^{\perp}\cdot\boldsymbol{p}_{kh} + \lambda^{-1}\boldsymbol{p}_{kh}^2\right)d\boldsymbol{x}\\ &\quad - \frac{C_F^2}{\underline{\mu_1}^2\lambda}\left(C_F\|\mathcal{R}_{3k}\|_\Omega + \|\mathcal{R}_{4k}\|_\Omega\right)^2 - \frac{1}{\underline{\mu_1}}\left(C_F\|\mathcal{R}_{1k}\|_\Omega + \|\mathcal{R}_{2k}\|_\Omega\right)\left(C_F\|\mathcal{R}_{3k}\|_\Omega + \|\mathcal{R}_{4k}\|_\Omega\right).\end{aligned} \tag{64}$$

6. Robust preconditioners for the minimal residual method

The saddle point systems (49), (50), (57) and (58) can be solved by using the preconditioned MinRes method, see [37]. A convergence result for the preconditioned MinRes method can be found in [13] stating that the convergence rate of the preconditioned MinRes method depends on the condition number of the preconditioned system. The derivation of preconditioners for problems (49) for $k = 1,\dots,N$ and (50) for $k = 0$ have already been presented and discussed in [17, 27] given by

$$\mathcal{P}_k = \operatorname{diag}(D_k, D_k, \lambda^{-1}D_k, \lambda^{-1}D_k) \quad\text{and}\quad \mathcal{P}_0 = \operatorname{diag}(D_0, \lambda^{-1}D_0), \tag{65}$$

respectively, where $D_k = \sqrt{\lambda}K_{h,\nu} + k\omega\sqrt{\lambda}M_{h,\sigma} + M_h$ and $D_0 = M_h + \sqrt{\lambda}K_{h,\nu}$. In [17], preconditioners are derived following the technique in [46] based on operator interpolation theory.

In this section, we present new robust preconditioners for the problem matrices in (57) for $k = 1,\dots,N$ and in (58) for $k = 0$ in order to solve optimization problem II. Here, we assume that σ and ν are constant, which we also choose in the numerical results in Section 7. Hence, $M_{h,\sigma} = \sigma M_h$

and $K_{h,\nu} = \nu K_h$. The block-diagonal preconditioners are practically implemented by the version of the algebraic multilevel iteration (AMLI) method from [20]. The AMLI preconditioned MinRes solver is robust and of optimal complexity which is proved in [21]. This can be also observed in the numerical results' section. Let us consider the saddle point system (57) for $k \geq N$. The derivation of preconditioners for system (58) in case of $k = 0$ is completely analogous. We refer the reader to [17] on details for the derivation of the preconditioners based on Schur complements.

Defining the matrices and vectors

$$A = \mathrm{diag}(K_h, K_h), \qquad B = \begin{pmatrix} -\nu K_h & -k\omega\sigma M_h \\ k\omega\sigma M_h & -\nu K_h \end{pmatrix}, \qquad C = \mathrm{diag}(\lambda^{-1} M_h, \lambda^{-1} M_h), \tag{66}$$

$\underline{f} = (K_h \underline{g}^c_{dk}, K_h \underline{g}^s_{dk})^T$, $\underline{y} = (\underline{y}^c_k, \underline{y}^s_k)^T$ and $\underline{p} = (\underline{p}^c_k, \underline{p}^s_k)^T$ leads to the following problem structure

$$\begin{pmatrix} A & B^T \\ B & -C \end{pmatrix} \begin{pmatrix} \underline{y} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ 0 \end{pmatrix} \tag{67}$$

with the symmetric and positive definite matrices A and C. We define the negative Schur complements $S = C + BA^{-1}B^T$ and $R = A + B^T C^{-1} B$ yielding the preconditioners for $k \geq N$ as follows

$$\tilde{\mathcal{P}} = \mathrm{diag}(A, S) \qquad \text{and} \qquad \tilde{\mathcal{Q}} = \mathrm{diag}(R, C). \tag{68}$$

The negative Schur complements are given by

$$S = \mathrm{diag}(\nu K_h + \lambda^{-1} M_h + k^2\omega^2\sigma^2 M_h K_h^{-1} M_h, \nu K_h + \lambda^{-1} M_h + k^2\omega^2\sigma^2 M_h K_h^{-1} M_h) \tag{69}$$

and

$$R = \mathrm{diag}(K_h + k^2\omega^2\sigma^2 \lambda M_h + \nu^2 \lambda K_h M_h^{-1} K_h, K_h + k^2\omega^2\sigma^2 \lambda M_h + \nu^2 \lambda K_h M_h^{-1} K_h). \tag{70}$$

Let us define

$$\tilde{D}^S_k = \nu K_h + \lambda^{-1} M_h + k^2\omega^2\sigma^2 M_h K_h^{-1} M_h \tag{71}$$

and

$$\tilde{D}^R_k = K_h + k^2\omega^2\sigma^2 \lambda M_h + \nu^2 \lambda K_h M_h^{-1} K_h. \tag{72}$$

Then S and R can be written as $S = \mathrm{diag}(\tilde{D}^S_k, \tilde{D}^S_k)$ and $R = \mathrm{diag}(\tilde{D}^R_k, \tilde{D}^R_k)$, respectively.

Remark 6. *Both Schur complement preconditioners $\tilde{\mathcal{P}}$ and $\tilde{\mathcal{Q}}$ in* (68) *can be chosen for computations of optimization problem II leading to fast and robust convergence rates, see* [22] *and* [35].

Inserting now A and C from (66) and S and R from (69) and (70), respectively, the Schur complement preconditioners in the form of $\tilde{\mathcal{P}}$ as presented in (68) in the case of $k \geq N$ as well as for $k = 0$ are given by

$$\tilde{\mathcal{P}}_k = \mathrm{diag}(K_h, K_h, \tilde{D}^S_k, \tilde{D}^S_k) \quad \text{and} \quad \tilde{\mathcal{P}}_0 = \mathrm{diag}(K_h, \nu K_h + \lambda^{-1} M_h) \tag{73}$$

Analogously in the form of $\tilde{\mathcal{Q}}$, they are given by

$$\tilde{\mathcal{Q}}_k = \mathrm{diag}(\tilde{D}^R_k, \tilde{D}^R_k, \lambda^{-1} M_h, \lambda^{-1} M_h) \quad \text{and} \quad \tilde{\mathcal{Q}}_0 = \mathrm{diag}(\lambda^{-1} M_h, K_h + \nu^2 \lambda K_h M_h^{-1} K_h), \tag{74}$$

for $k \geq N$ and $k = 0$ for the saddle point systems (57) and (58), respectively. We refer again to [17] for further details on the preconditioners' derivation. For the numerical experiments of this work, we simplify the preconditioner such that $\tilde{\tilde{\mathcal{P}}}_k = \mathrm{diag}(\tilde{\tilde{D}}^S_k, \tilde{\tilde{D}}^S_k, \lambda^{-1}\tilde{\tilde{D}}^S_k, \lambda^{-1}\tilde{\tilde{D}}^S_k)$ and $\tilde{\tilde{\mathcal{P}}}_0 = \mathrm{diag}(\tilde{\tilde{D}}^S_0, \lambda^{-1}\tilde{\tilde{D}}^S_0)$, where $\tilde{\tilde{D}}^S_k = \nu\sqrt{\lambda} K_h + (1 + k\omega\sigma\sqrt{\lambda}) M_h$ and $\tilde{\tilde{D}}^S_0 = M_h + \nu\sqrt{\lambda} K_h$ in a similar form as (65) in order to apply the AMLI preconditioned MinRes method analogously as for optimization problem I.

