ESAIM: COCV 30 (2024) 68
https://doi.org/10.1051/cocv/2024052

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

LOCAL QUADRATIC CONVERGENCE OF THE SQP METHOD FOR AN OPTIMAL CONTROL PROBLEM GOVERNED BY A REGULARIZED FRACTURE PROPAGATION MODEL

Andreas Hehl[1] and Ira Neitzel[1,*]

Abstract. We prove local quadratic convergence of the sequential quadratic programming (SQP) method for an optimal control problem of tracking type governed by one time step of the Euler-Lagrange equation of a time discrete regularized fracture or damage energy minimization problem. This lower-level energy minimization problem contains a penalization term for violation of the irreversibility condition in the fracture growth process and a viscous regularization term. Conditions on the latter, corresponding to a time step restriction, guarantee strict convexity and thus unique solvability of the Euler Lagrange equations. Nonetheless, these are quasilinear and the control problem is nonconvex. For the convergence proof with L^∞ localization of the SQP-method, we follow the approach from Tröltzsch [*SIAM J. Control Optim.* **38** (1999) 294–312], utilizing strong regularity of generalized equations and arguments from Hoppe and Neitzel [*Optim. Eng.* **22** (2021)] for L^2-localization.

Mathematics Subject Classification. 90C55, 49M41, 49M15.

Received July 29, 2023. Accepted July 12, 2024.

1. Introduction

In this work, we analyze the convergence behavior of the sequential quadratic programming (SQP) method applied to an optimal control problem for regularized fracture propagation including control constraints. The model problem is the same as in *e.g.* [1], and closely related to [2, 3]. It is of tracking-type, with a control q in a control set Q_{ad} acting as a boundary force, and associated state pair $\mathbf{u} = (u, \varphi)$ in a state space V consisting of a displacement u and a phase-field φ. As in *e.g.* [1] we consider a time discrete but spatially continuous model problem and for better readability restrict the presentation to one time step. In short notation, the problem reads as follows:

$$\begin{cases} \min\limits_{q\in Q_{\text{ad}},\, \mathbf{u}\in V} \quad J(q,\mathbf{u}) := \dfrac{1}{2}\|u-u_d\|^2_{L^2(\Omega,\mathbb{R}^2)} + \dfrac{\alpha}{2}\|q\|^2_{L^2(\Gamma)}, \\ \text{subject to :} \quad A(\mathbf{u}) + R(\varphi;\gamma) = Bq. \qquad (\mathrm{EL}^{\gamma,\eta}) \end{cases} \tag{NLP$^{\gamma,\eta}$}$$

Keywords and phrases: Optimal control, phase-field regularized fracture model, SQP method, local quadratic convergence.

[1] Institut für Numerische Simulation, Rheinische Friedrich-Wilhelms-Universität Bonn, Friedrich-Hirzebruch Allee 7, 53115 Bonn, Germany.

[*] Corresponding author: neitzel@ins.uni-bonn.de

© The authors. Published by EDP Sciences, SMAI 2024

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We will define the precise mathematical setting, including the operators A, R, and B in Section 2 below. The problem originates from a bi-level optimization problem with an upper-level optimal control problem subject to a lower-level energy minimization problem for variational fracture propagation. As mentioned above, we consider one time step of a time discrete and spatially continuous problem, where a regularized version of an energy minimization problem describing the lower-level fracture propagation is eventually replaced by its Euler-Lagrange equations. A nonregularized version of the fracture propagation problem has originally been considered in [4–6]. To avoid the irregular fracture set, an Ambrosio-Tortorelli regularization *cf.* [7] is used; *i.e.* an additional phase-field variable φ is introduced to replace the irregular Hausdorff measure. The phase-field variable φ has values in $[0, 1]$, and describes the condition of the material at every point in the domain, with $\varphi = 1$ where the material is completely sound, and $\varphi = 0$ where the material is fully broken, guaranteeing a smooth transition between those two states. Moreover, we apply a viscous regularization to guarantee strict convexity of the lower-level minimization problem; imposed conditions correspond to a reasonable time step restriction, see [8] for details.

As a consequence, the Euler-Lagrange equations are uniquely solvable, but remain quasilinear, making the overall control problem nonconvex. Finally, a violation of the irreversibility condition in the fracture growth process is penalized using a regularization with parameter γ in form of a (higher-order) penalization, as in [9]. The corresponding terms appear in the operator R in the differential equation, whereas the differential operator A stems from the actual (regularized) fracture propagation process. For a more detailed description of the mathematical and physical background of ($\mathrm{NLP}^{\gamma,\eta}$), we refer to the introductions of [1, 2]. We point out that for the purpose of this paper all mentioned regularization parameters remain fixed.

Let us give a brief summary of the current state of research for the model problem ($\mathrm{NLP}^{\gamma,\eta}$): Existence of solutions and first-order necessary optimality conditions for the model problem, under an additional trivial kernel assumption, but without a viscous approximation and control constraints, have been proven in [2]. In [3], convergence of regularized solutions with respect to γ was proven. Subsequently, convergence (w.r.t γ) of the dual variables was established in [10]. Finite element discretization error estimates have been derived for a linearized fracture control problem in [11], and algorithmic concepts, respectively the space-time formulation and time discretization, were studied in [12, 13]. Further, in [10] the sequential quadratic programming (SQP) method for ($\mathrm{NLP}^{\gamma,\eta}$) was described, and a preliminary analysis of the underlying quadratic subproblem was made, under an additional rather strong local coercivity condition, *cf.* Section 4. This is the starting point of our analysis. Utilizing second-order sufficient conditions we carry out a rigorous convergence analysis. We can rely on the results from [1], where we investigated second-order necessary and second-order sufficient conditions (SSC) with minimal gap and without two-norm discrepancy. It is well known that SSCs are commonly the basis for convergence proofs of the SQP method. For an introduction we refer the interested reader to *e.g.* the introduction of [14]. For SQP of control constrained problems governed by semilinear elliptic and parabolic equations we refer to [14–18], for semilinear problems with mixed control-state-constraints to [19, 20], and for the Navier-Stokes equation to [21–24]. In [25, 26], the SQP method for the optimal control of a (semilinear) phase-field equation was considered. Rather recently, convergence of the SQP method for an optimal control problem with quasilinear parabolic optimal equation was proven in [27]. The latter publication used L^2-type localization arguments instead of L^∞-closeness from earlier publications.

We will continue the work established in [1, 10], and analyze the SQP method applied to the quasilinear fracture control problem ($\mathrm{NLP}^{\gamma,\eta}$), utilizing the typical procedure of proving convergence of SQP methods in infinite dimensional spaces that goes back to [28]. We follow the ideas of *e.g.* [16, 20] and apply Newton's method to a generalized equation that corresponds to the necessary optimality conditions of the model problem. A Newton-Kantorovich like convergence theorem, *cf.* [28–30] will then ensure local quadratic convergence of the generated sequence. This theorem relies in particular on the so-called strong regularity property, *cf.* [31], which allows to generalize the implicit function theorem to generalized equations. It was later used to show convergence of Newton's method in the context of (unconstrained) optimal control in Banach spaces in [32]. Ensuring this strong regularity property and additional *e.g.* Lipschitz results for our model problem requires a careful, nontrivial analysis. We benefit from results that we have proven in the context of SSC in [1], since strong regularity is closely related to second-order sufficient conditions (SSCs). Let us therefore briefly comment on different types of second-order optimality conditions. On the one hand, it is preferable to keep the gap between the necessary and sufficient conditions as close as possible, which leads to SSC incorporating so-called strongly active constraints, *cf.* [33, 34]. We have established such a result for (NLP$^{\gamma,\eta}$) in [1]. Yet, this only ensures coercivity on some subspace of Q and it is not apriori clear that the directions generated by the SQP method belong to this subset. On the other hand, choosing SSC on the whole control space Q is a very strong assumption. In this work, we will use so called σ-strongly active constraints, see *e.g.* [16]. Following the ideas of [16], we will establish convergence of the SQP method for certain auxiliary quadratic subproblems that take into account the definition of the cone of critical directions. In a second step, we will show that the solutions of the auxiliary problems in fact correspond to the solution of the SQP method for (NLP$^{\gamma,\eta}$), restricted to an L^∞ neighborhood of a locally optimal solution. In a third and last step, we use L^2-neighborhoods as in [27].

To put our work into perspective, let us mention some further approaches related to fracture optimization. In [35, 36] control of a viscous damage model was considered in a continuous setting, shape optimization techniques were utilized in [37, 38], and an approach where the propagation of a crack was limited through controlling the release of the associated energy was used in [39, 40]. An optimal control problem of a two-field damage model, and a nonsmooth (viscous damage) coupled system, was analyzed in [41, 42]. For results concerning the lower-level fracture problem we refer to [43], where modelling and numerical analysis of multiphysics phase-field fracture models are addressed. Phase-field models are also applicable in other fields, like material science, medical applications and image segmentation. For the former, we refer to [44–46]. In the context of tumor growth, phase-field models have been used in *e.g.* [47, 48]. For the latter, the analysis of the Mumford-Shah image segmentation functional [49] through phase-field methods [7, 50, 51] is still an active field of research, see *e.g.* [52]. For an overview about numerical implementation of phase-field models, we refer to [53].

The outline of the present work is as follows: We start with a detailed description of the problem setting including all assumptions on the model problem and notational conventions in Section 2. In Section 3, we collect regularity and existence results for (EL$^{\gamma,\eta}$) and results on solvability as well as necessary and sufficient optimality conditions for (NLP$^{\gamma,\eta}$), respectively, and provide some auxiliary results. In Section 4, we describe the SQP method for (NLP$^{\gamma,\eta}$), and start with some preliminary considerations about the quadratic SQP subproblem. In Section 5 we develop convergence results for a first auxiliary problem *via* the strong regularity property, which we transfer to the SQP method for (NLP$^{\gamma,\eta}$) with classical L^∞-closeness condition for the controls as in [16] in Section 6, and finally, with L^2-closeness as in [27] in Section 7.

2. Problem setting and assumptions

In this section, we state the precise setting of the model problem, which is the same as in [1]. Let us recall the problem formulation

$$\begin{cases} \min\limits_{q\in Q_{\mathrm{ad}},\mathbf{u}\in V} \quad J(\mathbf{u},q) := \dfrac{1}{2}\|u-u_d\|^2_{L^2(\Omega,\mathbb{R}^2)} + \dfrac{\alpha}{2}\|q\|^2_{L^2(\Gamma)}, \\ \text{subject to :} \quad A(\mathbf{u}) + R(\varphi;\gamma) = Bq. \qquad\qquad (\mathrm{EL}^{\gamma,\eta}) \end{cases} \tag{NLP$^{\gamma,\eta}$}$$

from the introduction and collect our standing assumptions. We assume $\Omega \subset \mathbb{R}^2$ to be a polygonal Gröger regular domain, *cf.* [54], with boundary $\partial\Omega = \Gamma\dot\cup\Gamma_D$, where Γ is the Neumann part of the boundary on which the control q acts as a boundary force. The remaining part of $\partial\Omega$ is denoted by Γ_D, on which homogeneous Dirichlet boundary conditions are prescribed. The control space Q is given by $Q = L^2(\Gamma)$, and for $q_a, q_b \in L^\infty(\Gamma)$ with $q_a < q_b$ a.e. on Γ, the set of admissible controls is denoted by

$$Q_{\mathrm{ad}} := \{q \in Q \quad | \quad q_a \le q \le q_b \quad \text{a.e. on } \Gamma\}.$$

The state $\mathbf{u} := (u,\varphi) \in V$ consists of a pair of functions, with displacement u and phase-field φ. The state space V will be defined below. The given function $u_d \in L^2(\Omega,\mathbb{R}^2)$ denotes a desired displacement, and the Tikhonov cost parameter α is a fixed positive real number. Before we define the operators A, R, and B, we fix some general notation for function spaces. We define the spaces

$$
\begin{aligned}
V_u &:= H^1_D(\Omega;\mathbb{R}^2) := \{v \in H^1(\Omega;\mathbb{R}^2) \,|\, v = 0 \text{ on } \Gamma_D\}, & V_\varphi &:= H^1(\Omega), & V &:= V_u \times V_\varphi,\\
W_u &:= W^{1,p}_D(\Omega;\mathbb{R}^2), & W_\varphi &:= W^{2,q}(\Omega), & W &:= W_u \times W_\varphi,\\
W^\times &:= W^{-1,p}(\Omega;\mathbb{R}^2) \times L^q(\Omega)
\end{aligned}
$$

consistent with the notation in [1]. For our solution theory, we consider $p > 2$, but close to two, and $q := p/2 > 1$, hence q close to one, and assume Ω to be $W^{2,q}$-regular for the homogeneous Neumann-problem $-\varepsilon\Delta\varphi + \frac{1}{\varepsilon}\varphi = g$, $\partial_n\rho = 0$ on Γ, $\rho = 0$ on Γ_D, *i.e.* for right-hand side $g \in L^q(\Omega)$, the solution φ is in $W^{2,q}$, *cf.* also [3], Section 2. Note that for this choice of p, q and spatial dimensions $N = 2$, we have $W_u \hookrightarrow V_u$ and $W_\varphi \hookrightarrow V_\varphi$ by the Sobolev-Kondrachov theorem. Further, $W_u \hookrightarrow L^\infty(\Omega;\mathbb{R}^2)$ as well as $W_\varphi \hookrightarrow L^\infty(\Omega)$ hold, which will often be used without further notice. For better readability we introduce the short notations

$$
Y := W \times Q \times W, \quad Y_\infty := W \times L^\infty(\Gamma) \times W, \qquad Z := W^\times \times Q \times W^\times, \quad Z_\infty := W^\times \times L^\infty(\Gamma) \times W^\times \tag{2.1}
$$

and will frequently use $y := (\mathbf{u}, q, \mathbf{z})$ for functions triples in Y or Y_∞, consisting of a state $\mathbf{u}$, a control q, and an adjoint state $\mathbf{z}$ (to be introduced later). Note that by the definition of the admissible set, all admissible controls will have L^∞-regularity.

We understand that all appearing spaces are defined on the domain Ω unless otherwise stated, and often omit the dependency on Ω for the sake of readability. For norms and inner products, we agree that $(\cdot,\cdot)$ denotes the usual L^2-inner product on Ω with corresponding norm $\|\cdot\|$, and $(\cdot,\cdot)_Q$ as well as $\|\cdot\|_Q$ correspond to the inner product and norm of Q *i.e.* $L^2(\Gamma)$. For functions $\mathbf{v} = (v^u, v^\varphi) \in W$, the norm in the space W is given by

$$
\|\mathbf{v}\|_W = \|(v^u, v^\varphi)\|_W = \|v^u\|_{1,p} + \|v^\varphi\|_{2,q},
$$

where $\|\cdot\|_{1,p}$ and $\|\cdot\|_{2,q}$ are Sobolev norms, defined by $\|v^u\|_{1,p} = \sum_{|i|\le 1} \|D^i v^u\|_p$ and $\|v^\varphi\|_{2,q} = \sum_{|i|\le 2} \|D^i v^\varphi\|_q$, with $\|\cdot\|_p$ and $\|\cdot\|_q$ being the standard norms of the Lebesgue spaces $L^p(\Omega;\mathbb{R}^2)$ and $L^q(\Omega)$. We will denote dual spaces with a superscript *, *e.g.* V^* stands for the dual of V, and agree that $\langle\cdot,\cdot\rangle$ stands for a duality pairing where the spaces are omitted if obvious from the context, otherwise denoted by a subscript. Lastly, let us introduce $B^{\mathbb{V}}_r(v)$ as the open ball of radius r centered at $v \in \mathbb{V}$ w.r.t the norm of $\mathbb{V}$, where $\mathbb{V}$ can be any Banach space.

As explained in the introduction, the equation ($\mathrm{EL}^{\gamma,\eta}$) is in fact a necessary optimality condition of an energy minimization problem. We call $A\colon V \hookleftarrow W \to V^*$ the phase-field operator, $R\colon V_\varphi \to V_\varphi^*$ the penalization operator, and $B\colon Q \to V^*$ the control-action operator on the Neumann boundary Γ. For $\mathbf{u} = (u, \varphi) \in W$, with $0 \le \varphi \le 1$, they are defined as

$$
\begin{aligned}
\langle A(\mathbf{u}), \mathbf{v}\rangle &:= (g(\varphi)\mathbb{C}e(u), e(v^u)) + \varepsilon(\nabla\varphi, \nabla v^\varphi) - \frac{1}{\varepsilon}(1-\varphi, v^\varphi)\\
&\quad + \eta(\varphi - \varphi^-, v^\varphi) + (1-\kappa)(\varphi\mathbb{C}e(u) : e(u), v^\varphi),\\
\langle R(\varphi;\gamma), v^\varphi\rangle &:= \gamma([(\varphi-\varphi^-)^+]^3, v^\varphi),\\
\langle Bq, (v^u, v^\varphi)\rangle &:= (q, v^u)_Q,
\end{aligned}
$$

for all $\mathbf{v} = (v^u, v^\varphi) \in V$ and given phase-field $\varphi^- \in W_\varphi$ with $0 \le \varphi^- \le 1$. Here, $e(u) = \frac{1}{2}(\nabla u + \nabla u^T)$ denotes the symmetric gradient for $u \in V_u$, and $\mathbb{C}$ the rank-4 elasticity tensor with usual properties, *cf. e.g.* [55], Section 3.

The operators A and R will also be used as mappings into the more regular spaces $W^\times$ and L^q, where we will use test functions $\mathbf{v} \in V$, since $W^\times \hookrightarrow V^*$. Moreover, we will also use $B\colon Q \to W^\times$, which is well defined as long as $p \le 4$, in accordance with our assumptions.

The parameter $\varepsilon \in (0,1)$ is a fixed phase-field parameter. Further, let $\varepsilon \gg \kappa > 0$ and $g(x) := (1-\kappa)x^2 + \kappa$. Both κ and g appear in the problem due to an Ambrosio-Tortorelli regularization [7] and an additional regularization for the elastic energy degeneracy. For more details, we point to [2], Section 2. In the operator R, we abbreviate the max-function by $(\cdot)^+ := \max(0,\cdot)$, and the given parameter $\gamma > 0$ is called the penalization parameter. The term originates from a 4th-order regularization of the irreversibility condition of the fracture problem in the energy minimization problem, *i.e.* $\frac{\gamma}{4}\|(\varphi-\varphi^-)^+\|^4_{L^4(\Omega)}$, *cf.* [9]. This penalization leads to the third order term R in the Euler-Lagrange equations after differentiation. The high order is needed to ensure the second-order differentiability of ($\mathrm{EL}^{\gamma,\eta}$) needed for the SQP method. Note that $\varphi^- \in W_\varphi$ is actually the phase-field of the previous time step if more than one time step is considered. For the problem with one time-step φ^- is an a-priori given function or the initial phase field. Finally, a sufficiently large $\eta \geq 0$ will be referred to as the (viscosity) regularization parameter, *cf.* [8] and also [1, 3, 10] in the context of ($\mathrm{NLP}^{\gamma,\eta}$). For sufficiently large η, unique solvability of ($\mathrm{EL}^{\gamma,\eta}$) as well as differentiability of the corresponding control-to-state-operator are known. We refer to *e.g.* [1], Section 3 for more detailed calculations and dependence of η on $\varepsilon, 1/\varepsilon$ and κ. We will frequently make use of results from [1, 3], that hold under such a condition, we therefore tacitly assume:

Assumption 2.1 (Viscous approximation)**.** Let $\eta \geq 0$ be chosen large enough for all results and calculations of the following sections that depend on such a condition on η.

Before we start our analysis we point out again that throughout this paper the problem parameters $\varepsilon, \kappa, \eta$, and γ remain fixed.

3. Preliminary results for ($\mathrm{EL}^{\gamma,\eta}$) and ($\mathrm{NLP}^{\gamma,\eta}$)

In this section we give a quick overview about available theoretical results, *cf. e.g.* [1, 3, 10].

3.1. The Euler-Lagrange equation ($\mathrm{EL}^{\gamma,\eta}$)

We briefly summarize known results concerning solvability of ($\mathrm{EL}^{\gamma,\eta}$), its linearization, and the associated solution operators. The proofs can be found in [1] and the references therein, *cf.* [2, 3, 56]. By [1], Lemma 3.1, for $\eta \geq 0$ being sufficiently large and $0 \leq \varphi^- \in W_\varphi$, we know that ($\mathrm{EL}^{\gamma,\eta}$) has a unique weak solution $\mathbf{u} \in W \hookrightarrow V$ for every $q \in Q$, *i.e.* $\mathbf{u}$ satisfies

$$\langle A(\mathbf{u}), \mathbf{v}\rangle + \langle R(\varphi;\gamma), v^\varphi\rangle = \langle Bq, \mathbf{v}\rangle \qquad \forall\, \mathbf{v} \in V.$$

All boundary conditions are contained in the operators and solution spaces. We denote the associated solution operator, also frequently called the control-to-state-operator, by

$$G\colon Q \to W, \quad G(q) = \mathbf{u} = (u, \varphi). \tag{3.1}$$

By [1], Proposition 3.3, we know that again for $\eta \geq 0$ sufficiently large, G, A, and R, are twice continuously Fréchet-differentiable from Q into W, respectively from W into $W^\times$ and from W_φ into L^q. The first derivative of G, $\tilde{\mathbf{u}} := G'(q)\tilde{q}$, $\tilde{\mathbf{u}} = (\tilde{u}, \tilde{\varphi}) \in W$, is the solution of

$$A'(\mathbf{u})\tilde{\mathbf{u}} + R'(\varphi;\gamma)\tilde{\varphi} = B\tilde{q}, \tag{3.2}$$

for $\mathbf{u} = G(q)$, *cf.* also Lemma 3.1 with $\mathbf{f} := B\tilde{q}$ in the following. The second derivative $\mathbf{y} := G''(q)[\tilde{q}_1, \tilde{q}_2]$, $\mathbf{y} = (\mathrm{y}^u, \mathrm{y}^\varphi) \in W$ is the solution of

$$A'(\mathbf{u})\mathbf{y} + R'(\varphi;\gamma)\mathrm{y}^\varphi = -A''(\mathbf{u})[\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2] - R''(\varphi;\gamma)[\tilde{\varphi}_1, \tilde{\varphi}_2], \tag{3.3}$$

for $\mathbf{u} = G(q)$, and $\tilde{\mathbf{u}}_i = G'(q)\tilde{q}_i$, $i = 1, 2$. In the above, the operators $A'(\mathbf{u})\colon V \to V^*$, $A''(\mathbf{u})\colon W \times W \to V^*$, $R'(\varphi, \gamma)\colon V_\varphi \to V_\varphi^*$, and $R''(\varphi; \gamma)\colon W_\varphi \times W_\varphi \to V_\varphi^*$ are given by

$$\begin{aligned}
\langle A'(\mathbf{u})\tilde{\mathbf{u}}, \mathbf{v}\rangle &:= (g(\varphi)\mathbb{C}e(\tilde{u}), e(v^u)) + 2(1-\kappa)(\varphi\mathbb{C}e(u) : e(\tilde{u}), v^\varphi) + 2(1-\kappa)(\varphi\mathbb{C}e(u)\tilde{\varphi}, e(v^u)) \\
&\quad + \varepsilon(\nabla\tilde{\varphi}, \nabla v^\varphi) + \frac{1}{\varepsilon}(\tilde{\varphi}, v^\varphi) + \eta(\tilde{\varphi}, v^\varphi) + (1-\kappa)(\tilde{\varphi}\mathbb{C}e(u) : e(u), v^\varphi), \\
\langle A''(\mathbf{u})[\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2], \mathbf{v}\rangle &:= 2(1-\kappa)[(\tilde{\varphi}_2\mathbb{C}e(u)\tilde{\varphi}_1, e(v^u)) + (\tilde{\varphi}_2\mathbb{C}e(\tilde{u}_1)\varphi, e(v^u)) + (\tilde{\varphi}_2\mathbb{C}e(u) : e(\tilde{u}_1), v^\varphi) \\
&\quad + (\varphi\mathbb{C}e(\tilde{u}_2)\tilde{\varphi}_1, e(v^u)) + (\tilde{\varphi}_1\mathbb{C}e(\tilde{u}_2) : e(u), v^\varphi) + (\varphi\mathbb{C}e(\tilde{u}_2) : e(\tilde{u}_1), v^\varphi)], \\
\langle R'(\varphi;\gamma)\tilde{\varphi}, v^\varphi\rangle &:= 3\gamma([(\varphi-\varphi^-)^+]^2\tilde{\varphi}, v^\varphi), \\
\langle R''(\varphi;\gamma)[\tilde{\varphi}_1, \tilde{\varphi}_2], v^\varphi\rangle &:= 6\gamma([(\varphi-\varphi^-)^+]\tilde{\varphi}_1\tilde{\varphi}_2, v^\varphi)
\end{aligned}$$

for all $\mathbf{v} \in V$. As for A and R, we can also use them as mappings within more regular, spaces, *i.e.*

$$\begin{aligned}
A'(\mathbf{u})&\colon W \to W^\times, & R'(\varphi,\gamma)&\colon W_\varphi \to L^q, \\
A''(\mathbf{u})&\colon W \times W \to W^\times, & R''(\varphi,\gamma)&\colon W_\varphi \times W_\varphi \to L^q.
\end{aligned}$$

Note that for one time step the operators $A'(\mathbf{u})$ and $R'(\varphi;\gamma)$ are in fact self-adjoint, *cf.* [1], Section 3.3.

