# RELIABLE A POSTERIORI ERROR ESTIMATION FOR STATE-CONSTRAINED OPTIMAL CONTROL

ARND RÖSCH, KUNIBERT G. SIEBERT, SIMEON STEINIG

Abstract. We derive a reliable a posteriori error estimator for a state-constrained elliptic optimal control problem taking into account both regularisation and discretisation. The estimator is applicable to finite element discretisations of the problem with both discretised and non-discretised control. The performance of our estimator is illustrated by several numerical examples for which we also introduce an adaption strategy for the regularisation parameter.

## 1. Introduction

In many applications in science and engineering optimisation processes lead to optimal control problems where the sought after state is a solution to a partial differential equation (PDE) governed by a control. Often, there are additional constraints on the control and/or state. These kind of problems are frequently highly complex, that is why a numerical approximation of the true solution requires high computational effort. Thus, it is of great interest to construct efficient solution algorithms reducing this effort and still guaranteeing a precise approximation. One particular method is the adaptive finite element method based on the adaptive loop:

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}.$$

Let us briefly explain the different modules: The module 'SOLVE' consists of the solution of a finite element version of the original, continuous optimal control problem. This solution can then be used to compute local indicators and a global estimator for the true error in the module 'ESTIMATE'. Based on the local indicators, the module 'MARK' collects a subset of elements chosen for refinement whose execution is then the subject of the module 'REFINE' at the end of which we are in possession of a suitably refined grid so that the loop can potentially start anew.

In this paper, our sole focus lies on the module 'ESTIMATE', i.e. we will present a *reliable* a posteriori error estimator based on the discrete solution and given problem data, that is a numerically evalubale upper bound for the error between the true solution and the current discrete approximation up to constants depending solely on data.

Research into adaptive finite element methods for elliptic PDEs has already been going on for a fairly long time, starting with the paper [2]. Nowadays, we have every justification to claim that there exists an established and refined theory for different kind of discretisations, error estimators and general inf-sup stable problems, the interested reader is referred to the monographs [1, 3, 4, 45] and the included references.

In contrast, the investigation of adaptive finite element methods in the context of optimal control problems has only started fairly recently, to the best of the authors' knowlege [33] is very much the first. At present, there have been several results on unconstrained problems and control-constrained ones. We would like to mention e.g. [6, 32, 21, 19, 46, 26].

However, in this paper, we are concerned with state-constrained optimal control problems which are especially hard to treat both from an analytical and numerical viewpoint.

Analytically, the crucial difficulty lies in the irregularity of the Lagrange multiplier for the state

constraint (compare [12] for the general result and e.g. [14, 35, 34] for several examples), indeed it is in general only a measure, possibly non-unique and directly results in a low regularity for the adjoint state which is in general not an $H^1(\Omega)$-function. Numerically, this low regularity and non-uniqueness is a key reason why fast Newton-type algorithms are not available for the solution of the optimisation problem. Invariably, the optimal control problems need to be regularised. Several different techniques are available, the Moreau-Yosida approach, e.g. [18, 9], interior point methods (compare [24, 39]) or the virtual control approach employed by us (see e.g. [29, 28]). Yet all of them have in common that one encounters an additional error, the regularisation error. Hence, an a posteriori error estimator should be based on the discrete solution to a regularised problem, because this problem is ultimately the one which is numerically solved and thus necessarily the one from which we have to extract information to estimate the difference between the true solution and the current discrete, regularised one. Therefore our aim was to derive a numerically evaluable upper bound for the error between the true solution and the discrete, regularised one - that is a *reliable* a posteriori error estimator - and thus dispense with analysing any unregularised discrete problem, because these problems are almost never actually solved.
There have been several results on a posteriori error estimators for state-constrained problems, most operate in the dual weighted residual (DWR) framework (compare e.g [17],[7] or [47] and [20], the latter in a mixed control-state constrained setting). Naturally, no reliable a posteriori error estimator is provided. The onus of these works is different.
In a non-DWR framework, there have been - to the best of the authors' knowledge - precious few results: In [25] a residual error estimator based bound was derived, however, the bound in this paper is not reliable because terms associated with the discrete and continuous Lagrange multipliers for the state constraint are neglected. In contrast, in [38] an upper bound for the error between the true solution and an unregularised discrete one was proven, yet, aside from the fact that the error due to regularisation is not taken into account the key drawback perhaps is that $L_\infty(\Omega)$-estimators for the linear errors are needed which are known to be either efficient or reliable but not both. This is a severe obstacle when analysing convergence of the resulting estimator. Indeed, to the best of the authors' knowledge there is no general convergence result in the $L_\infty(\Omega)$-norm in the vein of [36] and [41] available for adaptive finite element discretisations of PDEs.
In this paper we now supplement these existing results by deriving a reliable a posteriori error estimator for the error between the true solution and a regularised discrete one taking into account both regularisation and discretisation error. In addition, the error estimator is based solely on $L_2(\Omega)$ and $H^1(\Omega)$-norm and scalar product terms which are best-suited to a convergence analysis in an adaptive finite element setting in the vein of [36] and [41].

The paper is organised as follows: First, we introduce the continuous and discrete problems analysed, then in Section 3, we state our main theorems: In Theorem 3.4 we derive a reliable a posteriori bound for the error between the true solution and the regularised discrete one for the variational discretisation approach due to [22], Theorem 3.5 does the same for the fully-discrete one, where the controls are discretised by piecewise constant functions. In Sections 4 and 5 these results are proven. In Section 6 we check the performance of our estimator numerically. In this section, we also introduce a possible way to adapt the regularisation parameter which works very well for the examples considered. We finish the paper by summarising our results and offering a view on future research possibilities.

## 2. Problem Setting: Continuous and Discrete Problem

2.1. **The Continuous Problem.** We examine the following elliptic model state-constrained optimal control problem on a polygonal respectively polyhedral domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, with boundary

$\partial\Omega$:

$$(P) \qquad \begin{cases} \min\limits_{u\in L_2(\Omega), y\in H_0^1(\Omega)} J(y,u) = \frac{1}{2}\,\|y-y_d\|^2_{L_2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L_2(\Omega)} \\ \text{s.t.} \\ -\Delta y = u \text{ in } \Omega \\ y = 0 \text{ on } \partial\Omega \\ \text{and} \\ a \le u \le b \text{ a.e in } \Omega,\ a,b \in \mathbb{R}\cup\{-\infty,\infty\} \\ y_c \le y \text{ a.e. in } \Omega,\ y_c|_{\partial\Omega} < 0 \end{cases}$$

Here, $y_d \in L_2(\Omega)$ and $y_c \in H^1(\Omega)\cap C(\bar{\Omega})$ with $\nabla y_c \in H(\mathrm{div},\Omega)$.
For the Poisson equation in $(P)$ we introduce a linear and continuous solution operator $S : L_2(\Omega) \to H_0^1(\Omega)$ mapping each right hand side $u \in L_2(\Omega)$ to the unique solution $y = Su$ of

$$(\nabla y, \nabla w)_{L_2(\Omega)} = (u, w)_{L_2(\Omega)} \qquad \forall w \in H_0^1(\Omega).$$

In addition, we observe that $\nabla y \in H(\mathrm{div},\Omega)$. For our analyses of Section 2-5 we do not need any higher regularity. In particular, $\Omega$ may possess very large maximal interior angles.
The adjoint operator to $S$ with respect to the $L_2(\Omega)$-scalar product, denoted by $S^*$, which we will frequently need throughout this paper, also maps $L_2(\Omega)$ to $H_0^1(\Omega)$ linear and continuously. For each $q \in L_2(\Omega)$ it is given by the unique solution $p = S^*q \in H_0^1(\Omega)$, $\nabla p \in H(\mathrm{div},\Omega)$, of

$$(\nabla p, \nabla w)_{L_2(\Omega)} = (q, w)_{L_2(\Omega)}.$$

We note that $S^*$ is self-adjoint.
With the solution operator $S$ one can reduce the objective $J$ in $(P)$ by eliminating the state $y$ in the following way:

$$f(u) := J(Su,u) = \frac{1}{2}\,\|Su - y_d\|^2_{L_2(\Omega)} + \frac{\nu}{2}\,\|u\|^2_{L_2(\Omega)}\,. \tag{2.1}$$

We assume that the feasible set $\mathbb{U}^{ad}$ defined by

$$\mathbb{U}^{ad} := \{u \in L_2(\Omega) \,:\, a \le u \le b \text{ a.e. in } \Omega,\ y_c \le Su \text{ a.e. in } \Omega\}$$

is non-empty. Then it is well-known that $(P)$ has a unique solution $\bar{u}$ with associated state $\bar{y} = S\bar{u}$, because $(P)$ is a strictly convex problem.
It is equally well-known that problems of type $(P)$ suffer from a lack of regularity for the Lagrange multiplier for the state constraint, in general it is only in $C(\bar{\Omega})^*$ and there are cases where it is not unique. Closely related to this lack of regularity is the fact that no fast Newton-type optimisation algorithms are available for the numerical solution of problems of type $(P)$. To remedy this obstacle, these problems are regularised and the resulting regularised problems are then discretised and numerically solved. Thus, the regularised and discretised problem has to be the foundation of every error estimator and as a consequence, we are not concerned with any unregularised discrete version of the original optimal control problem $(P)$.
Thus, naturally, a crucial tool for our analysis will be a sequence of continuous regularised problems. We employ the virtual control regularisation approach, e.g. [29], which in our formulation is equivalent to a Moreau-Yosida type regularisation (see e.g. [18]), and introduce the following problem:

$$(P^\varepsilon)\qquad \begin{cases} \min\limits_{u\in L_2(\Omega),y\in H_0^1(\Omega),v\in L_2(\Omega)} J^\varepsilon(y,u,v) = \frac{1}{2}\|y-y_d\|^2_{L_2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L_2(\Omega)} + \frac{1}{\varepsilon}\|v\|^2_{L_2(\Omega)} \\ \text{s.t.} \\ -\Delta y = u \text{ in } \Omega \\ y = 0 \text{ on } \partial\Omega \\ \text{and} \\ a \le u \le b \text{ a.e in } \Omega,\ a,b\in\mathbb{R}\cup\{-\infty,\infty\} \\ y_c \le y+\varepsilon v \text{ a.e. in } \Omega,\ y_c|_{\partial\Omega} < 0 \end{cases}$$

The feasible set to this problem is given by

$$\mathbb{U}^{\varepsilon,ad} := \{(u,v)\in L_2(\Omega)\times L_2(\Omega) \,:\, a\le u\le b \text{ a.e. in } \Omega,\ Su+\varepsilon v \ge y_c \text{ a.e in } \Omega\}$$

and the reduced goal functional to $J^\varepsilon$ by:

$$f^\varepsilon(u,v) := J^\varepsilon(y,u,v) = \frac{1}{2}\|Su-y_d\|^2_{L_2(\Omega)} + \frac{\nu}{2}\|u\|^2 + \frac{1}{\varepsilon}\|v\|^2_{L_2(\Omega)}. \tag{2.2}$$

Since $(P^\varepsilon)$ is a strictly convex problem, it is well-known that this problem possesses a unique solution couple $(\bar u^\varepsilon, \bar v^\varepsilon)$ with associated state $\bar y^\varepsilon = S\bar u^\varepsilon$ for every fixed $\varepsilon > 0$ fulfilling the following necessary and sufficient first order optimality condition with $\bar p_r^\varepsilon = S^*(\bar y^\varepsilon - y_d)$:

$$(\bar p_r^\varepsilon + \nu\bar u^\varepsilon, u-\bar u^\varepsilon)_{L_2(\Omega)} + \frac{1}{\varepsilon}(\bar v^\varepsilon, v-\bar v^\varepsilon)_{L_2(\Omega)} \ge 0\ \forall (u,v)\in\mathbb{U}^{ad,\varepsilon}. \tag{2.3}$$

Defining

$$\mathcal{U} := \{u\in L_2(\Omega) \,:\, a\le u\le b\}, \tag{2.4}$$

we realise that for every fixed $\varepsilon > 0$ the constraint mapping for the state constraint $M^\varepsilon(u,v) := y_c - Su - \varepsilon v$ is surjective as a mapping $M^\varepsilon : \mathcal{U}\times L_2(\Omega)\to L_2(\Omega)$. This ensures the existence of a Lagrange multiplier $\bar\theta^\varepsilon \in L_2(\Omega)$, $\bar\theta^\varepsilon \ge 0$ a.e. in $\Omega$, such that the following KKT system with $\bar p^\varepsilon = S^*(\bar y^\varepsilon - y_d - \bar\theta^\varepsilon)$ is fulfilled (compare [31]):

$$(\bar p^\varepsilon + \nu\bar u^\varepsilon, u-\bar u^\varepsilon)_{L_2(\Omega)} \ge 0\ \forall u\in\mathcal{U} \tag{2.5a}$$

$$\frac{1}{\varepsilon^2}\bar v^\varepsilon = \bar\theta^\varepsilon \tag{2.5b}$$

$$\bar\theta^\varepsilon \ge 0,\ (\bar\theta^\varepsilon, \bar y^\varepsilon - y_c + \varepsilon\bar v^\varepsilon)_{L_2(\Omega)} = 0 \tag{2.5c}$$

Note that (2.5a) is equivalent to the pointwise projection formula:

$$\bar u^\varepsilon(x) = \max(\min(-\frac{1}{\nu}\bar p^\varepsilon(x), b)a) =: \Pi(\bar p^\varepsilon)(x),\ \text{f.a.a. } x\in\Omega. \tag{2.6}$$

with

$$\Pi(f)(x) := \max(\min(-\frac{1}{\nu}f(x), b), a),\quad \text{f.a.a. } x\in\Omega,\ f\in L_2(\Omega) \tag{2.7}$$

Let us close this section with some remarks about the properties of an optimal value function associated to $(P)$ and $(P^\varepsilon)$. With the reduced goal functionals for the unregularised problem,(2.1), and the regularised one,(2.2), we set

$$\begin{aligned} a(\varepsilon) &:= f^\varepsilon(\bar u^\varepsilon, \bar v^\varepsilon),\ 0<\varepsilon\le 1, \\ a(0) &:= f(\bar u). \end{aligned} \tag{2.8}$$

This defines a continuous function $a : [0,1] \to 0$ which is also differentiable on $(0,1)$ with the derivative $a' \in L_1(0,1)$ given by:

$$a'(\varepsilon) = -\frac{3}{2\varepsilon^2} \|\bar{v}^\varepsilon\|^2_{L_2(\Omega)} \tag{2.9}$$

We will use both continuity and differentiability of $a$ as a tool in our future analyses. The detailed proofs can be found in [43], Section 3.
We can now turn to the discrete counterpart:

2.2. **Discrete Regularised Problem.** First, we introduce a **conforming** (compare [15], Section 3.2) triangulation of $\Omega$ to which we assign a natural number $k$. This single triangulation is denoted by $\mathcal{T}_k$.

