SIAM J. OPTIM.
Vol. 30, No. 1, pp. 290–318

GRADIENT-BASED SOLUTION ALGORITHMS FOR A CLASS OF BILEVEL OPTIMIZATION AND OPTIMAL CONTROL PROBLEMS WITH A NONSMOOTH LOWER LEVEL*

CONSTANTIN CHRISTOF†

Abstract. The aim of this paper is to explore a peculiar regularization effect that occurs in the sensitivity analysis of certain elliptic variational inequalities of the second kind. The effect causes the solution operator of the variational inequality at hand to be continuously Fréchet differentiable, although the problem itself contains nondifferentiable terms. Our analysis shows in particular that standard gradient-based algorithms can be used to solve bilevel optimization and optimal control problems that are governed by elliptic variational inequalities of the considered type—all without regularizing the nondifferentiable terms in the lower-level problem and without losing desirable properties of the solution such as, e.g., sparsity. Our results can, for instance, be used in the optimal control of Casson fluids and in bilevel optimization approaches for parameter learning in total variation image denoising models.

Key words. optimal control, nonsmooth optimization, bilevel optimization, elliptic variational inequality of the second kind, Casson fluid, total variation, machine learning, parameter identification

AMS subject classifications. 49J40, 49J52, 49N45, 68U10, 76A05, 90C33

DOI. 10.1137/18M1225707

1. Introduction and problem statement.

The aim of this paper is to study finite-dimensional optimization problems of the type

$$
\text{(P)}\quad \left\{\begin{aligned}
&\min J(y,u,\alpha,\beta)\\
&\text{s.t. } y,u\in\mathbb{R}^n,\ \alpha,\beta\in\mathbb{R}^m,\ (u,\alpha,\beta)\in U_{ad},\\
&\qquad \langle A(y),v-y\rangle+\sum_{k=1}^m\omega_k\big(\alpha_k\|G_kv\|+\beta_k\|G_kv\|^{1+\gamma}\big)\\
&\qquad\qquad -\sum_{k=1}^m\omega_k\big(\alpha_k\|G_ky\|+\beta_k\|G_ky\|^{1+\gamma}\big)\geq\langle Bu,v-y\rangle\ \ \forall v\in\mathbb{R}^n.
\end{aligned}\right.
$$

Our standing assumptions on the quantities in (P) are as follows.

ASSUMPTION 1.1 (standing assumptions and notation).

1. $l,m,n\in\mathbb{N}$, $\omega\in(0,\infty)^m$, $B\in\mathbb{R}^{n\times n}$, *and* $\gamma\in(0,1)$ *are given and fixed;*
2. $G_k\in\mathbb{R}^{l\times n}$, $k=1,\dots,m$, *are given matrices;*
3. $J:\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^m\to\mathbb{R}$ *is a continuously Fréchet differentiable function;*
4. U_{ad} *is a (sufficiently nice) nonempty subset of* $\mathbb{R}^n\times[0,\infty)^m\times(0,\infty)^m$*;*
5. $A:\mathbb{R}^n\to\mathbb{R}^n$ *is a continuously Fréchet differentiable operator that is strongly*

*Received by the editors November 7, 2018; accepted for publication (in revised form) September 30, 2019; published electronically January 23, 2020.
https://doi.org/10.1137/18M1225707
Funding: This research was conducted within the International Research Training Group IGDK 1754, funded by the German Science Foundation (DFG) and the Austrian Science Fund (FWF), and in association with project P16 of SPP 1962 "Non-Smooth and Complementarity-Based Distributed Parameter Systems: Simulation and Hierarchical Optimization."
†Chair of Optimal Control, Center for Mathematical Sciences, Technische Universität München, M17, 85748 Garching, Germany (christof@ma.tum.de).

monotone; i.e., there is a constant $c > 0$ with

$$\langle A(v_1) - A(v_2), v_1 - v_2 \rangle \geq c\|v_1 - v_2\|^2 \quad \forall v_1, v_2 \in \mathbb{R}^n;$$

6. $\|\cdot\|$ *denotes the Euclidean norm, and* $\langle \cdot, \cdot \rangle$ *denotes the Euclidean scalar product (we use the same symbols on* $\mathbb{R}^l$, $\mathbb{R}^m$, *and* $\mathbb{R}^n$*).*

Note that if the operator $A : \mathbb{R}^n \to \mathbb{R}^n$ can be identified with the gradient field of a convex and Fréchet differentiable function $a : \mathbb{R}^n \to \mathbb{R}$, i.e., if $a'(v) = A(v) \in \mathbb{R}^n$ holds for all $v \in \mathbb{R}^n$, then (P) is equivalent to the bilevel minimization problem

$$\text{(1.1)} \qquad \begin{cases} \min J(y, u, \alpha, \beta) \\ \text{s.t.} \;\; y, u \in \mathbb{R}^n, \;\; \alpha, \beta \in \mathbb{R}^m, \; (u, \alpha, \beta) \in U_{ad}, \\ \qquad y = \underset{v \in \mathbb{R}^n}{\arg\min}\, a(v) + \displaystyle\sum_{k=1}^m \omega_k \Big(\alpha_k \|G_k v\| + \beta_k \|G_k v\|^{1+\gamma} \Big) - \langle Bu, v \rangle \,. \end{cases}$$

For a proof of this equivalence, we refer the reader to [11, section 1.2]. Since elliptic variational inequalities involving nonconservative vector fields $A : \mathbb{R}^n \to \mathbb{R}^n$ appear naturally in some applications (cf. the references in [35, section II-2.1]), in this paper we work with the more general formulation (P) and not with (1.1).

Optimization problems of the type (P) arise, for instance, in the optimal control of non-Newtonian fluids, in glaciology, and in bilevel parameter learning approaches for variational image denoising models. See, e.g., [7, 8, 9, 16, 20, 21, 23, 24, 37, 38, 40, 42, 44, 47, 51, 52, 53, 55, 56] and the references therein, and also see the two tangible examples in section 2. The main difficulty in the study of the problem (P) is the nonsmoothness of the Euclidean norms present on the lower level. Because of these nondifferentiable terms, standard results and analytical tools are typically inapplicable, and one has to resort to either regularization techniques or concepts from nonsmooth analysis to derive, e.g., necessary optimality conditions for (P), or to devise feasible numerical solution algorithms. In this context, compare, for instance, with the results on bilevel problems with C^1-objective functions in [5, 25, 30, 49, 50, 67, 69, 70], with the analysis of nonsmooth problems in [27, 29, 68], and with the regularization approaches of [9, 16, 23, 38, 40, 44, 46, 56, 66]. For an overview on bilevel programming, see also [26, 28, 48].

The aim of this paper is to demonstrate that in the situation of Assumption 1.1, problems of the type (P) can also be solved without replacing the involved Euclidean norms with smooth approximations and without working with sophisticated instruments from nonsmooth analysis (and the associated constraint qualifications; cf. [29] and [70, sections 2, 3]). To be more precise, in what follows we prove the rather surprising fact that the solution operator $S : (u, \alpha, \beta) \mapsto y$ associated with the inner elliptic variational inequality in (P) is continuously Fréchet differentiable as a function $S : \mathbb{R}^n \times [0, \infty)^m \times (0, \infty)^m \to \mathbb{R}^n$ (see Theorem 3.3 for the main result). This quite counterintuitive behavior makes it possible to tackle minimization problems of the type (P) with standard gradient-based solution algorithms, even without regularizing the nonsmooth terms on the lower level. Avoiding such a regularization is highly desirable in many situations as the Euclidean norms in (P) typically cause the inner solutions y to have certain properties (such as sparsity) that are very important from the application point of view. Before we begin with our investigation, we give a short overview of the content and the structure of this paper.

In section 2, we first give two tangible examples of problems that fall under the scope of our analysis—one arising in the optimal control of non-Newtonian fluids and the other from the field of bilevel parameter learning in variational image denoising models. The examples found in this section illustrate that our results are not only of academic interest but also relevant in practice.

In section 3, we then address the sensitivity analysis of the inner elliptic variational inequality in (P). Here, we prove that the solution operator $S : (u, \alpha, \beta) \mapsto y$ is indeed continuously Fréchet differentiable as a function from $\mathbb{R}^n \times [0, \infty)^m \times (0, \infty)^m$ to $\mathbb{R}^n$, and we also give some comments, e.g., on the extension of our results to the infinite-dimensional setting.

Section 4 is concerned with the consequences of the results of the previous sections on the study and the numerical solution of bilevel optimization problems of the form (P). The main result of this section, Theorem 4.1, shows that gradients of the reduced objective function of (P) can be calculated via an adjoint calculus and by solving certain quadratic auxiliary problems with equality constraints.

Lastly, in section 5, we demonstrate by means of a numerical example that the differentiability of the solution map S indeed allows us to solve problems of the type (P) with standard gradient-based algorithms.

2. Two application examples.

In what follows, we discuss in more detail two application examples that are covered by the general problem formulation (P) and that may serve as motivation for the analysis in sections 3 and 4.

2.1. Optimal control of Casson fluids.

As a first example, we consider the following problem that arises in the optimal control of non-Newtonian fluids: Suppose that $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2\}$, is a simply connected, bounded, polyhedral domain; let $H_0^1(\Omega)$ and $L^p(\Omega)$ for $1 \leq p \leq \infty$ be defined as usual (see [1, 3] or other standard references); and let $(\bar{u}, \alpha) \in L^2(\Omega) \times (0, \infty)$ be arbitrary but fixed. Then, the so-called Mosolov problem for Casson fluids is given by

$$\bar{y} = \underset{\bar{v} \in H_0^1(\Omega)}{\arg\min} \int_\Omega \frac{1}{2}\|\nabla \bar{v}\|^2 + \frac{4}{3}\alpha^{1/2}\|\nabla \bar{v}\|^{3/2} + \alpha\|\nabla \bar{v}\| - \bar{u}\bar{v}\,\mathrm{d}x. \tag{2.1}$$

Here, ∇ denotes the weak gradient, and the bars indicate that we talk about functions and not about elements of the Euclidean space. In non-Newtonian fluid mechanics, the main interest in the minimization problem (2.1) stems from the fact that it models the unidirectional, stationary flow of a viscoplastic medium of Casson type between two plates with distance $\operatorname{diam}(\Omega)$ in the case $d = 1$ and in a cylindrical pipe with cross-section Ω in the case $d = 2$; cf. [42, 51, 52, 53] and the references therein. In this context, $\bar{u}$ is the pressure gradient parallel to the two enclosing plates or the pipe axis driving the fluid, $\bar{y}$ is the fluid velocity in the direction of $\bar{u}$ (i.e., perpendicular to Ω), and α is a material parameter (the generalized Oldroyd number); see Figure 1. Recall that the characteristic feature of a viscoplastic medium is that it behaves like a fluid everywhere, where the shear stress exceeds a certain threshold (the so-called yield stress), and that it behaves like a solid otherwise. For a Casson fluid, the behavior in fluid regions is additionally governed by a nonlinear relation between the shear rate and the shear stress; cf. [42, section 2.2]. In the model (2.1), the regions where rigid material behavior occurs are precisely those parts of the domain Ω where the gradient $\nabla\bar{y}$ vanishes, and the sudden change of the material behavior at the yield stress and the nonlinear material laws in the fluid regions are incorporated via the terms $\int_\Omega \alpha\|\nabla\bar{v}\|\mathrm{d}x$ and $\int_\Omega \frac{1}{2}\|\nabla\bar{v}\|^2 + \frac{4}{3}\alpha^{1/2}\|\nabla\bar{v}\|^{3/2}\mathrm{d}x$, respectively; see the derivation in [42, section 2]. Note that this implies in particular that the nondifferentiability of the objective in (2.1) is directly related to the underlying physical model, and that the nonsmoothness is of special importance in the above situation.

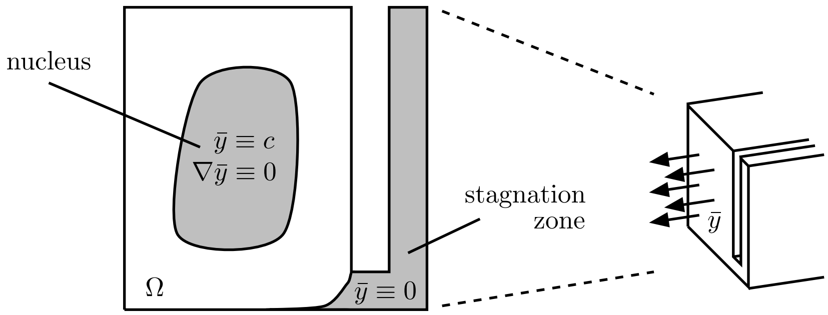


FIG. 1. *Typical flow behavior in the situation of the two-dimensional Mosolov problem with a constant pressure drop $\bar{u}$. The viscoplastic medium forms a solid nucleus in the middle of the fluid domain that moves with a constant velocity c along the pipe axis and sticks to the boundary in those regions of Ω where the pressure gradient $\bar{u}$ is too low to move the fluid (so-called stagnation zones). An analogous behavior can be observed in the case $d = 1$ for the flow of a Casson fluid between two plates, cf. the numerical results in section* 5.

Suppose now that we want to determine a pressure gradient $\bar{u} \in L^2(\Omega)$ and a material parameter $\alpha \in (0, \infty)$ such that the flow profile $\bar{y} \in H_0^1(\Omega)$ in Ω has a certain shape $\bar{y}_D \in C(\overline{\Omega}) \cap H_0^1(\Omega)$, where $C(\overline{\Omega})$ denotes the space of continuous functions on the closure of the domain Ω. Then, it is a natural approach to consider a tracking-type optimal control problem of the form

$$\text{(2.2)} \qquad \begin{cases} \min \dfrac{1}{2}\|\bar{y} - \bar{y}_D\|^2_{L^2(\Omega)} + \dfrac{\mu}{2}\left(\|\bar{u}\|^2_{L^2(\Omega)} + \alpha^2\right) \\ \text{s.t. } \ \bar{y} \in H_0^1(\Omega), \ \bar{u} \in L^2(\Omega), \ \alpha \in \tilde{U}_{ad}, \\ \qquad \bar{y} = \underset{\bar{v} \in H_0^1(\Omega)}{\arg\min} \displaystyle\int_\Omega \frac{1}{2}\|\nabla \bar{v}\|^2 + \frac{4}{3}\alpha^{1/2}\|\nabla \bar{v}\|^{3/2} + \alpha\|\nabla \bar{v}\| - \bar{u}\bar{v}\, \mathrm{d}x, \end{cases}$$

where $\tilde{U}_{ad}$ is some nonempty, convex, and closed subset of $(0, \infty)$, and where $\mu > 0$ is a fixed Tychonoff parameter. Let us briefly check that the above problem is indeed sensible.

