SIAM J. OPTIM.
Vol. 29, No. 2, pp. 1524–1557

CRITICALITY OF LAGRANGE MULTIPLIERS IN VARIATIONAL SYSTEMS*

BORIS MORDUKHOVICH† AND EBRAHIM SARABI‡

Abstract. The paper concerns the study of criticality of Lagrange multipliers in variational systems, which has been recognized in both theoretical and numerical aspects of optimization and variational analysis. In contrast to the previous developments dealing with polyhedral KKT systems and the like, we focus on general nonpolyhedral systems that are associated, in particular, with problems of conic programming. Developing a novel approach that is mainly based on advanced techniques and tools of second-order variational analysis and generalized differentiation allows us to overcome the principal challenges of nonpolyhedrality and to establish complete characterizations on noncritical multipliers in such settings. The results are illustrated by examples from semidefinite programming.

Key words. optimization and variational analysis, generalized KKT systems, critical and noncritical multipliers, second-order generalized differentiation, error bounds, calmness

AMS subject classifications. 90C31, 49J52, 49J53

DOI. 10.1137/18M1206862

1. Introduction.

This paper is devoted to investigating some core issues of optimization and variational analysis that revolve around *criticality* of dual elements (Lagrange multipliers) in the corresponding KKT systems. The motivation to study multiplier criticality came from applications to convergence rates of primal-dual algorithms for numerical optimization. Then it was realized that understanding these issues requires a careful theoretical investigation that reveals, in particular, close interrelations between criticality and other fundamental concepts of variational analysis and generalized differentiation that are themselves of interest.

The notion of criticality, i.e., critical and noncritical Lagrange multipliers, was introduced by Izmailov [13] for $\mathcal{C}^2$-smooth nonlinear programs (NLPs) with equality constraints. It was recognized from the very beginning that the existence of critical multipliers is the main reason to prevent superlinear convergence of primal iterations in Newtonian methods, since such multipliers persistently attract convergence of dual components. Theoretical and computational issues concerning this phenomenon in NLPs and related variational inequalities were analyzed in many publications and in the monograph by Izmailov and Solodov [15]. We also refer the reader to their excellent survey [16], which is devoted to various particular aspects of multiplier criticality in major primal-dual methods of nonlinear programming; see also the comments by Fischer, Martinez, Mordukhovich, and Robinson about this survey.[1]

*Received by the editors August 13, 2018; accepted for publication (in revised form) March 8, 2019; published electronically June 4, 2019.
http://www.siam.org/journals/siopt/29-2/M120686.html
Funding: Research of the first author was partly supported by the National Science Foundation under grants DMS-1512846 and DMS-1808978, and by the Air Force Office of Scientific Research under grant 15RT0462.
†Department of Mathematics, Wayne State University, Detroit, MI 48202 (boris@math.wayne.edu).
‡Department of Mathematics, Miami University, Oxford, OH 45065 (sarabim@miamioh.edu).
[1]See https://link.springer.com/journal/11750/23/1.

A striking property of noncritical Lagrange multipliers is that they yield a certain stability (calmness) property of solution maps to canonical perturbed KKT systems, which in turn helps to establish *superlinear* convergence for Newtonian methods. For instance, Izmailov and Solodov [14] prove in this way that, in the nonlinear programming framework, convergence to a noncritical Lagrange multiplier ensures a superlinear rate of convergence of primal-dual iterations in the stabilized sequential quadratic programming method even when the problem is degenerate, i.e., the corresponding set of Lagrange multipliers is not a singleton.

Our recent paper [20] conducts a systematic study of criticality for *polyhedral* variational systems (generalized KKT) that cover a significantly larger territory than NLPs. Employing advanced tools of second-order variational analysis and generalized differentiation, we obtain therein several characterizations of critical and noncritical multipliers and establish their connections with other fundamental or novel properties of variational systems. In particular, we show in [20] that the well-recognized and comprehensively characterized property of *full stability* of local minimizers in polyhedral problems of constrained optimization allows us to exclude the appearance of critical multipliers associated with such minimizers.

The current paper studies criticality for the class of *nonpolyhedral* variational systems described in the generalized KKT form

$$\Psi(x,\lambda) := f(x) + \nabla\Phi(x)^*\lambda = 0, \qquad \lambda \in N_\Theta(\Phi(x)), \tag{1.1}$$

where $f\colon \mathbb{X} \to \mathbb{X}$ is a differentiable mapping, while $\Phi\colon \mathbb{X} \to \mathbb{Y}$ is a twice differentiable mapping in the classical sense of between finite-dimensional spaces, where $\Theta \subset Y$ is a closed set with N_Θ standing for its (limiting) normal cone (2.3), and where the asterisk signifies the matrix transposition/adjoint operator.

A major source for the generalized KKT system (1.1) comes from the first-order necessary optimality conditions for constrained optimization problems. Indeed, consider a differentiable function $\varphi_0\colon \mathbb{X} \to \mathbb{R}$ and define the following constrained optimization problem:

$$\text{minimize } \ \varphi_0(x) \ \text{ subject to } \ \Phi(x) \in \Theta, \tag{1.2}$$

where Φ and Θ are taken from (1.1). It is well known that, under a certain constraint qualification, system (1.1) with $f := \nabla\varphi_0$ gives us necessary optimality conditions for (1.2).

Despite a good understanding of noncriticality for systems (1.1) with polyhedral sets Θ, not much has been done in the case of nonpolyhedrality. The results established recently in [25, Theorem 3.3] and [17, Proposition 4.2] do not provide a satisfactory picture in this regard. Indeed, the assumptions imposed therein are so strong that they may not be satisfied even for classical problems of nonlinear programming.

This paper develops a novel approach to the study of critical and noncritical Lagrange multipliers associated with (1.1), where Θ belongs to a rather general class of regular sets that includes, in particular, all the convex ones. The new notion of *semi-isolated calmness* is crucial for our characterizations of noncritical multipliers and subsequent applications. Prior to a detailed consideration of this property, let us emphasize the following: (1) it is strictly weaker than the isolated calmness used, e.g., in [2, 15] to justify superlinear convergence of the sequential quadratic programming (SQP) method for nonlinear programs; and (2) it allows us to deal with optimization problems admitting nonunique Lagrange multipliers.

It is important to realize that the generalized KKT systems (1.1) with nonpolyhedral sets Θ fail to satisfy some properties that are granted under polyhedrality. In particular, the semi-isolated calmness property for polyhedral systems (1.1) follows from the uniqueness and noncriticality of Lagrange multipliers. However, this is not the case for nonpolyhedral systems, as revealed by Example 5.8. This occurs due to the lack of a certain error bound, which is guaranteed by the Hoffman lemma in polyhedral settings. To overcome this challenge, we first establish new characterizations of *uniqueness* of Lagrange multipliers combined with some error bound. This plays a significant role in deriving our main result, Theorem 5.6, which provides a complete *characterization of noncriticality* under a general reducibility assumption.

The rest of the paper is organized as follows. Section 2 recalls some basic concepts of variational analysis and generalized differentiation utilized below. In section 3 we define critical and noncritical multipliers for system (1.1) together with an extended notion of $\mathcal{C}^2$-reducibility of Θ and then provide elaborations of these notions for major models of conic programming. Section 4 establishes new characterization of uniqueness of Lagrange multipliers in nonpolyhedral systems. In section 5 we develop a reduction approach for the study of criticality of multipliers in (1.1) under the $\mathcal{C}^2$-reducibility of Θ and in this way establish verifiable characterizations of noncritical multipliers with relationships to semi-isolated calmness. Furthermore, we show that the assumptions required for the obtained characterizations hold under the well-known strict complementarity condition.

Our notation and terminology are standard in variational analysis and generalized differentiation; see, e.g., [18, 22]. Recall that, given a nonempty set Ω in $\mathbb{X}$, the notation $\operatorname{bd}\Omega$, $\operatorname{int}\Omega$, $\operatorname{ri}\Omega$, $\operatorname{cl}\Omega$, Ω^*, $\operatorname{aff}\Omega$, and $\operatorname{span}\Omega$ stands for the boundary, interior, relative interior, closure, polar, affine hull of Ω, and the smallest linear subspace containing Ω, respectively. The notation $x \xrightarrow{\Omega} \bar{x}$ indicates that $x \to \bar{x}$ with $x \in \Omega$. By $\mathbb{B}$ we denote the closed unit ball in the space in question, while $\mathbb{B}_r(x) := x + r\mathbb{B}$ stands for the closed ball centered at x with radius $r > 0$. The indicator function of Ω is defined by $\delta_\Omega(x) := 0$ for $x \in \Omega$ and by $\delta_\Omega(x) := \infty$ otherwise. Denote by $\operatorname{diag}(a_1, \ldots, a_m)$ an $m \times m$ diagonal matrix whose diagonal entries are $a_1, \ldots, a_m$. We write $x = o(t)$ with $x \in \mathbb{X}$ and $t \in \mathbb{R}_+$ to indicate as usual that $\|x\|/t \to 0$ as $t \downarrow 0$. Finally, denote by $\mathbb{R}_+$ (respectively, $\mathbb{R}_-$) the set of nonnegative (respectively, nonpositive) real numbers.

2. Preliminaries from variational analysis. In this section we first briefly review basic constructions of variational analysis and generalized differentiation employed in the paper, which mainly follow those in the books [18, 22].

Given a set $\Omega \subset \mathbb{X}$, the (Bouligand–Severi) *tangent cone* $T_\Omega(\bar{x})$ to Ω at $\bar{x} \in \Omega$ is defined by

$$T_\Omega(\bar{x}) := \{w \in \mathbb{X} \mid \exists\, t_k \downarrow 0,\ w_k \to w \ \text{ as } \ k \to \infty \text{ with } \bar{x} + t_k w_k \in \Omega\}. \tag{2.1}$$

The (Fréchet) *regular normal cone* to Ω at $\bar{x} \in \Omega$ is

$$\widehat{N}_\Omega(\bar{x}) := \left\{ v \in \mathbb{X} \;\middle|\; \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle v, x - \bar{x}\rangle}{\|x - \bar{x}\|} \le 0 \right\}, \tag{2.2}$$

or, equivalently, $\widehat{N}_\Omega(\bar{x}) = T_\Omega(\bar{x})^*$. The (limiting/Mordukhovich) *normal cone* to Ω at $\bar{x} \in \Omega$ is defined by

$$N_\Omega(\bar{x}) = \{v \in \mathbb{X} \mid \exists\, x_k \to \bar{x},\ v_k \to v \ \text{ with } \ v_k \in \widehat{N}_\Omega(x_k)\}. \tag{2.3}$$

If Ω is convex, both constructions (2.2) and (2.3) reduce to the classical normal cone of convex analysis. The set Ω is called (normally) *regular* at $\bar{x} \in \Omega$ if $\widehat{N}_\Omega(\bar{x}) = N_\Omega(\bar{x})$.

In contrast to (2.2), the normal cone (2.3) and the associated constructions for functions and mappings enjoy comprehensive calculus rules based on variational/extremal principles of variational analysis.

Given an extended-real-valued function $f\colon \mathbb{X} \to \overline{\mathbb{R}} := (-\infty, \infty]$ that is finite at $\bar{x}$, the *subdifferential* of f at $\bar{x}$ is defined via the normal cone to its epigraph $\operatorname{epi} f := \{(x, \alpha) \in \mathbb{X} \times \mathbb{R} \mid \alpha \geq f(x)\}$ by

$$\partial f(\bar{x}) := \{v \in \mathbb{X} \mid (v, -1) \in N_{\operatorname{epi} f}(\bar{x}, f(\bar{x}))\}. \tag{2.4}$$

Considering next a set-valued mapping $F\colon \mathbb{X} \rightrightarrows \mathbb{Y}$ with its domain and graph given by

$$\operatorname{dom} F := \{x \in \mathbb{X} \mid F(x) \neq \emptyset\} \text{ and } \operatorname{gph} F := \{(x, y) \in \mathbb{X} \times \mathbb{Y} \mid x \in F(x)\},$$

we define the *graphical derivative* of F at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ by

$$DF(\bar{x}, \bar{y})(u) := \{v \in \mathbb{Y} \mid (u, v) \in T_{\operatorname{gph} F}(\bar{x}, \bar{y})\}, \quad u \in \mathbb{X}. \tag{2.5}$$

Finally, we recall the well-posedness properties of set-valued mappings used in what follows. The mapping $F\colon X \rightrightarrows Y$ is *metrically regular* around $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ if there exist $\ell \geq 0$ together with neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$d(x; F^{-1}(y)) \leq \ell\, d(y; F(x)) \quad \text{for all } (x, y) \in U \times V, \tag{2.6}$$

where $d(x; \Omega)$ stands for the distance between x and the set Ω. The *metric subregularity* of F at $(\bar{x}, \bar{y})$ corresponds to the validity of (2.6) with the fixed point $y = \bar{y}$. We say that F is *strongly metrically subregular* at $(\bar{x}, \bar{y})$ if there are $\ell \geq 0$ and a neighborhood U of $\bar{x}$ for which

$$\|x - \bar{x}\| \leq \ell\, d(y; F(x)) \text{ whenever } x \in U.$$

$F\colon \mathbb{X} \rightrightarrows \mathbb{Y}$ is *calm* at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ if there exist a number $\ell \geq 0$ and neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$F(x) \cap V \subset F(\bar{x}) + \ell\|x - \bar{x}\|\mathbb{B} \quad \text{for all } x \in U. \tag{2.7}$$

The *isolated calmness* property of F at $(\bar{x}, \bar{y})$ is defined by

$$F(x) \cap V \subset \{\bar{y}\} + \ell\|x - \bar{x}\|\mathbb{B} \text{ for all } x \in U$$

with some $\ell \geq 0$ and neighborhoods U of $\bar{x}$ and V of $\bar{y}$. It is well known that the calmness and isolated calmness of F at $(\bar{x}, \bar{y})$ are equivalent to the metric subregularity and strong metric subregularity of the inverse mapping F^{-1} at $(\bar{y}, \bar{x})$, respectively.

3. Criticality and reducibility. In this section we first define critical and noncritical multipliers associated with stationary solutions to variational systems of type (1.1). Then we discuss a modified notion of set reducibility under which criticality can be efficiently investigated in the framework of conic programming.

Given a point $\bar{x} \in \mathbb{X}$, we define the set of *Lagrange multipliers* associated with $\bar{x}$ by

$$\Lambda(\bar{x}) := \{\lambda \in \mathbb{Y} \mid \Psi(\bar{x}, \lambda) = 0,\ \lambda \in N_\Theta(\Phi(\bar{x}))\}. \tag{3.1}$$

Having $(\bar{x}, \bar{\lambda})$ as a solution to the variational system (1.1) always yields $\bar{\lambda} \in \Lambda(\bar{x})$. It is not hard to see that if $\bar{\lambda} \in \Lambda(\bar{x})$ and Θ is regular, then $\bar{x}$ is a stationary point of the variational system (1.1) in the sense that it satisfies the condition

$$0 \in f(\bar{x}) + \partial(\delta_\Theta \circ \Phi)(\bar{x}). \tag{3.2}$$

The following notions of criticality for (1.1) are taken from [20, Definition 3.1].

Definition 3.1 (critical and noncritical multipliers). *Let $(\bar{x}, \bar{\lambda})$ be a solution to* (1.1)*. The multiplier $\bar{\lambda} \in \Lambda(\bar{x})$ is* critical *for* (1.1) *if there is an $\xi \in \mathbb{X} \setminus \{0\}$ satisfying*

$$0 \in \nabla_x \Psi(\bar{x}, \bar{\lambda})\xi + \nabla\Phi(\bar{x})^* DN_\Theta(\Phi(\bar{x}), \bar{\lambda})(\nabla\Phi(\bar{x})\xi). \tag{3.3}$$

The Lagrange multiplier $\bar{\lambda} \in \Lambda(\bar{x})$ is noncritical *for* (1.1) *when the generalized equation* (3.3) *admits only the trivial solution $\xi = 0$.*

We can reformulate Definition 3.1 via the mapping $G\colon \mathbb{X} \times \mathbb{Y} \rightrightarrows \mathbb{X} \times \mathbb{Y}$ given by

$$G(x, \lambda) := \begin{bmatrix} \Psi(x, \lambda) \\ -\Phi(x) \end{bmatrix} + \begin{bmatrix} 0 \\ N_\Theta^{-1}(\lambda) \end{bmatrix}. \tag{3.4}$$

It follows from [20, Theorem 7.1] that $\bar{\lambda} \in \Lambda(\bar{x})$ is noncritical if and only if

$$(0,0) \in DG((\bar{x}, \bar{\lambda}), (0,0))(\xi, \eta) \Longrightarrow \xi = 0 \quad \text{for} \quad (\xi, \eta) \in \mathbb{X} \times \mathbb{Y}. \tag{3.5}$$

Observe that the stronger implication

$$(0,0) \in DG((\bar{x}, \bar{\lambda}), (0,0))(\xi, \eta) \Longrightarrow (\xi, \eta) = (0,0) \quad \text{for} \quad (\xi, \eta) \in \mathbb{X} \times \mathbb{Y}$$

ensures strong metric subregularity for the mapping G at $((\bar{x}, \bar{\lambda}), (0,0))$; see [20, Theorem 7.1] for more details and discussion.

The following property of the set Θ in (1.1) is crucial for our subsequent analysis.

Definition 3.2 (reducible sets). *A closed set $\Theta \subset \mathbb{Y}$ is said to be $\mathcal{C}^2$*-cone reducible *at $\bar{z} \in \Theta$ to a closed convex cone $C \subset \mathbb{E}$ of a finite-dimensional space E if there exist a neighborhood $\mathcal{O} \subset \mathbb{Y}$ of $\bar{z}$ and a $\mathcal{C}^2$-smooth mapping $h\colon \mathbb{Y} \to \mathbb{E}$ such that*

$$\Theta \cap \mathcal{O} = \{z \in \mathcal{O} \mid h(z) \in C\}, \quad h(\bar{z}) = 0, \quad \textit{and } \nabla h(\bar{z}) \textit{ is surjective.} \tag{3.6}$$

If this holds for all $\bar{z} \in \Theta$, then we say that Θ is $\mathcal{C}^2$-cone reducible.

Let us discuss this notion and its comparison with the known one in more detail.

Remark 3.3 (discussion on reducible sets). The conventional notion of reducibility from [5, Definition 3.135] requires that the convex cone C be pointed. The approach in this paper based on Definition 3.2 does not need this assumption. Moreover, in contrast to [5, Definition 3.135] we do not assume that the set Θ is convex; however, (3.6) implies that Θ is regular at any $z \in \mathcal{O}$. Another important point about reducible sets is the requirement that $h(\bar{z}) = 0$. This assumption plays a significant role in what follows and cannot be dropped. It helps to reduce our analysis at $\bar{z}$ in Θ to that at $h(\bar{z}) = 0$ in another convex cone C. Since $N_C(h(\bar{z})) = C^*$, the required inclusion holds automatically for C. Thus, our approach is to reduce the consideration to C, prove the claimed results for this cone, and then return to Θ.

The $\mathcal{C}^2$-cone reducibility of Θ allows us to deduce from the conventional first-order chain rules of variational analysis that for any $z \in \Theta \cap \mathcal{O}$ with $\mathcal{O}$ taken from (3.6) we have the normal and tangent cone representations

$$N_\Theta(z) = \nabla h(z)^* N_C(h(z)) \quad \text{and} \quad T_\Theta(z) = \{v \in \mathbb{Y} \mid \nabla h(z)v \in T_C(h(z))\}. \tag{3.7}$$

Let us now consider in more detail the three important cases of the variational system (1.1) where Θ therein is one of the following sets:

- a (convex) polyhedral set,
- the second-order cone,
- the cone of positive semidefinite symmetric matrices.