In [1], Lemma 3.6, 3.8 and 3.9, we have proven local Lipschitz continuity results for G and its first and second derivative. In particular, there exists a constant $c > 0$ such that

$$\|G(q_1) - G(q_2)\|_W \le c\|q_1 - q_2\|_Q, \quad \forall q_1, q_2 \in Q_{\mathrm{ad}}, \tag{3.4}$$
$$\|G'(q_1)d^q - G'(q_2)d^q\|_Q \le c\|q_1 - q_2\|_Q\|d^q\|_Q, \quad \forall q_1, q_2 \in Q_{\mathrm{ad}}, d^q \in Q \tag{3.5}$$

hold, since the set of admissible controls is bounded. Let us also collect some boundedness and Lipschitz results for A, R, and their linearizations from these proofs or calculations in *e.g.* [1, 2, 10] for later use. While in particular the results for the derivatives of the operators from [1] are sufficient for our purposes, we refine the estimates by using weaker norms in V instead of W where possible, *cf. e.g.* [10]. The estimates follow by direct calculations from the definition of the operatros. We will give some arguments for the proofs in the appendix. First, there is a constant $c > 0$ such that for all $\mathbf{u}, \mathbf{u}_1, \mathbf{u}_2 \in W$ the boundedness and Lipschitz result

$$\|A(\mathbf{u})\|_{W^\times} + \|R(\varphi;\gamma)\|_q \le c(\|\mathbf{u}\|_W + 1), \tag{3.6}$$
$$\|A(\mathbf{u}_1) - A(\mathbf{u}_2)\|_{W^\times} + \|R(\varphi_1;\gamma) - R(\varphi_2;\gamma)\|_q \le c\|\mathbf{u}_1 - \mathbf{u}_2\|_W, \tag{3.7}$$

holds. This readily follows as an auxiliary result from the proofs of [1], Sections 3.1 and 3.2, using techniques that have already been used in [2]. The definition of A' and R' ensures the existence of a constant $c > 0$, such that

$$\|A'(\mathbf{u})\tilde{\mathbf{u}}\|_{V^*} + \|R'(\varphi;\gamma)\tilde{\varphi}\|_{V_\varphi^*} \le c(\|\mathbf{u}\|_W^2 + 1)\|\tilde{\mathbf{u}}\|_V, \tag{3.8}$$
$$\|A'(\mathbf{u})\tilde{\mathbf{u}}\|_{W^\times} + \|R'(\varphi;\gamma)\tilde{\varphi}\|_q \le c(\|\mathbf{u}\|_W^2 + 1)\|\tilde{\mathbf{u}}\|_W, \tag{3.9}$$

for any $\mathbf{u}, \tilde{\mathbf{u}} \in W$. Analogously to the estimations made in the proof of [1], Lemma 3.9, we obtain Lipschitz continuity results for A' and R', *i.e.* there exists a constant $c > 0$ such that for all $\tilde{\mathbf{u}}, \mathbf{u}_1, \mathbf{u}_2 \in W$ the estimates

$$\|(A'(\mathbf{u}_1) - A'(\mathbf{u}_2))\tilde{\mathbf{u}}\|_{V^*} + \|(R'(\varphi_1;\gamma) - R'(\varphi_2;\gamma))\tilde{\varphi}\|_{V_\varphi^*} \le c\|\mathbf{u}_1 - \mathbf{u}_2\|_W(\|\mathbf{u}_1\|_W + \|\mathbf{u}_2\|_W)\|\tilde{\mathbf{u}}\|_V, \tag{3.10}$$
$$\|(A'(\mathbf{u}_1) - A'(\mathbf{u}_2))\tilde{\mathbf{u}}\|_{W^\times} + \|(R'(\varphi_1;\gamma) - R'(\varphi_2;\gamma))\tilde{\varphi}\|_q \le c\|\mathbf{u}_1 - \mathbf{u}_2\|_W(\|\mathbf{u}_1\|_W + \|\mathbf{u}_2\|_W)\|\tilde{\mathbf{u}}\|_W, \tag{3.11}$$

are satisfied. Next, for $\tilde{\mathbf{u}} := (\tilde{u}, \tilde{\varphi}) \in W$ and $\mathbf{z} := (z^u, z^\varphi) \in W$ we denote by $A''(\mathbf{u})[\tilde{\mathbf{u}}, \cdot]^*$ and $R''(\varphi;\gamma)[\tilde{\varphi}, \cdot]^*$ the adjoint operators defined by

$$\langle A''(\mathbf{u})[\tilde{\mathbf{u}}, \cdot]^* \mathbf{z}, \mathbf{v}\rangle = \langle A''(\mathbf{u})[\tilde{\mathbf{u}}, \mathbf{v}], \mathbf{z}\rangle, \quad \langle R''(\varphi;\gamma)[\tilde{\varphi}, \cdot]^* z^\varphi, v^\varphi\rangle = \langle R''(\varphi;\gamma)[\tilde{\varphi}, v^\varphi], z^\varphi\rangle,$$

for all test functions $\mathbf{v} \in W$. Note that $W \hookrightarrow (W^\times)^*$ under the assumption $p > 2$, so they are well defined for any $\mathbf{z} \in W$. Then, as in the proof of [1], Lemma 3.9 we find that

$$\|A''(\mathbf{u})[\tilde{\mathbf{u}}, \cdot]^* \mathbf{z}\|_{W^\times} \le c\|\mathbf{u}\|_W \|\mathbf{z}\|_W \|\tilde{\mathbf{u}}\|_W, \tag{3.12}$$

$$\|R''(\varphi;\gamma)[\tilde{\varphi}, \cdot]^* z^\varphi\|_q \le c(\|\varphi\|_{V_\varphi} + 1)\|\tilde{\varphi}\|_{V_\varphi}\|z^\varphi\|_{V_\varphi} \le c(\|\mathbf{u}\|_W + 1)\|\tilde{\mathbf{u}}\|_W\|\mathbf{z}\|_W, \tag{3.13}$$

hold for a constant $c > 0$. Note in particular that there exists a constant $c > 0$, such that for $\mathbf{u}, \tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2, \mathbf{z} \in W$ the boundedness results

$$|\langle A''(\mathbf{u})[\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2], \mathbf{z}\rangle| \le c\|\mathbf{u}\|_W \|\mathbf{z}\|_W \|\tilde{\mathbf{u}}_1\|_V \|\tilde{\mathbf{u}}_2\|_V, \tag{3.14}$$

$$|\langle R''(\varphi;\gamma)[\tilde{\varphi}_1, \tilde{\varphi}_2], z^\varphi\rangle| \le c(\|\mathbf{u}\|_W + 1)\|\mathbf{z}\|_W \|\tilde{\mathbf{u}}_1\|_V \|\tilde{\mathbf{u}}_2\|_V \tag{3.15}$$

hold. The estimates (3.14) and (3.15) are true as long as any two of the four involved functions are in the space W with improved regularity, while the other two are in V. This result has already been stated in [10], Lemma 2.2, and will be revisited in Appendix A.1. From (3.14) and (3.15), and the linearity of the function $\mathbf{u}$ in A'', as well as the Lipschitz continuity of $(\cdot)^+$, it also immediately follows that

$$|\langle (A''(\mathbf{u}_1) - A''(\mathbf{u}_2))[\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2], \mathbf{z}\rangle| + |\langle (R''(\varphi_1;\gamma) - R''(\varphi_2;\gamma))[\tilde{\varphi}_1, \tilde{\varphi}_2], z^\varphi\rangle| \le c\|\mathbf{u}_1 - \mathbf{u}_2\|_W \|\mathbf{z}\|_W \|\tilde{\mathbf{u}}_1\|_V \|\tilde{\mathbf{u}}_2\|_V. \tag{3.16}$$

Furthermore, there is a constant $c > 0$ such that for all $\mathbf{u}_1, \mathbf{u}_2, \tilde{\mathbf{u}}, \mathbf{z} \in W$ we have the estimate:

$$\|(A''(\mathbf{u}_1) - A''(\mathbf{u}_2))[\tilde{\mathbf{u}}, \cdot]^* \mathbf{z}\|_{W^\times} + \|(R''(\varphi_1;\gamma) - R''(\varphi_2;\gamma))[\tilde{\varphi}, \cdot]^* z^\varphi\|_q \le c\|\mathbf{u}_1 - \mathbf{u}_2\|_W \|\mathbf{z}\|_W \|\tilde{\mathbf{u}}\|_W. \tag{3.17}$$

For further explicit reference, we state a result from [1], Lemma 3.2 for linear equations as in (3.2) and (3.3) with arbitrary right-hand side data. Due to the self-adjointness of the operators A' and R', it is also applicable to adjoint type equations appearing in optimality conditions.

Lemma 3.1. *Let $\eta \ge 0$ sufficiently large and $\mathbf{u} \in W$ be given, then for every $\mathbf{f} = (f^u, f^\varphi) \in V^*$, the partial differential equations*

$$A'(\mathbf{u})\tilde{\mathbf{u}} + R'(\varphi;\gamma)\tilde{\varphi} = \mathbf{f} \quad \text{and} \quad (A'(\mathbf{u}))^* \mathbf{z} + (R'(\varphi;\gamma))^* z^\varphi = \mathbf{f} \tag{3.18}$$

have unique weak solutions $\tilde{\mathbf{u}}, \mathbf{z} \in V$. If further $\mathbf{f} \in W^\times$, then $\tilde{\mathbf{u}}, \mathbf{z} \in W$. Moreover, the estimates

$$\|\tilde{\mathbf{u}}\|_V + \|\mathbf{z}\|_V \le c\|\mathbf{f}\|_{V^*}, \quad \|\tilde{\mathbf{u}}\|_W + \|\mathbf{z}\|_W \le c\max(\|\mathbf{u}\|_W^1, \|\mathbf{u}\|_W^2, \|\mathbf{u}\|_W^3, \|\mathbf{u}\|_W^4)\|\mathbf{f}\|_{W^\times}, \tag{3.19}$$

are fulfilled with some constant $c > 0$ depending on the initial phase-field φ^- and on the model parameters.

We would like to point out again that our operators and solution spaces already incorporate homogeneous Dirichlet boundary conditions on Γ_D and, only for state and linearized state equations, Neumann boundary conditions *via* $B\tilde{q} \in W^{-1,p}(\Omega) \times L^q(\Omega) \subset V^*$. We can now introduce the solution operator $\mathcal{G}_\mathbf{u}$ corresponding to (3.18), for arbitrary right-hand sides $\mathbf{f} \in W^\times$ and given $\mathbf{u} \in W$, by

$$\mathcal{G}_\mathbf{u} \colon W^\times \to W, \quad \mathcal{G}_\mathbf{u}(\mathbf{f}) =: \tilde{\mathbf{u}}. \tag{3.20}$$

Note that the first and second derivatives G' and G'' of the control-to state operator G can be expressed by

$$G'(q)\tilde{q} = \mathcal{G}_{\mathbf{u}}(B\tilde{q}), \quad G''(q)[\tilde{q}_1, \tilde{q}_2] = \mathcal{G}_{\mathbf{u}}(-A''(\mathbf{u})[\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2] - (0,\ R''(\varphi;\gamma)[\tilde{\varphi}_1, \tilde{\varphi}_2])),$$

where again $\mathbf{u} = G(q) \in W$, and $\tilde{\mathbf{u}}_i = \mathcal{G}_{\mathbf{u}}(B\tilde{q}_i) = G'(q)\tilde{q}_i \in W$, $i = 1, 2$. We will use this in the description of the SQP subproblems in the following. In order to simplify the presentation of later proofs, we conclude this section with some immediate consequences of Lemma 3.1, the estimates (3.6)–(3.17), and the triangle inequalities

$$\|\mathbf{u}\|_W \le \|\mathbf{u} - \bar{\mathbf{u}}\|_W + \|\bar{\mathbf{u}}\|_W \le \hat{\omega} + c_u, \quad \|\mathbf{z}\|_W \le \|\mathbf{z} - \bar{\mathbf{z}}\|_W + \|\bar{\mathbf{z}}\|_W \le \hat{\omega} + c_z \tag{3.21}$$

in local neighborhoods with radius $\hat{\omega}$ of fixed functions $\bar{\mathbf{u}}, \bar{\mathbf{z}} \in W$.

Corollary 3.2. *Let $\hat{\omega} > 0$ and $\bar{\mathbf{u}} \in W$ be fixed. Moreover, let $\mathbf{u} \in W$ satisfy $\|\mathbf{u} - \bar{\mathbf{u}}\|_W \le \hat{\omega}$, and let $\mathbf{f} = (f^u, f^\varphi) \in W^\times$ be arbitrary. Then the solutions $\tilde{\mathbf{u}}$ and $\mathbf{z}$ of the partial differential equations (3.18) satisfy*

$$\|\tilde{\mathbf{u}}\|_V + \|\mathbf{z}\|_V \le (c + c_{\hat{\omega}})\|\mathbf{f}\|_{V^*}, \quad \|\tilde{\mathbf{u}}\|_W + \|\mathbf{z}\|_W \le (c + c_{\hat{\omega}})\|\mathbf{f}\|_{W^\times}. \tag{3.22}$$

for a constant $c > 0$ depending on $\bar{\mathbf{u}}$ but not on $\mathbf{u}$, and a constant $c_{\hat{\omega}} = \mathcal{O}(\hat{\omega})$ (as $\hat{\omega} \to 0$).

Corollary 3.3. *Let $\hat{\omega} > 0$ and $\bar{\mathbf{u}}, \bar{\mathbf{z}} \in W$ be given. There exist constants $c > 0$ depending on $\bar{\mathbf{u}}$ and $\bar{\mathbf{z}}$ and (generic) constants $\hat{c}_{\hat{\omega}} = \mathcal{O}(\hat{\omega})$, $c_{\hat{\omega}} = \mathcal{O}(\hat{\omega})$ (as $\hat{\omega} \to 0$) such that the estimates*

$$\begin{aligned}
\|A(\mathbf{u})\|_{W^\times} + \|R(\varphi;\gamma)\|_q &\le c + \hat{c}_{\hat{\omega}}, && (3.23)\\
\|A(\mathbf{u}_1) - A(\mathbf{u}_2)\|_{W^\times} + \|R(\varphi_1;\gamma) - R(\varphi_2;\gamma)\|_q &\le c_{\hat{\omega}}, && (3.24)\\
\|A'(\mathbf{u})\tilde{\mathbf{u}}\|_{V^*} + \|R'(\varphi;\gamma)\tilde{\varphi}\|_{V_\varphi^*} &\le (c + \hat{c}_{\hat{\omega}})\|\tilde{\mathbf{u}}\|_V, && (3.25)\\
\|A'(\mathbf{u})\tilde{\mathbf{u}}\|_{W^\times} + \|R'(\varphi;\gamma)\tilde{\varphi}\|_q &\le (c + \hat{c}_{\hat{\omega}})\|\tilde{\mathbf{u}}\|_W, && (3.26)\\
\|(A'(\mathbf{u}_1) - A'(\mathbf{u}_2))\tilde{\mathbf{u}}\|_{V^*} + \|(R'(\varphi_1;\gamma) - R'(\varphi_2;\gamma))\tilde{\varphi}\|_{V_\varphi^*} &\le (c + \hat{c}_{\hat{\omega}})\|\mathbf{u}_1 - \mathbf{u}_2\|_W\|\tilde{\mathbf{u}}\|_V\\
&\le c_{\hat{\omega}}\|\tilde{\mathbf{u}}\|_V, && (3.27)\\
\|(A'(\mathbf{u}_1) - A'(\mathbf{u}_2))\tilde{\mathbf{u}}\|_{W^\times} + \|(R'(\varphi_1;\gamma) - R'(\varphi_2;\gamma))\tilde{\varphi}\|_q &\le (c + \hat{c}_{\hat{\omega}})\|\mathbf{u}_1 - \mathbf{u}_2\|_W\|\tilde{\mathbf{u}}\|_W\\
&\le c_{\hat{\omega}}\|\tilde{\mathbf{u}}\|_W, && (3.28)\\
|\langle A''(\mathbf{u})[\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2], \mathbf{z}\rangle| + \langle R''(\varphi;\gamma)[\tilde{\varphi}_1, \tilde{\varphi}_2], z^\varphi\rangle| &\le (c + \hat{c}_{\hat{\omega}})\|\tilde{\mathbf{u}}_1\|_V\|\tilde{\mathbf{u}}_2\|_V, && (3.29)\\
\|(A''(\mathbf{u}_1) - A''(\mathbf{u}_2))[\tilde{\mathbf{u}}, \cdot]^*\mathbf{z}\|_{W^\times} + \|(R''(\varphi_1;\gamma) - R''(\varphi_2;\gamma))[\tilde{\varphi}, \cdot]^* z^\varphi\|_q &\le (c + \hat{c}_{\hat{\omega}})\|\mathbf{u}_1 - \mathbf{u}_2\|_W\|\tilde{\mathbf{u}}\|_W\\
&\le c_{\hat{\omega}}\|\tilde{\mathbf{u}}\|_W && (3.30)
\end{aligned}$$

hold for all $\mathbf{u}, \mathbf{z}, \mathbf{u}_i, \mathbf{z}_i \in W$ with $\|\mathbf{u} - \bar{\mathbf{u}}\|_W, \|\mathbf{z} - \bar{\mathbf{z}}\|_W\ \|\mathbf{u}_i - \bar{\mathbf{u}}\|_W, \|\mathbf{z}_i - \bar{\mathbf{z}}\|_W \le \hat{\omega}$ and all $\tilde{\mathbf{u}}, \tilde{\mathbf{u}}_i \in W$, $i = 1, 2$.

Remark 3.4. We omit the proofs of the last two Corollaries, but would like to comment on properties of the constant $c(\hat{\omega})$. The triangle inequalities (3.21) applied to $\mathbf{u}, \mathbf{u}_i, \mathbf{z}, \mathbf{z}_i$ inserted into the estimate (3.19) of Lemma 3.2 and estimates (3.6)–(3.17) immediately result in a precise dependence of $c(\hat{\omega})$ on (powers of) $\hat{\omega}$. As an example, we observe

$$\begin{aligned}
|(A'(\mathbf{u}_1) - A'(\mathbf{u}_2))\tilde{\mathbf{u}}\|_{W^\times} \le c\|\mathbf{u}_1 - \mathbf{u}_2\|_W(\|\mathbf{u}_1\|_W + \|\mathbf{u}_2\|_W)\|\tilde{\mathbf{u}}\|_W &\le 2c(c_u + \hat{\omega})\|\mathbf{u}_1 - \mathbf{u}_2\|_W\|\tilde{\mathbf{u}}\|_W\\
&\le 2cc_u\hat{\omega}(1 + \hat{\omega})\|\tilde{\mathbf{u}}\|_W.
\end{aligned}$$

In the following sections, we will use finite, but small neighborhoods with radius $\hat{\omega}$. It is clear from the above estimates that $c(\hat{\omega})$ is of order $\mathcal{O}(\hat{\omega})$ as $\hat{\omega} \to 0$. We can therefore guarantee that for any arbitrary but finite lower bound $c^- > 0$ there exists $\hat{\omega}^- > 0$ such that $0 < c(\hat{\omega}) < c^-$ for all $\hat{\omega} \le \hat{\omega}^-$. We will use this to make the following proofs less technical.

3.2. The optimization problem (NLP$^{\gamma,\eta}$)

Now, we gather some known results for the control problem (NLP$^{\gamma,\eta}$), see [1, 2, 10]. We first write (NLP$^{\gamma,\eta}$) in a usual reduced form. Utilizing the control-to-state operator G, and implicitly using the embedding $W \hookrightarrow L^2(\Omega,\mathbb{R}^2)\times L^2(\Omega)$, we define the reduced functional

$$f\colon Q\to\mathbb{R},\quad f(q):=J(G(q),q)=J(\mathbf{u},q). \tag{3.31}$$

Then (NLP$^{\gamma,\eta}$) is equivalent to

$$\min_q f(q),\quad \text{subject to}\quad q\in Q_{\mathrm{ad}}. \tag{NLP$^{\gamma,\eta}_{\mathrm{red}}$}$$

Existence of at least one global minimizer $\bar q\in Q_{\mathrm{ad}}$ of (NLP$^{\gamma,\eta}_{\mathrm{red}}$) with associated state $\bar{\mathbf{u}}\in W$ has been shown in [1], Proposition 4.1 by standard methods, see also [2], Theorem 4.3 for a model problem without control constraints. Due to the nonconvex structure of (NLP$^{\gamma,\eta}$), we call $\bar q\in Q_{\mathrm{ad}}$ a local minimizer of (NLP$^{\gamma,\eta}$) in the sense of $L^2(\Gamma)$, if there exists an $r>0$ such that

$$f(\bar q)\le f(q)\quad \forall\, q\in Q_{\mathrm{ad}}\quad\text{with}\quad \|\bar q-q\|_Q\le r. \tag{3.32}$$

First-order necessary optimality conditions for (NLP$^{\gamma,\eta}$) have been stated in [1], Lemma 4.3 by adapting the results of [2, 3] without control constraints. They read as follows:

Lemma 3.5. *Let $\bar q\in Q_{ad}$ be a local minimizer of (NLP$^{\gamma,\eta}_{\mathrm{red}}$) with associated state $\bar{\mathbf{u}}\in W$. Then there exists an adjoint state $\bar{\mathbf{z}}=(\bar z^u,\bar z^\varphi)\in W$ such that*

$$A(\bar{\mathbf{u}})+R(\bar\varphi;\gamma)=B\bar q, \tag{EL$^{\gamma,\eta}$}$$
$$(A'(\bar{\mathbf{u}}))^*\bar{\mathbf{z}}+(R'(\bar\varphi;\gamma))^*\bar z^\varphi=\bar u-u_d, \tag{AE$^{\gamma,\eta}$}$$
$$(B^*\bar{\mathbf{z}}+\alpha\bar q,q-\bar q)_Q\ge 0\qquad \forall q\in Q_{ad} \tag{VI$^{\gamma,\eta}$}$$

holds.

We omit the proof, but briefly recall that solvability and regularity of the adjoint equation (AE$^{\gamma,\eta}$) follow from Lemma 3.1. We point out that Lemma 3.5 already makes use of differentiability of f. In fact, for $\eta\ge 0$ sufficiently large, the reduced functional f is twice Fréchet-differentiable from Q into $\mathbb{R}$ by [1], Corollary 3.5 and Section 3.3, with derivatives

$$f'(q)d^q=(B^*\mathbf{z}+\alpha q,d^q)_Q, \tag{3.33}$$
$$f''(q)[d^q_1,d^q_2]=(d^u_1,d^u_2)+\alpha(d^q_1,d^q_2)_Q-\langle A''(\mathbf{u})[\mathbf{d}^{\mathbf{u}}_1,\mathbf{d}^{\mathbf{u}}_2],\mathbf{z}\rangle-\langle R''(\varphi;\gamma)[d^\varphi_1,d^\varphi_2],z^\varphi\rangle, \tag{3.34}$$

where $\mathbf{u}=G(q)$, $\mathbf{z}=G'(q)^*((u-u_d,0))$, $\mathbf{d}^{\mathbf{u}}_i=\mathcal{G}_{\mathbf{u}}(Bd^q_i)$, and $d^q,d^q_i\in Q$, for $i=1,2$. In [1], Theorem 4.6 and Remark 4.7, we have established second-order sufficient optimality conditions. If $\bar q\in Q_{\mathrm{ad}}$ with associated state $\bar{\mathbf{u}}\in W$ and adjoint state $\bar{\mathbf{z}}\in W$ satisfies the first-order necessary conditions from Lemma 3.5 as well as the coercivity condition

$$\exists\,\delta_{\mathrm{SSC}}>0\quad\text{such that}\quad f''(\bar q)(d^q,d^q)\ge\delta_{\mathrm{SSC}}\|d^q\|_Q^2\quad\forall\, d^q\in\mathcal{C}(\bar q), \tag{3.35}$$

with a cone of critical directions $\mathcal{C}(\bar q)$ defined by

$$\mathcal{C}(\bar q) := \{d^q \in Q \mid d^q(x) \begin{cases} \geq 0 & \text{if} \quad \bar q(x) = q_a(x), \\ \leq 0 & \text{if} \quad \bar q(x) = q_b(x), \\ = 0 & \text{if} \quad B^*\bar{\mathbf{z}}(x) + \alpha \bar q(x) \neq 0, \end{cases} \}, \tag{3.36}$$

then there exist constants $\epsilon > 0$ and $c > 0$ such that the quadratic growth condition

$$f(q) \geq f(\bar q) + c\|q - \bar q\|_Q^2 \quad \forall\, q \in Q_{\text{ad}} \quad \text{that satisfy} \quad \|q - \bar q\|_Q \leq \epsilon$$

holds without two-norm gap. However, for proving convergence of the SQP method using a second-order sufficient condition with a cone of critical directions, we will have to ensure that the descent directions determined by the quadratic subproblem stay within this cone of critical directions. If this cone is too small, we cannot guarantee this. Following the ideas of [16], we therefore use slightly stronger SSC that involve so-called σ-strongly active constraints, for an arbitrarily small but fixed parameter $\sigma > 0$. Let us therefore define the set

$$\mathcal{I}(\sigma) := \{x \in \Gamma \quad \Big| \quad |B^*\bar{\mathbf{z}} + \alpha \bar q| \geq \sigma\}, \quad \text{for} \quad \sigma > 0 \tag{3.37}$$

and the cone of (σ-)critical directions

$$\mathcal{C}_\sigma(\bar q) := \{d^q \in Q \,|\, d^q(x) = 0 \text{ on } \mathcal{I}(\sigma)\}, \tag{3.36'}$$

without sign conditions. Note that for the cone $\mathcal{C}(\bar q)$ defined in (3.36) we have the inclusion

$$\mathcal{C}(\bar q) \subset \mathcal{C}_\sigma(\bar q) \tag{3.38}$$

for $\sigma > 0$. Since in this paper we will frequently use Lagrange functionals, we now define

$$\mathcal{L}\colon Y \to \mathbb{R}, \quad \mathcal{L}(y) := J(\mathbf{u}, q) - \langle A(\mathbf{u}), \mathbf{z}\rangle - \langle R(\varphi;\gamma), z^\varphi\rangle + \langle Bq, \mathbf{z}\rangle, \tag{3.39}$$

with $y = (\mathbf{u}, q, \mathbf{z})$ as introduced in Section 2. We will state a second order sufficient condition in terms of the Lagrangian. With the differentiability results for A, R and the objective functional it is clear that second-order Fréchet-differentiability of $\mathcal{L}$ from $W \times Q$ into $\mathbb{R}$ holds with

$$\mathcal{L}''(y)[(\mathbf{d}_1^{\mathbf{u}}, d_1^q), (\mathbf{d}_2^{\mathbf{u}}, d_2^q)] = (d_1^u, d_2^u) + \alpha(d_1^q, d_2^q)_Q - \langle A''(\mathbf{u})[\mathbf{d}_1^{\mathbf{u}}, \mathbf{d}_2^{\mathbf{u}}], \mathbf{z}\rangle - \langle R''(\varphi;\gamma)[d_1^\varphi, d_2^\varphi], z^\varphi\rangle, \tag{3.40}$$

for all $y \in Y$, $\mathbf{d}_i^{\mathbf{u}} = (d_i^u, d_i^\varphi) \in W$, $d_i^q \in L^2(\Gamma)$, $i = 1, 2$, where $\mathcal{L}''(y)$ denotes the second derivative of the Lagrangian with respect to $(\mathbf{u}, q)$. A quick calculation verifies in particular that

$$f''(q)[d_1^q, d_2^q] = \mathcal{L}''(y)[(\mathbf{d}_1^{\mathbf{u}}, d_1^q), (\mathbf{d}_2^{\mathbf{u}}, d_2^q)] \tag{3.41}$$

holds for all $q, d_i^q \in Q$ and all $y \in Y$, $\mathbf{d}_i^{\mathbf{u}} \in W$, $i = 1, 2$, that satisfy $y = (G(q),\, q,\, G'(q)^*((G(q) - u_d, 0)))$ and $\mathbf{d}_i^{\mathbf{u}} = G'(q)d_i^q$, $i = 1, 2$. With the identity (3.41) and the inclusion (3.38) in mind, we now impose the following assumption:

Assumption 3.6. Let $\bar y = (\bar{\mathbf{u}}, \bar q, \bar{\mathbf{z}}) \in Y_\infty$ fulfill the first-order necessary conditions given in Lemma 3.5 and assume there exist constants $\sigma > 0$ and $\delta_{\text{SSC}} > 0$ such that

$$\mathcal{L}''(\bar y)[(\mathbf{d}^{\mathbf{u}}, d^q), (\mathbf{d}^{\mathbf{u}}, d^q)] \geq \delta_{\text{SSC}}\|d^q\|_Q^2 \quad \forall\, d^q \in \mathcal{C}_\sigma(\bar q),\, \mathbf{d}^{\mathbf{u}} = G'(\bar q)d^q. \tag{3.35'}$$

Corollary 3.7. *Let $\bar{y}$ fulfill Assumption 3.6. There exist constants $\epsilon > 0$ and $c > 0$ such that the quadratic growth condition*

$$f(q) \geq f(\bar{q}) + c\|q - \bar{q}\|_Q^2$$

holds for every $q \in Q_{ad}$ with $\|q - \bar{q}\|_Q \leq \epsilon$. In particular, $\bar{q}$ is a strict locally optimal control in the sense of L^2.

