On the SQH scheme to solve non-smooth PDE optimal control problems

Tim Breitenbach* Alfio Borzì†

Abstract

A sequential quadratic Hamiltonian (SQH) scheme for solving different classes of non-smooth and non-convex PDE optimal control problems is investigated considering seven different benchmark problems with increasing difficulty. These problems include linear and nonlinear PDEs with linear and bilinear control mechanisms, non-convex and discontinuous costs of the controls, L^1 tracking terms, and the case of state constraints.

The SQH method is based on the characterisation of optimality of PDE optimal control problems by the Pontryagin's maximum principle (PMP). For each problem, a theoretical discussion of the PMP optimality condition is given and results of numerical experiments are presented that demonstrate the large range of applicability of the SQH scheme.

This is a preprint of the paper

Tim Breitenbach and Alfio Borzì,
On the SQH scheme to solve non-smooth PDE optimal control problems
appeared in Journal of Numerical Functional Analysis and Optimization
https://www.tandfonline.com/doi/full/10.1080/01630563.2019.1599911

MSC2010: 35J25, 35K10, 49K20, 49M05, 65K05.

1 Introduction

The characterisation of solutions to optimal control problems with the Pontryagin's maximum principle (PMP) is a fundamental technique in both the ordinary differential equation (ODE) [17, 21, 22] and the partial differential equation (PDE) frameworks [14, 27, 38, 37, 44, 45]. However, once necessary optimality conditions are formulated according to the PMP, a numerical strategy for computing optimal controls is required, and the purpose of this paper is to discuss and validate a recently proposed PMP-based sequential quadratic Hamiltonian (SQH) scheme for solving non-smooth optimal control problems governed by partial differential equations (PDEs); see [12]. For this purpose, we present the results of a key study with 7 different challenging benchmarks that cannot be solved by other known methods that involve the use of (semi-smooth) gradients unless some sort of regularisation is employed.

In fact, one of our aims is to demonstrate that the present PMP - SQH framework seems appropriate to address the open topic of PDE optimization concerning the case of non-convex and discontinuous

*Institut für Mathematik, Universität Würzburg, Emil-Fischer-Strasse 30, 97074 Würzburg, Germany; tim.breitenbach@mathematik.uni-wuerzburg.de

†Institut für Mathematik, Universität Würzburg, Emil-Fischer-Strasse 30, 97074 Würzburg, Germany; alfio.borzi@mathematik.uni-wuerzburg.de

cost functionals. This discussion is also motivated by research work on discontinuous functionals that have real applications; see, e.g., [6, 29, 39], for which the problem of efficient numerical computation of solutions is less investigated and very challenging. In this respect, the purpose of our case study is to provide a contribution towards the realization of a general computational framework for solving PDE optimization problems with non-convex and discontinuous functionals.

In the past, the research effort towards the numerical realisation of the PMP has focused on optimal control problems governed by ODEs. For this purpose, well-known direct and indirect methods have been investigated; see [13] for a survey. However, the application of these schemes appear difficult in the higher space-time dimensional setting of PDE-control problems and in the case of large-size ODE problems. On the other hand, we would like to demonstrate that the iterative SQH scheme, thanks to its pointwise update strategy, proves effective in solving these problems. We remark that the less recent results in [10, 20, 40, 42, 46] and further in [30, 31, 33, 34, 41] play a major role in the formulation of the SQH scheme. Specifically, the approach of [40, 42], where a quadratic penalisation is added to the PMP Hamiltonian of ODE control problems, allows to design robust schemes with proved convergence under appropriate regularity conditions [10]. Unfortunately, this approach requires to update the state variable after each local control update that makes it infeasible for PDE control problems. On the other hand, in [30, 31] a pointwise successive iteration scheme for ODE problems is proposed with pointwise updates of the control without simultaneous update of the state variable and without penalisation of the Hamiltonian. In this way, the resulting scheme appears less robust but computationally fast. The SQH method, presented by the Authors of this paper in [12], combines the advantages of the optimisation schemes proposed in [40, 42] and in [30, 31]. Further, in [12] it is demonstrated that the SQH scheme is able to solve discontinuous and non-convex linear parabolic optimal control problems with distributed control. Moreover, in this reference, convergence of the SQH method for the linear control problem with a non-convex discontinuous cost functional is shown.

These encouraging results provide a motivation for the present work, where the SQH scheme is applied to solve seven different benchmark problems with increasing difficulty including linear, nonlinear, and non-smooth PDEs, linear and bilinear control mechanisms, non-smooth and discontinuous costs of the controls, L^1 tracking terms and the case of state constraints. In this way, we demonstrate the applicability of the SQH method to different classes of PDE optimal control problems with a setting that prevents the use of other (sub-)gradient-based optimisation schemes. Along with this case study, we also present new theoretical results concerning the convergence properties of the SQH scheme and provide for each problem a theoretical discussion concerning the PMP characterisation of optimality and the application of the SQH method. In particular, in Theorem 6 we prove that the limit of a converging subsequence within the iterates of the SQH method fulfils the PMP conditions. Further, in Theorem 8 we give sufficient conditions of optimality for the PMP framework.

In order to make the new SQH computational tool accessible to the scientific computing community, we provide a MATLAB implementation of the SQH scheme for each problem, giving a download link to the corresponding code.

In the following, we discuss the PMP characterisation and the SQH solution of the following PDE optimal control problems, where y denotes the state variable and u the control function. We have

P.1) **Linear elliptic problem**
$\min_{y,u} J(y,u) \coloneqq \int_\Omega \frac{1}{2}\left(y(x) - y_d(x)\right)^2 + g_1(u(x))\,dx$, subject to $(\nabla y, \nabla v) = (u, v)$, $u \in U_{ad}$;

P.2) **Linear parabolic problem**
$\min_{y,u} J(y,u) \coloneqq \int_Q \frac{1}{2}\left(y(x,t) - y_d(x,t)\right)^2 + g_1(u(x,t))\,dxdt$,
subject to $\left(y'(\cdot,t), v\right) + D\left(\nabla y(\cdot,t), \nabla v\right) = \left(u(\cdot,t), v\right)$, $y(\cdot,0) = y_0$, $u \in U_{ad}$;

P.3) **Bilinear elliptic problem**

$\min_{y,u} J(y,u) := \int_\Omega \frac{1}{2}(y(x) - y_d(x))^2 + g_1(u(x))\,dx$, subject to $(\nabla y, \nabla v) + (uy, v) = \left(\tilde{f}, v\right)$, $u \in U_{ad}$;

P.4) **Bilinear parabolic problem**
$\min_{y,u} J(y,u) := \int_Q \frac{1}{2}(y(x,t) - y_d(x,t))^2 + g_1(u(x,t))\,dxdt$,
subject to $(y'(\cdot,t), v) + D(\nabla y(\cdot,t), \nabla v) + (u(\cdot,t)\,y(\cdot,t), v) = (f(\cdot,t), v)$, $y(\cdot,0) = y_0$, $u \in U_{ad}$;

P.5) **Semilinear elliptic problem**
$\min_{y,u} J(y,u) := \int_\Omega \frac{1}{2}(y(x) - y_d(x))^2 + g_1(u(x))\,dx$, subject to $(\nabla y, \nabla v) + (y^3, v) = (u, v)$, $u \in U_{ad}$;

P.6) **$\mathbf{L^1}$-tracking term problem**
$\min_{y,u} J(y,u) := \int_\Omega |y(x) - y_d(x)| + g_2(u(x))\,dx$, subject to $(\nabla y, \nabla v) + (\max(y,0), v) = (u, v)$, $u \in U_{ad}$;

P.7) **State-constrained problem**
$\min_{y,u} J(y,u) := \int_\Omega \frac{1}{2}(y(x) - y_d(x))^2 + g_1(u(x))\,dx$, subject to $(\nabla y, \nabla v) = (u, v)$, $y \le \xi$, $u \in U_{ad}$.

In all cases, we choose homogeneous Dirichlet boundary conditions. In the cost functionals above, we make the following choice

$$g_1(z) := \begin{cases} |z| & \text{if } |z| > s \\ 0 & \text{else} \end{cases}, s > 0, \qquad \text{and} \qquad g_2(z) := \ln(1 + |z|), \tag{1}$$

This choice is motivated by our purpose to demonstrate the applicability of the SQH method to different non-smooth and non-convex problems.

In the following section, we discuss a PMP framework for PDE control problems. Further, we establish the notation and state our assumptions that are required for our theoretical discussion. In Section 3, we illustrate the SQH scheme and present new theoretical results concerning the convergence behaviour of the SQH scheme. These results extend those in [12]. The main result of this section is Theorem 6 that states convergence of the SQH iterates to a PMP optimal solution under some (restrictive) assumptions. Also in this section, with Theorem 8 we discuss a sufficient PMP optimality condition. Examples are given to illustrate the theoretical statements.

It is clear that the present framework should successfully apply to semi-smooth PDE control problems for which (sub-)gradient techniques are available. For this reason, in Section 4, we discuss the PMP-SQH framework for $L^2 - L^1$ optimization problems, where existence of optimal solutions and their characterization by the PMP optimality condition presents no difficulties. Moreover, in these cases we can prove that the SQH method provides a sequence of iterates converging to a PMP optimal solution that also satisfies the optimality conditions derived in the Lagrange framework.

In Section 5, we present results of numerical experiments for the different control problems P.1) to P.7) that validate the theoretical results and the computational performance of the SQH scheme. In these experiments, our main concern is to show that by choosing a more stringent tolerance in the SQH stopping criterion, the fulfilment of the PMP optimality condition improves. Moreover, we present the convergence history of SQH scheme showing that, in all cases, the values of the cost functional are monotonically reducing along the SQH iterations.

In Section 5.1, we consider the linear elliptic control problem P.1) and the parabolic control problem P.2), both with a discontinuous cost functional. In Section 5.2, we extend the discussion of Section 5.1 to the bilinear control problems P.3) and P.4). In Section 5.3, we present results with P.5), that is, an

optimal control problem with distributed control and a semi-linear elliptic model with cubic nonlinearity. Notice that, also in this case, we are able to prove the applicability of the PMP optimality condition.

In Section 5.4, we have the case P.6) of an elliptic PDE with a Lipschitz continuous term and distributed control and a cost functional with L^1 tracking term. In this case, we are not able to prove the PMP characterisation of optimality for this problem, and we explain where this difficulty arises. Nevertheless, it is possible to apply our SQH scheme also in this case and we obtain very encouraging results. In Section 5.5, we demonstrate the applicability of the SQH method to the state-constrained optimal control problem P.7). In this case, we consider a modification of this problem by an augmented Lagrangian technique, and discuss the PMP characterisation of optimality of the modified problem. Since the augmentation step consists in adding a weighted term to the cost functional, we prove Theorem 11 that states how the PMP solution to the augmented problem is related to the solution of P.7) as the augmentation weight increases.

To improve readability of this paper, we have postponed some theoretical discussion to the Appendix. In particular, notice that the results in Appendix B are essential in order to prove the PMP characterisation of optimality. However, in order to keep the exposition of our work to a reasonable size, we use a compact notation to discuss, in a unified way, the cases of space and space-time dependent PDEs.

A section of conclusion completes the exposition of our work.

2 General formulation of PDE optimal control problems

In this section, we formulate classes of elliptic and parabolic optimal control problems, and in order to distinguish between the elliptic and the parabolic case, we introduce the index $i \in \{\mathrm{e}, \mathrm{p}\}$, where e refers to the elliptic case and p to the parabolic case. We denote $Z_{\mathrm{e}} \coloneqq \Omega \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$ and $Z_{\mathrm{p}} \coloneqq \Omega \times (0,T)$, $T > 0$ where Ω is an open and bounded domain. In the parabolic case, we assume that Ω has a smooth boundary. Further, we define a vector of controls $u \coloneqq (u_1, ..., u_m)$, where each u_j, $j \in \{1, ..., m\}$, $m \in \mathbb{N}$, is an element of the following admissible set of controls

$$U_{ad}^j = \left\{ u_j \in L^q(Z_i) \,|\, u_j(z) \in K_U^j \text{ a.e. in } Z_i \right\},$$

with $K_U^j \subseteq \mathbb{R}$ compact, $q = 2$ for $n = 1$, $q \geq \frac{n}{2}+1$ for $n \geq 2$ and $U_{ad} \coloneqq U_{ad}^1 \times ... \times U_{ad}^m$, $K_U \coloneqq K_U^1 \times ... \times K_U^m$.

Next, we formulate our PDE constraint (governing model) in weak form as in [25, page 296] for the elliptic case and as in [25, page 351/352] for the parabolic case.

Consider the bilinear form for the parabolic case as follows: $B : \mathcal{H} \times \mathcal{H} \times [0,T] \to \mathbb{R}$, $(y, v; t) \mapsto B(y, v; t)$ where the function space $\mathcal{H}$, a set of functions mapping Z_i to $\mathbb{R}$, has to be chosen accordingly. Then, the weak formulation of a parabolic equation is given by

$$(y'(\cdot, t), v) + B(y, v; t) = \int_\Omega f(x, t, y, u)\, v(x)\, dx, \tag{2}$$

for almost every $t \in [0,T]$ and all $v \in \mathcal{H}$ where $(\cdot, \cdot)$ is the $L^2(\Omega)$ scalar product, $y' \coloneqq \frac{\partial}{\partial t} y$ and $f : \mathbb{R}^n \times \mathbb{R}_0^+ \times \mathbb{R} \times K_U \to \mathbb{R}$. Notice that $f(z, y, u) \coloneqq f(z, y(z), u(z))$ whenever an argument of f is a function instead of a number. We implicitly assume for the rest of this work that if functions do not depend on time, we refer to the elliptic case and if we consider an elliptic equation, then functions do to not depend on time t. This implies in the elliptic case that $(y'(\cdot, t), v) = 0$ for all $v \in \mathcal{H}$. We require that (2) is well defined and that there is a unique solution $y : Z_i \to \mathbb{R}$, $z \mapsto y(z)$ to (2), that means that y fulfils (2) for almost all $t \in [0,T]$ and all $v \in \mathcal{H}$.

Now, in view of this requirement, we consider our control problems P.1) to P.7) and define $\mathcal{H} \coloneqq H_0^1(\Omega)$ and $y_0 \in H_0^1(\Omega) \cap L^\infty(\Omega)$ in the parabolic case. For the governing model of P.1), we have the bilinear form $B_1(y, v) \coloneqq (\nabla y, \nabla v)$ and the right-hand side $f_1 = u$ where the constraint of P.1) has

a unique solution $y \in H_0^1(\Omega)$, see [25, 6.2 Theorem 1] and [1, Theorem 2.14]. For P.2), we have the bilinear form $B_2(y,v;t) \coloneqq D(\nabla y(\cdot,t),v)$ and the right-hand side $f_2 = u$ where the model of P.2) has a unique solution $y \in L^2(0,T,H^2(\Omega)) \cap L^\infty(0,T;H_0^1(\Omega))$, see [25, 7.1 Theorem 5]. Analogously for P.3) with $B_3 \coloneqq B_1$ and $f_3 \coloneqq \tilde{f} - uy$ and for P.4) with $B_4 \coloneqq B_2$ and $f_4 \coloneqq f_3$, we have a unique solution for the corresponding constraint. In P.5), the constraint for $B_5 \coloneqq B_1$ and $f_5 \coloneqq u - y^3$ has a unique solution $y \in H_0^1(\Omega)$, see [15]. The constraint of P.6) for $B_6 \coloneqq B_1$ and $f_6 \coloneqq u - \max(y,0)$ has a unique solution $y \in H_0^1(\Omega)$, see [16]. The case P.7) is analogous to P.1) where we additionally require that $y \leq \xi$, $\xi \in \mathbb{R}$.

In the next step, we formulate our general optimal control problem as follows

$$\begin{aligned} &\min_{y,u} J(y,u) \coloneqq \int_{Z_i} h(y(z)) + g(u(z))\,dz \\ &\text{s.t.} \int_\Omega y'(x,t)\,v(x)\,dx + B(y,v;t) = \int_\Omega f(x,t,y(x,t),u(x,t))\,v(x)\,dx \\ &u \in U_{ad} \end{aligned} \tag{3}$$

where $g(u(z)) \coloneqq \sum_{j=1}^m g_j(u_j(z))$, $z \coloneqq (x,t)$ for the parabolic case and $z \coloneqq x$ for the elliptic case. We assume J to be bounded from below and require that (3) is well posed. In particular, we assume $g_j : \mathbb{R} \to \mathbb{R}$, $z \mapsto g_j(z)$, $j \in \{1,...,m\}$ to be bounded from below and lower semi-continuous. This includes the case of the so called L^0-norm which is obtained with $g_{L^0} \coloneqq \begin{cases} 1 & \text{if } z \neq 0 \\ 0 & \text{else} \end{cases}$.

Remark 1. We have that the map $u \mapsto G(u) : U_{ad} \to \mathbb{R}$ with $G(u) \coloneqq \int_{z_i} g(u(z))\,dz$ is lower semi-continuous on U_{ad}. This means that for any sequence $(u_k)_{k\in\mathbb{N}}$ with $\lim_{k\to\infty} \|u_k - \bar{u}\|_{L^2(Z_i)} = 0$ we have that $\liminf_{k\to\infty} G(u_k) \geq G(\bar{u})$. If this was not the case, then there would be an index set $K_1 \subseteq \mathbb{N}$ with $\lim_{K_1 \ni k\to\infty} G(u_k) < G(\bar{u})$. Due to the boundedness of G from below by definition and from above by $G(\bar{u})$, there exists an index set $K_2 \subseteq K_1$ where $(G(u_k))_{k\in K_2}$ is converging, see [3, II Theorem 5.8]. Since $(u_k)_{k\in K_2}$ is converging to $\bar{u}$, we have the existence of another index set $K_3 \subseteq K_2$ with $\lim_{K_3 \ni k\to\infty} u_k(z) = \bar{u}(z)$ almost everywhere on Z_i, see [7, Proposition 3.6,Remark 3.7]. Now, similar to the proof of [12, Theorem A.2], by [3, II Theorem 1.15] and Lemma of Fatou [7, Lemma 2.15], we obtain the following contradiction

$$G(\bar{u}) > \lim_{K_1 \ni k\to\infty} G(u_k) = \lim_{K_3 \ni k\to\infty} G(u_k) \geq \liminf_{K_3 \ni k\to\infty} G(u_k) \geq G(\bar{u}).$$

This means that the smallest accumulation point of $(u_k)_{k\in\mathbb{N}}$ is greater or equal than $G(\bar{u})$ which proves that $\liminf_{\mathbb{N} \ni k\to\infty} G(u_k) \geq G(\bar{u})$, see [3, Theorem 5.5]. Assuming a continuous control-to-state map $S : U_{ad} \to L^2(Q)$ and a Lipschitz continuous h, we also have that the reduced cost functional $\hat{J}(u) \coloneqq J(S(u),u)$ is lower semi-continuous on U_{ad}, see the proof of [12, Theorem A.2].

