January 2020

Variational Analysis In Second-Order Cone Programming And Applications

Hang Thi Van Nguyen
Wayne State University

VARIATIONAL ANALYSIS IN SECOND-ORDER CONE PROGRAMMING WITH APPLICATIONS

by

HANG THI VAN NGUYEN

DISSERTATION

Submitted to the Graduate School,

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2020

MAJOR: APPLIED MATHEMATICS

Approved By:

____

Advisor Date

DEDICATION

To my mother and my late father

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude and appreciation for my advisor, Professor Boris Mordukhovich. His ceaseless devotion, support, and encouragement has been the driving force behind all my accomplishments during my doctoral studies. He is the inspiration for the kind of mathematician I aspire to be. It was my great honor to have him as my advisor during the years at Wayne State University.

I am also indebted to my scientific advisor from Hanoi Institute of Mathematics, Professor Nguyen Dong Yen, who has brought me to the beautiful world of optimization and variational analysis and given me a lot of thoughtful advice to follow this academic journey.

I would like to thank the rest of my thesis committee: Professors Alper Murat, Peiyong Wang, and George Yin for taking their time to serve on my dissertation committee.

My sincere thanks also go to the entire Department of Mathematics at Wayne State University for their kind support in a number of ways. Specially, I wish to thank Shereen Schultz and Christopher Leirstein who have trained me and supported me on teaching and many other educational aspects.

Most importantly, none of this would have been possible without the love and endless care of my family and friends. I would like to specially thank my mother for her love, encouragement, and support.

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION

1.1 Subjects

Second-order cone programs (SOCPs for brevity) are optimization problems given in the form

$$\text{minimize } \; f(x) \; \text{ subject to } \; \Phi(x) \in \mathcal{Q}, \tag{1.1}$$

where both function $f\colon \mathbb{R}^n \to \mathbb{R}$ and mapping $\Phi\colon \mathbb{R}^n \to \mathbb{R}^{m+1}$ are twice continuously differentiable ($\mathcal{C}^2$-smooth) around the reference points, and where the underlying set $\mathcal{Q}$ is the *second-order/Lorentz/ice-cream cone* in $\mathbb{R}^{m+1}$ defined by

$$\mathcal{Q} := \big\{ y = (y_0, y_r) \in \mathbb{R} \times \mathbb{R}^m \,\big|\, \|y_r\| \le y_0 \big\}. \tag{1.2}$$

Problems of this type are mathematically challenging while being important for various applications; see, e.g., [1, 5, 6, 42, 46, 47] and the bibliographies therein. A remarkable feature of SOCPs, which significantly distinguishes them from nonlinear programs (NLPs) and the like, is the *nonpolyhedrality* of the underlying second-order cone $\mathcal{Q}$ in the definition of the SOCP constraint system

$$\Gamma := \big\{ x \in \mathbb{R}^n \,\big|\, \Phi(x) \in \mathcal{Q} \big\}. \tag{1.3}$$

The *Karush-Kuhn-Tucker* (KKT) optimality system associated with (1.1) is given by

$$\nabla_x L(x, \lambda) = \nabla f(x) + \nabla \Phi(x)^* \lambda = 0, \quad \lambda \in N_{\mathcal{Q}}\big(\Phi(x)\big), \tag{1.4}$$

where $L(x,\lambda) := f(x) + \langle \lambda, \Phi(x) \rangle$ is the (standard) *Lagrangian* of problem (1.1) with $(x,\lambda) \in \mathbb{R}^n \times \mathbb{R}^{m+1}$. Assume that $\bar{x} \in \mathbb{R}^n$ is a *stationary point* of (1.1), i.e., there exists some $\bar{\lambda} \in \mathbb{R}^{m+1}$ such that $(\bar{x},\bar{\lambda})$ satisfies the KKT system (1.4). Such a vector $\bar{\lambda}$ is call a *Lagrange multiplier* associated with $\bar{x}$. For each $\bar{x} \in \mathbb{R}^n$, define the set of *Lagrange multipliers* associated with $\bar{x}$ by

$$\Lambda(\bar{x}) := \big\{ \lambda \in \mathbb{R}^{m+1} \;\big|\; \nabla_x L(\bar{x},\lambda) = 0,\ \lambda \in N_{\mathcal{Q}}\big(\Phi(\bar{x})\big) \big\}\,. \tag{1.5}$$

Thus $\bar{x}$ is a stationary point of (1.1) if and only if $\Lambda(\bar{x}) \neq \emptyset$.

The intention of this dissertation is to conduct a comprehensive second-order variational analysis for SOCPs by using appropriate tools of second-order generalized differentiation and to illustrate some applications of the obtained results in both stability analysis ans numerical analysis. Our main contribution is threefold:

- proving the *twice epi-differentiability* of the indicator function of $\mathcal{Q}$ and of the *augmented Lagrangian* associated with SOCP (1.1), and deriving explicit formulae for the calculation of the second epi-derivatives of both functions;

- establishing a precise formula–entirely via the initial data– for *calculating the graphical derivative* of the *normal cone mapping* generated by the constraint set Γ in (1.3) *without* imposing any *nondegeneracy* condition;

- conducting a complete convergence analysis of the *Augmented Lagrangian Method* (ALM) for SOCPs (1.1) with *solvability*, *stability* and *local convergence analysis* of both *exact* and *inexact* versions of the ALM under fairly *mild assumptions*.

1.2 Second-Order Generalized Differentiation

Our main devices in second-order variation analysis for SOCPs are *second epi-derivative* and *graphical derivative* of the *normal cone mapping*, see Sect. 2.1 for precise definitions of these constructions.

Rockafellar [55] introduced the concept of the twice epi-differentiability for nonconvex extended-real-valued functions. Second epi-derivative is proved to accumulate vital second-order information of such functions and therefor plays an important role in modern second-order variational analysis, see, e.g., [37]. In this dissertation, we pay a major attention to the second epi-derivative due to its ability to characterize the second-order growth condition and thus to provide a second-order sufficient condition for strict local minimizers of a given function. Twice epi-differentiability and it calculation of the second epi-derivatives of the indicator function $\delta_{\mathcal{Q}}$ and of the augmented Lagrangian function associated with SOCP (1.1) are established without imposing any assumption. By employing the geometry of the second-order cone $\mathcal{Q}$ in (1.2), we obtain explicit formulae for second epi-derivative for both functions, which are new to the best of our knowledge. The obtained results are then showed to have great applications in investigating the constraint system (1.3) and SOCPs (1.1).

Another necessary optimality of the SOCP (1.1) can be read as the following *variational system* via the constraint set Γ:

$$0 \in \nabla f(x) + N_{\Gamma}(x). \tag{1.6}$$

The optimality conditions (1.4) and (1.6) are equivalent under some suitable *qualification* *conditions*, see Sect. 3.1. These carry certain first-order information about SOCPs via the first-order derivative ∇f and the *limiting normal cones* $N_{\mathcal{Q}}$ and N_Γ, see Sect. 2.1 for definitions of normal cones. Therefore, generalized differentiation of normal cone mappings leads us to second-order construction. Study of the nonrobust, tangentially generated graphical derivative (of the normal cone mapping), which is of its own interest, has come to our attention by it application to characterizing the so-called isolated calmness property for parametric constraint or variational systems, see, e.g., [9, 11, 40] and the references therein. Developing calculation for such nonrobust object is a challenging issue, especially when the underlying cone $\mathcal{Q}$ is not a polyhedral, see Remark 3.12. We precisely calculate the graphical derivative of the normal cone mapping generated by (1.3) under merely the metric subregularity constraint qualification. The results obtained here seem to be the first one in the literature for nonpolyhedral problems without imposing any nodegeneracy assumptions.

1.3 Essence of the ALM

The *augmented Lagrangian* $\mathcal{L}\colon \mathbb{R}^n \times \mathbb{R}^{m+1} \times (0,\infty) \to \mathbb{R}$ associated with the SOCP (1.1) is defined by

$$\mathcal{L}(x,\lambda,\rho) := f(x) + \frac{\rho}{2}\mathrm{dist}^2\big(\Phi(x) + \rho^{-1}\lambda; \mathcal{Q}\big) - \frac{1}{2}\rho^{-1}\|\lambda\|^2, \quad (x,\lambda,\rho) \in \mathbb{R}^n \times \mathbb{R}^{m+1} \times (0,\infty), \tag{1.7}$$

where $\lambda \in \mathbb{R}^{m+1}$ is a (vector) multiplier, and where $\rho > 0$ is a penalty parameter of $\mathcal{L}$. The principal idea of the *augmented Lagrangian method* (ALM) for (1.1) is to solve a sequence of unconstrained problems which objectives are defined by the augmented Lagrangian

(1.7) at a given multiplier-parameter pair (λ, ρ); namely,

$$\text{minimize } \mathcal{L}(x, \lambda, \rho) \text{ over } x \in \mathbb{R}^n. \tag{1.8}$$

This means that, given a multiplier λ and a penalty parameter ρ, the ALM solves the unconstrained problem (1.8) for the primal variable x and uses the obtained value to update both the multiplier and penalty parameter in the next iteration.

The ALM was first proposed independently by Hestenes and Powell for nonlinear programming problems (NLPs) with equality constraints [25, 48] and was originally known as the method of multipliers. For the latter framework, Powell observed in [48] that the ALM converges locally with an arbitrarily linear rate if one started the method with a sufficiently high penalty factor (but without the requirement of driving the penalty parameter to infinity) and from a point sufficiently close to a primal-dual pair that satisfies the standard second-order sufficient conditions (SOSC). This is an appealing feature of the ALM, since it provides a numerical stability that cannot be achieved in the usual smooth penalty method.

The ALM was largely extended to various settings of NLPs as well as convex programming with both equality and inequality constraints by Rockafellar [52, 53, 54]; see also the monographs [4, 45, 61] and the references therein. The classical results for the linear convergence of the ALM in NLP framework impose the SOSC, the linear independence constraint qualification (LICQ), and the strict complementarity condition, which all together guarantee the uniqueness of the primal solution as well as the corresponding dual solution/multiplier.

More recently, the study of the ALM has been growing with important theoretical developments. On one hand, various attempts have been made to relax the restrictive assumptions for the convergence of this method in the NLP settings. In such a framework, Fernández and Solodov achieved in [14] a remarkable progress for NLPs by proving that the linear convergence of the primal-dual sequence in the ALM can be ensured if the SOSC alone is satisfied. This result significantly improved the classical ones for NLPs by verifying that neither the LICQ nor the strict complementarity condition is required for local convergence analysis of the ALM. A further improvement was obtained in Izmailov et al. [28] by showing that the conventional SOSC utilized in [14] can be replaced by the noncriticality of Lagrange multipliers for problems with equality constraints. On the other hand, the ALM has been studied for other major classes of constrained optimization including SOCPs [33] and semidefinite programming problems (SDPs) [64]. For $\mathcal{C}^2$-cone reducible problems of conic programming (in the sense of Bonnans and Shapiro [6]), Kanzow and Steck [30, 31] established the linear convergence of the primal-dual sequence generated by modified versions of the ALM under the SOSC and strong Robinson constraint qualification; the latter yields that the Lagrange multiplier is unique. However, the solvability of subproblems in the ALM was not addressed in these papers. We also refer the reader to the paper by Cui et al. [8] and the bibliography therein for recent developments on the ALM for particular classes of convex composite problems of conic programming.

The major goal of Chapters 4 and 5 is to develop both *exact* and *inexact* versions of the ALM for SOCPs under fairly *mild assumptions*. We aim first at establishing the *solvability* and *Lipschitzian stability* of the ALM subproblems by imposing merely the corresponding

SOSC for (1.1) in the general case of *nonunique* Lagrange multipliers. Having this, we verify a local *primal-dual convergence* of iterates with an arbitrary *linear rate* by assuming in addition the *uniqueness* of multipliers. Similarly to Fernández and Solodov [14], our approach revolves around the *second-order growth condition* for the augmented Lagrangian (1.7). To the best of our knowledge, the origin of such a second-order growth condition for NLPs goes back to Rockafellar in [59, Theorem 7.4] from which [14] significantly benefits. However, in contrast to [59], [14] as well as to the vast majority of other publications on numerical optimization, we achieve our goal for (1.1) by employing the concepts of the *second subderivative* and *twice epi-differentiability* of extended-real-valued functions in the framework of second-order variational analysis.

We next recall some properties of the augmented Lagrangian (1.7) that are used below; see, e.g., [60, Exercise 11.56].

Proposition 1.1 (properties of the augmented Lagrangian). *For* (1.7) *with* $(x, \lambda, \rho) \in \mathbb{R}^n \times \mathbb{R}^{m+1} \times (0, \infty)$ *the following hold:*

(i) *The function* $\rho \mapsto \mathcal{L}(x, \lambda, \rho)$ *is nondecreasing.*

(ii) *The function* $\lambda \mapsto \mathcal{L}(x, \lambda, \rho)$ *is concave.*

It follows from the direct differentiation of (1.7) that for any $\rho > 0$ we have

$$\begin{aligned} \nabla_x \mathcal{L}(x, \lambda, \rho) &= \nabla f(x) + \nabla \Phi(x)^* \Pi_{-\mathcal{Q}}\big(\rho\Phi(x) + \lambda\big)\big), \\ \nabla_\lambda \mathcal{L}(x, \lambda, \rho) &= \rho^{-1}\big[\Pi_{-\mathcal{Q}}\big(\rho\Phi(x) + \lambda\big) - \lambda\big], \end{aligned} \tag{1.9}$$

which allows us to readily deduce that $(\bar{x}, \bar{\lambda})$ is a solution to the KKT system (1.4) if and only if for any $\rho > 0$ this pair satisfies the equation

$$\big(\nabla_x \mathcal{L}(x,\lambda,\rho), \nabla_\lambda \mathcal{L}(x,\lambda,\rho)\big) = (0,0). \tag{1.10}$$

Finally, let us list some properties of the projection mapping for the second-order cone $\mathcal{Q}$ that are extensively exploited in studying of augmented Lagragians:

(P1) $p = \Pi_{\mathcal{Q}}(y)$ if and only if $p \in \mathcal{Q}$, $\langle y - p, p\rangle = 0$, and $y - p \in -\mathcal{Q}$.

