ESAIM: COCV 32 (2026) 37
https://doi.org/10.1051/cocv/2026011

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

NO-GAP SECOND-ORDER CONDITIONS FOR MINIMIZATION PROBLEMS IN SPACES OF MEASURES

Gerd Wachsmuth[1] and Daniel Walter[2,*]

Abstract. Over the last years, minimization problems over spaces of measures have received increased interest due to their relevance in the context of inverse problems, optimal control and machine learning. A fundamental role in their numerical analysis is played by the assumption that the optimal dual state admits finitely many global extrema and satisfies a second-order sufficient optimality condition in each one of them. In this work, we show the full equivalence of these structural assumptions to a no-gap second-order condition involving the second subderivative of the Radon norm as well as to a local quadratic growth property of the objective functional with respect to the bounded Lipschitz norm.

Mathematics Subject Classification. 46E27, 49K27, 49J52, 49J53.

Received March 18, 2025. Accepted February 2, 2026.

1. Introduction

In this paper, second-order necessary and sufficient optimality conditions for sparse minimization problems of the form

$$\min_u J(u) := L(Ku) + \alpha \|u\|_{\mathcal{M}} \tag{P}$$

are studied. Here, $\|\cdot\|_{\mathcal{M}}$ denotes the canonical norm on the space of Radon measures $\mathcal{M}(\Omega)$ on a compact spatial domain Ω, $\alpha > 0$, $L\colon Y \to \mathbb{R}$ is a smooth, but not necessarily convex, loss function on a Hilbert space Y of observations and $K\colon \mathcal{M}(\Omega) \to Y$ is a linear and continuous control-to-observation operator given by

$$Ku = \int_\Omega k(x)\,\mathrm{d}u(x) \qquad \forall u \in \mathcal{M}(\Omega) \tag{1.1}$$

for some sufficiently smooth $k\colon \Omega \to Y$. Problems of this or similar form appear in a variety of challenging settings in the context of inverse problems [1, 2], such as source identification in acoustics [3, 4] or microscopy [5, 6], optimal control [7–9], or the training of shallow neural networks [10], and optimal sensor placement [11, 12]. This is largely attributed to the observation that the appearance of the nonsmooth Radon norm in the objective

Keywords and phrases: Optimization in measure space, second-order optimality conditions, second subderivatives, Wasserstein distance.

[1] Institute of Mathematics, Brandenburgische Technische Universität Cottbus–Senftenberg, 03046 Cottbus, Germany.
[2] Institut für Numerische Mathematik, Johannes Kepler Universität Linz, 4040 Linz, Austria.

[*] Corresponding author: daniel.walter@jku.at

functional promotes *sparse minimizers*, *i.e.*, solutions given as a finite linear combination of Dirac measures

$$\bar{u} = \sum_{j=1}^{N} \bar{\lambda}_j \delta_{\bar{x}_j} \quad \text{with} \quad \bar{x}_j \in \Omega,\ \bar{\lambda}_j \in \mathbb{R} \setminus \{0\}, \tag{1.2}$$

where $\delta_{\bar{x}_j} \in \mathcal{M}(\Omega)$ is the Dirac measure in the point $\bar{x}_j$, *i.e.*, $\int_\Omega \varphi \, \mathrm{d}\delta_{\bar{x}_j} = \varphi(\bar{x}_j)$ holds for all continuous functions φ on Ω. From an analytic perspective, the sparsity promoting property of the Radon norm follows, *e.g.*, by convex representer theorems, if the image of K is finite dimensional, [13, 14], or it can be deduced from the first-order necessary optimality condition for the Jordan decomposition $\bar{u} = \bar{u}_+ - \bar{u}_-$,

$$\operatorname{supp} \bar{u}_\pm \subset \{x \in \Omega \mid \bar{p}(x) = \pm\alpha\} \quad \text{as well as} \quad |\bar{p}(x)| \le \alpha \quad \text{on } \Omega \tag{1.3}$$

provided that the *dual variable* $\bar{p} = -(k(\cdot), \nabla L(K\bar{u}))_Y \in C(\Omega)$ admits finitely many extrema, see [15], Proposition 3.8.

Up to now, previous works have focused on convex loss functions L. In this context, given a sparse control $\bar{u}$, represented as in (1.2), with associated dual variable $\bar{p}$ satisfying

$$\operatorname{supp} \bar{u} = \{x \in \Omega \mid |\bar{p}(x)| = \alpha\} = \{\bar{x}_j\}_{j=1}^N \subset \operatorname{int}(\Omega), \tag{1.4}$$

we call $\bar{p}$ *non-degenerate* if we have

$$\dim(\operatorname{span}\{k(\bar{x}_1), \dots, k(\bar{x}_N)\}) = N \quad \text{and} \quad \operatorname{sgn}(\bar{\lambda}_j) \nabla^2 \bar{p}(\bar{x}_j) \le_L -\theta \operatorname{Id} \quad \forall j = 1, \dots, N, \tag{1.5}$$

where $\theta > 0$ and "$\le_L$" denotes the Loewner order. Non-degeneracy plays a fundamental role in the analysis of convex sparse minimization problems allowing, *e.g.*, a fine analysis of the asymptotic regularization properties of the Radon norm [16], the derivation of sharp *a priori* error estimates for discretizations of (P), or the proof of fast convergence rates for problem-tailored numerical solution algorithms [15, 17, 18].

1.1. Contribution

Given a sparse measures $\bar{u}$, *i.e.* a measure in the form (1.2), which satisfies the first-order necessary optimality condition (1.3) together with its associated dual variable $\bar{p}$, we investigate *no-gap* second-order conditions for general, not necessarily convex sparse minimization problems of the form (P). By the latter, we understand a condition on suitable second-order derivatives at $\bar{u}$ which is equivalent to a local quadratic growth behavior of the objective functional J around $\bar{u}$. For this purpose, an equivalent "lifted" version of Problem (P) onto $C^1(\Omega)^*$, the topological dual space of the continuous differentiable functions, is considered,

$$\min_{h \in C^1(\Omega)^*} L(\mathcal{K}(h)) + G(h),$$

involving suitable extensions $\mathcal{K}\colon C^1(\Omega)^* \to Y$ as well as $G\colon C^1(\Omega)^* \to (-\infty, \infty]$ which satisfy

$$\mathcal{K}(\imath_{\mathcal{M}} u) = Ku, \quad G(\imath_{\mathcal{M}} u) = \alpha \|u\|_{\mathcal{M}} \quad \forall u \in \mathcal{M}(\Omega) \quad \text{and} \quad G(h) = \infty \quad \forall h \in C^1(\Omega)^* \setminus \mathcal{M}(\Omega),$$

where $\imath_{\mathcal{M}}$ denotes the compact embedding of $\mathcal{M}(\Omega)$ into $C^1(\Omega)^*$. Our abstract main result, Theorem 5.17, clarifies the connection between structural properties of the dual variable $\bar{p}$, in particular its curvature around $\operatorname{supp} \bar{u}$, the second order condition

$$(\mathcal{K}h, \nabla^2 L(K\bar{u})\mathcal{K}h)_Y + G''(\imath_{\mathcal{M}} \bar{u}, \bar{p}; h) > 0 \qquad \forall h \in C^1(\Omega)^* \setminus \{0\},$$

involving the *weak* second subderivative* $G''(\imath_{\mathcal{M}}\bar{u}, \bar{p}; \cdot)\colon C^1(\Omega)^* \to [-\infty, \infty]$ of G at $\imath_{\mathcal{M}}\bar{u}$ for $\bar{p}$, as well as local quadratic growth of the objective functional in the vicinity of $\bar{u}$ w.r.t. the bounded Lipschitz norm

$$\|u\|_{\mathrm{BL}} = \sup_{\|\varphi\|_{\mathrm{Lip}} \le 1} \int_\Omega \varphi \, \mathrm{d}u \quad \text{where} \quad \|\varphi\|_{\mathrm{Lip}} \coloneqq \max\{\|\varphi\|_C, \mathrm{lip}(\varphi)\},$$

and $\mathrm{lip}(\varphi)$ denotes the Lipschitz constant. While we emphasize that Theorem 5.17 covers nonconvex loss functions L and contains no restriction to settings in which $\bar{p}$ only admits a finite number of global extrema, our general analysis reveals non-degeneracy in the sense of (1.5) as a special case. Indeed, Corollary 5.19 shows the equivalence of the following statements under the assumption that L is strongly convex and that the strict complementarity condition (1.4) holds.

1. The dual variable $\bar{p}$ is non-degenerate, *i.e.*, it satisfies (1.5).
2. There holds

$$(\mathcal{K}h, \nabla^2 L(K\bar{u})\mathcal{K}h)_Y + G''(\imath_{\mathcal{M}}\bar{u}, \bar{p}; h) > 0 \qquad \forall h \in C^1(\Omega)^* \setminus \{0\}, \tag{1.6}$$

   and

$$\text{for all } (t_k) \subset (0,\infty),\ (h_k) \subset C^1(\Omega)^* \text{ with } t_k \searrow 0,\ h_k \overset{*}{\rightharpoonup} 0 \text{ and } \|h_k\|_{(C^1)^*} = 1, \text{ we have}$$
$$\liminf_{k\to\infty} \left(\frac{1}{t_k^2} (G(\imath_{\mathcal{M}}\bar{u} + t_k h_k) - G(\imath_{\mathcal{M}}\bar{u})) - \frac{1}{t_k} \langle\!\langle \bar{p}, h_k \rangle\!\rangle \right) > 0$$

   where $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ denotes the duality pairing between $C^1(\Omega)$ and $C^1(\Omega)^*$.
3. There exist $\gamma > 0$ and $\varepsilon > 0$ such that there holds

$$J(u) - J(\bar{u}) \ge \gamma \|u - \bar{u}\|_{\mathrm{BL}}^2 \qquad \forall u \in \mathcal{M}(\Omega), \|u - \bar{u}\|_{\mathrm{BL}} \le \varepsilon. \tag{1.7}$$

A main ingredient of our derivation of Theorem 5.17 is an explicit formula for $G''(\imath_{\mathcal{M}}\bar{u}, \bar{p}; \cdot)$ which we obtain as a byproduct of showing the weak* twice epi-differentiable of G. Moreover, we also conclude that the sparsity of $\bar{u}$ is a fundamental requirement for our results: Indeed quadratic growth w.r.t. the bounded Lipschitz norm around a measure $\bar{u}$ can only hold if $\bar{u}$ is a finite sum of Dirac measures, see Lemma 5.24. Given the prevalence and utility of regularity conditions in the spirit of non-degeneracy outlined above, we believe that our thorough characterization in terms of, both, no-gap second order conditions as well as quadratic growth will facilitate future work and will allow to extend previous results to more complex settings, *e.g.*, to nonconvex problems. Moreover, proving the connection to second-order conditions based on second subderivatives allows for the application of a variety of tools established for abstract minimization problems. For example, [19] provides abstract stability and sensitivity results for perturbed minimization problems under the assumption of no-gap SSCs.

1.2. Related work

The present work is mainly influenced by prior work on no-gap second-order conditions utilizing the notion of second subderivatives. This is a well-known concept in finite-dimensional variational analysis, see the classical treatment in [20], Definition 13.3 as well as the recent contribution [21]. In the infinite-dimensional setting, this derivative has been used in [22] to characterize the directional differentiability of projections. More recently, [23], second subderivatives have been utilized to derive second-order conditions for infinite-dimensional constrained minimization problems. Subsequently, this approach has been further developed in the contributions [24, 25].

In the present work, we follow an implicit approach to computing $G''(\imath_{\mathcal{M}}\bar{u}, \bar{p}; \cdot)$ in order to avoid the necessity of working on the dual space $C^1(\Omega)^*$. More in detail, we first prove the strong-strong twice epi-differentiability of the preadjoint of G, which is now a functional over $C^1(\Omega)$, and derive a candidate for $G''(\imath_{\mathcal{M}}\bar{u}, \bar{p}; \cdot)$ *via* arguments from convex duality. Conceptually, this is closest related to the exposition in [25] where the authors consider a similar treatment for integral functionals defined on $L^1(\Omega)$. Finally, we also mention [19] where the authors address the differentiability of parameter-to-solution maps associated to parameter-dependent variational inequalities by using second subderivatives.

The notion of non-degeneracy of certain dual variables for convex, sparse minimization problems was possibly first coined in the seminal paper [16] where it was used to address the recovery properties of Radon norm variational regularization in inverse problems. Similar applications can be found in, *e.g.* [26–28]. Subsequently, these structural assumptions have also become a central pillar for the analysis of finite element discretizations of PDE-constrained optimal control problems with sparse controls, see, *e.g.* [29], as well as the derivation of fast convergence results for numerical algorithms such as generalized conditional gradient or exchange type methods [15, 18], or over-parametrized gradient descent [17]. In this context, we also point out that [17], Proposition 3.2 deduces local quadratic growth w.r.t. the bounded Lipschitz norm. As outlined above, for strongly convex loss functions L, our general result provides the equivalence of non-degeneracy, quadratic growth and no-gap SSCs, thus complementing these earlier results.

If Y is finite dimensional and L is convex, we also note that Problem (P) can be interpreted as the Fenchel dual of a semi-infinite, state-constrained problem. From this perspective, $\bar{u}$ corresponds to the Lagrange multiplier for the state constraint imposed on the optimal state $\bar{p}$. Moreover, (1.4) can be interpreted as a strict complementarity condition. In this context, assumptions on the curvature of $\bar{p}$ in the vicinity of $\operatorname{supp}\bar{u}$, have been utilized, *e.g.*, for the derivation of approximation error estimates [30].

1.3. Outline

After introducing the necessary notation in Section 2, we collect some basic but important results on both, sparse minimization problems and second-order conditions based on second subderivatives, in Sections 3 and 4, respectively. Subsequently, the latter are applied to Problem (P). Our main abstract result, Theorem 5.17, as well as tangible conclusions based upon it are stated in Section 5. The following sections, Sections 6 and 7, are dedicated to its technical proof. Finally, Appendix A contains some required auxiliary results used throughout the paper.

2. Notation

Let d be a positive integer. In the following, let $\Omega \subset \mathbb{R}^d$ be compact and assume that Ω equals the closure of its interior. By $\mathcal{B}(\Omega)$ we denote the Borel σ-algebra of Ω. We further introduce the space of signed Radon measures $\mathcal{M}(\Omega)$ on Ω which we identify with the topological dual space of $C(\Omega)$ where the latter is equipped with the maximum norm

$$\|\varphi\|_C \coloneqq \max\{|\varphi(x)| \mid x \in \Omega\} \quad \forall \varphi \in C(\Omega).$$

The corresponding duality pairing is given by

$$\langle \varphi, u \rangle = \int_\Omega \varphi \, \mathrm{d}u \quad \forall \varphi \in C(\Omega), u \in \mathcal{M}(\Omega).$$

We equip $\mathcal{M}(\Omega)$ with the canonical dual norm

$$\|u\|_{\mathcal{M}} \coloneqq \max\{\langle \varphi, u \rangle \mid \|\varphi\|_C \le 1\} \quad \forall u \in \mathcal{M}(\Omega)$$

making it a Banach space. A measure $u \in \mathcal{M}(\Omega)$ is called positive, abbreviated by $u \ge 0$, if we have $\langle \varphi, u \rangle \ge 0$ for all $\varphi \in C(\Omega)$ with $\varphi(x) \ge 0$ for all $x \in \Omega$. The cone of positive measures is denoted by $\mathcal{M}(\Omega)^+$.

For a function $\varphi\colon \Omega \to \mathbb{R}$, we define its Lipschitz constant

$$\operatorname{lip}(\varphi) \coloneqq \sup\left\{ \frac{\varphi(x_1) - \varphi(x_2)}{|x_1 - x_2|} \;\middle|\; x_1, x_2 \in \Omega, x_1 \neq x_2 \right\}. \tag{2.1}$$

In particular, φ is Lipschitz continuous if and only if $\operatorname{lip}(\varphi) < \infty$.

The space $C^1(\Omega)$ consists of all functions $\varphi \in C(\Omega)$ which are continuously differentiable on $\operatorname{int}(\Omega)$ such that $\nabla\varphi$ can be continuously extended from $\operatorname{int}(\Omega)$ to Ω. Under the canonical norm

$$\|\varphi\|_{C^1} = \max\{\|\varphi\|_C, \|\nabla\varphi\|_{C^d}\} \qquad \text{where} \qquad \|\varphi\|_{C^d} \coloneqq \sup\{|\varphi(x)| \mid x \in \Omega\} \qquad \forall \varphi \in C^1(\Omega),$$

the space $C^1(\Omega)$ becomes a Banach space, see [31], Section 1.6. Here, $|\cdot|$ is the Euclidean norm on $\mathbb{R}^d$.

For all $x \in \mathbb{R}^d$ and $r > 0$, we denote by

$$B_r(x) := \{y \in \mathbb{R}^d \mid |x - y| \leq r\}$$

the closed ball around x with radius r.

Finally, we abbreviate $Av^2 \coloneqq v^\top Av$ for $A \in \mathbb{R}^{d\times d}$ and $v \in \mathbb{R}^d$. Similarly, if $b\colon X \times X \to \mathbb{R}$ is a bilinear mapping, we abbreviate $bh^2 := b[h, h]$. This will be used, in particular, for second derivatives.

3. Minimization problems in spaces of measures

In this section, we collect some preliminary results on minimization problems of the form (P). Where possible, proofs are omitted for the sake of brevity.

Assumption 3.1. Throughout the paper, we assume that:

- **A1** The space of observations Y is a separable Hilbert space with norm $\|\cdot\|_Y = \sqrt{(\cdot,\cdot)_Y}$.
- **A2** The functional $L\colon Y \to \mathbb{R}$ is of class $\mathcal{C}^2$.
- **A3** The kernel $k\colon \Omega \to Y$ satisfies $k \in C^{1,\gamma}(\Omega; Y)$ for some $0 < \gamma \leq 1$.
- **A4** The set $\Omega \subset \mathbb{R}^d$ is compact, it equals the closure of its interior $\operatorname{int}(\Omega)$ and $\operatorname{int}(\Omega)$ is uniformly locally quasiconvex, see Definition 5.1.

First, we note that the regularity of k implies weak*-to-strong continuity of the integral operator K.

Lemma 3.2. *Let Assumption 3.1 hold. Then the operator $K\colon \mathcal{M}(\Omega) \to Y$ as defined in (1.1) is linear and sequentially weak*-to-strong continuous. Moreover, we have $K = (K_*)^*$ where the preadjoint $K_*\colon Y \to C(\Omega)$ is given by*

$$[K_* y](x) = (k(x), y)_Y \quad \forall y \in Y,\ x \in \Omega.$$

Proof. Note that K_* is linear and bounded since $k(x) \in C(\Omega; Y)$. Moreover, we observe

$$\langle K_* y, u\rangle = \int_\Omega (k(x), y)_Y \,\mathrm{d}u(x) = \left(y, \int_\Omega k(x)\,\mathrm{d}u(x)\right)_Y = (y, Ku)_Y$$

for all $y \in Y$ and $u \in \mathcal{M}(\Omega)$. Thus, we have $(K_*)^* = K$. In particular, K is linear and bounded as well as sequentially weak*-to-weak continuous. Let $(u_k) \subset \mathcal{M}(\Omega)$ be a weak* convergent sequence with limit $\bar{u}$. Then, we have

$$\|Ku_k\|_Y^2 = \langle K_* K u_k, u_k\rangle.$$

We readily verify that $K_* K u_k$ converges weakly to $K_* K \bar{u}$ in $C(\Omega)$. Moreover, since $k \in C^{0,\gamma}(\Omega; Y)$, the sequence $(K_* K u_k)$ is uniformly bounded in $C^{0,\gamma}(\Omega)$. Hence, due to Arzelà–Ascoli theorem, we conclude $K_* K u_k \to K_* K \bar{u}$ in $C(\Omega)$ and, finally,

$$\lim_{k\to\infty} \|K u_k\|_Y^2 = \lim_{k\to\infty} \langle K_* K u_k, u_k \rangle = \langle K_* K \bar{u}, \bar{u} \rangle = \|K \bar{u}\|_Y^2$$

Together with $K u_k \rightharpoonup K \bar{u}$ in Y, this shows the strong convergence $K u_k \to K \bar{u}$ in Y. Consequently, K is weak*-to-strong continuous. □

In what follows, we denote by $\alpha \partial \|u\|_{\mathcal{M}}$ the convex subdifferential of $\alpha \|\cdot\|_{\mathcal{M}}$ at $u \in \mathcal{M}(\Omega)$. Next, we summarize necessary and sufficient first-order optimality conditions for minimizers of Problem (P).

Proposition 3.3. *Given $\bar{u} \in \mathcal{M}(\Omega)$, define*

$$\bar{p} = -K_* \nabla L(K \bar{u}) = -(k(\cdot), \nabla L(K \bar{u}))_Y \in C^1(\Omega).$$

If the measure $\bar{u}$ is a minimizer of (P), then the following equivalent conditions hold.