Notice that we assume existence of optimal solutions to (3), denoted with $(\bar{y},\bar{u})$, and focus on their characterisation in the PMP framework. However, see Appendix C for a discussion concerning the proof of existence of optimal solutions also in relation to the PMP framework. Further, we remark that, for the problems considered in this work, it is possible to prove existence of suboptimal solutions in the framework of Ekeland's variational principle [26].

Before we define the corresponding PMP necessary optimality conditions that a solution to (3) must fulfil, we introduce the adjoint bilinear form $B^* : \mathcal{H} \times \mathcal{H} \times [0,T] \to \mathbb{R}$, $(p,v,t) \mapsto B^*(p,v;t)$, where we require that

$$B^*(p,v;t) = B(v,p;t),$$

holds for almost every $t \in [0,T]$ and all $v \in \mathcal{H}$. The adjoint equation to (2) is, analogously to [38, Theorem 2.1], defined as follows

$$-(p'(\cdot,t),v) + B^*(p,v;t) = \int_\Omega \left(\frac{\partial}{\partial y} h(y)\,|_{y=y(x,t)} + \frac{\partial}{\partial y} f(x,t,y,u)\,|_{y=y(x,t)} p(x,t) \right) v(x)\,dx, \tag{4}$$

with $p(\cdot, T) = 0$ where y is the solution to (2) and $p' \coloneqq \frac{\partial}{\partial t} p$. We require that there exists a unique solution $p : Z_i \to \mathbb{R}$, $z \mapsto p(z)$ such that (4) holds for almost all $t \in [0, T]$ and all $v \in \mathcal{H}$.

Crucial for the PMP is the Hamiltonian $H : Z_i \times \mathbb{R} \times K_U \times \mathbb{R} \to \mathbb{R}$, $(z, y, u, p) \mapsto H(z, y, u, p)$ that is given by

$$H(z, y, u, p) \coloneqq h(y) + g(u) + pf(z, y, u). \tag{5}$$

Now, we can formulate the necessary optimality conditions given by the PMP. If $\bar{p}$ is the solution to (4) where $\bar{y}$ is inserted for y, we write $y \leftarrow \bar{y}$, and $\bar{u}$ is inserted for u, we write $u \leftarrow \bar{u}$, then we have that

$$H(z, \bar{y}, \bar{u}, \bar{p}) = \min_{w \in K_U} H(z, \bar{y}, w, \bar{p}), \tag{6}$$

for almost all $z \in Z_i$.

Notice that we use the notation $H(z, y, u, p) \coloneqq H(z, y(z), u(z), p(z))$ whenever an argument of H is a function instead of a number.

Similar to [38], we can prove the following theorem.

Theorem 2. *Let $y \mapsto h(y)$ be continuously differentiable, and let $(y, u) \mapsto f(y, u)$ be continuously differentiable, assume that f does not depend explicitly on z. Furthermore, let $|f(0, u)| \leq \tilde{c}|u|$, $\tilde{c} > 0$, and $(y, u) \mapsto \frac{\partial}{\partial y} f(y, u)$ be bounded by a constant. Then any solution $(\bar{y}, \bar{u})$ to (3) fulfils (6) for almost any $z \in Z_i$, where $\bar{p}$ solves (4) for $(\bar{y}, \bar{u})$ instead of (y, u).*

According to Theorem 2, optimal solutions to our problems P.1) to P.5) can be characterised by the PMP. On the other hand, this is not the case for P.6), and in the case P.7) additional work is required.

To conclude our general discussion concerning optimization PMP and lower semi-continuous functionals, we would like to point out that with the help of Ekeland's variational principle, see [36, Theorem 55], it is possible to define a weaker concept of a minimizer that is naturally accommodated in the PMP framework. To illustrate this fact, we endow U_{ad} with the metric $\delta_0(u_1, u_2) \coloneqq |\{t|\ u_1(z) \neq u_2(z)\}|$, giving the measure of the set where $u_1 \in U_{ad}$ differs from $u_2 \in U_{ad}$. The proof of [24, Lemma 7.2] also applies in this setting and provides that (U_{ad}, δ_0) is a complete metric space. We continue with the following remark.

Remark 3. Since $\hat{J}$ is bounded from below, there exists an element $\tilde{u} \in U_{ad}$ within any minimizing sequence such that, for any $\epsilon > 0$, it holds $\hat{J}(\tilde{u}) \leq \inf_{v \in U_{ad}} \hat{J}(v) + \epsilon$, see [3, II Theorem 4.1]. Consequently, by applying Ekeland's variational principle, there exists a $u^* \in U_{ad}$, with $\hat{J}(u^*) \leq \hat{J}(\tilde{u})$, that satisfies $\hat{J}(u^*) < \hat{J}(w) + \epsilon\, \delta_0(w, u^*)$ for all $w \in U_{ad} \backslash \{u^*\}$. In this sense, u^* can be named an ϵ-minimizer for our problem. This point has a natural PMP characterization as follows. Take in the inequality above w equal to the needle variation of u^* at any arbitrary point $z \in Z_i$, denoted it with u_k, see [12] for details. Then we obtain that $\hat{J}(u_k) - \hat{J}(u^*) > -\epsilon \delta_0(u_k, u^*) = -\epsilon |S_k(z)|$, where $|S_k(z)|$ is the measure of the ball centered at $z \in Z_i$, where u_k differs from u^*, and $|S_k|$ converges to zero for $k \to \infty$. Therefore we have $\frac{1}{|S_k(z)|} \left(\hat{J}(u_k) - \hat{J}(u^*) \right) > -\epsilon$. According to [12, Lemma 3.2], the limit for $k \to \infty$ provides

$$H(z, y^*, u^*, p^*) \leq H(z, y^*, u, p^*) + \epsilon,$$

for all $u \in K_U$ and for almost all $z \in Z_i$, see [3, II Theorem 2.7], where y^* is the solution to state equation for u^* and p^* is the solution to the adjoint equation for (y^*, u^*). In fact, the PMP characterization can be proved with Ekeland's variational principle; see, e.g., [32].

In the analysis of the SQH method, we make the following assumptions. For this purpose, we define the set $I \subseteq \mathbb{R}$ as the convex hull [8, Section 3.1] of the union of all images from each solution y to (2) for any $u \in U_{ad}$, given by

$$I \coloneqq \operatorname{conv} \{ y(Z_i) \subseteq \mathbb{R} |\ y \text{ solves (2) for } u \in U_{ad} \}.$$

A.1) The functions $h : I \to \mathbb{R}$, $v \mapsto h(v)$, and $f : I \to \mathbb{R}$, $v \mapsto f(z, v, u)$ are supposed to be twice continuously differentiable for all $u \in K_U$ and for almost all $z \in K_i$.

Furthermore, we require the existence of a constant $c > 0$ such that the following holds

A.2) $\|\delta y\|_{L^2(Z_i)} \leq c\|\delta u\|_{L^2(Z_i)}$, $\|\delta p\|_{L^2(Z_i)} \leq c\|\delta u\|_{L^2(Z_i)}$;

A.3) $\int_0^T |\left((\delta y)'(\cdot, t), \delta p(\cdot, t)\right) + B(\delta y, \delta p; t)|dt \leq c\|\delta u\|^2_{L^2(Z_i)}$;

A.4) $\|p\|_{L^\infty(Z_i)} \leq c$ for any solution (y, u) to (2) for $u \in U_{ad}$;

A.5) $\|\frac{\partial}{\partial y} f(\cdot, y, u)\|_{L^\infty(Z_i)} \leq c$, $\|\frac{\partial^2}{\partial y^2} f(\cdot, y, u)\|_{L^\infty(Z_i)} \leq c$, $\|\frac{\partial^2}{\partial y^2} h(y)\|_{L^\infty(Z_i)} \leq c$ for all solutions (y, u) to (2) with $u \in U_{ad}$;

where $\delta y := y_1 - y_2$, $\delta u := u_1 - u_2$, $\delta p := p_1 - p_2$ and $\|\delta u\|^2_{L^2(Z_i)} := \sum_{j=1}^m \|\delta u_j\|^2_{L^2(Z_i)}$ where (y_ℓ, u_ℓ), $\ell = 1, 2$, are solutions to (2) and (y_ℓ, u_ℓ, p_ℓ) are solutions to (4) . Additionally, we require

A.6) $f : \mathbb{R}^m \to \mathbb{R}$, $(\cdot, \cdot, u) \mapsto f(\cdot, \cdot, u)$ is bounded from below for all $u \in K_U$ and lower semi-continuous on K_U.

We remark that the Assumptions A.1) to A.6) are fulfilled by P.1) to P.5), see below. The cases P.6) and P.7) are not covered by the theory in [38]. Nevertheless, we show how to apply the SQH method to these problems.

3 The sequential quadratic Hamiltonian method

In this section, we illustrate the algorithm that implements the SQH method as in [12] and discuss its convergence to a PMP solution. For this purpose, we define the following augmented Hamiltonian

$$K_\epsilon(z, y, u, v, p) := H(z, y, u, p) + \epsilon(u(z) - v(z))^2, \tag{7}$$

where $K_\epsilon : Z_i \times \mathbb{R} \times K_U \times K_U \times \mathbb{R} \to \mathbb{R}$, $\epsilon > 0$ and $(u(z) - v(z))^2 := \sum_{j=1}^m (u_j(z) - v_j(z))^2$. We use the notation $K_\epsilon(z, y, u, v, p) := K_\epsilon(z, y(z), u(z), v(z), p(z))$ whenever an argument of K_ϵ is a function instead of a number.

The SQH scheme is implemented as follows.

Algorithm 1 (SQH method)

1. Choose $\epsilon > 0$, $\kappa \geq 0$, $\sigma > 1$, $\zeta \in (0,1)$, $\eta \in (0,\infty)$, $u^0 \in U_{ad}$, compute y^0 by (2) for $u \leftarrow u^0$ and p^0 by (4) for $y \leftarrow y^0$ and $u \leftarrow u^0$, set $k \leftarrow 0$

2. Choose $u \in K_U$ such that

$$K_\epsilon\left(z, y^k, u, u^k, p^k\right) \leq K_\epsilon\left(z, y^k, w, u^k, p^k\right)$$

   for all $w \in K_U$ and all $z \in Z_i$.

3. Calculate y by (2) for u and $\tau \coloneqq \|u - u^k\|^2_{L^2(Z_i)}$

4. If $J(y,u) - J\left(y^k, u^k\right) > -\eta\tau$: Choose $\epsilon \leftarrow \sigma\epsilon$
   Else:
   Choose $\epsilon \leftarrow \zeta\epsilon$, $y^{k+1} \leftarrow y$, $u^{k+1} \leftarrow u$, calculate p^{k+1} by (4) for $y \leftarrow y^{k+1}$ and $u \leftarrow u^{k+1}$, set $k \leftarrow k+1$

5. If $\tau < \kappa$: STOP and return u^k
   Else go to 2.

In [12] the Lebesgue measurability of u that is defined by minimizing the augmented Hamiltonian pointwise in Step 2 of Algorithm 1 is discussed, stating that for continuous g, the control u is always Lebesgue measurable. Further, a direct calculation is presented to show that in the case of $g(z) \coloneqq \begin{cases} |z| & \text{if } |z| > s \\ 0 & \text{else} \end{cases}$, $s > 0$, the control function u is measurable. For the purpose of this work, and in order to keep the analysis of the SQH method as general as possible, we assume that u is always measurable.

Next, we state a lemma concerning the minimising properties of the SQH iterates. Specifically, if in one iterate no sufficient decrease of J is achieved, then it is possible to improve descent by choosing a larger ϵ in K_ϵ. Consequently, the following lemma ensures that a cost-functional-reducing update to the control can be found in a finite number of steps. This is also an important fact in the proof of Theorem 5, see the proof of [12, Theorem 4.1] for details. A similar result can be found in [40, 10] and in [12, Lemma 4.2]. However, the proof in [12, Lemma 4.2] is not applicable to the presented general case. Therefore, we give the proof of the following lemma in Appendix A.

Lemma 4. *Let (y,u) and $\left(y^k, u^k\right)$ be generated by Algorithm 1, $k \in \mathbb{N}_0$, and u, u^k be measurable; denote $\delta u \coloneqq u - u^k$. Subject to assumptions A.1) to A.6), there is a $\theta > 0$ independent of ϵ, k, and u^k such that for the $\epsilon > 0$ currently chosen by Algorithm 1, the following holds*

$$J(y,u) - J\left(y^k, u^k\right) \leq -(\epsilon - \theta)\,\|\delta u\|^2_{L^2(Z_i)}.$$

In particular, $J(y,u) - J\left(y^k, u^k\right) \leq 0$ for $\epsilon \geq \theta$.

In the next two following theorems, we investigate the sequences $\left(y^k\right)_{k\in\mathbb{N}}$ and $\left(u^k\right)_{k\in\mathbb{N}}$ generated by the iterated Steps 2 to 4 of Algorithm 1 (no stopping criterion) assuming that u^k is not optimal for all $k \in \mathbb{N}_0$ but Lebesgue measurable. If there was a $\bar{k} \in \mathbb{N}_0$ such that $u^{\bar{k}}$ was optimal in the sense of (6), i.e. $H\left(z, y^{\bar{k}}, u^{\bar{k}}, p^{\bar{k}}\right) = \min_{w\in K_U} H\left(z, y^{\bar{k}}, w, p^{\bar{k}}\right)$ for almost all $z \in Z_i$, then we would have that $u^k = u^{\bar{k}}$ for all $k \geq \bar{k}$ in the $L^2(Z_i)$ sense, see the proof of [12, Lemma 4.3]. The following theorem is proved in [12, Theorem 4.1].

Theorem 5. *Let the sequence* $\left(y^k\right)_{k\in\mathbb{N}_0}$ *and* $\left(u^k\right)_{k\in\mathbb{N}_0}$ *be generated as in Algorithm 1 (loop over Step 2 to Step 4). Then, the sequence of cost functional values* $J\left(y^k,u^k\right)$ *monotonically decreases with*

$$\lim_{k\to\infty}\left(J\left(y^{k+1},u^{k+1}\right)-J\left(y^k,u^k\right)\right)=0,$$

and

$$\lim_{k\to\infty}\|u^{k+1}-u^k\|_{L^2(Z_i)}=0.$$

From Theorem 5, we obtain that Algorithm 1 is well defined for $\kappa>0$. This means that there is an iteration number $\bar{k}\in\mathbb{N}_0$ such that $\|u^{\bar{k}+1}-u^{\bar{k}}\|_{L^2(Z_i)}\leq\kappa$ and consequently Algorithm 1 stops in finitely many steps; see Step 4 in Algorithm 1 and Lemma 4.

Next, we discuss our novel theoretical result that is represented by the following theorem that states the convergence of the SQH method to the PMP solution, characterized by (6), without any differentiability assumptions on the Hamiltonian function with respect to the control argument u. Therefore this result can be applied to optimal control problems with discontinuous cost functionals. However, for our discussion of the convergence properties of the SQH scheme, we have the requirement that for any iterate u^k, $k\in\mathbb{N}_0$ and for any ϵ chosen by Algorithm 1 there exists a $r\geq\epsilon$ such that

$$K_\epsilon\left(z,y^k,u^{k+1},u^k,p^k\right)+r\left(w-u^{k+1}(z)\right)^2\leq K_\epsilon\left(z,y^k,w,u^k,p^k\right),\tag{8}$$

is fulfilled for all $w\in K_U$ and for all $z\in Z_i$. This condition ensures sufficient descent of the cost functional. In Example 7, we verify this condition for a non-smooth L^2-L^1-cost functional.

Theorem 6. *Let Assumptions A.1) to A.6) hold. In addition, let* $f(\cdot,y,u):I\times K_U\to\mathbb{R}$ *be a continuous function for almost all* $z\in Z_i$*, and require that (8) holds. Then for any accumulation point* $\bar{u}$ *in* $L^q(Z_i)$ *of the SQH sequence* $\left(u^k\right)_{k\in\mathbb{N}_0}$*, which has the property* $\lim_{\tilde{k}\to\infty}\|u^{\tilde{k}}-\bar{u}\|_{L^2(Z_i)}=0$*,* $\tilde{k}\in\tilde{K}\subseteq\mathbb{N}_0$*, it holds that there exists a subsequence* $\left(u^{\hat{k}}\right)_{\hat{k}\in\hat{K}}$*,* $\hat{K}\subseteq\tilde{K}$ *such that*

$$\lim_{\hat{k}\to\infty}u^{\hat{k}}(z)=\bar{u}(z),$$

for almost all $z\in Z_i$*, and*

$$H(z,\bar{y},\bar{u},\bar{p})=\min_{w\in K_U}H(z,\bar{y},w,\bar{p}),$$

for almost all $z\in Z_i$ *where* $\bar{y}(z)=\lim_{\hat{k}\to\infty}y^{\hat{k}}(z)$ *is the state corresponding to* $\bar{u}$ *and* $\bar{p}(z)=\lim_{\hat{k}\to\infty}p^{\hat{k}}(z)$ *is the corresponding adjoint variable.*

Furthermore, for almost each $z\in Z_i$ *and any* $\mu>0$*, there exists a* $\bar{k}\in\hat{K}$ *such that*

$$H\left(z,y^{\hat{k}+1},u^{\hat{k}+1},p^{\hat{k}+1}\right)\leq H\left(z,y^{\hat{k}+1},w,p^{\hat{k}+1}\right)+\mu,\tag{9}$$

for all $w\in K_U$ *and for all* $\hat{k}\geq\bar{k}$*.*

Proof. The control u^k is determined by Algorithm 1 such that due to (8) the following holds

$$K_\epsilon\left(z,y^k,u^{k+1},u^k,p^k\right)+r\left(w-u^{k+1}(z)\right)^2\leq K_\epsilon\left(z,y^k,w,u^k,p^k\right),$$

for all $w\in K_U$, for all $k\in\mathbb{N}_0$ and all $z\in Z_i$ which is equivalent to

$$H\left(z,y^k,u^{k+1},p^k\right)+\epsilon\left(u^{k+1}(z)-u^k(z)\right)^2+r\left(w-u^{k+1}(z)\right)^2\leq H\left(z,y^k,w,p^k\right)+\epsilon\left(w-u^k(z)\right)^2.\tag{10}$$

For the accumulation point $\bar{u}$ with $\lim_{\tilde{k}\to\infty}\|u^{\tilde{k}}-\bar{u}\|_{L^2(Z_i)}=0$, $\tilde{k}\in\tilde{K}$, we have that due to the strong $L^2(Z_i)$-convergence there is a subsequence $\left(u^{\tilde{k}_1}\right)_{\tilde{k}_1\in\tilde{K}_1}$, $\tilde{K}_1\subseteq\tilde{K}$ with $\lim_{\tilde{k}_1\to\infty}u^{\tilde{k}_1}(z)=\bar{u}(z)$ for almost all $z\in Z_i$, see [7, Proposition 3.6, Remark 3.7]. Due to our Assumption A.2), the strong $L^2(Z_i)$-convergence of the control implies strong $L^2(Z_i)$-convergence of the state and adjoint variable and thus analogously, there is a subsequence $\left(u^{\hat{k}}\right)_{\hat{k}\in\hat{K}}$, $\hat{K}\subseteq\tilde{K}_1$ with $\lim_{\hat{k}\to\infty}y^{\hat{k}}(z)=\bar{y}(z)$ and $\lim_{\hat{k}\to\infty}p^{\hat{k}}(z)=\bar{p}(z)$ for almost all $z\in Z_i$. Analogously, since we have that for the sequence $\left(u^k\right)_{k\in\mathbb{N}_0}$ (see Theorem 5), it holds that $\lim_{k\to\infty}\|u^{k+1}-u^k\|_{L^2(Z_i)}$, we obtain another subsequence, still denoted with $\left(u^{\hat{k}}\right)_{\hat{k}\in\hat{K}}$, with the following property $\lim_{\hat{k}\to\infty}u^{\hat{k}+1}(z)=\bar{u}(z)$, $\lim_{\hat{k}\to\infty}y^{\hat{k}+1}(z)=\bar{y}(z)$, $\lim_{\hat{k}\to\infty}p^{\hat{k}+1}(z)=\bar{p}(z)$.