7. Numerical results

In this section, we present numerical results performed in C++ for computing the minorants and majorants of the two optimal control problems with cost functionals (9) and (21) and for different cases of given data denoted by Examples 1–6. For the numerical results on the majorants of optimization problem I, we refer to [26] and for the numerical analysis including convergence results, to [17, 27]. However, all the numerical experiments on the majorants of optimization problem II as well as all on minorants for both problems I and II are new. We perform numerical experiments for the same three cases (applied on the problem data) as for problem I but in Examples 4–6 they are applied on the desired gradients.

The domain Ω is the two-dimensional unit square $\Omega = (0,1)^2$ with Friedrichs' constant $C_F = 1/(\sqrt{2}\pi)$. The triangulation of the domain is performed regular leading to a uniform grid. The finite elements are chosen as described in Section 5. The coefficients are chosen to be $\nu = \sigma = 1$. In Examples 1, 2, 4 and 5, the cost parameter is chosen to be $\lambda = 0.1$, $T = 2\pi/\omega$ and $\omega = 1$. In Examples 3 and 6, $\lambda = 0.01$, $T = 1$ and $\omega = 2\pi$. The MhFE approximations for η, ζ and $\boldsymbol{\tau}, \boldsymbol{\rho}$ are chosen as well as the fluxes are reconstructed by RT^0-extensions (lowest-order standard Raviart-Thomas) resulting in averaged fluxes which are now from $H(\text{div}, \Omega)$, see also [25, 26] for further details. The grid sizes range between 16×16 and 256×256 as well as 512×512 to obtain finer grid solutions for Examples 3 and 6 as a reference for the exact solution. The preconditioned MinRes iteration was stopped after 8 iteration steps in all computations using the AMLI preconditioner with 4 inner iterations. The numerical experiments for Examples 1, 2, 4 and 5 were performed on a laptop with Intel(R) Core(TM) i5-6267U CPU @ 2.90GHz processor and 16 GB 2133 MHz LPDDR3 memory. The numerical experiments for Examples 3 and 6 were performed on a CPU server with a Tumbleweed distribution having 64 cores and 1 terabyte memory in order to provide enough memory for computing the finer grid solutions in addition. All computational times t^{sec} are presented in seconds and include the CPU times also needed to derive the majorants and minorants. However, we want to highlight that these times are much smaller compared to the rest. The computational times of Examples 3 and 6 exclude the computation of the solution on the finer grid (512×512).

7.1. Numerical results for optimization problem I.

7.1.1. *Example 1.* The desired state is a time-periodic and time-analytic function

$$y_d(\boldsymbol{x}, t) = e^t \sin(t)\, 0.1 \left(\left(12 + 4\pi^4\right) \sin^2(t) - 6\cos(t)(\cos(t) - \sin(t)) \right) \sin(x_1\pi) \sin(x_2\pi), \tag{75}$$

which is however not time-harmonic. Note that the exact state function for this example is

$$y(\boldsymbol{x}, t) = e^t \sin(t)^3 \sin(x_1\pi) \sin(x_2\pi). \tag{76}$$

The truncation index for the multiharmonic approximations is chosen as $N = 8$ here. Table 1 presents for different grid sizes CPU times t^{sec}, values for the majorants $\mathcal{J}_0^{\oplus}$ and minorants $\mathcal{J}_0^{\ominus}$ as defined in (51) and (55), and efficiency indices $I_{\text{eff}}^{\mathcal{J}_0^{\oplus}} = \mathcal{J}_0^{\oplus}/\mathcal{J}_0$, $I_{\text{eff}}^{\mathcal{J}_0^{\ominus}} = \mathcal{J}_0^{\ominus}/\mathcal{J}_0$ and $I_{\text{eff}}^{\mathcal{J},0} = \mathcal{J}_0^{\oplus}/\mathcal{J}_0^{\ominus}$. Here, $\mathcal{J}_0 = \mathcal{J}_0(y_0^c, u_0^c) = \frac{1}{2}\|y_0^c - y_{d0}^c\|_\Omega^2 + \frac{\lambda}{2}\|u_0^c\|_\Omega^2$ as introduced in Subsection 2.2. In Table 2, the numerical results for the Fourier mode $k = 1$ are presented including $\mathcal{J}_k^{\oplus}$, $\mathcal{J}_k^{\ominus}$ as defined in (52) and (56) and the corresponding efficiency indices $I_{\text{eff}}^{\mathcal{J}_k^{\oplus}} = \mathcal{J}_k^{\oplus}/\mathcal{J}_k$, $I_{\text{eff}}^{\mathcal{J}_k^{\ominus}} = \mathcal{J}_k^{\ominus}/\mathcal{J}_k$ and $I_{\text{eff}}^{\mathcal{J},k} = \mathcal{J}_k^{\oplus}/\mathcal{J}_k^{\ominus}$. Moreover, we present the efficiency indices for $\mathcal{M}_1^{\oplus}$ given for the modes by

$$I_{\text{eff}}^{\mathcal{M}_1,0} = \sqrt{\frac{\mathcal{M}_{1,0}^{\oplus}(\alpha_0, \beta_0; y_{0h}^c, p_{0h}^c, \boldsymbol{\tau}_{0h}^c, \boldsymbol{\rho}_{0h}^c)}{|||y_0^c - y_{0h}^c|||_{1,0}^2}} \quad \text{and} \quad I_{\text{eff}}^{\mathcal{M}_1,k} = \sqrt{\frac{\mathcal{M}_{1,k}^{\oplus}(\alpha_k, \beta_k; \boldsymbol{y}_{kh}, \boldsymbol{p}_{kh}, \boldsymbol{\tau}_{kh}, \boldsymbol{\rho}_{kh})}{|||\boldsymbol{y}_k - \boldsymbol{y}_{kh}|||_{1,k}^2}}.$$