1. *We have $\bar{p} \in \alpha \partial \|\bar{u}\|_{\mathcal{M}}$.*
2. *There holds $|\bar{p}(x)| \le \alpha$ for all $x \in \Omega$ as well as $\langle \bar{p}, \bar{u} \rangle = \alpha \|\bar{u}\|_{\mathcal{M}}$.*
3. *The Jordan decomposition $\bar{u} = \bar{u}_+ - \bar{u}_-$ satisfies*

$$\operatorname{supp} \bar{u}_\pm \subset \{x \in \Omega \mid \bar{p}(x) = \pm\alpha\} \quad \textit{as well as} \quad |\bar{p}| \le \alpha \quad \textit{on } \Omega$$

If L is convex and if these conditions hold, then $\bar{u}$ is a minimizer of (P).

4. No-gap second-order conditions for abstract minimization problems

Next, we summarize tangible results on no-gap second-order conditions for nonsmooth minimization problems. Subsequently, these will be specialized to the problem at hand. More in detail, we consider

$$\min_{x\in X} \left[F(x) + G(x)\right]. \tag{4.1}$$

We start by fixing the setting.

Assumption 4.1. The following assumptions are made throughout this section.

1. There holds $X = Z^*$ for a separable Banach space Z.
2. The functional $G\colon X \to (-\infty, \infty]$ and the point $\bar{x} \in \operatorname{dom}(G)$ are given.
3. Associated with the functional $F\colon \operatorname{dom}(G) \to \mathbb{R}$, there exist $F'(\bar{x}) \in Z$ and a bounded bilinear form $F''(\bar{x})\colon X \times X \to \mathbb{R}$ such that

$$\lim_{k\to\infty} \frac{F(\bar{x} + t_k h_k) - F(\bar{x}) - t_k F'(\bar{x}) h_k - \frac{1}{2} t_k^2 F''(\bar{x}) h_k^2}{t_k^2} = 0 \tag{4.2}$$

   holds for all sequences $(t_k) \subset (0,\infty)$, $(h_k) \subset X$ satisfying $t_k \searrow 0$, $h_k \xrightharpoonup{*} h \in X$ and $\bar{x} + t_k h_k \in \operatorname{dom}(G)$.

Loosely speaking, the "smooth" part F of the objective admits a second-order Taylor expansion around $\bar{x}$. The bilinear form $F''(\bar{x})$ can be interpreted as its Hessian. Regarding the potentially nonsmooth term G, we rely on the notion of the weak* second subderivative.

Definition 4.2. Let $x \in \operatorname{dom}(G)$ and $p \in Z$ be given. The *weak* second subderivative*

$$G''(x,p;\cdot)\colon X \to [-\infty,\infty]$$

of G at x for p is given by

$$G''(x,p;h) \coloneqq \inf\left\{ \liminf_{k\to\infty} \frac{G(x+t_kh_k)-G(x)-t_k\langle p,h_k\rangle}{t_k^2/2} \;\middle|\; t_k \searrow 0,\ h_k \overset{*}{\rightharpoonup} h \right\}.$$

The functional G is called *strictly twice epi-differentiable* at x for p if for every $h \in X$ and every sequence $(t_k) \subset (0,\infty)$ with $t_k \searrow 0$ there exists a sequence $(h_k) \subset X$ with

$$G''(x,p;h) = \lim_{k\to\infty} \frac{G(x+t_kh_k)-G(x)-t_k\langle p,h_k\rangle}{t_k^2/2}, \quad h_k \overset{*}{\rightharpoonup} h,\ \|h_k\|_X \to \|h\|_X.$$

We use the following result, see [24], Theorem 2.20.

Theorem 4.3. *Let Assumption 4.1 hold and assume that $h \mapsto F''(\bar{x})h^2$ is sequentially weak* continuous. Then the following assertions are equivalent.*

(i) There exist $\varepsilon > 0$ and $\gamma > 0$ satisfying the second-order growth condition

$$(F+G)(x) - (F+G)(\bar{x}) \geq \frac{\gamma}{2}\|x-\bar{x}\|_X^2 \qquad \forall x \in X, \|x-\bar{x}\|_X \leq \varepsilon.$$

(ii) We have the sufficient second-order condition

$$F''(\bar{x})h^2 + G''(\bar{x},-F'(\bar{x});h) > 0 \qquad \forall h \in X \setminus \{0\} \tag{SSC}$$

as well as the non-degeneracy condition

$$\begin{gathered} \text{for all } (t_k) \subset (0,\infty),\ (h_k) \subset X \text{ with } t_k \searrow 0,\ h_k \overset{*}{\rightharpoonup} 0 \text{ and } \|h_k\|_X = 1, \text{ we have} \\ \liminf_{k\to\infty}\left(\frac{1}{t_k^2}\big(G(\bar{x}+t_kh_k)-G(\bar{x})\big) + \langle F'(\bar{x}), h_k/t_k\rangle + \frac{1}{2}F''(\bar{x})h_k^2\right) > 0. \end{gathered} \tag{NDC}$$

We point out that this equivalence still holds under slightly weaker assumptions, see, *e.g.*, the stated reference for more details. From a practical perspective and for particular examples, it is desirable to link the abstract conditions (SSC) and (NDC) to tangible structural properties of the problem under consideration. The following auxiliary results will prove useful in this regard. Loosely speaking, these allow to circumvent the necessity to work on the potentially complicated dual space X by studying related properties of suitable pre-conjugates $H\colon Z \to (-\infty,\infty]$, $H^* = G$. For this purpose, we require the following strong analogue of Definition 4.2 for H.

Definition 4.4. Let $p \in Z$ and $x \in X$ be given. The *strong second subderivative*

$$H''(p,x;\cdot)\colon Z \to [-\infty,\infty]$$

of H at p for $x \in X$ is given by

$$H''(p,x;z) = \inf\left\{ \liminf_{k\to\infty} \frac{H(p+t_kz_k)-H(p)-t_k\langle z_k,x\rangle}{t_k^2/2} \;\middle|\; t_k \searrow 0,\ z_k \to z \text{ in } Z \right\}.$$

Accordingly, the functional H is called strongly-strongly twice epi-differentiable at p for x if for every $z \in Z$ and every sequence $(t_k) \subset (0,\infty)$ with $t_k \searrow 0$, there exists a sequence $(z_k) \subset Z$ with

$$H''(p,x;z) = \lim_{k\to\infty} \frac{H(p+t_k z_k) - H(p) - t_k\langle z_k, x\rangle}{t_k^2/2}, \quad z_k \to z.$$

We start by recalling that the convex conjugate of $H''(p,x;\cdot)$ provides a lower bound on $G''(x,p;\cdot)$, see [25], Lemma 4.4.

Lemma 4.5. *Assume that $H\colon Z \to (-\infty,\infty]$ is proper, convex, lower semicontinuous and satisfies $H^* = G$. Moreover, let $x \in \operatorname{dom}(G)$ and $p \in Z$ with $p \in \partial G(x)$ be given. If H is strongly-strongly twice epi-differentiable at p for x, then there holds*

$$\frac{1}{2} G''(x,p;h) \ge \left(\frac{1}{2} H''(p,x;\cdot)\right)^*(h) \qquad \forall h \in X.$$

Second, we provide a sufficient condition for (NDC) involving H. Note that this generalizes [25], Lemma 4.9.

Lemma 4.6. *Assume that $H\colon Z \to (-\infty,\infty]$ is proper, convex, lower semicontinuous and satisfies $H^* = G$. Moreover, let $\bar p \in Z$ with $\bar p \in \partial G(\bar x)$ as well as $m \in \mathbb{N}$ be given. Fix some linearly independent $\zeta_1,\ldots,\zeta_m \in Z$ and define $H_2\colon Z \to (-\infty,\infty]$ via*

$$H_2\left(\sum_{i=1}^m \beta_i\zeta_i\right) = \frac{1}{2}\sum_{i=1}^m \beta_i^2 + \left\langle \sum_{i=1}^m \beta_i \zeta_i, \bar x\right\rangle$$

on $\operatorname{span}\{\zeta_1,\ldots,\zeta_m\}$ and $H_2(z) = \infty$ outside of this linear hull. Let $\tilde H = H \oplus H_2$ denote the infimal convolution and suppose that

$$\tilde H(p) \le H(\bar p) + \langle p - \bar p, \bar x\rangle + \frac{\Lambda}{2}\|p-\bar p\|_Z^2 \qquad \forall p \in Z, \|p-\bar p\|_Z \le \eta \tag{4.3}$$

holds for some $\Lambda, \eta > 0$. Then

$$G(x) \ge G(\bar x) + \langle \bar p, x - \bar x\rangle + \frac{1}{2\Lambda}\|x-\bar x\|_X^2 - \frac{1}{2}\sum_{i=1}^m |\langle \zeta_i, x-\bar x\rangle|^2 \qquad \forall x \in X, \|x-\bar x\|_X \le \eta\Lambda. \tag{4.4}$$

If further $X \ni h \mapsto F''(\bar x)h^2 \in \mathbb{R}$ is sequentially weak lower semicontinuous and $\bar p = -F'(\bar x)$, then (NDC) holds.*

Proof. We have $(H \oplus H_2)^* = G + H_2^*$, see [32], Exercise 4.5.7. The conjugate of H_2 is

$$H_2^*(x) = \sup\left\{ \left\langle \sum_{i=1}^m \beta_i\zeta_i, x\right\rangle - \left(\frac{1}{2}\sum_{i=1}^m \beta_i^2 + \left\langle \sum_{i=1}^m \beta_i\zeta_i, \bar x\right\rangle\right) \;\middle|\; \beta_i \in \mathbb{R}\right\} = \frac{1}{2}\sum_{i=1}^m |\langle \zeta_i, x-\bar x\rangle|^2.$$

Thus, we have $0 \in \partial H_2^*(\bar x)$ and $\bar p \in \partial(G + H_2^*)(\bar x) = \partial \tilde H^*(\bar x)$. Now, the Fenchel–Young identity gives

$$\tilde H(\bar p) = \langle \bar p, \bar x\rangle - (G + H_2^*)(\bar x) = \langle \bar p, \bar x\rangle - G(\bar x) = H(\bar p).$$

Hence, we can replace $H(\bar{p})$ in (4.3) by $\tilde{H}(\bar{p})$. Next, we follow the proof of [25], Lemma 4.9 which yields

$$\tilde{H}^*(x) \geq \tilde{H}^*(\bar{x}) + \langle \bar{p}, x - \bar{x} \rangle + \frac{1}{2\Lambda} \|x - \bar{x}\|_X^2 \qquad \forall x \in X, \|x - \bar{x}\|_X \leq \eta\Lambda.$$

The identity $\tilde{H}^* = G + H_2^*$ implies (4.4). Finally, let sequences $(t_k) \subset (0, \infty)$, $(h_k) \subset X$ with $t_k \searrow 0$, $h_k \stackrel{*}{\rightharpoonup} 0$ in X and $\|h_k\|_X = 1$ be given. Then, we utilize (4.4) to get

$$\liminf_{k \to \infty} \left(\frac{G(\bar{x} + t_k h_k) - G(\bar{x}) - \langle \bar{p}, t_k h_k \rangle}{t_k^2} + \frac{1}{2} F''(\bar{x}) h_k^2 \right) \geq \liminf_{k \to \infty} \left(\frac{1}{2\Lambda} - \frac{1}{2} \sum_{i=1}^m |\langle \zeta_i, h_k \rangle|^2 \right) = \frac{1}{2\Lambda},$$

where we used $h_k \stackrel{*}{\rightharpoonup} 0$ and the assumed property of $F''(\bar{x})$. This shows that (NDC) is satisfied. □

5. No-gap second-order conditions in spaces of measures

In the remainder of the paper, we will use the auxiliary results of Section 4 to investigate the abstract conditions (SSC) and (NDC) for a suitable interpretation of Problem (P) and an appropriate choice of the variable space. In this section, before formalizing our results, we therefore introduce an equivalent lifting of Problem (P) to $X = C^1(\Omega)^*$, the dual space of the continuous differentiable functions on Ω.

5.1. Uniform local Quasiconvexity

In what follows, we recall the notion of "uniform local quasiconvexity" appearing in Assumption 3.1 and discuss its consequences.

Definition 5.1. We say that a set $D \subset \mathbb{R}^d$ is *uniformly locally quasiconvex* with constants $r > 0$ and $C \geq 1$ if for any $x, y \in D$ with $|y - x| \leq r$ there exists a curve $\gamma \in C([0,1]; D)$ with $\gamma(0) = x$, $\gamma(1) = y$ and Lipschitz constant at most $C|y - x|$.

Remark 5.2. Typically, it is required that the curve $\gamma \in C([0,1]; D)$ from Definition 5.1 is rectifiable and of length at most $C|y - x|$ (instead of being Lipschitz continuous), see [33], p. 1223. However, one can perform an arc-length reparametrization to ensure Lipschitz continuity of γ.

It is clear that convex sets are uniformly locally quasiconvex. Further, the next result shows that Definition 5.1 is a rather weak requirement.

Lemma 5.3. *Let $D \subset \mathbb{R}^d$ be open and bounded. Further, we require that D is a Lipschitz set in the sense that for every point $p \in \partial D$, there exists $r > 0$ and a bijection $l_p \colon B_r(p) \to B_1(0)$ such that l_p and l_p^{-1} are Lipschitz,*

$$l_p(D \cap B_r(p)) = \{z \in B_1(0) \mid z_n > 0\} =: B_1^+(0) \quad \textit{and} \quad l_p(\partial D \cap B_r(p)) = \{z \in B_1(0) \mid z_n = 0\}.$$

Then, D is uniformly locally quasiconvex.

The proof of Lemma 5.3 can be found in Appendix A. We give two important consequences of Definition 5.1. For the proofs, we again refer to Appendix A. The first lemma shows that functions from $C^1(\Omega)$ are Lipschitz continuous.

Lemma 5.4. *Let $\operatorname{int}(\Omega)$ be uniformly locally quasiconvex with constants $r > 0$ and $C \geq 1$. Then, every $\varphi \in C^1(\Omega)$ is Lipschitz continuous and*

$$\operatorname{lip}(\varphi) \leq \max\left\{ \frac{2}{r} \|\varphi\|_C, C \|\nabla \varphi\|_{C^d} \right\} \leq \max\{2r^{-1}, C\} \|\varphi\|_{C^1}.$$

The second lemma shows that we have a uniform Taylor expansion.

Lemma 5.5. *Assume that* $\operatorname{int}(\Omega)$ *is uniformly locally quasiconvex and let* $\varphi \in C^1(\Omega)$ *be given. For any* $\varepsilon > 0$, *there exists* $\delta > 0$ *such that*

$$\left|\varphi(y) - \varphi(x) - \nabla\varphi(x)^\top (y - x)\right| \le \varepsilon |y - x| \qquad \forall x, y \in \Omega, |y - x| \le \delta.$$

Remark 5.6. In case that $\Omega \subset \mathbb{R}^d$ is a domain, [33], Theorem 7 suggests that the assertions of Lemmas 5.4 and 5.5 hold if and only if Ω is uniformly locally quasiconvex. For compact sets $\Omega \subset \mathbb{R}^d$, we have seen that it is sufficient that $\operatorname{int}(\Omega)$ is uniformly locally quasiconvex. It is a natural question whether it is also necessary in this context. This is indeed not the case: Consider $\Omega = [-1, 0]^2 \cup [0, 1]^2$. Clearly, $\operatorname{int}(\Omega) = (-1, 0)^2 \cup (0, 1)^2$ is not uniformly locally quasiconvex. However, we can apply Lemmas 5.4 and 5.5 on $[-1, 0]^2$ and $[0, 1]^2$ separately and, consequently, the desired properties hold on Ω. On the contrary, we provide another example in Appendix A, Example A.1, showing that both conclusions, Lemmas 5.4 and 5.5, can fail in the absence of local uniform quasiconvexity.

5.2. A primer on the dual space of $C^1(\Omega)$ and the relation to the bounded Lipschitz norm

We start by giving a precise characterization of the space $C^1(\Omega)^*$ and of its canonical dual norm. For this purpose, note that

$$C^1(\Omega) \cong C \coloneqq \left\{(\varphi, \nabla\varphi) \;\middle|\; \varphi \in C^1(\Omega)\right\} \subset C(\Omega) \times C(\Omega; \mathbb{R}^d),$$

where the product space on the right-hand side is equipped with the norm

$$\|(\varphi, \phi)\|_{C \times C^d} = \max\left\{\|\varphi\|_C, \|\phi\|_{C^d}\right\} \qquad \forall (\varphi, \phi) \in C(\Omega) \times C(\Omega; \mathbb{R}^d).$$

Note that the mapping $C^1(\Omega) \ni \varphi \mapsto (\varphi, \nabla\varphi) \in C(\Omega) \times C(\Omega; \mathbb{R}^d)$ is an isometry. One readily verifies that the dual space of $C(\Omega) \times C(\Omega; \mathbb{R}^d)$ is given by $\mathcal{M}(\Omega) \times \mathcal{M}(\Omega; \mathbb{R}^d)$ with associated duality pairing

$$\langle \varphi, u\rangle + \langle \phi, v\rangle \qquad \forall (\varphi, \phi) \in C(\Omega) \times C(\Omega; \mathbb{R}^d),\ (u, v) \in \mathcal{M}(\Omega) \times \mathcal{M}(\Omega; \mathbb{R}^d)$$

and norm

$$\|(u, v)\|_{\mathcal{M}(\Omega) \times \mathcal{M}(\Omega; \mathbb{R}^d)} = \|u\|_{\mathcal{M}} + \|v\|_{\mathcal{M}^d} = \|u\|_{\mathcal{M}} + \||v|\|_{\mathcal{M}},$$

where $|v| \in \mathcal{M}(\Omega)$ is the variation of the vector-valued measure v (w.r.t. the Euclidean norm on its image space $\mathbb{R}^d$).

The subset C is complete since it is isometrically isomorphic to the complete space $C^1(\Omega)$. Thus, C is a closed subspace of $C(\Omega) \times C(\Omega; \mathbb{R}^d)$. Consequently, its dual space can be characterized using the annihilator

$$C^\perp \coloneqq \left\{(\mu, \nu) \in \mathcal{M}(\Omega) \times \mathcal{M}(\Omega; \mathbb{R}^d) \;\middle|\; \langle \varphi, \mu\rangle + \langle \nabla\varphi, \nu\rangle = 0 \quad \forall (\varphi, \nabla\varphi) \in C\right\}.$$

Lemma 5.7 ([34], Proposition 11.10)**.** *There holds*

$$(\mathcal{M}(\Omega) \times \mathcal{M}(\Omega; \mathbb{R}^d))/C^\perp \cong C^1(\Omega)^*. \tag{5.1}$$

The isometric isomorphism is given by

$$(\mathcal{M}(\Omega) \times \mathcal{M}(\Omega; \mathbb{R}^d))/C^\perp \ni (u, v) + C^\perp \mapsto \left(C^1(\Omega) \ni \varphi \mapsto \langle \varphi, u\rangle + \langle \nabla\varphi, v\rangle\right) \in C^1(\Omega)^*.$$

Consequently, we identify the elements of $C^1(\Omega)^*$ with equivalence classes

$$[\![u,v]\!] := (u,v) + C^\perp = \{(u,v) + (\mu,\nu) \mid (\mu,\nu) \in C^\perp\} \in (\mathcal{M}(\Omega) \times \mathcal{M}(\Omega;\mathbb{R}^d))/C^\perp$$

of measures $(u,v) \in \mathcal{M}(\Omega) \times \mathcal{M}(\Omega;\mathbb{R}^d)$ which act on functions $\varphi \in C^1(\Omega)$ *via* the duality pairing

$$\langle\!\langle \varphi, [\![u,v]\!] \rangle\!\rangle := \langle \varphi, u\rangle + \langle \nabla\varphi, v\rangle.$$

In what follows, we will always employ the identification (5.1), *i.e.*, bounded functionals on $C^1(\Omega)$ are always identified with the corresponding equivalence classes. Consequently, the norm on $C^1(\Omega)^*$ is given by the quotient norm

$$\|[\![u,v]\!]\|_{(C^1)^*} = \sup_{\|\varphi\|_{C^1}\le 1} \langle\!\langle \varphi, [\![u,v]\!] \rangle\!\rangle = \inf\{\|u+\mu\|_{\mathcal{M}} + \|v+\nu\|_{\mathcal{M}^d} \mid (\mu,\nu) \in C^\perp\}.$$

In the subsequent sections, we frequently encounter elements of the form $[\![u,0]\!]$ with $u \in \mathcal{M}(\Omega) = C(\Omega)^*$. Note that $[\![u,0]\!]$ is the restriction of the functional $u\colon C(\Omega) \to \mathbb{R}$ to the domain $C^1(\Omega)$. In what follows, we give a simpler representation of the dual norm for such elements $[\![u,0]\!]$. For this purpose, we denote by $\operatorname{Lip}(\Omega)$ the space of Lipschitz continuous functions on Ω which we equip with the canonical norm

$$\|\varphi\|_{\mathrm{Lip}} := \max\{\|\varphi\|_C, \operatorname{lip}(\varphi)\},$$

where we used the Lipschitz constant $\operatorname{lip}(\varphi)$ of φ defined in (2.1). Naturally, every $u \in \mathcal{M}(\Omega)$ defines a linear function on $\operatorname{Lip}(\Omega)$. The associated dual norm

$$\|u\|_{\mathrm{BL}} = \sup_{\|\varphi\|_{\mathrm{Lip}}\le 1} \langle \varphi, u\rangle \tag{5.2}$$

will be referred to as the *bounded Lipschitz* norm. Due to our standing assumption that $\operatorname{int}(\Omega)$ is uniformly locally quasiconvex, the space $C^1(\Omega)$ is embedded into the Lipschitz functions, see Lemma 5.4, and we obtain the equivalence of the bounded Lipschitz norm with the natural norm of $C^1(\Omega)$.