Proof. This follows from [1], Theorem 4.6 and the inclusion (3.38). □

3.3. Properties of the Lagrangian

To end this section, we point out that by by [1], Lemma 3.10 and for $\eta \geq 0$ sufficiently large, f is second-order locally Lipschitz continuous in Q, *i.e.* for every $\rho > 0$ there exists a constant $c_L = c_L(\rho) > 0$ such that for all $q_1, q_2, d_1^q, d_2^q \in Q$ with $\|q_1 - q_2\|_Q \leq \rho$, it holds

$$\left|(f''(q_1) - f''(q_2))[d_1^q, d_2^q]\right| \leq c_L\|q_1 - q_2\|_Q\|d_1^q\|_Q\|d_2^q\|_Q. \tag{3.42}$$

By (3.41), this property transfers to the second derivative of the Lagrangian with $\mathbf{d}_i^{\mathbf{u}} = G'(q)d_i^q$ and $\mathbf{u} = G(q)$, but for future frequent use we will derive properties for arbitrary linearization points $y_1, y_2 \in Y$ and arbitrary directions.

Proposition 3.8. *There exists a constant $c > 0$ such that for all $y, y_i \in Y$, $\mathbf{d}^{\mathbf{u}}, \mathbf{d}_i^{\mathbf{u}} \in W$, $d^q, d_i^q \in Q$, $i = 1, 2$ the estimates*

$$\begin{aligned}&|[\mathcal{L}''(y_1) - \mathcal{L}''(y_2)][(\mathbf{d}_1^{\mathbf{u}}, d_1^q), (\mathbf{d}_2^{\mathbf{u}}, d_2^q)]| \\ &\quad \leq c(\|\mathbf{u}_1\|_W + \|\mathbf{u}_2\|_W + \|\mathbf{z}_1\|_W + \|\mathbf{z}_2\|_W)(\|\mathbf{u}_1 - \mathbf{u}_2\|_W + \|\mathbf{z}_1 - \mathbf{z}_2\|_W)\|\mathbf{d}_1^{\mathbf{u}}\|_V\|\mathbf{d}_2^{\mathbf{u}}\|_V,\end{aligned} \tag{3.43}$$

$$\begin{aligned}&|\mathcal{L}''(y)[(\mathbf{d}_1^{\mathbf{u}}, d_1^q), (\mathbf{d}_1^{\mathbf{u}}, d_1^q)] - \mathcal{L}''(y)[(\mathbf{d}_2^{\mathbf{u}}, d_2^q), (\mathbf{d}_2^{\mathbf{u}}, d_2^q)]| \\ &\quad \leq c\|\mathbf{u}\|_W\|\mathbf{z}\|_W(\|\mathbf{d}_1^{\mathbf{u}}\|_V + \|\mathbf{d}_2^{\mathbf{u}}\|_V)\|\mathbf{d}_1^{\mathbf{u}} - \mathbf{d}_2^{\mathbf{u}}\|_V + c(\|d_1^q\|_Q + \|d_2^q\|_Q)\|d_1^q - d_2^q\|_Q\end{aligned} \tag{3.44}$$

are satisfied.

Before we give a short proof, we want to point out that due to (3.14), (3.15) and (3.16), we can in fact use the V-norm for the directions $\mathbf{d}_i^{\mathbf{u}}$, $i = 1, 2$ on the right hand side of (3.43) and (3.44). Nevertheless, our results remain true using the norm in W.

Proof. We start with estimate (3.43). Using (3.40) and rearranging terms leads to

$$\begin{aligned}&[\mathcal{L}''(y_1) - \mathcal{L}''(y_2)][(d_1^{\mathbf{u}}, d_1^q), (d_2^{\mathbf{u}}, d_2^q)] \\ =&\langle A''(\mathbf{u}_2)[\mathbf{d}_1^{\mathbf{u}}, \mathbf{d}_2^{\mathbf{u}}], \mathbf{z}_2\rangle - \langle A''(\mathbf{u}_1)[\mathbf{d}_1^{\mathbf{u}}, \mathbf{d}_2^{\mathbf{u}}], \mathbf{z}_1\rangle + \langle R''(\varphi_2;\gamma)[d_1^\varphi, d_2^\varphi], z_2^\varphi\rangle - \langle R''(\varphi_1;\gamma)[d_1^\varphi, d_2^\varphi], z_1^\varphi\rangle.\end{aligned} \tag{3.45}$$

The first claim now follows from the Lipschitz properties of A'' and R'' stated in (3.16). For estimate (3.44) we calculate

$$\begin{aligned}&\mathcal{L}''(y)[(\mathbf{d}_1^{\mathbf{u}}, d_1^q), (\mathbf{d}_1^{\mathbf{u}}, d_1^q)] - \mathcal{L}''(y)[(\mathbf{d}_2^{\mathbf{u}}, d_2^q), (\mathbf{d}_2^{\mathbf{u}}, d_2^q)] \\ &\quad = \|d_1^u\|^2 - \|d_2^u\|^2 + \alpha(\|d_1^q\|_Q^2 - \|d_2^q\|_Q^2) \\ &\qquad + \langle A''(\mathbf{u})[\mathbf{d}_2^{\mathbf{u}}, \mathbf{d}_2^{\mathbf{u}}] - A''(\mathbf{u})[\mathbf{d}_1^{\mathbf{u}}, \mathbf{d}_1^{\mathbf{u}}], \mathbf{z}\rangle + \langle R''(\varphi;\gamma)[d_2^\varphi, d_2^\varphi] - R''(\varphi;\gamma)[d_1^\varphi, d_1^\varphi], z^\varphi\rangle \\ &\quad = \|d_1^u\|^2 - \|d_2^u\|^2 + \alpha(\|d_1^q\|_Q^2 - \|d_2^q\|_Q^2) \\ &\qquad + \langle A''(\mathbf{u})[\mathbf{d}_1^{\mathbf{u}} - \mathbf{d}_2^{\mathbf{u}}, \mathbf{d}_1^{\mathbf{u}} - \mathbf{d}_2^{\mathbf{u}}], \mathbf{z}\rangle + 2\langle A''(\mathbf{u})[\mathbf{d}_2^{\mathbf{u}}, \mathbf{d}_1^{\mathbf{u}} - \mathbf{d}_2^{\mathbf{u}}], \mathbf{z}\rangle\end{aligned}$$

$$+ \langle R''(\varphi;\gamma)[d_1^\varphi - d_2^\varphi, d_1^\varphi - d_2^\varphi], z^\varphi\rangle + 2\langle R''(\varphi;\gamma)[d_2^\varphi, d_1^\varphi - d_2^\varphi], z^\varphi\rangle. \tag{3.46}$$

Estimates (3.14) and (3.15) combined with the Lipschitz continuity of the squared norms finalize the proof. □

Again, we state simple consequences of Proposition 3.8 in certain neighborhoods of $\bar{\mathbf{u}}, \bar{\mathbf{z}}$, *cf.* Remark 3.4, that will be used in later proofs.

Corollary 3.9. *Let $\bar{y} = (\bar{\mathbf{u}}, \bar{q}, \bar{\mathbf{z}}) \in Y_\infty$ denote a fixed triple satisfying Assumption 3.6. Then, for all $\hat{\omega} > 0$ there exist (generic) constants $\hat{c}_{\hat{\omega}} = \mathcal{O}(\hat{\omega}), c_{\hat{\omega}} = \mathcal{O}(\hat{\omega})$ (as $\hat{\omega} \to 0$) such that for all $y_i \in Y_\infty$ with $\|\mathbf{u}_i - \bar{\mathbf{u}}\|_W, \|\mathbf{z}_i - \bar{\mathbf{z}}\|_W \leq \hat{\omega}$ and all $\mathbf{d}_i^{\mathbf{u}} \in W$, $d_i^q \in Q$ $i = 1, 2$, the following estimate is satisfied for a constant $c > 0$ depending on $\bar{y}$, only:*

$$\begin{aligned} |[\mathcal{L}''(y_1) - \mathcal{L}''(y_2)][(\mathbf{d}_1^{\mathbf{u}}, d_1^q), (\mathbf{d}_2^{\mathbf{u}}, d_2^q)]| \leq & (c + \hat{c}_{\hat{\omega}})(\|\mathbf{u}_1 - \mathbf{u}_2\|_W + \|\mathbf{z}_1 - \mathbf{z}_2\|_W)\|\mathbf{d}_1^{\mathbf{u}}\|_V \|\mathbf{d}_2^{\mathbf{u}}\|_V. & (3.47) \\ \leq & c_{\hat{\omega}} \|\mathbf{d}_1^{\mathbf{u}}\|_V \|\mathbf{d}_2^{\mathbf{u}}\|_V. & (3.48) \end{aligned}$$

Moreover, for all $\hat{\omega} > 0$ there exist (generic) constants $\hat{c}_{\hat{\omega}} = \mathcal{O}(\hat{\omega}), c_{\hat{\omega}} = \mathcal{O}(\hat{\omega})$ (as $\hat{\omega} \to 0$) such that for all $y \in Y_\infty$ with $\|\mathbf{u} - \bar{\mathbf{u}}\|_W, \|\mathbf{z} - \bar{\mathbf{z}}\|_W \leq \hat{\omega}$ and all $\mathbf{d}_i^{\mathbf{u}} \in W$, $i = 1, 2$, $d^q \in Q$, the following estimate is satisfied for a constant $c > 0$ depending on $\bar{y}$, only:

$$|\mathcal{L}''(y)[(\mathbf{d}_1^{\mathbf{u}}, d^q), (\mathbf{d}_1^{\mathbf{u}}, d^q)] - \mathcal{L}''(y)[(\mathbf{d}_2^{\mathbf{u}}, d^q), (\mathbf{d}_2^{\mathbf{u}}, d^q)]| \leq (c_y + c_{\hat{\omega}})(\|\mathbf{d}_1^{\mathbf{u}}\|_V + \|\mathbf{d}_2^{\mathbf{u}}\|_V)\|\mathbf{d}_1^{\mathbf{u}} - \mathbf{d}_2^{\mathbf{u}}\|_V. \tag{3.49}$$

Assumption 3.10. In the following sections, we tacitly assume $\bar{y} = (\bar{\mathbf{u}}, \bar{q}, \bar{\mathbf{z}}) \in Y_\infty \subset Y$ with $\bar{q} \in Q_{\text{ad}}$ will always denote a fixed triple that satisfies Assumption 3.6, *i.e.* the first order necessary and (strong) second order sufficient optimality conditions of Problem ($\text{NLP}^{\gamma,\eta}$).

4. The SQP method

Let us now describe the SQP method for ($\text{NLP}^{\gamma,\eta}$). As is well-known it involves the iterative solution of quadratic subproblems, which we denote by (QP_k). For our precise control problem, they read as follows: Given a current iterate $y^k = (\mathbf{u}^k, q^k, \mathbf{z}^k) \in Y$, find an update direction $\mathbf{d}^k = (\mathbf{d}^{\mathbf{u},k}, d^{q,k}) := (\mathbf{u} - \mathbf{u}^k, q - q^k) \in W \times Q$ (with associated adjoint state $\mathbf{z}^{k+1}$ still to be introduced) that solves

$$\begin{cases} \min\limits_{\mathbf{d}} J_k(\mathbf{d}) := J'(\mathbf{u}^k, q^k)\mathbf{d} + \dfrac{1}{2}\mathcal{L}''(y^k)[\mathbf{d}, \mathbf{d}], \\ \quad \text{s.t.} \quad A'(\mathbf{u}^k)\mathbf{d}^{\mathbf{u}} + R'(\varphi^k;\gamma)d^\varphi = Bd^q + Bq^k - A(\mathbf{u}^k) - R(\varphi^k;\gamma) \\ \quad \text{and} \quad d^q \in Q_{\text{ad}}^k, \end{cases} \tag{QP_k}$$

with

$$Q_{\text{ad}}^k := \{d^q \in Q \quad | \quad q_a - q^k \leq d^q \leq q_b - q^k \quad \text{a.e.} \quad \text{on} \quad \Gamma\}.$$

The optimal update directions $\mathbf{d}^k$ determine new iterates $\mathbf{u}^{k+1} = \mathbf{u}^k + \mathbf{d}^{\mathbf{u},k} \in W$ and $q^{k+1} = q^k + d^{q,k} \in Q_{\text{ad}}$, if this problem is solvable. This will be discussed later. Note first that $d^q \in Q_{\text{ad}}^k$ is equivalent to $q := q^k + d^q \in Q_{\text{ad}}$, and let us also point out that we can equivalently express the equation for $(\mathbf{d}^{\mathbf{u},k}, d^{q,k})$ by

$$\mathbf{u} - \mathbf{u}^k = \mathbf{d}^{\mathbf{u},k} = \mathcal{G}_{\mathbf{u}^k}(Bd^{q,k} + Bq^k - A(\mathbf{u}^k) - (0,\ R(\varphi^k;\gamma))), \tag{4.1}$$

or

$$\mathbf{u} = \mathcal{G}_{\mathbf{u}^k}(Bq - A(\mathbf{u}^k) - (0,\ R(\varphi^k;\gamma)) + A'(\mathbf{u}^k)\mathbf{u}^k + (0, R'(\varphi^k;\gamma)\varphi^k)), \tag{4.2}$$

with the operator $\mathcal{G}_{\mathbf{u}^k}$ defined in (3.20). Note in particular that the operator B is linear, and q^k will cancel out of the PDE and the constraints. We can therefore introduce a reduced version of (QP_k) for later use,

$$\min_{q\in Q_{ad}} f_k(q) := J_k(\mathcal{G}_{\mathbf{u}^k}(Bq - A(\mathbf{u}^k) - (0,\, R(\varphi^k;\gamma)) + A'(\mathbf{u}^k)\mathbf{u}^k + (0, R'(\varphi^k;\gamma)\varphi^k)) - \mathbf{u}^k, q - q^k). \quad (\mathrm{QP}_k^{\mathrm{red}})$$

As is typical for Newton-type methods, with the formulation (QP_k) one can first determine an update direction and then use this to determine new iterates. Nevertheless, the reduced formulation ($\mathrm{QP}_k^{\mathrm{red}}$) has theoretical advantages, which is why we will use it in the upcoming analysis. For further reference, we note that the first order derivative of f_k is given, for all $q, d^q \in Q$, by

$$f_k'(q)d^q = (\alpha q, d^q)_Q + (B^*\mathbf{z}_k(q), d^q)_Q, \quad (4.3)$$

where

$$\mathbf{z}_k(q) := \mathcal{G}_{\mathbf{u}^k}^*((u-u_d,0) - A''(\mathbf{u}^k)[\mathbf{d}^{\mathbf{u},k},\cdot]^*\mathbf{z}^k - (0, R''(\varphi^k;\gamma)[d^{\varphi,k},\cdot]^* z^{\varphi,k})) \quad (4.4)$$

is an adjoint state, depending on a linearized state associated with q by (4.2). Note that for $q = q^{k+1}$ this means $\mathbf{z}_k(q) = \mathbf{z}^{k+1}$. This follows by standard straight forward calculations and is used to eventually obtain a first order optimality system in Lemma 4.1. The second order derivative fulfills

$$f_k''(q)(d_1^q, d_2^q) = \mathcal{L}''(y^k)[(\tilde{\mathbf{u}}_{k,1}, d_1^q), (\tilde{\mathbf{u}}_{k,2}, d_2^q)], \quad (4.5)$$

with $\tilde{\mathbf{u}}_{k,i} := \mathcal{G}_{\mathbf{u}^k}(Bd_i^q)$, $i = 1,2$, for any $q, d_1^q, d_2^q \in Q$. We refer to the analogous expressions for f, (3.33) and (3.34). Also for later reference, we point out that for each linearization point y^k that fulfills $\|\mathbf{u}^k - \bar{\mathbf{u}}\|_W, \|\mathbf{z}^k - \bar{\mathbf{z}}\|_W \le \hat{\omega}$ for an $\hat{\omega} > 0$, there exist constants $c_{\hat{\omega},1}, c_{\hat{\omega},2} > 0$ such that

$$|f_k'(q)d^q| \le c_{\hat{\omega},1}\|d^q\|_Q, \quad |f_k''(q)(d_1^q, d_2^q)| \le c_{\hat{\omega},2}\|d_1^q\|_Q\|d_2^q\|_Q, \quad (4.6)$$

for all $q \in Q_{\mathrm{ad}}$ and $d^q, d_1^q, d_2^q \in Q$. This follows in a straight forward way from Corollary 3.2 and Corollary 3.3 and will be used in Section 7. Note in particular that Corollary 3.3 guarantees that $A(\mathbf{u}^k), R(\varphi^k;\gamma), A'(\mathbf{u}^k)\mathbf{u}^k, R'(\varphi^k;\gamma)\varphi^k$ appearing in (4.2), and $A''(\mathbf{u}^k)[\mathbf{d}^{\mathbf{u},k},\cdot]^*\mathbf{z}^k, \mathbb{R}''(\varphi^k;\gamma)[d^{\varphi,k},\cdot]^* z^{\varphi,k}$ appearing in (4.4) are bounded in the $W^\times$-norm by some constant $c + \hat{c}_{\hat{\omega}}$. Then, by applying Corollary 3.2 first to $\mathbf{u}$ from (4.2) and then to $\mathbf{z}_k$ from (4.4), (4.3) eventually implies

$$|f_k'(q)d^q| = |(\alpha q, d^q) + (z_k(q), d^q)| \le c(\alpha + c + \hat{c}_{\hat{\omega},1})\|d^q\|_Q$$

for all $q \in Q_{\mathrm{ad}}$.

Similarly, f_k' is locally Lipschitz continuous with respect to Q with a Lipschitz constant $L_{\hat{\omega},1} > 0$ depending on $\hat{\omega}$. This follows as for adjoint equations of ($\mathrm{NLP}^{\gamma,\eta}$) shown in [1], Corollary 3.11. More precisely, one first obtains a Lipschitz result for the states from (4.2) from the properties of $\mathcal{G}_{\mathbf{u}^k}$ and Corollary 3.3, and thereafter for the adjoint by analogous arguments. Analogously, the definition of (3.40) combined with the appropriate estimates from Corollary 3.3 yields Lipschitz-continuity for the second derivative. Overall, we subsume that there exists $L_{\hat{\omega},1}, L_{\hat{\omega},2} > 0$, depending on $\hat{\omega}$, such that for all $q_1, q_2 \in Q_{\mathrm{ad}}$ and $d^q, d_1^q, d_2^q \in Q$,

$$|(f_k'(q_1) - f_k'(q_2))d^q| \le L_{\hat{\omega},1}\|d^q\|_Q, \quad |(f_k''(q_1) - f_k''(q_2))[d_1^q, d_2^q]| \le L_{\hat{\omega},2}\|d_1^q\|_Q\|d_q^2\|_Q \quad (4.7)$$

is satisfied. This will also be used in Section 7. For now, assume that $\mathbf{d}^k$ is a solution of (QP_k) or, equivalently, $(\mathbf{u}^{k+1}, q^{k+1})$ is a solution of ($\mathrm{QP}_k^{\mathrm{red}}$). We state first-order necessary optimality conditions for ($\mathrm{QP}_k^{\mathrm{red}}$) for further reference, which are a straightforward adaption of [10], Section 4.2.

Lemma 4.1. *Let $y^k \in Y$ be given and $(\mathbf{u}^{k+1}, q^{k+1})$ be a minimizer of (QP$_k$). Then, there exists an adjoint state $\mathbf{z}^{k+1} = (z^{u,k+1}, z^{\varphi,k+1}) \in W$ such that*

$$A'(\mathbf{u}^k)\mathbf{d}^{\mathbf{u},k} + R'(\varphi^k;\gamma)d^{\varphi,k} = Bq^{k+1} - A(\mathbf{u}^k) - R(\varphi^k;\gamma), \tag{4.8a}$$

$$(A'(\mathbf{u}^k))^*\mathbf{z}^{k+1} + (R'(\varphi^k;\gamma))^* z^{\varphi,k+1} = u^{k+1} - u_d - A''(\mathbf{u}^k)[\mathbf{d}^{\mathbf{u},k},\cdot]^*\mathbf{z}^k - R''(\varphi^k;\gamma)[d^{\varphi,k},\cdot]^* z^{\varphi,k}, \tag{4.8b}$$

$$(B^*\mathbf{z}^{k+1} + \alpha q^{k+1}, q - q^{k+1})_Q \geq 0 \qquad \forall q \in Q_{ad} \tag{4.8c}$$

holds, where $\mathbf{d}^{\mathbf{u},k} = \mathbf{u}^{k+1} - \mathbf{u}^k$.

Remark 4.2. Note that the update direction for the state, $\mathbf{d}^{\mathbf{u},k}$, is used in Lemma 4.1 as a short notation used in the state equation (4.8a) and the adjoint equation (4.8b). Optimality conditions could of course also be written for the optimal update directions, observing that $Bq^{k+1} = Bd^{q,k} + Bq^k$ appears linearly in the state equation (4.8b), and writing the variational inequality in terms of $d^q, d^{q,k}$ and Q^k_{ad} to obtain an equivalent condition to (4.8c). With this in mind, we will, in slight abuse of terms, call q^{k+1} with associated state $\mathbf{u}^{k+1}$ and adjoint state $\mathbf{z}^{k+1}$ a solution of both the reduced problem formulation (QP$_k^{\text{red}}$) and the nonreduced formulation (QP$_k$), and say that a triple $y^{k+1} = (\mathbf{u}^{k+1}, q^{k+1}, \mathbf{z}^{k+1})$ satisfying (4.8a)–(4.8c) satisfies the necessary optimality conditions for both (QP$_k^{\text{red}}$) and (QP$_k$).

The SQP algorithm reads, *cf.* also [10], Algorithm 4.1:

Algorithm 4.3 (SQP algorithm for (NLP$^{\gamma,\eta}$))**.**

(0) Choose $y^0 = (\mathbf{u}^0, q^0, \mathbf{z}^0) \in Y$, and set $k = 0$.
(1) STOP, if $y^k = (\mathbf{u}^k, q^k, \mathbf{z}^k)$ satisfies the first-order necessary optimality conditions of Lemma 3.5.
(2) Solve (QP$_k$) to obtain $\mathbf{d}^k$ with adjoint state $\mathbf{z}^{k+1}$
(3) Set $(\mathbf{u}^{k+1}, q^{k+1}) := (\mathbf{u}^k, q^k) + \mathbf{d}^k$
(4) Set $k = k + 1$ and go to step 1.

We want to point out that convexity of the objective functional at every iterate and therefore unique solvability is still an open question at this point. Solvability of (QP$_k$) even without control bounds has been addressed in [10], Lemma 4.1, under the condition that a strong coercivity property is fulfilled in every iterate $y^k \in Y$, *i.e.* that

$$\exists\, c > 0 \quad \text{such that} \quad \mathcal{L}''(y^k)[\mathbf{d}^k, \mathbf{d}^k] \geq c\|d^{q,k}\|_Q^2 \tag{4.9}$$

holds for all $\mathbf{d}^k \in W \times Q$ that satisfy $\mathbf{d}^{\mathbf{u},k} = \mathcal{G}_{\mathbf{u}^k}(Bd^{q,k} + Bq^k - A(\mathbf{u}^k) - (0,\, R(\varphi^k;\gamma)))$. Transferring such a condition uniformly from one iterate to the next would be rather straight forward using the Lipschitz property of $\mathcal{L}''$, if the algorithm is initialized close to a local minimum fulfilling the strong condition

$$\mathcal{L}''(\bar{y})[\mathbf{d}^k, \mathbf{d}^k] \geq c\|d^{q,k}\|_Q^2, \tag{4.10}$$

with a constant $c > 0$ for all $\mathbf{d}^k \in W \times Q$ that satisfy $\mathbf{d}^{\mathbf{u},k} = \mathcal{G}_{\bar{\mathbf{u}}}(Bd^{q,k} + Bq^k - A(\mathbf{u}^k) - (0,\, R(\varphi^k;\gamma)))$ on the whole control space $Q = L^2(\Gamma)$. It is more involved using the weaker assumption $d^{q,k} \in \mathcal{C}_\sigma(\bar{q})$. To prove convergence under this weaker condition involving the cone of (σ-)critical directions $\mathcal{C}_\sigma(\bar{q})$ from (3.36') requires the discussion of several auxiliary results, a key result being Lemma 5.1 that ensures that (4.9) holds for the linearization points in a certain neighborhood of $\bar{y}$. The remainder of this paper now discusses the following steps: Section 5 contains a first important intermediate local convergence result. First-order necessary optimality conditions are reformulated as a generalized equation, and the SQP-method is interpreted as Newton's method for these conditions. The proof of convergence then relies on showing so called strong regularity of the generalized equations. Since *a priori*, this cannot be shown easily for the subproblems in Algorithm 4.3 under Assumption 3.6, this is done first for auxiliary subproblems. Section 6 then transfers these convergence results to a variant of Algorithm 4.3 where the admissible controls of the subproblems are confined to an $L^\infty-$neighborhood of the locally optimal control $\bar{q}$ satisfying Assumption 3.6. In both Section 5 and 6, we follow a meanwhile classical approach in the convergence of SQP methods, see *e.g.* [16, 28, 29], also [19, 20], yet the application for our specific problem requires the careful application of appropriate regularity results. Finally, in Section 7, we replace L^∞-localization by L^2-localization, following the ideas from [27], based on [33].