(P2) For every $y \in \mathbb{R}^{m+1}$ we have $y = \Pi_{\mathcal{Q}}(y) + \Pi_{-\mathcal{Q}}(y)$.

(P3) For every $y \in \mathbb{R}^{m+1}$ we have $\big\langle \Pi_{\mathcal{Q}}(y), \Pi_{-\mathcal{Q}}(y)\big\rangle = 0$.

(P4) $\lambda \in N_{\mathcal{Q}}(y)$ if and only if $\Pi_{\mathcal{Q}}(y + \lambda) = y$.

1.4 Overview of the Contents

Chapter 2 is mainly devoted to the study of *twice epi-differentiability* (in the sense of Rockafellar [57]) of the indicator function $\delta_{\mathcal{Q}}$ of the second-order cone (1.2). We start by reviewing some important notions of variational analysis and generalized differentiation that are broadly used throughout the whole dissertation. The main result here not only justifies the twice epi-differentiability of $\delta_{\mathcal{Q}}$, but also establishes a precise formula for calculating the second epi-derivative of this function in terms of the given data of $\mathcal{Q}$ without any additional assumptions. We conclude by presenting some of its consequences and related properties.

Chapter 3 concerns computation of *graphical derivative of the normal cone mapping* to the constraint set (1.3). The first section is devoted to the study of second-order properties of the SOCP constraint system (1.3) by using the twice epi-differentiability of $\delta_{\mathcal{Q}}$ and the *metric subregularity constraint qualification* (MSCQ) for (1.3), which seems to be the weakest constraint qualification that has been investigated and employed recently in the (polyhedral) NLP framework; see [19, 16, 7]. Among the most important results obtained in this first section we mention the following: (i) a constructive description of generalized normals to the critical cone at the point in question under MSCQ, and (ii) a characterization of the uniqueness of Lagrange multipliers *together* with an appropriate error bound estimate (automatic in the polyhedral case) at stationary points via a new constraint qualification in conic programming, which happens to be in the case of (1.3) a *dual* form of the *strict Robinson constraint qualification* (SRCQ) from [6, 9]. We also present here novel *approximate duality* relationships for a linear conic optimization problem associated with the second-order cone $\mathcal{Q}$ that play a significant role in establishing the main result of the paper. In the next section, we derive a new formula allowing us to precisely calculate the *graphical derivative* of the normal cone mapping generated by (1.3), merely under the validity of MSCQ. The obtained major result is the first in the literature for nonpolyhedral constraint systems without imposing nondegeneracy. As discussed below, its proof is significantly different from the recent ones given in [7, 16, 19] for polyhedral systems, even in the latter case. It is also largely different from the approaches developed in [20, 40, 41] for conic programs under nondegeneracy assumptions. We present in the end of this section a nontrivial example of a two-dimensional constraint system (1.3) with the three-dimensional second-order cone $\mathcal{Q}$ illustrating applications of the graphical derivative formula. In this example the MSCQ condition holds at any feasible point of (1.3) while the nondegeneracy and metric regularity/Robinson constraint qualification fail therein. Finally, we apply the obtained graphical derivative formula to deriving a complete *characterization* of the *isolated calmness* property for solution maps to canonically perturbed variational systems associated with SOCP and give a numerical example.

Chapter 4 conducts a comprehensive *second-order variational analysis* of the *augmented Lagrangian* associated with the SOCP (1.1), see (1.7) for definition of this function. Based on the obtained precise computation of the second subderivative of (1.7), we characterize here the *second-order growth* condition for (1.7) via the SOSC and then establish its *uniform* counterpart needed in the general case of nonunique Lagrange multipliers.

Chapter 5 focuses on convergence analysis of the *Augmented Lagrangian Method* (ALM) for SOCPs (1.1). In the first section, we provide an *error bound* estimate for the canonically perturbed KKT system associated with (1.1) under the SOSC and a certain *calmness* property of the multiplier mapping with respect to perturbations that automatically holds for NLPs. We also present here an example showing that the imposed calmness property is essential for the validity of the error bound in the SOCP setting and then discuss efficient conditions ensuring the fulfillment of this calmness for nonpolyhedral SOCPs. We then give a detailed *solvability*, *stability*, and *local convergence analysis* of the suggested ALM algorithm for SOCPs that strongly exploits the SOSC and obtained second-order growth conditions. Our analysis includes the proof of solvability of the ALM subproblems in both exact and inexact versions and then establishes the linear convergence of primal-dual iterates to the designated solution of the KKT systems under the SOSC by using the established *robust isolated calmness* and *upper Lipschitzian* properties of the corresponding perturbed multiplier mappings. In this way we obtain explicit relationships between the constants involved in the algorithm and the imposed assumptions on the given data.

1.5 Notation

Our notation and terminology are standard in variational analysis, conic programming, and generalized differentiation; see, e.g., [6, 38, 60]. Recall that $\mathbb{B}$ and $\mathbb{S}$ stand for the closed unit ball and the unit sphere, respectively, of the space in question, and that $\mathbb{B}_\gamma(x) := x + \gamma\mathbb{B}$ is the closed ball centered at x with radius $\gamma > 0$. A^* indicates the transpose of a matrix A, while φ^* and K^* signify respectively the conjugate of a function φ and the polar cone of a set K. Given a nonempty set $\Omega \subset \mathbb{R}^n$, the symbols $\operatorname{int}\Omega$, $\operatorname{ri}\Omega$, $\operatorname{bd}\Omega$, and $\Omega^\perp$ signify its interior, relative interior, boundary, and orthogonal complement space, respectively. The indicator function of Ω is defined by $\delta_\Omega(x) := 0$ for $x \in \Omega$ and $\delta_\Omega(x) := \infty$ otherwise, $\operatorname{dist}(x;\Omega)$ signifies the distance between $x \in \mathbb{R}^n$ and the set Ω, and the projection of x onto Ω is denoted by $\Pi_\Omega(x)$. The symbol $x \xrightarrow{\Omega} \bar{x}$ indicates that $x \to \bar{x}$ with $x \in \Omega$. As in (1.2), we often decompose a vector $y \in \mathcal{Q} \subset \mathbb{R}^{m+1}$ into $y = (y_0, y_r)$ with $y_0 \in \mathbb{R}$ and $y_r \in \mathbb{R}^m$. Taking this decomposition into account, denote $\widetilde{y} := (-y_0, y_r)$. Similarly, for a mapping $\Phi\colon \mathbb{R}^n \to \mathbb{R}^{m+1}$ with $\Phi = (\Phi_0, \ldots, \Phi_m)$, we often implement the decomposition of the vector $\Phi(x)$ into $(\Phi_0(x), \Phi_r(x)) \in \mathbb{R} \times \mathbb{R}^m$ and denote by $\widetilde{\Phi}(x)$ the vector $(-\Phi_0(x), \Phi_r(x))$ for any $x \in \mathbb{R}^n$.

CHAPTER 2 TWICE EPI-DIFFERENTIABILITY OF THE INDICATOR FUNCTION OF THE SECOND-ORDER CONE

2.1 Tools of Variational Analysis

In this first section, we briefly review constructions of variational analysis and generalized differentiation; see [6, 39, 60] for more details and references. Given a nonempty set $\Theta \subset \mathbb{R}^n$ with $\bar{x} \in \Theta$, the (Bouligand-Severi) *tangent/contingent cone* $T_\Theta(\bar{x})$ to Θ at $\bar{x}$ is defined by

$$T_\Theta(\bar{x}) := \big\{ w \in \mathbb{R}^n \;\big|\; \exists\, t_k \downarrow 0,\;\; w^k \to w \;\;\text{as}\;\; k \to \infty \;\;\text{with}\;\; \bar{x} + t_k w^k \in \Theta \big\}, \tag{2.1}$$

while the (Mordukhovich) *basic/limiting normal cone* $N_\Theta(\bar{x})$ to Θ at this point is given by

$$N_\Omega(\bar{x}) := \operatorname*{Lim\,sup}_{x \to \bar{x}} \big[\text{cone}\big(x - \Pi_\Omega(x)\big)\big], \tag{2.2}$$

where $\Pi_\Theta \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ stands for the *Euclidean projector* onto the set Θ. If the set Θ is convex, then constructions (2.1) and (2.2) reduce, respectively, to the classical tangent and normal cones of convex analysis. In this setting, it then holds that

$$N_\Theta(\bar{x}) := \big\{ v \in \mathbb{R}^n \;\big|\; \langle v, x - \bar{x} \rangle \le 0 \big\}] \quad \text{and} \quad T_\Theta(\bar{x}) = N_\Theta(\bar{x})^*.$$

n the case where $\Theta = \mathcal{Q}$, the second-order cone (1.2), we get, respectively, the expressions

$$T_{\mathcal{Q}}(y) = \begin{cases} \mathbb{R}^{m+1} & \text{if } \ y \in \operatorname{int} \mathcal{Q}, \\ \mathcal{Q} & \text{if } \ y = 0, \\ \big\{ y' \in \mathbb{R}^{m+1} \ \big| \ \langle \widetilde{y}, y' \rangle \le 0 \big\} & \text{if } \ y \in (\operatorname{bd} \mathcal{Q}) \setminus \{0\}, \end{cases}$$

$$N_{\mathcal{Q}}(y) = \begin{cases} \{0\} & \text{if } \ y \in \operatorname{int} \mathcal{Q}, \\ -\mathcal{Q} & \text{if } \ y = 0, \\ \mathbb{R}_+ \widetilde{y} & \text{if } \ y \in (\operatorname{bd} \mathcal{Q}) \setminus \{0\}. \end{cases} \tag{2.3}$$

Given further an extended-real-valued function $\varphi \colon \mathbb{R}^n \to \overline{\mathbb{R}} := (\infty, \infty]$, its domain and epigraph are defined, respectively, by

$$\operatorname{dom} \varphi := \big\{ x \in \mathbb{R}^n \ \big| \ \varphi(x) < \infty \big\} \quad \text{and} \quad \operatorname{epi} \varphi := \big\{ (x, \alpha) \in \mathbb{R}^{n+1} \ \big| \ \varphi(x) \le \alpha \big\}.$$

Given $\bar{x} \in \operatorname{dom} \varphi$, the (first-order) *subdifferential* of φ at $\bar{x}$ is defined via the epigraph $\operatorname{epi} \varphi$ by

$$\partial \varphi(\bar{x}) := \big\{ v \in \mathbb{R}^n \big| \ (v, -1) \in N_{\operatorname{epi} \varphi}(\bar{x}, \varphi(\bar{x})) \big\}. \tag{2.4}$$

Considering next a set-valued mapping $F \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ with its domain and graph given by

$$\operatorname{dom} F := \big\{ x \in \mathbb{R}^n \ \big| \ F(x) \ne \emptyset \big\} \ \text{ and } \ \operatorname{gph} F := \big\{ (x, y) \in \mathbb{R}^n \times \mathbb{R}^m \ \big| \ y \in F(x) \big\},$$

we define the following generalized differential notions for F induced by the above tangent and normal cones to its graph. Given $(\bar{x}, \bar{y}) \in \operatorname{gph} F$, the *graphical derivative* of F at $(\bar{x}, \bar{y})$ is

$$DF(\bar{x}, \bar{y})(u) := \big\{v \in \mathbb{R}^m \big| \ (u, v) \in T_{\operatorname{gph} F}(\bar{x}, \bar{y})\big\}, \quad u \in \mathbb{R}^n, \tag{2.5}$$

while the *limiting coderivative* to F at $(\bar{x}, \bar{y})$ is defined by

$$D^*F(\bar{x}, \bar{y})(v) := \big\{u \in \mathbb{R}^n \big| \ (u, -v) \in N_{\operatorname{gph} F}(\bar{x}, \bar{y})\big\}, \quad v \in \mathbb{R}^m. \tag{2.6}$$

Recall that $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is *metrically regular* around $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ if there is $\ell \geq 0$ such that we have the distance estimate

$$\operatorname{dist}\big(x; F^{-1}(y)\big) \leq \ell \operatorname{dist}\big(y; F(x)\big) \qquad \text{for all} \quad (x, y) \ \text{close to} \ (\bar{x}, \bar{y}). \tag{2.7}$$

We say that F is *metrically subregular* at $(\bar{x}, \bar{y})$ if the estimate in (2.7) holds for all x close to $\bar{x}$ and $y = \bar{y}$.

The mapping F is said to be *calm* at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ if there exist $\tau \geq 0$ and neighborhoods U of $\bar{x}$ and V of $\bar{y}$ for which

$$F(x) \cap V \subset F(\bar{x}) + \tau\|x - \bar{x}\|\mathbb{B} \ \text{ whenever } \ x \in U. \tag{2.8}$$

It is known that F is metrically subregular at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ if and only if it inverse F^{-1} is calm at $(\bar{y}, \bar{x}) \in \operatorname{gph} F^{-1}$.