With this triangulation $\mathcal{T}_k$ and its triangular ($d = 2$) respectively tetrahedral ($d = 3$) elements $T \in \mathcal{T}_k$ we can define the finite element spaces we need using for $X = L_2(\Omega)$ or $X = H^1_0(\Omega)$ and the space of polynomials up to degree $l \in \mathbb{N} \cup \{0\}$ on an element $T \in \mathcal{T}_k$, $\mathbb{P}_l(T)$, the notation

$$\mathbf{FES}(\mathcal{T}_k, \mathbb{P}_l, X) = \{W \in X \;:\; W|_T \in \mathbb{P}_l(T)\; \forall T \in \mathcal{T}_k\}\,.$$

The spaces we use are the space of piecewise linear $H^1_0(\Omega)$-conforming finite elements for the discretisation of the state

$$\mathbb{Y}_k = \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_1, H^1_0(\Omega)),$$

the space of continuous or discontinuous piecewise linear finite elements for the discretisation of the virtual control

$$\mathbb{V}_k = \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_1, H^1_0(\Omega)) \text{ or } \mathbb{V}_k = \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_1, L_2(\Omega))$$

and finally, for the control, either the **variational discretisation** approach of [22]

$$\mathbb{U}_k = L_2(\Omega)$$

or a **full discretisation** approach with piecewise constant controls:

$$\mathbb{U}_k = \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_0, L_2(\Omega)).$$

Reflecting the continuous solution operator $S$, the discrete one $S_k$, $S_k : L_2(\Omega) \to \mathbb{Y}_k$ maps each right hand side $u \in L_2(\Omega)$ linear and continuously to the solution $Y = S_k u$ of the equation:

$$(\nabla Y, \nabla W)_{L_2(\Omega)} = (u, W)_{L_2(\Omega)}\; \forall W \in \mathbb{Y}_k.$$

The adjoint operator with respect to the $L_2(\Omega)$-scalar product $S^*_k$, $S^*_k : L_2(\Omega) \to \mathbb{Y}_k$, maps each right hand side $q \in L_2(\Omega)$ linear and continuously to the solution $P = S^*_k q$ of the equation:

$$(\nabla P, \nabla W)_{L_2(\Omega)} = (q, W)_{L_2(\Omega)}\; \forall W \in \mathbb{Y}_k.$$

Finally, for the discretisation of the state constraint we use - for simplicity, other choices are possible, too - the Lagrange interpolant of $y_c$ denoted by $I_k y_c$.
After these preliminaries, we can finally lay out our discretisation:

Mirroring $(P^\varepsilon)$, we introduce the following discrete regularised problem:

$$(P_k^\varepsilon) \qquad \begin{cases} \min\limits_{U\in\mathbb{U}_k, Y\in\mathbb{Y}_k, V\in\mathbb{V}_k} J^\varepsilon(Y,U,V) = \frac{1}{2}\left\|Y-y_d\right\|^2_{L_2(\Omega)} + \frac{\nu}{2}\|U\|^2_{L_2(\Omega)} + \frac{1}{\varepsilon}\left\|V\right\|^2_{L_2(\Omega)} \\ \text{s.t.} \\ S_kU = Y \text{ in } \Omega \\ \text{and} \\ a \leq U \leq b \text{ a.e in } \Omega,\ a,b\in\mathbb{R}\cup\{-\infty,\infty\} \\ I_ky_c \leq Y + \varepsilon V \text{ a.e. in } \Omega, \end{cases}$$

For this problem we also define an admissible set

$$\mathbb{U}_k^{\varepsilon,ad} := \left\{(U,V)\in\mathbb{U}_k\times\mathbb{V}_k \,:\, a\leq U\leq b \text{ a.e. in } \Omega,\ I_ky_c \leq S_kU+\varepsilon V \text{ a.e. in } \Omega\right\}.$$

We observe that this set is always nonempty as long as there exist functions $U\in\mathbb{U}_k$ with $a\leq U\leq b$ (which is the case for our ansatz spaces) regardless of the question whether there exists a function $U\in\mathbb{U}_k$ with $a\leq U\leq b$ a.e. in $\Omega$ *and* which fulfills the unregularised state constraint $I_ky_c\leq S_kU$. In particular, for every $\varepsilon>0$ there exists a unique solution to $(P_k^\varepsilon)$ which we denote by $(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ with associated state $\bar{Y}_k^\varepsilon = S_k\bar{U}_k^\varepsilon$.
Because for every fixed $\varepsilon>0$ the discrete constraint mapping for the state constraint $M_k^\varepsilon(U,V) := I_ky_c - S_kU - \varepsilon V$ is surjective as a mapping $M_k^\varepsilon : \mathcal{U}\cap\mathbb{U}_k\times\mathbb{V}_k\to\mathbb{V}_k$ (for the definition of $\mathcal{U}$ compare (2.4)) we know for the same reason as on the continuous level that there exists a Lagrange multiplier $\bar{\Theta}_k^\varepsilon\in\mathbb{V}_k$ such that the unique solution $(\bar{U}_k^\varepsilon,\bar{V}_k^\varepsilon)$ fulfills the following KKT system with $\bar{P}_k^\varepsilon := S_k^*(\bar{Y}_k^\varepsilon - y_d - \bar{\Theta}_k^\varepsilon)$ reflecting the continuous one (2.5):

$$(\bar{P}_k^\varepsilon + \nu\bar{U}_k^\varepsilon, U-\bar{U}_k^\varepsilon)_{L_2(\Omega)} \geq 0\ \forall U\in\mathcal{U}\cap\mathbb{U}_k \tag{2.10a}$$

$$\frac{1}{\varepsilon^2}\bar{V}_k^\varepsilon = \bar{\Theta}_k^\varepsilon \tag{2.10b}$$

$$(\bar{\Theta}_k^\varepsilon, W)_{L_2(\Omega)} \geq 0,\ \forall W\in\mathbb{V}_k,\ W\geq 0 \text{ a.e. in } \Omega,\ (\bar{\Theta}_k^\varepsilon, \bar{Y}_k^\varepsilon - I_ky_c + \varepsilon\bar{V}_k^\varepsilon)_{L_2(\Omega)} = 0 \tag{2.10c}$$

In the case of the variational discretisation approach $\mathbb{U}_k = L_2(\Omega)$, (2.10a) is - as in the continuous case - equivalent to the pointwise projection formula

$$\bar{U}_k^\varepsilon(x) = \Pi(\bar{P}_k^\varepsilon)(x) \text{ f.a.a. } x\in\Omega \tag{2.11}$$

with $\Pi(\cdot)$ as in (2.7).
After these preliminaries we can now proceed to state the main results of this paper:

## 3. Statement of the Main Results

To state our main result, we make the following standard Slater-type assumption:

**Assumption 3.1** (Slater-Assumption)**.** *There exists $u_s\in L_2(\Omega)$ and $\tau>0$ such that*

$$y_c + \tau \leq Su_s \textit{ a.e. in } \Omega.$$

With this assumption we can formulate our main results, first the error estimator for the variational discretisation approach, $\mathbb{U}_k = L_2(\Omega)$, then for the full discretisation technique, $\mathbb{U}_k = \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_0, L_2(\Omega))$. First, though, let us quickly lay some notational groundwork:
For $e,f\in\mathbb{R}^+$ we write $e\lesssim f$, if there exists a constant $C$ solely depending on the data of the problem, i.e. $\Omega,\nu,a,b,y_d,y_c,S$ and the existence of a feasible point $u\in\mathbb{U}^{ad}$ such that $e\leq Cf$. For a particular type of data, though, we want to trace the dependency of the constants more closely. That is why we introduce the following notational convention:

Let $p' \geq 1$ be given such that the embedding $H^1(\Omega) \hookrightarrow L_{p'}(\Omega)$ holds. Besides let $\tau > 0$ be the $\tau$ from Assumption 3.1. With $s(\tau)$ and $c(p')$ we mark generic constants with the property that

$$\begin{aligned} s(\tau) \to& \infty, \text{ as } \tau \to 0 \\ c(p') \to& \infty, \text{ as } p' \to \infty, \end{aligned}$$

i.e. with the help of $c(p')$ and $s(\tau')$ estimates such as

$$e \lesssim \tau^{-3} f, e \lesssim p' f, \ e, f \in \mathbb{R}^+$$

can be shortened to

$$e \lesssim s(\tau) f, \ e \lesssim c(p') f.$$

In addition, we need a special projection:

**Definition 3.2** (definition of $P_k^{0+}$). *For an arbitrary function $z \in L_2(\Omega)$ we define the projection $P_k^{0+} : L_2(\Omega) \to \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_0, L_2(\Omega))$ by*

$$P_k^{0+} z = \operatorname*{arg\,min}_{W \in \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_0, L_2(\Omega)), W \geq 0 \ a.e.} \frac{1}{2} \|W - z\|_{L_2(\Omega)}^2 , \tag{3.1}$$

**Remark 3.3.** *It is also possible to use the following type of projection on the space of piecewise linear discontinuous finite elements denoted by $P_k^{1+}$ and defined by*

$$P_k^{1+} z = \operatorname*{arg\,min}_{W \in \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_1, L_2(\Omega)), W \geq 0 \ a.e.} \frac{1}{2} \|W - z\|_{L_2(\Omega)}^2 .$$

With these definitions we can formulate our main results: We start with the estimator for the variational discretisation technique $\mathbb{U}_k = L_2(\Omega)$:

**Theorem 3.4** (estimator for variational discretisation). *Let $N \in \mathbb{N}$ be fixed but arbitrary and $4 < p' < \infty$ in case $d = 2$ and $p' = 6$ in case $d = 3$ be fixed. Besides, let Assumption 3.1 be fulfilled, let $\bar{u}$ be the solution to $(P)$, $(\bar{u}^{\varepsilon^N}, \bar{v}^{\varepsilon^N})$ be the solution couple to $(P^\varepsilon)$ with regularisation parameter $\varepsilon = \varepsilon^N$ and $(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ be the solution couple to $(P_k^\varepsilon)$ with $\mathbb{U}_k = L_2(\Omega)$. Then there exists $\gamma > 0$ and constants $c(p')$ and $s(\tau)$ such that:*

$$\left\|\bar{U}_k^\varepsilon - \bar{u}\right\|_{L_2(\Omega)}^2 + \frac{4}{\nu\varepsilon^N} \left\|\bar{v}^{\varepsilon^N}\right\|_{L_2(\Omega)}^2 \lesssim \frac{1}{\nu} c(p') s(\tau) \varepsilon^{\gamma N} + \mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) + \mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) \tag{3.2}$$

*with*

$$\begin{aligned} \mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) = & \frac{1}{\nu} \left\|(S - S_k)\bar{U}_k^\varepsilon\right\|_{L_2(\Omega)}^2 + \frac{2}{\nu^2} \left\|(S - S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)}^2 \\ & - \frac{4}{\nu\varepsilon} \left\|\bar{V}_k^\varepsilon\right\|_{L_2(\Omega)}^2 + \frac{4}{\nu} (\bar{\Theta}_k^\varepsilon, I_k y_c - y_c)_{L_2(\Omega)} \end{aligned} \tag{3.3}$$

*and*

$$\begin{aligned} \mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) = & \frac{4}{\nu} \left\|\bar{Y}_k^\varepsilon - y_d\right\|_{L_2(\Omega)} \left\|(S - S_k)\bar{U}_k^\varepsilon\right\|_{L_2(\Omega)} \\ & + \frac{4}{\nu} \left\|\bar{U}_k^\varepsilon\right\| \left\|(S - S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)} \\ & + \frac{4}{\nu} c(p') s(\tau) \left\|P_k^{0+}\bar{\Theta}_k^\varepsilon\right\|_{L_2(\Omega)} \varepsilon^{N(\frac{3}{2}+\gamma)} + \frac{4}{\nu} \left\|\bar{\Theta}_k^\varepsilon - P_k^{0+}\bar{\Theta}_k^\varepsilon\right\|_{L_2(\Omega)} \\ & + \frac{4}{\nu} \min\Bigg\{ s(\tau) \left\|(\bar{U}_k^\varepsilon + \Delta y_c)^-\right\|_{L_2(\Omega)}, c(p') s(\tau) \varepsilon^{-3N/p'} \Bigg( |(S - S_k)\bar{U}_k^\varepsilon|_{H^1(\Omega)} \\ & + \|y_c - I_k y_c\|_{H^1(\Omega)} + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)} \Bigg) \Bigg\}, \end{aligned} \tag{3.4}$$

For the full discretisation $\mathbb{U}_k = \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_0, L_2(\Omega))$ the estimator takes the following shape:

**Theorem 3.5** (estimator for the full discretisation). *Let $N \in \mathbb{N}$ and $\varepsilon > 0$ be fixed but arbitrary and $4 < p' < \infty$ in case $d = 2$ and $p' = 6$ in case $d = 3$ be fixed. Besides, let Assumption 3.1 be fulfilled, let $\bar{u}$ be the solution to $(P)$, $(\bar{u}^{\varepsilon^N}, \bar{v}^{\varepsilon^N})$ be the solution couple to $(P^\varepsilon)$ with regularisation parameter $\varepsilon = \varepsilon^N$ and $(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ be the solution couple to $(P_k^\varepsilon)$ with $\mathbb{U}_k = \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_0, L_2(\Omega))$. Then there exists $\gamma > 0$ and constants $c(p')$ and $s(\tau)$ such that:*

$$\left\|\bar{U}_k^\varepsilon - \bar{u}\right\|^2_{L_2(\Omega)} + \frac{8}{\nu\varepsilon^N}\left\|\bar{v}^{\varepsilon^N}\right\|^2_{L_2(\Omega)} \lesssim \frac{1}{\nu}c(p')s(\tau)\varepsilon^{\gamma N} + \mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) + \mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) \tag{3.5}$$

*with*

$$\begin{aligned}
\mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) = {} & \frac{4}{\nu}(\left\|(S - S_k)\bar{U}_k^\varepsilon\right\|^2_{L_2(\Omega)} + \frac{2\nu + 4}{\nu}\left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|^2_{L_2(\Omega)} \\
& + \frac{4}{\nu^2}\left\|(S - S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\Theta}_k^\varepsilon)\right\|^2_{L_2(\Omega)} \\
& + \frac{8}{\nu}(\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon), \bar{P}_k^\varepsilon)_{L_2(\Omega)} - \frac{8}{\nu}\left\|\bar{V}_k^\varepsilon\right\|^2 \\
& + \frac{8}{\nu}(\bar{\Theta}_k^\varepsilon, I_k y_c - y_c)_{L_2(\Omega)},
\end{aligned} \tag{3.6}$$

*and*

$$\begin{aligned}
\mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) = {} & \frac{8}{\nu}\left\|\bar{Y}_k^\varepsilon - y_d\right\|_{L_2(\Omega)}(\left\|(S - S_k)\bar{U}_k^\varepsilon\right\|_{L_2(\Omega)} + \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|_{L_2(\Omega)}) \\
& + \frac{8}{\nu}\left\|\Pi(\bar{P}_k^\varepsilon)\right\|_{L_2(\Omega)}\left\|(S - S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)} \\
& + \frac{8}{\nu}c(p')s(\tau)\left\|P_k^{0+}(\bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)}\varepsilon^{N(\frac{3}{2}+\gamma)} + \frac{8}{\nu}\left\|\bar{\Theta}_k^\varepsilon - P_k^{0+}\bar{\Theta}_k^\varepsilon\right\|_{L_2(\Omega)} \\
& + \frac{8}{\nu}\min\Big\{s(\tau)\left\|(\Pi(\bar{P}_k^\varepsilon) + \Delta y_c)^-\right\|_{L_2(\Omega)}, c(p')s(\tau)\varepsilon^{-3N/p'}\big(|(S - S_k)\bar{U}_k^\varepsilon|_{H^1(\Omega)} \\
& + \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|_{L_2(\Omega)} + \|y_c - I_k y_c\|_{H^1(\Omega)} + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)}\big)\Big\}
\end{aligned} \tag{3.7}$$

*where $\Pi(\cdot)$ as defined in (2.7).*

**Remark 3.6.** *In [13], Theorem 3.1, a $H^{-1}(\Omega)$-regularity result for the Lagrange multiplier for the state constraint of the unregularised problem was proven under stronger regularity assumptions. If there is also a uniform bound on the sequence of regularised multipliers $\bar{\theta}^{\varepsilon^N}$, i.e. $\|\bar{\theta}^{\varepsilon^N}\|_{H^{-1}(\Omega)} \lesssim s(\tau)$, then the negative power $\varepsilon^{-3N/p'}$ in (3.4) and (3.7) respectively does not enter.*

Over the next two sections we will prove both Theorem 3.4 and Theorem 3.5. The proofs themselves can be found in Section 5.1 and Section 5.2 respectively. Let us dwell a little bit further on the steps which need to be taken:
The broad approach is to use the function $\Pi(\bar{P}_k^\varepsilon)$ (compare (2.7)) and split the error in the following way:

$$\left\|\bar{u} - \Pi(\bar{P}_k^\varepsilon)\right\|^2_{L_2(\Omega)} \lesssim \left\|\bar{u} - \bar{u}^{\varepsilon^N}\right\|^2_{L_2(\Omega)} + \left\|\bar{u}^{\varepsilon^N} - \Pi(\bar{P}_k^\varepsilon)\right\|^2_{L_2(\Omega)}. \tag{3.8}$$

The idea now is to treat the terms on the right separately. For the first we will derive an a priori estimate in terms of $\varepsilon^N$ in Section 4, specifically Theorem 4.5 and Corollary 4.7, i.e.

$$\left\|\bar{u} - \bar{u}^{\varepsilon^N}\right\|^2_{L_2(\Omega)} \lesssim \frac{1}{\nu}c(p')s(\tau)\varepsilon^{\gamma N}. \tag{3.9}$$

For the second term in (3.8) we will derive an a posteriori bound in Section 5 (compare Lemma 5.1, Lemma 5.2 and Lemma 5.3) of the type

$$\left\|\bar{u}^{\varepsilon^N} - \Pi(\bar{P}_k^\varepsilon)\right\|^2_{L_2(\Omega)} \lesssim \mathcal{E}_r^2(\Pi(\bar{P}_k^\varepsilon), \bar{V}_k^\varepsilon) + \mathcal{E}_s(\Pi(\bar{P}_k^\varepsilon), \bar{V}_k^\varepsilon). \tag{3.10}$$

To deduce the results of Theorem 3.4 we then just have to recall the relation $\bar{U}_k^\varepsilon = \Pi(\bar{P}_k^\varepsilon)$ and combine it with the bounds (3.10),(3.9) and (3.8).
To derive the bounds of Theorem 3.5, we additionally have to take into account the error incurred by the full discretisation technique, i.e. $\left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|_{L_2(\Omega)}$.