5. Convergence analysis for an auxiliary sequence

Following the ideas of [16, 29], we will first show that the SQP method of Algorithm 4.3 corresponds to iteratively applying Newton's method to a generalized equation. In order to do this, we will transform the optimality conditions from Lemma 3.5 and Lemma 4.1 into generalized equations and identify the associated Newton steps. We can then investigate convergence results for Newton's method with auxiliary subproblems.

5.1. Optimality conditions as generalized equation

Let us start by looking at the first-order optimality conditions of (NLP$^{\gamma,\eta}$) from Lemma 3.5. Following [16, 29], we define the generalized equation

$$0 \in F(\bar{y}) + N(\bar{y}), \tag{GE}$$

where the mapping $F\colon Y \longrightarrow Z$ and the set-valued map $N\colon Y \rightrightarrows Z$ are given by

$$F(y) := \begin{pmatrix} A(\mathbf{u}) + R(\varphi;\gamma) - Bq \\ (A'(\mathbf{u}))^*\mathbf{z} + (R'(\varphi;\gamma)^* z^\varphi - u + u_d \\ B^*\mathbf{z} + \alpha q \end{pmatrix}, \quad N(y) := \begin{pmatrix} 0 \\ 0 \\ N_{\text{nc}}(q) \end{pmatrix}. \tag{5.1}$$

Here, $N_{\text{nc}}(q)$ denotes the normal cone of Q_{ad} at a $q \in Q$, *i.e.*

$$N_{\text{nc}}(q) = \{d^q \in Q \quad | \quad (d^q, \tilde{q} - q)_Q \le 0 \quad \text{for all} \quad \tilde{q} \in Q_{\text{ad}}\}.$$

Note that due to the nonlinearity of A and A' with respect to $\mathbf{u}$, (GE) is also nonlinear. The operator F is Fréchet differentiable from Y into Z, with derivative

$$F'(y)\tilde{y} = \begin{pmatrix} A'(\mathbf{u})\tilde{\mathbf{u}} + R'(\varphi;\gamma)\tilde{\varphi} - B\tilde{q} \\ A''(\mathbf{u})[\tilde{\mathbf{u}},\cdot]^*\mathbf{z} + (A'(\mathbf{u}))^*\tilde{\mathbf{z}} + R''(\varphi;\gamma)[\tilde{\varphi},\cdot]^* z^\varphi + (R'(\varphi;\gamma))^*\tilde{z}^\varphi - \tilde{u} \\ B^*\tilde{\mathbf{z}} + \alpha\tilde{q} \end{pmatrix}, \tag{5.2}$$

where $\tilde{y} := (\tilde{\mathbf{u}}, \tilde{q}, \tilde{\mathbf{z}})$. This follows directly from the second-order continuous Fréchet differentiability of A and R and the linearity of B. We can thus apply Newton's method to (GE). Given the function triple $y^k = (\mathbf{u}^k, q^k, \mathbf{z}^k) \in Y$, the next iterate $y^{k+1} = (\mathbf{u}^{k+1}, q^{k+1}, \mathbf{z}^{k+1}) \in Y$ is determined by solving the generalized equation

$$0 \in F(y^k) + F'(y^k)(y^{k+1} - y^k) + N(y^{k+1}). \tag{NM}$$

Writing out the definitions of F, F', and N, we see that (NM) is equivalent to

$$\begin{aligned} A(\mathbf{u}^k) + R(\varphi^k;\gamma) - Bq^{k+1} + A'(\mathbf{u}^k)(\mathbf{u}^{k+1} - \mathbf{u}^k) + R'(\varphi^k;\gamma)(\varphi^{k+1} - \varphi^k) &= 0, \\ (A'(\mathbf{u}^k))^*\mathbf{z}^{k+1} + (R'(\varphi^k;\gamma))^* z^{k+1,\varphi} + A''(\mathbf{u}^k)[\mathbf{u}^{k+1} - \mathbf{u}^k,\cdot]^*\mathbf{z}^k & \\ + R''(\varphi^k;\gamma)[\varphi^{k+1} - \varphi^k,\cdot]^* z^{\varphi,k} - u^{k+1} + u_d &= 0, \\ (B^*\mathbf{z}^{k+1} + \alpha q^{k+1}, q - q^{k+1})_Q &\ge 0 \quad \forall q \in Q_{\text{ad}}, \end{aligned}$$

which is precisely the formulation of (4.8a)–(4.8c), recalling $\mathbf{d}^{\mathbf{u},k} = \mathbf{u}^{k+1} - \mathbf{u}^k$.

The main challenge in the following will be to ensure that the iterates produced by (NM) remain within the local neighborhood where (4.9) holds. This difficulty is tackled by means of auxiliary subproblems in the next subsection.

5.2. An auxiliary subproblem ($\widehat{\mathrm{QP}}_{\boldsymbol{k}}$)

We already pointed out that *a priori* it is not clear whether the directions $\mathbf{d}^k$ produced by Algorithm 4.3 lie in the cone of critical directions from (3.36') that Assumption 3.6 uses. This suggests to look at auxiliary subproblems. We follow the approach of [16] and change the definition of the admissible set in a first step. The auxiliary subproblem, reads

$$\begin{cases} \min\limits_{\mathbf{d}} J_k(\mathbf{d}) = J'(\mathbf{u}^k, q^k)\mathbf{d} + \dfrac{1}{2}\mathcal{L}''(y^k)[\mathbf{d},\mathbf{d}], \\ \quad \text{s. t. } A'(\mathbf{u}^k)\mathbf{d}^{\mathbf{u}} + R'(\varphi^k;\gamma)d^{\varphi} = Bd^q + Bq - A(\mathbf{u}^k) - R(\varphi^k;\gamma), \\ \quad \text{and } d^q + q^k = q \in \widehat{Q_{\mathrm{ad}}} := \{q \in Q_{\mathrm{ad}} \quad | \quad q = \bar{q} \text{ on } \mathcal{I}(\sigma)\}, \end{cases} \tag{$\widehat{\mathrm{QP}}_k$}$$

for which we recall $\mathcal{I}(\sigma) := \{x \in \Omega \;\big|\; |B^*\bar{\mathbf{z}} + \alpha\bar{q}\,| \geq \sigma\}$, *cf.* (3.37). Again we state a reduced version of this auxiliary problem that we will call equivalently. With f_k as in ($\mathrm{QP}_k^{\mathrm{red}}$) and $\mathbf{d} = (\mathbf{d}^{\mathbf{u}}, d^q) = (\mathbf{u} - \mathbf{u}^k, q - q^k) \in W \times Q$, it reads

$$\min_q f_k(q), \quad \text{s.t. } q \in \widehat{Q_{\mathrm{ad}}} := \{q \in Q_{\mathrm{ad}} \quad | \quad q = \bar{q} \text{ on } \mathcal{I}(\sigma)\}, \tag{$\widehat{\mathrm{QP}_k^{\mathrm{red}}}$}$$

The definition of the admissible set immediately implies $d^q \in \mathcal{C}_\sigma(\bar{q})$. To obtain a unique solution $\hat{\mathbf{d}}^k = (\hat{\mathbf{d}}^{\mathbf{u},k}, \hat{d}^{q,k}) = (\hat{\mathbf{u}}^{k+1} - \mathbf{u}^k, \hat{q}^{k+1} - q^k) \in W \times Q$ of ($\widehat{\mathrm{QP}}_k$), we prove a coercivity result, following [16], Lemma 6.2.

Lemma 5.1. *Let $\delta_{SSC} > 0$ be as in Assumption 3.6. Then, there exists a radius $\omega_1 > 0$ such that for all $y^k \in Y$ with $\|\mathbf{u}^k - \bar{\mathbf{u}}\|_W, \|\mathbf{z}^k - \bar{\mathbf{z}}\|_W \leq \omega_1$ the coercivity condition*

$$\mathcal{L}''(y^k)[(\mathbf{d}_k^{\mathbf{u}}, d^q), (\mathbf{d}_k^{\mathbf{u}}, d^q)] \geq \frac{\delta_{SSC}}{2}\|d^q\|_Q^2$$

holds for all $(\mathbf{d}_k^{\mathbf{u}}, d^q) \in W \times Q_{ad}$ that satisfy $\mathbf{d}_k^{\mathbf{u}} = \mathcal{G}_{\mathbf{u}^k}(Bd^q)$ and $d^q = 0$ on $\mathcal{I}(\sigma)$.

Before we start the proof, let us point out that the upcoming analysis, in particular Theorem 5.8, will guarantee that the solution $\hat{y}^{k+1}$ of ($\widehat{\mathrm{QP}}_k$) will again fulfill the prerequisites of Lemma 5.1.

Proof. Let $y^k \in Y$ be fixed but arbitrary with

$$\|\mathbf{u}^k - \bar{\mathbf{u}}\|_W, \|\mathbf{z}^k - \bar{\mathbf{z}}\|_W \leq \omega_1 \tag{5.3}$$

for an $\omega_1 > 0$ to be determined. Moreover, let d^q and $\mathbf{d}_k^{\mathbf{u}}$ be as assumed. We define the auxiliary linearized state $\mathbf{d}^{\bar{\mathbf{u}}} := \mathcal{G}_{\bar{\mathbf{u}}}(Bd^q) = G'(\bar{q})d^q$. For later use, we point out that Lemma 3.1 and Corollary 3.2 guarantee

$$\|\mathbf{d}^{\bar{\mathbf{u}}}\|_W \leq c\|d^q\|_Q, \quad \|\mathbf{d}_k^{\mathbf{u}}\|_W \leq (c + c_{\omega_1,1})\|d^q\|_Q, \tag{5.4}$$

for some $c > 0$ depending on $\bar{y}$ but not on y^k and a constant $c_{\omega_1,1} = \mathcal{O}(\omega_1)$. Moreover, note that

$$\mathbf{d}_k^{\mathbf{u}} - \mathbf{d}^{\bar{\mathbf{u}}} = \mathcal{G}_{\bar{\mathbf{u}}}([A'(\bar{\mathbf{u}}) - A'(\mathbf{u}^k)]\mathbf{d}_k^{\mathbf{u}} + (0,\, [R'(\bar{\varphi};\gamma) - R'(\varphi^k;\gamma)d_k^{\varphi}])).$$

Then Lemma 3.1 and the Lipschitz results from Corollary 3.3 for A' and R' lead to

$$\|\mathbf{d}_k^{\mathbf{u}} - \mathbf{d}^{\bar{\mathbf{u}}}\|_W \le c(\|[A'(\bar{\mathbf{u}}) - A'(\mathbf{u}^k)]\mathbf{d}_k^{\mathbf{u}}\|_{W^\times} + \|[R'(\bar{\varphi};\gamma) - R'(\varphi^k;\gamma)]d_k^{\varphi}\|_q) \le c_{\omega_1,2}\|\mathbf{d}_k^{\mathbf{u}}\|_W, \tag{5.5}$$

with a constant $c_{\omega_1,2} = \mathcal{O}(\omega_1)$. Now, a short calculation shows:

$$\begin{aligned}\mathcal{L}''(y^k)[(\mathbf{d}_k^{\mathbf{u}}, d^q), (\mathbf{d}_k^{\mathbf{u}}, d^q)] &= \mathcal{L}''(\bar{y})[(\mathbf{d}^{\bar{\mathbf{u}}}, d^q), (\mathbf{d}^{\bar{\mathbf{u}}}, d^q)] + [\mathcal{L}''(y^k) - \mathcal{L}''(\bar{y})][(\mathbf{d}^{\bar{\mathbf{u}}}, d^q), (\mathbf{d}^{\bar{\mathbf{u}}}, d^q)]\\ &\quad + \mathcal{L}''(y^k)[(\mathbf{d}_k^{\mathbf{u}}, d^q), (\mathbf{d}_k^{\mathbf{u}}, d^q)] - \mathcal{L}''(y^k)[(\mathbf{d}^{\bar{\mathbf{u}}}, d^q), (\mathbf{d}^{\bar{\mathbf{u}}}, d^q)].\end{aligned} \tag{5.6}$$

We bound all terms from the right-hand side of (5.6) from below. We first recognize that d^q lies in the cone of critical directions $\mathcal{C}_\sigma(\bar{q})$ from (3.36'), therefore Assumption 3.6 implies

$$\mathcal{L}''(\bar{y})[(\mathbf{d}^{\bar{\mathbf{u}}}, d^q), (\mathbf{d}^{\bar{\mathbf{u}}}, d^q)] \ge \delta_{\mathrm{SSC}}\|d^q\|_Q^2. \tag{5.7}$$

The Lipschitz results (3.47) and (3.48) of Corollary 3.9 combined with $W \hookrightarrow V$ and (5.4), lead to

$$[\mathcal{L}''(y^k) - \mathcal{L}''(\bar{y})][(\mathbf{d}^{\bar{\mathbf{u}}}, d^q), (\mathbf{d}^{\bar{\mathbf{u}}}, d^q)] \ge -(c + c_{\omega_1,3})\omega_1\|\mathbf{d}^{\bar{\mathbf{u}}}\|_W^2 \ge -c_{\omega_1,3}\|d^q\|_Q^2 \tag{5.8}$$

with a (generic) constant $c_{\omega_1,3} = \mathcal{O}(\omega_1)$. Estimate (3.49) of Corollary 3.9 guarantees

$$|\mathcal{L}''(y^k)[(\mathbf{d}_k^{\mathbf{u}}, d^q), (\mathbf{d}_k^{\mathbf{u}}, d^q)] - \mathcal{L}''(y^k)[(\mathbf{d}^{\bar{\mathbf{u}}}, d^q), (\mathbf{d}^{\bar{\mathbf{u}}}, d^q)]| \ge -(c + c_{\omega_1,4})(\|\mathbf{d}^{\bar{\mathbf{u}}}\|_W + \|\mathbf{d}_k^{\mathbf{u}}\|_W)\|\mathbf{d}_k^{\mathbf{u}} - \mathbf{d}^{\bar{\mathbf{u}}}\|_W \tag{5.9}$$

with a constant $c_{\omega_1,4} = \mathcal{O}(\omega_1)$. Estimates (5.4) and (5.5) immediately imply

$$|\mathcal{L}''(y^k)[(\mathbf{d}_k^{\mathbf{u}}, d^q), (\mathbf{d}_k^{\mathbf{u}}, d^q)] - \mathcal{L}''(y^k)[(\mathbf{d}^{\bar{\mathbf{u}}}, d^q), (\mathbf{d}^{\bar{\mathbf{u}}}, d^q)]| \ge -(c + c_{\omega_1,3})(c + c_{\omega_1,1})^2 c_{\omega_1,4}\|d^q\|_Q^2. \tag{5.10}$$

Finally, collecting the estimates (5.7), (5.8), and (5.10) and inserting them into (5.6) eventually leads to

$$\mathcal{L}''(y^k)[(d_k^{\mathbf{u}}, d^q), (d_k^{\mathbf{u}}, d^q)] \ge (\delta_{\mathrm{SSC}} - c(\omega_1))\,\|d^q\|_Q^2$$

with $c(\omega_1) = \mathcal{O}(\omega_1)$ (for $\omega_1 \to 0$). Choosing ω_1 sufficiently small yields $\delta_{\mathrm{SSC}} - c(\omega_1) \ge \frac{\delta_{\mathrm{SSC}}}{2}$, which concludes the proof. ☐

An existence and uniqueness result for solutions to ($\widehat{\mathrm{QP}}_k$) now follows from the properties of $\widehat{Q_{\mathrm{ad}}}$, *i.e.* bounded, closed and convex and therefore weakly sequentially compact, and the strict convexity of the objective functional of ($\widehat{\mathrm{QP}}_k$) that is implied by the coercivity condition shown in Lemma 5.1 as in [57], Lemma 4.1.

Corollary 5.2. *Let $y^k \in Y$ fulfill $\|\mathbf{u}^k - \bar{\mathbf{u}}\|_W, \|\mathbf{z}^k - \bar{\mathbf{z}}\|_W \le \omega_1$ for a sufficiently small $\omega_1 > 0$. Then ($\widehat{\mathrm{QP}}_k$) has a unique solution $\hat{q}^{k+1} \in \widehat{Q_{ad}}$ with associated $\hat{\mathbf{u}}^{k+1} \in W$.*

Proof. For any $q \in \widehat{Q_{\mathrm{ad}}}$ and associated direction $d^q = q - q^k$ we split $\mathbf{d}^{\mathbf{u}} = \mathcal{G}_{\mathbf{u}^k}(Bd^q + Bq^k - A(\mathbf{u}^k) - (0,\, R(\varphi^k;\gamma)))$ into $\mathbf{d}^{\mathbf{u}} = \hat{\mathbf{u}} + \tilde{\mathbf{u}}$, where $\hat{\mathbf{u}} := \mathcal{G}_{\mathbf{u}^k}(Bd^q)$ and $\tilde{\mathbf{u}} := \mathcal{G}_{\mathbf{u}^k}(Bq^k - A(\mathbf{u}^k) - (0,\, R(\varphi^k;\gamma)))$, *i.e.* $\tilde{\mathbf{u}}$ does not depend on the control direction d^q. The objective functional of ($\widehat{\mathrm{QP}}_k$) can now be rewritten into

$$\begin{aligned}J_k(\mathbf{u}, q) &= J'(\mathbf{u}^k, q^k)(\hat{\mathbf{u}} + \tilde{\mathbf{u}}, d^q) + \frac{1}{2}\mathcal{L}''(y^k)[(\hat{\mathbf{u}} + \tilde{\mathbf{u}}, d^q), (\hat{\mathbf{u}} + \tilde{\mathbf{u}}, d^q)]\\ &= \frac{1}{2}\mathcal{L}''(\bar{y})[(\hat{\mathbf{u}}, d^q), (\hat{\mathbf{u}}, d^q)]\\ &\quad + \mathcal{L}''(\bar{y})[(\hat{\mathbf{u}}, 0), (\tilde{\mathbf{u}}, 0)] + J'(\mathbf{u}^k, q^k)(\hat{\mathbf{u}}, d^q)\\ &\quad + \frac{1}{2}\mathcal{L}''(y^k)[(\tilde{\mathbf{u}}, 0), (\tilde{\mathbf{u}}, 0)] + J'(\mathbf{u}^k, q^k)(\tilde{\mathbf{u}}, 0).\end{aligned}$$

Note that $(\hat{\mathbf{u}}, d^q)$ belongs to the subspace where Lemma 5.1 applies since $d^q = 0$ on $\mathcal{I}(\sigma)$ for all admissible $q \in \widehat{Q_{\mathrm{ad}}}$. Thus, the term in the first line of the right-hand side is coercive. The terms in the second line build a linear functional in $(\mathbf{u}, d^q)$, and the terms in the third line are independent of d^q. Thus the objective functional J_k is strictly convex. The set $\widehat{Q_{\mathrm{ad}}}$ is uniformly compact in Q, the claim now follows from standard arguments, *cf.* [58], Theorem 2.14. □

First-order optimality conditions for ($\widehat{\mathrm{QP}_k}$) follow in a standard way as well. Under the closeness condition of Lemma 5.1 they are also sufficient due to convexity. We omit the proof.

Corollary 5.3. *Let $y^k \in Y$ with $\|\mathbf{u}^k - \bar{\mathbf{u}}\|_W \le \omega_1$ and $\|\mathbf{z}^k - \bar{\mathbf{z}}\|_W \le \omega_1$ with ω_1 as in Lemma 5.1 be given. The pair $(\hat{\mathbf{u}}^{k+1}, \hat{q}^{k+1}) = (\mathbf{u}^k, q^k) + \hat{\mathbf{d}}^k$ with $\hat{q}^{k+1} \in \widehat{Q_{ad}}$ is a minimizer of ($\widehat{\mathrm{QP}_k^{\mathrm{red}}}$) if and only if there exists an adjoint state $\mathbf{z}^{k+1} = (z^{u,k+1}, z^{\varphi,k+1}) \in W$ such that the optimality system*

$$A'(\mathbf{u}^k)\hat{\mathbf{d}}^{\mathbf{u},k} + R'(\varphi^k;\gamma)\hat{d}^{\varphi,k} = B\hat{q}^{k+1} - A(\mathbf{u}^k) - R(\varphi^k;\gamma), \tag{5.11a}$$

$$(A'(\mathbf{u}^k))^*\hat{\mathbf{z}}^{k+1} + (R'(\varphi^k;\gamma))^*\hat{z}^{\varphi,k+1} = \hat{u}^{k+1} - u_d - A''(\mathbf{u}^k)[\hat{\mathbf{d}}^{\mathbf{u},k},\cdot]^*\mathbf{z}^k - R''(\varphi^k;\gamma)[\hat{d}^{\varphi,k},\cdot]^* z^{\varphi,k}, \tag{5.11b}$$

$$(B^*\hat{\mathbf{z}}^{k+1} + \alpha\hat{q}^{k+1}, q - \hat{q}^{k+1})_Q \ge 0 \qquad \forall q \in \widehat{Q_{ad}}. \tag{5.11c}$$

is satisfied.

Analogously to the course of action at the end of Section 5.1, we observe the equivalence of the optimality system of Corollary 5.3 with a generalized equation, which is introduced as

$$0 \in F(y) + \hat{N}(y), \tag{$\widehat{\mathrm{GE}}$}$$

where F is given as in (5.1) and $\hat{N}(y) := (0, 0, \hat{N}_{\mathrm{nc}}(q))^T$, with $\hat{N}_{\mathrm{nc}}(q)$ defined by

$$\hat{N}_{\mathrm{nc}}(q) = \{d^q \in Q \quad | \quad (d^q, \tilde{q} - q)_Q \le 0 \quad \text{for all} \quad \tilde{q} \in \widehat{Q_{\mathrm{ad}}}\}.$$

Formally, the Newton subproblem associated to ($\widehat{\mathrm{GE}}$) reads: Given the function triple $y^k \in Y$, the next iterate $\hat{y}^{k+1}$ is determined by solving the generalized equation

$$0 \in F(y^k) + F'(y^k)(\hat{y}^{k+1} - y^k) + \hat{N}(\hat{y}^{k+1}). \tag{$\widehat{\mathrm{NM}}$}$$

Convergence of the sequence generated by ($\widehat{\mathrm{NM}}$) is addressed within the remainder of this chapter.

5.3. The strong regularity property

We now prove local quadratic convergence of the sequence of solutions generated by ($\widehat{\mathrm{NM}}$), or equivalently by ($\widehat{\mathrm{QP}_k}$). As in [16, 20], we use a Newton-Kantorovich like convergence theorem, following the approach of [59]. It ensures that the generated sequence $\{\hat{q}^k\} \subset \widehat{Q_{\mathrm{ad}}}$ of global solutions of ($\widehat{\mathrm{QP}_k}$) is well-defined and stays in the convergence radius of Newton's method, if started from a good initial guess, and ensures local quadratic convergence to $\bar{q}$.

Let us continue with some notation. Firstly, let $y_\delta = (\mathbf{u}_\delta, q_\delta, \mathbf{z}_\delta)$ and $y_{\delta'} = (\mathbf{u}_{\delta'}, q_{\delta'}, \mathbf{z}_{\delta'})$, and let $\delta := (\delta_1, \delta_2, \delta_3)$ and $\delta' := (\delta'_1, \delta'_2, \delta'_3)$ denote triples of perturbations that lie in the spaces Z or Z_∞, *cf.* (2.1). For radii $\rho > 0$ and $r > 0$, we write $B_r^Z(\delta')$, $B_r^{Z_\infty}(\delta')$, $B_\rho^Y(\tilde{y})$ and $B_\rho^{Y_\infty}(\tilde{y})$ for the open balls

$$B_r^Z(\delta') := \{\delta \in Z \quad | \quad \|\delta - \delta'\|_Z < r\}, \qquad B_r^{Z_\infty}(\delta') := \{\delta \in Z_\infty \quad | \quad \|\delta - \delta'\|_{Z_\infty} < r\},$$
$$B_\rho^Y(\tilde{y}) := \{y \in Y \quad | \quad \|y - \tilde{y}\|_Y < \rho\}, \qquad B_\rho^{Y_\infty}(\tilde{y}) := \{y \in Y_\infty \quad | \quad \|y - \tilde{y}\|_{Y_\infty} < \rho\}.$$

We will often set $\delta' = 0$. With this, we can state the definition of the strong regularity property, see [31, 32].

Definition 5.4.

- We say that the generalized equation ($\widehat{\text{GE}}$) has the strong regularity property in the space Y_∞ at $\bar{y}$ if there exist radii $\rho, r > 0$ and a constant $L_\infty > 0$, such that for all perturbations $\delta \in B_r^{Z_\infty}(0)$ the perturbed generalized equation
$$\delta \in F(\bar{y}) + F'(\bar{y})(y_\delta - \bar{y}) + \hat{N}(y_\delta) \tag{5.12}$$
fulfills the following properties:
  1. The perturbed generalized equation (5.12) has a solution $y_\delta \in B_\rho^{Y_\infty}(\bar{y})$.
  2. y_δ is the only solution of (5.12) in $B_\rho^{Y_\infty}(\bar{y})$.
  3. Let $y_\delta, y_{\delta'}$ be the unique solutions to (5.12) in $B_\rho^{Y_\infty}(\bar{y})$ for $\delta, \delta' \in B_r^{Z_\infty}(0)$. Then the Lipschitz condition
  $$\|y_\delta - y_{\delta'}\|_{Y_\infty} \leq L_\infty \|\delta - \delta'\|_{Z_\infty}$$
  holds.
- The strong regularity property in Y is defined analogously, with the spaces Y_∞ and Z_∞ replaced by Y and Z, and for a constant $L_2 > 0$, respectively.