The error norms for the modes are given by

$$|||y_0^c - y_{0h}^c|||_{1,0}^2 = \frac{1}{2}\|y_0^c - y_{0h}^c\|_\Omega^2 + \frac{\lambda {\mu_1}^2}{2C_F^2}\|\nabla y_0^c - \nabla y_{0h}^c\|_\Omega^2$$

and

$$|||\boldsymbol{y}_k - \boldsymbol{y}_{kh}|||_{1,k}^2 = \left(\frac{1}{2} + \frac{k\omega\lambda\underline{\mu_1}^2}{2C_F^2}\right)\|\boldsymbol{y}_k - \boldsymbol{y}_{kh}\|_\Omega^2 + \frac{\lambda\underline{\mu_1}^2}{2C_F^2}\|\nabla\boldsymbol{y}_k - \nabla\boldsymbol{y}_{kh}\|_\Omega^2$$

leading the representation

$$|||u - v|||_1^2 = T\,|||y_0^c - y_{0h}^c|||_{1,0}^2 + \frac{T}{2}\sum_{k=1}^{N}|||\boldsymbol{y}_k - \boldsymbol{y}_{kh}|||_{1,k}^2 + \mathcal{F}_N \tag{77}$$

with the remainder term $\mathcal{F}_N := \frac{T}{2}\sum_{k=N+1}^{\infty}|||\boldsymbol{y}_k|||_{1,k}^2$. For the numerical experiments, we can estimate the efficiency index for $\mathcal{M}_1^\oplus$ from above by estimating (77) from below ignoring the remainder term $\mathcal{F}_N$ leading to the overall efficiency index for $\mathcal{M}_1^\oplus$

$$I_{\text{eff}}^{\mathcal{M}_1} = \sqrt{\frac{\mathcal{M}_1^\oplus(\alpha,\beta;\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho},v)}{T\,|||y_0^c - y_{0h}^c|||_{1,0}^2 + \frac{T}{2}\sum_{k=1}^{N}|||\boldsymbol{y}_k - \boldsymbol{y}_{kh}|||_{1,k}^2}}. \tag{78}$$

The corresponding majorants $\mathcal{M}_{1,0}^\oplus = \mathcal{M}_{1,0}^\oplus(\alpha_0,\beta_0;y_{0h}^c,p_{0h}^c,\boldsymbol{\tau}_{0h}^c,\boldsymbol{\rho}_{0h}^c)$ and $\mathcal{M}_{1,k}^\oplus = \mathcal{M}_{1,k}^\oplus(\alpha_k,\beta_k;\boldsymbol{y}_{kh},\boldsymbol{p}_{kh},\boldsymbol{\tau}_{kh},\boldsymbol{\rho}_{kh})$ are given by

$$\begin{aligned}\mathcal{M}_{1,0}^\oplus = \mathcal{J}_0^\oplus - \mathcal{J}_0^\ominus + \frac{3\lambda}{4C_F^2}\left(C_F\|\mathcal{R}_{1}{}_0^c\|_\Omega + \|\mathcal{R}_{2}{}_0^c\|_\Omega\right)^2 = \frac{\alpha_0}{2}\|y_{0h}^c - y_{d}{}_0^c\|_\Omega^2 + \gamma_0\|\mathcal{R}_{2}{}_0^c\|_\Omega^2 \\ + \frac{\gamma_0 C_F^2}{\beta_0}\|\mathcal{R}_{1}{}_0^c\|_\Omega^2 + \int_\Omega\left(\nu\nabla y_{0h}^c\cdot\nabla p_{0h}^c + \lambda^{-1}(p_{0h}^c)^2\right)d\boldsymbol{x} + \frac{C_F^2}{\underline{\mu_1}^2\lambda}\left(C_F\|\mathcal{R}_{3}{}_0^c\|_\Omega + \|\mathcal{R}_{4}{}_0^c\|_\Omega\right)^2 \\ + \frac{1}{\underline{\mu_1}}\left(C_F\|\mathcal{R}_{1}{}_0^c\|_\Omega + \|\mathcal{R}_{2}{}_0^c\|_\Omega\right)\left(C_F\|\mathcal{R}_{3}{}_0^c\|_\Omega + \|\mathcal{R}_{4}{}_0^c\|_\Omega\right) + \frac{3\lambda}{4C_F^2}\left(C_F\|\mathcal{R}_{1}{}_0^c\|_\Omega + \|\mathcal{R}_{2}{}_0^c\|_\Omega\right)^2\end{aligned}$$

and also $\mathcal{M}_{1,k}^\oplus = \mathcal{J}_k^\oplus - \mathcal{J}_k^\ominus + \frac{3\lambda}{4C_F^2}\left(C_F\|\mathcal{R}_{1k}\|_\Omega + \|\mathcal{R}_{2k}\|_\Omega\right)^2$. Table 3 sums up the numerical results for Example 1 by presenting the minorants, majorants and efficiency indices on a grid of size 256×256 for all k up to $N=4$ and then for $k=6$ and $k=8$ (since their results were similar as for $k=5$ and $k=7$). For $N=3$ and $N=8$, the truncation's remainder terms can be precomputed

grid	t^{sec}	$\mathcal{J}_0^\ominus$	$I_{\text{eff}}^{\mathcal{J}_0^\ominus}$	$\mathcal{J}_0^\oplus$	$I_{\text{eff}}^{\mathcal{J}_0^\oplus}$	$I_{\text{eff}}^{\mathcal{J},0}$	$I_{\text{eff}}^{\mathcal{M}_1,0}$
16×16	0.02	1.13e+05	0.90	1.26e+05	1.01	1.12	1.53
32×32	0.07	1.14e+05	0.90	1.27e+05	1.00	1.11	1.47
64×64	0.24	1.14e+05	0.90	1.27e+05	1.00	1.11	1.44
128×128	1.16	1.14e+05	0.90	1.27e+05	1.00	1.11	1.43
256×256	4.51	1.14e+05	0.90	1.27e+05	1.00	1.11	1.42

TABLE 1. **Example 1**. Minorant $\mathcal{J}_0^\ominus$, majorant $\mathcal{J}_0^\oplus$ and their efficiency indices computed on grids of different sizes.

and are given by $\mathcal{E}_3 = 63694.86$ and $\mathcal{E}_8 = 106.06$, respectively. Since the overall efficiency indices in Table 3 stay all in approximately the same range, we observe that the method is robust. However, the efficiency indices for the combined error norm $I_{\text{eff}}^{\mathcal{M}_1}$ indicate that the modes $k=1$ and $k=4$ are the most significant to represent the solution by its multiharmonic approximation. Comparing the last two lines of Table 3 shows that the value for representing the cost functional of the exact solution is already sufficiently accurate for a truncation index $N=3$. One of the reasons for this is that the remainder term $\mathcal{E}_N$ can be precomputed exactly.