## 4. Estimating $\|\bar{u} - \bar{u}^{\varepsilon^N}\|_{L_2(\Omega)}$

To prove the bound (3.9) we will derive the estimate

$$\|\bar{u} - \bar{u}^\varepsilon\|^2_{L_2(\Omega)} + \frac{1}{\nu\varepsilon}\|\bar{v}^\varepsilon\|^2_{L_2(\Omega)} \lesssim \frac{1}{\nu}c(p')s(\tau)\varepsilon^\gamma \tag{4.1}$$

independent of $\varepsilon$ and then insert $\varepsilon = \varepsilon^N$. We take the following steps:

(1) Prove a penalty structure for $\bar{v}^\varepsilon$, i.e. $\bar{v}^\varepsilon = \frac{1}{\varepsilon}(y_c - \bar{y}^\varepsilon)^+$ a.e. in $\Omega$, Lemma 4.1
(2) Prove a bound for $\left\|\bar{\theta}^\varepsilon\right\|_{L_p(\Omega)}$, $1 \le p \le 2$, Lemma 4.3
(3) Derive an estimate for $|(y_c - \bar{y}^\varepsilon)^+|_{H^1(\Omega)}$, Lemma 4.4
(4) Using all the previous results, we demonstrate (4.1) and as corollary obtain (3.9), Theorem 4.5 and Corollary 4.7.

We start with the first step, the result can be found in [28], Lemma 2.

**Lemma 4.1** (penalty structure of virtual control)**.** *For every fixed $\varepsilon > 0$ we have $\bar{v}^\varepsilon = \frac{1}{\varepsilon}(y_c - \bar{y}^\varepsilon)^+$ a.e. in $\Omega$.*

Now let us proceed to the second step, the bound for $\left\|\bar{\theta}^\varepsilon\right\|_{L_p(\Omega)}$. Before though, we need the following auxiliary result:

**Lemma 4.2.** *For $\bar{u}^\varepsilon, \bar{v}^\varepsilon, \bar{y}^\varepsilon$ and $\bar{u}$ we have independent of $\varepsilon$:*

$$\|\bar{u}^\varepsilon\|_{L_2(\Omega)}, \|\bar{y}^\varepsilon\|_{L_2(\Omega)}, \|\bar{u}\|_{L_2(\Omega)}, \|\bar{y}\|_{L_2(\Omega)} \lesssim 1,\ \|\bar{v}^\varepsilon\|_{L_2(\Omega)} \lesssim \varepsilon^{\frac{1}{2}}.$$

*Proof.* Since we assumed that the feasible set for the unregularised problem $(P)$, $\mathbb{U}^{ad}$, is non-empty, we know that there is a function $u \in \mathbb{U}^{ad}$ such that:

$$f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) \le f(\bar{u}) \le f(u).$$

Due to the structure of the objective this immediately gives

$$\|\bar{u}^\varepsilon\|_{L_2(\Omega)}, \|\bar{u}\|_{L_2(\Omega)} \lesssim 1,\ \|\bar{v}^\varepsilon\|_{L_2(\Omega)} \lesssim \varepsilon^{\frac{1}{2}}.$$

The bound on $\bar{y}^\varepsilon$ (independent of $\varepsilon$) and $\bar{y}$ follow by continuity of $S$. □

We can now turn to proving a 'qualified' bound for $\left\|\bar{\theta}^\varepsilon\right\|_{L_p(\Omega)}$:

**Lemma 4.3.** *Let $1 \le p \le 2$ be given and let $p'$ be its conjugate exponent, i.e. $\frac{1}{p} + \frac{1}{p'} = 1$. Then*

$$\left\|\bar{\theta}^\varepsilon\right\|_{L_p(\Omega)} \lesssim s(\tau)\varepsilon^{-3/p'},\ 1 \le p \le 2,\ \frac{1}{p} + \frac{1}{p'} = 1. \tag{4.2}$$

*In particular, for $p = 1$ we have:*

$$\left\|\bar{\theta}^\varepsilon\right\|_{L_1(\Omega)} \lesssim s(\tau). \tag{4.3}$$

*Proof.* The proof of the bound (4.3) can be found in [23], Lemma 2.3.
In addition, thanks to Lemma 4.2 we know that

$$\|\bar{v}^\varepsilon\|_{L_2(\Omega)} \lesssim \varepsilon^{1/2} \tag{4.4}$$

Besides, we observe that due to (2.5b)

$$\frac{1}{\varepsilon^2}\bar{v}^\varepsilon = \bar{\theta}^\varepsilon \;\Leftrightarrow\; \frac{1}{\varepsilon^2}\|\bar{v}^\varepsilon\|_{L_2(\Omega)} = \left\|\bar{\theta}^\varepsilon\right\|_{L_2(\Omega)}$$

and thus thanks to (4.4)

$$\left\|\bar{\theta}^\varepsilon\right\|_{L_2(\Omega)} \lesssim \varepsilon^{-3/2}. \tag{4.5}$$

Using interpolation of $L_p(\Omega)$-spaces (compare Chapter 5, [8]) we finally obtain for $\sigma = 1 - \frac{2}{p'}$ with the help of (4.3) and (4.5)

$$\left\|\bar{\theta}^\varepsilon\right\|_{L_p(\Omega)} \lesssim \left\|\bar{\theta}^\varepsilon\right\|_{L_2(\Omega)}^{1-\sigma} \left\|\bar{\theta}^\varepsilon\right\|_{L_1(\Omega)}^{\sigma} \lesssim s(\tau)\varepsilon^{-3/p'},\; 1 \le p \le 2$$

□

We now can tackle the third step of the proof, the bound for $|(y_c - \bar{y}^\varepsilon)^+|_{H^1(\Omega)}$:

**Lemma 4.4.** *Let $2 \le p' \le \infty$ in case $d = 2$ and $2 \le p' \le 6$ in case $d = 3$. Furthermore, let $p$ denote its dual exponent, i.e. $\frac{1}{p'} + \frac{1}{p} = 1$. For the violation of the state constraint $(y_c - \bar{y}^\varepsilon)^+$ we have the following estimate:*

$$\left|(y_c - \bar{y}^\varepsilon)^+\right|_{H^1(\Omega)} \lesssim c(p')s(\tau)\varepsilon^{1-1/p'}. \tag{4.6}$$

*Proof.* Evidently, $(y_c - \bar{y}^\varepsilon)^+ \in H_0^1(\Omega)$, since $\bar{y}^\varepsilon = 0$ on $\partial\Omega$ and $y_c|_{\partial\Omega} < 0$ by assumption. Thus, with $\bar{y}^\varepsilon = S\bar{u}^\varepsilon$ we can deduce:

$$\begin{aligned}|(y_c - \bar{y}^\varepsilon)^+|_{H^1(\Omega)}^2 &= |(\nabla(y_c - \bar{y}^\varepsilon), \nabla(y_c - \bar{y}^\varepsilon)^+)_{L_2(\Omega)}| \\ &= |(\bar{u}^\varepsilon + \Delta y_c, (y_c - \bar{y}^\varepsilon)^+)_{L_2(\Omega)}|\end{aligned}$$

Now we can take advantage of the penalty structure of $\bar{v}^\varepsilon$, i.e $(y_c - \bar{y}^\varepsilon)^+ = \varepsilon\bar{v}^\varepsilon$ (compare Lemma 4.1) insert it in the equation above and get:

$$\begin{aligned}|(y_c - \bar{y}^\varepsilon)^+|_{H^1(\Omega)}^2 &= |(\bar{u}^\varepsilon + \Delta y_c, \varepsilon\bar{v}^\varepsilon)_{L_2(\Omega)}| \\ &\le \|\bar{u}^\varepsilon + \Delta y_c\|_{L_2(\Omega)} \|\varepsilon\bar{v}^\varepsilon\|_{L_2(\Omega)}.\end{aligned} \tag{4.7}$$

Presently, let us estimate $\|\bar{v}^\varepsilon\|_{L_2(\Omega)}$:
Recalling the complimentary slackness condition (2.5c), the embedding $H^1(\Omega) \hookrightarrow L_{p'}(\Omega)$, $1 \le p' < \infty$ in case $d = 2$ and $1 \le p' \le 6$ in case $d = 3$, we obtain

$$\frac{1}{\varepsilon}\|\bar{v}^\varepsilon\|_{L_2(\Omega)}^2 = (\bar{\theta}^\varepsilon, y_c - \bar{y}^\varepsilon)_{L_2(\Omega)} \le \left\|\bar{\theta}^\varepsilon\right\|_{L_p(\Omega)} \left\|(y_c - \bar{y}^\varepsilon)^+\right\|_{L_{p'}(\Omega)} \lesssim c(p')\left\|\bar{\theta}^\varepsilon\right\|_{L_p(\Omega)} \left\|(y_c - \bar{y}^\varepsilon)^+\right\|_{H^1(\Omega)}.$$

Using the Poincaré-Friedrich-inequality ($(y_c - \bar{y}^\varepsilon)^+ \in H_0^1(\Omega)$) and the results of Lemma 4.3 on the right, gain:

$$\frac{1}{\varepsilon}\|\bar{v}^\varepsilon\|_{L_2(\Omega)}^2 \lesssim c(p')s(\tau)\varepsilon^{-3/p'}|(y_c - \bar{y}^\varepsilon)^+|_{H^1(\Omega)}.$$

A short rearrangement of this bound yields

$$\|\bar{v}^\varepsilon\|_{L_2(\Omega)} \lesssim c(p')s(\tau)\varepsilon^{\frac{1}{2}-\frac{3}{2p'}}|(y_c - \bar{y}^\varepsilon)^+|_{H^1(\Omega)}^{\frac{1}{2}}.$$

Let us now insert this bound in (4.7) bearing in mind that $\|\bar{u}^\varepsilon + \Delta y_c\|_{L_2(\Omega)}$ is uniformly bounded (= independent of $\varepsilon$) in $L_2(\Omega)$ (compare Lemma 4.2):

$$|(y_c - \bar{y}^\varepsilon)^+|^2_{H^1(\Omega)} \lesssim \varepsilon \|\bar{v}^\varepsilon\|_{L_2(\Omega)} \lesssim c(p')s(\tau)\varepsilon^{\frac{3}{2}-\frac{3}{2p'}} |(y_c - \bar{y}^\varepsilon)^+|^{\frac{1}{2}}_{H^1(\Omega)}$$

Dividing by $|(y_c - \bar{y}^\varepsilon)^+|^{1/2}_{H^1(\Omega)}$ (if this term were 0, the postulated bound (4.6) trivially holds) yields:

$$|(y_c - \bar{y}^\varepsilon)^+|^{\frac{3}{2}}_{H^1(\Omega)} \lesssim c(p')s(\tau)\varepsilon^{\frac{3}{2}-\frac{3}{2p'}}.$$

(4.6) then readily follows. □

We will now deal with the fourth and last step, the a priori bound for $\|\bar{u} - \bar{u}^\varepsilon\|_{L_2(\Omega)}$ in terms of the regularisation parameter $\varepsilon$:

**Theorem 4.5.** *Let $\bar{u}$ be the solution to $(P)$ and $(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$ be the solution to $(P^\varepsilon)$ for $\varepsilon > 0$. Besides, let $4 < p' < \infty$ in case $d = 2$ and $p' = 6$ in case $d = 3$. Then the following estimates are valid:*

$$\|\bar{u} - \bar{u}^\varepsilon\|^2_{L_2(\Omega)} + \frac{1}{\nu\varepsilon}\|\bar{v}^\varepsilon\|^2_{L_2(\Omega)} \lesssim \begin{cases} \frac{1}{\nu}c(p')s(\tau)\varepsilon^{1-\frac{4}{p'}} & d = 2 \\ \frac{1}{\nu}s(\tau)\varepsilon^{\frac{1}{3}} & d = 3\ (p' = 6) \end{cases} \tag{4.8}$$

*Thus, $\gamma$ in (3.9) is defined by*

$$\gamma = \begin{cases} 1 - \frac{4}{p'} & d = 2 \\ \frac{1}{3} & d = 3 \end{cases} \tag{4.9}$$

*Proof.* The proof uses Taylor expansion of the reduced goal functional $f^\varepsilon$ at $(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$. Since $f^\varepsilon$ is quadratic, we gain

$$\begin{aligned} f(\bar{u}) - f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) &= f^\varepsilon(\bar{u}, 0) - f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) \\ &= D^1 f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) \cdot \left[\bar{u} - \bar{u}^\varepsilon, \quad -\bar{v}^\varepsilon\right] + D^2 f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) \cdot \left[\bar{u} - \bar{u}^\varepsilon, \quad -\bar{v}^\varepsilon\right]^2. \end{aligned} \tag{4.10}$$

Straightforward differentiating, $(\bar{u}, 0) \in \mathbb{U}^{\varepsilon,ad}$ and (2.3) yield:

$$f(\bar{u}) - f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) \geq \nu \|\bar{u} - \bar{u}^\varepsilon\|^2_{L_2(\Omega)} + \frac{1}{\varepsilon}\|\bar{v}^\varepsilon\|^2_{L_2(\Omega)}$$

Using the differentiability of the optimal value function $a$ (defined in (2.8)) on $(0,1)$, (2.9), and continuity at 0, we deduce

$$\begin{aligned} \nu \|\bar{u} - \bar{u}^\varepsilon\|^2_{L_2(\Omega)} + \frac{1}{\varepsilon}\|\bar{v}^\varepsilon\|^2_{L_2(\Omega)} &\leq f(\bar{u}) - f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) = a(0) - a(\varepsilon) \\ &= \int_0^\varepsilon -a'(t)\,dt = \int_0^\varepsilon \frac{1}{t^2}\|\bar{v}^t\|^2_{L_2(\Omega)}\,dt. \end{aligned} \tag{4.11}$$

Harnessing $\bar{\theta}^t \geq 0$ for all $t > 0$, Hölder's inequality and the complimentary slackness condition (2.5c) with $\varepsilon = t$, we arrive at:

$$\frac{1}{t^2}\|\bar{v}^t\|^2_{L_2(\Omega)} = \frac{1}{t}(\bar{\theta}^t, y_c - \bar{y}^t)_{L_2(\Omega)} \leq \frac{1}{t}\|\bar{\theta}^t\|_{L_p(\Omega)} \|(y_c - \bar{y}^t)^+\|_{L_{p'}(\Omega)}, \quad \frac{1}{p} + \frac{1}{p'} = 1$$

The embedding $H^1_0(\Omega) \hookrightarrow L_{p'}(\Omega)$, $1 \leq p' < \infty$ if $d = 2$ and $1 \leq p' \leq 6$ if $d = 3$, and the results of Lemma 4.3 and Lemma 4.4 for $\varepsilon = t$ enable us to continue our estimates in the ensuing vein:

$$\begin{aligned} \frac{1}{t^2}\|\bar{v}^t\|^2_{L_2(\Omega)} &\leq \frac{1}{t}\|\bar{\theta}^t\|_{L_p(\Omega)} \|(y_c - \bar{y}^t)^+\|_{L_{p'}(\Omega)} \\ &\lesssim c(p')s(\tau)\frac{1}{t}(t^{1-4/p'}) = c(p')s(\tau)t^{-4/p'}. \end{aligned}$$

Presently, we can pick up the thread we left off in (4.11) and continue our estimates taking advantage of the estimates above (remember that we assumed $p' > 4$ if $d = 2$ and $p' = 6$ if $d = 3$).

$$\nu \|\bar{u} - \bar{u}^\varepsilon\|^2_{L_2(\Omega)} + \frac{1}{\varepsilon} \|\bar{v}^\varepsilon\|^2_{L_2(\Omega)} \le \int_0^\varepsilon \frac{1}{t^2} \|\bar{v}^t\|^2_{L_2(\Omega)} \lesssim c(p')s(\tau) \int_0^\varepsilon t^{-4/p'}\, dt$$

$$= c(p')s(\tau)\frac{1}{1-4/p'}\varepsilon^{1-4/p'} \lesssim c(p')s(\tau)\varepsilon^{1-4/p'}, \quad p' > 4$$

In case $d = 2$, this gives the first estimate in (4.8). If $d = 3$ we just have to insert $p' = 6$ into the estimates above. Here, note that the dependence on the embedding constant $c(p')$ can be neglected since we fix it to $c(p') = c(6)$. Ultimately, we deduce in case $d = 3$

$$\nu \|\bar{u} - \bar{u}^\varepsilon\|^2_{L_2(\Omega)} \lesssim s(\tau)\varepsilon^{1/3}$$

This gives the second estimate in (4.8). □

**Remark 4.6.** *Following the technique employed in* [28], *Theorem 4, and defining* $0 < \alpha < 1$ *such that* $S \in \mathcal{L}(L_2(\Omega), C^{0,\alpha}(\bar{\Omega}))$ *one attains the estimate:*

$$\|\bar{u} - \bar{u}^\varepsilon\|^2_{L_2(\Omega)} \lesssim \|S\|_{\mathcal{L}(L_2(\Omega), C^{0,\alpha}(\bar{\Omega}))}\, \varepsilon^{\frac{3}{2}\frac{\alpha}{\alpha+1}}.$$

*In principle, one could also use this result, at the 'price' of the constant* $\|S\|_{\mathcal{L}(L_2(\Omega), C^{0,\alpha}(\bar{\Omega}))}$ *entering. For more regular solutions* $\alpha > \frac{1}{2}$ *the estimate above is better, but for small* $\alpha$ *it is worse. We chose our approach, because on the one hand we can do without additional regularity assumptions on* $\Omega$ *which may give rise to comparable estimates in less regular settings and on the other hand we gained 'qualified' bounds on the* $\|\bar{\theta}^\varepsilon\|_{L_p(\Omega)}$ *as additional valuable results.*

Now, inserting $\varepsilon = \varepsilon^N$ in the bound (4.8) gives the following corollary which completes this section:

**Corollary 4.7.** *In the setting of Theorem 4.5 we have the following estimate:*

$$\left\|\bar{u} - \bar{u}^{\varepsilon^N}\right\|^2_{L_2(\Omega)} + \frac{1}{\nu\varepsilon^N}\left\|\bar{v}^{\varepsilon^N}\right\|^2_{L_2(\Omega)} \lesssim \begin{cases} \frac{1}{\nu}c(p')s(\tau)\varepsilon^{N(1-\frac{4}{p'})} & d = 2 \\ \frac{1}{\nu}s(\tau)\varepsilon^{\frac{N}{3}} & d = 3\ (p' = 6) \end{cases}$$

*and* $\gamma$ *as in* (4.9).