THEOREM 2.1 (solvability of (2.2)). *Assume that $\bar{y}_D$, Ω, μ, and $\tilde{U}_{ad}$ are as before. Then,* (2.2) *admits at least one solution $(\bar{u}^*, \alpha^*) \in L^2(\Omega) \times \tilde{U}_{ad}$.*

Proof. From standard arguments (as found, e.g., in [35, Lemma 4.1]), it follows straightforwardly that the lower-level problem in (2.2) possesses a well-defined solution operator $S : L^2(\Omega) \times [0, \infty) \to H_0^1(\Omega)$, $(\bar{u}, \alpha) \mapsto \bar{y}$. Furthermore, it is easy to check (using the weak lower semicontinuity of convex and continuous functions) that this solution map is weak-to-weak continuous; i.e., for every sequence $\{(\bar{u}_i, \alpha_i)\} \subset L^2(\Omega) \times [0, \infty)$ with $\bar{u}_i \rightharpoonup \bar{u}$ in $L^2(\Omega)$ and $\alpha_i \to \alpha$ in $\mathbb{R}$ for $i \to \infty$, we have $S(\bar{u}_i, \alpha_i) \rightharpoonup S(\bar{u}, \alpha)$ in $H_0^1(\Omega)$. The claim now follows immediately from the direct method of calculus of variations and the compactness of the embedding $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$. □

To transform (2.2) into a problem that can be solved numerically, we consider a standard finite element discretization with piecewise linear ansatz functions. More precisely, we assume the following.

ASSUMPTION 2.2 (assumptions and notation for the discretization of (2.2)).

1. $\mathcal{T} = \{T_k\}_{k=1}^m$, $m \in \mathbb{N}$, *is a triangulation of Ω consisting of simplices T_k (see, e.g.,* [10, *Definition* 2] *for the precise definition of the term "triangulation");*

2. $\{x_i\}_{i=1}^n$, $n \in \mathbb{N}$, *are the nodes of* $\mathcal{T}$ *that are contained in* Ω*;*
3. $V_h := \{\bar v \in C(\overline{\Omega}) \mid \bar v$ *is affine on* T_k *for all* $k = 1, \ldots, m$ *and* $\bar v|_{\partial\Omega} = 0\}$*;*
4. $\{\varphi_i\}$ *is the nodal basis of* V_h*, i.e.,* $\varphi_i(x_i) = 1$ *for all* i*,* $\varphi_i(x_j) = 0$ *for* $i \neq j$*.*

By replacing the spaces $H_0^1(\Omega)$ and $L^2(\Omega)$ in (2.2) with V_h, we now arrive at a finite-dimensional minimization problem of the following form:

(2.3)
$$\begin{cases} \min \dfrac{1}{2}\left\langle B(y-y_D), y-y_D\right\rangle + \dfrac{\mu}{2}\left(\langle Bu,u\rangle + \alpha^2\right) \\ \text{s.t.}\ \ y,u \in \mathbb{R}^n,\ \alpha \in \tilde U_{ad}, \\ \qquad y = \operatorname*{arg\,min}_{v\in\mathbb{R}^n} \dfrac{1}{2}\langle Av,v\rangle + \displaystyle\sum_{k=1}^m |T_k|\left(\alpha\|G_k v\| + \frac{4}{3}\alpha^{1/2}\|G_k v\|^{3/2}\right) - \langle Bu, v\rangle. \end{cases}$$

Here, y, u, and y_D are the coordinate vectors of the discretized state, the discretized control, and the Lagrange interpolate of $\bar y_D$ w.r.t. the nodal basis $\{\varphi_i\}$, respectively; A and B denote the stiffness and the mass matrix, i.e.,

$$A := \left(\int_\Omega \langle\nabla\varphi_i, \nabla\varphi_j\rangle\,\mathrm{d}x\right)_{i,j=1,\ldots,n}, \qquad B := \left(\int_\Omega \varphi_i\varphi_j\,\mathrm{d}x\right)_{i,j=1,\ldots,n};$$

$|T_k|$ is the d-dimensional volume of the simplex T_k; and $G_k \in \mathbb{R}^{d\times n}$ is the matrix that maps a coordinate vector $v \in \mathbb{R}^n$ to the gradient of the associated finite element function on the cell T_k, i.e.,

$$G_k v = \nabla\left(\sum_{i=1}^n v_i\varphi_i\right)\Bigg|_{T_k} \in \mathbb{R}^d \quad \forall v \in \mathbb{R}^n \quad \forall k = 1,\ldots,m.$$

Note that (2.3) is precisely of the form (1.1) (with $\omega_k := |T_k|$ and an appropriately defined $U_{ad} \subset \mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$). This shows that, after a discretization, the minimization problem (2.2) indeed falls under the scope of the general setting introduced in section 1, and that our analysis indeed allows us to study optimal control problems for Casson fluids. We will get back to this topic in section 5, where (2.3) will serve as a model problem for our numerical experiments.

2.2. Bilevel optimization approaches for parameter learning. As a second application example, we consider a bilevel optimization problem that has been proposed in [44] as a framework for parameter learning in variational image denoising models (cf. also [7, 56] and the more general setting in [41] in this context). The problem takes the form

$$\text{(2.4)} \qquad \begin{cases} \min \|y-g\|^2 \\ \text{s.t.}\ \ y \in \mathbb{R}^n,\ \vartheta \in [0,\infty)^q, \\ \qquad y = \operatorname*{arg\,min}_{v\in\mathbb{R}^n} \dfrac{1}{2}\|v-f\|^2 + \displaystyle\sum_{i=1}^q \vartheta_i\left(\sum_{j=1}^r |(K_i v)_j|^p\right). \end{cases}$$

Here, n, q, and r are natural numbers; p is some exponent in $[1,\infty)$; $g \in \mathbb{R}^n$ is the given ground truth data; $f \in \mathbb{R}^n$ is the noisy image; the terms $\sum_{j=1}^r |(K_i v)_j|^p$, $i = 1,\ldots,q$, are so-called analysis-based priors involving matrices $K_i \in \mathbb{R}^{r\times n}$; ϑ is the learning parameter; and $\frac{1}{2}\|v-f\|^2$ and $\sum_{i=1}^q \vartheta_i(\sum_{j=1}^r |(K_i v)_j|^p)$ are the fidelity and the regularization term of the underlying denoising model, respectively (cf. the classical total variation denoising method). For more details on the background of (2.4), we refer the reader to [44] and the references therein.

Suppose now that we enrich the model (2.4) by allowing the exponent p to depend on i so that the framework also covers the case of composite $\ell_{p_1}^{p_1} + \ell_{p_2}^{p_2}$-regularizers on the lower level as used, e.g., in the context of elastic net regularization; see [61, 71]. Then, in the special situation $q = 2s$, $s \in \mathbb{N}$, $K_i = K_{i+s}$ for all $i = 1, \ldots, s$, $p_i = 1$ for all $i = 1, \ldots, s$, and $p_{i+s} = 1 + \gamma$ for all $i = 1, \ldots, s$, with an arbitrary but fixed $\gamma \in (0, 1)$, the problem (2.4) can also be written as

(2.5)
$$\begin{cases} \min \|y - g\|^2 \\ \text{s.t. } \ y \in \mathbb{R}^n, \ \vartheta, \tilde{\vartheta} \in [0, \infty)^s, \\ \qquad y = \underset{v \in \mathbb{R}^n}{\arg\min}\, \dfrac{1}{2}\|v - f\|^2 + \displaystyle\sum_{i=1}^{s} \vartheta_i \left(\sum_{j=1}^{r} |(K_i v)_j| \right) + \tilde{\vartheta}_i \left(\sum_{j=1}^{r} |(K_i v)_j|^{1+\gamma} \right). \end{cases}$$

Note that it makes sense to consider the exponents $p_i = 1$, $i = 1, \ldots, s$, here since this choice ensures that the priors are sparsity promoting (due to the induced nonsmoothness; cf. [64]). If we replace the constraint on $\tilde{\vartheta}$ in (2.5) with $\tilde{\vartheta} \in [\varepsilon, \infty)^s$ for some $0 < \varepsilon \ll 1$ (a modification that is also commonly made in learning problems to resolve solvability or stability issues; see [38, 40, 41]), define

$$\alpha_{ij} := \vartheta_i, \qquad \beta_{ij} := \tilde{\vartheta}_i, \qquad G_{ij} : \mathbb{R}^n \to \mathbb{R}, \ \ v \mapsto (K_i v)_j; \tag{2.6}$$

use the binomial identities; exploit the fact that terms which depend only on f are irrelevant in the lower-level problem; and identify f with u; then (2.5) takes the form

$$\begin{aligned} &\min \|y - g\|^2 \\ &\text{s.t. } \ y, u \in \mathbb{R}^n, \ \alpha, \beta \in \mathbb{R}^{rs}, \ (u, \alpha, \beta) \in U_{ad}, \\ &\qquad y = \underset{v \in \mathbb{R}^n}{\arg\min}\, \frac{1}{2}\|v\|^2 + \sum_{i=1}^{s}\sum_{j=1}^{r} \alpha_{ij}\, |G_{ij} v| + \beta_{ij}\, |G_{ij} v|^{1+\gamma} - \langle u, v \rangle \end{aligned}$$

with an appropriately defined admissible set $U_{ad} \subset \mathbb{R}^n \times [0, \infty)^{rs} \times (0, \infty)^{rs}$ which ensures the equality $u = f$ and enforces that α_{ij} and β_{ij} depend only on i (cf. (2.6)). The above problem is again of the type (1.1) and satisfies all conditions in Assumption 1.1. This shows that the general setting of section 1 is also of interest in the study of parameter learning problems for variational image denoising models.

3. Solvability and sensitivity analysis of the inner problem in (P).

Having demonstrated that the general framework of section 1 indeed covers situations that are relevant for practical applications, we now turn our attention to the inner elliptic variational inequality in (P), i.e., to the problem

$$\text{(V)} \quad \begin{cases} y \in \mathbb{R}^n, \\ \langle A(y), v - y \rangle + \displaystyle\sum_{k=1}^{m} \omega_k \Big(\alpha_k \|G_k v\| + \beta_k \|G_k v\|^{1+\gamma} \Big) \\ \qquad\qquad - \displaystyle\sum_{k=1}^{m} \omega_k \Big(\alpha_k \|G_k y\| + \beta_k \|G_k y\|^{1+\gamma} \Big) \geq \langle Bu, v - y \rangle \quad \forall v \in \mathbb{R}^n. \end{cases}$$

Here and in what follows, we always assume that $l, m, n, \omega, B, \gamma, G_k$, and A satisfy the conditions in Assumption 1.1. Let us first check that (V) is well-posed.

Proposition 3.1 (solvability). *The variational inequality* (V) *admits a unique solution $y \in \mathbb{R}^n$ for all $u \in \mathbb{R}^n$, $\alpha \in [0,\infty)^m$, and $\beta \in [0,\infty)^m$. This solution satisfies*

$$\text{(3.1)}\quad \begin{cases} \langle A(y), z\rangle + \displaystyle\sum_{k=1}^m \omega_k \Big(\alpha_k \|\cdot\|'(G_k y; G_k z) + \beta_k (\|\cdot\|^{1+\gamma})'(G_k y; G_k z)\Big) \geq \langle Bu, z\rangle \\ \hfill \forall z \in \mathbb{R}^n, \end{cases}$$

where $H'(x;h)$ denotes the directional derivative of a function $H : \mathbb{R}^l \to \mathbb{R}$ at a point $x \in \mathbb{R}^l$ in a direction $h \in \mathbb{R}^l$. Further, there exists a constant $C > 0$ independent of u, α, and β with $\|y\| \leq C\|u\|$ for all u, α, and β.

Proof. The unique solvability of (V) for all $u \in \mathbb{R}^n$, $\alpha \in [0,\infty)^m$, and $\beta \in [0,\infty)^m$ is a straightforward consequence of Browder's theorem; see [60, Theorem 3.43] and [11, Theorem 1.2.2]. To obtain the variational inequality (3.1), it suffices to choose vectors of the form $v = y + tz$, $t > 0$, $z \in \mathbb{R}^n$, in (V), to divide by t, and to pass to the limit $t \searrow 0$. The bound $\|y\| \leq C\|u\|$ finally follows from (V) when we choose $v = 0$ and exploit the strong monotonicity of A. ∎

As a starting point for our sensitivity analysis, we prove the following.

Proposition 3.2 (Lipschitz continuity of the solution map). *For every $M > 0$ there exists a constant $C > 0$ depending only on ω, M, A, B, and G_k such that the solution map $S : \mathbb{R}^n \times [0,\infty)^m \times [0,\infty)^m \to \mathbb{R}^n$, $(u, \alpha, \beta) \mapsto y$, associated with* (V) *satisfies*

$$\text{(3.2)}\qquad \|S(u_1, \alpha_1, \beta_1) - S(u_2, \alpha_2, \beta_2)\| \leq C\Big(\|u_1 - u_2\| + \|\alpha_1 - \alpha_2\| + \|\beta_1 - \beta_2\|\Big)$$

for all $(u_1, \alpha_1, \beta_1), (u_2, \alpha_2, \beta_2) \in \mathbb{R}^n \times [0,\infty)^m \times [0,\infty)^m$ with $\|u_1\|, \|u_2\| \leq M$.

Proof. We proceed along the lines of [13, Theorem 2.6]: Suppose that a constant $M > 0$ is given; consider some $(u_1, \alpha_1, \beta_1), (u_2, \alpha_2, \beta_2) \in \mathbb{R}^n \times [0,\infty)^m \times [0,\infty)^m$ with $\|u_1\|, \|u_2\| \leq M$; and denote the solutions of (V) associated with the triples (u_1, α_1, β_1) and (u_2, α_2, β_2) with y_1 and y_2, respectively. Then, (3.1) yields

$$\langle A(y_i), z\rangle + \sum_{k=1}^m \omega_k \Big(\alpha_{i,k} \|\cdot\|'(G_k y_i; G_k z) + \beta_{i,k} (\|\cdot\|^{1+\gamma})'(G_k y_i; G_k z)\Big) \geq \langle Bu_i, z\rangle$$

for all $z \in \mathbb{R}^n$ and $i = 1, 2$, and we may choose $z = y_2 - y_1$ in the inequality for y_1

and $z = y_1 - y_2$ in the inequality for y_2, and add the resulting inequalities to obtain

$$
\begin{aligned}
(3.3)\quad &\langle A(y_1) - A(y_2), y_1 - y_2\rangle \\
&\quad \le \langle B(u_1 - u_2), y_1 - y_2\rangle \\
&\quad + \sum_{k=1}^{m} \omega_k(\alpha_{1,k} - \alpha_{2,k})\Big(\|\cdot\|'(G_k y_1; G_k(y_2 - y_1))\Big) \\
&\quad + \sum_{k=1}^{m} \omega_k(\beta_{1,k} - \beta_{2,k})\Big((\|\cdot\|^{1+\gamma})'(G_k y_1; G_k(y_2 - y_1))\Big) \\
&\quad + \sum_{k=1}^{m} \omega_k \alpha_{2,k}\Big(\|\cdot\|'(G_k y_1; G_k(y_2 - y_1)) + \|\cdot\|'(G_k y_2; G_k(y_1 - y_2))\Big) \\
&\quad + \sum_{k=1}^{m} \omega_k \beta_{2,k}\Big((\|\cdot\|^{1+\gamma})'(G_k y_1; G_k(y_2 - y_1)) + (\|\cdot\|^{1+\gamma})'(G_k y_2; G_k(y_1 - y_2))\Big).
\end{aligned}
$$

Due to the convexity of the functions $\|\cdot\|$ and $\|\cdot\|^{1+\gamma}$ and the nonnegativity of the vectors α, β, and ω, the last two sums on the right-hand side of (3.3) are nonpositive and can be ignored (see [13, Lemma 2.3e)]. Further, we obtain from the Lipschitz continuity of $\|\cdot\|$ and $\|\cdot\|^{1+\gamma}$ on bounded subsets of $\mathbb{R}^l$ and from Proposition 3.1 that there exists a constant $C = C(\omega, M, G_k) > 0$ with

$$
\begin{aligned}
&\sum_{k=1}^{m} \omega_k(\alpha_{1,k} - \alpha_{2,k})\Big(\|\cdot\|'(G_k y_1; G_k(y_2 - y_1))\Big) \\
&\quad + \sum_{k=1}^{m} \omega_k(\beta_{1,k} - \beta_{2,k})\Big((\|\cdot\|^{1+\gamma})'(G_k y_1; G_k(y_2 - y_1))\Big) \\
&\qquad \le C\left(\|\alpha_1 - \alpha_2\| + \|\beta_1 - \beta_2\|\right)\|y_1 - y_2\|.
\end{aligned}
$$

The claim now follows immediately from (3.3), the strong monotonicity of A, and the Cauchy–Schwarz inequality. ∎

We are now in the position to prove the main result of this paper.