7.1.2. *Example 2.* We choose the time-analytic, not time-periodic, desired state function

$$y_d(\boldsymbol{x},t) = e^t\,0.2\left((5+2\pi^4)\sin(t) - \cos(t)\right)\sin(x_1\pi)\sin(x_2\pi) \tag{79}$$

having as exact solution the state function

$$y(\boldsymbol{x},t) = e^t\sin(t)\sin(x_1\pi)\sin(x_2\pi). \tag{80}$$

grid	t^{sec}	$\mathcal{J}_1^{\ominus}$	$I_{\text{eff}}^{\mathcal{J}_1^{\ominus}}$	$\mathcal{J}_1^{\oplus}$	$I_{\text{eff}}^{\mathcal{J}_1^{\oplus}}$	$I_{\text{eff}}^{\mathcal{J},1}$	$I_{\text{eff}}^{\mathcal{M}_1,1}$
16×16	0.02	4.27e+05	0.90	4.74e+05	1.00	1.11	6.18
32×32	0.07	4.31e+05	0.90	4.79e+05	1.00	1.11	6.01
64×64	0.25	4.32e+05	0.90	4.80e+05	1.00	1.11	5.93
128×128	1.13	4.32e+05	0.90	4.80e+05	1.00	1.11	5.90
256×256	4.66	4.32e+05	0.90	4.80e+05	1.00	1.11	5.89

TABLE 2. **Example 1**. Minorant $\mathcal{J}_1^{\ominus}$, majorant $\mathcal{J}_1^{\oplus}$ and their efficiency indices computed on grids of different sizes.

mode	t^{sec}	$\mathcal{J}^{\ominus}$	$I_{\text{eff}}^{\mathcal{J}^{\ominus}}$	$\mathcal{J}^{\oplus}$	$I_{\text{eff}}^{\mathcal{J}^{\oplus}}$	$I_{\text{eff}}^{\mathcal{J}}$	$I_{\text{eff}}^{\mathcal{M}_1}$
$k = 0$	4.51	1.14e+05	0.90	1.27e+05	1.00	1.11	1.42
$k = 1$	4.66	4.32e+05	0.90	4.80e+05	1.00	1.11	5.89
$k = 2$	4.75	1.79e+05	0.90	1.99e+05	1.00	1.11	1.64
$k = 3$	4.81	6.10e+04	0.90	6.74e+04	1.00	1.11	1.84
$k = 4$	4.74	7.68e+03	0.91	8.42e+03	1.00	1.10	11.06
$k = 6$	4.72	2.05e+02	0.90	2.29e+02	1.00	1.11	2.08
$k = 8$	4.81	1.97e+01	0.93	2.19e+01	1.04	1.12	1.22
overall ($N = 3$)	–	2.86e+06	0.90	3.17e+06	1.00	1.11	2.08
overall ($N = 8$)	–	2.86e+06	0.90	3.17e+06	1.00	1.11	2.09

TABLE 3. **Example 1**. Overall minorant $\mathcal{J}^{\ominus}$ and overall majorant $\mathcal{J}^{\oplus}$, their parts, and their efficiency indices computed on a grid of size 256×256.

The approximations by the MhFEM are computed for the truncation index $N = 10$. For Example 2, it suffices to present here only the overall results as in Table 3, now presented in Table 4. We compare the overall values of majorants and minorants for different truncation indices $N = 6$ and $N = 10$, for which the corresponding truncation's remainder terms are given by $\mathcal{E}_6 = 44094.84$ and $\mathcal{E}_{10} = 10597.20$. One can see from the last two lines that the truncation index $N = 6$ suffices already to provide an accurate enough approximate solution. Also efficiency indices being around 1 show that the majorants and minorants perform well for that example.

mode	t^{sec}	$\mathcal{J}^{\ominus}$	$I_{\text{eff}}^{\mathcal{J}^{\ominus}}$	$\mathcal{J}^{\oplus}$	$I_{\text{eff}}^{\mathcal{J}^{\oplus}}$	$I_{\text{eff}}^{\mathcal{J}}$	$I_{\text{eff}}^{\mathcal{M}_1}$
$k = 0$	4.55	3.20e+05	0.90	3.56e+05	1.00	1.11	1.43
$k = 1$	4.53	1.02e+06	0.90	1.14e+06	1.00	1.11	3.19
$k = 2$	4.62	2.57e+05	0.90	2.85e+05	1.00	1.11	3.17
$k = 3$	4.80	6.07e+04	0.91	6.69e+04	1.00	1.10	1.82
$k = 4$	4.75	1.99e+04	0.91	2.19e+04	1.00	1.10	1.55
$k = 6$	4.83	4.05e+03	0.92	4.38e+03	1.00	1.08	1.32
$k = 8$	4.90	1.30e+03	0.93	1.40e+03	1.00	1.07	1.20
$k = 10$	4.72	5.40e+02	0.94	5.75e+02	1.00	1.06	1.14
overall ($N = 6$)	–	6.34e+06	0.90	7.06e+06	1.00	1.11	2.09
overall ($N = 10$)	–	6.34e+06	0.90	7.06e+06	1.00	1.11	2.09

TABLE 4. **Example 2**. Overall minorant $\mathcal{J}^{\ominus}$ and overall majorant $\mathcal{J}^{\oplus}$, their parts, and their efficiency indices computed on a grid of size 256×256.

7.1.3. *Example 3.* The desired state is chosen to be a space-time non-smooth function

$$y_d(\boldsymbol{x}, t) = \chi_{[\frac{1}{2},1]^2}(\boldsymbol{x})\, \chi_{[\frac{1}{4},\frac{3}{4}]}(t). \tag{81}$$

Here, we denote by χ the space-time characteristic function. The desired state's Fourier cofficients are analytically computed and given by $y_{d0}^c(\boldsymbol{x}) = \chi_{[\frac{1}{2},1]^2}(\boldsymbol{x})/2$ and

$$y_{dk}^c(\boldsymbol{x}) = \chi_{[\frac{1}{2},1]^2}(\boldsymbol{x})\frac{\left(\sin(\frac{3k\pi}{2}) - \sin(\frac{k\pi}{2})\right)}{k\pi} \qquad \text{and} \qquad y_{dk}^s(\boldsymbol{x}) = 0 \qquad \forall k \in \mathbb{N}. \tag{82}$$

The desired state has jumps in space **and** time. In this example, the exact solution cannot be precomputed analytically. Hence, as approximation for it, we use a MhFE representation on the finer grid with size 512×512. Note also that the Fourier coefficients are zero for the even Fourier modes besides for $k = 0$. We can observe in Table 5 that the efficiency indices, especially regarding the combined norm, are similar for higher modes. We have computed the results up to $N = 11$ and also added the results for the modes $k = 21$, $k = 41$ and $k = 81$ to the table as examples. The majorants of the combined norm stay approximately in the same range for $k \geq 1$. The majorants and minorants for the cost functional are close to 1, which demonstrates their efficiency also in this numerical example, where the given data has jumps in space and time.

