SIAM J. OPTIM.
Vol. 33, No. 4, pp. 2625–2653

SECOND-ORDER OPTIMALITY CONDITIONS FOR GENERAL NONCONVEX OPTIMIZATION PROBLEMS AND VARIATIONAL ANALYSIS OF DISJUNCTIVE SYSTEMS*

MATÚŠ BENKO†, HELMUT GFRERER‡, JANE J. YE§, JIN ZHANG¶, AND JINCHUAN ZHOU‖

Abstract. In this paper, we propose second-order sufficient optimality conditions for a very general nonconvex constrained optimization problem, which covers many prominent mathematical programs. Unlike the existing results in the literature, our conditions prove to be sufficient, for an essential local minimizer of second order, under merely basic smoothness and closedness assumptions on the data defining the problem. In the second part, we propose a comprehensive first- and second-order variational analysis of disjunctive systems and demonstrate how the second-order objects appearing in the optimality conditions can be effectively computed in this case.

Key words. second-order variational analysis, second-order optimality conditions, essential local minimizer of second order, second subderivative, second-order tangent sets, lower generalized support function, disjunctive system

MSC codes. 49J53, 49J52, 90C26, 90C4

DOI. 10.1137/22M1484742

1. Introduction.

For decades, variational analysis has been recognized as an important tool for studying optimization problems; we refer the reader to the standard monographs [5, 8, 24, 30, 31, 41]. Recently, second-order variational analysis has been developed rapidly; see [14, 16, 19, 29, 39, 40] and the references therein.

In this paper, we will deal with some special aspects of second-order variational analysis, namely, second-order optimality conditions for an optimization problem in the form

$$\text{(GP)} \qquad \min \; f(x) \quad \text{s.t.} \quad g(x) \in C.$$

Here $C \subset \mathbb{R}^m$ is a closed set, and $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are twice continuously differentiable functions unless otherwise specified. This general model covers

*Received by the editors March 16, 2022; accepted for publication (in revised form) June 30, 2023; published electronically October 12, 2023.
https://doi.org/10.1137/22M1484742
Funding: The first author's research was supported by the Austrian Science Fund (FWF) under grant P32832-N. The research of the third author was partially supported by NSERC. The fourth author's work was supported by the National Natural Science Foundation of China (12222106) and the Guangdong Basic and Applied Basic Research Foundation (2022B1515020082). The fifth author's work was supported by National Natural Science Foundation of China (12371305), the Young Innovation Teams of Shandong Province (2019KJI013), and the Shandong Province Natural Science Foundation (ZR2023MA020, ZR2021MA066).

†Applied Mathematics and Optimization, University of Vienna, 1090 Vienna, Austria (matus.benko@univie.ac.at).
‡Institute of Computational Mathematics, Johannes Kepler University Linz, 4040 Linz, Austria (helmut.gfrerer@jku.at).
§Department of Mathematics and Statistics, University of Victoria, Victoria V8W 2Y2, BC, Canada (janeye@uvic.ca).
¶Department of Mathematics, Southern University of Science and Technology, National Center for Applied Mathematics Shenzhen, Peng Cheng Laboratory, Shenzhen 518055, People's Republic of China (zhangj9@sustech.edu.cn).
‖Department of Statistics, School of Mathematics and Statistics, Shandong University of Technology, Zibo 255049, People's Republic of China (jinchuanzhou@163.com).

many common optimization problems, including the very challenging ones with constraints expressed via complementarity relations, in which case not only the feasible set $g^{-1}(C) := \{x \in \mathbb{R}^n \mid g(x) \in C\}$ but also the set C are nonconvex; see the comments below.

We concentrate on the development of tight second-order optimality conditions; i.e., the difference between the necessary and sufficient conditions should be small. Note that there are also other intrinsic issues of second-order conditions like stability of solutions or the convergence of numerical algorithms. For example, Rockafellar [39, 40] has demonstrated the importance of second-order variational analysis in numerical optimization, but these topics are far beyond the scope of this paper.

Let us now provide a brief discussion on existing results dealing with second-order optimality conditions, both necessary and sufficient. If C is convex polyhedral, as in the case of the standard nonlinear programs, second-order optimality conditions can be expressed via the second derivative of the Lagrangian. If C lacks polyhedrality, however, an additional term is needed to capture the curvature of C, and there are various tools that can be utilized for that purpose. When C is convex, a comprehensive analysis of second-order conditions is available in Bonnans and Shapiro [5, sections 3.2 and 3.3]. There, the second-order necessary conditions are derived within the framework of convex analysis and are of the following form (cf. [5, Theorem 3.45]): If a suitable constraint qualification holds at a local minimizer $\bar{x}$, then for every critical direction u and every convex subset $K(u)$ of the second-order tangent set $T_C^2(g(\bar{x}); \nabla g(\bar{x})u)$, there is a multiplier λ fulfilling first-order optimality conditions such that

$$\nabla_{xx}^2 L(\bar{x}, \lambda)(u, u) - \sigma_{K(u)}(\lambda) \geq 0. \tag{1.1}$$

Here, L denotes the Lagrangian, and σ is the support function. In particular, if the second-order tangent set $T_C^2(g(\bar{x}); \nabla g(\bar{x})u)$ is convex, we arrive at the condition

$$\nabla_{xx}^2 L(\bar{x}, \lambda)(u, u) - \sigma_{T_C^2(g(\bar{x}); \nabla g(\bar{x})u)}(\lambda) \geq 0. \tag{1.2}$$

By [5, Proposition 3.46], this condition is also necessary at a local minimizer, provided that the multiplier λ fulfilling the first-order optimality condition is unique, regardless whether or not $T_C^2(g(\bar{x}); \nabla g(\bar{x})u)$ is convex.

In the very recent paper by Gfrerer, Ye, and Zhou [19], optimality conditions have been stated for nonconvex C. In this case, one has to consider different types of first-order optimality conditions involving strong (S-), Mordukhovich (M-), and Clarke (C-) multipliers, respectively. Moreover, the feasible region may behave quite differently when moving away from the minimizer $\bar{x}$ in different directions. This fact motivates the use of different constraint qualifications and different types of multipliers when considering different critical directions.

When a directional nondegeneracy condition for the critical direction u is satisfied, ensuring that directional S-, M-, and C-multipliers coincide and are unique, condition (1.2) remains valid; see [19, Corollary 5]. When relaxing the directional nondegeneracy to the directional Robinson constraint qualification, one can still show that for every convex subset $K(u)$ of $T_C^2(g(\bar{x}); \nabla g(\bar{x})u)$, there is some directional C-multiplier λ satisfying (1.1); cf. [19, Corollary 4]. As it is shown in [19, Proposition 8], this is a very strong second-order necessary condition. However, it has the disadvantage that the directional C-multiplier λdepends not only on the critical direction u but also on the convex set $K(u)$. This can be remediated by the use of the so-called lower generalized support function $\hat{\sigma}$. It was shown in [19] that under the directional metric subregularity constraint qualification, which is weaker than the directional Robinson constraint qualification, there is a directional M-multiplier λ such that

$$\nabla^2_{xx}L(\bar{x},\lambda)(u,u) - \hat{\sigma}_{T^2_C(g(\bar{x});\nabla g(\bar{x})u)}(\lambda) \geq 0.$$

This function $\hat{\sigma}$ is indeed an extension of the support function, as for any closed set D, we have $\hat{\sigma}_D \leq \sigma_D$, and the equality holds when D is closed and convex.

Now let us consider the second-order sufficient conditions. Let $L^\alpha(x,\lambda) := \alpha f(x) + \langle \lambda, g(x)\rangle$. If, at a feasible point $\bar{x}$, for every critical direction u, the set C is *outer second-order regular* at $g(\bar{x})$ in direction $\nabla g(\bar{x})u$ and there are $\alpha \geq 0$ and $\lambda \in \mathbb{R}^m$ such that $\alpha \nabla f(\bar{x})u = 0$, $\nabla_x L^\alpha(\bar{x},\lambda) = 0$, and

$$\nabla^2_{xx}L^\alpha(\bar{x},\lambda)(u,u) - \sigma_{T^2_C(g(\bar{x});\nabla g(\bar{x})u)}(\lambda) > 0,$$

then the point $\bar{x}$ is a local minimizer fulfilling the so-called *quadratic growth condition.* In a slightly different form, this result was first proved in [5, Theorem 3.86] for convex sets C and then extended in [19, Theorem 4] to the nonconvex case. These sufficient conditions were essentially improved in the recent work by Mohammadi, Mordukhovich, and Sarabi [29, Proposition 7.3], where the assumption of outer second-order regularity is dropped and the sigma term $-\sigma_{T^2_C(g(\bar{x});\nabla g(\bar{x})u)}(\lambda)$ is replaced by the second-order subderivative $\mathrm{d}^2\delta_C(g(\bar{x});\lambda)(\nabla g(\bar{x})u)$ of the indicator function δ_C. One of the main results of this paper is an improvement of [29, Proposition 7.3] in that the set C is assumed to be neither convex nor parabolically derivable. Moreover, we do not need the existence of an S-multiplier, and we can choose different multipliers for every critical direction in order to fulfill the second-order sufficient condition. Finally, we not only prove the quadratic growth condition but also show that the point in question is an essential local minimizer of second order.

Summing up these considerations, we see that, besides the imposed constraint qualification, the second-order optimality conditions rely on the three second-order objects $\hat{\sigma}_{T^2_C(\bar{z};w)}(\lambda)$, $\sigma_{T^2_C(\bar{z};w)}(\lambda)$, and $\mathrm{d}^2\delta_C(\bar{z};\lambda)(w)$, each of them describing in some way the curvature of the set C and linked together by the inequalities

$$\mathrm{d}^2\delta_C(\bar{z};\lambda)(w) \leq -\sigma_{T^2_C(\bar{z};w)}(\lambda) \leq -\hat{\sigma}_{T^2_C(\bar{z};w)}(\lambda),$$

which are valid for any closed set C, every tangent $w \in T_C(\bar{z})$, and every λ with $\langle \lambda, w\rangle \geq 0$; see Proposition 2.18 below.

When studying optimization problems via some elaborate machinery of variational analysis, interesting as it may be, the important question remains: Can the employed tools be effectively computed or estimated and the obtained results suitably applied?

While with convex C we can formulate several standard programs as problem (GP), such as nonlinear programs, second-order cone programs, etc., we are primarily interested in programs modeled with nonconvex C. Such programs are considered very challenging but also increasingly important by the optimization community. Among others, they include the bilevel programs (see, e.g., Dempe [9] and Ye and Zhu [45]), programs with constraints governed by quasi-variational inequalities (see, e.g., Mordukhovich and Outrata [32]), and the mathematical program with second-order cone complementarity constraints (SOC-MPCC) (see, e.g., Outrata and Sun [34] and Ye and Zhou [43]). All of these problem classes can be modeled as a special case of problem (GP) with set C possessing the following structure:

$$C = \left\{ \left(y, b(y)^T\eta\right) \mid (q(y),\eta) \in \operatorname{gph} N_P \right\}, \tag{1.3}$$

where b and q are sufficiently smooth mappings (b maps into the space of matrices of an appropriate dimension, and in many cases, there holds that $b = \nabla q$), P is a convex polyhedral set, and N_P is the associated normal cone mapping. By taking $z := (y, \eta)$, $B(z_1, z_2) := (z_1, b(z_1)^T z_2)$, $G(z_1, z_2) := (q(z_1), z_2)$, and $D := \operatorname{gph} N_P$, the set C given by (1.3) can be represented as

$$C = B(\Gamma), \text{ where } \Gamma = \{z \mid G(z) \in D\}.$$

Since P is assumed to be convex polyhedral, the set D is polyhedral; i.e., D is the finite union of convex polyhedral sets.

In this paper, we will calculate $\hat{\sigma}_{T^2_\Gamma(\bar{z};w)}(\lambda)$, $\sigma_{T^2_\Gamma(\bar{z};w)}(\lambda)$, and $\mathrm{d}^2\delta_\Gamma(\bar{z};\lambda)(w)$ and defer the calculation of these three quantities for Γ replaced by the set C defined by (1.3) to a forthcoming paper by Benko et al. [4]. To accomplish this goal, in the second part of the paper, we provide a comprehensive first- and second-order variational analysis of the disjunctive system

$$\Gamma := G^{-1}(D) = \{x \in \mathbb{R}^n \mid G(x) \in D\},$$

where $G : \mathbb{R}^n \to \mathbb{R}^d$ is twice continuously differentiable and $D \subset \mathbb{R}^d$ is assumed to be polyhedral. The obtained results are also of independent interest and may be useful in other applications.

We organize our paper as follows. Section 2 contains the preliminaries and auxiliary results. In section 3, we derive the second-order sufficient optimality condition for the general program (GP). Sections 4 and 5 are devoted to the first- and the second-order variational analysis of the set $G^{-1}(D)$, respectively. Finally, in section 6, we demonstrate how to recover the second-order necessary and sufficient conditions from [12] for the disjunctive program by means of our results.

2. Preliminaries and auxiliary results. In this section, we recall some background material from variational analysis and provide some preliminary results. Let us begin with the notation. The open unit ball is denoted by $\mathbb{B}$, and the open ball centered at z with radius δ is denoted by $\mathbb{B}(z, \delta)$. For a set $S \subset \mathbb{R}^n$, denote by span S, cl S, and conv S its linear span, closure, and convex hull, respectively. We call a subspace $L \subset \mathbb{R}^n$ the *generalized lineality space* of S and denote it by $\mathcal{L}(S)$ provided that it is the largest subspace satisfying $S + L \subset S$. Since any linear subspace includes 0, we actually have $S + \mathcal{L}(S) = S$, and if S is a closed convex cone, we get $\mathcal{L}(S) = S \cap (-S)$. The indicator function $\delta_S : \mathbb{R}^n \to \bar{\mathbb{R}} := [-\infty, +\infty]$ of S is given as $\delta_S(z) = 0$ for $z \in S$ and $\delta_S(z) = +\infty$ if $z \notin S$. Next, if S is closed, let S° and $\sigma_S : \mathbb{R}^n \to \bar{\mathbb{R}}$ stand for the polar cone to S and the support function of S, respectively, i.e., $S^\circ := \{z^* \in \mathbb{R}^n \mid \langle z^*, z\rangle \le 0,\ \forall z \in S\}$ and $\sigma_S(z^*) := \sup\{\langle z^*, z\rangle \mid z \in S\}$ for $z^* \in \mathbb{R}^n$. For an extended function $\varphi : \mathbb{R}^n \to \bar{\mathbb{R}}$, we define its effective domain by $\operatorname{dom}\varphi := \{z \,|\, \varphi(z) < +\infty\}$. For $w \in \mathbb{R}^n$, denote by $\{w\}^\perp$ the orthogonal complement of the linear space generated by w. Let $o : \mathbb{R}_+ \to \mathbb{R}^n$ stand for a mapping with the property that $o(t)/t \to 0$ when $t \downarrow 0$. The symbol $z' \xrightarrow{S} z$ indicates that $z' \in S$ and $z' \to z$. For a mapping $\psi : \mathbb{R}^n \to \mathbb{R}^d$ and $z \in \mathbb{R}^n$, we denote by $\nabla\psi(z) \in \mathbb{R}^{d\times n}$ its Jacobian at z and by $\nabla^2\psi(z)$ its second derivative at z as defined by

$$w^T\nabla^2\psi(z) := \lim_{t\to 0} \frac{\nabla\psi(z + tw) - \nabla\psi(z)}{t} \quad \forall w \in \mathbb{R}^n.$$

Hence,

$$\nabla^2\psi(z)(w, w) := w^T\nabla^2\psi(z)w = (w^T\nabla^2\psi_1(z)w, \ldots, w^T\nabla^2\psi_d(z)w)^T \quad \forall w \in \mathbb{R}^n.$$

2.1. Variational geometry. First we review the various classical concepts of tangent and normal cones.

DEFINITION 2.1 (tangent and normal cones [41]). *Given $S \subset \mathbb{R}^n$, $z \in S$, the regular/Clarke tangent cone and tangent/contingent cone to S at z are defined, respectively, by*

$$\widehat{T}_S(z) := \liminf_{\substack{z' \xrightarrow{S} z \\ t \downarrow 0}} \frac{S - z'}{t} = \left\{ w \in \mathbb{R}^n \,\middle|\, \forall\, t_k \downarrow 0,\, z_k \xrightarrow{S} z,\, \exists w_k \to w \ \text{ with } \ z_k + t_k w_k \in S \right\},$$

$$T_S(z) := \left\{ w \in \mathbb{R}^n \,\middle|\, \exists\, t_k \downarrow 0,\, w_k \to w \ \text{ with } \ z + t_k w_k \in S \right\}.$$

For $w \in T_S(z)$, the outer second-order tangent set to S at z in direction w is defined by

$$T_S^2(z;w) := \left\{ s \in \mathbb{R}^n \,\middle|\, \exists\, t_k \downarrow 0, s_k \to s \ \text{ with } \ z + t_k w + \frac{1}{2} t_k^2 s_k \in S \right\}.$$

The regular/Fréchet normal cone, the proximal normal cone, and the limiting/Mordukhovich normal cone to S at z are given, respectively, by

$$\widehat{N}_S(z) := \left\{ z^* \in \mathbb{R}^n \,\middle|\, \langle z^*, z' - z \rangle \le o\big(\|z' - z\|\big) \ \forall z' \in S \right\},$$
$$\widehat{N}_S^p(z) := \left\{ z^* \in \mathbb{R}^n \,\middle|\, \exists\, \gamma > 0 : \langle z^*, z' - z \rangle \le \gamma \|z' - z\|^2 \ \ \forall z' \in S \right\},$$
$$N_S(z) := \left\{ z^* \in \mathbb{R}^n \,\middle|\, \exists\, z_k \xrightarrow{S} z, z_k^* \to z^* \ \text{with} \ \ z_k^* \in \widehat{N}_S(z_k) \right\}.$$

Recall that for any set S, one always has $\widehat{N}_S(z) = T_S(z)^\circ$, and if S is closed, then $N_S(z)^\circ = \widehat{T}_S(z)$; cf. Rockafellar and Wets [41, Theorem 6.28].