THEOREM 3.3 (continuous Fréchet differentiability of the solution map). *The solution operator $S : (u, \alpha, \beta) \mapsto y$ associated with the variational inequality* (V) *is continuously Fréchet differentiable as a function $S : \mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m \to \mathbb{R}^n$; i.e., there exists a continuous map S' which maps $\mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$ into the space $L(\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$ of linear operators from $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$ to $\mathbb{R}^n$ such that, for every $w \in \mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$, we have*

$$
(3.4)\qquad \lim_{\substack{\|h\| \to 0,\\ w+h \in \mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m}} \frac{\|S(w+h) - S(w) - S'(w)h\|}{\|h\|} = 0.
$$

Moreover, for every triple $(u, \alpha, \beta) \in \mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$, the Fréchet derivative $S'(u, \alpha, \beta) \in L(\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$ is precisely the solution map $(h_1, h_2, h_3) \mapsto \delta$ of *the elliptic variational equality*

(3.5)
$$\begin{cases} \delta \in W(y), \\ \langle A'(y)\delta, z\rangle \\ \qquad + \displaystyle\sum_{k\,:\,G_k y \neq 0} \omega_k \Big(\alpha_k \|\cdot\|''(G_k y)(G_k\delta, G_k z) + \beta_k (\|\cdot\|^{1+\gamma})''(G_k y)(G_k \delta, G_k z)\Big) \\ \qquad = \langle B h_1, z\rangle - \displaystyle\sum_{k\,:\,G_k y\neq 0} \omega_k \Big(h_{2,k} \|\cdot\|'(G_k y)(G_k z) + h_{3,k} (\|\cdot\|^{1+\gamma})'(G_k y)(G_k z)\Big) \\ \hfill \forall z \in W(y). \end{cases}$$

Here, $y := S(u,\alpha,\beta)$ *is the solution of* (V) *associated with the triple* (u,α,β)*,* $A'(y)$ *is the Fréchet derivative of* A *in* y*,* $\|\cdot\|'$ *and* $(\|\cdot\|^{1+\gamma})'$ *(respectively,* $\|\cdot\|''$ *and* $(\|\cdot\|^{1+\gamma})''$*) are the first (respectively, second) Fréchet derivatives of the functions* $\|\cdot\|$ *and* $\|\cdot\|^{1+\gamma}$ *away from the origin, and* $W(y)$ *is the subspace of* $\mathbb{R}^n$ *defined by*

$$W(y) := \{z \in \mathbb{R}^n \mid G_k z = 0 \quad \forall\ k = 1, \dots, m \text{ with } G_k y = 0\}. \tag{3.6}$$

Note that by direct calculation, we obtain that the variational equality (3.5) can also be written in the following, more explicit form:

$$\begin{aligned} &\delta \in W(y), \\ &\langle A'(y)\delta, z\rangle + \sum_{k\,:\,G_k y\neq 0} \omega_k \alpha_k \frac{\|G_k y\|^2 \langle G_k \delta, G_k z\rangle - \langle G_k y, G_k\delta\rangle \langle G_k y, G_k z\rangle}{\|G_k y\|^3} \\ &\qquad + \sum_{k\,:\,G_k y\neq 0} \omega_k \beta_k (1+\gamma) \frac{\|G_k y\|^2 \langle G_k \delta, G_k z\rangle - \langle G_k y, G_k\delta\rangle \langle G_k y, G_k z\rangle}{\|G_k y\|^{3-\gamma}} \\ &\qquad + \sum_{k\,:\,G_k y\neq 0} \omega_k \beta_k (\gamma^2+\gamma) \frac{\langle G_k y, G_k\delta\rangle \langle G_k y, G_k z\rangle}{\|G_k y\|^{3-\gamma}} \\ &\qquad = \langle B h_1, z\rangle - \sum_{k\,:\,G_k y\neq 0} \omega_k \left(h_{2,k} \frac{\langle G_k y, G_k z\rangle}{\|G_k y\|} + h_{3,k}(1+\gamma) \frac{\langle G_k y, G_k z\rangle}{\|G_k y\|^{1-\gamma}} \right) \\ &\hfill \forall z \in W(y). \end{aligned}$$

Proof of Theorem 3.3. To prove the different claims in Theorem 3.3, we proceed in several steps as follows:

Step 1 *(Gâteaux differentiability).* We begin by showing that the solution map S associated with (V) is Gâteaux differentiable everywhere in $\mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$. The approach that we use in the following to establish the Gâteaux differentiability is fairly standard and relies heavily on the Lipschitz estimate (3.2) in Proposition 3.2. In this context, compare, e.g., with [22, 39, 45] and also with the more general theory for infinite-dimensional problems in [2, 11, 15].

Suppose that an arbitrary but fixed triple $w := (u,\alpha,\beta) \in \mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$ is given, and let $h := (h_1, h_2, h_3)$ be a vector such that $w + t_0 h \in \mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$ holds for some $t_0 > 0$. Then, the convexity of the set $\mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$ implies that $w + th$ is an element of $\mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$ for all $t \in (0, t_0)$, and we may define

$$\delta_t := \frac{S(w+th) - S(w)}{t} \in \mathbb{R}^n$$

for all $t \in (0, t_0)$. Due to the Lipschitz estimate (3.2), the difference quotients δ_t remain bounded as t tends to zero. This implies in particular that, for every arbitrary but fixed sequence $\{t_j\} \subset (0, t_0)$ with $t_j \searrow 0$, we can find a subsequence (still denoted by the same symbol) such that $\delta_j := \delta_{t_j}$ converges to some $\delta \in \mathbb{R}^n$ for $j \to \infty$. By defining $y := S(w)$ and choosing test vectors of the form $v := y + t_j z$, $z \in \mathbb{R}^n$, in the variational inequality for $S(w + t_j h) = y + t_j \delta_j$, we may now deduce that

$$\begin{aligned}
&\langle A(y + t_j\delta_j), t_j z - t_j \delta_j\rangle \\
&\quad + \sum_{k=1}^m \omega_k \Big((\alpha_k + t_j h_{2,k}) \|G_k y + t_j G_k z\| + (\beta_k + t_j h_{3,k}) \|G_k y + t_j G_k z\|^{1+\gamma} \Big) \\
&\quad - \sum_{k=1}^m \omega_k \Big((\alpha_k + t_j h_{2,k}) \|G_k y + t_j G_k \delta_j\| + (\beta_k + t_j h_{3,k}) \|G_k y + t_j G_k \delta_j\|^{1+\gamma} \Big) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \geq \langle Bu + t_j B h_1, t_j z - t_j \delta_j\rangle
\end{aligned}$$

holds for all $z \in \mathbb{R}^n$. If we divide by t_j^2 and rearrange terms, then the last inequality can be restated as

(3.7)

$$\begin{aligned}
&\left\langle \frac{A(y + t_j\delta_j) - A(y)}{t_j}, z - \delta_j \right\rangle \\
&\quad + \sum_{k=1}^m \omega_k \alpha_k \frac{1}{t_j} \left(\frac{\|G_k y + t_j G_k z\| - \|G_k y\|}{t_j} - \|\cdot\|'(G_k y; G_k z) \right) \\
&\quad + \sum_{k=1}^m \omega_k \beta_k \frac{1}{t_j} \left(\frac{\|G_k y + t_j G_k z\|^{1+\gamma} - \|G_k y\|^{1+\gamma}}{t_j} - (\|\cdot\|^{1+\gamma})'(G_k y; G_k z) \right) \\
&\quad - \sum_{k=1}^m \omega_k \alpha_k \frac{1}{t_j} \left(\frac{\|G_k y + t_j G_k \delta_j\| - \|G_k y\|}{t_j} - \|\cdot\|'(G_k y; G_k \delta_j) \right) \\
&\quad - \sum_{k=1}^m \omega_k \beta_k \frac{1}{t_j} \left(\frac{\|G_k y + t_j G_k \delta_j\|^{1+\gamma} - \|G_k y\|^{1+\gamma}}{t_j} - (\|\cdot\|^{1+\gamma})'(G_k y; G_k \delta_j) \right) \\
&\quad + \frac{1}{t_j} \left(\langle A(y) - Bu, z\rangle + \sum_{k=1}^m \omega_k \Big(\alpha_k \|\cdot\|'(G_k y; G_k z) + \beta_k (\|\cdot\|^{1+\gamma})'(G_k y; G_k z) \Big) \right) \\
&\quad - \frac{1}{t_j} \left(\langle A(y) - Bu, \delta_j\rangle + \sum_{k=1}^m \omega_k \Big(\alpha_k \|\cdot\|'(G_k y; G_k \delta_j) + \beta_k (\|\cdot\|^{1+\gamma})'(G_k y; G_k \delta_j) \Big) \right) \\
&\geq \langle B h_1, z - \delta_j\rangle \\
&\quad + \sum_{k=1}^m \omega_k h_{2,k} \left(\frac{\|G_k y + t_j G_k \delta_j\| - \|G_k y + t_j G_k z\|}{t_j} \right) \\
&\quad + \sum_{k=1}^m \omega_k h_{3,k} \left(\frac{\|G_k y + t_j G_k \delta_j\|^{1+\gamma} - \|G_k y + t_j G_k z\|^{1+\gamma}}{t_j} \right) \quad \forall z \in \mathbb{R}^n.
\end{aligned}$$

Note that we have added several terms here (e.g., the norms $\|G_k y\|$), so that the expressions on the left-hand side of (3.7) take the form of classical difference quotients. An important observation at this point is that all of the terms in the large brackets on the left-hand side of (3.7) are nonnegative (the second-order difference quotients of the functions $\|\cdot\|$ and $\|\cdot\|^{1+\gamma}$ due to convexity and the terms in the last two lines of the left-hand side of (3.7) due to (3.1)). This allows us to deduce the following from (3.7) when we choose z to be zero:

(3.8)
$$
\begin{aligned}
0 &\leq \sum_{k\,:\,G_k y=0} \omega_k \beta_k \frac{\|G_k \delta_j\|^{1+\gamma}}{t_j^{1-\gamma}} \\
&= \sum_{k\,:\,G_k y=0} \omega_k \beta_k \frac{1}{t_j}\left(\frac{\|G_k y + t_j G_k \delta_j\|^{1+\gamma} - \|G_k y\|^{1+\gamma}}{t_j} - (\|\cdot\|^{1+\gamma})'(G_k y; G_k \delta_j)\right) \\
&\leq \left\langle B h_1 - \frac{A(y + t_j \delta_j) - A(y)}{t_j}, \delta_j \right\rangle \\
&\quad - \sum_{k=1}^m \omega_k h_{2,k}\left(\frac{\|G_k y + t_j G_k \delta_j\| - \|G_k y\|}{t_j}\right) \\
&\quad - \sum_{k=1}^m \omega_k h_{3,k}\left(\frac{\|G_k y + t_j G_k \delta_j\|^{1+\gamma} - \|G_k y\|^{1+\gamma}}{t_j}\right).
\end{aligned}
$$

Since the right-hand side of (3.8) remains bounded for $t_j \searrow 0$, since $\gamma \in (0,1)$, and since $\omega_k \beta_k > 0$ for all k, the above implies that the limit δ of the difference quotients δ_j is contained in the set $W(y) = \{z \in \mathbb{R}^n \mid G_k z = 0 \text{ for all } k \text{ with } G_k y = 0\}$. From (3.7), the fact that (3.1) holds with equality for all $z \in W(y)$, and again the information about the signs of the terms in (3.7), we now obtain that δ_j satisfies

(3.9)
$$
\begin{aligned}
&\left\langle \frac{A(y + t_j \delta_j) - A(y)}{t_j}, z - \delta_j \right\rangle \\
&+ \sum_{k\,:\,G_k y \neq 0} \omega_k \alpha_k \frac{1}{t_j}\left(\frac{\|G_k y + t_j G_k z\| - \|G_k y\|}{t_j} - \|\cdot\|'(G_k y)(G_k z)\right) \\
&+ \sum_{k\,:\,G_k y \neq 0} \omega_k \beta_k \frac{1}{t_j}\left(\frac{\|G_k y + t_j G_k z\|^{1+\gamma} - \|G_k y\|^{1+\gamma}}{t_j} - (\|\cdot\|^{1+\gamma})'(G_k y)(G_k z)\right) \\
&- \sum_{k\,:\,G_k y \neq 0} \omega_k \alpha_k \frac{1}{t_j}\left(\frac{\|G_k y + t_j G_k \delta_j\| - \|G_k y\|}{t_j} - \|\cdot\|'(G_k y)(G_k \delta_j)\right) \\
&- \sum_{k\,:\,G_k y \neq 0} \omega_k \beta_k \frac{1}{t_j}\left(\frac{\|G_k y + t_j G_k \delta_j\|^{1+\gamma} - \|G_k y\|^{1+\gamma}}{t_j} - (\|\cdot\|^{1+\gamma})'(G_k y)(G_k \delta_j)\right) \\
&\geq \langle B h_1, z - \delta_j \rangle \\
&+ \sum_{k=1}^m \omega_k h_{2,k}\left(\frac{\|G_k y + t_j G_k \delta_j\| - \|G_k y\|}{t_j} - \frac{\|G_k y + t_j G_k z\| - \|G_k y\|}{t_j}\right) \\
&+ \sum_{k=1}^m \omega_k h_{3,k}\left(\frac{\|G_k y + t_j G_k \delta_j\|^{1+\gamma} - \|G_k y\|^{1+\gamma}}{t_j} - \frac{\|G_k y + t_j G_k z\|^{1+\gamma} - \|G_k y\|^{1+\gamma}}{t_j}\right)
\end{aligned}
$$

for all $z \in W(y)$. Note that all of the expressions in (3.9) are well-behaved for $j \to \infty$ (since the Euclidean norm $\|\cdot\|$ is smooth away from the origin and Hadamard directionally differentiable everywhere, and since A is Fréchet). We may thus pass to

the limit to arrive at the following variational inequality of the second kind:
(3.10)

$$\begin{aligned}
&\langle A'(y)\delta - Bh_1, z-\delta\rangle\\
&\quad + \frac{1}{2}\sum_{k\,:\,G_k y\neq 0}\omega_k\Big(\alpha_k\|\cdot\|''(G_k y)(G_k z, G_k z) + \beta_k(\|\cdot\|^{1+\gamma})''(G_k y)(G_k z, G_k z)\Big)\\
&\quad - \frac{1}{2}\sum_{k\,:\,G_k y\neq 0}\omega_k\Big(\alpha_k\|\cdot\|''(G_k y)(G_k \delta, G_k \delta) + \beta_k(\|\cdot\|^{1+\gamma})''(G_k y)(G_k \delta, G_k \delta)\Big)\\
&\qquad \geq -\sum_{k\,:\,G_k y\neq 0}\omega_k\left(h_{2,k}\frac{\langle G_k y, G_k(z-\delta)\rangle}{\|G_k y\|} + h_{3,k}(1+\gamma)\frac{\langle G_k y, G_k(z-\delta)\rangle}{\|G_k y\|^{1-\gamma}}\right)\\
&\hspace{28em}\forall z\in W(y).
\end{aligned}$$

Since the Fréchet derivative $A'(y) : \mathbb{R}^n \to \mathbb{R}^n$ inherits the strong monotonicity of the original operator A (see [11, Lemma 1.2.3]), the problem (3.10) can have at most one solution $\delta \in W(y)$ (cf. step 3 in the proof of [11, Theorem 1.2.2]), and we may deduce that the limit δ of the difference quotients δ_j is independent of the choice of the (sub)sequence $\{t_j\} \subset (0, t_0)$ that we started with. The latter implies, in combination with classical contradiction arguments, that the whole family of difference quotients $\{\delta_t\}$ converges to the unique solution $\delta \in W(y)$ of (3.10) for $t \searrow 0$, and that the solution operator S associated with (V) is directionally differentiable in the point w in every direction h with $w+t_0h \in \mathbb{R}^n\times[0,\infty)^m\times(0,\infty)^m$ for some $t_0 > 0$. By choosing test vectors of the form $\delta + sz$, $z \in \mathbb{R}^n$, $s > 0$, in (3.10); by dividing by s; by passing to the limit $s \searrow 0$; and by exploiting that $W(y)$ is a subspace, we obtain further that (3.10) can be rewritten as (3.5). Since (3.5) has a linear (and thus also continuous) solution operator $(h_1, h_2, h_3) \mapsto \delta$, it now follows immediately that the map S is Gâteaux differentiable in w and that the derivative $S'(w) \in L(\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$ is characterized by (3.5). This completes the first step of the proof.