Lemma 5.8. *The norms $\|[\![\cdot,0]\!]\|_{(C^1)^*}$ and $\|\cdot\|_{\mathrm{BL}}$ are equivalent on $\mathcal{M}(\Omega)$.*

Proof. From Lemma 5.4 and the definition of $\|\cdot\|_{\mathrm{Lip}}$ we have

$$\|\varphi\|_{\mathrm{Lip}} \le L\|\varphi\|_{C^1} \qquad \forall \varphi \in C^1(\Omega)$$

for some constant $L > 0$. This yields

$$\|[\![u,0]\!]\|_{(C^1)^*} = \sup_{\|\varphi\|_{C^1}\le 1} \langle\!\langle \varphi, [\![u,0]\!] \rangle\!\rangle = \sup_{\|\varphi\|_{C^1}\le 1} \langle \varphi, u\rangle \le \sup_{\|\varphi\|_{\mathrm{Lip}}\le L} \langle \varphi, u\rangle = L\|u\|_{\mathrm{BL}}.$$

For the reverse estimate, let $\varphi \in \operatorname{Lip}(\Omega)$ with $\|\varphi\|_{\mathrm{Lip}} \le 1$ be given. Using classical arguments, φ can be extended to all of $\mathbb{R}^d$ while preserving its Lipschitz constant. Now, let $\varepsilon > 0$ be given and denote by Φ_ε a mollifier supported on $B_\varepsilon(0) \subset \mathbb{R}^d$. Then, $\psi := \Phi_\varepsilon \star \varphi$ satisfies $\psi \in C^1(\Omega)$, $\|\psi\|_{C^1} \le 1$ and $\|\psi - \varphi\|_C \le \varepsilon$. For an arbitrary $u \in \mathcal{M}(\Omega)$, we get

$$\langle \varphi, u\rangle \le \varepsilon\|u\|_{\mathcal{M}} + \langle \psi, u\rangle = \varepsilon\|u\|_{\mathcal{M}} + \langle\!\langle \psi, [\![u,0]\!] \rangle\!\rangle \le \varepsilon\|u\|_{\mathcal{M}} + \|u\|_{(C^1)^*}.$$

Taking the limit $\varepsilon \searrow 0$ and the supremum w.r.t. φ, this shows $\|u\|_{\mathrm{BL}} \le \|u\|_{(C^1)^*}$. ☐

Moreover, we require the notion of the Wasserstein-1 distance between positive measures of the same mass.

Definition 5.9. Given two measures $\mu, \nu \in \mathcal{M}(\Omega)$ with $\mu, \nu \geq 0$ and $\mu(\Omega) = \nu(\Omega)$, we define the Wasserstein-1 distance between μ and ν as

$$W_1(\mu, \nu) = \min\left\{ \int_{\Omega\times\Omega} |x - y| \,\mathrm{d}\gamma(x, y) \;\middle|\; \gamma \in \Pi(\mu, \nu) \right\},$$

where the minimum is taken over all couplings $\gamma \in \Pi(\mu, \nu) \subset \mathcal{M}(\Omega \times \Omega)$ between μ and ν, *i.e.*, there holds $\gamma \geq 0$ as well as

$$\gamma(B \times \Omega) = \mu(B) \quad \text{and} \quad \gamma(\Omega \times B) = \nu(B)$$

for all Borel sets $B \in \mathcal{B}(\Omega)$.

The well known Kantorovich duality yields

$$W_1(\mu, \nu) = \sup\left\{ \int_\Omega \psi \,\mathrm{d}(\mu - \nu) \;\middle|\; \operatorname{lip}(\psi) \leq 1 \right\}.$$

Using this, we can derive an optimal transport based representation of the bounded Lipschitz norm.

Lemma 5.10. *Fix $u \in \mathcal{M}(\Omega)$ and let $u_1, u_2 \in \mathcal{M}(\Omega)^+$ be such that $u = u_1 - u_2$. Then, we have*

$$\|u\|_{\mathrm{BL}} = \min_{\substack{\mu,\nu\geq 0\\ \mu(\Omega)=\nu(\Omega)}} \left[W_1(\mu, \nu) + \|u_1 - \mu\|_{\mathcal{M}} + \|u_2 - \nu\|_{\mathcal{M}} \right]. \tag{5.3}$$

In particular, there holds

$$\|u\|_{\mathrm{BL}} = \min_{\substack{\mu,\nu\geq 0\\ \mu(\Omega)=\nu(\Omega)}} \left[W_1(\mu, \nu) + \|u_+ - \mu\|_{\mathcal{M}} + \|u_- - \nu\|_{\mathcal{M}} \right],$$

where $u = u_+ - u_-$ denotes the Jordan decomposition of u.

Proof. The formula (5.3) will be proven by Fenchel duality. We define the sets

$$\begin{aligned}
X &:= C(\Omega)^3, & D &:= \{(\psi_1, \psi_2, \psi_3) \in X \mid \operatorname{lip}(\psi_1) \leq 1,\ \|\psi_2\|_C \leq 1,\ \|\psi_3\|_C \leq 1\},\\
Y &:= C(\Omega)^2, & M &:= \{(\varphi_1, \varphi_2) \in Y \mid \varphi_1 \leq 0,\ \varphi_2 \leq 0\},
\end{aligned}$$

as well as the mappings $A\colon X \to Y$, $f\colon X \to (-\infty, \infty]$ and $g\colon Y \to (-\infty, \infty]$ *via*

$$\begin{aligned}
A(\psi_1, \psi_2, \psi_3) &:= (\psi_1 - \psi_2, -\psi_1 - \psi_3), & g(\varphi_1, \varphi_2) &:= I_M(\varphi_1, \varphi_2),\\
f(\psi_1, \psi_2, \psi_3) &:= I_D(\psi_1, \psi_2, \psi_3) + \langle \psi_2, u_1 \rangle + \langle \psi_3, u_2 \rangle.
\end{aligned}$$

The functionals f and g are convex and lower semicontinuous and A is linear and bounded. The set $A \operatorname{dom}(f)$ contains the constant function $(-1, -1)$ and this is a continuity point of g. Hence, Fenchel duality gives

$$j_1 := \inf_{(\psi_1,\psi_2,\psi_3)\in X} f(\psi_1, \psi_2, \psi_3) + g(A(\psi_1, \psi_2, \psi_3)) = -\min_{(\mu,\nu)\in Y^*} f^*(-A^*(\mu, \nu)) + g^*(\mu, \nu) =: j_2,$$

see [32], Theorem 4.4.3. It remains to compute j_1 and j_2. For j_1 we get

$$\begin{aligned} j_1 &= \inf_{(\psi_1,\psi_2,\psi_3)\in X} f(\psi_1,\psi_2,\psi_3) + g(A(\psi_1,\psi_2,\psi_3)) \\ &= \inf\{\langle \psi_2, u_1\rangle + \langle \psi_3, u_2\rangle \mid (\psi_1,\psi_2,\psi_3) \in D,\ -\psi_3 \le \psi_1 \le \psi_2\} \\ &= \inf\{\langle \psi_1, u_1\rangle + \langle -\psi_1, u_2\rangle \mid (\psi_1,\psi_1,-\psi_1) \in D\}, \end{aligned}$$

where we used that $u_1, u_2 \ge 0$. Now, $(\psi_1,\psi_1,-\psi_1) \in D$ is equivalent to $\|\psi_1\|_{\mathrm{Lip}} \le 1$ and, thus,

$$j_1 = -\|u\|_{\mathrm{BL}}.$$

To evaluate j_2, we first compute

$$g^*(\mu,\nu) = \sup\{\langle \varphi_1, \mu\rangle + \langle \varphi_2, \nu\rangle \mid (\varphi_1,\varphi_2) \in M\} = I_P(\mu,\nu),$$

where we abbreviate $P = \mathcal{M}(\Omega)^+ \times \mathcal{M}(\Omega)^+$. Moreover,

$$\begin{aligned} f^*(-A^*(\mu,\nu)) &= f^*(\nu-\mu,\mu,\nu) \\ &= \sup\{\langle \psi_1, \nu-\mu\rangle + \langle \psi_2, \mu\rangle + \langle \psi_3, \nu\rangle - f(\psi_1,\psi_2,\psi_3) \mid (\psi_1,\psi_2,\psi_3) \in X\} \\ &= \sup\{\langle \psi_1, \nu-\mu\rangle + \langle \psi_2, \mu - u_1\rangle + \langle \psi_3, \nu - u_2\rangle \mid (\psi_1,\psi_2,\psi_3) \in D\} \\ &= \begin{cases} W_1(\mu,\nu) + \|u_1-\mu\|_{\mathcal{M}} + \|u_2-\nu\|_{\mathcal{M}} & \text{if } \mu(\Omega) = \nu(\Omega) \\ \infty & \text{if } \mu(\Omega) \ne \nu(\Omega) \end{cases} \end{aligned}$$

for all $(\mu,\nu) \in P$. Consequently,

$$j_2 = -\min\{f^*(-A^*(\mu,\nu)) \mid (\mu,\nu) \in P\} = -\min_{\substack{\mu,\nu\ge 0 \\ \mu(\Omega)=\nu(\Omega)}} [W_1(\mu,\nu) + \|u_1-\mu\|_{\mathcal{M}} + \|u_2-\nu\|_{\mathcal{M}}].$$

This finishes the proof of (5.3). □

A different proof for (5.3) can be obtained by using the ideas from [35], Proposition 24.

Example 5.11. Let $u = \delta_{x_1} - \delta_{x_2}$ with $x_1, x_2 \in \Omega$ and $x_1 \ne x_2$ be given. Then there holds $\|u\|_{\mathcal{M}} = 2$ but

$$\|u\|_{\mathrm{BL}} = \min\{2, |x_1 - x_2|\},$$

see [35], Example 22. Similarly, we can check that for $x \in \Omega$, $h \in \mathbb{R}^d \setminus \{0\}$ with $x \pm h \in \Omega$ and $u = 2\delta_x - \delta_{x+h} - \delta_{x-h}$ we have $\|u\|_{\mathcal{M}} = 4$ but

$$\|u\|_{\mathrm{BL}} = 2\min\{2, |h|\}.$$

Remark 5.12. In the case that the diameter of Ω is small enough and $u(\Omega) = 0$, we even have

$$\|u\|_{\mathrm{BL}} = W_1(u_+, u_-).$$

Indeed, let $\psi \in C(\Omega)$ with $\mathrm{lip}(\psi) \le 1$ be given such that $W_1(u_+,u_-) = \int_\Omega \psi \,\mathrm{d}u$. Note that the existence of ψ follows from the Arzelà-Ascoli theorem. W.l.o.g., we can assume $|\psi| \le \mathrm{diam}(\Omega)/2$. Consequently,

$$W_1(\mu,\nu) - W_1(u_+,u_-) \ge \int_\Omega \psi \,\mathrm{d}(\mu - \nu - u) = \int_\Omega \psi \,\mathrm{d}(\mu - u_+) - \int_\Omega \psi \,\mathrm{d}(\nu - u_-)$$

$$\geq -\frac{\operatorname{diam}(\Omega)}{2}(\|\mu - u_+\|_{\mathcal{M}} + \|\nu - u_-\|_{\mathcal{M}}).$$

This shows that the choice $\mu = u_+$ and $\nu = u_-$ in (5.3) is optimal whenever $\operatorname{diam}(\Omega) \leq 2$.

5.3. A lifted setting

In the following, let the canonical injection of $\mathcal{M}(\Omega)$ into $C^1(\Omega)^*$ be denoted by

$$\Pi\colon \mathcal{M}(\Omega) \to C^1(\Omega)^* \quad \text{where} \quad \Pi u := [\![u, 0]\!].$$

Note that the continuity of this injection follows from the definitions. Moreover, we define the lifted penalty term and forward operator by

$$G\colon C^1(\Omega)^* \to [0, \infty] \quad \text{with} \quad G([\![u, v]\!]) := \begin{cases} \alpha\|\tilde{u}\|_{\mathcal{M}} & \exists \tilde{u} \in \mathcal{M}(\Omega)\colon [\![u, v]\!] = \Pi\tilde{u} \\ \infty & \text{else} \end{cases} \tag{5.4}$$

as well as

$$\mathcal{K}\colon C^1(\Omega)^* \to Y \quad \text{with} \quad \mathcal{K}([\![u, v]\!]) := \int_\Omega k(x)\,\mathrm{d}u(x) + \int_\Omega Dk(x)\,\mathrm{d}v(x) \tag{5.5}$$

where the second integral has to be understood as

$$\int_\Omega Dk(x)\,\mathrm{d}v(x) := \int_\Omega Dk(x)\frac{\mathrm{d}v}{\mathrm{d}|v|}(x)\,\mathrm{d}|v|(x).$$

The properties of G and $\mathcal{K}$ will be investigated in Lemmas 5.14 and 5.15 below.

The range of Π is not (weak*) closed. Thus, it might not be possible to separate elements $[\![u, v]\!] \in C^1(\Omega)^* \setminus \operatorname{Ran}(\Pi)$ from $\operatorname{Ran}(\Pi)$ by a continuous linear functional. However, we can utilize the following characterization.

Lemma 5.13. *An element $[\![u, v]\!] \in C^1(\Omega)^*$ belongs to the range of Π if and only if is bounded by $\|\cdot\|_C$, i.e., there exists $C > 0$ such that we have*

$$|\langle\!\langle \varphi, [\![u, v]\!] \rangle\!\rangle| \leq C\|\varphi\|_C \qquad \forall \varphi \in C^1(\Omega).$$

Proof. It is clear that "$\Rightarrow$" holds. In order to prove "$\Leftarrow$", let $[\![u, v]\!]$ be bounded by $\|\cdot\|_C$. Thus, this linear functional is a bounded functional on the dense subspace $C^1(\Omega)$ of the normed space $C(\Omega)$. Hence, it can be uniquely extended to a functional $\tilde{u} \in C(\Omega)^* = \mathcal{M}(\Omega)$, *i.e.*, we have

$$\langle\!\langle \varphi, [\![u, v]\!] \rangle\!\rangle = \langle \varphi, \tilde{u} \rangle_C = \langle\!\langle \varphi, [\![\tilde{u}, 0]\!] \rangle\!\rangle \qquad \forall \varphi \in C^1(\Omega).$$

This proves $[\![u, v]\!] = \Pi\tilde{u}$. □

Next, we show that G is a nice functional.

Lemma 5.14. *The function G from (5.4) is proper, convex and weak* lower semi-continuous on $C^1(\Omega)^*$. Its preconjugate $H\colon C^1(\Omega) \to [0, \infty]$ is given by*

$$\forall \varphi \in C^1(\Omega): \qquad H(\varphi) = \begin{cases} 0 & \text{if } \varphi \in \mathbb{B}_\alpha, \\ \infty & \text{else,} \end{cases}$$

where $\mathbb{B}_\alpha := \{\phi \in C(\Omega) \mid \|\phi\|_C \leq \alpha\}$.

Proof. We first show $H^* = G$. The claimed properties of G then follow immediately. Note that

$$H^*([\![u,v]\!]) = \sup\{\langle\!\langle \varphi, [\![u,v]\!]\rangle\!\rangle \mid \varphi \in \mathbb{B}_\alpha\}.$$

We argue by a distinction of cases. First, if $v = 0$, we immediately get $H^*([\![u,0]\!]) = \alpha\|u\|_{\mathcal{M}}$ due to the density of $\mathbb{B}_\alpha \cap C^1(\Omega)$ in $\mathbb{B}_\alpha \subset C(\Omega)$. This shows that H^* and G coincide on the range of Π. In case that $[\![u,v]\!] \notin \mathrm{Ran}(\Pi)$, we can utilize Lemma 5.13 and obtain that for every $n \in \mathbb{N}$, there exists $\varphi_n \in C^1(\Omega)$ with

$$|\langle\!\langle \varphi_n, [\![u,v]\!]\rangle\!\rangle| \geq n\|\varphi_n\|_C.$$

Since $\pm\alpha\varphi_n/\|\varphi_n\|_C \in \mathbb{B}_\alpha$, this implies $H^*([\![u,v]\!]) \geq n$. Thus, we conclude $H^*([\![u,v]\!]) = G([\![u,v]\!]) = \infty$ if $[\![u,v]\!] \notin \mathrm{Ran}(\Pi)$. □

Similarly, the operator $\mathcal{K}$ has nice properties since it possesses a preadjoint.

Lemma 5.15. *The operator $\mathcal{K}$ from (5.5) is well defined and sequentially weak*-to-strong continuous from $C^1(\Omega)^*$ to Y. Moreover, there holds $\mathcal{K} = (\mathcal{K}_*)^*$ where $\mathcal{K}_*\colon Y \to C^1(\Omega)$ satisfies $\mathcal{K}_* y = K_* y$.*

Proof. The proof follows by the same argument as Lemma 3.2 noting that $k \in C^{1,\gamma}(\Omega; Y)$. For the sake of brevity, further details are omitted. □

Now we consider the lifted minimization problem

$$\min_{[\![u,v]\!]\in C^1(\Omega)^*} \mathcal{J}([\![u,v]\!]) := L(\mathcal{K}([\![u,v]\!])) + G([\![u,v]\!]). \tag{$\mathcal{P}$}$$

The next proposition establishes its equivalence to Problem (P).

Proposition 5.16. *An equivalence class $[\![u,v]\!]$ is a minimizer of Problem ($\mathcal{P}$) if and only if there holds $[\![u,v]\!] = \Pi\bar{u}$ for a solution $\bar{u}$ of Problem (P).*

Proof. This follows immediately noting that $\mathrm{dom}(\mathcal{J}) = \mathrm{Ran}(\Pi)$. □

5.4. The main result

We recall that we assume Assumption 3.1 throughout the paper. In the following, we assume the existence of a stationary point

$$\bar{u} = \sum_{j=1}^{N} \bar{\lambda}_j \delta_{\bar{x}_j} \quad \text{where} \quad \bar{\lambda}_j \neq 0, \bar{x}_j \in \mathrm{int}(\Omega) \qquad \forall j = 1, \ldots, N \tag{5.6}$$

of (P), *i.e.*, the associated dual variable $\bar{p} = -K_*\nabla L(K\bar{u})$ satisfies $\bar{p} \in C^1(\Omega) \cap \alpha\partial\|\bar{u}\|_{\mathcal{M}}$. Moreover, we define the sets $\mathcal{A}$ as well as $\mathcal{I}_+$ and $\mathcal{I}_-$ by

$$\mathcal{A} := \mathrm{supp}\,\bar{u} = \{\bar{x}_j\}_{j=1}^N, \quad \mathcal{I}_\pm := \{x \in \Omega \mid \bar{p}(x) = \pm\alpha\} \setminus \mathcal{A}. \tag{5.7}$$

Finally, set $s_j = \mathrm{sgn}(\bar{p}(\bar{x}_j))$ for every $\bar{x}_j \in \mathcal{A}$. The following theorem provides a tangible no-gap second-order condition for Problem (P) at $\bar{u}$ as well as its equivalence to the quadratic growth of J w.r.t. $\|\cdot\|_{\mathrm{BL}}$ in the vicinity of $\bar{u}$. We emphasize that we do not assume that the sets $\mathcal{I}_\pm$ are finite, in particular, our setting is more general than the usual requirement (1.4), which is equivalent to $\mathcal{I}_\pm = \emptyset$.

Theorem 5.17. *Consider a measure $\bar{u}$ of the form (5.6) and assume that $\bar{p} = -K_*\nabla L(K\bar{u}) \in C^1(\Omega) \cap \alpha\partial\|\bar{u}\|_{\mathcal{M}}$ is two times continuously differentiable around all $\bar{x}_j \in \mathcal{A}$. Then the following statements are equivalent.*

(B1) *There exists $\theta > 0$ with*

$$s_j \nabla^2 \bar{p}(\bar{x}_j) \leq_L -\theta \operatorname{Id} \qquad \forall \bar{x}_j \in \mathcal{A}. \tag{5.8}$$

Moreover,

$$\nabla^2 L(K\bar{u})\left(\mathcal{K}\left[\!\left[\mu, \sum_{j=1}^{n} V_j \delta_{\bar{x}_j}\right]\!\right]\right)^2 - \sum_{j=1}^{N} \left[\frac{1}{\bar{\lambda}_j} \nabla^2 \bar{p}(\bar{x}_j) V_j^2\right] > 0 \tag{5.9}$$

holds for all $(\mu, V) \in \big(\mathcal{C} \times (\mathbb{R}^d)^N\big) \setminus \{0\}$, where the critical cone $\mathcal{C} \subset \mathcal{M}(\Omega)$ is given by

$$\mathcal{C} := \left\{ \sum_{j=1}^{N} \lambda_j \delta_{\bar{x}_j} + \widehat{\mu}_+ - \widehat{\mu}_- \;\middle|\; \lambda \in \mathbb{R}^N,\ \widehat{\mu}_+ \in \mathcal{M}(\mathcal{I}_+)^+,\ \widehat{\mu}_- \in \mathcal{M}(\mathcal{I}_-)^+ \right\}. \tag{5.10}$$

(B2) *There exist $\gamma > 0$ and $\varepsilon > 0$ such that there holds*

$$J(u) - J(\bar{u}) \geq \gamma \|u - \bar{u}\|_{\mathrm{BL}}^2 \qquad \forall u \in \mathcal{M}(\Omega), \|u - \bar{u}\|_{\mathrm{BL}} \leq \varepsilon. \tag{5.11}$$

Remark 5.18. At this point, we emphasize that Theorem 5.17 as well as the conclusions made from it in the following, also remain valid, mutatis mutandis, if we restrict minimization in equation (P) to $\mathcal{M}(\Omega)^+$. In this case, we have $s_j = 1$ for all $\bar{x}_j \in \mathcal{A}$ and set $\mathcal{I}_- = \emptyset$.