Recently, motivated by the formula $\widehat{T}_S(z) = \liminf_{z' \xrightarrow{S} z} T_S(z')$ (cf. [41, Theorem 6.26]), a directional variant of the regular tangent cone has been introduced.

DEFINITION 2.2 (directional regular tangent cone [19, Definition 2]). *Given $S \subset \mathbb{R}^n$, $z \in S$, and $w \in \mathbb{R}^n$, the regular/Clarke tangent cone to S at z in direction w is defined by*

$$\begin{aligned}\widehat{T}_S(z;w) &:= \liminf_{\substack{t \downarrow 0, w' \to w \\ z + tw' \in S}} T_S(z + tw') \\ &= \left\{ v \in \mathbb{R}^n \,\middle|\, \forall t_k \downarrow 0, w_k \to w, z + t_k w_k \in S, \exists v_k \to v \ \text{with} \ \ v_k \in T_S(z + t_k w_k) \right\}.\end{aligned}$$

It is easy to see that $\widehat{T}_S(z;0) = \widehat{T}_S(z)$. Similar to the regular tangent cone, the directional regular tangent cone $\widehat{T}_S(z;w)$ is a closed and convex cone; see [19, Proposition 3].

PROPOSITION 2.3 ([19, Proposition 1]). *Given a closed set $S \subset \mathbb{R}^n$, for every $z \in S$ and every $w \in T_S(z)$, one has*

$$T_{T_S(z)}(w) + \widehat{T}_S(z;w) = T_{T_S(z)}(w), \quad T_S^2(z;w) + \widehat{T}_S(z;w) = T_S^2(z;w).$$

DEFINITION 2.4 (directional normal cones [11, 19, 20]). *Given $S \subset \mathbb{R}^n$, $z \in S$ and a direction $w \in \mathbb{R}^n$, the limiting and the Clarke normal cone to S in direction w at z are given, respectively, by*

$$N_S(z;w) := \left\{ z^* \in \mathbb{R}^n \,\middle|\, \exists\, t_k \downarrow 0, w_k \to w, z_k^* \to z^* \ \text{with} \ \ z_k^* \in \widehat{N}_S(z + t_k w_k) \right\},$$
$$N_S^c(z;w) := \operatorname{cl}\operatorname{conv} N_S(z;w).$$

By [12, Lemma 2.1], when S is the union of finitely many convex polyhedral sets, we have that for any $w\in T_S(z)$,

$$N_S(z;w)\subset N_S(z)\cap\{w\}^{\perp}.$$

Moreover, if S is a closed convex set and $w\in T_S(z)$,

$$N_S(z;w)=N_{T_S(z)}(w)=N_S(z)\cap\{w\}^{\perp}. \tag{2.1}$$

PROPOSITION 2.5 (directional tangent-normal polarity [19, Proposition 3]). *For a closed set $S\subset\mathbb{R}^n$, $z\in S$, and $w\in\mathbb{R}^n$, one has*

$$\widehat{T}_S(z;w)=N_S(z;w)^{\circ}=N_S^c(z;w)^{\circ},\quad \widehat{T}_S(z;w)^{\circ}=N_S^c(z;w).$$

The definition of the lineality space $\mathcal{L}(S)$ readily yields

$$T_S(z+l)=T_S(z),\quad \widehat{N}_S(z+l)=\widehat{N}_S(z)\quad \forall z\in S,\quad \forall l\in\mathcal{L}(S). \tag{2.2}$$

By the previous proposition, we also get the following result.

PROPOSITION 2.6. *Let $S\subset\mathbb{R}^n$ be a closed set, $z\in S$, and $w\in T_S(z)$. Then*

$$\big(\operatorname{span}N_S(z;w)\big)^{\circ}=\big(\operatorname{span}N_S^c(z;w)\big)^{\circ}=\mathcal{L}(\widehat{T}_S(z;w))\subset\mathcal{L}(T_{T_S(z)}(w)). \tag{2.3}$$

Proof. Notice that

$$\begin{aligned}\big(\operatorname{span}N_S(z;w)\big)^{\circ}&=\big(\operatorname{span}N_S^c(z;w)\big)^{\circ}=\big(N_S^c(z;w)-N_S^c(z;w)\big)^{\circ}\\&=\big(N_S^c(z;w)\big)^{\circ}\cap-\big(N_S^c(z;w)\big)^{\circ}\\&=\mathcal{L}(N_S^c(z;w)^{\circ})=\mathcal{L}(\widehat{T}_S(z;w)),\end{aligned}$$

where the first equality holds obviously using the fact that the set $N_S^c(z;w)$ is a closed convex cone, the second equality follows from [38, Theorem 2.7], the third equality follows from the calculus rule for polar cones in [38, Corollary 16.4.2], the fourth equality hold by the fact that the set $N_S^c(z;w)^{\circ}$ is a closed convex cone, and the fifth equality holds by Proposition 2.5.

Proposition 2.3 yields $T_{T_S(z)}(w)+\widehat{T}_S(z;w)=T_{T_S(z)}(w)$, and since $\widehat{T}_S(z;w)$ is a closed convex cone by [19, Proposition 3], we have $\mathcal{L}(\widehat{T}_S(z;w))=\widehat{T}_S(z;w)\cap(-\widehat{T}_S(z;w))$. Thus,

$$T_{T_S(z)}(w)+\mathcal{L}(\widehat{T}_S(z;w))\subset T_{T_S(z)}(w)$$

holds as well, and the inclusion in (2.3) follows by the definition of the lineality space. ∎

2.2. Directional proximal normal cone. It turns out that we also need a directional version of the proximal normal cone; see Proposition 2.18. To this end, we need the following definition.

DEFINITION 2.7 (directional neighborhood [11]). *Let $w\in\mathbb{R}^n$. For $\delta,\rho>0$,*

$$V_{\delta,\rho}(w):=\big\{w'\in\delta\mathbb{B}\,\big|\,\big\|\|w\|w'-\|w'\|w\big\|\le\rho\|w'\|\|w\|\,\big\}$$

is called a directional neighborhood of direction w.

It is easy to see that $V_{\delta,\rho}(w) \subset V_{\delta,\rho}(0) = \delta\mathbb{B}$. Hence, the directional neighborhood is in general smaller than the classical neighborhood.

Recall that the proximal normal cone to a closed set S at a point $z \in S$ can be equivalently given by

$$\widehat{N}^p_S(z) := \{z^* \in \mathbb{R}^n \mid \exists\ \delta, \gamma > 0 : \langle z^*, z' - z\rangle \le \gamma\|z' - z\|^2 \ \ \forall z' \in S \cap \mathbb{B}(z, \delta)\};$$

see, e.g., [8, Proposition 1.5]. By replacing the standard neighborhood by the directional one, we arrive at the following directional version of the proximal normal cone.

Definition 2.8 (directional proximal normal cone). *Given a closed set $S \subset \mathbb{R}^n$, a point $z \in S$, and a direction $w \in T_S(z)$, we define the proximal prenormal cone to S in direction w at z as*

$$\hat{\mathcal{N}}^p_S(z; w) := \{z^* \in \mathbb{R}^n \mid \exists\ \delta, \rho, \gamma > 0 : \langle z^*, z' - z\rangle \le \gamma\|z' - z\|^2 \ \ \forall z' \in S \cap (z + V_{\delta,\rho}(w))\}$$

and the proximal normal cone to S at z in direction w as

$$\widehat{N}^p_S(z; w) := \hat{\mathcal{N}}^p_S(z; w) \cap \{w\}^\perp.$$

In the case when $w \notin T_S(z)$, we set $\hat{\mathcal{N}}^p_S(z; w) := \widehat{N}^p_S(z; w) := \emptyset$.

From the definition, we can see that the proximal prenormal cone is in general larger than the classical proximal normal cone, i.e., $\widehat{N}^p_S(z) \subset \hat{\mathcal{N}}^p_S(z; w)$, for any $w \in T_S(z)$. Moreover, from the definition, any vector z^* satisfying $\langle z^*, w\rangle < 0$ is always included in $\hat{\mathcal{N}}^p_S(z; w)$. In fact, we have

$$\{z^* | \langle z^*, w\rangle < 0\} \subset \hat{\mathcal{N}}^p_S(z; w) \subset \{z^* | \langle z^*, w\rangle \le 0\}. \tag{2.4}$$

Since the vectors z^* satisfying $\langle z^*, w\rangle < 0$ do not provide much useful information, it is natural to restrict the directional proximal prenormals by intersecting with the orthogonal complement of w. This restriction yields the concept of directional proximal normal cone and ensures that the directional proximal normal is contained in the directional limiting normal cone; see Proposition 2.9. In particular, when S is a closed convex set, combining (2.1) and (2.6) below, we get

$$\widehat{N}^p_S(z; w) = N_S(z; w) = N_S(z) \cap \{w\}^\perp = N_{T_S(z)}(w). \tag{2.5}$$

In the following proposition, we show convexity of the directional proximal normal cone and compare it with other normal cones.

Proposition 2.9. *Let $S \subset \mathbb{R}^n$ be closed, and let $w \in T_S(z)$ be given. Then both $\hat{\mathcal{N}}^p_S(z; w)$ and $\widehat{N}^p_S(z; w)$ are convex cones and*

$$\widehat{N}^p_S(z) \cap \{w\}^\perp \subset \widehat{N}^p_S(z; w) \subset \widehat{N}_{T_S(z)}(w) \subset N_{T_S(z)}(w) \subset N_S(z; w). \tag{2.6}$$

Proof. By definition, it is easy to show that $\hat{\mathcal{N}}^p_S(z; w)$ is a convex cone. Thus, $\widehat{N}^p_S(z; w)$ is also a convex cone as the intersection of two convex cones.

The first inclusion in (2.6) follows immediately from $\widehat{N}^p_S(z) \subset \hat{\mathcal{N}}^p_S(z; w)$, the third one is trivial, and the last one was proved in [19, Lemma 3]. Thus, it remains to show the second inclusion.

Since

$$\widehat{N}^p_S(z; 0) = \widehat{N}^p_S(z) \subset \widehat{N}_S(z) = \widehat{N}_{T_S(z)}(0),$$

where the last equation follows from [18, equation (3)], the inclusion holds true for $w=0$. Now let $w\neq 0$, and consider $z^*\in \widehat{N}^p_S(z;w)$. We wish to show that

$$z^*\in (T_{T_S(z)}(w))^\circ = \widehat{N}_{T_S(z)}(w). \tag{2.7}$$

By definition, we can find some $\delta>0,\gamma>0$ such that

$$\langle z^*, z'-z\rangle \le \gamma\|z'-z\|^2 \quad \forall z'\in (z+V_{\delta,\delta}(w))\cap S. \tag{2.8}$$

To show (2.7), we pick $v\in T_{T_S(z)}(w)$ together with sequences $t_k\downarrow 0$ and $v_k\to v$ satisfying $w+t_kv_k\in T_S(z)$. For every k, there exist sequences $\tau_j^k\downarrow 0$ and $s_j^k\to 0$ as $j\to\infty$ satisfying $z+\tau_j^k(w+t_kv_k+s_j^k)\in S\forall j$. For all k sufficiently large, we have $\left\|\frac{w+t_kv_k}{\|w+t_kv_k\|}-\frac{w}{\|w\|}\right\|<\frac{\delta}{2}$, and we can find an index $j(k)$ such that

$$\tau^k_{j(k)}<\frac{1}{k}t_k,\ \|s^k_{j(k)}\|<\frac{1}{k}t_k,\ \left\|\frac{w+t_kv_k+s^k_{j(k)}}{\left\|w+t_kv_k+s^k_{j(k)}\right\|}-\frac{w+t_kv_k}{\|w+t_kv_k\|}\right\|<\frac{\delta}{2},$$
$$\tau^k_{j(k)}\|w+t_kv_k+s^k_{j(k)}\|<\delta.$$

It follows by Definition 2.7 that $z+\tau^k_{j(k)}(w+t_kv_k+s^k_{j(k)})\in(z+V_{\delta,\delta}(w))\cap S$, and together with $\langle z^*,w\rangle=0$, we obtain from (2.8) that

$$\begin{aligned}\tau^k_{j(k)}t_k\left\langle z^*, v_k+\frac{s^k_{j(k)}}{t_k}\right\rangle &= \left\langle z^*, \left(z+\tau^k_{j(k)}\left(w+t_kv_k+s^k_{j(k)}\right)\right)-z\right\rangle\\ &\le \gamma\left(\tau^k_{j(k)}\right)^2\left\|w+t_kv_k+s^k_{j(k)}\right\|^2.\end{aligned}$$

Dividing this inequality by $\tau^k_{j(k)}t_k$, we conclude that

$$\langle z^*,v\rangle=\lim_{k\to\infty}\left\langle z^*, v_k+\frac{s^k_{j(k)}}{t_k}\right\rangle\le\lim_{k\to\infty}\gamma\frac{\tau^k_{j(k)}}{t_k}\left\|w+t_kv_k+s^k_{j(k)}\right\|^2=0.$$

Hence, (2.7) holds, and therefore the second inclusion in (2.6) follows. □

2.3. Polyhedral sets. Next, we provide formulas for tangents and normals to polyhedral sets. A set $D\subset\mathbb{R}^d$ is said to be *convex polyhedral* if it is the intersection of finitely many half-spaces, whereas it is said to be *polyhedral* whenever it is the union of finitely many convex polyhedral sets.

Polyhedral sets enjoy the following important property; see also [41, Exercise 6.47].

PROPOSITION 2.10 (exactness of tangential approximations [24, Proposition 8.24]). *If D is polyhedral and $z\in D$, then there is an open neighborhood W of* 0 *such that*

$$(D-z)\cap W = T_D(z)\cap W,$$

or, equivalently,

$$D\cap(z+W)=(z+T_D(z))\cap(z+W). \tag{2.9}$$

Equation (2.10) in the result below extends [41, Proposition 13.13] from convex polyhedral sets to polyhedral sets.

