SIAM J. OPTIM.
Vol. 22, No. 3, pp. 953–986

SECOND-ORDER SUBDIFFERENTIAL CALCULUS WITH APPLICATIONS TO TILT STABILITY IN OPTIMIZATION*

B. S. MORDUKHOVICH† AND R. T. ROCKAFELLAR‡

Abstract. This paper concerns the second-order generalized differentiation theory of variational analysis and new applications of this theory to some problems of constrained optimization in finite-dimensional spaces. The main focus is the so-called (full and partial) second-order subdifferentials of extended-real-valued functions, which are dual-type constructions generated by coderivatives of first-order subdifferential mappings. We develop an extended second-order subdifferential calculus and analyze the basic second-order qualification condition ensuring the fulfillment of the principal second-order chain rule for strongly and fully amenable compositions. We also calculate the second-order subdifferentials for some major classes of piecewise linear-quadratic functions. These results are applied to the study of tilt stability of local minimizers for important classes of problems in constrained optimization that include, in particular, problems of nonlinear programming and certain classes of extended nonlinear programs described in composite terms.

Key words. variational analysis, constrained optimization, nonlinear and extended nonlinear programming, second-order subdifferentials, calculus rules, qualification conditions, amenable functions, tilt-stable minimizers, strong regularity

AMS subject classifications. 49J52, 90C30, 90C31

DOI. 10.1137/110852528

1. Introduction.

Variational analysis has been recognized as a fruitful area of mathematics, which primarily deals with optimization-related problems while also applying variational principles and techniques to a broad spectrum of problems that may not be of a variational nature. We refer the reader to the books by Borwein and Zhu [4], Mordukhovich [31, 32], Rockafellar and Wets [49], and the references therein for the major results of variational analysis and its numerous applications.

Nonsmooth functions, sets with nonsmooth boundaries, and set-valued mappings naturally and frequently appear in the framework of variational theory and its applications via using variational principles and techniques, even for problems with smooth initial data. Thus tools of generalized differentiation play a crucial role in many aspects of variational analysis and optimization; see, e.g., the books [4, 8, 12, 19, 31, 32, 49, 50] and the references thererin.

Over the years, the first-order subdifferential theory of variational analysis has been well developed and understood in both finite-dimensional and infinite-dimensional settings; see [4, 31, 49] and the commentaries therein. In contrast, the *second-order* theory still requires much further development and implementation, although many second-order generalized differential constructions have been suggested and successfully applied to various optimization, sensitivity, and related problems; see, e.g., the

*Received by the editors October 10, 2011; accepted for publication (in revised form) May 29, 2012; published electronically August 9, 2012.
http://www.siam.org/journals/siopt/22-3/85252.html

†Department of Mathematics, Wayne State University, Detroit, MI 48202 (boris@math.wayne.edu). The research of this author was partially supported by the National Science Foundation under grant DMS-1007132, by the Australian Research Council under grant DP-12092508, and by the Portuguese Foundation of Science and Technologies under grant MAT/11109.

‡Department of Mathematics, University of Washington, Seattle, WA 98195 (rtr@math.washington.edu).

books [3, 31, 49] summarizing mainstream developments and trends in the second-order theory and its applications.

As is well known, there are two generally independent approaches to second-order differentiation in classical analysis. One of them is based on the Taylor expansion, while the other defines the second derivative of a function as the derivative of its first-order derivative.

In this paper we develop the latter "derivative-of-derivative" approach to the second-order generalized differentiation of extended-real-valued functions $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}} := (-\infty, \infty]$ finite at the reference points. The dual-space route in this vein suggested by Mordukhovich [26] is to treat a (first-order) *subdifferential* $\partial\varphi$ of φ as a set-valued analogue of the classical derivative for nonsmooth functions and then to define a *second-order subdifferential* $\partial^2\varphi$ of φ via a *coderivative* $D^*\partial\varphi$ of the subgradient mapping $\partial\varphi$; see section 2 for more details. This second-order construction was originally motivated by applications to sensitivity analysis of variational systems [26, 29] inspired by the coderivative characterization of Lipschitzian stability [26, 27], but then the second-order subdifferential and its modification were successfully employed in the study of a broad spectrum of other important issues in variational analysis and its applications; see, e.g., [1, 5, 11, 10, 13, 14, 15, 17, 19, 20, 21, 31, 32, 33, 34, 38, 39, 40, 43, 54, 55] and the references therein. We specifically mention a significant result by Poliquin and Rockafellar [43], who established a full characterization of *tilt-stable local minimizers* of functions as the positive-definiteness of their second-order subdifferential mapping. For $\mathcal{C}^2$ functions, the latter criterion reduces to the positive-definiteness of the classical Hessian matrix—a well-known sufficient condition for the standard optimality in unconstrained problems, which happens to be necessary and sufficient for tilt-stable local minimizers [43]. We also refer the reader to the recent papers by Chieu and coworkers [5, 6] providing complete characterizations of *convexity* and *strong convexity* of nonsmooth (in the second order) functions via positive-semidefiniteness and definiteness of their second-order subdifferentials $\partial^2\varphi$. Related characterizations of *monotonicity* and *submonotonicity* of continuous mappings can be found in [7].

Needless to say, efficient implementations and potential extensions of the latter result to constrained optimization problems, as well as any other valuable applications of the aforementioned second-order subdifferential construction and its modifications, largely depend on the possibility to develop a fairly rich *second-order subdifferential calculus* and on precisely *calculating* such constructions for attractive classes of nonsmooth functions overwhelmingly encountered in variational analysis and optimization. A certain amount of useful second-order calculus rules were developed in [20, 28, 30, 31, 33, 34, 36, 39]. On the other hand, precisely calculating the second-order subdifferential entirely in terms of the initial data was effected for the following major classes of extended-real-valued functions particularly important in various applications:

• For indicator functions of *convex polyhedra* and related settings, it was initiated by Dontchev and Rockafellar [11] and then developed in [2, 14, 15, 17, 18, 37, 44, 52] for more involved frameworks. The obtained calculations played a crucial role in deriving [11] verifiable characterizations of the Robinson strong regularity [45] for variational inequalities over (convex) polyhedral sets as well as their specifications for complementarity problems and the associated Karush–Kuhn–Tucker (KKT) conditions for nonlinear programs with $\mathcal{C}^2$ data. Further results in this vein on Lipschitzian stability of parametric variational systems were given in [2, 14, 37, 44, 53] and other publications in both finite and infinite dimensions. Applications to stationarity conditions for stochastic equilibrium problems with equilibrium constraints in electricity spot market modeling were developed by Henrion and Römisch [17].

• For the so-called *separable piecewise* $\mathcal{C}^2$ functions, it was done by Mordukhovich and Outrata [33]; see also [5] for further developments. It provided the basis for the efficient sensitivity analysis [33] of mathematical programs with equilibrium constraints (MPECs) including practical ones that arose in applications to certain contact problems of continuum mechanics. Similar calculations were applied in [1] for the qualitative stability analysis of nonmonotone variational inclusions with applications to practical models in electronics.

• For indicator functions to *smooth nonpolyhedral inequality systems*, it was done by Henrion, Outrata, and Surowiec [15] by employing and developing the transformation formulas from [33]. Then these calculations were applied in [16, 51] to deriving stationarity conditions for equilibrium problems with equilibrium constraints (EPECs), in both deterministic and stochastic frameworks, and to EPEC models of oligopolistic competition in electricity spot markets.

• For classes of functions associated with metric projection onto the *second-order/Lorentz cone*, it was done by Outrata and Sun [40]. The obtained calculations were applied in [40] to the study of Lipschitzian stability of parameterized second-order cone complementarity problems and to deriving optimality conditions for mathematical programs with second-order cone complementarity constraints. These results and calculations were further developed by Outrata and Ramírez [39] to completely characterize the Aubin/Lipschitz-like property of canonically perturbed KKT systems related to the *second-order cone programs* and to establish the equivalence of the latter property to the Robinson strong regularity in second-order cone programming.

• For the indicator function of the convex cone of $n \times n$ positive, symmetric, and *semidefinite matrices*, it was done by Ding, Sun, and Ye in [10]. The results obtained were applied in [10] to deriving various stationarity conditions for a class of mathematical programs with *semidefinite cone complementarity constraints*, which can be treated as a matrix counterpart of MPECs.

• For a particular class of functions arising in optimal control of the Moreau *sweeping process*, is was done in the paper by Colombo et al. [9]. These calculations played a significant role in deriving constructive optimality conditions for discontinuous differential inclusions generated by the sweeping process and have great potential for further applications.

Now we briefly describe the main *goals and achievements* of this paper. Our primary attention is focused on the following major issues new in second-order variational analysis:

• Developing refined *second-order chain rules* of the equality and inclusion (outer/upper estimate) types for the aforementioned second-order subdifferential and its partial modifications.

• Analyzing the basic *second-order qualification condition* ensuring the fulfillment of the extended second-order chain rules for *strongly amenable* compositions.

• *Calculating* precisely the second-order subdifferentials for some major classes of *fully amenable* compositions.

• Applying the obtained calculus and computational results to deriving necessary optimality conditions as well as to establishing *complete characterizations* of *tilt-stable minimizers* for broad classes of constrained optimization problems, including those in nonlinear programming (NLP) and extended nonlinear programming (ENLP) described via amenable compositions.

The rest of the paper is organized as follows. section 2 contains basic definitions and brief discussions of the first-order and second-order generalized differential constructions studied and used in the paper. We also review there some preliminary results widely employed in what follows.

In section 3 we deal with *second-order chain rules* of the equality and inclusion types for the basic second-order subdifferential and its partial counterparts. The equality-type results are established under the *full rank condition* on the Jacobian matrix of the inner mapping of the composition. Without imposing the latter assumption, we develop a new *quadratic penalty* approach that allows us to derive inclusion-type second-order chain rules for a broad class of strongly amenable compositions valid under certain second-order qualification conditions. The latter chain rules generally provide merely outer estimates of the second-order subgradient sets for compositions: we present an example showing that the chain rule inclusion may be strict even for linear inner mappings and piecewise linear outer functions in fully amenable compositions.

Section 4 is devoted to deriving *exact* (equality-type) second-order chain rules for major classes of *fully amenable* compositions without imposing the full rank assumption. Our approach is mainly based on precisely *calculating* the second-order subdifferential for *piecewise linear quadratic* functions (which is important for its own sake) and on a detailed analysis of the basic *second-order qualification condition* developed in section 3. In particular, we show that this condition reduces locally to the full rank requirement on the inner mapping Jacobian matrix if the outer function is *piecewise linear*.

The concluding section 5 concerns applications of the chain rules and calculation results developed in the previous sections to the study of *tilt-stable local minimizers* for some classes of constrained optimization problems represented in composite formats, which are convenient for developing both theoretical and computational aspects of optimization. Such classes include, besides standard nonlinear programs broader models of the so-called extended nonlinear programming (ENLP). Based on the second-order sum and chain rules with equalities, we derive *complete characterizations* of tilt-stable local minimizers for important problems of constrained optimization. The results obtained show, in particular, that for a general class of NLP problems, the well-recognized *strong second-order optimality condition* is *necessary and sufficient* for tilt stability of local minimizers, which therefore is equivalent to the *Robinson strong regularity* of the associated variational inequalities in such settings. Furthermore, the calculus rules derived in this paper for *partial* second-order subdifferentials lead us also to characterizations of *full stability* in optimization (see Remark 5.6), while a detailed elaboration of this approach is a subject of our ongoing research.

Although a number of the results obtained in this paper hold in (or can be naturally extended to) infinite-dimensional spaces, for definiteness we confine ourselves to the finite-dimensional setting. Throughout the paper we use standard notation of variational analysis; cf. [31, 49]. Recall that, given a set-valued mapping $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, the symbol

$$\operatorname*{Lim\,sup}_{x\to\bar{x}} F(x) := \Big\{ y \in \mathbb{R}^m \Big| \exists\, x_k \to \bar{x},\ \exists\, y_k \to y \ \text{ as } \ k \to \infty \ \text{ with } \ y_k \in F(x_k) \ \text{ for all } \ k \in I\!N := \{1, 2, \ldots\} \Big\} \tag{1.1}$$

signifies the *Painlevé–Kuratowski outer/upper limit* of F as $x \to \bar{x}$. Given a set $\Omega \subset \mathbb{R}^n$ and an extended-real-valued function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ finite at $\bar{x}$, the symbols $x \stackrel{\Omega}{\to} \bar{x}$ and

$x \xrightarrow{\varphi} \bar{x}$ stand for $x \to \bar{x}$ with $x \in \Omega$ and for $x \to \bar{x}$ with $\varphi(x) \to \varphi(\bar{x})$, respectively. As usual, $I\!B(x, r)$ denotes the closed ball of the space in question centered at x with radius $r > 0$.

2. Basic definitions and preliminaries. In this section we define and briefly discuss the basic generalized differential constructions of our study and review some preliminaries widely used in what follows; see [31, 49] for more details.

Let $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be an extended-real-valued function finite at $\bar{x}$. The *regular subdifferential* (known also as the presubdifferential and as the Fréchet or viscosity subdifferential) of φ at $\bar{x}$ is

$$\widehat{\partial}\varphi(\bar{x}) := \Big\{ v \in \mathbb{R}^n \Big| \ \liminf_{x \to \bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} \geq 0 \Big\}. \tag{2.1}$$

Each $v \in \widehat{\partial}\varphi(\bar{x})$ is a *regular subgradient* of φ at $\bar{x}$. While $\widehat{\partial}\varphi(\bar{x})$ reduces to a singleton $\{\nabla\varphi(\bar{x})\}$ if φ is Fréchet differentiable at $\bar{x}$ with the gradient $\nabla\varphi(\bar{x})$, and reduces to the classical subdifferential of convex analysis if φ is convex, the set (2.1) may often be empty for nonconvex and nonsmooth functions such as, e.g., for $\varphi(x) = -|x|$ at $\bar{x} = 0 \in \mathbb{R}$. Another serious disadvantage of the subdifferential construction (2.1) is the failure of standard calculus rules inevitably required in the theory and applications of variational analysis and optimization. In particular, the inclusion (outer estimate) sum rule $\widehat{\partial}(\varphi_1 + \varphi_2)(\bar{x}) \subset \widehat{\partial}\varphi_1(\bar{x}) + \widehat{\partial}\varphi_2(\bar{x})$ does not hold for the simplest nonsmooth functions $\varphi_1(x) = |x|$ and $\varphi_2(x) = -|x|$ at $\bar{x} = 0 \in \mathbb{R}$.

The picture dramatically changes when we employ a limiting "robust regularization" procedure over the subgradient mapping $\widehat{\partial}\varphi(\cdot)$ that leads us to the (basic first-order) *subdifferential* of φ at $\bar{x}$ defined by

$$\partial\varphi(\bar{x}) := \operatorname*{Lim\,sup}_{x \xrightarrow{\varphi} \bar{x}} \widehat{\partial}\varphi(x) \tag{2.2}$$

and known also as the general, or limiting, or Mordukhovich subdifferential; it was first introduced in [24] in an equivalent way. Each $v \in \partial\varphi(\bar{x})$ is called a (basic) *subgradient* of φ at $\bar{x}$. Thus, by taking into account definition (1.1) of Lim sup and the notation $x \xrightarrow{\varphi} \bar{x}$, we represent the basic subgradients $v \in \partial\varphi(\bar{x})$ as follows:

There are sequences $x_k \to \bar{x}$ with $\varphi(x_k) \to \varphi(\bar{x})$ and $v_k \in \widehat{\partial}\varphi(x_k)$ with $v_k \to v$.

In contrast to (2.1), the subgradient set (2.2) is generally nonconvex (e.g., $\partial\varphi(0) = \{-1, 1\}$ for $\varphi(x) = -|x|$) while enjoying comprehensive calculus rules ("full calculus"); this is based on *variational/extremal principles*, which replace separation arguments in the absence of convexity. Moreover, the basic subdifferential (2.2) is the *smallest* among any axiomatically defined subgradient sets satisfying certain natural requirements; see [35, Theorem 9.7].

In what follows, we also need another subdifferential construction effective for non-Lipschitzian extended-real-valued functions. Given $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ finite at $\bar{x}$, the *singular/horizon subdifferential* $\partial^\infty\varphi(\bar{x})$ of φ at $\bar{x}$ is defined by

$$\partial^\infty\varphi(\bar{x}) := \operatorname*{Lim\,sup}_{\substack{x \xrightarrow{\varphi} \bar{x} \\ \lambda \downarrow 0}} \lambda\widehat{\partial}\varphi(x). \tag{2.3}$$

If the function φ is lower semicontinuous (l.s.c.) around $\bar{x}$, then $\partial^\infty\varphi(\bar{x}) = \{0\}$ if and only if φ is locally Lipschitzian around this point.