It is said that F has the *isolated calmness* property at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ if (2.8) holds with the replacement of $F(\bar{x})$ by $\{\bar{y}\}$ on the right-hand side therein. Furthermore, F has the *robust isolated calmness* property at $(\bar{x},\bar{y})$ if

$$F(x)\cap V\subset\{\bar{y}\}+\ell\|x-\bar{x}\|\mathbb{B}\quad\text{with}\quad F(x)\cap V\neq\emptyset\quad\text{for all}\quad x\in U. \tag{2.9}$$

Properties of this type go back to Robinson [49] who introduced the *upper Lipschitzian* version of calmness corresponding to (2.8) with $V=\mathbb{R}^m$. Similarly to (2.9), we say that F has the *robust isolated upper Lipschitzian* property if (2.9) holds with $V=\mathbb{R}^m$. It is well known that (2.8) is equivalent to the metric subregularity of the inverse mapping F^{-1} at $(\bar{y},\bar{x})$. These "one point" properties are more subtle and essentially less investigated than their robust "two-points" counterparts (as metric regularity and Lipschitz-like/Aubin ones), while their importance for optimization theory, numerical algorithms, and applications has been broadly recognized in the literature; see, e.g., [8, 11, 12, 17, 24, 29, 39, 36, 65] with the references and discussions therein.

Turning now to the constructions of second-order variational analysis, for a function $\varphi\colon\mathbb{R}^n\to\overline{\mathbb{R}}$, define the parametric family of *second-order difference quotients* at $\bar{x}$ for $\bar{v}\in\mathbb{R}^n$ by

$$\Delta_t^2\varphi(\bar{x},\bar{v})(w)=\frac{\varphi(\bar{x}+tw)-\varphi(\bar{x})-t\langle\bar{v},w\rangle}{\frac{1}{2}t^2}\quad\text{with}\quad w\in\mathbb{R}^n,\ t>0. \tag{2.10}$$

If $\varphi(\bar{x})$ is finite, the *second subderivative* of φ at $\bar{x}$ for $\bar{v}$ and w is defined by

$$d^2\varphi(\bar{x},\bar{v})(w)=\liminf_{\substack{t\downarrow 0\\ w'\to w}}\Delta_t^2\varphi(\bar{x},\bar{v})(w'). \tag{2.11}$$

Following [60, Definition 13.6], a function $\varphi\colon\mathbb{R}^n\to\overline{\mathbb{R}}$ is said to be *twice epi-differentiable* at $\bar{x}$ for $\bar{v}$ if the sets $\operatorname{epi}\Delta_t^2\varphi(\bar{x},\bar{v})$ converge to $\operatorname{epi}\mathrm{d}^2\varphi(\bar{x},\bar{v})$ as $t\downarrow 0$. If in addition the second subderivative is a proper function (i.e., does not take the value $-\infty$ and is finite at some point), then we say that φ is *properly* twice epi-differentiable at $\bar{x}$ for $\bar{v}$. The twice epi-differentiability of φ at $\bar{x}$ for $\bar{v}$ can be understood equivalently by [60, Proposition 7.2] as that for every $w\in\mathbb{R}^n$ and every sequence $t_k\downarrow 0$ there exists a sequence $w^k\to w$ with

$$\Delta_{t_k}^2\varphi(\bar{x},\bar{v})(w^k)\to\mathrm{d}^2\varphi(\bar{x},\bar{v})(w).$$

Twice epi-differentiability, together with a precise calculation of the second subderivative (2.11) of the augmented Lagrangian (1.7) associated with (1.1), plays a major role in our developments. This property was introduced by Rockafellar in [57] who verified it for *fully amenable* compositions. Quite recently [34, 35, 37], the class of extended-real-valued functions satisfying this property has been dramatically enlarged by showing that twice epi-differentiability holds under *parabolic regularity*, which covers the SOCP setting; see more details in the cited papers.

2.2 Twice Epi-Differentiability of the Indicator Function of $\mathcal{Q}$

We begin our second-order analysis with the study of twice epi-differentiability of the indicator function $\delta_{\mathcal{Q}}$ of the second-order cone (1.2). The notions of first- and second-order epi-differentiability for extended-real-valued functions were introduced by Rockafellar in [57], where he proved the twice epi-differentiability of convex piecewise linear-quadratic functions in finite dimensions. This result was extended in [60, Theorem 14.14] to the class of fully amenable functions based on their polyhedral structure. Furthermore, Do [10, Example 2.10] established the twice epi-differentiability of the indicator function of convex polyhedric sets in reflexive Banach spaces while Levy [32, Theorem 2.1] proved this fact in the general nonreflexive Banach space setting. Note that polyhedric sets reduce to polyhedral ones in finite dimensions.

The following theorem justifies the twice epi-differentiability of the indicator function $\delta_{\mathcal{Q}}$ of the second-order cone (1.2) and calculates its second-order epi-derivative via the given data of $\mathcal{Q}$. Recall that

$$\overline{\mathcal{K}} := T_{\mathcal{Q}}(\bar{x}) \cap \{\bar{y}\}^{\perp} \tag{2.12}$$

defines the *critical cone* of $\mathcal{Q}$ at $\bar{x}$ for $\bar{y}$.

Theorem 2.1. (second-order epi-derivative of the indicator function of $\mathcal{Q}$). *Given any $\bar{x}$ from the second-order cone $\mathcal{Q}$ in* (1.2)*, we have that the indicator function $\delta_{\mathcal{Q}}$ is twice epi-differentiable at $\bar{x}$ for every $\bar{y} \in N_{\mathcal{Q}}(\bar{x})$ and its second-order epi-derivative is calculated by*

$$d^2\delta_{\mathcal{Q}}(\bar{x}|\bar{y})(v) = \begin{cases} 0 & \text{if } \bar{x} \in [\text{int}(\mathcal{Q}) \cup \{0\}],\ v \in \overline{\mathcal{K}}, \\ \dfrac{\|\bar{y}\|}{\|\bar{x}\|}(\|v_r\|^2 - v_0^2) & \text{if } \bar{x} \in \text{bd}\,(\mathcal{Q}) \setminus \{0\},\ v \in \overline{\mathcal{K}}, \\ \infty & \text{if } v \notin \overline{\mathcal{K}}. \end{cases} \tag{2.13}$$

Proof. Fix $\bar{x} \in \mathcal{Q}$, $\bar{y} \in N_{\mathcal{Q}}(\bar{x})$, and $v \in \mathbb{R}^{m+1}$ and denote by $\Delta(\bar{x}, \bar{y})(v)$ the right-hand side of (2.13). To verify formula (2.13), we apply [60, Proposition 7.2] that gives us the following description of the twice epi-differentiability of $\delta_{\mathcal{Q}}$ at $\bar{x}$ for $\bar{y}$:

- For every sequences $t_k \downarrow 0$ and $v^k \to v$ the second-order difference quotients (2.10) satisfy

$$\liminf_{k\to\infty} \Delta_{t_k}^2 \delta_{\mathcal{Q}}(\bar{x}|\bar{y})(v^k) \geq \Delta(\bar{x}, \bar{y})(v). \tag{2.14}$$

- For every sequence $t_k \downarrow 0$ there is some sequence $v^k \to v$ satisfying the inequality

$$\limsup_{k\to\infty} \Delta^2_{t_k}\delta_{\mathcal{Q}}(\bar{x}|\bar{y})(v^k) \le \Delta(\bar{x},\bar{y})(v). \tag{2.15}$$

We split the proof into considering the three cases for $\bar{x} \in \mathcal{Q}$ in representation (2.13).

Case 1: $\bar{x} \in \text{int}(\mathcal{Q})$. In this case we have $N_{\mathcal{Q}}(\bar{x}) = \{0\}$ and hence $\bar{y} = 0$. Fix $v \in \overline{\mathcal{K}} = \mathbb{R}^{m+1}$ and observe from (2.13) that $\Delta(\bar{x}, 0)(v) = 0$. Picking an arbitrary sequence $v^k \to v$ as $k \to \infty$, we arrive at

$$\Delta^2_{t_k}\delta_{\mathcal{Q}}(\bar{x}|0)(v^k) = \frac{\delta_{\mathcal{Q}}(\bar{x} + t_k v^k) - \delta_{\mathcal{Q}}(\bar{x}) - t_k \cdot 0}{\frac{1}{2}t_k^2} = 0$$

for all k sufficiently large. This tells us that

$$\lim_{k\to 0} \Delta^2_{t_k}\delta_{\mathcal{Q}}(\bar{x}|0)(v^k) = 0 = \Delta(\bar{x},0)(v),$$

which justifies conditions (2.14) and (2.15), and, therefore, formula (2.13) in this case.

Case 2: $\bar{x} = 0$. In this case we have $\bar{y} \in N_{\mathcal{Q}}(\bar{x}) = -\mathcal{Q}$. Pick $v \in \mathbb{R}^{m+1}$ and let $v^k \to v$ as $k \to \infty$. Using (2.10) gives us the representations

$$\Delta^2_{t_k}\delta_{\mathcal{Q}}(0|\bar{y})(v^k) = \frac{\delta_{\mathcal{Q}}(t_k v^k) - \delta_{\mathcal{Q}}(0) - t_k\langle \bar{y}, v^k\rangle}{\frac{1}{2}t_k^2} = \begin{cases} -\dfrac{\langle \bar{y}, v^k\rangle}{\frac{1}{2}t_k} \ge 0 & \text{if } v^k \in \mathcal{Q}, \\ \infty & \text{if } v^k \notin \mathcal{Q}. \end{cases} \tag{2.16}$$

If $v \in \overline{\mathcal{K}}$, we conclude from the above definition of $\Delta(\bar{x},\bar{y})(v)$ that $\Delta(0,\bar{y})(v) = 0$. Thus (2.14) comes directly from (2.16), while (2.15) can be justified by choosing $v^k = v$ for any

k. Pick now $v \notin \overline{\mathcal{K}} = \mathcal{Q} \cap \{\bar{y}\}^\perp$ and observe that it amounts to saying that either $v \notin \mathcal{Q}$ or $\langle \bar{y}, v \rangle < 0$. It follows from the definition of $\Delta(\bar{x}, \bar{y})(v)$ in this case that $\Delta(0, \bar{y})(v) = \infty$, and hence inequality (2.15) holds. To verify (2.14), pick an arbitrary sequence $v^k \to v$. If $v \notin \mathcal{Q}$, then we can assume without loss of generality that $v^k \notin \mathcal{Q}$ for all k, which together with (2.16) ensures (2.14). The verification of (2.14) for $\langle \bar{y}, v \rangle < 0$ is similar.

Case 3: $\bar{x} \in \mathrm{bd}\,(\mathcal{Q}) \setminus \{0\}$. Defining the mapping $\psi \colon \mathbb{R}^{m+1} \to \mathbb{R}^2$ by

$$\psi(x_0, x_r) := \big(\|x_r\|^2 - x_0^2, -x_0\big), \quad (x_0, x_r) \in \mathbb{R} \times \mathbb{R}^m, \tag{2.17}$$

observe the following representations of the Lorentz cone and its indicator function, respectively:

$$\mathcal{Q} = \big\{x \in \mathbb{R}^{m+1} \;\big|\; \psi(x) \in \mathbb{R}^2_-\big\} \qquad \text{and} \qquad \delta_{\mathcal{Q}} = \delta_{\mathbb{R}^2_-} \circ \psi. \tag{2.18}$$

For any $v \in \mathbb{R}^{m+1}$ and $t > 0$ we form the vector

$$w := \frac{\psi(\bar{x} + tv) - \psi(\bar{x})}{t} \tag{2.19}$$

and use it to write down the relationships

$$\delta_{\mathcal{Q}}(\bar{x} + tv) = \delta_{\mathbb{R}^2_-}(\psi(\bar{x}) + tw) \quad \text{and} \quad \delta_{\mathcal{Q}}(\bar{x}) = \delta_{\mathbb{R}^2_-}\big(\psi(\bar{x})\big). \tag{2.20}$$

It is easy to see that $\nabla\psi(\bar{x})$ is surjective due to $\bar{x} \in \mathrm{bd}\,(\mathcal{Q}) \setminus \{0\}$. Employing the first-order chain rule, we get $N_{\mathcal{Q}}(\bar{x}) = \nabla\psi(\bar{x})^* N_{\mathbb{R}^2_-}(\psi(\bar{x}))$. This together with $\bar{y} \in N_{\mathcal{Q}}(\bar{x})$ yields the existence of some $\bar\lambda \in N_{\mathbb{R}^2_-}(\psi(\bar x))$ for which $\bar y = \nabla\psi(\bar x)^*\bar\lambda$. This allows us to arrive at

$$-t\langle \bar y, v\rangle = -t\langle \nabla\psi(\bar x)^*\bar\lambda, v\rangle = -t\langle \bar\lambda, w\rangle + \langle \bar\lambda, t(w - \nabla\psi(\bar x)v)\rangle.$$

Furthermore, it follows from (2.19) that

$$t\big(w - \nabla\psi(\bar x)v\big) = \psi(\bar x + tv) - \psi(\bar x) - t\nabla\psi(\bar x)v = \frac{1}{2}t^2\big\langle \nabla^2\psi(\bar x)v, v\big\rangle + o(t^2),$$

which in turn leads us to the representation

$$-t\langle \bar y, v\rangle = -t\langle \bar\lambda, w\rangle + \frac{1}{2}t^2\left\langle \nabla^2\langle \bar\lambda, \psi\rangle(\bar x)v, v\right\rangle + o(t^2).$$