It is well known that these sets are $\mathcal{C}^2$-cone reducible; see [5, Examples 3.139 and 3.140]. Below we provide simplified and constructive proofs for these reductions. Our first example concerns polyhedral sets, where—in contrast to [5, Examples 3.139]—we explicitly construct h in (3.6) as an affine mapping, which is used in our subsequent analysis.

Example 3.4 (polyhedral sets). Let $\mathbb{Y} = \mathbb{R}^m$, and let Θ in (1.1) be a polyhedral set with $\bar{z} \in \Theta$. We intend to show that Θ is $\mathcal{C}^2$-cone reducible at $\bar{z}$. It follows from [8, Theorem 2E.3] that there is a neighborhood U of $0 \in \mathbb{R}^m$ for which $T_\Theta(\bar{z}) \cap U = (\Theta - \bar{z}) \cap U$. Define further the mapping $h(z) := z - \bar{z}$ for any $z \in \mathbb{R}^m$ and the set $O := \bar{z} + U$. We clearly get

$$\Theta \cap O = \{z \in O \mid h(z) \in C\} \ \text{ with } \ C := T_\Theta(\bar{z}).$$

It is easy to check that the constructed mapping h and the convex cone C satisfy (3.6), and thus the set Θ is $\mathcal{C}^2$-cone reducible.

The second example addresses a nonpolyhedral cone, which generates an important class of problems of second-order cone programming.

Example 3.5 (second-order cone). Let $\mathbb{Y} = \mathbb{R}^m$, and let $\Theta := \mathcal{Q} \subset \mathbb{Y}$, where $\mathcal{Q}$ is the *second-order/Lorentz/ice-cream cone* defined by

$$\mathcal{Q} := \{s = (s_r, s_m) \in \mathbb{R}^{m-1} \times \mathbb{R} \mid \|s_r\| \le s_m\}. \tag{3.8}$$

It follows from [3, Lemma 15] that the second-order cone $\mathcal{Q}$ is $\mathcal{C}^2$-cone reducible at $\bar{z} \in \mathcal{Q}$ to

$$C := \begin{cases} \mathcal{Q} & \text{if } \bar{z} = 0, \\ \{0\} & \text{if } \bar{z} \in (\operatorname{int} \mathcal{Q}) \setminus \{0\}, \\ \mathbb{R}_- & \text{if } \bar{z} \in (\operatorname{bd} \mathcal{Q}) \setminus \{0\}. \end{cases}$$

In what follows we represent an element $y \in \mathcal{Q}$ as $y = (y_r, y_m)$ with $y_m \in \mathbb{R}$ and $y_r \in \mathbb{R}^{m-1}$. The reduction mapping h can be defined as

$$h(z) := \begin{cases} z & \text{if } \bar{z} = 0, \\ 0 \in \mathbb{R} & \text{if } \bar{z} \in \operatorname{int} \mathcal{Q}, \\ \|z_r\|^2 - z_m^2 & \text{if } \bar{z} \in (\operatorname{bd} \mathcal{Q}) \setminus \{0\} \end{cases} \tag{3.9}$$

for all vectors z in a neighborhood of $\bar{z}$. Picking $z = (z_r, z_m) \in \mathcal{Q}$ and $\lambda = (\lambda_r, \lambda_m) \in N_\mathcal{Q}(z)$, we construct the matrix $\mathcal{H}(z, \lambda)$ as follows:

$$\mathcal{H}(z, \lambda) := \begin{cases} -\dfrac{\lambda_m}{z_m} \operatorname{diag}(\underbrace{1, \ldots, 1}_{m-1 \text{ times}}, -1) & \text{if } z = (z_r, z_m) \in (\operatorname{bd} \mathcal{Q}) \setminus \{0\}, \\ 0 & \text{if } z \in [(\operatorname{int} \mathcal{Q}) \cup \{0\}]. \end{cases} \tag{3.10}$$

This matrix appears as the curvature term of the second-order cone $\mathcal{Q}$ in Proposition 3.7.

Next we consider a more involved cone Θ in (1.1), which generates problems of *semidefinite programming* (SDP) that are highly important in applications.

Example 3.6 (positive semidefinite cone). Let $\mathbb{Y} := \mathcal{S}^m$ be the space of $m \times m$ symmetric matrices, which is conveniently treated via the inner product

$$\langle A, B\rangle := \operatorname{tr} AB$$

with $\operatorname{tr} AB$ standing for the sum of the diagonal entries of AB. This inner product induces a norm on $\mathcal{S}^m$ known as the *Frobenius/Hilbert–Schmidt norm* and defined by

$$\|A\| := \left(\sum_{i,j=1}^{m} a_{ij}^2\right)^{\frac{1}{2}} \quad \text{with} \quad A := (a_{ij}).$$

Given $A, B \in \mathcal{S}_+^m$, it is not hard to see that $\langle A, B\rangle = 0$ if and only if $AB = 0$. For a matrix $A \in \mathcal{S}^m$, denote by $A^\dagger$ the *Moore–Penrose pseudoinverse* of A. In this case we have $\Theta = \mathcal{S}_+^m$, where $\mathcal{S}_+^m$ is the cone of $m \times m$ positive semidefinite symmetric matrices. Denote $\operatorname{rank} A$ by p for $A \in \mathcal{S}_+^m$ and consider the following two cases. In the case where $p = m$ the matrix A is positive definite and hence belongs to the interior of $\mathcal{S}_+^m$. Then it is easy to observe that $\mathcal{S}_+^m$ is $\mathcal{C}^2$-cone reducible at A to $\{0\}$ with the reduction mapping $h\colon \mathcal{S}_+^m \to \{0\}$ defined by $h(B) := 0$ for B in a neighborhood of A. In the case where $p < m$ we know from [5, Example 3.140] that $\mathcal{S}_+^m$ is $\mathcal{C}^2$-cone reducible at A to $\mathcal{S}_+^{m-p}$ via the mapping $h\colon \mathcal{S}^m \to \mathcal{S}^{m-p}$ defined by $h(B) := U(B)^* B U(B)$; see [5, Example 3.140] for the definition of $U(B)$ and more details on this mapping. It follows from [5, Example 3.98] that $h(A) = U(A)^* A U(A) = \alpha I_{m-p}$, where α is the smallest eigenvalue of A and I_p stands for the $(m-p) \times (m-p)$ identity matrix. Since $p < m$, we have that $\alpha = 0$ and thus $h(A) = 0$, which indeed shows that h satisfies (3.6).

The next result calculates the graphical derivative of the normal cone mapping (which is a primal-dual construction of second-order variational analysis) generated by reducible sets Θ. This is instrumental for the study of multiplier criticality in such settings. Recall that the *critical cone* to Θ at $z \in \Theta$ for $\lambda \in N_\Theta(z)$ is defined by

$$K_\Theta(z, \lambda) := T_\Theta(z) \cap \{\lambda\}^\perp. \tag{3.11}$$

Proposition 3.7 (graphical derivative of normal cones to reducible sets). *Let $(\bar z, \bar\lambda) \in \operatorname{gph} N_\Theta$, and let Θ be $\mathcal{C}^2$-cone reducible at $\bar z$ to a closed convex cone C. Then the graphical derivative of the normal cone mapping N_Θ is calculated by*

$$DN_\Theta(\bar z, \bar\lambda)(u) = \nabla^2\langle\bar\mu, h\rangle(\bar z)u + N_{K_\Theta(\bar z,\bar\lambda)}(u) \quad \textit{for all } u \in \mathbb{Y} \tag{3.12}$$

via the critical cone (3.11)*, where $\bar\mu$ is the unique solution to the system*

$$\bar\lambda = \nabla h(\bar z)^*\bar\mu \quad \textit{and} \quad \bar\mu \in N_C(h(\bar z)) = C^*, \tag{3.13}$$

and where h is taken from (3.6)*. If Θ is a polyhedral set in $\mathbb{Y} = \mathbb{R}^m$, then we have $\nabla^2\langle\bar\mu, h\rangle(\bar z)u = 0$ as $u \in \mathbb{R}^m$ for the curvature term in* (3.12)*. If $\Theta = \mathcal{Q} \subset \mathbb{Y} = \mathbb{R}^m$, then*

$$\nabla^2\langle\bar\mu, h\rangle(\bar z)u = \mathcal{H}(\bar z, \bar\lambda)u \quad \textit{for all } u \in \mathbb{R}^m. \tag{3.14}$$

Finally, in the SDP case where $\mathbb{Y} = \mathcal{S}^m$ *and* $\Theta = \mathcal{S}^m_+$ *we have the representation*

$$\nabla^2\langle\bar\mu, h\rangle(\bar z)u = -2\bar\lambda u\bar z^\dagger \quad \textit{for all } u \in \mathcal{S}^m. \tag{3.15}$$

Proof. Since $\bar\lambda \in \Lambda(\bar x)$ and $\nabla h(\bar z)$ is surjective, the normal cone representation in (3.7) implies that there is a unique vector $\bar\mu \in N_C(h(\bar z))$ such that $\bar\lambda = \nabla h(\bar z)^*\bar\mu$. This allows us to deduce (3.12) from [9, Corollary 4.5]. If $\mathbb{Y} = \mathbb{R}^m$ and Θ is a polyhedral set, we know from Example 3.4 that $\nabla^2\langle\bar\mu, h\rangle(\bar z)u = 0$ for all $u \in \mathbb{R}^m$. To calculate the curvature term for the second-order cone $\mathcal{Q}$, we get from (3.9) that $\nabla^2\langle\bar\mu, h\rangle(\bar z)u = 0$ if $\bar z \in [\operatorname{int}\mathcal{Q}] \cup \{0\}$, which verifies (3.14) in this case due to (3.10). If $\bar z \in (\operatorname{bd}\mathcal{Q}) \setminus \{0\}$, it follows from (3.9) that

$$h(y) = \|y_r\|^2 - y_m^2 \quad \text{whenever} \quad y = (y_r, y_m) \in \mathbb{R}^{m-1} \times \mathbb{R}.$$

Since $\bar\mu \in N_C(h(\bar z))$ with $C = \mathbb{R}_-$, we get $\bar\mu \in \mathbb{R}_+$ and thus conclude from (3.13) that

$$\bar\lambda = \nabla h(\bar z)^*\bar\mu = \bar\mu \begin{pmatrix} 2\bar z_r \\ -2\bar z_m \end{pmatrix},$$

which in turn yields $\bar\mu = -\frac{\bar\lambda_m}{2\bar z_m}$. On the other hand, the direct calculations lead us to

$$\nabla^2\langle\bar\mu, h\rangle(\bar z) = \bar\mu \operatorname{diag}(\underbrace{2, \ldots, 2}_{m-1 \text{ times}}, -2) = -\frac{\bar\lambda_m}{\bar z_m} \operatorname{diag}(\underbrace{1, \ldots, 1}_{m-1 \text{ times}}, -1).$$

Using (3.10) now gives us (3.14) in the case where $\bar z \in (\operatorname{bd}\mathcal{Q}) \setminus \{0\}$. To calculate the curvature term for $\mathcal{S}^m_+$, we employ [4, equation (66)] and get

$$\langle\nabla^2\langle\bar\mu, h\rangle(\bar z)u, u\rangle = -2\langle\bar\lambda, u\bar z^\dagger u\rangle \quad \text{for all } u \in \mathcal{S}^m.$$

Differentiating both sides above with respect to u brings us to

$$\nabla^2\langle\bar\mu, h\rangle(\bar z)u = -\frac{\partial\langle\bar\lambda, u\bar z^\dagger u\rangle}{\partial u} = -\bar\lambda u\bar z^\dagger - \bar z^\dagger u\bar\lambda = -2\bar\lambda u\bar z^\dagger,$$

which justifies (3.15) and thus completes the proof of the proposition. ∎

As an immediate consequence of Definition 3.1 and Proposition 3.7, we arrive at the following equivalent description of critical multipliers for (1.1) when Θ is a $\mathcal{C}^2$-cone reducible set.

COROLLARY 3.8 (equivalent description of critical multipliers). *Let* $(\bar x, \bar\lambda)$ *be a solution to the variational system* (1.1) *and let* Θ *be* $\mathcal{C}^2$*-cone reducible at* $\bar z := \Phi(\bar x)$ *to a closed convex cone* C*, and let* $\bar\mu$ *be a unique solution to* (3.13). *Then* $\bar\lambda$ *is critical for* (1.1) *if and only if the system*

$$\nabla_x\Psi(\bar x, \bar\lambda)\xi + \nabla\Phi(\bar x)^*\eta = 0 \quad \textit{and} \quad \eta - \nabla^2\langle\bar\mu, h\rangle(\bar z)\nabla\Phi(\bar x)\xi \in N_{K_\Theta(\bar z, \bar\lambda)}(\nabla\Phi(\bar x)\xi)$$

admits a solution $(\xi, \eta) \in \mathbb{X} \times \mathbb{Y}$ *such that* $\xi \neq 0$*, where the mapping* h *is taken from* (3.6).

As mentioned in section 1, KKT systems corresponding to the problems of *constrained optimization* (1.2) clearly belong to class (1.1). The Lagrangian for (1.2) is defined by

$$L(x, \lambda) := \varphi_0(x) + \langle\Phi(x), \lambda\rangle,$$

while the set of Lagrange multipliers for (1.2) associated with a feasible solution $\bar{x}$ is given by

$$\Lambda_c(\bar{x}) := \{\lambda \in \mathbb{Y} \mid \nabla_x L(\bar{x}, \lambda) = 0,\ \lambda \in N_\Theta(\Phi(\bar{x}))\}.$$

Let $(\bar{z}, \bar{\lambda}) \in \operatorname{gph} N_\Theta$ with $\bar{z} = \Phi(\bar{x})$, and let Θ be $\mathcal{C}^2$-cone reducible at $\bar{z}$ to the closed convex cone C. Given $\bar{\lambda} \in \Lambda_c(\bar{x})$, we formulate the *second-order sufficient condition* for (1.2) as

$$\begin{cases} \langle \nabla^2_{xx} L(\bar{x}, \bar{\lambda})u, u\rangle + \langle \nabla^2 \langle \bar{\mu}, h\rangle(\bar{z})\nabla\Phi(\bar{x})u, \nabla\Phi(\bar{x})u\rangle > 0 \\ \text{for all } \ 0 \neq u \in \mathbb{X} \ \text{ with } \ \nabla\Phi(\bar{x})u \in K_\Theta(\bar{z}, \bar{\lambda}), \end{cases} \tag{3.16}$$

where h and $\bar{\mu}$ are taken from (3.6) and (3.13), respectively. Note that the second-order sufficient optimality condition for this framework expressed via the so-called "sigma term" is equivalent to (3.15); see [5, Page 242].

When $\mathbb{Y} = \mathbb{R}^m$ and $\Theta = \mathcal{Q}$, the curvature term in (3.16) is calculated in Proposition 3.7 as $\langle \nabla^2 \langle \bar{\mu}, h\rangle(\bar{z})u, u\rangle = \langle \mathcal{H}(\bar{z}, \bar{\lambda})u, u\rangle$ for all $u \in \mathbb{Y}$. If $\mathbb{Y} = \mathcal{S}^m$ and $\Theta = \mathcal{S}^m_+$, the curvature term in (3.16) reduces by Proposition 3.7 to $\langle \nabla^2 \langle \bar{\mu}, h\rangle(\bar{z})u, u\rangle = -2\langle \bar{\lambda}, u\Phi(\bar{x})^\dagger u\rangle$ for all $u \in \mathcal{S}^m_+$. Note that (3.16) can be stronger than the classical second-order sufficient condition for (1.2) given by

$$\begin{cases} \sup\limits_{\bar{\lambda} \in \Lambda_c(\bar{x})} \{\langle \nabla^2_{xx} L(\bar{x}, \bar{\lambda})u, u\rangle + \langle \nabla^2 \langle \bar{\mu}, h\rangle(\bar{z})\nabla\Phi(\bar{x})u, \nabla\Phi(\bar{x})u\rangle\} > 0 \\ \text{for all } \ 0 \neq u \in \mathbb{X} \ \text{ with } \ \nabla\Phi(\bar{x})u \in K_\Theta(\bar{z}, \bar{\lambda}) \end{cases}$$

if the set of Lagrange multipliers is not a singleton. However, an advantage of (3.16) is that it provides a sufficient condition for noncriticality of Lagrange multipliers. Example 3.10 confirms that it may be much easier to justify noncriticality by using the second-order sufficient condition (3.16) than working with definition (3.3) or its simplification from Corollary 3.8.

PROPOSITION 3.9 (sufficient condition for noncriticality of a Lagrange multiplier). *Let $\bar{x}$ be a feasible solution to* (1.2)*, let $\bar{\lambda} \in \Lambda_c(\bar{x})$, and let Θ be $\mathcal{C}^2$-cone reducible at $\bar{z} = \Phi(\bar{x})$ to a closed convex cone C. If the second-order sufficient condition* (3.16) *holds, then $\bar{x}$ is a strict local minimizer for* (1.2) *and the Lagrange multiplier $\bar{\lambda}$ is noncritical.*

Proof. The first fact is a well-known result, which follows, e.g., from [5, Theorem 3.86]. The noncriticality of $\bar{\lambda}$ under (3.16) can be verified directly arguing by contradiction. ∎

Let us present an SDP example borrowed from Shapiro [24, Example 4.5], who constructed it for different purposes. In our case it shows that the unique Lagrange multiplier is noncritical.

Example 3.10 (SDP). Consider the semidefinite program with $\mathbb{X} = \mathbb{R}^2$, $\mathbb{Y} = \mathcal{S}^2$, and $\Theta = \mathcal{S}^2_+$:

$$\text{minimize } x_1 + \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 \ \text{ subject to } \ \Phi(x_1, x_2) \in \Theta, \tag{3.17}$$

where $\Phi\colon \mathbb{R}^2 \to \mathbb{Y}$ is defined by $\Phi(x_1, x_2) := \operatorname{diag}(x_1, x_2)$. The feasible set of this problem can be written as $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \geq 0,\ x_2 \geq 0\}$. This shows that

$\bar{x} := (0,0)$ is a unique optimal solution to (3.17). Picking $\bar{\lambda} \in \Lambda_c(\bar{x})$, we see that $\bar{\lambda}$ satisfies the first-order optimality conditions

$$\nabla_x L(\bar{x},\bar{\lambda}) = 0, \quad \langle \bar{\lambda}, \Phi(\bar{x})\rangle = 0, \quad \text{and} \quad \bar{\lambda} \in S^2_-.$$

These imply that $\bar{\lambda} = \mathrm{diag}\,(-1,0)$, and so the set of Lagrange multipliers is a singleton. It follows from $\Phi(\bar{x}) = \mathrm{diag}\,(0,0)$ that $\Phi(\bar{x})^\dagger = \mathrm{diag}\,(0,0)$. We are going to show first via Definition 3.1 that $\bar{\lambda}$ is noncritical. It suffices to verify the implication

$$\nabla^2_{xx} L(\bar{x},\bar{\lambda})\xi + \nabla\Phi(\bar{x})^*\eta = 0, \ \ \eta \in N_{K_{\mathcal{S}^2_+}(\bar{z},\bar{\lambda})}(\nabla\Phi(\bar{x})\xi) \implies \xi = 0.$$

To proceed, pick any elements

$$\eta = \begin{pmatrix} \eta_{11} & \eta_{12} \\ \eta_{12} & \eta_{22} \end{pmatrix} \quad \text{and} \quad \xi = (\xi_1,\xi_2)$$

and deduce from $\nabla\Phi(\bar{x}) = \Big(\dfrac{\partial\Phi(\bar{x})}{\partial x_1}, \dfrac{\partial\Phi(\bar{x})}{\partial x_2}\Big) = (\mathrm{diag}\,(1,0), \mathrm{diag}\,(0,1))$ that

$$\begin{aligned} 0 &= \nabla^2_{xx} L(\bar{x},\bar{\lambda})\xi + \nabla\Phi(\bar{x})^*\eta = \mathrm{diag}\,(1,1)\xi + \left(\left\langle \eta, \frac{\partial\Phi(\bar{x})}{\partial x_1}\right\rangle, \left\langle \eta, \frac{\partial\Phi(\bar{x})}{\partial x_2}\right\rangle\right) \\ &= (\xi_1 + \eta_{11}, \xi_2 + \eta_{22}), \end{aligned}$$

which yields $\xi_1 = -\eta_{11}$ and $\xi_2 = \eta_{22}$. Using this together with $\eta \in N_{K_{\mathcal{S}^2_+}(\bar{z},\bar{\lambda})}(\nabla\Phi(\bar{x})\xi)$ yields

$$0 = \langle \eta, \nabla\Phi(\bar{x})\xi\rangle = \langle \nabla\Phi(\bar{x})^*\eta, \xi\rangle = \langle (\eta_{11},\eta_{22}), (\xi_1,\xi_2)\rangle = \xi_1\eta_{11} + \xi_2\eta_{22} = -\xi_1^2 - \xi_2^2,$$

which leads us to $\xi = 0$ and hence justifies the noncriticality of $\bar{\lambda}$.