Given further a nonempty subset $\Omega \subset \mathbb{R}^n$, consider its indicator function $\delta(x;\Omega)$ equal to 0 for $x \in \Omega$ and to ∞ otherwise. For any fixed $\bar{x} \in \Omega$, define the *regular normal cone* to Ω at $\bar{x}$ by

$$\widehat{N}(\bar{x};\Omega) := \widehat{\partial}\delta(\bar{x};\Omega) = \Big\{ v \in \mathbb{R}^n \Big| \ \operatorname*{Lim\,sup}_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle v, x - \bar{x}\rangle}{\|x - \bar{x}\|} \le 0 \Big\} \tag{2.4}$$

and similarly the (basic, limiting) *normal cone* to Ω at $\bar{x}$ by $N(\bar{x};\Omega) := \partial\delta(\bar{x};\Omega)$. It follows from (2.2) and (2.4) that the normal cone $N(\bar{x};\Omega)$ admits the limiting representation

$$N(\bar{x};\Omega) = \operatorname*{Lim\,sup}_{x \xrightarrow{\Omega} \bar{x}} \widehat{N}(x;\Omega), \tag{2.5}$$

meaning that the basic normals $v \in N(\bar{x};\Omega)$ are those vectors $v \in \mathbb{R}^n$ for which there are sequences $x_k \to \bar{x}$ and $v_k \to v$ with $x_k \in \Omega$ and $v_k \in \widehat{N}(x_k;\Omega)$, $k \in I\!N$. If Ω is locally closed around $\bar{x}$, (2.5) is equivalent to the original definition by Mordukhovich [24]:

$$N(\bar{x};\Omega) = \operatorname*{Lim\,sup}_{x \to \bar{x}} \Big[\operatorname{cone}\big(x - \Pi(x;\Omega)\big)\Big],$$

where $\Pi(x;\Omega)$ signifies the Euclidean projector of $x \in \mathbb{R}^n$ on the set Ω, and where "cone" stands for the conic hull of a set.

There is the duality/polarity correspondence

$$\widehat{N}(\bar{x};\Omega) = T(\bar{x};\Omega)^* := \Big\{ v \in \mathbb{R}^n \Big| \ \langle v, w\rangle \le 0 \ \text{ for all } \ w \in T(\bar{x};\Omega) \Big\} \tag{2.6}$$

between the regular normal cone (2.4) and the *tangent cone* to Ω at $\bar{x} \in \Omega$ defined by

$$T(\bar{x};\Omega) := \Big\{ w \in \mathbb{R}^n \Big| \ \exists\, x_k \xrightarrow{\Omega} \bar{x},\ \alpha_k \ge 0 \ \text{ with } \ \alpha_k(x_k - \bar{x}) \to w \ \text{ as } \ k \to \infty \Big\} \tag{2.7}$$

and known also as the Bouligand–Severi contingent cone to Ω at this point. Note that the basic normal cone (2.5) cannot be tangentially generated in a polar form (2.6) by using some set of tangents, since it is intrinsically nonconvex, while the polar T^* to any set T is automatically convex. In what follows, we may also use the subindex set notation like $N_\Omega(\bar{x})$, $T_\Omega(\bar{x})$, etc. for the constructions involved.

It is worth observing that the *convex closure*

$$\overline{N}(\bar{x};\Omega) := \operatorname{clco} N(\bar{x};\Omega) \tag{2.8}$$

of (2.5), known as the Clarke/convexified normal cone to Ω at $\bar{x}$ (see [8]), may dramatically enlarge the set of basic normals (2.5). Indeed, it is proved by Rockafellar [47] that for every vector function $f\colon \mathbb{R}^n \to \mathbb{R}^m$ locally Lipschitzian around $\bar{x}$ the convexified normal cone (2.8) to the graph of f at $(\bar{x}, f(\bar{x}))$ is in fact a *linear subspace* of dimension $d \ge m$ in $\mathbb{R}^n \times \mathbb{R}^m$, where the equality $d = m$ holds if and only if the function f is strictly differentiable at $\bar{x}$ with the derivative (Jacobian matrix) denoted for simplicity by $\nabla f(\bar{x})$, i.e.,

$$\lim_{x,u \to \bar{x}} \frac{f(x) - f(u) - \nabla f(\bar{x})(x - u)}{\|x - u\|} = 0,$$

which is automatic when f is $\mathcal{C}^1$ around $\bar{x}$. In particular, this implies that $\overline{N}((\bar{x}, f(\bar{x}));\operatorname{gph} f)$ is the whole space $\mathbb{R}^n \times \mathbb{R}^m$ whenever f is nonsmooth around $\bar{x}$, $n = 1$, and $m \geq 1$. Moreover, the aforementioned results were discovered by Rockafellar [47] not only for graphs of locally Lipschitzian functions, but also for the so-called *Lipschitzian manifolds* (or graphically Lipschitzian sets), which are locally homeomorphic to graphs of Lipschitzian vector functions. The latter class includes graphs of maximal monotone relations and subdifferential mappings for convex, saddle, lower-$\mathcal{C}^2$, and more general prox-regular functions typically encountered in variational analysis and optimization. In fact such *graphical sets* play a crucial role in the coderivative and second-order subdifferential constructions studied in this paper.

Given a set-valued mapping $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, define its *coderivative* at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ by [25]

$$D^*F(\bar{x}, \bar{y})(v) := \Big\{u \in \mathbb{R}^n \Big|\, (u, -v) \in N\big((\bar{x}, \bar{y}); \operatorname{gph} F\big)\Big\}, \quad v \in \mathbb{R}^m, \tag{2.9}$$

via the normal cone (2.5) to the graph $\operatorname{gph} F$. Clearly the mapping $D^*F(\bar{x}, \bar{y})\colon \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is positive-homogeneous; it reduces to the adjoint derivative

$$D^*F(\bar{x})(v) = \big\{\nabla F(\bar{x})^* v\big\}, \quad v \in \mathbb{R}^m, \tag{2.10}$$

where * stands for the matrix transposition, if F is single-valued (then we omit $\bar{y} = F(\bar{x})$ in the coderivative notation) and strictly differentiable at $\bar{x}$. Note that the coderivative values in (2.9) are often nonconvex sets due to the nonconvexity of the normal cone on the right-hand side. Furthermore, the latter cone is taken to a graphical set, and thus its convexification in (2.9) may create serious trouble; see above.

The main construction studied in the paper was introduced in [26] as follows.

Definition 2.1 (second-order subdifferential). *Let the function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite at $\bar{x}$, and let $\bar{y} \in \partial\varphi(\bar{x})$ be a basic first-order subgradient of φ at $\bar{x}$. Then the* SECOND-ORDER SUBDIFFERENTIAL *of φ at $\bar{x}$ relative to $\bar{y}$ is defined by*

$$\partial^2\varphi(\bar{x}, \bar{y})(u) := (D^*\partial\varphi)(\bar{x}, \bar{y})(u), \quad u \in \mathbb{R}^n, \tag{2.11}$$

via the coderivative (2.9) *of the first-order subdifferential mapping* (2.2).

Observe that if $\varphi \in \mathcal{C}^2$ around $\bar{x}$ (in fact, it is merely continuously differentiable around $\bar{x}$ with the strict differentiable first-order derivative at this point), then

$$\partial^2\varphi(\bar{x})(u) = \big\{\nabla^2\varphi(\bar{x})u\big\}, \quad u \in \mathbb{R}^n,$$

where $\nabla^2\varphi(\bar{x})$ is the (symmetric) Hessian of φ at $\bar{x}$. Sometimes the second-order construction (2.11) is called the "generalized Hessian" of φ at the reference point [43]. Note also that for the so-called $\mathcal{C}^{1,1}$ *functions* (i.e., continuously differentiable ones with locally Lipschitzian derivatives around $\bar{x}$; another notation is $\mathcal{C}^{1+}$), we have the representation

$$\partial^2\varphi(\bar{x})(u) = \partial\langle u, \nabla\varphi\rangle(\bar{x}), \quad u \in \mathbb{R}^n,$$

via the first-order subdifferential (2.2) of the derivative scalarization $\langle u, \nabla\varphi\rangle(x) := \langle u, \nabla\varphi(x)\rangle$ as $x \in \mathbb{R}^n$; see [31, Proposition 1.120]. It is worth emphasizing that the second-order subdifferential (2.11) as well as the generating coderivative and first-order subdifferential mappings are *dual-space* intrinsically nonconvex constructions, which cannot correspond by duality to any derivative-like objects in primal spaces studied, e.g., in [3, 49].

Following the scheme of Definition 2.1 and keeping the coderivative (2.9) as the underlying element of our approach while using different first-order subdifferentials in (2.11), we may define a variety of second-order constructions of type (2.11). In particular, for functions $\varphi\colon \mathbb{R}^n\times\mathbb{R}^d\to\overline{\mathbb{R}}$ of $(x,w)\in\mathbb{R}^n\times\mathbb{R}^d$ there are two reasonable ways of introducing partial second-order subdifferentials; cf. [21]. To proceed, define the *partial first-order* subgradient mapping $\partial_x\varphi\colon \mathbb{R}^n\times\mathbb{R}^d\rightrightarrows\mathbb{R}^n$ by

$$\partial_x\varphi(x,w):=\Big\{\text{set of subgradients } v \text{ of } \varphi_w:=\varphi(\cdot,w) \text{ at } x\Big\}=\partial\varphi_w(x).$$

Then given $(\bar x,\bar w)$ and $\bar y\in\partial_x\varphi(\bar x,\bar w)$, define the *partial second-order subdifferential* of φ with respect to x at $(\bar x,\bar w)$ relative to $\bar y$ by

$$\partial_x^2\varphi(\bar x,\bar w,\bar y)(u):=(D^*\partial\varphi_{\bar w})(\bar x,\bar y)(u)=\partial^2\varphi_{\bar w}(\bar x,\bar y)(u),\quad u\in\mathbb{R}^n, \tag{2.12}$$

with $\varphi_{\bar w}(x)=\varphi(x,\bar w)$. On the other hand, we can define the *extended partial second-order subdifferential* of φ with respect to x at $(\bar x,\bar w)$ relative to $\bar y$ by

$$\widetilde{\partial}_x^2\varphi(\bar x,\bar w,\bar y)(u):=(D^*\partial_x\varphi)(\bar x,\bar w,\bar y)(u),\quad u\in\mathbb{R}^n. \tag{2.13}$$

As argued in [21], constructions (2.12) and (2.13) are not the same even in the case of $\mathcal{C}^2$ functions when (2.12) reduces to $\nabla^2_{xx}\varphi(\bar x,\bar w)(u)$, while (2.13) comes out as $(\nabla^2_{xx}\varphi(\bar x,\bar w)u,\nabla^2_{xw}\varphi(\bar x,\bar w)u)$. This happens due to the involvement of $w\to\bar w$ in the limiting procedure to define the extended partial second-order subdifferential set $\widetilde{\partial}_x^2\varphi(\bar x,\bar w,\bar y)(u)$, which is hence larger than (2.12). Note that both partial second-order constructions (2.12) and (2.13) are proved to be useful in applications; see, e.g., [20, 21] for more details.

It has been well recognized and documented (see, e.g., [4, 31, 32, 49, 50] and the references therein) that the first-order limiting constructions (2.2), (2.5), and (2.9) enjoy full calculi, which are crucial for their numerous applications. Based on definitions (2.11) of the second-order subdifferential and its partial counterparts (2.12) and (2.13), it is natural to try to combine calculus results for first-order subgradients with those for coderivatives to arrive at the corresponding second-order calculus rules. However, there are nontrivial complications when proceeding in this way due to the fact that general results of the first-order subdifferential calculus hold as *inclusions* while the coderivative operation (2.9) does not generally preserve them. Thus the initial requirement arises on selecting classes of functions for which calculus rules for first-order subgradients hold as *equalities*. Proceeding in this direction, a number of second-order calculus rules have been established in [20, 28, 30, 31, 33, 34, 36, 39] for full, while *not for partial*, second-order subdifferentials in finite and infinite dimensions.

In the next section we obtain new second-order chain rules applied to full and partial second-order subdifferentials and develop, in particular, a direct approach based on quadratic penalties to derive general results for strongly amenable compositions.

3. General second-order subdifferential chain rules. Given a vector function $h\colon \mathbb{R}^n\times\mathbb{R}^d\to\mathbb{R}^m$ with $m\le n$ and a proper extended-real-valued function $\theta\colon\mathbb{R}^m\to\overline{\mathbb{R}}$, consider the composition

$$\varphi(x,w)=(\theta\circ h)(x,w):=\theta\big(h(x,w)\big) \tag{3.1}$$

with $x \in \mathbb{R}^n$ and $w \in \mathbb{R}^d$. Our first theorem provides *exact formulas* for calculating the partial second-order subdifferentials (2.12) and (2.13) of composition (3.1) under the *full rank condition* on the partial derivative (Jacobian matrix) $\nabla_x h(\bar{x}, \bar{w})$ at the reference point.

THEOREM 3.1 (exact second-order chain rules with full rank condition). *Given a point $(\bar{x}, \bar{w}) \in \mathbb{R}^n \times \mathbb{R}^d$, suppose that θ in (3.1) is finite at $\bar{z} := h(\bar{x}, \bar{w})$, that $h(\cdot, \bar{w})\colon \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable around $\bar{x}$ with the full row rank condition*

$$\operatorname{rank} \nabla_x h(\bar{x}, \bar{w}) = m, \tag{3.2}$$

and that the mapping $\nabla_x h(\cdot, \bar{w})\colon \mathbb{R}^n \to \mathbb{R}^m$ is strictly differentiable at $\bar{x}$. Pick any $\bar{y} \in \partial_x \varphi(\bar{x}, \bar{w})$ and denote by $\bar{v}$ a unique vector satisfying the relationships

$$\bar{v} \in \partial\theta(\bar{z}) \quad \textit{and} \quad \nabla_x h(\bar{x}, \bar{w})^* \bar{v} = \bar{y}.$$

Then we have the chain rule equality for the partial second-order subdifferential (2.12)*:*

$$\partial_x^2 \varphi(\bar{x}, \bar{w}, \bar{y})(u) = \nabla_{xx}^2 \langle \bar{v}, h \rangle(\bar{x}, \bar{w})u + \nabla_x h(\bar{x}, \bar{w})^* \partial^2\theta(\bar{z}, \bar{v})(\nabla_x h(\bar{x}, \bar{w})u), \quad u \in \mathbb{R}^n. \tag{3.3}$$

If in addition the mapping $h\colon \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^m$ is continuously differentiable around $(\bar{x}, \bar{w})$ and its derivative $\nabla h\colon \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^m$ is strictly differentiable at $(\bar{x}, \bar{w})$, then we have

$$\begin{aligned} \widetilde{\partial}_x^2 \varphi(\bar{x}, \bar{w}, \bar{y})(u) = &\Big(\nabla_{xx}^2 \langle \bar{v}, h \rangle(\bar{x}, \bar{w})u, \nabla_{xw}^2 \langle \bar{v}, h \rangle(\bar{x}, \bar{w})u\Big) \\ &+ \Big(\nabla_x h(\bar{x}, \bar{w}), \nabla_w h(\bar{x}, \bar{w})\Big)^* \partial^2\theta(\bar{z}, \bar{v})(\nabla_x h(\bar{x}, \bar{w})u) \end{aligned} \tag{3.4}$$

whenever $u \in \mathbb{R}^n$ for the extended partial second-order subdifferential (2.13).

Proof. We derive the chain rule (3.4) for the extended partial second-order subdifferential; the proof of (3.3) is just a simplification of the one given below.

On the first-order subdifferential level we have from [31, Proposition 1.112] under the assumptions made (and from [49, Exercise 10.7] under some additional assumptions) that there is a neighborhood U of $(\bar{x}, \bar{w})$ such that

$$\partial_x \varphi(x, w) = \Big\{ y \in \mathbb{R}^n \,\Big|\, \exists\, v \in \partial\theta\big(h(x, w)\big) \ \text{ with } \ \nabla_x h(x, w)^* v = y \Big\} \ \text{ for all } \ (x, w) \in U.$$

For any fixed $\bar{y} \in \partial_x \varphi(\bar{x}, \bar{w})$, this gives us locally around $(\bar{x}, \bar{w}, \bar{y})$ the graph representation

$$\begin{aligned} \operatorname{gph} \partial_x \varphi = \Big\{ (x, w, y) \in \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n \,\Big|\, &\exists\, (p, v) \in \operatorname{gph} \partial\theta \ \text{ such that} \\ &h(x, w) = p,\ \nabla_x h(x, w)^* v = y \Big\}. \end{aligned} \tag{3.5}$$

Consider now the following two possible cases in the graph representation (3.5): (i) the "square" case when $m = n$, and (ii) the "general" case when $m < n$.

