Optimality Conditions for Bilevel Imaging Learning Problems with Total Variation Regularization

Juan Carlos De los Reyes* David Villacís†

July 20, 2021

Abstract

We address the problem of optimal scale-dependent parameter learning in total variation image denoising. Such problems are formulated as bilevel optimization instances with total variation denoising problems as lower-level constraints. For the bilevel problem, we are able to derive M-stationarity conditions, after characterizing the corresponding Mordukhovich generalized normal cone and verifying suitable constraint qualification conditions. We also derive B-stationarity conditions, after investigating the Lipschitz continuity and directional differentiability of the lower-level solution operator. A characterization of the Bouligand subdifferential of the solution mapping, by means of a properly defined linear system, is provided as well. Based on this characterization, we propose a two-phase non-smooth trust-region algorithm for the numerical solution of the bilevel problem and test it computationally for two particular experimental settings.

Keywords: Bilevel optimization, machine learning, variational models, total variation.
AMS: 49K99, 90C33, 68U10, 68T99, 65K10

1 Introduction

In this work we address the bilevel optimal selection of parameters in total variation image denoising models, from a variational and nonsmooth analysis perspectives. Bilevel techniques were proposed for optimal parameter selection of variational models in the seminal work [11], provided a training dataset to learn from. The variational models considered there were based on the total variation (TV) seminorm and different noise statistics models were taken into account. Thereafter, apart of variational denoising problems, the bilevel learning framework has been considered for other imaging applications such as image restoration [10, 9], blind image deconvolution [18], image segmentation [24, 26], nonlocal models [7] and kernel parameter estimation for support vector machines [20]. The bilevel methodology was also extended by Cao et al. [31, 3], Hintermüller et al. [18, 13] and Strong et al. [30] for learning different optimal *scale-dependent* total variation (TV) and total generalized variation (TGV) denoising parameters. Either using a training set or noise statistics of the image, scale-dependent problems face the risk of leading to overfitted optimal parameters (see, e.g., [12]). To gain generalization, alternative parameter functions may be considered that lie between a scalar and a fully scale-dependent weight. Examples of those intermediate functions are patch-dependent parameters, dictionary weights or basis functions coefficients.

One of the most challenging aspects in bilevel imaging learning, either with scalar or scale-dependent parameters, is the derivation of necessary optimality conditions that characterize the optimal solutions sharply and that may be used for numerical purposes. The approach pursued in [11, 21, 31] was based on a local regularization of the nonsmooth terms and an asymptotic analysis thereafter, yielding a *C-stationarity system* that does not fully characterize optimal parameters. A related approach, based on a dual reformulation of the lower level instance, was studied in [16, 17]. Alternatively, a direct nonsmooth approach was considered in [18] and [12] to learn point spread functions in blind deconvolution models and the weight in front of the fidelity term in denoising models, respectively. In both such cases, the parameter

*Research Center for Mathematical Modelling (MODEMAT), Escuela Politécnica Nacional, Quito, Ecuador (juan.delosreyes@epn.edu.ec, david.villacis01@epn.edu.ec, `http://www.modemat.epn.edu.ec/`).

affects the fidelity term and does not alter the structure of the primal-dual reformulation of the lower-level problem. Based on variational analysis tools, an M-stationarity system was then obtained.

Our aim in this paper consists in deriving sharp optimality conditions of *Mordukhovich (M-)* and *Bouligand (B-)* type for total variation bilevel learning problems, when the scale-dependent parameter appears within the regularizer. In such cases, differently from [18, 12], the nonnegativity constraint of the parameter alters the structure of the primal-dual characterization of the variational denoising model, leading to the appearance of not only a biactive, but also a *triactive set*. This involved nonsmoothness significantly complicates the derivation of the corresponding Fréchet and Mordukhovich normal cones, as well as the nonsmooth analysis of the lower-level solution mapping.

Firstly, after introducing the bilevel problem, we are able to reformulate it as a *generalized mathematical program with equilibrium constraints (GMPEC)*. We rigorously characterize the corresponding Mordukhovich generalized normal cone (Theorem 3.4) and verify a constraint qualification condition. This leads to the derivation of an *M-stationarity* system for the characterization of local minima (Theorem 3.5). Thereafter, we focus on *B-stationarity* conditions, which require the directional differentiability of the solution mapping. In that respect, we are able to prove the Lipschitz continuity and directional differentiability of the solution operator, and characterize the directional derivative by means of a variational inequality (Theorem 4.6). In case of strict complementarity, the directional derivative becomes the Fréchet one and is characterized by a linear system. Also the Bouligand subdifferential of the solution operator is characterized by means of a general linear problem, based on a splitting of the biactive set and a properly defined subspace (Theorem 5.1).

A further challenge is related with the numerical solution of the bilevel learning problems. Based on the properties of a local regularization technique, a numerical optimization algorithm of second-order type was considered in [11] and further explored in [2], where dynamic sampling techniques were taken into account to accelerate the algorithm in the case of large datasets. A new approach handling the bilevel problem using a non-smooth setting can be found in the work of Ochs et. al. [24], where the authors make use of algorithmic differentiation to find optimal parameters in variational models. An additional contribution of this paper is related to the solution of the bilevel instances by means of a nonsmooth trust-region algorithm. Specifically, based on the results in [6] and the Bouligand subdifferential characterization, we devise a two-phase trust-region algorithm for the solution of the problems (Algorithm 1). In the first phase, the Bouligand subdifferential is used in the quadratic trust-region subproblem, while in the second one, when the trust-region radius becomes small enough, a regularized gradient is utilized. We verify the performance of the proposed algorithm by means of scalar and scale-dependent experiments.

The paper is organized as follows. We will describe the bilevel learning problem in Section 2 along with the standing assumptions. In Section 3 we will characterize the Mordhukovich generalized normal cone and verify an appropriate constraint qualification condition, that will enable us to derive an M-stationary optimality system for the bilevel learning problem. In Section 4 we study the Lipschitz continuity and directional differentiability of the solution operator, with a proper characterization of the directional derivative, to obtain B-stationarity conditions for the bilevel problem. Thereafter, in Section 5 we characterize the Bouligand subdifferential of the solution mapping by means of a generalized linear system. This characterization of the Bouligand subdifferential will be of use in Section 6, where we propose a non-smooth trust-region solution algorithm. Finally, detailed numerical experiments will be provided in Section 7.

2 Problem Formulation

Variational image denoising is a known *ill-posed* inverse problem. Indeed, by considering an image as a $m_1 \times m_2$ pixel matrix and mapping this matrix into a vector of length $m = m_1 m_2$, a variational denoising model can be formulated as follows:

$$\min_{u \in \mathbb{R}^m} \mathcal{E}(u) = \phi(u, f) + \mathcal{R}(u, \alpha), \tag{1}$$

where f is the noise contaminated image, α is a parameter that balances a *data fidelity* term ϕ and a *regularization* term $\mathcal{R}$.

This paper will focus on the case where the regularization term corresponds to the *Isotropic Total Variation Semi-norm* (TV). To characterize the TV regularizer, let us introduce the discrete gradient operator

$\mathbb{K} : \mathbb{R}^m \to \mathbb{R}^{n\times 2}$

$$\mathbb{K}u = (\mathbb{K}_x u, \mathbb{K}_y u) \in \mathbb{R}^n \times \mathbb{R}^n.$$

The matrix $\mathbb{K}_x$ computes the difference in all neighboring pixel intensities in the x-direction and $\mathbb{K}_y$ in the y-direction. Taking $u \in \mathbb{R}^m$ as a given image, the isotropic discrete total variation is then defined by $\|\mathbb{K}u\|_{2,1} := \sum_{j=1}^{n} \|(\mathbb{K}u)_j\|$, where $\|\cdot\|$ is the euclidean norm and $(\mathbb{K}u)_j$ corresponds to the j-th row of $\mathbb{K}u$. This regularizer induces sparsity on the gradients of the image, which favors piecewise constant images with sparse edges [29].

Now, regarding the balance parameter α, it can be considered as a scalar value in the case we assume a uniform noise distribution across the entire image. However, such assumption is known to be unrealistic for most applications. Therefore, we will explore different types of balance parameters, starting with a *scale-dependent* parameter $\alpha \in \mathbb{R}^n_+$ and the following form of the regularization term:

$$\mathcal{R}(u, \alpha) = \sum_{j=1}^{n} \alpha_j \|(\mathbb{K}u)_j\|. \tag{2}$$

With these considerations in mind, and assuming a C^2 and strongly convex data fidelity term $\phi : \mathbb{R}^m \to \mathbb{R}$, our variational denoising model reads as follows:

$$\min_{u\in\mathbb{R}^n} \ \mathcal{E}(u) := \phi(u, f) + \sum_{j=1}^{n} \alpha_j \|(\mathbb{K}u)_j\|. \tag{3}$$

In certain circumstances, such scale-dependent parameter may lead to an overfitted solution, which is why an intermediate approach is often desirable. One of such approaches considers a parameter $\alpha \in \mathbb{R}^p$ with $p << n$ as piecewise constant in different patches of the image. Along the paper, we will derive the theoretical results only for the scale-dependent model, since they easily extend to the patch-based and scalar models.

We will focus on bilevel strategies for finding optimal regularization parameters for the image denoising model eq. (3). That is, given a *training set* containing N pairs of damaged and "noise-free" pairs of images, we will find an optimal regularization parameter $\alpha \in \mathbb{R}^n_+$ that minimizes a continuously differentiable, proper and strongly convex loss function $J : \mathbb{R}^m \to \mathbb{R}$, that measures the quality of the reconstruction with respect to the ground truth images in the training set.

Using training set pairs $(\bar{u}_i, f_i)$, $i = 1, \ldots, N$, with noisy images f_i and ground truth images $\bar{u}_i$, our bilevel parameter learning problem reads as follows:

$$\min_{\alpha\in\mathbb{R}^n, \alpha\geq 0} \quad \sum_{i=1}^{N} J(u_i, \bar{u}_i) \tag{4a}$$

$$\text{s.t.} \qquad u_i \in \arg\min_{u\in\mathbb{R}^m} \phi(u, f_i) + \sum_{j=1}^{n} \alpha_j \|(\mathbb{K}u)_j\|, \qquad i = 1, \ldots, N \tag{4b}$$

Regarding the spatially dependent TV denoising problem eq. (3), we can guarantee the existence of a unique solution as well as characterize it through a necessary and sufficient optimality condition.

Theorem 2.1. Problem eq. (3) has a unique solution $u^* \in \mathbb{R}^m$. Moreover, a necessary and sufficient condition is given by the following variational inequality of the second kind

$$\langle \phi'(u^*), v - u^* \rangle + \sum_{j=1}^{n} \alpha_j \|(\mathbb{K}v)_j\| - \sum_{j=1}^{n} \alpha_j \|(\mathbb{K}u^*)_j\| \geq 0, \ \forall v \in \mathbb{R}^m. \tag{5}$$

Proof. The strong convexity of ϕ, along with the convexity of the total variation semi-norm, yields the strong convexity of the lower level optimization problem. Consequently this problem has a unique optimizer u^*. The variational inequality then follows by standard arguments (see, e.g., [8, Theorem 6.1]) $\square$

Using duality techniques [14], the variational inequality of the second kind (5) can be equivalently written in primal-dual form, yielding the following reformulation of the lower-level problem:

$$\phi'(u) + \mathbb{K}^\top q = 0, \tag{6a}$$
$$\langle q_j, (\mathbb{K}u)_j \rangle - \alpha_j \|(\mathbb{K}u)_j\| = 0, \qquad \forall j = 1, \ldots, n, \tag{6b}$$
$$\|q_j\| - \alpha_j \leq 0, \qquad \forall j = 1, \ldots, n, \tag{6c}$$

Consequently, the bilevel parameter learning problem, for a single training pair, can be written as:

$$\min_{\alpha \in \mathbb{R}^n} \quad J(u, \bar{u}) \tag{7a}$$
$$\text{s.t.} \quad \phi'(u) + \mathbb{K}^\top q = 0, \tag{7b}$$
$$\langle q_j, (\mathbb{K}u)_j \rangle - \alpha_j \|(\mathbb{K}u)_j\| = 0, \qquad \forall j = 1, \ldots, n, \tag{7c}$$
$$\|q_j\| - \alpha_j \leq 0, \qquad \forall j = 1, \ldots, n, \tag{7d}$$
$$\alpha_j \geq 0, \qquad \forall j = 1, \ldots, n. \tag{7e}$$

For the sake of clarity in the exposition, we restrict hereafter the analysis to the case of a single training pair. The results are, however, easily extendable to larger training sets.

3 Mordukhovich Stationarity

In this section we address stationarity conditions for the bilevel problem eq. (7) using variational analysis tools. Specifically, we reformulate the bilevel problem as a generalized mathematical program with equilibrium constraints (GMPEC) and verify a constraint qualification condition based on the variational geometry of the solution set.

A generalized mathematical program with equilibrium constraints may be formulated in the following general form:

$$\begin{aligned} \min \quad & f(x,y) \\ \text{s.t.} \quad & 0 \in F_1(x,y) + Q(F_2(x,y)), \\ & (x,y) \in \omega, \end{aligned} \tag{8}$$

where f is locally Lipschitz on $\mathbb{R}^n \times \mathbb{R}^m$, $F_1 : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ and $F_2 : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^l$ are continuously differentiable, $\omega \subset \mathbb{R}^n \times \mathbb{R}^m$ is non-empty and closed and $Q : \mathbb{R}^l \rightrightarrows \mathbb{R}^m$ is a multifunction with closed graph. A constraint qualification condition for GMPEC that guarantee the existence of KKT multipliers and its corresponding stationarity system is privede in the next result.

Theorem 3.1 (Outrata [25])**.** *Let (x^*, y^*) be a local solution of eq. (8) and the following constraint qualification*

$$\left.\begin{aligned} \begin{bmatrix} (\nabla_x F_2(x^*,y^*))^\top & -(\nabla_x F_1(x^*,y^*))^\top \\ (\nabla_y F_2(x^*,y^*))^\top & -(\nabla_y F_1(x^*,y^*))^\top \end{bmatrix} \begin{bmatrix} w \\ z \end{bmatrix} \in -N^M_\omega(x^*,y^*), \\ (w,z) \in N^M_{\operatorname{gph} Q}(F_2(x^*,y^*), -F_1(x^*.y^*)) \end{aligned}\right\} \text{implies} \begin{cases} w = 0, \\ z = 0 \end{cases} \tag{9}$$

holds true. Then there exist a pair $(\xi, \eta) \in \partial f(x^*, y^*)$, a pair $(\gamma, \delta) \in N^M_\omega(x^*, y^*)$, and a KKT pair $(w^*, z^*) \in N^M_{\operatorname{gph} Q}(F_2(x^*, y^*), -F_1(x^*, y^*))$ such that

$$0 = \xi + (\nabla_x F_2(x^*,y^*))^\top w^* - (\nabla_x F_1(x^*,y^*))^\top z^* + \gamma,$$
$$0 = \eta + (\nabla_y F_2(x^*,y^*))^\top w^* - (\nabla_y F_1(x^*,y^*))^\top z^* + \delta.$$

where $N^M_{\operatorname{gph} Q}$ stands for Mordukhovich generalized normal cone to the graph of Q.

In our case, using the primal-dual formulation eq. (6), the constraints may be written as

$$0 \in \phi'(u) + Q(\alpha, u), \tag{10}$$

where $Q : \mathbb{R}^n_+ \times \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ is the set-valued operator associated to the subdifferential of the Euclidean norm, i.e.,

$$Q(\alpha, u) := \left\{ \mathbb{K}^\top q : q \in \mathbb{R}^{n\times 2}, \begin{cases} q_j = \alpha_j \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, & \text{if } \|(\mathbb{K}u)_j\| \neq 0, \alpha_j \geq 0, \\ \|q_j\| \leq \alpha_j, & \text{if } \|(\mathbb{K}u)_j\| = 0, \alpha_j \geq 0. \end{cases} \right\} \tag{11}$$

Equivalently, by making use of the definition of the graph of the multifunction Q we can rewrite eq. (10) as follows

$$\phi'(u) + \mathbb{K}^\top q = 0, \tag{12a}$$

$$(\alpha, u, \mathbb{K}^\top q) \in \text{gph } Q, \tag{12b}$$

$$(\alpha, u) \in \omega := \mathbb{R}^n_+ \times \mathbb{R}^m, \tag{12c}$$

where $\operatorname{gph} Q := \{(\alpha, u, \mathbb{K}^\top q) \in \mathbb{R}^n_+ \times \mathbb{R}^m \times \mathbb{R}^m : \mathbb{K}^\top q \in Q(\alpha, u)\}$.

The constraint qualification in theorem 3.1 guarantees the existence of multipliers that allow the derivation of a stationarity system. In [18] such constraint qualification condition was circumvented by using Robinson strong regularity of the set valued equation. In contrast, the structure of the multifunction eq. (11), depending both on the regularization parameter α and the image u, prevent us from using the same shortcut.

Using the structure of the set valued operator Q presented in eq. (11), let us introduce the following notation for the inactive, strongly active, biactive, zero-inactive and triactive sets, respectively:

$$\begin{aligned}
\mathcal{I}(\alpha, u) &:= \{j \in \{1, \dots, n\} : (\mathbb{K}u)_j \neq 0,\ \alpha_j > 0\}, \\
\mathcal{A}_s(\alpha, u) &:= \{j \in \{1, \dots, n\} : \|q_j\| < \alpha_j\}, \\
\mathcal{B}(\alpha, u) &:= \{j \in \{1, \dots, n\} : \|q_j\| = \alpha_j,\ (\mathbb{K}u)_j = 0,\ \alpha_j > 0\}, \\
\mathcal{I}_0(\alpha, u) &:= \{j \in \{1, \dots, n\} : (\mathbb{K}u)_j \neq 0,\ \alpha_j = 0\}, \\
\mathcal{T}(\alpha, u) &:= \{j \in \{1, \dots, n\} : \|q_j\| = \alpha_j,\ (\mathbb{K}u)_j = 0,\ \alpha_j = 0\}.
\end{aligned}$$

We will omit the arguments in the set notation whenever they can be inferred from the context. Let us note that the condition in the strongly active set $\mathcal{A}_s$ implies a strict positive parameter $\alpha_j > 0$ for this index set.

The constraint qualification condition theorem 3.1 makes use of key fundamental definitions from Mordukhovich's generalized calculus. In particular, the Mordukhovich normal cone to the graph of the mutifunction Q. For the reader's convenience we will lay some basic concepts in variational geometry that will be used through this section. For a more rigorous review refer to, e.g., [27]

Definition 3.1 (Bouligand Tangent Cone)**.** Let $C \subset \mathbb{R}^n$ and $\bar{x} \in clC$. The Bouligand Tangent Cone to C at x is defined as

$$T_C(\bar{x}) := \{d \in \mathbb{R}^n : \exists t_k \to 0, \exists x_k \in C : \frac{x_k - \bar{x}}{t_k} \to d\}. \tag{13}$$

Definition 3.2 (Fréchet Normal Cone)**.** Let $C \subset \mathbb{R}^n$ and $\bar{x} \in clC$. The Fréchet Normal Cone is defined as the polar to $T_C(\bar{x})$, i.e.,

$$N^F_C(\bar{x}) = \left\{ d \in \mathbb{R}^n : \limsup_{x \xrightarrow{C} \bar{x}} \frac{\langle d, x - \bar{x}\rangle}{\|x - \bar{x}\|} \leq 0 \right\} = [T_C(\bar{x})]^\circ. \tag{14}$$

Definition 3.3 (Mordukhovich Normal Cone)**.** Let $C \subset \mathbb{R}^n$ and $\bar{x} \in clC$. The Mordukhovich Normal Cone is defined as

$$N^M_C(\bar{x}) := \{d \in \mathbb{R}^n : \exists x_k \to \bar{x}, d_k \in N^F_C(x_k) \text{ s.t. } d_k \to d\}. \tag{15}$$

Remark 1. *If the set C in definition 3.3 is convex, $N^M_C(\bar{x})$ amounts to the standard normal cone to C at $\bar{x}$. Otherwise, this cone is in general non-convex.*

In the following lemmata we obtain the Bouligand tangent cone, the Fréchet normal cone and the Mordukhovich normal cone to the graph of the multifunction Q.