At this point, we remind the reader of the definitions and embeddings of the spaces Y, Y_∞ as well as Z, Z_∞. In particular, note that functions in the space W admit L^∞-regularity, and perturbations in L^2 are contained in $W^\times$. Let us put on record that $\delta \in Z$ or $\delta \in Z_\infty$, respectively, only enters (5.12) linearly. Further, note that $\bar{y}$ is a solution to both ($\widehat{\text{GE}}$) as well as (5.12) for $\delta = 0$. Closely following [16], we point out that (5.12) is exactly the perturbation of the linearization of ($\widehat{\text{GE}}$) in $\bar{y}$, *i.e.* a perturbation of ($\widehat{\text{NM}}$) in this function triple. Let us recall this generalized equation for further reference, and from now on also call it

$$\delta \in F(\bar{y}) + F'(\bar{y})(y_\delta - \bar{y}) + \hat{N}(y_\delta), \tag{$\widehat{\text{NM}_\delta}$}$$

for a $\delta \in Z$ or $\delta \in Z_\infty$, respectively. It is clear that ($\widehat{\text{NM}_\delta}$) is the first-order necessary condition for an auxiliary perturbed subproblem with perturbations $\delta = (\delta_1, \delta_2, \delta_3) \in Z$ given (again with a short notation for directions $\mathbf{d} := (\mathbf{d^u}, d^q) := (\mathbf{u} - \bar{\mathbf{u}}, q - \bar{q})$) by:

$$\begin{cases} \min\limits_{\mathbf{d}} J_\delta(\mathbf{u}, q) := J'(\bar{\mathbf{u}}, \bar{q})\mathbf{d} + \dfrac{1}{2}\mathcal{L}''(\bar{y})[\mathbf{d}, \mathbf{d}] \\ \qquad - \langle \delta_1, \mathbf{d^u} \rangle_{(W^{1,p'} \times L^{q'})^*, W^{1,p} \times L^q} - (\delta_2, d^q)_Q, \\ \quad \text{s. t. } A'(\bar{\mathbf{u}})\mathbf{d^u} + R'(\bar{\varphi}; \gamma)d^\varphi = Bq + \delta_3 - A(\bar{\mathbf{u}}) - R(\bar{\varphi}; \gamma) \\ \quad \text{and } q \in \widehat{Q_{\text{ad}}}. \end{cases} \tag{$\widehat{\text{QP}}_\delta$}$$

Unique solvability can be shown precisely as in [57], Lemma 4.1 and demonstrated in Corollary 5.2 by splitting $\mathbf{d^u} = \hat{\mathbf{u}} + \tilde{\mathbf{u}}$, where $\hat{\mathbf{u}} = \mathcal{G}_{\bar{\mathbf{u}}}(Bd^q)$ and $\tilde{\mathbf{u}} = \mathcal{G}_{\bar{\mathbf{u}}}(B\bar{q} - A(\bar{\mathbf{u}}) - (0,\ R(\bar{\varphi}; \gamma)) + \delta_3)$, *i.e.* $\tilde{\mathbf{u}}$ does not depend on the control d^q. The objective functional of ($\widehat{\text{QP}}_\delta$) can then also be split into a constant part, a part that depends linearly on d^q, and a quadratic, coercive part. We omit the details, referring to techniques in *e.g. cf.* [20], Lemma 5.1.

Lemma 5.5. *For each $\delta \in Z$, ($\widehat{\text{QP}}_\delta$) has a unique solution $(\hat{\mathbf{d}}_\delta^{\mathbf{u}}, \hat{d}_\delta^q) \in W \times Q_{ad}$ with $\hat{q}_\delta = \bar{q} + \hat{d}^\delta \in \widehat{Q_{ad}}$, depending on δ.*

First-order necessary conditions for ($\widehat{\text{QP}}_\delta$) can be shown in a standard way.

Corollary 5.6. *Let $\delta \in Z$ be given. A control direction $\hat{d}^q_\delta \in \widehat{Q_{ad}}$ (with control $\hat{q}_\delta = \hat{d}^q_\delta + \bar{q}$) and associated optimal state $\hat{\mathbf{u}}_\delta = \hat{\mathbf{d}}^{\mathbf{u}}_\delta + \bar{u}$ and adjoint state $\hat{\mathbf{z}}_\delta = (\hat{z}^u_\delta, \hat{z}^\varphi_\delta)$, is optimal for the subproblem ($\widehat{\mathrm{QP}_\delta}$) if and only if the optimality system*

$$A'(\bar{\mathbf{u}})\hat{\mathbf{d}}^{\mathbf{u}}_\delta + R'(\bar{\varphi};\gamma)\hat{d}^\varphi_\delta = B\hat{q}_\delta - A(\bar{\mathbf{u}}) - R(\bar{\varphi};\gamma) + \delta_3, \tag{5.13a}$$

$$(A'(\bar{\mathbf{u}}))^*\hat{\mathbf{z}}_\delta + (R'(\bar{\varphi};\gamma))^*\hat{z}^\varphi_\delta = \hat{u}_\delta - u_d - A''(\bar{\mathbf{u}})[\hat{\mathbf{d}}^{\mathbf{u}}_\delta,\cdot]^*\bar{\mathbf{z}} - R''(\bar{\varphi};\gamma)[\hat{d}^\varphi_\delta,\cdot]^*\bar{z}^\varphi - \delta_1, \tag{5.13b}$$

$$(B^*\hat{\mathbf{z}}_\delta + \alpha\hat{q}_\delta - \delta_2, q - \hat{q}_\delta) \geq 0 \qquad \forall q \in \widehat{Q_{ad}}. \tag{5.13c}$$

is satisfied.

Obviously Lemma 5.5 and Corollary 5.6 remain valid for perturbations $\delta \in Z_\infty \subset Z$. We will now show the Lipschitz condition from Definition 5.4 in both the L^2 and the L^∞-setting. We closely follow the proof of [57], Theorems 4.2 and 5.2.

Lemma 5.7. *Let $\hat{q}_\delta$ and $\hat{q}_{\delta'}$ be the the unique optimal controls of ($\widehat{\mathrm{QP}_\delta}$) for perturbations $\delta, \delta' \in Z$, with $\hat{\mathbf{u}}_\delta$, $\hat{\mathbf{u}}_{\delta'}$ and $\hat{\mathbf{z}}_{\delta'}, \hat{\mathbf{z}}_\delta$ satisfying (5.13a) and (5.13b), respectively. Let $\hat{y}_\delta$ and $\hat{y}_{\delta'}$ denote the associated function triples. There exists a constant $L_2 > 0$, such that*

$$\|\hat{y}_\delta - \hat{y}_{\delta'}\|_Y \leq L_2\|\delta - \delta'\|_Z. \tag{5.14}$$

If further $\delta, \delta' \in Z_\infty$, then there exists a constant $L_\infty > 0$ such that

$$\|\hat{y}_\delta - \hat{y}_{\delta'}\|_{Y_\infty} \leq L_\infty\|\delta - \delta'\|_{Z_\infty}. \tag{5.15}$$

Proof. We start with the proof of (5.14) for perturbations $\delta, \delta' \in Z$. Let $\mathbf{u} := \hat{\mathbf{u}}_\delta - \hat{\mathbf{u}}_{\delta'}$, $q := \hat{q}_\delta - \hat{q}_{\delta'}$, $\mathbf{z} := \hat{\mathbf{z}}_\delta - \hat{\mathbf{z}}_{\delta'}$, and note that (5.13a) and (5.13b) imply

$$A'(\bar{\mathbf{u}})\mathbf{u} + R'(\bar{\varphi};\gamma)\varphi = Bq + \delta_3 - \delta_3', \tag{5.16}$$

$$(A'(\bar{\mathbf{u}}))^*\mathbf{z} + R'(\bar{\varphi};\gamma)^*z^\varphi = -A''(\bar{\mathbf{u}})[\mathbf{u},\cdot]^*\bar{\mathbf{z}} - R''(\bar{\varphi};\gamma)[\varphi,\cdot]^*\bar{z}^\varphi + u - (\delta_1 - \delta_1'). \tag{5.17}$$

We apply Lemma 3.1 and obtain

$$\|\mathbf{u}\|_W \leq c\|q\|_Q + c\|\delta_3 - \delta_3'\|_{W^\times} \leq c\|q\|_Q + c\|\delta - \delta'\|_Z, \tag{5.18}$$

$$\begin{aligned} \|\mathbf{z}\|_W &\leq \| - A''(\bar{\mathbf{u}})[\mathbf{u},\cdot]^*\bar{\mathbf{z}} - R''(\bar{\varphi};\gamma)[\varphi,\cdot]^*\bar{z}^\varphi + u - (\delta_1 - \delta_1')\|_{W^\times} \\ &\leq 2c\|\bar{\mathbf{u}}\|_W\|\mathbf{u}\|_W\|\bar{\mathbf{z}}\|_W + \|\mathbf{u}\|_W + \|\delta_1 - \delta_1'\|_{W^\times}. \\ &\leq c\|\mathbf{u}\|_W + c\|\delta_1 - \delta_1'\|_{W^\times} \leq c\|q\|_Q + c\|\delta - \delta'\|_Z, \end{aligned} \tag{5.19}$$

due to (3.12) and (3.13). The generic constant $c > 0$ depends on $\bar{\mathbf{u}}$ and $\bar{\mathbf{z}}$ but not on $\mathbf{u}$. We test (5.16) with $\mathbf{z}$ and (5.17) with $\mathbf{u}$ and take the sum to obtain

$$(q, B^*\mathbf{z}) + \langle\delta_3 - \delta_3', \mathbf{z}\rangle = -\langle A''(\bar{\mathbf{u}})[\mathbf{u},\mathbf{u}], \bar{\mathbf{z}}\rangle - \langle R''(\bar{\varphi};\gamma)[\varphi,\varphi], \bar{z}^\varphi\rangle + (u,u) - \langle\delta_1 - \delta_1', \mathbf{u}\rangle. \tag{5.20}$$

Testing (5.13c) once in $(\hat{\mathbf{z}}_\delta, \hat{q}_\delta)$ with $\hat{q}_{\delta'}$ and once in $(\hat{\mathbf{z}}_{\delta'}, \hat{q}_{\delta'})$ with $\hat{q}_\delta$, and adding both inequalities, leads to

$$(\delta_2 - \delta_2', q) - \alpha(q,q) \geq (B^*\mathbf{z}, q). \tag{5.21}$$

Inserting (5.20) and (5.21) into the second derivative of the Lagrangian (3.40) leads to

$$\mathcal{L}''(\bar{y})[(\mathbf{u},q),(\mathbf{u},q)] \leq \langle \delta_1 - \delta_1', \mathbf{u}\rangle + (\delta_2 - \delta_2', q) + \langle \delta_3 - \delta_3', \mathbf{z}\rangle. \tag{5.22}$$

We split $\mathbf{u}$ into $\mathbf{u} = \hat{\mathbf{u}} + \mathbf{u}_\delta$, where

$$A'(\bar{\mathbf{u}})\hat{\mathbf{u}} + R'(\bar{\varphi};\gamma)\hat{\varphi} = Bq, \qquad A'(\bar{\mathbf{u}})\mathbf{u}_\delta + R'(\bar{\varphi};\gamma)\varphi_\delta = \delta_3 - \delta_3'$$

and note that Lemma 3.1 ensures

$$\|\hat{\mathbf{u}}\|_W \leq c\|q\|_Q, \qquad \|\mathbf{u}_\delta\|_W \leq c\|\delta_3 - \delta_3'\|_{W^\times}, \tag{5.23}$$

where again $c > 0$ is a constant that depends on $\bar{\mathbf{u}}$ and $\bar{\mathbf{z}}$. A calculation similar to the proof of Lemma 5.1 yields

$$\begin{aligned}\mathcal{L}''(\bar{y})[(\hat{\mathbf{u}}+\mathbf{u}_\delta,q),(\hat{\mathbf{u}}+\mathbf{u}_\delta,q)] = &\mathcal{L}''(\bar{y})[(\hat{\mathbf{u}},q),(\hat{\mathbf{u}},q)] + 2(\hat{u},u_\delta) + \|u_\delta\|^2 - 2\langle A''(\bar{\mathbf{u}})[\mathbf{u}_\delta,\hat{\mathbf{u}}],\bar{\mathbf{z}}\rangle \\ &- \langle A''(\bar{\mathbf{u}})[\mathbf{u}_\delta,\mathbf{u}_\delta],\bar{\mathbf{z}}\rangle - 2\langle R''(\bar{\varphi};\gamma)[\varphi_\delta,\hat{\varphi}],\bar{z}^\varphi\rangle - \langle R''(\bar{\varphi};\gamma)[\varphi_\delta,\varphi_\delta]\bar{z}^\varphi\rangle, \\ \geq &\,\delta_{\mathrm{SSC}}\|q\|_Q^2 - c\|\delta_3 - \delta_3'\|_{W^\times}\|q\|_Q - c\|\delta_3 - \delta_3'\|_{W^\times}^2,\end{aligned}$$

where we have used the definition of $\mathcal{L}''$ from (3.40), Assumption 3.6, and the boundedness estimates for A'' and R'' from (3.12) and (3.13) combined with (3.1). Combining the last estimate with (5.22) leads to

$$\begin{aligned}\delta_{\mathrm{SSC}}\|q\|_Q^2 &\leq c\|\delta_3 - \delta_3'\|_{W^\times}\|q\|_Q + c\|\delta_3 - \delta_3'\|_{W^\times}^2 + \langle \delta_1 - \delta_1', \mathbf{u}\rangle + (\delta_2 - \delta_2', q) + \langle \delta_3 - \delta_3', \mathbf{z}\rangle \\ &\leq c\|\delta - \delta'\|_Z(\|\mathbf{u}\|_W + \|q\|_Q + \|\mathbf{z}\|_W) + c\|\delta - \delta'\|_Z^2.\end{aligned}$$

Inserting the estimates (5.18) and (5.19) for $\mathbf{u}$ and $\mathbf{z}$ and using Young's inequality finally results in

$$\|q\|_Q \leq c\|\delta - \delta'\|_Z, \tag{5.24}$$

for a constant $c > 0$ independent of δ, δ'. Applying (5.18) and (5.19) now concludes the proof of (5.14).

Let now $\delta, \delta' \in Z_\infty$. Since $Z_\infty \hookrightarrow Z$, $W_u \hookrightarrow L^\infty(\Omega;\mathbb{R}^2)$, $W_\varphi \hookrightarrow L^\infty(\Omega)$ we conclude from (5.14) that

$$\|\mathbf{u}\|_\infty \leq c\|\mathbf{u}\|_W \leq c\|\delta - \delta'\|_Z, \quad \text{as well as} \quad \|\mathbf{z}\|_\infty \leq c\|\mathbf{z}\|_W \leq c\|\delta - \delta'\|_Z \tag{5.25}$$

hold. We then recognize that $q_\delta = q_{\delta'} = \bar{q}$ on $\mathcal{I}(\sigma)$, and that the variational inequalities of ($\widehat{\mathrm{QP}}_\delta$) for q_δ and $\hat{q}_\delta$ imply

$$q_\delta = \frac{1}{\alpha}\mathrm{Proj}_{Q_{ad}}(B^*\mathbf{z}_\delta - \delta_2), \quad \hat{q}_\delta = \frac{1}{\alpha}\mathrm{Proj}_{Q_{ad}}(B^*\hat{\mathbf{z}}_\delta - \delta_2) \text{ on } \Gamma\setminus\mathcal{I}(\sigma),$$

with $B^*\mathbf{z}_\delta, B^*\mathbf{z}_{\delta'} \in L^\infty(\Gamma)$. Due to the Lipschitz continuity of the pointwise projection this leads to

$$\|q\|_{L^\infty(\Gamma)} \leq c\|B^*\mathbf{z}\|_{L^\infty(\Gamma)} + c\|\delta_2 - \delta_2'\|_{L^\infty(\Gamma)} \leq c(\|\mathbf{z}\|_W + \|\delta_2 - \delta_2'\|_{L^\infty(\Gamma)}). \tag{5.26}$$

Combining this with (5.25) finalizes the proof. □

At this point, we have established all properties of Definition 5.4 for ($\widehat{\mathrm{NM}}_\delta$) in both Y and Y_∞. In summary, we have shown:

Theorem 5.8. *The generalized equation* ($\widehat{\mathrm{GE}}$) *is strongly regular at* $\bar{y}$ *in both* Y *and* Y_∞.

5.4. Convergence of ($\widehat{\text{NM}}$)

Let us now turn to the proof of convergence for the sequences $\{\hat{q}^k\}$ generated by ($\widehat{\text{NM}}$), or equivalently by solving ($\widehat{\text{QP}}_k$). Due to the strong regularity of ($\widehat{\text{GE}}$) in both Y and Y_∞, we can make use of a generalization of the implicit function theorem. The proof relies on standard arguments, see *e.g.* [15, 19, 20, 59]. For completeness, we recapitulate the main arguments of the proof given in [19], Theorem 7.1, adapting them to the operator F and its properties, some of them postponed to the Appendix. Before doing so, let us point out that most classical proofs guarantee convergence if the guess for control and state is taken close to the local minimizer $\bar{q}$ with associated state $\bar{\mathbf{u}}$ and adjoint state $\bar{\mathbf{z}}$, in a norm implied by the strong regularity property in Y or Y_∞. Yet in fact, as observed in [27] the control $\hat{q}^k$ from the previous iteration cancels out in the update directions resulting from solving the subproblems ($\widehat{\text{QP}}_k$), *cf.* Corollary 5.3, so in particular the first iterate $\hat{y}^1$ is independent of the choice of q^0.

Moreover, while closeness condition on $\mathbf{u}^0$ and $\mathbf{z}^0$ in the W-norm seems rather strong, the Lipschitz results (3.4) of the control to state operator actually guarantee that if $\|q^0 - \bar{q}\|_Q$ is chosen small enough, $\mathbf{u}^0 := G(q^0)$ will automatically fulfill a closeness property in the space W. The Lipschitz result for linearized states (3.5) can be applied to adjoint equations, see also [1], Corollary 3.11, and $\mathbf{z}^0 := G'(q^0)^*((u^0 - u_d, 0))$ will automatically fulfill the required closeness property $\|\bar{\mathbf{z}} - \mathbf{z}^0\|_W \le c\|\bar{\mathbf{u}} - \mathbf{u}^0\|_W \le c\|q^0 - \bar{q}\|_Q$ in the space W.

Theorem 5.9. *There exists a radius $\omega_2 > 0$ and a constant $C_N > 0$, such that for each starting point $y^0 \in Y_\infty$ with $\|\mathbf{u}^0 - \bar{\mathbf{u}}\|_W, \|\mathbf{z}^0 - \bar{\mathbf{z}}\|_W \le \omega_2$ and $q^0 \in Q_{ad}$, the auxiliary subproblem ($\widehat{\text{QP}}_k$) generates a unique sequence of iterates $\{\hat{y}^k\}_{k\in\mathbb{N}}$ with $\|\hat{y}^k - \bar{y}\|_{Y_\infty} \le \omega_2$ and $\hat{q}^k \in \widehat{Q_{ad}}$ that satisfies*

$$\|\hat{y}^{k+1} - \bar{y}\|_{Y_\infty} \le C_N \|\hat{y}^k - \bar{y}\|_{Y_\infty}^2 \qquad \text{for all } k \ge 1. \tag{5.27}$$

In particular, we can choose $q^0 \in Q_{ad}$ and $\|q^0 - \bar{q}\|_Q \le \hat{\epsilon}$ sufficiently small, and $\mathbf{u}^0 := G(q^0)$ and $\mathbf{z}^0 := G'(q^0)^*((u^0 - u_d, 0))$.

An analogous convergence result holds for a radius $\omega_{2,2} > 0$ and a $C_{N,2} > 0$ if we replace Y_∞ with Y and Z_∞ with Z.

Proof. By Theorem 5.8, ($\widehat{\text{GE}}$) is strongly regular in Y_∞. Further, note the auxiliary Lipschitz result for F from Lemma A.1. As a result, we can utilize Dontchev's implicit function theorem for generalized equations, *cf.* [32], Theorem 2.4, which ensures the existence of $\rho_1, \rho_2 > 0$, such that for any $y^k = \hat{y}^k \in B_{\rho_1}^{Y_\infty}(\bar{y})$, there exists a unique solution $\hat{y}^{k+1} \in B_{\rho_2}^{Y_\infty}(\bar{y})$ to ($\widehat{\text{NM}}$). Note again that the update direction from Problem ($\widehat{\text{QP}}_k$) is independent of q^k, *cf.* [27], Theorem 6.2 (3) and the Lipschitz properties of G and $G'(\bar{q})^*$.

If ρ is chosen such that $0 < \rho \le \rho_1$, we obtain

$$0 \in F(\bar{y}) + F'(\bar{y})(\bar{y} - \bar{y}) + N(\bar{y}), \tag{5.28}$$

$$0 \in F(\hat{y}^k) + F'(\hat{y}^k)(\hat{y}^{k+1} - \hat{y}^k) + N(\hat{y}^{k+1}). \tag{5.29}$$

Adding and subtracting $F(\bar{y})$ and $F'(\bar{y})(\hat{y}^{k+1} - \bar{y})$ to (5.29), leads to

$$\delta^{k+1} \in F(\bar{y}) + F'(\bar{y})(\hat{y}^{k+1} - \bar{y}) + N(\hat{y}^{k+1}), \tag{5.30}$$

where δ^{k+1} is defined as

$$\delta^{k+1} := F(\bar{y}) - F(\hat{y}^k) + F'(\bar{y})(\hat{y}^{k+1} - \bar{y}) - F'(\hat{y}^k)(\hat{y}^{k+1} - \hat{y}^k). \tag{5.31}$$

Note that we can apply the Lipschitz result of Lemma A.1 to (5.31), which then yields

$$\|\delta^{k+1}\|_{Z_\infty} \le L\left(\|\hat{\mathbf{u}}^k - \bar{\mathbf{u}}\|_W + \|\hat{\mathbf{z}}^k - \bar{\mathbf{z}}\|_W\right) \le L\,\rho, \tag{5.32}$$

for a constant $L > 0$ depending only on the radii ρ_1 and ρ_2. Next, we recognize that (5.28) and (5.30) are equivalent to the first-order necessary conditions of ($\widehat{\mathrm{QP}}_\delta$) for $\delta = 0$ and $\delta = \delta^{k+1}$, respectively. Therefore, from Lemma 5.7 we see

$$\|\hat{y}^{k+1} - \bar{y}\|_{Y_\infty} \leq L_\infty \|\delta^{k+1}\|_{Z_\infty}, \tag{5.33}$$

and combined with (5.32) this leads to

$$\|\hat{y}^{k+1} - \bar{y}\|_{Y_\infty} \leq \; L_\infty L\, \rho. \tag{5.34}$$

To obtain a quadratic convergence result, we estimate $\|\delta^{k+1}\|_{Z_\infty}$ further. Its definition (5.31) yields

$$\|\delta^{k+1}\|_{Z_\infty} \leq \|F(\bar{y}) - F(\hat{y}^k) - F'(\hat{y}^k)(\bar{y} - \hat{y}^k)\|_{Z_\infty} + \|(F'(\bar{y}) - F'(\hat{y}^k))(\hat{y}^{k+1} - \bar{y})\|_{Z_\infty}. \tag{5.35}$$

The estimation of the right-hand side is postponed to the appendix. Note that $\|\hat{y}^k - \bar{y}\|_{Y_\infty} \leq \rho_1$, thus applying Lemma A.3 and Lemma A.2 yields

$$\begin{aligned}\|\delta^{k+1}\|_{Z_\infty} &\leq c_1(\|\hat{\mathbf{u}}^k - \bar{\mathbf{u}}\|_W^2 + \|\hat{\mathbf{z}}^k - \bar{\mathbf{z}}\|_W^2) + c_2(\|\hat{\mathbf{u}}^k - \bar{\mathbf{u}}\|_W + \|\hat{\mathbf{z}}^k - \bar{\mathbf{z}}\|_W)(\|\hat{\mathbf{u}}^{k+1} - \bar{\mathbf{u}}\|_W + \|\hat{\mathbf{z}}^{k+1} - \bar{\mathbf{z}}\|_W) \\ &\leq c_1\rho^2 + c_2\rho\|\hat{y}^{k+1} - \bar{y}\|_{Y_\infty}\end{aligned} \tag{5.36}$$

for constants $c_1, c_2 > 0$ depending on the radius ρ_1. Combining this with (5.33) leads to

$$\|\hat{y}^{k+1} - \bar{y}\|_{Y_\infty} \leq L_\infty c_1 \rho^2 + L_\infty c_2 \rho \|\hat{y}^{k+1} - \bar{y}\|_{Y_\infty}, \tag{5.37}$$

since $\hat{y}^k \in B_\rho^{Y_\infty}(\bar{y})$. Let us additionally demand $\rho \leq \frac{1}{L_\infty\, c_1 + L_\infty^2 c_2\, L}$. From (5.34) and (5.37) we can then conclude

$$\|\hat{y}^{k+1} - \bar{y}\|_{Y_\infty} \leq L_\infty\, c_1\rho^2 + L_\infty\, c_2\rho\, L_\infty\, L\, \rho \leq \rho^2[L_\infty c_1 + L_\infty^2 c_2\, L] \leq \rho, \tag{5.38}$$

which guarantees that $\hat{y}^{k+1} \in B_\rho^{Y_\infty}(\bar{y})$. Finally, let us demand $\rho \leq \frac{1}{2L_\infty c_2}$, and set $C_N = \frac{L_\infty c_1}{1 - \rho c_2 L_\infty} > 0$. Since (5.36) shows

$$\|\hat{y}^{k+1} - \bar{y}\|_{Y_\infty} \leq L_\infty c_1 \|\hat{y}^k - \bar{y}\|_{Y_\infty}^2 + L_\infty c_2 \rho \|\hat{y}^{k+1} - \bar{y}\|_{Y_\infty}.$$

Finally, we set $\rho := \min(\rho_1, \frac{1}{L_\infty\, c_1 + L_\infty^2 c_2\, L}, \frac{1}{2L_\infty\, c_2}) > 0$ to conclude the proof.