Now, we consider again (10) where it also holds due to our assumption $r\geq\epsilon$ that

$$H\left(z,y^k,u^{k+1},p^k\right)+\epsilon\left(u^{k+1}(z)-u^k(z)\right)^2+\epsilon\left(w-u^{k+1}(z)\right)^2\leq H\left(z,y^k,w,p^k\right)+\epsilon\left(w-u^k(z)\right)^2,$$

and thus by inserting

$$\left(w-u^{k+1}(z)\right)^2=\left(w-u^k(z)\right)^2+\left(u^k(z)-u^{k+1}(z)\right)^2+2\left(w-u^k(z)\right)\left(u^k(z)-u^{k+1}(z)\right),$$

we obtain

$$H\left(z,y^k,u^{k+1},p^k\right)+2\epsilon\left(u^{k+1}(z)-u^k(z)\right)^2+2\epsilon\left(w-u^k(z)\right)\left(u^k(z)-u^{k+1}(z)\right)\leq H\left(z,y^k,w,p^k\right). \quad (11)$$

Then (11) is equivalent to

$$\begin{aligned}&h\left(y^k(z)\right)+g\left(u^{k+1}(z)\right)+p^k(z)\,f\left(z,y^k,u^{k+1}\right)+2\epsilon\left(u^{k+1}(z)-u^k(z)\right)^2\\&+2\epsilon\left(w-u^k(z)\right)\left(u^k(z)-u^{k+1}(z)\right)\leq h\left(y^k(z)\right)+g(w)+p^k(z)\,f\left(z,y^k,w\right).\end{aligned} \quad (12)$$

Next, we have that ϵ is bounded from below by 0 and from above by $\sigma(\eta+\theta)$ due to [12, (16)] and Step 4 in Algorithm 1. The boundedness of ϵ guarantees that the corresponding terms go to zero for $\hat{k}$ to infinity, see [3, Theorem 2.4, Theorem 6.1] since $u^{\hat{k}}$ converges pointwise and $\left(w-u^k(z)\right)$ is also bounded as $w,u^k\in K_U$ for all $k\in\mathbb{N}_0$. This connection is exploited in the next step.

Now, as g is lower semi-continuous, we apply the lim inf on both sides of the last inequality and recalling that whenever the lim exists the lim inf equals lim, see [3, Theorem 5.7] and the calculation rules for a sum of lim inf [23, Theorem 3.127], we obtain for the left-hand side of (12) the following

$$\begin{aligned}&\liminf_{\hat{k}\to\infty}\left(h\left(y^{\hat{k}}(z)\right)+g\left(u^{\hat{k}+1}(z)\right)+p^{\hat{k}}(z)\,f\left(z,y^{\hat{k}},u^{\hat{k}+1}\right)+2\epsilon\left(u^{\hat{k}+1}(z)-u^{\hat{k}}(z)\right)^2\right.\\&\left.+2\epsilon\left(w-u^{\hat{k}}(z)\right)\left(u^{\hat{k}}(z)-u^{\hat{k}+1}(z)\right)\right)\geq h\left(\bar{y}(z)\right)+g\left(\bar{u}(z)\right)+\bar{p}(z)\,f(z,\bar{y},\bar{u})=H(z,\bar{y},\bar{u},\bar{p}).\end{aligned}$$

For the right-hand side of (12), we have

$$\begin{aligned}&\liminf_{\hat{k}\to\infty}\left(h\left(y^{\hat{k}}(z)\right)+g(w)+p^{\hat{k}}(z)\,f\left(z,y^{\hat{k}},w\right)\right)=\lim_{\hat{k}\to\infty}\left(h\left(y^{\hat{k}}(z)\right)+g(w)+p^{\hat{k}}(z)\,f\left(z,y^{\hat{k}},w\right)\right)\\&=h\left(\bar{y}(z)\right)+g(w)+\bar{p}(z)\,f(z,\bar{y},w)=H(z,\bar{y},w,\bar{p}),\end{aligned}$$

where we use the continuity assumptions for f, and recall that differentiable functions are continuous. Further, we set $u^{\hat{k}+1}(z)=:a^{\hat{k}+1}\to\bar{a}:=\bar{u}(z)$ for $\hat{k}\to\infty$, we have

$$\liminf_{\hat{k}\to\infty}g\left(u^{\hat{k}+1}(z)\right)=\liminf_{\hat{k}\to\infty}g\left(a^{\hat{k}+1}\right)\geq g(\bar{a})=g\left(\bar{u}(z)\right),$$

for almost all $z \in Z_i$. Consequently, we obtain the optimality condition

$$H\left(z, \bar{y}, \bar{u}, \bar{p}\right) \leq H\left(z, \bar{y}, w, \bar{p}\right),$$

for all $w \in K_U$ and almost all $z \in Z_i$.

In order to prove (9), we consider (10) inserting the assumption $r \geq \epsilon$ and obtain

$$\begin{aligned} &H\left(z, y^{k+1}, u^{k+1}, p^{k+1}\right) \leq H\left(z, y^{k+1}, w, p^{k+1}\right) \\ &+ \left|p^k\left(z\right) f\left(z, y^k, u^{k+1}\right) - p^{k+1}\left(z\right) f\left(z, y^{k+1}, u^{k+1}\right)\right| + \left|p^k\left(z\right) f\left(z, y^k, w\right) - p^{k+1}\left(z\right) f\left(z, y^{k+1}, w\right)\right| \\ &+ \epsilon\left|\left(\left(w - u^k\left(z\right)\right)^2 - \left(u^{k+1}\left(z\right) - u^k\left(z\right)\right)^2 - \left(w - u^{k+1}\left(z\right)\right)^2\right)\right| \end{aligned} \tag{13}$$

by adding and subtracting corresponding terms. If we choose $\hat{k} \in \hat{K}$ where $y^{\hat{k}}\left(z\right) \to \bar{y}\left(z\right)$, $u^{\hat{k}}\left(z\right) \to \bar{u}\left(z\right)$ and $p^{\hat{k}}\left(z\right) \to \bar{p}\left(z\right)$ for $\hat{k} \to \infty$ instead of $k \in \mathbb{N}_0$ in (13), then by continuity, especially by the required continuity of f, it follows the result (9) for almost all $z \in Z_i$ if $\hat{k}$ is sufficiently large using the boundedness of ϵ and [3, Theorem 2.4, Theorem 6.1]. □

If we consider (9) on a finite number of elements $\tilde{Z}_i \subsetneq Z_i$ and a fixed $\mu > 0$, then there is a $\bar{k} \in \hat{K}$ such that (9) holds for all $z \in \tilde{Z}_i$. This can be seen by applying the calculation following (13) for each $z \in \tilde{Z}_i$ and then choosing the largest $\bar{k}$.

We discuss the following example to illustrate (8).

Example 7. In this example, we consider a linear elliptic optimal control problem on an open and bounded domain $\Omega \subseteq \mathbb{R}^n$ for $u_a \leq u \leq u_b$, $u_a < 0 < u_b$, with the cost functional

$$J\left(y, u\right) \coloneqq \frac{1}{2}\|y - y_d\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}^2 + \beta\|u\|_{L^1},$$

where $\alpha, \beta \geq 0$, and the PDE constraint $-\Delta y = u$. The corresponding Hamiltonian is given by $H\left(t, y, u, p\right) \coloneqq \frac{1}{2}\left(y - y_d\right)^2 + \frac{\alpha}{2}u^2 + \beta|u| + pu$. Then the augmented Hamiltonian is given by

$$K_\epsilon\left(t, y, u, v, p\right) \coloneqq \frac{1}{2}\left(y - y_d\right)^2 + \frac{\alpha}{2}u^2 + \beta|u| + pu + \epsilon\left(u - v\right)^2.$$

In Algorithm 1, we have $K_\epsilon\left(t, y^k, u^{k+1}, u^k, p^k\right)$ with $K_\epsilon\left(t, y^k, u^{k+1}, u^k, p^k\right) \leq K_\epsilon\left(t, y^k, w, u^k, p^k\right)$ for all $w \in K_U$. Now, we show that (8) holds, and for this purpose we consider the cases $0 < u^{k+1} \leq u_b$, $u_a \leq u^{k+1} < 0$, and $u^{k+1} = 0$, separately.

We start with the case that $0 < u^{k+1} < u_b$. In this case, we have

$$\alpha u^{k+1} + \beta + p^k + 2\epsilon\left(u^{k+1} - u^k\right) = 0,$$

thus

$$p^k = -\alpha u^{k+1} - \beta - 2\epsilon\left(u^{k+1} - u^k\right), \tag{14}$$

and

$$u^{k+1} = \frac{2\epsilon u^k - \beta - p^k}{\alpha + 2\epsilon}. \tag{15}$$

We insert (14) into (8) and obtain that

$$\begin{aligned} &\frac{\alpha}{2}\left(u^{k+1}\right)^2 + \beta u^{k+1} + p^k u^{k+1} + \epsilon\left(u^{k+1} - u^k\right)^2 + r\left(w - u^{k+1}\right)^2 \\ &\leq \frac{\alpha}{2}w^2 + \beta|w| + p^k w + \epsilon\left(w - u^k\right)^2 \end{aligned} \tag{16}$$

which equivalently gives by inserting (14) the following

$$r\left(w-u^{k+1}\right) \leq \left(\epsilon+\frac{\alpha}{2}\right)\left(w-u^{k+1}\right)^2+\beta\left(|w|-w\right). \tag{17}$$

For $r=\epsilon+\frac{\alpha}{2} \geq \epsilon$, we have that (17) is fulfilled since if $w>0$, then it is $\beta\left(w-w\right)=0$ and if $w<0$, then we have that $\beta\left(-w-w\right) \geq 0$.

In the case that $u^{k+1}=u_b$, we have from (15) that $u^{k+1} \leq \frac{2\epsilon u^k-\beta-p^k}{\alpha+2\epsilon}$. Consequently, we have that

$$p^k \leq 2\epsilon u^k-\beta-\left(\alpha+2\epsilon\right)u^{k+1}.$$

Since $w-u^{k+1}=w-u_b \leq 0$ and p^k is replaced by an expression that is greater, the expression $\left(w-u^{k+1}\right)p^k$ is replaced by a term that is smaller. However, it is the same as discussed for (16) where p^k is replaced by (14). Therefore we obtain again (17), which holds for $r=\epsilon+\frac{\alpha}{2}$, and thus (16) also holds.

The case $u_a<u^{k+1}<0$ is analogous to the case $0<u^{k+1}<u_b$, where we have $\beta\left(|w|+w\right) \geq 0$ instead of $\beta\left(|w|-w\right) \geq 0$. Furthermore, the same reasoning as in the case $u^{k+1}=u_b$ holds for the case that $u^{k+1}=u_a$ where $u^{k+1} \geq \frac{2\epsilon u^k+\beta-p^k}{\alpha+2\epsilon}$ and thus

$$p^k \geq 2\epsilon u^k+\beta-\left(\alpha+2\epsilon\right)u^{k+1}.$$

In contrast to the case above where $u^{k+1}=u_b$, we have that $w-u^{k+1}=w-u_a \geq 0$ and consequently the expression $\left(w-u^{k+1}\right)p^k$ is again replaced by a term that is smaller and the argument is analogous.

Next, we have the case where $u^{k+1}=0$. The fact that $u^{k+1}=0$ means that

$$K_\epsilon\left(t,y^k,0,u^k,p^k\right) \leq K_\epsilon\left(t,y^k,w,u^k,p^k\right), \tag{18}$$

for all $w \in K_U$. Now, we perform a preliminary discussion that shows that the minimum

$$w_1=\frac{2\epsilon u^k-\beta-p^k}{\alpha+2\epsilon}$$

of the parabola

$$w \mapsto K_\epsilon\left(t,y^k,w,u^k,p^k\right)=\frac{\alpha}{2}w^2+\beta w+p^k w+\epsilon\left(w-u^k\right)^2 \tag{19}$$

with the property (18) is not positive, that means $w_1 \leq 0$ and that the minimum

$$w_2=\frac{2\epsilon u^k+\beta-p^k}{\alpha+2\epsilon}$$

of the parabola

$$w \mapsto K_\epsilon\left(t,y^k,w,u^k,p^k\right)=\frac{\alpha}{2}w^2-\beta w+p^k w+\epsilon\left(w-u^k\right)^2, \tag{20}$$

with the property (18) is not negative, that means $w_2 \geq 0$.

According to (18), the situation corresponding to (19) can be translated into $ax^2+bx+c \geq c$ for all $x>0$ with $a>0$, and the situation corresponding to (20) can be translated into $ax^2+bx+c \geq c$ for all $x<0$ with $a>0$. We start with the case associated with (20). Then we have for the minimum $\tilde{x}$ that it holds $\tilde{x} \leq 0$. This can be seen es follows. The inequality $ax^2+bx+c \geq c$ for all $x>0$ is equivalent to $ax+b \geq 0$ for all $x>0$. The minimum of the function $x \mapsto ax^2+bx+c$ is characterized by the root $\tilde{x}$ of the first derivative such that it holds $2a\tilde{x}+b=0$. If we now assumed that $\tilde{x}>0$, then it would follow that $b<0$ in order to fulfill this equality. In addition we have that $ax+b \geq 0$ holds for all $x>0$, that means in particular it has to hold $a\tilde{x}+b \geq 0$, and $2a\tilde{x}+b=0$ at the same time. Inserting the equation into the inequality provides the contradiction $\frac{b}{2} \geq 0$ to $b<0$ as discussed before. Analogous

for the case (20) where it holds that $ax^2 + bx + c \geq c$ for all $x < 0$ with $a > 0$. Then we have that the minimum is not negative.

Inserting the definition of K_ϵ, Inequality (18) is given by

$$\epsilon \left(u^k\right)^2 \leq \left(\frac{\alpha}{2} + \epsilon\right) w^2 + \beta|w| + p^k w + -2\epsilon u^k w + \epsilon \left(u^k\right)^2.$$

Now, we conclude the preliminary discussion in the situation (19) with $w > 0$ for the corresponding function given by

$$w \mapsto \left(\frac{\alpha}{2} + \epsilon\right) w^2 + \left(\beta + p^k + -2\epsilon u^k\right) w + \epsilon \left(u^k\right)^2,$$

whose minimum is given by $w_1 = \frac{2\epsilon u^k - \beta - p^k}{\alpha + 2\epsilon}$, and it holds

$$\frac{2\epsilon u^k - \beta - p^k}{\alpha + 2\epsilon} \leq 0.$$

Consequently, we have that

$$p^k \geq 2\epsilon u^k - \beta. \tag{21}$$

Analogously, in the situation (20) with $w < 0$ for the function

$$w \mapsto \left(\frac{\alpha}{2} + \epsilon\right) w^2 + \left(-\beta + p^k + -2\epsilon u^k\right) w + \epsilon \left(u^k\right)^2,$$

we have that $p^k \leq 2\epsilon u^k + \beta$.

We have to show that $K_\epsilon \left(t, y^k, 0, u^k, p^k\right) + rw^2 \leq K_\epsilon \left(t, y^k, w, u^k, p^k\right)$, or equivalently that

$$\epsilon \left(u^k\right)^2 + rw^2 \leq \frac{\alpha}{2} w^2 + \beta|w| + p^k w + \epsilon \left(w - u^k\right)^2 \tag{22}$$

is fulfilled for an $r \geq \epsilon$ in order to show (8). For the case that $w > 0$, we have the following

$$rw^2 \leq \left(\frac{\alpha}{2} + \epsilon\right) w^2 + \beta w + 2\epsilon u^k w - \beta w - 2\epsilon u^k w,$$

which is true for $r = \epsilon + \frac{\alpha}{2}$, and thus (22) holds. Analogous is the argument for the case where $w < 0$.

We can conclude stating that in all cases we can choose $r = \epsilon + \frac{\alpha}{2}$ and thus (8) is fulfilled for our cost functional.

We remark that the same result as in Example 7 can be proved for g_1 of (1), however with further assumptions on the iterates u^k of the SQH method.