Lemma 3.2. The Bouligand tangent cone to the graph of Q, described in eq. (12), is given by

$$
T_{\operatorname{gph} Q}(\alpha, u, \mathbb{K}^\top q) =
$$
$$
\left\{ (\delta_\alpha, \delta_u, \mathbb{K}^\top \delta_q) : \begin{cases}
(\delta_q)_j - (\delta_\alpha)_j \dfrac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|} - \alpha_j T_j (\mathbb{K}\delta_u)_j = 0, & \text{if } j \in \mathcal{I}, \\
(\mathbb{K}\delta_u)_j = 0, & \text{if } j \in \mathcal{A}_s, \\
\left.\begin{array}{l} (\mathbb{K}\delta_u)_j = 0, \langle (\delta_q)_j, q_j \rangle \le \alpha_j (\delta_\alpha)_j \vee \\ (\mathbb{K}\delta_u)_j = \tilde{c} q_j (\tilde{c} \ge 0), \langle (\delta_q)_j, q_j \rangle = \alpha_j (\delta_\alpha)_j \end{array}\right\} & \text{if } j \in \mathcal{B}, \\
(\delta_q)_j - (\delta_\alpha)_j \dfrac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|} = 0, \ (\delta_\alpha)_j \ge 0 & \text{if } j \in \mathcal{I}_0, \\
\left.\begin{array}{l} (\delta_\alpha)_j \ge 0, (\mathbb{K}\delta_u)_j \in \mathbb{R}^2 \backslash \{0\}, (\delta_q)_j - (\delta_\alpha)_j \dfrac{(\mathbb{K}\delta_u)_j}{\|(\mathbb{K}\delta_u)_j\|} = 0 \vee \\ (\delta_\alpha)_j \ge 0, (\mathbb{K}\delta_u)_j = 0, \|(\delta_q)_j\| - (\delta_\alpha)_j \le 0 \end{array}\right\} & \text{if } j \in \mathcal{T},
\end{cases} \right\}
$$

where

$$
T_j(\mathbb{K}v)_j = \frac{(\mathbb{K}v)_j}{\|(\mathbb{K}u)_j\|} - \frac{(\mathbb{K}u)_j (\mathbb{K}u)_j^\top (\mathbb{K}v)_j}{\|(\mathbb{K}u)_j\|^3}, \text{ for } v \in \mathbb{R}^n.
$$

Proof. The tangent cone to the graph of the multifunction Q is defined as

$$
T_{\operatorname{gph} Q}(\alpha, u, \mathbb{K}^\top q) = \{(\delta_\alpha, \delta_u, \mathbb{K}^\top \delta_q) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m : \exists t_k \to 0, (\alpha_k, u_k, \mathbb{K}^\top q_k) \in \operatorname{gph} Q :
$$
$$
\frac{1}{t_k}((\alpha_k, u_k, \mathbb{K}^\top q_k) - (\alpha, u, \mathbb{K}^\top q)) \to (\delta_\alpha, \delta_u, \mathbb{K}^\top \delta_q)\}.
$$

In order to calculate this cone, we split the analysis into different cases, according to the definition of the multifunction Q.

Case 1: $j \in \mathcal{I}$. In this index set the dual variable can be uniquely characterized. According to eq. (11), the following equation is fulfilled:

$$
h_j(\alpha, u, \mathbb{K}^\top q) = q_j - \alpha_j \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|} = 0. \tag{16}
$$

Using Lyusternik's theorem [19, Theorem 4.21], the $j-$th component of the tangent direction then satisfies

$$
(\delta_q)_j - (\delta_\alpha)_j \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|} - \alpha_j T_j (\mathbb{K}\delta_u)_j = 0.
$$

Case 2: $j \in \mathcal{A}_s$. In this index set we know $\|q_j\| < \alpha_j$. Therefore, this point can only be approximated by taking sequences in the strongly active set. For n sufficiently large we then take sequences such that $(\mathbb{K}u_k)_j = 0$, $\|(q_k)_j\| < (\alpha_k)_j$. Taking the limit in $(\mathbb{K}u_k)_j = 0$, as $k \to \infty$, yields $(\mathbb{K}\delta_u)_j = 0$. For the dual variable, let us take the sequence $(q_k)_j = q_j + t_k d$ with arbitrary $d \in \mathbb{R}^2$. It then follows that

$$
(\delta_q)_j = \lim_{t_k \to 0} \frac{(q_k)_j - q_j}{t_k} = d.
$$

Since we took an arbitrary direction d it yields $(\delta_q)_j \in \mathbb{R}^2$. Similarly, we get that $(\delta_\alpha)_j \in \mathbb{R}$.

Case 3: $j \in \mathcal{B}$. There are three possible approximations to a point in this index set, via inactive, strongly active or biactive sequences. Since $\alpha_j > 0$ in all these three cases, we can approximate it by sequences $(\alpha_k)_j < \alpha_j$ or $(\alpha_k)_j > \alpha_j$. Consequently, we get $(\delta_\alpha)_j \in \mathbb{R}$.

Now, if we approach a biactive point using a sequence in the inactive set, this sequence satisfies $(\mathbb{K}u_k)_j \neq 0$. Taking in particular $(\mathbb{K}u_k)_j = (\mathbb{K}u)_j + t_k(\mathbb{K}\delta_{u_k})_j = t_k(\mathbb{K}\delta_{u_k})_j$, then the sequence of dual variables has the following form

$$
(q_k)_j = (\alpha_k)_j \frac{(\mathbb{K}u_k)_j}{\|(\mathbb{K}u_k)_j\|} = (\alpha_k)_j \frac{(\mathbb{K}\delta_{u_k})_j}{\|(\mathbb{K}\delta_{u_k})_j\|}. \tag{17}
$$

Furthermore, considering the following product

$$\langle (q_k)_j, (\mathbb{K}\delta_{u_k})_j\rangle = \left\langle (\alpha_k)_j \frac{(\mathbb{K}\delta_{u_k})_j}{\|(\mathbb{K}\delta_{u_k})_j\|}, (\mathbb{K}\delta_{u_k})_j \right\rangle = (\alpha_k)_j \|(\mathbb{K}\delta_{u_k})_j\|,$$

and taking the limit as $k \to \infty$ we get $\langle q_j, (\mathbb{K}\delta_u)_j\rangle = \alpha\|(\mathbb{K}\delta_u)_j\|$. Recalling that in this index set $\|q_j\| = \alpha_j > 0$, we know both vectors are collinear, i.e.,

$$(\mathbb{K}\delta_u)_j = \tilde{c} q_j, \quad \text{for some } \tilde{c} \geq 0. \tag{18}$$

Using that $\|q_j\| = \alpha_j$, the following product holds

$$\begin{aligned}\left\langle \frac{(q_k)_j - q_j}{t_k}, q_j \right\rangle &= \frac{1}{t_k}(\langle (q_k)_j, q_j\rangle - \langle q_j, q_j\rangle) = \frac{1}{t_k}(\langle (q_k)_j, q_j\rangle - \langle (q_k)_j, (q_k)_j\rangle + \langle (q_k)_j, (q_k)_j\rangle - \alpha_j^2),\\ &= \left\langle (q_k)_j, \frac{q_j - (q_k)_j}{t_k} \right\rangle + \frac{(\alpha_k)_j^2 - \alpha_j^2}{t_k}.\end{aligned}$$

The last equality holds since from eq. (17), we get the product $\langle (q_k)_j, (q_k)_j\rangle = (\alpha_k)_j^2$. Taking the limit as $t_k \to 0$ we get

$$\langle (\delta_q)_j, q_j\rangle = \alpha_j (\delta_\alpha)_j. \tag{19}$$

Now, if the approximation is done through a sequence of strongly active points, we know the sequence satisfies $(\mathbb{K}u_k)_j = 0$ and $\|q_k\| < (\alpha_k)_j$. In this case we know $(\mathbb{K}\delta_u)_j = 0$ and, using the Cauchy-Schwarz inequality, we get

$$\left\langle \frac{(q_k)_j - q_j}{t_k}, q_j \right\rangle \leq \frac{\alpha_j}{t_k}(\|(q_k)_j\| - \|q_j\|) < \alpha_j \left(\frac{(\alpha_k)_j - \alpha_j}{t_k} \right),$$

which implies that $\langle (\delta_q)_j, q_j\rangle \leq \alpha_j(\delta_\alpha)_j$.

Finally, approximating through a biactive set sequence, we know again $(\mathbb{K}\delta_u)_j = 0$ and, estimating the product

$$\left\langle \frac{(q_k)_j - q_j}{t_k}, q_j \right\rangle \leq \frac{\alpha_j}{t_k}(\|(q_k)_j\| - \|q_j\|) = \alpha_j \left(\frac{(\alpha_k)_j - \alpha_j}{t_k} \right),$$

we get

$$\langle (\delta_q)_j, q_j\rangle \leq \alpha_j (\delta_\alpha)_j. \tag{20}$$

Case 4: $j \in \mathcal{I}_0$. We can approximate a point in the zero-inactive set by sequences in the inactive set. Therefore, considering a sequence $(\mathbb{K}u_k)_j = (\mathbb{K}u)_j + t_k v$ with $v \in \mathbb{R}^2$ arbitrary we have

$$(\mathbb{K}\delta_u)_j = \lim_{t_k \to 0} \frac{(\mathbb{K}u_k)_j - (\mathbb{K}u)_j}{t_k} = v. \tag{21}$$

Since we took $v \in \mathbb{R}^2$ arbitrary, it follows that $(\mathbb{K}\delta_u) \in \mathbb{R}^2$. Furthermore, since $q_j = \alpha_j((\mathbb{K}u)_j/\|(\mathbb{K}u)_j\|)$ and $\alpha_j = 0$, then $q_j = 0$ in this index set. Considering the sequence $(\alpha_k)_j = \alpha_j + t_k h$ we obtain

$$(\delta_q)_j = \lim_{t_k \to 0} \frac{(q_k)_j}{t_k} = \lim_{t_k \to 0} h \frac{(\mathbb{K}u_k)_j}{\|(\mathbb{K}u_k)_j\|} = (\delta_\alpha)_j \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}. \tag{22}$$

Since $\alpha_j = 0$, the only valid approximations are the ones coming from positive elements, thus $(\delta_\alpha)_j \geq 0$. Another possible approximation can be done through zero-inactive points, meaning $(q_k)_j = 0$ and $(\alpha_k)_j = 0$. This case implies $(\delta_q)_j = 0$, $(\delta_\alpha)_j = 0$ and $(\mathbb{K}\delta_u)_j \in \mathbb{R}^2$, which is a particular case of eq. (22).

Case 5: $j \in \mathcal{T}$. There are four ways to approach a triactive point. As in the zero-inactive case, all valid approximations come from $(\alpha_k)_j \geq 0$, which again implies $(\delta_\alpha)_j \geq 0$. Similarly to the zero-inactive case eq. (21), we also get $(\mathbb{K}\delta_u)_j \in \mathbb{R}^2$.

Approximating through a sequence in the inactive set, we obtain

$$(\delta_q)_j = \lim_{t_k \to 0} \frac{1}{t_k} \left((\alpha_k)_j \frac{(\mathbb{K}u_k)_j}{\|(\mathbb{K}u_k)_j\|} \right) = (\delta_\alpha)_j \frac{(\mathbb{K}\delta_u)_j}{\|(\mathbb{K}\delta_u)_j\|}. \tag{23}$$

Likewise, the approximation can be made using zero-inactive points. In this case $(\mathbb{K}\delta_u)_j \neq 0$, $(q_k)_j = 0$ and $(\alpha_k)_j = 0$. From this sequence we can derive $(\delta_q)_j = 0$, $(\delta_\alpha)_j = 0$ and $(\mathbb{K}\delta_u)_j \in \mathbb{R}^2$, which is included in eq. (23).

Moving forward, we can also approximate through strongly active points, i.e., $(\mathbb{K}u_k)_j = 0$ and $\|(q_k)_j\| < (\alpha_k)_j$. From this sequence we know $(\mathbb{K}\delta_u)_j = 0$ and $(\delta_\alpha)_j \geq 0$ and the dual variable will have the following form

$$\|(\delta_q)_j\| = \left\| \lim_{t_k \to 0} \frac{(q_k)_j}{t_k} \right\| = \lim_{t_k \to 0} \frac{1}{t_k} \|(q_k)_j\| \leq \lim_{t_k \to 0} \frac{1}{t_k} (\alpha_k)_j = (\delta_\alpha)_j, \tag{24}$$

yielding $\|(\delta_q)_j\| \leq (\delta_\alpha)_j$.

Finally, we consider an approximation through biactive points, meaning $(\mathbb{K}u_k)_j = 0$ and $\|(q_k)_j\| = (\alpha_k)_j$. We the obtain $(\mathbb{K}\delta_u)_j = 0$, $(\delta_\alpha)_j \geq 0$ and

$$\|(\delta_q)_j\| = \left\| \lim_{t_k \to 0} \frac{(q_k)_j}{t_k} \right\| = \lim_{t_k \to 0} \frac{1}{t_k} \|(q_k)_j\| = \lim_{t_k \to 0} \frac{1}{t_k} (\alpha_k)_j = (\delta_\alpha)_j,$$

which is a particular case of eq. (24). Consequently, $\|(\delta_q)_j\| = (\delta_\alpha)_j$.

$\square$

Lemma 3.3. The Fréchet normal cone to the graph of Q, described in eq. (12), is given by

$$N^F_{\operatorname{gph} Q}(\alpha, u, \mathbb{K}^\top q) = \left\{ (\vartheta, \mathbb{K}^\top \mu, p) : \begin{cases} \mu_j + \alpha_j T_j (\mathbb{K}p)_j = 0, & \text{if } j \in \mathcal{I}, \\ \vartheta_j + \dfrac{\langle (\mathbb{K}u)_j, (\mathbb{K}p)_j \rangle}{\|(\mathbb{K}u)_j\|} = 0, & \text{if } j \in \mathcal{I}, \\ \vartheta_j = 0, \ (\mathbb{K}p)_j = 0 & \text{if } j \in \mathcal{A}_s, \\ \vartheta_j + c\alpha_j = 0, (\mathbb{K}p)_j = cq_j (c \geq 0), \langle \mu_j, q_j \rangle \leq 0, & \text{if } j \in \mathcal{B}, \\ \vartheta_j + \dfrac{\langle (\mathbb{K}u)_j, (\mathbb{K}p)_j \rangle}{\|(\mathbb{K}u)_j\|} \leq 0, \quad \mu_j = 0, & \text{if } j \in \mathcal{I}_0, \\ \vartheta_j \leq 0, \ (\mathbb{K}p)_j = 0, \ \mu_j = 0, & \text{if } j \in \mathcal{T}. \end{cases} \right\} \tag{25}$$

Proof. Using the definition of the Fréchet normal cone for this problem

$$N^F_{\operatorname{gph} Q}(\alpha, u, \mathbb{K}^\top q) = \{ (\vartheta, \mathbb{K}^\top \mu, p) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m : \langle (\vartheta, \mathbb{K}^\top \mu, p), (\delta_\alpha, \delta_u, \mathbb{K}^\top \delta_q) \rangle \leq 0 \}.$$

Indeed, we can rewrite this inequality as

$$\sum_{j=1}^{n} ((\delta_\alpha)_j \vartheta_j + \langle (\mathbb{K}\delta_u)_j, \mu_j \rangle + \langle (\delta_q)_j, (\mathbb{K}p)_j \rangle) \leq 0.$$

Using this representation, along with the tangent cone characterization from theorem 3.2, we analyze the different cases according to their index set.

Case 1: $j \in \mathcal{I}$. Using the characterization of the elements in the tangent cone, we have

$$(\delta_\alpha)_j \vartheta_j + \langle (\mathbb{K}\delta_u)_j, \mu_j \rangle + \left\langle (\delta_\alpha)_j \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, (\mathbb{K}p)_j \right\rangle + \langle \alpha_j T_j (\mathbb{K}\delta_u)_j, (\mathbb{K}p)_j \rangle \leq 0.$$

Rearranging the terms and using the symmetry of T_j,

$$(\delta_\alpha)_j \left(\vartheta_j + \frac{\langle (\mathbb{K}u)_j, (\mathbb{K}p)_j \rangle}{\|(\mathbb{K}u)_j\|} \right) + \langle (\mathbb{K}\delta_u)_j, \mu_j + \alpha_j T_j (\mathbb{K}p)_j \rangle \leq 0.$$

Since $(\mathbb{K}\delta_u)_j \in \mathbb{R}^2$ and $(\delta_\alpha)_j \in \mathbb{R}$, it necessarily must hold

$$\vartheta_j + \frac{\langle (\mathbb{K}u)_j, (\mathbb{K}p)_j \rangle}{\|(\mathbb{K}u)_j\|} = 0, \qquad \mu_j + \alpha_j T_j (\mathbb{K}p)_j = 0.$$

Case 2: $j \in \mathcal{A}_s$. In this index set we know that $(\mathbb{K}\delta_u)_j = 0$, $(\delta_\alpha)_j \in \mathbb{R}$ and $(\delta_q)_j \in \mathbb{R}^2$. Consequently the product reads

$$(\delta_\alpha)_j \vartheta_j + \langle (\delta_q)_j, (\mathbb{K}p)_j \rangle \leq 0,$$

and we obtain that $(\mathbb{K}p)_j = 0$ and $\vartheta_j = 0$.