PROPOSITION 2.11. *Let D be a polyhedral set, $z \in D$, and $w \in T_D(z)$. Then*

$$T_D^2(z;w) = T_{T_D(z)}(w), \tag{2.10}$$

$$\widehat{N}_{T_D(z)}(w) = \widehat{N}_{T_D^2(z;w)}(0) = (T_D^2(z;w))^\circ, \tag{2.11}$$

$$N_D(z;w) = N_{T_D(z)}(w) = N_{T_D^2(z;w)}(0). \tag{2.12}$$

Proof. Let $D := \cup_{i=1}^s D_i$, where each $D_i (i = 1, \ldots, s)$ is convex polyhedral and $z \in D$. By (2.9), we get

$$T_D(z') = T_{z+T_D(z)}(z') = T_{T_D(z)}(z' - z) \quad \forall z' \in D \cap (z + W), \tag{2.13}$$

where W is an open neighborhood of 0. Consider a tangent direction $w \in T_D(z)$. Then, by [41, Proposition 13.13], we have $T_{D_i}^2(z;w) = T_{T_{D_i}(z)}(w)$ whenever $z \in D_i$ and $w \in T_{D_i}(z)$. Since we have $T_{D_i}^2(z;w) = T_{T_{D_i}(z)}(w) = \emptyset$ for the remaining i by definition, we obtain (2.10) by

$$T_D^2(z;w) = \bigcup_{i=1}^s T_{D_i}^2(z;w) = \bigcup_{i=1}^s T_{T_{D_i}(z)}(w) = T_{T_D(z)}(w), \tag{2.14}$$

where the first and third equations are due to [5, Proposition 3.37]. Polarization of both sides yields $(T_D^2(z;w))^\circ = (T_{T_D(z)}(w))^\circ = \widehat{N}_{T_D(z)}(w)$. Since $T_D^2(z;w) = T_{T_D(z)}(w)$, we have

$$\widehat{N}_{T_D^2(z;w)}(0) = \widehat{N}_{T_{T_D(z)}(w)}(0) = \widehat{N}_{T_D(z)}(w),$$

where the last equality follows from the fact that $T_D(z)$ is a closed cone; see, e.g., [18, equation (3)]. Hence, (2.11) holds. It remains to show (2.12). For all z' sufficiently close to z, we have $\widehat{N}_D(z') = \widehat{N}_{T_D(z)}(z' - z)$ by virtue of (2.13). Hence, for every $w \in T_D(z)$, we have

$$\begin{aligned} N_D(z;w) &= \{z^* \mid \exists t_k \downarrow 0, w_k \to w, z_k^* \to z^* \text{ with } z_k^* \in \widehat{N}_D(z + t_k w_k) = \widehat{N}_{T_D(z)}(w_k)\} \\ &= N_{T_D(z)}(w). \end{aligned} \tag{2.15}$$

For $w = 0$, we particularly have $N_D(z) = N_{T_D(z)}(0)$. Since $T_D(z)$ is also polyhedral, the same formula applies, and taking into account (2.14), we get $N_{T_D(z)}(w) = N_{T_{T_D(z)}(w)}(0) = N_{T_D^2(z;w)}(0)$. Combining this equation with (2.15), we obtain (2.12). ∎

2.4. Variational geometry of constraint systems under metric subregularity. Let us mention some basic facts about the tangents and the normals to a set S described by constraints as $S := g^{-1}(C) = \{x \in \mathbb{R}^n \mid g(x) \in C\}$, where $g : \mathbb{R}^n \to \mathbb{R}^d$ and $C \subset \mathbb{R}^d$. We will need to use the following concept of directional metric subregularity, which we introduce only in the special case of constraint mappings.

DEFINITION 2.12 (directional metric subregularity [11, Definition 1]). *Let $g : \mathbb{R}^n \to \mathbb{R}^d$, $C \subset \mathbb{R}^d$, and $\bar{x} \in S := g^{-1}(C)$. We say that the set-valued constraint mapping $x \rightrightarrows g(x) - C$ is metrically subregular at $(\bar{x}, 0)$ in direction $u \in \mathbb{R}^n$ or that the metric subregularity constraint qualification (MSCQ) holds at $\bar{x}$ in direction u if there exist $\kappa, \delta, \rho > 0$ such that*

$$\operatorname{dist}(x, S) \le \kappa \operatorname{dist}(g(x), C) \quad \forall x \in \bar{x} + V_{\delta,\rho}(u). \tag{2.16}$$

The infimum of all κ for which there are $\delta, \rho > 0$ satisfying (2.16) *is called the subregularity modulus. In the case $u = 0$, we simply say that the constraint mapping is metrically subregular at $(\bar{x}, 0)$ or that MSCQ holds at $\bar{x}$.*

If g is continuously differentiable, then by [12, Theorem 2.6], a sufficient condition for MSCQ at $\bar{x}$ in direction u is the condition

$$\nabla g(\bar{x})^T y^* = 0,\ y^* \in N_C(g(\bar{x}); \nabla g(\bar{x})u) \implies y^* = 0. \tag{2.17}$$

Asking (2.17) to be satisfied for all nonzero $u \in \mathbb{R}^n$ corresponds to the so-called first-order sufficient condition for metric subregularity (FOSCMS), which implies MSCQ at $\bar{x}$. If in addition the graph of the constraint mapping is a closed cone, then the metric subregularity holds locally if and only if it holds globally.

PROPOSITION 2.13 ([13, Lemma 3]). *Let $g : \mathbb{R}^n \to \mathbb{R}^d$, $C \subset \mathbb{R}^d$, and assume that $\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^d \mid g(x) - y \in C\}$, the graph of the constraint mapping $x \rightrightarrows g(x) - C$, is a closed cone. Then $0 \in S$, and if MSCQ holds at 0, then there is some $\kappa > 0$ such that* (2.16) *holds for all x.*

In the following proposition, we collect the basic results about tangents, normals, and second-order tangents to set S.

PROPOSITION 2.14. *Let $g : \mathbb{R}^n \to \mathbb{R}^d$ be continuously differentiable, let $C \subset \mathbb{R}^d$ be a closed set, and consider $\bar{x} \in S := g^{-1}(C)$. Suppose that the constraint mapping $x \rightrightarrows g(x) - C$ is metrically subregular at $(\bar{x}, 0)$ in direction $\bar{u} \in \mathbb{R}^n$. Then there is a neighborhood U of $\bar{u}$ such that for every $u \in U$, one has*

$$T_S(\bar{x}) \cap U = \nabla g(\bar{x})^{-1}\big(T_C(g(\bar{x}))\big) \cap U, \quad T_{T_S(\bar{x})}(u) = \nabla g(\bar{x})^{-1}\big(T_{T_C(g(\bar{x}))}(\nabla g(\bar{x})u)\big), \tag{2.18a}$$

$$N_S(\bar{x}; u) \subset \nabla g(\bar{x})^T N_C(g(\bar{x}); \nabla g(\bar{x})u), \quad N_{T_S(\bar{x})}(u) \subset \nabla g(\bar{x})^T N_{T_C(g(\bar{x}))}(\nabla g(\bar{x})u). \tag{2.18b}$$

Additionally, if g is twice continuously differentiable and $u \in T_S(\bar{x}) \cap U$, one has

$$T_S^2(\bar{x}; u) = \{p \in \mathbb{R}^n \mid \nabla g(\bar{x})p + \nabla^2 g(\bar{x})(u, u) \in T_C^2(g(\bar{x}); \nabla g(\bar{x})u)\}$$

and, denoting the subregularity modulus by κ,

$$\operatorname{dist}(p, T_S^2(\bar{x}; u)) \le \kappa \operatorname{dist}(\nabla g(\bar{x})p + \nabla^2 g(\bar{x})(u, u), T_C^2(g(\bar{x}); \nabla g(\bar{x})u)) \quad \forall p \in \mathbb{R}^n.$$

Moreover, if there exists a subspace $L \subset \mathbb{R}^d$ such that

$$T_{T_C(g(\bar{x}))}(\nabla g(\bar{x})u) + L \subset T_{T_C(g(\bar{x}))}(\nabla g(\bar{x})u) \ \text{ and } \ \nabla g(\bar{x})\mathbb{R}^n + L = \mathbb{R}^d, \tag{2.19}$$

then

$$\widehat{N}_{T_S(\bar{x})}(u) = \nabla g(\bar{x})^T \widehat{N}_{T_C(g(\bar{x}))}(\nabla g(\bar{x})u).$$

The subregularity assumption as well as the existence of the subspace L satisfying (2.19) *with $u = \bar{u}$ are fulfilled particularly under the following directional nondegeneracy condition:*

$$\nabla g(\bar{x})^T y^* = 0,\ y^* \in \operatorname{span} N_C(g(\bar{x}); \nabla g(\bar{x})\bar{u}) \implies y^* = 0. \tag{2.20}$$

Proof. By Definition 2.12, there exists a neighborhood U of $\bar{u}$ such that $x \rightrightarrows g(x)-C$ is metrically subregular at $\bar{x}$ in every direction $u \in U$ with the same modulus. Thus, the estimate for the directional limiting normal cone comes from [2, Theorem 3.1]. In [19, Proposition 5], one can find all the statements regarding the second-order tangents as well as the first formula in (2.18a), which means that, locally around any $u \in U$, the set $T_S(\bar{x})$ has the same preimage structure as S. By [19, Lemma 1], however, we infer that the corresponding constraint mapping $u' \rightrightarrows \nabla g(\bar{x})u' - T_C(g(\bar{x}))$ is metrically subregular at $(u,0)$, and so the remaining two estimates for $T_{T_S(\bar{x})}(u)$ and $N_{T_S(\bar{x})}(u)$ are results of the standard, nondirectional calculus. Moreover, the formula for the regular normal cone is from [16, Theorem 4].

Note that the nondegeneracy condition (2.20) is clearly stronger than FOSCMS, so it obviously implies MSCQ at $\bar{x}$ in direction $\bar{u}$.

Let us now show that the subspace $\mathcal{L}(\widehat{T}_C(g(\bar{x}); \nabla g(\bar{x})\bar{u}))$ satisfies (2.19) with $u = \bar{u}$. The first property follows immediately from Proposition 2.6. By the nondegeneracy, we get

$$\mathbb{R}^d = \left(\ker \nabla g(\bar{x})^T \cap \operatorname{span} N_C(g(\bar{x}); \nabla g(\bar{x})\bar{u})\right)^{\perp} = \nabla g(\bar{x})\mathbb{R}^n + \mathcal{L}(\widehat{T}_C(g(\bar{x}); \nabla g(\bar{x})\bar{u})),$$

and (2.19) follows. $\square$

For more information about the directional nondegeneracy (2.20), we the reader refer to [3, section 2.4], where this condition was first introduced for convex polyhedral set C. Particularly, [3, Example 2.15] clarifies that for a nonzero direction, it is a strictly milder assumption than the standard nondegeneracy [5, Formula 4.172], which corresponds to the case $\bar{u} = 0$. We will further utilize directional nondegeneracy in sections 4 and 5 in the case of polyhedral set C, showing that under (2.20), all the four sets in (2.18b) actually coincide (see Theorem 4.1) and certain directional multipliers are unique (see Corollary 5.8).

2.5. Generalized support function and second subderivative.

In this final preliminary part, we recall the definitions of the lower generalized support function and state some basic properties.

DEFINITION 2.15 (lower generalized support function [19]). *Given a nonempty closed set $S \subset \mathbb{R}^n$, we define the lower generalized support function to S as the mapping $\hat{\sigma}_S : \mathbb{R}^n \to \bar{\mathbb{R}}$ by*

$$\hat{\sigma}_S(z^*) := \liminf_{\tilde{z}^* \to z^*} \inf_z \{\langle \tilde{z}^*, z\rangle \mid \tilde{z}^* \in N_S(z)\} = \liminf_{\tilde{z}^* \to z^*} \inf_z \{\langle \tilde{z}^*, z\rangle \mid \tilde{z}^* \in \widehat{N}_S(z)\} \quad \forall z^* \in \mathbb{R}^n.$$

If $S = \emptyset$, then we define $\hat{\sigma}_S(z^*) := -\infty$ for all z^*.

It was shown in [19] that, in general, $\hat{\sigma}_S(z^*) \le \sigma_S(z^*)$ for all z^*, and the equality holds when S is convex. If $S = z + K$ is a translation of a cone K, we get the following formula, which will come in handy in section 5.

PROPOSITION 2.16. *For every nonempty closed cone $K \subset \mathbb{R}^n$ (not necessarily convex) and every $z \in \mathbb{R}^n$, we have*

$$\hat{\sigma}_{z+K}(z^*) = \begin{cases} \langle z^*, z\rangle & \text{if } z^* \in N_K(0), \\ \infty & \textit{otherwise}. \end{cases}$$

Particularly, $\operatorname{dom} \hat{\sigma}_{z+K} = N_K(0)$.

Proof. First, note that since K is assumed to be a cone, for all $q \in K$, we have

$$\widehat{N}_{z+K}(z+q) = \widehat{N}_K(q) = \widehat{N}_K(\alpha q) = \widehat{N}_{z+K}(z+\alpha q) \quad \forall \alpha > 0, \tag{2.21a}$$

$$\langle z^*, q\rangle = 0 \quad \forall z^* \in \widehat{N}_K(q). \tag{2.21b}$$

We shall show that $\hat{\sigma}_{z+K}(z^*) < \infty$ if and only if $z^* \in N_K(0)$, and in this case, we have $\hat{\sigma}_{z+K}(z^*) = \langle z^*, z\rangle$. If $\hat{\sigma}_{z+K}(z^*) < \infty$, then there exist sequences $z_k^* \to z^*$ and $q_k \in K$ with $z_k^* \in \widehat{N}_{z+K}(z+q_k)$ for all k such that $\hat{\sigma}_{z+K}(z^*) = \lim_{k\to\infty}\langle z_k^*, z+q_k\rangle$. By (2.21a), we have $z_k^* \in \widehat{N}_{z+K}(z+q_k) = \widehat{N}_K(q_k)$ and hence $\langle z_k^*, q_k\rangle = 0$ by (2.21b). It follows that

$$\hat{\sigma}_{z+K}(z^*) = \lim_{k\to\infty}\langle z_k^*, z\rangle = \langle z^*, z\rangle.$$

Moreover, by (2.21a), we have $z_k^* \in \widehat{N}_K(\alpha_k q_k)$ for $\alpha_k := 1/(k(\|q_k\|+1))$. Taking the limit as k goes to ∞, we obtain $z^* \in N_K(0)$. Conversely, let $z^* \in N_K(0)$, and consider sequences $q_k \in K$ and $z_k^* \in \widehat{N}_K(q_k)$ such that $q_k \to 0$ and $z_k^* \to z^*$. Then $z_k^* \in \widehat{N}_K(q_k) = \widehat{N}_{z+K}(z+q_k)$ by (2.21a) and $\langle z_k^*, q_k\rangle = 0$ by (2.21b). Hence, by Definition 2.15, we obtain that

$$\hat{\sigma}_{z+K}(z^*) \le \liminf_{k\to\infty}\langle z_k^*, z+q_k\rangle = \langle z^*, z\rangle < \infty. \qquad \square$$

DEFINITION 2.17 (second subderivative [41, Definition 13.3]). *Let $\varphi : \mathbb{R}^n \to \bar{\mathbb{R}}$, $\varphi(z)$ be finite and $z^* \in \mathbb{R}^n$. The second subderivative of φ at z for z^* is a function defined by*

$$\mathrm{d}^2\varphi(z;z^*)(w) := \liminf_{\substack{t\downarrow 0\\ w'\to w}} \frac{\varphi(z+tw') - \varphi(z) - t\langle z^*, w'\rangle}{\frac{1}{2}t^2} \qquad \forall w \in \mathbb{R}^n.$$

According to [41, Example 13.8], if φ is twice differentiable at z and $z^* = \nabla\varphi(z)$, one has

$$\mathrm{d}^2\varphi(z;z^*)(w) = w^T\nabla\varphi^2(z)w.$$

By definition, the second subderivative of the indicator function δ_S of a set S at $z \in S$ for z^* is

$$\mathrm{d}^2\delta_S(z;z^*)(w) = \liminf_{\substack{t\downarrow 0\\ w'\to w}} \frac{\delta_S(z+tw') - \delta_S(z) - t\langle z^*, w'\rangle}{\frac{1}{2}t^2} = \liminf_{\substack{t\downarrow 0,\, w'\to w\\ z+tw'\in S}} \frac{-2\langle z^*, w'\rangle}{t}. \tag{2.22}$$

The second subderivative of the indicator function is extended-real-valued and, by definition, a function of the direction w. However, when dealing with second-order optimality conditions, it also makes sense to consider its dependence on z^*. In the following proposition, we will investigate the set of all (z^*, w) such that $\mathrm{d}^2\delta_S(z;z^*)(w)$ is finite and the relationship between $\mathrm{d}^2\delta_S(z;z^*)(w)$ and the support function of the second-order tangent set $\sigma_{T_S^2(z;w)}(z^*)$ as well as with the lower generalized support function $\hat{\sigma}_{T_S^2(z;w)}(z^*)$. It turns out that the directional proximal normal cones are useful in characterizing the points where $\mathrm{d}^2\delta_S(z;z^*)(w)$ is finite.