Now, we show that the condition $H\left(x, \bar{y}, \nabla\bar{y}, \bar{u}, \bar{p}\right) + r\left(w - \bar{u}\right)^2 \leq H\left(x, \bar{y}, \nabla\bar{y}, w, \bar{p}\right)$ for a triple $(\bar{y}, \bar{u}, \bar{p})$ can serve as a sufficient condition for an optimal solution to (3). The idea for the present formulation can be found in [9, 43].

Theorem 8. *Let Assumptions A.1) to A.6) hold. Let $(\bar{y}, \bar{u})$ solve (2) and $\bar{p}$ solve the corresponding adjoint equation (4) for $y \leftarrow \bar{y}$ and $u \leftarrow \bar{u}$. Assume that $|\partial_y f\left(z, \bar{y}, u\right) - \partial_y f\left(z, \bar{y}, \bar{u}\right)| \leq \tilde{c}|u - \bar{u}|$ holds for almost all $z \in Z_i$ where $\tilde{c} > 0$ is a constant. Let $(\bar{y}, \bar{u}, \bar{p})$ fulfil*

$$H\left(z, \bar{y}, \bar{u}, \bar{p}\right) + r\left(w - \bar{u}\right)^2 \leq H\left(z, \bar{y}, w, \bar{p}\right), \tag{23}$$

for all $w \in K_U$ and for almost all $z \in Z_i$ with $r \geq 0$ sufficiently large. Then $(\bar{y}, \bar{u})$ is an optimal solution to (3), that is, $J\left(y, u\right) \geq J\left(\bar{y}, \bar{u}\right)$ for all (y, u) solving (2) with $u \in U_{ad}$.

Proof. Notice that the notation is analogous to the one of the proof of Lemma 4 with $\delta y \coloneqq y - \bar{y}$ and $\delta u \coloneqq u - \bar{u}$ where we do not show the functions dependency on z. We have

$$
\begin{aligned}
&J(y,u) - J(\bar{y},\bar{u}) = \int_{Z_i} h(y) + g(u) - h(\bar{y}) - g(\bar{u})\, dz \\
&= \int_{Z_i} H(z,y,u,\bar{p}) - \bar{p} f(z,y,u) - H(z,\bar{y},\bar{u},\bar{p}) + \bar{p} f(z,\bar{y},\bar{u})\, dz \\
&= \int_{Z_i} H(z,\bar{y},u,\bar{p}) + \partial_y H(z,\bar{y},u,\bar{p})\,\delta y + \frac{1}{2}\left(\partial_{yy} H(z,\bar{y},u,\bar{p})\,(\delta y)^2\right) dz + \int_{Z_i} R(H,\bar{y};\delta y)\, dz \\
&- \int_{Z_i} H(z,\bar{y},\bar{u},\bar{p})\, dz - \int_0^T (\delta y'(\cdot,t), \bar{p}(\cdot,t)) + B(\delta y,\bar{p};t)\, dt \\
&\geq r\int_{Z_i} \delta u^2 dz + \partial_y h(\bar{y})\,\delta y + \bar{p}\partial_y f(\bar{y},u)\,\delta y + \frac{1}{2}\left(\partial_{yy} h(\bar{y})\,(\delta y)^2 + \bar{p}\partial_{yy} f(\bar{y},u)\,(\delta y)^2\right) \\
&+ \int_{Z_i} R(H,\bar{y};\delta y)\, dz - \int_0^T (\delta y'(\cdot,t), \bar{p}(\cdot,t)) + B(\delta y,\bar{p};t)\, dt \\
&= r\int_{Z_i} \delta u^2 dz + \bar{p}\partial_y f(\bar{y},u)\,\delta y - \bar{p}\partial_y f(\bar{y},\bar{u})\,\delta y + \frac{1}{2}\left(\partial_{yy} h(\bar{y})\,(\delta y)^2 + \bar{p}\partial_{yy} f(\bar{y},u)\,(\delta y)^2\right) \\
&+ \int_{Z_i} R(H,\bar{y};\delta y)\, dz - \int_0^T (\delta y'(\cdot,t), \bar{p}(\cdot,t)) + B(\delta y,\bar{p};t)\, dt + \int_0^T -(\bar{p}'(\cdot,t), \delta y(\cdot,t),) + B^*(\bar{p},\delta y;t)\, dt \\
&\geq r\|\delta u\|^2_{L^2(Z_i)} - d_1\|\delta u\|_{L^2(Z_i)}\|\delta y\|_{L^2(Z_i)} - d_2\|\delta u\|^2_{L^2(Z_i)} - d_3\|\delta u\|_{L^2(Z_i)}
\end{aligned}
\tag{24}
$$

for all $u \in U_{ad}$, $d_1, d_2, d_3 > 0$, where we use the partial integration rule [45, Theorem 3.11], the Cauchy-Schwarz inequality [2, Lemma 2.2] in the last inequality and the estimation of $R(H,\bar{y};\delta y)$ as in the proof of Lemma 4 in Appendix A, and $\int_0^T B^*(\bar{p},\delta y;t) - B(\delta y,\bar{p};t)\, dt = \int_0^T B(\delta y,\bar{p};t) - B(\delta y,\bar{p};t)\, dt = 0$. □

If we consider $g(z) \coloneqq \frac{\alpha}{2} z^2 + \beta|z|$, then, with an analogous calculation as in Example 7 for $\epsilon = 0$, we obtain that (23) holds for any $r \in \left[0, \frac{\alpha}{2}\right]$. Consequently, we know that if α is sufficiently large that r can be chosen sufficiently large and we have that any solution which is PMP optimal, that means fulfils (6) ($r = 0$ in (23)), also fulfils (23). We have the following corollary for a special case.

Corollary 9. *Let the assumptions of Theorem 8 hold. If the function f does not depend on y and if h is a quadratic function with $\frac{\partial^2}{\partial y^2} h \geq 0$, then we can choose $r = 0$ in (23) and thus the necessary condition*

$$
H(z,\bar{y},\bar{u},\bar{p}) = \min_{w \in K_U} H(z,\bar{y},w,\bar{p}) \tag{25}
$$

is sufficient for $(\bar{y},\bar{u})$ to be optimal for (3).

4 The SQH method in the case of $L^2 - L^1$ control costs

In this section, we discuss our SQH framework in the case of a semi-smooth PDE control problem. In this case, existence of an optimal solution [45] and its PMP characterization (Theorem 2) can be proved. Moreover, all assumptions for proving convergence of the SQH method (see Theorem 6) are fulfilled.

Consider the following elliptic problem with distributed control given as follows. Find $y \in H_0^1(\Omega)$ and $u \in U_{ad}$, with $K_U = [-100, 100]$, such that

$$\begin{aligned} \min_{y,u} J(y,u) \coloneqq \int_\Omega \frac{1}{2} \left(y(x) - y_d(x)\right)^2 + g(u(x))\, dx \\ (\nabla y, \nabla v) = (u, v) \\ u \in U_{ad}, \end{aligned} \tag{26}$$

for all $v \in H_0^1(\Omega)$. For the cost functional, we choose $g(z) \coloneqq \frac{\alpha}{2} z^2 + \beta |z|$, $\alpha = 0, \beta = 10^{-3}$, $y_d(x) \coloneqq \sin(2\pi x_1)\cos(2\pi x_2) + 1$.

We have $h(y) \coloneqq \frac{1}{2}(y - y_d)^2$ and $f(x, y, u) \coloneqq u$, and the following Hamiltonian

$$H(x, y, u, p) \coloneqq \frac{1}{2}\left(y(x) - y_d(x)\right)^2 + g(u(x)) + p(x)\, u(x).$$

Corresponding to (4), we have the following adjoint problem

$$(\nabla p, \nabla v) = (y - y_d, v),$$

for all $v \in H_0^1(\Omega)$ where $p \in H_0^1(\Omega)$.

We remark that $\|\delta y\|_{L^2(\Omega)} \leq \tilde{c} \|\nabla \delta y\|$, $\|\delta p\|_{L^2(\Omega)} \leq \tilde{c} \|\nabla \delta p\|$, $\tilde{c} \geq 0$ because of the Poincaré inequality [2, 6.7] and thus $\|\delta y\|_{L^2(\Omega)} \leq c \|\delta u\|_{L^2(\Omega)}$, $\|\delta p\|_{L^2(\Omega)} \leq c \|\delta u\|_{L^2(\Omega)}$ and $|(\nabla \delta y, \nabla \delta p)| \leq c^2 \|\delta u\|^2_{L^2(\Omega)}$ because of the Cauchy-Schwarz inequality, see [2, Lemma 2.2] for some constant $c > 0$. Furthermore, we have that $y \in L^\infty(\Omega)$, see Appendix B, and thus applying Theorem 12 to the adjoint equation considering the pointwise boundedness of the control, we have $\|p\|_{L^\infty(\Omega)} \leq c$ for any solution (y, u) to the state equation with $u \in U_{ad}$. As the derivatives $\frac{\partial^2}{\partial y^2} h = 1$, $\frac{\partial}{\partial y} f = \frac{\partial^2}{\partial y^2} f = 0$ and f are continuous, we can apply Theorem 5. In view of Example 7, the assumptions of Theorem 6 are fulfilled which means that any limit of the converging SQH subsequence fulfils the PMP.

To measure PMP optimality, we define the function

$$\triangle H(z) \coloneqq \left(H(z, y, u, p) - \min_{w \in K_U} H(z, y, w, p) \right),$$

where y, u, and p are the return values from the SQH method upon convergence. Furthermore, we report the number $N_{\%}^l$ that is the percentage of the grid points at which the inequality $0 \leq \triangle H \leq 10^{-l}$, $l \in \mathbb{N}$, is fulfilled. This is to verify the PMP optimality (6) up to a tolerance, at least on a subset of grid points.

For the numerical solution, we use finite differences and the following analytical formulas, which are given by a case study, to determine the pointwise minimum of the augmented Hamiltonian in Step 2 of Algorithm 1.

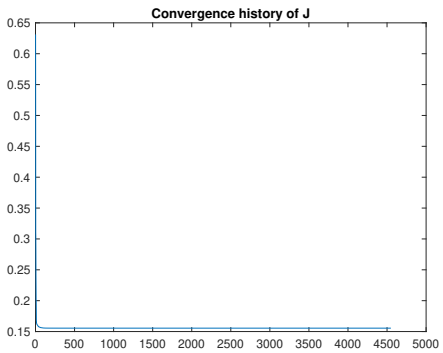
We have that the pointwise minimum is either given by

$$u_1 = \min\left(\max\left(0, \frac{2\epsilon u^k - \beta - p^k}{2\epsilon + \alpha} \right), 100 \right),$$

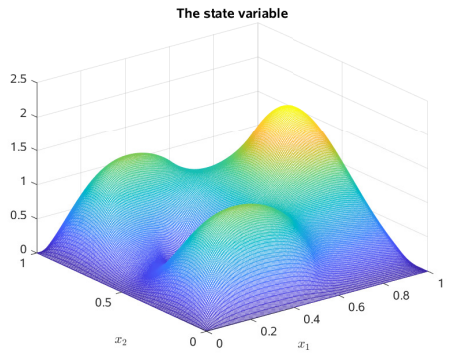
or by

$$u_1 = \min\left(\max\left(-100, \frac{2\epsilon u^k + \beta - p^k}{2\epsilon + \alpha} \right), 0 \right).$$

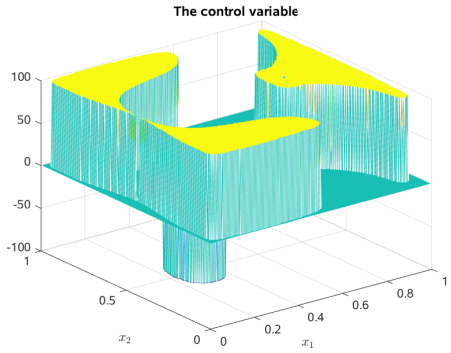
Consequently, the pointwise update of the control is given by the one of these two values that makes the augmented Hamiltonian smaller. (Indeed, this minimum can also be calculated by a secant method.)



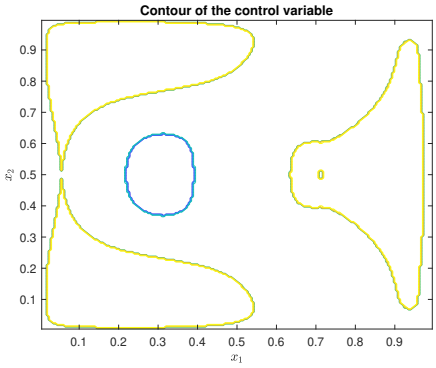
(a) Convergence history of the cost functional for the updates to the initial guess of the control.



(b) The state y.



(c) The optimal control u.



(d) The optimal control u as a contour plot.

Figure 1: Results for the elliptic optimal control problem with distributed control and L^1 cost term.

In Figure 1, we depict the optimal solution obtained with Algorithm 1 for the elliptic optimal control problem (26). In this case, the parameters are as follows. The initial guess for ϵ equals $\frac{1}{150}$ and for $u^0 = 0$, $\kappa = 10^{-6}$, $\sigma = 50$, $\zeta = \frac{3}{20}$, $\eta = 10^{-9}$. The domain Ω is discretised with an uniform mesh with size $\triangle x = \frac{1}{200}$. In this experiment, choosing $\kappa = 10^{-6}$, we have that $N_{\%}^{16} = 96.2\%$ and for $\kappa = 10^{-12}$, we have $N_{\%}^{16} = 99.9\%$, which very well validates the statement of Theorem 6.

The numerical results obtained in this section have been computed with the code SQH_P1_L1.zip:

https://opus.bibliothek.uni-wuerzburg.de/frontdoor/index/index/docId/17858

(We the same link the reader can find the code for the same optimization problem with a bilinear control mechanism.)

Based on our experience, we can state that similar theoretical and numerical results are obtained in the case elliptic and parabolic optimal control problems with linear and bilinear distributed controls and semi-smooth cost functionals as in (26). Moreover, our numerical experience shows that the SQH algorithm may be competitive with respect to other state-of-the-art methods (e.g., semi-smooth Newton methods), in the case where globalization is usually required.

Although this is on itself a very valuable result, our main aim is to demonstrate that our PMP - SQH framework appears appropriate to address a presently wide open topic of PDE optimization, that is, the case of non-convex and discontinuous cost functionals. This discussion is not only of academic interest but it is motivated by a collection of outstanding works on discontinuous functionals with real applications; see, e.g., [6, 29, 39]. For these problems, the proof of existence of minimizers is a great

challenge that has been addressed in specific cases with ad-hoc techniques (e.g., relaxation). However, also the characterization and efficient numerical computation of these solutions is equally challenging and the purpose of our case study below is to contribute towards the formulation of a PMP - SQH framework that allows to accommodate non-convex and discontinuous PDE optimization problems.

5 The SQH scheme: a case study

This section is devoted to a detailed discussion of seven non-convex and discontinuous PDE control problems focusing on the PMP characterization of minimizers (assuming they exist) and on their computation by our SQH method. Our different PDE control problems are named P.1) to P.7), as given in the list (1), and for all these problems, the purpose of the numerical experiments is to demonstrate the applicability and computational performance of the SQH scheme in solving these problems.

In particular, we show that by choosing a more stringent tolerance in the SQH stopping criterion, the fulfilment of the PMP optimality condition improves as expected. Furthermore, we plot the optimal solutions to the given problems and show the convergence history of the SQH scheme in terms of reduction of the value of the cost functional, which demonstrates that in all cases the SQH scheme provides a minimizing sequence.

In the case of elliptic problems, we choose the domain $\Omega = (0,1) \times (0,1)$ and use a 5-point finite-difference discretisation of the Laplacian on a grid with mesh size $\triangle x = \frac{1}{50}$. In the parabolic cases, we have $Q = (0,1) \times (0,1)$, and we use the implicit Euler scheme with finite differences on a space-time grid with $\triangle x = \frac{1}{50}$ and $\triangle t = \frac{1}{100}$.

To solve linear problems (or subproblems), we use the Matlab backslash operator. Non-linear problems as P.5) and P.6) are solved by a Gauss-Seidel-Picard iteration [11] with a tolerance on the discrete L^2-norm of the residuum of 10^{-8}. Specifically, for P.5) the state variable is updated pointwise within a loop over the interior points of the domain with

$$y(i,j) = \frac{1}{4}\left(y(i+1,j) + y(i-1,j) + y(i,j+1) + y(i,j-1) + h^2\left(u(i-1,j-1) - y^3(i,j)\right)\right).$$

For P.6) we use

$$y(i,j) = \frac{y(i+1,j) + y(i-1,j) + y(i,j+1) + y(i,j-1)}{4} + u(i-1,j-1) \cdot \begin{cases} \frac{\triangle x^2}{4+\triangle x^2} & \text{if } y(i,j) \geq 0 \\ \frac{\triangle x^2}{4} & \text{else} \end{cases}.$$

The minimization of the augmented Hamiltonian in Step 2 of Algorithm 1 can be performed by a secant method or by an analytical formula that solves the one-dimensional (since u is scalar valued) minimization problem.

In the attempt to provide an overall view and comparison of the SQH performance with all test cases, we present results concerning PMP optimality of the SQH solution in Table 1. In this table, N_{up} denotes the total number of updates that are made by the SQH method on the given grid starting with the same initial guess of the control and "iter" is the number of total sweeps, which means Step 2 to Step 5 in Algorithm 1.

Table 1 provides an overview for the optimality results of the SQH method for the problems P.1) to P.7) . In most cases, PMP optimality of over 90% is achieved with very stringent tolerance ($l = 8, 12$). We remark that in the case P.7), we solve an augmented problem, see Section 5.5 for details.

The results of Table 1 correspond to the following choice of parameters: $\sigma = 50$, $\zeta = \frac{3}{20}$, $\eta = 10^{-9}$, $\kappa = 10^{-8}$, $u^0 = 0$, and the initial guess for $\epsilon = \frac{1}{150}$.

	N_{up}	iter	$\max_{z\in Z_i} \triangle H(z)$	$\frac{N^2_{\%}}{\%}$	$\frac{N^4_{\%}}{\%}$	$\frac{N^6_{\%}}{\%}$	$\frac{N^8_{\%}}{\%}$	$\frac{N^{12}_{\%}}{\%}$
P.1)	26	40	$3.42 \cdot 10^{-3}$	100	99.3336	99.2503	99.1670	99.1670
P.2)	40	58	$2.74 \cdot 10^{-3}$	100	99.8776	99.8776	99.8776	99.8776
P.3)	436	645	$2.88 \cdot 10^{-3}$	100	90.9204	90.8372	90.8372	90.8372
P.4)	173	255	$3.11 \cdot 10^{-5}$	100	100	97.2245	96.8776	96.8776
P.5)	235	348	$7.75 \cdot 10^{-3}$	100	97.0012	91.5452	90.9621	90.9621
P.6)	283	419	$3.00 \cdot 10^{-1}$	76.2371	75.3393	75.3393	75.3393	75.3393
P.7)	864	1281	$2.05 \cdot 10^{-2}$	93.7526	87.4636	83.7151	83.3819	83.2153

Table 1: Numerical investigation of optimality of the SQH solution to the problems P.1) for P.7) with $\kappa = 10^{-8}$.