Case 3: $j \in \mathcal{B}$. In this index set there are two conditions on the normal directions. For the first one we take $(\mathbb{K}\delta_u)_j = 0$ and the cone inequality reads

$$(\delta_\alpha)_j \vartheta_j + \langle (\delta_q)_j, (\mathbb{K}p)_j \rangle \leq 0, \quad \forall (\delta_\alpha)_j, (\delta_q)_j \text{ s.t. } \langle (\delta_q)_j, q_j \rangle \leq \alpha_j (\delta_\alpha)_j. \tag{26}$$

Taking in particular, $(\delta_\alpha)_j = 0$ we get

$$\langle (\delta_q)_j, (\mathbb{K}p)_j \rangle \leq 0, \forall (\delta_q)_j \text{ s.t. } \langle (\delta_q)_j, q_j \rangle \leq 0.$$

Therefore, $(\mathbb{K}p)_j = cq_j$ with $c \geq 0$. Using this result in eq. (26) for the particular case $(\delta_\alpha)_j = \frac{1}{\alpha_j} \langle (\delta_q)_j, q_j \rangle$, we obtain

$$0 \geq (\delta_\alpha)_j \vartheta_j + \langle (\delta_q)_j, cq_j \rangle = (\delta_\alpha)_j (\vartheta_j + c\alpha_j),$$

from where $\vartheta_j + c\alpha_j = 0$ holds. The resulting cone reads

$$\vartheta_j + c\alpha_j = 0, \quad (\mathbb{K}p)_j = cq_j, \quad c \geq 0, \quad \mu_j \in \mathbb{R}^2. \tag{27}$$

For the second case we take $(\mathbb{K}\delta_u) = cq_j$ $(c \geq 0)$ in eq. (26),

$$(\delta_\alpha)_j \vartheta_j + \langle cq_j, \mu_j \rangle + \langle (\mathbb{K}p)_j, (\delta_q)_j \rangle \leq 0, \forall (\delta_\alpha)_j \in \mathbb{R},\ (\delta_q)_j \text{ s.t. } \langle (\delta_q)_j, q_j \rangle = \alpha_j (\delta_\alpha)_j, \tag{28}$$

Again, considering $(\delta_\alpha)_j = 0$ and $(\mathbb{K}\delta_u)_j = 0$, we get

$$\langle (\delta_q)_j, (\mathbb{K}p)_j \rangle \leq 0, \forall (\delta_q)_j \text{ s.t. } \langle (\delta_q)_j, q_j \rangle = 0.$$

Consequently, $(\mathbb{K}p)_j = cq_j$ with $c \in \mathbb{R}$. Using this result in eq. (28), while keeping $(\delta_\alpha)_j = 0$, yields $\tilde{c}\langle q_j, \mu_j \rangle \leq 0$. Thanks to the positiveness of $\tilde{c}$ we then get $\langle q_j, \mu_j \rangle \leq 0$. Now, applying all previous results in eq. (28) we get

$$0 \geq (\delta_\alpha)_j \vartheta_j + \langle cq_j, (\delta_q)_j \rangle = (\delta_\alpha)_j \vartheta_j + c\alpha_j (\delta_\alpha)_j, \forall (\delta_\alpha)_j \in \mathbb{R}, c \in \mathbb{R},$$

yielding $\vartheta_j \in \mathbb{R}$. Therefore, the resulting cone for the second case reads

$$\vartheta_j \in \mathbb{R}, \quad (\mathbb{K}p)_j = cq_j, \quad c \in \mathbb{R}. \tag{29}$$

Finally, considering both cases we obtain

$$\vartheta_j + c\alpha_j = 0 \wedge \langle \mu_j, q_j \rangle \leq 0 \wedge (\mathbb{K}p)_j = cq_j \wedge c \geq 0.$$

Case 4: $j \in \mathcal{I}_0$. By using the characterization of the tangent cone in this index set, we have

$$(\delta_\alpha)_j \left(\vartheta_j + \frac{\langle (\mathbb{K}u)_j, (\mathbb{K}p)_j \rangle}{\|(\mathbb{K}u)_j\|} \right) + \langle (\mathbb{K}\delta_u)_j, \mu_j \rangle \leq 0.$$

This relationship must hold for all $(\mathbb{K}\delta_u)_j \in \mathbb{R}^2$ and $(\delta_\alpha)_j \geq 0$, which implies

$$\vartheta_j + \frac{\langle (\mathbb{K}u)_j, (\mathbb{K}p)_j \rangle}{\|(\mathbb{K}u)_j\|} \leq 0, \quad \langle (\mathbb{K}\delta_u)_j, \mu_j \rangle \leq 0.$$

Since $(\mathbb{K}\delta_u)_j \in \mathbb{R}^2$ we get $\mu_j = 0$.

Case 5: $j \in \mathcal{T}$. For the first case in this index set we know the elements of the tangent cone satisfy

$$(\delta_q)_j = (\delta_\alpha)_j \frac{(\mathbb{K}\delta_u)_j}{\|(\mathbb{K}\delta_u)_j\|}, \ (\delta_\alpha)_j \geq 0, \ (\mathbb{K}\delta_u)_j \in \mathbb{R}^2 \backslash \{0\}.$$

Replacing these terms into the normal cone inequality,

$$(\delta_\alpha)_j \left(\vartheta_j + \frac{\langle (\mathbb{K}p)_j, (\mathbb{K}\delta_u)_j \rangle}{\|(\mathbb{K}\delta_u)_j\|} \right) + \langle \mu_j, (\mathbb{K}\delta_u)_j \rangle \leq 0, \ \forall (\delta_\alpha)_j \geq 0, (\mathbb{K}\delta_u)_j \in \mathbb{R}^2 \backslash \{0\}.$$

In particular, for $(\delta_\alpha)_j = 0$, we get that $\langle \mu_j, (\mathbb{K}\delta_u)_j \rangle \leq 0$, for all $(\mathbb{K}\delta_u)_j \in \mathbb{R}^2 \backslash \{0\}$, which implies that $\mu_j = 0$. Moreover, thanks to the positiveness of $(\delta_\alpha)_j \geq 0$, we get

$$\vartheta_j + \frac{\langle (\mathbb{K}p)_j, (\mathbb{K}\delta_u)_j \rangle}{\|(\mathbb{K}\delta_u)_j\|} \leq 0, \forall (\mathbb{K}\delta_u)_j \in \mathbb{R}^2 \backslash \{0\}. \tag{30}$$

Testing this inequality with $\pm(\mathbb{K}\delta_u)_j$, we get $\vartheta_j \leq \langle (\mathbb{K}p)_j, (\mathbb{K}\delta_u)_j \rangle \leq -\vartheta_j$, which implies that the product $\langle (\mathbb{K}p)_j, (\mathbb{K}\delta_u)_j \rangle = 0$ for all $(\mathbb{K}\delta_u)_j \in \mathbb{R}^2 \backslash \{0\}$, consequently $(\mathbb{K}p)_j = 0$. Using this result in eq. (30) yields also $\vartheta_j \leq 0$.

Now, regarding the second condition, we know $\|(\delta_q)_j\| - (\delta_\alpha)_j \leq 0$, $(\mathbb{K}\delta_u)_j = 0$ and $(\delta_\alpha)_j \geq 0$. Since $(\mathbb{K}\delta_u)_j = 0$, it follows $\mu_j \in \mathbb{R}^2$. Using the normal cone inequality,

$$0 \geq (\delta_\alpha)_j \vartheta_j + \langle (\delta_q)_j, (\mathbb{K}p)_j \rangle \geq (\delta_\alpha)_j \vartheta_j - \|(\delta_q)_j\| \|(\mathbb{K}p)_j\| \geq (\delta_\alpha)_j \vartheta_j - (\delta_\alpha)_j \|(\mathbb{K}p)_j\|,$$

from where we derive $(\delta_\alpha)_j(\vartheta_j - \|(\mathbb{K}p)_j\|) \leq 0$. Along with the positivity of $(\delta_\alpha)_j$, this implies $\vartheta_j \leq \|(\mathbb{K}p)_j\|$. Finally, considering both conditions, the result is obtained.

$\square$

Lemma 3.4. The Mordukhovich normal cone to the graph of Q, described in eq. (12), is given by

$$N^M_{\operatorname{gph} Q}(\alpha, u, \mathbb{K}^\top q) = \left\{ (\vartheta, \mathbb{K}^\top \mu, p) : \begin{cases} \mu_j + \alpha_j T_j (\mathbb{K}p)_j = 0, & \text{if } j \in \mathcal{I}, \\ \vartheta_j + \dfrac{\langle (\mathbb{K}u)_j, (\mathbb{K}p)_j \rangle}{\|(\mathbb{K}u)_j\|} = 0, & \text{if } j \in \mathcal{I}, \\ \vartheta_j = 0, (\mathbb{K}p)_j = 0, & \text{if } j \in \mathcal{A}_s, \\ \left.\begin{array}{l} \vartheta_j = 0, (\mathbb{K}p)_j = 0, \ \vee \\ (\mathbb{K}p)_j = c q_j (c \in \mathbb{R}), \ \langle \mu_j, q_j \rangle = 0, \ \vee \\ \vartheta_j + c\alpha_j = 0, (\mathbb{K}p)_j = c q_j (c \geq 0), \langle \mu_j, q_j \rangle \leq 0. \end{array}\right\} & \text{if } j \in \mathcal{B}, \\ \vartheta_j + \dfrac{\langle (\mathbb{K}u)_j, (\mathbb{K}p)_j \rangle}{\|(\mathbb{K}u)_j\|} \leq 0, \quad \mu_j = 0, & \text{if } j \in \mathcal{I}_0, \\ \left.\begin{array}{l} \vartheta_j = 0, (\mathbb{K}p)_j = 0, \ \vee \\ \vartheta_j \leq \|(\mathbb{K}p)_j\|, \mu_j = 0, \ \vee \\ |\vartheta_j| \leq \|(\mathbb{K}p)_j\|, \langle \mu_j, (\mathbb{K}p)_j \rangle \leq 0 \end{array}\right\} & \text{if } j \in \mathcal{T}. \end{cases} \right\}$$

Proof. Let us recall the definition of the Mordukhovich normal cone for our problem

$$N^M_{\operatorname{gph} Q}(\alpha, u, \mathbb{K}^\top q) = \{(\vartheta, \mathbb{K}^\top \mu, p) : (\vartheta_k, \mathbb{K}^\top \mu_k, p_k) \in N^F_{\operatorname{gph} Q}(\alpha_k, u_k, \mathbb{K}^\top q_k) : \\ (\vartheta_k, \mathbb{K}^\top \mu_k, p_k) \to (\vartheta, \mathbb{K}^\top \mu, p), (\alpha_k, u_k, \mathbb{K}^\top q_k) \to (\alpha, u, \mathbb{K}^\top q)\}.$$

Considering limiting sequences to the inactive, strongly active and zero-inactive sets, the same directions as for the Fréchet normal cone are obtained. The differences lie in the biactive and triactive sets, where several approximations may be considered.

Case 1: $j \in \mathcal{B}$. By taking approximation sequences in the inactive set, from theorem 3.3 we know

$$0 = (\mu_k)_j + (\alpha_k)_j \frac{(\mathbb{K}p_k)_j}{\|(\mathbb{K}u_k)_j\|} - (\alpha_k)_j \frac{(\mathbb{K}u_k)\langle((\mathbb{K}u_k)_j), (\mathbb{K}p_k)_j\rangle}{\|(\mathbb{K}u_k)_j\|^3}. \tag{31}$$

Multiplying eq. (31) with $(\mathbb{K}p_k)_j$ yields

$$\langle(\mu_k)_j, (\mathbb{K}p_k)_j\rangle = (\alpha_k)_j \frac{\langle(\mathbb{K}u_k)_j, (\mathbb{K}p_k)_j\rangle^2}{\|(\mathbb{K}u_k)_j\|^3} - (\alpha_k)_j \frac{\langle(\mathbb{K}p_k)_j, (\mathbb{K}p_k)_j\rangle}{\|(\mathbb{K}u_k)_j\|}$$

Again, multiplying by $(\alpha_k)_j\|(\mathbb{K}u_k)_j\|$ on both sides and recalling that $(q_k)_j = (\alpha_k)_j \frac{(\mathbb{K}u_k)_j}{\|(\mathbb{K}u_k)_j\|}$, we get

$$(\alpha_k)_j\|(\mathbb{K}u_k)_j\|\langle(\mu_k)_j, (\mathbb{K}p_k)_j\rangle = \langle(q_k)_j, (\mathbb{K}p_k)_j\rangle^2 - (\alpha_k)_j^2\|(\mathbb{K}p_k)_j\|^2.$$

Taking the limit as $k \to \infty$ and recalling $(\alpha_k)_j = \|(q_k)_j\|$ in this index set, we obtain

$$\langle q_j, (\mathbb{K}p)_j\rangle^2 = \|q_j\|^2\|(\mathbb{K}p)_j\|^2,$$

which implies that $(\mathbb{K}p)_j = cq_j (c \in \mathbb{R})$. Now, multiplying eq. (31) with $(q_k)_j$ we get the following product

$$\begin{aligned}\langle(\mu_k)_j, (q_k)_j\rangle &= (\alpha_k)_j \frac{\langle(q_k)_j, (\mathbb{K}u_k)_j\rangle\langle(\mathbb{K}u_k)_j, (\mathbb{K}p_k)_j\rangle}{\|(\mathbb{K}u_k)_j\|^3} - (\alpha_k)_j \frac{\langle(q_k)_j, (\mathbb{K}p_k)_j\rangle}{\|(\mathbb{K}u_k)_j\|}, \\ &= (\alpha_k)_j^2 \frac{\langle(\mathbb{K}u_k)_j, (\mathbb{K}p_k)_j\rangle}{\|(\mathbb{K}u_k)_j\|^2} - (\alpha_k)_j^2 \frac{\langle(\mathbb{K}u_k)_j, (\mathbb{K}p_k)_j\rangle}{\|(\mathbb{K}u_k)_j\|^2} = 0.\end{aligned} \tag{32}$$

Taking the limit we get that $\langle\mu_j, q_j\rangle = 0$.

Regarding ϑ_j we have $(\alpha_k)_j(\vartheta_k)_j + \langle(q_k)_j, (\mathbb{K}p)_j\rangle = 0$. Taking the limit as $k \to \infty$ we obtain

$$0 = \alpha_j\vartheta_j + \langle q_j, (\mathbb{K}p)_j\rangle = \alpha_j\vartheta_j + c\|q_j\|^2, (c \in \mathbb{R}),$$

which implies that $\vartheta_j = -c\alpha_j \in \mathbb{R}$.

Finally, when taking the approximation through the strongly active and biactive sets, the cone directions coincide with the Fréchet normal ones.

Case 2: $j \in \mathcal{T}$. This index set can be approximated by sequences belonging either to the inactive, biactive, strongly active or zero-inactive sets. Considering strongly active sequences, $(\mathbb{K}p_k)_j = 0$ and $(\vartheta_k)_j = 0$. Taking the limit as $k \to \infty$ we get $(\mathbb{K}p)_j = 0$ and $\vartheta_j = 0$ as well.

Likewise, when taking biactive sequences we get $(\vartheta_k)_j + c(\alpha_k)_j = 0$, $(\mathbb{K}p_k)_j = c(q_k)_j (c \geq 0)$ and $\langle(\mu_k)_j, (q_k)_j\rangle \leq 0$. Again, taking the limit as $k \to \infty$ we get, since $q_j = 0$ and $\alpha_j = 0$, that $\mu_j \in \mathbb{R}^2$, $(\mathbb{K}p)_j = 0$ and $\vartheta_j = 0$.

Furthermore, taking sequences in the zero-inactive set we have $(\mu_k)_j = 0$, which implies that $\mu_j = 0$. Using the Cauchy-Schwarz inequality we get

$$0 \geq (\vartheta_k)_j + \frac{\langle(\mathbb{K}u_k)_j, (\mathbb{K}p_k)_j\rangle}{\|(\mathbb{K}u_k)_j\|} \geq (\vartheta_k)_j - \|(\mathbb{K}p_k)_j\|.$$

Taking the limit as $k \to \infty$ yields $\vartheta_j \leq \|(\mathbb{K}p)_j\|$.

When taking inactive sequences, we know that

$$(\mu_k)_j + (\alpha_k)_j \left(\frac{I}{\|(\mathbb{K}u_k)_j\|} - \frac{(\mathbb{K}u_k)_j(\mathbb{K}u_k)_j^\top}{\|(\mathbb{K}u_k)_j\|^3} \right) (\mathbb{K}p_k)_j = 0, \tag{33}$$

$$(\vartheta_k)_j + \frac{\langle (\mathbb{K}u_k)_j, (\mathbb{K}p_k)_j \rangle}{\|(\mathbb{K}u_k)_j\|} = 0. \tag{34}$$

Applying Cauchy-Schwarz in eq. (34), yields $|(\vartheta_k)_j| \leq \|(\mathbb{K}p_k)_j\|$. Now, multiplying eq. (33) with $(\mathbb{K}p_k)_j$ we obtain

$$\langle (\mu_k)_j, (\mathbb{K}p_k)_j \rangle \leq (\alpha_k)_j \left(\frac{\|(\mathbb{K}p_k)_j\|^2 \|(\mathbb{K}u_k)_j\|^2}{\|(\mathbb{K}u_k)_j\|^3} - \frac{\|(\mathbb{K}p_k)_j\|^2}{\|(\mathbb{K}u_k)_j\|} \right) = 0. \tag{35}$$

Taking the limit as $k \to \infty$ we get $\langle \mu_j, (\mathbb{K}p)_j \rangle \leq 0$, finishing the proof.