PROPOSITION 2.18. *Consider a closed set $S \subset \mathbb{R}^n$, $z \in S$ and a pair $(w, z^*) \in \mathbb{R}^n \times \mathbb{R}^n$. The following statements hold:*

(i) *If $w \notin T_S(z)$ or $\langle z^*, w\rangle < 0$, then $\mathrm{d}^2\delta_S(z; z^*)(w) = \infty$.*

(ii) *For $w \in T_S(z)$, we have $\mathrm{d}^2\delta_S(z; z^*)(w) > -\infty$ if and only if $z^* \in \hat{\mathcal{N}}^p_S(z; w)$.*

(iii) *If $\mathrm{d}^2\delta_S(z; z^*)(w)$ is finite, then $z^* \in \widehat{N}^p_S(z; w)$.*

(iv) *We have*

$$\mathrm{d}^2\delta_S(z; z^*)(w) \le -\sigma_{T^2_S(z;w)}(z^*) \le -\hat{\sigma}_{T^2_S(z;w)}(z^*)$$

if and only if $w \in T_S(z)$ and $\langle z^*, w\rangle \ge 0$ or $T^2_S(z; w) = \emptyset$.

Proof. (i) The statement follows from the definition of tangent cone and (2.22).

(ii) In order to show the if-part of the statement, let $w \in T_S(z)$, and consider $z^* \in \hat{\mathcal{N}}^p_S(z; w)$. Then we can find some $\delta > 0, \gamma > 0$ such that

$$\langle z^*, z' - z\rangle \le \gamma \|z' - z\|^2 \quad \forall z' \in (z + V_{\delta,\delta}(w)) \cap S. \tag{2.23}$$

On the other hand, by (2.22), we can also find sequences $t_k \downarrow 0, w_k \to w$ such that $z + t_k w_k \in S$ and $\mathrm{d}^2\delta_S(z; z^*)(w) = \lim_{k\to\infty} \frac{-2\langle z^*, w_k\rangle}{t_k}$. Since $w_k \to w$, for all k sufficiently large, we have $t_k w_k \in V_{\delta,\delta}(w)$. By (2.23), we have $\langle z^*, t_k w_k\rangle \le \gamma t_k^2 \|w_k\|^2$, from which we obtain the desired inequality

$$\mathrm{d}^2\delta_S(z; z^*)(w) = \lim_{k\to\infty} \frac{-2\langle z^*, w_k\rangle}{t_k} \ge \lim_{k\to\infty} -2\gamma\|w_k\|^2 = -2\gamma\|w\|^2 > -\infty.$$

In order to show the only if-part, assume on the contrary that $w \in T_S(z)$ and $z^* \notin \hat{\mathcal{N}}^p_S(z; w)$. Then there are sequences $t_k \downarrow 0$ and $w_k \to w$ such that $z_k := z + t_k w_k \in S$ and

$$\limsup_{k\to\infty} \frac{\langle z^*, z_k - z\rangle}{\|z_k - z\|^2} = \limsup_{k\to\infty} \frac{\langle z^*, w_k\rangle}{t_k\|w_k\|^2} = \infty.$$

If $w \ne 0$, then the contradiction

$$\infty = \limsup_{k\to\infty} \frac{2\langle z^*, w_k\rangle}{t_k} \le \limsup_{\substack{t\downarrow 0, w'\to w \\ z+tw'\in S}} \frac{2\langle z^*, w'\rangle}{t} = -\mathrm{d}^2\delta_S(z; z^*)(w)$$

follows. In the case when $w = 0$, after possibly passing to a subsequence, we can assume that $\|w_k\| < \frac{1}{k}$ and $\langle z^*, w_k\rangle/(t_k\|w_k\|^2) > k^3$ holds for all k. Defining $\tilde{t}_k := t_k\|w_k\|k$, $\tilde{w}_k := w_k/(k\|w_k\|)$, we have $z_k = z + \tilde{t}_k\tilde{w}_k \in S$, $\tilde{t}_k \downarrow 0$, and $\tilde{w}_k \to 0$, and therefore we obtain once more the contradiction

$$\begin{aligned}\mathrm{d}^2\delta_S(z; z^*)(0) &= \liminf_{\substack{t\downarrow 0, w'\to 0 \\ z+tw'\in S}} \frac{-2\langle z^*, w'\rangle}{t} \le \liminf_{k\to\infty} \frac{-2\langle z^*, \tilde{w}_k\rangle}{\tilde{t}_k} = \liminf_{k\to\infty} \frac{-2\langle z^*, w_k\rangle}{k^2 t_k\|w_k\|^2} \\ &\le \liminf_{k\to\infty} -k = -\infty.\end{aligned}$$

The above arguments show that $z^* \in \hat{\mathcal{N}}^p_S(z; w)$.

(iii) If $\mathrm{d}^2\delta_S(z; z^*)(w)$ is finite, then by statement (i), we must have $w \in T_S(z)$ and $\langle z^*, w\rangle \ge 0$. It then follows by statement (ii) that $z^* \in \hat{\mathcal{N}}^p_S(z; w)$. Since $\mathrm{d}^2\delta_S(z; z^*)(w) = -\infty$ as $\langle z^*, w\rangle > 0$, by definition, we obtain $\langle z^*, w\rangle = 0$. Thus, $z^* \in \widehat{N}^p_S(z; w)$, and the third assertion is shown.

(iv) According to [19, Proposition 6], we know $-\sigma_{T_S^2(z;w)}(z^*) \le -\hat{\sigma}_{T_S^2(z;w)}(z^*)$ for all z^*. Hence, it remains to show

$$\mathrm{d}^2\delta_S(z;z^*)(w) \le -\sigma_{T_S^2(z;w)}(z^*) \tag{2.24}$$

if and only if $w \in T_S(z)$ and $\langle z^*, w\rangle \ge 0$ or $T_S^2(z;w) = \emptyset$. For necessity, if $T_S^2(z;w) = \emptyset$, then $-\sigma_{T_S^2(z;w)}(z^*) = \infty$, and hence (2.24) holds. Let $w \in T_S(z)$. If $\langle z^*, w\rangle > 0$, then (2.24) follows from $\mathrm{d}^2\delta_S(z;z^*)(w) = -\infty$, while if $\langle z^*, w\rangle = 0$, then it holds by [29, Proposition 3.2]. We prove the sufficiency by contradiction. Suppose that (2.24) holds but $T_S^2(z;w) \ne \emptyset$ and either $w \notin T_S(z)$ or $\langle z^*, w\rangle < 0$. In this case, we must have $\mathrm{d}^2\delta_S(z;z^*)(w) = \infty$ by statement (i). On the other hand, since $T_S^2(z;w) \ne \emptyset$, $-\sigma_{T_S^2(z;w)}(z^*) < +\infty$. Hence, $\mathrm{d}^2\delta_S(z;z^*)(w) > -\sigma_{T_S^2(z;w)}(z^*)$, contradicting (2.24). ◻

3. Second-order optimality conditions for (GP). Recall the general problem

$$\text{(GP)} \qquad \min \ f(x) \quad \text{s.t.} \quad g(x) \in C$$

from the introduction. At a feasible point $\bar{x}$ of (GP), the critical cone is defined as

$$\mathcal{C}(\bar{x}) := \{u \in \mathbb{R}^n \mid \nabla g(\bar{x})u \in T_C(g(\bar{x})), \nabla f(\bar{x})u \le 0\},$$

and the generalized Lagrangian $L^\alpha : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ with $\alpha \ge 0$ is given as

$$L^\alpha(x,\lambda) := \alpha f(x) + g(x)^T\lambda,$$

where for $\alpha = 1$, we get the standard Lagrangian $L := L^1$. To study optimality conditions for (GP), we define various multiplier sets as follows, where $u \in \mathcal{C}(\bar{x})$ denotes a critical direction:

$$\Lambda(\bar{x};u) := \{\lambda \in N_C(g(\bar{x});\nabla g(\bar{x})u) \mid \nabla_x L(\bar{x},\lambda) = 0\} \text{ (directional M-multipliers)},$$
$$\Lambda^s(\bar{x};u) := \{\lambda \in \widehat{N}_{T_C(g(\bar{x}))}(\nabla g(\bar{x})u) \mid \nabla_x L(\bar{x},\lambda) = 0\} \text{ (directional S-multipliers)},$$
$$\Lambda^p(\bar{x};u) := \{\lambda \in \widehat{N}^p_C(g(\bar{x});\nabla g(\bar{x})u) \mid \nabla_x L(\bar{x},\lambda) = 0\} \text{ (directional proximal multipliers)}.$$

For $u = 0$, we speak of just M-, S-, and proximal multipliers, which we denote by $\Lambda(\bar{x}) := \Lambda(\bar{x};0)$, $\Lambda^s(\bar{x}) := \Lambda^s(\bar{x};0)$, and $\Lambda^p(\bar{x}) := \Lambda^s(\bar{x};0)$, respectively. For every $u \in \mathcal{C}(\bar{x})$ and every $\lambda \in \Lambda^p(\bar{x}) \subset \widehat{N}^p_C(g(\bar{x})) \subset \widehat{N}_C(g(\bar{x})) = (T_C(g(\bar{x})))^\circ$, we have

$$0 \le -\nabla f(\bar{x})u = \lambda^T\nabla g(\bar{x})u \le 0$$

implying that $\lambda^T\nabla g(\bar{x})u = 0$. Hence, by virtue of Proposition 2.9, the relations

$$\Lambda^p(\bar{x}) \subset \Lambda^p(\bar{x};u) \subset \Lambda^s(\bar{x};u) \subset \Lambda(\bar{x};u)$$

hold, and the inclusions become equalities provided that C is convex by (2.5). In general, we only have the inclusion $\Lambda^p(\bar{x}) \subset \Lambda^s(\bar{x})$, but they are equal for many nonconvex and nonpolyhedral sets important in applications, e.g., the second-order cone complementarity set [44] and the semidefinite complementarity cone [10].

Recall first the following second-order necessary optimality condition for (GP).

THEOREM 3.1 ([19, Theorem 2 and Corollary 5]). *Let $\bar{x}$ be a local optimal solution for problem (GP). Then for every critical direction $u \in \mathcal{C}(\bar{x})$, the following necessary optimality conditions hold:*

(i) *Suppose the constraint mapping $x \rightrightarrows g(x) - C$ is metrically subregular at $(\bar{x}, 0)$ in direction u. Then there exists a directional M-multiplier $\lambda \in \Lambda(\bar{x}; u)$ such that*

$$\nabla^2_{xx} L(\bar{x}, \lambda)(u, u) - \hat{\sigma}_{T^2_C(g(\bar{x}); \nabla g(\bar{x})u)}(\lambda) \geq 0. \tag{3.1}$$

(ii) *Suppose that the directional nondegeneracy condition*

$$\nabla g(\bar{x})^T y^* = 0, \ y^* \in \operatorname{span} N_C(g(\bar{x}); \nabla g(\bar{x})u) \implies y^* = 0$$

holds. Then $\Lambda^s(\bar{x}; u) = \Lambda(\bar{x}; u) = \{\lambda_0\}$ is a singleton, and the second-order condition

$$\nabla^2_{xx} L(\bar{x}, \lambda_0)(u, u) - \sigma_{T^2_C(g(\bar{x}); \nabla g(\bar{x})u)}(\lambda_0) \geq 0$$

holds.

We now derive second-order sufficient optimality conditions for (GP). We state our result in terms of the following notion introduced by Penot [35].

DEFINITION 3.2 (essential local minimizer of second order). *A point $\bar{x}$ is said to be an essential local minimizer of second order for problem (GP) if $\bar{x}$ is feasible and there exist $\varepsilon > 0$ and $\delta > 0$ such that*

$$\max\{f(x) - f(\bar{x}), \operatorname{dist}(g(x), C)\} \geq \varepsilon \|x - \bar{x}\|^2 \quad \forall x \in \mathbb{B}(\bar{x}, \delta).$$

THEOREM 3.3. *Let $\bar{x}$ be a feasible point of problem (GP). Suppose that for every $u \in \mathcal{C}(\bar{x}) \setminus \{0\}$, there is $\alpha \geq 0$ and $\lambda \in \mathbb{R}^m$ such that*

$$\nabla_x L^\alpha(\bar{x}, \lambda) = 0, \tag{3.2a}$$

$$\nabla^2_{xx} L^\alpha(\bar{x}, \lambda)(u, u) + \mathrm{d}^2 \delta_C(g(\bar{x}); \lambda)(\nabla g(\bar{x})u) > 0. \tag{3.2b}$$

Then $\bar{x}$ is an essential local minimizer of second order.

Proof. By contradiction, if $\bar{x}$ is not an essential local minimizer of second order, then there exists a sequence x_k converging to $\bar{x}$ such that

$$f(x_k) - f(\bar{x}) \leq o(\|x_k - \bar{x}\|^2), \tag{3.3a}$$

$$\operatorname{dist}(g(x_k), C) \leq o(\|x_k - \bar{x}\|^2). \tag{3.3b}$$

Let $t_k := \|x_k - \bar{x}\|$ and $u_k := (x_k - \bar{x})/t_k$. We assume without loss of generality that u_k is converging to u. From (3.3a), it readily follows that $\nabla f(\bar{x})u \leq 0$ and

$$\liminf_k -\frac{f(x_k) - f(\bar{x})}{\frac{1}{2}t_k^2} \geq 0. \tag{3.4}$$

By (3.3b), there exists $r_k \in \mathbb{R}^d$ such that $\|r_k\| \to 0$ and $g(x_k) + t_k^2 r_k \in C$. By Taylor's expansion, since $x_k = \bar{x} + t_k u_k$, we have

$$\begin{aligned} v_k &:= \frac{g(x_k) + t_k^2 r_k - g(\bar{x})}{t_k} = \frac{\nabla g(\bar{x}) t_k u_k + o(t_k) + t_k^2 r_k}{t_k} \\ &= \nabla g(\bar{x}) u_k + \frac{o(t_k)}{t_k} + t_k r_k \quad \to \ \nabla g(\bar{x}) u. \end{aligned}$$

Moreover, $g(\bar{x}) + t_k v_k = g(x_k) + t_k^2 r_k \in C$, and so $\nabla g(\bar{x})u \in T_C(g(\bar{x}))$ follows. Thus, $u \in \mathcal{C}(\bar{x}) \setminus \{0\}$, and the assumption of the theorem yields the existence of $\alpha \geq 0$ and $\lambda \in \mathbb{R}^m$ satisfying (3.2a) and (3.2b). Using (2.22), (3.4), $\|r_k\| \to 0$, and (3.2a), however, we obtain

$$\begin{aligned}
\mathrm{d}^2\delta_C(g(\bar{x});\lambda)(\nabla g(\bar{x})u) &= \liminf_{\substack{t\downarrow 0, v'\to\nabla g(\bar{x})u\\ g(\bar{x})+tv'\in C}} \frac{-\langle\lambda, v'\rangle}{\frac{1}{2}t} \leq \liminf_k \frac{-t_k\langle\lambda, v_k\rangle}{\frac{1}{2}t_k^2}\\
&= \liminf_k \frac{-\langle\lambda, g(x_k) + t_k^2 r_k - g(\bar{x})\rangle}{\frac{1}{2}t_k^2}\\
&\leq \liminf_k -\alpha\frac{f(x_k) - f(\bar{x})}{\frac{1}{2}t_k^2} + \liminf_k -\frac{\langle\lambda, g(x_k) - g(\bar{x})\rangle}{\frac{1}{2}t_k^2}\\
&\leq \liminf_k -\frac{L^\alpha(x_k,\lambda) - L^\alpha(\bar{x},\lambda)}{\frac{1}{2}t_k^2}\\
&= \liminf_k -\frac{\nabla_x L^\alpha(\bar{x},\lambda)(t_k u_k) + \frac{1}{2}\nabla^2_{xx}L^\alpha(\bar{x},\lambda)(t_k u_k, t_k u_k) + o(t_k^2)}{\frac{1}{2}t_k^2}\\
&= \liminf_k -\nabla^2_{xx}L^\alpha(\bar{x},\lambda)(u_k,u_k) = -\nabla^2_{xx}L^\alpha(\bar{x},\lambda)(u,u),
\end{aligned}$$

which contradicts (3.2b). This completes the proof. ∎

Additional requirements on λ are hidden in conditions (3.2a) and (3.2b).