5.1 Application to the linear control problems P.1) and P.2)

In this section, we consider linear elliptic and parabolic optimal control problems, given by P.1) and P.2), with distributed control and a discontinuous cost functional.

We start with the elliptic case P.1), and choose $Z_{\mathrm{e}} = \Omega := (0,1) \times (0,1)$. Consider the following optimal control problem:

Find $y \in H_0^1(\Omega)$ and $u \in U_{ad}$ with $K_U = [0, 100]$ such that

$$\begin{aligned} \min_{y,u} J(y,u) := \int_\Omega \frac{1}{2} \left(y(x) - y_d(x)\right)^2 + g(u(x))\, dx \\ (\nabla y, \nabla v) = (u, v) \\ u \in U_{ad}, \end{aligned} \tag{27}$$

for all $v \in H_0^1(\Omega)$ where $g(z) := \begin{cases} \beta|z| & \text{if } |z| > 20 \\ 0 & \text{else} \end{cases}$, $\beta = 10^{-3}$, $y_d(x) := \sin(2\pi x_1)\cos(2\pi x_2) + 1$.

We have $h(y) := \frac{1}{2}(y - y_d)^2$ and $f(x,y,u) := u$ and the following Hamiltonian

$$H(x,y,u,p) := \frac{1}{2}\left(y(x) - y_d(x)\right)^2 + g(u(x)) + p(x)\, u(x).$$

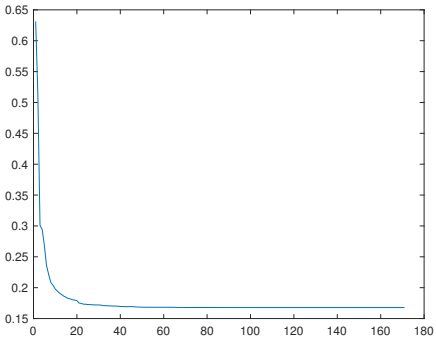
Corresponding to (4), the adjoint problem is given by

$$(\nabla p, \nabla v) = (y - y_d, v),$$

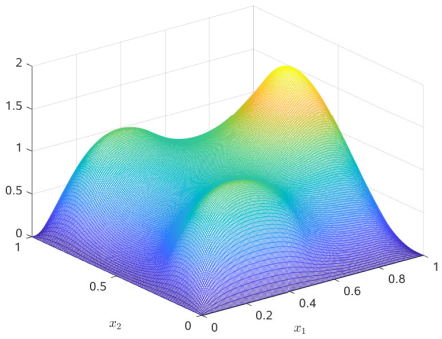
for all $v \in H_0^1(\Omega)$ where $p \in H_0^1(\Omega)$. The discussion that the Assumptions A.1) to A.6) are fulfilled is analogous to the L^1 case of Section 4.

In Figure 2, we depict the optimal solution obtained with Algorithm 1 for the elliptic optimal control problem (27). In this case, the parameters are as follows. The initial guess for ϵ equals $\frac{1}{150}$ and for $u^0 = 0$, $\kappa = 10^{-6}$, $\sigma = 50$, $\zeta = \frac{3}{20}$, $\eta = 10^{-9}$. The domain Ω is discretised with an equidistant mesh with size $\triangle x = \frac{1}{200}$.

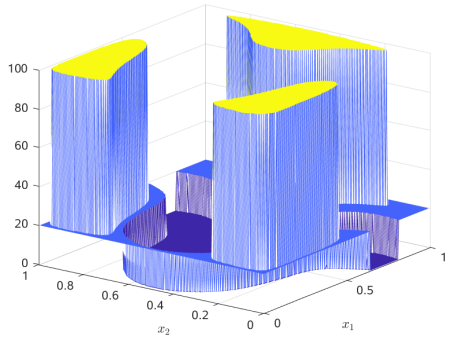
Now, we numerically validate that the solution of the SQH method fulfils the PMP. We denote $(\bar{y}, \bar{u}, \bar{p})$ the solution to which Algorithm 1 converges. The inequality $H(x, \bar{y}, \bar{u}, \bar{p}) - \min_{w \in K_U} H(x, \bar{y}, w, \bar{p}) \leq \mathrm{eps}$, $\mathrm{eps} = 2.2 \cdot 10^{-16}$, is fulfilled at 37.39% of the grid points for $\kappa = 10^{-1}$, at 86.23% of the grid points for $\kappa = 10^{-3}$, at 99.15% of the grid points for $\kappa = 10^{-4}$, at 99.93% of the grid points for $\kappa = 10^{-6}$ and at 99.95% of the grid points for $\kappa = 10^{-8}$. Further results are given in Table 1.



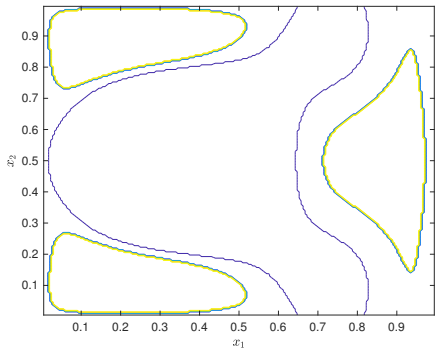
(a) Convergence history of the cost functional for the updates to the initial guess of the control.



(b) The state y.



(c) The optimal control u.



(d) The optimal control u as a contour plot.

Figure 2: Results for the elliptic optimal control problem P.1).

The results presented in this section are obtained with the code SQH_P1.zip available at https://opus.bibliothek.uni-wuerzburg.de/frontdoor/index/index/docId/16566

Similar results for the parabolic control problem P.2) can be found in [12].

The corresponding code SQH_P2.zip is available at:

https://opus.bibliothek.uni-wuerzburg.de/frontdoor/index/index/docId/16566

5.2 Application to bilinear optimal control problems

In this section, we consider the bilinear control problems P.3) and P.4) with $K_U \subseteq \mathbb{R}_0^+$. We start with the elliptic problem P.3) given by

$$\begin{aligned} \min_{y,u} J(y,u) := \int_\Omega \frac{1}{2}\left(y(x) - y_d(x)\right)^2 + g(u(x))\,dx \\ (\nabla y, \nabla v) + (uy, v) = \left(\tilde{f}, v\right) \\ u \in U_{ad}, \end{aligned} \tag{28}$$

for all $v \in H_0^1(\Omega)$ and $y_d \in L^q(\Omega)$ and g is specified below.

Now, we check the Assumptions A.1) to A.6). From (5), the corresponding Hamiltonian for (28) is given by

$$H(x,y,u,p) = \frac{1}{2}\left(y(x) - y_d(x)\right)^2 + g(u(x)) + p(x)\,\tilde{f}(x) - u(x)\,y(x)\,p(x). \tag{29}$$

According to (4), the adjoint problem is as follows: Find $p \in H_0^1(\Omega)$ such that

$$(\nabla p, \nabla v) + (up, v) = (y - y_d, v), \tag{30}$$

holds for all $v \in H_0^1(\Omega)$. Notice that all solutions to (30) are essentially bounded by a constant independent of y and u, see Theorem 13 in Appendix B.

We remark that $\|\nabla \delta y\|_{L^2(\Omega)} \leq \tilde{c}\|\delta u\|_{L^2(\Omega)}$, $\tilde{c} \geq 0$ and $\|\delta y\|_{L^2(\Omega)} \leq c\|\delta u\|_{L^2(\Omega)}$. If we define $\delta u \coloneqq u_1 - u_2$ and $\delta y \coloneqq y_1 - y_2$ and use the Poincaré inequality [2, 6.7], then the previous inequalities follow from taking the difference of (28) for two different pairs (y_ℓ, u_ℓ), $\ell \in \{1, 2\}$. Consequently, we obtain

$$(\nabla (y_1 - y_2), \nabla v) + (u_1 y_1, v) - (u_2 y_2, v) = 0,$$

equivalently

$$(\nabla (y_1 - y_2), \nabla v) + (u_2 (y_1 - y_2), v) = (-(u_1 - u_2) y_1, v).$$

By choosing $v = y_1 - y_2$, we have

$$\|\nabla (y_1 - y_2)\|_{L^2(\Omega)}^2 \leq \|y_1\|_{L^\infty(\Omega)} \|u_1 - u_2\|_{L^2(\Omega)} \|y_1 - y_2\|_{L^2(\Omega)} \tag{31}$$

with the Cauchy-Schwarz inequality, see [2, 2.2].

To discuss the boundedness of the adjoint, we first subtract (30) for (y_2, u_2, p_2) from (30) for (y_1, u_1, p_1) and obtain

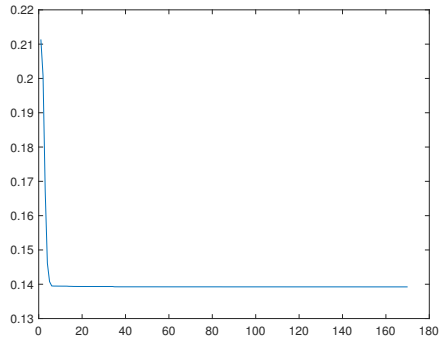
$$(\nabla \delta p, \nabla v) + (u_2 \delta p, v) = ((y_1 - y_2)\, \delta y, v) - (p_1 (u_1 - u_2), v),$$

where $\delta p \coloneqq p_1 - p_2$. Because y_1 and y_2 are essentially bounded, see Theorem 13 and [35, Corollary 37], and since $u_2 \geq 0$ almost everywhere, $p_1 \in L^\infty(\Omega)$, we have that $\|\nabla \delta p\|_{L^2(\Omega)} \leq \tilde{c}\|\delta u\|_{L^2(\Omega)}$, $\tilde{c} > 0$. With an analogous calculation as for (31) we obtain $\|\delta p\|_{L^2(\Omega)} \leq c\|\delta u\|_{L^2(\Omega)}$ with the Poincaré inequality [2, 6.7]. From this, we also have

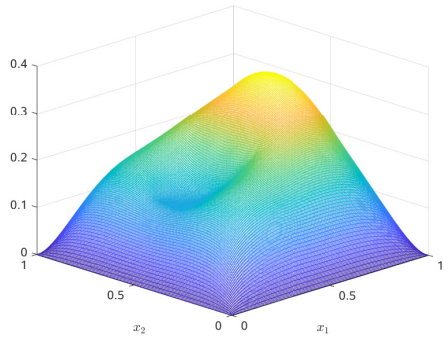
$$|B(\delta y, \delta p)| = |(\nabla \delta y, \nabla \delta p)| \leq \|\nabla \delta y\|_{L^2(\Omega)} \|\nabla \delta p\|_{L^2(\Omega)}.$$

As $f = \tilde{f} - uy$ is continuous with bounded first and second derivatives with respect to y and u, Assumptions A.1) to A.6) are fulfilled and we can apply the Theorems 5.

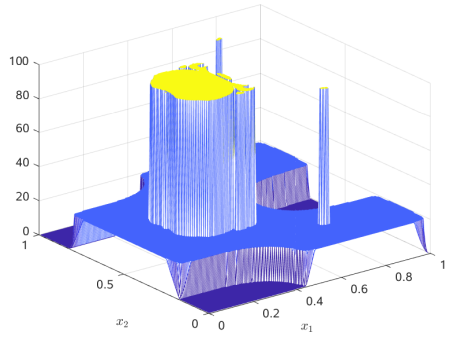
Now, we choose $\Omega = (0,1) \times (0,1)$ and $g(z) \coloneqq \begin{cases} \beta|z| & \text{if } |z| > 20 \\ 0 & \text{else} \end{cases}$, $\beta = 10^{-3}$ and $y_d(x) \coloneqq \sin(2\pi x_1)\cos(2\pi x_2)$. In Figure 3, one we depict the solution obtained with Algorithm 1 solving (28) with $\tilde{f} = 10$. The parameters are as follows. The initial guess for ϵ equals $\frac{1}{150}$ and for $u^0 = 0$, $\kappa = 10^{-6}$, $\sigma = 50$, $\zeta = \frac{3}{20}$, $\eta = 10^{-9}$, $K_U = [0, 100]$. The domain Ω is discretised with an equidistant mesh with size $\triangle x = \frac{1}{200}$.



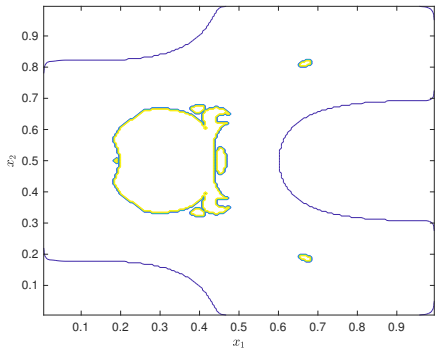
(a) Convergence history of the cost functional for the updates to the initial guess of the control.



(b) The state y.



(c) The optimal control u.



(d) The optimal control u as a contour plot.

Figure 3: Solution to the elliptic bilinear optimal control problem P.3).

Notice that the two peaks of the control appear also at finer discretisation, e.g., $\triangle x = \frac{1}{500}$, and thus they are not numerical artefacts.

Next, we numerically validate that the solution of the SQH method fulfils the PMP. We denote with $(\bar{y}, \bar{u}, \bar{p})$ the solution obtained with Algorithm 1. The inequality $H(x, \bar{y}, \bar{u}, \bar{p}) - \min_{w \in K_U} H(x, \bar{y}, w, \bar{p}) \leq$ eps, eps $= 2.2 \cdot 10^{-16}$, is fulfilled at 73.31% of the grid points for $\kappa = 10^{-1}$, at 84.28% of the grid points for $\kappa = 10^{-3}$, at 90.80% of the grid points for $\kappa = 10^{-6}$, at 94.21% of the grid points for $\kappa = 10^{-10}$. See also Table 1 for additional results.

The results for P.3) are obtained with the code SQH_P3.zip:

https://opus.bibliothek.uni-wuerzburg.de/frontdoor/index/index/docId/16566

Next, we present numerical results with the bilinear parabolic control problem P.4). We have

$$\min_{y,u} \int_0^T \int_\Omega \frac{1}{2} \left(y(x,t) - y_d(x,t)\right)^2 + g(u(x,t))\, dxdt$$

$$\left(y'(\cdot,t), v\right) + D\left(\nabla y(\cdot,t), \nabla v\right) + \left(u(\cdot,t)\, y(\cdot,t), v\right) = \left(\tilde{f}(\cdot,t), v\right) \text{ in } Q \text{ for all } v \in H_0^1(\Omega)$$

$$u \in U_{ad},$$

with $y(x,t) = 0$ for $x \in \partial\Omega$ and all $t \in [0,T]$, $y(0,x) = 0$ for all $x \in \Omega$, $D = \frac{1}{5}$. Further, $T = 1$, $\Omega = (0,1)$, $Q = (0,1) \times (0,1)$, $y_d(x,t) = \begin{cases} \frac{1}{2} & \text{if } \bar{x}(t) - \frac{7}{100} \leq x \leq \bar{x}(t) + \frac{7}{100} \\ 0 & \text{else} \end{cases}$, where $\bar{x}(t) := \frac{1}{2} + \frac{2}{5}\sin(2\pi t)$,

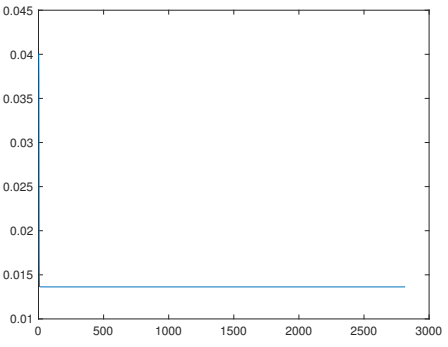
$g(z) := \begin{cases} \beta|z| & \text{if } |z| > 10 \\ 0 & \text{else} \end{cases}$, $\beta = 10^{-4}$, $\tilde{f}$ is a constant function with value 1 and $K_U = [0, 15]$.

Also in this parabolic case, Assumptions A.1) to A.6) are fulfilled and we can apply Theorem 5. In fact $\|y\|_{L^\infty(Q)} \leq c$, independent of the control. Following we can adopt the proof of [12, Theorem 6.1] due to $uy^2 \geq 0$ and thus obtain [12, (28)].

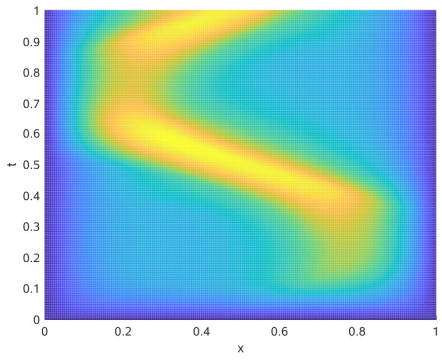
In this case the Hamiltonian is given by

$$H(x,t,y,u,p) := \frac{1}{2}(y(x,t) - y_d(x,t))^2 + g(u(x,t)) + p(x,t)\,\tilde{f}(x,t) - u(x,t)\,y(x,t)\,p(x,t).$$

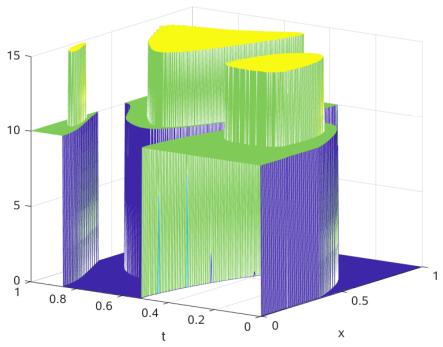
The parameters for the numerical experiment are as follows. The initial guess for $\epsilon = \frac{3}{5}$ and for u is the zero function. The parameters $\sigma = 50$, $\zeta = \frac{3}{20}$, $\eta = 10^{-12}$, $\kappa = 10^{-12}$. The discretisation is equidistant in time and space with $\triangle t = \frac{1}{400}$ and $\triangle x = \frac{1}{200}$. The results are presented in Figure 4.