The convergence result in Y follows in the same way, noting that all auxiliary results, *cf.* Lemma A.1, Lemma A.2 and Lemma A.3, hold also in the Y-Z-setting, *cf.* the independence of the control update directions, see the introduction of this section. □

6. Local convergence of Algorithm 4.3 with L^∞-localization

In this section, we will show a first main result, *i.e.* that the sequence $\{q^k\}$ of iterates produced by the SQP method from Algorithm 4.3 converges locally quadratically to $\bar{y} = (\bar{\mathbf{u}}, \bar{q}, \bar{\mathbf{z}})$ under a closeness condition for the controls in L^∞. We have already proven that the auxiliary subproblem ($\widehat{\mathrm{QP}}_k$) produces feasible iterates, exploiting the strong regularity property, and that these iterates converge quadratically to $\bar{q}$ in the Y_∞-norm. To carry the results obtained for ($\widehat{\mathrm{QP}}_k$) over to (QP_k), we introduce yet another auxiliary problem in Section 6.1 still following the ideas of [16]. We then show equivalence results for the intermediate subproblems and ($\widehat{\mathrm{QP}}_k$) in Section 6.2, which relies on L^∞-techniques that are applicable due to Thereom 5.9.

Assumption 6.1. Let us note that in the following $y^k = (\mathbf{u}^k, q^k, \mathbf{z}^k) \in Y_\infty$ with $q^k \in Q_{\text{ad}}$ will always refer to a fixed function triple that lies in a neighborhood of $\bar{y}$ to be determined in Assumption 6.5 below.

6.1. The intermediate subproblem ($\text{QP}_k^{\omega,\infty}$)

As in [16] we introduce a localization of (QP_k), in the sense that the admissible control set Q_{ad} is restricted to a local L^∞-neighborhood of $\bar{q}$. We therefore define an admissible set with L^∞-closeness of $\bar{q}$ for an $\omega > 0$ by

$$Q_{\text{ad}}^{\omega,\infty} := \{q \in Q_{\text{ad}} \quad \big| \quad \|q - \bar{q}\|_{L^\infty(\Gamma)} \leq \omega\},$$

and the auxiliary subproblem

$$\min_q f_k(q), \quad \text{s.t. } q \in Q_{\text{ad}}^{\omega,\infty}. \tag{$\text{QP}_k^{\omega,\infty}$}$$

This problem will serve as an intermediate problem between (QP_k) and ($\widehat{\text{QP}}_k$). Let us start with existence of at least one solution.

Lemma 6.2. *Let $\omega > 0$ be sufficiently small, $y^k \in Y_\infty$ be given, and $\|\mathbf{u}^k - \bar{\mathbf{u}}\|_W, \|\mathbf{z}^k - \bar{\mathbf{z}}\|_W \leq \omega$. The auxiliary subproblem ($\text{QP}_k^{\omega,\infty}$) has at least one solution q_ω^{k+1} with $\mathbf{u}_\omega^{k+1} \in W$ satisfying (6.1a), see immediately below, and directions $d_\omega^{q,k} = q_\omega^{k+1} - q^k$ and $\mathbf{d}_\omega^{\mathbf{u},k} = \mathbf{u}_\omega^{k+1} - \mathbf{u}^k$.*

The proof follows standard methods and will not be carried out. First-order optimality conditions follow again in a standard way:

Corollary 6.3. *Let q_ω^{k+1} with state $\mathbf{u}_\omega^{k+1}$ be a solution to ($\text{QP}_k^{\omega,\infty}$) for given $y^k \in Y_\infty$. Then there exists an adjoint state pair $\mathbf{z}_\omega^{k+1} = (z_\omega^{u,k+1}, z_\omega^{\varphi,k+1}) \in W$, such that*

$$A'(\mathbf{u}^k)\mathbf{d}_\omega^{\mathbf{u},k+1} + R'_\varphi(\varphi^k;\gamma)d_\omega^{\varphi,k+1} = Bq_\omega^{k+1} - A(\mathbf{u}^k) - R(\varphi^k;\gamma), \tag{6.1a}$$

$$(A'(\mathbf{u}^k))^*\mathbf{z}_\omega^{k+1} + (R'(\varphi^k;\gamma))^* z_\omega^{\varphi,k+1} = u_\omega^{k+1} - u_d - A''(\mathbf{u}^k)[\mathbf{d}_\omega^{\mathbf{u},k+1},\cdot]^*\mathbf{z}^k - R''(\varphi^k;\gamma)[d_\omega^{\varphi,k+1},\cdot]^* z^{\varphi,k}, \tag{6.1b}$$

$$(B^*\mathbf{z}_\omega^{k+1} + \alpha q_\omega^{k+1}, q - q_\omega^{k+1}) \geq 0 \qquad \forall q \in Q_{ad}^{\omega,\infty}. \tag{6.1c}$$

Note that the optimality conditions of ($\widehat{\text{QP}}_k$) from Corollary 5.3 and the optimality conditions of ($\text{QP}_k^{\omega,\infty}$) from Corollary 6.3 only differ in the definition of the admissible sets in the variational inequality.

6.2. Equivalence of ($\text{QP}_k^{\omega,\infty}$) and ($\widehat{\text{QP}}_k$)

We will show that the (unique) solution $\hat{q}^{k+1}$ of ($\widehat{\text{QP}}_k$), together with $\hat{\mathbf{u}}^{k+1}$ and $\hat{\mathbf{z}}^{k+1}$ satisfying (5.11a) and (5.11b), satisfies the first-order necessary conditions (6.1a)–(6.1c) of ($\text{QP}_k^{\omega,\infty}$), and that these optimality conditions have a unique solution if y^k lies sufficiently close to $\bar{y}$ in the Y_∞-sense. In particular, we will also show that q_ω^{k+1} lies in $\widehat{Q_{\text{ad}}}$, if $\omega > 0$ is sufficiently small. We start with a technical auxiliary lemma, analogously to [16], Lemma 6.5.

Lemma 6.4. *There exists an $\omega_3 > 0$ with the following properties: Suppose $\omega \leq \omega_3$, $y^k \in Y$ with $\|\mathbf{u}^k - \bar{\mathbf{u}}\|_W, \|\mathbf{z}^k - \bar{\mathbf{z}}\|_W \leq \omega_3$, and let the triple $y = (\mathbf{u}, q, \mathbf{z})$ satisfy*

$$\begin{aligned} &q \in Q_{ad}^{\omega,\infty}, \\ &\mathbf{u} = \mathcal{G}_{\mathbf{u}^k}\big(B(q) + A'(\mathbf{u}^k)\mathbf{u}^k + (0,\, R'(\varphi^k;\gamma)\varphi^k) - A(\mathbf{u}^k) - (0, R(\varphi^k;\gamma))\big), \\ &\mathbf{z} = \mathcal{G}_{\mathbf{u}^k}\big((u - u_d,\, 0) - A''(\mathbf{u}^k)[\mathbf{u} - \mathbf{u}^k,\cdot]^*\mathbf{z}^k - (0,\, R''(\varphi^k;\gamma)[\varphi - \varphi^k,\cdot]^* z^{\varphi,k})\big). \end{aligned}$$

Then it holds

$$\begin{aligned} sign(B^*\mathbf{z}+\alpha q)(x) &= sign(B^*\bar{\mathbf{z}}+\alpha\bar{q})(x) \quad a.e.\ on \quad \mathcal{I}(\sigma),\\ \big|(B^*\mathbf{z}+\alpha q)(x)\big| &\geq \frac{\sigma}{2} \quad a.e.\ on \quad \mathcal{I}(\sigma). \end{aligned}$$

Before we start the proof, let us point out that the assumptions of Lemma 6.4 mean that y satisfies the state equation (6.1a) and adjoint equation (6.1b) of ($\mathrm{QP}_k^{\omega,\infty}$) from Corollary 6.3, but it is not yet clear that the variational inequality (6.1c) holds. This is shown in the following.

Proof. Analogously to [16], Lemma 6.5, the function $\mathbf{d^u} = (d^u, d^\varphi) := \mathbf{u}-\bar{\mathbf{u}}$ satisfies

$$\begin{aligned} A'(\mathbf{u}^k)\mathbf{d^u} + R'(\varphi^k;\gamma)d^\varphi &= A'(\mathbf{u}^k)(\mathbf{u}^k-\bar{\mathbf{u}}) + R'(\varphi^k)(\varphi^k-\bar{\varphi})\\ &\quad - (A(\mathbf{u}^k)-A(\bar{\mathbf{u}})) - (R(\varphi^k;\gamma)-R(\bar{\varphi};\gamma)) + B(q-\bar{q}). \end{aligned} \tag{6.2}$$

Taylor expansion of the auxiliary functional $T\colon [0,1]\to W^\times$, $T(\theta) := A(\mathbf{u}^k+\theta(\bar{\mathbf{u}}-\mathbf{u}^k))$ yields $T(1)-T(0) = T'(\theta)$, for $\theta\in(0,1)$, hence

$$A(\bar{\mathbf{u}}) - A(\mathbf{u}^k) = A'(\mathbf{u}^k+\theta(\bar{\mathbf{u}}-\mathbf{u}^k))(\bar{\mathbf{u}}-\mathbf{u}^k).$$

The operator R can be handled analogously. We obtain:

$$\begin{aligned} A'(\mathbf{u}^k)\mathbf{d^u} + R'(\varphi^k;\gamma)d^\varphi &= A'(\mathbf{u}^k+\theta(\bar{\mathbf{u}}-\mathbf{u}^k) - A'(\mathbf{u}^k))(\bar{\mathbf{u}}-\mathbf{u}^k)\\ &\quad R'(\varphi^k+\theta(\bar{\varphi}-\varphi^k) - R'(\varphi^k)\bar{\varphi})(\bar{\varphi}-\varphi^k) + B(q-\bar{q}). \end{aligned} \tag{6.3}$$

As in the proof of Lemma 5.7 or Lemma 3.1 we use the embeddings $W_u \hookrightarrow L^\infty(\Omega;\mathbb{R}^2)$ and $W_\varphi \hookrightarrow L^\infty(\Omega)$,

$$\|\mathbf{d^u}\|_\infty \leq c\|\mathbf{d^u}\|_W, \tag{6.4}$$

and estimate this further analogously to the proof of Lemma 5.1 or Lemma 5.7. Application of Lemma 3.1 and Corollary 3.3 yield

$$\|\mathbf{d^u}\|_W \leq (c+c_{\omega_3})\|\mathbf{u}^k-\bar{\mathbf{u}}\|_W + c\|q-\bar{q}\|_Q \leq c_{\omega_3}, \tag{6.5}$$

with a (generic) constant $c_{\omega_3} = \mathcal{O}(\omega_3)$, using $q\in Q_{\mathrm{ad}}^{\omega,\infty}$ in combination with $\omega\leq\omega_3$ for an ω_3 to be determined. Analogously, we obtain $\|\mathbf{z}-\bar{\mathbf{z}}\|_W \leq c_{\omega_3}$, from which we conclude,

$$|B^*(\mathbf{z}-\bar{\mathbf{z}}) + \alpha(q-\bar{q})| \leq c_{\omega_3}. \tag{6.6}$$

Therefore

$$B^*\mathbf{z}+\alpha q = B^*\bar{\mathbf{z}}+\alpha\bar{q} + B^*(\mathbf{z}-\bar{\mathbf{z}}) + \alpha(q-\bar{q}) \geq \sigma - c_{\omega_3} \text{ a.e. on } \mathcal{I}(\sigma),$$

where we used (6.6) and $|B^*\bar{\mathbf{z}}+\alpha\bar{q}|\geq\sigma$ due to Assumption 3.6. Since $c_{\omega_3} = \mathcal{O}(\omega_3)$ for $\omega_3\to 0$, choosing $\omega_3>0$ sufficiently small completes the proof. □

Let us now summarize all requirements for the different ω_i, $i=1,2,3$, that allow to apply all previously proven statements.

Assumption 6.5. In all that follows, we chose

$$\omega := \min(\omega_1, \omega_2, \omega_3).$$

Before showing uniqueness of the solution of ($\mathrm{QP}_k^{\omega,\infty}$), we need another auxiliary lemma, which shows $q_\omega^{k+1} \in \widehat{Q_{\mathrm{ad}}}$, *i.e.* feasibility of q_ω^{k+1} for ($\widehat{\mathrm{QP}_k}$).

Lemma 6.6. *Any control $q_\omega^{k+1} \in Q_{ad}^{\omega,\infty}$ of (QP_k^ω), that satisfies the optimality conditions from Corollary 6.3, together with the $\mathbf{u}_\omega^{k+1}$ and $\mathbf{z}_\omega^{k+1}$ satisfying (6.1a), and (6.1b) respectively, fulfills*

$$q_\omega^{k+1}(x) = \bar{q}(x) \quad a.e.\ on \quad \mathcal{I}(\sigma).$$

Proof. The proof works in the same way as [16], Corollary 6.6. For convenience we will recapitulate it: Let x be on $\mathcal{I}(\sigma)$. We have $\bar{q}(x) = q_b$ where $(B^*\bar{\mathbf{z}} + \alpha\bar{q})(x) \leq -\sigma$, and $\bar{q}(x) = q_a$ where $(B^*\bar{\mathbf{z}} + \alpha\bar{q})(x) \geq \sigma$. For any $q \in Q_{\mathrm{ad}}^{\omega,\infty}$, therefore either $q(x) \in [q_b - \omega, q_b]$ or $q(x) \in [q_a, q_a + \omega]$. By Lemma 6.4, either $B^*\mathbf{z}_\omega^{k+1} + \alpha q_\omega^{k+1} \geq \frac{\sigma}{2}$ or $B^*\mathbf{z}_\omega^{k+1} + \alpha q_\omega^{k+1} \leq -\frac{\sigma}{2}$, thus by (6.1c) it holds either $q_\omega^{k+1} = q_b$ or $q_\omega^{k+1} = q_a$. Thus on $\mathcal{I}(\sigma)$, we have shown $q_\omega^{k+1} = \bar{q}$. ☐

By means of Lemma 6.6, we will now show that solutions of the optimality system from Corollary 6.3 are unique under Assumption 3.6 and Assumption 6.5, using the ideas of [16], Theorem 6.12.

Lemma 6.7. *Let $y_k \in Y_\infty$ satisfy $\|\mathbf{u}_k - \bar{\mathbf{u}}\|_W, \|\mathbf{z}_k - \bar{\mathbf{z}}\|_W \leq \omega$. The optimality system from Corollary 6.3 for ($\mathrm{QP}_k^{\omega,\infty}$) admits a unique KKT triple $y_\omega^{k+1} \in Y_\infty$.*

Proof. According to Lemma 6.2, there is at least one solution $q_{\omega,1}^{k+1}$ of ($\mathrm{QP}_k^{\omega,\infty}$), hence at least one triple $y_{\omega,1}^{k+1}$ has to satisfy the optimality conditions of ($\mathrm{QP}_k^{\omega,\infty}$) from Corollary 6.3. We assume that $y_{\omega,2}^{k+1}$ also satisfies the optimality system from Corollary 6.3. Testing the variational inequality (6.1c) once in $q_{\omega,1}^{k+1}$ with $q_{\omega,2}^{k+1}$, and once in $q_{\omega,2}^{k+1}$ with $q_{\omega,1}^{k+1}$, taking the sum of the resulting inequalities, we obtain

$$0 \leq (B^*\mathbf{z}_{\omega,1}^{k+1} + \alpha q_{\omega,1}^{k+1}, q_{\omega,2}^{k+1} - q_{\omega,1}^{k+1}) + (B^* z_{\omega,2}^{k+1} + \alpha q_{\omega,2}^{k+1}, q_{\omega,1}^{k+1} - q_{\omega,2}^{k+1}).$$

Introducing the notation $\mathbf{u} := \mathbf{u}_{\omega,2}^{k+1} - \mathbf{u}_{\omega,1}^{k+1}, q := q_{\omega,2}^{k+1} - q_{\omega,1}^{k+1}$, $\mathbf{z} := \mathbf{z}_{\omega,2}^{k+1} - \mathbf{z}_{\omega,1}^{k+1}$, this leads to

$$\langle \mathbf{z}, Bq \rangle - \alpha(q, q)_Q \geq 0. \tag{6.7}$$

The functions $\mathbf{z}$ and $\mathbf{u}$ satisfy

$$\begin{aligned}(A'(\mathbf{u}^k))^*\mathbf{z} + (R'(\varphi^k;\gamma))^* z^\varphi &= -A''(\mathbf{u}^k)[\mathbf{u}, \cdot]^*\mathbf{z}^k - R''(\varphi^k;\gamma)[\varphi, \cdot]^* z^{\varphi,k} + u,\\ A'(\mathbf{u}^k)\mathbf{u} + R'(\varphi^k;\gamma)\varphi &= Bq.\end{aligned}$$

Testing the weak formulation of the first equation with $\mathbf{u}$ and the weak formulation of the second equation with $\mathbf{z}$ and again taking the sum of both equations, leads to

$$\langle Bq, \mathbf{z} \rangle = -\langle A''(\mathbf{u}^k)[\mathbf{u}, \mathbf{u}], \mathbf{z}^k \rangle - \langle R''(\varphi^k;\gamma)[\varphi, \varphi], z^{\varphi,k} \rangle + (u, u). \tag{6.8}$$

Combining (6.7) with (6.8), and using the definition of the Lagrangian function, we obtain

$$0 \geq (u, u) + \alpha(q, q)_Q - \langle A''(\mathbf{u}^k)[\mathbf{u}, \mathbf{u}], \mathbf{z}^k \rangle - \langle R''(\varphi^k;\gamma)[\varphi, \varphi], z^{\varphi,k} \rangle = \mathcal{L}''(y^k)[(\mathbf{u}, q), (\mathbf{u}, q)].$$

Due to Assumption 6.5, by Lemma 6.6 we have $q = 0$ on $\mathcal{I}(\sigma)$. Thus, by Lemma 5.1 it holds

$$\frac{\delta}{2}\|q\|_Q^2 \le \mathcal{L}''(y^k)[(\mathbf{u}, q), (\mathbf{u}, q)] \le 0.$$

However, this means $q = 0$, thus $q_{\omega,1}^{k+1} = q_{\omega,2}^{k+1}$. □

Note that Lemma 6.2 ensures the existence of at least one solution of ($\mathrm{QP}_k^{\omega,\infty}$) and Lemma 6.7 implies uniqueness of solutions of the optimality system associated to ($\mathrm{QP}_k^{\omega,\infty}$). We immediately conclude:

Corollary 6.8. *The subproblem* ($\mathrm{QP}_k^{\omega,\infty}$) *has a unique solution* $q_\omega^{k+1} \in Q_{ad}^{\omega,\infty}$.

The last crucial step in this section is to prove that $\hat{q}^{k+1}$ satisfies the optimality conditions of ($\mathrm{QP}_k^{\omega,\infty}$) and is thus the unique solution. We also see that the necessary conditions of (QP_k) are fulfilled. The next result works similarly to [16], Corollary 6.9. Note that we now also include a closeness assumption on q^k in addition to conditions on $\mathbf{u}^k, \mathbf{z}^k$.

Lemma 6.9. *Let* $\|y_k - \bar{y}\|_{Y_\infty} \le \omega$*, where* $\omega > 0$ *satisfies Assumption 6.5. The unique solution* $\hat{q}^{k+1}$ *of* ($\widehat{\mathrm{QP}}_k$)*, with associated* $\hat{\mathbf{u}}^{k+1}$ *satisfying* (5.11a)*, and* $\hat{\mathbf{z}}^{k+1}$ *satisfying* (5.11b)*, satisfies the optimality conditions of* ($\mathrm{QP}_k^{\omega,\infty}$) *from Corollary 6.3 and the optimality conditions of* (QP_k) *from Lemma 3.5.*

Proof. Observe first that $\hat{q}^{k+1} \in Q_{\mathrm{ad}}^{\omega,\infty} \subset Q_{\mathrm{ad}}$ under Assumption 6.5, as an immediate consequence of Theorem 5.9. To show the remaining properties, observe that the state equation of ($\widehat{\mathrm{QP}}_k$), ($\mathrm{QP}_k^{\omega,\infty}$), and ($\mathrm{QP}_k$) as well as the adjoint equation of ($\widehat{\mathrm{QP}}_k$), ($\mathrm{QP}_k^{\omega,\infty}$), and ($\mathrm{QP}_k$) are identical. It remains to prove that $\hat{q}^{k+1}$ and $\hat{\mathbf{z}}^{k+1}$ satisfy the variational inequality (6.1c) of ($\mathrm{QP}_k^{\omega,\infty}$), and (4.8c) of ($\mathrm{QP}_k$). We know that $\hat{q}^{k+1}$ and $\hat{\mathbf{z}}^{k+1}$ fulfill the variational inequality (5.11c) from Corollary 5.3, which reads

$$(B^*\hat{\mathbf{z}}^{k+1} + \alpha\hat{q}^{k+1}, q - \hat{q}^{k+1}) \ge 0 \qquad \forall \quad q \in \widehat{Q_{\mathrm{ad}}}.$$

Similarly to [16], Corollary 6.9, we recognize that on $\mathcal{I}(\sigma)$, there are two cases: If $B^*\hat{\mathbf{z}}^{k+1} + \alpha\hat{q}^{k+1} \ge \sigma$, we have $q_a = \bar{q} = \hat{q}^{k+1}$, recalling that since $\hat{q}^{k+1} \in \widehat{Q_{\mathrm{ad}}}$, it holds $\hat{q}^{k+1} = \bar{q}$. Likewise, if $B^*\hat{\mathbf{z}}^{k+1} + \alpha\hat{q}^{k+1} \le -\sigma$, then $q_b = \bar{q} = \hat{q}^{k+1}$. We already know that $\hat{q}^{k+1}$, with associated $\hat{\mathbf{u}}^{k+1}$ and $\hat{\mathbf{z}}^{k+1}$, is feasible for ($\mathrm{QP}_k^{\omega,\infty}$), therefore we can utilize Lemma 6.4 for the triple $\hat{y}^{k+1} = (\hat{\mathbf{u}}^{k+1}, \hat{q}^{k+1}, \hat{\mathbf{z}}^{k+1})$ and conclude that either $B^*\hat{\mathbf{z}}^{k+1} + \alpha\hat{q}^{k+1} \ge \frac{\sigma}{2}$ or $B^*\hat{\mathbf{z}}^{k+1} + \alpha\hat{q}^{k+1} \le -\frac{\sigma}{2}$. Therefore, $(B^*\hat{\mathbf{z}}^{k+1} + \alpha\hat{q}^{k+1})(q - \hat{q}^{k+1}) \ge 0$ holds on $\mathcal{I}(\sigma)$ for all $q \in [q_a, q_b]$. On $Q\backslash\mathcal{I}(\sigma)$, the controls $q \in \widehat{Q_{\mathrm{ad}}}$ fulfill the constraint $q \in [q_a, q_b]$. Overall, we obtain

$$\begin{aligned}(B^*\hat{\mathbf{z}}^{k+1} + \alpha\hat{q}^{k+1}, q - \hat{q}^{k+1})_Q = &(B^*\hat{\mathbf{z}}^{k+1} + \alpha\hat{q}^{k+1}, q - \hat{q}^{k+1})_{Q\backslash\mathcal{I}(\sigma)} \\ &+ (B^*\hat{\mathbf{z}}^{k+1} + \alpha\hat{q}^{k+1}, q - \hat{q}^{k+1})_{\mathcal{I}(\sigma)} \ge 0 \quad \forall\, q \in Q_{\mathrm{ad}}.\end{aligned} \tag{6.9}$$

Since $Q_{\mathrm{ad}}^{\omega,\infty} \subset Q_{\mathrm{ad}}$, the last inequality in particularly also holds for all $q \in Q_{\mathrm{ad}}^{\omega,\infty}$, which concludes the proof. □

We immediately conclude:

Corollary 6.10. *The unique solution* $\hat{q}^{k+1}$ *for* ($\widehat{\mathrm{QP}}_k$) *and the unique solution* q_ω^{k+1} *of* ($\mathrm{QP}_k^{\omega,\infty}$) *coincide.*

In particular, this means that $\hat{y}^{k+1} = y_\omega^{k+1}$ and that $\hat{q}^{k+1} = q_\omega^{k+1}$ is the unique (global) solution of both the subproblem ($\widehat{\mathrm{QP}}_k$) and the subproblem ($\mathrm{QP}_k^{\omega,\infty}$) if $\|y_k - \bar{y}\|_{Y_\infty} \le \omega$ for an $\omega > 0$ sufficiently small.

6.3. Local convergence of Algorithm 4.3 with L^∞-localization

As a consequence of Corollary 6.10 and Theorem 5.9 we obtain local quadratic convergence of Algorithm 4.3 if (QP_k) is replaced by ($\mathrm{QP}_k^{\omega,\infty}$), *i.e.* if we enforce that controls stay in an L^∞-neighborhood of $\bar{q}$.

Theorem 6.11. *Let $\bar{y}$ satisfy Assumption 3.6 and let Assumption 6.5 hold, i.e. let $\omega > 0$ be sufficiently small. In particular, we may take $q^0 \in Q_{ad}$ satisfying $\|q^0 - \bar{q}\|_Q \leq \epsilon_\infty$, and set $\mathbf{u}^0 = G(q^0), \mathbf{z}^0 = G'(q^0)^*((u^0 - u_d, 0))$ for a constant $\epsilon_\infty > 0$ sufficiently small. Then the sequence $\{y_\omega^k\}_{k\in\mathbb{N}}$ generated by Algorithm 4.3 with quadratic subproblem ($\mathrm{QP}_k^{\omega,\infty}$) converges quadratically in Y_∞ to $\bar{y}$.*

Proof. Note that for $\epsilon_\infty > 0$ small the associated state $\mathbf{u}^0$ and adjoint state $\mathbf{z}^0$ fulfill $\|\mathbf{u}^0 - \bar{\mathbf{u}}\|_W, \|\mathbf{z}^0 - \bar{\mathbf{z}}\|_W \leq \omega$ with ω as in Assumption 6.5 due to (3.4) and (3.5), since the latter estimates holds for adjoint equations. By Theorem 5.9, the sequence of iterates $\{\hat{y}^k\}_{k\in\mathbb{N}}$ generated by ($\widehat{\mathrm{QP}}_k$) fulfills $\|\hat{y}^k - \bar{y}\|_{Y_\infty} \leq \tilde{\omega}$ for all $k \geq 1$, and $\hat{y}^k$ converges to $\bar{y}$ in Y_∞. Due to Corollary 6.10, the controls $\hat{q}^{k+1}$ are the unique solution of the problem ($\mathrm{QP}_k^{\omega,\infty}$). This concludes the proof. □

Note that the size of the initial local neighborhood for the convergence result will depend on the physical problem parameters.