mode	t^{sec}	$\mathcal{J}_k^{\ominus}$	$I_{\text{eff}}^{\mathcal{J}_k^{\ominus}}$	$\mathcal{J}_k^{\oplus}$	$I_{\text{eff}}^{\mathcal{J}_k^{\oplus}}$	$I_{\text{eff}}^{\mathcal{J},k}$	$I_{\text{eff}}^{\mathcal{M}_1,k}$
$k=0$	6.44	5.71e+04	0.92	8.20e+04	1.32	1.44	6.48
$k=1$	6.45	9.30e+04	0.92	1.31e+05	1.30	1.41	3.59
$k=3$	6.46	1.06e+04	0.95	1.35e+04	1.20	1.27	2.86
$k=5$	6.39	3.90e+03	0.97	4.63e+03	1.15	1.19	3.16
$k=7$	6.46	2.01e+03	0.98	2.28e+03	1.11	1.13	3.28
$k=9$	6.55	1.23e+03	0.98	1.35e+03	1.08	1.10	3.30
$k=11$	6.48	8.23e+02	0.99	8.89e+02	1.07	1.08	3.23
$k=21$	6.44	2.28e+02	1.00	2.37e+02	1.03	1.04	3.88
$k=41$	6.46	5.99e+01	1.00	6.19e+01	1.03	1.03	4.81
$k=81$	6.42	1.53e+01	1.00	1.63e+01	1.06	1.06	3.19

TABLE 5. **Example 3**. Minorants, majorants, and their efficiency indices as well as the efficiency indices of the combined norm computed on a grid of size 256×256.

7.2. **Numerical results for optimization problem II.** We compute the numerical results for the three same cases as for problem I but now applied on the desired gradient.

7.2.1. *Example 4.* We set the desired gradient to be time-periodic and time-analytic

$$\boldsymbol{g}_d(\boldsymbol{x},t) = \frac{e^t \sin(t)(-3\cos(t)(\cos(t)+\sin(t)) + (10\pi^2 + 1 + 2\pi^4)\sin(t)^2)}{10\pi}\begin{pmatrix}\cos(x_1\pi)\sin(x_2\pi)\\ \sin(x_1\pi)\cos(x_2\pi)\end{pmatrix}.$$

The exact solution for the state function is given by (76). Moreover, we present the efficiency indices for $\tilde{\mathcal{M}}_1^{\oplus}$ given for the modes by

$$I_{\text{eff}}^{\tilde{\mathcal{M}}_1,0} = \sqrt{\frac{\tilde{\mathcal{M}}_{1,0}^{\oplus}(\alpha_0,\beta_0; y_{0h}^c, p_{0h}^c, \boldsymbol{\tau}_{0h}^c, \boldsymbol{\rho}_{0h}^c)}{||||y_0^c - y_{0h}^c||||_{1,0}^2}} \qquad \text{and} \qquad I_{\text{eff}}^{\tilde{\mathcal{M}}_1,k} = \sqrt{\frac{\tilde{\mathcal{M}}_{1,k}^{\oplus}(\alpha_k,\beta_k; \boldsymbol{y}_{kh}, \boldsymbol{p}_{kh}, \boldsymbol{\tau}_{kh}, \boldsymbol{\rho}_{kh})}{||||\boldsymbol{y}_k - \boldsymbol{y}_{kh}||||_{1,k}^2}}.$$

The error norms for the modes are given by

$$||||y_0^c - y_{0h}^c||||_{1,0}^2 = \left(\frac{1}{2} + \frac{\lambda\underline{\mu_1}^2}{2C_F^2}\right)\|\nabla y_0^c - \nabla y_{0h}^c\|_\Omega^2$$

and

$$||||\boldsymbol{y}_k - \boldsymbol{y}_{kh}||||_{1,k}^2 = \frac{k\omega\lambda\underline{\mu_1}^2}{2C_F^2}\|\boldsymbol{y}_k - \boldsymbol{y}_{kh}\|_\Omega^2 + \left(\frac{1}{2} + \frac{\lambda\underline{\mu_1}^2}{2C_F^2}\right)\|\nabla\boldsymbol{y}_k - \nabla\boldsymbol{y}_{kh}\|_\Omega^2$$

leading the representation

$$|||u-v|||_1^2 = T\,|||y_0^c - y_{0h}^c|||_{1,0}^2 + \frac{T}{2}\sum_{k=1}^{N} |||\boldsymbol{y}_k - \boldsymbol{y}_{kh}|||_{1,k}^2 + \tilde{\mathcal{F}}_N \tag{83}$$

with the remainder term $\tilde{\mathcal{F}}_N := \frac{T}{2}\sum_{k=N+1}^{\infty} |||\boldsymbol{y}_k|||_{1,k}^2$. For the numerical experiments, the efficiency index for $\tilde{\mathcal{M}}_1^{\oplus}$ from above by estimating (83) from below ignoring the remainder term $\tilde{\mathcal{F}}_N$ leading to the overall efficiency index for $\tilde{\mathcal{M}}_1^{\oplus}$ given by

$$I_{\text{eff}}^{\tilde{\mathcal{M}}_1} = \sqrt{\frac{\tilde{\mathcal{M}}_1^{\oplus}(\alpha,\beta;\eta,\zeta,\boldsymbol{\tau},\boldsymbol{\rho},v)}{T\,|||y_0^c - y_{0h}^c|||_{1,0}^2 + \frac{T}{2}\sum_{k=1}^{N} |||\boldsymbol{y}_k - \boldsymbol{y}_{kh}|||_{1,k}^2}}.$$

The corresponding majorants are given by $\tilde{\mathcal{M}}_{1,0}^{\oplus} = \tilde{\mathcal{M}}_{1,0}^{\oplus}(\alpha_0,\ \beta_0; y_{0h}^c, p_{0h}^c, \boldsymbol{\tau}_{0h}^c, \boldsymbol{\rho}_{0h}^c) = \tilde{\mathcal{J}}_0^{\oplus} - \tilde{\mathcal{J}}_0^{\ominus} + \frac{3\lambda}{4C_F^2}(C_F\|\mathcal{R}_{1}{}_0^c\|_\Omega + \|\mathcal{R}_{2}{}_0^c\|_\Omega)^2$ and $\tilde{\mathcal{M}}_{1,k}^{\oplus} = \tilde{\mathcal{M}}_{1,k}^{\oplus}(\alpha_k,\ \beta_k; \boldsymbol{y}_{kh}, \boldsymbol{p}_{kh}, \boldsymbol{\tau}_{kh}, \boldsymbol{\rho}_{kh}) = \tilde{\mathcal{J}}_k^{\oplus} - \tilde{\mathcal{J}}_k^{\ominus} + \frac{3\lambda}{4C_F^2}\left(C_F\|\mathcal{R}_{1k}\|_\Omega + \|\mathcal{R}_{2k}\|_\Omega\right)^2$. We present the numerical results for the modes $k=0$ and $k=1$ for different grid sizes in Tables 6 and 7. The efficiency indices for the majorants and minorants are very close to 1.00. Also the efficiency indices for $\tilde{\mathcal{M}}_{1,0}$ show a good accuracy. Table 8 compares

grid	t^{sec}	$\tilde{\mathcal{J}}_0^{\ominus}$	$I_{\text{eff}}^{\tilde{\mathcal{J}}_0^{\ominus}}$	$\tilde{\mathcal{J}}_0^{\oplus}$	$I_{\text{eff}}^{\tilde{\mathcal{J}}_0^{\oplus}}$	$I_{\text{eff}}^{\tilde{\mathcal{J}},0}$	$I_{\text{eff}}^{\tilde{\mathcal{M}}_1,0}$
16×16	0.02	8.92e+03	0.99	9.85e+03	1.09	1.10	2.04
32×32	0.06	9.24e+03	0.99	9.95e+03	1.07	1.08	1.97
64×64	0.24	9.32e+03	0.99	9.97e+03	1.05	1.07	1.94
128×128	1.04	9.34e+03	0.98	9.98e+03	1.05	1.07	1.93
256×256	4.39	9.35e+03	0.98	9.98e+03	1.05	1.07	1.92