PROPOSITION 3.4. *Let $\bar{x}$ be a feasible point of problem (GP), let $u \in \mathcal{C}(\bar{x})$ be a critical direction, and let $\alpha \geq 0$, $\lambda \in \mathbb{R}^m$ satisfy conditions* (3.2a) *and* (3.2b). *Then α and λ are not both zero, and $\lambda \in \widehat{N}^p_C(g(\bar{x});\nabla g(\bar{x})u)$. Particularly, if $\alpha \neq 0$, then $\tilde{\lambda} := \lambda/\alpha \in \Lambda^p(\bar{x};u)$, and conditions* (3.2a) *and* (3.2b) *hold with $\tilde{\alpha} := 1$ and $\tilde{\lambda}$.*

Proof. Note that α and λ cannot be simultaneously zero because otherwise

$$\nabla^2_{xx}L^\alpha(\bar{x},\lambda)(u,u) = \mathrm{d}^2\delta_C(g(\bar{x});\lambda)(\nabla g(\bar{x})u) = 0,$$

contradicting (3.2b). Since $\mathrm{d}^2\delta_C(g(\bar{x});\lambda)(\nabla g(\bar{x})u) > -\infty$, we conclude that $\lambda \in \hat{\mathcal{N}}^p_C(g(\bar{x});\nabla g(\bar{x})u)$ by Proposition 2.18(ii) and $\langle\lambda, \nabla g(\bar{x})u\rangle \leq 0$ by (2.4). Meanwhile, λ satisfies $\nabla_x L^\alpha(\bar{x},\lambda) = 0$, i.e., $\alpha\nabla f(\bar{x}) + \nabla g(\bar{x})^T\lambda = 0$, implying that $\langle\lambda, \nabla g(\bar{x})u\rangle = -\alpha\nabla f(\bar{x})u \geq 0$ due to $u \in \mathcal{C}(\bar{x})$. Thus, $\langle\lambda, \nabla g(\bar{x})u\rangle = 0$, and we get

$$\lambda \in \hat{\mathcal{N}}^p_C(g(\bar{x});\nabla g(\bar{x})u) \cap \{\nabla g(\bar{x})u\}^\perp = \widehat{N}^p_C(g(\bar{x});\nabla g(\bar{x})u).$$

Since $\widehat{N}^p_C(g(\bar{x});\nabla g(\bar{x})u)$ is a cone, it also contains λ/α if $\alpha \neq 0$. Thus, dividing (3.2a)–(3.2b) by α yields the last claim, taking into account (2.22). ∎

Let us now compare the concept of essential local minimizers with the more common notion that the quadratic growth condition for (GP) holds at $\bar{x}$; i.e., there exist $\varepsilon > 0$ and $\delta > 0$ such that

$$f(x) \geq f(\bar{x}) + \varepsilon\|x - \bar{x}\|^2 \qquad \forall x \in \mathbb{B}(\bar{x},\delta) \;\text{ s.t. }\; g(x) \in C. \tag{3.5}$$

LEMMA 3.5. *Consider the following statements:*

(i) *$\bar{x}$ is an essential local minimizer of second order.*

(ii) *The quadratic growth condition holds at $\bar{x}$.*

Then the implication (i) $\Rightarrow$ (ii) *always holds. Conversely, if the constraint mapping $x \rightrightarrows g(x) - C$ is metrically subregular at $(\bar{x}, 0)$ in every critical direction $u \in \mathcal{C}(\bar{x}) \setminus \{0\}$, then the reverse implication* (ii) $\Rightarrow$ (i) *is also valid.*

Proof. The validity of the implication (i) $\Rightarrow$ (ii) follows immediately from the definitions. We show the second assertion by contraposition. Assume that the quadratic growth condition and the stated constraint qualification hold, and assume on the contrary that there is a sequence $x_k \to \bar{x}$ with $\max\{f(x_k) - f(\bar{x}), \operatorname{dist}(g(x_k), C)\}/\|x_k - \bar{x}\|^2 \to 0$. By passing to a subsequence, we may assume that $(x_k - \bar{x})/\|x_k - \bar{x}\|$ converges to some u, and the arguments already employed in the proof of Theorem 3.3 show that $u \in \mathcal{C}(\bar{x}) \setminus \{0\}$. By the assumed directional metric subregularity, there is some $\kappa > 0$ such that for all k sufficiently large, we can find some $\tilde{x}_k$ with $g(\tilde{x}_k) \in C$ and $\|\tilde{x}_k - x_k\| \le \kappa \operatorname{dist}(g(x_k), C) = o(\|x_k - \bar{x}\|^2)$. Since f is Lipschitz continuous in a neighborhood of $\bar{x}$ with some constant l, we obtain from (3.5) the contradiction

$$\begin{aligned} 0 < \varepsilon &\le \liminf_{k\to\infty} \frac{f(\tilde{x}_k) - f(\bar{x})}{\|\tilde{x}_k - \bar{x}\|^2} \le \liminf_{k\to\infty} \frac{f(x_k) - f(\bar{x}) + l\|\tilde{x}_k - x_k\|}{(\|x_k - \bar{x}\| - \|\tilde{x}_k - x_k\|)^2} \\ &= \liminf_{k\to\infty} \frac{f(x_k) - f(\bar{x}) + o(\|x_k - \bar{x}\|^2)}{\|x_k - \bar{x}\|^2 - o(\|x_k - \bar{x}\|^2)} = \liminf_{k\to\infty} \frac{f(x_k) - f(\bar{x})}{\|x_k - \bar{x}\|^2} \le 0. \qquad \square \end{aligned}$$

Note that Theorem 3.3 improves Mohammadi, Mordukhovich, and Sarabi [29, Proposition 7.3] in that the set C is not required to be convex and parabolically derivable, α can be zero, we can choose different multipliers for different critical directions, and the concept of local minimizer is stronger. For the sake of completeness, we also state the following corollary.

COROLLARY 3.6. *Let $\bar{x}$ be a feasible point of problem (GP). Suppose that for every $u \in \mathcal{C}(\bar{x})\setminus\{0\}$, there is a directional proximal multiplier $\lambda \in \Lambda^p(\bar{x}; u)$ such that*

$$\nabla^2_{xx} L(\bar{x}, \lambda)(u, u) + \mathrm{d}^2 \delta_C(g(\bar{x}); \lambda)(\nabla g(\bar{x})u) > 0.$$

Then the quadratic growth condition (3.5) *holds for problem (GP).*

4. First-order variational analysis of disjunctive systems.

In this section, we begin with first-order variational analysis of the disjunctive system $\Gamma := \{x \in \mathbb{R}^n \mid G(x) \in D\}$, where $G : \mathbb{R}^n \to \mathbb{R}^d$ is continuously differentiable and $D \subset \mathbb{R}^d$ is polyhedral. Note that many first-order results are valid for any closed set D and were already stated in Proposition 2.14. In the following theorem, we show that since D is polyhedral, some results from Proposition 2.14 can be improved. Namely, under the directional nondegeneracy (2.20), the inclusions in (2.18b) become equalities, making all four sets equal. Note also that [41, Theorem 6.14] cannot be applied directly since D is not regular in the sense of Clarke (cf. [41, Definition 6.4]).

THEOREM 4.1. *Consider a feasible point $\bar{x} \in \Gamma$ and a direction $u \in \mathbb{R}^n$ and assume that the directional nondegeneracy condition*

$$\nabla G(\bar{x})^T y^* = 0,\ y^* \in \operatorname{span} N_D(G(\bar{x}); \nabla G(\bar{x})u) \implies y^* = 0 \tag{4.1}$$

is fulfilled. Then

$$N_\Gamma(\bar{x}; u) = \nabla G(\bar{x})^T N_D(G(\bar{x}); \nabla G(\bar{x})u) = \nabla G(\bar{x})^T N_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u) = N_{T_\Gamma(\bar{x})}(u). \tag{4.2}$$

Proof. First note that condition (4.1) implies FOSCMS (2.17), and hence $x \rightrightarrows G(x) - D$ is metrically subregular at $(\bar{x}, 0)$ in direction u. From [19, Lemma 3], Proposition 2.14, and Proposition 2.11, respectively, we know that

$$N_{T_\Gamma(\bar{x})}(u) \subset N_\Gamma(\bar{x}; u) \subset \nabla G(\bar{x})^T N_D(G(\bar{x}); \nabla G(\bar{x})u) = \nabla G(\bar{x})^T N_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u).$$

Thus, it remains to show $\nabla G(\bar{x})^T N_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u) \subset N_{T_\Gamma(\bar{x})}(u)$. Consider $y^* \in N_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)$ together with sequences $v_k \in T_D(G(\bar{x}))$ and $y_k^* \in \widehat{N}_{T_D(G(\bar{x}))}(v_k)$ with $v_k \to \nabla G(\bar{x})u$ and $y_k^* \to y^*$. The nondegeneracy condition and Proposition 2.6 yield

$$\mathbb{R}^d = \left(\ker \nabla G(\bar{x})^T \cap \operatorname{span} N_D(G(\bar{x}); \nabla G(\bar{x})u)\right)^\perp = \nabla G(\bar{x})\mathbb{R}^n + \mathcal{L}(T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)).$$

Consider the $d \times (n+d)$ matrix

$$A := \begin{bmatrix} \nabla G(\bar{x}) & P \end{bmatrix},$$

where P is the symmetric $d \times d$ matrix representing the orthogonal projection onto $\mathcal{L}(T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u))$. By the equation above, A has full row rank d, and therefore the solution set S of the linear equation $A\binom{d}{r} = v_k - \nabla G(\bar{x})u$ is an n-dimensional affine subspace. Denoting by $A^\dagger$ the pseudoinverse (also known as Moore–Penrose inverse) of A,

$$\begin{pmatrix} d_k \\ r_k \end{pmatrix} := A^\dagger (v_k - \nabla G(\bar{x})u)$$

selects from S the solution with minimal Euclidean norm; cf. [36]. Since $P^2 = P$, the vector $\binom{d_k}{Pr_k}$ is another solution, and due to the minimal norm property of $\binom{d_k}{r_k}$, we obtain

$$\|Pr_k\|^2 \geq \|r_k\|^2 = \|Pr_k\|^2 + \|(I-P)r_k\|^2,$$

which is only possible if $(I-P)r_k = 0$. Thus, we conclude that $r_k = Pr_k \in \mathcal{L}(T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u))$. Further, since $v_k - \nabla G(\bar{x})u \to 0$, we also have $d_k \to 0$ and $r_k \to 0$ as $k \to \infty$. Hence, $u_k := u + d_k \to u$ and $\nabla G(\bar{x})u_k = v_k - r_k \to \nabla G(\bar{x})u$ follows.

We obtain

$$\begin{aligned} y_k^* \in \widehat{N}_{T_D(G(\bar{x}))}(v_k) &= \widehat{N}_{T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)}(v_k - \nabla G(\bar{x})u) \\ &= \widehat{N}_{T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)}(v_k - \nabla G(\bar{x})u - r_k) = \widehat{N}_{T_D(G(\bar{x}))}(v_k - r_k) \\ &= \widehat{N}_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u_k), \end{aligned} \tag{4.3}$$

where the first and the third equalities follow from (2.12) and the second equality comes from (2.2). Moreover, since $\widehat{N}_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u_k) \neq \emptyset$, we must have $\nabla G(\bar{x})u_k \in T_D(G(\bar{x}))$. Thus, taking into account $u_k \to u$, from (2.18a), we get $u_k \in T_\Gamma(\bar{x})$ for sufficiently large k, and $\nabla G(\bar{x})^T y_k^* \in \widehat{N}_{T_\Gamma(\bar{x})}(u_k)$ follows by (4.3) and [41, Theorem 6.14]. Taking limits as $k \to \infty$, we conclude that $\nabla G(\bar{x})^T y^* \in N_{T_\Gamma(\bar{x})}(u)$, proving

$$\nabla G(\bar{x})^T N_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u) \subset N_{T_\Gamma(\bar{x})}(u).$$ ∎

5. Second-order variational analysis of disjunctive systems. In this section, we continue with second-order variational analysis of the disjunctive system $\Gamma := \{x \in \mathbb{R}^n \mid G(x) \in D\}$, where G is now twice continuously differentiable and D is polyhedral. We will study the domain of the support functions and the connection between the second-order objects $\mathrm{d}^2\delta_\Gamma(\bar{x}; \cdot)(u), \sigma_{T_\Gamma^2(\bar{x};u)}(\cdot), \hat{\sigma}_{T_\Gamma^2(\bar{x};u)}(\cdot)$ and the second derivative $\nabla^2 G(\bar{x})(u,u)$. In the first part, we provide very general results, assuming only that the constraint mapping $x \rightrightarrows G(x) - D$ is metrically subregular at $(\bar{x}, 0)$ in direction u (MSCQ holds at $\bar{x}$ in direction u), and in the second part, we show

how everything gets simpler under the directional nondegeneracy (2.20) or a relaxed version of this property.

Given a feasible point $\bar{x} \in \Gamma$ and a pair $(u, x^*) \in \mathbb{R}^n \times \mathbb{R}^n$, we denote the set of S- and M-multipliers associated with $(\bar{x}, x^*)$ in direction u, respectively, by

$$\Lambda^s_{x^*}(\bar{x}; u) := \{y^* \in \widehat{N}_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u) \mid x^* = \nabla G(\bar{x})^T y^*\},$$
$$\Lambda_{x^*}(\bar{x}; u) := \{y^* \in N_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u) \mid x^* = \nabla G(\bar{x})^T y^*\}.$$

Consider a direction u belonging to the linearization cone $L_\Gamma(\bar{x})$ defined by

$$L_\Gamma(\bar{x}) := \{u \in \mathbb{R}^n \mid \nabla G(\bar{x})u \in T_D(G(\bar{x}))\}.$$

If MSCQ holds at $\bar{x}$ in direction u, then $u \in T_\Gamma(\bar{x})$ by Proposition 2.14, and we also get

$$T^2_\Gamma(\bar{x}; u) = \{p \in \mathbb{R}^n \mid \nabla G(\bar{x})p + \nabla^2 G(\bar{x})(u, u) \in T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)\} \tag{5.1}$$

from Propositions 2.11 and 2.14 (see also [28, 29]).