□

Theorem 3.5 (M-Stationarity)**.** Let $J : \mathbb{R}^m \to \mathbb{R}$ be continuously differentiable, $\phi : \mathbb{R}^m \to \mathbb{R}$ twice continuously differentiable and strongly convex, and (α^*, u^*, q^*) be a local solution to eq. (7). Then there exist KKT multipliers $(\vartheta, \mathbb{K}^\top \mu, p)$ such that

$$\begin{aligned}
\phi'(u^*) + \mathbb{K}^\top q^* &= 0, && && \text{(36a)}\\
\langle q_j^*, (\mathbb{K}u^*)_j \rangle - \alpha_j^* \|(\mathbb{K}u^*)_j\| &= 0, && \forall j = 1, \dots, n, && \text{(36b)}\\
\|q_j^*\| &\leq \alpha_j^*, && \forall j = 1, \dots, n, && \text{(36c)}\\
\phi''(u^*)^\top p - \mathbb{K}^\top \mu - \nabla J(u^*) &= 0, && && \text{(36d)}\\
\vartheta + \rho &= 0, && && \text{(36e)}\\
0 \leq \alpha^* \perp \rho &\geq 0, && && \text{(36f)}\\
(\vartheta, \mathbb{K}^\top \mu, p) &\in N^M_{\operatorname{gph} Q}(\alpha^*, u^*, \mathbb{K}^\top q^*) && && \text{(36g)}
\end{aligned}$$

Proof. Referring to theorem 3.1 let us take $F_1(\alpha, u) = \phi'(u) \in \mathbb{R}^m$ and $F_2(\alpha, u) = (\alpha, u) \in \mathbb{R}^n_+ \times \mathbb{R}^m$. Existence of the KKT multipliers is guaranteed if the following constraint qualification condition holds for $(\vartheta, \mathbb{K}^\top \mu, p) \in N^M_{\operatorname{gph} Q}(\alpha^*, u^*, \mathbb{K}^\top q^*)$

$$\begin{bmatrix} I & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I & -\phi''(u^*)^\top \end{bmatrix} \begin{bmatrix} \vartheta \\ \mathbb{K}^\top \mu \\ p \end{bmatrix} \in -N^M_{\mathbb{R}^n_+}(\alpha^*) \times \{0\} \text{ implies } \vartheta = 0,\ \mathbb{K}^\top \mu = 0,\ p = 0. \tag{37}$$

Recalling remark 1 and using the expression of the Mordukhovich normal cone $N^M_{R^n_+}(\alpha^*) = N_{R^n_+}(\alpha^*) = \{v \in \mathbb{R}^2 : \langle v, \alpha^* \rangle = 0,\ v \leq 0\}$, condition eq. (37) can also be written as

$$\mathbb{K}^\top \mu - \phi''(u^*)^\top p = 0, \tag{38}$$

$$\langle \alpha^*, \vartheta \rangle = 0, \tag{39}$$

$$\vartheta \geq 0. \tag{40}$$

Let us take $(\vartheta, \mathbb{K}^\top \mu, p) \in N^M_{\operatorname{gph} Q}(\alpha^*, u^*, \mathbb{K}^\top q^*)$ and let us multiply eq. (38) by p on the left. Recalling $(\mathbb{K}p)_j = 0$ in $\mathcal{A}_s$ and $\mu_j = 0$ in $\mathcal{I}_0$, we have for each remaining index set

$$\begin{aligned}
\langle p, \phi''(u^*)^\top p \rangle = \langle p, \mathbb{K}^\top \mu \rangle &= \sum_{j \in \mathcal{I}} \langle \mu_j, (\mathbb{K}p)_j \rangle + \sum_{j \in \mathcal{B}} \langle \mu_j, (\mathbb{K}p)_j \rangle + \sum_{j \in \mathcal{B}_0} \langle \mu_j, (\mathbb{K}p)_j \rangle, \\
&= \sum_{j \in \mathcal{I}} -\alpha_j \langle (\mathbb{K}p)_j, T_j (\mathbb{K}p)_j \rangle + \sum_{j \in \mathcal{B}} c \underbrace{\langle \mu_j, q_j \rangle}_{\leq 0} + \sum_{j \in \mathcal{B}_0} \underbrace{\langle \mu_j, (\mathbb{K}p)_j \rangle}_{\leq 0} \leq 0,
\end{aligned}$$

where we used the characterization of the Mordukhovich normal cone. Furthermore, using the strong convexity of the function ϕ we have $\langle p, \phi''(u^*)^\top p \rangle \geq 0$. Both inequalities imply $p = 0$ and, according to eq. (38), it also yields $\mathbb{K}^\top \mu = 0$. Moreover, if we consider the index set $\mathcal{I} \cup \mathcal{A}_s \cup \mathcal{B}$, we know in all these sets $\alpha_j^* > 0$, and therefore, to satisfy equation eq. (39) it must hold $\vartheta_j = 0$. Since $p = 0$ in $\mathcal{T}$ we know $\vartheta_j = 0$ or $\vartheta_j \leq \|(\mathbb{K}p)_j\|$ for this index set, in both cases it leads to $\vartheta_j = 0$. In $\mathcal{I}_0$, we have $\vartheta_j \leq -\frac{\langle(\mathbb{K}u)_j,(\mathbb{K}p)_j\rangle}{\|(\mathbb{K}u)_j\|} = 0$ and eq. (40) yields $\vartheta_j = 0$. Therefore, $\vartheta_j = 0$ for all j. Consequently, the existence of multipliers is guaranteed and there exists a vector $\rho \in N^M_{R^n_+}(\alpha^*)$ and KKT multipliers $(\vartheta, \mathbb{K}^\top \mu, p) \in N^M_{\operatorname{gph} Q}(\alpha^*, u^*, \mathbb{K}^\top q^*)$ such that

$$0 = \nabla_u J(u^*) + (\nabla_u F_2(\alpha, u))^\top \begin{bmatrix} \vartheta \\ \mathbb{K}^\top \mu \end{bmatrix} - (\nabla_u F_1(\alpha, u))^\top p, \tag{41}$$

$$0 = (\nabla_\alpha F_2(\alpha, u))^\top \begin{bmatrix} \vartheta \\ \mathbb{K}^\top \mu \end{bmatrix} - (\nabla_\alpha F_1(\alpha, u))^\top p + \rho. \tag{42}$$

□

4 Bouligand Stationarity

In this section we will study the Bouligand stationarity condition for eq. (4). With this goal in mind, let us introduce the solution operator for the lower-level problem $S : \mathbb{R}^n_+ \ni \alpha \to u \in \mathbb{R}^m$ that maps each parameter $\alpha \in \mathbb{R}^n_+$ to the corresponding reconstruction $u \in \mathbb{R}^n$. If this mapping is bijective, we can make use of it to formulate eq. (4) as a reduced optimization problem

$$\min_{\alpha \in \mathbb{R}^n_+} \quad j(\alpha) := J(S(\alpha)). \tag{43}$$

Furthermore, if the solution operator is Bouligand (B)-differentiable, we can make use of the chain rule for B-differentiable functions to conclude that the composite mapping J, as a function of α, is B-differentiable as well. In this case, its directional derivative in a direction h is given by

$$j'(\alpha; h) = \langle \nabla J(u), S'(\alpha; h) \rangle, \tag{44}$$

where $S'(\alpha; h)$ is the directional derivative of the solution operator in direction h. Moreover, if α^* is a local optimal solution and $u^* = S(\alpha^*)$ its corresponding reconstruction, then it satisfies the following necessary condition

$$j'(\alpha^*; \alpha - \alpha^*) = \langle \nabla J(u^*), S'(\alpha^*; \alpha - \alpha^*) \rangle \geq 0, \ \forall \alpha \in \mathbb{R}^n_+. \tag{45}$$

A point α^* satisfying the necessary condition eq. (45) is called *Bouligand (B)-stationary.* This type of stationarity condition is based on the tangent cone to our feasible parameter set and can be interpreted as the counterpart of the implicit programming approach in the discussion of finite-dimensional MPECs, see [23, Lemma 4.2.5].

To obtain such B-stationarity condition for our problem, a sensitivity analysis of the solution mapping must be carried out, to prove that it is indeed Bouligand differentiable, i.e., locally Lipschitz continuous and directionally differentiable. Let us recall that in Section 2 we already argued the existence and uniqueness of the solution to the lower-level problem eq. (4b), implying that $S : \mathbb{R}^n_+ \to \mathbb{R}^m$ is singled valued.

Theorem 4.1. The solution operator for the lower-level problem eq. (4b) $S : \mathbb{R}^n_+ \ni \alpha \to u \in \mathbb{R}^m$ is Lipschitz continuous.

Proof. Thanks to Theorem 2.1, we know the lower-level problem has a unique solution. Moreover, $\alpha_1, \alpha_2 \in \mathbb{R}^n_+$ and its corresponding solutions u_1, u_2 satisfy

$$\langle \phi'(u_1), v - u_1 \rangle + \sum_{j=1}^n (\alpha_1)_j \|(\mathbb{K}v)_j\| - \sum_{j=1}^n (\alpha_1)_j \|(\mathbb{K}u_1)_j\| \geq 0, \ \forall v \in \mathbb{R}^m$$

$$\langle \phi'(u_2), w - u_2 \rangle + \sum_{j=1}^n (\alpha_2)_j \|(\mathbb{K}w)_j\| - \sum_{j=1}^n (\alpha_2)_j \|(\mathbb{K}u_2)_j\| \geq 0, \ \forall w \in \mathbb{R}^m.$$

Taking in particular $v = u_2$ and $w = u_1$ and adding the inequalities, it yields

$$\langle \phi'(u_2) - \phi'(u_1), u_2 - u_1 \rangle \leq \sum_{j=1}^{n} ((\alpha_1)_j - (\alpha_2)_j)(\|(\mathbb{K}u_2)_j\| - \|(\mathbb{K}u_1)_j\|),$$

Moreover, given that ϕ is strongly convex and using the Cauchy-Schwarz inequality, it yields

$$\begin{aligned} c\|u_2 - u_1\|^2 &\leq \sum_{j=1}^{n} (\alpha_{2,j} - \alpha_{1,j})\|(\mathbb{K}(u_2 - u_1))_j\| \leq \|\alpha_2 - \alpha_1\| \sum_{j=1}^{n} \|(\mathbb{K}(u_2 - u_1))_j\|, \\ &\leq \|\alpha_2 - \alpha_1\|\|\mathbb{K}\|\|u_2 - u_1\|, \end{aligned}$$

where $\|\mathbb{K}\|$ is the operator norm of the linear operator $\mathbb{K}$. □

4.1 Directional Differentiability

Now, we are interested in the differentiability properties of the solution operator for the lower-level problem eq. (4b). This will require a sensitivity analysis of the solution operator with respect to the regularization parameter. By taking a perturbed regularization parameter α^t in the primal-dual formulation for the lower-level problem eq. (6) such that $\alpha_j^t = \alpha_j + th_j \geq 0$ we get the following perturbed lower-level problem

$$\phi'(u^t) + \mathbb{K}^\top q^t = 0, \tag{46a}$$

$$\langle q_j^t, (\mathbb{K}u^t)_j \rangle - (\alpha_j + th_j)\|(\mathbb{K}u^t)_j\| = 0, \quad \forall j = 1, \ldots, n, \tag{46b}$$

$$\|q_j^t\| - (\alpha_j + th_j) \leq 0, \quad \forall j = 1, \ldots, n. \tag{46c}$$

Thanks to the local Lipschitz continuity of the solution operator proved in theorem 4.1 and the boundedness of q^t, see eq. (46), there exist a subsequence, denoted the same, so that $q^t \to \tilde{q} \in \mathbb{R}^{n \times 2}$, to some $\tilde{q}$. Additionally, we can guarantee the existence of a subsequence of u^t, denoted w.l.o.g. with the same symbol, satisfying the following limit

$$\lim_{t \to 0} \frac{u^t - u}{t} \to \eta \in \mathbb{R}^m. \tag{47}$$

Theorem 4.2. The limit described in eq. (47) satisfies $\eta \in \mathcal{C} = \mathcal{C}(\alpha, u)$ where

$$\mathcal{C}(\alpha, u) := \left\{ v \in \mathbb{R}^n : \begin{cases} (\mathbb{K}v)_j = 0, & \forall j \in \mathcal{A}_s, \\ \langle q_j, (\mathbb{K}v)_j \rangle = \alpha_j \|(\mathbb{K}v)_j\|, & \forall j \in \mathcal{B}. \end{cases} \right\} \tag{48}$$

Proof. By adding the complementarity relationships in eqs. (6) and (46), and dividing by t, we get

$$\left\langle \frac{q_j^t - q_j}{t}, (\mathbb{K}u)_j \right\rangle + \left\langle q_j^t, \frac{(\mathbb{K}u^t)_j - (\mathbb{K}u)_j}{t} \right\rangle - \alpha_j \left(\frac{\|(\mathbb{K}u^t)_j\| - \|(\mathbb{K}u)_j\|}{t} \right) - h_j \|(\mathbb{K}u^t)_j\| = 0.$$

For $j \in \mathcal{A}_s \cup \mathcal{B}$, taking the limit as $t \to 0$ and using the boundedness of the sequence q^t along with the Bouligand differentiability of the Euclidean norm, it yields

$$\langle \tilde{q}_j, (\mathbb{K}\eta)_j \rangle - \alpha_j \|(\mathbb{K}\eta)_j\| - h_j \|(\mathbb{K}u)_j\| = 0.$$

Since $(\mathbb{K}u)_j = 0$ for $j \in \mathcal{A}_s \cup \mathcal{B}$, we get that

$$\langle \tilde{q}_j, (\mathbb{K}\eta)_j \rangle - \alpha_j \|(\mathbb{K}\eta)_j\| = 0.$$

Moreover, for $j \in \mathcal{A}_s$, recalling $\alpha_j > 0$ in this index set, we get

$$\alpha_j \|(\mathbb{K}\eta)_j\| = \langle \tilde{q}_j, (\mathbb{K}\eta)_j \rangle \leq \|\tilde{q}_j\| \|(\mathbb{K}\eta)_j\| < \alpha_j \|(\mathbb{K}\eta)_j\|,$$

which only holds if $(\mathbb{K}\eta)_j = 0$ in this index set, finishing the proof. □

Remark 2. *If q^1 and q^2 are two different slack variables associated with the solution u in eq. (6), then the two sets*

$$\mathcal{C}_i := \left\{ v \in \mathbb{R}^n : \begin{cases} (\mathbb{K}v)_j = 0, & \text{if } \|q^i_j\| < \alpha_j, \\ \langle q^i_j, (\mathbb{K}v)_j \rangle = \alpha_j \|(\mathbb{K}v)_j\|, & \text{if } \ (\mathbb{K}u)_j = 0,\ \alpha_j > 0,\ \|q^i_j\| = \alpha_j. \end{cases} \right\},$$

coincide, since $\mathbb{K}^\top q^1 = -\phi'(u) = \mathbb{K}^\top q^2$. As a consequence, the set $\mathcal{C}(\alpha, u)$ does not depend on the slack variable, only on the solution u and the parameter α.

Lemma 4.3. The cone $\mathcal{C}(\alpha, u)$ can alternatively be written as

$$\mathcal{C}(\alpha, u) = \left\{ v \in \mathbb{R}^n : \langle \mathbb{K}^\top q, v \rangle \geq \sum_{j \in \mathcal{I}} \left\langle \alpha_j \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, (\mathbb{K}v)_j \right\rangle + \sum_{j \in \mathcal{B}} \alpha_j \|(\mathbb{K}v)_j\| \right\} \tag{49}$$

Proof. Let us denote the set in eq. (49) as $\mathcal{M}$. Taking $v \in \mathcal{C}$, as in eq. (48), and using its definition, we obtain

$$\begin{aligned} \langle \mathbb{K}^\top q, v \rangle &= \sum_{j \in \mathcal{I}} \langle q_j, (\mathbb{K}v)_j \rangle + \sum_{j \in \mathcal{A}_s} \langle q_j, (\mathbb{K}v)_j \rangle + \sum_{j \in \mathcal{B}} \langle q_j, (\mathbb{K}v)_j \rangle, \\ &= \sum_{j \in \mathcal{I}} \left\langle \alpha_j \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, (\mathbb{K}v)_j \right\rangle + \sum_{j \in \mathcal{A}_s} \langle q_j, \underbrace{(\mathbb{K}v)_j}_{=0} \rangle + \sum_{j \in \mathcal{B}} \alpha_j \|(\mathbb{K}v)_j\|, \end{aligned}$$

and, consequently, $\mathcal{C} \subset \mathcal{M}$.

To prove the reverse inclusion, let us take $v \in \mathcal{M}$. For $j \in \mathcal{A}_s \cup \mathcal{B}$ the following relation holds true

$$\sum_{j \in \mathcal{B}} \alpha_j \|(\mathbb{K}v)_j\| \leq \sum_{j \in \mathcal{A}_s \cup \mathcal{B}} \langle q_j, (\mathbb{K}v)_j \rangle \leq \sum_{j \in \mathcal{A}_s \cup \mathcal{B}} \alpha_j \|(\mathbb{K}v)_j\|. \tag{50}$$

where we used the Cauchy-Schwarz inequality and $\|q_j\| \leq \alpha_j$. Regarding the left inequality in eq. (50) we know

$$\sum_{j \in \mathcal{A}_s \cup \mathcal{B}} \langle q_j, (\mathbb{K}v)_j \rangle - \sum_{j \in \mathcal{B}} \alpha_j \|(\mathbb{K}v)_j\| \geq 0, \tag{51}$$

using the Cauchy-Schwarz inequality we also know

$$\sum_{j \in \mathcal{A}_s \cup \mathcal{B}} \langle q_j, (\mathbb{K}v)_j \rangle - \sum_{j \in \mathcal{A}_s \cup \mathcal{B}} \alpha_j \|(\mathbb{K}v)_j\| \leq 0. \tag{52}$$

Multiplying eq. (51) by -1 and adding it to eq. (52) we get $\sum_{j \in \mathcal{A}_s} \alpha_j \|(\mathbb{K}v)_j\| = 0$, which implies $(\mathbb{K}v)_j = 0$, for all $j \in \mathcal{A}_s$. Now, using this result in eq. (50) we get

$$\sum_{j \in \mathcal{B}} (\alpha_j \|(\mathbb{K}v)_j\| - \langle q_j, (\mathbb{K}v)_j \rangle) = 0$$

and, consequently, it holds $\langle q_j, (\mathbb{K}v)_j \rangle = \alpha_j \|(\mathbb{K}v)_j\|$. □

Now, to prove the directional differentiability of the solution operator for the lower-level problem eq. (4b) we will first demonstrate the following lemmata.

Lemma 4.4. Let $\mathbb{R}^n_+ \ni \alpha \geq 0$ and $\mathbb{R}^n_+ \ni \alpha + th \geq 0$. Then for every $v \in \mathcal{C}$ it holds

$$\begin{aligned} \left\langle \mathbb{K}^\top \left(\frac{q^t - q}{t} \right), v \right\rangle \leq \sum_{j \in \mathcal{I}} \frac{\alpha_j}{t} \left\langle \frac{(\mathbb{K}u^t)_j}{\|(\mathbb{K}u^t)_j\|} - \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, (\mathbb{K}v)_j \right\rangle \\ + \sum_{j \in \mathcal{I}} h_j \left\langle \frac{(\mathbb{K}u^t)}{\|(\mathbb{K}u^t)_j\|}, (\mathbb{K}v)_j \right\rangle + \sum_{j \in \mathcal{B}} h_j \|(\mathbb{K}v)_j\| + \sum_{j \in \mathcal{T} \cup \mathcal{I}_0} h_j \|(\mathbb{K}v)_j\|. \end{aligned} \tag{53}$$

Proof. Given that $v \in \mathcal{C}$, let us first bound the following product

$$\begin{aligned}\langle \mathbb{K}^\top q^t, v\rangle &= \sum_{j\in\mathcal{I}} \langle q_j^t, (\mathbb{K}v)_j\rangle + \sum_{j\in\mathcal{A}_s} \langle q_j^t, \underbrace{(\mathbb{K}v)_j}_{=0}\rangle + \sum_{j\in\mathcal{B}} \langle q_j^t, (\mathbb{K}v)_j\rangle + \sum_{j\in\mathcal{T}\cup\mathcal{I}_0} \langle q_j^t, (\mathbb{K}v)_j\rangle,\\ &\leq \sum_{j\in\mathcal{I}} \left\langle (\alpha_j + th_j)\frac{(\mathbb{K}u^t)_j}{\|(\mathbb{K}u^t)_j\|}, (\mathbb{K}v)_j \right\rangle + \sum_{j\in\mathcal{B}} (\alpha_j + th_j)\|(\mathbb{K}v)_j\| + \sum_{j\in\mathcal{T}\cup\mathcal{I}_0} th_j\|(\mathbb{K}v)_j\|,\end{aligned}$$

for t sufficiently small, since $u^t \to u$ implies $\mathcal{I}(\alpha, u) \subset \mathcal{I}(\alpha + th, u^t)$, where we used the property $(\mathbb{K}v)_j = 0$ for $j \in \mathcal{A}_s$ and $\alpha_j = 0$ for $j \in \mathcal{T} \cup \mathcal{I}_0$, along with Cauchy-Schwarz inequality and $\|q_j^t\| \leq \alpha_j + th_j$. Now, as $v \in \mathcal{C}$ we know the bound in eq. (49) holds, i.e.,

$$\langle \mathbb{K}^\top q, v\rangle \geq \sum_{j\in\mathcal{I}} \left\langle \alpha_j \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, (\mathbb{K}v)_j \right\rangle + \sum_{j\in\mathcal{B}} \alpha_j \|(\mathbb{K}v)_j\|.$$

Therefore,

$$\begin{aligned}\langle \mathbb{K}^\top (q^t - q), v\rangle \leq \sum_{j\in\mathcal{I}} \alpha_j &\left\langle \frac{(\mathbb{K}u^t)_j}{\|(\mathbb{K}u^t)_j\|} - \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, (\mathbb{K}v)_j \right\rangle\\ &+ \sum_{j\in\mathcal{I}} th_j \left\langle \frac{(\mathbb{K}u^t)_j}{\|(\mathbb{K}u^t)_j\|}, (\mathbb{K}v)_j \right\rangle + \sum_{j\in\mathcal{B}} th_j\|(\mathbb{K}v)_j\| + \sum_{j\in\mathcal{T}\cup\mathcal{I}_0} th_j\|(\mathbb{K}v)_j\|.\end{aligned}$$

Finally, dividing both sides by t yields the result. ☐

Lemma 4.5. Let $\mathbb{R}^n \ni \alpha \geq 0$ and $\mathbb{R}^n \ni \alpha + th \geq 0$. Then, it holds

$$\begin{aligned}\left\langle \mathbb{K}^\top \left(\frac{q^t - q}{t}\right), \frac{u^t - u}{t} \right\rangle &\geq \sum_{j\in\mathcal{I}} \frac{\alpha_j}{t} \left\langle \frac{(\mathbb{K}u^t)_j}{\|(\mathbb{K}u^t)_j\|} - \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, \frac{(\mathbb{K}u^t)_j - (\mathbb{K}u)_j}{t} \right\rangle\\ &+ \sum_{j\in\mathcal{I}} h_j \left\langle \frac{(\mathbb{K}u^t)}{\|(\mathbb{K}u^t)_j\|}, \frac{(\mathbb{K}u^t)_j - (\mathbb{K}u)_j}{t} \right\rangle + \frac{1}{t} \sum_{j\in\mathcal{A}\cup\mathcal{I}_0} h_j(\|(\mathbb{K}u^t)_j\| - \|(\mathbb{K}u)_j\|),\end{aligned}$$

where $\mathcal{A} = \mathcal{A}_s \cup \mathcal{B} \cup \mathcal{T}$.