TABLE 6. **Example 4**. Minorant $\tilde{\mathcal{J}}_0^{\ominus}$, majorant $\tilde{\mathcal{J}}_0^{\oplus}$ and their efficiency indices computed on grids of different sizes.

grid	t^{sec}	$\tilde{\mathcal{J}}_1^{\ominus}$	$I_{\text{eff}}^{\tilde{\mathcal{J}}_1^{\ominus}}$	$\tilde{\mathcal{J}}_1^{\oplus}$	$I_{\text{eff}}^{\tilde{\mathcal{J}}_1^{\oplus}}$	$I_{\text{eff}}^{\tilde{\mathcal{J}},1}$	$I_{\text{eff}}^{\tilde{\mathcal{M}}_1,1}$
16×16	0.02	3.22e+04	0.95	3.51e+04	1.03	1.09	1.52
32×32	0.06	3.39e+04	0.96	3.58e+04	1.02	1.05	1.33
64×64	0.26	3.45e+04	0.97	3.60e+04	1.01	1.04	1.25
128×128	1.08	3.48e+04	0.97	3.62e+04	1.01	1.04	1.21
256×256	4.31	3.51e+04	0.98	3.65e+04	1.01	1.04	1.19

TABLE 7. **Example 4**. Minorant $\tilde{\mathcal{J}}_1^{\ominus}$, majorant $\tilde{\mathcal{J}}_1^{\oplus}$ and their efficiency indices computed on grids of different sizes.

the results for different Fourier modes up to $N=8$ computed on a grid of size 256×256. Here, the overall minorants, majorants and efficiency indices are presented, where the remainder terms $\mathcal{E}_N$ for $N=8$ and also $N=6$ have been precomputed exactly. The values of the efficiency indices vary for different modes k. For example, the results for $\tilde{\mathcal{M}}_{1,4}$ indicate that the mode $k=4$ is essential to represent the solution accurately. The values for $I_{\text{eff}}^{\mathcal{J}_7^{\ominus}}$ and $I_{\text{eff}}^{\mathcal{J}_8^{\ominus}}$ indicate that the minorants require a different, higher refinement for a more accurate representation of the overall solution. An adaptive scheme is the natural choice. On the other hand, in these cases the majorants give a good representation for the cost functional. Finally, comparing the last two lines of Table 8 again shows that the overall value for representing the cost functional of the exact solution is already sufficiently accurate for a truncation index $N=6$.

mode	t^{sec}	$\tilde{\mathcal{J}}^{\ominus}$	$I_{\text{eff}}^{\tilde{\mathcal{J}}^{\ominus}}$	$\tilde{\mathcal{J}}^{\oplus}$	$I_{\text{eff}}^{\tilde{\mathcal{J}}^{\oplus}}$	$I_{\text{eff}}^{\tilde{\mathcal{J}}}$	$I_{\text{eff}}^{\tilde{\mathcal{M}}_1}$
$k = 0$	4.39	9.35e+03	0.98	9.98e+03	1.05	1.07	1.92
$k = 1$	4.31	3.51e+04	0.98	3.65e+04	1.01	1.04	1.19
$k = 2$	4.42	9.23e+03	0.63	1.57e+04	1.06	1.70	1.65
$k = 3$	4.43	2.85e+03	0.58	5.06e+03	1.03	1.78	1.08
$k = 4$	4.44	5.89e+02	0.98	6.47e+02	1.07	1.10	5.91
$k = 6$	4.36	1.74e+01	0.97	2.82e+01	1.58	1.62	2.32
$k = 8$	4.33	2.99e-01	0.23	1.37e+00	1.05	4.59	1.97
overall ($N = 6$)	–	2.09e+05	0.89	2.45e+05	1.04	1.17	2.46
overall ($N = 8$)	–	2.09e+05	0.89	2.45e+05	1.04	1.17	2.46

TABLE 8. **Example 4**. Overall minorant $\tilde{\mathcal{J}}^{\ominus}$ and overall majorant $\tilde{\mathcal{J}}^{\oplus}$, their parts, and their efficiency indices computed on a grid of size 256×256.

7.2.2. *Example 5.* We choose the non time-periodic but time-analytic desired gradient

$$\boldsymbol{g}_d(\boldsymbol{x}, t) = \frac{-e^t \sin(t)(0.1\cos(t) - \pi^2(1 + 2\pi^2 0.1))}{\pi} \begin{pmatrix} \cos(x_1\pi)\sin(x_2\pi) \\ \sin(x_1\pi)\cos(x_2\pi) \end{pmatrix}$$

leading to the time-analytic, but not time-periodic exact state (80). We compute the MhFE approximation of the desired gradient and solve the systems (49) and (50) for modes up to $N = 10$ on a 256×256-mesh and present the results in Table 9. The remainder terms for $N = 6$ and $N = 10$ are $\mathcal{E}_6 = 4796.54$ and $\mathcal{E}_{10} = 1149.65$, respectively. The efficiency indices for the overall majorant and minorant show that a truncation index of $N = 6$ already gives a sufficiently accurate approximation for the overall cost functional. Note that the efficiency index for $\tilde{\mathcal{M}}_{1,2}$ indicates that the mode $k = 2$ is essential for the multiharmonic approximation giving an accurate representation of the solution.

mode	t^{sec}	$\tilde{\mathcal{J}}^{\ominus}$	$I_{\text{eff}}^{\tilde{\mathcal{J}}^{\ominus}}$	$\tilde{\mathcal{J}}^{\oplus}$	$I_{\text{eff}}^{\tilde{\mathcal{J}}^{\oplus}}$	$I_{\text{eff}}^{\tilde{\mathcal{J}}}$	$I_{\text{eff}}^{\tilde{\mathcal{M}}_1}$
$k = 0$	4.31	2.63e+04	1.00	2.79e+04	1.06	1.06	1.36
$k = 1$	4.30	8.49e+04	1.00	8.60e+04	1.02	1.01	1.00
$k = 2$	4.39	2.08e+04	0.98	2.21e+04	1.04	1.06	2.83
$k = 4$	4.35	1.58e+03	0.96	1.75e+03	1.06	1.11	1.74
$k = 6$	4.36	2.93e+02	0.87	3.53e+02	1.05	1.20	1.63
$k = 8$	4.37	8.95e+01	0.82	1.20e+02	1.10	1.34	1.08
$k = 10$	4.34	3.27e+01	0.71	5.23e+01	1.14	1.60	1.19
overall ($N = 6$)	–	5.23e+05	1.00	5.43e+05	1.04	1.04	2.00
overall ($N = 10$)	–	5.22e+05	1.00	5.42e+05	1.04	1.04	2.00

TABLE 9. **Example 5**. Overall minorant $\tilde{\mathcal{J}}^{\ominus}$ and overall majorant $\tilde{\mathcal{J}}^{\oplus}$, their parts, and efficiency indices of them and the combined norm computed on a grid of size 256×256.