In the square case (i) we have by the full rank condition (3.2) that the matrix $\nabla_x h(x, w)$ is invertible for (x, w) near $(\bar{x}, \bar{w})$, and hence (3.5) can be rewritten as

$$\operatorname{gph} \partial_x \varphi = \Big\{ (x, w, y) \in \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n \,\Big|\, f(x, w, y) \in \operatorname{gph} \partial\theta \Big\} \tag{3.6}$$

via the mapping $f\colon \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^{2n}$ given by

$$(3.7)\qquad f(x,w,y) := \Big(h(x,w), (\nabla_x h(x,w)^*)^{-1} y\Big) \ \text{ for } \ (x,w,y) \ \text{ near } \ (\bar{x},\bar{w},\bar{y}).$$

In other words, representation (3.6) can be expressed via the preimage/inverse image of the set $\operatorname{gph} \partial\theta$ under the mapping f as follows:

$$(3.8)\qquad \operatorname{gph} \partial_x \varphi = f^{-1}\big(\operatorname{gph} \partial\theta\big).$$

Since $\nabla_x h$ is assumed to be strictly differentiable at $(\bar{x},\bar{w})$, the mapping f in (3.7) is strictly differentiable at $(\bar{x},\bar{w},\bar{y})$ and, by (3.2) with $m=n$, its Jacobian matrix $\nabla f(\bar{x},\bar{w},\bar{y})$ has full row rank $2n$. Employing [31, Theorem 1.17] to (3.8) gives us

$$(3.9)\qquad N\big((\bar{x},\bar{w},\bar{y}); \operatorname{gph} \partial_x\varphi\big) = \nabla f(\bar{x},\bar{w},\bar{y})^* N\big(f(\bar{x},\bar{w},\bar{y}); \operatorname{gph} \partial\theta\big).$$

Now we calculate the derivative/Jacobian matrix of f at $(\bar{x},\bar{w},\bar{y})$ by using the particular structure of f in (3.7), the classical chain rule, and the well-known Leach inverse function theorem for strictly differentiable mappings; see, e.g., [12, 31]. Define the mappings $f_1\colon \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^n$ and $f_2\colon \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^n$ with $f=(f_1,f_2)$ by $f_1(x,w,y) := h(x,w)$ and

$$f_2(x,w,y) := (\nabla_x h(x,w)^{-1})^* y \ \text{ for } \ (x,w,y) \in \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n.$$

It is clear that $\nabla f_1(\bar{x},\bar{w},\bar{y}) = (\nabla h(\bar{x},\bar{w}),0)$, while for calculating $\nabla f_2(\bar{x},\bar{w},\bar{y})$ we introduce two auxiliary mappings $g\colon \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^n$ and $q\colon \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}$ by

$$g(x,w,p) := \nabla_x h(x,w)^* p \ \text{ and } \ q(x,w,p) := \langle p, h(x,w)\rangle \ \text{ for } \ (x,w,p) \in \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n.$$

Note that $g(x,w,p) = \nabla q_x(x,w,p)^*$ and that $g(x,w,f_2(x,w,y)) - y = 0$. Differentiating the latter equality gives us

$$(3.10)\qquad \nabla_x g(x,w,f_2(x,w,y)) + \nabla_p g(x,w,f_2(x,w,y)) \nabla_x f_2(x,w,y) = 0.$$

Observing further that $\nabla_x g(x,w,p) = \nabla_x(\nabla_x q(x,w,p))^* = \nabla^2_{xx} q(x,w,p)$, we get from (3.10) and the definitions above that the equation

$$\nabla^2_{xx} \langle (\nabla_x h(x,w)^{-1})^* y, h(x,w)\rangle + \nabla_x h(x,w)^* \nabla f_2(x,w,y) = 0$$

is satisfied, which implies in turn the representation of the partial derivative

$$(3.11)\qquad \nabla_x f_2(x,w,y) = -(\nabla_x h(x,w)^{-1})^* \cdot \nabla^2_{xx} \langle (\nabla_x h(x,w)^{-1})^* y, h(x,w)\rangle.$$

Similarly we have the following representation of the other partial derivative of f_2:

$$(3.12)\qquad \nabla_w f_2(x,w,y) = -(\nabla_x h(x,w)^{-1})^* \cdot \nabla^2_{xw} \langle (\nabla_x h(x,w)^{-1})^* y, h(x,w)\rangle.$$

Taking into account that $\nabla_y f_2(x,w,y) = (\nabla_x h(x,w)^{-1})^*$ gives us finally

$$(3.13)\qquad \nabla f(\bar{x},\bar{w},\bar{y}) = \begin{bmatrix} \nabla_x h(\bar{x},\bar{w}) & \nabla_w h(\bar{x},\bar{w}) & 0 \\ \nabla_x f_2(\bar{x},\bar{w},\bar{y}) & \nabla_w f_2(\bar{x},\bar{w},\bar{y}) & (\nabla_x h(\bar{x},\bar{w})^{-1})^* \end{bmatrix},$$

where $\nabla_x f_2(\bar{x},\bar{w},\bar{y})$ and $\nabla_w f_2(\bar{x},\bar{w},\bar{y})$ are as calculated in (3.11) and (3.12), respectively. The second-order chain rule (3.4) in the square case (i) follows now from substituting (3.13) into (3.9) and then by using the definitions of the constructions involved and elementary transformations.

It remains to consider the general case (ii) with $m < n$. This case can be reduced to the previous one by introducing a linear mapping $\widetilde{h}\colon \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^{n-m}$ such that the mapping

$$\overline{h}(x,w) := \big(h(x,w), \widetilde{h}(x,w)\big) \text{ from } \mathbb{R}^n \times \mathbb{R}^d \text{ to } \mathbb{R}^n$$

has full rank. It can be done, e.g., by choosing a basis $\{a_1, \ldots, a_{n-m}\}$ for the $(n-m)$-dimensional spaces $\{u \in \mathbb{R}^n |\ \nabla h_x(\bar{x}, \bar{w})u = 0\}$ and letting $\overline{h}(x,w) := (h(x,w), \langle a_1, x\rangle, \ldots, \langle a_{n-m}, x\rangle)$; cf. [49, Exercise 6.7] for a first-order setting. Then viewing φ as $\overline{\theta} \circ \overline{h}$ with $\overline{\theta}(z,p) := \theta(z)$ for all $z \in \mathbb{R}^m$ and $p \in \mathbb{R}^{n-m}$ reduces (ii) to (i) and thus completes the proof of the theorem. □

Some remarks on the results related to those obtained in Theorem 3.1 are in order.

Remark 3.2 (discussions on second-order chain rules with full rank/surjectivity conditions). Previously known second-order chain rules of type (3.3) were derived for the full second-order subdifferential (2.11), where condition (3.2) was written as $\operatorname{rank}\nabla h(\bar{x}) = m$. To the best of our knowledge, the first result in this direction was obtained in [33, Theorem 3.4] with the inclusion "$\subset$" in (3.3). Various infinite-dimensional extensions of (3.3) in the inclusion and equality forms were derived in [30, 36] and [31, Theorem 1.127] by imposing the *surjectivity condition* on the derivative $\nabla h(\bar{x})$ as the counterpart of (3.2) in infinite-dimensional spaces. The most recent second-order subdifferential chain rule for (2.11) was derived in [39, Theorem 7] in the framework of cone programming with $\mathcal{C}^2$-reducible constraints [3] under a certain *nondegeneracy* qualification condition. Observe finally that the proof of equality (3.4) given above in case (i) corresponding to the invertible partial derivative $\nabla_x h(\bar{x}, \bar{w})$ holds in any Banach space, while the device in case (ii) is finite-dimensional.

Next we explore the possibility of deriving second-order chain rules for (3.1) when the rank condition (3.2) may not be satisfied. This can be done for broad classes of amenable functions defined in the way originated in [41], which are overwhelmingly encountered in finite-dimensional parametric optimization. Recall [22] that a proper function $\varphi\colon \mathbb{R}^n \times \mathbb{R}^d \to \overline{\mathbb{R}}$ is *strongly amenable* in x at $\bar{x}$ with *compatible parameterization* in w at $\bar{w}$ if there is a neighborhood V of $(\bar{x}, \bar{w})$ on which φ is represented in the composition form (3.1), where h is of class $\mathcal{C}^2$, while θ is a proper, l.s.c., convex function such that the first-order qualification condition

$$\partial^\infty \theta\big(h(\bar{x}, \bar{w})\big) \cap \ker \nabla_x h(\bar{x}, \bar{w})^* = \{0\} \tag{3.14}$$

involving the singular subdifferential (2.3) is satisfied. The latter qualification condition automatically holds if either θ is locally Lipschitzian around $h(\bar{x}, \bar{w})$ or the full rank condition (3.2) is fulfilled, since it is equivalent to

$$\ker \nabla_x h(\bar{x}, \bar{w})^* := \big\{v \in \mathbb{R}^n \big|\ 0 = \nabla_x h(\bar{x}, \bar{w})^* v\big\} = \{0\}.$$

Properties of strongly amenable compositions $\varphi(x) = \theta(h(x))$ and related functions are largely investigated in [41, 42, 49]; most of them hold also for strongly amenable compositions (3.1) with compatible parameterization [21, 22]. Strong amenability is a property that bridges the gap between smoothness and convexity while at the same time covering a great many of the functions that are of interest as the essential objective in minimization problems; see [49] for more details.

The next theorem establishes second-order subdifferential chain rules of the inclusion type for strongly amenable compositions with no full rank requirement (3.2).

THEOREM 3.3 (second-order chain rules for strongly amenable compositions). *Let* $\varphi\colon \mathbb{R}^n \times \mathbb{R}^d \to \overline{\mathbb{R}}$ *be strongly amenable in* x *at* $\bar{x}$ *with compatible parameterization in* w *at* $\bar{w}$*, and let* $\bar{y} \in \partial_x\varphi(\bar{x}, \bar{w})$*. Denote* $\bar{z} := h(\bar{x}, \bar{w})$ *and consider the nonempty set*

$$M(\bar{x}, \bar{w}, \bar{y}) := \Big\{ v \in \mathbb{R}^m \Big| \, v \in \partial\theta(\bar{z}) \ \text{ with } \ \nabla_x h(\bar{x}, \bar{w})^* v = \bar{y} \Big\}.$$

Assume the fulfillment of the second-order qualification condition:

$$\partial^2\theta(\bar{z}; v)(0) \cap \ker \nabla_x h(\bar{x}, \bar{w})^* = \{0\} \ \text{ for all } \ v \in M(\bar{x}, \bar{w}, \bar{y}). \tag{3.15}$$

Then we have the following chain rules for the partial second-order subdifferentials (2.12) *and* (2.13)*, respectively, valued for all* $u \in \mathbb{R}^n$*:*

(3.16)
$$\partial_x^2\varphi(\bar{x}, \bar{w}, \bar{y})(u) \subset \bigcup_{v \in M(\bar{x}, \bar{w}, \bar{y})} \nabla^2_{xx}\langle v, h\rangle(\bar{x}, \bar{w})u + \nabla_x h(\bar{x}, \bar{w})^*\partial^2\theta(\bar{z}, v)(\nabla_x h(\bar{x}, \bar{w})u),$$

(3.17)
$$\begin{aligned}\widetilde{\partial}_x^2\varphi(\bar{x}, \bar{w}, \bar{y})(u) \subset \bigcup_{v \in M(\bar{x}, \bar{w}, \bar{y})} & \Big(\nabla^2_{xx}\langle v, h\rangle(\bar{x}, \bar{w})u, \nabla^2_{xw}\langle v, h\rangle(\bar{x}, \bar{w})u\Big) \\ & + \Big(\nabla_x h(\bar{x}, \bar{w})^*\partial^2\theta(\bar{z}, v)(\nabla_x h(\bar{x}, \bar{w})u), \nabla_w h(\bar{x}, \bar{w})^*\partial^2\theta(\bar{z}, v)(\nabla_x h(\bar{x}, \bar{w})u)\Big).\end{aligned}$$

Proof. For brevity and simplicity of the arguments and notation, we present a detailed proof just for the full second-order subdifferential (2.11) of the strongly amenable nonparameterized compositions $\varphi(x) = \theta(h(x))$, in which case both formulas (3.16) and (3.17) reduce to

$$\partial^2\varphi(\bar{x}, \bar{y})(u) \subset \bigcup_{\substack{v \in \partial\theta(\bar{z}) \\ \nabla h(\bar{x})^* v = \bar{y}}} \Big(\nabla^2\langle v, h\rangle(\bar{x})u + \nabla h(\bar{x})^*\partial^2\theta(\bar{z}, v)(\nabla h(\bar{x})u)\Big) \tag{3.18}$$

with $\bar{z} = h(\bar{x})$ under the *basic second-order qualification condition*

$$\partial^2\theta(\bar{z}, v)(0) \cap \ker \nabla h(\bar{x})^* = \{0\} \ \text{ whenever } \ v \in \partial\theta(\bar{z}) \ \text{ and } \ \nabla h(\bar{x})^* v = \bar{y}. \tag{3.19}$$

The reader can readily check that the method of quadratic penalties developed below works perfectly for the case of partial second-order subdifferentials to produce the chain rule inclusions (3.16) and (3.17) under the "partial" second-order qualification condition (3.15).

We begin with observing that the first-order chain rule

$$\partial\varphi(x) = \nabla h(x)^*\partial\theta(h(x)) \ \text{ whenever } \ x \in U \tag{3.20}$$

holds as equality for strongly amenable compositions on some neighborhood U of $\bar{x}$. Indeed, it follows from the more general chain rule of [31, Theorem 3.41(iii)] due to (2.10) and the particular properties of strongly amenable functions summarized in [49, Exercise 10.25].

Now we proceed with calculating the second-order subdifferential $\partial^2\varphi(\bar{x}, \bar{y})$ for the given first-order subgradient $\bar{y} \in \partial\varphi(\bar{x})$. The definitions in (2.1), (2.9), and (2.5) suggest calculating the regular normal cone $\widehat{N}((x,y);\Omega)$ to the subdifferential graph $\Omega := \operatorname{gph}\partial\varphi$ of φ at points $(x,y) \in \operatorname{gph}\partial\varphi$ near $(\bar{x},\bar{y})$ and then passing to the limit therein as $(x,y) \to (\bar{x},\bar{y})$. To simplify notation, let us focus first on calculating $\widehat{N}((\bar{x},\bar{y});\Omega)$ for the graphical set Ω. Developing a *variational approach* to subdifferential calculus and employing the *smooth variational description* of regular normals from [49, Theorem 6.11] and [31, Theorem 1.30], we have that $(\omega,-\xi) \in \widehat{N}((\bar{x},\bar{y});\Omega)$ if and only if there is a smooth function $\vartheta\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ such that

$$\operatorname{argmin}_{(x,y)\in\Omega} \vartheta(x,y) = \{(\bar{x},\bar{y})\} \ \text{ and } \ \nabla\vartheta(\bar{x},\bar{y}) = (-\omega,\xi). \tag{3.21}$$

Using the first-order chain rule formula (3.20) allows us to transform the minimization problem in (3.21) into the following:

$$\begin{cases} \text{minimize } \vartheta(x,\nabla h(x)^* v) \ \text{ over all} \\ x \in \mathbb{R}^n, \ (z,v) \in \operatorname{gph}\partial\theta \ \text{ with } \ h(x) - z = 0. \end{cases} \tag{3.22}$$

We know from (3.21) that (x,z,v) is an optimal solution to (3.22) if and only if

$$x = \bar{x}, \ z = \bar{z} \ \text{ and } \ \nabla h(\bar{x})^* v = \bar{y}, \tag{3.23}$$

which is therefore the unique optimal solution to this problem. Denote $G := \operatorname{gph}\partial\theta$ and observe that this set is closed due to the convexity and lower semicontinuity of θ. Observe further that the mapping $M\colon \Omega \rightrightarrows \mathbb{R}^m$ defined on the set $\Omega = \operatorname{gph}\partial\varphi$ with the values

$$M(x,y) := \Big\{ v \in \partial\theta(h(x)) \Big| \ \nabla h(x)^* v = y \Big\} \tag{3.24}$$

is uniformly bounded near $(\bar{x},\bar{y})$. This can be easily verified, arguing by contradiction due to the qualification condition (3.14) in the definition of amenable functions.

By the aforementioned uniform boundedness of (3.24), we find $s, r > 0$ such that the inclusion $M(x,y) \subset I\!B_s(0)$ holds for all $(x,y) \in I\!B_r(\bar{x},\bar{y})$. Taking any sequence $\varepsilon_k \downarrow 0$ as $k \to \infty$, consider the family of *quadratic penalty problems*

$$\begin{cases} \text{minimize } \ \vartheta(x,\nabla h(x)^* v) + \frac{1}{2\varepsilon_k}\|h(x) - z\|^2 \\ \text{over all } \ (x,v) \in I\!B_r(\bar{x}) \times I\!B_{2s}(0) \ \text{ and } \ (z,v) \in G. \end{cases} \tag{3.25}$$

It is easy to see that the level set of each problem (3.25) is bounded. Thus this problem admits optimal solutions (x_k,z_k,v_k). To show that the triples (x_k,z_k,v_k) are uniformly bounded as $k \in I\!N$, it suffices to check the boundedness of the sequence $\{z_k\}$; indeed, this follows from

$$\begin{aligned} \|z_k\| &\le \|h(x_k) - z_k\| + \|h(x_k)\| \\ &\le \sqrt{2\varepsilon_k\Big[\vartheta(\bar{x},\nabla h(\bar{x})^*\bar{v}) - \vartheta\Big(x_k,\nabla h(x_k)^* v_k\Big)\Big]} + \|h(x_k)\| \\ &\le \sqrt{2\varepsilon_k\Big[\vartheta(\bar{x},\nabla h(\bar{x})^*\bar{v}) - \min_{(x,v)\in I\!B_r(\bar{x})\times I\!B_s(0)} \vartheta\Big(x,\nabla h(x)^* v\Big)\Big]} \\ &\quad + \max_{x\in I\!B_r(\bar{x})} \|h(x)\|, \quad k \in I\!N. \end{aligned}$$

By passing to a subsequence of (x_k,z_k,v_k) and taking into account the uniqueness of solution (3.23) to the unperturbed problem (3.22), we get

$$x_k \to \bar{x}, \ z_k \to \bar{z} \ \text{ and } \ y_k := \nabla h(x_k)^* v_k \to \nabla h(\bar{x})^* \bar{v} = \bar{y} \tag{3.26}$$

with $(\bar{z},\bar{v}) \in G$ and $(x_k,v_k) \in \operatorname{int}\big(I\!B_r(\bar{x}) \times I\!B_{2s}(0)\big)$ for all k sufficiently large. Now, applying the first-order necessary optimality conditions from [49, Theorem 6.12] to the solution (x_k,z_k,v_k) of the penalized problem (3.25) with a smooth cost function and a geometric constraint gives us

$$\nabla_x\Big[\vartheta\Big(x,\nabla h(x)^*v_k\Big)+\frac{1}{2\varepsilon_k}\|h(x)-z_k\|^2\Big]\Big|_{x=x_k}=0,$$

$$-\nabla_{z,v}\Big[\vartheta\Big(x_k,\nabla h(x_k)^*v\Big)+\frac{1}{2\varepsilon_k}\|h(x_k)-z\|^2\Big]\Big|_{(z,v)=(z_k,v_k)}\in\widehat{N}((z_k,v_k);G)$$

for all $k\in I\!N$. Denoting $p_k:=[h(x_k)-z_k]/\varepsilon_k$, these conditions calculate out to

$$\nabla_x\vartheta(x_k,y_k)+\nabla^2\langle v_k,h\rangle(x_k)\nabla_y\vartheta(x_k,y_k)+\nabla\langle p_k,h\rangle(x_k)=0, \tag{3.27}$$

$$\Big(p_k,-\nabla h(x_k)\nabla_y\vartheta(x_k,y_k)\Big)\in\widehat{N}((z_k,v_k);G). \tag{3.28}$$

By passing above to subsequences as $k\to\infty$ if needed, we can reduce the situation to considering one of the following two cases:

Case 1*:* $\{p_k\}$ converges to some $\bar{p}$.