## 5. Estimating $\|\bar{u}^{\varepsilon^N} - \Pi(\bar{P}^\varepsilon_k)\|^2_{L_2(\Omega)}$

Let us now recall our split (3.8):

$$\left\|\bar{u} - \Pi(\bar{P}^\varepsilon_k)\right\|^2_{L_2(\Omega)} \lesssim \left\|\bar{u} - \bar{u}^{\varepsilon^N}\right\|^2_{L_2(\Omega)} + \left\|\bar{u}^{\varepsilon^N} - \Pi(\bar{P}^\varepsilon_k)\right\|^2_{L_2(\Omega)}.$$

To estimate $\|\bar{u}^{\varepsilon^N} - \Pi(\bar{P}^\varepsilon_k)\|^2_{L_2(\Omega)}$, we pursue the following course:

(1) Estimate for arbitrary $P \in H^1_0(\Omega)$ the difference $\|\bar{u}^{\varepsilon^N} - \Pi(P)\|^2_{L_2(\Omega)}$. Lemma 5.1.
(2) Estimate the resulting terms involving the Lagrange multipliers $\bar{\Theta}^\varepsilon_k$, $\bar{\theta}^{\varepsilon^N}$, Lemmas 5.2 and 5.3.
(3) Combining both, derive an estimate for the variational discretisation approach $\bar{U}^\varepsilon_k = \Pi(\bar{P}^\varepsilon_k)$ (compare (2.11)), this gives our first main result Theorem 3.4.
(4) Deduce an estimate for the full discretisation by splitting the difference in the following way

$$\left\|\bar{U}^\varepsilon_k - \bar{u}^{\varepsilon^N}\right\|^2_{L_2(\Omega)} \le 2\left\|\Pi(\bar{P}^\varepsilon_k) - \bar{U}^\varepsilon_k\right\|^2_{L_2(\Omega)} + 2\left\|\Pi(\bar{P}^\varepsilon_k) - \bar{u}^{\varepsilon^N}\right\|^2_{L_2(\Omega)}.$$

The first term can be evaluated exactly (compare [26], Remark 4.3) the second one can be dealt with as in the variational discretisation setting. This gives our second main result Theorem 3.5.

Let us start with the first step:

**Lemma 5.1.** *Let $(P, Y) \in H_0^1(\Omega) \times H_0^1(\Omega)$ be arbitrary. Furthermore, we define $\hat{y} := S\Pi(P), \hat{p} := S^*(Y - y_d - \bar{\Theta}_k^\varepsilon)$ and $U := \Pi(P)$. Then, with a fixed $N \geq 1$, we have*

$$\begin{aligned} \left\|\Pi(P) - \bar{u}^{\varepsilon^N}\right\|^2_{L_2(\Omega)} \leq \frac{1}{2\nu} \|\hat{y} - Y\|^2_{L_2(\Omega)} + \frac{1}{\nu^2} \|\hat{p} - P\|^2_{L_2(\Omega)} \\ + \frac{2}{\nu}(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - \hat{y})_{L_2(\Omega)} + \frac{2}{\nu}(\bar{\Theta}_k^\varepsilon, \hat{y} - \bar{y}^{\varepsilon^N})_{L_2(\Omega)}, \end{aligned} \tag{5.1}$$

*Proof.* $\Pi(P)$ is the best-approximation of $P$ in the closed and convex set $\mathcal{U} = \{u \in L_2(\Omega) \ : \ a \leq u \leq b \text{ a.e. in } \Omega\}$. Hence:

$$(P + \nu\Pi(P), \Pi(P) - u)_{L_2(\Omega)} \leq 0 \quad \forall u \in \mathcal{U}. \tag{5.2}$$

Inserting $u = \bar{u}^{\varepsilon^N} \in \mathcal{U}$ we gain:

$$(P + \nu\Pi(P), \Pi(P) - \bar{u}^{\varepsilon^N})_{L_2(\Omega)} \leq 0,$$

Similarly, employing (2.5a) into which we insert $u = \Pi(P) \in \mathcal{U}$, we obtain

$$-(\bar{p}^{\varepsilon^N} + \nu\bar{u}^{\varepsilon^N}, \Pi(P) - \bar{u}^{\varepsilon^N})_{L_2(\Omega)} \leq 0$$

From these observations we can infer that:

$$\begin{aligned} \nu\left\|\Pi(P) - \bar{u}^{\varepsilon^N}\right\|^2_{L_2(\Omega)} &= (P + \nu\Pi(P), \Pi(P) - \bar{u}^{\varepsilon^N})_{L_2(\Omega)} - (\bar{p}^{\varepsilon^N} + \nu\bar{u}^{\varepsilon^N}, \Pi(P) - \bar{u}^{\varepsilon^N})_{L_2(\Omega)} \\ &+ (\bar{p}^{\varepsilon^N} - P, \Pi(P) - \bar{u}^{\varepsilon^N})_{L_2(\Omega)} \leq (\bar{p}^{\varepsilon^N} - P, \Pi(P) - \bar{u}^{\varepsilon^N})_{L_2(\Omega)} \end{aligned} \tag{5.3}$$

Inserting $\hat{p}$, we can split the last term in the following fashion:

$$(\bar{p}^{\varepsilon^N} - P, \Pi(P) - \bar{u}^{\varepsilon^N})_{L_2(\Omega)} = (\bar{p}^{\varepsilon^N} - \hat{p}, \Pi(P) - \bar{u}^{\varepsilon^N})_{L_2(\Omega)} + (\hat{p} - P, \Pi(P) - \bar{u}^{\varepsilon^N})_{L_2(\Omega)} \tag{5.4}$$

For the second term on the right hand side we use Cauchy-Schwarz and then Young's inequality to obtain:

$$(\hat{p} - P, \Pi(P) - \bar{u}^{\varepsilon^N})_{L_2(\Omega)} \leq \frac{1}{2\nu} \|\hat{p} - P\|^2_{L_2(\Omega)} + \frac{\nu}{2} \left\|\Pi(P) - \bar{u}^{\varepsilon^N}\right\|^2_{L_2(\Omega)}. \tag{5.5}$$

For the first term on the right hand side in (5.4) we use $\bar{p}^{\varepsilon^N}, \hat{p} \in H_0^1(\Omega)$ for every fixed $\varepsilon > 0$, as well as $\hat{y} = S\Pi(P)$ and $\hat{p} = S^*(Y - y_d - \bar{\Theta}_k^\varepsilon)$ to deduce:

$$\begin{aligned} (\bar{p}^{\varepsilon^N} - \hat{p}, \Pi(P) - \bar{u}^{\varepsilon^N})_{L_2(\Omega)} &= (\nabla(\hat{y} - \bar{y}^{\varepsilon^N}), \nabla(\bar{p}^{\varepsilon^N} - \hat{p}))_{L_2(\Omega)} \\ &= (\bar{y}^{\varepsilon^N} - Y, \hat{y} - \bar{y}^{\varepsilon^N})_{L_2(\Omega)} + (\bar{\Theta}_k^\varepsilon - \bar{\theta}^{\varepsilon^N}, \hat{y} - \bar{y}^{\varepsilon^N})_{L_2(\Omega)} \end{aligned} \tag{5.6}$$

The second term already forms a part of (5.1), thus, at this stage, we content ourselves with estimating the first using Cauchy-Schwarz's and then Young's inequality:

(5.7)

$$\begin{aligned} (\bar{y}^{\varepsilon^N} - Y, \hat{y} - \bar{y}^{\varepsilon^N})_{L_2(\Omega)} &= (\bar{y}^{\varepsilon^N} - \hat{y}, \hat{y} - \bar{y}^{\varepsilon^N})_{L_2(\Omega)} + (\hat{y} - Y, \hat{y} - \bar{y}^{\varepsilon^N})_{L_2(\Omega)} \\ &= -\left\|\hat{y} - \bar{y}^{\varepsilon^N}\right\|^2_{L_2(\Omega)} + \|\hat{y} - Y\|_{L_2(\Omega)} \left\|\hat{y} - \bar{y}^{\varepsilon^N}\right\|_{L_2(\Omega)} \leq \frac{1}{4} \|\hat{y} - Y\|^2_{L_2(\Omega)} \end{aligned}$$

Combining all our previous estimates, i.e (5.4),(5.5),(5.6) and (5.7) and inserting them in (5.3) we discern

$$\nu \left\|\Pi(P) - \bar{u}^{\varepsilon^N}\right\|^2_{L_2(\Omega)} \le \frac{1}{2\nu} \|\hat{p} - P\|^2_{L_2(\Omega)} + \frac{\nu}{2} \left\|\Pi(P) - \bar{u}^{\varepsilon^N}\right\|^2_{L_2(\Omega)} + \frac{1}{4} \|\hat{y} - Y\|^2_{L_2(\Omega)}$$
$$+ (\bar{\Theta}^\varepsilon_k - \bar{\theta}^{\varepsilon^N}, \hat{y} - \bar{y}^{\varepsilon^N})_{L_2(\Omega)}$$

Subtracting the term $\frac{\nu}{2}\left\|\Pi(P) - \bar{u}^{\varepsilon^N}\right\|^2_{L_2(\Omega)}$ in the inequality above and multiplying the resulting inequality by $\frac{2}{\nu}$, we obtain the desired result 5.1. □

Next we tackle Step 2, namely estimating the remaining multiplier terms

$$(\bar{\Theta}^\varepsilon_k, \hat{y} - \bar{y}^{\varepsilon^N})_{L_2(\Omega)},\ (\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - \hat{y})_{L_2(\Omega)} \tag{5.8}$$

starting with the first:

**Lemma 5.2** (estimate for $(\bar{\Theta}^\varepsilon_k, \hat{y} - \bar{y}^{\varepsilon^N})_{L_2(\Omega)}$)**.** *Let $(U, P, Y) \in L_2(\Omega) \times H^1_0(\Omega) \times \mathbb{Y}_k$ be arbitrary. Furthermore, let $\hat{y} = S\Pi(P)$ and $\hat{p} = S^*(Y - y_d - \bar{\Theta}^\varepsilon_k)$ be given. Finally, let $R_k$ denote the Ritz-projection on $\mathbb{Y}_k$ defined by*

$$(\nabla(R_k q - q), \nabla W)_{L_2(\Omega)} = 0\ \ \forall W \in \mathbb{Y}_k,\ q \in H^1_0(\Omega). \tag{5.9}$$

*Then*

$$\begin{aligned}(\bar{\Theta}^\varepsilon_k, \hat{y} - \bar{y}^{\varepsilon^N})_{L_2(\Omega)} \le{}& (\Pi(P), P - \hat{p})_{L_2(\Omega)} - (Y - y_d, R_k(SU) - \hat{y})_{L_2(\Omega)} \\ &+ (U - \Pi(P), P)_{L_2(\Omega)} + (\bar{\Theta}^\varepsilon_k, R_k(SU) - Y)_{L_2(\Omega)} + (\bar{\Theta}^\varepsilon_k, Y - I_k y_c)_{L_2(\Omega)} \\ &+ (\bar{\Theta}^\varepsilon_k, I_k y_c - y_c)_{L_2(\Omega)} + (\bar{\Theta}^\varepsilon_k, y_c - \bar{y}^{\varepsilon^N})_{L_2(\Omega)}.\end{aligned} \tag{5.10}$$

*Let $4 < p' < \infty$ in case $d = 2$ and $p' = 6$ in case $d = 3$. Besides, let $\gamma$ be as in (4.9). Then, for the last term in (5.10), we have the estimate*

$$(\bar{\Theta}^\varepsilon_k, y_c - \bar{y}^\varepsilon)_{L_2(\Omega)} \lesssim \left\|\bar{\Theta}^\varepsilon_k - P^{0+}_k \bar{\Theta}^\varepsilon_k\right\|_{L_2(\Omega)} + \left\|P^{0+}_k \bar{\Theta}^\varepsilon_k\right\|_{L_2(\Omega)} \varepsilon^{N(\frac{3}{2}+\gamma)}, \tag{5.11}$$

*where $P^{0+}_k$ is defined as in Definition 3.2.*

*Proof.* Let us first split the left hand side in (5.10) in the following way

$$(\bar{\Theta}^\varepsilon_k, \hat{y} - \bar{y}^{\varepsilon^N})_{L_2(\Omega)} = (\bar{\Theta}^\varepsilon_k, \hat{y} - y_c)_{L_2(\Omega)} + (\bar{\Theta}^\varepsilon_k, y_c - \bar{y}^{\varepsilon^N})_{L_2(\Omega)}. \tag{5.12}$$

The proof is now divided into two parts: First, we will prove (5.11) for the second term on the right in the equation above which already appears in (5.10) as the last term. Secondly, we will derive the rest of the bound in (5.10) by investigating the first term on the right in (5.12):

**1st part of the proof**:
We now turn our focus towards the second term on the right hand side in (5.12). To prove its additional property (5.11), we first split it in the following way with the help of the projection $P^{0+}_k$ defined in Definition 3.2:

$$(\bar{\Theta}^\varepsilon_k, y_c - \bar{y}^{\varepsilon^N})_{L_2(\Omega)} = (\bar{\Theta}^\varepsilon_k - P^{0+}_k \bar{\Theta}^\varepsilon_k, y_c - \bar{y}^{\varepsilon^N})_{L_2(\Omega)} + (P^{0+}_k \bar{\Theta}^\varepsilon_k, y_c - \bar{y}^{\varepsilon^N})_{L_2(\Omega)}. \tag{5.13}$$

Let us first tackle the second term on the right in the equation above. Since $(P^{0+}_k \bar{\Theta}^\varepsilon_k)(x) \ge 0$ f.a.a. $x \in \Omega$ by construction, we obtain using the penalty structure, Lemma 4.1:

$$(P^{0+}_k \bar{\Theta}^\varepsilon_k, y_c - \bar{y}^{\varepsilon^N})_{L_2(\Omega)} \le (P^{0+}_k \bar{\Theta}^\varepsilon_k, (y_c - \bar{y}^{\varepsilon^N})^+)_{L_2(\Omega)} \le \varepsilon^N \left\|P^{0+}_k \bar{\Theta}^\varepsilon_k\right\|_{L_2(\Omega)} \left\|\bar{v}^{\varepsilon^N}\right\|_{L_2(\Omega)} \tag{5.14}$$

Corollary 4.7 entails

$$\left\|\bar v^{\varepsilon^N}\right\|_{L_2(\Omega)} \lesssim \varepsilon^{N(\frac{1}{2}+\gamma)}$$

with $\gamma$ as in (4.9). Inserting this estimate in (5.14), we obtain:

$$(P_k^{0+}\bar\Theta_k^\varepsilon, y_c - \bar y^{\varepsilon^N})_{L_2(\Omega)} \lesssim \left\|P_k^{0+}\bar\Theta_k^\varepsilon\right\|_{L_2(\Omega)} \varepsilon^{N(\frac{3}{2}+\gamma)}$$

Going back to (5.13) and employing Cauchy-Schwarz's inequality, we have thus gained the bound

$$(\bar\Theta_k^\varepsilon, y_c - \bar y^{\varepsilon^N})_{L_2(\Omega)} \lesssim \left\|\bar\Theta_k^\varepsilon - P_k^{0+}\bar\Theta_k^\varepsilon\right\|_{L_2(\Omega)} \|y_c - \bar y^{\varepsilon^N}\|_{L_2(\Omega)} + \left\|P_k^{0+}\bar\Theta_k^\varepsilon\right\|_{L_2(\Omega)} \varepsilon^{N(\frac{3}{2}+\gamma)} \tag{5.15}$$

Because $\left\|y_c - \bar y^{\varepsilon^N}\right\|_{L_2(\Omega)}$ is bounded independent of $\varepsilon$ (compare Lemma 4.2) we finally derive (5.11).