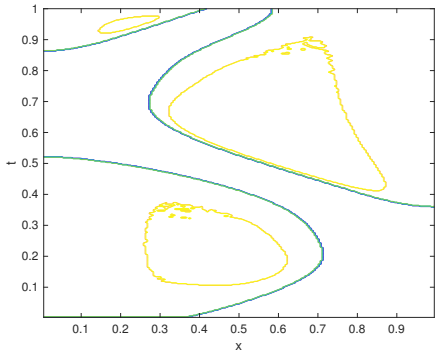
(a) Convergence history of the cost functional for the updates to the initial guess of the control.



(b) The state y viewed from above.



(c) The control u.



(d) A contour plot of the control u.

Figure 4: Numerical results with the parabolic bilinear optimal control problem P.4).

Also in the present parabolic case, we numerically validate that the solution of the SQH method fulfils the PMP. We denote with $(\bar{y}, \bar{u}, \bar{p})$ the solution obtained by Algorithm 1. The inequality $H(x,t,\bar{y},\bar{u},\bar{p}) - \min_{w \in K_U} H(x,t,\bar{y},w,\bar{p}) \leq \text{eps}$, $\text{eps} = 2.2 \cdot 10^{-16}$, is fulfilled at 89.27% of the grid points for $\kappa = 10^{-4}$, at 93.64% of the grid points for $\kappa = 10^{-6}$, at 97.06% of the grid points for $\kappa = 10^{-8}$, at 97.59% of the grid points for $\kappa = 10^{-10}$ and at 98.04% of the grid points for $\kappa = 10^{-12}$.

The results for P.4) are obtained with the code SQH_P4.zip:

https://opus.bibliothek.uni-wuerzburg.de/frontdoor/index/index/docId/16566

5.3 Application to a non-linear elliptic optimal control problem

In this section, we discuss P.5) that is given by

$$\begin{aligned} \min_{y,u} J(y,u) \coloneqq \int_\Omega \frac{1}{2}\left(y(x)-y_d(x)\right)^2 + g(u(x))\,dx \\ (\nabla y, \nabla v) + \left(y^3, v\right) = (u,v) \\ u \in U_{ad}, \end{aligned} \tag{32}$$

where we choose $\Omega \coloneqq (0,1)\times(0,1)$, $g(z) \coloneqq \begin{cases} \beta|z| & \text{if } |z|>20 \\ 0 & \text{else} \end{cases}$, $\beta = 10^{-3}$, $K_U \coloneqq [-100,100]$, and $y_d(x) \coloneqq \sin(2\pi x_1)\cos(2\pi x_2)$.

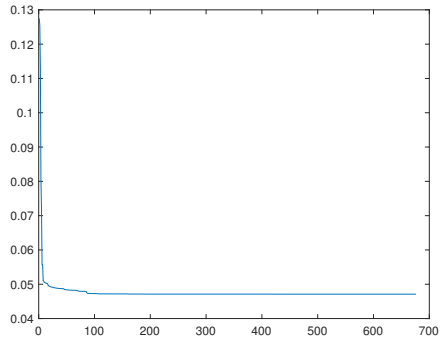
We have $h(y) \coloneqq \frac{1}{2}(y-y_d)^2$ and $f(x,y,u) \coloneqq u - y^3$ and we define the following Hamiltonian

$$H(x,y,u,p) \coloneqq \frac{1}{2}\left(y(x)-y_d(x)\right)^2 + g(u(x)) + p(x)\left(u(x)-y^3(x)\right).$$

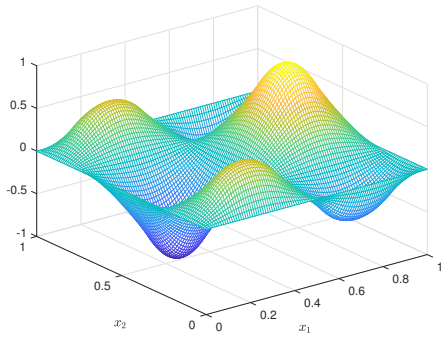
Corresponding to (4), we have the following adjoint problem for $p \in H_0^1(\Omega)$

$$(\nabla p, \nabla v) + \left(3y^2 p, v\right) = (y - y_d, v),$$

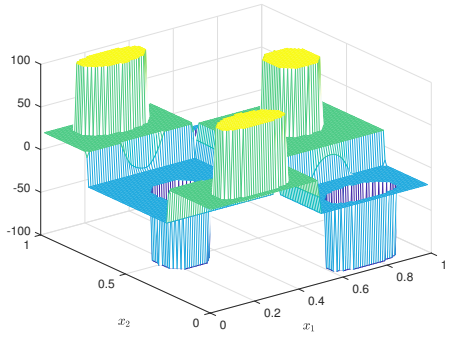
for all $v \in H_0^1(\Omega)$. We remark that $\|\delta y\|_{L^2(\Omega)} \le c\|\delta u\|_{L^2(\Omega)}$, $\|\delta p\|_{L^2(\Omega)} \le c\|\delta u\|_{L^2(\Omega)}$ and $|(\nabla\delta y, \nabla\delta p)| \le c^2\|\delta u\|^2_{L^2(\Omega)}$ for some constant $c>0$ analogously to the elliptic case in Section 5.1. Furthermore, we have that $y \in L^\infty(\Omega)$, see Appendix B, and thus applying Theorem 12 to the adjoint equation considering the pointwise boundedness of the control, we have $\|p\|_{L^\infty(\Omega)} \le c$ for any solution (y,u) to the state equation with $u \in U_{ad}$. The function $f(\cdot,y,u) = u - y^3$ is continuous with respect to y and u, the derivatives $\frac{\partial^2}{\partial y^2}h = 1$, $\frac{\partial}{\partial y}f = 3y^2$, $\frac{\partial}{\partial y}f = 6y$ are continuous and bounded as y is essentially bounded by a constant independent of $u \in U_{ad}$. Therefore, Assumptions A.1) to A.6) are fulfilled and we can apply Theorem 5. The parameters in Algorithm 1 are chosen as follows. The initial guess for ϵ equals $\frac{1}{150}$ and for $u^0 = 0$, $\kappa = 10^{-12}$, $\sigma = 50$, $\zeta = \frac{3}{20}$, $\eta = 10^{-9}$. The domain Ω is discretized with a mesh with size $\triangle x = \frac{1}{100}$. The non-linear discrete PDE is solved until the L^2-norm of its residuum is less than 10^{-6}. The results are shown in Figure 5 where one can see the action of the bounds on the control and of the discontinuous control costs.



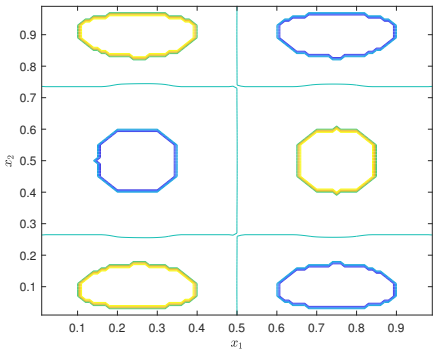
(a) Convergence history of the cost functional for the updates to the initial guess of the control.



(b) The state y.



(c) The optimal control u.



(d) The optimal control u as a contour plot.

Figure 5: Solution to the non-linear elliptic optimal control problem P.5).

Next, we numerically validate that the solution of the SQH method fulfils the PMP. We denote with $(\bar{y}, \bar{u}, \bar{p})$ the solution obtained with Algorithm 1. The inequality $H(x, t, \bar{y}, \bar{u}, \bar{p}) - \min_{w \in K_U} H(x, t, \bar{y}, w, \bar{p}) \leq 10^{-7}$, is fulfilled at 37.68% of the grid points for $\kappa = 10^{-2}$, is fulfilled at 79.88% of the grid points for $\kappa = 10^{-4}$, at 94.34% of the grid points for $\kappa = 10^{-6}$, at 94.83% of the grid points for $\kappa = 10^{-8}$ and at 95.25% of the grid points for $\kappa = 10^{-12}$. Notice that we use the smaller tolerance 10^{-7} instead of the machine precision $2.2 \cdot 10^{-16}$ since the state equation is only solved to a tolerance of 10^{-6} and not exactly as in the linear case.

The results for P.5) are obtained with the code SQH_P5.zip:

https://opus.bibliothek.uni-wuerzburg.de/frontdoor/index/index/docId/16566

5.4 Application to an elliptic optimal control problem with L^1 tracking term

Consider the non-smooth optimal control problem P.6) that is given by

$$\begin{aligned} \min_{y,u} J(y,u) \coloneqq \int_\Omega |y(x) - y_d(x)| + g(u(x))\, dx \\ (\nabla y, \nabla v) + (\max(0, y), v) = (u, v) \\ u \in U_{ad} \end{aligned} \tag{33}$$

for all $v \in H_0^1(\Omega)$ where $\Omega \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, open and bounded, $y_d \in L^1(\Omega)$, $g : \mathbb{R} \to \mathbb{R}$, $z \mapsto g(z)$ lower semi-continuous and non-negative with $\int_\Omega g(u(x))\, dx < \infty$ for all $u \in U_{ad}$. In the experiment, we

choose $g(z) \coloneqq \beta \ln(1+|z|)$, $\beta > 0$.

Remark 10. The characterisation of a solution to (33) with the PMP is in general not possible with the technique in [38] for the following reason. We define $h(y) \coloneqq |y - y_d|$. Then, because of the Lipschitz continuity of h, we have the existence of a function h' such that

$$\begin{aligned} &h(y_1(x)) - h(y_2(x)) = h(y_2(x) + \eta(y_1(x) - y_2(x)))\,\big|_{\eta=0}^{1} \\ &= \int_0^1 h'(y)\,\big|_{y=y_2(x)+\eta(y_1(x)-y_2(x))} d\eta\,(y_1(x) - y_2(x)) \end{aligned} \tag{34}$$

almost everywhere on Ω, see [7, Theorem 7.3], and we define $\tilde{h}(y_1, y_2) \coloneqq \int_0^1 h'(y)\,\big|_{y=y_2(x)+\eta(y_1(x)-y_2(x))} d\eta$. In order to apply the technique used in [38, Proposition 4.4], we need the existence of a function $p : \Omega \to \mathbb{R}$, $z \mapsto p(z)$ such that $\lim_{k\to\infty} \|p_k - p\|_{L^\infty(\Omega)} = 0$, where p_k is the solution to the so-called intermediate adjoint equation, see [38] for details. This is proved by subtracting the adjoint equation from the intermediate adjoint equation, where it is necessary to define a pointwise limit $\lim_{k\to\infty} \tilde{h}(y_k, y)(x) = h'(y)(x)$ almost everywhere on Ω, for $y_k \to y$ pointwise for $k \to \infty$ almost everywhere on Ω. In the case of $y(x) = y_d(x)$ on a set M of measure non zero and $\lim_{k\to\infty} y_k(x) = y(x)$ for $x \in M$, then we do not know if the sign of $y_k(x) - y(x)$, $x \in M$ changes along the sequence $(y_k)_{k\in\mathbb{N}}$ and we cannot extract a subsequence with constant sign of $y_k(x) - y(x)$, as there are uncountable many elements in M. Consequently, the required limit $\lim_{k\to\infty} \tilde{h}(y_k, y)(x)$ does not exist in general and the proof used so far does not work.

Nevertheless, we apply our SQH Algorithm 1 to P.6). For this purpose, we consider the following Hamiltonian

$$H(x, y, u, p) = |y(x) - y_d(x)| + \beta \ln(1 + |u(x)|) + p(x)\,u(x) - p(x)\max(0, y(x)),$$

and the following adjoint equation

$$\int_\Omega \nabla p(x)\,\nabla v(x) + h_2(y(x))\,p(x)\,v(x)\,dx = \int_\Omega h_1(y(x))\,v(x)\,dx, \tag{35}$$

where $h_1(y(x)) \coloneqq \begin{cases} 1 & \text{if } y(x) \geq y_d(x) \\ -1 & \text{else} \end{cases}$ and $h_2(y) \coloneqq \begin{cases} 1 & \text{if } y \geq 0 \\ 0 & \text{else} \end{cases}$ are bounded and measurable, see [18, 2.1 Measurable Functions] and therefore elements of $L^\infty(\Omega)$. Thus (35) is uniquely solvable, see [25, Theorem 3 on page 301] with $\|p\|_{L^\infty(\Omega)} \leq c$, $c > 0$ as $h_2 \geq 0$, see Theorem 12 in the appendix.

The motivation for (35) is related to (34) that is needed for the characterization of a solution with the PMP. Based on the definition of h', we obtain that

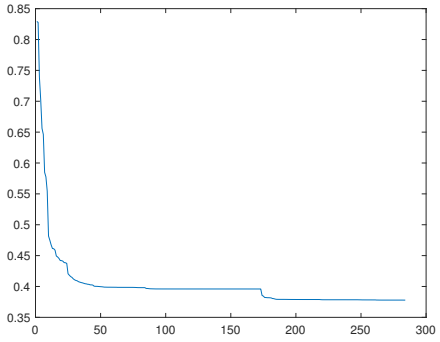
$$|y_k - y_d| - |y^* - y_d| = \tilde{h}(y_k, y^*)(y_k - y^*).$$

Although, as discussed above, there is in general no pointwise limit of $\tilde{h}(y_k, y^*)$, we proceed formally by taking $h' = h_1$, and similarly we take h_2 as above.

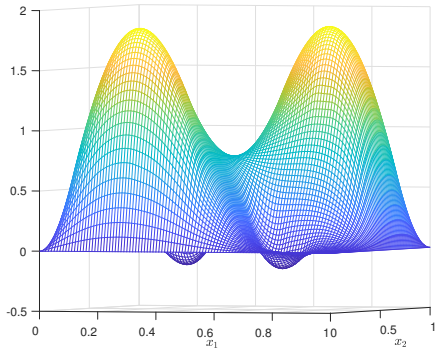
We consider P.6) with $\Omega \coloneqq (0,1) \times (0,1)$, $y_d(x) \coloneqq \sin(2\pi x_1)\sin(2\pi x_2) + \frac{8}{10}$, $K_U = [-100, 100]$ and $\beta = 9 \cdot 10^{-2}$. In our finite differences framework, the non-linear equation $-\Delta y + \max(0, y) = u$ is solved by a Picard iteration until the L^2-norm of its residuum is less than 10^{-6}. The initial guess for the control equals zero and ϵ equals $\frac{1}{150}$. The parameter $\sigma = 50$, $\zeta = \frac{3}{20}$, $\eta = 10^{-9}$ and $\kappa = 10^{-8}$ where Ω is equidistantly discretised with $\triangle x = \frac{1}{100}$.

Results of this experiment are shown in Figure 6. Notice the fast reduction of the value of the cost functional in the first few iterations. This shows that the SQH method also works well in the case of a problem with L^1 tracking term and non-smooth PDE constraints.

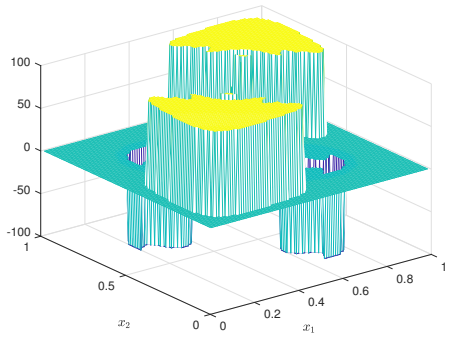
Notice that, although we cannot prove that (6) is necessary for a solution to (33), the numerical optimality of the solution obtained with the SQH scheme is fulfilled even for small tolerances in more than 75% of the grid points, see Table 1.



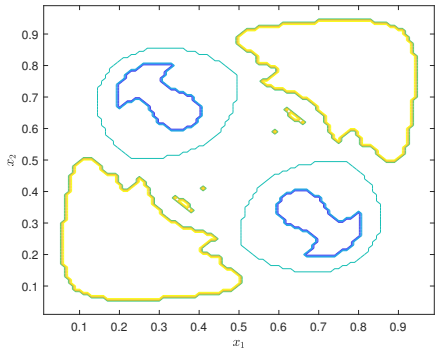
(a) Convergence history of the cost functional for the updates to the initial guess of the control.



(b) The state y.



(c) The control u.



(d) The control u as a contour plot.

Figure 6: Results of the non-smooth optimal control problem P.6) for $\kappa = 10^{-8}$.

The results for P.6) are obtained with the code SQH_P6.zip:
https://opus.bibliothek.uni-wuerzburg.de/frontdoor/index/index/docId/16758

5.5 Application to a state-constraint optimal control problem

In this section, we discuss P.7) that is given by

$$\begin{aligned} &\min_{y,u} J(y,u) := \int_\Omega h(y(x)) + g(u(x))\,dx \\ &\text{s.t. } (\nabla y, \nabla v) = (u, v) \\ &y \le \xi \\ &u \in U_{ad} \end{aligned} \tag{36}$$

for all $v \in H_0^1(\Omega)$ where $h(y) := \frac{1}{2}(y - y_d)^2$, $g(u) := \begin{cases} \beta|u| & \text{if } |u| > 20 \\ 0 & \text{else} \end{cases}$, $\beta = 10^{-3}$, $K_U := [-100, 100]$, $\xi \in \mathbb{R}$ and we assume that (36) admits a solution denoted by $(\bar{y}, \bar{u})$.

In this case, PMP optimality involves multipliers that are only implicitly characterised by inequalities; see, e.g., [14]. For this reason, the computation of solutions to (36) is a delicate issue. We follow the idea of augmented Lagrangian [28, Section 3] and transform the optimal control problem (36) into the following

$$\begin{aligned} &\min_{y,u} J(y,u;\xi,\gamma) \coloneqq \int_\Omega h_\xi(y(x);\gamma) + g(u(x))\,dx \\ &\text{s.t. } (\nabla y, \nabla v) = (u,v) \\ &u \in U_{ad} \end{aligned} \tag{37}$$

where $h_\xi(y;\gamma) \coloneqq h(y) + \gamma(\max(0, y-\xi))^3$, $\gamma \geq 0$.

Since we require $h_\xi(y;\gamma)$ to be twice differentiable in order to fulfil Assumption A.1), we choose $(\max(0, y-\xi))^3$. We assume that (37) admits a solution for any $\gamma \geq 0$. Analogously to (27), and the fact that $h_\xi(y;\gamma)$ is twice continuously differentiable with respect to y, we have that a solution to (37) is characterised by the PMP, see Theorem 2. Consequently, we can apply the SQH method to (37) and the results from Section 3 hold.