Step 2 *(Fréchet differentiability).* The Fréchet differentiability of the solution map S on $\mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$ follows immediately from the Gâteaux differentiability of S, the Lipschitz estimate (3.2), and standard arguments. We include the proof for the convenience of the reader: Suppose that there exists a $w \in \mathbb{R}^n\times[0,\infty)^m\times(0,\infty)^m$ such that S is not Fréchet differentiable in w in the sense of (3.4). Then, there exist an $\varepsilon > 0$ and sequences $\{h_j\} \subset \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$, $\{t_j\} \subset (0,\infty)$ such that $t_j \searrow 0$ for $j \to \infty$, $\|h_j\| = 1$ for all j, $w + t_jh_j \in \mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$ for all j, and

$$\|S(w + t_jh_j) - S(w) - t_jS'(w)h_j\| \geq \varepsilon t_j \quad \forall j.$$

Since $\|h_j\| = 1$, we may assume w.l.o.g. that $h_j \to h$ holds for some $h \in \mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^m$, and from the properties of the sequence $\{h_j\}$ (or the signs of the components of h_j, to be more precise), it follows straightforwardly that there has to be an $s > 0$ with $w + sh \in \mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$. The local Lipschitz continuity of S now yields

$$\varepsilon \leq \frac{\|S(w + t_jh_j) - S(w) - t_jS'(w)h_j\|}{t_j} = \frac{\|S(w + t_jh) - S(w) - t_jS'(w)h\|}{t_j} + o(1),$$

where the Landau symbol refers to the limit $j \to \infty$. This contradicts the Gâteaux differentiability that we established in Step 1 of the proof. Thus, S is Fréchet differentiable, and the second step of the proof is complete.

Step 3 *(continuity of the Fréchet derivative).* It remains to prove that the map $S' : \mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m \to L(\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$ is continuous. To this end, we consider an arbitrary but fixed sequence $\{w_j\} \subset \mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$, which satisfies $w_j = (u_j, \alpha_j, \beta_j) \to w = (u, \alpha, \beta)$ for some $w \in \mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$. Note that to prove the continuity of S', it suffices to show that $S'(w_j)h \to S'(w)h$ holds for all $h \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$ (since this convergence already implies $S'(w_j) \to S'(w)$ in the operator norm). So let us assume that an $h = (h_1, h_2, h_3) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$ is given. Then, (3.5) yields that the vectors $\eta_j := S'(w_j)h$ satisfy

$$\eta_j \in W(y_j) = \{z \in \mathbb{R}^n \mid G_k z = 0 \quad \forall\, k = 1, \ldots, m \text{ with } G_k y_j = 0\}$$

and

$$\begin{aligned}
&\langle A'(y_j)\eta_j, \eta_j\rangle \\
&\quad + \sum_{k\,:\, G_k y_j \neq 0} \omega_k \alpha_{j,k} \frac{\|G_k y_j\|^2 \|G_k \eta_j\|^2 - \langle G_k y_j, G_k \eta_j\rangle^2}{\|G_k y_j\|^3} \\
&\quad + \sum_{k\,:\, G_k y_j \neq 0} \omega_k \beta_{j,k} (1+\gamma) \frac{\|G_k y_j\|^2 \|G_k \eta_j\|^2 - (1-\gamma)\langle G_k y_j, G_k \eta_j\rangle^2}{\|G_k y_j\|^{3-\gamma}} \\
&\qquad = \langle Bh_1, \eta_j\rangle - \sum_{k\,:\, G_k y_j \neq 0} \omega_k \left(h_{2,k} \frac{\langle G_k y_j, G_k \eta_j\rangle}{\|G_k y_j\|} + h_{3,k}(1+\gamma) \frac{\langle G_k y_j, G_k \eta_j\rangle}{\|G_k y_j\|^{1-\gamma}} \right).
\end{aligned} \tag{3.11}$$

Here, y_j is short for $S(w_j)$. From the strong monotonicity of $A'(y_j)$ (which is uniform in j by our assumptions and [11, Lemma 1.2.3]), the Cauchy–Schwarz inequality, the nonnegativity of α, β, and ω, $\gamma \in (0,1)$, and the boundedness of the sequence $y_j = S(w_j)$, we obtain that there exist constants $c, C > 0$ independent of j with

$$\begin{aligned}
\langle A'(y_j)\eta_j, \eta_j\rangle &+ \sum_{k\,:\, G_k y_j \neq 0} \omega_k \alpha_{j,k} \frac{\|G_k y_j\|^2 \|G_k \eta_j\|^2 - \langle G_k y_j, G_k \eta_j\rangle^2}{\|G_k y_j\|^3} \\
&+ \sum_{k\,:\, G_k y_j \neq 0} \omega_k \beta_{j,k} (1+\gamma) \frac{\|G_k y_j\|^2 \|G_k \eta_j\|^2 - (1-\gamma)\langle G_k y_j, G_k \eta_j\rangle^2}{\|G_k y_j\|^{3-\gamma}} \\
&\geq c\|\eta_j\|^2 + \sum_{k\,:\, G_k y_j \neq 0} \omega_k \beta_{j,k} (\gamma + \gamma^2) \frac{\|G_k \eta_j\|^2}{\|G_k y_j\|^{1-\gamma}}
\end{aligned}$$

and

$$\begin{aligned}
\langle Bh_1, \eta_j\rangle &- \sum_{k\,:\, G_k y_j \neq 0} \omega_k \left(h_{2,k} \frac{\langle G_k y_j, G_k \eta_j\rangle}{\|G_k y_j\|} + h_{3,k}(1+\gamma) \frac{\langle G_k y_j, G_k \eta_j\rangle}{\|G_k y_j\|^{1-\gamma}} \right) \\
&\leq \|Bh_1\|\|\eta_j\| + \sum_{k\,:\, G_k y_j \neq 0} \omega_k \Big(|h_{2,k}|\|G_k \eta_j\| + |h_{3,k}|(1+\gamma)\|G_k y_j\|^{\gamma} \|G_k \eta_j\| \Big) \\
&\leq C\|\eta_j\|.
\end{aligned}$$

Using the above in (3.11) gives

$$c\|\eta_j\|^2 + \sum_{k\,:\, G_k y_j \neq 0} \omega_k \beta_{j,k} (\gamma + \gamma^2) \frac{\|G_k \eta_j\|^2}{\|G_k y_j\|^{1-\gamma}} \leq C\|\eta_j\|$$

and, by Young's inequality,

$$\frac{c}{2}\|\eta_j\|^2 + \sum_{k\,:\,G_k y_j \neq 0} \omega_k \beta_{j,k}(\gamma+\gamma^2)\frac{\|G_k\eta_j\|^2}{\|G_k y_j\|^{1-\gamma}} \le \frac{C^2}{2c}. \tag{3.12}$$

The sequence $\{\eta_j\}$ is thus bounded, and we may pass over to a subsequence (still denoted by the same symbol) with $\eta_j \to \eta$ for some η. We claim that this η satisfies $\eta \in W(y)$, where $y := S(w)$ is the solution associated with the limit point w. To see this, we consider an arbitrary but fixed $k \in \{1, \dots, m\}$ with $G_k y = 0$ and distinguish between two cases: If we can find a subsequence j_i such that $G_k y_{j_i} = 0$ holds for all i, then we trivially have $G_k \eta_{j_i} = 0$ for all i (since $\eta_{j_i} \in W(y_{j_i})$), and we immediately obtain from the convergence $\eta_{j_i} \to \eta$ that $G_k\eta = 0$. If, on the other hand, we can find a subsequence j_i such that $G_k y_{j_i} \neq 0$ holds for all i, then (3.12) yields

$$0 \le \omega_k \beta_{j_i,k}(\gamma+\gamma^2)\|G_k\eta_{j_i}\|^2 \le \frac{C^2}{2c}\|G_k y_{j_i}\|^{1-\gamma},$$

and we may use the convergences $y_{j_i} \to y$ and $\beta_{j_i} \to \beta \in (0,\infty)^m$ to conclude that $G_k\eta = 0$. This shows that $G_k\eta = 0$ holds for all k with $G_k y = 0$ and that η is indeed an element of $W(y)$. Suppose now that j is so large that $G_k y_j \neq 0$ holds for all $k \in \{1, \dots, m\}$ with $G_k y \neq 0$ (this is the case for all large enough j due to the convergence $y_j \to y$). Then, it clearly holds that $W(y) \subset W(y_j)$, and we may deduce the following from the variational equality (3.5) for η_j:

$$\begin{aligned}
&\langle A'(y_j)\eta_j, z\rangle + \sum_{k\,:\,G_k y\neq 0} \omega_k \alpha_{j,k} \frac{\|G_k y_j\|^2 \langle G_k\eta_j, G_k z\rangle - \langle G_k y_j, G_k \eta_j\rangle \langle G_k y_j, G_k z\rangle}{\|G_k y_j\|^3} \\
&\quad + \sum_{k\,:\,G_k y\neq 0} \omega_k \beta_{j,k}(1+\gamma) \frac{\|G_k y_j\|^2 \langle G_k\eta_j, G_k z\rangle - \langle G_k y_j, G_k \eta_j\rangle \langle G_k y_j, G_k z\rangle}{\|G_k y_j\|^{3-\gamma}} \\
&\quad + \sum_{k\,:\,G_k y\neq 0} \omega_k \beta_{j,k}(\gamma^2+\gamma) \frac{\langle G_k y_j, G_k \eta_j\rangle \langle G_k y_j, G_k z\rangle}{\|G_k y_j\|^{3-\gamma}} \\
&\quad = \langle B h_1, z\rangle - \sum_{k\,:\,G_k y\neq 0} \omega_k \left(h_{2,k}\frac{\langle G_k y_j, G_k z\rangle}{\|G_k y_j\|} + h_{3,k}(1+\gamma)\frac{\langle G_k y_j, G_k z\rangle}{\|G_k y_j\|^{1-\gamma}}\right) \quad \forall z \in W(y).
\end{aligned}$$

If we pass to the limit $j \to \infty$ in the above, then it follows that η solves the variational problem (3.5), which characterizes $S'(w)h$. This shows that $\eta = S'(w)h$ has to hold and that $S'(w_j)h$ converges to $S'(w)h$ for $j \to \infty$. Using the same arguments as in Step 1 of the proof, we obtain that this convergence also holds for the whole original sequence $S'(w_j)h$ and not just for the subsequence that we chose after (3.12). This proves the continuity of S' and completes the proof. ∎

Some remarks are in order regarding Theorem 3.3.

Remark 3.4.

1. It seems to be a common belief that minimization problems and elliptic variational inequalities which involve nondifferentiable terms necessarily also have nondifferentiable solution operators. Theorem 3.3 shows that there is, in fact, no such mechanism, and that it is perfectly possible that the solution map of a nonsmooth problem is a C^1-function. In the fields of bilevel optimization and optimal control, this observation is, of course, very valuable.

2. We would like to point out that the solution map S associated with (V) is typically not Fréchet differentiable in points (u, α, β) with $\beta_k = 0$ for some k. Theorem 3.3 thus no longer holds in general when we replace the set $\mathbb{R}^n \times [0, \infty)^m \times (0, \infty)^m$ with $\mathbb{R}^n \times [0, \infty)^m \times [0, \infty)^m$. Similarly, we can no longer expect Fréchet differentiability when the exponent γ is equal to zero or one. The fact that S is not Fréchet differentiable for $\gamma = 0$ and $\gamma = 1$ but for all γ between these values is quite counterintuitive. Note that, in all of the above cases, the solution operator is still Hadamard directionally differentiable in the sense of [6, Definition 2.45], as one may easily check using the same arguments as in Step 1 of the proof of Theorem 3.3.
3. What we observe in Theorem 3.3 can be interpreted as a nonstandard regularization effect. Consider, for instance, the simple model problem

$$\begin{cases} \min \|y - y_D\|^2 + \dfrac{\mu}{2}\left(\|u\|^2 + \alpha^2\right) \\ \text{s.t.}\ \ y \in \mathbb{R}^n,\ u \in \mathbb{R}^n,\ \alpha \in [0, \infty), \\ \qquad y = \operatorname*{arg\,min}_{v \in \mathbb{R}^n} \dfrac{1}{2}\|v - u\|^2 + \alpha\|v\|_1, \end{cases} \tag{3.13}$$

where $y_D \in \mathbb{R}^n$ and $\mu > 0$ are given, and where $\|\cdot\|_1$ denotes the 1-norm on the Euclidean space. Then, it is easy to check that the solution operator $S : \mathbb{R}^n \times [0, \infty) \to \mathbb{R}^n$, $(u, \alpha) \mapsto y$, associated with the inner minimization problem in (3.13) is nonsmooth. In fact, in the special case $n = 1$, we can derive the following closed formula for the solution map:

$$S(u, \alpha) = \begin{cases} u + \alpha & \text{if } u \le -\alpha, \\ 0 & \text{if } u \in (-\alpha, \alpha), \\ u - \alpha & \text{if } u \ge \alpha. \end{cases} \tag{3.14}$$