The proof of this result is quite involved. The implication "(B2)⇒(B1)" is given in Lemma 5.23. The reverse implication is given in Sections 6 and 7, see also the comment after Lemma 5.23.

In case that L is convex, the condition (5.9) simplifies.

Corollary 5.19. *In addition to the assumptions of Theorem 5.17, we assume that L is convex. Consider the following statements.*

(C1) *Condition (5.8) holds for some $\theta > 0$ and*

$$(K\mu, \nabla^2 L(K\bar{u}) K\mu)_Y > 0 \qquad \forall \mu \in \mathcal{C} \setminus \{0\} \tag{5.12}$$

is satisfied, where the critical cone $\mathcal{C}$ is defined in (5.10).

(C2) *There exist $\gamma > 0$ and $\varepsilon > 0$ such that the quadratic growth condition (5.11) holds.*

(C3) *Condition (5.8) holds for some $\theta > 0$ and the matrix*

$$\big((k(\bar{x}_i), \nabla^2 L(K\bar{u}) k(\bar{x}_j))_Y\big)_{i,j=1}^{N}$$

is positive definite.

(C4) *Condition (5.8) holds for some $\theta > 0$ and $\dim(\operatorname{span}\{k(\bar{x}_1), \ldots, k(\bar{x}_N)\}) = N$.*

Then, (C1) and (C2) are equivalent. In case that sets $\mathcal{I}_+$ and $\mathcal{I}_-$ are empty, i.e., if $|\bar{p}(x)| < \alpha$ holds for all $x \in \Omega \setminus \mathcal{A}$, we get the equivalence of (C1)–(C3). If, additionally, L is strongly convex, we get the equivalence of (C1)–(C4).

The equivalence of (C1), (C3) and (C4) (under the respective assumptions) are easy to check. It thus remains to check the equivalence of (C1) and (C2), and this will be verified by showing that (C1) is equivalent to condition (B1) from Theorem 5.17, see Lemma 6.7 below. Note that (C4) corresponds to the non-degeneracy of $\bar{p}$, see (1.5).

Remark 5.20. In order to prove Theorem 5.17, we use the lifted setting and work in the space $C^1(\Omega)^*$. This is also apparent from the condition (5.9). In the convex setting of Corollary 5.19, however, this is no longer the case and its characterizations can be stated purely in the measure space $\mathcal{M}(\Omega)$, see (5.12).

In order to prove Theorem 5.17, we are going to apply the theory from Section 4. To this end, we first verify the Taylor formula (4.2).

Lemma 5.21. *The functional $F(\llbracket u, v\rrbracket) \coloneqq L(\mathcal{K}(\llbracket u, v\rrbracket))$ satisfies (4.2) with the choices*

$$\begin{aligned} F'(\llbracket \bar u, 0\rrbracket)\llbracket u, v\rrbracket &\coloneqq (\nabla L(K\bar u), \mathcal{K}\llbracket u, v\rrbracket)_Y = \langle\!\langle -\bar p, \llbracket u, v\rrbracket\rangle\!\rangle, \\ F''(\llbracket \bar u, 0\rrbracket)\llbracket u, v\rrbracket^2 &\coloneqq \nabla^2 L(K\bar u)(\mathcal{K}\llbracket u, v\rrbracket)^2 \end{aligned}$$

for arbitrary $\bar u \in \mathcal{M}(\Omega)$ and associated $\bar p = -K_*\nabla L(K\bar u) \in C^1(\Omega)$.

Proof. This is clear as L is assumed to be twice continuously Fréchet differentiable and $\mathcal{K}$ is linear and bounded. □

We first point out the following consequence of the abstract result in Theorem 4.3.

Corollary 5.22. *Let a measure $\bar u \in \mathcal{M}(\Omega)$ be given and set $\bar p \coloneqq -K_*\nabla L(K\bar u)$. Then the following statements are equivalent.*

1. *There holds the second-order condition*

$$(\mathcal{K}h, \nabla^2 L(K\bar u)\mathcal{K}h)_Y + G''(\llbracket \bar u, 0\rrbracket, \bar p; h) > 0 \qquad \forall h \in C^1(\Omega)^* \setminus \{0\} \tag{SSC-$\mathcal{M}$}$$

   and the non-degeneracy condition

$$\begin{aligned} &\text{for all } (t_k) \subset (0,\infty),\ (h_k) \subset C^1(\Omega)^* \text{ with } t_k \searrow 0,\ h_k \overset{*}{\rightharpoonup} 0 \text{ and } \|h_k\|_{(C^1)^*} = 1, \text{ we have} \\ &\liminf_{k\to\infty}\left(\frac{1}{t_k^2}(G(\llbracket \bar u, 0\rrbracket + t_k h_k) - G(\llbracket \bar u, 0\rrbracket)) - \frac{1}{t_k}\langle\!\langle \bar p, h_k\rangle\!\rangle\right) > 0. \end{aligned} \tag{NDC-$\mathcal{M}$}$$

2. *There exist $\gamma > 0$ and $\varepsilon > 0$ such that there holds*

$$J(u) - J(\bar u) \ge \gamma\|u - \bar u\|_{\mathrm{BL}}^2 \qquad \forall u \in \mathcal{M}(\Omega), \|u - \bar u\|_{\mathrm{BL}} \le \varepsilon.$$

Proof. Due to Lemma 5.21, Assumption 4.1 is satisfied. By definition, there holds

$$F''(\llbracket \bar u, 0\rrbracket)h^2 = (\mathcal{K}h, \nabla^2 L(K\bar u)\mathcal{K}h)_Y = (\mathcal{K}h, \nabla^2 L(\mathcal{K}(\llbracket(\bar u, 0)\rrbracket))\mathcal{K}h)_Y.$$

Moreover, due to Lemma 5.15, the mapping

$$h \mapsto (\mathcal{K}h, \nabla^2 L(K\bar u)\mathcal{K}h)_Y$$

is sequentially weak*-continuous in $C^1(\Omega)^*$. This shows that the second assumption of Theorem 4.3 is satisfied and this also allows to drop the last term

$$\frac{1}{2}(\mathcal{K}h_k, \nabla^2 L(K\bar u)\mathcal{K}h_k)_Y$$

which would appear in the NDC. Consequently, see Theorem 4.3, (SSC-$\mathcal{M}$) and (NDC-$\mathcal{M}$) are equivalent to

$$\mathcal{J}([\![u,v]\!]) - \mathcal{J}([\![\bar{u},0]\!]) \geq \tilde{\gamma} \|[\![u-\bar{u},v]\!]\|^2_{(C^1)^*} \qquad \forall [\![u,v]\!] \in C^1(\Omega)^*, \|[\![u-\bar{u},v]\!]\|_{(C^1)^*} \leq \varepsilon$$

for some $\tilde{\gamma} > 0$ and $\varepsilon > 0$. The claimed statement now follows from $\operatorname{dom}(\mathcal{J}) = \operatorname{Ran}(\Pi)$ as well as Lemma 5.8. $\square$

Thus, it suffices to show that (SSC-$\mathcal{M}$) together with (NDC-$\mathcal{M}$) are equivalent to (B1). Since this is rather technical, it will be split into several parts. In this section, we show that (SSC-$\mathcal{M}$) and (NDC-$\mathcal{M}$) imply (B1).

Lemma 5.23. *Let $\bar{u}$ and $\bar{p}$ satisfy the assumptions of Theorem 5.17 and assume that (SSC-$\mathcal{M}$) and (NDC-$\mathcal{M}$) hold. Then (B1) is satisfied.*

Proof. We start by showing that (NDC-$\mathcal{M}$) implies the definiteness of the Hessian at all $\bar{x}_j \in \mathcal{A}$. For this purpose, we construct a suitable sequence $(h_k) \subset C^1(\Omega)^*$ with $\|h_k\|_{(C^1)^*} = 1$, $k \in \mathbb{N}$, as well as $h_k \rightharpoonup^* 0$ in $C^1(\Omega)^*$ and insert it into (NDC-$\mathcal{M}$). More in detail, let $(t_k) \subset (0,\infty)$ denote an arbitrary sequence with $t_k \searrow 0$ and fix $\bar{x}_j \in \mathcal{A}$ as well as $v \in \mathbb{R}^d$, $|v| = 1$. Define

$$\tilde{h}_k = \Pi \mu_k \quad \text{where} \quad \mu_k = \bar{\lambda}_j \frac{\left(\delta_{\bar{x}_j - (t_k/\bar{\lambda}_j)v} - 2\delta_{\bar{x}_j} + \delta_{\bar{x}_j + (t_k/\bar{\lambda}_j)v}\right)}{2t_k}$$

where $\bar{\lambda}_j \neq 0$ denotes coefficient associated to $\delta_{\bar{x}_j}$. For k large enough, we have $(t_k/\bar{\lambda}_j)|v| \leq 1$. Thus, Example 5.11 yields $\|\mu_k\|_{\mathrm{BL}} = 1$. Now, Lemma 5.8 implies $1 \leq \|\tilde{h}_k\|_{(C^1)^*} \leq L$ for some $L > 0$, since $\tilde{h}_k = \Pi \mu_k$. Consequently, $\sigma_k := \|\tilde{h}_k\|^{-1}_{(C^1)^*}$ satisfies $\sigma_k \in [L^{-1}, 1]$ and we set

$$h_k := \sigma_k \tilde{h}_k.$$

For arbitrary $\varphi \in C^1(\Omega)$, we can use a Taylor expansion at $\bar{x}_j$ to obtain

$$\lim_{k\to\infty} \langle\!\langle \varphi, \tilde{h}_k \rangle\!\rangle = \lim_{k\to\infty} \bar{\lambda}_j \frac{\left(\varphi(\bar{x}_j - (t_k/\bar{\lambda}_j)v) - 2\varphi(\bar{x}_j) + \varphi(\bar{x}_j + (t_k/\bar{\lambda}_j)v)\right)}{2t_k} = 0.$$

Consequently, $\tilde{h}_k \rightharpoonup^* 0$ and $h_k \rightharpoonup^* 0$ in $C^1(\Omega)^*$.

Finally, observe that there holds

$$G([\![\bar{u},0]\!] + t_k h_k) - G([\![\bar{u},0]\!]) = 0$$

for all $k \in \mathbb{N}$ large enough. Since $\bar{p}$ is assumed to be twice differentiable at $\bar{x}_j$, we can perform a second-order Taylor expansion to get

$$\begin{aligned}\lim_{k\to\infty} -\frac{1}{t_k} \langle\!\langle \bar{p}, \tilde{h}_k \rangle\!\rangle &= \lim_{k\to\infty} -\bar{\lambda}_j \frac{\left(\bar{p}(\bar{x}_j - (t_k/\bar{\lambda}_j)v) - 2\bar{p}(\bar{x}_j) + \bar{p}(\bar{x}_j + (t_k/\bar{\lambda}_j)v)\right)}{2t_k^2} \\ &= -\frac{\nabla^2 \bar{p}(\bar{x}_j) v^2}{2\bar{\lambda}_j} = -\operatorname{sgn}(\bar{p}(\bar{x}_j)) \frac{\nabla^2 \bar{p}(\bar{x}_j) v^2}{2|\bar{\lambda}_j|}.\end{aligned}$$

Summarizing the previous observations, (NDC-$\mathcal{M}$) implies

$$-\operatorname{sgn}(\bar{p}(\bar{x}_j)) \frac{\nabla^2 \bar{p}(\bar{x}_j) v^2}{2|\bar{\lambda}_j|} = \liminf_{k\to\infty} \sigma_k^{-1} \left(\frac{1}{t_k^2} \left(G([\![\bar{u},0]\!] + t_k h_k) - G([\![\bar{u},0]\!]) \right) - \frac{1}{t_k} \langle\!\langle \bar{p}, h_k \rangle\!\rangle \right) > 0.$$

Since $v \in \mathbb{R}^d$ with $|v| = 1$ was arbitrary, this shows that the Hessian is definite at $\bar{x}_j \in \mathcal{A}$.

Next, we prove that (SSC-$\mathcal{M}$) implies the definiteness of $\nabla^2 L(K\bar{u})$ in the sense of (5.12). We do not argue by contradiction. Let a measure

$$\mu = \sum_{j\in\mathcal{A}} \lambda_j \delta_{x_j} + \widehat{\mu}_+ - \widehat{\mu}_- \in \mathcal{C} \setminus \{0\}$$

be given and denote by $h = \Pi\mu$ the corresponding equivalence class. Then there holds $h \neq 0$ and

$$0 \leq G''(\llbracket \bar{u}, 0 \rrbracket, \bar{p}; h) \leq \liminf_{k\to\infty} \frac{G(\llbracket \bar{u}, 0 \rrbracket + t_k h) - G(\llbracket \bar{u}, 0 \rrbracket) - t_k \langle\!\langle \bar{p}, h \rangle\!\rangle}{t_k^2/2}$$

for an arbitrary but fixed sequence $(t_k) \subset (0, \infty)$ with $t_k \searrow 0$. Now, for k large enough, we conclude

$$\begin{aligned} &G(\llbracket \bar{u}, 0 \rrbracket + t_k h) - G(\llbracket \bar{u}, 0 \rrbracket) - t_k \langle\!\langle \bar{p}, h \rangle\!\rangle \\ &\qquad = \sum_{j\in\mathcal{A}} \left[\alpha|\bar{\lambda}_j + t_k \lambda_j| - \alpha|\bar{\lambda}_j| - t_k \lambda_j \bar{p}(\bar{x}_j)\right] + \alpha t_k \|\widehat{\mu}_+\|_{\mathcal{M}} + \alpha t_k \|\widehat{\mu}_-\|_{\mathcal{M}} - t_k \langle\!\langle \bar{p}, \widehat{\mu}_+ - \widehat{\mu}_- \rangle\!\rangle \\ &\qquad = 0 \end{aligned}$$

noting that $\bar{p}(\bar{x}_j) = s_j \alpha$, $\bar{\lambda}_j \neq 0$ and $|\bar{\lambda}_j| = s_j \bar{\lambda}_j$ for all $\bar{x}_j \in \mathcal{A}$ as well as

$$\langle \bar{p}, \widehat{\mu}_+ \rangle = \alpha \|\widehat{\mu}_+\|_{\mathcal{M}}, \quad \langle \bar{p}, \widehat{\mu}_- \rangle = -\alpha \|\widehat{\mu}_-\|_{\mathcal{M}}$$

by construction of $\mathcal{I}_\pm$. Consequently, (SSC-$\mathcal{M}$) implies

$$(\mathcal{K}h, \nabla^2 L(K\bar{u})\mathcal{K}h)_Y = (\mathcal{K}h, \nabla^2 L(K\bar{u})\mathcal{K}h)_Y + G''(\llbracket \bar{u}, 0 \rrbracket, \bar{p}; h) > 0$$

for this particular choice of h, finishing the proof. ☐

It remains to prove (B1) $\Rightarrow$ (SSC-$\mathcal{M}$) + (NDC-$\mathcal{M}$). For the sake of readability this will be done in two steps. In Section 6, we compute the weak* second subderivative of G and this enables us to prove that (B1) implies (SSC-$\mathcal{M}$), see Lemma 6.7. As a byproduct, we also show the equivalence of (B1) and (C1), see Lemma 6.7. The implication of (NDC-$\mathcal{M}$) by (B1) is addressed in Section 7.

We close this section with a final observation. Note that Theorem 5.17 assumes the structure (5.6) of $\bar{u}$. The next lemma shows that (5.6) is "almost" necessary for (NDC-$\mathcal{M}$) and, consequently, for the quadratic growth condition (B2). In fact, whenever $\bar{u}$ is not a finite sum of Dirac measures, these conditions (NDC-$\mathcal{M}$) and (B2) cannot be true.

Lemma 5.24. *Let $\bar{u} \in \mathcal{M}(\Omega)$ and $\bar{p} \in C(\Omega) \cap \alpha\partial\|\bar{u}\|_{\mathcal{M}}$ be given. If (NDC-$\mathcal{M}$) holds, then all points in $\operatorname{supp}(\bar{u})$ are isolated. In particular,*

$$\bar{u} = \sum_{j=1}^{N} \bar{\lambda}_j \delta_{\bar{x}_j}$$

for some $N \in \mathbb{N}$, $\bar{\lambda}_j \in \mathbb{R}$ and $\bar{x}_j \in \Omega$.

Proof. We prove the contraposition. We suppose that $x \in \operatorname{supp}(\bar{u})$ is not isolated. W.l.o.g., we assume $x \in \operatorname{supp}(\bar{u}_+)$. The case $x \in \operatorname{supp}(\bar{u}_-)$ can be handled analogously.

From $\bar{p} \in \alpha\partial\|\bar{u}\|_{\mathcal{M}}$, we get $\bar{p}(x) = \alpha$. Due to the continuity of $\bar{p}$, $\operatorname{supp}(\bar{u}_-)$ has a positive distance $\delta > 0$ from x.

Since x is not an isolated point of $\operatorname{supp}(\bar u)$, there exists a sequence $(x_k) \subset \operatorname{supp}(\bar u) \setminus \{x\}$ with $x_k \to x$. We set $s_k := |x_k - x|$ and, w.l.o.g., we assume $(x_k - x)/s_k \to v \in \mathbb{R}^d$ and $s_k \le s_{k-1}/2$. Next, we choose

$$\begin{aligned} r_k &:= s_k/k, & \lambda_k &:= s_k/s_{k-1} \in [0, 1/2], \\ \mu_k &:= \bar u_{|B_{r_k}(x_k)}, & m_k &:= \|\mu_k\|_{\mathcal{M}} > 0, \\ \hat h_k &:= ((1-\lambda_k)\delta_x + \lambda_k \delta_{x_{k-1}})m_k - \mu_k, & t_k &:= \|\hat h_k\|_{\mathrm{BL}}, & h_k &:= \hat h_k / t_k. \end{aligned}$$

We want to give estimates for t_k. For deriving a lower bound, we use the definition (5.2) with $\varphi(\cdot) := |\cdot - x_k|$. Note that $\|\varphi\|_{\mathrm{Lip}} = 1$. Together with $|x_{k-1} - x_k| \ge s_{k-1} - s_k \ge s_k$ we arrive at

$$t_k = \|\hat h_k\|_{\mathrm{BL}} \ge \langle \varphi, \hat h_k \rangle \ge ((1-\lambda_k)|x - x_k| + \lambda_k |x_{k-1} - x_k| - r_k) m_k \ge \left(1 - \frac{1}{k}\right) s_k m_k \ge \frac{1}{2} s_k m_k$$

for all $k \ge 2$. For the upper bound, we use (5.3) with $\mu = (\hat h_k)_+$, $\nu = (\hat h_k)_-$. Together with Definition 5.9 and the coupling $\gamma = (\mu \otimes \nu)/m_k$, we arrive at

$$t_k \le W_1(\mu, \nu) \le \int_{\Omega \times \Omega} |y - z| \,\mathrm{d}\gamma(y, z) \le ((1-\lambda_k)(|x - x_k| + r_k) + \lambda_k(|x_{k-1} - x_k| + r_k)) m_k.$$

In particular, $t_k \searrow 0$.