Henceforth, we use our SQH scheme to solve (37) for increasing γ denoted by $\gamma = \gamma_k$ for increasing k. Let (y_k, u_k) be a corresponding solution to (37) and $(\bar{y}, \bar{u})$ be a solution to (36). We show in the next theorem that increasing γ_k improves the solution to (37) with respect to the original task of solving the state-constrained optimal control problem (36). Specifically, we have that the measure of the violation of the state constraint by the corresponding solution goes to zero for increasing γ_k. Summarizing, we show that solving (37) with the SQH method provides a solution that fulfils the state constraint up to a tolerance depending on γ_k and results in a value $J(y_k, u_k)$ that is smaller than $J(\bar{y}, \bar{u})$. In addition if g is a square function, it can be proven that for increasing γ_k the sequence (y_k, u_k) converges to $(\bar{y}, \bar{u})$, see for example [28, Lemma 3.6] for details.

Theorem 11. *Let $\lim_{k\to\infty} \gamma_k = \infty$, let (y_k, u_k) be the corresponding solution to (37) with $M_k \coloneqq \{x \in \Omega |\ y_k(x) > \xi\}$ and let $(\bar{y}, \bar{u})$ be a solution to (36). Then, we have that $\lim_{k\to\infty} \int_{M_k} (y_k(x) - \xi)^3\,dx = 0$ and $J(y_k, u_k) = \int_\Omega h(y_k(x)) + g(u_k(x))\,dx \leq J(\bar{y}, \bar{u})$ for all $k \in \mathbb{N}$.*

Proof. First, the set M_k is measurable as y_k is measurable, see [5, X Theorem 1.9] for details. Thus integration over M_k is well defined. We have that

$$J(\bar{y}, \bar{u}) = \int_\Omega h(\bar{y}(x)) + g(\bar{u}(x)) + \gamma(\max(0, \bar{y}(x) - \xi))^3\,dx = J(\bar{y}, \bar{u}; \xi, \gamma),$$

as $\bar{y} \leq \xi$ and thus for an optimal solution (y_k, u_k) to (37) it holds that $J(y_k, u_k; \xi, \gamma) \leq J(\bar{y}, \bar{u})$. We equivalently have

$$\int_\Omega h(y_k(x)) + g(u_k(x)) + \gamma_k(\max(0, y_k(x) - \xi))^3\,dx \leq J(\bar{y}, \bar{u}). \tag{38}$$

Now if we assume that there is an $\epsilon > 0$ such that $\int_\Omega (\max(0, y_k(x) - \xi))^3\,dx = \int_{M_k} (y_k(x) - \xi)^3\,dx > \epsilon$ for all $k \in \mathbb{N}$, then we have a contradiction to (38) due to the lower boundedness of h and g. Also from (38) we have that

$$J(\bar{y}, \bar{u}) \geq \int_\Omega h(y_k(x)) + g(u_k(x)) + \gamma_k(\max(0, y_k(x) - \xi))^3\,dx \geq \int_\Omega h(y_k(x)) + g(u_k(x))\,dx.$$

□

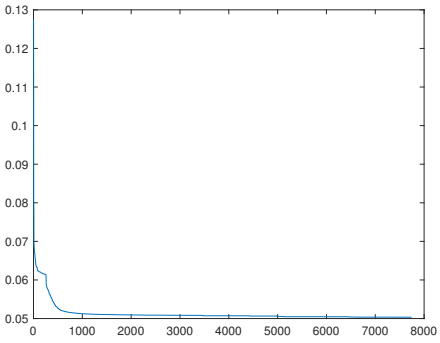
We remark that the arguments of the proof of Theorem 11 are not restricted to the elliptic optimal control problem (36) but also hold in the general framework of Section 2. That means that they also hold for a state constraint optimal control problem corresponding to (3).

For our numerical experiment, we choose the initial guess $\epsilon = \frac{1}{150}$ and $u^0 = 0$, $\sigma = 50$, $\zeta = \frac{3}{20}$, $\eta = 10^{-9}$, $\xi = \frac{3}{5}$, $\triangle x = \frac{1}{100}$, $\kappa = 10^{-10}$, $K_U = [-100, 100]$ and $y_d(x) \coloneqq \sin(2\pi x_1)\cos(2\pi x_2)$. The minimum of the augmented Hamiltonian K_ϵ in Step 2 of Algorithm 1 is determined pointwise with an exact formula as follows. The candidates at which a minimum of the augmented Hamiltonian is located are given by $u_1 = \min\left(\max\left(20, \frac{2\epsilon u^k(x) - p^k(x) - \beta}{\alpha + 2\epsilon}\right), 100\right)$, $u_2 = \min\left(\max\left(-100, \frac{2\epsilon u^k(x) - p^k(x) + \beta}{\alpha + 2\epsilon}\right), -20\right)$, $u_3 = \min\left(\max\left(-20, \frac{2\epsilon u^k(x) - p^k(x)}{\alpha + 2\epsilon}\right), 20\right)$. Consequently, the update for the control is pointwise given by $u(x) = \arg\min_{w \in \{u_1, u_2, u_3\}} K_\epsilon\left(x, y^k, w, u^k, p^k\right)$.

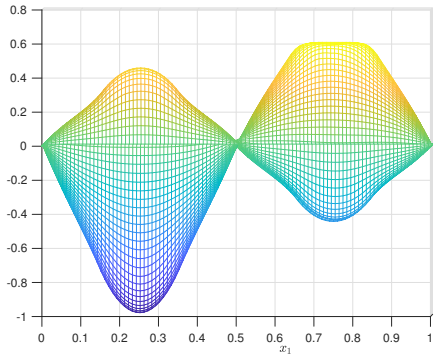
In Table 2, we show results that validate Theorem 11. We can see that for increasing γ the maximum of the state variable y converges to the upper bound of the state. Additionally, the measure of the set M_k where the state variable violates the upper bound becomes smaller when γ increases. According to Theorem 11 the quantity $\int_{M_k} \left(y_k(x) - \xi\right)^3 dx$ converges to zero for increasing γ. In Figure 7, we depict the state and control for $\gamma = 100000$.

γ	$\max_{x \in \Omega} y(x)$	$\int_{M_k} \left(y_k(x) - \frac{3}{5}\right)^3 dz$	$\lvert M_k \rvert$	$J\left(y, u; \frac{3}{5}\right)$
1	0.8218	$1.7843 \cdot 10^{-4}$	0.0461	0.0474
10	0.7125	$2.1182 \cdot 10^{-5}$	0.0383	0.0480
100	0.6543	$1.3580 \cdot 10^{-6}$	0.0289	0.0484
1000	0.6237	$7.9157 \cdot 10^{-8}$	0.0185	0.0487
10000	0.6081	$2.9827 \cdot 10^{-9}$	0.0137	0.0489
100000	0.6032	$1.1971 \cdot 10^{-10}$	0.0104	0.0503

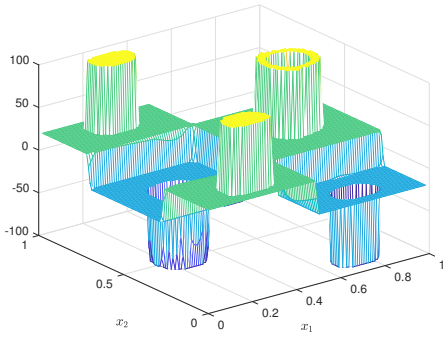
Table 2: Results that numerically validate Theorem 11 where (y, u) is obtained with the SQH method.



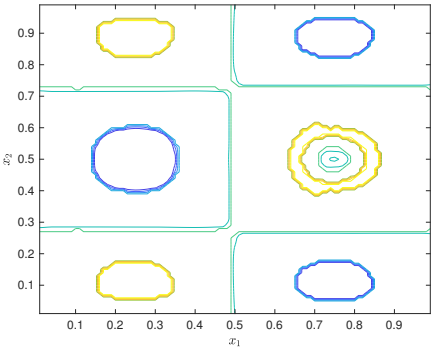
(a) Convergence history of the cost functional for the updates to the initial guess of the control.



(b) The state y from the side.



(c) The control u.



(d) The control u as a contour plot.

Figure 7: Solution to the optimal control problem (37) corresponding to P.7) for $\gamma = 100000$.

Similar as for the validation of the PMP optimality for P.1) to P.5), we have the following for the solution of (37) with $\gamma = 100000$. The inequality $H(x, t, \bar{y}, \bar{u}, \bar{p}) - \min_{w \in K_U} H(x, t, \bar{y}, w, \bar{p}) \leq 2.2 \cdot 10^{-16}$ is fulfilled at 6.45% of the grid points for $\kappa = 10^{-4}$, at 76.53% of the grid points for $\kappa = 10^{-8}$ and at 81.16% of the grid points for $\kappa = 10^{-12}$.

The results for P.7) are obtained with the code SQH_P7.zip:

https://opus.bibliothek.uni-wuerzburg.de/frontdoor/index/index/docId/16566

Conclusion

A sequential quadratic Hamiltonian (SQH) scheme for non-smooth and non-convex PDE optimal control problems was investigated considering seven different benchmark problems with increasing difficulty. For each problem, a theoretical discussion on the characterisation of optimality by the Pontryagin's maximum principle that is fundamental for the formulation of the SQH scheme was given and results of numerical experiments were presented that demonstrated the large range of applicability of this new method.

Acknowledgments

We would like to thank Daniel Steck and Daniel Wachsmuth for helpful comments concerning the Ekeland's variational principle. For the same reason and for many constructive remarks, we are grateful to the anonymous Reviewer and to Otmar Scherzer. The continued support of Richard Greiner is gratefully acknowledged.

A Proof of Lemma 1

In the following, we present the proof of Lemma 4.

Proof. We define $(\delta u)^2 := \sum_{j=1}^{m} (\delta u_j(z))^2$, $\delta y := y(z) - y^k(z)$ and $\delta p := p(z) - p^k(z)$ where p is calculated by (4) for y and u. Furthermore, to save notational effort, we note $H := H(z, y, u, p)$ or $H^k := H\left(z, y^k, u^k, p^k\right)$ and drop the functional dependency of the functions y, y^k, u, u^k, p and p^k as well as $f := f(z, y, u)$, $f^k := f\left(z, y^k, u^k\right)$, $h^k := h\left(y^k\right)$ and $h := h(y)$ for all $k \in \mathbb{N}_0$. We use from Algorithm 1 that u is determined such that $K_\epsilon\left(z, y^k, u, u^k, p^k\right) \leq K_\epsilon\left(z, y^k, w, u^k, p^k\right)$ for all $w \in K_U$ and thus especially

$$K_\epsilon\left(z, y^k, u, u^k, p^k\right) \leq K_\epsilon\left(z, y^k, u^k, u^k, p^k\right) = H\left(z, y^k, u^k, p^k\right),$$

for almost all $z \in Z_i$. We start the proof as follows

$$\begin{aligned}
&J(y,u) - J\left(y^k, u^k\right) = \int_{Z_i} h(y) + g(u) - h\left(y^k\right) - g\left(u^k\right) dz = \int_{Z_i} H - pf - H^k + p^k f^k dz \\
&= \int_{Z_i} H - H\left(z, y^k, u, p^k\right) + H\left(z, y^k, u, p^k\right) + \epsilon(\delta u)^2 - H^k - \epsilon(\delta u)^2 dz \\
&- \int_{Z_i} \delta p f dz - \int_0^T \left((\delta y)', p^k\right) + B\left(\delta y, p^k; t\right) dt \\
&\leq \int_{Z_i} H - H\left(z, y^k, u, p^k\right) - \epsilon(\delta u)^2 dz - \int_{Z_i} \delta p f dz - \int_0^T \left((\delta y)', p^k\right) + B\left(\delta y, p^k; t\right) dt.
\end{aligned} \tag{39}$$

Next, we estimate the term $|\int_{Z_i} H - H\left(z, y^k, u, p^k\right) dz - \int_{Z_i} \delta p f dz - \int_0^T \left((\delta y)', p^k\right) + B\left(\delta y, p^k; t\right) dt|$. Using the Taylor formula [4, Chapter VII, Theorem 5.8] and with the symmetry of the second derivative

[4, Chapter VII, Theorem 5.2], we obtain

$$
\begin{aligned}
&|\int_{Z_i} H - H\left(z, y^k, u, p^k\right) dz - \int_{Z_i} \delta p f dz - \int_0^T \left((\delta y)', p^k\right) + B\left(\delta y, p^k; t\right) dt| \\
&= |\int_{Z_i} H - H\left(z, y - \delta y, u, p - \delta p\right) dz - \int_{Z_i} \delta p f dz - \int_0^T \left((\delta y)', p^k\right) + B\left(\delta y, p^k; t\right) dt| \\
&= |\int_{Z_i} \partial_y H \delta y + \partial_p H \partial p dz - \frac{1}{2}\int_{Z_i} \partial_{yy} H \left(\delta y\right)^2 + 2\partial_{yp} H \delta y \delta p dz + \int_{Z_i} R_2\left(H, y, p; \delta y, \delta p\right) dz \\
&- \int_{Z_i} \delta p f dz - \int_0^T \left((\delta y)', p^k\right) + B\left(\delta y, p^k; t\right) dt| \\
&= |\int_{Z_i} \partial_y h \delta y + p \partial_y f \delta y + f \delta p dz - \frac{1}{2}\int_{Z_i} \left(\partial_{yy} h + p \partial_{yy} f\right)\left(\delta y\right)^2 + 2\partial_y f \delta y \delta p dz \\
&+ \int_{Z_i} R_2\left(H, y, p; \delta y, \delta p\right) dz - \int_{Z_i} \delta p f dz - \int_0^T \left((\delta y)', p^k\right) + B\left(\delta y, p^k; t\right) dt| \\
&= |\int_0^T - \left((p)', \delta y\right) + B^*\left(p, \delta y; t\right) dt - \int_0^T \left((\delta y)', p^k\right) + B\left(\delta y, p^k; t\right) dt \\
&- \frac{1}{2}\int_{Z_i} \left(\partial_{yy} h + p \partial_{yy} f\right)\left(\delta y\right)^2 + 2\partial_y f \delta y \delta p dz + \int_{Z_i} R_2\left(H, y, p; \delta y, \delta p\right) dz| \\
&\leq \theta \|\delta u\|^2_{L^2(Z_i)},
\end{aligned}
\tag{40}
$$

where we use the partial integration rule [45, Theorem 3.11], A.2), A.5) and A.3), the linearity of H in the p-argument, the boundedness of p if $\partial_{yy} f \neq 0$ (see A.4)), the Cauchy-Schwarz inequality [2, Lemma 2.2] in the last inequality for the term $\int_{Z_i} 2\partial_y f \delta y \delta p dz$ and that the Taylor remainder $R_2\left(H, y, p; \delta y, \delta p\right)$ is estimated by the remainder formula [4, Chapter VII, Theorem 5.8] and the boundedness of the second derivatives analogously to the calculation which are done for the second derivatives in (40). Combining (39) and (40), we obtain

$$
\begin{aligned}
&J\left(y, u\right) - J\left(y^k, u^k\right) \\
&\leq \int_{Z_i} H - H\left(z, y^k, u, p^k\right) - \epsilon\left(\delta u\right)^2 dz - \int_{Z_i} \delta p f dz - \int_0^T \left(\delta y', p^k\right) + B\left(\delta y, p^k\right) dt \\
&\leq |\int_{Z_i} H - H\left(z, y^k, u, p^k\right) dz - \int_{Z_i} \delta p f dz - \int_0^T \left(\delta y', p^k\right) + B\left(\delta y, p^k\right) dt| - \int_{Z_i} \epsilon\left(\delta u\right)^2 dz \\
&\leq \theta \|\delta u\|^2_{L^2(Z_i)} - \int_{Z_i} \epsilon\left(\delta u\right)^2 dz = \left(\theta - \epsilon\right)\|\delta u\|^2_{L^2(Z_i)}.
\end{aligned}
$$

□

B L^∞ estimates

In the PMP framework including the numerical treatment with the SQH method, the $L^\infty\left(\Omega\right)$ boundedness of the solution of the corresponding PDE is crucial. Therefore, we provide the corresponding results for the PDEs of our examples P.1) to P.7). We first consider the elliptic case for an open and bounded domain Ω. We have the following

$$
\begin{aligned}
B\left(y, v\right) &= \left(h, v\right) \text{ in } \Omega \\
y &= 0 \text{ on } \partial\Omega
\end{aligned}
\tag{41}
$$

where $B(y,v) : H_0^1 \times H_0^1 \to \mathbb{R}$ is a bilinear map with the coercivity condition $\beta\|y\|^2_{H_0^1(\Omega)} \leq B(y,y)$, $\beta > 0$ and $B(-k,v) \leq 0$ for $k \geq 0$ if $v \geq 0$ and $h \in L^q(\Omega)$, $q \geq \frac{n}{2} + 1$. We assume that (41) has a unique solution $y \in H_0^1(\Omega)$. Then the following theorem holds.

Theorem 12. *The initial value problem (41) has an essential bounded solution for which holds*

$$\|y\|_{L^\infty(\Omega)} \leq C\|h\|_{L^q(\Omega)},$$

where $C > 0$.