Suppose now that we modify (3.13) by adding a term of the form $\varepsilon\|v\|_p^p$ for some $p \in (1, 2)$ in the lower-level problem, where $\varepsilon > 0$ is an arbitrary but fixed small number and where $\|\cdot\|_p$ denotes the p-norm on $\mathbb{R}^n$. Then, the resulting bilevel minimization problem

$$\begin{cases} \min \|y - y_D\|^2 + \dfrac{\mu}{2}\left(\|u\|^2 + \alpha^2\right) \\ \text{s.t.}\ \ y \in \mathbb{R}^n,\ u \in \mathbb{R}^n,\ \alpha \in [0, \infty), \\ \qquad y = \operatorname*{arg\,min}_{v \in \mathbb{R}^n} \dfrac{1}{2}\|v - u\|^2 + \alpha\|v\|_1 + \varepsilon\|v\|_p^p \end{cases} \tag{3.15}$$

can also be written in the form (P) (cf. the second example in section 2), and we obtain from Theorem 3.3 that the solution operator $S_\varepsilon : \mathbb{R}^n \times [0, \infty) \to \mathbb{R}^n$, $(u, \alpha) \mapsto y$, associated with the lower level of (3.15) is continuously Fréchet differentiable. By adding the term $\varepsilon\|\cdot\|_p^p$, we have thus indeed regularized the original problem (3.13). What is appealing about the above method of regularization is that it is "minimally invasive." It produces an approximate problem whose reduced objective function possesses C^1-regularity (and is thus amenable to gradient-based solution algorithms; see section 4) while preserving the nonsmooth features and, e.g., the sparsity-promoting nature on the lower level. Note that the addition of the term $\varepsilon\|\cdot\|_p^p$ in particular does not change the subdifferential at zero of the objective function of the inner problem in (3.13). We would like to emphasize at this point that the lower-level problems in (3.13) and (3.15) can be solved easily with various standard algorithms (e.g., semismooth Newton, subgradient, or bundle methods). The major difficulty in (3.13) is handling the nonsmoothness of the solution map $S : (u, \alpha) \mapsto y$ on the upper level. In view of these facts, the regularization effect observed above is the best that we can hope for: By adding the term $\varepsilon \| \cdot \|_p^p$, we regularize the solution operator of the inner problem in (3.13) without regularizing the nondifferentiable terms in the inner problem. This removes the nonsmoothness where it is problematic (in the solution map) while preserving it where it can be handled (in the lower-level problem). To the best of the author's knowledge, similar effects have not been documented so far in the literature (where primarily Huber-type regularizations are used, which do not preserve sparsity-promoting effects; see [9, 16, 23, 44, 56]).

4. In the context of the general sensitivity analysis for elliptic variational inequalities of the first and second kinds developed in [2, 11, 59], the differentiability result in Theorem 3.3 can be explained as follows: The singular curvature properties of the terms $\|G_k(\cdot)\|^{1+\gamma}$ at the origin enforce the fact that the second subderivative of the nonsmooth functional in (V) is generated by a symmetric bilinear form defined on a subspace of $\mathbb{R}^n$ (namely, the space $W(y)$ in (3.6)). This, in combination with the second-order epidifferentiability of the involved terms, yields the Fréchet differentiability of the solution operator to (V). For details on this topic and the underlying theory, we refer the reader to [11, Chapters 1 and 4, Theorem 1.4.1, Corollary 1.4.4].
5. The regularization effect in Theorem 3.3 can also be exploited in the infinite-dimensional setting (see, for instance, [11, section 4.3.3] for a simple example). However, in infinite dimensions, one typically requires additional Lipschitz continuity/compactness properties to establish the directional differentiability of the solution map S, and the analysis becomes much more involved (cf. the approach in [13], where superposition operators are considered). In particular, it does not seem to be possible to derive a "general purpose" result analogous to Theorem 3.3 for elliptic variational inequalities in arbitrary Hilbert spaces.
6. Results analogous to Theorem 3.3 can also be obtained for problems which involve nonsmooth functions whose properties are similar to those of the Euclidean norm (e.g., the maximum function $\max(0, \cdot)$).
7. It should be noted that the variational problem (3.5) that characterizes the operators $S'(w)$ arises from the original variational inequality (V) by termwise differentiation (where, at the origin, the missing second derivative of the Euclidean norm is replaced with the conditions in (3.6)). In this context, compare also with the alternative formulation (3.10). An analogous behavior can be observed for smooth problems; cf., e.g., the results in [6].
8. To the best of the author's knowledge, it is currently unknown whether effects similar to those observed in Theorem 3.3 are also present when the sensitivities of the solution y of (P) w.r.t. perturbations of the matrices G_k are considered. (A differentiability result for the map $\{G_k\}_{k=1}^m \mapsto y$ would be interesting, e.g., in the context of the application in subsection 2.2.)

4. Consequences for the applicability of gradient-based algorithms. The consequences of Theorem 3.3 on the analysis and the numerical solution of the bilevel optimization and optimal control problems in sections 1 and 2 are obvious: Since the solution operator $S : (u, \alpha, \beta) \mapsto y$ associated with the elliptic variational inequality (V) is continuously Fréchet differentiable on $\mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$, we can tackle every problem of type (P) that satisfies the conditions in Assumption 1.1 with standard gradient-based algorithms. Depending on the precise nature of the problem at hand, possible choices could be, for instance, trust-region methods (see [12, 19, 54]), (projected) gradient methods (see [4, 33, 43]), or nonlinear conjugated gradient methods; see [33, 54, 57]. For a tangible example of a solution algorithm, we refer the reader to section 5. Note that our standing assumption $U_{ad} \subset \mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$ is indispensable at this point since Theorem 3.3 does not yield any information about S on the cone $\{(u,\alpha,\beta) \in \mathbb{R}^n \times [0,\infty)^m \times [0,\infty)^m \mid \beta_k = 0 \text{ for some } k\}$. (The author suspects that it is still possible to identify subgradients on this critical set by exploiting (3.1); cf. the analysis in [14, 58].) Further, it should be noted that all of the above-mentioned algorithms require evaluations of the derivative of the reduced objective function $F(u,\alpha,\beta) := J(S(u,\alpha,\beta),u,\alpha,\beta)$ associated with the problem (P). To calculate the gradients of F efficiently, we can use an adjoint calculus, as the following theorem shows.

THEOREM 4.1 (calculation of gradients). *In the situation of Assumption 1.1, the Fréchet derivative $F'(u,\alpha,\beta) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$ of the reduced objective function $F(u,\alpha,\beta) := J(S(u,\alpha,\beta),u,\alpha,\beta)$ in a point $(u,\alpha,\beta) \in \mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$ is given by*

$$F'(u,\alpha,\beta) = \Big((\partial_u J)(y,u,\alpha,\beta) + B^* p_1, (\partial_\alpha J)(y,u,\alpha,\beta) + p_2, (\partial_\beta J)(y,u,\alpha,\beta) + p_3\Big).$$

Here, y is short for $S(u,\alpha,\beta)$; $p_1 = p_1(u,\alpha,\beta) \in \mathbb{R}^n$ is the unique solution of

$$\begin{cases} p_1 \in W(y), \\ \langle A'(y)^* p_1, z\rangle \\ \qquad + \sum\limits_{k\,:\,G_k y \neq 0} \omega_k \Big(\alpha_k \|\cdot\|''(G_k y)(G_k p_1, G_k z) + \beta_k (\|\cdot\|^{1+\gamma})''(G_k y)(G_k p_1, G_k z)\Big) \\ \qquad = \langle (\partial_y J)(y,u,\alpha,\beta), z\rangle \qquad\qquad \forall z \in W(y); \end{cases} \tag{4.1}$$

the vectors $p_2, p_3 \in \mathbb{R}^m$ are defined by

$$(p_2)_k := \begin{cases} -\omega_k \dfrac{\langle G_k y, G_k p_1\rangle}{\|G_k y\|} & \text{if } G_k y \neq 0, \\ 0 & \text{else,} \end{cases} \qquad k = 1,\dots,m,$$

and

$$(p_3)_k := \begin{cases} -(1+\gamma)\omega_k \dfrac{\langle G_k y, G_k p_1\rangle}{\|G_k y\|^{1-\gamma}} & \text{if } G_k y \neq 0, \\ 0 & \text{else,} \end{cases} \qquad k = 1,\dots,m;$$

$W(y)$ denotes the space in (3.6)*; B^* and $A'(y)^*$ are the adjoints of B and $A'(y)$ (w.r.t. the Euclidean scalar product); and $\partial_y J$, $\partial_u J$, $\partial_\alpha J$, and $\partial_\beta J$ denote the partial derivatives of the function J w.r.t. the first, second, third, and fourth arguments, respectively.*

Proof. From the chain rule (see [6, Proposition 2.47]), it follows straightforwardly that for every point $(u,\alpha,\beta) \in \mathbb{R}^n \times [0,\infty)^m \times (0,\infty)^m$ and every $h = (h_1,h_2,h_3) \in$

$\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$, we have

$$\langle F'(u,\alpha,\beta), h\rangle = \langle (\partial_y J)(y,u,\alpha,\beta), S'(u,\alpha,\beta)h\rangle + \langle (\partial_u J)(y,u,\alpha,\beta), h_1\rangle \\ + \langle (\partial_\alpha J)(y,u,\alpha,\beta), h_2\rangle + \langle (\partial_\beta J)(y,u,\alpha,\beta), h_3\rangle\,.$$

Further, we obtain from the variational equality (3.5) for $\delta := S'(u,\alpha,\beta)h \in W(y)$ and the definitions of the vectors p_1, p_2, and p_3 that

$$\begin{aligned}
&\langle (\partial_y J)(y,u,\alpha,\beta), \delta\rangle \\
&\quad = \langle A'(y)^* p_1, \delta\rangle \\
&\qquad + \sum_{k\,:\,G_k y\neq 0} \omega_k \Big(\alpha_k \|\cdot\|''(G_k y)(G_k p_1, G_k\delta) + \beta_k (\|\cdot\|^{1+\gamma})''(G_k y)(G_k p_1, G_k \delta)\Big) \\
&\quad = \langle B h_1, p_1\rangle - \sum_{k\,:\,G_k y\neq 0} \omega_k \Big(h_{2,k}\|\cdot\|'(G_k y)(G_k p_1) + h_{3,k}(\|\cdot\|^{1+\gamma})'(G_k y)(G_k p_1)\Big) \\
&\quad = \langle B^* p_1, h_1\rangle + \langle p_2, h_2\rangle + \langle p_3, h_3\rangle\,.
\end{aligned}$$

The claim now follows immediately. ◻

Note that every evaluation of the derivative F' requires the solution of the non-smooth elliptic variational inequality (V) (since we need y). This, however, is not a major problem. As we have already mentioned in Remark 3.4, the inequality (V) is comparatively well-behaved and can be tackled with various standard algorithms (especially when it can be identified with a minimization problem of the type (1.1)). In this context, compare also with the approach in section 5.

5. An example of a solution algorithm and a numerical experiment.

In what follows, we demonstrate by means of a tangible example that the results in Theorems 3.3 and 4.1 indeed allow us to solve problems of the type (P) with standard gradient-based algorithms. As a model problem, we consider a special instance of the optimal control problem for Casson fluids that we have derived in subsection 2.1, namely,

(5.1)
$$\begin{cases}
\min \dfrac{1}{2}\langle B(y-y_D), y-y_D\rangle + \dfrac{\mu}{2}\left(\langle Bu,u\rangle + \alpha^2\right) \\
\text{s.t.}\ \ y,u\in\mathbb{R}^n,\ \ \alpha\in[\kappa,\infty), \\
\qquad y = \underset{v\in\mathbb{R}^n}{\arg\min}\, \dfrac{1}{2}\langle Av, v\rangle + \displaystyle\sum_{k=1}^{m} |T_k| \left(\alpha|G_k v| + \frac{4}{3}\alpha^{1/2}|G_k v|^{3/2}\right) - \langle Bu, v\rangle\,.
\end{cases}$$

Here, we have chosen the dimension d to be one (so that the Euclidean norms in (2.1) are just absolute value functions), $\kappa \in (0,\infty)$ is a given constant (a lower bound for the Oldroyd number α), and the quantities A, B, etc. are defined as in subsection 2.1. To solve the problem (5.1), we will employ a standard gradient projection method in the spirit of [4, 33, 43]; see Algorithm 5.4 below. We would like to emphasize that the subsequent analysis should be understood as a feasibility study. With Theorems 3.3 and 4.1 at our disposal, at this point we could also consider more complicated problems and more sophisticated algorithms (e.g., nonlinear conjugated gradient or trust-region methods). To avoid overloading this paper, we leave a detailed discussion of the various possible applications of Theorems 3.3 and 4.1 (e.g., in the fields of parameter learning and identification; see subsection 2.2) for future research. For the sake of simplicity, we further leave aside the influence of numerical errors in the solution of the lower-level problem in (5.1) and other sources of inexactness in the convergence analysis of this section. Strategies for the incorporation of such effects can be found, e.g., in [17, 18, 31, 36, 62].

Before we state the algorithm that we use for the solution of the optimization problem (5.1), we prove some auxiliary results.

LEMMA 5.1 (multiplier system for the lower-level problem in (5.1)). *A vector $y \in \mathbb{R}^n$ solves the lower-level problem in* (5.1) *for a given tuple $(u, \alpha) \in \mathbb{R}^n \times (0, \infty)$ if and only if there exist multipliers $\lambda_1, \ldots, \lambda_m, \eta_1, \ldots, \eta_m \in \mathbb{R}$ such that*

$$\begin{cases} Ay + \sum_{k=1}^{m} |T_k| \left(\alpha G_k^* \lambda_k + 2\alpha^{1/2} G_k^* \eta_k \right) - Bu = 0, & \\ \max \left(\lambda_k^2 - 1, |G_k y| - (G_k y)\lambda_k \right) = 0, & k = 1, \ldots, m, \\ \max \left(\eta_k^2 - |G_k y|, |G_k y|^{3/2} - (G_k y)\eta_k \right) = 0, & k = 1, \ldots, m. \end{cases} \tag{5.2}$$

Proof. From standard calculus rules for the convex subdifferential (see, e.g., [32]), we obtain that a vector $y \in \mathbb{R}^n$ solves the lower-level problem in (5.1) if and only if there exist $\lambda_1, \ldots, \lambda_m, \eta_1, \ldots, \eta_m \in \mathbb{R}$ with

$$\begin{cases} Ay + \sum_{k=1}^{m} |T_k| \left(\alpha G_k^* \lambda_k + 2\alpha^{1/2} G_k^* \eta_k \right) - Bu = 0, & \\ \lambda_k \in \partial |\cdot|(G_k y) & \forall k = 1, \ldots, m, \\ \eta_k \in \partial \left(\frac{2}{3} |\cdot|^{3/2} \right) (G_k y) & \forall k = 1, \ldots, m. \end{cases} \tag{5.3}$$

If we plug in explicit formulas for the convex subdifferentials of the functions $|\cdot|$ and $|\cdot|^{3/2}$ in the above, then it follows straightforwardly that y and the multipliers λ_k and η_k satisfy the system (5.2). This proves the first implication. If, conversely, we start with the system (5.2), then the conditions on λ_k yield $|\lambda_k| \le 1$ and

$$0 \ge |G_k y| - (G_k y)\lambda_k \ge |G_k y||\lambda_k| - (G_k y)\lambda_k \ge 0$$

for all $k = 1, \ldots, m$. The above entails

$$|\lambda_k| \le 1 \quad \forall\, k \qquad \text{and} \qquad \lambda_k = \operatorname{sgn}(G_k y) \quad \forall\, k \text{ with } G_k y \neq 0,$$

which is equivalent to $\lambda_k \in \partial |\cdot|(G_k y)$ for all k. For the multipliers η_k, we obtain along the same lines that $|\eta_k| \le |G_k y|^{1/2}$ and that

$$0 \ge |G_k y|^{3/2} - (G_k y)\eta_k \ge |G_k y||\eta_k| - (G_k y)\eta_k \ge 0$$

holds for all $k = 1, \ldots, m$. This yields

$$\eta_k = \begin{cases} 0 & \forall\, k \text{ with } G_k y = 0, \\ \dfrac{G_k y}{|G_k y|^{1/2}} & \forall\, k \text{ with } G_k y \neq 0 \end{cases}$$

and, as a consequence, $\eta_k \in \partial \left(\frac{2}{3} |\cdot|^{3/2} \right) (G_k y)$ for all $k = 1, \ldots, m$. This shows that (5.2) is equivalent to (5.3) and completes the proof. ∎

Note that the system (5.2) is amenable to numerical solution by a semismooth Newton method. This will be exploited in Algorithm 5.4 below. To compute the gradients of the reduced objective function associated with (5.1), we formulate the following corollary of Theorem 4.1.