By construction of $\hat h_k$, we get

$$G([\![\bar u + t_k h_k, 0]\!]) = G([\![\bar u + \hat h_k, 0]\!]) = G([\![\bar u, 0]\!])$$

and from $x, x_{k-1} \in \operatorname{supp}(\bar u_+)$, we get $\bar p(x) = \bar p(x_{k-1}) = \alpha$ and, consequently,

$$\langle\!\langle \bar p, [\![\hat h_k, 0]\!] \rangle\!\rangle = 0.$$

In order to check that (NDC-$\mathcal{M}$) does not hold, it remains to check that $[\![h_k, 0]\!] \xrightarrow{*} 0$ in $C^1(\Omega)^*$. Let $\varphi \in C^1(\Omega)$ be arbitrary. We have

$$\langle\!\langle \varphi, [\![h_k, 0]\!] \rangle\!\rangle = \frac{((1-\lambda_k)\varphi(x) + \lambda_k \varphi(x_{k-1}))m_k - \int_{B_{r_k}(x_k)} \varphi \,\mathrm{d}\mu_k}{t_k}.$$

From the Lipschitz continuity of φ, see Lemma 5.4, we get the estimate

$$\left| \varphi(x_k) m_k - \int_{B_{r_k}(x_k)} \varphi \,\mathrm{d}\mu_k \right| \le r_k m_k \|\varphi\|_{\mathrm{Lip}}.$$

Consequently,

$$|\langle\!\langle \varphi, [\![h_k, 0]\!] \rangle\!\rangle| \le \left| \frac{((1-\lambda_k)\varphi(x) + \lambda_k \varphi(x_{k-1}))m_k - \varphi(x_k) m_k}{t_k} \right| + \frac{r_k m_k}{t_k} \|\varphi\|_{\mathrm{Lip}}.$$

Due to the choice of r_k and the lower bound on t_k, the last term goes to zero and, for brevity, we will replace it by $o(1)$. Together with the lower bound for t_k, we get

$$\frac{1}{2} |\langle\!\langle \varphi, [\![h_k, 0]\!] \rangle\!\rangle| \le \left| \frac{(1-\lambda_k)\varphi(x) + \lambda_k \varphi(x_{k-1}) - \varphi(x_k)}{s_k} \right| + o(1).$$

Owing to Lemma 5.5, we get

$$\big|\varphi(x_k) - \varphi(x) - \nabla\varphi(x)^\top (x_k - x)\big| \le \eta_k |x_k - x| = \eta_k s_k$$

for some sequence $(\eta_k) \subset [0, \infty)$ with $\eta_k \searrow 0$. Consequently,

$$\frac{1}{2}|\langle\!\langle \varphi, [\![h_k, 0]\!]\rangle\!\rangle| \le \left| \frac{\nabla\varphi(x)^\top (\lambda_k(x_{k-1} - x) - (x_k - x))}{s_k} \right| + \frac{\lambda_k \eta_{k-1} s_{k-1} + \eta_k s_k}{s_k} + o(1).$$

Now, we insert $\lambda_k = s_k / s_{k-1}$ and get

$$\frac{1}{2}|\langle\!\langle \varphi, [\![h_k, 0]\!]\rangle\!\rangle| \le |\nabla\varphi(x)| \left| \frac{x_{k-1} - x}{s_{k-1}} - \frac{x_k - x}{s_k} \right| + \eta_{k-1} + \eta_k + o(1) \to 0.$$

This shows $[\![h_k, 0]\!] \overset{*}{\rightharpoonup} 0$. Consequently, (NDC-$\mathcal{M}$) is violated. $\square$

6. Structural assumptions imply SSC

In this section, we prove that the structural assumption (B1) implies the second-order condition (SSC-$\mathcal{M}$). To this end, we explicitly characterize the weak* second subderivative of G at $[\![\bar u, 0]\!]$ for $\bar p$,

$$G''([\![\bar u, 0]\!], \bar p; h) = \inf\left\{ \liminf_{k\to\infty} \frac{G([\![\bar u, 0]\!] + t_k h_k) - G([\![\bar u, 0]\!]) - t_k \langle\!\langle \bar p, h\rangle\!\rangle}{t_k^2/2} \;\middle|\; t_k \searrow 0,\ h_k \overset{*}{\rightharpoonup} h \right\},$$

in all directions $h \in C^1(\Omega)^*$. For this purpose, and to avoid working with the weak* topology on $C^1(\Omega)^*$, we argue similarly to [25] and start by showing that the preconjugate H, see Lemma 5.14, is twice strong-strong epi-subdifferentiable. A candidate for $G''([\![\bar u, 0]\!], \bar p; h)$ is then found by applying Lemma 4.5.

6.1. Twice strong-strong epi-differentiability of *H*

We recall that the functional H on $C^1(\Omega)$ is the indicator function of $\{\phi \in C^1(\Omega) \mid \|\phi\|_C \le \alpha\}$, see Lemma 5.14. We start by proving an auxiliary result concerning the tangent cone of this set.

Lemma 6.1. *We set $\mathcal{S} \coloneqq \{\phi \in C^1(\Omega) \mid \|\phi\|_C \le \alpha\}$. For every $p \in \mathcal{S}$, the tangent cone of $\mathcal{S}$ at p (in the sense of convex analysis) is given by*

$$\mathcal{T}_{\mathcal{S}}(p) = \big\{ z \in C^1(\Omega) \;\big|\; z \le 0 \text{ on } \{p = \alpha\},\ z \ge 0 \text{ on } \{p = -\alpha\} \big\}.$$

Proof. The inclusion "⊂" is clear and we just have to check "⊃". Thus, let z from the right-hand side be given. For $t > 0$ we define the optimal values

$$\bar d_\pm \coloneqq \max\{\pm z(x) \mid x \in \Omega,\ p(x) = \pm\alpha\}, \qquad \bar d_{\pm,t} \coloneqq \max\big\{ t^{-1}(\pm p(x) - \alpha) \pm z(x) \;\big|\; x \in \Omega \big\}.$$

By assumption, there holds $\bar d_\pm \le 0$. We readily verify $\bar d_{\pm,t} \to \bar d$ and thus $d_{\pm,t} \coloneqq \max\{0, \bar d_{\pm,t}\} \to 0$ as $t \to 0$. By construction, we further have

$$t^{-1}(\pm p(x) - \alpha) \pm z(x) \le \bar d_{\pm,t} \le d_{\pm,t} \qquad \forall x \in \Omega.$$

Rearranging, this implies

$$\pm(p(x) + t z(x) \mp t d_{\pm,t}) \le \alpha \qquad \forall x \in \Omega. \tag{6.1}$$

Due to the continuity of p, the sets $\{p \le -\alpha/2\} = \{-p \ge \alpha/2\}$ and $\{p \ge \alpha/2\}$ have a positive distance. Thus, we can find $\varphi_\pm \in C^1(\Omega)$ with disjoint support, $0 \le \varphi_\pm \le 1$ and $\varphi_\pm = 1$ on $\{\pm p \ge \alpha/2\}$.

Now, we consider

$$\xi_t \coloneqq -d_{+,t}\varphi_+ + d_{-,t}\varphi_-. \tag{6.2}$$

The convergence of ξ_t towards 0 in $C^1(\Omega)$ follows immediately and it remains to check $p + t(z + \xi_t) \in \mathcal{S}$ for $t > 0$ small enough. For $x \in \{p \ge \alpha/2\}$, we have

$$(p + t(z + \xi_t))(x) = p(x) + tz(x) - td_{+,t} \le \alpha$$

due to (6.1). For $x \notin \{p \ge \alpha/2\}$, we get

$$(p + t(z + \xi_t))(x) \le \frac{\alpha}{2} + t\|z + \xi_t\|_C \le \alpha$$

for $t > 0$ small enough. Similarly, we can handle the lower bound and this finishes the proof. □

It is interesting to note that tangent cone in $C^1(\Omega)$ does not see the derivative of the direction z on the sets $\{p = \pm\alpha\}$.

In order to prepare the computation of the strong second subderivative of H, we give an auxiliary lemma.

Lemma 6.2. *We assume that $\bar{u} \in \mathcal{M}(\Omega)$ is of the form (5.6). Further, $\bar{p} \in C^1(\Omega) \cap \alpha\partial\|\bar{u}\|_{\mathcal{M}}$ is given such that (5.8) holds, where $\mathcal{A}$ is defined in (5.7).*

Then, there exists $r_0 > 0$ such that the following holds.

(1) The balls $B_{r_0}(\bar{x}_j)$, $j = 1, \ldots, N$, are pairwise disjoint and subsets of Ω.
(2) For all $j \in \{1, \ldots, N\}$ we have

$$\frac{\alpha}{2} \le s_j\bar{p}(x) \le \alpha - \frac{\theta}{4}|x - \bar{x}_j|^2 \qquad \forall x \in B_{r_0}(\bar{x}_j). \tag{6.3}$$

Proof. It is clear that (1) holds for $r_0 > 0$ small enough. In order to verify (2) we can combine a Taylor expansion

$$\bar{p}(x) = \bar{p}(\bar{x}_j) + \nabla\bar{p}(\bar{x}_j)^\top(x - \bar{x}_j) + \frac{1}{2}\nabla^2\bar{p}(\bar{x}_j)(x - \bar{x}_j)^2 + o(|x - \bar{x}_j|^2)$$

with $\bar{p}(\bar{x}_j) = s_j\alpha$, $\nabla\bar{p}(\bar{x}_j) = 0$ and (5.8). □

Now we are in position to give a precise characterization of $H''(\bar{p}, [\![\bar{u}, 0]\!]; z)$.

Proposition 6.3. *Under the assumptions of Lemma 6.2, the functional H is strongly-strongly twice epi-differentiable at $\bar{p}$ for $[\![\bar{u}, 0]\!]$ with*

$$H''(\bar{p}, [\![\bar{u}, 0]\!], z) = -\sum_{j=1}\bar{\lambda}_j\nabla^2\bar{p}(\bar{x}_j)^{-1}\nabla z(\bar{x}_j)^2 + I_{\mathcal{Z}}(z) \qquad \forall z \in C^1(\Omega) \tag{6.4}$$

where

$$\mathcal{Z} = \left\{z \in C^1(\Omega) \;\middle|\; z \le 0 \text{ on } \mathcal{I}_+,\ z \ge 0 \text{ on } \mathcal{I}_-,\ z(\bar{x}_j) = 0\ \forall j = 1, \ldots, N\right\}.$$

Proof. Note that $\bar{p} \in \mathbb{B}_\alpha$ and thus $H(\bar{p}) = 0$. For abbreviation, given $t_k > 0$ and $z_k \in C^1(\Omega)$, set

$$D(t_k, z_k) := \frac{H(\bar{p} + t_k z_k) - t_k \langle\!\langle z_k, [\![\bar{u}, 0]\!] \rangle\!\rangle}{t_k^2/2} = \frac{H(\bar{p} + t_k z_k) - t_k \sum_{j=1}^N \bar{\lambda}_j z_k(\bar{x}_j)}{t_k^2/2}.$$

Thus,

$$H''(\bar{p}, [\![\bar{u}, 0]\!]; z) = \inf\left\{ \liminf_{k\to\infty} D(t_k, z_k) \;\middle|\; t_k \searrow 0,\ z_k \to z \right\} \qquad \forall z \in C^1(\Omega).$$

The proof is divided into several steps.

Step 1: We start by showing that $H''(\bar{p}, [\![\bar{u}, 0]\!]; z) = \infty$ if $z \notin \mathcal{Z}$. For this purpose, we first note that

$$\mathcal{Z} = \mathcal{T}_{\mathcal{S}}(\bar{p}) \cap \left\{ z \in C^1(\Omega) \;\middle|\; z(\bar{x}_j) = 0 \ \forall j = 1, \ldots, N \right\},$$

cf. Lemma 6.1. In case $z \notin \mathcal{T}_{\mathcal{S}}(\bar{p})$, for all sequences $(t_k) \subset (0, \infty)$ and $(z_k) \subset C^1(\Omega)$ with $t_k \searrow 0$ and $z_k \to z$ in $C^1(\Omega)$, we have $\bar{p} + t_k z_k \notin \mathcal{S}$ for all k large enough. Consequently, $H(\bar{p} + t_k z_k) = \infty$ for all k large enough and, therefore, $H''(p, [\![\bar{u}, 0]\!]; z) = \infty$ for all $z \notin \mathcal{T}_{\mathcal{S}}(\bar{p})$.

Now, for $z \in \mathcal{T}_{\mathcal{S}}(\bar{p}) \setminus \mathcal{Z}$ there exists an index $\bar{\jmath} \in \{1, \ldots, N\}$ with $z(\bar{x}_{\bar{\jmath}}) \neq 0$. This implies $\operatorname{sgn}(\bar{p}(\bar{x}_{\bar{\jmath}}))z(\bar{x}_{\bar{\jmath}}) < 0$ and, thus, $\bar{\lambda}_{\bar{\jmath}} z(\bar{x}_{\bar{\jmath}}) < 0$. Moreover, $\bar{\lambda}_j z(\bar{x}_j) \leq 0$ for all $j \in \{1, \ldots, N\}$. For all sequences $(t_k) \subset (0, \infty)$ and $(z_k) \subset C^1(\Omega)$ with $t_k \searrow 0$ and $z_k \to z$ in $C^1(\Omega)$, we have

$$\liminf_{k\to\infty} D(t_k, z_k) \geq \liminf_{k\to\infty} \frac{-\sum_{j=1}^N \bar{\lambda}_j z_k(\bar{x}_j)}{t_k^2/2} \geq \liminf_{k\to\infty} \frac{-\bar{\lambda}_{\bar{\jmath}} z_k(\bar{x}_{\bar{\jmath}})}{t_k^2/2} = \infty.$$

Thus, $H''(\bar{p}, [\![\bar{u}, 0]\!]; z) = \infty$.

Summarizing the previous observations, we conclude $H''(\bar{p}, [\![\bar{u}, 0]\!]; z) = \infty$ if $z \notin \mathcal{Z}$.

Step 2: Next, we will show that $H''(\bar{p}, [\![\bar{u}, 0]\!]; z)$ can be bounded from below by the right-hand side of (6.4) for all $z \in \mathcal{Z}$. Let sequences $(t_k) \subset (0, \infty)$ and $(z_k) \subset C^1(\Omega)$ with $t_k \searrow 0$ and $z_k \to z$ in $C^1(\Omega)$ be given. We have to show that $\liminf_{k\to\infty} D(t_k, z_k)$ can be bounded from below by the right-hand side of (6.4). It is clear that we only have to consider sequences with

$$\bar{p} + t_k z_k \in \mathbb{B}_\alpha \qquad \forall k \in \mathbb{N}$$

in the following. We have

$$\liminf_{k\to\infty} D(t_k, z_k) = \liminf_{k\to\infty} \frac{-\sum_{j=1}^N \bar{\lambda}_j t_k z(\bar{x}_j)}{t_k^2/2} \geq 2 \sum_{j=1}^N \liminf_{k\to\infty} \frac{-\bar{\lambda}_j z(\bar{x}_j)}{t_k}. \tag{6.5}$$

For all $j \in \{1, \ldots, N\}$ we define the perturbation

$$\xi_{j,k} = -t_k \nabla^2 \bar{p}(\bar{x}_j)^{-1} \nabla z_k(\bar{x}_j).$$

Note that $|\xi_{j,k}| = O(t_k)$ as $k \to \infty$. Since $\bar{x}_j \in \operatorname{int}(\Omega)$ there holds $\bar{x}_j + \xi_{j,k} \in \Omega$ for k sufficiently large. Using the feasibility $\bar{p} + t_k z_k \in \mathbb{B}_\alpha$, we get

$$\alpha \geq s_j(\bar{p}(\bar{x}_j + \xi_{j,k}) + t_k z_k(\bar{x} + \xi_{j,k})) \tag{6.6}$$

for all sufficiently large k. By Taylor expansions of $\bar p$ and z_k, respectively, at $\bar x_j$, we get

$$\begin{aligned}&\bar p(\bar x_j+\xi_{j,k})+t_k z_k(\bar x_j+\xi_{j,k})\\&\quad=\bar p(\bar x_j)+t_k z_k(\bar x_j)+\nabla\bar p(\bar x_j)^\top\xi_{j,k}+t_k\nabla z_k(\bar x_j)^\top\xi_{j,k}+\frac12\nabla^2\bar p(\bar x_j)\xi_{j,k}^2+o(t_k^2)\\&\quad=s_j\alpha+t_k z_k(\bar x_j)-\frac{t_k^2}{2}\nabla^2\bar p(\bar x_j)^{-1}\nabla z_k(\bar x_j)^2+o(t_k^2).\end{aligned}$$

Plugging this into (6.6), we arrive at

$$\frac{-\bar\lambda_j z_k(\bar x_j)}{t_k}\ge-\frac12\bar\lambda_j\nabla^2\bar p(\bar x_j)^{-1}\nabla z_k(\bar x_j)^2+o(1).$$

By combining this estimate with (6.5) and using $z_k\to z$ in $C^1(\Omega)$ yields the desired result.

Step 3: Finally, given $z\in\mathcal Z$ as well as an arbitrary sequence $(t_k)\subset(0,\infty)$ with $t_k\searrow0$, we construct a sequence $(z_k)\subset C^1(\Omega)$ with $z_k\to z$ and which achieves the right-hand side of (6.4).

By using the ξ_k from (6.2) (with $t=t_k$), we get $\bar p+t_k(z+\xi_k)\in\mathbb B_\alpha$. However, we have to modify this function in the neighborhoods of $\bar x_j$ in order to make the value of $\bar\lambda_j z_k(\bar x_j)$, which appears in $D(t_k,z_k)$, as large as possible.

Let $r_0>0$ from Lemma 6.2 be given. We define $r_k:=\sqrt{4t_k\|z\|_C/\theta}$. In the sequel, k is large enough, such that $r_k\le r_0$. Now, we consider an arbitrary $j\in\{1,\dots,N\}$ with $\bar p(\bar x_j)=\alpha$, *i.e.*, $s_j=1$. Combining the definition of r_k with (6.3), we get

$$(\bar p+t_k z)(x)\le\alpha-\frac\theta4 r_k^2+t_k\|z\|_C=\alpha\qquad\forall x\in B_{r_0}(\bar x_j)\setminus B_{r_k}(\bar x_j).\tag{6.7}$$

For all $x\in B_{r_k}(\bar x_j)$ we have the Taylor estimates

$$\begin{aligned}\bar p(x)&\le\bar p(\bar x_j)+\nabla\bar p(\bar x_j)^\top(x-\bar x_j)+\frac12\nabla^2\bar p(\bar x_j)(x-\bar x_j)^2+\frac{\varepsilon_k}{2}|x-\bar x_j|^2,\\ z(x)&\le z(\bar x_j)+\nabla z(\bar x_j)^\top(x-\bar x_j)+\varepsilon_k|x-\bar x_j|\end{aligned}$$

with $\varepsilon_k\searrow0$. Multiplying the second inequality by t_k and adding the first inequality, we obtain

$$(\bar p+t_k z)(x)-\alpha\le\frac12\nabla^2\bar p(\bar x_j)(x-\bar x_j)^2+t_k\nabla z(\bar x_j)^\top(x-\bar x_j)+\varepsilon_k t_k|x-\bar x_j|+\frac{\varepsilon_k}{2}|x-\bar x_j|^2.$$

Using Young's inequality and (5.8), we get

$$\begin{aligned}(\bar p+t_k z)(x)-\alpha&\le\frac12\nabla^2\bar p(\bar x_j)(x-\bar x_j)^2+t_k\nabla z(\bar x_j)^\top(x-\bar x_j)+\frac{\varepsilon_k}{2}t_k^2+\varepsilon_k|x-\bar x_j|^2\\&\le\frac{1-2\varepsilon_k/\theta}{2}\nabla^2\bar p(\bar x_j)(x-\bar x_j)^2+t_k\nabla z(\bar x_j)^\top(x-\bar x_j)+\frac{\varepsilon_k}{2}t_k^2\end{aligned}$$

for all $x\in B_{r_k}(\bar x_j)$. For k large enough, $1-2\varepsilon_k/\theta\ge0$. Since $\nabla^2\bar p(\bar x_j)$ is negative definite, all vectors $a,b\in\mathbb R^d$ obey the inequality $2a^\top b\le-\nabla^2\bar p(\bar x_j)a^2-\nabla^2\bar p(\bar x_j)^{-1}b^2$. Thus,

$$(\bar p+t_k z)(x)-\alpha\le-\frac{t_k^2}{2(1-2\varepsilon_k/\theta)}\nabla^2\bar p(\bar x_j)^{-1}\nabla z(\bar x_j)^2+\frac{\varepsilon_k}{2}t_k^2=:-t_k\tilde\xi_{j,k}$$

for all $x \in B_{r_k}(\bar{x}_j)$. Note that $\tilde{\xi}_{j,k} \le 0$. Together with (6.7) we infer

$$\forall x \in B_{r_0}(\bar{x}_j): \quad (\bar{p} + t_k(z + \tilde{\xi}_{j,k}))(x) \le \alpha.$$

For k large enough, we also get a lower bound from (6.3) and this results in

$$\forall x \in B_{r_0}(\bar{x}_j): \quad |(\bar{p} + t_k(z + \tilde{\xi}_{j,k}))(x)| \le \alpha.$$

For $j \in \{1, \ldots, N\}$ with $s_j = -1$, we can argue similarly and obtain the same estimate, with an appropriately defined $\tilde{\xi}_{j,k}$.