This completes the first part of the proof, let us now tackle the second.

**2nd part of the proof**:
Let us have a look at the following term, its significance will become evident later on:

$$(\nabla(R_k(\hat p) - \hat p), \nabla(Y - \hat y))_{L_2(\Omega)}. \tag{5.16}$$

Using Galerkin-orthogonality, the Ritz-projection $R_k$, (5.9), $\hat y = S\Pi(P)$ and $\hat p = S^*(Y - y_d - \bar\Theta_k^\varepsilon)$ we can conclude - after some calculations - that:

$$\begin{aligned}(\nabla(R_k(\hat p) - \hat p), \nabla(Y - \hat y))_{L_2(\Omega)} &= (\nabla(R_k(\hat p) - \hat p), \nabla(Y - R_k(SU)))_{L_2(\Omega)}\\ &\quad + (\nabla(R_k(\hat p) - \hat p), \nabla(R_k(SU) - \hat y))_{L_2(\Omega)}\\ &= (R_k(\hat p), U)_{L_2(\Omega)} - (R_k(\hat p), \Pi(P))_{L_2(\Omega)}\\ &\quad - (Y - y_d - \bar\Theta_k^\varepsilon, R_k(SU) - \hat y)_{L_2(\Omega)}\end{aligned} \tag{5.17}$$

Now, let us look at (5.16) from another point of view. Harnessing once again the Ritz projection $R_k$, (5.9), and $\hat y = S\Pi(P)$ we immediately arrive at

$$(\nabla(R_k(\hat p) - \hat p), \nabla(Y - \hat y))_{L_2(\Omega)} = -(\nabla(R_k(\hat p) - \hat p), \nabla\hat y)_{L_2(\Omega)} = -(\Pi(P), R_k(\hat p)_{L_2(\Omega)} - \hat p)_{L_2(\Omega)}.$$

Rearranging (5.17), we derive

$$(\bar\Theta_k^\varepsilon, \hat y - R_k(SU))_{L_2(\Omega)} = (\Pi(P), R_k(\hat p) - \hat p)_{L_2(\Omega)} - (Y - y_d, R_k(SU) - \hat y)_{L_2(\Omega)} + (U - \Pi(P), R_k(\hat p))_{L_2(\Omega)}$$

and thus, combining this with our previous deductions, we obtain

$$\begin{aligned}(\bar\Theta_k^\varepsilon, \hat y - y_c)_{L_2(\Omega)} &= (\bar\Theta_k^\varepsilon, \hat y - R_k(SU))_{L_2(\Omega)} + (\bar\Theta_k^\varepsilon, R_k(SU) - Y)_{L_2(\Omega)} + (\bar\Theta_k^\varepsilon, Y - I_k y_c)_{L_2(\Omega)}\\ &\quad + (\bar\Theta_k^\varepsilon, I_k y_c - y_c)_{L_2(\Omega)}\\ &= (\Pi(P), R_k(\hat p) - \hat p)_{L_2(\Omega)} - (Y - y_d, R_k(SU) - \hat y)_{L_2(\Omega)} + (U - \Pi(P), R_k(\hat p))_{L_2(\Omega)}\\ &\quad + (\bar\Theta_k^\varepsilon, R_k(SU) - Y)_{L_2(\Omega)} + (\bar\Theta_k^\varepsilon, Y - I_k y_c)_{L_2(\Omega)} + (\bar\Theta_k^\varepsilon, I_k y_c - y_c)_{L_2(\Omega)}.\end{aligned}$$

This can now be inserted in (5.12) to gain the estimate (5.10) completing the proof. □

The next lemma deals with the second term in (5.8):

**Lemma 5.3** (estimate for $(\bar\theta^{\varepsilon^N}, \bar y^{\varepsilon^N} - \hat y)_{L_2(\Omega)}$). *Let $(Y, P) \in H_0^1(\Omega) \times H_0^1(\Omega)$ be arbitrary, $\hat y = S\Pi(P)$ and $4 < p' < \infty$ in case $d = 2$ and $p' = 6$ in case $d = 3$. Then the following estimate is valid*

$$\begin{aligned}(\bar\theta^{\varepsilon^N}, \bar y^\varepsilon - \hat y)_{L_2(\Omega)} \lesssim \min\Big\{ &s(\tau)\left\|(\Pi(P) + \Delta y_c)^-\right\|_{L_2(\Omega)}, c(p')s(\tau)\varepsilon^{-3N/p'}\big(|\hat y - Y|_{H^1(\Omega)}\\ &+ |(Y - I_k y_c)^-|_{H^1(\Omega)} + \|y_c - I_k y_c\|_{H^1(\Omega)}\big)\Big\} - \frac{1}{\varepsilon^N}\left\|\bar v^{\varepsilon^N}\right\|^2.\end{aligned} \tag{5.18}$$

*Proof.* We use the harmonic extension $H\iota$ for $\iota = y_c|_{\partial\Omega} \in H^{1/2}(\Omega)$ to split the critical termn in the following way:

$$(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon} - \hat{y})_{L_2(\Omega)} = (\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon} - (y_c - H\iota))_{L_2(\Omega)} + (\bar{\theta}^{\varepsilon^N}, y_c - H\iota - \hat{y})_{L_2(\Omega)}. \tag{5.19}$$

We note that due to the maximum principle for elliptic operators and $y_c|_{\partial\Omega} < 0$ we have $H\iota \leq 0$ a.e. in $\Omega$. Now, investigating the first term in the equation above, we take advantage of the complimentary slackness condition (2.5c) and $H\iota \leq 0$ a.e. to deduce:

$$(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - (y_c - H\iota))_{L_2(\Omega)} = -\frac{1}{\varepsilon^N} \left\| \bar{v}^{\varepsilon^N} \right\|^2_{L_2(\Omega)} + \underbrace{(\bar{\theta}^{\varepsilon^N}, H\iota)_{L_2(\Omega)}}_{\leq 0} \leq 0.$$

We still need to estimate the second term on the right in (5.19), though.
Using the 'singular part' $\bar{p}_s^{\varepsilon^N}$ of the adjoint state defined by $\bar{p}_s^{\varepsilon^N} := -S^*\bar{\theta}^{\varepsilon^N}$, $\nabla\hat{y}, \nabla y_c, \nabla H\iota \in H(\mathrm{div}, \Omega)$, $y_c - H\iota \in H_0^1(\Omega)$ and Green's formula, we can conclude:

$$(\bar{\theta}^{\varepsilon^N}, y_c - H\iota - \hat{y})_{L_2(\Omega)} = (-\nabla\bar{p}_s^{\varepsilon^N}, \nabla(y_c - H\iota - \hat{y}))_{L_2(\Omega)} = (\bar{p}_s^{\varepsilon^N}, \Delta y_c - \Delta H\iota - \Delta\hat{y})_{L_2(\Omega)}.$$

Now, observe that $\Delta H\iota = 0$ and $-\Delta\hat{y} = \Pi(P)$ a.e in $\Omega$. This allows us to conclude:

$$(\bar{\theta}^{\varepsilon^N}, y_c - H\iota - \hat{y})_{L_2(\Omega)} = (\bar{p}_s^{\varepsilon^N}, \Delta y_c - \Delta H\iota - \Delta\hat{y})_{L_2(\Omega)} = (\bar{p}^{\varepsilon^N}, \Delta y_c + \Pi(P))_{L_2(\Omega)}.$$

Thanks to $-\bar{\theta}^{\varepsilon^N} \leq 0$ a.e. in $\Omega$ (compare (2.5c)) and the maximum principle applied to $\bar{p}_s^{\varepsilon^N} = -S^*\bar{\theta}_s^{\varepsilon^N}$ we know that $\bar{p}^{\varepsilon^N} = -S^*\bar{\theta}^{\varepsilon^N} \leq 0$ a.e. Together with continuity of $S^*$ as a mapping $S^* : L_1(\Omega) \to W_s^1(\Omega) \hookrightarrow L_2(\Omega)$, $s < \frac{d}{d-1}$ (compare [42], Theorem 9.1) and boundedness of $\bar{\theta}^{\varepsilon^N}$ in $L_1(\Omega)$, cf Lemma 4.3, we have:

$$\left\| \bar{p}_s^{\varepsilon^N} \right\|_{L_2(\Omega)} = \left\| S^*\bar{\theta}^{\varepsilon^N} \right\|_{L_2(\Omega)} \lesssim \left\| \bar{\theta}^{\varepsilon^N} \right\|_{L_1(\Omega)} \lesssim s(\tau),$$

With the observation above we can deduce that

$$(\bar{p}_s^{\varepsilon^N}, \Delta y_c + \Pi(P))_{L_2(\Omega)} \lesssim s(\tau) \left\| (\Delta y_c + \Pi(P))^- \right\|_{L_2(\Omega)} \tag{5.20}$$

All in all, we have proven the bound:

$$(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon} - \hat{y})_{L_2(\Omega)} \leq \left\| \bar{p}_s^{\varepsilon^N} \right\|_{L_2(\Omega)} \left\| (\Delta y_c + \Pi(P))^- \right\|_{L_2(\Omega)} - \frac{1}{\varepsilon^N} \left\| \bar{v}^{\varepsilon^N} \right\|^2_{L_2(\Omega)} =: min_1. \tag{5.21}$$

Neglecting the shift $-\frac{1}{\varepsilon^N} \left\| \bar{v}^{\varepsilon^N} \right\|^2_{L_2(\Omega)}$, we realise that the inequality above gives the first argument in the min operation in (5.18).
To derive the other bound given by the second argument in the min operation in (5.18), we return to the start, (5.19), and split the critical term in the following different way:

$$(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon} - \hat{y})_{L_2(\Omega)} = (\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon} - y_c)_{L_2(\Omega)} + (\bar{\theta}^{\varepsilon^N}, y_c - \hat{y})_{L_2(\Omega)} \tag{5.22}$$

For the first term, we once again take advantage of the complimentary slackness condition (2.5c)

$$(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon} - y_c)_{L_2(\Omega)} = -\frac{1}{\varepsilon^N} \left\| \bar{v}^{\varepsilon^N} \right\|^2_{L_2(\Omega)}.$$

For the other term on the right in (5.22) we discern that

$$(\bar{\theta}^{\varepsilon^N}, y_c - \hat{y})_{L_2(\Omega)} = (\bar{\theta}^{\varepsilon^N}, y_c - I_k y_c)_{L_2(\Omega)} + (\bar{\theta}^{\varepsilon^N}, I_k y_c - Y)_{L_2(\Omega)} + (\bar{\theta}^{\varepsilon^N}, Y - \hat{y})_{L_2(\Omega)}$$

Utilising the bound for $\bar{\theta}^{\varepsilon^N}$ derived in Lemma 4.3, $\bar{\theta}^{\varepsilon^N} \geq 0$ (compare (2.5c)) and the embedding $H_0^1(\Omega) \hookrightarrow L_p(\Omega)$ as well as Poincaré -Friedrich's inequality, we obtain for $1 \leq p \leq 2$ and $\frac{1}{p} + \frac{1}{p'} = 1$:

$$
\begin{aligned}
(\bar{\theta}^{\varepsilon^N}, y_c - I_k y_c)_{L_2(\Omega)} &\leq \left\| \bar{\theta}^{\varepsilon^N} \right\|_{L_p(\Omega)} \|y_c - I_k y_c\|_{L_{p'}(\Omega)} \lesssim c(p')s(\tau)\varepsilon^{-3N/p'} \|y_c - I_k y_c\|_{H^1(\Omega)} \\
(\bar{\theta}^{\varepsilon^N}, I_k y_c - Y)_{L_2(\Omega)} &\leq \left\| \bar{\theta}^{\varepsilon^N} \right\|_{L_p(\Omega)} \left\|(Y - I_k y_c)^-\right\|_{L_{p'}(\Omega)} \lesssim c(p')s(\tau)\varepsilon^{-3N/p'} |(Y - I_k y_c)^-|_{H^1(\Omega)} \\
(\bar{\theta}^{\varepsilon^N}, Y - \hat{y})_{L_2(\Omega)} &\leq \left\| \bar{\theta}^{\varepsilon^N} \right\|_{L_p(\Omega)} \|Y - \hat{y}\|_{L_{p'}(\Omega)} \lesssim c(p')s(\tau)\varepsilon^{-3N/p'} |Y - \hat{y}|_{H^1(\Omega)},
\end{aligned}
$$

Reviewing the previous estimates, we have proven the following bound given by the second argument of the min argument in (5.18):

$$
\begin{aligned}
(\bar{\theta}^{\varepsilon^N}, \bar{y}^\varepsilon - \hat{y})_{L_2(\Omega)} \lesssim\ & c(p')s(\tau)\varepsilon^{-3N/p'} \big( \|y_c - I_k y_c\|_{H^1(\Omega)} + |(Y - I_k y_c)^-|_{H^1(\Omega)} + |Y - \hat{y}|_{H^1(\Omega)} \big) \\
& - \frac{1}{\varepsilon^N} \left\| \bar{v}^{\varepsilon^N} \right\|_{L_2(\Omega)}^2 =: min_2
\end{aligned}
$$

Combining this with (5.21), we obtain

$$
(\bar{\theta}^{\varepsilon^N}, \bar{y}^\varepsilon - \hat{y})_{L_2(\Omega)} \lesssim \min(min_1, min_2)
$$

which is the desired result. □

With these auxiliary results we are now in a position to prove a reliable estimator for the variational discretisation approach, i.e. Theorem 3.4. This is the third step of our 'roadmap', presented at the beginning of this section:

### 5.1. **Proof of Theorem 3.4.** :

*Proof of Theorem 3.4.* First of all, we split the error $\left\|\bar{u} - \bar{U}_k^\varepsilon\right\|_{L_2(\Omega)}^2$ with the help of the triangle and weighted Young's inequality in the following way

$$
\|\bar{u} - \bar{U}_k^\varepsilon\|_{L_2(\Omega)}^2 \leq 2 \left\|\bar{u} - \bar{u}^{\varepsilon^N}\right\|_{L_2(\Omega)}^2 + 2 \left\|\bar{u}^{\varepsilon^N} - \bar{U}_k^\varepsilon\right\|_{L_2(\Omega)}^2. \tag{5.23}
$$

For the first term on the right in the inequality above we can use Theorem 4.5 and Corollary 4.7 to deduce

$$
\left\|\bar{u} - \bar{u}^{\varepsilon^N}\right\|_{L_2(\Omega)}^2 \lesssim \frac{1}{\nu} c(p')s(\tau)\varepsilon^{\gamma N}.
$$

with $\gamma$ defined in (4.9).
For the second term on the right in (5.23) we employ Lemma 5.1 with $U = \bar{U}_k^\varepsilon = \Pi(\bar{P}_k^\varepsilon)$, $P = \bar{P}_k^\varepsilon$ and $Y = \bar{Y}_k^\varepsilon$ to gain

$$
\begin{aligned}
2 \left\|\bar{u}^{\varepsilon^N} - \bar{U}_k^\varepsilon\right\|_{L_2(\Omega)}^2 \leq\ & \frac{1}{\nu} \left\|S\bar{U}_k^\varepsilon - \bar{Y}_k^\varepsilon\right\|_{L_2(\Omega)}^2 + \frac{2}{\nu^2} \left\|S^*(\bar{Y}_k^\varepsilon - y_d - \bar{\Theta}_k^\varepsilon) - \bar{P}_k^\varepsilon\right\|_{L_2(\Omega)}^2 \\
& + \frac{4}{\nu}(\bar{\Theta}_k^\varepsilon, S\bar{U}_k^\varepsilon - \bar{y}^{\varepsilon^N})_{L_2(\Omega)} + \frac{4}{\nu}(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - S\bar{U}_k^\varepsilon)_{L_2(\Omega)}.
\end{aligned} \tag{5.24}
$$