Proof. For t small enough, we can split the product by their index set

$$\begin{aligned}&\left\langle \mathbb{K}^\top \left(\frac{q^t - q}{t}\right), \frac{u^t - u}{t} \right\rangle =\\ &\sum_{j\in\mathcal{I}} \frac{\alpha_j}{t} \left\langle \frac{(\mathbb{K}u^t)_j}{\|(\mathbb{K}u^t)_j\|} - \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, \frac{(\mathbb{K}u^t)_j - (\mathbb{K}u)_j}{t} \right\rangle + \sum_{j\in\mathcal{I}} h_j \left\langle \frac{(\mathbb{K}u^t)_j}{\|(\mathbb{K}u^t)_j\|}, \frac{(\mathbb{K}u^t)_j - (\mathbb{K}u)_j}{t} \right\rangle\\ &\qquad + \frac{1}{t} \sum_{j\in\mathcal{A}_s\cup\mathcal{B}} \left\langle q_j^t - q_j, \frac{(\mathbb{K}u^t)_j - (\mathbb{K}u)_j}{t} \right\rangle + \frac{1}{t} \sum_{j\in\mathcal{I}_0\cup\mathcal{T}} \left\langle q_j^t - q_j, \frac{(\mathbb{K}u^t)_j - (\mathbb{K}u)_j}{t} \right\rangle.\end{aligned}$$

Focusing, on the index set $\mathcal{A}_s \cup \mathcal{B}$, the complementarity relations in eqs. (6) and (46) yield

$$\begin{aligned}&\frac{1}{t^2} \sum_{j\in\mathcal{A}_s\cup\mathcal{B}} \langle q_j^t - q_j, (\mathbb{K}u^t)_j - (\mathbb{K}u)_j\rangle\\ &= \frac{1}{t^2} \sum_{j\in\mathcal{A}_s\cup\mathcal{B}} \langle q_j^t, (\mathbb{K}u^t)_j\rangle - \langle q_j^t, (\mathbb{K}u)_j\rangle - \langle q_j, (\mathbb{K}u^t)_j\rangle + \langle q_j, (\mathbb{K}u)_j\rangle,\\ &\geq \frac{1}{t^2} \sum_{j\in\mathcal{A}_s\cup\mathcal{B}} (\alpha_j + th_j)\|(\mathbb{K}u^t)_j\| - \underbrace{\|q_j^t\|}_{\leq \alpha_j + th_j} \|(\mathbb{K}u)_j\| - \underbrace{\|q_j\|}_{\leq \alpha_j} \|(\mathbb{K}u^t)_j\| + \alpha_j\|(\mathbb{K}u)_j\|,\\ &\geq \frac{1}{t} \sum_{j\in\mathcal{A}_s\cup\mathcal{B}} h_j(\|(\mathbb{K}u^t)_j\| - \|(\mathbb{K}u)_j\|).\end{aligned}$$

Using the same analysis over the set $\mathcal{I}_0 \cup \mathcal{T}$ we get

$$\begin{aligned}
\frac{1}{t^2} \sum_{j\in\mathcal{I}_0\cup\mathcal{T}} \langle q_j^t - q_j, (\mathbb{K}u^t)_j - (\mathbb{K}u)_j\rangle &= \frac{1}{t^2} \sum_{j\in\mathcal{I}_0\cup\mathcal{T}} \langle q_j^t, (\mathbb{K}u^t)_j\rangle - \langle q_j^t, (\mathbb{K}u)_j\rangle, \\
\geq \frac{1}{t^2} \sum_{j\in\mathcal{I}_0\cup\mathcal{T}} t h_j \|(\mathbb{K}u^t)_j\| - \underbrace{\|q_j^t\|}_{\leq t h_j} \|(\mathbb{K}u)_j\| &\geq \frac{1}{t} \sum_{j\in\mathcal{I}_0\cup\mathcal{T}} h_j (\|(\mathbb{K}u^t)_j\| - \|(\mathbb{K}u)_j\|).
\end{aligned}$$

□

Theorem 4.6. Let $\alpha \in \mathbb{R}_+^n$ and $h \in \mathbb{R}^n$ be a direction such that $\alpha + th \geq 0$ for t small enough. The solution operator $S : \alpha \to S(\alpha) = u \in \mathbb{R}^m$ is directionally differentiable and its directional derivative $\eta \in \mathcal{C}(\alpha, u)$ at u, in direction h, is given by the solution of the following variational inequality

$$\begin{aligned}
\langle \phi''(u)\eta, v - \eta\rangle + \sum_{j\in\mathcal{I}} \alpha_j \langle T_j(\mathbb{K}\eta)_j, (\mathbb{K}v)_j - (\mathbb{K}\eta)_j\rangle + h_j \left\langle \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, (\mathbb{K}v)_j - (\mathbb{K}\eta)_j \right\rangle \\
+ \sum_{j\in\mathcal{B}} \frac{h_j}{\alpha_j} \langle q_j, (\mathbb{K}v)_j - (\mathbb{K}\eta)_j\rangle + \sum_{j\in\mathcal{I}_0\cup\mathcal{T}} h_j (\|(\mathbb{K}v)_j\| - \|(\mathbb{K}\eta)_j\|) \geq 0, \ \forall v \in \mathcal{C}, \quad (54)
\end{aligned}$$

where $T_j(\mathbb{K}v)_j = \frac{(\mathbb{K}v)_j}{\|(\mathbb{K}u)_j\|} - \frac{(\mathbb{K}u)_j (\mathbb{K}u)_j^\top (\mathbb{K}v)_j}{\|(\mathbb{K}u)_j\|^3}$ for $v \in \mathbb{R}^m$.

Proof. To verify the variational inequality, let us take eq. (46), eq. (6) and test them with $v - \frac{u^t - u}{t}$, with $v \in \mathcal{C}(\alpha, u)$

$$\begin{aligned}
0 &= \left\langle \frac{\phi'(u^t) - \phi'(u)}{t}, v - \frac{u^t - u}{t} \right\rangle + \left\langle \mathbb{K}^\top \left(\frac{q^t - q}{t}\right), v - \frac{u^t - u}{t} \right\rangle, \\
&= \left\langle \frac{\phi'(u^t) - \phi'(u)}{t}, v - \frac{u^t - u}{t} \right\rangle + \left\langle \mathbb{K}^\top \left(\frac{q^t - q}{t}\right), v \right\rangle - \left\langle \mathbb{K}^\top \left(\frac{q^t - q}{t}\right), \frac{u^t - u}{t} \right\rangle
\end{aligned}$$

Now, applying the bounds in theorems 4.4 and 4.5 we have

$$\begin{aligned}
0 \leq &\left\langle \frac{\phi'(u^t) - \phi'(u)}{t}, v - \frac{u^t - u}{t} \right\rangle \\
&+ \sum_{j\in\mathcal{I}} \frac{\alpha_j}{t} \left\langle \frac{(\mathbb{K}u^t)_j}{\|(\mathbb{K}u^t)_j\|} - \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, (\mathbb{K}v)_j - \frac{(\mathbb{K}u^t)_j - (\mathbb{K}u)_j}{t} \right\rangle \\
&+ h_j \left\langle \frac{(\mathbb{K}u^t)}{\|(\mathbb{K}u^t)_j\|}, (\mathbb{K}v)_j - \frac{(\mathbb{K}u^t)_j - (\mathbb{K}u)_j}{t} \right\rangle + \sum_{j\in\mathcal{B}} h_j \|(\mathbb{K}v)_j\| \\
&+ \sum_{j\in\mathcal{I}_0\cup\mathcal{T}} h_j \|(\mathbb{K}v)_j\| - \frac{1}{t} \sum_{j\in\mathcal{A}_s\cup\mathcal{B}\cup\mathcal{I}_0\cup\mathcal{T}} h_j (\|(\mathbb{K}u^t)_j\| - \|(\mathbb{K}u)_j\|).
\end{aligned}$$

Taking the limit $t \to 0$, as well as, the differentiability of the term $x/\|x\|$ in the inactive set, and given that $(\mathbb{K}\eta)_j = 0$ in the strongly active set $\mathcal{A}_s$, it yields

$$\begin{aligned}
0 \leq \langle \phi''(u)\eta, v - \eta\rangle + \sum_{j\in\mathcal{I}} \alpha_j \left\langle \left(\frac{I}{\|(\mathbb{K}u)_j\|} - \frac{(\mathbb{K}u)_j (\mathbb{K}u)_j^\top}{\|(\mathbb{K}u)_j\|^3} \right) (\mathbb{K}\eta)_j, (\mathbb{K}v)_j - (\mathbb{K}\eta)_j \right\rangle \\
+ h_j \left\langle \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, (\mathbb{K}v)_j - (\mathbb{K}\eta)_j \right\rangle + \sum_{j\in\mathcal{B}} h_j (\|(\mathbb{K}v)_j\| - \|(\mathbb{K}\eta)_j\|) + \sum_{j\in\mathcal{I}_0\cup\mathcal{T}} h_j (\|(\mathbb{K}v)_j\| - \|(\mathbb{K}\eta)_j\|).
\end{aligned}$$

Using the definition for T_j and recalling $v, \eta \in \mathcal{C}$, the inequality takes the form in eq. (54).

Now it is required to verify the uniqueness of the limit. For this purpose, let us note that eq. (54) is a variational inequality

$$\langle \phi''(u)\eta, v-\eta\rangle + \sum_{j\in\mathcal{I}} \alpha_j \langle T_j(\mathbb{K}\eta)_j, (\mathbb{K}v)_j - (\mathbb{K}\eta)_j\rangle + \sum_{j\in\mathcal{B}} \frac{h_j}{\alpha_j}\langle q_j, (\mathbb{K}v)_j - (\mathbb{K}\eta)_j\rangle \\ + \sum_{j\in\mathcal{I}_0\cup\mathcal{T}} h_j(\|(\mathbb{K}v)_j\| - \|(\mathbb{K}\eta)_j\|) \geq -\sum_{j\in\mathcal{I}} h_j \left\langle \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, (\mathbb{K}v)_j - (\mathbb{K}\eta)_j \right\rangle, \forall v\in\mathcal{C}$$

Now, given that the function $f(z) := \sum_{j=1}^{n} \|(\mathbb{K}z)_j\|$ is indeed convex, lower semicontinuous and proper, the right hand side is continuous and linear, and finally, the bilinear form in the left hand side is V-elliptic, we know by [15], that there exists a unique solution for this variational inequality. $\square$

Using the demonstrated Bouligand differentiability of the solution operator, and the cooresponding characterization of the directional derivative described in this section, we have proven the following result.

Theorem 4.7. Let $\alpha^* \in \mathbb{R}^n_+$ be a local optimal solution of eq. (43) and $u^* = S(\alpha^*)$. Then α^* is a B-stationary point, i.e., it satisfies the following inequality

$$\langle \nabla J(u^*), S'(\alpha^*; \alpha-\alpha^*)\rangle \geq 0, \quad \forall \alpha\in\mathbb{R}^n_+, \tag{55}$$

where $S'(\alpha^*;\alpha-\alpha^*) =: \eta$ is the unique solution to Equation (54).

4.2 Strict Complementarity

The characterization of the directional differentiability can take different formulations if any of the active sets becomes empty. For instance, assuming the zero-inactive and triactive sets empty, i.e., $\mathcal{I}_0\cup\mathcal{T} = \emptyset$, then the directional derivative of the solution operator can be written as the following variational inequality of the first kind

$$\langle \phi''(u)\eta, v-\eta\rangle + \sum_{j\in\mathcal{I}} \alpha_j \langle T_j(\mathbb{K}\eta)_j, (\mathbb{K}v)_j - (\mathbb{K}\eta)_j\rangle \\ + h_j \left\langle \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, (\mathbb{K}v)_j - (\mathbb{K}\eta)_j \right\rangle + \sum_{j\in\mathcal{B}} \frac{h_j}{\alpha_j}(\langle q_j, (\mathbb{K}v)_j - (\mathbb{K}\eta)_j\rangle) \geq 0, \ \forall v\in\mathcal{C}. \tag{56}$$

Furthermore, assuming an empty biactive set and $\alpha_j = 0$, for all j, we obtain that the solution operator is Fréchet differentiable as stated in the following theorem.

Theorem 4.8. Let us assume the index set $\mathcal{B}\cup\mathcal{I}_0\cup\mathcal{T}$ is empty. Then, the solution operator is Fréchet differentiable and the derivative can be computed as the solution of the following system of equations

$$\phi''(u)\eta + \mathbb{K}^\top\lambda = 0, \tag{57a}$$

$$\lambda_j - \alpha_j T_j(\mathbb{K}\eta)_j - \frac{h_j}{\alpha_j} q_j = 0, \qquad \forall j\in\mathcal{I}, \tag{57b}$$

$$(\mathbb{K}\eta)_j = 0, \qquad \forall j\in\mathcal{A}_s. \tag{57c}$$

Proof. Using the empty biactive set assumption, we get that the cone $\mathcal{C}$ becomes the following linear subspace $\mathcal{C}(\alpha,u) = \{v\in\mathbb{R}^m : (\mathbb{K}v)_j = 0 \text{ if } (\mathbb{K}u)_j = 0\}$. Thus, the variational inequality in eq. (56) becomes the following variational equation

$$\langle \phi''(u)\eta, v-\eta\rangle + \sum_{j\in\mathcal{I}} \alpha_j \langle T_j(\mathbb{K}\eta)_j, (\mathbb{K}v)_j - (\mathbb{K}\eta)_j\rangle + h_j \left\langle \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, (\mathbb{K}v)_j - (\mathbb{K}\eta)_j \right\rangle = 0, \ \forall v\in\mathcal{C}. \tag{58}$$

Equation (58) guarantees that the directional derivative of the solution operator is a linear mapping w.r.t. the direction h. Since S is Bouligand differentiable, it implies the Fréchet differentiability [28, Proposition 3.1.2]. Furthermore, eq. (58) is equivalent to the following optimization problem

$$\min_{\eta \in \mathcal{C}} \quad \frac{1}{2}\langle \eta, \phi''(u)\eta \rangle + \sum_{j\in\mathcal{I}} \alpha_j \left(\frac{\|(\mathbb{K}\eta)_j\|^2}{\|(\mathbb{K}u)_j\|} - \frac{\langle (\mathbb{K}u)_j, (\mathbb{K}\eta)_j \rangle^2}{\|(\mathbb{K}u)_j\|^3} \right) + h_j \left\langle (\mathbb{K}\eta)_j, \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|} \right\rangle \tag{59}$$

Then the KKT-optimality conditions for this problem look as follows

$$\langle \phi''(u)\eta, v\rangle + \sum_{j\in\mathcal{I}} \alpha_j \langle T_j(\mathbb{K}\eta)_j, (\mathbb{K}v)_j\rangle + h_j \left\langle \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, (\mathbb{K}v)_j \right\rangle + \sum_{j\in\mathcal{A}_s} \langle \nu_j, (\mathbb{K}v)_j\rangle = 0,$$
$$(\mathbb{K}\eta)_j = 0, \forall j \in \mathcal{A}_s,$$

with Lagrange multipliers $\nu_j \in \mathbb{R}^2$. Since all the constraints are linear, the Abadie constraint qualification condition is satisfied. By introducing $\lambda \in \mathbb{R}^{n\times 2}$ as

$$\lambda_j := \begin{cases} \nu_j, & \forall j \in \mathcal{A}_s \\ \alpha_j T_j(\mathbb{K}\eta)_j + \frac{h_j}{\alpha_j} q_j, & \forall j \in \mathcal{I} \end{cases}$$

the result is obtained. □

5 Bouligand Subdifferential

Even though the Bouligand stationarity condition presented in Section 4 holds for any local optimal solution without requiring any constraint qualification, its purely primal form is in general not amenable for algorithmic purposes; this limitation is related to the non-linearity of the directional derivative. As a remedy, in this section we will focus on the study of the Bouligand subdifferential of the solution operator S. The characterization of the linear elements of this subdifferential turns out to be useful when devising a numerical algorithm to solve the bilevel problem.

Thanks to the local Lipschitz continuity of S, showed in section 4, and Rademacher's theorem, we know the solution operator is differentiable almost everywhere. Denoting the set of points where this function is differentiable as D_S, the Bouligand subdifferential $\partial_B S(\alpha)$ is defined as follows.

Definition 5.1 (Bouligand subdifferential)**.** Let $S : \mathbb{R}^n \to \mathbb{R}^m$ be a locally Lipschitz function, and $\alpha \in \mathbb{R}^n$ arbitrary but fixed. The set

$$\partial_B S(\alpha) := \{G \in \mathbb{R}^{m\times n} : \exists \{\alpha_k\} \subset D_S,\ \alpha_k \to \alpha \wedge S'(\alpha_k) \to G\}, \tag{60}$$

is called *Bouligand subdifferential* of S at α.

In the next result a characterization of the elements of the Bouligand subdifferential is provided. We assume along this section that $\alpha_j > 0$, i.e., $\mathcal{T} \cup \mathcal{I}_0 = \emptyset$.

Theorem 5.1. Let $G \in \partial_B S(\alpha)$. There exists a partition of the biactive set $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ such that, for any h such that $\alpha + th \geq 0$, $\tilde{\eta} := Gh$ is the unique solution of the system

$$\langle \phi''(u)\tilde{\eta}, v\rangle + \sum_{j\in\mathcal{I}\cup\mathcal{B}_2} \langle \tilde{\lambda}_j, (\mathbb{K}v)_j\rangle = 0, \qquad \forall v \in V \tag{61a}$$

$$\tilde{\lambda}_j - \alpha_j T_j(\mathbb{K}\tilde{\eta})_j - \frac{h_j}{\alpha_j} q_j = 0, \qquad \forall j \in \mathcal{I}, \tag{61b}$$

$$\tilde{\lambda}_j - \frac{h_j}{\alpha_j} q_j = 0, \qquad \forall j \in \mathcal{B}_2 \tag{61c}$$

where $V := \{v \in \mathbb{R}^m :\ (\mathbb{K}v)_j = 0,\ \forall j \in \mathcal{A}_s \cup \mathcal{B}_1; (\mathbb{K}v)_j \in \operatorname{span}(q_j), \forall j \in \mathcal{B}_2\}$.