Case 2*:* $p_k\to\infty$ while $\{p_k/\|p_k\|\}$ converges to some $\bar{p}\neq 0$.

In Case 1 it follows from (3.26), (3.27), and (3.28) that

$$\nabla_x\vartheta(\bar{x},\bar{y})+\nabla^2\langle\bar{v},h\rangle(\bar{x})\nabla_y\vartheta(\bar{x},\bar{y})+\nabla\langle\bar{p},h\rangle(\bar{x})=0, \tag{3.29}$$

$$\Big(\bar{p},-\nabla h(\bar{x})\nabla_y\vartheta(\bar{x},\bar{y})\Big)\in N((\bar{z},\bar{v});G), \tag{3.30}$$

where $\nabla_x\vartheta(\bar{x},\bar{y})=-\omega$ and $\nabla_y\vartheta(\bar{x},\bar{y})=\xi$ by the second equality in (3.21).

In Case 2 by, first dividing both parts of (3.27) and (3.28) by $\|p_k\|$ and then passing to the limit therein as $k\to\infty$, we get that

$$\nabla h(\bar{x})^*\bar{p}=0\ \text{ and }\ (\bar{p},0)\in N((\bar{z},\bar{v});G)\ \text{ with }\ \|\bar{p}\|=1. \tag{3.31}$$

Thus, by taking into account our choice of $(\omega,-\xi)\in\widehat{N}((\bar{x},\bar{y});\Omega)$ and the construction of M in (3.24), we deduce from (3.29)–(3.31) the existence of $\bar{v}\in M(\bar{x},\bar{y})$ and $\bar{p}$ satisfying either (3.29) and (3.30) or (3.31). Since the arguments above equally hold for every point $(x,y)\in\operatorname{gph}\partial\varphi$ near $(\bar{x},\bar{y})$, they ensure the following description of the regular normal cone to $\Omega=\operatorname{gph}\varphi$ at points $(x,y)\in\Omega$ in a neighborhood of the reference one $(\bar{x},\bar{y})$, where $G=\operatorname{gph}\partial\theta$:

$$\left[\begin{array}{l}(\omega,-\xi)\in\widehat{N}((x,y);\Omega)\Longrightarrow\ \exists\, v\in M(x,y),\ p\in\mathbb{R}^m\ \text{ such that}\\[2ex] \text{either }\ \left\{\begin{array}{l}\omega=\nabla^2\langle v,h\rangle(x)\xi+\nabla h(x)^*p\\ \text{with }\ (p,-\nabla h(x)\xi)\in N((h(x),v);G)\end{array}\right.\\[2ex] \text{or }\ \nabla h(x)^*p=0\ \text{ with }\ (p,0)\in N((h(x),v);G),\ \|p\|=1.\end{array}\right. \tag{3.32}$$

Next we take any basic normal $(\omega,-\xi)\in N((\bar{x},\bar{y});\Omega)$ (not just a regular one) and by (2.5) find sequences $(x_k,y_k)\to(\bar{x},\bar{y})$ and $(\omega_k,-\xi_k)\to(\omega,-\xi)$ as $k\to\infty$ satisfying

$$(x_k,y_k)\in\Omega\ \text{ and }\ (\omega_k,-\xi_k)\in\widehat{N}((x_k,y_k);\Omega)\ \text{ for all }\ k\in I\!N.$$

Employing the description of regular normals (3.32) ensures the existence of $v_k\in M(x_k,y_k)$ and $p_k\in\mathbb{R}^m$ such that the either/or alternative in (3.32) holds for each

$k \in I\!N$. Due to the established local boundedness of the mapping M, suppose, with no loss of generality, that

$$v_k \to v \text{ as } k \to \infty \text{ for some } v \in M(\bar{x}, \bar{y}).$$

By a further passage to subsequences, we can reduce the situation to where just one of the "either/or" parts of the alternative in (3.32) holds for all k. Consider first the "or" part of this alternative, i.e., the validity of

$$\nabla h(x_k)^* p_k = 0 \text{ with } (p_k, 0) \in N((h(x_k), v_k); G), \ \|p_k\| = 1 \text{ for all } k.$$

In this case the sequence $\{p_k\}$ has a cluster point p. Thus we get

$$\nabla h(\bar{x})^* p = 0 \text{ with } (p, 0) \in N((h(\bar{x}), v); G), \ \|p\| = 1, \text{ and } v \in M(\bar{x}, \bar{y}) \tag{3.33}$$

by passing to the limit as $k \to \infty$ and by taking into account the continuity of f and ∇h, as well as the robustness property

$$N(\bar{u}; G) = \operatorname*{Lim\,sup}_{u \xrightarrow{G} \bar{u}} N(u; G)$$

of the normal cone (2.5), which follows from its definition in finite dimensions.

When the "either" part holds, we proceed similarly to Cases 1 and 2 above. In the first case there is $p \in \mathbb{R}^m$ such that $p_k \to p$. Then the passage to the limit in

$$\omega_k = \nabla^2 \langle v_k, h \rangle (x_k) \xi_k + \nabla h(x_k)^* p_k, \quad (p_k, -\nabla h(x_k)\xi_k) \in N((h(x_k), v_k); G), \tag{3.34}$$

with taking into account the continuity assumptions and the convergence above, leads to

$$\omega = \nabla^2 \langle v, h \rangle (\bar{x}) \xi + \nabla h(\bar{x})^* p, \quad (p, -\nabla h(\bar{x})\xi) \in N((h(\bar{x}), v); G). \tag{3.35}$$

In the remaining case we have $\|p_k\| \to \infty$ and thus find p such that $p_k / \|p_k\| \to p$ with $\|p\| = 1$. Divide now both sides of (3.34) by $\|p_k\|$ for any large k and take the limit therein as $k \to \infty$. Then we again arrive at (3.33). Unifying (3.33) and (3.35) gives the description of basic normals to $\Omega = \operatorname{gph} \partial\varphi$ via the following alternative:

$$\left[\begin{array}{l} (\omega, -\xi) \in N((\bar{x}, \bar{y}); \Omega) \Longrightarrow \exists\, v \in M(\bar{x}, \bar{y}), \ p \in \mathbb{R}^m \text{ such that} \\ \\ \text{either } \left\{\begin{array}{l} \omega = \nabla^2 \langle v, h \rangle (\bar{x}) \xi + \nabla h(\bar{x})^* p \\ \text{with } (p, -\nabla h(\bar{x})\xi) \in N((h(\bar{x}), v); G) \end{array}\right. \\ \\ \text{or } \nabla h(\bar{x})^* p = 0 \text{ with } (p, 0) \in N((h(\bar{x}), v); G), \ \|p\| = 1. \end{array}\right. \tag{3.36}$$

Recalling the notation introduced in the theorem and in the definitions of the constructions used, we see that the "either" part of (3.36) amounts to the second-order subdifferential inclusion (3.18), while the "or" part of (3.36) means the negation of the basic second-order qualification condition (3.19). Thus the assumed fulfillment of (3.19) shows that the "or" part of (3.36) does not hold, which justifies the validity of the second-order chain rule (3.18).

By repeating finally the arguments above while taking into account that the partial counterparts of the first-order chain rule equality (3.20) are satisfied due to the results of [21, Proposition 3.4] (see also [49, Corollary 10.11] and [31, Corollary 3.44] in more general settings), we get the partial second-order subdifferential chain rules (3.16) and (3.17) under the partial second-order qualification condition (3.15). ∎

Remark 3.4 (second-order chain rules with inclusions). A chain rule in form (3.16) for the *full* second-order subgradient sets of strongly amenable compositions with compatible parameterization in finite dimensions was derived in [20] under the second-order subdifferential condition of type (3.15) with $\nabla h(\bar{x},\bar{w})$ replacing $\nabla_x h(\bar{x},\bar{w})$. The proof in [20] was based on applying a coderivative chain rule to full first-order subdifferential mappings. A similar approach was employed in [30] and [31, Theorem 3.74] to derive second-order chain rules of type (3.18) in infinite dimensions under an appropriate infinite-dimensional counterpart of the second-order qualification condition (3.19). Although the results of [30, 31] are applied to a more general class of subdifferential regular functions θ in (3.18), they require a number of additional assumptions in both finite and infinite dimensions. Finally, we mention a second-order chain rule of the inclusion type (3.18) obtained in [34, Theorem 3.1] for a special kind of strongly amenable composition with the indicator function $\theta = \delta(\cdot;\Theta)$ of a set Θ in finite dimensions, which does not generally require the fulfillment of the second-order qualification condition (3.19) while imposing instead a certain *calmness* assumption on some auxiliary multifunction. The latter holds, in particular, in the case of polyhedral sets Θ due to seminal results of [46].

Next we show that the second-order chain rule formula (3.18), and hence those in (3.16) and (3.17), cannot be generally used for the *precise calculation* of the second-order subdifferentials for strongly amenable compositions: the inclusion therein may be *strict* even for fairly simple functions θ and h in $\varphi = \theta \circ h$ without a kind of full rank condition.

Example 3.5 (strict inclusion in the second-order chain rule formula). The inclusion in (3.18) can be strict even when h is linear, while θ is piecewise linear and convex. Moreover, the set on the right-hand side of (3.18) can be nonempty when the set on the left-hand side is empty.

Proof. Let the functions $h\colon \mathbb{R}^2 \to \mathbb{R}^4$ and $\theta\colon \mathbb{R}^4 \to \mathbb{R}$ be given by

$$h(x_1,x_2) := (x_1,-x_1,x_2,-x_2) = Ax \ \text{ with } \ A := \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix},$$

$$\theta(z_1,z_2,z_3,z_4) := \max\Big\{z_1,z_2,z_3,z_4\Big\} = \sigma_M(z),$$

where $M := \{v = (v_1,v_2,v_3,v_4) \in \mathbb{R}^4 |\ v_i \geq 0,\ \sum_{i=1}^4 v_i = 1\}$ is the unit simplex in $\mathbb{R}^4$, and where σ_Ω stands for the support function of the set Ω. Considering the composition $\varphi(x) := \theta(h(x))$ on $\mathbb{R}^2$, observe that it can be represented as

$$\varphi(x) = \sigma_B(x) \ \text{ for } \ B := \Big\{(y_1,y_2) \in \mathbb{R}^2 \Big|\ |y_1| + |y_2| \leq 1\Big\}. \tag{3.37}$$

Note that the outer function θ in the strongly amenable (in fact fully amenable) composition $\theta \circ h$ is *convex piecewise linear* in the terminology of [49]; it can be equivalently described by [49, Theorem 2.49] as a function with the polyhedral epigraph.

Using the explicit form (3.37) of φ allows us to compute its second-order subdifferential $\partial^2\varphi(\bar{x},\bar{y})$ with $\bar{x} = (0,0)$ and $\bar{y} = (0,0)$ directly by Definition 2.1. Indeed, we get from (3.37) that $\partial\varphi(\bar{x}) = B$, and hence $\bar{y} \in \operatorname{int}\partial\varphi(\bar{x})$. This tells us that

$$\partial^2\varphi(\bar{x},\bar{y})(u) = \begin{cases} \mathbb{R}^2 & \text{if } \ u = (0,0), \\ \emptyset & \text{if } \ u \neq (0,0). \end{cases} \tag{3.38}$$

On the other hand, formula (3.18) reads as the inclusion $\partial^2\varphi(\bar{x},\bar{y})(u) \subset Q(u)$ with

$$Q(u) := \bigcup \Big\{ A^*\partial^2\theta(0,v)(Au) \Big|\, v \in M,\ A^*v = 0 \Big\}.$$

Take $\bar{u} = (0,1)$, and $\bar{v} = (1/2,1/2,0,0)$ and then check that $A\bar{u} = (0,0,1,-1)$, $\bar{v} \in M$, and $A^*\bar{v} = 0$. This ensures the converse inclusion

$$Q(\bar{u}) \supset A^*\partial^2\theta(0,\bar{v})(A\bar{u}),$$

which shows that we have $Q(\bar{u}) \neq \emptyset$ provided that $\partial^2\theta(0,\bar{v})(A\bar{u}) \neq \emptyset$. To check the latter, recall the representation of the outer function $\theta = \sigma_M = \delta_M^*$ via the indicator function $\delta_M = \delta(\cdot; M)$ of M. Thus we get the description

$$\omega \in \partial^2\theta(0,\bar{v})(A\bar{u}) \Longleftrightarrow -A\bar{u} \in \partial^2\delta_M(\bar{v},0)(-\omega).$$

Since the set M is a convex polyhedron, an exact formula for $\partial^2\delta_M(0,\bar{v})$ is available from [11].

In order to state this formula, we need to deal with the *critical cone* for a convex polyhedron Ω at $x \in \Omega$ with respect to $p \in \partial\delta_\Omega(x) = N_\Omega(x)$; this is a polyhedral cone defined by

$$K(x,p) := \big\{ w \in T_\Omega(x) \big|\, w \perp p \big\},$$

where $T_\Omega(x)$ is the tangent cone (2.7) to Ω at x. Recall that a *closed face* C of a polyhedral cone K is a polyhedral cone of the form

$$C := \big\{ x \in K |\ x \perp v \big\} \ \text{ for some } \ v \in K^*,$$

where K^* denotes the polar of the cone K. By the proof of [11, Theorem 2] (see also [43, Proposition 4.4]) we have the following description of the second-order subdifferential of the indicator function for a convex polyhedron:

$$w \in \partial^2\delta_\Omega(x,p)(u) \Longleftrightarrow \begin{cases} \text{there exist closed faces } \ C_1 \subset C_2 \ \text{ of } \ K(x,p) \\ \text{with } \ u \in C_1 - C_2,\ w \in (C_2 - C_1)^*. \end{cases} \tag{3.39}$$

Applying this to our setting with the simplex $\Omega = M$, we get the critical cone

$$K = T_M(\bar{v}) \cap 0^\perp = T_M(\bar{v}) = \left\{ (\omega_1,\omega_2,\omega_3,\omega_4) \,\middle|\, \begin{array}{l} \omega_1 + \omega_2 = 0 \\ \omega_3 \geq 0,\ \omega_4 \geq 0 \end{array} \right\}.$$

It follows from the second-order subdifferential formula (3.39) that

$$-A\bar{u} \in \partial^2\delta_M(\bar{v},0)(-\omega) \Longleftrightarrow \begin{cases} \text{there exist closed faces } \ C_1 \subset C_2 \ \text{ of } \ K \\ \text{with } \ \omega \in C_2 - C_1,\ -A\bar{u} \in (C_2 - C_1)^*. \end{cases}$$

Observe that the closed faces of K have the form

$$\{(\omega_1,\omega_2) |\ \omega_1 + \omega_2 = 0\} \times I \times J, \ \text{ where } \ I, J \ \text{ can be either } \ \mathbb{R}_+ \ \text{ or } \ \{0\}.$$

Denoting $L := \{(\omega_1,\omega_2) |\ \omega_1 + \omega_2 = 0\}$, we deduce from above the following possibilities:

$$C_1 - C_2 = L \times (I_1 - I_2) \times (J_1 - J_2) \ \text{ and } \ (C_1 - C_2)^* = L^\perp \times (I_1 - I_2)^* \times (J_1 - J_2)^*,$$

where $I_1 - I_2$ and $J_1 - J_2$ can be $\mathbb{R}$, $\mathbb{R}_+$, and $\{0\}$ while, respectively, $(I_1 - I_2)^*$ and $(J_1 - J_2)^*$ can be $\{0\}$, $\mathbb{R}_-$, and $\mathbb{R}$. Setting now $(I_1 - I_2)^* = \mathbb{R}_-$ and $(J_1 - J_2)^* = \mathbb{R}$, we get

$$-A\bar{u} \in \partial^2\delta_M(\bar{v}, 0)(-\omega) \text{ or, } \text{ equivalently, } \omega \in \partial^2\theta(0, \bar{v})(A\bar{u})$$

whenever $\omega \in L \times \mathbb{R}_+ \times \{0\}$. Taking, e.g., $\omega = (0, 0, 1, 0)$, gives us $\partial^2\theta(0, \bar{v})(A\bar{u}) \neq \emptyset$. Hence the set $Q(\bar{u})$ on the right-hand side of (3.18) is nonempty while $\partial^2\varphi(\bar{x}, \bar{y})(\bar{u}) = \emptyset$ by (3.38). □

It is not hard to check that the second-order qualification condition (3.19) does not hold in Example 3.5. Therefore, besides the emphasis above, this example can be considered as a counterexample to the equality in the second-order chain rule (3.18) with no full rank condition on the derivative and also as an illustration of the possible validity of the inclusion in (3.18) without the second-order qualification condition (3.19). In the next section we show that *if* condition (3.19) holds, then in fact it yields that the full rank condition must be satisfied for a large class of outer functions θ in amenable compositions $\theta \circ h$, including the one in Example 3.5. Thus the validity of the second-order qualification condition (3.19) in such settings ensures the *equality* in the second-order chain rule formula (3.18).