5.1. Subregular systems. We begin with the main result of this section.

THEOREM 5.1. *Let $\bar{x} \in \Gamma$ and $u \in L_\Gamma(\bar{x})$, and suppose that MSCQ holds at $\bar{x}$ in direction u with the subregularity modulus κ. Then the following statements hold:*

(i) *For every $x^* \in \{u\}^\perp$, we have*

$$\mathrm{d}^2\delta_\Gamma(\bar{x}; x^*)(u) = -\sigma_{T^2_\Gamma(\bar{x};u)}(x^*). \tag{5.2}$$

(ii) *We have $\operatorname{dom}\sigma_{T^2_\Gamma(\bar{x};u)} = \widehat{N}^p_\Gamma(\bar{x}; u) = \widehat{N}_{T_\Gamma(\bar{x})}(u)$. For every $x^* \in \operatorname{dom}\sigma_{T^2_\Gamma(\bar{x};u)}$, we have $x^* \in \{u\}^\perp$, the equality (5.2) holds, and*

$$\inf_{y^* \in \Lambda_{x^*}(\bar{x};u) \cap \kappa\|x^*\| cl\,\mathbb{B}} \langle y^*, \nabla^2 G(\bar{x})(u, u)\rangle \le d^2\delta_\Gamma(\bar{x}; x^*)(u) \le \sup_{y^* \in \Lambda_{x^*}(\bar{x};u) \cap \kappa\|x^*\| cl\,\mathbb{B}} \langle y^*, \nabla^2 G(\bar{x})(u, u)\rangle, \tag{5.3a}$$

$$d^2\delta_\Gamma(\bar{x}; x^*)(u) \ge \sup_{y^* \in \Lambda^s_{x^*}(\bar{x};u)} \langle y^*, \nabla^2 G(\bar{x})(u, u)\rangle, \tag{5.3b}$$

and, moreover, there exists $y^* \in \Lambda_{x^*}(\bar{x}; u)$ such that $\mathrm{d}^2\delta_\Gamma(\bar{x}; x^*)(u) = \langle y^*, \nabla^2 G(\bar{x})(u, u)\rangle$.

(iii) *We have $\operatorname{dom}\hat{\sigma}_{T^2_\Gamma(\bar{x};u)} \subset \{x^* \mid \Lambda_{x^*}(\bar{x}; u) \neq \emptyset\}$, and for every $x^* \in \operatorname{dom}\hat{\sigma}_{T^2_\Gamma(\bar{x};u)}$, it holds that*

$$\inf_{y^* \in \Lambda_{x^*}(\bar{x};u) \cap \kappa\|x^*\| \mathrm{cl}\,\mathbb{B}} \langle y^*, \nabla^2 G(\bar{x})(u, u)\rangle \le -\hat{\sigma}_{T^2_\Gamma(\bar{x};u)}(x^*) \le \sup_{y^* \in \Lambda_{x^*}(\bar{x};u) \cap \kappa\|x^*\| \mathrm{cl}\,\mathbb{B}} \langle y^*, \nabla^2 G(\bar{x})(u, u)\rangle, \tag{5.4}$$

and there exists $y^* \in \Lambda_{x^*}(\bar{x}; u)$ such that $-\hat{\sigma}_{T^2_\Gamma(\bar{x};u)}(x^*) = \langle y^*, \nabla^2 G(\bar{x})(u, u)\rangle$. Moreover, the upper bound in (5.4) is valid for all $x^* \in \mathbb{R}^n$.

Proof. (i) Consider $x^* \in \{u\}^\perp$. Taking into account (2.22), consider sequences $t_k \downarrow 0$ and $u_k \to u$ such that $\bar{x} + t_k u_k \in \Gamma$ and

$$\mathrm{d}^2\delta_\Gamma(\bar{x}; x^*)(u) = \lim_{k\to\infty} -\frac{2\langle x^*, u_k\rangle}{t_k} = \lim_{k\to\infty} -\frac{2\langle x^*, u_k - u\rangle}{t_k}. \tag{5.5}$$

Since $G(\bar{x}+t_k u_k)=G(\bar{x})+t_k\nabla G(\bar{x})u_k+\frac{1}{2}t_k^2(\nabla^2 G(\bar{x})(u,u)+r_k)\in D$ with $r_k\to 0$, we obtain

$$\nabla G(\bar{x})u_k+\frac{1}{2}t_k(\nabla^2 G(\bar{x})(u,u)+r_k)\in T_D(G(\bar{x})).$$

Consequently, Proposition 2.10 yields

$$\nabla G(\bar{x})\frac{2(u_k-u)}{t_k}+\nabla^2 G(\bar{x})(u,u)+r_k\in T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u). \tag{5.6}$$

From Proposition 2.14 and (5.1), we get that the mapping

$$\Phi(p):=\nabla G(\bar{x})p+\nabla^2 G(\bar{x})(u,u)-T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u) \tag{5.7}$$

satisfies

$$\operatorname{dist}(p,\Phi^{-1}(0))\le\kappa\operatorname{dist}(0,\Phi(p))\quad\forall p\in\mathbb{R}^n. \tag{5.8}$$

By (5.6), we have $-r_k\in\Phi(2(u_k-u)/t_k)$. Hence, for every k, we can find some $p_k\in\Phi^{-1}(0)$ satisfying $\|\frac{2(u_k-u)}{t_k}-p_k\|\le\kappa\|r_k\|$ and

$$\nabla G(\bar{x})p_k+\nabla^2 G(\bar{x})(u,u)\in T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u),$$

and $p_k\in T^2_\Gamma(\bar{x};u)$ follows from (5.1). Thus, by definition of the support function, we have $\langle x^*,p_k\rangle\le\sigma_{T^2_\Gamma(\bar{x};u)}(x^*)$. Moreover, by (5.5), we have

$$\mathrm{d}^2\delta_\Gamma(\bar{x};x^*)(u)=\lim_{k\to\infty}-\frac{2\langle x^*,u_k-u\rangle}{t_k}=\lim_{k\to\infty}-\langle x^*,p_k\rangle\ge-\sigma_{T^2_\Gamma(\bar{x};u)}(x^*).$$

Since the opposite inequality holds by Proposition 2.18(iv), (5.2) is established.

(ii) We have already shown in Proposition 2.9 the inclusion $\widehat{N}^p_\Gamma(\bar{x};u)\subset\widehat{N}_{T_\Gamma(\bar{x})}(u)$. Further, since $u\in T_\Gamma(\bar{x})$, by Proposition 2.14, we get $\nabla G(\bar{x})u\in T_D(G(\bar{x}))$, and $T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)\neq\emptyset$ follows. This, however, ensures that $T^2_\Gamma(\bar{x};u)\neq\emptyset$ due to (5.8) and $\Phi^{-1}(0)=T^2_\Gamma(\bar{x};u)$ by (5.1). Hence, $\sigma_{T^2_\Gamma(\bar{x};u)}(x^*)>-\infty$ holds for all $x^*\in\mathbb{R}^n$. Now let $x^*\in\widehat{N}_{T_\Gamma(\bar{x})}(u)$, and we first show that $x^*\in\operatorname{dom}\sigma_{T^2_\Gamma(\bar{x};u)}$. Assume on the contrary that $\sigma_{T^2_\Gamma(\bar{x};u)}(x^*)=\infty$, and consider a sequence $p_k\in T^2_\Gamma(\bar{x};u)$ with $\langle x^*,p_k\rangle\to\infty$ as $k\to\infty$. By (5.1), we get

$$\nabla G(\bar{x})p_k+\nabla^2 G(\bar{x})(u,u)\in T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u). \tag{5.9}$$

The mapping

$$p\rightrightarrows\nabla G(\bar{x})p-T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)$$

is polyhedral; i.e., its graph is a polyhedral set and is therefore metrically subregular at $(0,0)$ by Robinson's result [37]. Since its graph is also a closed cone, Proposition 2.13 yields the existence of $\kappa'>0$ such that for every k, we can find some $\tilde{p}_k$ with $\nabla G(\bar{x})\tilde{p}_k\in T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)$ and

$$\|\tilde{p}_k-p_k\|\le\kappa'\operatorname{dist}(\nabla G(\bar{x})p_k,T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u))\le\kappa'\|\nabla^2 G(\bar{x})(u,u)\|,$$

where the second inequality follows from (5.9). Since $\nabla G(\bar{x})\tilde{p}_k\in T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)$, by Proposition 2.14, we obtain $\tilde{p}_k\in T_{T_\Gamma(\bar{x})}(u)$, implying that $\langle x^*,\tilde{p}_k\rangle\le 0$ due to $x^*\in$

$\widehat{N}_{T_\Gamma(\bar{x})}(u) = (T_{T_\Gamma(\bar{x})}(u))^\circ$. This, however, contradicts the assumption that $\langle x^*, p_k\rangle \to \infty$ as $k \to \infty$ since the sequence $\{\tilde{p}_k - p_k\}$ is bounded, showing $x^* \in \operatorname{dom} \sigma_{T^2_\Gamma(\bar{x};u)}$.

Consider now $x^* \in \operatorname{dom} \sigma_{T^2_\Gamma(\bar{x};u)}$. Note that in order to show that $x^* \in \widehat{N}^p_\Gamma(\bar{x};u)$, it suffices to prove that $x^* \in \{u\}^\perp$ since then we get (5.2) and Proposition 2.18(iii) gives the claim. As we will see, however, $x^* \in \{u\}^\perp$ comes as a by-product of the following arguments.

Since $T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)$ is a polyhedral cone, it can be written as the union of finitely many convex polyhedral cones, say, K_i, $i = 1, \ldots, s$, and therefore

$$\sigma_{T^2_\Gamma(\bar{x};u)}(x^*) = \max_{i=1,\ldots,s} \sup_p \{\langle x^*, p\rangle \mid \nabla G(\bar{x})p + \nabla^2 G(\bar{x})(u,u) \in K_i\}.$$

Taking into account that $\sigma_{T^2_\Gamma(\bar{x};u)}(x^*)$ is finite, every linear program $\sup_p\{\langle x^*, p\rangle \mid \nabla G(\bar{x})p + \nabla^2 G(\bar{x})(u,u) \in K_i\}$ either is infeasible, resulting in the optimal value $-\infty$, or has a finite optimal value, and this optimal value is attained; see, e.g., [5, Theorem 2.198]. Hence, the program

$$\max \langle x^*, p\rangle \text{ subject to } \nabla G(\bar{x})p + \nabla^2 G(\bar{x})(u,u) \in T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u) \tag{5.10}$$

has an optimal solution $\bar{p}$, and $\sigma_{T^2_\Gamma(\bar{x};u)}(x^*) = \langle x^*, \bar{p}\rangle$ follows. The corresponding constraint mapping is precisely Φ from (5.7), and it is metrically subregular at $(\bar{p}, 0)$ by (5.8). Thus, by [17, Theorem 3], there exists a multiplier y^* fulfilling the first-order optimality conditions

$$-x^* + \nabla G(\bar{x})^T y^* = 0, \ \|y^*\| \le \kappa \|x^*\|, \tag{5.11a}$$

$$y^* \in N_{T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)}(\nabla G(\bar{x})\bar{p} + \nabla^2 G(\bar{x})(u,u)) \subset N_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u), \tag{5.11b}$$

where in (5.11b) we used [41, Proposition 6.27(a)]. Particularly, since $T_D(G(\bar{x}))$ and $T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)$ are cones, we conclude that

$$\langle y^*, \nabla G(\bar{x})u\rangle = 0 \quad \text{and} \quad \langle y^*, \nabla G(\bar{x})\bar{p} + \nabla^2 G(\bar{x})(u,u)\rangle = 0.$$

This means, however, that $\langle x^*, u\rangle = \langle \nabla G(\bar{x})^T y^*, u\rangle = \langle y^*, \nabla G(\bar{x})u\rangle = 0$ and

$$\sigma_{T^2_\Gamma(\bar{x};u)}(x^*) = \langle x^*, \bar{p}\rangle = -\langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle, \tag{5.12}$$

and we indeed get (5.2) as claimed. Moreover, conditions (5.11a) and (5.11b) ensure that $y^* \in \Lambda_{x^*}(\bar{x}; u)$, and so (5.3a) follows from (5.2).

In order to show (5.3b), consider $y^* \in \Lambda^s_{x^*}(\bar{x}; u)$, i.e.,

$$y^* \in \widehat{N}_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u) = [T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)]^\circ, \quad x^* = \nabla G(\bar{x})^T y^*.$$

Since $\bar{p}$ is an optimal solution of program (5.10), it is feasible, i.e.,

$$\nabla G(\bar{x})\bar{p} + \nabla^2 G(\bar{x})(u,u) \in T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u).$$

It follows that $\langle y^*, \nabla G(\bar{x})\bar{p} + \nabla^2 G(\bar{x})(u,u)\rangle \le 0$, showing $\sigma_{T^2_\Gamma(\bar{x};u)}(x^*) = \langle x^*, \bar{p}\rangle \le -\langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle$. Again, (5.3b) follows from (5.2). Finally, by (5.2) and (5.12), we obtain the last conclusion of (ii).

(iii) Let x^* satisfy

$$\hat{\sigma}_{T^2_\Gamma(\bar{x};u)}(x^*) := \liminf_{\tilde{x}^* \to x^*} \inf_{p'} \{\langle \tilde{x}^*, p'\rangle | \tilde{x}^* \in \widehat{N}_{T^2_\Gamma(\bar{x};u)}(p')\} < \infty.$$

Consider the sequences $x_k^* \to x^*$ and p_k such that $x_k^* \in \widehat{N}_{T^2_\Gamma(\bar{x};u)}(p_k)$ and $\hat{\sigma}_{T^2_\Gamma(\bar{x};u)}(x^*) = \lim_{k\to\infty}\langle x_k^*, p_k\rangle$. By (5.1), we have $T^2_\Gamma(\bar{x};u) = \Phi^{-1}(0)$, where the mapping Φ is given by (5.7). Since Φ is metrically subregular with modulus κ at $(p_k, 0)$ and $x_k^* \in \widehat{N}_{T^2_\Gamma(\bar{x};u)}(p_k) \subset N_{T^2_\Gamma(\bar{x};u)}(p_k) = N_{\Phi^{-1}(0)}(p_k)$, [15, Proposition 4.1] yields the existence of some y_k^* satisfying $\|y_k^*\| \le \kappa\|x_k^*\|$ and $(x_k^*, -y_k^*) \in N_{\operatorname{gph}\Phi}(p_k, 0)$. Applying the change of coordinates formula (see, e.g., [41, Exercise 6.7]) to $N_{\operatorname{gph}\Phi}(p_k, 0)$, we obtain

$$x_k^* = \nabla G(\bar{x})^T y_k^*, \tag{5.13a}$$

$$y_k^* \in N_{T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)}(\nabla G(\bar{x})p_k + \nabla^2 G(\bar{x})(u,u)) \subset N_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u), \tag{5.13b}$$

taking into account [41, Proposition 6.27(a)] as before. Since $T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)$ is a cone, by (5.13b), we have $\langle y_k^*, \nabla G(\bar{x})p_k + \nabla^2 G(\bar{x})(u,u)\rangle = 0$, which together with (5.13a) implies that

$$\langle x_k^*, p_k\rangle = -\langle y_k^*, \nabla^2 G(\bar{x})(u,u)\rangle. \tag{5.14}$$

Since the sequence y_k^* is bounded due to $\|y_k^*\| \le \kappa\|x_k^*\|$, we can assume that it converges to some y^* with $\|y^*\| \le \kappa\|x^*\|$. Taking limits in (5.14), (5.13a), and (5.13b), we obtain $y^* \in \Lambda_{x^*}(\bar{x};u)$ and $\hat{\sigma}_{T^2_\Gamma(\bar{x};u)}(x^*) = -\langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle$, proving (5.4). The upper bound in (5.4) is obviously valid if $\hat{\sigma}_{T^2_\Gamma(\bar{x};u)}(x^*) = \infty$. ∎

Remark 5.2. Recall that by Proposition 2.11, we have

$$N_{T_D(z)}(w) = N_D(z;w),\ \forall z \in D,\ w \in T_D(z).$$

Applying Theorem 5.1(ii) with G being the identity mapping and $\Gamma = D$ yields the counterpart

$$\widehat{N}_{T_D(z)}(w) = \widehat{N}^p_D(z;w)\ \forall z \in D,\ w \in T_D(z).$$

Note that the bounds for the second subderivative and the lower generalized support function have the same structure, the only difference being the range of validity. Further note that although the inclusion $\operatorname{dom}\hat{\sigma}_{T^2_\Gamma(\bar{x};u)} \subset \{x^* \mid \Lambda_{x^*}(\bar{x};u) \ne \emptyset\}$ holds, it might be strict in general. However, the equality can be obtained under the directional nondegeneracy condition, as shown in Corollary 5.8(i) below.