7. Convergence of Algorithm 4.3 with L^2-localization

In this final section we will prove a convergence result for the SQP method of Algorithm 4.3 with L^2-closeness conditions in the quadratic subproblems, *i.e.* we define $Q_{\mathrm{ad}}^{\omega,2}$ around $\bar{q}$ for an $\omega > 0$ by

$$Q_{\mathrm{ad}}^{\omega,2} := \{q \in Q_{\mathrm{ad}} \quad | \quad \|q - \bar{q}\|_Q \leq \omega\},$$

and consider subproblems

$$\min_q f_k(q), \quad \text{s.t. } q \in Q_{\mathrm{ad}}^{\omega,2}, \tag{$\mathrm{QP}_k^{\omega,2}$}$$

with f_k as in ($\mathrm{QP}_k^{\mathrm{red}}$). We follow the idea in [27], where an analogous result has been obtained for a tracking type control problem governed by a quasilinear parabolic equation. Before doing so, let us mention some considerations from [33] in the context of second order sufficient conditions. First, it is clear that a local solution in the sense of L^2 is also a local solution in the sense of L^∞, yet it has been stated in [33] that under certain conditions many results that are only expected for L^∞-neighborhoods even hold in L^2-neighborhoods; in particular, quadratic growth conditions and strict local optimality in the sense of L^2 were shown, under certain conditions, even in the context of the so called two-norm discrepancy, were the coercivity property of the second derivative of the objective functional and differentiability hold in different spaces. The parabolic setting of [27] is such a setting, since the control-to-state operator is in general not differentiable with respect to L^2. Nevertheless, a quadratic growth condition with respect to L^2 was already obtained in [60], and eventually convergence of the SQP-method with respect to L^2-closeness in the quadratic subproblems in [27]. For our model problem ($\mathrm{NLP}^{\gamma,\eta}$) we do not have to deal with a two-norm discrepancy since G and f are differentiable with respect to L^2. Applying the ideas of [27] will lead to an analogous result, even though the arguments of the last section cannot be transferred directly to L^2-neighborhoods, as the proof of Lemma 6.6 unfortunately cannot be transferred directly.

Nevertheless, following [27], the main idea is, in essence, that the unique solution of ($\widehat{\mathrm{QP}}_k$), which we already know to be a solution of ($\mathrm{QP}_k^{\omega,\infty}$) with L^∞-localization, is in fact also a local L^2-solution of (QP_k), satisfying a quadratic growth condition. The challenges are best illustrated by starting the discussion with the following result:

Proposition 7.1. *Let $\bar{y}$ satisfy Assumption 3.6. There exist constants $\omega_4 > 0$, satisfying Assumption 6.5, and $c_{k+1} > 0, \epsilon_{k+1} > 0$ such that for all $y^k \in Y_\infty$ with $\|y^k - \bar{y}\|_{Y_\infty} \leq \omega_4$ the unique solution $\hat{q}^{k+1}$ of ($\widehat{\mathrm{QP}}_k$) satisfies the quadratic growth condition*

$$f_k(q) \geq f_k(\hat{q}^{k+1}) + c_{k+1}\|q - \hat{q}^{k+1}\|_Q^2, \tag{7.1}$$

for every $q \in Q_{ad}$ *with* $\|q - \hat{q}^{k+1}\|_Q \leq \epsilon_{k+1}$*. In particular,* $\hat{q}^{k+1}$ *is then also a strict* L^2*-local solution of* (QP_k)*. Moreover,* $\hat{y}^{k+1}$ *is the only stationary point of* (QP_k) *in* $B^Q_{\epsilon_{k+1}}(\hat{q}^{k+1})$*.*

Before we start the proof, let us point out that this is an analogue of [27], Proposition 6.13, with the important difference that up to now it is not clear whether or not the size of the radius ϵ_{k+1} in Proposition 7.1 can be chosen independently of the solution $\hat{q}^{k+1}$ itself. We will explain how this issue was tackled in [27] later on in this section. The proof of [27], Proposition 6.13 was based on [33], Theorem 2.3 and Corollary 2.6, [61], Theorem 3.22, using similar arguments as in [62]. Here, we verify the Assumptions of [62], Theorem 2.5 directly.

Proof. As already pointed out, we verify the Assumptions of [62], Theorem 2.5, setting $U_2 = U_\infty = Q$ and $\mathcal{K} = Q_{\mathrm{ad}}$ in the latter. Note first that Assumption $(A.1)$ is fulfilled since f_k is twice continuously differentiable from Q into $\mathbb{R}$. Moreover, in Lemma 6.9, we have already verified that $\hat{q}^{k+1}$ satisfies the first order necessary conditions for (QP_k) and $(\mathrm{QP}_k^{\omega,\infty})$. By analogous arguments, *i.e.* $Q_{\mathrm{ad}}^{\omega,2} \subset Q_{\mathrm{ad}}$, it satisfies the first order necessary conditions for $(\mathrm{QP}_k^{\omega,2})$. Moreover, note that $\mathcal{C}(\bar{q}) \subset \mathcal{C}_\sigma(\bar{q})$, *cf.* (3.38). We also recall the representations (4.3) for f_k' and (4.5) for f_k'' and the boundedness (4.6) and the Lipschitz result (4.7). Therefore, we only need to prove the Legendre form properties of f_k'', which is rather clear due to the quadratic structure with Tikhonov term. Nevertheless, let $\{q_n\}_{n\in\mathbb{N}}$ denote a sequence converging strongly to $\hat{q}^{k+1}$ in Q and $\{v_n\}_{n\in\mathbb{N}}$ a sequence converging weakly in Q to some $v \in Q$. With the identity (4.5) we obtain

$$\begin{aligned} f_k''(\hat{q}^{k+1})[v_n, v_n] =& \mathcal{L}''(y^k)[(\tilde{\mathbf{u}}_n, v_n), (\tilde{\mathbf{u}}_n, v_n)] \\ =& \mathcal{L}''(y^k)[(\tilde{\mathbf{u}}, v), (\tilde{\mathbf{u}}, v)] + \mathcal{L}''(y^k)[(\tilde{\mathbf{u}}_n, v_n), (\tilde{\mathbf{u}}_n, v_n)] - \mathcal{L}''(y^k)[(\tilde{\mathbf{u}}, v), (\tilde{\mathbf{u}}, v)] \end{aligned} \tag{7.2}$$

with $\tilde{\mathbf{u}} := \mathcal{G}_{\mathbf{u}^k}(Bv)$ and $\tilde{\mathbf{u}}_n = \mathcal{G}_{\mathbf{u}^k}(Bv_n)$. Estimates (3.48) and (3.49) from Corollary 3.9 then yield

$$f_k''(\hat{q}^{k+1})[v_n, v_n] \geq f_k''(\hat{q}^{k+1})(v, v) - (c + c_{\omega_4})(\|\tilde{\mathbf{u}}_n\|_V + \|\tilde{\mathbf{u}}\|_V)\|\tilde{\mathbf{u}}_n - \tilde{\mathbf{u}}\|_V \to \mathcal{L}''(y^k)[(\tilde{\mathbf{u}}, v), (\tilde{\mathbf{u}}, v)],$$

as $n \to \infty$ since $\tilde{\mathbf{u}}_n \to \tilde{\mathbf{u}}$ strongly in V. Last, assume that

$$f_k''(\hat{q}^{k+1})[v_n, v_n] \to f_k''(\hat{q}^{k+1})[v, v] = \mathcal{L}''(y^k)[(\tilde{\mathbf{u}}, v), (\tilde{\mathbf{u}}, v)]$$

as $n \to \infty$. Then (7.2) implies $\mathcal{L}''(y^k)[(\tilde{\mathbf{u}}_n, v_n), (\tilde{\mathbf{u}}_n, v_n)] - \mathcal{L}''(y^k)[(\tilde{\mathbf{u}}, v), (\tilde{\mathbf{u}}, v)] \to 0$ as $n \to \infty$, and inspection of (3.46) in the proof of Proposition 3.8, combined with (3.14) and (3.15) implies $\|v_n\|_Q^2 - \|v\|_Q^2 \to 0$, since $\tilde{\mathbf{u}}_n \to \tilde{\mathbf{u}}$ strongly in V. Since weak convergence together with norm convergence implies strong convergence, this implies $v_n \to v$ in L^2 in this case. Overall, [62], Theorem 2.5 is applicable and concludes the proof of the quadratic growth condition. Uniqueness of stationary points can be shown as in [33], Corollary 2.6. □

Let us now discuss why the radius ϵ_{k+1} can be chosen independently of $\hat{q}^{k+1}$. We still follow [27] and prove that in fact the unique solution $\hat{q}_\delta$ of $(\widehat{\mathrm{QP}}_\delta)$ is a unique global solution of

$$\min_q f_\delta(q), \quad \text{s.t. } q \in Q_{\mathrm{ad}}^{\omega,2}, \tag{$\mathrm{QP}_\delta^{\omega,2}$}$$

and also a local solution of

$$\min_q f_\delta(q), \quad \text{s.t. } q \in Q_{\mathrm{ad}}, \tag{QP_δ}$$

as long as the perturbations fulfill a smallness condition in Z_∞. Here, we tacitly use the notation f_δ for the reduced objective function for $(\widehat{\mathrm{QP}}_\delta)$. Note that these problems differ in the definition of the admissible set.

This is the key idea from [27], Proposition 6.7. It is important to notice that the linearization point in (QP_δ) and $(\mathrm{QP}_\delta^{\omega,2})$ is the point $\bar{y}$ as opposed to a current iterate y^k in $(\mathrm{QP}_k^{\omega,2})$. After proving these solution properties, it will be clear that the strong regularity property of ($\widehat{\text{GE}}$) transfers to the generalized equation (GE) and its localized analogue

$$0 \in F(\bar{y}) + N_2(\bar{y}), \tag{GE$^{\omega,2}$}$$

with $N_2(y) := \big(0,\, 0,\, N_{\text{nc}}^{\omega,2}(q)\big)^T$, where $N_{\text{nc}}^{\omega,2}(q) = \{d^q \in Q \mid (d^q, \tilde{q} - q)_Q \le 0 \text{ for all } \tilde{q} \in Q_{\text{ad}}^{\omega,2}\}$. This guarantees local uniqueness of stationary points of (QP_k) and ($\text{QP}_k^{\omega,2}$)) within fixed neighborhoods of $\bar{y}$. The smallness condition on the perturbations and the local neighborhood where $\hat{q}_\delta$ is a strict solution therefore only depend on $\bar{y}$ as has been pointed out in the context of [27], Proposition 6.7. It is clear from standard arguments that ($\text{QP}_\delta^{\omega,2}$) admits at least one solution for every perturbation vector $\delta \in Z_\infty$. Likewise, first order optimality conditions can be derived in a straight forward manner, and could be identified with solving a generalized equation ($\text{NM}_\delta^{\omega,2}$) analogously to (NM). We omit the details. The following result is the analogue of [27], Proposition 6.7 and Corollary 6.8, where the key idea of this approach was proven.

Proposition 7.2. *Let $\bar{y}$ satisfy Assumption 3.6 and let σ, L_∞ be the constants from Assumption 3.6 and Lemma 5.7. There exists constants constants $\tilde{r} \in [0, \frac{\sigma}{2L_\infty}]$ and $c > 0, \tilde{\epsilon} > 0$ s.t. the solution $\hat{q}_\delta$ of ($\widehat{\text{QP}}_\delta$) satisfies the quadratic growth condition*

$$f_\delta(q) \ge f_\delta(\hat{q}_\delta) + c\|q - \hat{q}_\delta\|_Q^2,$$

for every $q \in Q_{ad}$ with $\|q - \hat{q}_\delta\|_Q \le \tilde{\epsilon}$ for all $\|\delta\|_{Z_\infty} \le \tilde{r}$. In particular, $\hat{q}_\delta$ is a strict L^2-local solution of (QP_δ). Moreover $\hat{y}_\delta = (\hat{\mathbf{u}}_\delta, \hat{q}_\delta, \hat{\mathbf{z}}_\delta)$ is the only stationary point for (QP_δ) in $\overline{B_{\tilde{\epsilon}}^Q(\hat{q}_\delta)}$.

Proof. The proof is the same as for Proposition 7.1, *i.e.* again we check the prerequisites of [62], Theorem 2.5. Note that compared to f_k, the linearization point in f_δ is $\bar{y}$ instead of y^k and all perturbations enter the formulation linearly, so f_δ will fulfill the required properties that we have proven in detail for f_k. We only need to discuss that $\hat{q}_\delta$ satisfies the first order necessary conditions of ($\text{QP}_\delta^{\omega,2}$). This follows similarly to Lemma 6.9, see in particular [24], but only for perturbations δ that are suffficiently small with respect to the Z_∞ norm. The reason is that the strongly active set $\mathcal{I}_\delta(\sigma) := \{x \in \Gamma \mid |B^*\bar{\mathbf{z}} + \alpha\bar{q} - \delta_2| > 0\}$ behaves sufficiently well for small perturbations $\delta \in Z_\infty$, according to Lemma 6.4, *cf.* [27], Lemma 6.6 and [24], Corollary 5.3. In particular, we cannot take δ_2 small with respect to L^2, only. Again, noting (3.38) concludes the proof. ☐

The same arguments as in [27], Corollary 6.8 let us conclude that $\hat{q}_\delta$ is in fact the unique global solution of ($\text{QP}_\delta^{\omega,2}$), but also a local solution of Problem (QP_δ).

Proposition 7.3. *Let $\bar{y}$ satisfy Assumption 3.6. Then, there exists a constant $\omega_5 > 0$ and a radius $r > 0$ such that the unique triple $(\hat{\mathbf{u}}_\delta, \hat{q}_\delta, \hat{\mathbf{z}}_\delta)$ solving ($\widehat{\text{QP}}_\delta$) is also the unique solution of ($\text{QP}_\delta^{\omega,2}$) as well as the unique solution of (QP_δ) that is contained in the set $Q_{ad}^{\omega,2}$ for $\omega = \omega_5$, for all $\|\delta\|_{Z_\infty} \le r$.*

Proof. The proof is exactly the same as in [27], Corollary 6.8. We state the main idea for convenience of the reader. Set $\omega_5 = \frac{2\tilde{\epsilon}}{3}$ and $r < \min(\tilde{r}, \frac{\tilde{\epsilon}}{3c_e L_\infty})$ with $\tilde{\epsilon}, \tilde{r}$ chosen as in Proposition 7.2, c_e the constant of the embedding $L^\infty \hookrightarrow L^2$, and L_∞ the Lipschitz-constant from Lemma 5.7. Then $\|\hat{q}_\delta - \bar{q}\|_Q < c_e L_\infty \|\delta\|_{Z_\infty} < \tilde{\epsilon}/3$ due to Lemma 5.7, and therefore

$$\|q - \hat{q}_\delta\|_Q \le \|q - \bar{q}\|_Q + \|\bar{q} - \hat{q}_\delta\|_Q < \omega_5 + \frac{\tilde{\epsilon}}{3} \le \frac{2\tilde{\epsilon}}{3} + \frac{\tilde{\epsilon}}{3} = \tilde{\epsilon} \quad \forall q \in B_{\omega_5}^Q(\bar{q}). \tag{7.3}$$

Together, this verifies $\hat{q}_\delta \in \Big(Q_{\text{ad}} \cap \overline{B_{\omega_5}^Q(\bar{q})}\Big) \subset \Big(Q_{\text{ad}} \cap \overline{B_{\tilde{\epsilon}}^Q(\hat{q}_\delta)}\Big)$. Observing that $\hat{q}_\delta$ satisfies a quadratic growth condition in the larger set and is contained in the smaller set, and that Proposition 7.3 also guarantees that there is no further stationary point inside the larger set concludes the proof. Note that [27], Proposition 6.7

and Corollary 6.8 distinguish between different neighborhoods with respect to stationary points and solutions. Without loss of generality, we choose the same radius for both. □

As mentioned earlier, it is now a straight forward conclusion from Theorem 5.8 and Proposition 7.3 that the generalized equations (GE) and ($\mathrm{GE}^{\omega,2}$) are strongly regular in Y_∞.

Theorem 7.4. *The generalized equation (GE) is strongly regular at $\bar{y}$ in Y_∞. Moreover, there is $\omega > 0$ such that ($\mathrm{GE}^{\omega,2}$) is strongly regular at $\bar{y}$ in Y_∞.*

Note that, unlike ($\widehat{\mathrm{QP}}_k$), the quadratic subproblem (QP_k) may still admit multiple solutions. Solving the latter is therefore not equivalent to solving the generalized equation (NM). For ($\mathrm{QP}_k^{\omega,2}$) and a localized version of (NM), however, this is true.

We can finally establish our second main result with the help of the strong regularity property for ($\mathrm{GE}^{\omega,2}$), analogously to [27], Theorem 6.15.

Theorem 7.5. *Let $\bar{y}$ satisfy Assumption 3.6 and let $\|\mathbf{u}_0 - \bar{\mathbf{u}}\|_W, \|\mathbf{z}_0 - \bar{\mathbf{z}}\|_W \le \omega$ with ω sufficiently small. In particular, we may take $q^0 \in Q_{ad}$ with $\|q^0 - \bar{q}\|_Q$ sufficiently small, and set $\mathbf{u}^0 = G(q^0), \mathbf{z}^0 = G'(q^0)^*(u^0 - u_d)$. Then the sequence $\{y^k\}_{k\in\mathbb{N}}$ generated by Algorithm 4.3 with quadratic subproblem ($\mathrm{QP}_k^{\omega,2}$) converges quadratically in Y_∞ to $\bar{y}$.*

Proof. First, we point out that Proposition 7.1 and Theorem 7.4 guarantee that there exists a constant $\omega_6 > 0$ that satisfies Assumption 6.5 and a radius $\epsilon > 0$ independent of $\hat{q}^{k+1}$ such that for all $y^k \in Y_\infty$ with $\|y^k - \bar{y}\|_{Y_\infty} \le \omega_6$ the unique solution $\hat{q}^{k+1}$ of ($\widehat{\mathrm{QP}}_k$) is also the unique solution of (QP_k) with admissible set $\underline{Q_{\mathrm{ad}}^{\epsilon,2}}$. Moreover, $\hat{q}^{k+1}$ is also the unique L^2-local solution of the global problem (QP_k) that is contained in $Q_{\mathrm{ad}} \cap \overline{B_\epsilon^Q}(\bar{q})$, *cf.* also [27], Proposition 6.7. Then the proof is a simple consequence of Theorem 5.9. □

Compared to Theorem 6.11 we can now replace the L^∞ localization in the subproblems by an L^2-localization as in [27]. The price we pay for considering weaker second-order sufficient conditions compared to *e.g.* [20] is, that we cannot avoid localization completely. In [16], there are some hints how the local neighborhoods might be chosen in the L^∞ localized setting.

Appendix A. Auxiliary results

A.1 Auxiliary estimates for the differential operators

For convenience of the reader, we demonstrate the main techniques that are used in the proof of (3.6)–(3.17). These results haven been showed in the context of [1–3, 10], and go back to ideas from [2]. This section is concerned with the Lipschitz estimates (3.14) and (3.10) for the differential operators A' and A''. Note that we do not discuss the (easier) estimations for R' and R''. We recall for convenience:

$$\begin{aligned}\langle A'(\mathbf{u})\tilde{\mathbf{u}}, \mathbf{v}\rangle :=& (g(\varphi)\mathbb{C}e(\tilde{u}), e(v^u)) + 2(1-\kappa)(\varphi\mathbb{C}e(u) : e(\tilde{u}), v^\varphi) + 2(1-\kappa)(\varphi\mathbb{C}e(u)\tilde{\varphi}, e(v^u)) \\ &+ \varepsilon(\nabla\tilde{\varphi}, \nabla v^\varphi) + \frac{1}{\varepsilon}(\tilde{\varphi}, v^\varphi) + \eta(\tilde{\varphi}, v^\varphi) + (1-\kappa)(\tilde{\varphi}\mathbb{C}e(u) : e(u), v^\varphi), \qquad (A.1)\\ \langle A''(\mathbf{u})[\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2], \mathbf{z}\rangle :=& 2(1-\kappa)[(\tilde{\varphi}_2\mathbb{C}e(u)\tilde{\varphi}_1, e(z^u)) + (\tilde{\varphi}_2\mathbb{C}e(\tilde{u}_1)\varphi, e(z^u)) + (\tilde{\varphi}_2\mathbb{C}e(u) : e(\tilde{u}_1), z^\varphi) \\ &+ (\varphi\mathbb{C}e(\tilde{u}_2)\tilde{\varphi}_1, e(z^u)) + (\tilde{\varphi}_1\mathbb{C}e(\tilde{u}_2) : e(u), v^\varphi) + (\varphi\mathbb{C}e(\tilde{u}_2) : e(\tilde{u}_1), z^\varphi)]. \qquad (A.2)\end{aligned}$$

A key argument in the calculations are the embeddings $V_u, V_\varphi \hookrightarrow L^s(\Omega)$ for all $s < \infty$, in our setting with spatial dimension $N = 2$. We prove the boundedness estimate (3.14) in detail, which we recall as

$$|\langle A''(\mathbf{u})[\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2], \mathbf{z}\rangle| \le c\|\mathbf{u}\|_W \|\mathbf{z}\|_W \|\tilde{\mathbf{u}}_1\|_V \|\tilde{\mathbf{u}}_2\|_V.$$

We prove that the terms in (A.2) can be bounded accordingly. For the first term, note that with the above regularities $e(u) \in L^p$ with $p > 2$, $e(z^u) \in L^2$, and $\tilde{\varphi}_2, \tilde{\varphi}_2 \in L^s$ for all $s < \infty$. In particular, $s > 1$ can be chosen such $2/s + 1/p + 1/2 = 1$, since $p > 2$. Then, by Hölders inequality

$$|(\tilde{\varphi}_2 \mathbb{C} e(u) \tilde{\varphi}_1, e(z^u))| \leq c \|e(u)\|_p \|e(z^u)\|_2 \|\tilde{u}_1\|_s \|\tilde{u}_2\|_s$$

holds. Using $V_u \hookrightarrow L^s(\Omega)$ completes the proof for the first term. In *e.g.* the last term of (A.2), we make use of higher regularity of $\mathbf{z}$: With $e(\tilde{u}_1), e(\tilde{u}_2) \in L^2(\Omega)$, both φ and z^φ need to be bounded in the L^∞-norm in Hölder's inequality

$$|(\varphi \mathbb{C} e(\tilde{u}_2) : e(\tilde{u}_1), z^\varphi)| \leq c \|\varphi\|_\infty \|e(\tilde{u}_1)\|_2 \|e(\tilde{u}_1)\|_2 \|z^\varphi\|_\infty,$$

which can be done using W-regularity of both $\mathbf{u}$ and $\mathbf{z}$ by the embedding $W_\varphi \hookrightarrow L^\infty$. Similarly, for the second term of (A.2), we use $\mathbf{u}, \mathbf{z} \in W$, hence $e(u), e(z^u) \in L^p(\Omega)$ to prove

$$|(\tilde{\varphi}_2 \mathbb{C} e(\tilde{u}_1) \varphi, e(z^u))| \leq c \|\tilde{\varphi}_2|_s \|e(\tilde{u}_1)\|_2 \|e(z^u)\|_p \|z^\varphi\|_\infty.$$

The remaining terms can be bounded by similar calculations, and the Lipschitz results with respect to the corresponding norms are a straight forward application of these results, since only once component of the linearization point $\mathbf{u}$ enters these terms linearly.

Instead, we demonstrate the calculations needed for the Lipschitz result (3.10) of A' with the help of the first term in (A.1). Note that for A', we only use $\mathbf{v} \in V$, hence $e(v^u) \in L^2(\Omega)$ requires to bound the product of the remaining terms in $L^2(\Omega)$ We calculate

$$g(\varphi_1) \mathbb{C} e(\tilde{u}) - g(\varphi_2) \mathbb{C} e(\tilde{u}) = (g(\varphi_1) - g(\varphi_2)) \mathbb{C} e(\tilde{u}).$$

Then $\varphi_1, \varphi_2 \in L^\infty(\Omega)$, $e(\tilde{u}) \in L^2(\Omega)$ and the Lipschitz continuity of g complete the calculations. The remaining terms do not pose additional difficulties. Note that explicit calculations also give a precise dependence on the physical model parameters κ, ϵ, γ, that we consider fixed.

A.2 Auxiliary results for Theorem 5.9

We establish some auxiliary results that are necessary for the proof of Theorem 5.9. We start with an auxiliary result that is required for the application of Dontchev's implicit function theorem [32], Theorem 2.4, and follows along the lines of [20], Lemma 6.2.