Combining the latter with (2.20) and (2.10) readily yields

$$\begin{aligned} \Delta_t^2\delta_{\mathcal Q}(\bar x|\bar y)(v) &= \frac{\delta_{\mathbb{R}^2_-}(\psi(\bar x) + tw) - \delta_{\mathbb{R}^2_-}(\psi(\bar x)) - t\langle \bar\lambda, w\rangle}{\frac{1}{2}t^2} + \big\langle \nabla^2\langle \bar\lambda, \psi\rangle(\bar x)v, v\big\rangle + \frac{o(t^2)}{t^2} \\ &= \Delta_t^2\delta_{\mathbb{R}^2_-}\big(\psi(\bar x)|\bar\lambda\big)(w) + \big\langle \nabla^2\langle \bar\lambda, \psi\rangle(\bar x)v, v\big\rangle + \frac{o(t^2)}{t^2}. \end{aligned} \tag{2.21}$$

Pick next arbitrary sequences $v^k \to v$ and $t_k \downarrow 0$, and define $w^k := \dfrac{\psi(\bar x + t_k v^k) - \psi(\bar x)}{t_k}$

similarly to (2.19). Since $w^k \to \nabla\psi(\bar{x})v$ as $k \to \infty$, we conclude from (2.21) that

$$\begin{aligned}
\liminf_{k\to\infty} \Delta^2_{t_k}\delta_{\mathcal{Q}}(\bar{x}|\bar{y})(v^k) &= \liminf_{k\to\infty} \Big\{\Delta^2_{t_k}\delta_{\mathbb{R}^2_-}\big(\psi(\bar{x})|\bar{\lambda}\big)(w^k) + \big\langle\nabla^2\langle\bar{\lambda},\psi\rangle(\bar{x})v^k, v^k\big\rangle + \frac{o(t_k^2)}{t_k^2}\Big\} \\
&\geq \big\langle\nabla^2\langle\bar{\lambda},\psi\rangle(\bar{x})v, v\big\rangle + \inf_{u^k\to\nabla\psi(\bar{x})v} \liminf_{k\to\infty} \Delta^2_{t_k}\delta_{\mathbb{R}^2_-}\big(\psi(\bar{x})|\bar{\lambda}\big)(u^k) \\
&\geq \begin{cases} \big\langle\nabla^2\langle\bar{\lambda},\psi\rangle(\bar{x})v, v\big\rangle & \text{if } \nabla\psi(\bar{x})v \in T_{\mathbb{R}^2_-}\big(\psi(\bar{x})\big) \cap \{\bar{\lambda}\}^\perp, \\ \infty & \text{otherwise,} \end{cases}
\end{aligned}$$

where the last inequality comes from [60, Proposition 13.9] in which the twice epi-differentiability of the indicator function of a convex polyhedron was established. On the other hand, it follows from the surjectivity of $\nabla\psi(\bar{x})$ and (2.18) that

$$v \in T_{\mathcal{Q}}(\bar{x}) \cap \{\bar{y}\}^\perp \Longleftrightarrow \nabla\psi(\bar{x})v \in T_{\mathbb{R}^2_-}\big(\psi(\bar{x})\big) \cap \{\bar{\lambda}\}^\perp,$$

which in turn leads us to the estimate

$$\liminf_{k\to\infty} \Delta^2_{t_k}\delta_{\mathcal{Q}}(\bar{x}|\bar{y})(v^k) \geq \begin{cases} \big\langle\nabla^2\langle\bar{\lambda},\psi\rangle(\bar{x})v, v\big\rangle & \text{if } v \in T_{\mathcal{Q}}(\bar{x}) \cap \{\bar{y}\}^\perp, \\ \infty & \text{otherwise.} \end{cases} \tag{2.22}$$

To finish the proof of (2.14), recall that $\bar{\lambda} \in N_{\mathbb{R}^2_-}(\psi(\bar{x}))$ with $\bar{x} = (\bar{x}_0, \bar{x}_r) \in \text{bd}\,(\mathcal{Q}) \setminus \{0\}$. Therefore we get the representation $\bar{\lambda} = (\bar{\alpha}, 0)$ with some $\bar{\alpha} \geq 0$ and so deduce from here and the notation $\widetilde{\bar{x}}$ introduced in Sect. 2.1 the following equalities:

$$\bar{y} = \nabla\psi(\bar{x})^*\bar{\lambda} = \begin{bmatrix} -2\bar{x}_0 & -1 \\ 2\bar{x}_r & 0 \end{bmatrix} \begin{pmatrix} \bar{\alpha} \\ 0 \end{pmatrix} = 2\bar{\alpha}\widetilde{\bar{x}},$$

which yield $\bar{\alpha} = \dfrac{\|\bar{y}\|}{2\|\widetilde{\bar{x}}\|} = \dfrac{\|\bar{y}\|}{2\|\bar{x}\|}$. Employing now (2.17) brings us to the relationships

$$\langle \bar{\lambda}, \psi \rangle(\bar{x}) = \bar{\alpha}(-\bar{x}_0^2 + \|\bar{x}_r\|^2), \quad \nabla^2 \langle \bar{\lambda}, \psi \rangle(\bar{x}) = 2\bar{\alpha} \begin{bmatrix} -1 & 0 \\ 0 & I \end{bmatrix},$$

$$\left\langle \nabla^2 \langle \bar{\lambda}, \psi \rangle(\bar{x})v, v \right\rangle = 2\bar{\alpha}(-v_0^2 + \|v_r\|^2) = \frac{\|\bar{y}\|}{\|\bar{x}\|}(-v_0^2 + \|v_r\|^2). \tag{2.23}$$

Unifying it with (2.22) verifies the first condition (2.14) in the second-order epi-differentiability.

It remains to prove the other condition (2.15) in the framework of Case 3. The latter inequality clearly holds when the right-hand side of it equals infinity. Thus we only need to consider the situation where $v \in \overline{\mathcal{K}}$ with the critical cone $\overline{\mathcal{K}}$ described by

$$\overline{\mathcal{K}} = T_{\mathcal{Q}}(\bar{x}) \cap \{\bar{y}\}^{\perp} = \begin{cases} \left\{ u \in \mathbb{R}^{m+1} \;\middle|\; \langle u, \widetilde{\bar{x}} \rangle \le 0 \right\} & \text{if } \bar{y} = 0, \\ \left\{ u \in \mathbb{R}^{m+1} \;\middle|\; \langle u, \widetilde{\bar{x}} \rangle = 0 \right\} & \text{if } \bar{y} \neq 0. \end{cases}$$

Construct a sequence $v^k \to v$ satisfying (2.15) based on the position of v in $\overline{\mathcal{K}}$ as follows:

Case 3(i): $v \in \text{bd}\,(\overline{\mathcal{K}}) \cap \mathcal{Q}$ or $v \in \text{int}(\overline{\mathcal{K}})$. Having $v = (v_0, v_r) \in \mathbb{R} \times \mathbb{R}^m$, define $v^k := v$ for any k and claim that $\bar{x} + tv = (\bar{x}_0 + tv_0, \bar{x}_r + tv_r) \in \mathcal{Q}$ when $t > 0$ is small enough. This is clear if $v \in \text{bd}\,(\overline{\mathcal{K}}) \cap \mathcal{Q}$. To justify the claim, it suffices to show that

$$\bar{x}_0 + tv_0 \ge \|\bar{x}_r + tv_r\| \tag{2.24}$$

for all small $t > 0$ provided that $v \in \text{int}(\overline{\mathcal{K}})$. We easily derive that $\langle \widetilde{\bar{x}}, v \rangle < 0$ and $\|\bar{x}_r\| = \bar{x}_0 > 0$ from the facts that $v \in \text{int}(\overline{\mathcal{K}})$ and $\bar{x} = (\bar{x}_0, \bar{x}_r) \in \text{bd}\,(\mathcal{Q}) \setminus \{0\}$, respectively. This yields

$$\bar{x}_0 + tv_0 > 0 \qquad \text{and} \qquad \langle v_r, \bar{x}_r\rangle - \bar{x}_0 v_0 + t(\|v_r\|^2 - v_0^2) < 0$$

for t sufficiently small. The above inequalities tell us that $(\bar{x}_0 + tv_0)^2 > \|\bar{x}_r + tv_r\|^2$, which thus verifies (2.24). Letting $t_k \downarrow 0$, we deduce from $\bar{x} + t_k v \in \mathcal{Q}$ and $v \in \{\bar{y}\}^\perp$ that

$$\Delta^2_{t_k} \delta_{\mathcal{Q}}(\bar{x}|\bar{y})(v^k) = \frac{\delta_{\mathcal{Q}}(\bar{x} + t_k v) - \delta_{\mathcal{Q}}(\bar{x}) - t_k\langle \bar{y}, v\rangle}{\frac{1}{2}t_k^2} = 0 \tag{2.25}$$

for k sufficiently large. It is not hard to see furthermore that

$$\Delta(\bar{x}, \bar{y})(v) = \frac{\|\bar{y}\|}{\|\bar{x}\|}(-v_0^2 + \|v_r\|^2) = 0.$$

Combining this with (2.25) justifies (2.15) under the imposed conditions on v.

Case 3(ii): $v = (v_0, v_r) \in \mathrm{bd}\,(\overline{\mathcal{K}}) \setminus Q$. Assume without loss of generality that $\|\bar{x}\| = \|v\| = 1$. Remembering that $\widetilde{\bar{x}} = (-\bar{x}_0, \bar{x}_r)$ according to the notation of Sect. 2.1, we conclude from $-\bar{x}_0 v_0 + \langle \bar{x}_r, v_r\rangle = \langle \widetilde{\bar{x}}, v\rangle = 0$ and $\bar{x} = (\bar{x}_0, \bar{x}_r) \in \mathrm{bd}\,(\mathcal{Q}) \setminus \{0\}$ that

$$\|v_r\|^2 - v_0^2 \geq 0. \tag{2.26}$$

Letting $t_k \downarrow 0$ and employing (2.21) and (2.23) yield

$$\limsup_{k\to\infty} \Delta^2_{t_k} \delta_{\mathcal{Q}}(\bar{x}|\bar{y})(v^k) = \limsup_{k\to\infty} \Delta^2_{t_k} \delta_{\mathbb{R}^2_-}\big(\psi(\bar{x})|\bar{\lambda}\big)(w^k) + \frac{\|\bar{y}\|}{\|\bar{x}\|}(-v_0^2 + \|v_r\|^2). \tag{2.27}$$

Define further the sequence of vectors v^k by

$$v^k := \frac{x^k - \bar{x}}{t_k} \qquad \text{with} \qquad x^k := \bar{x} + \alpha_k v - \beta_k \widetilde{\bar{x}} \quad \text{and} \quad \beta_k = \frac{\alpha_k^2(-v_0^2 + \|v_r\|^2)}{4\bar{x}_0(\bar{x}_0 + \alpha_k v_0)}, \tag{2.28}$$

where $\alpha_k > 0$ is chosen—we will show in the *claim* below that such a number α_k does exist for each k—so that $\|x^k - \bar{x}\| = t_k$ and $x^k \in \operatorname{bd}(\mathcal{Q})$. It follows from construction (2.28) of $v^k = (v_0^k, v_r^k) \in \mathbb{R} \times \mathbb{R}^m$ that the vectors w^k defined in (2.19) admit the representations

$$w^k = \frac{\psi(\bar{x} + t_k v^k) - \psi(\bar{x})}{t_k} = \frac{1}{t_k}\Big((0, -\bar{x}_0 - t_k v_0^k,) - (0, -\bar{x}_0)\Big) = (0, -v_0^k,),$$

This tells us that $\langle \bar{\lambda}, w^k \rangle = \langle (\bar{\alpha}, 0), (0, -v_0^k,) \rangle = 0$ and implies in turn that

$$\Delta_{t_k}^2 \delta_{\mathbb{R}_-^2}\big(\psi(\bar{x})|\bar{\lambda}\big)(w^k) = \frac{\delta_{\mathbb{R}_-^2}\big(\psi(\bar{x} + t_k v^k)\big) - \delta_{\mathbb{R}_-^2}\big(\psi(\bar{x})\big) - t_k \langle \bar{\lambda}, w^k \rangle}{\frac{1}{2} t_k^2} = 0 \quad \text{for all} \ \ k \in \mathbb{N}.$$

It allows us to arrive at the equality

$$\limsup_{k \to \infty} \Delta_{t_k}^2 \delta_{\mathbb{R}_-^2}\big(\psi(\bar{x})|\bar{\lambda}\big)(w^k) = 0,$$

which together with (2.27) justifies the second twice epi-differentiability requirement (2.15).

Let us now verify the aforementioned claim formulated as follows.