Investigating the second to last term on the right, we harness Lemma 5.2 and $R_k(S\bar{U}_k^\varepsilon) = \bar{Y}_k^\varepsilon$ to conclude with the additional help of (5.11):

$$\begin{aligned}
(\bar{\Theta}_k^\varepsilon, S\bar{U}_k^\varepsilon - \bar{y}^{\varepsilon^N})_{L_2(\Omega)} &\le (\bar{U}_k^\varepsilon, \bar{P}_k^\varepsilon - S^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon))_{L_2(\Omega)} - (\bar{Y}_k^\varepsilon - y_d, \bar{Y}_k^\varepsilon - \hat{y})_{L_2(\Omega)} \\
&\quad + (\bar{\Theta}_k^\varepsilon, \bar{Y}_k^\varepsilon - I_k y_c)_{L_2(\Omega)} + (\bar{\Theta}_k^\varepsilon, I_k y_c - y_c)_{L_2(\Omega)} + (\bar{\Theta}_k^\varepsilon, y_c - \bar{y}^{\varepsilon^N})_{L_2(\Omega)} \\
&\lesssim \left\|\bar{U}_k^\varepsilon\right\|_{L_2(\Omega)} \left\|\bar{P}_k^\varepsilon - S^*(\bar{Y}_k^\varepsilon - y_d - \bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)} \\
&\quad + \left\|\bar{Y}_k^\varepsilon - y_d\right\|_{L_2(\Omega)} \left\|\bar{Y}_k^\varepsilon - \hat{y}\right\|_{L_2(\Omega)} + (\bar{\Theta}_k^\varepsilon, \bar{Y}_k^\varepsilon - I_k y_c)_{L_2(\Omega)} \\
&\quad + \left\|\bar{\Theta}_k^\varepsilon - P_k^{0+}(\bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)} + c(p')s(\tau)\left\|P_k^{0+}(\bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)} \varepsilon^{N(\frac{3}{2}+\gamma)}
\end{aligned} \tag{5.25}$$

Combining (5.24) and (5.25) and the complimentary slackness condition for the discrete problem (2.10c), we obtain

$$\begin{aligned}
\left\|\bar{U}_k^\varepsilon - \bar{u}^{\varepsilon^N}\right\|_{L_2(\Omega)}^2 &\lesssim \mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) + \frac{4}{\nu}\left\|\bar{U}_k^\varepsilon\right\|_{L_2(\Omega)} \left\|\bar{P}_k^\varepsilon - S^*(\bar{Y}_k^\varepsilon - y_d - \bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)} \\
&\quad + \frac{4}{\nu}\left\|\bar{Y}_k^\varepsilon - y_d\right\|_{L_2(\Omega)} \left\|\bar{Y}_k^\varepsilon - \hat{y}\right\|_{L_2(\Omega)} + \frac{4}{\nu}\left\|\bar{\Theta}_k^\varepsilon - P_k^{0+}(\bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)} \\
&\quad + \frac{4}{\nu}c(p')s(\tau)\left\|P_k^{0+}(\bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)} \varepsilon^{N(\frac{3}{2}+\gamma)} + \frac{4}{\nu}(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - S\bar{U}_k^\varepsilon)_{L_2(\Omega)}
\end{aligned} \tag{5.26}$$

The last term remains to be estimated.
Recalling Lemma 5.3 with $Y = \bar{Y}_k^\varepsilon$, $P = \bar{P}_k^\varepsilon$ and $\bar{U}_k^\varepsilon = \Pi(\bar{P}_k^\varepsilon)$, we deduce

$$\begin{aligned}
\frac{4}{\nu}(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - S\bar{U}_k^\varepsilon)_{L_2(\Omega)} &\lesssim \frac{4}{\nu}\min\Bigg\{ s(\tau)\left\|(\bar{U}_k^\varepsilon + \Delta y_c)^-\right\|_{L_2(\Omega)}, \\
&\quad + c(p')s(\tau)\varepsilon^{-3N/p'}\Bigg( |S\bar{U}_k^\varepsilon - \bar{Y}_k^\varepsilon|_{H^1(\Omega)} + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)} \\
&\quad + \|y_c - I_k y_c\|_{H^1(\Omega)}\Bigg)\Bigg\} - \frac{4}{\nu\varepsilon^N}\left\|\bar{v}^{\varepsilon^N}\right\|_{L_2(\Omega)}^2
\end{aligned}$$

The term $\frac{4}{\nu\varepsilon^N}\left\|\bar{v}^{\varepsilon^N}\right\|_{L_2(\Omega)}^2$ can be shifted to the left in (5.24) to complete the left hand side in (3.2). Presently, inserting the estimate above in (5.26) and recalling the definition of $\mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$, (3.4), we deduce the the right hand side in (3.2), i.e.:

$$\left\|\bar{U}_k^\varepsilon - \bar{u}\right\|_{L_2(\Omega)}^2 + \frac{4}{\nu\varepsilon^N}\left\|\bar{v}^{\varepsilon^N}\right\|_{L_2(\Omega)}^2 \lesssim \frac{1}{\nu}c(p')s(\tau)\varepsilon^{\gamma N} + \mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) + \mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$$

□

Let us now prove the analogous result for the full discretisation technique, i.e. Theorem 3.5. Here, $\mathbb{U}_k = \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_0, L_2(\Omega))$.

5.2. **Proof of Theorem 3.5.** :

*Proof of Theorem 3.5.* As in the proof of Theorem 3.4 we harness the results of Lemmas 5.1, 5.2 and 5.3 to prove the bound (3.5).
A similar split as in (5.23) and Corollary 4.7 yield

$$\left\|\bar{U}_k^\varepsilon - \bar{u}\right\|_{L_2(\Omega)}^2 \lesssim 2\left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|_{L_2(\Omega)}^2 + 4\left\|\Pi(\bar{P}_k^\varepsilon) - \bar{u}^{\varepsilon^N}\right\|_{L_2(\Omega)}^2 + \frac{4}{\nu}c(p')s(\tau)\varepsilon^{\gamma N}. \tag{5.27}$$

The first term $\left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|^2_{L_2(\Omega)}$ can be evaluated exactly (compare again [26], Remark 4.3). Thus, the aim now is to control the term $\left\|\Pi(\bar{P}_k^\varepsilon) - \bar{u}^{\varepsilon^N}\right\|^2_{L_2(\Omega)}$.
Inserting $(\bar{Y}_k^\varepsilon, \bar{P}_k^\varepsilon)$ for $(Y, P)$ in Lemma 5.1, we obtain

$$
\begin{aligned}
\left\|\Pi(\bar{P}_k^\varepsilon) - \bar{u}^{\varepsilon^N}\right\|^2_{L_2(\Omega)} \leq{}& \frac{1}{2\nu}\left\|S\Pi(\bar{P}_k^\varepsilon) - \bar{Y}_k^\varepsilon\right\|^2_{L_2(\Omega)} + \frac{1}{\nu^2}\left\|S^*(\bar{Y}_k^\varepsilon - y_d - \bar{\Theta}_k^\varepsilon) - \bar{P}_k^\varepsilon\right\|^2_{L_2(\Omega)} \\
&+ \frac{2}{\nu}(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - S\Pi(\bar{P}_k^\varepsilon))_{L_2(\Omega)} + \frac{2}{\nu}(\bar{\Theta}_k^\varepsilon, S\Pi(\bar{P}_k^\varepsilon) - \bar{y}^{\varepsilon^N})_{L_2(\Omega)}
\end{aligned}
\tag{5.28}
$$

We will derive the regular part $\mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$, (3.6), first: To begin with, we discern with $\bar{Y}_k^\varepsilon = S_k\bar{U}_k^\varepsilon$ and continuity of $S$ that

$$
\left\|S\Pi(\bar{P}_k^\varepsilon) - \bar{Y}_k^\varepsilon\right\|^2_{L_2(\Omega)} \lesssim 2\left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|^2_{L_2(\Omega)} + 2\left\|(S - S_k)\bar{U}_k^\varepsilon\right\|^2_{L_2(\Omega)}
\tag{5.29}
$$

Let us now tackle the scalar product terms in (5.28) beginning with $(\bar{\Theta}_k^\varepsilon, S\Pi(\bar{P}_k^\varepsilon) - \bar{y}^{\varepsilon^N})_{L_2(\Omega)}$:
To gain an estimate for $(\bar{\Theta}_k^\varepsilon, S\Pi(\bar{P}_k^\varepsilon) - \bar{y}^{\varepsilon^N})_{L_2(\Omega)}$, we employ Lemma 5.2 with $U = \bar{U}_k^\varepsilon$, $P = \bar{P}_k^\varepsilon$ and $Y = \bar{Y}_k^\varepsilon$, Cauchy-Schwarz inequality and (5.11); also recollect that $R_k(S\bar{U}_k^\varepsilon) = \bar{Y}_k^\varepsilon$:

$$
\begin{aligned}
(\bar{\Theta}_k^\varepsilon, S\Pi(\bar{P}_k^\varepsilon) - \bar{y}^{\varepsilon^N})_{L_2(\Omega)} \lesssim{}& \left\|\Pi(\bar{P}_k^\varepsilon)\right\|_{L_2(\Omega)}\left\|(S - S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)} \\
&+ \left\|\bar{Y}_k^\varepsilon - y_d\right\|_{L_2(\Omega)}(\left\|(S - S_k)\bar{U}_k^\varepsilon\right\|_{L_2(\Omega)} + \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|_{L_2(\Omega)}) \\
&+ \left\|\bar{\Theta}_k^\varepsilon - P_k^{0+}(\bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)} + c(p')s(\tau)\left\|P_k^{0+}(\bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)}\varepsilon^{N(\frac{3}{2}+\gamma)} \\
&+ (\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon), \bar{P}_k^\varepsilon)_{L_2(\Omega)} + (\bar{\Theta}_k^\varepsilon, I_k y_c - y_c)_{L_2(\Omega)} + (\bar{\Theta}_k^\varepsilon, \bar{Y}_k^\varepsilon - I_k y_c)_{L_2(\Omega)}
\end{aligned}
\tag{5.30}
$$

This bound can now be inserted in (5.28): We first set

$$
\begin{aligned}
\mathcal{X} :={}& \left\|\Pi(\bar{P}_k^\varepsilon)\right\|_{L_2(\Omega)}\left\|(S - S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)} \\
&+ \left\|\bar{Y}_k^\varepsilon - y_d\right\|_{L_2(\Omega)}(\left\|(S - S_k)\bar{U}_k^\varepsilon\right\|_{L_2(\Omega)} + \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|_{L_2(\Omega)}) \\
&+ \left\|\bar{\Theta}_k^\varepsilon - P_k^{0+}(\bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)} + c(p')s(\tau)\left\|P_k^{0+}(\bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)}\varepsilon^{N(\frac{3}{2}+\gamma)}
\end{aligned}
\tag{5.31}
$$

and observe that $\mathcal{X}$ already contains many of the terms given in (3.7).
With this definition of $\mathcal{X}$ we plug (5.30) in (5.28) and afterwards (5.28) in (5.27). We obtain:

$$
\begin{aligned}
\left\|\bar{U}_k^\varepsilon - \bar{u}\right\|^2_{L_2(\Omega)} \lesssim{}& \frac{8}{\nu}c(p')s(\tau)\varepsilon^{\gamma N} + (2 + \frac{4}{\nu})\left\|\Pi(\bar{P}_k^\varepsilon) - \bar{U}_k^\varepsilon\right\|^2_{L_2(\Omega)} \\
&+ \frac{4}{\nu}\left\|(S - S_k)\bar{U}_k^\varepsilon\right\|^2_{L_2(\Omega)} + \frac{4}{\nu^2}\left\|S^*(\bar{Y}_k^\varepsilon - y_d - \bar{\Theta}_k^\varepsilon) - \bar{P}_k^\varepsilon\right\|^2_{L_2(\Omega)} \\
&+ \frac{8}{\nu}(\bar{\Theta}_k^\varepsilon, I_k y_c - y_c)_{L_2(\Omega)} + \frac{8}{\nu}(\bar{\Theta}_k^\varepsilon, \bar{Y}_k^\varepsilon - I_k y_c)_{L_2(\Omega)} + \frac{8}{\nu}(\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon), \bar{P}_k^\varepsilon)_{L_2(\Omega)} \\
&+ \frac{8}{\nu}\mathcal{X} + \frac{8}{\nu}(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - S\Pi(\bar{P}_k^\varepsilon))_{L_2(\Omega)}
\end{aligned}
$$

In short, bearing in mind the complimentary slackness condition on the discrete level (2.10c):

$$
\left\|\bar{U}_k^\varepsilon - \bar{u}\right\|^2_{L_2(\Omega)} \lesssim \frac{1}{\nu}c(p')s(\tau)\varepsilon^{\gamma N} + \mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) + \frac{8}{\nu}\mathcal{X} + \frac{8}{\nu}(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - S\Pi(\bar{P}_k^\varepsilon))_{L_2(\Omega)}
\tag{5.32}
$$

The last term remains to be estimated:
We employ Lemma 5.3 with $Y = \bar{Y}_k^\varepsilon$ and $P = \bar{P}_k^\varepsilon$ to obtain:

$$
\begin{aligned}
(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - S\Pi(\bar{P}_k^\varepsilon))_{L_2(\Omega)} \lesssim \min \Big\{ & s(\tau) \left\| (\Pi(\bar{P}_k^\varepsilon) + \Delta y_c)^- \right\|_{L_2(\Omega)}, \\
& + c(p')s(\tau)\varepsilon^{-3N/p'} \Big( \left| S\Pi(\bar{P}_k^\varepsilon) - S_k \bar{U}_k^\varepsilon \right|_{H^1(\Omega)} \\
& + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)} + \|y_c - I_k y_c\|_{H^1(\Omega)} \Big) \Big\} - \frac{8}{\nu \varepsilon^N} \left\| \bar{v}^{\varepsilon^N} \right\|^2 .
\end{aligned} \tag{5.33}
$$

Except for $\left| S\Pi(\bar{P}_k^\varepsilon) - S_k \bar{U}_k^\varepsilon \right|_{H^1(\Omega)}$ all terms already appear in (3.7). Let us therefore further estimate this term. Using continuity of $S$ and setting $|S| = \|S\|_{\mathcal{L}(L_2(\Omega), H^1(\Omega))}$, we derive

$$
\begin{aligned}
(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - S\Pi(\bar{P}_k^\varepsilon))_{L_2(\Omega)} \lesssim \min \Big\{ & s(\tau) \left\| (\Pi(\bar{P}_k^\varepsilon) + \Delta y_c)^- \right\|_{L_2(\Omega)}, \\
& c(p')s(\tau)\varepsilon^{-3N/p'} \Big( \left| (S - S_k) \bar{U}_k^\varepsilon \right|_{H^1(\Omega)} + \left\| \Pi(\bar{P}_k^\varepsilon) - \bar{U}_k^\varepsilon \right\| \\
& + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)} + \|y_c - I_k y_c\|_{H^1(\Omega)} \Big) \Big\} - \frac{8}{\nu \varepsilon^N} \left\| \bar{v}^{\varepsilon^N} \right\|^2 =: \mathcal{Y}
\end{aligned} \tag{5.34}
$$

We observe that by definition (compare (5.31)):

$$
\mathcal{X} + \mathcal{Y} = \mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) - \frac{8}{\nu \varepsilon^N} \left\| \bar{v}^{\varepsilon^N} \right\|_{L_2(\Omega)}^2 .
$$

Bearing this relation in mind, we can now insert the bound derived in (5.34) in (5.32) to deduce:

$$
\left\| \bar{U}_k^\varepsilon - \bar{u} \right\|_{L_2(\Omega)}^2 \lesssim c(p')s(\tau)\varepsilon^{\gamma N} + \mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) + \mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) - \frac{8}{\nu \varepsilon^N} \left\| \bar{v}^{\varepsilon^N} \right\|_{L_2(\Omega)}^2 .
$$

Shifting the term $-\frac{8}{\nu \varepsilon^N} \left\| \bar{v}^{\varepsilon^N} \right\|^2$ to the left, we obtain the desired inequality (3.5). This completes the proof. □

## 6. Numerical Experiments

Let us first present our discretisation ansatz for the examples considered in Sections 6.3.1 and 6.3.2:

### 6.1. **Discrete Setting & The Adaptive Algorithm.**

For our discrete spaces we make the following ansatz:

$$
\begin{aligned}
\mathbb{U}_k &= \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_0, L_2(\Omega)) \\
\mathbb{Y}_k = \mathbb{V}_k &= \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_1, H_0^1(\Omega))
\end{aligned}
$$

To present the adaptive algorithm, we will follow the structure provided to us by the adaptive loop

$$
\text{SOLVE} \to \text{ESTIMATE} \to \text{MARK} \to \text{REFINE}
$$

6.1.1. *SOLVE.* The optimisation problems for fixed $\varepsilon$ and fixed $k$, Problem $(P_k^\varepsilon)$, were solved with the help of a primal-dual-active-set-strategy (see e.g. [10] and [30]).