Appealing now to Proposition 3.9 provides an easier way to justify the noncriticality of $\bar{\lambda}$. Indeed, it is not hard to see that

$$\begin{aligned} &\langle \nabla^2_{xx} L(\bar{x},\bar{\lambda})u, u\rangle + \langle \nabla^2\langle \bar{\mu}, h\rangle(\bar{z})\nabla\Phi(\bar{x})u, \nabla\Phi(\bar{x})u\rangle \\ &\quad = \langle \nabla^2_{xx} L(\bar{x},\bar{\lambda})u, u\rangle + 2\langle \bar{\lambda}, \nabla\Phi(\bar{x})u\Phi(\bar{x})^\dagger\nabla\Phi(\bar{x})u\rangle \\ &\quad = \langle \mathrm{diag}\,(1,1)u, u\rangle \\ &\quad = \|u\|^2 > 0 \quad \text{for all} \ \ 0 \neq u \in \mathbb{R}^2, \end{aligned}$$

which verifies that the second-order sufficient condition (3.16) holds for $\bar{\lambda}$. Employing Proposition 3.9 now tells us that the unique Lagrange multiplier $\bar{\lambda}$ is noncritical.

When the set Θ is $\mathcal{C}^2$-cone reducible at $\bar{z} = \Phi(\bar{x})$ to a closed convex cone C, it is useful to consider a counterpart of (1.1) for the closed convex cone C from (3.6) written as

$$\Psi^r(x,\mu) := f(x) + \nabla(h\circ\Phi)(x)^*\mu = 0 \ \ \text{and} \ \ \mu \in N_C((h\circ\Phi)(x)) \tag{3.18}$$

with $(x,\mu) \in \mathbb{X}\times\mathbb{E}$. The set of Lagrange multipliers for the *reduced variational system* (3.18) associated with a point $\bar{x} \in \mathbb{R}^n$ is defined by

$$\Lambda^r(\bar{x}) := \{\mu \in \mathbb{E} \mid \Psi^r(\bar{x},\mu) = 0, \ \mu \in N_C((h\circ\Phi)(\bar{x}))\}.$$

Since $\nabla h(\bar{z})$ is surjective, we get the relationship

$$\Lambda(\bar{x}) = \nabla h(\bar{z})^*\Lambda^r(\bar{x}), \tag{3.19}$$

which is largely exploited below.

4. Uniqueness and stability of Lagrange multipliers. This section is devoted to establishing necessary and sufficient conditions for the uniqueness of Lagrange multipliers in nonpolyhedral systems (1.1) combined with their certain error bound. Besides being of its own interest, this issue is instrumental for characterizing noncritical multipliers in the next section. Given a point $\bar{x} \in \mathbb{X}$, we define the *Lagrange multiplier mapping* $M_{\bar{x}}\colon \mathbb{X} \times \mathbb{Y} \rightrightarrows \mathbb{Y}$ associated with $\bar{x}$ by

$$M_{\bar{x}}(v, w) := \{\lambda \in \mathbb{Y} \mid (v, w) \in G(\bar{x}, \lambda)\} \quad \text{for all } \ (v, w) \in \mathbb{X} \times \mathbb{Y}, \tag{4.1}$$

where G is taken from (3.4). It is easy to see that $M_{\bar{x}}(0, 0) = \Lambda(\bar{x})$, where $\Lambda(\bar{x})$ is the set of Lagrange multipliers at $\bar{x}$ defined in (3.1).

The following theorem provides characterizations of the uniqueness of Lagrange multipliers in (1.1) together with some error bound and calmness properties, which are automatic for polyhedral systems. In particular, in the case of NLPs the obtained characterizations of uniqueness reduce to the strong Mangasarian–Fromovitz constraint qualification (SMFCQ); see [15, Page 11] for more details. A similar result has been recently been established in [12, Theorem 4.5] in the case when $\mathbb{Y} = \mathbb{R}^m$ and the set Θ is the second-order cone $\mathcal{Q}$. Further discussions are given in Remark 4.2.

THEOREM 4.1 (characterizations of uniqueness and stability of Lagrange multipliers). *Let* $(\bar{x}, \bar{\lambda})$ *be a solution to the variational system* (1.1) *and let* Θ *be regular at* $\bar{x}$*. Then we have the following equivalent assertions.*

(i) *The Lagrange multiplier* $\bar{\lambda}$ *is unique and there exist constants* $\ell \geq 0$ *and* $\varepsilon > 0$ *ensuring the error bound estimate*

$$\|\lambda - \bar{\lambda}\| = d(\lambda; \Lambda(\bar{x})) \leq \ell \left(\|\Psi(\bar{x}, \lambda)\| + d(\Phi(\bar{x}); N_\Theta^{-1}(\lambda))\right) \quad \textit{for all } \ \lambda \in \mathbb{B}_\varepsilon(\bar{\lambda}). \tag{4.2}$$

(ii) *The Lagrange multiplier* $\bar{\lambda}$ *is unique and the mapping* $M_{\bar{x}}$ *from* (4.1) *is calm at* $((0, 0), \bar{\lambda})$.

(iii) *The Lagrange multiplier mapping* $M_{\bar{x}}$ *is isolatedly calm at* $((0, 0), \bar{\lambda})$.

(iv) *The dual qualification condition is satisfied as follows:*

$$DN_\Theta(\Phi(\bar{x}), \bar{\lambda})(0) \cap \ker \nabla\Phi(\bar{x})^* = \{0\}. \tag{4.3}$$

Proof. Assertions (i) and (ii) are equivalent by definition. To proceed further, define $G_{\bar{x}}(\lambda) := G(\bar{x}, \lambda)$ and see that $G_{\bar{x}}^{-1} = M_{\bar{x}}$. Then (i) amounts to saying that the mapping $G_{\bar{x}}$ is strongly metrically subregular at $(\bar{\lambda}, (0, 0))$. Indeed, the validity of (i) clearly yields the strong subregularity property of $G_{\bar{x}}$ at $(\bar{\lambda}, (0, 0))$. Conversely, the latter property tells us that (4.2) holds and that for some $\varepsilon > 0$ we get the equalities

$$M_{\bar{x}}(0, 0) \cap \mathbb{B}_\varepsilon(\bar{\lambda}) = G_{\bar{x}}^{-1}(0, 0) \cap \mathbb{B}_\varepsilon(\bar{\lambda}) = \{\bar{\lambda}\}.$$

It follows from the regularity of Θ at $\bar{x}$ that $M_{\bar{x}}$ is convex-valued. Thus, $\Lambda(\bar{x}) = M_{\bar{x}}(0, 0) = \{\bar{\lambda}\}$, which gives us (i). Since $G_{\bar{x}}^{-1} = M_{\bar{x}}$, the strong metric subregularity of $G_{\bar{x}}$ at $(\bar{\lambda}, (0, 0))$ means the isolated calmness of $M_{\bar{x}}$ at $((0, 0), \bar{\lambda})$, and therefore we have (i) $\Longleftrightarrow$ (iii).

It remains to verify the equivalence between (iii) and (iv). Calculating the graphical derivative of $G_{\bar{x}}$ due to structure (3.4) gives us

$$DG_{\bar{x}}(\bar{\lambda}, (0, 0))(\eta) = \begin{bmatrix} \nabla\Phi(\bar{x})^*\eta \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ DN_\Theta^{-1}(\bar{\lambda}, \Phi(\bar{x}))(\eta) \end{bmatrix} \quad \text{for all } \ \eta \in \mathbb{Y}.$$

Since the graph of $G_{\bar{x}}$ is closed, we deduce from [8, Theorem 4E.1] that $G_{\bar{x}}$ is strongly metrically subregular at $(\bar{\lambda},(0,0))$ if and only if the implication

$$(0,0) \in DG_{\bar{x}}(\bar{\lambda},(0,0))(\eta) \Longrightarrow \eta = 0$$

holds. This amounts to saying that

$$\eta \in DN_\Theta(\Phi(\bar{x}),\bar{\lambda})(0) \cap \ker \nabla\Phi(\bar{x})^* \Longrightarrow \eta = 0.$$

The latter verifies the equivalence between (iii) and (iv), and thus completes the proof. ∎

Remark 4.2 (discussion on error bounds). It can be checked by direct calculation that in the case of NLPs in (1.1) the dual qualification condition (4.3) reduces to the SMFCQ. In the latter framework the error bound estimate (4.2) always holds and can be derived by applying the classical Hoffman lemma (see, e.g., [8, Lemma 3C.4]) to the Lagrange multiplier mapping $M_{\bar{x}}$ from (4.1). This explains why for nonlinear programming problems the uniqueness of Lagrange multipliers and the SMFCQ are equivalent. More broadly, if Θ is a polyhedral set, we can show that (4.2) holds automatically. Indeed, we know from convex analysis that $N_\Theta^{-1} = \partial\delta_\Theta^*$. Thus, it follows from [22, Theorem 11.14] that δ_Θ^* is convex piecewise linear in the sense of [22, Definition 2.47], and so its subdifferential mapping is outer/upper Lipschitzian due to Robinson's seminal result [21]. This allows us to justify the error bound estimate (4.2) when Θ is a (convex) *polyhedron*. It is not hard to go further and show that if the normal cone N_Θ is replaced by the subdifferential mapping of a *convex piecewise linear-quadratic function* from [22, Definition 10.20], then estimate (4.2) also automatically holds.

The result of [5, Proposition 4.50] tells us that the *strong Robinson constraint qualification* (SRCQ) defined in primal terms by

$$\nabla\Phi(\bar{x})\mathbb{X} + T_\Theta(\Phi(\bar{x})) \cap \{\bar{\lambda}\}^\perp = \mathbb{Y} \tag{4.4}$$

(this terminology was suggested in [7]) provides a sufficient condition for the uniqueness of Lagrange multipliers in constrained optimization with Θ being a closed, convex, not necessarily $\mathcal{C}^2$-cone reducible set. On the other hand, the novel dual qualification condition (4.3) addresses the generalized KKT systems (1.1) that appear in a broader framework than constrained optimization and seem to be sufficient for the uniqueness of multipliers therein for reducible sets Θ. As we recently proved in [12, Theorem 4.5], both constraint qualifications are equivalent when $\mathbb{Y} = \mathbb{R}^m$ and Θ is the second-order cone $\mathcal{Q}$. Now we extend this result to the general case where Θ is any $\mathcal{C}^2$-cone reducible set that may not even be convex.

Proposition 4.3 (equivalence between the SRCQ and dual constraint qualifications under reducibility). *Let* $(\bar{x},\bar{\lambda})$ *be a solution to the variational system* (1.1) *and let* Θ *be* $\mathcal{C}^2$*-cone reducible at* $\bar{z} = \Phi(\bar{x})$ *to a closed convex cone* C. *Then the dual qualification condition* (4.3) *is equivalent to SRCQ* (4.4).

Proof. It follows from (3.12) that

$$DN_\Theta(\bar{z},\bar{\lambda})(0) = N_{K_\Theta(\bar{z},\bar{\lambda})}(0) \quad \text{with } \bar{z} = \Phi(\bar{x}). \tag{4.5}$$

Assuming the validity of the SRCQ, we get the equalities

$$\begin{aligned} K_\Theta(\bar{z},\bar{\lambda})^* \cap \ker \nabla\Phi(\bar{x})^* &= (T_\Theta(\bar{z}) \cap \{\bar{\lambda}\}^\perp)^* \cap \ker \nabla\Phi(\bar{x})^* \\ &= (T_\Theta(\bar{z}) \cap \{\bar{\lambda}\}^\perp + \nabla\Phi(\bar{x})\mathbb{X})^* = \{0\}. \end{aligned}$$

Combining this with (4.5) clearly yields (4.3). Conversely, assuming (4.3) and appealing again to (4.5) tells us that

$$\begin{aligned}\mathrm{cl}\,(\nabla\Phi(\bar{x})\mathbb{X} + T_\Theta(\bar{z}) \cap \{\bar{\lambda}\}^\perp) &= (K_\Theta(\bar{z},\bar{\lambda})^* \cap \ker \nabla\Phi(\bar{x})^*)^* \\ &= (DN_\Theta(\bar{z},\bar{\lambda})(0) \cap \ker \nabla\Phi(\bar{x})^*)^* = \mathbb{Y}.\end{aligned}$$

Since the set $\nabla\Phi(\bar{x})\mathbb{X} + T_\Theta(\Phi(\bar{x})) \cap \{\bar{\lambda}\}^\perp$ is convex, it has nonempty relative interior. Hence, it follows from [22, Proposition 2.40] that the relationships

$$\begin{aligned}\mathbb{Y} = \mathrm{ri}\,(\mathbb{Y}) &= \mathrm{ri}\,[\mathrm{cl}(\nabla\Phi(\bar{x})\mathbb{X} + T_\Theta(\Phi(\bar{x})) \cap \{\bar{\lambda}\}^\perp)] \\ &= \mathrm{ri}\,(\nabla\Phi(\bar{x})\mathbb{X} + T_\Theta(\Phi(\bar{x})) \cap \{\bar{\lambda}\}^\perp) \\ &\subset (\nabla\Phi(\bar{x})\mathbb{X} + T_\Theta(\Phi(\bar{x})) \cap \{\bar{\lambda}\}^\perp)\end{aligned}$$

are satisfied, which therefore completes the proof. ∎

We highlight here that Theorem 4.1 closes a gap that can be observed in [5, Proposition 4.47(ii)] between the uniqueness of Lagrange multipliers in the general framework of (1.1) and the SRCQ, which is equivalent to the dual condition (4.3) for $\mathcal{C}^2$-cone reducible sets. To elaborate further, Bonnans and Shapiro showed in [5, Proposition 4.50] that the SRCQ yields the uniqueness of Lagrange multipliers for constrained optimization problems. Furthermore, they proved in [5, Proposition 4.47(ii)] that the uniqueness of Lagrange multipliers together with the upper/outer Lipschitzian property of the multiplier mapping and some closedness assumptions imply a stronger version of the SRCQ, called the strict constraint qualification; see [5, Definition 4.40]. Theorem 4.1 closes the aforementioned gap between the uniqueness of Lagrange multipliers and the SRCQ by replacing the outer Lipschitzian of the multiplier mapping by a weaker condition, namely the calmness of the same mapping. On doing so, Theorem 4.1 confirms that the extra closedness assumptions utilized in [5, Proposition 4.47(ii)] can be dropped without harm.

As mentioned above, the uniqueness of Lagrange multipliers for NLPs is fully characterized by the SMFCQ. However, it follows from Theorem 4.1 that in the general setting of (1.1) the validity of such a result demands that the Lagrange multiplier mapping $M_{\bar{x}}$ be calm. Is the calmness of the latter mapping essential for the validity of Theorem 4.1? The next example confirms that this is the case, particularly for SDPs.

Example 4.4 (failure of the dual qualification condition for SDPs with unique Lagrange multipliers). Consider SDP (3.17) from Example 3.10, where $\Theta = \mathcal{S}^2_+$ is $\mathcal{C}^2$-cone reducible. To verify that the dual qualification condition (4.3) fails, observe from (4.5) that

$$DN_\Theta(\Phi(\bar{x}),\bar{\lambda})(0) \cap \ker \nabla\Phi(\bar{x})^* = K_{\mathcal{S}^2_+}(\bar{z},\bar{\lambda})^* \cap \ker \nabla\Phi(\bar{x})^*,$$

where $\bar{z} := \Phi(\bar{x}) = \mathrm{diag}\,(0,0)$ and $\bar{\lambda} = \mathrm{diag}\,(-1,0)$. We calculate the critical cone $K_{\mathcal{S}^2_+}(\bar{z},\bar{\lambda})$ from

$$K_{\mathcal{S}^2_+}(\bar{z},\bar{\lambda}) = \{u \in \mathcal{S}^2_+ \mid \langle u,\bar{\lambda}\rangle = 0\} = \{u \in \mathcal{S}^2_+ m \mid u\bar{\lambda} = 0\} = \{\mathrm{diag}\,(0,a)m \mid a \geq 0\}.$$

It follows from

$$\nabla\Phi(\bar{x}) = \left(\frac{\partial\Phi(\bar{x})}{\partial x_1}, \frac{\partial\Phi(\bar{x})}{\partial x_2}\right) = (\mathrm{diag}\,(1,0), \mathrm{diag}\,(0,1))$$

that

$$\begin{aligned}&\ker\nabla\Phi(\bar{x})^*\\&\quad=\left\{a=\begin{pmatrix}a_{11}&a_{12}\\a_{12}&a_{22}\end{pmatrix}\in\mathcal{S}^2\;\middle|\;\left(\left\langle a,\frac{\partial\Phi(\bar{x})}{\partial x_1}\right\rangle,\left\langle a,\frac{\partial\Phi(\bar{x})}{\partial x_2}\right\rangle\right)=\nabla\Phi(\bar{x})^*a=0\right\}\\&\quad=\left\{a=\begin{pmatrix}0&a_{12}\\a_{12}&0\end{pmatrix}\in\mathcal{S}^2\;\middle|\;a_{12}\in\mathbb{R}\right\}.\end{aligned}$$

In this way we arrive at the representation

$$\begin{aligned}&K_{\mathcal{S}^2_+}(\bar{z},\bar{\lambda})^*\cap\ker\nabla\Phi(\bar{x})^*\\&\quad=\left\{\begin{pmatrix}b_{11}&b_{12}\\b_{12}&b_{22}\end{pmatrix}\in\mathcal{S}^2\;\middle|\;b_{22}\le 0\right\}\cap\left\{\begin{pmatrix}0&a_{12}\\a_{12}&0\end{pmatrix}\in\mathcal{S}^2\;\middle|\;a_{12}\in\mathbb{R}\right\}\\&\quad=\left\{\begin{pmatrix}0&a_{12}\\a_{12}&0\end{pmatrix}\in\mathcal{S}^2\;\middle|\;a_{12}\in\mathbb{R}\right\},\end{aligned}\tag{4.6}$$

which shows that the dual qualification condition (4.3) does not hold for SDP (3.17). On the other hand, we get from Example 3.10 that $\Lambda_c(\bar{x})=\{\bar{\lambda}\}$. Let us now check that the multiplier mapping $M_{\bar{x}}$ is not calm at $((0,0),\bar{\lambda})$. Observe that $M_{\bar{x}}$ admits the representation

$$\begin{aligned}M_{\bar{x}}(v,w)&=\left\{\lambda=\begin{pmatrix}\lambda_{11}&\lambda_{12}\\\lambda_{12}&\lambda_{22}\end{pmatrix}\in\mathcal{S}^2_-\;\middle|\;v=\nabla_xL(\bar{x},\lambda),\ \lambda\in N_{\mathcal{S}^2_+}(w)\right\}\\&=\left\{\lambda=\begin{pmatrix}\lambda_{11}&\lambda_{12}\\\lambda_{12}&\lambda_{22}\end{pmatrix}\in\mathcal{S}^2_-\;\middle|\;v=(1+\lambda_{11},\lambda_{22}),\ \langle\lambda,w\rangle=0\right\}\end{aligned}$$

with $(v,w)\in\mathbb{R}^2\times\mathcal{S}^2$. Pick an arbitrary $t>0$ and define $v_t:=(-\frac{t^2}{2},-\frac{t^2}{2})$, $w_t:=\operatorname{diag}(0,0)$, and

$$\lambda_t:=\begin{pmatrix}-1-\frac{t^2}{2}&\frac{t}{2}\\\frac{t}{2}&-\frac{t^2}{2}\end{pmatrix}.$$

It is easy to see that $\lambda_t\in M_{\bar{x}}(v_t,w_t)\cap\mathbb{B}_t(\bar{\lambda})$ when t is sufficiently small. However, we have the limit calculation

$$\lim_{t\downarrow 0}\frac{\|\lambda_t-\bar{\lambda}\|}{\|v_t\|+\|w_t\|}=\lim_{t\downarrow 0}\frac{\sqrt{\frac{t^4}{2}+\frac{t^2}{2}}}{\frac{t^2}{\sqrt{2}}}=\infty,$$

which shows that the mapping $M_{\bar{x}}$ is not calm at $((0,0),\bar{\lambda})$.