7.2.3. *Example 6.* We set the space-time non-smooth desired gradient

$$\boldsymbol{g}_d(\boldsymbol{x}, t) = (\chi_{[\frac{1}{2},1]^2}(\boldsymbol{x})\, \chi_{[\frac{1}{4},\frac{3}{4}]}(t), \chi_{[\frac{1}{2},1]^2}(\boldsymbol{x})\, \chi_{[\frac{1}{4},\frac{3}{4}]}(t))^T. \tag{84}$$

Also the coefficients of the Fourier expansion associated with $\boldsymbol{g}_d$ can be found analytically. They are as in Example 3 given by (82) for each direction of the gradient (84). Again the exact solution cannot be computed analytically and hence we use its MhFE approximations on the finer mesh of size 512×512 as a reference. Table 10 presents the results for modes up to truncation index $N = 11$ as well as for $k = 21$, $k = 41$ and $k = 81$ analogously to Example 3. The results reflected by the efficiency indices show the good representation by using the minorants and majorants, especially, considering the efficiency indices in the last two columns of Table 10. This again demonstrates the efficiency of the minorants and majorants for data having jumps in space and time but now for optimization problem II.

mode	t^{sec}	$\mathcal{J}_k^{\ominus}$	$I_{\text{eff}}^{\mathcal{J}_k^{\ominus}}$	$\mathcal{J}_k^{\oplus}$	$I_{\text{eff}}^{\mathcal{J}_k^{\oplus}}$	$I_{\text{eff}}^{\mathcal{J},k}$	$I_{\text{eff}}^{\mathcal{M}_1,k}$
$k=0$	8.12	1.55e+01	0.79	2.05e+01	1.04	1.32	2.23
$k=1$	8.03	7.70e+00	0.86	1.13e+01	1.26	1.47	2.87
$k=3$	8.51	1.46e+01	0.93	1.73e+01	1.10	1.18	2.56
$k=5$	8.25	1.44e+01	0.99	1.65e+01	1.14	1.15	2.59
$k=7$	8.08	8.92e+00	0.98	9.80e+00	1.08	1.10	1.18
$k=9$	8.35	4.51e+00	0.90	5.14e+00	1.03	1.14	1.26
$k=11$	8.36	2.57e+00	0.96	3.13e+00	1.17	1.22	1.51
$k=21$	8.51	1.36e+00	0.99	2.27e+00	1.65	1.67	3.16
$k=41$	7.78	4.54e+00	0.86	6.26e+00	1.18	1.38	3.09
$k=81$	7.81	3.19e+00	0.79	7.09e+00	1.75	2.22	3.44

TABLE 10. **Example 6**. Minorants, majorants, and their efficiency indices as well as the efficiency indices of the combined norm computed on a grid of size 256×256.

8. Conclusions and outlook

In this work, the a posteriori error analysis started in [26] has been extended now by deriving new lower bounds, called minorants, for the cost functional leading to an upper estimate for the error norm of the state and control or equivalently in state and adjoint state. These lower bounds are guaranteed and computable. Together with using the results from [44] as well as [16] one can apply the method also to time-periodic optimal control problems, where box constraints are being imposed on the Fourier coefficients of the control. The estimates are derived for two different cost functionals, where the second one is now new in this context.

Since in the linear case the problems are decoupled, the solutions on the Fourier coefficients could easily be computed on grids of different sizes depending on the accuracy needed, which could be exactly determined by using the a posteriori estimates presented in this work leading to an adaptive method in time. Together with the adaptive finite element method we then obtain a space-time adaptive method, the adaptive multiharmonic finite element method, which we call AMhFEM, as mentioned for the first time in [26].

In this work, a first derivation of preconditioners for the MinRes method for the second optimization problem has been presented as well as a preconditoner for applying AMLI has been suggested. Several numerical tests for optimization problem I and II have been presented showing the efficiency of the upper and especially – with regard to the article – lower bounds for the cost functionals in practice.

Acknowledgments

The author gratefully acknowledges the financial support by the Academy of Finland under the grant 295897. The author would like to thank the anonymous referees for their valuable comments improving the article.

References

[1] K. Altmann, S. Stingelin, and F. Tröltzsch, *On some optimal control problems for electrical circuits*, Int. J. Circuit Theory Appl., 42 (2014), pp. 808–830.

[2] O. Axelsson and D. Lukáš, *Preconditioning methods for eddy-current optimally controlled time-harmonic electromagnetic problems*, J. Numer. Math., 27 (2019), pp. 1–21.

[3] F. Bachinger, M. Kaltenbacher, and S. Reitzinger, *An efficient solution strategy for the HBFE method*, Proceedings of the IGTE '02 Symposium Graz, Austria, 2 (2002), pp. 385–389.

[4] F. Bachinger, U. Langer, and J. Schöberl, *Numerical analysis of nonlinear multiharmonic eddy current problems*, Numer. Math., 100 (2005), pp. 593–616.

[5] ———, *Efficient solvers for nonlinear time-periodic eddy current problems*, Comput. Vis. Sci., 9 (2006), pp. 197–207.

[6] D. Braess, *Finite elements: Theory, fast solvers, and applications in solid mechanics*, Cambridge University Press, third edition, 2007.

[7] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, vol. 4 of Studies in Mathematics and its Applications, North-Holland, Amsterdam, 1978. Republished by SIAM in 2002.

[8] C. Clason, B. Kaltenbacher, and D. Wachsmuth, *Functional error estimators for the adaptive discretization of inverse problems*, Inverse Probl., 32 (2016), p. 104004.

[9] D. M. Copeland and U. Langer, *Domain decomposition solvers for nonlinear multiharmonic finite element equations*, J. Numer. Math., 18 (2010), pp. 157–175.

[10] A. Gaevskaya, R. H. W. Hoppe, and S. Repin, *A posteriori estimates for cost functionals of optimal control problems*, Numerical Mathematics and Advanced Applications, Proceedings of the ENUMATH, 2005 (2006), pp. 308–316.

[11] ———, *Functional approach to a posteriori error estimation for elliptic optimal control problems with distributed control*, J. Math. Sci., 144 (2007), pp. 4535–4547.