Claim. *For any $v_0 \geq 0$ in Case 3(ii) and any $k \in \mathbb{N}$ there is $\alpha_k > 0$ satisfying (2.28) such that $x^k \in \operatorname{bd}(\mathcal{Q})$ and $\|x^k - \bar{x}\| = t_k$. If $v_0 < 0$ in this case, then we can select $\alpha_k \in (0, -\frac{\bar{x}_0}{v_0})$ as $k \in \mathbb{N}$ so that the above conditions on x^k from (2.28) are also satisfied.*

We prove this claim by arguing in parallel for both cases of $v_0 \geq 0$ and $v_0 < 0$. Pick $v_0 \geq 0$ (resp. $v_0 < 0$) satisfying (2.26) and observe that $\beta_k \geq 0$ when $\alpha_k > 0$ (resp. when $\alpha_k \in (0, -\frac{\bar{x}_0}{v_0})$) in (2.28). Employing $\bar{x}_0^2 = \|\bar{x}_r\|^2$ and $\bar{x}_0 v_0 = \langle \bar{x}_r, v_r \rangle$, we obtain by the direct calculation that the relationship

$$-\big((1+\beta_k)\bar{x}_0 + \alpha_k v_0\big)^2 + \|(1-\beta_k)\bar{x}_r + \alpha_k v_r\|^2 = 0$$

is valid in both cases and yields in turn the inequality

$$\|(1-\beta_k)\bar{x}_r + \alpha_k v_r\| = (1+\beta_k)\bar{x}_0 + \alpha_k v_0 > 0.$$

This confirms that if $v_0 \geq 0$ (resp. $v_0 < 0$), then for any $\alpha_k > 0$ (resp. $\alpha_k \in (0, -\frac{\bar{x}_0}{v_0})$) we have

$$x^k = \big((1+\beta_k)\bar{x}_0 + \alpha_k v_0, (1-\beta_k)\bar{x}_r + \alpha_k v_r\big) \in \mathrm{bd}\,(\mathcal{Q}).$$

Now it remains to show that for each $k \in \mathbb{N}$ there exists α_k from the intervals above such that $\|x^k - \bar{x}\| = t_k$. To proceed, consider the polynomial

$$p(\alpha) = \big((-v_0^2 + \|v_r\|^2)^2 + 16\bar{x}_0^2 v_0^2\big)\,\alpha^4 + 32\bar{x}_0^3 v_0 \alpha^3 + 16(\bar{x}_0^4 - t_k^2 \bar{x}_0^2 v_0^2)\alpha^2 - 32 t_k^2 \bar{x}_0^3 v_0 \alpha - 16 t_k^2 \bar{x}_0^4.$$

Since $p(0) = -16 t_k^2 \bar{x}_0^4 < 0$ and the leading coefficient of $p(\alpha)$ is positive, this polynomial has a positive zero, which we denote by α_k. It follows from

$$t_k^2 = \|x^k - \bar{x}\|^2 = \|\alpha_k v - \beta_k \widehat{\bar{x}}\|^2 = \alpha_k^2 + \beta_k^2 = \alpha_k^2 + \frac{\alpha_k^4(-v_0^2 + \|v_r\|^2)^2}{16\bar{x}_0^2(\bar{x}_0 + \alpha_k v_0)^2} \tag{2.29}$$

that any root $\alpha_k > 0$ satisfies all our requirements in (2.28) provided that $v_0 \geq 0$. If $v_0 < 0$, we need to show in addition that there is a root of $p(\alpha)$ belonging to the interval $(0, -\frac{\bar{x}_0}{v_0})$. But it is an immediate consequence of the conditions

$$p\big(-\frac{\bar{x}_0}{v_0}\big) = \frac{(-v_0^2 + \|v_r\|^2)^2 \bar{x}_0^4}{v_0^4} > 0 \qquad \text{and} \qquad p(0) = -16t_k^2\bar{x}_0^4 < 0,$$

which therefore finish the proof of this claim.

Let us finally show that $v^k \to v$ as $k \to \infty$. From (2.29) we get that $\alpha_k \to 0$ since $t_k \downarrow 0$ as $k \to \infty$. Remembering that $\|v^k\| = 1 = \|v\|$, it follows directly from (2.28) and (2.29) that

$$\|v^k - v\|^2 = 2 - 2\langle v^k, v\rangle = 2 - \frac{2\alpha_k}{t_k} = 2 - \frac{2\alpha_k}{\sqrt{\alpha_k^2 + \beta_k^2}} = 2 - \frac{2}{\sqrt{1 + \frac{\beta_k^2}{\alpha_k^2}}} \to 2 - 2 = 0$$

as $k \to \infty$, and hence $v^k \to v$. The the proof of the theorem is complete. □

Remark 2.2. (comparison with known results). Twice epi-differentiability of $\delta_{\mathcal{Q}}$ in Theorem 2.1 can be obtained by combining some known results about the second-order cone $\mathcal{Q}$. Indeed, it has been realized that the projection mapping $\Pi_{\mathcal{Q}}$ to the second-order cone is always directionally differentiable; see, e.g., [47, Lemma 2]. Thus we can conclude from [60, Corollary 13.43] that the indicator function $\delta_{\mathcal{Q}}$ is twice epi-differentiable at any $\bar{x} \in \mathcal{Q}$ for every $\bar{y} \in N_{\mathcal{Q}}(\bar{x})$. However, the established formula (2.13) for the second epi-derivative formula for $\delta_{\mathcal{Q}}$ cannot be obtained from the aforementioned arguments, and therefore is new to the best of our knowledge.

In the rest of this section we present some immediate consequences of Theorem 2.1

important in second-order variational analysis of SOCPs. The first one uses the established twice epi-differentiability of $\delta_{\mathcal{Q}}$ to verify a derivative-coderivative relationship for the normal cone to $\mathcal{Q}$.

Corollary 2.3. (derivative-coderivative relationship between the normal cone to $\mathcal{Q}$). *Let $\bar{x} \in \mathcal{Q}$ and $\bar{y} \in N_{\mathcal{Q}}(\bar{x})$. Then we have the inclusion*

$$(DN_{\mathcal{Q}})(\bar{x},\bar{y})(v) \subset (D^*N_{\mathcal{Q}})(\bar{x},\bar{y})(v) \qquad \textit{for all } \ v \in \mathbb{R}^{m+1}.$$

Proof. It follows from [60, Theorem 13.57] that the claimed inclusion holds for any convex set whose indicator function is twice epi-differentiable at the reference point. The latter is the case for the second-order cone $\mathcal{Q}$ due to Theorem 2.1. □

The next corollary provides a precise calculation for the graphical derivative (2.5) of the normal cone to $\mathcal{Q}$ that is significant for the subsequent material of the paper. The tangent cone to the graph of $N_{\mathcal{Q}}$ has been calculated before by using different approaches; see, e.g., [66, Lemma 6.6]. Based on such calculations, it is possible to compute the graphical derivative of $N_{\mathcal{Q}}$. Here we present another device that employs on the new second-order formula (2.13).

Corollary 2.4. (graphical derivative of the normal cone to $\mathcal{Q}$). *Let $\bar{x} \in \mathcal{Q}$ and $\bar{y} \in N_{\mathcal{Q}}(\bar{x})$. Then for all $v = (v_0, v_r) \in \mathbb{R} \times \mathbb{R}^m$ the graphical derivative of $N_{\mathcal{Q}}$ admits the representation*

$$(DN_{\mathcal{Q}})(\bar{x},\bar{y})(v) = \begin{cases} N_{\overline{\mathcal{K}}}(v) & \textit{if } \ \bar{x} \in [\mathrm{int}(\mathcal{Q}) \cup \{0\}], \\ \dfrac{\|\bar{y}\|}{\|\bar{x}\|}(-v_0, v_r) + N_{\overline{\mathcal{K}}}(v) & \textit{if } \ \bar{x} \in \mathrm{bd}\,(\mathcal{Q}) \setminus \{0\}, \end{cases}$$

where the critical cone $\overline{\mathcal{K}}$ is defined in (2.12).

Proof. It follows from [60, Theorem 13.40] and from the twice epi-differentiability of $\delta_{\mathcal{Q}}$ established in Theorem 2.1 that for all $v \in \mathbb{R}^{m+1}$ we have

$$(DN_{\mathcal{Q}})(\bar{x},\bar{y})(v) = \partial\Big(\frac{1}{2}d^2\delta_{\mathcal{Q}}(\bar{x}|\bar{y})\Big)(v),$$

where the subdifferential on the right-hand side is defined in (2.4). Combining this with the second epi-derivative formula from Theorem 2.1 verifies the claimed representation. □

Now we discuss relationships between the obtained results and a major condition introduced and employed in [41] for representing the graphical derivative of the normal cone mappings in conic programming under the nondegeneracy condition. Let us first recall this notion.

Definition 2.5. (projection derivative condition). *Given a closed set $\Omega \subset \mathbb{R}^n$, assume that the projection operator $\Pi_\Omega\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ admits the classical directional derivative $\Pi'_\Omega(x;h)$ at each $x \in \mathbb{R}^n$ in any direction h. We say that Ω satisfies the* PROJECTION DERIVATION CONDITION (*PDC*) *at $x \in \Omega$ if*

$$\Pi'_\Omega(x+y;h) = \Pi_{\mathcal{K}(x,y)}(h) \;\; \textit{whenever} \;\; y \in N_\Omega(x) \;\; \textit{and} \;\; h \in \mathbb{R}^n,$$

where $\mathcal{K}(x,y) := T_\Omega(x) \cap \{y\}^\perp$ signifies the critical cone of Ω at x for y.

It is proved in [41] that PDC is valid for any convex set Ω satisfying the extended polyhedrality condition from [6, Definition 3.52] (this includes convex polyhedra) and may also hold in nonpolyhedral settings. Furthermore, PDC holds at the vertex of any convex cone Ω. On the other hand, we show below that PDC *fails* at every nonzero boundary point of the nonpolyhedral second-order cone $\mathcal{Q}$ despite its second-order regularity [6] and other nice properties.

To proceed, we first present a useful characterization of PDC important for its own sake.

Proposition 2.6. (graphical derivative description of the projection derivation condition). *Let $\Omega\subset\mathbb{R}^n$ be a convex set. Then PDC holds at $\bar{x}\in\Omega$ if and only if*

$$(DN_\Omega)(\bar{x},\bar{y})(v)=N_{\mathcal{K}(\bar{x},\bar{y})}(v)\qquad \textit{for all}\ \ \bar{y}\in N_\Omega(\bar{x})\quad \textit{and}\ \ v\in\mathbb{R}^n. \tag{2.30}$$

Proof. Assuming that PDC holds at $\bar{x}$, take $\bar{y}\in N_\Omega(\bar{x})$ and $v\in\mathbb{R}^n$. To verify the inclusion "$\subset$" in (2.30), pick $w\in(DN_\Omega)(\bar{x},\bar{y})(v)$ and get by definition (2.5) that $(v,w)\in T_{\mathrm{gph}\,N_\Omega}(\bar{x},\bar{y})$. Then it follows from the projection representation in [60, Proposition 6.17] that

$$\Pi_\Omega(x)=\big(I+N_\Omega\big)^{-1}(x)\qquad \text{for any}\ \ x\in\mathbb{R}^n. \tag{2.31}$$

Employing elementary tangent cone calculus gives us the representation

$$\begin{aligned} T_{\mathrm{gph}\,N_\Omega}(\bar{x},\bar{y}) &= \big\{(v,w)\big|\ (v+w,v)\in T_{\mathrm{gph}\,\Pi_\Omega}(\bar{x}+\bar{y},\bar{y})\big\}\\ &= \big\{(v,w)\big|\ v=\Pi'_\Omega(\bar{x}+\bar{y};v+w)\big\}\ \ \text{whenever}\ \ \bar{y}\in N_\Omega(\bar{x}). \end{aligned} \tag{2.32}$$

The above relationships readily imply that

$$v=\Pi'_\Omega(\bar{x}+\bar{y};v+w)=\Pi_{\mathcal{K}(\bar{x},\bar{y})}(v+w)=\big(I+N_{\mathcal{K}(\bar{x},\bar{y})}\big)^{-1}(v+w).$$

This leads us in turn to $w \in N_{\mathcal{K}(\bar{x},\bar{y})}(v)$ and hence justifies the inclusion "$\subset$" in (2.30). The opposite inclusion can be verified similarly.

Conversely, suppose that equality (2.30) is satisfied. Pick $h \in \mathbb{R}^n$, $\bar{y} \in N_\Omega(\bar{x})$, and $v = \Pi'_\Omega(\bar{x}+\bar{y};h)$. Employing (2.32) tells us that $(v, h-v) \in T_{\text{gph}\,N_\Omega}(\bar{x},\bar{y})$, and hence we get $h - v \in N_{\mathcal{K}(\bar{x},\bar{y})}(v)$ due to (2.30). Combining the latter with (2.31) gives us $v = \Pi_{\mathcal{K}(\bar{x},\bar{y})}(h)$, which verifies PDC. □

Now we are ready to demonstrate the aforementioned failure of PDC for the second-order cone $\mathcal{Q} \subset \mathbb{R}^{m+1}$ with $m \geq 2$ on its entire boundary off the origin. If $m = 1$, then $\mathcal{Q}$ is a convex polyhedron, and hence it satisfies the PDC condition.

Corollary 2.7. (failure of PDC for the second-order cone at its nonzero boundary points). *Given* $\bar{x} \in \mathcal{Q} \subset \mathbb{R}^{m+1}$ *with* $m \geq 2$*, PDC fails whenever* $\bar{x} \in \text{bd}\,(\mathcal{Q}) \setminus \{0\}$.