Proof. We know the solution operator is locally Lipschitz continuous (see theorem 4.1), which implies it is differentiable almost everywhere. Let us consider a sequence $\{\alpha_k\} \subset D_S$ such that $\alpha_k \to \alpha$ and $S'(\alpha_k) \to G$.

Due to the differentiability in the elements of this sequence, the parameter sequence fulfills $(\alpha_k)_j > 0$, for all $j = 1, \ldots, n$. Moreover, thanks to the Lipschitz continuity of S we know that

$$u_k = S(\alpha_k) \to S(\alpha) = u,$$
$$\mathbb{K}^\top q_k = -\phi'(u_k) \to -\phi'(u) = \mathbb{K}^\top q.$$

This last statement follows from the fact that $\{q_k\}$ is also bounded and, therefore, has a converging subsequence. Now, each of this subsequence elements (u_k, q_k) define their respective inactive $\mathcal{I}^k := \mathcal{I}(\alpha_k, u_k)$ and strongly active $\mathcal{A}_s^k := \mathcal{A}_s(\alpha_k, u_k)$ sets.
By continuity, we know that $\mathcal{I} \subset \mathcal{I}^k$ and $\mathcal{A}_s \subset \mathcal{A}_s^k$, for k sufficiently large. Since $\{\alpha_k\} \subset D_S$, it then follows that $\eta_k := S'(\alpha_k)h$ satisfies the system

$$\phi''(u_k)\eta_k + \mathbb{K}^\top \lambda_k = 0, \tag{62a}$$
$$(\lambda_k)_j - (\alpha_k)_j (T_k)_j (\mathbb{K}\eta_k)_j = \frac{h_j}{(\alpha_k)_j} \mathbb{K}_j^\top (q_k)_j, \qquad \forall j \in \mathcal{I}^k, \tag{62b}$$
$$(\mathbb{K}\eta_k)_j = 0, \qquad \forall j \in \mathcal{A}_s^k, \tag{62c}$$

or equivalently,

$$\langle \phi''(u_k)\eta_k, v\rangle + \sum_{j \in \mathcal{I}^k} \langle (\alpha_k)_j (T_k)_j (\mathbb{K}\eta_k)_j, (\mathbb{K}v)_j\rangle + \frac{h_j}{(\alpha_k)_j} \langle (q_k)_j, (\mathbb{K}v)_j\rangle = 0, \forall v \in V^k, \tag{63}$$

with $V^k := \{v \in \mathbb{R}^m : (\mathbb{K}v)_j = 0, \forall j \in \mathcal{A}_s^k\}$. From the definition of the Bouligand subdifferential it follows that $\tilde{\eta} = \lim_{k\to\infty} \eta_k$. Moreover, since for $j \in \mathcal{I}$ the sequence $\{(\lambda_k)_j\}$ is bounded, then there exists a subsequence that converges to a limit point $\tilde{\lambda}_j$. Therefore, up to a subsequence, by passing to the limit we get

$$\tilde{\lambda}_j - \alpha_j T_j (\mathbb{K}\tilde{\eta})_j - \frac{h_j}{\alpha_j} q_j = 0, \ \forall j \in \mathcal{I},$$
$$(\mathbb{K}\tilde{\eta})_j = 0, \ \forall j \in \mathcal{A}_s.$$

Let us now consider a partition of the biactive set $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$, with

$$\mathcal{B}_1(\alpha, u) := \{j \in \mathcal{B}(\alpha, u) : \exists \{u_{k_l}\} : (\mathbb{K}u_{k_l})_j = 0, \forall l\} \quad \text{and} \quad \mathcal{B}_2(\alpha, u) := \mathcal{B}(\alpha, u) \backslash \mathcal{B}_1(\alpha, u).$$

For the index set $\mathcal{B}_1$ we know the subsequence $(\mathbb{K}\eta_{k_l})_j = 0$, for all k. Since $\eta_k \to \tilde{\eta}$, we get

$$(\mathbb{K}\tilde{\eta})_j = 0, \ \forall j \in \mathcal{A}_s \cup \mathcal{B}_1.$$

Considering the partition $\mathcal{B}_2$, we approach a biactive point by a sequence of points such that $(\mathbb{K}u_k)_j \neq 0$, i.e., $j \in \mathcal{I}^k$. Let us first notice that the term on the right hand side of Equation (62b) is uniformly bounded and, therefore, as $k \to \infty$,

$$\sum_{j \in \mathcal{I}^k} \frac{h_j}{(\alpha_k)_j} \langle (q_k)_j, (\mathbb{K}v)_j\rangle \to \sum_{j \in \mathcal{I} \cup \mathcal{B}_2} \frac{h_j}{\alpha_j} \langle q_j, (\mathbb{K}v)_j\rangle, \quad \forall v \in V.$$

In addition, defining $(\zeta_k)_j = (\alpha_k)_j (T_k)_j (\mathbb{K}\eta_k)_j$, for $j \in \mathcal{I}^k$, we get that

$$\langle (\zeta_k)_j, (\mathbb{K}\eta_k)_j\rangle = \frac{(\alpha_k)_j}{\|(\mathbb{K}u_k)_j\|} \left(\|(\mathbb{K}\eta_k)_j\|^2 - \frac{1}{\|(\mathbb{K}u_k)_j\|^2} \langle (\mathbb{K}\eta_k)_j, (\mathbb{K}u_k)_j\rangle^2 \right) \geq 0, \ \forall j \in \mathcal{I}^k.$$

Thanks to Equation (63), we also get that

$$0 \leq \langle (\zeta_k)_j, (\mathbb{K}\eta_k)_j\rangle \leq \langle \phi''(u_k)\eta_k, \eta_k\rangle + \sum_{j \in \mathcal{I}^k} \langle (\zeta_k)_j, (\mathbb{K}\eta_k)_j\rangle \leq \sum_{j \in \mathcal{I}^k} |h_j| \|(\mathbb{K}\eta_k)_j\|,$$

which, since $\eta_k \to \tilde{\eta}$, as $k \to \infty$, implies that $\langle (\zeta_k)_j, (\mathbb{K}\eta_k)_j \rangle$ is uniformly bounded. Since for $j \in \mathcal{B}_2$ we know that $(\mathbb{K}u_k)_j \to 0$, it follows from the previous relations that

$$\alpha_j^2 \|(\mathbb{K}\tilde{\eta})_j\|^2 - \langle q_j, (\mathbb{K}\tilde{\eta})_j \rangle^2 = \lim_{k\to\infty} (\alpha_k)_j^2 \|(\mathbb{K}\eta_k)_j\|^2 - \langle (q_k)_j, (\mathbb{K}\eta_k)_j \rangle^2 = 0,$$

which implies that $(\mathbb{K}\tilde{\eta})_j \in \operatorname{span}(q_j), \forall j \in \mathcal{B}_2$.

Finally, for any $v \in V$, we obtain that

$$\begin{aligned}
\lim_{k\to\infty} \langle (\zeta_k)_j, (\mathbb{K}v)_j \rangle = \lim_{k\to\infty} \langle (\zeta_k)_j, c(q_k)_j \rangle &= \lim_{k\to\infty} \left\langle (\zeta_k)_j, c(\alpha_k)_j \frac{(\mathbb{K}u_k)_j}{\|(\mathbb{K}u_k)_j\|} \right\rangle, \\
&= c \lim_{k\to\infty} \left\langle (\alpha_k)_j \frac{(\mathbb{K}\eta_k)_j}{\|(\mathbb{K}u_k)_j\|}, (\alpha_k)_j \frac{(\mathbb{K}u_k)_j}{\|(\mathbb{K}u_k)_j\|} \right\rangle \\
&\qquad - \left\langle (\alpha_k)_j \frac{\langle (\mathbb{K}\eta_k)_j, (\mathbb{K}u_k)_j \rangle (\mathbb{K}u_k)_j}{\|(\mathbb{K}u_k)_j\|^3}, (\alpha_k)_j \frac{(\mathbb{K}u_k)_j}{\|(\mathbb{K}u_k)_j\|} \right\rangle, \\
&= c \lim_{k\to\infty} \frac{(\alpha_k)_j^2}{\|(\mathbb{K}u_k)_j\|^2} \langle (\mathbb{K}\eta_k)_j, (\mathbb{K}u_k)_j \rangle \\
&\qquad - \frac{(\alpha_k)_j^2}{\|(\mathbb{K}u_k)_j\|^4} \langle (\mathbb{K}\eta_k)_j, (\mathbb{K}u_k)_j \rangle \langle (\mathbb{K}u_k)_j, (\mathbb{K}u_k)_j \rangle = 0.
\end{aligned}$$

By passing to the limit in eq. (63) the result is obtained. $\square$

Corollary 1. *Let $G \in \partial_B S(\alpha)$. There exists a partition of the biactive set $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ and a multiplier $\theta \in \mathbb{R}^m$ such that, for any h such that $\alpha + th \geq 0$, $\tilde{\eta} := Gh$ is the unique solution of the system*

$$\phi''(u)\tilde{\eta} + \mathbb{K}^T\theta = 0 \tag{65a}$$

$$\theta_j - \alpha_j T_j(\mathbb{K}\tilde{\eta})_j - \frac{h_j}{\alpha_j} q_j = 0, \qquad \forall j \in \mathcal{I}, \tag{65b}$$

$$\langle \theta_j, q_j \rangle - \alpha_j h_j = 0, \qquad \forall j \in \mathcal{B}_2. \tag{65c}$$

Proof. Let us consider the functional $\mathcal{F} \in \mathbb{R}^n$ defined by

$$(\mathcal{F}, v) := (\phi''(u)\tilde{\eta}, v) + \sum_{j\in\mathcal{I}} \langle \alpha_j T_j(\mathbb{K}\tilde{\eta})_j, (\mathbb{K}v)_j \rangle + \sum_{j\in\mathcal{I}\cup\mathcal{B}_2} \frac{h_j}{\alpha_j} \langle q_j, (\mathbb{K}v)_j \rangle, \ \forall v \in V.$$

Equation (61a) can then be written as $\mathcal{F} \in V^\perp$. Thanks to the structure of the linear subspace V, it can be represented in a separated way as $V = \left(\bigcap_{j\in\mathcal{A}_S\cup\mathcal{B}_1} V_j^1\right) \cap \left(\bigcap_{j\in\mathcal{B}_2} V_j^2\right)$, where

$$\begin{aligned}
V_j^1 &:= \{v \in \mathbb{R}^n : (\mathbb{K}v)_j = 0\}, && j \in \mathcal{A}_S \cup \mathcal{B}_1, \\
V_j^2 &:= \{v \in \mathbb{R}^n : (\mathbb{K}v)_j \in \operatorname{span}(q_j)\}, && j \in \mathcal{B}_2.
\end{aligned}$$

Consequently, $V^\perp = \sum_{j\in\mathcal{A}_S\cup\mathcal{B}_1} (V_j^1)^\perp + \sum_{j\in\mathcal{B}_2} (V_j^2)^\perp$.

For $j \in \mathcal{A}_S \cup \mathcal{B}_1$, we get that $(V_j^1)^\perp = \ker(\mathbb{K}_j)^\perp$. Thanks to the orthogonality relations, it follows that $\ker(\mathbb{K}_j)^\perp = \operatorname{range}(\mathbb{K}_j^T)$. Hence, for any $\xi_j \in (V_j^1)^\perp$, there exist π_j such that $\xi_j = \mathbb{K}_j^T \pi_j$. Consequently,

$$\sum_{j\in\mathcal{A}_S\cup\mathcal{B}_1} (V_j^1)^\perp = \sum_{j\in\mathcal{A}_S\cup\mathcal{B}_1} \mathbb{K}_j^T \pi_j, \quad \pi_j \in \mathbb{R}^2.$$

For $j \in \mathcal{B}_2$, any $v \in V_j^2$ can be represented as

$$v = \phi + \varphi, \quad \text{with } (\mathbb{K}_j\varphi) = 0 \text{ and } \phi \in \operatorname{range}(\mathbb{K}_j^T).$$

Since $(\mathbb{K}v)_j \in \operatorname{span}(q_j)$ and $(\mathbb{K}_j\varphi) = 0$, it follows that $(\mathbb{K}v)_j \in \operatorname{span}(q_j)$ as well. Let us now consider $w_j \in (V_j^2)^\perp$, which can be represented as $w_j = \tilde{w}_j + \hat{w}_j$, where $\tilde{w}_j \in \operatorname{range}(\mathbb{K}_j^T)$ and $\hat{w}_j \in \operatorname{range}(\mathbb{K}_j^T)^\perp = \ker(\mathbb{K}_j)$. Consequently, there exists ψ_j such that

$$w_j = \mathbb{K}_j^T \psi_j + \hat{w}_j, \quad \text{with } \mathbb{K}_j \hat{w}_j = 0.$$

Taking the scalar product with $v_j \in V_j^2$ we get

$$(w_j, v_j) = (\mathbb{K}_j^T \psi_j + \hat{w}_j, \phi + \varphi) = \langle \psi_j, \mathbb{K}_j \phi \rangle + (\hat{w}_j, \mathbb{K}_j^T \psi) + (\hat{w}_j, \varphi) = c\langle \psi_j, q_j \rangle + (\hat{w}_j, \varphi),$$

since $\mathbb{K}_j \varphi = \mathbb{K}_j \hat{w}_j = 0$. For the product to be zero, it is then required that $(\hat{w}_j, \varphi) = 0, \forall \varphi \in \ker(\mathbb{K}_j)$ and $\langle \psi_j, q_j \rangle = 0$. Since $\hat{w}_j$ belongs to $\ker(\mathbb{K}_j)$ as well, it follows that $\hat{w}_j = 0$. Consequently,

$$\sum_{j \in \mathcal{B}_2} (V_j^2)^\perp = \sum_{j \in \mathcal{B}_2} \mathbb{K}_j^T \psi_j, \quad \psi_j \in \mathbb{R}^2 : \langle \psi_j, q_j \rangle = 0.$$

Altogether, we then obtain that there exist multipliers π_j and ψ_j such that

$$\mathcal{F} + \sum_{j \in \mathcal{A}_S \cup \mathcal{B}_1} \mathbb{K}_j^T \pi_j + \sum_{j \in \mathcal{B}_2} \mathbb{K}_j^T \psi_j = 0,$$

with $\langle \psi_j, q_j \rangle = 0$. Defining

$$\theta_j := \begin{cases} \alpha_j T_j(\mathbb{K}\tilde{\eta})_j + \frac{h_j}{\alpha_j} q_j, & j \in \mathcal{I}, \\ \pi_j, & j \in \mathcal{A}_S \cup \mathcal{B}_1, \\ \psi_j + \frac{h_j}{\alpha_j} q_j, & j \in \mathcal{B}_2, \end{cases}$$

the result is obtained. ☐

Next we verify that, along a given direction, there exists a solution of system eq. (61) which coincides with the directional derivative. When properly characterized, this enables us to use of a linear representative of the (otherwise nonlinear) directional derivative within a solution algorithm.

Theorem 5.2. For any $\alpha \in \mathbb{R}^n_+$ and $h \in \mathbb{R}^n$ such that $\alpha + th \geq 0$, there exists a linearized element $\tilde{\eta} = Gh$ such that $S'(\alpha)h = Gh$.

Proof. Let us recall that, since by assumption $\mathcal{T} \cup \mathcal{I}_0 = \emptyset$, the directional derivative of the solution mapping, in direction h, is given by the unique $\eta \in \mathcal{C}(\alpha, u)$ solution of

$$\langle \phi''(u)\eta, v - \eta \rangle + \sum_{j \in \mathcal{I}} \langle \alpha_j T_j(\mathbb{K}\eta)_j, (\mathbb{K}v)_j - (\mathbb{K}\eta)_j \rangle \geq$$
$$- \sum_{j \in \mathcal{I}} h_j \left\langle \frac{(\mathbb{K}u)_j}{\|(\mathbb{K}u)_j\|}, (\mathbb{K}v)_j - (\mathbb{K}\eta)_j \right\rangle - \sum_{j \in \mathcal{B}} \frac{h_j}{\alpha_j} \langle q_j, (\mathbb{K}v)_j - (\mathbb{K}\eta)_j \rangle, \quad (66)$$

for all $v \in \mathcal{C}(u)$. Considering the sets $\mathcal{B}_1 := \{j \in \mathcal{B} : (\mathbb{K}\eta)_j = 0\}$ and $\mathcal{B}_2 := \mathcal{B} \backslash \mathcal{B}_1$, and since $\eta \in \mathcal{C}(u)$, it also follows that $(\mathbb{K}\eta) = c_j q_j$, for all $j \in \mathcal{B}_2$, for some $c_j > 0$. Consequently η belongs to the subspace

$$V := \{v \in \mathbb{R}^n : (\mathbb{K}v)_j = 0, \ \forall j \in \mathcal{A}_s \cup \mathcal{B}_1; \ (\mathbb{K}v)_j \in \operatorname{span}(q_j), \ \forall j \in \mathcal{B}_2\}.$$

Moreover, for any $w \in V$ it follows that, for t sufficiently small, $\eta \pm tw \in \mathcal{C}(u)$. Testing (66) with these vectors we then get that

$$\langle \phi''(u)\eta, w \rangle + \sum_{j \in \mathcal{I}} \langle \alpha_j T_j(\mathbb{K}\eta)_j, (\mathbb{K}w)_j \rangle = - \sum_{j \in \mathcal{I} \cup \mathcal{B}_2} \frac{h_j}{\alpha_j} \langle q_j, (\mathbb{K}w)_j \rangle, \quad \forall w \in V,$$

and, consequently, the directional derivative takes the form $\eta = Gh$, solution of eq. (61), with $\mathcal{B}_2$ as defined above. ☐

6 Trust Region Algorithm

In this section we will describe the numerical algorithm used for finding optimal parameters of eq. (4). Thanks to the Bouligand subdifferential characterization given in section 5, it is possible to compute a linearized representative of the directional derivative via the linear system eq. (61). Using this information we can make use of a descent-like algorithm for its numerical solution. Indeed, by using the uniqueness properties of the solution operator, we can write a *reduced optimization problem*

$$\min_{\alpha \geq 0} \quad j(u(\alpha)) \tag{67}$$

where $u(\alpha)$ is the image reconstruction corresponding to a particular value of α. With this reduced problem, we can make use of the stationarity condition for the bilevel problem described in eq. (44) and the directional derivative characterization eq. (57). By using the definition of the directional derivative for the reduced optimization problem, we get

$$\langle j'(\alpha), h \rangle = \langle \nabla J(u), S'(\alpha; h) \rangle = \langle \nabla J(u), \tilde{\eta} \rangle. \tag{68}$$

where $\tilde{\eta}$ is a solution of system eq. (61) for a particular partition of the biactive set $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$. Let us now define a *generalized adjoint* $p \in \mathbb{R}^m$ as the solution of the following system

$$\begin{aligned} \langle \phi''(u)^\top p, v \rangle + \sum_{j \in \mathcal{I}} \langle \mu_j, (\mathbb{K}v)_j \rangle - \langle \nabla J(u), v \rangle &= 0, && \forall v \in V, \\ \mu_j - \alpha_j T_j (\mathbb{K}p)_j &= 0, && \forall j \in \mathcal{I}, \end{aligned}$$

where V is defined as in Theorem 5.1. Moreover, using the results in Theorem 5.2, we know that $\tilde{\eta} \in V$ is a linear representative of the directional derivative. Consequently, eq. (68) reads

$$\langle j'(\alpha), h \rangle = \langle \nabla J(u), \tilde{\eta} \rangle = \langle \phi''(u)^\top p, \tilde{\eta} \rangle + \sum_{j \in \mathcal{I}} \langle \alpha_j T_j (\mathbb{K}p)_j, (\mathbb{K}\tilde{\eta})_j \rangle,$$

Rearranging the terms we get

$$\langle j'(\alpha), h \rangle = \langle p, \phi''(u)\tilde{\eta} \rangle + \sum_{j \in \mathcal{I}} \langle (\mathbb{K}p)_j, \alpha_j T_j (\mathbb{K}\tilde{\eta})_j \rangle = \langle p, \phi''(u)\tilde{\eta} \rangle + \sum_{j \in \mathcal{I}} \langle (\mathbb{K}p)_j, \tilde{\lambda}_j \rangle,$$

Finally, using eq. (57) we get

$$\langle j'(\alpha), h \rangle = - \sum_{j \in \mathcal{I} \cup \mathcal{B}_2} \frac{h_j}{\alpha_j} \langle q_j, (\mathbb{K}p)_j \rangle \tag{69}$$

In this work, we will rely on a nonsmooth trust region method in the spirit of [6]. In general, a trust-region algorithm works by defining for an iteration k a radius Δ_k and replaces the function with a *model function* $m_k(\alpha_k)$ within a *trust-region.* This region in this paper will be represented by a l_∞ ball, and the radius Δ_k is updated according to a *quality measure* based on the predicted decrease of the model function *pred* and the actual cost function reduction *ared*.