Observe to this end that in the nonlinear programming polyhedral framework we do not have the situation in Example 4.4, since the calmness of $M_{\bar{x}}$ is a direct consequence of the Hoffman lemma. In section 5 we reveal a similar phenomenon telling us that $M_{\bar{x}}$ is automatically calm in general nonpolyhedral systems under the strict complementarity condition formulated therein.

Remark 4.5 (another characterization of uniqueness of Lagrange multipliers). In [23, Proposition 2.1] Shapiro established a slightly different characterization of the uniqueness of Lagrange multipliers in the framework of the constrained optimization (1.2) when Θ is a convex cone. However, it can be stated for the variational system

(1.1) as given below. It is shown in [23] that the Lagrange multiplier $\bar\lambda \in \Lambda(\bar x)$ is unique if and only if

$$(N_\Theta(\Phi(\bar x)) + \mathbb{R}\bar\lambda) \cap \ker \nabla\Phi(\bar x)^* = \{0\}. \tag{4.7}$$

This observation can be justified simply by taking a vector from the left-hand side of (4.7) and checking that it must be zero. It is not hard to see that this can be extended to the framework of Theorem 4.1, i.e., when Θ is merely regular. Supposing that $\bar\lambda \in \Lambda(\bar x)$ is indeed unique and picking η from the left-hand side of (4.7), we have $\eta = \lambda + t\bar\lambda$ for some $\lambda \in N_\Theta(\Phi(\bar x))$ and $t \in \mathbb{R}$. Since Θ is regular, it follows that $\bar\lambda + \eta \in \Lambda(\bar x)$ if $t \geq 0$ and $\bar\lambda - \frac{1}{2t}\eta \in \Lambda(\bar x)$ otherwise; hence, $\eta = 0$ by the uniqueness of $\bar\lambda$. The opposite conclusion is easily justified by a contraposition. Observe also that when Θ is $\mathcal{C}^2$-cone reducible the dual condition (4.3) can be written as

$$\mathrm{cl}\,(N_\Theta(\Phi(\bar x)) + \mathbb{R}\bar\lambda) \cap \ker \nabla\Phi(\bar x)^* = \{0\}.$$

In subsequent sections (see, in particular, the proof of Theorem 5.6) we show that, in contrast to (4.7), condition (4.3) appears in our analysis naturally and can be used as a constraint qualification in subproblems of the SQP method for constrained optimization.

5. Characterizations of noncritical multipliers. In this section we establish the main result of the paper, which gives us a complete characterization of the noncriticality of Lagrange multipliers in general variational systems (1.1). Our previous result in this direction [20, Theorem 4.1] addresses KKT systems of type (1.1) with N_Θ replaced by the subdifferential mapping of a convex piecewise linear function. The proof therein is strongly based on the polyhedral structure of the latter systems and cannot be extended to a nonpolyhedral case. Here we develop a new approach that works for the general $\mathcal{C}^2$-cone reducible sets Θ.

First, we present several lemmas, each of independent interest.

Lemma 5.1 (closed images under surjectivity). *Let $A\colon \mathbb{X} \to \mathbb{Y}$ be a surjective linear operator. Then $D \subset \mathbb{X}$ is closed if and only if $A^*(D)$ has this property, where A^* stands for the adjoint operator associated with A.*

Proof. The "if" part comes as a direct consequence of the surjectivity condition $\ker A^* = \{0\}$. The "only if" part follows from [18, Lemma 1.18]. ∎

Lemma 5.2 (propagation of closedness). *Let the pair $(\bar x, \bar\lambda)$ be a solution to the variational system* (1.1), *and let Θ be $\mathcal{C}^2$-cone reducible at $\bar z = \Phi(\bar x)$ to a closed convex cone C. Then the following assertions are equivalent:*
- (i) *the set $K_\Theta(\bar z, \bar\lambda)^* - [K_\Theta(\bar z, \bar\lambda)^* \cap \ker \nabla\Phi(\bar x)^*]$ is closed;*
- (ii) *the set $K_C(h(\bar z), \bar\mu)^* - [K_C(h(\bar z), \bar\mu)^* \cap \ker \nabla(h \circ \Phi)(\bar x)^*]$ is closed, where h is taken from* (3.6), *$\bar\mu$ is a unique solution to* (3.13), *and $K_C(h(\bar z), \bar\mu) := T_C(h(\bar z)) \cap \{\bar\mu\}^\perp$ is the critical cone to C at $h(\bar z)$ for $\bar\mu$.*

Proof. It follows from (3.12) that $DN_\Theta(\bar z, \bar\lambda)(0) = K_\Theta(\bar z, \bar\lambda)^*$. Thus, the set in (i) can equivalently be represented as

$$DN_\Theta(\bar z, \bar\lambda)(0) - [DN_\Theta(\bar z, \bar\lambda)(0) \cap \ker \nabla\Phi(\bar x)^*].$$

Since C is a closed convex cone with $h(\bar z) = 0 \in C$, we conclude that C is $\mathcal{C}^2$-cone reducible to itself at $h(\bar z)$ in the sense of (3.6) with $h = I\colon \mathbb{E} \to \mathbb{E}$ being the identity mapping. This yields

$$DN_C(h(\bar z), \bar\mu)(v) = N_{K_C(h(\bar z),\bar\mu)}(v) \quad \text{for all } v \in \mathbb{E}. \tag{5.1}$$

Using the equivalent local representation (3.6) for Θ and the surjectivity/full rank of $\nabla h(\bar{z})$, we deduce from (5.1) and the second-order chain rule in [10, Theorem 2] that

$$DN_\Theta(\bar{z},\bar{\lambda})(u) = \nabla^2\langle\bar{\mu},h\rangle(\bar{z})u + \nabla h(\bar{z})^* N_{K_C(h(\bar{z}),\bar{\mu})}(\nabla h(\bar{z})u) \quad \text{for all } u \in \mathbb{Y}, \tag{5.2}$$

which in turn implies the equalities

$$DN_\Theta(\bar{z},\bar{\lambda})(0) = \nabla h(\bar{z})^* N_{K_C(h(\bar{z}),\bar{\mu})}(0) = \nabla h(\bar{z})^* DN_C(h(\bar{z}),\bar{\mu})(0).$$

The latter leads us to the representation

$$\begin{aligned} &DN_\Theta(\bar{z},\bar{\lambda})(0) - [DN_\Theta(\bar{z},\bar{\lambda})(0) \cap \ker \nabla\Phi(\bar{x})^*] \\ &\qquad = \nabla h(\bar{z})^*\{DN_C(h(\bar{z}),\bar{\mu})(0) - [DN_C(h(\bar{z}),\bar{\mu})(0) \cap \ker \nabla(h\circ\Phi)(\bar{x})^*]\}. \end{aligned} \tag{5.3}$$

Thus, the claimed result amounts to saying that the following assertions are equivalent:

(a) the set $DN_\Theta(\bar{z},\bar{\lambda})(0) - [DN_\Theta(\bar{z},\bar{\lambda})(0) \cap \ker \nabla\Phi(\bar{x})^*]$ is closed,

(b) the set $DN_C(h(\bar{z}),\bar{\mu})(0) - [DN_C(h(\bar{z}),\bar{\mu})(0) \cap \ker \nabla(h\circ\Phi)(\bar{x})^*]$ is closed.

Now employing (5.3) together with Lemma 5.1 for the linear operator $A := \nabla h(\bar{z})$ readily verifies the equivalence between (a) and (b), and consequently between (i) and (ii). ∎

Consider next the set-valued mapping $S\colon \mathbb{X}\times\mathbb{Y} \rightrightarrows \mathbb{X}\times\mathbb{Y}$ given by

$$S(v,w) := \{(x,\lambda) \in \mathbb{X}\times\mathbb{Y} \mid (v,w) \in G(x,\lambda)\} \quad \text{for } (v,w) \in \mathbb{X}\times\mathbb{Y}, \tag{5.4}$$

where the mapping G is taken from (3.4). We can see that (5.4) defines the *solution map* to the *canonical perturbation* of the original variational system (1.1). The counterpart of (5.4) for the reduced generalized equation (3.18) is

$$S^r(v,w) := \{(x,\mu) \in \mathbb{X}\times\mathbb{E} \mid (v,w) \in G^r(x,\mu)\} \quad \text{with } (v,w) \in \mathbb{X}\times\mathbb{E}, \tag{5.5}$$

where the corresponding mapping G^r for (3.18) is defined by

$$G^r(x,\mu) := \begin{bmatrix} \Psi^r(x,\mu) \\ -(h\circ\Phi)(x) \end{bmatrix} + \begin{bmatrix} 0 \\ N_C^{-1}(\mu) \end{bmatrix}. \tag{5.6}$$

The following lemma establishes the equivalence between an important stability property for the mappings S and S^r we introduced in [20] under the name of *semi-isolated calmness.*

Lemma 5.3 (propagation of semi-isolated calmness for solution mappings). *Let $(\bar{x},\bar{\lambda})$ be a solution to the variational system* (1.1)*, where Θ is $\mathcal{C}^2$-cone reducible at $\bar{z} = \Phi(\bar{x})$ to a closed convex cone C. Then the following assertions are equivalent:*

(i) *there are numbers $\varepsilon > 0$ and $\ell \geq 0$ as well as neighborhoods V of $0 \in \mathbb{X}$ and W of $0 \in \mathbb{Y}$ such that for any $(v,w) \in V\times W$ we have*

$$S(v,w) \cap \mathbb{B}_\varepsilon(\bar{x},\bar{\lambda}) \subset [\{\bar{x}\}\times\Lambda(\bar{x})] + \ell(\|v\|+\|w\|)\mathbb{B}; \tag{5.7}$$

(ii) *there are numbers $\varepsilon' > 0$ and $\ell' \geq 0$ as well as neighborhoods V of $0 \in \mathbb{X}$ and W of $0 \in \mathbb{E}$ such that for any $(v,w) \in V\times W$ we have*

$$S^r(v,w) \cap \mathbb{B}_{\varepsilon'}(\bar{x},\bar{\mu}) \subset [\{\bar{x}\}\times\Lambda^r(\bar{x})] + \ell'(\|v\|+\|w\|)\mathbb{B}, \tag{5.8}$$

where $\bar{\mu}$ is a unique solution to (3.13).

Proof. Since $\nabla h(\bar{z})$ is surjective, there is a $\delta > 0$ such that for any $z \in \mathbb{B}_\delta(\bar{z})$ the derivative $\nabla h(z)$ is surjective. Pick $z \in U$ and find by [18, Lemma 1.18] a constant $\kappa_z > 0$ for which

$$\kappa_z \|y\| \le \|\nabla h(z)^* y\| \quad \text{whenever} \quad y \in \mathbb{E}.$$

Define $\bar{\kappa} := \inf\{\kappa_z \mid z \in \mathbb{B}_{\delta/2}(\bar{z})\}$ and observe that $\bar{\kappa} > 0$. Let us show then that

$$\bar{\kappa}\|y\| \le \|\nabla h(z)^* y\| \quad \text{for all} \;\; z \in \mathbb{B}_{\delta/2}(\bar{z}) \;\; \text{and} \;\; y \in \mathbb{E}. \tag{5.9}$$

Indeed, it follows from [18, Lemma 1.18] that $\kappa_z = \inf\{\|\nabla h(z)^* y\| \mid \|y\| = 1\}$ whenever $z \in \mathbb{B}_{\delta/2}(\bar{z})$. If $\bar{\kappa} = 0$, we find a sequence of $z_k \in \mathbb{B}_{\delta/2}(\bar{z})$ with $\kappa_{z_k} \to 0$ as $k \to \infty$. This implies that there is a sequence of y_k with $\|y_k\| = 1$ such that

$$\|\nabla h(z_k)^* y_k\| \le \kappa_{z_k} + k^{-1}, \quad k \in \mathbb{N}.$$

Passing to subsequences if necessary, assume without loss of generality that $z_k \to \tilde{z}$ and $y_k \to \tilde{y}$ with $\tilde{z} \in \mathbb{B}_{\delta/2}(\bar{z})$ and $\|\tilde{y}\| = 1$. Thus, we arrive at $\nabla h(\tilde{z})^* \tilde{y} = 0$, and hence $\tilde{y} = 0$ due to the surjectivity of $\nabla h(\tilde{z})$. The obtained contradiction verifies (5.9).

Assume now that (i) holds. Taking ε from (i), suppose without loss of generality that $\ell > 0$ is a Lipschitz constant for the mappings ∇h on $\mathbb{B}_\varepsilon(\bar{z})$ and Φ on $\mathbb{B}_\varepsilon(\bar{x})$. Let $M > 0$ be an upper bound for the values of $\|\nabla h(\cdot)\|$ on $\mathbb{B}_\varepsilon(\bar{z})$ and of $\|\nabla \Phi(\cdot)\|$ on $\mathbb{B}_\varepsilon(\bar{x})$. It follows from [18, Theorem 1.57] and the surjectivity of $\nabla h(\bar{z})$ that h is metrically regular around $(\bar{z}, 0)$, i.e., there exist constants $\alpha > 0$ and $\rho \ge 0$ such that we have the estimate

$$d(z; h^{-1}(y)) \le \rho\, \|h(z) - y\| \quad \text{for all} \;\; (z, y) \in \mathbb{B}_\alpha(\bar{z}) \times \mathbb{B}_\alpha(0). \tag{5.10}$$

We can always suppose that $\mathbb{B}_\alpha(\bar{z}) \subset \mathcal{O}$ with $\mathcal{O}$ taken from (3.6). To prove the semi-isolated calmness of the mapping S^r at $((0,0), (\bar{x}, \bar{\mu}))$, we claim that inclusion (5.8) holds with

$$0 < \varepsilon' \le \min\left\{\frac{\varepsilon}{4\rho}, \frac{\varepsilon}{4\rho\ell\|\bar{\mu}\|}, \frac{\varepsilon}{4\ell}, \frac{\varepsilon}{4\ell^2\|\bar{\mu}\|}, \frac{\alpha}{1+\ell^2}, \frac{\varepsilon}{4M}, \frac{\varepsilon}{4}, \frac{\alpha}{\ell}, \frac{\delta}{2\rho + 2\ell}\right\}, \tag{5.11}$$

$V := \mathbb{B}_{\varepsilon'}(0)$, and $W := \mathbb{B}_{\varepsilon'}(0)$. To proceed, picking $(v, w) \in \mathbb{B}_{\varepsilon'}(0) \times \mathbb{B}_{\varepsilon'}(0)$ and $(x, \mu) \in S^r(v, w) \cap \mathbb{B}_{\varepsilon'}(\bar{x}, \bar{\mu})$ we get the relationships

$$v = \Psi^r(x, \mu) \;\; \text{and} \;\; w + (h \circ \Phi)(x) \in N_C^{-1}(\mu). \tag{5.12}$$

Let $y_w := w + (h \circ \Phi)(x)$ and observe from (5.11) that $(\Phi(x), y_w) \in \mathbb{B}_\alpha(\bar{z}) \times \mathbb{B}_\alpha(0)$. Setting $z := \Phi(x)$ and $y := y_w$ in (5.10) gives us $z_w \in \mathbb{Y}$ such that

$$\|\Phi(x) - z_w\| \le \rho\|w\| \;\; \text{and} \;\; h(z_w) = y_w. \tag{5.13}$$

This, together with (3.7) and (5.12), tells us that

$$\begin{aligned}
&v' = \Psi(x, \lambda), \;\; w' + \Phi(x) \in N_\Theta^{-1}(\lambda) \;\; \text{with} \\
&\lambda := \nabla h(z_w)^* \mu, \;\; w' := z_w - \Phi(x), \;\; v' := v + \nabla\Phi(x)^*(\nabla h(z_w) - \nabla h(\Phi(x)))^* \mu.
\end{aligned}$$

Using (5.11), we have the estimates

$$\|z_w - \bar{z}\| \le \|z_w - \Phi(x)\| + \|\Phi(x) - \Phi(\bar{x})\| \le \rho\|w\| + \ell\|x - \bar{x}\| \le \min\left\{\frac{\varepsilon}{2}, \frac{\varepsilon}{2\ell\|\bar{\mu}\|}\right\},$$

which in turn yield the following inequalities:

$$\begin{aligned}\|\lambda-\bar{\lambda}\| &\le \|\nabla h(z_w)\|\cdot\|\mu-\bar{\mu}\|+\|\nabla h(z_w)-\nabla h(\bar{z})\|\cdot\|\bar{\mu}\|\\ &\le M\|\mu-\mu\|+\ell\|\bar{\mu}\|\cdot\|z_w-\bar{z}\|\le\frac{3\varepsilon}{4}.\end{aligned}$$

This implies that $(x,\lambda)\in S(v',w')\cap\mathbb{B}_\varepsilon(\bar{x},\bar{\lambda})$. It follows from (i) that there is a multiplier $\lambda'\in\Lambda(\bar{x})$ such that $\|x-\bar{x}\|+\|\lambda-\lambda'\|\le\ell(\|v'\|+\|w'\|)$. Using (3.19) gives us $\mu'\in\Lambda^r(\bar{x})$ such that $\lambda'=\nabla h(\bar{z})^*\mu'$. Then we get from (5.11) that $z_w\in\mathbb{B}_{\delta/r}(\bar{z})$, which ensures by (5.9) that

$$\begin{aligned}\bar{\kappa}\|\mu-\mu'\| &\le \|\nabla h(z_w)^*\mu-\nabla h(z_w)^*\mu'\|\\ &\le \|\nabla h(z_w)^*\mu-\nabla h(\bar{z})^*\mu'\|+\|\nabla h(z_w)-\nabla h(\bar{z})\|\cdot\|\mu\|\\ &\le \|\lambda-\lambda'\|+\ell\|z_w-\bar{z}\|(\varepsilon+\|\bar{\mu}\|).\end{aligned}$$

This allows us to obtain the relationships

$$\begin{aligned}&\|x-\bar{x}\|+\|\mu-\bar{\mu}\|\\ &\quad\le\|x-\bar{x}\|+\frac{1}{\bar{\kappa}}\|\lambda-\lambda'\|+\frac{\ell(\varepsilon+\|\bar{\mu}\|)}{\bar{\kappa}}\|z_w-\bar{z}\|\\ &\quad\le\|x-\bar{x}\|+\frac{1}{\bar{\kappa}}\|\lambda-\lambda'\|+\frac{\ell(\varepsilon+\|\bar{\mu}\|)}{\bar{\kappa}}(\rho\|w\|+\ell\|x-\bar{x}\|)\\ &\quad\le\max\left\{\frac{1}{\bar{\kappa}},1+\frac{\ell^2(\varepsilon+\|\bar{\mu}\|)}{\bar{\kappa}}\right\}(\|x-\bar{x}\|+\|\lambda-\lambda'\|)+\frac{\ell\rho(\varepsilon+\|\bar{\mu}\|)}{\bar{\kappa}}\|w\|\\ &\quad\le\max\left\{\frac{1}{\bar{\kappa}},1+\frac{\ell^2(\varepsilon+\|\bar{\mu}\|)}{\bar{\kappa}}\right\}\ell\left(\|v'\|+\|w'\|\right)+\frac{\ell\rho(e+\|\bar{\mu}\|)}{\bar{\kappa}}\|w\|\\ &\quad\le\max\left\{\frac{1}{\bar{\kappa}},1+\frac{\ell^2(\varepsilon+\|\bar{\mu}\|)}{\bar{\kappa}}\right\}\ell\left(\|v\|+M(\varepsilon+\|\bar{\mu}\|)\ell\rho\|w\|+\rho\|w\|\right)\\ &\qquad+\frac{\ell\rho(\varepsilon+\|\bar{\mu}\|)}{\bar{\kappa}}\|w\|,\end{aligned}$$

which therefore verify the claimed inclusion (5.17).