Proof. Suppose on the contrary that PDC holds at some $\bar{x} \in \text{bd}\,(\mathcal{Q}) \setminus \{0\}$. Thus for every $\bar{y} \in N_\mathcal{Q}(\bar{x})$ condition (2.30) is satisfied. Pick $\bar{y} = (\bar{y}_0, \bar{y}_r) \in N_\mathcal{Q}(\bar{x}) = \{t\tilde{\bar{x}} \,|\, t \geq 0\}$ with $\bar{y} \neq 0$. It tells us that $\bar{y}_0 \neq 0$ and $\bar{y}_r \neq 0$. Employing the graphical derivative formula from Corollary 2.4 together with the PDC description in Proposition 2.6 as $\Omega = \mathcal{Q}$ and $\mathcal{K}(\bar{x},\bar{y}) = \overline{\mathcal{K}}$ shows that

$$N_{\overline{\mathcal{K}}}(v) = (DN_\mathcal{Q})(\bar{x},\bar{y})(v) = \frac{\|\bar{y}\|}{\|\bar{x}\|}(-v_0, v_r) + N_{\overline{\mathcal{K}}}(v) \qquad \text{for all } \; v = (v_0, v_r) \in \mathbb{R} \times \mathbb{R}^m.$$

Since $\bar{y} \neq 0$, we obtain $(-v_0, v_r) \in N_{\overline{\mathcal{K}}}(v) = \overline{\mathcal{K}}^* \cap \{v\}^\perp$ for all $v \in \overline{\mathcal{K}} = T_\mathcal{Q}(\bar{x}) \cap \{\bar{y}\}^\perp = \{v \in \mathbb{R}^{m+1} |\; \langle \bar{y}, v\rangle = 0\}$. It says, in particular, that for all $v \in \overline{\mathcal{K}}$ with $v = (v_0, v_r)$ we should have $(-v_0, v_r) \in \overline{\mathcal{K}}^* = \mathbb{R}\bar{y}$. Pick a vector $a \in \mathbb{R}^m$ with $a \neq 0$ and $\langle a, \bar{y}_r \rangle = 0$ (such a vector always exists by $m \geq 2$) and put $v := (0, a)$. It is clear that $v \in \overline{\mathcal{K}}$ while $v \notin \mathbb{R}\bar{y}$, which is a contradiction that justifies the claimed statement. □

CHAPTER 3 COMPUTATION OF GRAPHICAL DERIVATIVE OF THE NORMAL CONE MAPPING

3.1 Remarkable Properties of Second-Order Cone Constraints

In this section we derive new properties of the second-order cone $\mathcal{Q}$, which are important in what follows while being also of their own interest. The derivation of some of the results below employs those obtained in the previous section.

Our first result here provides a complete description of the set of Lagrange multipliers associated with stationary points of the constraint system Γ in (1.3). Given a stationary pair $(x, x^*) \in \operatorname{gph} N_\Gamma$, define the set of *Lagrange multipliers* associated with (x, x^*) by

$$\Lambda(x, x^*) := \big\{\lambda \in N_{\mathcal{Q}}\big(\Phi(x)\big) \,\big|\, \nabla\Phi(x)^*\lambda = x^*\big\} \tag{3.1}$$

and the *critical cone* to Γ at (x, x^*) by

$$K(x, x^*) := T_\Gamma(x) \cap \{x^*\}^\perp. \tag{3.2}$$

If $\Phi(\bar{x}) = 0$ for some $\bar{x}$ with $(\bar{x}, \bar{x}^*) \in \operatorname{gph} N_\Gamma$, then the Lagrange multiplier set reduces to

$$\Lambda(\bar{x}, \bar{x}^*) = \big\{\lambda \in -\mathcal{Q} \,\big|\, \nabla\Phi(\bar{x})^*\lambda = \bar{x}^*\big\}. \tag{3.3}$$

Following [6, Definition 4.74], we say that the *strict complementarity condition* holds for $\Lambda(\bar{x}, \bar{x}^*)$ from (3.3) if there is a multiplier $\lambda \in \text{int}(-\mathcal{Q})$ such that $\nabla\Phi(\bar{x})^*\lambda = \bar{x}^*$. The next result provides a precise description of the Lagrange multiplier set (3.3) that plays a significant role in our method of conducting the second-order analysis of Γ. A part of this analysis is inspired by the unpublished work of Shapiro and Nemirovski [62] about the “no duality gap" property in linear conic programs generated by convex cones; see, in particular, the proof of [62, Proposition 3] and the discussion after it.

Proposition 3.1. (description of Lagrange multipliers for the second-order cone). *Let* $(\bar{x},\bar{x}^*)\in\mathrm{gph}\,N_\Gamma$ *with* $\Phi(\bar{x})=0$*, and let* $\Lambda(\bar{x},\bar{x}^*)\neq\emptyset$ *for the set of Lagrange multipliers* (3.3). *Then one of the following alternatives holds for* $\Lambda(\bar{x},\bar{x}^*)$*:*

(LMS1) *The strict complementarity condition holds for* $\Lambda(\bar{x},\bar{x}^*)$ *from* (3.3). *In this case we get that for any* $\bar{\lambda}\in\Lambda(\bar{x},\bar{x}^*)$ *there are numbers* $\ell,\varepsilon>0$ *ensuring the error bound estimate*

$$\mathrm{dist}\big(\lambda;\Lambda(\bar{x},\bar{x}^*)\big)\le\ell\big(\mathrm{dist}(\lambda;-\mathcal{Q})+\|\nabla\Phi(\bar{x})^*\lambda-\bar{x}^*\|\big)\quad\textit{whenever}\quad\lambda\in\mathbb{B}_\varepsilon(\bar{\lambda}).\tag{3.4}$$

(LMS2) $\Lambda(\bar{x},\bar{x}^*)=\{\bar{\lambda}\}$ *for some multiplier* $\bar{\lambda}\in\mathrm{bd}\,(-\mathcal{Q})\setminus\{0\}$.

(LMS3) $\Lambda(\bar{x},\bar{x}^*)=\big\{t\bar{\lambda}\,\big|\,t\ge0\big\}$ *for some* $\bar{\lambda}\in\mathrm{bd}\,(-\mathcal{Q})$. *In this case we have* $\bar{x}^*=0$.

Proof. The validity of (3.4) in (LMS1) follows from [2, Corollary 5]. Suppose that the strict complementarity condition fails. If $\Lambda(\bar{x},\bar{x}^*)$ is a singleton, then either (LMS2) or (LMS3) with $\bar{\lambda}=0$ holds. Suppose now that $\Lambda(\bar{x},\bar{x}^*)$ is not a singleton and pick $\bar{\lambda}\in\Lambda(\bar{x},\bar{x}^*)$ such that $\bar{\lambda}\neq0$. We claim that $\Lambda(\bar{x},\bar{x}^*)\subset\mathbb{R}_+\bar{\lambda}$. Assuming the contrary allows us to find $0\neq\lambda\in\Lambda(\bar{x},\bar{x}^*)$ such that $\lambda\notin\mathbb{R}_+\bar{\lambda}$. Since the strict complementarity condition fails, we have $\bar{\lambda},\ \lambda\in\mathrm{bd}\,(-\mathcal{Q})\setminus\{0\}$. Define $\lambda_\alpha:=\alpha\bar{\lambda}+(1-\alpha)\lambda$ with $\alpha\in(0,1)$ and observe that $\lambda_\alpha\in$ int$(-\mathcal{Q})$; otherwise $\lambda\in\mathbb{R}_+\bar{\lambda}$. This observation amounts to saying that the strict complementarity condition holds for $\Lambda(\bar{x},\bar{x}^*)$, which is a contradiction. Thus we arrive at the inclusion $\Lambda(\bar{x},\bar{x}^*)\subset\mathbb{R}_+\bar{\lambda}$, which together with $\Lambda(\bar{x},\bar{x}^*)$ not being a singleton results in $0\in\Lambda(\bar{x},\bar{x}^*)$. It follows from the latter that $\Lambda(\bar{x},\bar{x}^*)=\mathbb{R}_+\bar{\lambda}$, telling us that (LMS3) is satisfied. Since $0 \in \Lambda(\bar{x}, \bar{x}^*)$ in case (LMS3), we get $\bar{x}^* = 0$ in this case and hence complete the proof of the proposition. $\triangle$

To proceed with our further analysis, we introduce an appropriate (very weak) constraint qualification for the second-order cone constraint system (1.3). This condition has been recently employed in the polyhedral framework of NLPs to conduct a second-order analysis of the classical equality and inequality constraint systems with $\mathcal{C}^2$-smooth data; see [7, 16, 19]. It has also been studied in [18] in nonpolyhedral settings via first-order and second-order constructions of variational analysis. However, to the best of our knowledge, it has never been implemented before for the second-order variational analysis of nonpolyhedral systems as we do in this paper.

Definition 3.2. (metric subregularity constraint qualification). *We say that system* (1.3) *satisfies the* METRIC SUBREGULARITY CONSTRAINT QUALIFICATION (*MSCQ*) *at* $\bar{x} \in \Gamma$ *with modulus* $\kappa > 0$ *if the mapping* $x \mapsto \Phi(x) - \mathcal{Q}$ *is metrically subregular at* $(\bar{x}, 0)$ *with modulus* κ.

Using (2.7) with the fixed vector $y = \bar{y} = 0$, observe that the introduced MSCQ with modulus κ for (1.3) can be equivalently described as the existence of a neighborhood U of $\bar{x}$ such that

$$\operatorname{dist}(x; \Gamma) \leq \kappa \operatorname{dist}\big(\Phi(x); \mathcal{Q}\big) \qquad \text{for all } \ x \in U. \tag{3.5}$$

Note that the defined MSCQ property of (1.3) is *robust* in the sense that its validity at $\bar{x} \in \Gamma$ yields this property at any $x \in \Gamma$ near $\bar{x}$. Furthermore, it is clear (Example 3.13 below) that the MSCQ from Definition 3.2 is strictly weaker than the qualification condition corresponding to the *metric regularity* of the mapping $x \mapsto \Phi(x) - \mathcal{Q}$ around $(\bar{x}, 0)$ therein.

The latter is well known to be equivalent to the *Robinson constraint qualification* (RCQ), which is the basic qualification condition in conic programming:

$$N_{\mathcal{Q}}\big(\Phi(\bar{x})\big) \cap \ker \nabla\Phi(\bar{x})^* = \{0\}. \tag{3.6}$$

An important role of MSCQ and its calmness equivalent for inverse mappings has been recognized in generalized differential calculus of variational analysis. In particular, it follows from [23, Theorem 4.1] and the convexity of $\mathcal{Q}$ that there is a neighborhood U of $\bar{x}$ such that

$$N_\Gamma(x) = \widehat{N}_\Gamma(x) = \nabla\Phi(x)^* N_{\mathcal{Q}}\big(\Phi(x)\big) \qquad \text{for all } \; x \in \Gamma \cap U, \tag{3.7}$$

where $\widehat{N}_\Omega(\bar{x})$ stands for the *regular/Fréchet normal cone* to Ω at $\bar{x} \in \Omega$ defined by

$$\widehat{N}_\Omega(\bar{x}) := \Big\{ v \in \mathbb{R}^n \Big|\ \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle v, x - \bar{x}\rangle}{\|x - \bar{x}\|} \le 0 \Big\},$$

which is *dual* to the tangent cone (2.1), i.e, $\widehat{N}_\Omega(\bar{x}) = T^*_\Omega(\bar{x})$. The first equality in (3.7) postulates the *normal regularity* of Γ at any point $x \in \Gamma$ near $\bar{x}$. Note also that the validity of MSCQ for Γ at $\bar{x} \in \Gamma$ ensures by [24, Proposition 1] the tangent cone calculus rule

$$T_\Gamma(x) = \big\{ v \in \mathbb{R}^n \big|\ \nabla\Phi(x)v \in T_{\mathcal{Q}}\big(\Phi(x)\big)\big\} \qquad \text{for all } \; x \in \Gamma \cap U. \tag{3.8}$$

To proceed further, recall that the second-order cone $\mathcal{Q}$ is *reducible* at its *nonzero boundary points* to a convex polyhedron in the sense of [6, Definition 3.135]; this was first shown in [5, Lemma 15]. In what follows we use a different reduction of $\mathcal{Q}$ via the mapping ψ

from (2.17) that allows us to simplify the subsequent calculations. Indeed, the alternative representation (2.18) of the second-order cone $\mathcal{Q}$ via the mapping ψ from (2.17) in the proof of Case 3 of Theorem 2.1 is instrumental to furnish the reduction of $\mathcal{Q}$ to $\mathbb{R}^2_-$ at its nonzero boundary points. Observe that the Jacobian matrix $\nabla\psi(x)$ has full rank and

$$\Gamma = \big\{x \in \mathbb{R}^n \,\big|\, (\psi \circ \Phi)(x) \in \mathbb{R}^2_-\big\} \qquad \text{whenever } \; \Phi(x) \in \mathrm{bd}\,(\mathcal{Q}) \setminus \{0\}. \tag{3.9}$$

By showing below that the metric subregularity of the mapping $x \mapsto \Phi(x) - \mathcal{Q}$ at nonzero boundary points yields the one for $x \mapsto (\psi \circ \Phi)(x) - \mathbb{R}^2_-$, we open the door to the usage in this case the results for convex polyhedra established in [19].

Lemma 3.3. (propagation of metric subregularity for nonzero boundary points of $\mathcal{Q}$). *Let $\bar{x} \in \Gamma$ be such that $\Phi(\bar{x}) \in \mathrm{bd}\,(\mathcal{Q}) \setminus \{0\}$. Then the metric subregularity of the mapping $x \mapsto \Phi(x) - \mathcal{Q}$ at $(\bar{x}, 0)$ ensures the one for $x \mapsto (\psi \circ \Phi)(x) - \mathbb{R}^2_-$ at $(\bar{x}, 0)$ with $\psi \colon \mathbb{R}^{m+1} \to \mathbb{R}^2$ taken from* (2.17).