4. Exact second-order subdifferential chain rules for fully amenable compositions. This section is mainly devoted to deriving *exact* (equality-type) second-order subdifferential chain rules in the framework of (3.18) for major classes of *fully* amenable compositions $\varphi = \theta \circ h$ *without imposing* the full rank condition

$$\text{rank}\,\nabla h(\bar{x}) = m \Longleftrightarrow \ker \nabla h(\bar{x})^* = \{0\}. \tag{4.1}$$

This is done below on the basis of *calculating* the second-order subdifferential (2.11) for piecewise linear quadratic outer functions θ in the representation of φ.

Note that assuming the full rank condition (4.1) clearly ensures the second-order qualification condition (3.19) and, moreover, the validity by Theorem 3.1 of the second-order chain rule

$$\partial^2\varphi(\bar{x}, \bar{y})(u) = \nabla^2\langle\bar{v}, h\rangle(\bar{x})u + \nabla h(\bar{x})^*\partial^2\theta(\bar{z}, \bar{v})(\nabla h(\bar{x})u) \text{ for all } u \in \mathbb{R}^n \tag{4.2}$$

with a unique vector $\bar{v} \in \partial\theta(\bar{z})$ satisfying $\nabla h(\bar{x})^*\bar{v} = \bar{y}$. In what follows we show that the full rank assumption (4.1) is not needed for the validity of (4.2) if φ belongs to some favorable classes of composite functions widely encountered in variational analysis and optimization.

Recall [41, 49] that a strongly amenable function φ is *fully amenable* at $\bar{x}$ if the outer function $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ in its composite representation $\varphi = \theta \circ h$ can be chosen as *piecewise linear-quadratic*. The latter class includes piecewise linear functions discussed in Example 3.5. Their domains as well as their subgradient sets (2.2) and (2.3) are polyhedral; see [49, section10E].

We first present a general result on calculating the second-order subdifferential of piecewise linear-quadratic functions in fully amenable compositions, which is of independent interest while playing a significant role in deriving the exact second-order chain rules in the rest of this section.

THEOREM 4.1 (second-order subdifferentials of piecewise linear-quadratic functions). *Let $\varphi = \theta \circ h$ be a fully amenable composition at $\bar{x}$, let $M\colon \mathbb{R}^n \times \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be the set-valued mapping defined in* (3.24)*, and let $S(z)$ be the subspaces of $\mathbb{R}^m$ parallel* *to the affine hull* $\operatorname{aff}\partial\theta(z)$ *for all* z *near the point* $\bar z := h(\bar x)$. *Then for any sufficiently small neighborhood* O *of* $\bar z$ *we have the second-order subdifferential representation*

$$\partial^2\theta(\bar z,\bar v)(0) = \bigcup_{z\in O} S(z) \quad \textit{whenever} \;\; \bar v \in M(\bar x,\bar y), \tag{4.3}$$

where the union is taken only over finitely many subspaces $S(z)$.

Proof. As mentioned in the proof of Theorem 3.3, the mapping M from (3.24) is closed-graph around $(\bar x,\bar y,\bar z)$ and uniformly bounded around $(\bar x,\bar y)$ for strongly amenable compositions. Since $\theta\circ h$ is fully amenable at $\bar x$, the set $M(x,y)$ is also polyhedral for all (x,y) sufficiently close to $(\bar x,\bar y)$. This follows from the structure of M, the preservation of full amenability under small perturbations of the reference point, and the polyhedrality of subgradient sets for fully amenably functions; see [49, Exercise 10.25(a,b)].

Fix any $\bar v\in M(\bar x,\bar y)$. Since θ is piecewise linear-quadratic, its graph $G := \operatorname{gph}\partial\theta$ is *piecewise polyhedral*; i.e., it is the union of finitely many polyhedral sets in $\mathbb{R}^{2m}$. Using this and taking into account formulas (2.5) and (2.6), we find a neighborhood W of $(\bar z,\bar v)$ such that

$$N_G(\bar z,\bar v) = \bigcup\Big\{\widehat N_G(z,v)\Big|\,(z,v)\in G\cap W\Big\} = \bigcup\Big\{T_G(z,v)^*\Big|\,(z,v)\in G\cap W\Big\}, \tag{4.4}$$

where only *finitely many* cones (all of them polyhedral) occur in the unions. Therefore

$$w\in\partial^2\theta(\bar z,\bar v)(0) \Longleftrightarrow \exists\,(z,v)\in G\cap W \;\text{ with }\; (w,0)\in T_G(z,v)^*. \tag{4.5}$$

On the other hand, we have $T_G(z,v) = \operatorname{gph}(D\partial\theta)(z,v)$ by definition of the *graphical derivative* D of a set-valued mapping, and furthermore

$$(D\partial\theta)(z,v) = \partial\left(\frac{1}{2}d^2\theta(z,v)\right)$$

via the *second subderivative* of the function θ under consideration; see [49, Theorem 13.40 and Proposition 13.32] for more details. Hence it ensures that

$$\operatorname{dom}(D\partial\theta)(z,v) = \operatorname{dom} d^2\theta(z,v) = N_{\partial\theta(z)}(v)$$

by [49, Theorem 13.14]. Employing now (4.4) and (4.5) gives us the representations

$$\begin{aligned}\partial^2\theta(\bar z,\bar v)(0) &= \bigcup_{(z,v)\in G\cap W}\Big[\operatorname{dom}(D\partial\theta)(z,v)\Big]^* = \bigcup_{(z,v)\in G\cap W}\Big[N_{\partial\theta(z)}(v)\Big]^*\\ &= \bigcup_{(z,v)\in G\cap W} T_{\partial\theta(z)}(v),\end{aligned} \tag{4.6}$$

where only finitely many sets are taken in the unions. Pick $v\in\partial\theta(z)$ and find, by the polyhedrality of the subgradient sets $\partial\theta(z)$ and the construction of the subspaces $S(z)$, a vector $v'\in\operatorname{ri}\partial\theta(z)$ arbitrarily close to v, and get for all such v' the relationships

$$T_{\partial\theta(z)}(v') = S(z) \supset T_{\partial\theta(z)}(v),$$

which imply by (4.6) the equality

$$\partial^2\theta(\bar z,\bar v)(0) = \bigcup S(z),$$

where the finite union of *subspaces* is taken over z such that $(z,v) \in G \cap W$ for some v.

So far we have focused our analysis on a particular point $\bar v \in M(\bar x,\bar y)$ and an associated neighborhood W of $(\bar z,\bar v)$. Since the mapping M is of closed graph and locally bounded, the set $M(\bar x,\bar y)$ can be covered by *finitely many* of such neighborhoods. This allows us to obtain (4.3) and complete the proof of the theorem. □

Our next step is to justify the exact second-order chain rule (4.2) in the case when the outer function θ in the fully amenable composition is (convex) *piecewise linear*; see [49] and Example 3.5 above. This can be done by using Theorem 4.1 and the following *reduction lemma*, which describes a general setting when the second-order chain rule can be obtained by a local reduction to the full rank case.

LEMMA 4.2 (local reduction to full rank condition). *Let $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be a strongly amenable composition at $\bar x$ represented as $\varphi = \theta \circ h$ near $\bar x$ with $h\colon \mathbb{R}^n \to \mathbb{R}^m$ and $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}}$. Let L be any subspace of $\mathbb{R}^m$ satisfying the conditions*

$$L \cap \ker \nabla h(\bar x)^* = \{0\}, \tag{4.7}$$

$$\partial\theta(z) \subset L \ \text{ for all } z \in \mathbb{R}^m \text{ sufficiently close to } \ \bar z = h(\bar x). \tag{4.8}$$

Then we have the exact second-order chain rule formula (4.2).

Proof. Let $\dim L = s \le m$. It is easy to see that the case of $s = m$ corresponds by (4.7) to the full rank condition on $\nabla h(\bar x)$, and thus the second-order chain rule (4.2) follows in this case from Theorem 3.1. Suppose now that $s < m$, and let A be the matrix of a linear isometry from $\mathbb{R}^m$ into $\mathbb{R}^s \times \mathbb{R}^{m-s}$ under which $A^*L = \mathbb{R}^s \times \{0\}$. Denoting $P := A^{-1}h$ and $\vartheta := \theta A$ gives us the representation $\varphi = \vartheta \circ P$. It follows from the chain rule of convex analysis that

$$\partial\vartheta(w) = A^*\partial\theta(Aw) \ \text{ for } \ w = A^{-1}z \ \text{ and all } \ z \in \mathbb{R}^m.$$

This implies by (4.8) that for all z sufficiently close to $\bar x$ we have

$$\partial\vartheta(w) \subset A^*L = \mathbb{R}^s \times \{0\} \ \text{ with } \ w = A^{-1}z. \tag{4.9}$$

Hence the initial framework of the lemma can be reduced to one in which we have, in terms of $P(x) = (p_1(x),\ldots,p_m(x))$ and $w = Az$, the implication

$$\left.\begin{array}{l} z = Aw \ \text{ sufficiently close to } \ \bar z \\ v = (v_1,\ldots,v_m) \in \partial\vartheta(w) \end{array}\right\} \Longrightarrow v_{s+1} = 0,\ldots,v_m = 0.$$

This means that in analyzing $\partial\varphi$ locally via ϑ and P it is possible to pass with no loss of generality to the "submapping"

$$P_0\colon x \mapsto \big(p_1(x),\ldots,p_s(x)\big),$$

since only $p_1,\ldots,p_s$ are active locally while $p_{s+1},\ldots,p_m$ do not matter in the implication

$$y \in \partial\varphi(x) \Longrightarrow \exists\, v \in \partial\vartheta(P(x)) \ \text{ such that } \ \nabla P(x)^*v = y.$$

To reduce the conclusion of the lemma to that of Theorem 3.1, it remains to show that the Jacobian matrix $\nabla P_0(\bar x)$ is of full rank ($= s$). To proceed, consider the equation

$$\nabla P_0(\bar x)^*u = 0_n \ \text{ for } \ u \in \mathbb{R}^s, \tag{4.10}$$

where 0_n stands for the origin in $\mathbb{R}^n$. It follows from the construction of P_0 that (4.10) is equivalent to the equation

$$\nabla P(\bar{x})^*(u, 0_{m-s}) = 0_n.$$

Since $\nabla P(\bar{x}) = A^{-1}\nabla h(\bar{x})$, we get from the above equation that

$$\nabla h(\bar{x})^*\left[(A^*)^{-1}(u, 0_{m-s})\right] = 0_n,$$

which amounts to the inclusion

$$(A^*)^{-1}(u, 0_{m-s}) \in \ker \nabla h(\bar{x})^*.$$

The choice of A ensures that $(A^*)^{-1}(u, 0_{m-s}) \in L$. This gives by assumption (4.7) that $(A^*)^{-1}(u, 0_{m-s}) = 0_m$, and so $u = 0_s$. Thus (4.10) has only the trivial solution, which justifies the full rank condition for $\nabla P_0(\bar{x})$ and completes the proof of the lemma. □

Next we show that in the case of *piecewise linear* outer functions in fully amenable compositions $\varphi = \theta \circ h$ the second-order qualification condition (3.19) allows us to obtain the exact second-order chain rule (4.2) by reducing the situation to the setting considered in Lemma 4.2.

THEOREM 4.3 (second-order chain rule for fully amenable compositions with piecewise linear outer functions). *Let $\varphi = \theta \circ h$ be a fully amenable composition at $\bar{x}$, where $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ is* (*convex*) *piecewise linear. Picking any $(\bar{x}, \bar{y}) \in \operatorname{gph} \partial\varphi$, we get that the set $M(\bar{x}, \bar{y})$ in* (3.24) *is a singleton $\{\bar{v}\}$ and impose the second-order qualification condition* (3.19) *at $v = \bar{v}$. Then the exact second-order chain rule formula* (4.2) *holds.*

Proof. It is not hard to check that for convex piecewise linear functions θ we have the inclusion $\partial\theta(z) \subset \partial\theta(\bar{z})$ for any neighborhood O of $\bar{z}$ sufficiently small. This implies that $S(z) \subset S(\bar{z})$ whenever $z \in O$, where $S(z)$ stands for the affine hull of $\partial\theta(z)$. Hence representation (4.3) in Theorem 4.1 reduces in this case to the equality

$$\partial^2\theta(\bar{z}, \bar{v})(0) = S(\bar{z}) \;\text{ whenever }\; \bar{v} \in M(\bar{x}, \bar{y}). \tag{4.11}$$

Substituting (4.11) into the second-order qualification condition (3.19) gives us

$$S(z) \cap \ker \nabla h(\bar{x})^* = \{0\} \;\text{ for all }\; z \in O. \tag{4.12}$$

Recall that the subspace $S(z)$ consists of all vectors $\lambda(v' - v)$ such that $\lambda \in \mathbb{R}$ and $v, v' \in \partial\theta(z)$. Hence the second-order qualification condition (4.12) is equivalent to the following: There exist neighborhoods U of $\bar{z}$ and V of $\bar{y}$ such that

$$\Big[x \in U,\ y \in V,\ v, v' \in M(x, y)\Big] \Longrightarrow v = v'.$$

In other words, the latter means that the mapping M from (3.24) is *single-valued* on the subdifferential graph $\operatorname{gph} \partial\varphi$ around $(\bar{x}, \bar{y})$, and thus it is continuous as well.

To reduce the situation to the case considered in Lemma 4.2, we proceed as follows. Denote $L := S(\bar{z})$, and by the construction of $S(\bar{z})$ find $b \in \mathbb{R}^m$ such that $\partial\theta(\bar{z}) \subset L - b$. Consider now the convex piecewise linear function $\widetilde{\theta}(z) = \theta(z) - \langle b, z\rangle$ and the fully amenable composition $\widetilde{\varphi} := \widetilde{\theta} \circ h$. Recalling that $\partial\theta(z) \subset \partial\theta(\bar{z})$ for all z near $\bar{z}$, we get that $0 \in \partial\widetilde{\theta}(z)$ for such z, and thus Lemma 4.2 can be applied to the composition $\widetilde{\varphi} = \widetilde{\theta} \circ h$. It is easy to see that $\bar{y} - \nabla h(\bar{x})^* b \in \partial\varphi(\bar{x}) - \nabla h(\bar{x})^* b = \partial\widetilde{\varphi}(\bar{x})$ and that, by the second-order subdifferential sum rule from [31, Proposition 1.121], we have

$$\partial^2\varphi(\bar{x},\bar{y})(u) = \partial^2\widetilde{\varphi}\big(\bar{x},\bar{y} - \nabla h(\bar{x})^* b\big)(u) + \nabla^2\langle b,h\rangle(\bar{x})u \ \text{ for all } \ u \in \mathbb{R}^n. \tag{4.13}$$

Applying now the second-order chain rule (4.2) to the composition $\widetilde{\varphi} = \widetilde{\theta} \circ h$ at $(\bar{x}, \bar{y} - \nabla h(\bar{x})^* b)$ allows us to find $\widetilde{v} \in \partial\widetilde{\theta}(\bar{z})$ such that $\nabla h(\bar{x})^* \widetilde{v} = \bar{y} - \nabla h(\bar{x})^* b$ and

$$\partial^2\widetilde{\varphi}(\bar{x},\bar{y} - \nabla h(\bar{x})^* b) = \nabla^2\langle\widetilde{v},h\rangle(\bar{x}) + \nabla h(\bar{x})^* \partial^2\widetilde{\theta}(\bar{z},\widetilde{v})(\nabla h(\bar{x})u), \quad u \in \mathbb{R}^n. \tag{4.14}$$

Define $\bar{v} := \widetilde{v} + b$ and observe that $\bar{v} \in \partial\widetilde{\theta}(\bar{z}) + b = \partial\theta(\bar{z})$ and $\nabla h(\bar{x})^*\bar{v} = \nabla h(\bar{x})^*\widetilde{v} + \nabla h(\bar{x})^* b = \bar{y}$. Combining finally (4.13) and (4.14) gives us

$$\begin{aligned}\partial^2\varphi(\bar{x},\bar{y})(u) &= \nabla^2\langle b,h\rangle(\bar{x})u + \nabla^2\langle\widetilde{v},h\rangle(\bar{x}) + \nabla h(\bar{x})^*\partial^2\widetilde{\theta}(\bar{z},\widetilde{v})(\nabla h(\bar{x})u)\\ &= \nabla^2\langle\bar{v},h\rangle(\bar{x})u + \nabla h(\bar{x})^*\partial^2\theta(\bar{z},\bar{v})(\nabla h(\bar{x})u), \quad u \in \mathbb{R}^n.\end{aligned}$$

The uniqueness of $\bar{v} \in \partial\theta(\bar{z})$ satisfying $\nabla h(\bar{x})^*\bar{v} = \bar{y}$ follows from the arguments above based on the second-order qualification condition. This justifies the exact second-order chain rule (4.2) for $\varphi = \theta \circ h$ at $(\bar{x},\bar{y})$ and thus completes the proof of the theorem. ∎

Next we consider a major subclass of piecewise linear-quadratic outer functions in fully amenable compositions given by

$$\theta(z) := \sup_{v \in C}\left\{\langle v,z\rangle - \frac{1}{2}\langle v,Qv\rangle\right\}, \tag{4.15}$$

where $C \subset \mathbb{R}^m$ is a nonempty polyhedral set, and where $Q \in \mathbb{R}^{m \times m}$ is a symmetric positive-semidefinite matrix. Functions of this class are useful in many aspects of variational analysis and optimization, in particular, as penalty expressions in composite formats of optimization; see, e.g., [49] and the references therein. By definition (4.15) we see that θ is proper, convex, and piecewise linear-quadratic (piecewise linear when $Q = 0$) with the conjugate representation

$$\theta(z) = (\delta_C + j_Q)^*(z) \ \text{ for } \ j_Q(v) := \frac{1}{2}\langle v,Qv\rangle. \tag{4.16}$$

Prior to deriving the exact second-order chain rule for fully amenable compositions $\varphi = \theta \circ h$ with θ given in (4.15), we calculate the second-order subdifferential of such functions θ, which is undoubtedly of interest for its own sake while playing a significant role in the subsequent proof of the aforementioned second-order chain rule.