Suppose next that the mapping S^r is semi-isolatedly calm at $((0,0),(\bar{x},\bar{\mu}))$ and thus find constants $\ell'\ge 0$ and $\varepsilon'>0$ for which (5.8) is satisfied. We can always assume that ℓ is a Lipschitz constant for the mappings ∇h on $\mathbb{B}_{\varepsilon'}(\bar{z})$ and Φ on $\mathbb{B}_{\varepsilon'}(\bar{x})$ and that M is an upper bound for $\|\nabla\Phi(\cdot)\|$ on $\mathbb{B}_{\varepsilon'}(\bar{x})$. To prove (5.7), take $\varepsilon>0$ such that

$$\varepsilon\le\min\left\{\frac{\varepsilon'}{4(\ell+1)},\frac{\delta}{4(\ell+1)},\frac{\bar{\kappa}\varepsilon'}{2(1+(\ell+\ell^2)\|\bar{\mu}\|)},\frac{\varepsilon'}{4}\right\},\tag{5.14}$$

where δ is taken from (5.9), and suppose that $\mathbb{B}_{\varepsilon'}(\bar{z})\subset\mathcal{O}$ with $\mathcal{O}$ taken from (3.6). Picking $(v,w)\in\mathbb{B}_\varepsilon(0)\times\mathbb{B}_\varepsilon(0)$, we get $(x,\lambda)\in S(v,w)\cap\mathbb{B}_\varepsilon(\bar{x},\bar{\lambda})$, and hence

$$v=\Psi(x,\lambda)\quad\text{and}\quad w+\Phi(x)\in N_\Theta^{-1}(\lambda).$$

Let $z_w:=w+\Phi(x)$ and deduce from (5.14) that $z_w\in\mathbb{B}_{\varepsilon'}(\bar{z})\subset\mathcal{O}$. This tells us by (3.7) that

$$\lambda:=\nabla h(z_w)^*\mu\quad\text{for some}\quad\mu\in N_C(h(z_w)),$$

which therefore ensures that

$$\begin{aligned}&v' = \Psi^r(x,\mu), \quad w' + h(\Phi(x)) \in N_C^{-1}(\mu) \quad \text{with}\\ &w' := \nabla h(\Phi(x))w + o(\|w\|), \quad v' := v + \nabla\Phi(x)^*(\nabla h(z_w) - \nabla h(\Phi(x)))^*\mu.\end{aligned}$$

It follows from (5.14) that $z_w \in \mathbb{B}_{\delta/2}(\bar{z})$, and thus (5.9) leads us to the estimates

$$\begin{aligned}\|\mu - \bar{\mu}\| &\le \frac{1}{\bar{\kappa}}\|\nabla h(z_w)^*(\mu - \bar{\mu})\|\\ &\le \frac{1}{\bar{\kappa}}(\|\nabla h(z_w)^*\mu - \nabla h(\bar{z})^*\bar{\mu}\| + \|\nabla h(z_w) - \nabla h(\bar{z})\| \cdot \|\bar{\mu}\|)\\ &\le \frac{1}{\bar{\kappa}}(\|\lambda - \bar{\lambda}\| + \ell\|\bar{\mu}\| \cdot \|z_w - \bar{z}\|)\\ &\le \frac{1}{\bar{\kappa}}(\|\lambda - \bar{\lambda}\| + \ell\|\bar{\mu}\|(\|w\| + \ell\|x - \bar{x}\|)) \le \frac{\varepsilon(1 + (\ell + \ell^2)\|\bar{\mu}\|)}{\bar{\kappa}} \le \frac{\varepsilon'}{2},\end{aligned}$$

which yield $(x,\mu) \in S^r(w',v') \cap \mathbb{B}_{\varepsilon'}(\bar{x},\bar{\mu})$. Appealing to (5.8) now gives us $\mu' \in \Lambda^r(\bar{x})$ such that $\|x - \bar{x}\| + \|\mu - \mu'\| \le \ell'(\|v'\| + \|w'\|)$. By (3.19) we find $\lambda' \in \Lambda(\bar{x})$ with $\lambda' = \nabla h(\bar{z})^*\mu'$ and

$$\begin{aligned}\|\lambda - \lambda'\| &\le \|\nabla h(z_w) - \nabla h(\bar{z})\| \cdot \|\mu\| + \|\nabla h(\bar{z})\| \cdot \|\mu - \mu'\|\\ &\le (\|\bar{\mu}\| + \varepsilon')\ell\|z_w - \bar{z}\| + \|\nabla h(\bar{z})\| \cdot \|\mu - \mu'\|\\ &\le (\|\bar{\mu}\| + \varepsilon')\ell(\|w\| + \ell\|x - \bar{x}\|) + \|\nabla h(\bar{z})\| \cdot \|\mu - \mu'\|.\end{aligned}$$

Therefore, we arrive at the inequalities

$$\begin{aligned}&\|x - \bar{x}\| + \|\lambda - \lambda'\|\\ &\le \|x - \bar{x}\| + (\|\bar{\mu}\| + \varepsilon')\ell(\|w\| + \ell\|x - \bar{x}\|) + \|\nabla h(\bar{z})\| \cdot \|\mu - \mu'\|\\ &\le \max\{1 + \ell^2(\|\bar{\mu}\| + \varepsilon'), \|\nabla h(\bar{z})\|\}(\|x - \bar{x}\| + \|\mu - \mu'\|) + (\|\bar{\mu}\| + \varepsilon')\ell\|w\|\\ &\le \max\{1 + \ell^2(\|\bar{\mu}\| + \varepsilon'), \|\nabla h(\bar{z})\|\}\ell'(\|v'\| + \|w'\|) + (\|\bar{\mu}\| + \varepsilon')\ell\|w\|\\ &\le \max\{1 + \ell^2(\|\bar{\mu}\| + \varepsilon'), \|\nabla h(\bar{z})\|\}\ell'(M\|w\| + \|o(\|w\|)\| + \|v\| + M\ell(\|\bar{\mu}\| + \varepsilon')\|w\|)\\ &\qquad + (\|\bar{\mu}\| + \varepsilon')\ell\|w\|,\end{aligned}$$

which verify (5.16) and thus complete the proof of this lemma. ☐

Next we establish relationships between the calmness property (2.7) for the original system (1.1) and its reduced counterpart (3.6). To proceed, pick a stationary point $\bar{x}$ from (3.2) and define the *reduced multiplier mapping* $M^r_{\bar{x}}\colon \mathbb{X} \times \mathbb{E} \rightrightarrows \mathbb{E}$ by

$$M^r_{\bar{x}}(v,w) := \{\mu \in \mathbb{E} \mid (v,w) \in G^r(\bar{x},\mu)\} \quad \text{with} \quad (v,w) \in \mathbb{X} \times \mathbb{E}. \tag{5.15}$$

LEMMA 5.4 (propagation of calmness for multiplier mappings). *Let* $(\bar{x},\bar{\lambda})$ *be a solution to the variational system* (1.1), *where* Θ *is* $\mathcal{C}^2$*-cone reducible at* $\bar{z} = \Phi(\bar{x})$ *to a closed convex cone* C. *Then the calmness of the mapping* $M_{\bar{x}}$ *from* (4.1) *at* $((0,0),\bar{\lambda})$ *is equivalent to that of the mapping* $M^r_{\bar{x}}$ *from* (5.15) *at* $((0,0),\bar{\mu})$, *where* $\bar{\mu}$ *is a unique solution to* (3.13).

Proof. The calmness property of $M_{\bar{x}}$ at $((0,0),\bar{\lambda})$ gives us $\ell \ge 0$ and $\varepsilon > 0$ with

$$M_{\bar{x}}(v,w) \cap \mathbb{B}_\varepsilon(\bar{\lambda}) \subset M_{\bar{x}}(0,0) + \ell(\|v\| + \|w\|)\mathbb{B} \quad \text{for all} \quad (v,w) \in \mathbb{B}_\varepsilon(0,0). \tag{5.16}$$

To verify the calmness of $M^r_{\bar{x}}$ at $((0,0),\bar{\mu})$, we show that

$$M^r_{\bar{x}}(v,w) \cap \mathbb{B}_{\varepsilon'}(\bar{\mu}) \subset M^r_{\bar{x}}(0,0) + \ell'(\|v\| + \|w\|)\mathbb{B} \quad \text{whenever} \quad (v,w) \in \mathbb{B}_{\varepsilon'}(0,0) \tag{5.17}$$

by (5.9). To proceed, pick $(v, w) \in \mathbb{B}_{\varepsilon'}(0,0)$ and $(v, w) \in M^r_{\bar{x}}(v, w) \cap \mathbb{B}_{\varepsilon'}(\bar{\mu})$, which tell us that

$$v = \Psi^r(\bar{x}, \mu) \;\text{ and }\; w + h(\bar{z}) \in N_C^{-1}(\mu).$$

Since $h(\bar{z}) = 0$, we have $N_C(y) \subset C^* = N_C(h(\bar{z}))$ for any $y \in \mathbb{E}$. Defining $\lambda := \nabla h(\bar{z})^*\mu$, deduce from (3.7) that the above conditions yield

$$v = \Psi(\bar{x}, \lambda) \;\text{ and }\; \lambda \in N_\Theta(\bar{z}),$$

and thus $\lambda \in M_{\bar{x}}(v, 0)$. It follows from $\mu \in \mathbb{B}_{\varepsilon'}(\bar{\mu})$ that $\lambda \in \mathbb{B}_\varepsilon(\bar{\lambda})$. Combining this with (5.16), we find $\lambda' \in M_{\bar{x}}(0,0) = \Lambda(\bar{x})$ such that $\|\lambda - \lambda'\| \le \ell\|v\|$. Invoking (3.19) gives us $\mu' \in \Lambda^r(\bar{x}) = M^r_{\bar{x}}(0,0)$ with $\lambda' = \nabla h(\bar{z})^*\mu'$. Remembering (5.9), we arrive at the relationships

$$\kappa\|\mu - \mu'\| \le \|\nabla h(\bar{z})^*\mu - \nabla h(\bar{z})^*\mu'\| = \|\lambda - \lambda'\| \le \ell\|v\|,$$

which justify the claimed inclusion (5.17).

Assume now that the mapping $M^r_{\bar{x}}$ is calm at $((0,0), \bar{\mu})$ and find constants $\ell' \ge 0$ and $\varepsilon' > 0$ for which (5.17) is satisfied. To prove (5.16) for the mapping $M_{\bar{x}}$, select $\varepsilon > 0$ so that

$$\varepsilon \le \min\left\{\frac{\varepsilon'}{4}, \frac{\varepsilon'}{4\ell\|\nabla\Phi(\bar{x})\|(\|\bar{\mu}\| + \varepsilon')}, \frac{\bar{\kappa}\varepsilon'}{4}, \frac{\bar{\kappa}\varepsilon'}{4\ell(\|\bar{\mu}\| + \varepsilon)}\right\},$$

where ℓ is a Lipschitz constant for ∇h around $\bar{z}$. Picking $(v, w) \in \mathbb{B}_\varepsilon(0,0)$ and $\lambda \in M_{\bar{x}}(v, w) \cap \mathbb{B}_\varepsilon(\bar{\lambda})$, we arrive at the conditions

$$v = \Psi(\bar{x}, \lambda) \;\text{ and }\; w + \bar{z} \in N_\Theta^{-1}(\lambda).$$

Suppose without loss of generality that $w + \bar{z} \in \mathcal{O}$, where the neighborhood $\mathcal{O}$ is taken from (3.6). This allows us to deduce from (3.7) that $\lambda = \nabla h(w + \bar{z})^*\mu$ for some $\mu \in N_C(h(w + \bar{z})) \subset N_C(h(\bar{z}))$, and therefore to get

$$v + \nabla\Phi(\bar{x})^*(\nabla h(\bar{z}) - \nabla h(w + \bar{z}))^*\mu = \Psi^r(\bar{x}, \mu) \;\text{ and }\; h(\bar{z}) \in N_C^{-1}(\mu).$$

This means that $\mu \in M^r_{\bar{x}}(v', 0)$ with $v' = v + \nabla\Phi(\bar{x})^*(\nabla h(\bar{z}) - \nabla h(w + \bar{z}))^*\mu$. By using (5.9) and the selection of ε we obtain the inequalities

$$\begin{aligned}\|\mu - \bar{\mu}\| &\le \frac{1}{\bar{\kappa}}\|\nabla h(\bar{z})^*(\mu - \bar{\mu})\| \\ &\le \frac{1}{\bar{\kappa}}\|\lambda - \bar{\lambda}\| + \frac{\|\nabla h(w + \bar{z}) - \nabla h(\bar{z})\|(\|\bar{\mu}\| + \varepsilon)}{\bar{\kappa}} \le \frac{\varepsilon'}{4} + \frac{\varepsilon'}{4} < \varepsilon',\end{aligned}$$

which show that $\mu \in M^r_{\bar{x}}(v', 0) \cap \mathbb{B}_{\varepsilon'}(\bar{\mu})$ with v' satisfying

$$\|v'\| \le \|v\| + \ell\|\nabla\Phi(\bar{x})\| \cdot \|w\| \cdot \|\mu\| \le \frac{\varepsilon'}{4} + \ell\|\nabla\Phi(\bar{x})\| \cdot \|w\|(\|\bar{\mu}\| + \varepsilon') \le \frac{\varepsilon'}{4} + \frac{\varepsilon'}{4} < \varepsilon'.$$

Appealing now to (5.17) gives us $\mu' \in M^r_{\bar{x}}(0,0) = \Lambda^r(\bar{x})$ with $\|\mu - \mu'\| \le \ell'\|v'\|$. Employing (3.19) again, we find $\lambda' \in \Lambda(\bar{x}) = M_{\bar{x}}(0,0)$ such that $\lambda' = \nabla h(\bar{z})^*\mu'$ and

$$\begin{aligned}\|\lambda - \lambda'\| &= \|\nabla h(w + \bar{z})^*\mu - \nabla h(\bar{z})^*\mu'\| \\ &\le \|\nabla h(w + \bar{z}) - \nabla h(\bar{z})\| \cdot \|\mu\| + \|\nabla h(\bar{z})\| \cdot \|\mu - \mu'\|\end{aligned}$$

$$\le (\|\bar{\mu}\| + \varepsilon')\ell\|w\| + \ell'\|\nabla h(\bar{z})\| \cdot \|v'\|$$
$$\le (\|\bar{\mu}\| + \varepsilon')\ell\|w\| + \ell'\|\nabla h(\bar{z})\|(\|v\| + \ell\|\nabla\Phi(\bar{x})\|(\|\bar{\mu}\| + \varepsilon')\|w\|),$$

which verifies (5.16) and thus completes the proof. ∎

The last lemma in this section establishes an equivalence between the noncriticality of Lagrange multipliers of the original and reduced systems.

LEMMA 5.5 (propagation of noncriticality). *Let $(\bar{x}, \bar{\lambda})$ be a solution to the variational system* (1.1)*, and let Θ be $\mathcal{C}^2$-cone reducible at $\bar{z} = \Phi(\bar{x})$ to the closed convex cone C. Then the Lagrange multiplier $\bar{\lambda} \in \Lambda(\bar{x})$ from* (3.1) *is noncritical for* (1.1) *if and only if the unique solution $\bar{\mu} \in \Lambda^r(\bar{x})$ to* (3.13) *is noncritical for* (3.18)*.*

Proof. Employing the classical chain rule, we get

$$\begin{aligned}
\nabla^2\langle\bar{\mu}, h \circ \Phi\rangle(\bar{x}) &= \nabla(\nabla(h \circ \Phi)(\bar{x})^*\bar{\mu}) = \nabla[\nabla\Phi(\cdot)^*\nabla h(\Phi(\cdot))^*\bar{\mu}]|_{x=\bar{x}}\\
&= \nabla[\nabla\Phi(\cdot)^*\nabla h(\Phi(\bar{x}))^*\bar{\mu}]|_{x=\bar{x}} + \nabla[\nabla\Phi(\bar{x})^*\nabla h(\Phi(\cdot))^*\bar{\mu}]|_{x=\bar{x}}\\
&= \nabla[\nabla\Phi(\cdot)^*(\nabla h(\bar{z})^*\bar{\mu})]|_{x=\bar{x}} + \nabla\Phi(\bar{x})^*\nabla^2\langle\bar{\mu}, h\rangle(\bar{z})\nabla\Phi(\bar{x})\\
&= \nabla^2\langle\bar{\lambda}, \Phi\rangle(\bar{x}) + \nabla\Phi(\bar{x})^*\nabla^2\langle\bar{\mu}, h\rangle(\bar{z})\nabla\Phi(\bar{x}).
\end{aligned}$$

Combining this with (3.3), (5.2), and (5.1) yields the relationships

$$\begin{aligned}
&\nabla_x\Psi(\bar{x}, \bar{\lambda})\xi + \nabla\Phi(\bar{x})^*DN_\Theta(\Phi(\bar{x}), \bar{\lambda})(\nabla\Phi(\bar{x})\xi)\\
&= \nabla_x\Psi(\bar{x}, \bar{\lambda})\xi + \nabla\Phi(\bar{x})^*\{\nabla^2\langle\bar{\mu}, h\rangle(\bar{z})\nabla\Phi(\bar{x})\xi + \nabla h(\bar{z})^*N_{K_C(h(\bar{z}),\bar{\mu})}(\nabla h(\bar{z})\nabla\Phi(\bar{x})\xi)\}\\
&= \nabla f(\bar{x})\xi + \nabla^2\langle\bar{\lambda}, \Phi\rangle(\bar{x})\xi + \nabla\Phi(\bar{x})^*\nabla^2\langle\bar{\mu}, h\rangle(\bar{z})\nabla\Phi(\bar{x})\xi\\
&\quad + \nabla(h \circ \Phi)(\bar{x})^*DN_C(h(\bar{z}), \bar{\mu})(\nabla(h \circ \Phi)(\bar{x})\xi)\\
&= \nabla f(\bar{x})\xi + \nabla^2\langle\bar{\mu}, (h \circ \Phi)\rangle(\bar{x})\xi + \nabla(h \circ \Phi)(\bar{x})^*DN_C(h(\bar{z}), \bar{\mu})(\nabla(h \circ \Phi)(\bar{x})\xi)\\
&= \nabla_x\Psi^r(\bar{x}, \bar{\mu})\xi + \nabla(h \circ \Phi)(\bar{x})^*DN_C(h(\bar{z}), \bar{\mu})(\nabla(h \circ \Phi)(\bar{x})\xi),
\end{aligned}$$

which justify the claimed equivalence for noncritical Lagrange multipliers. ∎

Now we are ready to establish the main result of the paper, which provides a complete characterization of noncriticality of Lagrange multipliers for nonpolyhedral variational systems (1.1).