In order to glue these estimates together, let $(\varphi_j)_{j=1}^N \subset C^1(\Omega)$ denote Urysohn functions with

$$\varphi_j(x) \in [0,1] \quad \forall x \in \Omega, \qquad \varphi_j(x) = 1 \quad \forall x \in B_{r_0/2}(\bar{x}_j), \qquad \varphi_j(x) = 0 \quad \forall x \in \Omega \setminus B_{r_0}(\bar{x}_j)$$

for all $j = 1, \ldots, N$. We set

$$z_k \coloneqq z + \xi_k + \sum_{j=1}^N \varphi_j(\tilde{\xi}_{j,k} - \xi_k).$$

Note that $\bar{p} + t_k(z + \xi_k) \in \mathbb{B}_\alpha$ already implies that $|\bar{p} + t_k z_k| \le \alpha$ outside of the balls $B_{r_0}(\bar{x}_j)$. For $j \in \{1, \ldots, N\}$, we get on the set $B_{r_0}(\bar{x}_j)$

$$\bar{p} + t_k z_k = (1 - \varphi_j)(\bar{p} + t_k(z + \xi_k)) + \varphi_j(\bar{p} + t_k(z + \tilde{\xi}_{j,k})) \in [-\alpha, \alpha].$$

Consequently, $\bar{p} + t_k z_k \in \mathbb{B}_\alpha$, *i.e.*, $H(\bar{p} + t_k z_k) = 0$. Further, $z_k \to z$ in $C^1(\Omega)$ is clear since $\xi_k \to 0$ and $\tilde{\xi}_{j,k} \to 0$. Finally,

$$D(t_k, z_k) = -\sum_{j=1}^N \frac{\bar{\lambda}_j z_k(\bar{x}_j)}{t_k/2} = -\sum_{j=1}^N \frac{\bar{\lambda}_j \tilde{\xi}_{j,k}}{t_k/2} \to -\sum_{j=1}^N \bar{\lambda}_j \nabla^2 \bar{p}(\bar{x}_j)^{-1} \nabla z(\bar{x}_j)^2.$$

Hence, $(z_k) \subset C^1(\Omega)$ is the desired recovery sequence. $\square$

We remark that expressions similar to (6.4) arise when one studies second-order tangent sets (w.r.t. the norm of $C(\Omega)$) to the set of non-negative functions in $C(\Omega)$, see [36], Example 4.152. This is quite surprising, since we have worked in $C^1(\Omega)$ in order to arrive at (6.4).

6.2. Twice weak* epi-differentiability of G

Now, we are in position to characterize the weak* second subderivative of G. For this purpose, we exploit the derived representation of $H''(\bar{p}, [\![\bar{u}, 0]\!]; \cdot)$ as well as Lemma 4.5.

Proposition 6.4. *Under the assumptions of Lemma 6.2, we have*

$$G''([\![\bar{u}, 0]\!], \bar{p}, [\![\mu, \nu]\!]) \ge -\sum_{j=1}^N \frac{1}{\bar{\lambda}_j} \nabla^2 \bar{p}(\bar{x}_j) V_j^2 \tag{6.8}$$

whenever there holds

$$[\![\mu,\nu]\!] = \left[\!\!\left[\sum_{j=1}^N c_j\delta_{\bar{x}_j} + \mu_+ + \mu_-, \sum_{j=1}^N V_j\delta_{\bar{x}_j}\right]\!\!\right] \quad \textit{for some} \quad \begin{array}{r}(V_j,c_j)\in\left(\mathbb{R}^d\times\mathbb{R}\right),\ j=1,\dots,N,\\ \mu_+\in\mathcal{M}(\mathcal{I}_+)^+, \mu_-\in\mathcal{M}(\mathcal{I}_-)^-.\end{array}$$

Otherwise $G''([\![\bar{u},0]\!],\bar{p},[\![\mu,\nu]\!])=\infty$.

Proof. We utilize Proposition 6.3 and Lemma 4.5. In order to compute the conjugate, observe that $H''(\bar{p},[\![\bar{u},0]\!];\cdot)$ can be rewritten as a composition,

$$\frac{1}{2}H''(\bar{p},[\![\bar{u},0]\!];z) = -\sum_{j=1}^N \bar{\lambda}_j\nabla^2\bar{p}(\bar{x}_j)^{-1}\nabla z(\bar{x}_j)^2 + I_{\mathcal{Z}}(z) = g(Az),$$

where we define

$$A\colon C^1(\Omega)\to\left(\mathbb{R}^d\times\mathbb{R}\right)^N\times C(\mathcal{I}_+)\times C(\mathcal{I}_-),\quad Az=\left(\nabla z(\bar{x}_1),z(\bar{x}_1),\dots,\nabla z(\bar{x}_N),z(\bar{x}_N),z_{|\mathcal{I}_+},z_{|\mathcal{I}_-}\right)$$

as well as $g\colon \left(\mathbb{R}^d\times\mathbb{R}\right)^N\times C(\mathcal{I}_+)\times C(\mathcal{I}_-)\to(-\infty,\infty]$ by

$$g(V_1,c_1,\dots,V_N,c_N,w_+,w_-)=\sum_{j=1}^N\left[-\frac{\bar{\lambda}_j}{2}\nabla^2\bar{p}(\bar{x}_j)^{-1}V_j^2+I_{\{0\}}(c_j)\right]+I_{\le 0}(w_+)+I_{\ge 0}(w_-).$$

where

$$I_{\le 0}(w)=\begin{cases}0 & w\le 0,\\ \infty & \text{else,}\end{cases}\qquad I_{\ge 0}(w)=\begin{cases}0 & w\ge 0,\\ \infty & \text{else,}\end{cases}\qquad I_0(c)=\begin{cases}0 & c=0\\ \infty, & \text{else.}\end{cases}$$

We claim that there holds

$$\operatorname{dom} g-\operatorname{Ran}A=\left(\mathbb{R}^d\times\mathbb{R}\right)^N\times C(\mathcal{I}_+)\times C(\mathcal{I}_-). \tag{6.9}$$

Indeed, consider an arbitrary element

$$\Psi=(V_1,c_1,\dots,V_N,c_N,w_1,w_2)\in\left(\mathbb{R}^d\times\mathbb{R}\right)^N\times C(\mathcal{I}_+)\times C(\mathcal{I}_-)$$

and let $r>0$ be small enough such that the sets $(O_j)_{j=1}^{N+2}$, where

$$O_j=B_r(\bar{x}_j),\ j=1,\dots,N,\quad O_{N+1}=\mathcal{I}_+,\quad O_{N+2}=\mathcal{I}_-,$$

are pairwise disjoint. Furthermore, let $(\varphi_j)_{j=1}^{N+2}$ denote a family of Urysohn functions subordinate to $(O_j)_{j=1}^{N+2}$. Define

$$\Phi=(V_1,0,V_2,0,\dots,V_N,0,w_1-\|w_1\|_C,w_2+\|w_2\|_C)\in\operatorname{dom} g$$

as well as

$$z = -\sum_{j=1}^{N} c_j \varphi_j - \|w_1\|_C \varphi_{N+1} + \|w_2\|_C \varphi_{N+2}.$$

We readily verify $\Psi = \Phi - A(z)$, hence (6.9) holds. In particular, this implies $0 \in \operatorname{int}(\operatorname{dom} g - \operatorname{Ran} A)$. Thus, the conjugate of $g \circ A$ is given by the infimal postcomposition, cf. [37], Theorem 15.27[1], *i.e.*,

$$\begin{aligned}&\left(\frac{1}{2} H''(\bar{p}, [\![\bar{u}, 0]\!], \cdot)\right)^* ([\![\mu, \nu]\!]) \\ &\qquad = \inf\bigl\{ g^*(\bar{V}_1, \bar{c}_1, \ldots, \bar{V}_N, \bar{c}_N, \bar{\mu}_+, \bar{\mu}_-) \bigm| A^*(\bar{V}_1, \bar{c}_1, \ldots, \bar{V}_N, \bar{c}_N, \bar{\mu}_+, \bar{\mu}_-) = [\![\mu, \nu]\!] \bigr\},\end{aligned}$$

where we again adapt the convention of $\inf \emptyset = \infty$. We readily verify

$$A^*(\bar{V}_1, \bar{c}_1, \ldots, \bar{V}_N, \bar{c}_N, \bar{\mu}_+, \bar{\mu}_-) = \left[\!\!\left[\sum_{j=1}^{N} \bar{c}_j \delta_{\bar{x}_j} + \bar{\mu}_+ + \bar{\mu}_-, \sum_{j=1}^{N} \bar{V}_j \delta_{\bar{x}_j} \right]\!\!\right]. \tag{6.10}$$

This immediately implies $G''([\![\bar{u}, 0]\!], \bar{p}, [\![\mu, \nu]\!]) = \infty$ if $[\![\mu, \nu]\!]$ is not of the form

$$[\![\mu, \nu]\!] = \left[\!\!\left[\sum_{j=1}^{N} \bar{c}_j \delta_{\bar{x}_j} + \bar{\mu}_+ + \bar{\mu}_-, \sum_{j=1}^{N} \bar{V}_j \delta_{\bar{x}_j} \right]\!\!\right]$$

for some $(\bar{V}_j, \bar{c}_j) \in \mathbb{R}^d \times \mathbb{R}$ and $\bar{\mu}_\pm \in \mathcal{M}(\mathcal{I}_\pm)$. Next, we observe that the function g is given by a sum of independent terms. Consequently, we can compute the conjugate by summing the conjugates of the summands. This yields

$$g^*(\bar{V}_1, \bar{c}_1, \ldots, \bar{V}_N, \bar{c}_N, \bar{\mu}_+, \bar{\mu}_-) = -\sum_{j=1}^{N} \frac{1}{2\bar{\lambda}_j} \nabla^2 \bar{p}(\bar{x}_j) \bar{V}_j^2 + I_{\geq 0}(\bar{\mu}_+) + I_{\leq 0}(\bar{\mu}_-).$$

One can check that $\operatorname{Ran} A$ is dense, consequently, the adjoint A^* is injective. The claimed statement follows. $\square$

Now, in order to show the twice weak* epi-differentiability of G, we need to construct a suitable recovery sequence which achieves equality in (6.8). We start with a lemma in which we consider the contribution at a single point.

Lemma 6.5. *Let the assumptions of Lemma 6.2 be satisfied. Let $(V, c) \in \mathbb{R}^d \times \mathbb{R}$ and $j \in \{1, \ldots, N\}$ be given. We define*

$$[\![\mu, \nu]\!] \coloneqq [\![c\delta_{\bar{x}_j}, V\delta_{\bar{x}_j}]\!].$$

For an arbitrary sequence $(t_k) \subset (0, \infty)$ with $t_k \searrow 0$, we define

$$\mu_k \coloneqq c\delta_{\bar{x}_j} + (\bar{\lambda}_j + t_k c) \frac{\delta_{\bar{x}_j + (t_k/(\bar{\lambda}_j + t_k c))V} - \delta_{\bar{x}_j}}{t_k}$$

[1]This reference only deals with Hilbert spaces, but the proof can be applied to Banach spaces as well.

for k large enough such that $\bar\lambda_j + t_k c \neq 0$. Then there holds

$$[\![\mu_k, 0]\!] \rightharpoonup^* [\![\mu, \nu]\!] \quad \textit{in} \quad C^1(\Omega)^*$$

as well as

$$\lim_{k\to\infty} \frac{G\big([\![\bar\lambda_j \delta_{\bar x_j} + t_k\mu_k, 0]\!]\big) - G\big([\![\bar\lambda_j \delta_{\bar x_j}, 0]\!]\big) - t_k \langle\!\langle \bar p, [\![\mu_k, 0]\!]\rangle\!\rangle}{t_k^2/2} = -\frac{1}{\bar\lambda_j} \nabla^2 \bar p(\bar x_j) V^2. \tag{6.11}$$

Proof. In the sequel, k is large enough such that $t_k < |\bar\lambda_j|/|c|$ and $\bar x_j + \tilde t_k V \in \Omega$. In particular, $\bar\lambda_j + t_k c$ has the same sign as $\bar\lambda_j$ and, consequently, is not zero. For abbreviation, we set

$$\tilde t_k = t_k/(\bar\lambda_j + t_k c) \to 0.$$

In order to show the weak* convergence, let $\varphi \in C^1(\Omega)$ be arbitrary. Then we have

$$\langle\!\langle \varphi, [\![\mu_k, 0]\!]\rangle\!\rangle = c\varphi(\bar x_j) + \big(\varphi(\bar x_j + \tilde t_k V) - \varphi(\bar x_j)\big)/\tilde t_k.$$

Due to $\varphi \in C^1(\Omega)$, we arrive at

$$\lim_{k\to\infty} \langle\!\langle \varphi, [\![\mu_k, 0]\!]\rangle\!\rangle = c\varphi(\bar x_j) + V^\top \nabla\varphi(\bar x_j) = \langle\!\langle \varphi, [\![\mu, \nu]\!]\rangle\!\rangle.$$

Next, we show (6.11) in case $V = 0$. We have

$$\frac{G\big([\![\bar\lambda_j \delta_{\bar x_j} + t_k\mu_k, 0]\!]\big) - G\big([\![\bar\lambda_j \delta_{\bar x_j}, 0]\!]\big) - t_k \langle\!\langle \bar p, [\![\mu_k, 0]\!]\rangle\!\rangle}{t_k^2/2} = \frac{\alpha|\bar\lambda_j + t_k c| - \alpha|\bar\lambda_j| - t_k s_j \alpha c}{(t_k^2/2)} = 0,$$

where we used

$$|\bar\lambda_j + t_k c| = s_j(\bar\lambda_j + t_k c) = |\bar\lambda_j| + s_j t_k c \quad \text{as well as} \quad \bar p(\bar x_j) = s_j \alpha.$$

Finally, if $V \neq 0$, we have

$$t_k \mu_k = -\bar\lambda_j \delta_{\bar x_j} + (\bar\lambda_j + t_k c)\delta_{\bar x_j + \tilde t_k V}, \qquad \bar\lambda_j \delta_{\bar x_j} + t_k \mu_k = (\bar\lambda_j + t_k c)\delta_{\bar x_j + \tilde t_k V}.$$

Thus, we arrive at

$$\begin{aligned}
&\frac{G\big([\![\bar\lambda_j \delta_{\bar x_j} + t_k\mu_k, 0]\!]\big) - G\big([\![\bar\lambda_j \delta_{\bar x_j}, 0]\!]\big) - t_k \langle\!\langle \bar p, [\![\mu_k, 0]\!]\rangle\!\rangle}{t_k^2/2} \\
&\quad = \frac{\alpha|\bar\lambda_j + t_k c| - \alpha|\bar\lambda_j| + \bar\lambda_j \bar p(\bar x_j) - (\bar\lambda_j + t_k c)\bar p(\bar x_j + \tilde t_k V)}{t_k^2/2} \\
&\quad = \frac{(\bar\lambda_j + t_k c)\bar p(\bar x_j) - (\bar\lambda_j + t_k c)\bar p(\bar x_j + \tilde t_k V)}{t_k^2/2} = \frac{2}{(\bar\lambda_j + t_k c)} \frac{\bar p(\bar x_j) - \bar p(\bar x_j + \tilde t_k V)}{\tilde t_k^2} \\
&\quad \to -\frac{1}{\bar\lambda_j} \nabla^2 \bar p(\bar x_j) V^2,
\end{aligned}$$

where we used a second-order Taylor expansion of $\bar p$ at $\bar x_j$ together with $\nabla \bar p(\bar x_j) = 0$ in the last step. This finishes the proof. $\square$

Theorem 6.6. *We assume that $\bar u \in \mathcal{M}(\Omega)$ is of the form (5.6). Further, $\bar p \in C^1(\Omega) \cap \alpha\partial\|\bar u\|_{\mathcal{M}}$ is given such that (5.8) holds, where $\mathcal{A}$ is defined in (5.7).*

Then, the functional G is twice weakly epi-differentiable at $[\![\bar u, 0]\!]$ for $\bar p$. We have*

$$G''([\![\bar u, 0]\!], \bar p, [\![\mu, \nu]\!]) = -\sum_{j=1}^{N} \frac{1}{\bar\lambda_j} \nabla^2 \bar p(\bar x_j) V_j^2, \tag{6.12}$$

whenever there holds

$$[\![\mu, \nu]\!] = \left[\!\!\left[\sum_{j=1}^{N} c_j \delta_{\bar x_j} + \mu_+ + \mu_-, \sum_{j=1}^{N} V_j \delta_{\bar x_j}\right]\!\!\right] \quad \textit{for some} \quad \begin{array}{c} (V_j, c_j) \in \left(\mathbb{R}^d \times \mathbb{R}\right),\ j = 1, \dots, N, \\ \mu_+ \in \mathcal{M}(\mathcal{I}_+)^+, \mu_- \in \mathcal{M}(\mathcal{I}_-)^-. \end{array} \tag{6.13}$$

Otherwise $G''([\![\bar u, 0]\!], \bar p, [\![\mu, \nu]\!]) = \infty$.

Proof. From Proposition 6.4 we already know that the second subderivative is ∞ if $[\![\mu, \nu]\!]$ is not as in (6.13) and that "$\ge$" in (6.12) holds. It remains to construct a recovery sequence for $[\![\mu, \nu]\!]$ as in (6.13). To this end, let a sequence $(t_k) \subset (0, \infty)$ with $t_k \searrow 0$ be given. In view of Lemma 6.5, we define

$$\mu_k \coloneqq \sum_{j=1}^{N} \mu_{j,k} + \mu_+ + \mu_- \coloneqq \sum_{j=1}^{N} \left[c_j \delta_{\bar x_j} + (\bar\lambda_j + t_k c_j) \frac{\delta_{\bar x_j + (t_k/(\bar\lambda_j + t_k c_j))V_j} - \delta_{\bar x_j}}{t_k} \right] + \mu_+ + \mu_-$$

for k large enough. From Lemma 6.5 we get $[\![\mu_k, 0]\!] \overset{*}{\rightharpoonup} [\![\mu, \nu]\!]$. Moreover, for k large enough, we have

$$\begin{aligned} G\big([\![\bar u + t_k \mu_k, 0]\!]\big) &= \sum_{j=1}^{N} G\big([\![\bar\lambda_j \delta_{\bar x_j} + t_k \mu_{j,k}, 0]\!]\big) + t_k G(\mu_+) + t_k G(\mu_-), \\ G\big([\![\bar u, 0]\!]\big) &= \sum_{j=1}^{N} G\big([\![\bar\lambda_j \delta_{\bar x_j}, 0]\!]\big), \\ \langle\!\langle \bar p, [\![\mu_k, 0]\!] \rangle\!\rangle &= \sum_{j=1}^{N} \langle\!\langle \bar p, [\![\mu_{j,k}, 0]\!] \rangle\!\rangle + \alpha \mu_+(\mathcal{I}_+) + \alpha |\mu_-(\mathcal{I}_-)|. \end{aligned}$$

Note that $G(\mu_+) = \mu_+(\mathcal{I}_+)$ and $G(\mu_-) = |\mu_-(\mathcal{I}_-)|$. Consequently, summing over (6.11) yields

$$\lim_{k\to\infty} \frac{G\big([\![\bar u + t_k \mu_k, 0]\!]\big) - G\big([\![\bar u, 0]\!]\big) - t_k \langle\!\langle \bar p, [\![\mu_k, 0]\!] \rangle\!\rangle}{t_k^2/2} = -\sum_{j=1}^{N} \frac{1}{\bar\lambda_j} \nabla^2 \bar p(\bar x_j) V^2.$$

This proves the claim. ☐

Now, we can finally prove that (B1) implies (SSC-$\mathcal{M}$).

Lemma 6.7. *Let the assumptions of Theorem 5.17 be satisfied. Then, (B1) implies (SSC-$\mathcal{M}$). In case that L is convex, (B1) is equivalent to (C1).*

Proof. According to Theorem 6.6, it suffices to show that

$$(\mathcal{K}h, \nabla^2 L(K\bar u)\mathcal{K}h)_Y + G''([\![\bar u, 0]\!], \bar p; h) > 0 \tag{6.14}$$

for all $h = [\![\mu, \nu]\!] \neq 0$ as in (6.13). Together with the expression (6.12) for the second subderivative, this is precisely condition (5.9).

It remains to consider the situation with a convex L. We have to prove the equivalence of (5.9) and (5.12). It is clear that (5.9) implies (5.12).

Now, let (5.12) be satisfied. For an $h = [\![\mu, \sum_{j=1}^N V_j \delta_{\bar{x}_j}]\!] \neq 0$ with $\mu \in \mathcal{C}$ and $V \in (\mathbb{R}^d)^N$, we note that

$$\begin{aligned}(\mathcal{K}h, \nabla^2 L(K\bar{u})\mathcal{K}h)_Y + G''([\![\bar{u}, 0]\!], \bar{p}; h) &\geq (\mathcal{K}h, \nabla^2 L(K\bar{u})\mathcal{K}h)_Y - \sum_{j=1}^N \left[\frac{1}{\bar{\lambda}_j} \nabla^2 \bar{p}(\bar{x}_j) V_j^2\right] \\ &\geq (\mathcal{K}h, \nabla^2 L(K\bar{u})\mathcal{K}h)_Y + \theta \sum_{j=1}^N |V_j|^2.\end{aligned}$$

Due to the convexity of L, the first addend on the right-hand side is nonnegative. Thus, if there exists $j \in \{1, \dots, N\}$ with $V_j \neq 0$, the right-hand side is positive. Otherwise, $V_j = 0$ for all $j \in \{1, \dots, N\}$. Hence, we can apply (5.12) and, again, obtain the positivity of the right-hand side. This shows that (5.9) holds. ☐

7. Structural assumptions imply NDC

It remains to show that the structural assumption (B1) implies the non-degeneracy condition (NDC-$\mathcal{M}$) as stated in the next result.

Lemma 7.1. *Let the assumptions of Theorem 5.17 be satisfied. Then,* (B1) *implies* (NDC-$\mathcal{M}$).