6.1.2. *ESTIMATE.* For the linear errors in our estimators $\mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ and $\mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ we used standard residual type error estimators in the respective $L_2(\Omega)$- and $H_0^1(\Omega)$-norms (compare e.g [1] and [37], Chapter 6). We would like to stress that other reliable a posteriori error estimators for the linear errors $(S - S_k)$ such as hierarchical ones, e.g. [44], or other, [5], are applicable too.
With these residual error estimators we obtain the following upper bounds - provided $S$ maps continuously to $H^2(\Omega) \cap H_0^1(\Omega)$ -

$$
\begin{aligned}
\left\|(S - S_k)(\bar{U}_k^\varepsilon + f)\right\|_{L_2(\Omega)}^2 &\lesssim \sum_{T \in \mathcal{T}_k} h_T^4 \left\|\bar{U}_k^\varepsilon + f\right\|_{L_2(T)}^2 + h_T^3 \left\|[\![S_k(\bar{U}_k^\varepsilon + f)]\!]\right\|_{L_2(\partial T)}^2 \\
\left|(S - S_k)(\bar{U}_k^\varepsilon + f)\right|_{H^1(\Omega)}^2 &\lesssim \sum_{T \in \mathcal{T}_k} h_T^2 \left\|\bar{U}_k^\varepsilon + f\right\|_{L_2(T)}^2 + h_T \left\|[\![S_k(\bar{U}_k^\varepsilon + f)]\!]\right\|_{L_2(\partial T)}^2 \\
\left\|(S - S_k)(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon)\right\|_{L_2(\Omega)}^2 &\lesssim \sum_{T \in \mathcal{T}_k} h_T^4 \left\|\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon\right\|_{L_2(T)}^2 \\
&\quad + h_T^3 \left\|[\![S_k(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon)]\!]\right\|_{L_2(\partial T)}^2 ,
\end{aligned} \tag{6.1}
$$

where $[\![\cdot]\!]$ denotes the normal jump across two elements $T^+, T^- \in \mathcal{T}_k$ sharing a common side $\partial T$. In addition, the each summand on the right in (6.1) serves as a local indicator on every $T$ for future marking purposes.
The estimator $\mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ does not lend itself to immediate localisation, because it does not contain scalar products or squared $L_2(\Omega), H_0^1(\Omega)$-norms. Hence, we have to square it once again to obtain:

$$
\begin{aligned}
\mathcal{E}_s^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) \lesssim{} & \left\|\bar{Y}_k^\varepsilon - y_d\right\|_{L_2(\Omega)}^2 \left(\left\|(S - S_k)(\bar{U}_k^\varepsilon + f)\right\|_{L_2(\Omega)}^2 + \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|_{L_2(\Omega)}^2\right) \\
& + \left\|\Pi(\bar{P}_k^\varepsilon)\right\|_{L_2(\Omega)}^2 \left\|(S - S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)}^2 \\
& + c(p')s(\tau) \left\|P_k^{0+}(\bar{\Theta}_k^\varepsilon)\right\|_{L_2(\Omega)}^2 \varepsilon^{N(\frac{6}{2}+2\gamma)} + \left\|\bar{\Theta}_k^\varepsilon - P_k^{0+}\bar{\Theta}_k^\varepsilon\right\|_{L_2(\Omega)}^2 \\
& + \min\Big\{ s(\tau) \left\|(\Pi(\bar{P}_k^\varepsilon) + \Delta y_c)^-\right\|_{L_2(\Omega)}^2 , c(p')s(\tau)\varepsilon^{-3N/p'} \big(|(S - S_k)(\bar{U}_k^\varepsilon + f)|_{H^1(\Omega)}^2 \\
& + \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|_{L_2(\Omega)}^2 + \|y_c - I_k y_c\|_{H^1(\Omega)}^2 + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)}^2\big) \Big\}
\end{aligned}
$$

Now, we have the desired structure with squared $L_2(\Omega)$- and $H_0^1(\Omega)$-norms which can easily be localised on each $T \in \mathcal{T}_k$ using (6.1) for the linear errors $S - S_k$ or plain numerical integration on each $T \in \mathcal{T}_k$. Naturally, the different exponents in $\mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ and $\mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ need to be reflected in the marking algorithm. We will expound on this in the next section, let us now first make it clear that at the end of the 'ESTIMATE' module of the adaptive loop we are in possession of computable estimates

$$
\begin{aligned}
\mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) &\lesssim \sum_{T \in \mathcal{T}_k} E_r^2(T) \\
\mathcal{E}_s^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) &\lesssim \sum_{T \in \mathcal{T}_k} E_s^2(T)
\end{aligned} \tag{6.2}
$$

with local indicators $E_r^2(T)$ and $E_s^2(T)$ respectively.

6.1.3. *MARK and REFINE.* With the local indicators $E_r^2(T)$ and $E_s^2(T)$, recall (6.2), we now define the following marking algorithm which basically is a modified maximum strategy: Choose paramteres $\eta_r, \eta_s \in (0,1)$ and

$$
\textbf{MARK } T \in \mathcal{T}_k \textbf{ iff } E_r(T) \geq \eta_r \max_{T \in \mathcal{T}_k} \vee E_s(T) \geq \eta_s \max_{T \in \mathcal{T}_k} E_s(T). \tag{6.3}
$$

For the marked elements, we then perform a recursive refinement algorithm (compare [27]).
In some cases, the regularisation error may have a significant influence on the total error. To remedy this disadvantage, we performed an $\varepsilon$-adaptation in some examples.

6.2. **$\varepsilon$-Adaptation.** It was motivated by the fact that under certain conditions - the crucial one being the existence of a sequence of discrete Slater points $U_k^s$ with $S_k U_k^s - I_k y_c \geq \kappa$, $\kappa > 0$ - one can show that

$$\left\|\bar{U}_k - \bar{U}_k^\varepsilon\right\|^2_{L_2(\Omega)} \lesssim \varepsilon^{1-3/p'} |\bar{V}_k^\varepsilon|_{H^1(\Omega)},\ p' \geq 2,$$

where $\bar{U}_k$ denotes the (unknown) unique discrete solution to the discrete unregularised problem. We now take $\left\|\bar{U}_k - \bar{U}_k^\varepsilon\right\|^2$ as an indicator for the error solely incurred by regularisation and adapt $\varepsilon$ according to the following scheme: Given a tolerance $TOL > 0$ and parameters $c_a$, $0 < \beta < 1$ as well as the current value of the regularisation parameter $\varepsilon_{curr}$ and a minimal one $\varepsilon_{\min}$, we set the new regularisation parameter $\varepsilon_{new}$ by:

$$\begin{aligned} &\textbf{if } \varepsilon^{1-3/p'} |\bar{V}_k^\varepsilon|_{H^1(\Omega)} > c_a^2 TOL^2 \textbf{ do } \varepsilon_{new} = \max(\beta\varepsilon_{curr}, \varepsilon_{\min}) \\ &\textbf{else } \varepsilon_{new} = \varepsilon_{curr} \end{aligned} \tag{6.4}$$

6.3. **Numerical Results.** The numerical experiments were done with the help of the finite element toolbox ALBERTA, see [40].
For our numerical experiments we constructed an analytical solution to check the performance of our adaptive algorithm. To construct such a solution to a problem of type $(P)$, we include an additional source term $f$ in the PDE for the state $\bar{y}$. This slight modification does not change the analyses of the previous sections. Thus, the solution $\bar{u}, \bar{y}, \bar{p}, \bar{\theta}$ to the following optimality system solves problem $(P)$ with the additional source term $f$ included in the equation for the state $\bar{y}$:

$$\begin{aligned} -\Delta\bar{y} &= \bar{u} + f \text{ in } \Omega \\ \bar{y} &= 0 \qquad \text{on } \partial\Omega \\ -\Delta\bar{p} &= \bar{y} - y_d - \bar{\theta} \text{ in } \Omega \\ \bar{p} &= 0 \qquad\qquad \text{on } \partial\Omega \\ \bar{u}(x) &= \min(\max(-\frac{1}{\nu}\bar{p}(x), a), b) \text{ f.a.a. } x \in \Omega \\ \langle\bar{\theta}, \bar{y} - y_c\rangle &= 0 \end{aligned} \tag{6.5}$$

We deliberately leave some ambiguity as to the duality product in the last line: In Section 6.3.1 we have $\langle\bar{\theta}, \bar{y} - y_c\rangle = \langle\bar{\theta}, \bar{y} - y_c\rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})}$ while in Section 6.3.2 we have $\langle\bar{\theta}, \bar{y} - y_c\rangle = (\bar{\theta}, \bar{y} - y_c)_{L_2(\Omega)}$.

6.3.1. *A Dirac Multiplier Example.* This example represents to a certain extent a 'worst case' setting for state-constrained problems in two space dimensions. However, such a 'worst-case' is certainly not unusual in practice and thus not only of theoretical interest. The true solution to $(P)$ is constructed in such a way that the Lagrange multiplier for the state constraint is a Dirac measure, i.e. solely an element in $C(\bar{\Omega})^*$.

The ensuing problem setting is given: With $\Omega = B_1(0) \subset \mathbb{R}^2$, $a = -10^{12}$ and $b = 10^{12}$ we define:

$$\begin{aligned} \bar{y}(x) &= \sin(\pi|x|^2) \\ \bar{p}(x) &= -35 \ln|x| \\ \bar{u}(x) &= \min(\max(-\frac{1}{\nu}\bar{p}(x), a), b) \\ y_c(x) &= \sin(\pi|x|^2) - |x|^2 \\ \bar{\theta} &= 35\delta(0) \end{aligned}$$

With $\bar{y} = y_d$ and $f$ adjusted in such a way that the state equation in (6.5) is fulfilled one can show that $\bar{u}, \bar{y}, \bar{p}, \bar{\theta}$ solve the optimality system (6.5).
In these low-regularity settings it has been frequently observed that the error due to regularisation has a significant influence on the overall error, that is why we performed an $\varepsilon$-adaptation according to our strategy (6.4). We obtain the following results for $N = 12$, $p' = 3$, parameters $\eta_r = 0.5$ and $\eta_s = 0.8$ for our refinement strategy (6.3), as well as $\beta = 0.93$ and $c_a = 6.5$ for our $\varepsilon$-adaptation, (6.4):

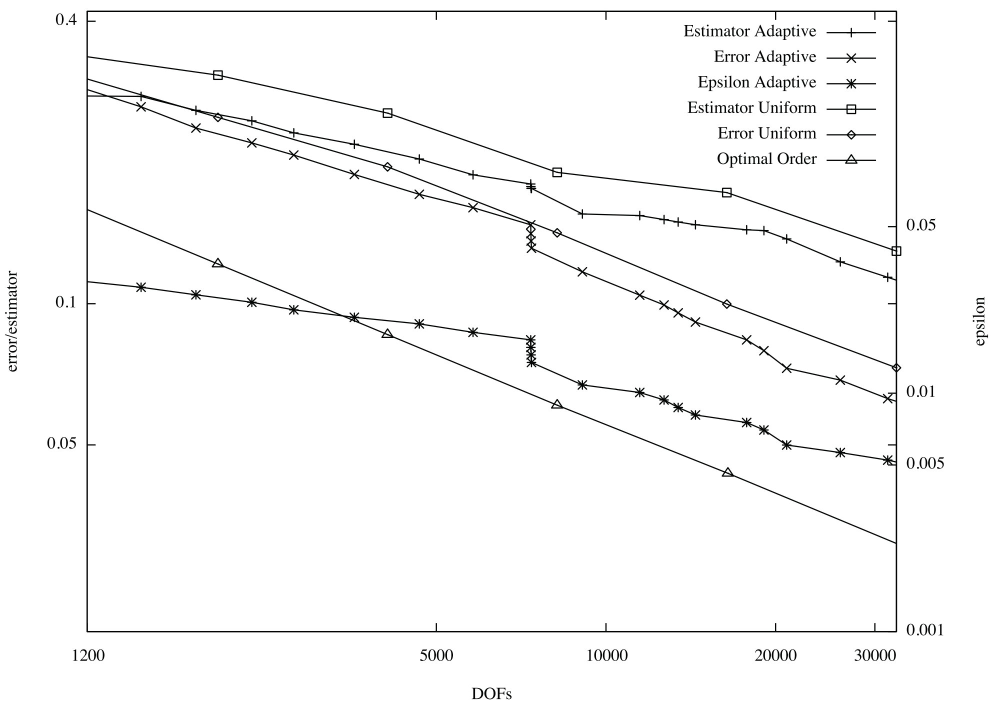


FIGURE 1. smooth example: performance of adaptive and uniform refinement

Let us elucidate on this chart a little further. As all the other ensuing plots it has a logarithmic scale on all axes. The error curves plot the error $\|\bar{U}_k^\varepsilon - \bar{u}\|_{L_2(\Omega)}$ against the number of control degrees of freedom (DOFs), the estimator curves do the same for our estimators. The optimal order curve represents the theoretically best possible approximation which can be achieved by a piecewise constant approximation: If $\dim(\mathbb{U}_k)$ is the number of control DOFs it is given by $(\dim(\mathbb{U}_k))^{-1/2}$

(see e.g. Theorem 2.3, [16] and [11]).
To make a valid comparison we also adapted $\varepsilon$ in the uniform refinement case by setting it manually to match the $\varepsilon$ for the adaptive case at the same amount of DOFs. We see that eventually the estimator captures the error in the adaptive and uniform case quite well, this is of course the desired effect in this 'worst-case'-setting. In addition adaptive refinement significantly outperforms uniform refinement. Our $\varepsilon$-adaptation leads to a continuous decrease in the regularisation parameter which is to be expected in this setting.

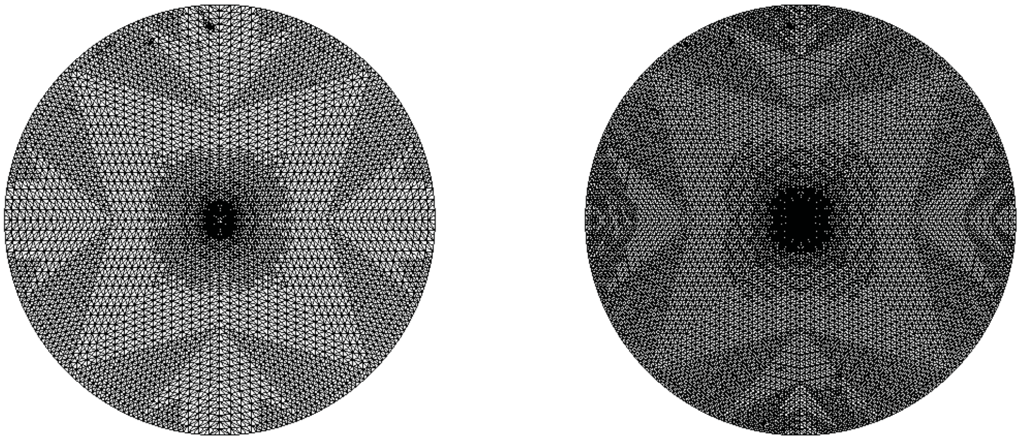

FIGURE 2. adaptive refinement: high resolution at the Dirac source

6.3.2. *A Smooth Example.* With this example we intend to test our estimator and our $\varepsilon$-adaptation in a case

- which is very much not a 'worst case' setting, because here, we will have a multiplier for the state constraint which is a regular $L_\infty(\Omega)$-function
- where the regularisation error does not have a significant influence on the total error

The following data is given: $\Omega = [0,1]^2$, $a = -5$, $b = 5$; with $x = (x_1, x_2)$ and $y_d$ and $f$ properly adjusted the following functions solve (6.5):

$$\begin{aligned}
\bar{y}(x) &= \sin(\pi x_1)\sin(\pi x_2) \\
\bar{p}(x) &= 10\sin(\pi x_1)\sin(\pi x_2) \\
\bar{u}(x) &= \min(\max(-\frac{1}{\nu}\bar{p}(x), a), b) \\
y_c(x) &= \begin{cases} \bar{y}(x) & \text{if } x \in [\frac{1}{4}, \frac{3}{4}]^2 \\ \bar{y}(x) - (x_1 - \frac{1}{4})^2(x_1 - \frac{3}{4})^2(x_2 - \frac{1}{4})^2(x_2 - \frac{3}{4})^2 & \text{else} \end{cases} \\
\bar{\theta}(x) &= \begin{cases} 10 & \text{if } x \in [\frac{1}{4}, \frac{3}{4}]^2 \\ 0 & \text{else} \end{cases}
\end{aligned}$$

We observe that the control is adjusted to be constant on a subdomain of $\Omega$, besides, the state constraint is chosen in such a way that the bi-active set, i.e. the set where both control and state constraint are active, does not have measure zero. This can be numerically disadvantageous. Nevertheless, we see that our adaptive algorithm performs quite well as the following figure shows.