Remark 5.3. Inspired by [29, Proposition 5.4], we further show that in (5.3b), the supremum over $\Lambda^s_{x^*}(\bar{x};u)$ provides a tight lower bound for $\mathrm{d}^2\delta_\Gamma(\bar{x};x^*)(u)$ from the point of view of weak duality.

In fact,

$$\begin{aligned}\mathrm{d}^2\delta_\Gamma(\bar{x};x^*)(u) &= -\sigma_{T^2_\Gamma(\bar{x};u)}(x^*)\\ &= \min_p\{\langle p, -x^*\rangle | \nabla G(\bar{x})p + \nabla^2 G(\bar{x})(u,u) \in T^2_D(G(\bar{x}); \nabla G(\bar{x})u)\}\\ &= \min_p\langle p, -x^*\rangle + \delta_{T^2_D(G(\bar{x});\nabla G(\bar{x})u)}\big(\nabla G(\bar{x})p + \nabla^2 G(\bar{x})(u,u)\big).\end{aligned}$$

The conjugate dual problem of the above minimization problem takes the form

$$\max_{y^*}\left\{\min_p\langle p, -x^*\rangle + \langle y^*, \nabla G(\bar{x})p + \nabla^2 G(\bar{x})(u,u)\rangle - \sigma_{T^2_D(G(\bar{x});\nabla G(\bar{x})u)}(y^*)\right\};$$

see, e.g., [5, equation (2.298)]. Note that by Proposition 2.11,

$$\sigma_{T^2_D(G(\bar{x});\nabla G(\bar{x})u)}(y^*) = \sigma_{T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)}(y^*) = \begin{cases} 0 & y^* \in \widehat{N}_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u) \\ +\infty & \text{otherwise.} \end{cases}$$

Hence, the dual problem can be rewritten equivalently as

$$\begin{array}{ll} \sup_{y^*} & \langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle \\ \text{s.t.} & x^* = \nabla G(\bar{x})^T y^*,\ y^* \in \widehat{N}_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u) \end{array} \iff \sup_{y^* \in \Lambda^s_{x^*}(\bar{x};u)} \langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle.$$

Consequently, by the weak duality [5], we have

$$\mathrm{d}^2\delta_\Gamma(\bar{x};x^*)(u) \geq \sup_{y^* \in \Lambda^s_{x^*}(\bar{x};u)} \langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle.$$

According to Theorem 5.1(ii), we observe that when the directional S- and M-multipliers coincide, the following equality holds:

$$\mathrm{d}^2\delta_\Gamma(\bar{x};x^*)(u) = \max_{y^* \in \Lambda_{x^*}(\bar{x};u)} \langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle.$$

Now we further require that D is convex polyhedral. Consider the critical cone defined by $K(\bar{x};x^*) := \{u \in \{x^*\}^\perp \mid \nabla G(\bar{x})u \in T_D(G(\bar{x}))\}$. Then for any critical direction $u \in K(\bar{x};x^*)$, it turns out that all S- and M-multipliers coincide with the set of (nondirectional) multipliers $\Lambda_{x^*}(\bar{x}) := \{y^* \in N_D(G(\bar{x})) \mid x^* = \nabla G(\bar{x})^T y^*\}$. Consequently, in the following corollary, by using Theorem 5.1(ii), we can recover the result [41, Exercise 13.17] under a weaker condition with the metric regularity replaced by the metric subregularity. It should be noted that the following result can also be obtained from [29, Example 3.4 and Theorem 5.6] by specifying the general convex set considered therein to be convex polyhedral.

Corollary 5.4. *Assume that D is convex polyhedral and that* MSCQ *holds at $\bar{x} \in \Gamma = G^{-1}(D)$. Then for any $x^* \in N_\Gamma(\bar{x})$ and any $u \in \mathbb{R}^n$, one has*

$$\mathrm{d}^2\delta_\Gamma(\bar{x};x^*)(u) = \delta_{K(\bar{x};x^*)}(u) + \max_{y^* \in \Lambda_{x^*}(\bar{x})} \langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle. \tag{5.15}$$

Proof. According to Proposition 2.14 together with [41, Theorem 6.14], we have

$$N_\Gamma(\bar{x}) \subset \nabla G(\bar{x})^T N_D(G(\bar{x})) = \nabla G(\bar{x})^T \widehat{N}_D(G(\bar{x})) \subset \widehat{N}_\Gamma(\bar{x}),$$

showing $N_\Gamma(\bar{x}) = \widehat{N}_\Gamma(\bar{x})$ as well as $\Lambda_{x^*}(\bar{x}) \neq \emptyset$ for $x^* \in N_\Gamma(\bar{x})$. Further, by Theorem 5.1(ii), we know that

$$\widehat{N}^p_\Gamma(\bar{x}) = \widehat{N}^p_\Gamma(\bar{x};0) = \widehat{N}_{T_\Gamma(\bar{x})}(0) = \widehat{N}_\Gamma(\bar{x}) = N_\Gamma(\bar{x}).$$

If $u \notin K(\bar{x};x^*)$, we claim that $\mathrm{d}^2\delta_\Gamma(\bar{x};x^*)(u) = \infty$. Indeed, we have either $u \notin T_\Gamma(\bar{x})$ or $u \in T_\Gamma(\bar{x})$ but $\langle x^*, u\rangle \neq 0$, in which case we must have $\langle x^*, u\rangle < 0$ since $\langle x^*, u\rangle \leq 0$ due to $x^* \in N_\Gamma(\bar{x}) = \widehat{N}_\Gamma(\bar{x}) = (T_\Gamma(\bar{x}))^\circ$. In either case, however, Proposition 2.18(i) yields $\mathrm{d}^2\delta_\Gamma(\bar{x};x^*)(u) = \infty$. Since $\Lambda_{x^*}(\bar{x}) \neq \emptyset$ as shown above, we get $\sup_{y^* \in \Lambda_{x^*}(\bar{x})} \langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle > -\infty$, and thus (5.15) holds.

If $u \in K(\bar{x};x^*)$, then

$$\begin{aligned} \Lambda_{x^*}(\bar{x};u) = \Lambda^s_{x^*}(\bar{x};u) &= \{y^* \mid y^* \in \widehat{N}_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u),\ x^* = \nabla G(\bar{x})^T y^*\} \\ &= \{y^* \mid y^* \in N_D(G(\bar{x})),\ \langle y^*, \nabla G(\bar{x})u\rangle = 0,\ x^* = \nabla G(\bar{x})^T y^*\} \\ &= \{y^* \mid y^* \in N_D(G(\bar{x})),\ \langle x^*, u\rangle = 0,\ x^* = \nabla G(\bar{x})^T y^*\} \\ &= \{y^* \mid y^* \in N_D(G(\bar{x})),\ x^* = \nabla G(\bar{x})^T y^*\} \\ &= \Lambda_{x^*}(\bar{x}), \end{aligned}$$

where the first equality is due to the convexity of D (see (2.1)) and the fifth equality comes from the fact that $\langle x^*, u\rangle = 0$ holds automatically as $u \in K(\bar{x}; x^*)$. Since $x^* \in N_\Gamma(\bar{x}) = \widehat{N}^p_\Gamma(\bar{x}) \subset \widehat{\mathcal{N}}^p_\Gamma(\bar{x}; u)$ and $\langle x^*, u\rangle = 0$, we have $x^* \in \widehat{N}^p_\Gamma(\bar{x}; u)$. It then follows from (5.3a) and (5.3b) that

$$\mathrm{d}^2\delta_\Gamma(\bar{x}; x^*)(u) = \sup_{y^* \in \Lambda_{x^*}(\bar{x})} \langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle = \sup_{y^* \in \Lambda_{x^*}(\bar{x}) \cap \kappa\|x^*\| \mathrm{cl}\,\mathbb{B}} \langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle.$$

Since $\Lambda_{x^*}(\bar{x}) \cap \kappa\|x^*\| \mathrm{cl}\,\mathbb{B}$ is a compact set, the supremum can be attained, and so it can be replaced by the maximum. This completes the proof. ∎

5.2. Nondegenerate systems. As we have seen in Corollary 5.4, the results from Theorem 5.1 get considerably simpler if the set D is convex polyhedral. Here we continue with simplifications, but we keep D arbitrary polyhedral and strengthen the assumptions on constraints instead. We start with the condition

$$\nabla G(\bar{x})^T y^* = 0,\ y^* \in \mathrm{span}\, N_D(G(\bar{x}); \nabla G(\bar{x})u) \implies \langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle = 0, \tag{5.16}$$

which does not yield uniqueness of $y^* \in \Lambda_{x^*}(\bar{x}; u)$, but it implies that the value $\langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle$ is the same for all the multipliers, making the lower and upper bounds in (5.3a) and (5.4) equal. Since (5.16) is in general weaker than the directional nondegeneracy condition (4.1), we refer to it as the generalized directional nondegeneracy condition. Moreover, we present the second-order tangent cone $T^2_\Gamma(\bar{x}; u)$ as a translation of a cone.

Proposition 5.5. *Let $\bar{x} \in \Gamma$ and $u \in L_\Gamma(\bar{x})$, and suppose that MSCQ holds at $\bar{x}$ in direction u. Assume that the generalized directional nondegeneracy condition (5.16) holds. Then for any x^*, the quantity $\langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle$ is the same for all $y^* \in \Lambda_{x^*}(\bar{x}; u)$, i.e.,*

$$\sup_{y^* \in \Lambda_{x^*}(\bar{x};u)} \langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle = \inf_{y^* \in \Lambda_{x^*}(\bar{x};u)} \langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle \tag{5.17}$$

and there is some p_0 satisfying

$$\nabla G(\bar{x})p_0 + \nabla^2 G(\bar{x})(u,u) \in \mathcal{L}(T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)). \tag{5.18}$$

For every such p_0, we have the representation

$$T^2_\Gamma(\bar{x}; u) = p_0 + K_{\bar{x};u}, \tag{5.19}$$

where $K_{\bar{x};u} := \{p \mid \nabla G(\bar{x})p \in T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)\}$ is a cone.

Proof. Consider $y_1^*, y_2^* \in \Lambda_{x^*}(\bar{x}; u)$. By definition, we have $\nabla G(\bar{x})^T(y_1^* - y_2^*) = x^* - x^* = 0$ and $y_1^* - y_2^* \in N_D(G(\bar{x}); \nabla G(\bar{x})u) - N_D(G(\bar{x}); \nabla G(\bar{x})u) \subset \mathrm{span}\, N_D(G(\bar{x}); \nabla G(\bar{x})u)$. It follows by (5.16) that

$$\langle y_1^* - y_2^*, \nabla^2 G(\bar{x})(u,u)\rangle = 0.$$

Consequently, (5.17) holds. From (5.16) and Proposition 2.6, we infer that

$$\begin{aligned}\nabla^2 G(\bar{x})(u,u) &\in \left(\ker \nabla G(\bar{x})^T \cap \mathrm{span}\, N_D(G(\bar{x}); \nabla G(\bar{x})u)\right)^\perp \\ &\subset \nabla G(\bar{x})\mathbb{R}^n + \mathcal{L}(T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)),\end{aligned}$$

which yields the existence of p_0 satisfying (5.18). In fact, by (5.1), we know that $p_0 \in T^2_\Gamma(\bar{x};u)$.

For every $p \in K_{\bar{x};u}$, we have

$$\begin{aligned}\nabla G(\bar{x})(p_0+p) + \nabla^2 G(\bar{x})(u,u) &\in T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u) + \mathcal{L}(T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u))\\ &= T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u),\end{aligned}$$

showing $p_0 + p \in T^2_\Gamma(\bar{x};u)$ by (5.1). On the other hand, if $p \in T^2_\Gamma(\bar{x};u)$, then by (5.1), we obtain

$$\nabla G(\bar{x})(p-p_0) \in T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u) - \mathcal{L}(T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)) = T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u),$$

i.e., $p - p_0 \in K_{\bar{x};u}$. This verifies $T^2_\Gamma(\bar{x};u) = p_0 + K_{\bar{x};u}$, and the proof is complete. □

Combining Propositions 5.5 and 2.16 together yields the following result.

COROLLARY 5.6. *Let $\bar{x} \in \Gamma$ and $u \in L_\Gamma(\bar{x})$, and suppose that MSCQ holds at $\bar{x}$ in direction u. If the generalized directional nondegeneracy condition* (5.16) *is fulfilled, then*

$$\mathrm{d}^2\delta_\Gamma(\bar{x};x^*)(u) = -\sigma_{T^2_\Gamma(\bar{x};u)}(x^*) = \langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle \;\; \forall x^* \in \operatorname{dom}\sigma_{T^2_\Gamma(\bar{x};u)} = \widehat{N}^p_\Gamma(\bar{x};u), \tag{5.20a}$$

$$-\hat{\sigma}_{T^2_\Gamma(\bar{x};u)}(x^*) = \langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle = -\langle x^*, p_0\rangle \;\; \forall x^* \in \operatorname{dom}\hat{\sigma}_{T^2_\Gamma(\bar{x};u)} = N_{K_{\bar{x};u}}(0) \tag{5.20b}$$

where y^* is an arbitrary element from $\Lambda_{x^*}(\bar{x};u)$ and p_0 is an arbitrary vector satisfying (5.18)*, respectively.*

Proof. Note that (5.20a) holds by Theorem 5.1(i)(ii) and Proposition 5.5. Let $x^* \in \operatorname{dom}\hat{\sigma}_{T^2_\Gamma(\bar{x};u)}$. Then, by Theorem 5.1(iii) and Proposition 5.5, for any $y^* \in \Lambda_{x^*}(\bar{x};u)$, we have

$$-\hat{\sigma}_{T^2_\Gamma(\bar{x};u)}(x^*) = \langle y^*, \nabla^2 G(\bar{x})(u,u)\rangle.$$

Moreover, by Proposition 5.5, there exists p_0 satisfying (5.18) such that $T^2_\Gamma(\bar{x};u) = p_0 + K_{\bar{x};u}$. Hence, (5.20b) follows from Proposition 2.16. □

Remark 5.7. Under the assumptions of Corollary 5.6, we have

$$\mathrm{d}^2\delta_\Gamma(\bar{x};x^*)(u) = -\sigma_{T^2_\Gamma(\bar{x};u)}(x^*) = -\hat{\sigma}_{T^2_\Gamma(\bar{x};u)}(x^*) \quad \forall x^* \in \operatorname{dom}\sigma_{T^2_\Gamma(\bar{x};u)} \subset \operatorname{dom}\hat{\sigma}_{T^2_\Gamma(\bar{x};u)}.$$

Thus, whenever $\mathrm{d}^2\delta_\Gamma(\bar{x};x^*)(u)$ differs from $-\hat{\sigma}_{T^2_\Gamma(\bar{x};u)}(x^*)$, there holds that $\sigma_{T^2_\Gamma(\bar{x};u)}(x^*) = \infty$.

In general, we only know an inclusion

$$\begin{aligned}N_{K_{\bar{x};u}}(0) \subset \nabla G(\bar{x})^T N_{T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)}(0) &= \nabla G(\bar{x})^T N_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)\\ &= \{x^* \mid \Lambda_{x^*}(\bar{x};u) \neq \emptyset\}.\end{aligned}$$

If we strengthen (5.16) to (4.1), however, this inclusion holds with equality, the multipliers become unique, and we are also able to give an alternative representation of the set $\widehat{N}^p_\Gamma(\bar{x};u)$.