THEOREM 5.6 (characterizations of noncritical Lagrange multipliers). *Let $(\bar{x}, \bar{\lambda})$ be a solution to the variational system* (1.1)*. Consider the following properties of* (1.1) *and the solution map S taken from* (5.4)*.*

(i) *The Lagrange multiplier $\bar{\lambda} \in \Lambda(\bar{x})$ from* (3.1) *is noncritical for* (1.1)*.*

(ii) *There are numbers $\varepsilon > 0$, $\ell \ge 0$ and neighborhoods V of $0 \in \mathbb{X}$ and W of $0 \in \mathbb{Y}$ such that for any $(v, w) \in V \times W$ the semi-isolated calmness inclusion* (5.7) *holds.*

(iii) *There are numbers $\varepsilon > 0$ and $\ell \ge 0$ such that the estimate*

$$\|x - \bar{x}\| + d(\lambda; \Lambda(\bar{x})) \le \ell(\|\Psi(x, \lambda)\| + d(\Phi(x); N_\Theta^{-1}(\lambda))) \tag{5.18}$$

is satisfied for all pairs $(x, \lambda) \in \mathbb{B}_\varepsilon(\bar{x}, \bar{\lambda})$.

Then we have the following assertions.

(a) *Implications* (iii) $\Longleftrightarrow$ (ii) $\Longrightarrow$ (i) *always hold.*

(b) *If Θ is $\mathcal{C}^2$-cone reducible at $\bar{z} = \Phi(\bar{x})$ to a closed convex cone C, if the set*

$$K_\Theta(\bar{z}, \bar{\lambda})^* - [K_\Theta(\bar{z}, \bar{\lambda})^* \cap \ker\nabla\Phi(\bar{x})^*] \tag{5.19}$$

is closed, and if the Lagrange multiplier mapping $M_{\bar{x}}$ from (4.1) *is calm at $((0, 0), \bar{\lambda})$, then the converse implication* (i) $\Longrightarrow$ (ii) *is also satisfied.*

Proof. The equivalence between (ii) and (iii) can be verified similarly to [20, Theorem 4.1]. To prove (ii) $\Longrightarrow$ (i), it suffices to show that (3.5) holds. Picking $(\xi,\eta)\in\mathbb{X}\times\mathbb{Y}$ satisfying $(0,0)\in DG((\bar{x},\bar{\lambda}),(0,0))(\xi,\eta)$ we get $((\xi,\eta),(0,0))\in T_{\mathrm{gph}\,G}((\bar{x},\bar{\lambda}),(0,0))$. By the definition of the graphical derivative, find sequences $t_k\downarrow 0$ and $((\xi_k,\eta_k),(v_k,w_k))\to((\xi,\eta),(0,0))$ with

$$((\bar{x},\bar{\lambda}),(0,0))+t_k((\xi_k,\eta_k),(v_k,w_k))\in\mathrm{gph}\,G\quad\text{for all}\ \ k\in\mathbb{N}.$$

Recalling the definition of S in (5.4) gives us the inclusions

$$(\bar{x}+t_k\xi_k,\bar{\lambda}+t_k\eta_k)\in S(t_kv_k,t_kw_k),\quad k\in\mathbb{N}.$$

It follows from (5.7) that for all k sufficiently large we have

$$t_k\|\xi_k\|=\|x_t-\bar{x}\|\le\ell t_k(\|v_t\|+\|w_t\|).$$

Dividing this by t_k and then letting $k\to\infty$ implies that $\xi=0$, and thus (a) holds.

Turning to (b), we appeal to Lemma 5.5, which tells us that $\bar{\mu}$ from (3.13) is a noncritical multiplier for (3.18). Let us show that the mapping S^r from (5.5) is semi-isolatedly calm at $((0,0),(\bar{x},\bar{\mu}))$, i.e., inclusion (5.8) holds for some constants $\varepsilon'>0$ and $\ell'\ge 0$ and for some neighborhoods V of $0\in\mathbb{X}$ and W of $0\in\mathbb{E}$. To furnish this, we first verify the following result.

Claim. There are numbers $\varepsilon'>0$, $\ell'\ge 0$ and neighborhoods V of $0\in\mathbb{X}$ and W of $0\in\mathbb{E}$ such that for any $(v,w)\in V\times W$ and any $(x_{vw},\mu_{vw})\in S^r(v,w)\cap\mathbb{B}_{\varepsilon'}(\bar{x},\bar{\mu})$ we have the estimate

$$\|x_{vw}-\bar{x}\|\le\ell'(\|v\|+\|w\|). \tag{5.20}$$

To prove this claim, suppose on the contrary that (5.20) fails, i.e., for any $k\in\mathbb{N}$ there are $(v_k,w_k)\in\mathbb{B}_{1/k}(0)\times\mathbb{B}_{1/k}(0)$ and $(x_k,\mu_k)\in S^r(v_k,w_k)\cap\mathbb{B}_{1/k}(\bar{x},\bar{\mu})$ satisfying

$$\frac{\|x_k-\bar{x}\|}{\|v_k\|+\|w_k\|}\to\infty\ \text{ as }\ k\to\infty,$$

which yields $v_k=o(\|x_k-\bar{x}\|)$ and $w_k=o(\|x_k-\bar{x}\|)$. Letting $y_k:=(h\circ\Phi)(x_k)+w_k$, observe from (5.5) that $(y_k,\mu_k)\in\mathrm{gph}\,N_C$. We know from Lemma 5.4 that the calmness property for $M_{\bar{x}}$ at $((0,0),\bar{\lambda})$ amounts to that for $M^r_{\bar{x}}$ at $((0,0),\bar{\mu})$. The latter is equivalent to the metric subregularity of $(M^r_{\bar{x}})^{-1}$ at $(\bar{\mu},(0,0))$, which gives us $\rho\ge 0$ and $\alpha>0$ such that

$$d(\mu;\Lambda^r(\bar{x}))\le\rho(\|\Psi^r(\bar{x},\mu)\|+d(h(\bar{z});N_C^{-1}(\mu)))\quad\text{for all}\ \ \mu\in\mathbb{B}_\alpha(\bar{\mu}). \tag{5.21}$$

This, together with $h(\bar{z})=0$, allows us to get, for all k sufficiently large, the estimates

$$\begin{aligned}
d(\mu_k;\Lambda^r(\bar{x}))&\le\rho(\|\Psi^r(\bar{x},\mu_k)\|+d(h(\bar{z});N_C^{-1}(\mu_k)))\\
&=\rho(\|f(\bar{x})+\nabla(h\circ\Phi)(\bar{x})^*\mu_k\|)\\
&\le\rho(\|\mu_k\|\cdot\|\nabla(h\circ\Phi)(x_k)-\nabla(h\circ\Phi)(\bar{x})\|+\|\nabla(h\circ\Phi)(x_k)^*\mu_k+f(x_k)\|\\
&\qquad+\|f(x_k)-f(\bar{x})\|)\\
&\le\rho(\ell''\|\mu_k\|\cdot\|x_k-\bar{x}\|+\|v_k\|+\ell''\|x_k-\bar{x}\|),
\end{aligned} \tag{5.22}$$

where ℓ'' is a calmness constant for the mappings f and $\nabla(h\circ\Phi)$ at $\bar{x}$. Thus, there is $\mu'_k \in \Lambda^r(\bar{x})$ such that the sequence $\frac{\mu_k-\mu'_k}{\|x_k-\bar{x}\|}$ is bounded and so contains a convergent subsequence

$$\eta_k := \frac{\mu_k-\mu'_k}{\|x_k-\bar{x}\|} \to \tilde{\eta} \;\text{ as }\; k\to\infty \;\text{ with some }\; \tilde{\eta}\in\mathbb{E}. \tag{5.23}$$

Passing to a subsequence if necessary, we get that

$$\xi_k := \frac{x_k-\bar{x}}{\|x_k-\bar{x}\|} \to \xi \;\text{ as }\; k\to\infty \;\text{ with some }\; 0\neq\xi\in\mathbb{X}. \tag{5.24}$$

Denote $t_k := \|x_k-\bar{x}\|$ and deduce from $(x_k,\mu_k)\in S^r(v_k,w_k)$ that

$$o(t_k) = \Psi^r(x_k,\mu_k) \;\text{ and }\; \mu_k\in N_C(y_k).$$

Taking this into account and using (5.23) lead us to

$$\begin{aligned} o(t_k) = v_k = \Psi^r(x_k,\mu_k) &= \Psi^r(x_k,\bar{\mu}) - \Psi^r(\bar{x},\bar{\mu}) + \nabla(h\circ\Phi)(x_k)^*(\mu_k-\bar{\mu})\\ &= \nabla_x\Psi^r(\bar{x},\bar{\mu})(x_k-\bar{x}) + \nabla(h\circ\Phi)(\bar{x})^*(\mu_k-\bar{\mu}) + o(t_k)\\ &= \nabla_x\Psi^r(\bar{x},\bar{\mu})(x_k-\bar{x}) + \nabla(h\circ\Phi)(\bar{x})^*(\mu_k-\mu'_k) + o(t_k), \end{aligned}$$

which in turn yields the equality

$$\nabla_x\Psi^r(\bar{x},\bar{\mu})\xi + \nabla(h\circ\Phi)(\bar{x})^*\tilde{\eta} = 0. \tag{5.25}$$

Since C is a closed convex cone, it follows from $(y_k,\mu_k)\in \operatorname{gph} N_C$ that $y_k\in C$ and $\langle y_k,\mu_k\rangle = 0$. The latter, together with $h(\bar{z})=0$, leads us to

$$0 = \langle y_k,\mu_k\rangle = \langle w_k + h(\Phi(x_k)) - h(\bar{z}),\mu_k\rangle = \langle \nabla(h\circ\Phi)(\bar{x})(x_k-\bar{x}) + o(t_k),\mu_k\rangle,$$

and hence $\langle\nabla(h\circ\Phi)(\bar{x})\xi,\bar{\mu}\rangle = 0$. We have that $(h\circ\Phi)(\bar{x})+t_k[\nabla(h\circ\Phi)(\bar{x})\xi_k+o(t_k)/t_k]\in C$, which implies that$\nabla(h\circ\Phi)(\bar{x})\xi\in T_C(h(\bar{z}))$ and so

$$\nabla(h\circ\Phi)(\bar{x})\xi \in T_C(h(\bar{z}))\cap\{\bar{\mu}\}^{\perp}\cap \operatorname{rge}\nabla(h\circ\Phi)(\bar{x}) = K_C(h(\bar{z}),\bar{\mu})\cap\operatorname{rge}\nabla(h\circ\Phi)(\bar{x}). \tag{5.26}$$

It follows from $\mu'_k,\bar{\mu}\in\Lambda^r(\bar{x})$ and $\mu_k\in N_C(y_k)$ that

$$\begin{aligned} \mu_k-\mu'_k = \mu_k-\bar{\mu}+\bar{\mu}-\mu'_k &\in C^* + \mathbb{R}\bar{\mu} - [\ker\nabla(h\circ\Phi)(\bar{x})^*\cap(C^*+\mathbb{R}\bar{\mu})]\\ &= N_C(h(\bar{z})) + \mathbb{R}\bar{\mu} - [\ker\nabla(h\circ\Phi)(\bar{x})^*\cap(N_C(h(\bar{z}))+\mathbb{R}\bar{\mu})]\\ &\subset \operatorname{cl}(N_C(h(\bar{z}))+\mathbb{R}\bar{\mu}) - [\ker\nabla(h\circ\Phi)(\bar{x})^*\cap\operatorname{cl}(N_C(h(\bar{z}))+\mathbb{R}\bar{\mu})]\\ &= K_C(h(\bar{z}),\bar{\mu})^* - [\ker\nabla(h\circ\Phi)(\bar{x})^*\cap(K_C(h(\bar{z}),\bar{\mu}))^*]\\ &= (K_C(h(\bar{z}),\bar{\mu})\cap\mathcal{D}^*)^*, \end{aligned}$$

where $\mathcal{D} := -[\ker\nabla(h\circ\Phi)(\bar{x})^*\cap(K_C(h(\bar{z}),\bar{\mu}))^*]$, and where the last equality comes from the closedness assumptions (5.19), Lemma 5.2, and [1, Proposition 20]. This leads us to

$$\tilde{\eta}\in(K_C(h(\bar{z}),\bar{\mu})\cap\mathcal{D}^*)^*. \tag{5.27}$$

On the other hand, we have $\operatorname{rge}\nabla(h\circ\Phi)(\bar{x})\subset\mathcal{D}^*$, which, together with (5.26), yields

$$\nabla(h\circ\Phi)(\bar{x})\xi\in K_C(h(\bar{z}),\bar{\mu})\cap\mathcal{D}^*. \tag{5.28}$$

Recall that $\mu_k' \in N_C(h(\bar z))$ and $\mu_k \in N_C(y_k)$. It follows from the monotonicity of normal cone mappings to convex sets that

$$0 \le \left\langle \frac{\mu_k - \mu_k'}{t_k}, \frac{y_k - h(\bar z)}{t_k} \right\rangle.$$

This therefore implies that

$$\langle \nabla(h \circ \Phi)(\bar x)\xi, \tilde\eta\rangle \ge 0. \tag{5.29}$$

Taking this into account together with (5.27) and (5.28) implies that

$$\tilde\eta \in N_{K_C(h(\bar z),\bar\mu)\cap \mathcal{D}^*}(\nabla(h \circ \Phi)(\bar x)\xi). \tag{5.30}$$

Appealing again to the intersection rule from [1, Proposition 20] to (5.30) gives us

$$\tilde\eta \in N_{K_C(h(\bar z),\bar\mu)}(\nabla(h \circ \Phi)(\bar x)\xi) + N_{\mathcal{D}^*}(\nabla(h \circ \Phi)(\bar x)\xi).$$

Thus, there exist vectors $\eta \in N_{K_C(h(\bar z),\bar\mu)}(\nabla(h \circ \Phi)(\bar x)\xi)$ and $\eta' \in N_{\mathcal{D}^*}(\nabla(h \circ \Phi)(\bar x)\xi)$ such that $\tilde\eta = \eta+\eta'$. Since $\mathcal{D}$ is a closed convex cone, we get $\eta' \in (\mathcal{D}^*)^* = \mathcal{D}$ and hence $\eta' \in \ker \nabla(h \circ \Phi)(\bar x)^*$. It follows from (5.1) that $\eta \in DN_C(h(\bar z),\bar\mu)(\nabla(h \circ \Phi)(\bar x)\xi)$. Employing this together with (5.25), we arrive at the relationships

$$\nabla_x \Psi^r(\bar x,\bar\mu)\xi + \nabla(h\circ\Phi)(\bar x)^*\eta = 0 \ \text{ and } \ \eta \in DN_C(h(\bar z),\bar\mu)(\nabla(h\circ\Phi)(\bar x)\xi) \ \text{ with } \ \xi \ne 0,$$

which contradict the noncriticality of $\bar\mu$ and hence verify (5.20).

To finalize the proof, take the obtained constant ε' and the neighborhoods V and W from the claim above and suppose without loss of generality that $\varepsilon' < \alpha/2$ with α taken from (5.21). Observe that there is a constant $\kappa \ge 0$ such that for any $(v,w) \in V \times W$ and any $(x_{vw},\mu_{vw}) \in S^r(v,w) \cap \mathbb{B}_{\varepsilon'}(\bar x,\bar\mu)$ we have the estimate

$$d(\mu_{vw};\Lambda^r(\bar x)) \le \kappa(\|x_{vw} - \bar x\| + \|v\| + \|w\|). \tag{5.31}$$

Indeed, (5.31) can be justified by the same arguments as (5.22). Combining (5.31) and (5.20) gives us (5.8) and thus verifies that the mapping S^r from (5.5) is semi-isolatedly calm at $((0,0),(\bar x,\bar\mu))$. Invoking Lemma 5.4 tells that the semi-isolated calmness of the mapping S^r yields the one for the mapping S from (5.4). This completes the proof of the theorem. ∎

Next we provide detailed discussions of our main result, Theorem 5.6, and its proof.

Remark 5.7 (discussing the obtained characterizations of noncriticality). Our approach to characterize noncriticality of Lagrange multipliers for general variational systems (1.1) developed above largely departs from those used in [15, Theorem 1.43] and [20, Theorem 4.1] in polyhedral settings. Indeed, the proof of implication (ii) $\Longrightarrow$ (i) in Theorem 5.6 is significantly simplified due to the better translation of noncriticality via implication (3.5) that holds for any closed set Θ. The proof of (i) $\Longrightarrow$ (ii) starts with a similar device to that in the polyhedral case but departs from the latter in several steps. A new idea here is to deal with $\mu_k - \mu_k'$ instead of $\mu_k - \bar\mu$ to bypass the nonpolyhedrality of Θ. The term $\mu_k - \bar\mu$ works well in the proofs of [15, Theorem 1.43] and [20, Theorem 4.1] due to intrinsic properties of polyhedral sets, while using the same idea in nonpolyhedral cases of [25, Theorem 3.3] and [17, Proposition 4.2] requires imposing strong assumptions, which may not hold even for the polyhedral settings of [15, 20]. Our new proof of (i) $\Longrightarrow$ (ii) resolves this issue by considering $\mu_k - \mu'_k$ and appealing to the calculus of normal cones for convex cones under weak assumptions that hold in our setting due to the closedness assumption (5.19). In this way a new term appears in our proof, namely,

$$K_\Theta(\bar{z}, \bar{\lambda})^* \cap \ker \nabla\Phi(\bar{x})^*, \tag{5.32}$$

which is equivalent to $DN_\Theta(\bar{z}, \bar{\lambda})(0) \cap \ker \nabla\Phi(\bar{x})^*$ due to the calculation of the graphical derivative of the normal cone mapping taken from Proposition 3.7. As follows from Theorem 4.1, this condition relates to the uniqueness of the Lagrange multipliers. It appears naturally in our analysis and allows us to address generalized KKT systems with nonunique multipliers.

Observe further that the closedness assumption (5.19) is automatic if the set of Lagrange multipliers is a singleton and the mapping $M_{\bar{x}}$ is calm at $((0,0), \bar{\lambda})$. In this case we get from Theorem 4.1 that the set in (5.32) is $\{0\}$, and thus (5.19) reduces to the closed set $K_\Theta(\bar{z}, \bar{\lambda})^*$. Another important case where the assumed closedness holds is when Θ is a polyhedral set, which ensures the polyhedrality and hence closedness of $K_\Theta(\bar{z}, \bar{\lambda})^*$. It is currently unclear whether the closedness of (5.19) is essential for the validity of (i) $\Longrightarrow$ (ii) in Theorem 5.6.

Note also that the calmness of the Lagrange multiplier mapping $M_{\bar{x}}$ at $((0,0), \bar{\lambda})$ assumed in Theorem 5.6(b) always holds when Θ is a polyhedral set. This condition is equivalent to the validity of (4.2) being a consequence of the Hoffman lemma; cf. Remark 4.2. The following example shows that the calmness assumption on $M_{\bar{x}}$ cannot be dropped in nonpolyhedral settings even in the case of unique Lagrange multipliers.