We give two fundamentally different proofs for this fact, one based on the transport-based representation of the bounded Lipschitz norm, see Lemma 5.10, and another utilizing the abstract result of Lemma 4.6.

7.1. Proof of Lemma 7.1 based on Lemma 5.10

In the following, let assumption (B1) hold but assume that (NDC-$\mathcal{M}$) does not. W.l.o.g., and recalling the definition of the lifted functional G, (5.4), the latter implies that there exists sequences $(t_k) \subset (0, \infty)$, $(\mu_k) \subset \mathcal{M}(\Omega)$ with

$$\begin{gathered} t_k \searrow 0,\ [\![\mu_k, 0]\!] \rightharpoonup^* 0,\ \|[\![\mu_k, 0]\!]\|_{(C^1)^*} = 1, \text{ and} \\ \lim_{k\to\infty} \frac{\alpha\|\bar{u} + t_k\mu_k\|_{\mathcal{M}} - \alpha\|\bar{u}\|_{\mathcal{M}} - t_k\langle \bar{p}, \mu_k\rangle}{t_k^2} = \lim_{k\to\infty} \frac{\alpha\|\bar{u} + t_k\mu_k\|_{\mathcal{M}} - \langle \bar{p}, \bar{u} + t_k\mu_k\rangle}{t_k^2} = 0. \end{gathered} \tag{7.1}$$

We want to lead this to a contradiction. For this purpose, let $r_0 > 0$ denote the radius from Lemma 6.2. The following lemma is a direct consequence of Lemma 6.2, the definition of $\mathcal{I}_\pm$ as well as the continuity of $\bar{p}$.

Lemma 7.2. *Let Assumption* (B1) *hold and denote by $r_0 > 0$ the radius from Lemma 6.2. Then there exists a compact set $\mathcal{I}$ such that:*

(1) The sets $B_{r_0}(\bar{x}_j)$, $j = 1, \dots, N$, and $\mathcal{I}$ are pairwise disjoint.
(2) There holds

$$(\mathcal{I}_+ \cup \mathcal{I}_-) \subset \mathcal{I} \qquad \text{as well as} \qquad |\bar{p}(x)| \leq \alpha - \sigma \qquad \forall x \in \Omega \setminus \left(\mathcal{I} \cup \bigcup_{j=1}^N B_{r_0}\right) \tag{7.2}$$

for some $\sigma > 0$.

We define by

$$\mu_{k,\mathcal{I}} \coloneqq \mu_k(\cdot \cap \mathcal{I}), \quad \mu_{k,j} \coloneqq \mu_k(\cdot \cap B_{r_0}(\bar{x}_j)),$$

the restrictions of μ_k to $\mathcal{I}$ and $B_{r_0}(\bar{x}_j)$, $j = 1, \ldots, N$, respectively, as well as

$$\tilde{\mu}_k \coloneqq \mu_k - \mu_{k,\mathcal{I}} - \sum_{j=1}^{N} \mu_{k,j} \qquad \text{and} \qquad \mu_{k,j}^{\perp} \coloneqq \mu_{k,j} - \mu_{k,j}(\{\bar{x}_j\})\delta_{\bar{x}_j}.$$

Furthermore, we denote by $\mu_{k,j}^{\perp,\pm}$ the positive and negative part of $\mu_{k,j}^{\perp}$, respectively, *i.e.*, there holds

$$\mu_{k,j}^{\perp} = \mu_{k,j}^{\perp,+} - \mu_{k,j}^{\perp,-}, \quad \|\mu_{k,j}^{\perp}\|_{\mathcal{M}} = \|\mu_{k,j}^{\perp,+}\|_{\mathcal{M}} + \|\mu_{k,j}^{\perp,-}\|_{\mathcal{M}},$$

and recall that $s_j = \operatorname{sgn}(\bar{p}(\bar{x}_j))$, $j = 1, \ldots, N$. Finally, we abbreviate

$$D_1^k = \sum_{j=1}^{N} \left[\alpha|\bar{\lambda}_j + t_k\mu_{k,j}(\{\bar{x}_j\})| - s_j\alpha(\bar{\lambda}_j + t_k\mu_{k,j}(\{\bar{x}_j\}))\right],\ D_2^k = \sum_{j=1}^{N} \left[\alpha\|\mu_{k,j}^{\perp,-s_j}\|_{\mathcal{M}} + s_j\langle \bar{p}, \mu_{k,j}^{\perp,-s_j}\rangle\right]$$

as well as

$$D_3^k = \alpha\|\mu_{k,\mathcal{I}}\|_{\mathcal{M}} - \langle \bar{p}, \mu_{k,\mathcal{I}}\rangle,\ D_4^k = \alpha\|\tilde{\mu}_k\|_{\mathcal{M}} - \langle \bar{p}, \tilde{\mu}_k\rangle, \quad D_5^k = \sum_{j=1}^{N} \left[\alpha\|\mu_{k,j}^{\perp,s_j}\|_{\mathcal{M}} - s_j\langle \bar{p}, \mu_{k,j}^{\perp,s_j}\rangle\right].$$

The following estimates are immediate.

Lemma 7.3. *Let Assumption (B1) hold and let (μ_k), (t_k) satisfy (7.1). Then there holds*

$$t_k^{-2}(\alpha\|\bar{u} + t_k\mu_k\|_{\mathcal{M}} - \langle \bar{p}, \bar{u} + t_k\mu_k\rangle) = t_k^{-2}D_1^k + t_k^{-1}\sum_{i=2}^{5} D_i^k$$

as well as

$$D_1^k \geq -2\alpha\sum_{j=1}^{N} \min\{s_j(\bar{\lambda}_j + t_k\mu_{k,j}(\{\bar{x}_j\})), 0\} \geq 0, \qquad D_2^k \geq \alpha\sum_{j=1}^{N} \|\mu_{k,j}^{\perp,-s_j}\|_{\mathcal{M}} \geq 0,$$

$$D_4^k \geq \sigma\|\tilde{\mu}_k\|_{\mathcal{M}} \geq 0, \qquad D_5^k \geq \frac{\theta}{4}\sum_{j=1}^{N} \int_{\Omega} |\bar{x}_j - x|^2 \,\mathrm{d}\mu_{k,j}^{\perp,s_j} \geq 0.$$

Proof. We derive the lower bound

$$t_k^{-2}(\alpha\|\bar{u} + t_k\mu_k\|_{\mathcal{M}} - \langle \bar{p}, \bar{u} + t_k\mu_k\rangle) \geq t_k^{-2}D_1^k + t_k^{-1}\sum_{i=2}^{5} D_i^k$$

noting that there holds

$$\|\mu_k\|_{\mathcal{M}} = \|\tilde{\mu}_k\|_{\mathcal{M}} + \sum_{j=1}^{N} \left[|\mu_{k,j}(\{\bar{x}_j\})| + \|\mu_{k,j}^{\perp,s_j}\|_{\mathcal{M}} + \|\mu_{k,j}^{\perp,-s_j}\|_{\mathcal{M}} \right] + \|\mu_{k,\mathcal{I}}\|_{\mathcal{M}}.$$

The lower estimates on D_i^k follow immediately noting that

$$D_2^k = \sum_{j=1}^{N} \left[\alpha \|\mu_{k,j}^{\perp,-s_j}\|_{\mathcal{M}} + \langle |\bar{p}|, \mu_{k,j}^{\perp,-s_j} \rangle \right] \geq \alpha \sum_{j=1}^{N} \|\mu_{k,j}^{\perp,-s_j}\|_{\mathcal{M}}, \quad D_4^k = \alpha \|\tilde{\mu}_k\|_{\mathcal{M}} - \langle \bar{p}, \tilde{\mu}_k \rangle \geq \sigma \|\tilde{\mu}_k\|_{\mathcal{M}}$$

as well as

$$D_5^k = \sum_{j=1}^{N} \left[\alpha \|\mu_{k,j}^{\perp,s_j}\|_{\mathcal{M}} - s_j \langle \bar{p}, \mu_{k,j}^{\perp,s_j} \rangle \right] = \sum_{j=1}^{N} \left[\alpha \|\mu_{k,j}^{\perp,s_j}\|_{\mathcal{M}} - \langle |\bar{p}|, \mu_{k,j}^{\perp,s_j} \rangle \right] \geq \frac{\theta}{4} \sum_{j=1}^{N} \int_{\Omega} |\bar{x}_j - x|^2 \, \mathrm{d}\mu_{k,j}^{\perp,s_j}$$

due to Lemma 6.2. $\square$

Using this auxiliary lemma, we argue that, for increasing k, μ_k strongly approximates the surrogate sequence

$$\widehat{\mu}_k := \sum_{j=1}^{N} \left[-s_j \mu_{k,j}^{\perp,s_j}(\Omega) \delta_{\bar{x}_j} + s_j \mu_{k,j}^{\perp,s_j} \right]$$

in $\mathcal{M}(\Omega)$.

Proposition 7.4. *Let assumption (B1) hold and let (μ_k), (t_k) satisfy (7.1). Then there holds*

$$\lim_{k\to\infty} \left[\|\tilde{\mu}_k\|_{\mathcal{M}} + \|\mu_{k,\mathcal{I}}\|_{\mathcal{M}} + \sum_{j=1}^{N} \left[\min\{s_j(\bar{\lambda}_j + t_k \mu_{k,j}(\{\bar{x}_j\})), 0\} + \|\mu_{k,j}^{\perp,-s_j}\|_{\mathcal{M}} \right] \right] = 0 \tag{7.3}$$

as well as

$$\lim_{k\to\infty} \left[\sum_{j=1}^{N} |s_j \mu_{k,j}(\{\bar{x}_j\}) + \mu_{k,j}^{\perp,s_j}(\Omega)| \right] = 0. \tag{7.4}$$

In particular, we have

$$\lim_{k\to\infty} \left[\|[\![\mu_k - \widehat{\mu}_k, 0]\!]\|_{(C^1)^*} + |\|[\![\widehat{\mu}_k, 0]\!]\|_{(C^1)^*} - 1| \right] = 0$$

and the sequence $\left(\sum_{j=1}^{N} \|\mu_{k,j}^{\perp,s_j}\|_{\mathcal{M}}\right)$ is unbounded.

Proof. We immediately conclude

$$\lim_{k\to\infty} \left[\|\tilde{\mu}_k\|_{\mathcal{M}} + \sum_{j=1}^{N} \left[\min\{s_j(\bar{\lambda}_j + t_k \mu_{k,j}(\{\bar{x}_j\})), 0\} + \|\mu_{k,j}^{\perp,-s_j}\|_{\mathcal{M}} \right] \right] = 0$$

from the lower bounds on D_1^k, D_2^k and D_4^k in Lemma 7.3, respectively. Moreover, due to (7.1) as well as Lemma 7.3, we have

$$\lim_{k\to\infty} D_3^k = \lim_{k\to\infty} [\alpha\|\mu_{k,\mathcal{I}}\|_{\mathcal{M}} - \langle \bar{p}, \mu_{k,\mathcal{I}}\rangle] = 0.$$

Let $\{\varphi_j\}_{j=1}^N \cup \{\varphi_{\mathcal{I}}\} \subset C^1(\Omega)$ denote a family of smooth Urysohn functions subordinate to $B_{r_0}(\bar{x}_j)$, $j = 1, \dots, N$, and $\mathcal{I}$. Then we have

$$0 = \lim_{k\to\infty} \langle\!\langle \bar{p}\varphi_{\mathcal{I}}, [\![\mu_k, 0]\!]\rangle\!\rangle = \lim_{k\to\infty} [\langle \bar{p}, \mu_{k,\mathcal{I}}\rangle + \langle \bar{p}\varphi_{\mathcal{I}}, \tilde{\mu}_k\rangle] = \lim_{k\to\infty} \langle \bar{p}, \mu_{k,\mathcal{I}}\rangle.$$

and consequently $\lim_{k\to\infty}\|\mu_{k,\mathcal{I}}\|_{\mathcal{M}} = 0$. Similarly, observe that

$$0 = \lim_{k\to\infty} s_j \langle\!\langle \varphi_j, [\![\mu_k, 0]\!]\rangle\!\rangle = \lim_{k\to\infty} \left[s_j \mu_{k,j}(\{\bar{x}_j\}) + \mu_{k,j}^{\perp,s_j}(\Omega) - \mu_{k,j}^{\perp,-s_j}(\Omega) + s_j \langle \varphi_j, \tilde{\mu}_k\rangle \right]$$

for all $j = 1, \dots, N$. Using (7.3), we conclude (7.4). The statements

$$\lim_{k\to\infty} \left[\|[\![\mu_k - \widehat{\mu}_k, 0]\!]\|_{(C^1)^*} + |\|[\![\widehat{\mu}_k, 0]\!]\|_{(C^1)^*} - 1| \right] = 0$$

immediately follow from (7.3) and (7.4), respectively, $\|[\![\mu_k, 0]\!]\|_{(C^1)^*} = 1$, as well as noting that

$$\mu_k - \widehat{\mu}_k = \tilde{\mu}_k + \mu_{k,\mathcal{I}} + \sum_{j=1}^N \left[-s_j \mu_{k,j}^{\perp,-s_j} + (\mu_{k,j}(\{\bar{x}_j\}) + s_j \mu_{k,j}^{\perp,s_j}(\Omega))\delta_{\bar{x}_j} \right]$$

and thus

$$\lim_{k\to} \|\mu_k - \widehat{\mu}_k\|_{\mathcal{M}} \le \lim_{k\to\infty} \left[\|\tilde{\mu}_k\|_{\mathcal{M}} + \|\mu_{k,\mathcal{I}}\|_{\mathcal{M}} + \sum_{j=1}^N \left[\|\mu_{k,j}^{\perp,-s_j}\|_{\mathcal{M}} + |\mu_{k,j}(\{\bar{x}_j\}) + s_j \mu_{k,j}^{\perp,s_j}(\Omega)| \right] \right] = 0$$

Finally, assume that $\sum_{j=1}^N \|\mu_{k,j}^{\perp,s_j}\|_{\mathcal{M}}$ is bounded. In view of (7.3) and (7.4), we then conclude that (μ_k) is bounded in $\mathcal{M}(\Omega)$. Thus we have $\mu_k \rightharpoonup^* 0$ in $\mathcal{M}(\Omega)$. By a compactness argument, this contradicts $\|[\![\mu_k, 0]\!]\|_{(C^1)^*} = 1$. ☐

Thus, w.l.o.g, we restrict ourselves to sequences (μ_k) with $\sum_{j=1}^N \|\mu_{k,j}^{\perp,s_j}\|_{\mathcal{M}} > 0$ in the following. The next corollary shows that the latter grows at most at a rate of t_k^{-1}.

Corollary 7.5. *Let Assumption (B1) hold and let (μ_k), (t_k) satisfy (7.1). For all $k \in \mathbb{N}$ large enough as well as all $j \in \{1, \dots, N\}$, there holds*

$$\frac{|\bar{\lambda}_j| + 1/2}{t_k} \ge -s_j \mu_{k,j}(\{\bar{x}_j\}) \quad \textit{as well as} \quad \|\mu_{k,j}^{\perp,s_j}\|_{\mathcal{M}} \le 1/2 + t_k^{-1}(|\bar{\lambda}_j| + 1/2).$$

In particular, this implies

$$\sum_{j=1}^N \|\mu_{k,j}^{\perp,s_j}\|_{\mathcal{M}} \le N/2 + t_k^{-1} \left(\sum_{j=1}^N |\bar{\lambda}_j| + N/2 \right).$$

Proof. According to (7.3) and (7.4), there holds

$$\|\mu_{k,j}^{\perp,s_j}\|_{\mathcal{M}} = \mu_{k,j}^{\perp,s_j}(\Omega) \le |s_j\mu_{k,j}(\{\bar{x}_j\}) + \mu_{k,j}^{\perp,s_j}(\Omega)| - s_j\mu_{k,j}(\{\bar{x}_j\}) \le 1/2 - s_j\mu_{k,j}(\{\bar{x}_j\})$$

as well as

$$s_j(\bar{\lambda}_j + t_k\mu_{k,j}(\{\bar{x}_j\})) = |\bar{\lambda}_j| + s_j t_k \mu_{k,j}(\{\bar{x}_j\}) \ge -1/2$$

for all $k \in \mathbb{N}$ large enough. Rearranging yields the desired results. □

Finally, we argue that the lower bound on D_5^k in Lemma 7.3 decreases at most at a linear rate.

Corollary 7.6. *Let Assumption (B1) hold and let (μ_k), (t_k) satisfy (7.1). Moreover, let $c_1 > 0$ be such that*

$$\|\mu\|_{\mathrm{BL}} \ge c_1 \|[\![\mu, 0]\!]\|_{(C^1)^*} \quad \forall \mu \in \mathcal{M}(\Omega).$$

For all $k \in \mathbb{N}$ large enough, there holds

$$\frac{c_1}{2}\left(\sum_{j=1}^{N} \|\mu_{k,j}^{\perp,s_j}\|_{\mathcal{M}}\right)^{-1} \le \sum_{j=1}^{N} \int_\Omega |\bar{x}_j - x|^2 \,\mathrm{d}\mu_{k,j}^{\perp,s_j}$$

Proof. Note that the positive and negative part of the surrogate $\widehat{\mu}_k$ are given by

$$\widehat{\mu}_k^+ = \sum_{j=1}^{N} \mu_{k,j}^{\perp,s_j} + \sum_{j=1}^{N} \mu_{k,j}^{\perp,s_j}(\Omega)\delta_{\bar{x}_j}, \quad \widehat{\mu}_k^- = \sum_{j=1}^{N} \mu_{k,j}^{\perp,s_j} + \sum_{j=1}^{N} \mu_{k,j}^{\perp,s_j}(\Omega)\delta_{\bar{x}_j}.$$

By construction, there holds $\|\widehat{\mu}_k^+\|_{\mathcal{M}} = \|\widehat{\mu}_k^-\|_{\mathcal{M}}$ and the canonical product measure

$$\gamma_k = \sum_{j=1}^{N} \mu_{k,j}^{\perp,s_j} \otimes \delta_{\bar{x}_j} + \sum_{j=1}^{N} \delta_{\bar{x}_j} \otimes \mu_{k,j}^{\perp,s_j}$$

satisfies

$$\gamma_k(B \times \Omega) = \widehat{\mu}_k^+(B) \quad \text{as well as} \quad \gamma_k(\Omega \times B) = \widehat{\mu}_k^-(B) \tag{7.5}$$

for all $B \in \mathcal{B}(\Omega)$ as well as

$$\|\gamma_k\|_{\mathcal{M}(\Omega\times\Omega)} = \sum_{j=1}^{N} \|\mu_{k,j}^{\perp,s_j}\|_{\mathcal{M}} > 0. \tag{7.6}$$

Hence γ_k is an admissible transport plan between $\widehat{\mu}_k^+$ and $\widehat{\mu}_k^-$. By Jensen's inequality as well as the definition of the bounded Lipschitz norm, we further conclude

$$\sum_{j=1}^{N} \int_\Omega |\bar{x}_j - x|^2 \,\mathrm{d}\mu_{k,j}^{\perp,s_j} = \int_{\Omega\times\Omega} |y-x|^2 \,\mathrm{d}\gamma_k(x,y) \ge \|\gamma_k\|_{\mathcal{M}(\Omega\times\Omega)}^{-1} \left(\int_{\Omega\times\Omega} |y-x| \,\mathrm{d}\gamma_k\right)^2 \ge \|\gamma_k\|_{\mathcal{M}(\Omega\times\Omega)}^{-1} \|\widehat{\mu}_k\|_{\mathrm{BL}}^2.$$

The claimed statement now follows by definition of c_1 as well as $\lim_{k\to\infty} \|[\![\widehat{\mu}_k, 0]\!]\|_{(C^1)^*} = 1$ and (7.6). □

Combining Corollaries 7.5 and 7.6, we are now prepared for the proof of Lemma 7.1.

Proof of Lemma 7.1. Assume that (NDC-$\mathcal{M}$) does not hold, *i.e.*, there are sequences (μ_k) and (t_k) satisfying (7.1). By Lemma 7.3, this also implies

$$\lim_{k\to 0} t_k^{-1} \sum_{j=1}^{N} \int_\Omega |\bar{x}_j - x|^2 \, \mathrm{d}\mu_{k,j}^{\perp,s_j} = 0.$$

However, following Corollaries 7.5 and 7.6, we also have

$$\liminf_{k\to\infty} t_k^{-1} \sum_{j=1}^{N} \int_\Omega |\bar{x}_j - x|^2 \, \mathrm{d}\mu_{k,j}^{\perp,s_j} \geq \liminf_{k\to\infty} c_1/2 \left(t_k N/2 + \sum_{j=1}^{N} |\bar{\lambda}_j| + N/2 \right) > 0$$

yielding a contradiction. □

7.2. Proof of Lemma 7.1 based on Lemma 4.6

We give a second proof of Lemma 7.1 which is based on Lemma 4.6.