Proof. To verify the lemma, we need to establish the existence of a positive number κ and a neighborhood V of $\bar{x}$ such that the metric estimate

$$\mathrm{dist}(x;\Gamma) \le \kappa\, \mathrm{dist}\big((\psi \circ \Phi)(x); \mathbb{R}^2_-\big) \qquad \text{for all } \; x \in V \tag{3.10}$$

holds. Let us first show that there are a constant $c > 0$ and a neighborhood U of $\bar{x}$ for which

$$\mathrm{dist}\big(\Phi(x); \mathcal{Q}\big) \le c\, \mathrm{dist}\big((\psi \circ \Phi)(x); \mathbb{R}^2_-\big) \qquad \text{for all } \; x \in U. \tag{3.11}$$

Indeed, employing the decomposition of $\Phi(x) = (\Phi_0(x), \Phi_r(x))$ together with the direct calculations tells us that

$$\operatorname{dist}\big(\Phi(x);\mathcal{Q}\big)=\begin{cases}0 & \text{if } \Phi(x)\in\mathcal{Q},\\ \|\Phi(x)\| & \text{if } \Phi(x)\in-\mathcal{Q},\\ \dfrac{\sqrt{2}}{2}\big(\|\Phi_r(x)\|-\Phi_0(x)\big) & \text{if } \Phi(x)\notin\mathcal{Q}\cup(-\mathcal{Q});\end{cases} \tag{3.12}$$

$$\operatorname{dist}\big((\psi\circ\Phi)(x);\mathbb{R}^2_-\big)=\begin{cases}0 & \text{if } \Phi(x)\in\mathcal{Q},\\ -\Phi_0(x) & \text{if } \Phi(x)\in-\mathcal{Q},\\ \|\Phi_r(x)\|^2-\Phi_0^2(x) & \text{if } \Phi(x)\notin\mathcal{Q}\cup(-\mathcal{Q}) \text{ and } \Phi_0(x)\ge 0,\\ \sqrt{(\|\Phi_r(x)\|^2-\Phi_0^2(x))^2+\Phi_0^2(x)} & \text{if } \Phi(x)\notin\mathcal{Q}\cup(-\mathcal{Q}) \text{ and } \Phi_0(x)<0.\end{cases}$$

It follows from $\bar{x}\in\Gamma$ and $\Phi_0(\bar{x})=\|\Phi_r(\bar{x})\|\neq 0$ that there exists a neighborhood U of $\bar{x}$ such that the inequality $\Phi_0(x)>\dfrac{1}{2}\Phi_0(\bar{x})$ holds whenever $x\in U$. Pick $x\in U$ and observe that the two cases may occur: either (a) $\Phi(x)\in\mathcal{Q}$ for which we have $\operatorname{dist}(\Phi(x);\mathcal{Q})=\operatorname{dist}((\psi\circ\Phi)(x);\mathbb{R}^2_-)=0$, and hence estimate (3.11) is clearly satisfied, or (b) $\Phi(x)\notin\mathcal{Q}$, which means that $\|\Phi_r(x)\|>\Phi_0(x)$. This yields

$$\begin{aligned}\operatorname{dist}\big((\psi\circ\Phi)(x);\mathbb{R}^2_-\big) &= \big(\|\Phi_r(x)\|-\Phi_0(x)\big)\big(\|\Phi_r(x)\|+\Phi_0(x)\big)\\ &\ge 2\sqrt{2}\,\Phi_0(x)\operatorname{dist}\big(\Phi(x);\mathcal{Q}\big)\\ &\ge \sqrt{2}\,\Phi_0(\bar{x})\operatorname{dist}\big(\Phi(x);\mathcal{Q}\big),\end{aligned}$$

which justifies estimate (3.11) with $c:=\big(\sqrt{2}\Phi_0(\bar{x})\big)^{-1}$. Combining this and estimate (3.5) leads us to (3.10) and thus completes the proof of the proposition. △

The next result is of its own interest while being important for calculating the graphical derivative of the normal cone mapping given in the next section.

Theorem 3.4. (normal cone to the critical cone of ice-cream constraint systems). *Let* $(\bar{x},\bar{x}^*)\in\operatorname{gph}N_\Gamma$ *and let MSCQ hold at* $\bar{x}\in\Gamma$. *Then for any* $\lambda\in\Lambda(\bar{x},\bar{x}^*)$ *and* $v\in K(\bar{x},\bar{x}^*)$ *the normal cone to the critical cone* $K(\bar{x},\bar{x}^*)$ *is represented by*

$$N_{K(\bar{x},\bar{x}^*)}(v)=\widehat{N}_{K(\bar{x},\bar{x}^*)}(v)=\nabla\Phi(\bar{x})^*\big[T_{N_{\mathcal{Q}}(\Phi(\bar{x}))}(\lambda)\cap\{\nabla\Phi(\bar{x})v\}^\perp\big]. \tag{3.13}$$

Proof. It follows from [51, Corollary 16.4.2], (3.7), and the normal-tangent duality that

$$\big(K(\bar{x},\bar{x}^*)\big)^*=\big(T_\Gamma(\bar{x})\cap\{\bar{x}^*\}^\perp\big)^*=\operatorname{cl}\big(N_\Gamma(\bar{x})+\mathbb{R}\bar{x}^*\big). \tag{3.14}$$

We proceed with verifying the following statement:

Claim. *If* $\Phi(\bar{x})\in\mathcal{Q}\setminus\{0\}$, *then*

$$\operatorname{cl}\left(N_\Gamma(\bar{x})+\mathbb{R}\bar{x}^*\right)=N_\Gamma(\bar{x})+\mathbb{R}\bar{x}^*. \tag{3.15}$$

Furthermore, (3.15) *is also valid if* $\Phi(\bar{x})=0$ *and if either* (LMS1) *or* (LMS3) *holds.*

To justify the claim, we split the arguments into the three cases depending on the position of the vector $\Phi(\bar{x})$ in the second-order cone $\mathcal{Q}$:

Case 1: $\Phi(\bar{x})\in\operatorname{int}\mathcal{Q}$. This gives us $\bar{x}^*=0$, which immediately yields (3.15).

Case 2: $\Phi(\bar{x})\in\operatorname{bd}\mathcal{Q}\setminus\{0\}$. Then the normal cone to Γ at $\bar{x}$ is a convex polyhedron. Using this together with [51, Corollary 19.3.2] ensures the validity of (3.15).

Case 3: $\Phi(\bar{x}) = 0$ and either (LMS1) or (LMS3) holds. If the strict complementarity condition in (LMS1) is satisfied, we have $\lambda \in \text{int}(-\mathcal{Q})$ such that $\nabla\Phi(\bar{x})^*\lambda = \bar{x}^*$, which shows together with (3.7) that

$$N_\Gamma(\bar{x}) + \mathbb{R}\bar{x}^* = \nabla\Phi(\bar{x})^* N_{\mathcal{Q}}\big(\Phi(\bar{x})\big) + \mathbb{R}\nabla\Phi(\bar{x})^*\lambda = \nabla\Phi(\bar{x})^*\big(-\mathcal{Q} + \mathbb{R}\lambda\big).$$

Pick $\eta \in \mathbb{R}^{m+1}$ and find $t > 0$ sufficiently small so that $\lambda + t\eta \in -\mathcal{Q}$. This leads us to

$$t\eta = \lambda + t\eta - \lambda \in -\mathcal{Q} + \mathbb{R}\lambda,$$

and therefore we get $\eta \in -\mathcal{Q} + \mathbb{R}\lambda$. It tells us that $-\mathcal{Q} + \mathbb{R}\lambda = \mathbb{R}^{m+1}$, which results in

$$N_\Gamma(\bar{x}) + \mathbb{R}\bar{x}^* = \nabla\Phi(\bar{x})^*\big(-\mathcal{Q} + \mathbb{R}\lambda\big) = \nabla\Phi(\bar{x})^*\mathbb{R}^{m+1}$$

and hence verifies (3.15) in this setting. To finish the proof of the claim, it remains to recall that under (LMS3) we have $\bar{x}^* = 0$, and thus (3.15) is satisfied.

To proceed with the proof of the theorem, we check first that (3.13) holds for all the cases in the above claim. Picking any $\lambda \in \Lambda(\bar{x}, \bar{x}^*)$ and $v \in K(\bar{x}, \bar{x}^*)$, deduce from (3.15) that

$$N_{K(\bar{x},\bar{x}^*)}(v) = \widehat{N}_{K(\bar{x},\bar{x}^*)}(v) = \big(K(\bar{x}, \bar{x}^*)\big)^* \cap \{v\}^\perp = (N_\Gamma(\bar{x}) + \mathbb{R}\bar{x}^*) \cap \{v\}^\perp. \tag{3.16}$$

For each $v^* \in N_{K(\bar{x},\bar{x}^*)}(v)$ we find by (3.7) and (3.16) some $\widetilde{\mu} \in N_Q(\bar{y})$ and $\alpha \in \mathbb{R}$ with

$$v^* = \nabla\Phi(\bar{x})^*\widetilde{\mu} + \alpha\bar{x}^* = \nabla\Phi(\bar{x})^*(\widetilde{\mu} + \alpha\lambda).$$

Letting $\mu := \widetilde{\mu} + \alpha\lambda$, we get $\lambda + \varepsilon\mu = (1 + \varepsilon\alpha)\lambda + \varepsilon\widetilde{\mu} \in N_Q(\bar{y})$ for any small $\varepsilon \geq 0$, which leads us to the inclusion $\mu \in T_{N_Q(\bar{y})}(\lambda)$. Taking it into account and using (3.16) give us $\langle \mu, \nabla\Phi(\bar{x})v \rangle = \langle v^*, v \rangle = 0$, and thus show that v^* belongs to the set on the right-hand side of (3.13).

To verify the opposite inclusion in (3.13), pick $\mu \in T_{N_Q(\bar{y})}(\lambda)$ with $\langle \mu, \nabla\Phi(\bar{x})v \rangle = 0$ and find sequences $t_k \downarrow 0$ and $\mu_k \to \mu$ with $\lambda + t_k\mu_k \in N_Q(\bar{y})$ for all $k \in \mathbb{N}$. It follows from (3.7) that

$$\nabla\Phi(\bar{x})^*(\lambda + t_k\mu_k) \in N_\Gamma(\bar{x}) = \big(T_\Gamma(\bar{x})\big)^*.$$

Using this, for any $w \in K(\bar{x}, \bar{x}^*)$ we get

$$t_k\langle \mu_k, \nabla\Phi(\bar{x})w \rangle = \langle \bar{x}^*, w \rangle + t_k\langle \mu_k, \nabla\Phi(\bar{x})w \rangle = \langle \lambda + t_k\mu_k, \nabla\Phi(\bar{x})w \rangle \leq 0.$$

The passage to the limit as $k \to \infty$ gives us the relationships

$$\langle \nabla\Phi(\bar{x})^*\mu, w \rangle = \langle \mu, \nabla\Phi(\bar{x})w \rangle \leq 0,$$

which imply that $\nabla\Phi(\bar{x})^*\mu \in \big(K(\bar{x}, \bar{x}^*)\big)^*$. Combining it with (3.16) and $\langle \mu, \nabla\Phi(\bar{x})v \rangle = 0$ leads us to $\nabla\Phi(\bar{x})^*\mu \in \widehat{N}_{K(\bar{x},\bar{x}^*)}(v)$, and thus justifies the inclusion "$\supset$ " in (3.13) and the equality therein under the assumptions of the above claim.

Continuing the proof of the theorem, we need to justify (3.13) in the setting where $\Phi(\bar{x}) = 0$ and (LMS2) hold. Since $\Lambda(\bar{x}, \bar{x}^*) = \{\bar{\lambda}\}$ with $\bar{\lambda} = (\bar{\lambda}_0, \bar{\lambda}_r) \in \mathrm{bd}\,(-\mathcal{Q}) \setminus \{0\}$ in this case, and since MSCQ is satisfied at $\bar{x}$, we have by using (3.8) that

$$\begin{aligned} K(\bar{x}, \bar{x}^*) &= T_\Gamma(\bar{x}) \cap \{\bar{x}^*\}^\perp = \big\{v \in \mathbb{R}^n \;\big|\; \nabla\Phi(\bar{x})v \in \mathcal{Q} \;\text{ and }\; \langle v, \nabla\Phi(\bar{x})^*\bar{\lambda}\rangle = 0\big\} \\ &= \big\{v \in \mathbb{R}^n \;\big|\; \nabla\Phi(\bar{x})v \in \mathcal{Q} \;\text{ and }\; \langle \nabla\Phi(\bar{x})v, \bar{\lambda}\rangle = 0\big\} \\ &= \big\{v \in \mathbb{R}^n \;\big|\; \nabla\Phi(\bar{x})v \in \mathcal{Q} \cap \{\bar{\lambda}\}^\perp\big\} = \big\{v \in \mathbb{R}^n \;\big|\; \nabla\Phi(\bar{x})v \in \mathbb{R}_+\widehat{\bar{\lambda}}\big\}, \end{aligned}$$

where $\widehat{\bar{\lambda}} = (-\bar{\lambda}_0, \bar{\lambda}_r)$. Pick now $v \in K(\bar{x}, \bar{x}^*)$ and observe that

$$\begin{aligned} N_{K(\bar{x},\bar{x}^*)}(v) &= \nabla\Phi(\bar{x})^* N_{\mathbb{R}_+\widehat{\bar{\lambda}}}\big(\nabla\Phi(\bar{x})v\big) = \nabla\Phi(\bar{x})^*\big[\big(\mathbb{R}_+\widehat{\bar{\lambda}}\big)^* \cap \{\nabla\Phi(\bar{x})v\}^\perp\big] \\ &= \nabla\Phi(\bar{x})^*\big[T_{-\mathcal{Q}}(\bar{\lambda}) \cap \{\nabla\Phi(\bar{x})v\}^\perp\big] = \nabla\Phi(\bar{x})^*\big[T_{N_{\mathcal{Q}}(\bar{z})}(\bar{\lambda}) \cap \{\nabla\Phi(\bar{x})v\}^\perp\big], \end{aligned}$$

where the first equality (chain rule) holds by Robinson's seminal result from [50] since $\mathbb{R}_+\widehat{\bar{\lambda}}$ is a convex polyhedron and the constraint mapping $\nabla\Phi(\bar{x})v$ is linear. This justifies (3.13) in the case under consideration and thus completes the proof of the theorem. $\triangle$

A similar result to Theorem 3.4 was established in [19, Lemma 1] for polyhedral constraint systems with equality and inequality constraints coming from problems of nonlinear programming. The nonpolyhedral nature of the second-order cone $\mathcal{Q}$ creates significant difficulties in comparison with the polyhedral NLP structure that are successfully overcome in the proof above.