The model function we will be using has the following form

$$m_k(\alpha_k) = j(\alpha_k) + g_k^\top d_k + \frac{1}{2} d_k^\top B_k d_k,$$

where B_k is a second order matrix built using BFGS updates, and g_k is obtained either by using the Bouligand subdifferential element obtained using eq. (69) or a gradient obtained from the regularized model, $g_{\gamma,k}$, if the radius falls below a threshold value, see Section 6.2. The idea of differentiating two phases for the trust region subproblem approximation was taken from [6], where this strategy is presented for optimizing nonsmooth, non-convex and locally-Lipschitz functions. By defining an appropriate model function, dependent on the radius value, convergence to a Clarke (C)-stationary points may be verified [6, Proposition 2.10]. The complete steps are provided in algorithm 1.

6.1 Trust Region Subproblem

Regarding step 10 in algorithm 1, we need to approximate the solution of the following trust region subproblem

$$\begin{aligned} \min_{s \in \mathbb{R}^n} \quad & j(\alpha) + g^\top s + \frac{1}{2} s^\top B s \\ \text{s.t.} \quad & \|s\|_\infty \leq \Delta, \\ & s + \alpha \geq 0, \end{aligned} \tag{70}$$

Algorithm 1 Non-smooth Trust-Region Algorithm

Choose initial parameter α_0, radius Δ_0, $0 < \eta_1 \leq \eta_2 < 1$, $0 < \gamma_1 \leq \gamma_2 < 1$ and $tol > 0$
Choose initial second order matrix B_0 and a threshold radius Δ_t
$k = 0$
while $\Delta_k > tol$ **do**
  if $\Delta_k >= \Delta_t$ **then**
    $m_k(\alpha_k) = j(\alpha_k) + g_k^\top d_k + \frac{1}{2} d_k^\top B_k d_k$
  else
    $m_k(\alpha_k) = j(\alpha_k) + g_{\gamma,k}^\top d_k + \frac{1}{2} d_k^\top B_k d_k$
  end if
  Compute a step s_k that "sufficiently" reduces the model m_k such that $\alpha_k + s_k \in B_{\Delta_k}$
  Update second order matrix B_k using limited memory BFGS.
  Compute the quality measure ρ_k
  $\alpha_{k+1} = \begin{cases} \alpha_k & \text{if } \rho_k \leq \eta_1, \\ \alpha_k + s_k & \text{otherwise.} \end{cases}$
  $\Delta_{k+1} = \begin{cases} [\Delta_k, \infty) & \text{if } \rho_k \geq \eta_2, \\ [\gamma_2 \Delta_k, \Delta_k] & \text{if } \rho_k \in (\eta_1, \eta_2), \\ [\gamma_1 \Delta_k, \gamma_2 \Delta_k] & \text{if } \rho_k < \eta_1. \end{cases}$
  $k \leftarrow k + 1$
end while
return α_k

which corresponds to a classical trust-region sub-problem with additional positivity constraints (see, e.g., [32, 33]). The main idea is to reformulate the problem taking advantage of the l_∞ norm used for the ball at the point x_k

$$\begin{aligned} \min_{s \in \mathbb{R}^n} \quad & j(\alpha) + g^\top s + \frac{1}{2} s^\top B s \\ \text{s.t.} \quad & \max(-\alpha_j, -\Delta) \leq s_j \leq \Delta, \ \forall j = 1, \dots, n. \end{aligned} \tag{71}$$

For performance purposes it is desirable to solve this problem approximately in such a way that we can guarantee a descent on the cost function. With that goal in mind we will make use of a dogleg strategy that takes into account a Newton step s_N and a Cauchy step s_C. In the context of this constrained problem, let us take $\tilde{B}_\Delta = B_\Delta \cap \mathbb{R}^n_+$ and distinguish the following three cases

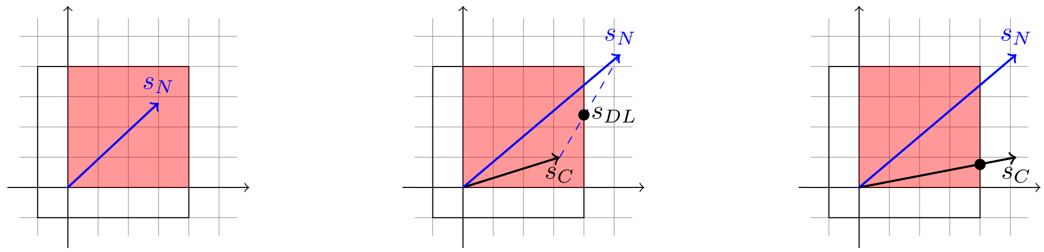
1. $s_N \in \tilde{B}_\Delta$,
2. $s_C \in \tilde{B}_\Delta$ and $s_N \notin \tilde{B}_\Delta$,
3. $s_C \notin \tilde{B}_\Delta$ and $s_N \notin \tilde{B}_\Delta$.

For case 1 we take the Newton step; for case 2 a dogleg strategy for box constraints is used; for case 3 we make use of a scaled Cauchy direction as described in algorithm 2.

6.2 Local Regularization

As a safeguard, for the cases where the trust-region radius fall below a threshold value Δ_t we will rely on a switching strategy between the nonsmooth model presented earlier and a regularized model. Let us consider the following optimization problem

$$\begin{aligned} \min_{\alpha \in \mathbb{R}^n_+} \quad & J(\bar{u}(\alpha), u^{train}) \\ \text{s.t.} \quad & \bar{u}(\alpha) = \arg\min_u \phi(u) + \sum_{j=1}^n \alpha_j \|(\mathbb{K}u)_j\|_\gamma \end{aligned} \tag{72}$$



(a) $s_N \in \tilde{B}_\Delta$ (b) $s_C \in \tilde{B}_\Delta$, $s_N \notin \tilde{B}_\Delta$ (c) $s_C \notin \tilde{B}_\Delta$, $s_N \notin \tilde{B}_\Delta$

Figure 1: Projected Cauchy Step: The figure depicts a two dimensional example for the three different possible cases when approximating the trust region subproblem using a dogleg strategy with l_∞ norm and positivity constraints.

Algorithm 2 Dogleg Step for Box Constraints

Calculate Newton step by solving the linear system $B_k s_N = -g_k$.
if $s_N \in \tilde{B}_\Delta$ **then**
  return s_N
end if
Calculate $s_C = -\frac{\|g_k\|^2}{g_k^\top * B_k * g_k} g_k$
if $s_C \in \tilde{B}_\Delta$ **then**
  Calculate the intersection of $s_N - s_C$ with $\tilde{B}_\Delta$ and obtain s_{DL}
  return s_{DL}
end if
Find t such that $t * s_C/\|s_C\|$ remains in $\tilde{B}_\Delta$.
return $t * \frac{s_C}{\|s_C\|}$

---

where $\|\cdot\|_\gamma$ is a local regularization of the euclidean norm given by

$$\|z\|_\gamma = \begin{cases} \|z\| - \frac{1}{2\gamma} & \text{if } \|z\| \geq \frac{1}{\gamma}, \\ \frac{\gamma}{2}\|z\|^2 & \text{if } \|z\| < \frac{1}{\gamma}. \end{cases}$$

Since the problem is now differentiable, we can define the following optimality condition for the lower level problem is obtained

$$\langle \phi'(u), v \rangle + \sum_{j=1}^{n} \alpha_j \langle h_\gamma((\mathbb{K}u)_j), (\mathbb{K}v)_j \rangle = 0, \ \forall v \in \mathbb{R}^m.$$

Here, h_γ correspond to the first derivative of the regularized euclidean norm. Moreover, regarding the bilevel problem, it is possible to formulate the KKT conditions for this problem. Let $(\alpha^*, u^*) \in \mathbb{R}^n_+ \times \mathbb{R}^m$ be a stationary point for the regularized bilevel problem eq. (72). Then there exist Lagrange multipliers $(p, \sigma) \in \mathbb{R}^m \times \mathbb{R}^n$ such that the following optimality system holds true

$$\langle \phi'(u), v \rangle + \sum_{j=1}^{n} \alpha_j^* \langle h_\gamma((\mathbb{K}u^*)_j), (\mathbb{K}v)_j \rangle = 0, \qquad \forall v \in \mathbb{R}^m, \tag{73}$$

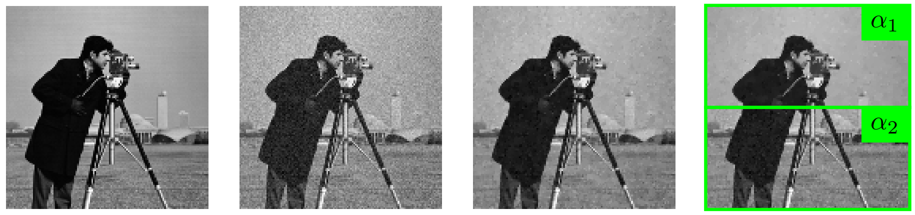
$$\langle \phi''(u^*)^\top p, v \rangle + \sum_{j=1}^{n} \alpha_j \langle h_\gamma'^*((\mathbb{K}u^*)_j)(\mathbb{K}p)_j, (\mathbb{K}v)_j \rangle = -\langle J'(u), v \rangle, \qquad \forall v \in \mathbb{R}^m, \tag{74}$$

$$\langle h_\gamma((\mathbb{K}u^*)_j), (\mathbb{K}p)_j \rangle = \sigma_j, \qquad \forall j = 1, \ldots, n, \tag{75}$$

$$0 \leq \sigma \perp \alpha^* \geq 0. \tag{76}$$

Moreover, with help of the adjoint equation eq. (74) it is possible to derive a gradient formula for the reduced cost function $j(\alpha) = J(u_\gamma(\alpha))$ as

$$(j'(\alpha))_j = \langle h_\gamma((\mathbb{K}u)_j), (\mathbb{K}p)_j \rangle.$$



Original
.

Noisy
SSIM=0.6091

$\alpha^* = 0.0155$
SSIM=0.7711

$\alpha^* = [0.0233; 0.0126]$
SSIM=0.8007

Figure 2: Cameraman Dataset
Optimal reconstructions using a scalar regularization parameter and a 2 dimensional regularization parameter.

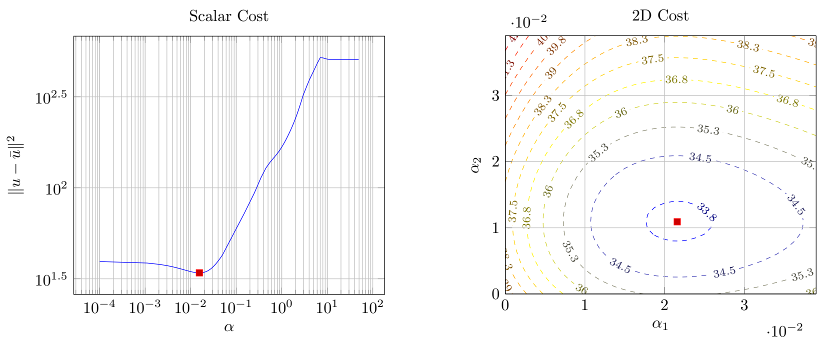


Figure 3: Cost Function - Cameraman Dataset
Values for the l_2 squared cost function using a scalar regularization parameter and a two dimensional regularization parameter using the Cameraman single image dataset.

According to algorithm 1, if the value for the radius runs below a predetermined threshold, the algorithm will make use of this gradient for its calculations.

7 Experiments

In this section we report on the performance of algorithm 1 presented in Section 6 to find optimal parameters. With this goal in mind we prepared two training image datasets.

7.1 Single Training Pair

The first dataset we will explore is a single 128 by 128 pixel image pair dataset based on the cameraman image and a corrupted version, obtained by adding gaussian noise with zero mean and standard deviation $\sigma = 0.05$. Figure 2 shows this training image pair along with the optimal parameter obtained using the trust-region algorithm for both a scalar and a two-dimensional patch parameter. The improvement when using a two-dimensional patch parameter, according to the SSIM value of the image reconstructions, in Figure 3. Moreover, the cost function when using a scalar and a two dimensional parameter is shown, respectively. These plots show the non-convexity of the cost function, and the cost corresponding to the optimal parameter.

In order to explore the behavior of the algorithm with more degrees of freedom we will make use of a so called *patch-based* parameter. This parameter will be a piecewise constant parameter. We split the image size into a grid of different size according to the number of patches, i.e., the patch size will be smaller as we require more patches within the image domain. For instance a 4x4 patch will calculate an optimal parameter of size 16.

Regarding the behavior of the algorithm, Table 1 shows the performance on the cameraman training dataset when using different number of patches, along with the number of iterations, cost, quality measures and residue. Indeed, the improvement on the reconstructed images when using more patches can be verified

Item				Reconstruction	
Patch	Iterations	$\|\alpha_{k+1} - \alpha_k\|$	COST	SSIM	PSNR
Scalar	19	7.629e-7	34.1206	0.7712	26.3794
2x1	20	4.883e-5	33.6915	0.8007	25.6222
2x2	19	4.639e-5	33.6625	0.8034	25.8982
4x4	22	1.461e-4	33.0175	0.8189	25.8303
8x8	28	5.614e-5	32.2040	0.8283	24.4066
16x16	36	9.119e-4	31.6976	0.8455	25.8264

Table 1: Trust Region Algorithm behavior on the Cameraman dataset.

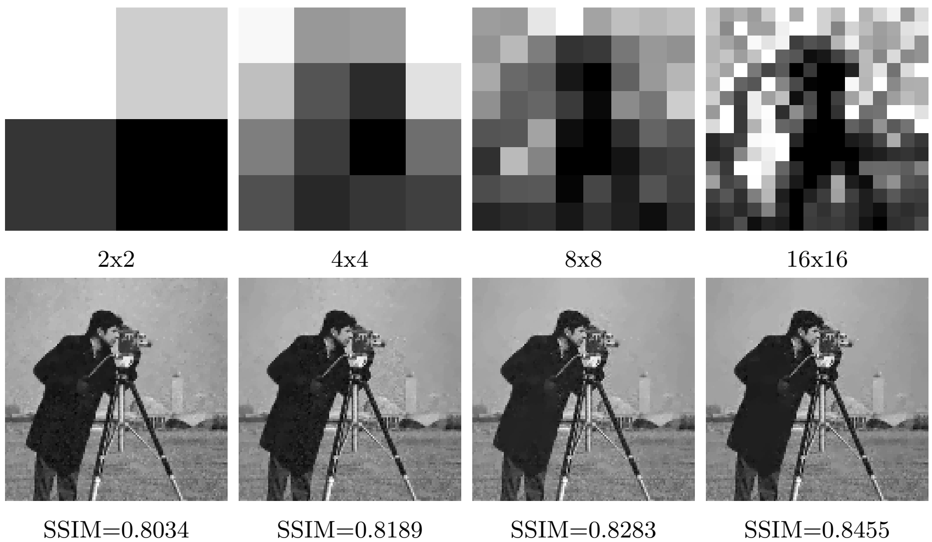


Figure 4: Optimal Patch Parameter - Cameraman Dataset
Values for the l_2 squared cost function using a scalar regularization parameter and a two dimensional regularization parameter using the Cameraman dataset.

from the data. Furthermore, Figure 4 shows the optimal reconstructions along with the optimal parameter for different number of patches. In this experiment it can be seen how the regularization parameter can adjust different regularization levels according to the training pair.

7.2 Multiple Training Pairs

For the second experiment, we used ten image pairs containing images of faces to generate a training dataset and ten different image pairs to generate a testing dataset; both datasets were based on the CelebA dataset [22]. These images are of size 128 by 128 pixels and in both datasets, the degenerated pairs were generated by adding gaussian noise with zero-mean and standard deviation $\sigma = 0.1$. A subset of the training dataset is depicted in Figure 5.

In fig. 6 we plot the cost function corresponding to a scalar parameter and two-dimensional patch parameter along with the cost function corresponding to the optimal value calculated by the algorithm. It is worth mentioning that when considering a patch-dependent parameter, as it was the case with the single image dataset, a scale dependent structure appears (see fig. 7).

For the training dataset proposed, the quality reconstruction results for different number of patches is shown in Table 2. This table shows the mean values for the SSIM and PSNR quality metric for the reconstructed images from the training dataset. Again, an improvement on the reconstruction quality can be seen as the degrees of freedom for the regularization parameter increases.

Finally, we can estimate the denoiser performance in images from the testing dataset. This experiment

Figure 5: Faces Dataset
A subset of the CelebA dataset corrupted with gaussian noise.

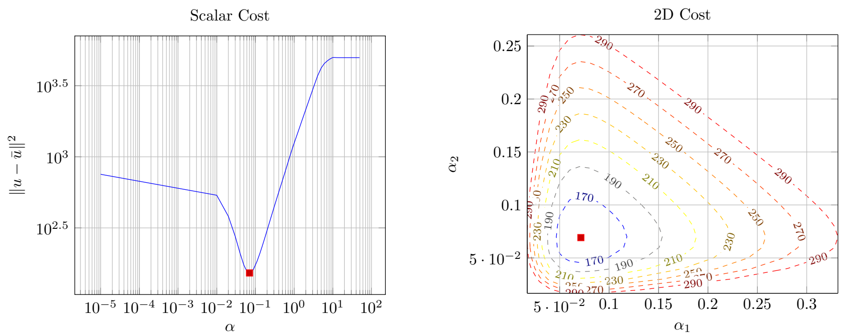


Figure 6: Cost Function - Faces Dataset
Values for the l_2 squared cost function using a scalar regularization parameter and a two dimensional regularization parameter using the Faces dataset.