LEMMA 4.4 (calculating the second-order subdifferential for a major subclass of piecewise linear-quadratic functions). *Let $\theta\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be a linear-quadratic function given in* (4.15) *under the assumptions made above. Fix any $(\bar{z},\bar{v}) \in \operatorname{gph}\partial\theta$ and define*

$$K := T_C(\bar{v}) \cap (\bar{z} - Q\bar{v})^{\perp}, \tag{4.17}$$

where $T_C(\bar{v})$ is the tangent cone (2.7) *to C at $\bar{v}$. Then the second-order subdifferential of θ at $(\bar{z},\bar{v})$ is calculated by*

$$w \in \partial^2\theta(\bar{z},\bar{v})(u) \Longleftrightarrow \begin{cases}\exists \textit{ closed faces } \ K_1 \supset K_2 \ \textit{ of } \ K\\ \textit{with } \ w \in K_1 - K_2, \ Qw - u \in (K_1 - K_2)^*.\end{cases} \tag{4.18}$$

Furthermore, the set $\partial^2\theta(\bar{z},\bar{v})(0)$ *is actually a subspace given by*

$$\partial^2\theta(\bar{z},\bar{v})(0) = (\ker Q) \cap (K - K). \tag{4.19}$$

Proof. It follows from (4.16) that the conjugate function to θ is $\theta^* = \delta_C + j_Q$. Observe that

$$w \in \partial^2\theta(\bar{z},\bar{v})(u) \Longleftrightarrow -u \in \partial^2\theta^*(\bar{v},\bar{z})(-w). \tag{4.20}$$

Furthermore, we have by the calculations in [49, Example 11.18] that

$$\partial\theta^*(v) = N_C(v) + Qv, \quad v \in C, \tag{4.21}$$

and hence $\bar{z} \in \partial\theta^*(\bar{v}) \Longleftrightarrow \bar{z} - Q\bar{v} \in N_C(\bar{v})$. Using this and definition (2.11) of the second-order subdifferential and then applying the coderivative sum rule [31, Theorem 1.62] to (4.21) gives us

$$\partial^2\theta^*(\bar{v},\bar{z})(-w) = D^*N_C(\bar{v},\bar{z} - Q\bar{v})(-w) - Qw,$$

which implies in turn the representation

$$-u \in \partial^2\theta^*(\bar{v},\bar{z})(-w) \Longleftrightarrow Qw - u \in \partial^2\delta_C(\bar{v},\bar{z} - Q\bar{v})(-w). \tag{4.22}$$

Employing (4.20) and (4.22) and proceeding similarly to the consideration in Example 3.5 above, we derive from (4.22) the exact formula (4.18) for calculating the second-order subdifferential of $\partial^2\theta(\bar{z},\bar{v})$, where K in (4.17) is the *critical cone* for C at $\bar{v}$ with respect to $\bar{z} - Q\bar{v}$. The positive-semidefiniteness of the matrix Q yields that

$$0 \geq \langle w, Qw\rangle \Longleftrightarrow Qw = 0 \Longleftrightarrow w \in \ker Q,$$

which allows us to deduce from (4.18) that

$$\begin{aligned} w \in \partial^2\theta(\bar{z},\bar{v})(0) \quad &\Longleftrightarrow \ \exists\, K_1 \supset K_2 \ \text{ with } \ w \in K_1 - K_2,\ Qw \in (K_1 - K_2)^* \\ &\Longleftrightarrow \ \exists\, K_1 \supset K_2 \ \text{ with } \ w \in (\ker Q) \cap (K_1 - K_2) \\ &\Longleftrightarrow \ w \in (\ker Q) \cap (K - K), \end{aligned}$$

where the last equivalence holds since the cone K is a closed face itself. Thus the set $\partial^2\theta(\bar{z},\bar{v})(0)$ is a *subspace* in $\mathbb{R}^m$ calculated by formula (4.19). This completes the proof of the lemma. ∎

Now we are ready to justify the exact second-order chain rule (4.2) for fully amenable compositions with outer functions of type (4.15). The proof below essentially employs the second-order subdifferential calculations given Theorem 4.1 and Lemma 4.4, while it is independent from the reduction to the full rank reduction in Lemma 4.2. To proceed, recall that a mapping $h\colon \mathbb{R}^n \to \mathbb{R}^m$ is *open* around $\bar{x}$ if for any neighborhood U of $\bar{x}$ there is some neighborhood V of $h(\bar{x})$ such that $V \subset h(U)$. It is well known that the openness property is essentially different from the *linear* openness/openness at linear rate, which is characterized for smooth mappings between finite-dimensional spaces by the full rank condition (3.2); see [31, 49]. A simple example of a function $h\colon \mathbb{R} \to \mathbb{R}$ that is open but not linear open around $\bar{x} = 0$ is given by $h(x) = x^3$.

THEOREM 4.5 (second-order calculus rule for a major subclass of fully amenable compositions). *Let* $\varphi = \theta \circ h$ *be a fully amenable composition at* $\bar{x}$ *with* θ *of class* (4.15)*, and let* h *be open around* $\bar{x}$*. Assuming in addition that* Q *is positive-definite* *and picking any* $(\bar{x},\bar{y}) \in \operatorname{gph}\partial\varphi$*, we get that the set* $M(\bar{x},\bar{y})$ *in* (3.24) *is a singleton* $\{\bar{v}\}$ *and impose the second-order qualification condition* (3.19) *at* $v = \bar{v}$*. Then the exact second-order chain rule formula* (4.2) *holds.*

Proof. Since Q is positive-definite, we have $\ker Q = \{0\}$, and hence $\partial^2\theta(\bar{z},\bar{v})(0) = \{0\}$ by (4.19), where $\bar{z} = h(\bar{x})$ and $\bar{v} \in \partial\theta(\bar{z})$. Then it follows from representation (4.3) of Theorem 4.1 that $S(z) = \{0\}$ for all z around z sufficiently close to $\bar{z}$. This implies by the definition of $S(z)$ that the subdifferential $\partial\theta(z)$ is a *singleton* around $\bar{z}$, and hence so is $M(\bar{x},\bar{y}) = \{\bar{v}\}$. Thus we have the inclusion "$\subset$" in the second-order chain rule (4.2) by Theorem 3.3. Since $\partial\theta(z) = (N_C + Q)^{-1}(z)$ by [49, Example 11.18], it follows from [11, Theorem 2E.6] that the mapping $z \mapsto \partial\theta(z)$ is *Lipschitz continuous* around $\bar{z}$ with some constant ℓ_1. Taking into account this and the Lipschitz continuity of h around $\bar{x}$ with some constant ℓ_2, we prove that the opposite inclusion "$\supset$" holds in (4.2).

First let us show that, for all $u \in \mathbb{R}^n$ and $q \in \mathbb{R}^m$ sufficiently close to $\nabla h(\bar{x})u$, we have the inclusion

$$\nabla^2\langle\bar{v},h\rangle(\bar{x})u + \nabla h(\bar{x})^*\widehat{D}^*\partial\theta(\bar{z},\bar{v})(q) \subset \widehat{D}^*_{\ell_1\ell_2\|q-\nabla h(\bar{x})u\|}\partial\varphi(\bar{x},\bar{y})(u), \tag{4.23}$$

where $\widehat{D}^*$ stands for the "regular" version of the coderivative defined by scheme (2.9) with the replacement of N by the regular normal cone (2.4), while the right-hand side of (4.23) employs the ε-enlargement of the regular coderivative with $\varepsilon > 0$ replacing 0 in the construction of (2.4). Indeed, take any p in the left-hand side set of (4.23) and find $w \in \widehat{D}^*\partial\theta(\bar{z},\bar{v})(q)$ such that $p = \nabla^2\langle\bar{v},h\rangle(\bar{x})u + \nabla h(\bar{x})^*w$. Since $\partial\theta(z)$ is a singleton for all z around $\bar{z}$, we have that $\bar{v} = \partial\theta(h(\bar{x}))$ and that

$$\begin{aligned}
0 &\ge \operatorname*{Lim\,sup}_{z\to\bar{z}} \frac{\langle w, z-\bar{z}\rangle - \langle q, \partial\theta(z)-\partial\theta(\bar{z})\rangle}{\|z-\bar{z}\|} \\
&= \operatorname*{Lim\,sup}_{z\to\bar{z}} \frac{\langle w, z-\bar{z}\rangle - \langle \nabla h(\bar{x})u, \partial\theta(z)-\partial\theta(\bar{z})\rangle - \langle q-\nabla h(\bar{x})u, \partial\theta(z)-\partial\theta(\bar{z})\rangle}{\|z-\bar{z}\|} \\
&\ge \operatorname*{Lim\,sup}_{z\to\bar{z}} \frac{\langle w, z-\bar{z}\rangle - \langle \nabla h(\bar{x})u, \partial\theta(z)-\partial\theta(\bar{z})\rangle - \|q-\nabla h(\bar{x})u\|\cdot\|\partial\theta(z)-\partial\theta(\bar{z})\|}{\|z-\bar{z}\|} \\
&\ge \operatorname*{Lim\,sup}_{z\to\bar{z}} \frac{\langle w, z-\bar{z}\rangle - \langle \nabla h(\bar{x})u, \partial\theta(z)-\partial\theta(\bar{z})\rangle}{\|z-\bar{z}\|} - \ell_1\|q-\nabla h(\bar{x})u\|.
\end{aligned}$$

This implies, by choosing $z = h(x)$ and using $\|h(x)-h(\bar{x})\| \le \ell_2\|x-\bar{x}\|$, that

$$\begin{aligned}
\ell_1\ell_2\|q-\nabla h(\bar{x})u\| &\ge \operatorname*{Lim\,sup}_{x\to\bar{x}} \frac{\langle w, h(x)-h(\bar{x})\rangle - \langle \nabla h(\bar{x})u, \partial\theta(h(x))-\partial\theta(h(\bar{x}))\rangle}{\|x-\bar{x}\|} \\
&\ge \operatorname*{Lim\,sup}_{x\to\bar{x}} \frac{\langle \nabla h(\bar{x})^*w, x-\bar{x}\rangle - \langle \nabla h(\bar{x})u, \partial\theta(h(x))-\partial\theta(h(\bar{x}))\rangle}{\|x-\bar{x}\|}.
\end{aligned}$$

By the continuity of $z \mapsto \partial\theta(z)$ around $\bar{z}$ it gives that

$$\begin{aligned}
&\ell_1\ell_2\|q-\nabla h(\bar{x})u\| \\
&\quad\ge \operatorname*{Lim\,sup}_{x\to\bar{x}} \frac{\langle p-\nabla^2\langle\bar{v},h\rangle(\bar{x})u, x-\bar{x}\rangle - \langle \nabla h(\bar{x})u, \partial\theta(h(x))-\partial\theta(h(\bar{x}))\rangle}{\|x-\bar{x}\|} \\
&\quad\ge \operatorname*{Lim\,sup}_{x\to\bar{x}} \frac{\langle p, x-\bar{x}\rangle - \langle \bar{v}, (\nabla h(x)-\nabla h(\bar{x}))u\rangle - \langle \nabla h(\bar{x})u, \partial\theta(h(x))-\partial\theta(h(\bar{x}))\rangle}{\|x-\bar{x}\|}
\end{aligned}$$

$$\geq \operatorname*{Lim\,sup}_{x\to\bar{x}} \frac{\langle p, x-\bar{x}\rangle - \langle \partial\theta(h(x)), \nabla h(x)u\rangle + \langle \nabla h(\bar{x})u, \partial\theta(h(\bar{x}))\rangle}{\|x-\bar{x}\|}$$
$$+ \frac{\langle \partial\theta(h(x)) - \partial\theta(h(\bar{x})), (\nabla h(x)-\nabla h(\bar{x}))u\rangle}{\|x-\bar{x}\|}$$
$$\geq \operatorname*{Lim\,sup}_{x\to\bar{x}} \frac{\langle p, x-\bar{x}\rangle - \langle u, \nabla h(x)^*\partial\theta(h(x)) - \nabla h(\bar{x})^*\partial\theta(h(\bar{x}))\rangle}{\|x-\bar{x}\|}$$
$$\geq \operatorname*{Lim\,sup}_{x\to\bar{x}} \frac{\langle p, x-\bar{x}\rangle - \langle u, \partial\varphi(x) - \partial\varphi(\bar{x})\rangle}{\|x-\bar{x}\|},$$

where the mapping $x \mapsto \partial\varphi(x) = \nabla h(x)^*\partial\theta(h(x))$ is also single-valued and continuous around $\bar{x}$. This ensures that $p \in \widehat{D}^*_{\ell_1\ell_2\|q-\nabla h(\bar{x})u\|}\partial\varphi(\bar{x},\bar{y})(u)$ and thus justifies (4.23). Similarly we get

$$\text{(4.24)}\quad \nabla^2\langle v,h\rangle(x)u + \nabla h(x)^*\widehat{D}^*\partial\theta(z,v)(q) \subset \widehat{D}^*_{\ell_1\ell_2\|q-\nabla h(x)u\|}\partial\varphi(x,y)(u), \quad u \in \mathbb{R}^n,$$

for all x around $\bar{x}$ as well as $v = \partial\theta(z)$, $z = h(x)$, and $y = \partial\varphi(x)$. Next we show that

$$\text{(4.25)}\qquad \nabla^2\langle \bar{v},h\rangle(\bar{x})u + \nabla h(\bar{x})^*\partial^2\theta(\bar{z},\bar{v})(\nabla h(\bar{x})u) \subset \partial^2\varphi(\bar{x},\bar{y})(u)$$

whenever $u \in \mathbb{R}^n$, which is actually the inclusion "$\supset$" in (4.2). Indeed, take any p on the left-hand side of (4.25) to get

$$p - \nabla^2\langle \bar{v},h\rangle(\bar{x})u \in \nabla h(\bar{x})^*\partial^2\theta(\bar{z},\bar{v})(\nabla h(\bar{x})u).$$

Hence there are sequences (w_k, q_k) and $(z_k, v_k) \overset{\text{gph}\,\partial\theta}{\to} (\bar{z},\bar{v})$ as $k \to \infty$ satisfying

$$\text{(4.26)}\quad w_k \in \widehat{D}^*\partial\theta(z_k,v_k)(q_k) \text{ and } (\nabla h(\bar{x})^*w_k, q_k) \to (p - \nabla^2\langle \bar{v},h\rangle(\bar{x})u, \nabla h(\bar{x})u).$$

It follows from the assumed openness of h and from inclusion (4.24) that there is a sequence $(x_k, y_k) \to (\bar{x},\bar{y})$ with $y_k := \partial\varphi(x_k)$ such that $z_k = h(x_k)$ and that

$$\text{(4.27)}\quad \nabla^2\langle v_k,h\rangle(x_k)u + \nabla h(x_k)^*w_k \in \widehat{D}^*_{\ell_1\ell_2\|q_k-\nabla h(x_k)u\|}\partial\varphi(z_k,y_k)(u), \quad u \in \mathbb{R}^n.$$

Passing to the limit as $k \to \infty$ in (4.27) while $\|q_k - \nabla h(x_k)u\| \downarrow 0$ and using (4.26) give us

$$p \in D^*(\partial\varphi)(\bar{x},\bar{y})(u) = \partial^2\varphi(\bar{x},\bar{y})(u), \quad u \in \mathbb{R}^n,$$

which ensures (4.25) and thus completes the proof of the theorem. ∎

5. Applications to tilt stability in nonlinear and extended nonlinear programming. The second-order chain rules and subdifferential calculations obtained in sections 3 and 4 are undoubtedly useful in any settings where the second-order subdifferential (2.11) and its partial counterparts are involved; see the discussions and references in section 1. In this section we confine ourselves to the usage of second-order chain rules for deriving *full characterizations* of *tilt-stable local minimizers* in some important classes of constrained optimization problems. It requires applying *equality-type* formulas of the second-order subdifferential calculus.