Example 5.8 (failure of noncriticality in the absence of calmness of Lagrange multipliers). Consider the semidefinite problem (3.17) and recall from Example 3.10 that $\Lambda_c(\bar{x}) = \{\bar{\lambda}\}$. It follows from Example 4.4 that the Lagrange multiplier mapping $M_{\bar{x}}$ is not calm at $((0,0), \bar{\lambda})$. Further, we can conclude from (4.6) that

$$\begin{aligned}
&K_{\mathcal{S}^2_+}(\bar{z}, \bar{\lambda})^* - K_{\mathcal{S}^2_+}(\bar{z}, \bar{\lambda})^* \cap \ker \nabla\Phi(\bar{x})^* \\
&= \left\{ \begin{pmatrix} b_{11} & b_{12} \\ b_{12} & b_{22} \end{pmatrix} \in \mathcal{S}^2 \;\middle|\; b_{22} \le 0 \right\} - \left\{ \begin{pmatrix} 0 & a_{12} \\ a_{12} & 0 \end{pmatrix} \in \mathcal{S}^2 \;\middle|\; a_{12} \in \mathbb{R} \right\} \\
&= \left\{ \begin{pmatrix} b_{11} & b_{12} \\ b_{12} & b_{22} \end{pmatrix} \in \mathcal{S}^2 \;\middle|\; b_{22} \le 0 \right\},
\end{aligned}$$

which ensures that the closedness assumption (5.19) of Theorem 5.6 is satisfied. Moreover, we know from Example 3.10 that the unique Lagrange multiplier $\bar{\lambda}$ is noncritical. Our major goal is to show that the mapping S from (5.4) for this problem is not semi-isolatedly calm at $((0,0), (\bar{x}, \bar{\lambda}))$, which therefore demonstrates that characterization (ii) of noncriticality of Lagrange multipliers in Theorem 5.6 fails without the calmness assumption on $M_{\bar{x}}$. Observing that in the SDP framework (3.17) the solution map S reads as

$$S(v, w) = \{(x, \lambda) \in \mathbb{R}^2 \times \mathcal{S}^2 \mid v = \nabla_x L(x, \lambda),\ \lambda \in N_{\mathcal{S}^2_+}(\Phi(x) + w)\}$$

with $(v, w) \in \mathbb{R}^2 \times \mathcal{S}^2$, we will actually get more: for any arbitrary small $t > 0$ there are $(v_t, w_t) \in \mathbb{B}_t(0,0) \subset \mathbb{R}^2 \times \mathcal{S}^2$ and $(x_t, \lambda_t) \in S(v_t, w_t) \cap \mathbb{B}_t(\bar{x}, \bar{\lambda})$ such that *both* terms $\|\lambda_t - \bar{\lambda}\|$ and $\|x_t - \bar{x}\|$ are not of order $O(\|v_t\| + \|w_t\|)$; each of these properties yields the failure of the semi-isolated calmness of S at $((0,0), (\bar{x}, \bar{\lambda}))$.

Considering first the λ-term, define $v_t := (-\frac{t^2}{2}, -\frac{t^2}{2})$, $w_t := \operatorname{diag}(0,0)$, $x_t := \bar{x}$, and

$$\lambda_t := \begin{pmatrix} -1 - \frac{t^2}{2} & \frac{t}{2} \\ \frac{t}{2} & -\frac{t^2}{2} \end{pmatrix}$$

in the framework of Example 4.4. As demonstrated therein, we have $\|\lambda_t - \bar{\lambda}\| = O(t)$, while $O(\|v_t\| + \|w_t\|) = O(t^2)$. This verifies the claimed assertion on $\|\lambda_t - \bar{\lambda}\|$ and confirms the failure of the semi-isolated calmness property for S at $((0,0),(\bar{x},\bar{\lambda}))$.

Next we show that the term $\|x_t - \bar{x}\|$ also cannot be of order $O(\|v_t\| + \|w_t\|)$ in the absence of the calmness of the multiplier mapping $M_{\bar{x}}$. This fact is instructive in understanding the importance of the latter calmness property for superlinear convergence of primal iterations of SQP and related algorithms for solving nonpolyhedral conic programs. To proceed, define $v_t := (0,0)$ and $w_t := \left(\begin{smallmatrix} 0 & t^2 \\ t^2 & 0 \end{smallmatrix}\right)$ for which $O(\|v_t\|+\|w_t\|) = O(t^2)$ and then observe that S can be considered as the KKT system for the parameterized semidefinite problem $P(t)$ given by

$$P(t): \quad \text{minimize } x_1 + \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 \ \text{ subject to } \ \Phi(x) + w_t \in \mathcal{S}^2_+. \tag{5.33}$$

It is proved in [24, Example 4.5] (see also [5, Example 4.54]) that the optimal solution mapping for (5.33) is not outer Lipschitzian. Now we are going to verify the failure of the essentially more delicate semi-isolated calmness property of the solution map S, meaning that for the above pair (v_t, w_t) there exists $(x_t, \lambda_t) \in S(v_t, w_t) \cap \mathbb{B}_t(\bar{x}, \bar{\lambda})$ whenever $t > 0$ is small enough. The latter task requires a significantly more involved analysis in comparison with [24]. We provide it below along with the verification of the aforementioned growth condition for $\|x_t - \bar{x}\|$.

First, observe that the parametric optimization problem (5.33) is equivalent to

$$\text{minimize } \varphi_t(x_1, x_2) := \frac{1}{2}((x_1+1)^2 + x_2^2 - 1) + \delta_{\mathcal{S}^2_+}(\Phi(x) + w_t) \qquad \text{subject to } \ x = (x_1, x_2) \in \mathbb{R}^2.$$

It is easy to see that the level sets of φ_t are uniformly bounded, which ensures the existence of minimizers for (5.33) by the parametric version of the Weierstrass theorem; see, e.g., [22, Theorem 1.17(a)]. Denote by $x_t = (x_{t1}, x_{2t})$ such a minimizer for $P(t)$ and notice that the family $\{x_t\}$ for $t > 0$ is uniformly bounded due to this property for the level sets of φ_t.

Recall from Example 3.10 that $\bar{x}$ is a unique minimizer for $P(0)$. Furthermore, it is clear from (5.36) that (t^2, t^2) is a feasible solution to $P(t)$, and so

$$0 \le x_{t1}^2 \le t^4 + t^2 \ \text{ and } \ 0 \le x_{t2}^2 \le t^4 + t^2,$$

which yields $x_t \to \bar{x}$ as $t \downarrow 0$. Note that the *Robinson constraint qualification* (RCQ)

$$N_{\mathcal{S}^2_+}(\Phi(\bar{x})) \cap \ker \nabla\Phi(\bar{x})^* = \{0\}$$

is satisfied for $P(0)$ and hence for $P(t)$ with small t due to the robustness of the RCQ. This ensures that the set of Lagrange multipliers for $P(t)$ associated with x_t is nonempty and uniformly bounded if t is sufficiently small. Thus, there are $\varepsilon > 0$ and $l \ge 0$ with

$$\|\lambda_t\| \le l \ \text{ whenever } \ |t| \le \varepsilon \tag{5.34}$$

for such Lagrange multipliers. It follows from $\Lambda_c(\bar{x}) = \{\bar{\lambda}\}$ that $\lambda_t \to \bar{\lambda}$ as $t \downarrow 0$ and $(x_t, \lambda_t) \in S(v_t, w_t)$. Letting

$$\lambda_t := \begin{pmatrix} \lambda_{11}^t & \lambda_{12}^t \\ \lambda_{12}^t & \lambda_{22}^t \end{pmatrix},$$

we obtain from the first-order optimality conditions that

$$v_t = \nabla_x L(x_t, \lambda_t) \iff \lambda_{11}^t = -x_{t1} - 1,\ \lambda_{22}^t = -x_{t2} \ \text{ and} \tag{5.35}$$
$$\lambda_t \in N_{\mathcal{S}_+^2}(\Phi(x_t) + w_t) \iff \Phi(x_t) + w_t \in \mathcal{S}_+^2,\ \lambda_t \in \mathcal{S}_-^2,\ \lambda_t(\Phi(x_t) + w_t) = \mathrm{diag}\,(0,0).$$

The latter tells us by elementary linear algebra that

$$\Phi(x_t) + w_t = \begin{pmatrix} x_{t1} & t^2 \\ t^2 & x_{t2} \end{pmatrix} \in \mathcal{S}_+^2 \iff x_{t1} \ge 0,\ x_{t2} \ge 0,\ x_{t1}x_{t2} \ge t^4 \ \text{ and} \tag{5.36}$$

$$\lambda_t = \begin{pmatrix} \lambda_{11}^t & \lambda_{12}^t \\ \lambda_{12}^t & \lambda_{22}^t \end{pmatrix} \in \mathcal{S}_-^2 \iff \lambda_{11}^t \le 0,\ \lambda_{22}^t \le 0,\ \lambda_{11}^t \lambda_{22}^t \ge 3(\lambda_{12}^t)^2. \tag{5.37}$$

Moreover, it follows from $\lambda_t(\Phi(x_t) + w_t) = \mathrm{diag}\,(0,0)$ that

$$\lambda_{11}^t x_{t1} + t^2 \lambda_{12}^t = 0,\ \lambda_{22}^t x_{t2} + t^2 \lambda_{12}^t = 0,\ t^2 \lambda_{11}^t + x_{t2}\lambda_{12}^t = 0, \ \text{ and } \ t^2 \lambda_{22}^t + x_{t1}\lambda_{12}^t = 0. \tag{5.38}$$

Using the first two equations in (5.38) together with (5.36) and (5.37) implies that

$$x_{t1}x_{t2} = t^4 \ \text{ and } \ \lambda_{11}^t \lambda_{22}^t = (\lambda_{12}^t)^2. \tag{5.39}$$

The latter, being combined with the last two equations in (5.38), tells us that $\lambda_{22}^t x_{t2} = \lambda_{11}^t x_{t1}$, which yields in turn the relationship

$$x_{t2}^3 = -\lambda_{11}^t t^4. \tag{5.40}$$

This, along with (5.34), verifies that $|x_{t2}| = O(t^{\frac{4}{3}})$ and hence allows us to deduce from $\lambda_t \to \bar{\lambda}$ as $t \downarrow 0$ that $|\lambda_{11}^t| \ge \frac{1}{2}$ for all t sufficiently small. Using this and the first equation in (5.39) together with (5.40), we get $x_{t1}\sqrt[3]{-\lambda_{11}^t} = t^{\frac{8}{3}}$ and so arrive at $|x_{t1}| = O(t^{\frac{8}{3}})$. Employing the latter condition together with (5.40) again brings us to

$$\|x_t - \bar{x}\| = \|x_t\| = O(t^{\frac{4}{3}}) \ \text{ and } \ x_t \in \mathbb{B}_{t/2}(\bar{x}) \ \text{ for all small } \ t > 0.$$

Combining the above with (5.35) and the second equation in (5.39) shows that

$$\|\lambda_t - \bar{\lambda}\| = O(t^{\frac{4}{3}}) \ \text{ and } \ \lambda_t \in \mathbb{B}_{\frac{t}{2}}(\bar{\lambda}) \ \text{ whenever } \ t \ \text{ is sufficiently small.}$$

This tells us that $(x_t, \lambda_t) \in S(v_t, w_t) \cap \mathbb{B}_t(\bar{x}, \bar{\lambda})$, that both terms $\|x_t - \bar{x}\|$ and $\|\lambda_t - \bar{\lambda}\|$ are of order $O(t^{\frac{4}{3}})$, and therefore

$$\lim_{t\downarrow 0} \frac{\|x_t - \bar{x}\|}{\|v_t\| + \|w_t\|} = \lim_{t\downarrow 0} \frac{\|\lambda_t - \bar{\lambda}\|}{\|v_t\| + \|w_t\|} = \lim_{t\downarrow 0} \frac{O(t^{\frac{4}{3}})}{\sqrt{2}\, t^2} = \infty.$$

This verifies all the claims made above and thus confirms that the calmness of the Lagrange multiplier mapping is essential for the characterizations of noncritical multipliers obtained in nonpolyhedral variational systems.

The next result strongly relates to Theorem 5.6 but gives us significant additional information. It shows that a new second-order condition, which strengthens noncriticality, yields the semi-isolated calmness property of the solution map (5.4) at $((0,0),(\bar{x},\bar{\lambda}))$ without imposing the closedness assumption, providing that the multiplier mapping $M_{\bar{x}}$ is calm at $((0,0),\bar{\lambda})$. The new *second-order condition* for (1.1) reads as follows:

$$
\begin{cases} \langle \nabla_x \Psi(\bar{x},\bar{\lambda})\xi,\xi\rangle + \langle \nabla^2\langle \bar{\mu}, h\rangle(\bar{z})\nabla\Phi(\bar{x})\xi, \nabla\Phi(\bar{x})\xi\rangle > 0 \\ \text{for all } \ 0 \neq \xi \in \mathbb{X} \ \text{ with } \ \nabla\Phi(\bar{x})\xi \in K_\Theta(\bar{z},\bar{\lambda}), \end{cases} \tag{5.41}
$$

where h and $\bar{\mu}$ are taken from (3.6) and (3.13), respectively. When $\Phi = \nabla_x L$ with L standing for the standard Lagrangian in constrained optimization (1.2), condition (5.41) reduces to the second-order sufficient condition (3.16).

THEOREM 5.9 (semi-isolated calmness from second-order condition). *Let* $(\bar{x},\bar{\lambda})$ *be a solution to* (1.1)*, let* Θ *be* $\mathcal{C}^2$*-cone reducible at* $\bar{z} = \Phi(\bar{x})$ *to a closed convex cone* C*, and let the multiplier mapping* $M_{\bar{x}}$ *from* (4.1) *be calm at* $((0,0),\bar{\lambda})$*. If the second-order condition* (5.41) *holds, then the solution map* S *from* (5.4) *is semi-isolatedly calm at* $((0,0),(\bar{x},\bar{\lambda}))$.

Proof. We utilize a reduction procedure similar to the device of Theorem 5.6 and thus present just a sketch of the proof. Considering the reduced system (3.18), observe that (5.41) corresponds to the reduced second-order condition

$$
\langle \nabla_x \Psi^r(\bar{x},\bar{\mu})\xi,\xi\rangle > 0 \quad \text{for all } \ 0 \neq \xi \in \mathbb{X} \ \text{ with } \ \nabla(h\circ\Phi)(\bar{x})\xi \in K_C(h(\bar{z}),\bar{\mu}) \tag{5.42}
$$

for (3.18); see [5, equation (3.272)] for more detail. By Lemma 5.3 it suffices to show that the solution map S^r from (5.5) is semi-isolatedly calm at $((0,0),(\bar{x},\bar{\mu}))$. To this end, we proceed as in the proof of Theorem 5.6 and show first that (5.20) holds. Arguing by contradiction and proceeding as in the proof of Theorem 5.6 gives us (5.25), (5.28), and (5.29) without using the closedness condition (5.19). This in turn implies that

$$
0 = \langle 0,\xi\rangle = \langle \nabla_x \Psi^r(\bar{x},\bar{\mu})\xi,\xi\rangle + \langle \tilde{\eta}, \nabla(h\circ\Phi)(\bar{x})\xi\rangle \geq \langle \nabla_x \Psi^r(\bar{x},\bar{\mu})\xi,\xi\rangle
$$

with $\xi \neq 0$ due to (5.24) and $\tilde{\eta}$ taken from (5.23). Employing (5.28) along with (5.42) yields $\xi = 0$, a contradiction, which verifies (5.20). Finally, we can justify (5.31) as in the proof of Theorem 5.6 using the calmness of the multiplier mapping $M_{\bar{x}}$ at $((0,0),\bar{\lambda})$. ∎

In the constrained optimization framework (1.2), the obtained result provides an important extension of a well-known fact for NLPs. Indeed, it can be distilled from [11, Lemma 2] that the second-order sufficient condition (3.16) yields the semi-isolated calmness of S. Theorem 5.9 reveals that such a result can be guaranteed in the general framework of (1.1) if, in addition to the second-order condition (5.42), the Lagrange multiplier mapping $M_{\bar{x}}$ is calm. Recall that the latter property is automatic for NLPs. Moreover, combining Examples 3.10 and 5.8 tells us that the calmness of $M_{\bar{x}}$ is essential in Theorem 5.9.

The final result of this section provides an efficient condition ensuring the validity of both assumptions on closedness (5.19) and calmness of Lagrange multipliers imposed in Theorem 5.6(b). In this way we get complete characterizations of noncriticality of Lagrange multipliers via the error bound and semi-isolated calmness of solution maps to nonpolyhedral systems as in the case of polyhedrality. The condition we are going to use is known as *strict complementarity* [5, Definition 4.74] for (1.1) at $\bar{x}$ meaning that there is $\lambda \in \Lambda(\bar{x})$ such that $\lambda \in \operatorname{ri} N_\Theta(\Phi(\bar{x}))$.

THEOREM 5.10 (characterizations of the noncriticality of multipliers under strict complementarity). *Let* $(\bar{x}, \bar{\lambda})$ *be a solution to the variational system* (1.1)*, let* Θ *be* $\mathcal{C}^2$*-cone reducible at* $\bar{z} = \Phi(\bar{x})$ *to a closed convex cone* C*, and let the strict complementarity condition be satisfied for* (1.1) *at* $\bar{x}$*. Then a Lagrange multiplier* $\bar{\lambda}$ *is noncritical if and only if either of the conditions* (ii) *and* (iii) *of Theorem* 5.6 *is satisfied.*

Proof. This theorem follows from Theorem 5.6 provided that the strict complementarity imposed implies both the closedness condition (5.19) and the calmness of the multiplier mapping $M_{\bar{x}}$ assumed in Theorem 5.6(b). We split the proof into the following three steps.

Step 1. *The strict complementarity condition holds for* (1.1) *if and only if it holds for the reduced KKT system* (3.18). To verify this claim, suppose that the strict complementarity condition holds at $\bar{x}$ for (1.1) and then find a multiplier $\lambda \in \Lambda(\bar{x})$ such that $\lambda \in \operatorname{ri} N_\Theta(\Phi(\bar{x}))$. It follows from the normal cone calculus (3.7) and from [22, Proposition 2.44] that

$$\lambda \in \operatorname{ri} N_\Theta(\bar{z}) = \operatorname{ri}\left(\nabla h(\bar{z})^* N_C(h(\bar{z}))\right) = \nabla h(\bar{z})^* (\operatorname{ri} N_C(h(\bar{z}))).$$

This ensures the existence of a vector $\mu \in \operatorname{ri} N_C(h(\bar{z}))$ such that $\lambda = \nabla h(\bar{z})^* \mu$. Unifying this with $\lambda \in \Lambda(\bar{x})$ gives us $\mu \in \Lambda^r(\bar{x})$ and therefore shows that the strict complementarity condition holds for (3.18). The opposite implication is proved similarly.