Proof of Lemma 7.1. In order to apply Lemma 4.6, we verify that the structural assumptions imply (4.3). Let $r_0 > 0$ from Lemma 6.2 be given. Let $(\varphi_j)_{j=1}^N \subset C^1(\Omega)$ denote Urysohn functions with

$$\varphi_j(x) \in [0,1] \quad \forall x \in \Omega, \qquad \varphi_j(x) = 1 \quad \forall x \in B_{r_0/2}(\bar{x}_j), \qquad \varphi_j(x) = 0 \quad \forall x \in \Omega \setminus B_{r_0}(\bar{x}_j)$$

for all $j = 1, \ldots, N$. We further set $\varphi_{N+1} := \left(1 - \sum_{j=1}^N \varphi_j\right) \alpha^{-1} \bar{p}$ and $s_{N+1} := 1$.

We set $\eta := \min\{\alpha/8, \theta r_0^2/16\} > 0$. Let $p \in C^1(\Omega)$ with $\|p - \bar{p}\|_{C^1} \leq \eta$ be arbitrary. We define

$$\beta_{N+1} := \|p - \bar{p}\|_C, \qquad \beta_j := \sup\{s_j p(x) \mid x \in B_{r_0/2}(\bar{x}_j)\} - \alpha \qquad \forall j = 1, \ldots, N.$$

Let us check that $\hat{p} := p - \sum_{j=1}^{N+1} s_j \beta_j \varphi_j \in \mathbb{B}_\alpha$. For $x \in B_{r_0/2}(\bar{x}_j)$, $j \in \{1, \ldots, N\}$, we have

$$|\hat{p}(x)| = |p(x) - s_j \beta_j| = s_j p(x) - \beta_j \leq \alpha.$$

For $x \in B_{r_0}(\bar{x}_j) \setminus B_{r_0/2}(\bar{x}_j)$, $j \in \{1, \ldots, N\}$, we have

$$\begin{aligned} |\hat{p}(x)| &= |p(x) - s_j \beta_j \varphi_j(x) - \beta_{N+1} \varphi_{N+1}(x)| = s_j p(x) - \beta_j \varphi_j(x) - s_j \|p - \bar{p}\|_C (1 - \varphi_j(x)) \alpha^{-1} \bar{p}(x) \\ &\leq s_j p(x) \leq s_j \bar{p}(x) + \|p - \bar{p}\|_C \leq \alpha - \frac{\theta}{4} \left(\frac{r_0}{2}\right)^2 + \|p - \bar{p}\|_C \leq \alpha, \end{aligned}$$

where we used (6.3). Finally, for $x \in \Omega \setminus \bigcup_{j=1}^N B_{r_0}(\bar{x}_j)$ we have

$$\begin{aligned} |\hat{p}(x)| &= |p(x) - \beta_{N+1} \alpha^{-1} \bar{p}(x)| \leq \|p - \bar{p}\|_C + |\bar{p}(x)| |1 - \beta_{N+1} \alpha^{-1}| \\ &\leq \|p - \bar{p}\|_C + \alpha |1 - \|p - \bar{p}\|_C \alpha^{-1}| = \alpha. \end{aligned}$$

This shows $\hat{p} = p - \sum_{j=1}^{N+1} s_j \beta_j \varphi_j \in \mathbb{B}_\alpha = \mathrm{dom}(H)$.

Now, let $\tilde H$ and H_2 be as in Lemma 4.6 with $K \coloneqq N+1$, $\bar x \coloneqq [\![\bar u, 0]\!]$ and $\zeta_i \coloneqq \varphi_i$. Thus,

$$\tilde H(p) \le H\left(p - \sum_{j=1}^{N+1} s_j\beta_j\varphi_j\right) + H_2\left(\sum_{j=1}^{N+1} s_j\beta_j\varphi_j\right) = H_2\left(\sum_{j=1}^{N+1} s_j\beta_j\varphi_j\right) = \frac{1}{2}\sum_{j=1}^{N+1}\beta_j^2 + \sum_{j=1}^{N}\beta_j|\bar\lambda_j|,$$

where we used $\langle\!\langle \varphi_j, [\![\bar u, 0]\!]\rangle\!\rangle = \langle \varphi_j, \bar u\rangle = \bar\lambda_j$, for all $j = 1, \ldots, N$ and $\langle\!\langle \varphi_{N+1}, [\![\bar u, 0]\!]\rangle\!\rangle = \langle \varphi_{N+1}, \bar u\rangle = 0$. Moreover,

$$\langle p - \bar p, \bar u\rangle = \sum_{j=1}^{N}(p(\bar x_j) - \alpha s_j)\bar\lambda_j = \sum_{j=1}^{N}(s_j p(\bar x_j) - \alpha)|\bar\lambda_j|.$$

Hence,

$$\begin{aligned}\tilde H(p) - H(\bar p) - \langle p - \bar p, \bar u\rangle &\le \frac{1}{2}\sum_{j=1}^{N+1}\beta_j^2 + \sum_{j=1}^{N}\beta_j|\bar\lambda_j| - \sum_{j=1}^{N}(s_j p(\bar x_j) - \alpha)|\bar\lambda_j| \\ &= \frac{1}{2}\sum_{j=1}^{N+1}\beta_j^2 + \sum_{j=1}^{N}(\beta_j - s_j p(\bar x_j) + \alpha)|\bar\lambda_j|.\end{aligned}$$

For the first addend, we can directly use $|\beta_j| \le \|p - \bar p\|_C$. For the second addend, we consider

$$\begin{aligned}\beta_j - s_j p(\bar x_j) + \alpha &= \sup\{s_j p(x) \mid x \in B_{r_0}(\bar x_j)\} - s_j p(\bar x_j) \\ &= \sup\{s_j(p(x) - p(\bar x_j)) \mid x \in B_{r_0}(\bar x_j)\}.\end{aligned}$$

For any $x \in B_{r_0}(\bar x_j)$ we use (6.3) to obtain

$$\begin{aligned}s_j(p(x) - p(\bar x_j)) &= s_j\left(\int_0^1 \nabla p(\bar x_j + t(x - \bar x_j))^\top (x - \bar x_j)\,\mathrm{d}t\right) \\ &\le \|p - \bar p\|_{C^1}|x - \bar x_j| + s_j\left(\int_0^1 \nabla \bar p(\bar x_j + t(x - \bar x_j))^\top (x - \bar x_j)\,\mathrm{d}t\right) \\ &= \|p - \bar p\|_{C^1}|x - \bar x_j| + s_j(\bar p(x) - \bar p(\bar x_j)) \\ &\le \|p - \bar p\|_{C^1}|x - \bar x_j| - \frac{\theta}{4}|x - \bar x_j|^2 \le \frac{1}{\theta}\|p - \bar p\|_{C^1}^2.\end{aligned}$$

By collecting the inequalities from above, we obtain that (4.3) is satisfied for some $\Lambda > 0$. Thus, the proof is finalized by the invocation of Lemma 4.6. □

Acknowledgments

The authors would like to thank Hannes Meinlschmidt for pointing out the notion of "uniform local quasiconvexity".

Data availability statement

No new data/code were created or analyzed in this study.

References

[1] K. Bredies and H.K. Pikkarainen, Inverse problems in spaces of measures. *ESAIM: Control Optim. Calc. Var.* **19** (2013) 190–218.

[2] O. Scherzer and B. Walch, Sparsity regularization for radon measures, in *Scale Space and Variational Methods in Computer Vision*, edited by X.-C. Tai, K. Mørken, M. Lysaker and K.-A. Lie, editors. Springer Berlin Heidelberg, Berlin, Heidelberg (2009) 452–463.

[3] K. Kunisch, P. Trautmann and B. Vexler, Optimal control of the undamped linear wave equation with measure valued controls. *SIAM J. Control Optim.* **54** (2016) 1212–1244.

[4] K. Pieper, B.Q. Tang, P. Trautmann and D. Walter, Inverse point source location with the Helmholtz equation on a bounded domain. *Computat. Optim. Appl.* **77** (2020) 213–249.

[5] Q. Denoyelle, V. Duval, G. Peyré and E. Soubies, The sliding Frank-Wolfe algorithm and its application to super-resolution microscopy. *Inverse Probl.* **36** (2020) 014001.

[6] C.W. McCutchen, Superresolution in microscopy and the abbe resolution limit. *J. Opt. Soc. Am.* **57** (1967) 1190–1192.

[7] E. Casas, C. Clason and K. Kunisch, Approximation of elliptic control problems in measure spaces with sparse solutions. *SIAM J. Control Optim.* **50** (2012) 1735–1752.

[8] C. Clason and K. Kunisch, A measure space approach to optimal source placement. *Comput. Optim. Appl.* **53** (2012) 155–171.

[9] K. Pieper and B. Vexler, A priori error analysis for discretization of sparse elliptic optimal control problems in measure space. *SIAM J. Control Optim.* **51** (2013) 2788–2808.

[10] F. Bach. Breaking the curse of dimensionality with convex neural networks. *J. Mach. Learn. Res.* **18** (2017) 53.

[11] J. Kiefer and J. Wolfowitz, Optimum designs in regression problems. *Ann. Math. Stat.* **30** (1959) 271–294.

[12] I. Neitzel, K. Pieper, B. Vexler and D. Walter, A sparse control approach to optimal sensor placement in PDE-constrained parameter estimation problems. *Numer. Math.* **143** (2019) 943–984.

[13] C. Boyer, A. Chambolle, Y. De Castro, V. Duval, F. de Gournay and P. Weiss, On representer theorems and convex regularization. *SIAM J. Optim.* **29** (2019) 1260–1281.

[14] K. Bredies and M. Carioni, Sparsity of solutions for variational inverse problems with finite-dimensional data. *Calc. Var. Part. Differ. Equ.* **59** (2020) Paper No. 14.

[15] K. Pieper and D. Walter, Linear convergence of accelerated conditional gradient algorithms in spaces of measures. *ESAIM Control Optim. Calc. Var.* **27** (2021) Paper No. 38, 37.

[16] V. Duval and G. Peyré, Exact support recovery for sparse spikes deconvolution. *Found. Computat. Math.* **15** (2014) 1315–1355.

[17] L. Chizat, Sparse optimization on measures with over-parameterized gradient descent. *Math. Program.* **194** (2022) 487–532.

[18] A. Flinth, F. de Gournay and P. Weiss, On the linear convergence rates of exchange and continuous methods for total variation minimization. *Math. Program.* **190** (2021) 221–257.

[19] C. Christof and G. Wachsmuth, Differential sensitivity analysis of variational inequalities with locally Lipschitz continuous solution operators. *Appl. Math. Optim.* **81** (2020) 23–62.

[20] R.T. Rockafellar and R.J.B. Wets, Variational Analysis. Springer Berlin Heidelberg (1998).

[21] M. Benko and P. Mehlitz, Why second-order sufficient conditions are, in a way, easy – or – revisiting calculus for second subderivatives. *J. Convex Anal.* **30** (2023) 541–589.

[22] C.N. Do, Generalized second-order derivatives of convex functions in reflexive Banach spaces. *Trans. Am. Math. Soc.* **334** (1992) 281–301.

[23] C. Christof and G. Wachsmuth, No-gap second-order conditions via a directional curvature functional. *SIAM J. Optim.* **28** (2018) 2097–2130.

[24] N. Borchard and G. Wachsmuth, Second-order conditions for spatio-temporally sparse optimal control via second subderivatives **5** (2024) 1–36.

[25] D. Wachsmuth and G. Wachsmuth, Second-order conditions for non-uniformly convex integrands: quadratic growth in L^1. *J. Nonsmooth Anal. Optim.* **3** (2022) 1-36.

[26] P.-T. Huynh, K. Pieper and D. Walter, Towards optimal sensor placement for inverse problems in spaces of measures. *Inverse Probl.* **40** (2024) 1–43.

[27] C. Poon, N. Keriven and G. Peyré, Support localization and the fisher metric for off-the-grid sparse Regularization, in *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, vol. 89 of *Proceedings of Machine Learning Research*, edited by K. Chaudhuri and M. Sugiyama. PMLR (2019), 1341–1350.

[28] C. Poon, N. Keriven and G. Peyré, The geometry of off-the-grid compressed sensing. *Found. Comput. Math.* **23** (2023) 241–327.

[29] D. Leykekhman, B. Vexler and D. Walter, Numerical analysis of sparse initial data identification for parabolic problems. *ESAIM Math. Model. Numer. Anal.* **54** (2020) 1139–1180.

[30] P. Merino, I. Neitzel and F. Tröltzsch, On linear-quadratic elliptic control problems of semi-infinite type. *Appl. Anal.* **90** (2011) 1047–1074.

[31] H.W. Alt, Lineare Funktionalanalysis, 6th edn. Springer, Berlin (2011).

[32] J. Borwein and Q. Zhu, Techniques of Variational Analysis. Springer-Verlag (2005).

[33] P. Hajłasz, P. Koskela and H. Tuominen, Sobolev embeddings, extensions and measure density condition. *J. Funct. Anal.* **254** (2008) 1217–1234.

[34] H. Brezis, Functional Analysis, Sobolev Spaces and Partial Differential Equations. Springer New York (2011).

[35] B. Piccoli, F. Rossi and M. Tournus, A Wasserstein norm for signed measures, with application to non-local transport equation with source term. *Commun. Math. Sci.* **21** (2023) 1279–1301.

[36] J. Frédéric Bonnans and A. Shapiro, Perturbation Analysis of Optimization Problems. Springer, Berlin (2000).

[37] H.H. Bauschke and P.L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, Berlin (2011).

Appendix A. Uniform local quasiconvexity

In this appendix, we give the proofs of the technical lemmas from Section 5.1.

Proof of Lemma 5.3. Since ∂D is compact, we find $n \in \mathbb{N}$, $(p_i)_{i=1}^n \subset \partial D$ and associated radii $(r_i)_{i=1}^n \subset (0,\infty)$ (as in the assumption) such that $\partial D \subset \bigcup_{i=1}^n U_{r_i/2}(p_i)$. The set $D \setminus \bigcup_{i=1}^n U_{r_i/2}(p_i)$ is compact and does not intersect ∂D. Thus

$$r_0 := \inf\left\{ |x-p| \;\middle|\; x \in D \setminus \bigcup_{i=1}^n U_{r_i/2}(p_i),\ p \in \partial D \right\}$$

is positive. We set $r := \min\{r_0, r_1/2, \dots, r_n/2\}$. Moreover, let $L > 0$ be a common Lipschitz constant for l_{p_i} and $l_{p_i}^{-1}$, $i \in \{1,\dots,n\}$. We define $C := L^2$.

We check that D is uniformly locally quasiconvex with parameters r and C. To this end, let $x, y \in D$ with $|x-y| \le r$ be given.

Case 1: $x \in B_{r_i/2}(p_i)$ for some $i \in \{1,\dots,n\}$. Due to $|x-y| \le r \le r_i/2$, we have $x, y \in B_{r_i}(p_i)$. Consequently, $l_{p_i}(x), l_{p_i}(y) \in B_1^+(0)$ can be joined by a line segment in $B_1^+(0)$ of length $|l_{p_i}(x) - l_{p_i}(y)| \le L|x-y|$. Under the map $l_{p_i}^{-1}$, this becomes a Lipschitz curve in D of length at most $L^2|x-y|$ joining x and y, as desired.

Case 2: $x \notin B_{r_i/2}(p_i)$ for all $i \in \{1,\dots,n\}$. In this case, $x \in D \setminus \bigcup_{i=1}^n U_{r_i/2}(p_i)$ and $y \in U_r(x)$. As $U_r(x)$ does not intersect ∂D, we can simply join x and y by a line segment of length $|x-y|$. $\square$

Now, we discuss the consequences of $\operatorname{int}(\Omega)$ being uniformly locally quasiconvex. Let $\gamma \in C([0,1]; \operatorname{int}(\Omega))$ be a Lipschitz continuous curve as in Definition 5.1. Moreover, let $\varphi \in C^1(\Omega)$ be given. Then, $\varphi \circ \gamma$ is absolutely continuous and its derivative at t is given by $\nabla\varphi(\gamma(t))^\top\gamma'(t)$ for almost all $t \in [0,1]$. Consequently,

$$\varphi(\gamma(1)) - \varphi(\gamma(0)) = \int_0^1 \nabla\varphi(\gamma(t))^\top\gamma'(t)\,\mathrm{d}t. \tag{A.1}$$

This identity will be crucial in the next two proofs.

The next proof is similar to a part of the proof of [33], Theorem 7. We recall that the Lipschitz constant $\operatorname{lip}(\cdot)$ was defined in (2.1).

Proof of Lemma 5.4. Let $x, y \in \operatorname{int}(\Omega)$ with $|y - x| > r$ be given. Then,

$$|\varphi(y) - \varphi(x)| \le 2\|\varphi\|_C \le \frac{2}{r}\|\varphi\|_C|y - x|.$$

On the other hand, if $x, y \in \operatorname{int}(\Omega)$ satisfy $|y - x|$, we find $\gamma \in C([0,1];\operatorname{int}(\Omega))$ with $\gamma(0) = x$, $\gamma(1) = y$ and the Lipschitz constant of γ is bounded by $C|y - x|$. Owing to (A.1), we have

$$|\varphi(y) - \varphi(x)| = \left|\int_0^1 \nabla\varphi(\gamma(t))^\top\gamma'(t)\,\mathrm{d}t\right| \le C\|\nabla\varphi\|_{C^d}|y - x|.$$

This shows

$$|\varphi(y) - \varphi(x)| \le \max\left\{\frac{2}{r}\|\varphi\|_C, C\|\nabla\varphi\|_{C^d}\right\}|y - x| \qquad \forall x, y \in \operatorname{int}(\Omega).$$

Since φ is continuous and since Ω is assumed to be the closure of $\operatorname{int}(\Omega)$, a limiting argument shows that the same inequality holds for all $x, y \in \Omega$. □

Next, we address the uniform Taylor expansion.

Proof of Lemma 5.5. Let $\varepsilon > 0$ be arbitrary. Let $r > 0$ and $C \ge 1$ be the constants from Definition 5.1. Since $\nabla\varphi$ is uniformly continuous, there exists $D > 0$ such that

$$|\nabla\varphi(x) - \nabla\varphi(y)| \le \frac{\varepsilon}{C} \qquad \forall x, y \in \Omega, |y - x| \le D.$$

We set $\delta := \min\{D/C, r\}/2$. Now, let $x, y \in \operatorname{int}(\Omega)$ with $|y - x| \le 2\delta$ be given. Since $\operatorname{int}(\Omega)$ is uniformly locally quasiconvex, there exists $\gamma \in C([0,1];\operatorname{int}(\Omega))$ with $\gamma(0) = x$, $\gamma(1) = y$ and Lipschitz constant at most $C|y - x|$. Note that $C|y - x| \le D$. Consequently, the range of γ belongs to $\operatorname{int}(\Omega) \cap B_D(x)$. Using (A.1), we get

$$\left|\varphi(y) - \varphi(x) - \nabla\varphi(x)^\top(y - x)\right| = \left|\int_0^1 (\nabla\varphi(\gamma(t)) - \nabla\varphi(x))^\top\gamma'(t)\,\mathrm{d}t\right| \le \varepsilon|y - x|.$$

This inequality holds for all $x, y \in \operatorname{int}(\Omega)$ with $|y - x| \le 2\delta$. Using a limiting argument yields the claim. □

Finally, we provide a counterexample which shows that the conclusions of Lemmas 5.4 and 5.5 may fail if $\operatorname{int}(\Omega)$ is not uniformly locally quasiconvex.

Example A.1. We define

$$U := \bigcup_{n=1}^\infty \left(\frac{1}{3^n}, \frac{2}{3^n}\right) \subset \mathbb{R} \quad \text{and} \quad \Omega := \operatorname{cl} U = \{0\} \cup \bigcup_{n=1}^\infty \left[\frac{1}{3^n}, \frac{2}{3^n}\right] \subset \mathbb{R}.$$

Note that U is the interior of Ω, *i.e.*, Ω is the closure of its interior, as required. Further, let $c\colon [0,1] \to [0,1]$ be the (continuous) Cantor function and f its restriction to Ω, *i.e.*, $f(0) = 0$ and $f(x) = 1/2^n$ for all $x \in [1/3^n, 2/3^n]$. Obviously, f is continuous.

On each connected subset of U, the function f is constant. Consequently, f is differentiable on U with derivative 0. Since the derivative can be continuously extended to Ω, we have $f \in C^1(\Omega)$.

However, f is not Lipschitz continuous on Ω since

$$\frac{|f(1/3^n) - f(0)|}{|1/3^n - 0|} = \frac{3^n}{2^n} \to \infty \quad \text{as } n \to \infty.$$

Similarly, the Taylor expansion

$$f(y) = f(0) + f'(0)(y - 0) + o(|y|) \qquad \text{as } \Omega \ni y \to 0$$

fails.

Note that Ω is not connected. However, it is quite easy to modify the example to obtain connectedness. We set $\hat{\Omega} := \Omega \times [-1, 0] \cup [0, 1]^2$ which is the closure of $\hat{U} := U \times (-1, 0] \cup (0, 1)^2$. We define the function $\hat{f} \colon \hat{\Omega} \to \mathbb{R}$ via

$$\hat{f}(x) := \begin{cases} 0 & \text{if } x_2 \geq 0, \\ f(x_1) x_2^2 & \text{if } x_2 < 0. \end{cases}$$

It is straightforward to check that $\hat{f} \in C^1(\hat{\Omega})$, but it fails to be Lipschitz continuous and the Taylor expansion fails at $(0, -1) \in \Omega$.