Lemma A.1. *Let $\bar{y} \in Y_\infty$ be given. For any $\tilde{\omega}_1, \tilde{\omega}_2 > 0$, there exists a constant $L(\tilde{\omega}_1, \tilde{\omega}_2) > 0$ such that for all $y_i = (\mathbf{u}_i, q_i, \mathbf{z}_i) \in Y$ with $\|\mathbf{u}_i - \bar{\mathbf{u}}\|_W, \|\mathbf{z}_i - \bar{\mathbf{z}}\|_W \leq \tilde{\omega}_1$, $i = 1, 2$, and for all $y = (\mathbf{u}, q, \mathbf{z}) \in Y_\infty$ with $\|\mathbf{u} - \bar{\mathbf{u}}\|_W, \|\mathbf{z} - \bar{\mathbf{z}}\|_W \leq \tilde{\omega}_2$, the following Lipschitz condition holds:*

$$\|F(y_1) + F'(y_1)(y - y_1) - F(y_2) - F'(y_2)(y - y_2)\|_{Z_\infty} \leq L(\tilde{\omega}_1, \tilde{\omega}_2) \left(\|\mathbf{u}_1 - \mathbf{u}_2\|_W + \|\mathbf{z}_1 - \mathbf{z}_2\|_W \right).$$

Proof. Let $\tilde{\omega}_1, \tilde{\omega}_2 > 0$, y_i and y be as assumed. A short calculation shows that

$$F(y_1) + F'(y_1)(y - y_1) - F(y_2) + F'(y_2)(y - y_2) = (f_1(\mathbf{u}_1) - f_1(\mathbf{u}_2),\ f_1(\mathbf{u}_1, \mathbf{z}_1) - f_1(\mathbf{u}_2, \mathbf{z}_2),\ 0)^T,$$

with

$$\begin{aligned} f_1(\mathbf{u}_i) &:= A(\mathbf{u}_i) + R(\varphi_i; \gamma) + A'(\mathbf{u}_i)(\mathbf{u} - \mathbf{u}_i) + R'(\varphi_i; \gamma)(\varphi - \varphi_i) \\ f_2(\mathbf{u}_i, \mathbf{z}_i) &:= (A'(\mathbf{u}_i))^* \mathbf{z} + R'(\varphi_i; \gamma)^* z^\varphi + A''(\mathbf{u}_i)[\mathbf{u} - \mathbf{u}_i, \cdot]^* \mathbf{z}_i + R''(\varphi_i; \gamma)[\varphi - \varphi_i, \cdot]^* z_i^\varphi. \end{aligned}$$

We observe that the difference in the last component is always zero, and that neither f_1 nor f_2 depends on q or q_i, $i = 1, 2$. We calculate

$$\begin{aligned} &f_2(\mathbf{u}_1, \mathbf{z}_1) - f_2(\mathbf{u}_2, \mathbf{z}_2) \\ &\quad = (A'(\mathbf{u}_1) - A'(\mathbf{u}_2)^*\mathbf{z} + (A''(\mathbf{u}_1) - A''(\mathbf{u}_2))[\mathbf{u} - \mathbf{u}_1, \cdot]^*\mathbf{z}_1 + A''(\mathbf{u}_2)[\mathbf{u} - \mathbf{u}_1, \cdot]^*(\mathbf{z}_1 - \mathbf{z}_2) \\ &\qquad + A''(\mathbf{u}_2)[\mathbf{u}_2 - \mathbf{u}_1, \cdot]^*\mathbf{z}_2 + (R'(\varphi_1; \gamma) - R'(\varphi_2; \gamma))^* z^\varphi + (R''(\varphi_1; \gamma) - R''(\varphi_2; \gamma))[\varphi - \varphi_1, \cdot]^* z_1^\varphi \\ &\qquad + R''(\varphi_2; \gamma)[\varphi - \varphi_1, \cdot\,]^*(z_1^\varphi - z_2^\varphi) + R''(\varphi_2; \gamma)[\varphi_2 - \varphi_1, \cdot\,]^* z_2^\varphi. \end{aligned}$$

We point out again that there is no explicit appearance of q, q_1, or q_2 in the above estimate. Applying the boundedness and Lipschitz results (3.27)–(3.30), we obtain

$$\|f_2(\mathbf{u}_1, \mathbf{z}_1) - f_2(\mathbf{u}_2, \mathbf{z}_2)\|_{W^\times} \le L(\tilde{\omega}_1, \tilde{\omega}_2)(\|\mathbf{u}_1 - \mathbf{u}_2\|_W + \|\mathbf{z}_1 - \mathbf{z}_2\|_W),$$

for an $L(\tilde{\omega}_1, \tilde{\omega}_2) > 0$. Here, we also used $\|\mathbf{u} - \mathbf{u}_i\|_W \le \|\mathbf{u} - \bar{\mathbf{u}}\|_W + \|\bar{\mathbf{u}} - \mathbf{u}_i\|_W \le \tilde{\omega}_1 + \tilde{\omega}_2$, for $i = 1, 2$. Estimating the difference for f_1 in a similar, comparably less technical way, concludes the proof. We omit further details. ☐

Lemma A.2. *The operator F' from (5.2) is locally Lipschitz continuous w.r.t to y as a mapping from Y_∞ into Z_∞ : For any $\tilde{\omega}_1, \tilde{\omega}_2 > 0$, there exists a constant $L(\tilde{\omega}_1, \tilde{\omega}_2) > 0$ such that for all $y_i = (\mathbf{u}_i, q_i, \mathbf{z}_i) \in Y$ with $\|\mathbf{u}_i - \bar{\mathbf{u}}\|_W, \|\mathbf{z}_i - \bar{\mathbf{z}}\|_W \le \tilde{\omega}_1$, $i = 1, 2$, and for all $y = (\mathbf{u}, q, \mathbf{z}) \in Y_\infty$ with $\|\mathbf{u} - \bar{\mathbf{u}}\|_W, \|\mathbf{z} - \bar{\mathbf{z}}\|_W \le \tilde{\omega}_2$, the following estimate holds*

$$\|(F'(y_1) - F'(y_2))y\|_{Z_\infty} \le L(\tilde{\omega}_1, \tilde{\omega}_2)\,(\|\mathbf{u}_1 - \mathbf{u}_2\|_W + \|\mathbf{z}_1 - \mathbf{z}_2\|_W)\,(\|\mathbf{u}\|_W + \|\mathbf{z}\|_W).$$

Proof. Similarly to the proof of Lemma A.1, we obtain

$$(F'(y_1) - F'(y_2))y = (f_1(\mathbf{u}_1) - f_1(\mathbf{u}_2),\ f_2(\mathbf{u}_1, \mathbf{z}_1) - f_2(\mathbf{u}_2, \mathbf{z}_2),\ 0)^T,$$

with

$$f_1(\mathbf{u}_i) := A'(\mathbf{u}_i)\mathbf{u} + R'(\varphi_i; \gamma)\varphi, \quad f_2(\mathbf{u}_i, \mathbf{z}_i) :=\ A''(\mathbf{u}_i)[\mathbf{u}, \cdot]^*\mathbf{z}_i + A'(\mathbf{u}_i)^*\mathbf{z} + R''(\varphi_i; \gamma)[\varphi, \cdot]^* z_i^\varphi + R'(\varphi_i; \gamma)^* z^\varphi.$$

Again, we observe that the difference in the last component is always zero. As in the previous Lemma, there is also no explicit dependence on q, q_1 or q_2 in either of the remaining components. We calculate again for f_2:

$$\begin{aligned} &f_2(\mathbf{u}_1, \mathbf{z}_1) - f_2(\mathbf{u}_2, \mathbf{z}_2) \\ &\quad = (A''(\mathbf{u}_1) - A''(\mathbf{u}_2))[\mathbf{u}, \cdot]^*\mathbf{z}_1 + A''(\mathbf{u}_2)[\mathbf{u}, \cdot]^*(\mathbf{z}_1 - \mathbf{z}_2) + (R''(\varphi_1; \gamma) - R''(\varphi_2; \gamma))[\varphi, \cdot]^* z_1^\varphi \\ &\qquad + R''(\varphi_2; \gamma)[\varphi, \cdot]^*(z_1^\varphi - z_2^\varphi) + (A'(\mathbf{u}_1) - A'(\mathbf{u}_2))^*\mathbf{z} + (R'(\varphi_1; \gamma) - R'(\varphi_2; \gamma))^* z^\varphi. \end{aligned}$$

The claim now follows analogously to Lemma A.1. Estimating the difference for f_1 in a similar way concludes the proof. We again omit the details. ☐

Finally, we need a quadratic bound for the second-order remainder of the derivative of F, that is used in the proof of Theorem 5.9. Note that in [19], Theorem 7.1, this bound immediately follows from second-order Fréchet-differentiability of F. However, in our case F is not twice Freçhet-differentiable due to the fact that the operator R is not three times Fréchet-differentiable. This requires some additional calculations.

Lemma A.3. *Let $\bar{y} \in Y_\infty$ and $y^k \in Y_\infty$ with $\|\mathbf{u}^k - \bar{\mathbf{u}}\|_W, \|\mathbf{z}^k - \bar{\mathbf{z}}\|_W, \le \tilde{\omega}_3$ for some $\tilde{\omega}_3 > 0$ be given. There exists a constant $c(\tilde{\omega}_3) > 0$ such that*

$$\|F(\bar{y}) - F(y^k) - F'(y^k)(\bar{y} - y^k)\|_{Z_\infty} \le c(\tilde{\omega}_3)\left(\|\mathbf{u}^k - \bar{\mathbf{u}}\|_W^2 + \|\mathbf{z}^k - \bar{\mathbf{z}}\|_W^2\right).$$

Proof. We observe

$$F(\bar{y}) - F(y^k) - F'(y^k)(\bar{y} - y^k) = (f_1(\bar{y}, y^k),\ f_2(\bar{y}, y^k),\ 0)^T,$$

with

$$f_1(\bar{y}, y^k) := A(\bar{\mathbf{u}}) - A(\mathbf{u}^k) - A'(\mathbf{u}^k)(\bar{\mathbf{u}} - \mathbf{u}^k) + R(\bar{\varphi}; \gamma) - R(\varphi^k; \gamma) - R'(\varphi^k; \gamma)(\bar{\varphi} - \varphi^k), \tag{A.3}$$

$$\begin{aligned} f_2(\bar{y}, y^k) := {} & (A'(\bar{\mathbf{u}}))^* \bar{\mathbf{z}} + R'(\bar{\varphi}; \gamma)^* \bar{z}^{\varphi} - (A'(\mathbf{u}^k))^* \mathbf{z}^k - R'(\varphi^k; \gamma)^* z^{\varphi,k} - A''(\mathbf{u}^k)[\bar{\mathbf{u}} - \mathbf{u}^k, \cdot]^* \mathbf{z}^k \\ & - A'(\mathbf{u}^k)^*(\bar{\mathbf{z}} - \mathbf{z}^k) - R''(\varphi^k; \gamma)[\bar{\varphi} - \varphi^k, \cdot]^* z^{\varphi,k} - R'(\varphi^k; \gamma)^*(\bar{z}^{\varphi} - z^{\varphi,k}). \end{aligned} \tag{A.4}$$

For the third term, we observe that the difference in the last component is always zero, and there is no explicit dependence on $\bar{q}$ or q^k. We again omit the estimation for f_1, and start with f_2 in the following way: For the first order terms in R we introduce the auxiliary functional $T\colon [0,1] \to \mathbb{R}$, $T(\theta) := R'(\varphi^k + \theta(\bar{\varphi} - \varphi^k); \gamma)^* \bar{z}^{\varphi}$. By Taylor's expansion $T(1) = T(0) + T'(\theta)$, for $\theta \in (0,1)$ we obtain

$$R'(\bar{\varphi}; \gamma)^* \bar{z}^{\varphi} - R'(\varphi^k; \gamma)^* \bar{z}^{\varphi} = R''(\varphi^k + \theta(\bar{\varphi} - \varphi^k); \gamma)[\cdot, \bar{\varphi} - \varphi^k])^* \bar{z}^{\varphi},$$

and analogously

$$A'(\bar{\mathbf{u}})^* \bar{z} - A'(\mathbf{u}^k)^* \bar{\mathbf{z}} = A''(\mathbf{u}^k + \theta(\bar{\mathbf{u}} - \mathbf{u}^k))[\cdot, \bar{\mathbf{u}} - \mathbf{u}^k])^* \bar{\mathbf{z}}.$$

Inserting this into (A.4) yields

$$\begin{aligned} f_2(\bar{y}, y^k) &= A''(\mathbf{u}^k)[\bar{\mathbf{u}} - \mathbf{u}^k, \cdot]^*(\bar{\mathbf{z}} - \mathbf{z}^k) + A''(\mathbf{u}^k + \theta(\bar{\mathbf{u}} - \mathbf{u}^k))[\cdot, \bar{\mathbf{u}} - \mathbf{u}^k])^* \bar{\mathbf{z}} \\ &\quad - R''(\varphi^k; \gamma)[\bar{\varphi} - \varphi^k, \cdot]^* z^{\varphi,k} + R''(\varphi^k + \theta(\bar{\varphi} - \varphi^k); \gamma)[\cdot, \bar{\varphi} - \varphi^k])^* \bar{z}^{\varphi} \\ &= (-A''(\mathbf{u}^k + \theta(\bar{\mathbf{u}} - \mathbf{u}^k)) - A''(\mathbf{u}^k))[\bar{\mathbf{u}} - \mathbf{u}^k, \cdot]^*(\bar{\mathbf{z}} - \mathbf{z}^k) + A''(\mathbf{u}^k + \theta(\bar{\mathbf{u}} - \mathbf{u}^k))[\cdot, \bar{\mathbf{u}} - \mathbf{u}^k])^* \bar{\mathbf{z}} \\ &\quad + (R''(\varphi^k + \theta(\bar{\varphi} - \varphi^k); \gamma) - R''(\varphi^k; \gamma))[\bar{\varphi} - \varphi^k, \cdot]^*(\bar{z}^{\varphi} - z^{\varphi,k}) + R''(\varphi^k + \theta(\bar{\varphi} - \varphi^k); \gamma)[\cdot, \bar{\varphi} - \varphi^k])^* \bar{z}^{\varphi}. \end{aligned}$$

Applying (3.12), (3.13), and (3.17) leads to

$$f_1(\bar{y}, y^k) = c(\tilde{\omega}_3)(\|\bar{\mathbf{u}} - \mathbf{u}^k\|_W^2 + \|\bar{\mathbf{z}} - \mathbf{z}^k\|_W^2).$$

The first component is estimated analogously to the second. We omit the details. ☐

Acknowledgements

We would like to thank the anonymous referees for valuable comments on the first version of this manuscript. This research was supported by the German Research Foundation (DFG) under grant number NE 1941/1-2 within the priority program Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization (SPP 1962).

References

[1] A. Hehl and I. Neitzel, Second-order optimality conditions of an optimal control problem governed by a regularized phase-field fracture propagation model. *Optimization* **72** (2023) 1665–1689.
[2] I. Neitzel, T. Wick and W. Wollner, An optimal control problem governed by a regularized phase-field fracture propagation model. *SIAM J. Control Optim.* **55** (2017) 2271–2288.
[3] I. Neitzel, T. Wick and W. Wollner, An optimal control problem governed by a regularized phase-field fracture propagation model. Part II: The regularization limit. *SIAM J. Control Optim.* **57** (2019) 1672–1690.

[4] B. Bourdin, G.A. Francfort and J.-J. Marigo, Numerical experiments in revisited brittle fracture. *J. Mech. Phys. Solids* **48** (2000) 797–826.
[5] B. Bourdin, G.A. Francfort and J.-J. Marigo, The variational approach to fracture. *J. Elast.* **91** (2008) 1–148.
[6] G.A. Francfort and J.-J. Marigo, Revisiting brittle fracture as an energy minimization problem. *J. Mech. Phys. Solids* **46** (1998) 1319–1342.
[7] L. Ambrosio and V.M. Tortorelli, On the approximation of free discontinuity problems. *Boll. Unione Math. Ital.* **1** (1992) 105–123.
[8] D. Knees, R. Rossi and C. Zanini, A vanishing viscosity approach to a rate-independent damage model. *Math. Models Methods Appl. Sci.* **23** (2013) 565–616.
[9] C. Meyer, A. Rademacher and W. Wollner, Adaptive optimal control of the obstacle problem. *SIAM J. Sci. Comput.* **37** (2015) A918–A945.
[10] A. Hehl, M. Mohammadi, I. Neitzel and W. Wollner, Optimizing fracture propagation using a phase-field approach, in *Non-Smooth and Complementarity-Based Distributed Parameter Systems: Simulation and Hierarchical Optimization*, edited by M. Hintermüller *et al.* Springer International Publishing, Cham (2022) 329–351.
[11] M. Mohammadi and W. Wollner, A priori error estimates for a linearized fracture control problem. *Optim. Eng.* **22** (2021) 2127–2149.
[12] D. Khimin, M.C. Steinbach and T. Wick, Space-time formulation, discretization, and computational performance studies for phase-field fracture optimal control problems. *J. Comput. Phys.* **470** (2022) 111554.
[13] D. Khimin, M.C. Steinbach and T. Wick, Optimal control for phase-field fracture: algorithmic concepts and Computations, in *Current Trends and Open Problems in Computational Mechanics*, edited by F. Aldakheel, B. Hudobivnik, M. Soleimani, H. Wessels, C. Weißenfels and M. Marino. Springer International Publishing, Cham (2022) 247–255.
[14] H. Goldberg and F. Tröltzsch, On a Lagrange–Newton method for a nonlinear parabolic boundary control problem. *Optim. Methods Softw.* **8** (1998) 225–247.
[15] F. Tröltzsch, An SQP method for optimal control of a nonlinear heat equation. *Control Cybern.* **23** (1994) 267–288.
[16] F. Tröltzsch, On the Lagrange–Newton-SQP method for the optimal control of semilinear parabolic equations. *SIAM J. Control Optim.* **38** (1999) 294–312.
[17] F. Tröltzsch and S. Volkwein, The SQP method for control constrained optimal control of the burgers equation. *ESAIM: Cont. Optim. Calc. Var.* **6** (2001) 649–674.
[18] A. Unger, Hinreichende Optimalitätsbedingungen 2. Ordnung und Konvergenz des SQP-Verfahrens für semilineare elliptische Randsteuerprobleme. Ph.D. thesis, Tech. U. Chemnitz (1997).
[19] R. Griesse, N. Metla and A. Rösch, Convergence analysis of the SQP method for nonlinear mixed-constrained elliptic optimal control problems. *ZAMM Z. Angew. Math. Mech.* **88** (2008) 776–792.
[20] R. Griesse, N. Metla and A. Rösch, Local quadratic convergence of SQP for elliptic optimal control problems with mixed control-state constraints. *Control Cybern.* **39** (2010) 717–738.
[21] M. Heinkenschloss, Formulation and analysis of a sequential quadratic programming method. *Optim. Control Appl. Optim.* **15** (1998) 178–203.
[22] M. Hintermüller and M. Hinze, A SQP-semismooth Newton-type algorithm applied to control of the instationary Navier–Stokes system subject to control constraints. *SIAM J. Optim.* **16** (2006) 1177–1200.
[23] M. Hinze and K. Kunisch, Second order methods for optimal control of time-dependent fluid flow. *SIAM J. Optim.* **40** (2001) 925–946.
[24] D. Wachsmuth, Analysis of the SQP-method for optimal control problems governed by the non-stationary Navier–Stokes equations. *ESAIM Cont. Optim. Calc. Var.* **12** (2007) 93–119.
[25] M. Heinkenschloss, The numerical solution of a control problem governed by a phase field model. *Optim. Methods Softw.* **7** (1997) 178–203.
[26] M. Heinkenschloss and F. Tröltzsch, Analysis of the Lagrange-SQP-Newton method for the control of a phase-field equation. *Control Cybern.* **28** (1999) 177–211.
[27] F. Hoppe and I. Neitzel, Convergence of the SQP method for quasilinear parabolic optimal control problems. *Optim. Eng.* **22** (2021).
[28] W. Alt, The Lagrange–Newton method for infinite-dimensional optimization problems. *Numer. Funct. Anal. Optim.* **11** (1990) 201–224.
[29] W. Alt, Local convergence of the Lagrange–Newton method with applications to optimal control. *Control Cybern.* **1–2** (1994) 87–105.

[30] N.H. Josephy, Newton's method for generalized equations. Tech. Summary Report, Mathematics Research Center, University of Wisconsin, Madison, WI (1979).

[31] S. Robinson, Strongly regular generalized equations. *Math. Oper. Res.* **5** (1980) 43–62.

[32] A. Dontchev, Implicit function theorems for generalized equations. *Math. Program.* **70** (1995) 91–106.

[33] E. Casas and F. Tröltzsch, Second order analysis for optimal control problems: Improving results expected from abstract theory. *SIAM J. Optim.* **22** (2012) 261–279.

[34] E. Casas and F. Tröltzsch, Second order optimality conditions and their role in PDE control. *Jahresber. Dtsch. Math.-Ver.* **117** (2015) 3–44.

[35] M.H. Farshbaf-Shaker and C. Heinemann, A phase-field approach for optimal boundary control of damage processes in two-dimensional viscoelastic media. *Math. Models Methods Appl. Sci.* **25** (2015) 2749–2793.

[36] M.H. Farshbaf-Shaker and C. Heinemann, Necessary conditions of first-order for an optimal boundary control problem for viscous damage processes in 2D. *ESAIM Control Optim. Calc. Var.* **24** (2018) 479–603.

[37] G. Allaire, F. Jouve and N. van Goethem, Damage and fracture evolution in brittle materials by shape optimization methods. *J. Comput. Phys.* **230** (2011) 5010–5044.

[38] A. Münch and P. Pedregal, Relaxation of an optimal design problem in fracture mechanic: The anti-plane case. *ESAIM: Control Optim. Calc. Var.* **16** (2010) 719–743.

[39] P. Destuynder, Remarks on a crack propagation control for stationary loaded structures. *CR Acad. Sci. Paris, Sér. IIb* **308** (1989) 697–701.

[40] P. Hild, A. Münch and Y. Ousset, On the active control of crack growth in elastic media. *Comput. Methods Appl. Mech. Eng.* **198** (2008) 407–419.

[41] L.M. Betz, Second-order sufficient optimality conditions for optimal control of nonsmooth, semilinear parabolic equations. *SIAM J. Control Optim.* **57** (2019) 4033–4062.

[42] L.M. Susu, Optimal control of a viscous two-field gradient damage model. *GAMM-Rep.* **40** (2018) 287–311.

[43] T. Wick, Multiphysics Phase-Field Fracture: Modeling, Adaptive Discretizations, and Solvers. De Gruyter, Berlin (2020).

[44] L. Blank, M.H. Farshbaf-Shaker, H. Garcke, C. Rupprecht and V. Styles, Multi-material phase field approach to structural topology Optimization, in *Trends in PDE Constrained Optim.*, edited by G. Leugering, P. Benner, S. Engell, A. Griewank, H. Harbrecht, M. Hinze, R. Rannacher and S. Ulbrich. Springer International Publishing, Cham (2014) 231–246.

[45] L. Blank, M.H. Farshbaf-Shaker, C. Hecht, J. Michl and C. Rupprecht, Optimal control of Allen–Cahn systems, in *Trends in PDE Constrained Optim.*, edited by G. Leugering, P. Benner, S. Engell, A. Griewank, H. Harbrecht, M. Hinze, R. Rannacher and S. Ulbrich. Springer International Publishing, Cham (2014) 11–26.

[46] L. Blank, H. Garcke, C. Hecht and C. Rupprecht, Sharp interface limit for a phase field model in structural optimization. *SIAM J. Control Optim.* **54** (2016) 1559–1584.

[47] H. Garcke, K.F. Lam and A. Signori, On a phase field model of Cahn–Hilliard type for tumour growth with mechanical effects. *Nonlinear Anal.: Real World Appl.* **57** (2021) 103192.

[48] H. Garcke and D. Trautwein, Numerical analysis for a Cahn–-Hilliard system modelling tumour growth with chemotaxis and active transport. *J. Num. Math.* **30** (2022) 295–324.

[49] D. Mumford and J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems. *Commun. Pure Appl. Math.* **42** (1989) 577–685.

[50] B. Bourdin, Image segmentation with a finite element method. *ESAIM: M2AN* **33** (1999) 229–244.

[51] B. Bourdin and A. Chambolle, Implementation of an adaptive finite-element approximation of the Mumford-Shah functional. *Numer. Math.* **85** (2000) 609–646.

[52] O. Scherzer, editor, *Handbook of Mathematical Methods in Imaging.* Springer Berlin Heidelberg, Berlin, Heidelberg (2020).

[53] L. De Lorenzis and T. Gerasimov, Numerical implementation of phase-field models of brittle fracture, in *Modeling in Engineering Using Innovative Numerical Methods for Solids and Fluids*, edited by L. De Lorenzis and A. Düster. Springer International Publishing, Cham (2020) 75–101.

[54] K. Gröger, A $W^{1,p}$-estimate for solutions to mixed boundary value problems for second order elliptic differential equations. *Math. Ann.* **283** (1989) 679–687.

[55] R. Rannacher, *Probleme der Kontinuumsmechanik und ihre numerische Behandlung.* Heidelberg University Publishing, Heidelberg, Germany (2017).

[56] R. Haller-Dintelmann, H. Meinlschmidt and W. Wollner, Higher regularity for solutions to elliptic systems in divergence form subject to mixed boundary conditions. *Ann. Math. Pura Appl.* **198** (2018) 1227–1241.

[57] F. Tröltzsch, Lipschitz stability of solutions to linear-quadratic parabolic control problems with respect to perturbations. *Dynam. Cont. Disc. Impul. Syst.* **7** (2000).

[58] F. Tröltzsch, Optimal control of partial differential equations: theory, methods and applications. Vol. 112 of *Grad. Stud. Math.* American Math. Soc, Providence, RI (2010).

[59] W. Alt, Discretization and mesh independence of Newton's method for generalized equations. *Math. Program. Data Perturbations* (1990) 1–30.

[60] L. Bonifacius and I. Neitzel, Second order optimality conditions for optimal control of quasilinear parabolic equations. *Math. Control Relat. Fields* **8** (2018) 1–34.

[61] F. Tröltzsch and D. Wachsmuth, Second-order sufficient optimality conditions for the optimal control of Navier–Stokes equations. *ESAIM: Control Optim. Calc. Var.* **12** (2006) 93–119.

[62] E. Casas and F. Tröltzsch, A general theorem on error estimates with application to a quasilinear elliptic control problem. *Comput. Optim. Appl.* **53** (2012) 173–206.

Please help to maintain this journal in open access!

This journal is currently published in open access under the Subscribe to Open model (S2O). We are thankful to our subscribers and supporters for making it possible to publish this journal in open access in the current year, free of charge for authors and readers.

Check with your library that it subscribes to the journal, or consider making a personal donation to the S2O programme by contacting subscribers@edpsciences.org.

More information, including a list of supporters and financial transparency reports, is available at https://edpsciences.org/en/subscribe-to-open-s2o.