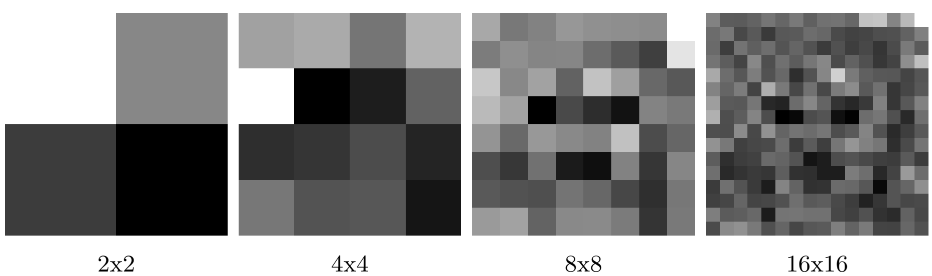


Figure 7: Optimal Patch Parameter - Faces Dataset
Values for the optimal parameters calculated for different parameter patch sizes.

Item				Reconstruction	
Patch	Iterations	$\|\alpha_{k+1}-\alpha_k\|$	COST	MSSIM	MPSNR
Scalar	14	1.562e-3	152.3726	0.8141	27.0689
2x1	17	1.049e-3	152.3209	0.8142	27.0589
2x2	21	1.762e-4	152.3120	0.8146	27.0739
4x4	21	7.421e-4	152.0911	0.8154	27.0766
8x8	24	7.050e-4	151.2011	0.8173	27.0787
16x16	25	1.410e-3	149.9673	0.8193	27.1047
32x32	35	3.285e-4	147.6573	0.8223	27.2141

Table 2: Trust Region Algorithm behavior on the Faces dataset.

img num	noisy	scalar	2x2	4x4	8x8	16x16	32x32
1	0.5247	0.7951	0.7948	0.7937	0.7956	0.7940	0.7930
2	0.4588	0.6614	0.6618	0.6616	0.6625	0.6625	0.6592
3	0.4267	0.7719	0.7721	0.7723	0.7747	0.7743	0.7710
4	0.3836	0.7237	0.7229	0.7214	0.7235	0.7233	0.7202
5	0.4580	0.7812	0.7799	0.7783	0.7798	0.7799	0.7768
6	0.4263	0.7400	0.7403	0.7403	0.7420	0.7428	0.7423
7	0.4547	0.6321	0.6322	0.6308	0.6305	0.6298	0.6262
8	0.4117	0.7381	0.7386	0.7405	0.7416	0.7413	0.7398
9	0.4655	0.7430	0.7409	0.7399	0.7405	0.7388	0.7354
10	0.5081	0.8210	0.8210	0.8204	0.8211	0.8197	0.8181
MSSIM		**0.7408**	**0.7405**	**0.7400**	**0.7412**	**0.7406**	**0.7382**

Table 3: Faces Dataset SSIM Quality Measures in the validation dataset.

will show the *overfitting* phenomena that may occur when dealing with large number of patches, including the case of scale-dependent parameters ($\alpha \in \mathbb{R}^n$). Indeed, it can be seen in the testing dataset an increment on the mean SSIM (MSSIM) for the reconstructed images from the testing dataset up to a 8x8 patch size. Any higher number of patches results in quality degradation. This is indeed the expected behavior when dealing with overfitting problems.

7.3 Learning Optimal Total Variation Discretization

The selection of an adequate discretization of the total variation seminorm in the context of image reconstruction problems is still an open problem [4, 1]. In [5], the authors propose a methodology for finding optimal discretization where instead of using hand-crafted discretization schemes for the total variation seminorm, a learning strategy is proposed.

The bilevel framework presented in this work, can also be used to learn optimal gradient discretization, by using different discretization schemes and their corresponding regularization parameters. We will make use of a training dataset to estimate the optimal regularization parameters for the contributions of each discretization scheme into the final solution. Let us define the following variational denoising model

$$\min_{u \in \mathbb{R}^n} \mathcal{E}(u) := \frac{1}{2}\|u-f\|^2 + \sum_{j=1}^{n} (\alpha_1)_j \|(\mathbb{K}_1 u)_j\| + \sum_{j=1}^{n} (\alpha_2)_j \|(\mathbb{K}_2 u)_j\| + \sum_{j=1}^{n} (\alpha_3)_j \|(\mathbb{K}_3 u)_j\|, \tag{77}$$

where $\mathbb{K}_1, \mathbb{K}_2$ and $\mathbb{K}_3$ are the forward, backward and centered finite differences discretization of the gradient operator respectively. The goal is to determine optimal parameters $(\alpha_1, \alpha_2, \alpha_3)^\top \in \mathbb{R}_+^{3\times n}$ that lead to an optimal discretization of the total variation operator. Indeed, we consider the following bilevel learning strategy

$$\begin{aligned} \min_{(\alpha_1,\alpha_2,\alpha_3)^\top \in \mathbb{R}_+^n \times \mathbb{R}_+^n \times \mathbb{R}_+^n} \quad & \frac{1}{2}\|\bar{u}(\alpha) - u^{\text{train}}\|^2 \\ \text{s.t.} \quad & \bar{u} = \operatorname*{arg\,min}_{u \in \mathbb{R}^m} \mathcal{E}(u), \end{aligned} \tag{78}$$

By extending the model presented previously, we can make use of a similar analysis using the Bouligand candidate presented in theorem 5.1 by defining the following adjoint state

$$\langle p, v\rangle + \sum_{i=1}^{3} \sum_{j \in \mathcal{I}^i} \langle \mu_j^i, (\mathbb{K}^i v)_j\rangle - \langle \nabla J(u), v\rangle = 0, \forall v \in \mathcal{V}$$

$$\mu_j^i - \alpha_j^i T_j^i (\mathbb{K}^i p)_j = 0,\ \forall j \in \mathcal{I}^i,\ i = 1,2,3,$$

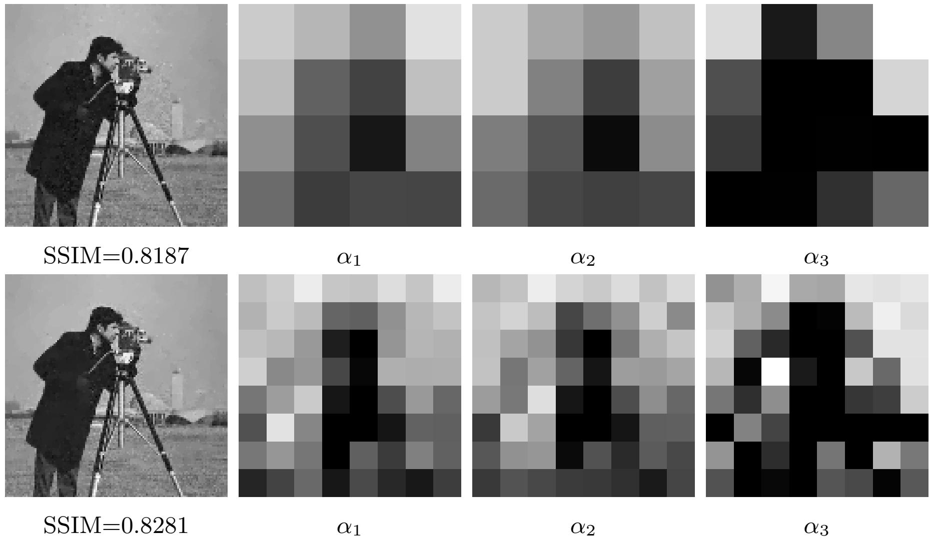


Figure 8: Optimal Patch Parameters - Cameraman Dataset
Values for the optimal parameter calculated for the Cameraman dataset for different patch sizes.

where $\mathcal{V} := \{v \in \mathbb{R}^m : (\mathbb{K}^i v)_j = 0,\ \forall \mathcal{A}_s^i \cup \mathcal{B}_1^i,\ (\mathbb{K}^i v)_j \in span(q_j^i),\ \forall j \in \mathcal{B}_2^i,\ i = 1, 2, 3\}$, and the following gradient form

$$\langle j'(\alpha^i), h^i \rangle = - \sum_{j \in \mathcal{I}^i \cup \mathcal{B}_2^i} \frac{h_j^i}{\alpha_j^i} \langle q_j^i, (\mathbb{K}^i p)_j \rangle \text{ for } i = 1, 2, 3.$$

The same procedure applies for the regularized problem. Using the KKT conditions for the differentiable problem we obtain

$$\langle p, v \rangle + \sum_{i=1}^{3} \sum_{j=1}^{n} \alpha_j^i \langle h_\gamma'^* ((\mathbb{K}^i u)_j)(\mathbb{K}^i p)_j, (\mathbb{K}^i v)_j \rangle = -\langle \nabla J(u), v \rangle,\ \forall v \in \mathbb{R}^m,$$

and its corresponding gradient characterization

$$(j'(\alpha^i))_j = \langle h_\gamma((\mathbb{K}^i u)_j), (\mathbb{K}^i p)_j \rangle,\ \forall i = 1, 2, 3.$$

The optimal regularization parameters for 4x4 and 8x8 patch-parameters for each of the three regularizers considered is presented in fig. 8 for the cameraman dataset and fig. 9 for the faces dataset. Again, we can see an improvement on the image quality by using more patches in the training datasets. Finally, in table 4 we can see a comparison between different patch sizes when using these three regularizers. Again, it is worth noticing that in the validation dataset the best reconstruction quality corresponds to an 8x8 patch-parameter size, and any further enlargement of the patch dimension drives a lower reconstruction quality.

References

[1] Corentin Caillaud and Antonin Chambolle. "Error estimates for finite differences approximations of the total variation". In: (2020).

[2] Luca Calatroni, Juan Carlos De Los Reyes, and Carola-Bibiane Schönlieb. "Dynamic sampling schemes for optimal noise learning under multiple nonsmooth constraints". In: *IFIP Conference on System Modeling and Optimization*. Springer. 2013, pp. 85–95.

[3] Luca Calatroni et al. "Bilevel approaches for learning of variational imaging models". In: *Variational Methods*. Walter de Gruyter GmbH, 2017, pp. 252–290.

[4] Antonin Chambolle and Thomas Pock. "Crouzeix–Raviart approximation of the total variation on simplicial meshes". In: *Journal of Mathematical Imaging and Vision* 62.6 (2020), pp. 872–899.

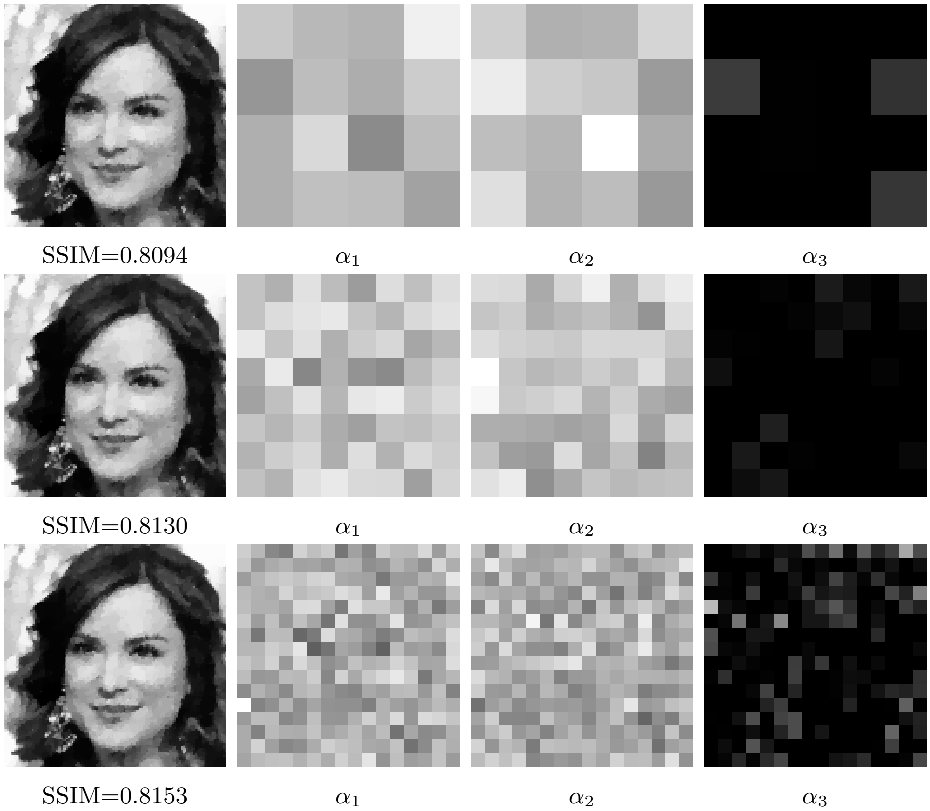


Figure 9: Optimal Patch Parameters - Faces Dataset
Values for the optimal parameters calculated for different parameter patch sizes.

img num	noisy	scalar	2x2	4x4	8x8	16x16
1	0.5247	0.7922	0.7952	0.7951	0.7977	0.7965
2	0.4588	0.6568	0.6676	0.6663	0.6649	0.6653
3	0.4267	0.7745	0.7742	0.7729	0.7748	0.7741
4	0.3836	0.7237	0.7239	0.7226	0.7244	0.7206
5	0.4580	0.7814	0.7796	0.7779	0.7806	0.7809
6	0.4263	0.7394	0.7406	0.7414	0.7432	0.7441
7	0.4547	0.6262	0.6372	0.6357	0.6338	0.6333
8	0.4117	0.7377	0.7380	0.7413	0.7429	0.7421
9	0.4655	0.7438	0.7481	0.7445	0.7432	0.7420
10	0.5081	0.8185	0.8225	0.8217	0.8219	0.8211
MSSIM		**0.7395**	**0.7420**	**0.7419**	**0.7428**	**0.7424**

Table 4: Faces Dataset SSIM Quality Measures - Optimal Gradient Discretization in the validation dataset

[5] Antonin Chambolle and Thomas Pock. "Learning Consistent Discretizations of the Total Variation". In: (2020).

[6] Constantin Christof, Juan Carlos De los Reyes, and Christian Meyer. "A Nonsmooth Trust-Region Method for Locally Lipschitz Functions with Application to Optimization Problems Constrained by Variational Inequalities". In: *SIAM Journal on Optimization* 30.3 (2020), pp. 2163–2196.

[7] M D'Elia, JC De Los Reyes, and A Miniguano-Trujillo. "Bilevel Parameter Learning for Nonlocal Image Denoising Models". In: *Journal of Mathematical Imaging and Vision* (2021), pp. 1–23.

[8] Juan Carlos De los Reyes. *Numerical PDE-Constrained Optimization.* Springer, 2015.

[9] Juan Carlos De los Reyes, C-B Schönlieb, and Tuomo Valkonen. "Bilevel parameter learning for higher-order total variation regularisation models". In: *Journal of Mathematical Imaging and Vision* 57.1 (2017), pp. 1–25.

[10] Juan Carlos De los Reyes, C-B Schönlieb, and Tuomo Valkonen. "The structure of optimal parameters for image restoration problems". In: *Journal of Mathematical Analysis and Applications* 434.1 (2016), pp. 464–500.

[11] Juan Carlos De los Reyes and Carola-Bibiane Schönlieb. "Image Denoising: Learning the noise model via nonsmooth pde-constrained optimization." In: *Inverse Problems & Imaging* 7.4 (2013).

[12] Juan Carlos De los Reyes and David Villacís. "Bilevel Optimization Methods in Imaging". In: *Handbook of Mathematical Models and Algorithms in Computer Vision and Imaging* 33.7 (2021), p. 074005.

[13] Yiqiu Dong, Michael Hintermüller, and M Monserrat Rincon-Camacho. "Automated regularization parameter selection in multi-scale total variation models for image restoration". In: *Journal of Mathematical Imaging and Vision* 40.1 (2011), pp. 82–104.

[14] Ivar Ekeland and Roger Temam. *Convex analysis and variational problems.* Vol. 28. SIAM, 1999.

[15] Roland Glowinski and J Tinsley Oden. "Numerical methods for nonlinear variational problems". In: *Journal of Applied Mechanics* 52 (1985), p. 739.

[16] Michael Hintermüller, Konstantinos Papafitsoros, and Carlos N Rautenberg. "Analytical aspects of spatially adapted total variation regularisation". In: *Journal of Mathematical Analysis and Applications* 454.2 (2017), pp. 891–935.

[17] Michael Hintermüller and Kostas Papafitsoros. "Generating structured nonsmooth priors and associated primal-dual methods". In: *Handbook of Numerical Analysis.* Vol. 20. Elsevier, 2019, pp. 437–502.

[18] Michael Hintermüller and Tao Wu. "Bilevel optimization for calibrating point spread functions in blind deconvolution". In: (2015).

[19] Johannes Jahn. *Introduction to the theory of nonlinear optimization.* Springer Nature, 2020.

[20] Teresa Klatzer and Thomas Pock. "Continuous hyper-parameter learning for support vector machines". In: *Computer Vision Winter Workshop (CVWW).* 2015, pp. 39–47.

[21] Karl Kunisch and Thomas Pock. "A bilevel optimization approach for parameter learning in variational models". In: *SIAM Journal on Imaging Sciences* 6.2 (2013), pp. 938–983.

[22] Ziwei Liu et al. "Deep Learning Face Attributes in the Wild". In: *Proceedings of International Conference on Computer Vision (ICCV).* Dec. 2015.

[23] Zhi-Quan Luo, Jong-Shi Pang, and Daniel Ralph. *Mathematical programs with equilibrium constraints.* Cambridge University Press, 1996.

[24] Peter Ochs et al. "Techniques for gradient-based bilevel optimization with non-smooth lower level problems". In: *Journal of Mathematical Imaging and Vision* 56.2 (2016), pp. 175–194.

[25] Jirí V Outrata. "A generalized mathematical program with equilibrium constraints". In: *SIAM Journal on Control and Optimization* 38.5 (2000), pp. 1623–1638.

[26] René Ranftl and Thomas Pock. "A deep variational model for image segmentation". In: *German Conference on Pattern Recognition.* Springer. 2014, pp. 107–118.

[27] R Tyrrell Rockafellar and Roger J-B Wets. *Variational analysis*. Vol. 317. Springer Science & Business Media, 2009.

[28] Stefan Scholtes. *Introduction to piecewise differentiable equations*. Springer Science & Business Media, 2012.

[29] David Strong and Tony Chan. "Edge-preserving and scale-dependent properties of total variation regularization". In: *Inverse problems* 19.6 (2003), S165.

[30] David M Strong and Tony F Chan. "Spatially and scale adaptive total variation based regularization and anisotropic diffusion in image processing". In: *Diusion in Image Processing, UCLA Math Department CAM Report*. Citeseer. 1996.

[31] Cao Van Chung, JC De los Reyes, and CB Schönlieb. "Learning optimal spatially-dependent regularization parameters in total variation image denoising". In: *Inverse Problems* 33.7 (2017), p. 074005.

[32] C Voglis and IE Lagaris. "A rectangular trust region dogleg approach for unconstrained and bound constrained nonlinear optimization". In: *WSEAS International Conference on Applied Mathematics*. Vol. 7. Citeseer. 2004.

[33] L Xu and J Burke. "ASTRAL: An active set linfinity-trust-region algorithm for box constrained optimization". In: *Optimization online, July* (2007).