[12] A. V. Gaevskaya and S. I. Repin, *A posteriori error estimates for approximate solutions of linear parabolic problems*, Differ. Equ., 41 (2005), pp. 970–983.

[13] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, Frontiers in Applied Mathematics 17, SIAM, Philadelphia, 1997.

[14] M. Hinze and A. Rösch, *Discretization of optimal control problems*, in Constrained Optimization and Optimal Control for Partial Differential Equations, Springer, 2012, pp. 391–430.

[15] B. Houska, F. Logist, J. Van Impe, and M. Diehl, *Approximate robust optimization of time-periodic stationary states with application to biochemical processes*, in Proceedings of the 48th IEEE Conference on Decision and Control, Shanghai, China, 2009, pp. 6280–6285.

[16] M. Kollmann and M. Kolmbauer, *A preconditioned MinRes solver for time-periodic parabolic optimal control problems*, Numer. Linear Algebra Appl., 20 (2013), pp. 761–784.

[17] M. Kollmann, M. Kolmbauer, U. Langer, M. Wolfmayr, and W. Zulehner, *A robust finite element solver for a multiharmonic parabolic optimal control problem*, Comput. Math. Appl., 65 (2013), pp. 469–486.

[18] M. Kolmbauer and U. Langer, *A robust preconditioned MinRes solver for distributed time-periodic eddy current optimal control problems*, SIAM J. Sci. Comput., 34 (2012), pp. B785–B809.

[19] ———, *Efficient solvers for some classes of time-periodic eddy current optimal control problems*, Numerical Solution of Partial Differential Equations: Theory, Algorithms, and Their Applications, 45 (2013), pp. 203–216.

[20] J. Kraus, *Additive Schur complement approximation and application to multilevel preconditioning*, SIAM J. Sci. Comput., 34 (2012), pp. A2872–A2895.

[21] J. Kraus and M. Wolfmayr, *On the robustness and optimality of algebraic multilevel methods for reaction-diffusion type problems*, Comput. Vis. Sci., 16 (2013), pp. 15–32.

[22] Y. A. Kuznetsov, *Efficient iterative solvers for elliptic finite element problems on nonmatching grids*, Russ. J. Numer. Anal. Math. Model., 10 (1995), pp. 187–211.

[23] O. A. Ladyzhenskaya, *The Boundary Value Problems of Mathematical Physics*, Nauka, Moscow, 1973. In Russian. Translated in *Appl. Math. Sci.* 49, Springer, 1985.

[24] O. A. Ladyzhenskaya, V. A. Solonnikov, and N. N. Ural'ceva, *Linear and Quasilinear Equations of Parabolic Type*, AMS, Providence, RI, 1968.

[25] U. Langer, S. Repin, and M. Wolfmayr, *Functional a posteriori error estimates for parabolic time-periodic boundary value problems*, Comput. Methods Appl. Math., 15 (2015), pp. 353–372.

[26] U. Langer, S. Repin, and M. Wolfmayr, *Functional a posteriori error estimates for time-periodic parabolic optimal control problems*, Num. Func. Anal. Opt., 37 (2016), pp. 1267–1294.

[27] U. Langer and M. Wolfmayr, *Multiharmonic finite element analysis of a time-periodic parabolic optimal control problem*, J. Numer. Math., 21 (2013), pp. 265–300.

[28] Z. Z. Liang, O. Axelsson, and M. Neytcheva, *A robust structured preconditioner for time-harmonic parabolic optimal control problems*, Numer. Algorithms, 79 (2018), pp. 575–596.

[29] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations (Grundlehren der Mathematischen Wissenschaften)*, vol. 170, Springer Berlin, 1971.

[30] O. Mali, *Error estimates for a class of elliptic optimal control problems*, Num. Func. Anal. Opt., 38 (2017), pp. 58–79.

[31] O. Mali, P. Neittaanmäki, and S. Repin, *Accuracy Verification Methods. Theory and Algorithms*, vol. 32 of Computational Methods in Applied Sciences, Springer, Netherlands, 2014.

[32] S. Matculevich and S. Repin, *Explicit constants in Poincaré-type inequalities for simplicial domains and application to a posteriori estimates*, Comput. Methods Appl. Math., 16 (2016), pp. 277–298.

[33] S. Matculevich and M. Wolfmayr, *On the a posteriori error analysis for linear Fokker–Planck models in convection-dominated diffusion problems*, Appl. Math. Comput., 339 (2018), pp. 779–804.

[34] S. G. Mikhlin, *Variational methods in mathematical physics*, Pergamon Press Oxford, 1964.

[35] M. F. Murphy, G. H. Golub, and A. J. Wathen, *A note on preconditioning for indefinite linear systems*, SIAM J. Sci. Comput., 21 (2000), pp. 1969–1972.

[36] P. Neittaanmäki, J. Sprekels, and D. Tiba, *Optimization of Elliptic Systems: Theory and Applications*, Springer, 2006.
[37] C. C. Paige and M. A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.
[38] P. A. Raviart and J. M. Thomas, *A mixed finite element method for 2-nd order elliptic problems*, Mathematical Aspects of Finite Element Methods, Lect. Notes Math., 606 (1977), pp. 292–315.
[39] S. Repin, *Estimates of deviation from exact solutions of initial-boundary value problems for the heat equation*, Rend. Mat. Acc. Lincei, 13 (2002), pp. 121–133.
[40] ———, *A Posteriori Estimates for Partial Differential Equations*, Radon Series on Computational and Applied Mathematics 4, Walter de Gruyter, Berlin, 2008.
[41] S. Repin and T. Rossi, *On quantitative analysis of an ill-posed elliptic problem with Cauchy boundary conditions*, in Numerical Methods for Differential Equations, Optimization, and Technological Problems, Springer, 2013, pp. 133–144.
[42] P. Shakya and R. K. Sinha, *A priori and a posteriori error estimates of H1-Galerkin mixed finite element method for parabolic optimal control problems*, Optim. Control Appl. Methods, 38 (2017), pp. 1056–1070.
[43] O. Steinbach, *Numerical approximation methods for elliptic boundary value problems: finite and boundary elements*, Springer Science & Business Media, 2007.
[44] M. Wolfmayr, *A note on functional a posteriori estimates for elliptic optimal control problems*, Numer. Methods Partial Differ. Equ., 33 (2017), pp. 403–424.
[45] S. Yamada and K. Bessho, *Harmonic field calculation by the combination of finite element analysis and harmonic balance method*, IEEE Trans. Magn., 24 (1988), pp. 2588–2590.
[46] W. Zulehner, *Nonstandard norms and robust estimates for saddle point problems*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 536–560.

(M. Wolfmayr) Faculty of Information Technology; University of Jyväskylä; P.O. Box 35 (Agora); FIN-40014 Jyväskylä; Finland
E-mail address: monika.k.wolfmayr@jyu.fi