The notion of tilt-stable minimizers was introduced by Poliquin and Rockafellar [43] in order to characterize strong manifestations of optimality that support computational work via the study of how local optimal solutions react to shifts (tilt perturbations) of the data. Recall that a point $\bar{x}$ is a *tilt-stable local minimizer* of the function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ finite at $\bar{x}$ if there is $\gamma > 0$ such that mapping

$$M\colon y \mapsto \text{argmin}\Big\{\varphi(x) - \varphi(\bar{x}) - \langle y, x - \bar{x}\rangle \Big|\ \|x - \bar{x}\| \le \gamma\Big\}$$

is single-valued and Lipschitz continuous on some neighborhood of $y = 0$ with $M(0) = \bar{x}$.

It is proved in [43, Theorem 1.3] that for $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ having $0 \in \partial\varphi(\bar{x})$ and such that φ is both prox-regular and subdifferentially continuous at $\bar{x}$ for $\bar{y} = 0$, the point $\bar{x}$ is a tilt-stable local minimizer of φ if and only if the second-order subdifferential mapping $\partial^2\varphi(\bar{x}, 0)\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is *positive-definite* in the sense that

$$\langle w, u\rangle > 0 \ \text{ whenever } \ w \in \partial^2\varphi(\bar{x}, 0)(u) \ \text{ with } \ u \ne 0. \tag{5.1}$$

The aforementioned properties of prox-regularity and subdifferential continuity introduced in [42] (see also [49, Definitions 13.27 and 13.28]) hold for broad classes of "nice" functions encountered in variational analysis and optimization. In particular, both properties are satisfied at all points of a neighborhood of $\bar{x}$ for any function strongly amenable at $\bar{x}$; see [49, Proposition 13.32].

Our subsequent goal is to extend the characterization of tilt-stable local minimizers from [43] to favorable classes of *constrained optimization* problems. To proceed, we use the following *composite format* of optimization known as *extended nonlinear programming* (ENLP); see [48, 49]:

$$\text{Minimize } \ \varphi(x) := \varphi_0(x) + \theta\big(\varphi_1(x), \ldots, \varphi_m(x)\big) = \varphi_0(x) + (\theta \circ \Phi)(x) \ \text{ over } \ x \in \mathbb{R}^n, \tag{5.2}$$

where $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ is an extended-real-valued function, and where $\Phi(x) :=(\varphi_1(x), \ldots \varphi_m(x))$ is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$. Written in the unconstrained format, problem (5.2) is actually a problem of constrained optimization with the set of feasible solutions given by

$$X := \{x \in \mathbb{R}^n |\ (\varphi_1(x), \ldots, \varphi_m(x)) \in Z\} \ \text{ for } \ Z := \{z \in \mathbb{R}^m |\ \theta(z) < \infty\}.$$

As argued in [48], the composite format (5.2) is a convenient general framework from both theoretical and computational viewpoints to accommodate a variety of particular models in constrained optimization. Note that the conventional problem of NLP with s inequality constraints and $m - s$ equality constraints can be written in the form

$$\text{minimize } \ \varphi_0(x) + \delta_Z(\Phi(x)) \ \text{ over } \ x \in \mathbb{R}^m \tag{5.3}$$

via the indicator functions of the set $Z = \mathbb{R}^s_- \times \{0\}^{m-s}$.

Our first result provides a complete second-order characterization of tilt-stable minimizers $\bar{x}$ for a general class of problems (5.2) under full rank of the Jacobian matrix $\nabla\Phi(\bar{x})$.

THEOREM 5.1 (characterization of tilt-stable minimizers for constrained problems with full rank condition). *Let $\bar{x} \in X$ be a feasible solution to* (5.2) *such that φ_0 and Φ are smooth around $\bar{x}$ with their derivatives strictly differentiable at $\bar{x}$, that* $\operatorname{rank} \nabla\Phi(\bar{x}) = m$*, and that θ is prox-regular and subdifferentially continuous at $\bar{z} := \Phi(\bar{x})$ for the* (*unique*) *vector $\bar{v} \in \mathbb{R}^m$ satisfying the relationships*

$$\bar{v} \in \partial\theta(\bar{z}) \ \textit{ and } \ \nabla\Phi(\bar{x})^*\bar{v} = -\nabla\varphi_0(\bar{x}). \tag{5.4}$$

Then $\bar{x}$ with $-\nabla\varphi_0(\bar{x}) \in \nabla\Phi(\bar{x})^*\partial\theta(\bar{z})$ is a tilt-stable local minimizer of (5.2) *if and only if the mapping $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ given by*

$$T(u) := \nabla^2\varphi_0(\bar{x})u + \nabla^2\langle\bar{v},\Phi\rangle(\bar{x})u + \nabla\Phi(\bar{x})^*\partial^2\theta(\bar{z},\bar{v})(\nabla\Phi(\bar{x})u), \quad u \in \mathbb{R}^n, \tag{5.5}$$

is positive-definite in the sense of (5.1).

Proof. Since φ_0 and Φ are smooth around $\bar{x}$ and $\nabla\Phi(\bar{x})$ has full rank m, it follows from the first-order subdifferential sum and chain rules of [49, Corollary 10.9 and Exercise 10.7] that

$$0 \in \partial\varphi(\bar{x}) \Longleftrightarrow -\nabla\varphi_0(\bar{x}) \in \nabla\Phi(\bar{x})^*\partial\theta(\bar{z})$$

for φ in (5.2). Furthermore, these rules and the definitions of prox-regularity and subdifferential continuity in [49] imply that the latter properties of φ at $\bar{x}$ for $0 \in \partial\varphi(\bar{x})$ are equivalent to the corresponding properties of θ at $\bar{z}$ for $\bar{v}$ satisfying (5.4).

It remains to check therefore that the positive-definiteness (5.1) of $\partial^2\varphi(\bar{x},0)$ is equivalent to that of T in (5.5). We show in fact that $\partial^2\varphi(\bar{x},0)(u) = T(u)$ for all $u \in \mathbb{R}^n$. Indeed, using the second-order chain rule from [31, Proposition 1.121] in (5.2) gives us

$$\partial^2\varphi(\bar{x},0)(u) = \nabla^2\varphi_0(\bar{x})u + \partial^2(\theta\circ\Phi)(\bar{x},-\nabla\varphi_0(\bar{x}))(u), \quad u \in \mathbb{R}^n. \tag{5.6}$$

To complete the proof of the theorem, we finally apply the exact second-order chain rule from Theorem 3.1 to the composition $\theta\circ\Phi$ in the latter equality. ◻

Next we address the conventional model of *nonlinear programming* (NLP) with smooth data:

$$\text{Minimize } \varphi_0(x) \text{ subject to } \varphi_i(x) = \begin{cases} \le 0 & \text{for } i = 1,\ldots,s,\\ = 0 & \text{for } i = s+1,\ldots,m. \end{cases} \tag{5.7}$$

As mentioned above, problem (5.7) can be written in form (5.3) with $Z = \mathbb{R}^s_- \times \{0\}^{m-s}$. For this problem, the full rank condition of Theorem 5.1 corresponds to the following: The gradients

$$\nabla\varphi_1(\bar{x}),\ldots,\nabla\varphi_m(\bar{x}) \text{ are linearly independent.} \tag{5.8}$$

Since our analysis is local, and since Lagrange multipliers corresponding to inactive inequality constraints disappear due to the complementarity conditions, in what follows we can drop any inactive inequality constraints from the picture; see also the remark after the proof of Theorem 5.2. Thus the situation is reduced to

$$\varphi_i(\bar{x}) = 0 \text{ for all } i = 1,\ldots,m. \tag{5.9}$$

Then the full rank condition (5.8) in case (5.9) is the classical *linear independence constraint qualification* (LICQ): The active constraint gradients at $\bar{x}$ are linearly independent.

To proceed further, consider the Lagrangian function in (5.7) given by

$$L(x,\lambda) := \varphi_0(x) + \sum_{i=1}^m \lambda_i\varphi_i(x) \text{ with } \lambda = (\lambda_1,\ldots,\lambda_m) \in \mathbb{R}^m$$

and recall that, for any local optimal solution $\bar{x}$ to (5.7), the LICQ at $\bar{x}$ ensures the existence of a *unique* multiplier vector $\bar{\lambda} = (\bar{\lambda}_1,\ldots,\bar{\lambda}_m) \in \mathbb{R}^s_+ \times \mathbb{R}^{m-s}$ such that

$$\nabla_x L(\bar{x},\bar{\lambda}) = \nabla\varphi_0(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i\nabla\varphi_i(\bar{x}) = 0. \tag{5.10}$$

Define the index sets for the inequality and equality constraints in (5.7) by

$$I_1 := \{i \in \{1, \ldots, s\} \mid \bar{\lambda}_i > 0\},$$
$$I_2 := \{i \in \{1, \ldots, s\} \mid \bar{\lambda}_i = 0\},$$
$$I_3 := \{s+1, \ldots, m\}$$

and recall that the *strong second-order optimality condition* (SSOC) holds at $\bar{x}$ if

$$\langle u, \nabla^2_{xx} L(\bar{x}, \bar{\lambda}) u \rangle > 0 \ \text{ for all } \ 0 \neq u \in S, \tag{5.11}$$

where the subspace $S \subset \mathbb{R}^n$ is given by

$$S := \{u \in \mathbb{R}^n \mid \langle \nabla \varphi_i(\bar{x}), u \rangle = 0 \ \text{ whenever } \ i \in I_1 \cup I_3\}. \tag{5.12}$$

Note that (5.11) is also known as the "strong second-order sufficient condition" for local optimality; see, e.g., [3]. The following theorem shows that, in the setting under consideration, the SSOC is *necessary and sufficient* for the tilt stability of local minimizers.

THEOREM 5.2 (characterization of tilt-stable local minimizers for NLP). *Let $\bar{x}$ be a feasible solution to* (5.7) *such that all of the functions φ_i for $i = 0, \ldots, m$ are smooth around $\bar{x}$ with their derivatives strictly differentiable at $\bar{x}$ and that the LICQ is satisfied at this point. Then we have the following assertions:*

(i) *If $\bar{x}$ is a tilt-stable local minimizer of* (5.7)*, then SSOC* (5.11) *holds at $\bar{x}$ with a unique multiplier vector $\bar{\lambda} \in \mathbb{R}^s_+ \times \mathbb{R}^{m-s}$ satisfying* (5.10).

(ii) *Conversely, the validity of SSOC at $\bar{x}$ with $\bar{\lambda} \in \mathbb{R}^s_+ \times \mathbb{R}^{m-s}$ satisfying* (5.10) *implies that $\bar{x}$ is a tilt-stable local minimizer of* (5.7).

Proof. As mentioned above, the LICQ corresponds to the full rank condition of Theorem 5.1. The prox-regularity and subdifferential continuity of $\theta = \delta_Z$ with $Z = \mathbb{R}^s_- \times \{0\}^{m-s}$ follow from its convexity [49, Example 13.30]. Let us next represent the mapping T in (5.5) via the initial data of problem (5.7). It is easy to see that $T(u)$ reduces in this case to

$$T(u) = \nabla^2_{xx} L(\bar{x}, \bar{\lambda}) u + \nabla \Phi(\bar{x})^* \partial^2 \delta_Z(0, \bar{\lambda})(\nabla \Phi(\bar{x}) u),$$

with $\Phi = (\varphi_1, \ldots, \varphi_m)$ and $Z = \mathbb{R}^s_- \times \{0\}^{m-s}$, provided that the first-order condition (5.10) is satisfied. This is of course the case when $\bar{x}$ is a tilt-stable local minimizer of (5.7), since it is a standard local minimizer as well. Thus the positive-definiteness of $T(u)$ amounts to

$$u \neq 0, \quad w \in \partial^2 \delta_Z(0, \bar{\lambda})(\nabla \Phi(\bar{x}) u) \Longrightarrow \langle u, \nabla^2_{xx} L(\bar{x}, \bar{\lambda}) u \rangle + \langle w, \nabla \Phi(\bar{x}) u \rangle > 0. \tag{5.13}$$

To proceed, we calculate the second-order subdifferential $\partial^2 \delta_Z(0, \bar{\lambda})$ in (5.13) by using formula (3.39) presented and discussed in Example 3.5. Observe that the critical cone in this situation is $K = Z \cap \bar{\lambda}^\perp$. It follows directly from (3.39) that

$$w \in \partial^2 \delta_Z(0, \bar{\lambda})(\nabla \Phi(\bar{x}) u) \Longleftrightarrow \begin{cases} \text{there exist closed faces } \ K_1 \subset K_2 \ \text{ of } \ K \\ \text{with } \ \nabla \Phi(\bar{x}) u \in K_1 - K_2, \ w \in (K_2 - K_1)^*. \end{cases} \tag{5.14}$$

The latter implies in turn that

$$\min\nolimits_{w \in \partial^2 \delta_Z(0, \bar{\lambda})(\nabla \Phi(\bar{x}) u)} \langle w, \nabla \Phi(\bar{x}) u \rangle = 0 \ \text{ for all } \ u \in \operatorname{dom} \partial^2 \delta_Z(0, \bar{\lambda})(\nabla \Phi(\bar{x})(\cdot)) \tag{5.15}$$

with the subdifferential domain representation

$$\operatorname{dom}\partial^2\delta_Z(0,\bar\lambda)(\nabla\Phi(\bar x)(\cdot)) = \bigcup\Big\{(K_1-K_2)\Big|\, K_1\subset K_2 \text{ closed faces of } K\Big\} = K-K.$$

To further elaborate on condition (5.15), we observe that for all vectors $v = (v_1,\ldots,v_m)\in K = Z\cap\bar\lambda^\perp$ we have the relationships

$$\langle v,\bar\lambda\rangle = \sum_{i=1}^m v_i\bar\lambda_i = \sum_{i=1}^s v_i\bar\lambda_i = 0 \text{ and } v_i = 0 \text{ whenever } i\in I_3.$$

This implies therefore the critical cone representation

$$K = \big\{v\in\mathbb{R}^m\big|\, v_i = 0 \text{ for } i\in I_1\cup I_3 \text{ and } v_i\le 0 \text{ for } i\in I_2\big\}. \tag{5.16}$$

Taking now any pair (u,w) with $w\in\partial^2\delta_Z(0,\bar\lambda)(\nabla\Phi(\bar x)u)$ and $u\neq 0$, we find by (5.14) two closed faces $K_1\subset K_2$ of K such that

$$\nabla\Phi(\bar x)u\in K_1-K_2 \text{ and } w\in(K_2-K_1)^*.$$

It follows from representation (5.16) that

$$v\in K_1-K_2\Longrightarrow v_i = 0 \text{ for all } i\in I_1\cup I_3.$$

Hence we get from $\nabla\Phi(\bar x)u\in K_1-K_2$ that

$$\langle\nabla\varphi_i(\bar x),u\rangle = 0 \text{ for all } i\in I_1\cup I_3$$

and thus conclude from (5.13), (5.14), and (5.15) that SSOC (5.11) holds. This completes the proof of assertion (i) in the theorem.

To prove assertion (ii), assume that SSOC (5.11) is satisfied together with (5.10) under the validity of LICQ at $\bar x$. To show that $\bar x$ is a tilt-stable local minimizer for (5.7), we need to check by Theorem 5.1 that the positive-definiteness condition (5.13) holds. Indeed, taking $w\in\partial^2\delta_Z(0,\bar\lambda)(\nabla\Phi(\bar x)u)$ with $u\neq 0$ and using representations (5.14) and (5.16) established above, we have by the SSOC at $\bar x$ that

$$\begin{aligned}\langle u,\nabla^2_{xx}L(\bar x,\bar\lambda)u\rangle + \langle w,\nabla\Phi(\bar x)u\rangle &\ge \langle u,\nabla^2_{xx}L(\bar x,\bar\lambda)u\rangle\\ &\quad+\min\big\{\langle w,\nabla\Phi(\bar x)u\rangle\big|\, w\in\partial^2\delta_Z(0,\bar\lambda)(\nabla\Phi(\bar x)u)\big\}\\ &\ge \langle u,\nabla^2_{xx}L(\bar x,\bar\lambda)u\rangle + 0 > 0.\end{aligned}$$

This justifies the positive-definiteness of $T(u)$ and ends the proof of the theorem. $\square$

It follows from the proof of Theorem 5.2 that assertion (i) holds true if we modify the SSOC condition (5.11) by narrowing the subspace S therein while considering only those vectors $u\in S$ satisfying in addition the relationships

$$\langle\nabla\varphi_i(\bar x),u\rangle\le 0 \text{ for } i\in I_2.$$

However, such a narrowing of S in (5.12) is not suitable to justify assertion (ii), since it does not allow us to ensure the validity of the inclusion $\nabla\Phi(\bar x)u\in K_1-K_2$.

Note also that the presence of *inactive* inequality constraints at $\bar x$ in problem (5.7) does not change the result of Theorem 5.2 and its proof given above. The only modification to make is to observe that the critical cone in this case is $K = T_Z(\Phi(\bar{x})) \cap \bar{\lambda}^\perp$ (cf. Example 3.5), and it admits the representation

$$K = \big\{ v \in \mathbb{R}^m \big|\, v_i \in \mathbb{R} \ \text{ for } \ i \in I_0,\ v_i = 0 \ \text{ for } \ i \in I_1 \cup I_3, \ \text{ and } \ v_i \le 0 \ \text{ for } \ i \in I_2 \big\},$$

where the index sets I_1, I_2, and I_3 are defined as in the proof of Theorem 5.2 for equalities and *active* inequality constraints, and where the set I_0 consists of indexes corresponding to inequality constraints that are *inactive* at $\bar{x}$.