Once again, we employed a logarithmic scale for both axes, the following parameters were chosen: For the estimator $N = 5$, $p' = 12$, for the refinement strategy $\eta_r = 0.5$, $\eta_s = 0.8$ and $\varepsilon$ was fixed to $\varepsilon = 0.01$:

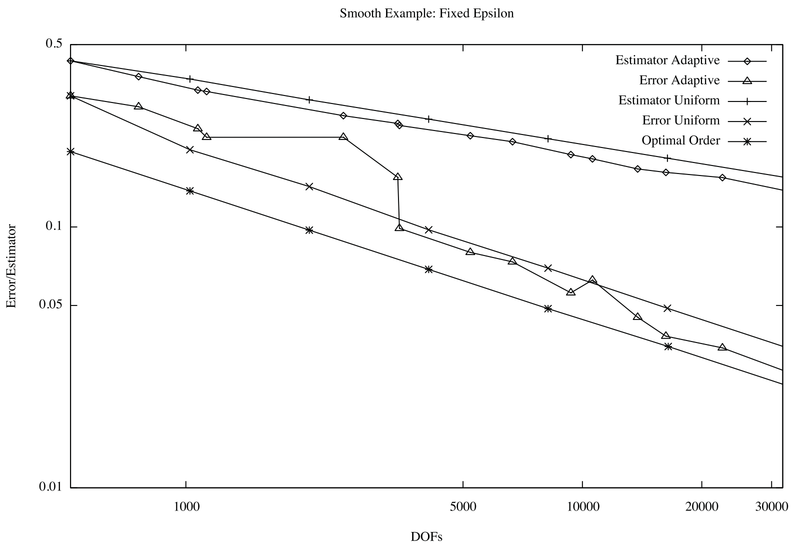


FIGURE 3. smooth example: performance of adaptive and uniform refinement

The curves reflect those of Figure 6.3.1.
We realise that after a slight initial 'wobble' the adaptive refinement strategy outperforms the uniform one in the end needing around $10,000$ fewer DOFs for the same precision. Unfortunately, the estimator curves betray a slight overestimation of the actual error, the given data is too benign for a reliable estimator which necessarily has to take into account the 'worst-case' setting of our first example.
We can also discern that the mesh is fairly coarse in those areas where the control is constant as the following images show. Naturally, this is the desired adaptive mesh refinement:

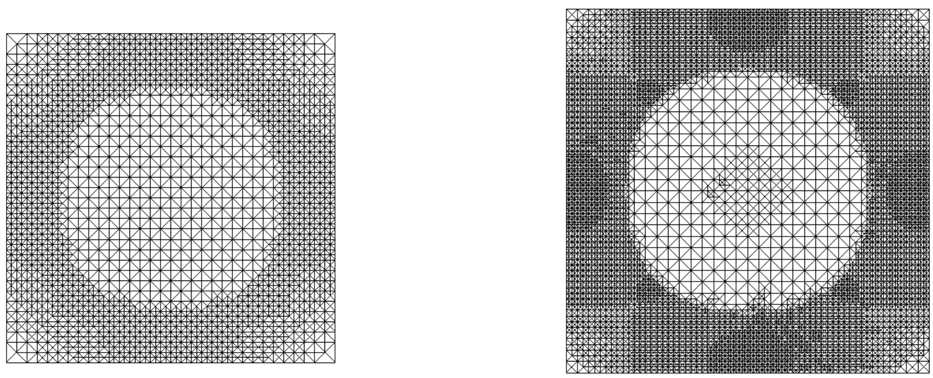

FIGURE 4. adaptive refinement: coarse meshes where $\bar{u}$ is constant

In this example, the error due to regularisation does not have any significant influence on the total error as the following table for uniformly refined meshes shows:

Error $\|\bar{u} - \bar{U}_k^\varepsilon\|$ for $\varepsilon = 0.04$

| DOFs | Error |
|---|---|
| 8192 | 0.069501 |
| 16384 | 0.050021 |
| 32322 | 0.035669 |

Error $\|\bar{u} - \bar{U}_k^\varepsilon\|$ for $\varepsilon = 0.01$

| DOFs | Error |
|---|---|
| 8192 | 0.069519 |
| 16384 | 0.0485 |
| 32768 | 0.034497 |

Error $\|\bar{u} - \bar{U}_k^\varepsilon\|$ for $\varepsilon = 0.003$

| DOFs | Error |
|---|---|
| 8192 | 0.069937 |
| 16384 | 0.048821 |
| 32768 | 0.034621 |

FIGURE 5. results for different regularisation parameters

Naturally, we are now interested in the question whether our $\varepsilon$-adaptation strategy with parameters $\beta = 0.93$ and $c_a = 6.5$, (6.4), observes this lack of influence of the regularisation error. As the following figure shows, it does:

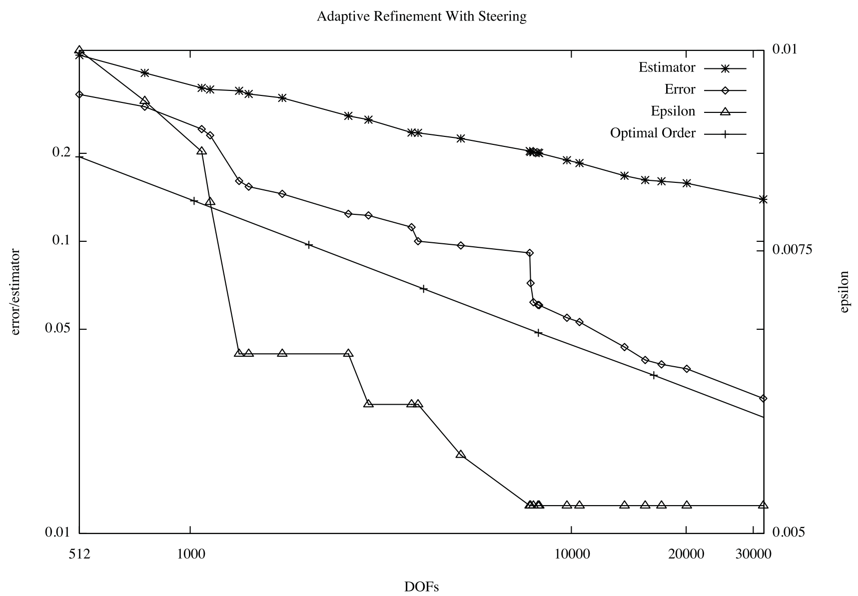


FIGURE 6. smooth example: steering of $\varepsilon$ coupled with adaptive refinement

We discern that after several iterates the regularisation parameter $\varepsilon$ stays constant. Thus, our algorithm (6.4) recognises the fact that for this problem the error incurred by regularisation does not have a significant influence on the overall error and hence keeps the parameter $\varepsilon$ unchanged. Besides, we can once again see that the adaptive algorithm achieves the optimal rate expected for a piecewise constant discretisation, while our estimator slightly overestimates the error, as previously experienced in this smooth setting.

## 7. Conclusion & Outlook

In this paper we have derived a reliable a posteriori error estimator for a state-constrained elliptic optimal control problem. Extensions to more general elliptic problems are easily possible, for the sake of brevity we focussed on the model problem $(P)$. Numerically, we have seen that the estimator captures the true error quite well in the 'worst-case' Dirac setting of Section 6.3.1. This 'worst-case' setting, however, is definitely not infrequent in practical applications. Yet, there is an over-estimation in the smooth problem of Section (6.3.2). This disparity is testament to the fact that in problems of type $(P)$ an estimator that is both reliable and efficient seems to be out of reach at the moment.

In addition, we developed a strategy for the adaptation of the regularisation parameter in the numerical solution of state-constrained optimal control problems, (6.4), which works well and which, in addition, can be swiftly transferred to other regularisation methods such as the Moreau-Yosida approach, because in essence $\bar{V}_k^\varepsilon \approx -\frac{1}{\varepsilon}(\bar{Y}_k^\varepsilon - I_k y_c)^-$.

In the imminent future, our a posteriori error estimator will be supplemented by a basic convergence result in the vein of [41] and a convergence result for the estimator itself in greater generality than already described in [43], Section 4.3. Taken together, both will then constitute a convergent adaptive finite element scheme for state-constrained optimal control problems extending already existing results for adaptive finite element discretisations for 'pure' PDEs and control-constrained problems.

Besides, extending these results to gradient-constrained problems or instationary ones should be another focus of research.

## References


[1] M. Ainsworth and J.T. Oden. *A Posteriori Error Estimation in Finite Element Analysis*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2000.

[2] I. Babuška and W. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.*, 15:736–754, 1978.

[3] I. Babuška, T. Strouboulis, and J.R. Whiteman. *Finite Elements: An Introduction to the Method and Error Estimation*. Oxford University Press, 2011.

[4] W. Bangerth and R. Rannacher. *Adaptive Finite Element Methods for Partial Differential Equations*. Number 1 in Lectures in Mathematics. ETH Zürich. Birkhäuser, 2003.

[5] S. Bartels and C. Carstensen. Each averaging technique yields reliable a posteriori control in FEM on unstructured grids. Part I: Low-order conforming, non-conforming and mixed FEM. *Math. Comp.*, 71:945–969, 2002.

[6] R. Becker, H. Kapp, and R. Rannacher. Adaptive finite element methods for optimal control of partial differential equations: Basic concept. *SIAM Journal Control and Optimization*, 39(1):113–132, 2000.

[7] O. Benedix and B. Vexler. A posteriori error estimation and adaptivity for elliptic optimal control problems with state constraints. *Computational Optimization and Applications*, 44(1):3–25, 2009.

[8] J. Bergh and J. Löfström. *Interpolation Spaces*. Springer, Berlin, 1976.

[9] M. Bergounioux, M. Haddou, M. Hintermüller, and K. Kunisch. A comparison of a Moreau-Yosida-based active set strategy and interior point methods for constrained optimal control problems. *SIAM Journal on Optimization*, 11(2):495–521, 2000.

[10] M. Bergounioux, K. Ito, and K. Kunisch. Primal-dual strategy for constrained optimal control problems. *SIAM J. Control Optim.*, 37(4):1176–1194, 1999.

[11] A. Bonito and R. Nochetto. Quasi-optimal convergence rate of an adaptive discontinuous Galerkin method. *SIAM J. Numer. Anal.*, 48:734–771, 2010.

[12] E. Casas. Control of an elliptic problem with pointwise state constraints. *SIAM J. Control and Optimization*, 4:1309–1322, 1986.

[13] E. Casas, M. Mateos, and B. Vexler. New regularity results and improved error estimates for optimal control problems with state constraints. *ESAIM COCV*, 3(20):803–822, 2014.

[14] S. Cherednichenko, K. Krumbiegel, and A. Rösch. Error estimates for the Lavrentiev regularization of elliptic optimal control problems. *Inverse Problems*, 24(6), 2008.

[15] P.G. Ciarlet. *The Finite Element Method for Elliptic Problems*. SIAM Classics In Applied Mathematics, Philadelphia, 2002.

[16] F. Gaspoz and P. Morin. Convergence rates for adaptive finite elements. *IMA J. Numer. Anal.*, 29(4):917–936, 2009.

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

[18] M. Hintermüller and M. Hinze. Moreau-Yosida regularization in state constrained elliptic control problems: Error estimates and parameter adjustment. *SIAM J. Numer. Anal.*, 47(3):1666–1683, 2009.

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

[20] M. Hintermüller and R.H.W. Hoppe. Goal-oriented mesh adaptivity for mixed control-state constrained elliptic optimal control problems. In W. Fitzgibbon, Y.A. Kuznetsov, P. Neittaanmäki, J. Périaux, and O. Pironneau, editors, *Applied and Numerical Partial Differential Equations*, volume 15 of *Computational Methods in Applied Sciences*, pages 97–111. Springer Netherlands, 2010.

[21] M. Hintermüller, R.H.W. Hoppe, Y. Iliash, and M. Kieweg. An a posteriori error analysis of adaptive finite element methods for distributed elliptic optimal control problems with control constraints. *ESAIM COCV*, 14:865–888, 2008.

[22] M. Hinze. A variational discretization concept in control constrained optimization: The linear-quadratic case. *Computational Optimization and Applications*, 30(1):45–61, 2005.

[23] M. Hinze and C. Meyer. Variational discretization of Lavrentiev-regularized state constrained elliptic control problems. *Comp. Optim. Appl.*, 46:487–510, 2010.

[24] M. Hinze and A. Schiela. Discretization of interior point methods for state constrained elliptic optimal control problems: optimal error estimates and parameter adjustment. *Computional Optimization and Applications*, 48(3):581–600, 2011.

[25] R.H.W. Hoppe and M. Kieweg. Adaptive finite element methods for mixed control-state constrained optimal control problems for elliptic boundary value problems. *Computational Optimization and Applications*, 46(3):511–533, 2010.

[26] K. Kohls, A. Rösch, and K. Siebert. A posteriori error analysis of optimal control problems with control constraints. *SIAM Journal on Control and Optimization*, 52(3):1832–1861, 2014.

[27] I. Kossacký. A recursive approach to local mesh refinement in two and three dimensions. *J. Comput. Appl. Math.*, 55:275–288, 1994.

[28] K. Krumbiegel and A. Rösch. On the regularization error of state constrained Neumann control problems. *Control and Cybernetics*, 37(2):369–392, 2008.

[29] K. Krumbiegel and A. Rösch. A virtual control concept for state constrained optimal control problems. *Computational Optimization and Applications*, 43(2):213–233, 2012.

[30] K. Kunisch and A. Rösch. Primal-dual active set strategy for a general class of constrained optimal control problems. *SIAM Journal Optimization*, 13(2):321–334, 2002.

[31] S. Kurcyusz and J. Zowe. Regularity and stability for the mathematical programming problem in Banach spaces. *Applied Mathematics and Optimization*, 5(1):49–62, 1979.

[32] R. Li, W. Liu, and N. Yan. A posteriori error estimates of recovery type for distributed convex optimal control problems. *J. Sci. Comput.*, 33:155–182, 2007.

[33] W. Liu and N. Yan. A posteriori error estimates for distributed convex optimal control problems. *Adv. Comput. Math*, 15(1-4):285–309, 2001.

[34] C. Meyer. Error estimates for the finite-element approximation of an elliptic control problem with pointwise state and control constraints. *Control and Cybernetics*, 37:51–85, 2008.

[35] C. Meyer, U. Prüfert, and F. Tröltzsch. On two numerical methods for state-constrained elliptic control problems. *Optimization Methods and Software*, 22(6):871–899, 2007.

[36] P. Morin, K. G. Siebert, and A Veeser. A basic convergence result for conforming adaptive finite elements. *Mathematical Models and Methods in Applied Sciences*, 18(5):707–737, 2008.

[37] R. H. Nochetto, K. G. Siebert, and A. Veeser. Theory of adaptive finite element methods: An introduction. In Ronald DeVore and Angela Kunoth, editors, *Multiscale, Nonlinear and Adaptive Approximation*, pages 409–542. Springer Berlin Heidelberg, 2009.

[38] A. Rösch and D. Wachsmuth. A posteriori error estimates for optimal control problems with state and control constraints. *Numerische Mathematik*, 120(4):733–762, 2012.

[39] A. Schiela. Barrier methods for optimal control problems with state constraints. *SIAM J. Optim*, 20(2):1002–1031, 2009.

[40] A. Schmidt and K.G. Siebert. *Design of Adaptive Finite Element Software. The Finite Element Toolbox ALBERTA*, volume 42. Springer, 2005.

[41] K.G Siebert. A convergence proof for adaptive finite elements without lower bounds. *IMA Journal of Numerical Analysis*, 31(3):947–970, 2011.

[42] G. Stampacchia. Le problème de Dirichlet pour les équations elliptiques du second order à coefficients discontinus. *Annales de l'Institut Fourier*, 15(1):189–257, 1965.

[43] S. Steinig. *Adaptive Finite Elements for State-Constrained Optimal Control Problems - Convergence Analysis and A Posteriori Error Estimation*. PhD thesis, Universität Stuttgart, 2014.

[44] A Veeser and R. Verfürth. Explicit upper bounds for dual norms of residuals. *SIAM J. Numer. Anal.*, 47(3):2387–2405, 2009.

[45] R. Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Advanced Numerical Mathematics. John Wiley, Chichester, UK, 1996.

[46] B. Vexler and W. Wollner. Adaptive finite elements for elliptic optimization problems with control constraints. *SIAM J. Contr. Opt.*, 47:509–534, 2008.

[47] W. Wollner. A posteriori error estimates for a finite element discretization of interior point methods for an elliptic optimization problem with state constraints. *Computational Optimization and Applications*, 47(1):133–159, 2010.