LEMMA 5.2 (calculation of gradients). *Denote the solution operator of the lower-level problem in* (5.1) *and the reduced objective function associated with* (5.1) *with* S *and* F, *respectively, i.e.,* $S : \mathbb{R}^n \times (0, \infty) \to \mathbb{R}^n$, $(u, \alpha) \mapsto y$, *and*

$$F : \mathbb{R}^n \times (0, \infty) \to \mathbb{R}, \qquad (u, \alpha) \mapsto \frac{1}{2} \langle B(y - y_D), y - y_D\rangle + \frac{\mu}{2} \left(\langle Bu, u\rangle + \alpha^2\right).$$

Then, the gradient $F'(u, \alpha) \in \mathbb{R}^n \times \mathbb{R}$ *at a point* $(u, \alpha) \in \mathbb{R}^n \times (0, \infty)$ *with associated state* $y := S(u, \alpha)$ *is given by*

$$F'(u,\alpha) = \left(B(\mu u + p), \mu\alpha - \sum_{k\,:\,G_k y\neq 0} |T_k| \left(\frac{(G_k y)(G_k p)}{|G_k y|} + \frac{(G_k y)(G_k p)}{\alpha^{1/2}|G_k y|^{1/2}}\right)\right). \tag{5.4}$$

Here, $p \in \mathbb{R}^n$ *is the unique solution of the variational equality*

$$\begin{cases} p \in W(y), \\ \langle Ap, z\rangle + \displaystyle\sum_{k\,:\,G_k y\neq 0} |T_k|\alpha^{1/2} \frac{(G_k p)(G_k z)}{|G_k y|^{1/2}} = \langle B(y - y_D), z\rangle \qquad \forall z \in W(y), \end{cases} \tag{5.5}$$

with $W(y) := \{z \in \mathbb{R}^n \mid G_k z = 0$ *for all* $k = 1, \ldots, m$ *with* $G_k y = 0\}$.

Proof. The claim follows immediately from Theorem 4.1, the self-adjointness of the operators A and B, and the chain rule in [6, Proposition 2.47]. ∎

Finally, we observe the following.

LEMMA 5.3. *The discretized optimal control problem* (5.1) *admits at least one solution* $(u^*, \alpha^*) \in \mathbb{R}^n \times [\kappa, \infty)$. *Further, every solution* (u^*, α^*) *of* (5.1) *satisfies* $\zeta(u^*, \alpha^*) = 0$, *where* $\zeta : \mathbb{R}^n \times [\kappa, \infty) \to [0, \infty)$ *is the function defined by*

$$\zeta(u,\alpha) := \begin{cases} \|F'(u,\alpha)\| & \text{if } \alpha > \kappa, \\ \left\|\left((\partial_u F)(u,\alpha), \min(0, (\partial_\alpha F)(u,\alpha))\right)\right\| & \text{if } \alpha = \kappa. \end{cases} \tag{5.6}$$

Proof. To show that the problem (5.1) admits a solution, we can use exactly the same argumentation as in the proof of Theorem 2.1. It remains to prove that $\zeta(u, \alpha) = 0$ is a necessary optimality condition. This, however, follows immediately from the equivalence

$$\zeta(u,\alpha) = 0 \qquad \Longleftrightarrow \qquad -F'(u,\alpha) \in \mathcal{N}_{\mathbb{R}^n\times[\kappa,\infty)}(u,\alpha),$$

for all $(u, \alpha) \in \mathbb{R}^n \times [\kappa, \infty)$, where $\mathcal{N}_{\mathbb{R}^n\times[\kappa,\infty)}(u, \alpha)$ denotes the normal cone to the set $\mathbb{R}^n \times [\kappa, \infty)$ at (u, α). ∎

We are now in the position to state the algorithm that we use for the solution of the problem (5.1).

Algorithm 5.4 (gradient projection method for the solution of (5.1)).

Choose an initial guess $(u_0, \alpha_0) \in \mathbb{R}^n \times [\kappa, \infty)$ and parameters $\sigma > 0$, $\nu, \theta \in (0, 1)$.
for $i = 0, 1, 2, 3, \ldots$ **do**
  Calculate $y_i := S(u_i, \alpha_i)$ by solving (5.2) with a semismooth Newton method.
  Calculate $F_i := F(u_i, \alpha_i)$ and solve (5.5) for $p_i := p(u_i, \alpha_i)$.
  Use the relation (5.4) to assemble the gradient $F'(u_i, \alpha_i)$.
  Calculate the stationarity measure $\zeta_i := \zeta(u_i, \alpha_i)$ (with ζ as in (5.6)).
  if $\zeta_i = 0$ **then**
    break
  end if
  Define $g_i := (\partial_u F)(u_i, \alpha_i)/\zeta_i$ and $h_i := (\partial_\alpha F)(u_i, \alpha_i)/\zeta_i$.
  Initialize $\sigma_0 := \sigma$ and calculate a step size as follows:
  for $j = 0, 1, 2, 3, \ldots$ **do**
    Use (5.2) to calculate the quantity

$$\begin{aligned} e_j &:= F_i - F(u_i - \sigma_j g_i, \max(\kappa, \alpha_i - \sigma_j h_i)) \\ &\quad - \theta \sigma_j \zeta_i \left(\|g_i\|^2 + \min(0, h_i)^2 + \min\left(\max(0, h_i)^2, \max(0, h_i) \frac{(\alpha_i - \kappa)}{\sigma_j} \right) \right). \end{aligned}$$

    if $e_j < 0$ **then**
      Define $\sigma_{j+1} := \nu \sigma_j$.
    else
      Define $\tau_i := \sigma_j$ and **break**.
    end if
  end for
  Define $u_{i+1} := u_i - \tau_i g_i$ and $\alpha_{i+1} := \max(\kappa, \alpha_i - \tau_i h_i)$.
end for

We remark that, in combination with smoothing approaches, gradient projection methods similar to the one above have already been used for the solution of bilevel optimization problems in [37, 38, 40, 46, 66]. In this context, compare also with the algorithms in [63, 65], and with [34] where heuristic gradient descent methods and line-search strategies for nonsmooth problems are considered.

To see that Algorithm 5.4 is sensible, we note the following.

LEMMA 5.5. *For every arbitrary but fixed tuple* $(u_i, \alpha_i) \in \mathbb{R}^n \times [\kappa, \infty)$ *that satisfies* $\zeta(u_i, \alpha_i) \neq 0$ *and for every choice of parameters* $\sigma > 0$ *and* $\nu, \theta \in (0, 1)$, *the Armijo-type line-search in Algorithm* 5.4 *(i.e., the inner for-loop with index* j*) terminates after finitely many steps.*

Proof. From the Fréchet differentiability of the reduced objective function F, the definitions $\zeta_i := \zeta(u_i, \alpha_i)$, $g_i := (\partial_u F)(u_i, \alpha_i)/\zeta_i$, and $h_i := (\partial_\alpha F)(u_i, \alpha_i)/\zeta_i$, and simple distinctions of cases, it follows straightforwardly that

$$
\begin{aligned}
&F(u_i,\alpha_i) - F(u_i - sg_i, \max(\kappa, \alpha_i - sh_i)) \\
&\quad = s\zeta_i\|g_i\|^2 + \zeta_i h_i\Big(\alpha_i - \max(\kappa, \alpha_i - sh_i)\Big) + o(s) \\
&\quad = \begin{cases} s\zeta_i\|g_i\|^2 + s\zeta_i h_i^2 + o(s) & \text{if } \alpha_i - sh_i \geq \kappa, \\ s\zeta_i\|g_i\|^2 + \zeta_i h_i(\alpha_i - \kappa) + o(s) & \text{if } \alpha_i - sh_i < \kappa \end{cases} \\
&\quad = \begin{cases} s\zeta_i\|g_i\|^2 + s\zeta_i \min(0,h_i)^2 + s\zeta_i \max(0,h_i)^2 + o(s) & \text{if } \alpha_i - sh_i \geq \kappa, \\ s\zeta_i\|g_i\|^2 + s\zeta_i \min(0,h_i)^2 + \zeta_i \max(0,h_i)(\alpha_i - \kappa) + o(s) & \text{if } \alpha_i - sh_i < \kappa \end{cases} \\
&\quad \geq s\zeta_i\|g_i\|^2 + s\zeta_i \min(0,h_i)^2 + \zeta_i \min\Big(s\max(0,h_i)^2, \max(0,h_i)(\alpha_i - \kappa)\Big) + o(s)
\end{aligned}
$$

holds for all $s \in (0,\infty)$, where the Landau symbol refers to the limit $s \searrow 0$. Since $\zeta_i \neq 0$ and, as a consequence,

$$
\liminf_{s \searrow 0} \left(\|g_i\|^2 + \min(0,h_i)^2 + \min\left(\max(0,h_i)^2, \max(0,h_i)\frac{(\alpha_i - \kappa)}{s}\right)\right) > 0,
$$

the above yields that there exists an $s_0 > 0$ such that, for all $s \in (0, s_0)$, we have

$$
\begin{aligned}
&F(u_i,\alpha_i) - F(u_i - sg_i, \max(\kappa, \alpha_i - sh_i)) \\
&\qquad \geq \theta s\zeta_i \left(\|g_i\|^2 + \min(0,h_i)^2 + \min\left(\max(0,h_i)^2, \max(0,h_i)\frac{(\alpha_i - \kappa)}{s}\right)\right).
\end{aligned}
$$

This establishes the claim. ∎

We may now prove the following.

THEOREM 5.6 (convergence properties of Algorithm 5.4). *For every choice of the initial guess* $(u_0, \alpha_0) \in \mathbb{R}^n \times [\kappa, \infty)$ *and the parameters* $\sigma > 0$ *and* $\nu, \theta \in (0,1)$, *Algorithm* 5.4 *either terminates after finitely many steps with an iterate which satisfies the stationarity condition in Lemma* 5.3 *or produces an infinite sequence of iterates* $\{(u_i, \alpha_i)\}$ *with the following properties:*

1. *The sequence of function values* $\{F(u_i, \alpha_i)\}$ *is monotonously decreasing.*
2. *The sequence* $\{(u_i, \alpha_i)\}$ *is bounded and has at least one accumulation point.*
3. *Every accumulation point* (u^*, α^*) *of the sequence* $\{(u_i, \alpha_i)\}$ *is stationary in the sense of Lemma* 5.3.

Proof. The proof is fairly standard. We include it for the convenience of the reader and to demonstrate that, in the situation of the problem (5.1), we do not require the assumption of $C^{1,1}$-regularity made, e.g., in [4].

First, we note that Algorithm 5.4 can only terminate after finitely many steps if the exit condition $\zeta_i = 0$ is triggered (cf. Lemma 5.5). This shows that, if only a finite number of iterates is generated, then the last of these iterates is necessarily stationary in the sense of Lemma 5.3. It remains to study the case where Algorithm 5.4 produces an infinite sequence $\{(u_i, \alpha_i)\}$. In this situation, it follows from the sufficient decrease condition used for the calculation of the step sizes τ_i that the sequence $\{F(u_i, \alpha_i)\}$ is monotonously decreasing, and we obtain from the structure of the objective function in (5.1) that there exists a constant $C > 0$ independent of i with

$$
0 \leq \|(u_i, \alpha_i)\|^2 \leq CF(u_i, \alpha_i) \leq CF(u_0, \alpha_0).
$$

The above implies that the sequence $\{(u_i, \alpha_i)\}$ is bounded, that the function values $F(u_i, \alpha_i)$ converge for $i \to \infty$, and that the sequence $\{(u_i, \alpha_i)\}$ possesses at least one accumulation point. To prove that every accumulation point of the iterates is stationary in the sense of Lemma 5.3, we argue by contradiction: Suppose that there exists an accumulation point $(u^*, \alpha^*) \in \mathbb{R}^n \times [\kappa, \infty)$ of the sequence $\{(u_i, \alpha_i)\}$ which satisfies $\zeta(u^*, \alpha^*) > 0$, and let (u_{i_j}, α_{i_j}) be a subsequence with $(u_{i_j}, \alpha_{i_j}) \to (u^*, \alpha^*)$ for $j \to \infty$. Then, it follows from the continuous Fréchet differentiability of F, the definition of the stationarity measure ζ, and the boundedness of $\{(u_i, \alpha_i)\}$ that the sequence $\{\zeta(u_{i_j}, \alpha_{i_j})\} \subset [0, \infty)$ is bounded, and we may assume w.l.o.g. that $\zeta_{i_j} := \zeta(u_{i_j}, \alpha_{i_j}) \to \zeta^*$ holds for some $\zeta^* \geq 0$. Note that we can ignore the case $\zeta^* = 0$ here since this equality would imply $\zeta(u^*, \alpha^*) = 0$ (see (5.6), the continuity of F', and a simple distinction of cases). Thus, w.l.o.g. $\zeta^* > 0$ and $\zeta_{i_j} \geq \varepsilon > 0$ for some constant $\varepsilon > 0$. We now consider two different situations: If there exists a subsequence of i_j (still denoted by the same symbol) such that the step sizes $\tau_{i_j} \in (0, \sigma]$ satisfy $\tau_{i_j} \to \tau^*$ for some $\tau^* > 0$, then we obtain from our line-search procedure, the definitions of g_{i_j} and h_{i_j}, the convergence of $\{F(u_i, \alpha_i)\}$, and the continuity of the derivative F' that

$$
\begin{aligned}
0 &= \lim_{j\to\infty} \big(F(u_{i_j}, \alpha_{i_j}) - F(u_{i_j+1}, \alpha_{i_j+1})\big) \\
&\geq \lim_{j\to\infty} \theta \tau_{i_j} \zeta_{i_j} \left(\|g_{i_j}\|^2 + \min(0, h_{i_j})^2 + \min\left(\max(0, h_{i_j})^2, \max(0, h_{i_j}) \frac{(\alpha_{i_j} - \kappa)}{\tau_{i_j}} \right) \right) \\
&\geq \frac{\theta \tau^*}{\zeta^*} \Bigg(\|(\partial_u F)(u^*, \alpha^*)\|^2 + \min(0, (\partial_\alpha F)(u^*, \alpha^*))^2 \\
&\qquad\qquad + \min\left(\max(0, (\partial_\alpha F)(u^*, \alpha^*))^2, \max(0, (\partial_\alpha F)(u^*, \alpha^*)) \frac{(\alpha^* - \kappa)\zeta^*}{\tau^*} \right) \Bigg).
\end{aligned}
$$