The obtained characterization of tilt-stable minimizers for NLP leads us to compare this notion with the classical Robinson's notion of *strong regularity* [45] of parameterized variational inequalities associated with the KKT conditions for NLP (5.7); see, e.g., [3, 11, 12] for the exact definition and more discussions. Complete characterizations of strong regularity for NLP are derived in [11]; see also the references therein.

COROLLARY 5.3 (comparing tilt stability and strong regularity). *Under the assumptions of Theorem* 5.2*, the tilt stability of local minimizers for* (5.7) *is equivalent to the strong regularity of the variational inequality associated with the KKT conditions for* (5.7)*.*

Proof. The proof follows directly from Theorem 5.2 and the characterization of strong regularity obtained in [11, Theorem 5 and Theorem 6]. ∎

It is not hard to check that the strong regularity of the KKT system directly implies the LICQ at the corresponding solution of (5.7). On the other hand, the LICQ requirement arising from the full rank condition of Theorem 5.1 is essential for the SSOC characterization of tilt stability in Theorem 5.2. Furthermore, even imposing the seemingly less restrictive second-order qualification condition (3.19) needed for deriving the second-order chain rule *unavoidably* leads us to the LICQ requirement for NLP, since the latter class is represented via fully amenable compositions with piecewise linear outer functions θ in the composite format (5.2). This follows from the results of section 4 and is reflected in the next theorem.

THEOREM 5.4 (characterizing tilt-stable minimizers for constrained problems described by fully amenable compositions). *Let $\bar{x}$ be a feasible solution to* (5.2) *such that φ_0 is smooth around $\bar{x}$ with the strictly differentiable derivative at $\bar{x}$ and that the composition $\theta \circ \Phi$ is fully amenable at $\bar{x}$ with the outer function $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ of the following types:*

- *either θ is piecewise linear,*
- *or θ is of class* (4.15) *under the assumptions of Theorem* 4.5*.*

Assume further that the second-order qualification condition (3.19) *holds at $\bar{x}$ with $h = \Phi$ therein, where $v = -\bar{v}$ is a unique vector satisfying* (5.4) *with $\bar{z} = \Phi(\bar{x})$. Then $\bar{x}$ with $-\nabla\varphi_0(\bar{x}) \in \nabla\Phi(\bar{x})^*\partial\theta(\bar{z})$ is a tilt-stable local minimizer of* (5.2) *if and only if the mapping $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ defined in* (5.5) *is positive-definite in the sense of* (5.1)*, where the second-order subdifferential $\partial^2\theta(\bar{z}, \bar{v})$ is calculated by formulas* (4.11) *and* (4.18)*, respectively.*

Proof. Observe first that in both cases under consideration the composition $\theta \circ \Phi$ is prox-regular and subdifferentially continuous at any point x around $\bar{x}$ by [49, Proposition 13.32]; hence the same holds for the function φ from (5.2). It follows from Theorems 4.3 and 4.5 that, under the validity of the second-order qualification condition (3.19), we have the unique vector $\bar{v}$ satisfying (5.4) and the second-order chain rule

$$\partial^2(\theta \circ \Phi)(\bar{x}, -\nabla\varphi_0(\bar{x}))(u) = \nabla^2\langle \bar{v}, \Phi\rangle(\bar{x})u + \nabla\Phi(\bar{x})^*\partial^2\theta(\bar{z}, \bar{v})(\nabla\Phi(\bar{x})u) \tag{5.17}$$

for all $u \in \mathbb{R}^n$ when θ belongs to one of the classes considered in this theorem. Substituting (5.17) into formula (5.6) due to the second-order sum rule from [31, Proposition 1.121] allows us to justify that

$$\partial^2\varphi(\bar{x},0) = T(u) \text{ whenever } u \in \mathbb{R}^n,$$

and thus the positive-definiteness of the mapping T from (5.5) fully characterizes the tilt stability of the local minimizer $\bar{x}$ of (3.1) in both cases of θ under consideration with the formulas for calculating $\partial^2\theta(\bar{z},\bar{v})$ derived in the proofs of Theorems 4.3 and 4.5, respectively. □

We conclude the paper with the following three final remarks.

Remark 5.5 (sufficient conditions for tilt-stable local minimizers). The second-order chain rule (3.18) of the *inclusion* type derived in Theorem 3.3 for *strongly amenable* compositions and the second-order sum rule inclusions obtained in [31, Theorem 3.73] allow us to establish general *sufficient* conditions for tilt-stable local minimizers in large classes of constrained optimization problems written in the composite format (5.2). Indeed if, in addition to the hypotheses of Theorem 3.3 for the composition $\theta \circ \Phi$ in (5.2), we assume that the function φ_0 is, e.g., of class $\mathcal{C}^{1,1}$ around $\bar{x}$, then we have by the second-order sum rule from [31, Theorem 3.73(i)] and the chain rule of Theorem 3.3 the fulfillment of the inclusion

$$\partial^2\varphi(\bar{x},0)(u) \subset T(u), \quad u \in \mathbb{R}^n, \tag{5.18}$$

for φ from (5.2) and T from (5.5). The prox-regularity and subdifferential continuity of such functions φ follow, under the assumptions made, from [49, Proposition 13.32 and Proposition 13.34] and first-order subdifferential calculus rules. Thus inclusion (5.18) ensures that the positive-definiteness of T implies that of $\partial^2\varphi(\bar{x},0)$, and the former is therefore a sufficient condition for tilt stability of local minimizers of (5.2).

Remark 5.6 (full stability of local minimizers). Developing the concept of tilt stability, Levy, Poliquin, and Rockafellar [21] introduced the notion of *fully stable* local minimizers of general optimization problems of the type

$$\text{minimize } \varphi(x,u) - \langle v,x\rangle \text{ over } x \in \mathbb{R}^n \tag{5.19}$$

with respect to both "basic" perturbations u and "tilt" perturbations v. The main result of that paper [21, Theorem 2.3] establishes a *complete characterization* of fully stable local minimizers of (5.19) via the positive-definiteness of the *extended partial second-order subdifferential* (2.13) of φ. Similarly to the results of this section for tilt-stable minimizers of constraint optimization problems written in the composite format (5.2), we can derive characterizations as well as sufficient conditions for fully stable minimizers of (5.19) based on [21, Theorem 2.3] and the second-order chain rules for the partial second-order counterpart (2.13) established above. Our ongoing research project is to comprehensively elaborate these developments on full stability in constrained optimization and its applications. We have already obtained some results in this direction, while further research is required to complete the project.

Remark 5.7 (tilt stability and partial smoothness). After completing this paper, we became aware of the concurrent work by Lewis and Zhang [23] related to second-order subdifferentials (generalized Hessians) and tilt stability. The main results of [23] provide calculations of the basic second-order construction from Definition 2.1 for $\mathcal{C}^2$–partly smooth functions on $\mathcal{C}^2$-smooth manifolds and then characterize tilt stability in such settings via strong criticality and local quadratic growth. To this end, note that a certain uniform second-order growth condition is used in [3, Theorem 5.36] for characterizing tilt stability of local minimizers in some parametric settings different from [23]. Related growth conditions for tilt stability are also studied in [13].

Acknowledgments. The authors are indebted to the anonymous referees, to the handling Associate Editor Jong-Shi Pang, and also to Tran Nghia, Jiří Outrata, Héctor Ramírez, and Alexander Shapiro for their valuable remarks that allowed us to improve the original presentation.

REFERENCES

[1] S. Adly and J. V. Outrata, *Qualitative stability of a class of non-monotone variational inclusions. Applications in electronics*, J. Convex Anal., to appear.

[2] L. Ban, B. S. Mordukhovich, and W. Song, *Lipschitzian stability of parametric variational inequalities over generalized polyhedra in Banach spaces*, Nonlinear Anal., 74 (2011), pp. 441–461.

[3] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Ser. Oper. Res., Springer, New York, 2000.

[4] J. M. Borwein and Q. J. Zhu, *Techniques of Variational Analysis*, CMS Books Math. 20, Springer, New York, 2005.

[5] N. H. Chieu, T. D. Chuong, J.-C. Yao, and N. D. Yen, *Characterizing convexity of a function by its Fréchet and limiting second-order subdifferentials*, Set-Valued Var. Anal., 19 (2011), pp. 75–96.

[6] N. H. Chieu and N. Q. Huy, *Second-order subdifferentials and convexity of real-valued functions*, Nonlinear Anal., 74 (2011), pp. 154–160.

[7] N. H. Chieu and N. T. Q. Trang, *Coderivatives and monotonicity of continuous mappings*, Taiwan. J. Math., 16 (2012), pp. 353–365.

[8] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Canad. Math. Soc. Ser. Monogr. Adv. Texts, John Wiley, New York, 1983.

[9] G. Colombo, R. Henrion, N. D. Hoang, and B. S. Mordukhovich, *Optimal control of the sweeping process*, Dynam. Cont. Disc. Impul. Syst. Ser. B, 19 (2012), pp. 117–159.

[10] C. Ding, D. Sun, and J. J. Ye, *First order optimality conditions for mathematical programs with semidefinite cone complementarity constraints*, Optimization-Online e-print 11/2820, 2010. Available online at www.optimization-online.org/DB_HTML/2010/11/2820.html.

[11] A. L. Dontchev and R. T. Rockafellar, *Characterizations of strong regularity for variational inequalities over polyhedral convex sets*, SIAM J. Optim., 6 (1996), pp. 1087–1105.

[12] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings: A View from Variational Analysis*, Springer Monogr. in Math., Springer, Dordrecht, 2009.

[13] A. C. Eberhard and R. Wenczel, *A study of tilt-stable optimality and sufficient conditions*, Nonlinear Anal., 75 (2012), pp. 1240–1281.

[14] R. Henrion, B. S. Mordukhovich, and N. M. Nam, *Second-order analysis of polyhedral systems in finite and infinite dimensions with applications to robust stability of variational inequalities*, SIAM J. Optim., 20 (2010), pp. 2199–2227.

[15] R. Henrion, J. V. Outrata, and T. Surowiec, *On the coderivative of normal cone mappings to inequality systems*, Nonlinear Anal., 71 (2009), pp. 1213–1226.

[16] R. Henrion, J. V. Outrata, and T. Surowiec, *Analysis of M-stationary points to an EPEC modeling oligopolistic competition in an electricity spot market*, ESAIM Contr. Optim. Calc. Var., to appear.

[17] R. Henrion and W. Römisch, *On M-stationary points for a stochastic equilibrium problem under equilibrium constraints in electricity spot market modeling*, Appl. Math., 52 (2007), pp. 473–494.

[18] N. Q. Huy and J.-C. Yao, *Exact formulae for coderivatives of normal cone mappings to perturbed polyhedral convex sets*, J. Optim. Theory Appl., to appear.

[19] V. Jeyakumar and D. T. Luc, *Nonsmooth Vector Functions and Continuous Optimization*, Optim. Appl., 10, Springer, New York, 2008.

[20] A. B. Levy and B. S. Mordukhovich, *Coderivatives in parametric optimization*, Math. Program., 99 (2004), pp. 311–327.

[21] A. B. Levy, R. A. Poliquin, and R. T. Rockafellar, *Stability of locally optimal solutions*, SIAM J. Optim., 10 (2000), pp. 580–604.

[22] A. B. Levy and R. T. Rockafellar, *Variational conditions and the proto-differentiation of partial subgradient mappings*, Nonlinear Anal., 26 (1996), pp. 1951–1964.
[23] A. S. Lewis and S. Zhang, *Partial smoothness, tilt stability, and generalized Hessians*, Optimization-online e-print, 2011. Available online at www.optimization-online.org/DB_HTML/2011/10/3209.html.
[24] B. S. Mordukhovich, *Maximum principle in problems of time optimal control with nonsmooth constraints*, J. Appl. Math. Mech., 40 (1976), pp. 960–969.
[25] B. S. Mordukhovich, *Metric approximations and necessary optimality conditions for general classes of extremal problems*, Soviet Math. Dokl., 22 (1980), pp. 526–530.
[26] B. S. Mordukhovich, *Sensitivity analysis in nonsmooth optimization*, in Theoretical Aspects of Industrial Design, D. A. Field and V. Komkov, eds., Proc. Appl. Math., 58, SIAM, Philadelphia, 1992, pp. 32–46.
[27] B. S. Mordukhovich, *Complete characterizations of openness, metric regularity, and Lipschitzian properties of multifunctions*, Trans. Amer. Math. Soc., 340 (1993), pp. 1–35.
[28] B. S. Mordukhovich, *Generalized differential calculus for nonsmooth and set-valued mappings*, J. Math. Anal. Appl., 183 (1994), pp. 250–288.
[29] B. S. Mordukhovich, *Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis*, Trans. Amer. Math. Soc., 343 (1994), pp. 609–658.
[30] B. S. Mordukhovich, *Calculus of second-order subdifferentials in infinite dimensions*, Control Cybernet., 31 (2002), pp. 557–573.
[31] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation,* I: *Basic Theory*, Grundlehren Math. Wiss. 330, Springer, Berlin, 2006.
[32] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation,* II: *Applications*, Grundlehren math. Wiss. 331, Springer, Berlin, 2006.
[33] B. S. Mordukhovich and J. V. Outrata, *On second-order subdifferentials and their applications*, SIAM J. Optim., 12 (2001), pp. 139–169.
[34] B. S. Mordukhovich and J. V. Outrata, *Coderivative analysis of quasi-variational inequalities with applications to stability and optimization*, SIAM J. Optim., 18 (2007), pp. 389–412.
[35] B. S. Mordukhovich and Y. Shao, *Nonsmooth sequential analysis in Asplund spaces*, Trans. Amer. Math. Soc., 348 (1996), pp. 1235–1280.
[36] B. S. Mordukhovich and B. Wang, *Restrictive metric regularity and generalized differential calculus in Banach spaces*, Int. J. Math. Math. Sci., 50 (2004), pp. 2650–2683.
[37] N. M. Nam, *Coderivatives of the normal cone mappings and Lipschitzian stability of parametric variational inequalities*, Nonlinear Anal., 73 (2010), pp. 2271–2282.
[38] J. V. Outrata, *Mathematical programs with equilibrium constraints: Theory and numerical methods*, in Nonsmooth Mechanics of Solids, CISM Lecture Notes, 485, J. Haslinger and G. E. Stavroulakis, eds., Springer, New York, 2006, pp. 221–274.
[39] J. V. Outrata and H. Ramírez C., *On the Aubin property of critical points to perturbed second-order cone programs*, SIAM J. Optim., 21 (2011), pp. 798–823.
[40] J. V. Outrata and D. Sun, *On the coderivative of the projection operator onto the second-order cone*, Set-Valued Anal., 16 (2008), pp. 999–1014.
[41] R. A. Poliquin and R. T. Rockafellar, *Amenable functions in optimization*, in Nonsmooth Optimization Methods and Applications, F. Giannessi, ed., Gordon and Breach, Philadelphia, 1992, pp. 338–353.
[42] R. A. Poliquin and R. T. Rockafellar, *Prox-regular functions in variational analysis*, Trans. Amer. Math. Soc., 348 (1996), pp. 1805–1838.
[43] R. A. Poliquin and R. T. Rockafellar, *Tilt stability of a local minimum*, SIAM J. Optim., 8 (1998), pp. 287–299.
[44] N. T. Qui, *Nonlinear perturbations of polyhedral normal cone mappings and affine variational inequalities*, J. Optim. Theory Appl., 153 (2012), pp. 98–122.
[45] S. M. Robinson, *Strongly regular generalized equations*, Math. Oper. Res., 5 (1980), pp. 43–62.
[46] S. M. Robinson, *Some continuity properties of polyhedral multifunctions*, Math. Programming Stud., 14 (1981), pp. 206–214.
[47] R. T. Rockafellar, *Maximal monotone relations and the second derivatives of nonsmooth functions*, Ann. Inst. H. Poincaré: Anal. Non Linéaire, 2 (1985), pp. 167–184.
[48] R. T. Rockafellar, *Extended nonlinear programming*, in Nonlinear Optimization and Related Topics, G. Di Pillo and F. Giannessi, eds., Appl. Optim., 36, Kluwer Academic Publishers, Dordrecht, 2000, pp. 381–399.
[49] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Grundlehren Math. Wiss. 317, Springer, Berlin, 2006.
[50] W. Schirotzek, *Nonsmooth Analysis*, Universitext, Springer, Berlin, 2007.

[51] T. Surowiec, *Explicit Stationarity Conditions and Solution Characterization for Equilibrium Problems with Equilibrium Constraints*, Ph.D. Dissertation, Humboldt University, Berlin, 2010.

[52] J.-C. Yao and N. D. Yen, *Coderivative calculation related to a parametric affine variational inequality. Part* 1: *Basic calculation*, Acta Math. Vietnam., 34 (2009), pp. 157–172.

[53] J.-C. Yao and N. D. Yen, *Coderivative calculation related to a parametric affine variational inequality. Part* 2: *Applications*, Pac. J. Optim., 5 (2009), pp. 493–506.

[54] J. J. Ye, *Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints*, SIAM J. Optim., 10 (2000), pp. 943–962.

[55] R. Zhang, *Problems of hierarchical optimization in finite dimensions*, SIAM J. Optim., 4 (1994), pp. 521–536.