COROLLARY 5.8. *Let $\bar{x} \in \Gamma$ and $u \in L_\Gamma(\bar{x})$. Under the directional nondegeneracy condition* (4.1)*, the following statements hold:*

(i) *We have* $\operatorname{dom}\sigma_{T^2_\Gamma(\bar{x};u)} = \nabla G(\bar{x})^T \widehat{N}_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u) = \widehat{N}^p_\Gamma(\bar{x};u)$, *and for every* $x^* \in \operatorname{dom}\sigma_{T^2_\Gamma(\bar{x};u)}$, *the set* $\Lambda^s_{x^*}(\bar{x};u)$ *is a singleton* $\{y^*_0\}$ *and* $\mathrm{d}^2\delta_\Gamma(\bar{x};x^*)(u) = -\sigma_{T^2_\Gamma(\bar{x};u)}(x^*) = \langle y^*_0, \nabla^2 G(\bar{x})(u,u)\rangle$.

(ii) *We have* $\operatorname{dom}\hat{\sigma}_{T^2_\Gamma(\bar{x};u)} = \nabla G(\bar{x})^T N_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u) = N_\Gamma(\bar{x};u)$, *and for every* $x^* \in \operatorname{dom}\hat{\sigma}_{T^2_\Gamma(\bar{x};u)}$, *the set* $\Lambda_{x^*}(\bar{x};u)$ *is a singleton* $\{y^*_0\}$ *and* $-\hat{\sigma}_{T^2_\Gamma(\bar{x};u)}(x^*) = \langle y^*_0, \nabla^2 G(\bar{x})(u,u)\rangle$.

Proof. Recall that (4.1) implies (2.17), which further ensures MSCQ at $\bar{x}$ in direction u.

(i) It follows from Theorem 5.1 and Proposition 2.14 that

$$\operatorname{dom}\sigma_{T^2_\Gamma(\bar{x};u)} = \widehat{N}^p_\Gamma(\bar{x};u) = \widehat{N}_{T_\Gamma(\bar{x})}(u) = \nabla G(\bar{x})^T \widehat{N}_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u).$$

Particularly, this shows that $\Lambda^s_{x^*}(\bar{x};u) \neq \emptyset$. Taking into account $\Lambda^s_{x^*}(\bar{x};u) \subset \Lambda_{x^*}(\bar{x};u)$, the remaining claims follow from Corollary 5.6 once we prove that $\Lambda_{x^*}(\bar{x};u)$ is a singleton in the next step.

(ii) Note that $N_{T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)}(0) = N_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u) = N_D(G(\bar{x});\nabla G(\bar{x})u)$ by Proposition 2.11. Hence, applying Theorem 4.1 to the set $K_{\bar{x};u} := \{p \mid \nabla G(\bar{x})p \in T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)\}$ at $p = 0$ yields

$$\begin{aligned}\operatorname{dom}\hat{\sigma}_{T^2_\Gamma(\bar{x};u)} &= N_{K_{\bar{x};u}}(0) = \nabla G(\bar{x})^T N_{T_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u)}(0)\\ &= \nabla G(\bar{x})^T N_{T_D(G(\bar{x}))}(\nabla G(\bar{x})u) = N_\Gamma(\bar{x};u),\end{aligned}$$

where the first equality comes from (5.20b) and the last equality follows from (4.2). Let $x^* \in \operatorname{dom}\hat{\sigma}_{T^2_\Gamma(\bar{x};u)} = N_\Gamma(\bar{x};u)$. Suppose that $y^*_1, y^*_2 \in \Lambda_{x^*}(\bar{x};u)$. Then $\nabla G(\bar{x})^T(y^*_1 - y^*_2) = 0$ and

$$y^*_1 - y^*_2 \in N_D(G(\bar{x});\nabla G(\bar{x})u) - N_D(G(\bar{x});\nabla G(\bar{x})u) \subset \operatorname{span} N_D(G(\bar{x});\nabla G(\bar{x})u).$$

The nondegeneracy condition (4.1) yields $y^*_1 = y^*_2$, which implies that the set $\Lambda_{x^*}(\bar{x};u)$ contains a unique element, say, y^*_0. Corollary 5.6 now completes the proof. ∎

6. Application: Second-order conditions for disjunctive programs.

Let us first reiterate that our ultimate goal is to investigate problems (GP) with set C having the complex structure (1.3). As explained, this has to be postponed until we complete the full analysis in the forthcoming paper [4]. In this section, we only provide a simple application of our results to the disjunctive program defined as

$$\text{(DP)} \qquad \begin{aligned} &\min \ f(x)\\ &\text{s.t.} \ g(x) \in D,\end{aligned}$$

where g is twice continuously differentiable and D is polyhedral. Several classes of interesting mathematical programs of practical interest can be reformulated as a (DP), including the mathematical program with equilibrium constraints (MPEC) (cf. [21, 27, 33, 42]), the mathematical program with vanishing constraints (cf. [1, 22, 23]), the mathematical program with switching constraints (cf. [25, 26]), and the mathematical program with cardinality constraints (cf. [6, 7]). A discussion on constraint qualifications of disjunctive programming can be found in [12, 28] and references therein.

Using the second-order variational analysis of the disjunctive system, we can now recover the second-order optimality conditions for the (DP) derived by Gfrerer in [12].

Moreover, using the calculations for directional S- and M-multiplier sets for MPECs in [12], we can easily obtain the corresponding second-order optimality conditions for MPECs from Theorem 6.1.

THEOREM 6.1 ([12, Theorems 3.3 and 3.17]). *Let $\bar{x}$ be a local optimal solution of the disjunctive program (DP). Then the following necessary optimality conditions hold:*

(i) *For $u \in \mathcal{C}(\bar{x})$, suppose that $x \rightrightarrows g(x) - D$ is metrically subregular in direction u at $(\bar{x}, 0)$. Then there exists $\lambda \in \Lambda(\bar{x}; u)$ such that*

$$\nabla^2_{xx} L(\bar{x}, \lambda)(u, u) \geq 0.$$

(ii) *For $u \in \mathcal{C}(\bar{x})$, assume that the nondegeneracy condition in direction u,*

$$\nabla g(\bar{x})^T y^* = 0,\ y^* \in \operatorname{span} N_D(g(\bar{x}); \nabla g(\bar{x})u) \implies y^* = 0,$$

is fulfilled. Then $\nabla^2_{xx} L(\bar{x}, \lambda)(u, u) \geq 0$ holds with the unique directional S-multiplier $\lambda \in \Lambda^s(\bar{x}; u)$.

Conversely, suppose that $\bar{x}$ is a feasible solution of the disjunctive program (DP). Suppose that for each nonzero $u \in \mathcal{C}(\bar{x})$, there are α and λ, not both equal to zero, with $\alpha \geq 0, \lambda \in \widehat{N}^p_D(g(\bar{x}); \nabla g(\bar{x})u) = \widehat{N}_{T_D(g(\bar{x}))}(\nabla g(\bar{x})u)$, such that

$$\nabla^2_{xx} L^\alpha(\bar{x}, \lambda)(u, u) > 0.$$

Then $\bar{x}$ is an essential local minimizer of second order.

Proof. To obtain the necessary optimality conditions, it suffices to calculate $\hat{\sigma}_{T^2_D(g(\bar{x}); \nabla g(\bar{x})u)}(\lambda)$ and apply Theorem 3.1. Under the assumptions, the second-order necessary optimality condition (3.1) holds with $C := D$. It follows that $\hat{\sigma}_{T^2_D(g(\bar{x}); \nabla g(\bar{x})u)}(\lambda) < \infty$, which ensures that $\lambda \in \operatorname{dom} \hat{\sigma}_{T^2_D(g(\bar{x}); \nabla g(\bar{x})u)}$. Taking G to be the identity mapping in Theorem 5.1(iii), we obtain $\Gamma = D$ and $\hat{\sigma}_{T^2_D(g(\bar{x}); \nabla g(\bar{x})u)}(\lambda) = 0$. Hence, the necessary optimality conditions (i) and (ii) hold for (DP).

To obtain the sufficient optimality condition, it suffices to calculate $\mathrm{d}^2 \delta_D(g(\bar{x}); \lambda)(\nabla g(\bar{x})u)$ and apply Theorem 3.3 with $C := D$. Again, applying Theorem 5.1(ii) with G being the identity mapping, we have $\mathrm{d}^2 \delta_D(g(\bar{x}); \lambda)(\nabla g(\bar{x})u) = 0$ since $\lambda \in \widehat{N}^p_D(g(\bar{x}); \nabla g(\bar{x})u)$, and the result follows. ∎

Note that we have only used the results from section 5 for the trivial identity mapping. Their full potential will be seen when applied to sets of the form (1.3).

7. Concluding remarks. In this paper, we have reviewed the second-order necessary optimality conditions and derived second-order sufficient optimality conditions for the general problem (GP). Since these conditions involve some second-order objects that need to be calculated or estimated, we have conducted second-order variational analysis of disjunctive systems. As an illustration, we have shown that one can recover second-order optimality conditions for disjunctive programs. In the forthcoming work [4], using the analysis of disjunctive systems from this paper as a tool, we will develop the variational analysis of the set given by (1.3), which will enable us to apply our second-order optimality conditions from Theorems 3.1 and 3.3.

Acknowledgments. The authors are indebted to the anonymous referees for their valuable suggestions that helped us to improve the original presentation of the paper.

REFERENCES

[1] W. Achtziger and C. Kanzow, *Mathematical programs with vanishing constraints: Optimality conditions and constraint qualifications*, Math. Program., 114 (2008), pp. 69–99.

[2] M. Benko, H. Gfrerer, and J. V. Outrata, *Calculus for directional limiting normal cones and subdifferentials*, Set-Valued Var. Anal., 27 (2019), pp. 713–745.

[3] M. Benko, H. Gfrerer, and J. V. Outrata, *Stability analysis for parameterized variational systems with implicit constraints*, Set-Valued Var. Anal., 28 (2020), pp. 167–193.

[4] M. Benko, H. Gfrerer, J. J. Ye, J. Zhang, and J. C. Zhou, *Second-Order Optimality Conditions for Mathematical Programs with Generalized Equation Constraints*, in preparation.

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

[6] M. Bucher and A. Schwartz, *Second-order optimality conditions and improved convergence results for regularization methods for cardinality-constrained optimization problems*, J. Optim. Theory Appl., 178 (2018), pp. 383–410.

[7] M. Červinka, C. Kanzow, and A. Schwartz, *Constraint qualifications and optimality conditions for optimization problems with cardinality constraints*, Math. Program., 160 (2016), pp. 353–377.

[8] F. H. Clarke, *Nonsmooth Analysis and Control Theory*, Wiley-Interscience, New York, 1983.

[9] S. Dempe, *Foundations of Bilevel Programming,* Nonconvex Optimization and Its Applications 61, Springer, New York, 2002.

[10] C. Ding, D. F. Sun, and J. J. Ye, *First order optimality conditions for mathematical programs with semidefinite cone complementarity constraints*, Math. Program., 147 (2014), pp. 539–579.

[11] H. Gfrerer, *On directional metric regularity, subregularity and optimality conditions for nonsmooth mathematical programs*, Set-Valued Var. Anal., 21 (2013), pp. 151–176.

[12] H. Gfrerer, *Optimality conditions for disjunctive programs based on generalized differentiation with application to mathematical programs with equilibrium constraints*, SIAM J. Optim., 24 (2014), pp. 898–931.

[13] H. Gfrerer, *Linearized M-stationarity conditions for general optimization problems*, Set-Valued Var. Anal., 27 (2019), pp. 819–840.

[14] H. Gfrerer and B. S. Mordukhovich, *Second-order variational analysis of parametric constraint and variational systems*, SIAM J. Optim., 29 (2019), pp. 423–453.

[15] H. Gfrerer and J. V. Outrata, *On Lipschitzian properties of implicit multifunctions*, SIAM J. Optim., 26 (2016), pp. 2160–2189.

[16] H. Gfrerer and J. V. Outrata, *On computation of generalized derivatives of the normal-cone mapping and their applications*, Math. Oper. Res., 41 (2016), pp. 1535–1556.

[17] H. Gfrerer and J. J. Ye, *New constraint qualifications for mathematical programs with equilibrium constraints via variational analysis*, SIAM J. Optim., 27 (2017), pp. 842–865.

[18] H. Gfrerer and J. J. Ye, *New sharp necessary optimality conditions for mathematical programs with equilibrium constraints*, Set-Valued Var. Anal., 28 (2020), pp. 395–426.

[19] H. Gfrerer, J. J. Ye, and J. C. Zhou, *Second-order optimality conditions for non-convex set-constrained optimization problems*, Math. Oper. Res., 47 (2022), pp. 2344–2365.

[20] I. Ginchev and B. S. Mordukhovich, *On directionally dependent subdifferentials*, C. R. Bulgare Acad. Sci., 64 (2011), pp. 497–508.

[21] L. Guo, G.-H. Lin, and J. J. Ye, *Second-order optimality conditions for mathematical programs with equilibrium constraints*, J. Optim. Theory Appl., 158 (2013), pp. 33–64.

[22] T. Hoheisel and C. Kanzow, *First- and second-order optimality conditions for mathematical programs with vanishing constraints*, Appl. Math., 52 (2007), pp. 495–514.

[23] T. Hoheisel and C. Kanzow, *Stationary conditions for mathematical programs with vanishing constraints using weak constraint qualifications*, J. Math. Anal. Appl., 337 (2008), pp. 292–310.

[24] A. D. Ioffe, *Variational Analysis of Regular Mappings*, Springer Monographs in Mathematics, Springer, Cham, Switzerland, 2017.

[25] C. Kanzow, P. Mehlitz, and D. Steck, *Relaxation schemes for mathematical programmes with switching constraints*, Optim. Methods Softw., 36 (2021), pp. 1223–1258.

[26] Y. C. Liang and J. J. Ye, *New optimality conditions and exact penalty for mathematical programs with switching constraints*, J. Optim. Theory Appl., 190 (2021), pp. 1–31.

[27] Z.-Q. Luo, J.-S. Pang, and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, 1996.

[28] P. Mehlitz, *On the linear independence constraint qualification in disjunctive programming*, Optimization, 69 (2020), pp. 2241–2277.

[29] A. Mohammadi, B. S. Mordukhovich, and E. Sarabi, *Parabolic regularity in geometric variational analysis*, Trans. Amer. Math. Soc., 374 (2021), pp. 1711–1763.
[30] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation*, Vols. 1 and 2, Springer, Berlin, 2006.
[31] B. S. Mordukhovich, *Variational Analysis and Applications,* Springer Monographs in Mathematics, Springer, Cham, Switzerland, 2018.
[32] 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.
[33] J. V. Outrata, M. Kocvara, and J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints,* Nonconvex Optimization and Its Applications 28, Springer, New York, 1998.
[34] J. V. Outrata and D. F. Sun, *On the coderivative of the projection operator onto the second-order cone*, Set-Valued Anal., 16 (2008), pp. 999–1014.
[35] J.-P. Penot, *Second-order conditions for optimization problems with constraints*, SIAM J. Control Optim., 37 (1998), pp. 303–318.
[36] R. Penrose, *On best approximate solutions of linear matrix equations*, Proc. Cambridge Philos. Soc., 52 (1956), pp. 17–19.
[37] S. M. Robinson, *Some continuity properties of polyhedral multifunctions*, Math. Program. Stud., 14 (1981), pp. 206–214.
[38] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
[39] R. T. Rockafellar, *Augmented Lagrangians and hidden convexity in sufficient conditions for local optimality*, Math. Program., 198 (2023), pp. 159–194.
[40] R. T. Rockafellar, *Convergence of augmented Lagrangian methods in extensions beyond nonlinear programming*, Math. Program., 199 (2023), pp. 375–420, https://doi.org/10.1007/s10107-022-01832-5.
[41] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.
[42] S. Scheel and S. Scholtes, *Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity*, Math. Oper. Res., 25 (2000), pp. 1–22.
[43] J. J. Ye and J. C. Zhou, *First-order optimality conditions for mathematical programs with second-order cone complementarity constraints*, SIAM J. Optim., 26 (2016), pp. 2820–2846.
[44] J. J. Ye and J. C. Zhou, *Exact formulas for the proximal/regular/limiting normal cone of the second-order cone complementarity set*, Math. Program., 162 (2017), pp. 33–50.
[45] J. J. Ye and D. L. Zhu, *Optimality conditions for bilevel programming problems*, Optimization, 33 (1995), pp. 9–27.