The above implies $\zeta(u^*, \alpha^*) = 0$, which is a contradiction. It remains to consider the case when the step sizes τ_{i_j} tend to zero for $j \to \infty$. In this situation, it follows from the convergence $\tau_{i_j} \to 0$ that the step widths τ_{i_j}/ν have been rejected by the line-search procedure in Algorithm 5.4 for all large enough j, and, as a consequence, that

$$
\begin{aligned}
0 &> F(u_{i_j}, \alpha_{i_j}) - F\left(u_{i_j} - \frac{\tau_{i_j}}{\nu} g_{i_j}, \max\left(\kappa, \alpha_{i_j} - \frac{\tau_{i_j}}{\nu} h_{i_j}\right)\right) \\
&\quad - \theta \frac{\tau_{i_j}}{\nu} \zeta_{i_j} \left(\|g_{i_j}\|^2 + \min(0, h_{i_j})^2 + \min\left(\max(0, h_{i_j})^2, \max(0, h_{i_j}) \frac{(\alpha_{i_j} - \kappa)\nu}{\tau_{i_j}} \right) \right).
\end{aligned}
$$

Rearranging the last estimate and employing the fundamental theorem of calculus yields

(5.7)

$$
\begin{aligned}
&\int_0^1 \left\langle F'\left(u_{i_j} - s \frac{\tau_{i_j}}{\nu} g_{i_j}, \alpha_{i_j} - s\left(\alpha_{i_j} - \max\left(\kappa, \alpha_{i_j} - \frac{\tau_{i_j}}{\nu} h_{i_j}\right)\right)\right), \right. \\
&\qquad\qquad\qquad \left. \left(\frac{\tau_{i_j}}{\nu} g_{i_j}, \alpha_{i_j} - \max\left(\kappa, \alpha_{i_j} - \frac{\tau_{i_j}}{\nu} h_{i_j}\right)\right) \right\rangle \mathrm{d}s \\
&\quad \leq \frac{\theta \tau_{i_j} \zeta_{i_j}}{\nu} \left(\|g_{i_j}\|^2 + \min(0, h_{i_j})^2 + \min\left(\max(0, h_{i_j})^2, \max(0, h_{i_j}) \frac{(\alpha_{i_j} - \kappa)\nu}{\tau_{i_j}} \right) \right).
\end{aligned}
$$

Suppose now that there exists a subsequence of i_j (again not relabeled) such that $\alpha_{i_j} - \tau_{i_j} h_{i_j}/\nu \geq \kappa$ holds for all j. Then, we may divide the left- and right-hand sides of (5.7) by τ_{i_j}, employ the elementary estimate $\min(a, b) \leq a$ for all $a, b \in \mathbb{R}$, and pass to the limit $j \to \infty$ (using the continuity of the derivative F', the boundedness of the iterates $\{(u_i, \alpha_i)\}$, the definitions of g_{i_j} and h_{i_j}, the convergence $\tau_{i_j} \searrow 0$, and the dominated convergence theorem) to obtain

$$\frac{1}{\zeta^* \nu} \|F'(u^*, \alpha^*)\|^2 \leq \frac{\theta}{\zeta^* \nu} \|F'(u^*, \alpha^*)\|^2.$$

This inequality again contradicts our assumption $\zeta(u^*, \alpha^*) > 0$. If, on the other hand, we can find a subsequence of i_j with $\alpha_{i_j} - \tau_{i_j} h_{i_j}/\nu \leq \kappa$, then, along this subsequence, it necessarily holds that

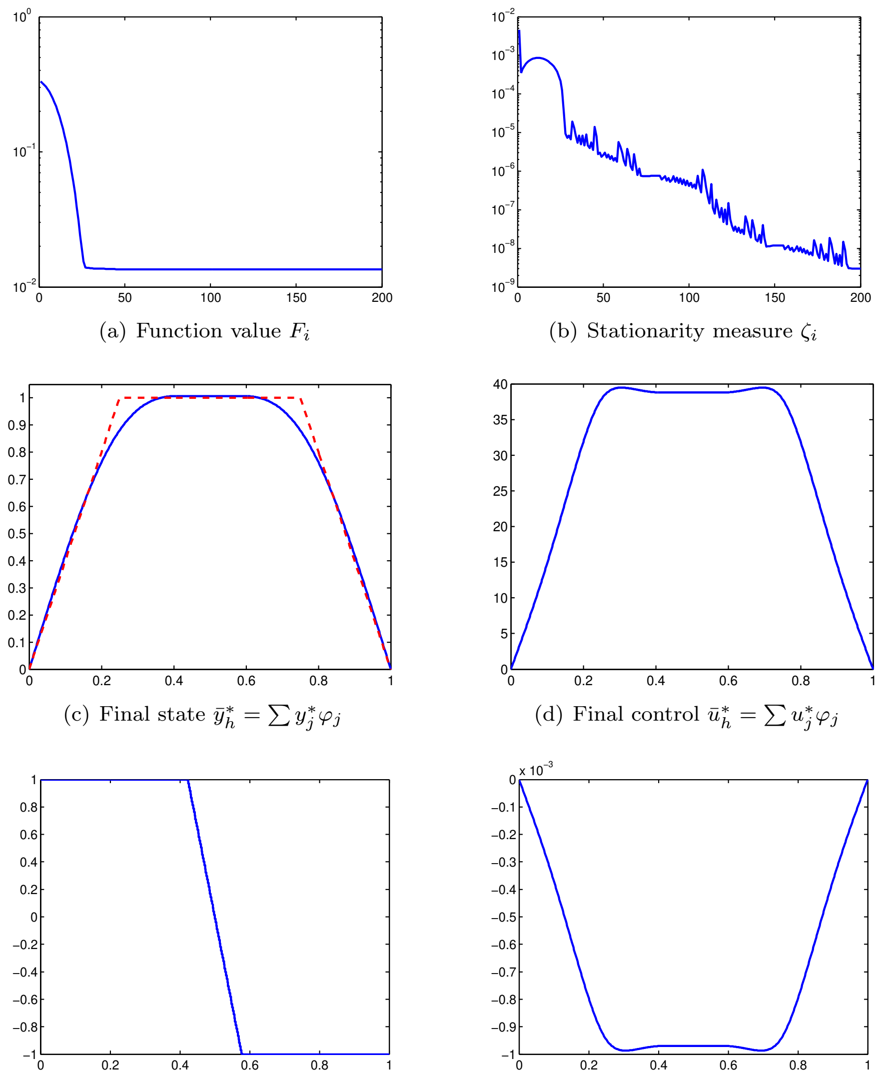
$$0 \leq \frac{\alpha_{i_j} - \kappa}{\tau_{i_j}} \leq \frac{1}{\nu} h_{i_j}, \qquad \alpha_{i_j} \to \kappa = \alpha^*, \quad \text{and} \quad h_{i_j} \to \frac{(\partial_\alpha F)(u^*, \alpha^*)}{\zeta^*} \geq 0, \tag{5.8}$$

and we may assume w.l.o.g. that $(\alpha_{i_j} - \kappa)/\tau_{i_j} \to \xi$ holds for some $\xi \geq 0$. By dividing by τ_{i_j} in (5.7) and by passing to the limit $j \to \infty$, we now obtain, analogously to the case $\alpha_{i_j} - \tau_{i_j} h_{i_j}/\nu \geq \kappa$, that

$$\begin{aligned}
&\frac{1}{\zeta^* \nu} \|(\partial_u F)(u^*, \alpha^*)\|^2 + \xi (\partial_\alpha F)(u^*, \alpha^*) \\
&\leq \frac{\theta}{\zeta^* \nu} \|(\partial_u F)(u^*, \alpha^*)\|^2 + \min\left(\frac{\theta}{\zeta^* \nu} (\partial_\alpha F)(u^*, \alpha^*)^2, \theta \xi (\partial_\alpha F)(u^*, \alpha^*) \right) \\
&\leq \theta \left(\frac{1}{\zeta^* \nu} \|(\partial_u F)(u^*, \alpha^*)\|^2 + \xi (\partial_\alpha F)(u^*, \alpha^*) \right).
\end{aligned}$$

The above implies $\|(\partial_u F)(u^*, \alpha^*)\|^2 = 0$ and yields, in combination with (5.8) and the definition of ζ, that $\zeta(u^*, \alpha^*) = 0$. This is again a contradiction. Accumulation points with $\zeta(u^*, \alpha^*) > 0$ thus cannot exist, and the proof is complete. ∎

The results of a numerical experiment conducted with Algorithm 5.4 can be seen in Figure 2 below. As the plots show, the behavior of the iterates generated by our gradient projection method agrees very well with the predictions of Theorem 5.6. In particular, we observe that the quantity $\zeta_i := \zeta(u_i, \alpha_i)$, which measures the degree of stationarity of the current iterate (u_i, α_i), converges to zero as i tends to infinity. This demonstrates that Theorems 3.3 and 4.1 indeed make it possible to solve bilevel optimization problems of the type (P) with standard gradient-based algorithms. We would like to emphasize at this point that Algorithm 5.4 solves (5.1) "as is," i.e., in the presence of the absolute value functions on the lower level and without any kind of regularization or modification of the problem or the solution procedure (in contrast to the methods in [9, 16, 23, 44, 55, 56]). The latter implies in particular that the sparsity-promoting effects of the nonsmooth terms $|G_k(\cdot)|$ in (5.1) on the gradient of the finite element functions $\bar{y}_h^* := \sum y_j^* \varphi_j$, $\bar{p}_h^* := \sum p_j^* \varphi_j$ and $\bar{u}_h^* := \sum u_j^* \varphi_j$ associated with a solution $(u^*, \alpha^*) \in \mathbb{R}^n \times [\kappa, \infty)$ of (5.1) (cf. subsection 2.1) are preserved in our approach. This can also be seen in Figure 2, where the optimal state $\bar{y}_h^* \in V_h$ has a distinct "flat" region in the middle of the fluid domain. Recall that in the context of the optimal control problem (5.1), the set $\{\nabla \bar{y}_h^* = 0\}$ is exactly that part of the domain Ω where the viscoplastic medium under consideration behaves like a solid (the nucleus). Our solution method thus allows us to identify precisely where rigid material behavior occurs in the fluid domain when we optimize the objective function in (5.1). Such an identification is no longer possible when regularization approaches are used,



(a) Function value F_i

(b) Stationarity measure ζ_i

(c) Final state $\bar{y}_h^* = \sum y_j^* \varphi_j$

(d) Final control $\bar{u}_h^* = \sum u_j^* \varphi_j$

(e) Final multipliers λ_k^*

(f) Final adjoint state $\bar{p}_h^* = \sum p_j^* \varphi_j$

FIG. 2. *Numerical results obtained for the problem* (5.1) *on the interval* $\Omega = (0,1)$ *with an equidistant partition* $\mathcal{T}$ *of width* $1/500$ *and* $\mu = 0.000025$, $\kappa = 3$, $\sigma = 40$, $\nu = 0.25$, $\theta = 0.5$, $\alpha_0 = 4$, *and* $u_0 = (10, 10, \dots, 10)$. *Panels* (a) *and* (b) *show the reduction of the function value* F_i *and the stationarity measure* ζ_i *during the first* 200 *iterations of Algorithm* 5.4. *The state, the control, the multipliers, and the adjoint state of the approximate solution at iteration* 200 *can be seen in panels* (c)–(f). *The considered desired state* $\bar{y}_D \in H_0^1(\Omega)$ *is plotted as a dashed line in* (c). *The multipliers* λ_k^* *are depicted as a step function, whose value on a cell* T_k *of* $\mathcal{T}$ *is* λ_k. *As a tolerance for the residue of the semismooth Newton method used for the solution of* (5.2), *we chose* 10^{-10}.

which necessarily remove the sparsity-promoting effects from the problem. Note that the functions $\bar{p}_h^* \in V_h$ and $\bar{u}_h^* \in V_h$ associated with a solution (u^*, α^*) of (5.1) directly inherit the "flatness" properties of $\bar{y}_h^*$ due to (5.5) and since the optimality condition $\zeta(u^*, \alpha^*) = 0$ implies $\mu u^* + p^* = 0$; see (5.4).

We conclude this paper with some additional remarks on Algorithm 5.4, the numerical results in Figure 2, and the analysis in sections 3 and 4.

Remark 5.7.

1. As the graph in Figure 2(a) shows, in our numerical experiment we no longer observe a significant decrease in the function value F_i after approximately 25 gradient steps. This number of iterations is thus sufficient if we are primarily interested in determining a tuple (u, α) for which the value of the reduced objective function is small. We would like to point out that, although F_i remains nearly constant for $i \geq 25$, the stationarity measure ζ_i still changes in the later iterations of the algorithm. The same is true for the quantities u_i and p_i. The two global maxima of the control seen in Figure 2(d), for example, are not visible until $i \approx 60$.
2. In the numerical experiment of Figure 2, the Oldroyd number α is decreased from the initial guess $\alpha_0 = 4$ to the lower bound $\kappa = 3$ in the first three iterations of Algorithm 5.4 and afterwards remains constant. This makes sense since a low α is preferable in the situation of the discrete tracking-type optimal control problem (5.1). (The lower the material parameter α, the lower the yield stress and the smaller the pressure gradient that is needed to create a desired flow profile.) If we replace the term α^2 on the upper level of (5.1) with, e.g., $(\alpha - \alpha_D)^2$ for some sufficiently large $\alpha_D > 0$, then this behavior changes, and we observe convergence to an $\alpha^* > \kappa$.
3. It is easy to check that the solution map $S : \mathbb{R}^n \times (0, \infty) \to \mathbb{R}^n$ associated with the problem on the lower level of (5.1) is constant zero in an open neighborhood of the cone $\{0\} \times (0, \infty)$. (This is precisely the set where the pressure gradient is not large enough to move the fluid under consideration; compare also with (3.14).) Because of this behavior, the tuple $(0, \kappa)$ is always a local minimizer of (5.1) and a stationary point in the sense of Lemma 5.3. To avoid converging to the point $(0, \kappa)$, which is typically not globally optimal and thus only of limited interest, one has to choose an initial guess u_0 that is sufficiently far away from the origin when an iterative method analogous to Algorithm 5.4 is used for the solution of (5.1). Note that in the situation of Figure 2, we have indeed found a point that is better than $(0, \kappa)$ since the final iterate achieves a function value that is far smaller than $F(0, \kappa) \approx 0.333446$.
4. Note that (4.1) and (5.5) can also be formulated as quadratic minimization problems with linear equality constraints. This makes it possible to use standard methods from quadratic programming for the calculation of the adjoint state and the gradient of the reduced objective function. In the numerical experiment of Figure 2, we determined p with the MATLAB routine `quadprog`.
5. We would like to point out that if an iterative scheme is used for the solution of the lower-level problem in (5.1), then the successive evaluations of the maps S and F in the line-search procedure of Algorithm 5.4 can be performed quite effectively since the last trial iterate can always be used as an initial guess for the calculation of the next required state $S(u_i - \sigma_j g_i, \max(\kappa, \alpha_i - \sigma_j h_i))$. In the situation of Figure 2, it can be observed that Algorithm 5.4 requires on average approximately four evaluations of the solution operator S per gradient step over the first 200 iterations. The majority of these calculations are needed for large i.
6. In this section, we have considered Algorithm 5.4 as a stand-alone solution procedure for the bilevel optimization problem (5.1). This is, of course, not

necessary. We could have also combined our algorithm with an inaccurate but cheap method (e.g., a regularization approach) that provides a good initial guess (u_0, α_0). Such a technique has been used, e.g., in [12, section 5].