Proof. The proof is based on [48, Theorem 4.2.1]. We assume that h is not the zero function. In the case of $h = 0$, the solution $y = 0$ solves (41) and thus the statement is true. We choose the constant $k \geq 0$. As $y - k \in H^1(\Omega)$, we have that $(y-k)_+ \coloneqq \max(y-k, 0) \in H_0^1(\Omega)$, see [19, Chapter 4, Proposition 6]. Then, we choose $v = (y-k)_+$ in (41) and obtain the following $B\left(y-k, (y-k)_+\right) \leq \left(h, (y-k)_+\right)$ where we use that

$$B\left(y, (y-k)_+\right) \geq B\left(y, (y-k)_+\right) + B\left(-k, (y-k)_+\right) = B\left(y-k, (y-k)_+\right)$$

and thus

$$\beta\| (y-k)_+ \|^2_{H_0^1(\Omega)} \leq \left(h, (y-k)_+\right) \tag{42}$$

as $(y-k)_+ = 0$ if $y - k \leq 0$ and $B\left(y-k, (y-k)_+\right) = B\left((y-k)_+, (y-k)_+\right)$ if $y - k > 0$ and $\beta\| (y-k)_+ \|^2_{H_0^1(\Omega)} \leq B\left((y-k)_+, (y-k)_+\right)$. We remark that the function $(y-k)_+ \in H_0^1(\Omega)$ is also an element of $L^p(\Omega)$ with $\| (y-k)_+ \|_{L^p(\Omega)} \leq M\| (y-k)_+ \|_{H_0^1(\Omega)}$, $M > 0$ where

$$2 \leq p \begin{cases} \leq \infty & \text{for } n = 1 \\ < \infty & \text{for } n = 2 \\ \leq \frac{2n}{n-2} & \text{for } n \geq 3 \end{cases} \tag{43}$$

, see the Sobolev embedding theorem [1, Theorem 4.12], especially [1, Theorem 4.12 Part III]. This implies

$$\| (y-k)_+ \|^2_{L^p(\Omega)} \leq \tilde{\beta} \int_\Omega h(x) (y-k)_+ (x)\, dx \tag{44}$$

with $\tilde{\beta} > 0$. Next, we define $A_k \coloneqq \{x \in \Omega |\ y(x) > k\}$ which is measurable, see [5, X, Theorem 1.9] and $|A_k(t)|$ is the measure of $A_k(t)$. Due to $(y(x) - k)_+ = 0$ for $x \in \Omega \backslash A_k$, we consequently have from (44) the following

$$\| (y-k)_+ \|^2_{L^p(A_k)} \leq \tilde{\beta} \int_{A_k} h(x) (y-k)_+ (x)\, dx. \tag{45}$$

In the next step, we have the estimate by Hölder's inequality see, [25, page 622]

$$\| (y-k)_+ \|^2_{L^p(A_k)} \leq \tilde{\beta} \left(\int_{A_k} |h(x)|^{\frac{n}{2}+1} dx \right)^{\frac{1}{\frac{n}{2}+1}} \left(\int_{A_k} (y-k)_+^{\frac{2+n}{n}} (x)\, dx \right)^{\frac{n}{2+n}}$$

which can be applied as $(y-k)_+ \in L^{\frac{n+2}{n}}(\Omega)$. This is true because in the case $n = 1$ and $n = 2$, we have $(y-k)_+ \in L^p$, $2 \leq p < \infty$ and in the case $n \geq 3$, we have $\frac{2n}{n-2} \geq \frac{2+n}{n}$ equivalently $n^2 \geq -4$, [1, Theorem 2.14], and consequently

$$\| (y-k)_+ \|^2_{L^p(A_k)} \leq \tilde{\beta} \|h\|_{L^{\frac{n}{2}+1}(\Omega)} \left(\int_{A_k} (y-k)_+^{\frac{2+n}{n}} (x)\, dx \right)^{\frac{n}{2+n}}. \tag{46}$$

We apply Hölder's inequality again with $\frac{1}{\tilde{p}}+\frac{1}{\tilde{q}}=1$, thus for a given $\tilde{p}$ we have $\tilde{q}=\frac{\tilde{p}}{\tilde{p}-1}$, and we obtain the following

$$\| (y-k)_+ \|^2_{L^p(A_k)} \leq \tilde{\beta}\|h\|_{L^{\frac{n}{2}+1}(\Omega)} \left(\int_{A_k} 1dx\right)^{\frac{\tilde{p}-1}{\tilde{p}}\frac{n}{2+n}} \left(\int_{A_k} (y-k)_+^{\frac{2+n}{n}\tilde{p}}(x)\,dx\right)^{\frac{n}{\tilde{p}(2+n)}} \tag{47}$$

We choose $p=\frac{2+n}{n}\tilde{p}$ and conclude from (47) for $\| (y-k)_+ \|_{L^{\frac{2+n}{n}\tilde{p}}(A_k)} > 0$ the following

$$\left(\int_{A_k} |(y-k)_+(x)|^{\frac{2+n}{n}\tilde{p}}dx\right)^{\frac{n}{\tilde{p}(2+n)}} = \| (y-k)_+ \|_{L^{\frac{2+n}{n}\tilde{p}}(A_k)} \leq \tilde{\beta}\|h\|_{L^{\frac{n}{2}+1}(\Omega)} \left(\int_{A_k} 1dx\right)^{\frac{\tilde{p}-1}{\tilde{p}}\frac{n}{2+n}}, \tag{48}$$

which is also true in the case $\| (y-k)_+ \|_{L^{\frac{2n}{n-2}}(A_k)} = 0$.

Furthermore, for $m > k$, we have that $A_m \subseteq A_k$. Additionally it is $y > m$ on A_m and thus $y \geq y-k > m-k$ on A_m due to $k \geq 0$. Since $y-k=(y-k)_+$ on A_m, we obtain

$$\int_{A_k} |(y-k)_+(x)|^{\frac{2+n}{n}\tilde{p}}dx \geq \int_{A_m} (y-k)^{\frac{2+n}{n}\tilde{p}}(x)\,dx \geq (m-k)^{\frac{2+n}{n}\tilde{p}} \int_{A_m} 1dx. \tag{49}$$

We combine (49) with (48) and obtain $(m-k)\,|A_m|^{\frac{n}{\tilde{p}(2+n)}} \leq \tilde{\beta}\|h\|_{L^{\frac{n}{2}+1}(\Omega)}|A_k|^{\frac{\tilde{p}-1}{\tilde{p}}\frac{n}{2+n}}$ and equivalently

$$|A_m| \leq \left(\frac{\tilde{\beta}\|h\|_{L^{\frac{n}{2}+1}(\Omega)}}{m-k}\right)^{\frac{2+n}{n}\tilde{p}} |A_k|^{\tilde{p}-1}. \tag{50}$$

In order to apply [48, Lemma 4.1.1], we need that $\tilde{p}-1>1$ and that p fulfils (43). For the case $n=1$ and $n=2$, we can choose any $\tilde{p}>2$, for example $\tilde{p}=3$. For the case that $n \geq 3$, we have to ensure that $\frac{2+n}{n}\tilde{p} \leq \frac{2n}{n-2}$, which is $\tilde{p} \leq \frac{2n^2}{n^2-4}$. Since the expression $\frac{2n^2}{n^2-4} > 2$ for $n \geq 3$ is equivalent to $0 > -8$ and thus always true, we can choose $\tilde{p}=\frac{2n^2}{n^2-4}$ in the case of $n \geq 3$. Then we also have that $\frac{2+n}{n}\tilde{p} > 0$. By applying [48, Lemma 4.1.1], we obtain that $|A_m|=0$ for $m \geq \tilde{\beta}\|h\|_{L^{\frac{n}{2}+1}(\Omega)}2^{\frac{\tilde{p}-1}{\tilde{p}-2}}|\Omega|^{\frac{n(\tilde{p}-2)}{\tilde{p}(2+n)}}$ where $|\Omega|$ is the measure of Ω. This means that the set where

$$y > \tilde{\beta}\|h\|_{L^{\frac{n}{2}+1}(\Omega)}2^{\frac{\tilde{p}-1}{\tilde{p}-2}}|\Omega|^{\frac{n(\tilde{p}-2)}{\tilde{p}(2+n)}}$$

is of measure zero. With the same arguments, we have for $(y+k)_- := \min(y+k,0)$ and $A_k := \{x \in \Omega |\ y < -k\}$ that the set where $y < -\tilde{\beta}\|h\|_{L^{\frac{n}{2}+1}(\Omega)}2^{\frac{\tilde{p}-1}{\tilde{p}-2}}|\Omega|^{\frac{n(\tilde{p}-2)}{\tilde{p}(2+n)}}$ is of measure zero. Therefore, we obtain that $\|y\|_{L^\infty(\Omega)} \leq C\|h\|_{L^{\frac{n}{2}+1}(\Omega)}$ with $C := \tilde{\beta}2^{\frac{\tilde{p}-1}{\tilde{p}-2}}|\Omega|^{\frac{n(\tilde{p}-2)}{\tilde{p}(2+n)}}$. □

Next, we check that a similar $L^\infty(Z_i)$ result holds for our examples. For Example P.1) and P.7), Theorem 12 holds immediately. For Example P.3) this result holds also immediately if we assume $K_U \subseteq \mathbb{R}_0^+$ because then we have that $-ukv \leq 0$ for $v \geq 0$ and we can continue the proof of Theorem 12 from (42) to obtain the following theorem.

Theorem 13. *For the solution y of an elliptic boundary value problem as in P.3) with $K_U \subseteq \mathbb{R}_0^+$, we have*

$$\|y\|_{L^\infty(\Omega)} \leq d\|\tilde{f}\|_{L^q(\Omega)},$$

with $d>0$ for any right-hand side $\tilde{f} \in L^q(\Omega)$.

For P.5), we have the following consideration such that the proof of Theorem 12 can be followed from (42). We have that $\left(\nabla y, \nabla (y-k)_+\right) \leq \left(\nabla y, \nabla (y-k)_+\right) + \left(y^3, (y-k)_+\right)$ as $y^3 (y-k)_+ \geq 0$ due to $(y-k)_+ = 0$ if $y \leq k \geq 0$. With similar considerations the $L^\infty(\Omega)$-result is proved for P.6) as $\max(0, y) \geq 0$.

For the parabolic case, i.e. P.2) and P.4), we have a result analogous to Theorem 12 proved in [12, Appendix]. With this result, a similar $L^\infty(Q)$-result to Theorem 13 holds immediately, assuming that $K_U \subseteq \mathbb{R}_0^+$, in the bilinear case, i.e. P.4).

C Discussion of admissible sets and PMP optimal solutions

It is well-known that existence of a solution to an optimizationl problem with a lower semi-continuous cost functional can be proven considering an admissible set that is compact in the optimization space; see [12, A.2]. However, in order to apply the technique of needle variation, it is necessary that all needle variations of all admissible controls, which means the needle variation of any admissible control at any point of the domain and any radius of the ball, are included into the admissible set. But the compactness of the admissible set does not allow the accommodation of all needle variations of all admissible controls.

A similar problem is faced considering the case where H^1 is compactly embedded into a L^p space. In fact, the single needle variations in general are not elements of H^1. Moreover, as shown in this paper and also in [47], optimal controls can be obtained that are discontinuous functions, which do not belong to a H^1 space. These results show that it is not reasonable to elaborate a proof of existence of an optimal solution over an admissible set that is a subset of H^1. Furthermore, in [47], an 'unsolvable problem' is constructed that has no PMP optimal solution over the admissible set that consists of pointwise bounded L^2 functions.

On the other hand, our numerical results indicate that the SQH scheme converges to functions that are PMP optimal. Consequently, we think that a proof of existence of an optimal solution over an appropriately chosen admissible set that also allows the PMP characterization might be successful.

References

[1] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces*, volume 140 of *Pure and applied mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003.

[2] H. W. Alt. *Linear Functional Analysis: An Application-Oriented Introduction*. Springer, 2016.

[3] H. Amann and J. Escher. *Analysis I*. Birkhäuser Basel, 2006.

[4] H. Amann and J. Escher. *Analysis II*. Birkhäuser Basel, 2008.

[5] H. Amann and J. Escher. *Analysis III*. Birkhäuser Basel, 2009.

[6] A. Ambrosetti and R. E. Turner. Some discontinuous variational problems. *Diff. & Integral Equat*, 1:341–349, 1988.

[7] L. Ambrosio, G. Da Prato, and A. C. Mennucci. *Introduction to Measure Theory and Integration*. Edizioni della Normale, 2011.

[8] H. H. Bauschke and P. L. Combettes. Convex analysis and monotone operator theory in hilbert spaces. 2017.

[9] T. Bayen and F. J. Silva. Second order analysis for strong solutions in the optimal control of parabolic equations. *SIAM Journal on Control and Optimization*, 54(2):819–844, 2016.

[10] J. F. Bonnans. On an algorithm for optimal control using Pontryagin's maximum principle. *SIAM Journal on Control and Optimization*, 24(3):579–588, 1986.

[11] A. Borzì and K. Kunisch. The numerical solution of the steady state solid fuel ignition model and its optimal control. *SIAM Journal on Scientific Computing*, 22(1):263–284, 2000.

[12] T. Breitenbach and A. Borzì. A sequential quadratic Hamiltonian method for solving parabolic optimal control problems with discontinuous cost functionals. *Journal of Dynamical and Control Systems*, pages 1–33, 2018.

[13] R. Bulirsch, E. Nerz, H. J. Pesch, and O. von Stryk. *Combining Direct and Indirect Methods in Optimal Control: Range Maximization of a Hang Glider*, pages 273–288. Birkhäuser Basel, Basel, 1993.

[14] E. Casas. Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations. *SIAM Journal on Control and Optimization*, 35(4):1297–1327, 1997.

[15] E. Casas. Second order analysis for bang-bang control problems of pdes. *SIAM Journal on Control and Optimization*, 50(4):2355–2372, 2012.

[16] C. Christof, C. Meyer, S. Walther, and C. Clason. Optimal control of a non-smooth semilinear elliptic equation. *Mathematical Control & Related Fields*, 8(1):247–276, 2018.

[17] F. Clarke. *Functional analysis, calculus of variations and optimal control*, volume 264. Springer Science & Business Media, 2013.

[18] D. L. Cohn. *Measure theory*. Springer, 2013.

[19] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Vol. 2*. Springer-Verlag, Berlin, 1988. Functional and variational methods, With the collaboration of Michel Artola, Marc Authier, Philippe Bénilan, Michel Cessenat, Jean Michel Combes, Hélène Lanchon, Bertrand Mercier, Claude Wild and Claude Zuily, Translated from the French by Ian N. Sneddon.

[20] D. Devadze and V. Beridze. Methods of numerical solution of optimal control problems based on the Pontryagin maximum principle. *Journal of Mathematical Sciences*, 206(4):348–356, 2015.

[21] A. V. Dmitruk. On the development of Pontryagin's maximum principle in the works of A. Ya. Dubovitskii and A.A. Milyutin. *Control and Cybernetics*, 38(4A):923–957, 2009.

[22] A. V. Dmitruk and N. P. Osmolovskii. On the proof of Pontryagin's maximum principle by means of needle variations. *Journal of Mathematical Sciences*, 218(5):581–598, 2016.

[23] C. M. Dunn. *Introduction to Analysis*. CRC Press, 2017.

[24] I. Ekeland. On the variational principle. *Journal of Mathematical Analysis and Applications*, 47(2):324–353, 1974.

[25] L. C. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 1998.

[26] A. Hamel. Suboptimality theorems in optimal control. In W. H. Schmidt, K. Heier, L. Bittner, and R. Bulirsch, editors, *Variational Calculus, Optimal Control and Applications*, pages 61–68, Basel, 1998. Birkhäuser Basel.

[27] K. Ito and K. Kunisch. Optimal control with $L^p(\Omega)$, $p \in (0, 1)$, control cost. *SIAM Journal on Control and Optimization*, 52(2):1251–1275, 2014.

[28] V. Karl and D. Wachsmuth. An augmented Lagrange method for elliptic state constrained optimal control problems. *Computational Optimization and Applications*, 69(3):857–880, 2018.

[29] R. V. Kohn and G. Strang. Optimal design and relaxation of variational problems, i. *Communications on Pure and Applied Mathematics*, 39(1):113–137, 1986.

[30] I. A. Krylov and F. L. Chernous'ko. On a method of successive approximations for the solution of problems of optimal control. *USSR Computational Mathematics and Mathematical Physics*, 2(6):1371–1382, 1963.

[31] I. A. Krylov and F. L. Chernous'ko. An algorithm for the method of successive approximations in optimal control problems. *USSR Computational Mathematics and Mathematical Physics*, 12(1):15–38, 1972.

[32] X. Li and J. Yong. *Optimal Control Theory for Infinite Dimensional Systems*. Optimal Control Theory for Infinite Dimensional Systems. Birkhäuser, 1995.

[33] A. A. Lyubushin. Modifications of the method of successive approximations for solving optimal control problems. *USSR Computational Mathematics and Mathematical Physics*, 22(1):29–34, 1982.

[34] M. S. Nikol'ski. On the convergence problem of a certain variant of the successive-approximation method for solving optimal control problems. *Journal of Mathematical Sciences*, 139(5):6902–6908, 2006.

[35] C. C. Pugh. *Real Mathematical Analysis*. Undergraduate Texts in Mathematics. Springer, Cham, second edition, 2015.

[36] V. D. Radulescu and D. D. Repovs. *Partial differential equations with variable exponents: variational methods and qualitative analysis*. Chapman and Hall/CRC, 2015.

[37] J.-P. Raymond and H. Zidani. Pontryagin's principle for state-constrained control problems governed by parabolic equations with unbounded controls. *SIAM Journal on Control and Optimization*, 36(6):1853–1879, 1998.

[38] J.-P. Raymond and H. Zidani. Hamiltonian Pontryagin's principles for control problems governed by semilinear parabolic equations. *Applied Mathematics and Optimization*, 39(2):143–177, 1999.

[39] G. Rozvany. Optimal plastic design with discontinuous cost functions. *Journal of Applied Mechanics*, 41(1):309–310, 1974.

[40] Y. Sakawa and Y. Shindo. On global convergence of an algorithm for optimal control. *IEEE Transactions on Automatic Control*, 25(6):1149–1153, 1980.

[41] E. Scheiber. Chernousko-Lyubushin version of the succesive approximation method for optimal control problem revisited. *Bulletin of the Transilvania University of Brasov Vol*, 6(55), 2013.

[42] Y. Shindo and Y. Sakawa. Local convergence of an algorithm for solving optimal control problems. *Journal of optimization theory and applications*, 46(3):265–293, 1985.

[43] I. Shvartsman and Z. Mingaleeva. Second-order optimality conditions for singular Pontryagin local minimizers. *Numerical Functional Analysis and Optimization*, 35(7-9):1245–1257, 2014.

[44] M. I. Sumin. The first variation and Pontryagin's maximum principle in optimal control for partial differential equations. *Computational Mathematics and Mathematical Physics*, 49(6):958–978, 2009.

[45] F. Tröltzsch. *Optimal Control of Partial Differential Equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. Theory, methods and applications.

[46] O. V. Vasiliev and A. I. Tyatyushkin. A method for solving optimal control problems that is based on the maximum principle. *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, 21(6):1376–1384, 1981.

[47] D. Wachsmuth. Iterative hard-thresholding applied to optimal control problems with $L^0(\Omega)$ control cost. *SIAM Journal on Control and Optimization*, 57(2):854–879, 2019.

[48] Z. Wu, J. Yin, and C. Wang. *Elliptic & Parabolic Equations*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2006.