Step 2. *The strict complementarity condition for* (1.1) *at* $\bar{x}$ *yields the closedness condition in Theorem* 5.6(b). It follows from Step 1 that we need to verify the closedness of the set

$$K_C(h(\bar{z}), \bar{\mu})^* - [K_C(h(\bar{z}), \bar{\mu})^* \cap \ker \nabla(h \circ \Phi)(\bar{x})^*] \tag{5.43}$$

from Lemma 5.2(ii) under the validity of the strict complementarity condition for the reduced system (3.18). To furnish this, recall that $h(\bar{z}) = 0$, and hence $K_C(h(\bar{z}), \bar{\mu})^* = \operatorname{cl}(C^* + \mathbb{R}\bar{\mu})$. Since $\bar{\mu} \in C^*$ and $\operatorname{span} C^*$ is closed, we have $\operatorname{cl}(C^* + \mathbb{R}\bar{\mu}) \subset \operatorname{span} C^*$. This leads us to

$$\begin{aligned} K_C(h(\bar{z}), \bar{\mu})^* - [K_C(h(\bar{z}), \bar{\mu})^* \cap \ker \nabla(h \circ \Phi)(\bar{x})^*] &\subset K_C(h(\bar{z}), \bar{\mu})^* - K_C(h(\bar{z}), \bar{\mu})^* \\ &= \operatorname{cl}(C^* + \mathbb{R}\bar{\mu}) - \operatorname{cl}(C^* + \mathbb{R}\bar{\mu}) \\ &\subset \operatorname{span} C^* - \operatorname{span} C^* = \operatorname{span} C^*. \end{aligned}$$

On the other hand, it follows from the strict complementarity condition for (3.18) that there is a vector $\mu \in \operatorname{ri} N_C(h(\bar{z})) = \operatorname{ri} C^*$ such that $\mu \in \Lambda^r(\bar{x})$. Pick $w \in \operatorname{span} C^*$ and observe that $\operatorname{aff} C^* = \operatorname{span} C^*$. By $\mu \in \operatorname{ri} C^*$ we find a small number $t > 0$ for which $\mu + tw \in C^*$. Combining the above facts brings us to the relationships

$$\begin{aligned} tw = (\mu + tw - \bar{\mu}) - (\mu - \bar{\mu}) &\subset \operatorname{cl}(C^* + \mathbb{R}\bar{\mu}) - [\operatorname{cl}(C^* + \mathbb{R}\bar{\mu}) \cap \ker \nabla(h \circ \Phi)(\bar{x})^*] \\ &\subset K_C(h(\bar{z}), \bar{\mu})^* - [K_C(h(\bar{z}), \bar{\mu})^* \cap \ker \nabla(h \circ \Phi)(\bar{x})^*], \end{aligned} \tag{5.44}$$

which readily imply the inclusion

$$\operatorname{span} C^* \subset K_C(h(\bar{z}), \bar{\mu})^* - [K_C(h(\bar{z}), \bar{\mu})^* \cap \ker \nabla(h \circ \Phi)(\bar{x})^*].$$

Since the opposite inclusion also holds by the above discussion, we come up with the equality

$$K_C(h(\bar{z}),\bar{\mu})^* - [K_C(h(\bar{z}),\bar{\mu})^* \cap \ker \nabla(h\circ\Phi)(\bar{x})^*] = \operatorname{span} C^*,$$

which verifies the closedness of the set in (5.43). Appealing now to Lemma 5.2 tells us that the set in (5.19) is closed as well.

Step 3. *The strict complementarity condition for* (1.1) *at* $\bar{x}$ *implies that the Lagrange multiplier mapping* $M_{\bar{x}}$ *is calm at* $((0,0),\bar{\lambda})$. By Step 1 it suffices to prove that estimate (5.21) holds under the strict complementarity condition for (3.18). Recall that $h(\bar{z}) = 0$ gives us $\Lambda^r(\bar{x}) = \{\mu \in \mathbb{E} \mid \Psi^r(\bar{x},\mu) = 0,\ \mu \in C^*\}$. Set $D := \{\mu \in \mathbb{E} \mid \Psi^r(\bar{x},\mu) = 0\}$ and $\Lambda^r(\bar{x}) = D \cap C^*$. Since the strict complementarity condition is satisfied for (3.18) at $\bar{x}$, we obtain $\operatorname{ri} C^* \cap D \neq \emptyset$. This, together with [1, Corollary 3], ensures the existence of $\varepsilon > 0$ and $\ell' \geq 0$ such that

$$d(\mu;\Lambda^r(\bar{x})) \leq \ell' \left(d(\mu;D) + d(\mu;C^*)\right) \quad \text{for all } \ \mu \in \mathbb{B}_\varepsilon(\bar{\mu}). \tag{5.45}$$

On the other hand, since D is a polyhedral set, the Hoffman lemma gives us $\ell'' > 0$, for which

$$d(\mu;D) \leq \ell'' \,\|\Psi^r(\bar{x},\mu)\| \ \text{ whenever } \ \mu \in \mathbb{E}.$$

Combining this and (5.45) ensures the existence of $\ell \geq 0$ such that the estimate

$$d(\mu;\Lambda^r(\bar{x})) \leq \ell \left(\|\Psi^r(\bar{x},\mu)\| + d(\mu;C^*)\right) \quad \text{for all } \ \mu \in \mathbb{B}_\varepsilon(\bar{\mu}) \tag{5.46}$$

holds. Pick $\mu \in \mathbb{E}$ and let $y := P_C(\mu)$, where $P_C(\mu)$ stands for the projection of μ onto the convex cone C. It implies that $\mu - y \in N_C(y)$ and so $\mu - y \in C^*$, which brings us to

$$d(\mu;C^*) \leq \|\mu - (\mu - y)\| = \|y\| = \|P_C(\mu)\| \quad \text{for all } \ \mu \in \mathbb{E}. \tag{5.47}$$

Now, observing that $P_C(\mu) = 0$ if and only if $\mu \in C^*$ allows us to deduce from $\mu \in C^*$ that

$$\|P_C(\mu)\| = 0 = d(0;N_{C^*}(\mu)) = d(0;N_C^{-1}(\mu)) = d(h(\bar{z});N_C^{-1}(\mu)).$$

If $\mu \notin C^*$, then $\|P_C(\mu)\| < d(h(\bar{z});N_{C^*}(\mu)) = d(h(\bar{z});N_C^{-1}(\mu)) = \infty$, and so

$$\|P_C(\mu)\| \leq d(h(\bar{z});N_C^{-1}(\mu)) \quad \text{for all } \ \mu \in \mathbb{E}. \tag{5.48}$$

Combining (5.46)–(5.48) verifies estimate (5.21), which, by Lemma 5.4, yields the calmness of the Lagrange multiplier mapping $M_{\bar{x}}$ at $((0,0),\bar{\lambda})$ and thus completes the proof. ∎

Remark 5.11 (comparison with related results on calmness of KKT solution mappings). Let us compare the results obtained in Theorems 5.6, 5.9, and 5.10 with the very recent results established in [6]. We begin with [6, Example 1], which shows that the solution map S from (5.4) may fail to be semi-isolated calm despite the validity of the strict complementarity condition. (Note that [6] refers to calmness instead of semi-isolated calmness, while these properties are equivalent for this example.) On the other hand, our Theorem 5.10 reveals that noncriticality, together with strict complementarity, ensures the semi-isolated calmness of the solution map S. To explain the consistency of [6, Example 1] with Theorem 5.10, we now show that the Lagrange multiplier in the latter example is indeed critical, and so this yields the failure of the semi-isolated calmness of S. To verify this, consider, by using [6, Example 1], the following linear semidefinite program with $\mathbb{X} = \mathbb{R}$, $\mathbb{Y} = \mathcal{S}^2$, and $\Theta = \mathcal{S}^2_+$:

$$\text{minimize } \varphi(x) \text{ subject to } \Phi(x) := x \begin{pmatrix} 0 & -1 \\ -1 & 1 \end{pmatrix} + \text{diag}\,(0,1) \in \Theta, \tag{5.49}$$

where $\varphi(x) \equiv 0$ on $\mathbb{R}$. It is easy to see that $\bar{x} = 0 \in \mathbb{R}$ is the only feasible solution to this problem. Furthermore, the corresponding set of Lagrange multipliers at $\bar{x}$ is

$$\begin{aligned} \Lambda_c(\bar{x}) &= \left\{ \lambda = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{12} & \lambda_{22} \end{pmatrix} \,\middle|\, \nabla\Phi(\bar{x})^*\lambda = 0,\ \lambda \in N_{\mathcal{S}^2_+}(\text{diag}\,(0,1)) \right\} \\ &= \left\{ \lambda = \begin{pmatrix} \lambda_{11} & 0 \\ 0 & 0 \end{pmatrix} \,\middle|\, \lambda_{11} \le 0 \right\}. \end{aligned}$$

As argued in [6, Example 1], this implies that the strict complementarity condition holds at $\bar{x}$. Moreover, it is shown therein that the solution map S for this problem is given by

$$S(v,w) = \{(x,\lambda) \in \mathbb{R} \times \mathcal{S}^2 m \mid v = \nabla\Phi(x)^*\lambda,\ \lambda \in N_{\mathcal{S}^2_+}(\Phi(x) + w)\}$$

with $(v,w) \in \mathbb{R} \times \mathcal{S}^2$ and fails the semi-isolated calmness property at $((\bar{x},\bar{\lambda}),(0,0)) \in \text{gph}\, S$ with $\bar{\lambda} = \text{diag}\,(0,0)$. In what follows we show that the Lagrange multiplier $\bar{\lambda}$ is critical. Due to (3.14), Corollary 3.8, and $\bar{\lambda} = \text{diag}\,(0,0)$ it suffices to prove that the generalized equation

$$\nabla^2_{xx} L(\bar{x},\bar{\lambda})\xi + \nabla\Phi(\bar{x})^*\eta = 0 \text{ and } \eta \in N_{K_{\mathcal{S}^2_+}(\Phi(\bar{x}),\bar{\lambda})}(\nabla\Phi(\bar{x})\xi),$$

where L is the Lagrangian corresponding to (5.49), admits a solution $(\xi,\eta) \in \mathbb{R} \times \mathcal{S}^2$ with $\xi \neq 0$. Let us show that in fact there exists a vector $\xi \in \mathbb{R} \setminus \{0\}$ satisfying

$$\nabla\Phi(\bar{x})\xi = \xi \begin{pmatrix} 0 & -1 \\ -1 & 1 \end{pmatrix} \in K_{\mathcal{S}^2_+}(\Phi(\bar{x}),\bar{\lambda}) = T_{\mathcal{S}^2_+}(\text{diag}\,(0,1)) = \text{cl}\,(\mathcal{S}^2_+ + \mathbb{R}\text{diag}\,(0,1)). \tag{5.50}$$

To justify this, observe that for any $k \in \mathbb{N}$ we have

$$\begin{pmatrix} \frac{1}{k} & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{k} & -1 \\ -1 & 2k^2 \end{pmatrix} - (2k^2 - 1)\text{diag}\,(0,1) \in \mathcal{S}^2_+ + \mathbb{R}\text{diag}\,(0,1).$$

Letting $k \to \infty$ yields $\nabla\Phi(\bar{x})\xi \in K_{\mathcal{S}^2_+}(\Phi(\bar{x}),\bar{\lambda})$ for all $\xi \ge 0$ and thus verifies claim (5.50). Further, setting $\eta := \text{diag}\,(0,0)$ gives us $\nabla^2_{xx} L(\bar{x},\bar{\lambda}) = 0$, which, together with (5.50), shows that $\bar{\lambda}$ is critical. This tells us that, to ensure the semi-isolated calmness of the solution map S under the strict complementarily condition in Theorem 5.10, the noncriticality of Lagrange multipliers is *essential*. If it is dropped, we may expect the failure of semi-isolated calmness.

Next we compare our results with [6, Theorem 3.2], which provides sufficient conditions for the (semi-isolated) calmness of the solution map S for the KKT system of a special class of semidefinite programs. The latter theorem employs the second-order sufficient condition as Theorem 5.9, as well as a partial strict complementarity condition as in Theorem 5.10. Moreover, it uses a closedness assumption that seems to be different from the one exploited in Theorem 5.6. Adopting the closedness assumption of [6] for the case of (1.1) gives us the closedness of the set $\nabla\Phi(\bar{x})^* K_\Theta(\Phi(\bar{x}), \bar{\lambda})^*$. Note to this end the recent result of [17, Lemma 5.1] showing that the simultaneous validity of metric subregularity of the mapping $x \mapsto \Phi(x) - \Theta$ at $(\bar{x}, 0)$ and the strict complementarity condition for (1.1) at $\bar{x}$ yields the closedness of $\nabla\Phi(\bar{x})^* K_\Theta(\Phi(\bar{x}), \bar{\lambda})^*$. However, we conclude from Theorem 5.10 that the closedness of the set in (5.19) is implied by merely the strict complementarity condition. It is unclear so far whether the closedness in [6] can be ensured if we drop the metric subregularity of the mapping $x \mapsto \Phi(x) - \Theta$ at $(\bar{x}, 0)$.

Finally, observe that [6, Theorem 3.2] imposes another condition (see assumption (ii) in [6, Theorem 3.2]) that has no counterpart in our results. Let us show that the assumed partial complementarity condition in [6, Theorem 3.2] yields the calmness of the multiplier mapping. This allows us to verify via Theorem 5.9 the semi-isolated calmness of the solution map S for the KKT system associated with the semidefinite program (27) in [6] when the second-order sufficient conditions and the partial strict complementarity conditions hold simultaneously. Thus, it improves [6, Theorem 3.2], which provides a similar result under two additional assumptions (assumptions (i) and (ii) in [6, Theorem 3.2]).

To proceed with verifying the statement above, note first that the semidefinite program (27) in [6] can be considered as a particular case of the constrained optimization problem

$$\text{minimize } \varphi(x, y) \ \text{ subject to } \ \Phi(x, y) := (x, g(x, y)) \in \Theta := D \times \mathcal{S}^m_+, \tag{5.51}$$

where $\varphi\colon \mathbb{X} \times \mathbb{Y} \to \mathbb{R}$ and $g\colon \mathbb{X} \times \mathbb{Y} \to \mathcal{S}^n$ are twice differentiable mappings, and where $D \subset \mathbb{X}$ is a polyhedral set. For $(\bar{x}, \bar{y}) \in \mathbb{X} \times \mathbb{Y}$, the set of Lagrange multipliers is represented by

$$\begin{aligned}\Lambda(\bar{x}, \bar{y}) = \{(\lambda_1, \lambda_2) \in \mathbb{X} \times \mathcal{S}^n \mid \nabla\varphi(\bar{x}, \bar{y}) + \nabla\Phi(\bar{x}, \bar{y})^*(\lambda_1, \lambda_2) = 0,&\\ \lambda_1 \in N_D(\bar{x}),\ \lambda_2 \in N_{\mathcal{S}^m_+}(g(\bar{x}, \bar{y}))\}.&\end{aligned}$$

It is said that the *partial strict complementarity condition* is satisfied at $(\bar{x}, \bar{y})$ for problem (5.51) if there exists a pair $(\lambda_1, \lambda_2) \in \Lambda(\bar{x}, \bar{y})$ such that $\lambda_2 \in \operatorname{ri} N_{\mathcal{S}^m_+}(g(\bar{x}, \bar{y}))$. Define the sets

$$\begin{aligned}A &:= \{(\lambda_1, \lambda_2) \in \mathbb{X} \times \mathcal{S}^n \mid \nabla\varphi(\bar{x}, \bar{y}) + \nabla\Phi(\bar{x}, \bar{y})^*(\lambda_1, \lambda_2) = 0\},\\ B &:= N_D(\bar{x}) \times \mathcal{S}^n, \quad C := \mathbb{X} \times N_{\mathcal{S}^m_+}(g(\bar{x}, \bar{y}))\end{aligned}$$

and observe that $\Lambda(\bar{x}, \bar{y}) = A \cap B \cap C$. The partial strict complementarity condition ensures that $A \cap B \cap \operatorname{ri} C \neq \emptyset$. Arguing similarly to the proof of Theorem 5.9 via [1, Corollary 3] and [6, Proposition 3.3] verifies the calmness of the multiplier mapping associated with (5.51).

We conclude the paper by recalling the result of [23] showing that the strict complementarity condition ensures the equivalence between the uniqueness of Lagrange multipliers and the SRCQ (4.4) for problems of semidefinite programming. Theorem 5.10 allows us to extend Shapiro's result to the general $\mathcal{C}^2$-cone reducible setting of (1.1).

Corollary 5.12 (uniqueness of Lagrange multipliers under the strict complementarity condition). *Let $(\bar{x}, \bar{\lambda})$ be a solution to the variational system* (1.1)*, where Θ is $\mathcal{C}^2$-cone reducible at $\bar{z} = \Phi(\bar{x})$ to a closed convex cone C. Assume that the strict* *complementarity condition holds at* $\bar{x}$ *for* (1.1)*. Then the Lagrange multiplier set* $\Lambda(\bar{x})$ *is a singleton if and only if the equivalent qualification conditions* (4.3) *and* (4.4) *are satisfied.*

Proof. The proof is obtained by the combination of Theorems 4.1 and 5.10 and Proposition 4.3. ∎

Acknowledgment. We thank the two anonymous referees for carefully reading the paper and for insightful comments that allowed us to improve the original presentation. In particular, we greatly appreciated reference [6] being brought to our attention by one of the referees.

REFERENCES

[1] H. H. Bauschke, J. M. Borwein, and W. Li, *Strong conical hull intersection property, bounded linear regularity, Jameson's property (G), and error bounds in convex optimization*, Math. Program., 86 (1999), pp. 135–160.

[2] J. F. Bonnans, *Local analysis of Newton-type methods for variational inequalities and nonlinear programming*, Appl. Math. Optim., 29 (1994), pp. 161–186.

[3] J. F. Bonnans and H. Ramírez C., *Perturbation analysis of second-order cone programming problems*, Math. Program., 104 (2005), pp. 205–227.

[4] J. F. Bonnans and H. Ramírez C., *Strong Regularity of Semidefinite Programming Problems*, Technical report DIM–CMM 137, Universidad de Chile, Santiago, Chile, 2005.

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

[6] Y. Cui, D. Sun, and K.-C. Toh, *On the asymptotic superlinear convergence of the augmented Lagrangian method for semidefinite programming with multiple solutions*, preprint, https://arxiv.org/abs/1610.00875, 2016.

[7] C. Ding, D. Sun and L. Zhang, *Characterization of the robust isolated calmness for a class of conic programming problems*, SIAM J. Optim., 27 (2017), pp. 67–90.

[8] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings: A View from Variational Analysis*, 2nd ed., Springer, New York, 2014.

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

[10] H. Gfrerer and J. V. Outrata, *On the Aubin property of a class of parameterized variational systems*, Math. Methods Oper. Res., 86 (2017), pp. 443–467.

[11] W. Hager and M. S. Gowda, *Stability in the presence of degeneracy and error estimation*, Math. Program., 85 (1999), pp. 181–192.

[12] N. T. V. Hang, B. S. Mordukhovich, and M. E. Sarabi, *Second-order variational analysis in second-order cone programming*, Math. Program. (2018), https://doi.org/10.1007/s10107-018-1345-6.

[13] A. F. Izmailov, *On the analytical and numerical stability of critical Lagrange multipliers*, Comput. Math. Math. Phys., 45 (2005), pp. 930–946.

[14] A. F. Izmailov and M. V. Solodov, *Stabilized SQP revisited*, Math. Program., 133 (2012), pp. 93–120.

[15] A. F. Izmailov and M. V. Solodov, *Newton-Type Methods for Optimization and Variational Problems*, Springer, New York, 2014.

[16] A. F. Izmailov and M. V. Solodov, *Critical Lagrange multipliers: What we currently know about them, how they spoil our life, and what we can do about it*, TOP, 23 (2015), pp. 1–26.

[17] Y. Liu and S. Pan, *Strong Calmness of Perturbed KKT System for a Class of Conic Programming with Degenerate Solutions*, preprint, https://arxiv.org/abs/1802.01277.

[18] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation,* I: *Basic Theory*, Grundlehren Math. Wiss. 330, Springer, Berlin, 2006.

[19] B. S. Mordukhovich and M. E. Sarabi, *Second-order analysis of piecewise linear functions with applications to optimization and stability*, J. Optim. Theory Appl., 171 (2016), pp. 1–23.

[20] B. S. Mordukhovich and M. E. Sarabi, *Critical multipliers in variational systems via second-order generalized differentiation*, Math. Program., 169 (2018), pp. 605–648.

[21] S. M. Robinson, *Some continuity properties of polyhedral multifunctions*, in Mathematical Programming at Oberwolfach, H. König, B. Korte, and K. Ritter, eds., Math. Program.

Stud. 14, Springer, Berlin, 1981, pp. 206–214.
[22] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.
[23] A. Shapiro, *On uniqueness of Lagrange multipliers in optimization problems subject to cone constraints*, SIAM J. Optim., 7 (1997), pp. 508–518.
[24] A. Shapiro, *Duality, optimality conditions and perturbation analysis*, in Semidefinite Programming and Applications Handbook, Kluwer Academic, Boston, MA, 2000, pp. 67–92.
[25] T. Y. Zhang and L. W. Zhang, *Critical Multipliers in Semidefinite Programming*, preprint, https://arxiv.org/abs/1801.02218v1, 2018.