7. Alternatively to the approach we pursued in this section, one could also try to tackle the necessary optimality condition $\zeta(u^*, \alpha^*) = 0$ in Lemma 5.3 directly, e.g., with a Newton-type method or a primal-dual-active-set–type algorithm. (Note that some care has to be taken here since Theorem 3.3 only provides a first derivative but not a second, so that a classical Newton algorithm is out of the question.) We leave this topic for further research.

REFERENCES

[1] R. A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.

[2] S. Adly and L. Bourdin, *Sensitivity analysis of variational inequalities via twice epi-differentiability and proto-differentiability of the proximity operator*, SIAM J. Optim., 28 (2018), pp. 1699–1725, https://doi.org/10.1137/17M1135013.

[3] H. Attouch, G. Buttazzo, and G. Michaille, *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*, 2nd ed., MOS/SIAM Ser. Optim. 17, SIAM, Philadelphia, 2014, https://doi.org/10.1137/1.9781611973488.

[4] A. Beck, *Introduction to Nonlinear Optimization: Theory, Algorithms, and Applications with MATLAB*, MOS-SIAM Ser. Optim. 19, SIAM, Philadelphia, 2014, https://doi.org/10.1137/1.9781611973655.

[5] F. Benita, S. Dempe, and P. Mehlitz, *Bilevel optimal control problems with pure state constraints and finite-dimensional lower level*, SIAM J. Optim., 26 (2016), pp. 564–588, https://doi.org/10.1137/141000889.

[6] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, Berlin, 2000.

[7] L. Calatroni, C. Chung, J. C. De los Reyes, J. Carlos, C.-B. Schönlieb, and T. Valkonen, *Bilevel approaches for learning of variational imaging models*, in Variational Methods: In Imaging and Geometric Control, M. Bergounioux et al., eds., Radon Ser. Comput. Appl. Math. 18, De Gruyter, Berlin, 2017, pp. 252–290.

[8] C. Carstensen, B. D. Reddy, and M. Schedensack, *A natural nonconforming FEM for the Bingham flow problem is quasi-optimal*, Numer. Math., 133 (2016), pp. 37–66.

[9] Y. J. Chen, T. Pock, and H. Bischof, *Learning ℓ_1-based analysis and synthesis sparsity priors using bi-level optimization*, Poster Session Presented at the 26th Annual Neural Information Processing Systems Conference (NIPS 2012), Lake Tahoe, 2012.

[10] C. Christof, *L^∞-error estimates for the obstacle problem revisited*, Calcolo, 54 (2017), pp. 1243–1264.

[11] C. Christof, *Sensitivity Analysis of Elliptic Variational Inequalities of the First and the Second Kind*, Ph.D. thesis, Technische Universität Dortmund, 2018.

[12] C. Christof, J. C. De los Reyes, and C. Meyer, *A Non-Smooth trust-Region Method for Locally Lipschitz Functions with Application to Optimization Problems Constrained by Variational Inequalities*, preprint, https://arxiv.org/abs/1711.03208, 2018.

[13] C. Christof and C. Meyer, *Sensitivity analysis for a class of H_0^1-elliptic variational inequalities of the second kind*, Set-Valued Var. Anal., 27 (2018), pp. 469–502, https://doi.org/10.1007/s11228-018-0495-2.

[14] C. Christof, C. Meyer, S. Walther, and C. Clason, *Optimal control of a non-smooth semilinear elliptic equation*, Math. Control Relat. Fields, 8 (2018), pp. 247–276.

[15] C. Christof and G. Wachsmuth, *Differential sensitivity analysis of variational inequalities with locally Lipschitz continuous solution operators*, Appl. Math. Optim., to appear; published online January 10, 2019.

[16] C. Chung, J. C. De los Reyes, and C.-B. Schönlieb, *Learning optimal spatially-dependent regularization parameters in total variation image denoising*, Inverse Problems, 33 (2017), pp. 1–31.

[17] M. B. Cohen, J. Diakonikolas, and L. Orecchia, *On Acceleration with Noise-Corrupted Gradients*, in Proceedings of the 35th International Conference on Machine Learning (PMLR 80), Stockholm, Sweden, 2018, pp. 1019–1028.

[18] A. R. Conn, N. Gould, A. Sartenaer, and Ph. L. Toint, *Global convergence of a class of trust region algorithms for optimization using inexact projections on convex constraints*, SIAM J. Optim., 3 (1993), pp. 164–221, https://doi.org/10.1137/0803009.

[19] A. R. Conn, N. I. M. Gould, and Ph. L. Toint, *Trust-Region Methods*, MOS-SIAM Ser. Optim. 1, SIAM, Philadelphia, 2000, https://doi.org/10.1137/1.9780898719857.
[20] J. C. De los Reyes, *Optimal control of a class of variational inequalities of the second kind*, SIAM J. Control Optim., 49 (2011), pp. 1629–1658, https://doi.org/10.1137/090764438.
[21] J. C. De los Reyes, *Optimization of mixed variational inequalities arising in flow of viscoplastic materials*, Comput. Optim. Appl., 52 (2012), pp. 757–784.
[22] J. C. De los Reyes and C. Meyer, *Strong stationarity conditions for a class of optimization problems governed by variational inequalities of the second kind*, J. Optim. Theory Appl., 168 (2016), pp. 375–409.
[23] J. C. De los Reyes, C.-B. Schönlieb, and T. Valkonen, *Bilevel parameter learning for higher-order total variation regularisation models*, J. Math. Imaging Vision, 57 (2017), pp. 1–25.
[24] E. J. Dean, R. Glowinski, and G. Guidoboni, *On the numerical simulation of Bingham visco-plastic flow: Old and new results*, J. Non-Newton. Fluid Mech., 142 (2007), pp. 36–62.
[25] S. Dempe, *A necessary and a sufficient optimality condition for bilevel programming problems*, Optimization, 25 (1992), pp. 341–354.
[26] S. Dempe, *Foundations of Bilevel Programming*, Nonconvex Optim. Appl. 61, Springer, New York, 2002.
[27] S. Dempe, J. Dutta, and B. S. Mordukhovich, *New necessary optimality conditions in optimistic bilevel programming*, Optimization, 56 (2007), pp. 577–604.
[28] S. Dempe, V. Kalashnikov, G. A. Pérez-Valdés, and N. Kalashnykova, *Bilevel Programming Problems: Theory, Algorithms and Applications to Energy Networks*, Energy Systems, Springer-Verlag, Berlin, 2015.
[29] S. Dempe and A. B. Zemkoho, *KKT reformulation and necessary conditions for optimality in nonsmooth bilevel optimization*, SIAM J. Optim., 24 (2014), pp. 1639–1669, https://doi.org/10.1137/130917715.
[30] S. Dempe and A. B. Zemkoho, *The bilevel programming problem: Reformulations, constraint qualifications and optimality conditions*, Math. Program., 138 (2013), pp. 447–473.
[31] P. Dvurechensky, *Gradient Method with Inexact Oracle for Composite Non-Convex Optimization*, preprint, https://arxiv.org/abs/1703.09180, 2017.
[32] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, translated from the French, corrected reprint of the 1976 English edition, Classics Appl. Math. 28, SIAM, Philadelphia, 1999, https://doi.org/10.1137/1.9781611971088.
[33] W. Forst and D. Hoffman, *Optimization - Theory and Practice*, Springer Undergrad. Texts Math. Technol., Springer, New York, 2010.
[34] T. L. Friesz, R. L. Tobin, H.-J. Cho, and N. J. Mehta, *Sensitivity analysis based heuristic algorithms for mathematical programs with variational inequality constraints*, Math. Program., 48 (1990), pp. 265–284.
[35] R. Glowinski, *Lectures on Numerical Methods for Non-Linear Variational Problems*, Tata Institute of Fundamental Research, Bombay, 1980.
[36] M. Heinkenschloss and L. N. Vicente, *Analysis of inexact trust-region SQP algorithms*, SIAM J. Optim., 12 (2001), pp. 283–302, https://doi.org/10.1137/S1052623499361543.
[37] M. Hintermüller and C. N. Rautenberg, *Optimal selection of the regularization function in a weighted total variation model: Part* I*: Modelling and theory*, J. Math. Imaging Vision, 59 (2017), pp. 498–514.
[38] M. Hintermüller, C. N. Rautenberg, T. Wu, and A. Langer, *Optimal selection of the regularization function in a weighted total variation model. Part* II*: Algorithm, its analysis and numerical tests*, J. Math. Imaging Vision, 59 (2017), pp. 515—533.
[39] M. Hintermüller and T. M. Surowiec, *On the directional differentiability of the solution mapping for a class of variational inequalities of the second kind*, Set-Valued Var. Anal., 26 (2018), pp. 631–642.
[40] M. Hintermüller and T. Wu, *Bilevel optimization for calibrating point spread functions in blind deconvolution*, Inverse Probl. Imaging, 9 (2015), pp. 1139–1169.
[41] G. Holler, K. Kunisch, and R. C. Barnard, *A bilevel approach for parameter learning in inverse problems*, Inverse Problems, 34 (2018), pp. 1–28.
[42] R. R. Huilgol and Z. You, *Application of the augmented Lagrangian method to steady pipe flows of Bingham, Casson and Herschel-Bulkley fluids*, J. Non-Newton. Fluid Mech., 128 (2005), pp. 126–143.
[43] C. T. Kelley, *Iterative Methods for Optimization*, Frontiers Appl. Math. 18, SIAM, Philadelphia, 1999, https://doi.org/10.1137/1.9781611970920.

[44] K. Kunisch and T. Pock, *A bilevel optimization approach for parameter learning in variational models*, SIAM J. Imaging Sci., 6 (2013), pp. 938–983, https://doi.org/10.1137/120882706.

[45] A. B. Levy, *Implicit multifunction theorems for the sensitivity analysis of variational conditions*, Math. Program., 74 (1996), pp. 333–350.

[46] G.-H. Lin, M. Xu, and J. J. Ye, *On solving simple bilevel programs with a nonconvex lower level program*, Math. Program., 144 (2014), pp. 277–305.

[47] P. Lindqvist, *Stability of solutions of* $\operatorname{div}(|\nabla u|^{p-2}\nabla u) = f$ *with varying* p, J. Math. Anal. Appl., 127 (1987), pp. 93–102.

[48] P. Mehlitz, *Contributions to Complementarity and Bilevel Programming in Banach Spaces*, Ph.D. thesis, Technische Universität Bergakademie Freiberg, 2017.

[49] P. Mehlitz, *Necessary optimality conditions for a special class of bilevel programming problems with unique lower level solution*, Optimization, 66 (2017), pp. 1533–1562.

[50] P. Mehlitz and G. Wachsmuth, *Weak and strong stationarity in generalized bilevel programming and bilevel optimal control*, Optimization, 65 (2016), pp. 907–935.

[51] P. P. Mosolov and V. P. Miasnikov, *Variational methods in the theory of the fluidity of a viscous-plastic medium*, J. Appl. Math. Mech., 29 (1965), pp. 545–577.

[52] P. P. Mosolov and V. P. Miasnikov, *On stagnant flow regions of a viscous-plastic medium in pipes*, PMM, 30 (1966), pp. 705–717 (in Russian); English translation in J. Appl. Math. Mech., 30 (1966), pp. 841–854, https://doi.org/10.1016/0021-8928(66)90035-9.

[53] P. P. Mosolov and V. P. Miasnikov, *On qualitative singularities of the flow of a viscoplastic medium in pipes*, J. Appl. Math. Mech., 31 (1967), pp. 609–613.

[54] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed., Springer Ser. Oper. Res. Financ. Eng., Springer, New York, 2006.

[55] P. Ochs, R. Ranftl, T. Brox, and T. Pock, *Bilevel optimization with nonsmooth lower level problems*, in Scale Space and Variational Methods in Computer Vision, Lecture Notes Comput. Sci. 9087, Springer, Cham, 2015, pp. 654–665.

[56] G. Peyré and J. Fadili, *Learning analysis sparsity priors*, in Proceedings of the 9th International Conference on Sampling Theory and Applications (Sampta'11), Singapore, 2011, https://hal.archives-ouvertes.fr/hal-00542016/.

[57] R. Pytlak, *Conjugate Gradient Algorithms in Nonconvex Optimization*, Nonconvex Optim. Appl. 89, Springer-Verlag, Berlin, 2009.

[58] A.-T. Rauls and G. Wachsmuth, *Generalized derivatives for the solution operator of the obstacle problem*, Set-Valued Var. Anal., to appear; published online February 7, 2019.

[59] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Grundlehren Math. Wiss. 317, Springer, Berlin, 1998.

[60] M. Ruzicka, *Nichtlineare Funktionalanalysis*, Springer, Berlin, 2004.

[61] C. Sabett, A. Hafftka, K. Sexton, and R. G. Spencer, L_1*,* L_p*,* L_2*, and elastic net penalties for regularization of Gaussian component distributions in magnetic resonance relaxometry*, Concepts Mag. Res., 46A (2017), e21427.

[62] E. W. Sachs, *Convergence of algorithms for perturbed optimization problems*, Ann. Oper. Res., 27 (1990), pp. 311–341.

[63] M. Solodov, *An explicit descent method for bilevel convex optimization*, J. Convex Anal., 14 (2007) pp. 227–237.

[64] S. Vaiter, G. Peyre, C. Dossal, and J. Fadili, *Robust sparse analysis regularization*, IEEE Trans. Inform. Theory, 59 (2013), pp. 2001–2016.

[65] L. Vicente, G. Savard, and J. Júdice, *Descent approaches for quadratic bilevel programming*, J. Optim. Theory Appl., 81 (1994), pp. 379–399.

[66] M. Xu, S.-Y. Wu, and J. J. Ye, *Solving semi-infinite programs by smoothing projected gradient method*, Comput. Optim. Appl., 59 (2014), pp. 591–616.

[67] J. J. Ye, *Constraint qualifications and KKT conditions for bilevel programming problems*, Math. Oper. Res., 31 (2006), pp. 811–824.

[68] J. J. Ye and X. Y. Ye, *Necessary optimality conditions for optimization problems with variational inequality constraints*, Math. Oper. Res., 22 (1997), pp. 977–997.

[69] J. J. Ye and D. Zhu, *New necessary optimality conditions for bilevel programs by combining the MPEC and value function approaches*, SIAM J. Optim., 20 (2010), pp. 1885–1905, https://doi.org/10.1137/080725088.

[70] J. J. Ye and D. L. Zhu, *Optimality conditions for bilevel programming problems*, Optimization, 33 (1995), pp. 9–27.

[71] H. Zou and T. Hastie, *Regularization and variable selection via the elastic net*, J. R. Stat. Soc. Ser. B. Stat. Methodol., 67 (2005), pp. 301–320.