Now we present the main result of this section giving a characterization of the simultaneous fulfillment of the *uniqueness* of Lagrange multipliers associated with stationary points of (1.3) and a certain *error bound* estimate, which is automatic for polyhedral systems. Both properties are algorithmically important; see, e.g., the book [29] that strongly employs the uniqueness of Lagrange multipliers in polyhedral NLP systems and its characterization via the *strict Mangasarian-Fromovitz constraint qualification condition* (SMFCQ) for Newton-type methods.

While dealing with the set Γ in the next theorem, the only point $\bar{x}$ that needs to be taken care of is the one for which $\Phi(\bar{x}) = 0$. This comes from the observation made right before Lemma 3.3 on the reducibility of $\mathcal{Q}$ at its nonzero boundary points to the convex polyhedron $\mathbb{R}^2_-$.

Theorem 3.5. (characterization of uniqueness of Lagrange multipliers with error bound estimate for second-order cone constraints). *Let* $(\bar{x}, \bar{x}^*) \in \operatorname{gph} N_\Gamma$*, and let* $\bar{\lambda} \in \Lambda(\bar{x}, \bar{x}^*)$ *with* $\Phi(\bar{x}) = 0$*. Then the following statements are equivalent:*

(i) $\bar{\lambda}$ *is a unique multiplier, and for some* $\ell > 0$ *the error bound estimate holds:*

$$\operatorname{dist}(\lambda; \Lambda(\bar{x}, \bar{x}^*)) \le \ell \, \|\nabla\Phi(\bar{x})^*\lambda - \bar{x}^*\| \qquad \text{for all} \quad \lambda \in -\mathcal{Q}. \tag{3.17}$$

(ii) *The dual qualification condition is satisfied:*

$$(DN_{\mathcal{Q}})\big(\Phi(\bar{x}), \bar{\lambda}\big)(0) \cap \ker \nabla\Phi(\bar{x})^* = \{0\}. \tag{3.18}$$

If in this case $\bar{\lambda} \in \operatorname{bd}(-\mathcal{Q}) \setminus \{0\}$*, then* (3.18) *implies that the matrix* $\nabla\Phi(\bar{x})$ *has full rank.*

(iii) *The strict Robinson constraint qualification holds:*

$$\nabla\Phi(\bar{x})\mathbb{R}^n - T_{\mathcal{Q}}\big(\Phi(\bar{x})\big) \cap \{\bar{\lambda}\}^\perp = \mathbb{R}^{m+1}. \tag{3.19}$$

Proof. Assume that (ii) is satisfied and pick any $\lambda \in \Lambda(\bar{x}, \bar{x}^*)$. We first show that $\lambda = \bar{\lambda}$, which verifies the uniqueness of Lagrange multipliers. It readily follows from (3.3) that

$$\lambda - \bar{\lambda} \in \ker \nabla\Phi(\bar{x})^* \quad \text{and} \quad \lambda - \bar{\lambda} \in -\mathcal{Q} + \mathbb{R}\bar{\lambda}. \tag{3.20}$$

Then Corollary 2.4 tells us that $(DN_{\mathcal{Q}})(\bar{z}, \bar{\lambda})(0) = N_{\overline{\mathcal{K}}}(0) = \overline{\mathcal{K}}^*$ with $\overline{\mathcal{K}} = T_{\mathcal{Q}}(\Phi(\bar{x})) \cap \{\bar{\lambda}\}^\perp = \mathcal{Q} \cap \{\bar{\lambda}\}^\perp$. Therefore we arrive at the relationships

$$\lambda - \bar{\lambda} \in -\mathcal{Q} + \mathbb{R}\bar{\lambda} \subset \big(\mathcal{Q} \cap \{\bar{\lambda}\}^\perp\big)^* = (DN_{\mathcal{Q}})(\Phi(\bar{x}), \bar{\lambda})(0). \tag{3.21}$$

Using them together with (3.18) and the first inclusion in (3.20), we get $\lambda = \bar{\lambda}$.

To verify now the error bound (3.17) in (i), we use $\Lambda(\bar{x}, \bar{x}^*) = \{\bar{\lambda}\}$ and arguing by contradiction. So for any $k \in \mathbb{N}$ there is $\lambda_k \in -\mathcal{Q}$ satisfying the conditions

$$\|\lambda_k - \bar{\lambda}\| > k\|\nabla\Phi(\bar{x})^*\lambda_k - \bar{x}^*\| = k\|\nabla\Phi(\bar{x})^*(\lambda_k - \bar{\lambda})\|.$$

Assume without loss of generality that $\frac{\lambda_k - \bar{\lambda}}{\|\lambda_k - \bar{\lambda}\|} \to \eta$ as $k \to \infty$ with $\|\eta\| = 1$. Thus passing to the limit in the above inequality brings us to

$$\nabla\Phi(\bar{x})^*\eta = 0. \tag{3.22}$$

On the other hand, we have the inclusions

$$\frac{\lambda_k - \bar{\lambda}}{\|\lambda_k - \bar{\lambda}\|} \in -\mathcal{Q} + \mathbb{R}\bar{\lambda} \subset \big(\mathcal{Q} \cap \{\bar{\lambda}\}^{\perp}\big)^*,$$

which together with (3.21) ensure the relationships

$$\eta \in \big(\mathcal{Q} \cap \{\bar{\lambda}\}^{\perp}\big)^* = (DN_{\mathcal{Q}})\big(\Phi(\bar{x}), \bar{\lambda}\big)(0).$$

Combining the latter with (3.22) and taking into account (ii) lead us to $\eta = 0$, which contradicts the fact that $\|\eta\| = 1$ and thus justifies the error bound estimate (3.17) in (i).

To verify next the converse implication (i) $\Longrightarrow$ (ii), take $\eta \in (DN_{\mathcal{Q}})(\Phi(\bar{x}), \bar{\lambda})(0) \cap \ker \nabla\Phi(\bar{x})^*$ and get by the definition of the graphical derivative that $(0, \eta) \in T_{\mathrm{gph}\, N_{\mathcal{Q}}}(\Phi(\bar{x}), \bar{\lambda})$. This allows us to find sequences $t_k \downarrow 0$ and $(v^k, \eta^k) \to (0, \eta)$ as $k \to \infty$ such that $(\Phi(\bar{x}), \bar{\lambda}) + t_k(v^k, \eta^k) \in \mathrm{gph}\, N_{\mathcal{Q}}$ and therefore $\bar{\lambda} + t_k\eta^k \in N_{\mathcal{Q}}(\Phi(\bar{x}) + t_k v^k) \subset -\mathcal{Q}$. Employing estimate (3.17) brings us to

$$\|\bar{\lambda} + t_k\eta^k - \bar{\lambda}\| = \mathrm{dist}\big(\bar{\lambda} + t_k\eta^k; \Lambda(\bar{x}, \bar{x}^*)\big) \le \ell\|\nabla\Phi(\bar{x})^*(\bar{\lambda} + t_k\eta^k) - \bar{x}^*\|,$$

which implies in turn that $\|\eta^k\| \le \ell\|\nabla\Phi(\bar{x})^*\eta^k\|$. Passing to the limit as $k \to \infty$ tells us that $\|\eta\| \le \ell\|\nabla\Phi(\bar{x})^*\eta\|$. By $\eta \in \ker \nabla\Phi(\bar{x})^*$ we get $\eta = 0$ and thus arrive at (3.18).

To finish the proof of (ii), suppose that $\bar{\lambda} = (\bar{\lambda}_0, \bar{\lambda}_r) \in \mathrm{bd}\,(-\mathcal{Q}) \setminus \{0\}$ and conclude from the graphical derivative formula in Corollary 2.4 that

$$(DN_{\mathcal{Q}})\big(\Phi(\bar{x}), \bar{\lambda}\big)(0) = \big(\mathcal{Q} \cap \{\bar{\lambda}\}^{\perp}\big)^* = \big(\mathbb{R}_+\widehat{\bar{\lambda}}\big)^* = \big\{(w_0, w_m) \in \mathbb{R} \times \mathbb{R}^m \big|\, \langle w_r, \bar{\lambda}_r\rangle - w_0\bar{\lambda}_0 \le 0\big\}.$$

It gives us by (3.18) that $\ker\nabla\Phi(\bar{x})^* = \{0\}$, and thus the matrix $\nabla\Phi(\bar{x})$ is of full rank.

To complete the proof of the theorem, it remains to show that the qualification conditions (3.18) and (3.19) are equivalent for the case of (1.3). Indeed, it follows from (3.19) that

$$\big(T_{\mathcal{Q}}(\Phi(\bar{x}))\cap\{\bar{\lambda}\}^{\perp}\big)^*\cap\ker\nabla\Phi(\bar{x})^* = \big(T_{\mathcal{Q}}(\Phi(\bar{x}))\cap\{\bar{\lambda}\}^{\perp}-\nabla\Phi(\bar{x})\mathbb{R}^n\big)^* = \{0\},$$

and hence the dual qualification condition (3.18) holds by Corollary 2.4. To verify the converse implication, we deduce from (3.18) that

$$\mathrm{cl}\big(\nabla\Phi(\bar{x})\mathbb{R}^n - T_{\mathcal{Q}}(\Phi(\bar{x}))\cap\{\bar{\lambda}\}^{\perp}\big) = \mathbb{R}^{m+1}.$$

Since $\nabla\Phi(\bar{x})\mathbb{R}^n - T_{\mathcal{Q}}(\Phi(\bar{x}))\cap\{\bar{\lambda}\}^{\perp}$ is convex, it has nonempty relative interior. Hence it follows from [60, Proposition 2.40] that the relationships

$$\begin{aligned}\mathbb{R}^{m+1} = \mathbf{ri}\,(\mathbb{R}^{m+1}) &= \mathbf{ri}\,\big[\mathrm{cl}\big(\nabla\Phi(\bar{x})\mathbb{R}^n - T_{\mathcal{Q}}\big(\Phi(\bar{x})\big)\cap\{\bar{\lambda}\}^{\perp}\big)\big]\\ &= \mathbf{ri}\,\big(\nabla\Phi(\bar{x})\mathbb{R}^n - T_{\mathcal{Q}}\big(\Phi(\bar{x})\big)\cap\{\bar{\lambda}\}^{\perp}\big)\\ &\subset \big(\nabla\Phi(\bar{x})\mathbb{R}^n - T_{\mathcal{Q}}\big(\Phi(\bar{x})\big)\cap\{\bar{\lambda}\}^{\perp}\big)\end{aligned}$$

are satisfied. This justifies (3.19) and thus ends the proof of the theorem. △

Remark 3.6. (discussions on constraint qualifications for second-order cone systems).

(i) Condition (3.19) was introduced in [6] as “strict constraint qualification" in conic programming and then was called “strict Robinson constraint qualification" (SRCQ) in [9].

In the case of NLPs this condition reduces to the strict Mangasarian-Fromovitz constraint qualification (SMFCQ) discussed before the formulation of Theorem 3.5. But in contrast to NLPs, where SMFCQ is well known as a characterization of the uniqueness of Lagrange multipliers, it is not the case for nonpolyhedral conic programs (including SOCPs), where SRCQ fails to be a characterization of this property; cf. [6, Propositions 4.47 and 4.50]. As proved in Theorem 3.5, SRCQ characterizes the uniqueness of Lagrange multipliers for the second-order cone constraint system (1.3) *along with* the *error bound* estimate (3.17), which is automatic for polyhedral systems as in NLPs due to the classical Hoffman lemma. Observe that, while being equivalent to SRCQ in the framework under consideration, the obtained form of *dual qualification condition* (3.18) seems to be new in conic programming.

(ii) It is worth highlighting the result of Theorem 3.5(ii) showing that the dual qualification condition (3.18) yields the full rank of $\nabla\Phi(\bar{x})$ in (1.3) if $\bar{\lambda} \in \mathrm{bd}\,(-\mathcal{Q}) \setminus \{0\}$. This is *not the case* for NLP constraint systems while reflecting the “fattiness" of the second-order cone $\mathcal{Q}$.

(iii) Note that the equivalence between (3.18) and (3.19) holds true if we replace $\mathcal{Q}$ with any closed *convex* sets that is $\mathcal{C}^2$*-cone reducible* in the sense of [6, Definition 1.135]. This can be shown by observing that the left-hand side of (3.19) is convex in this case, and therefore it has a nonempty relative interior in finite dimensions; cf. the proof of [6, Proposition 2.97]. Note also that Theorem 3.5 can be extended to any $\mathcal{C}^2$-cone reducible with the corresponding modifications of the error bound estimate (3.17). It is beyond the scope of this paper to provide a proof for such a general framework, and thus we postpone it to our future publications.

To proceed further, define the mapping $\mathcal{H}\colon \mathbb{R}^n \times \mathbb{R}^{m+1} \to \mathbb{R}^{n\times n}$ by

$$\mathcal{H}(x;\lambda) := \begin{cases} -\dfrac{\lambda_0}{\Phi_0(x)}\nabla\widehat{\Phi}(x)^*\nabla\Phi(x) & \text{if } \Phi(x) = \big(\Phi_0(x), \Phi_r(x)\big) \in \mathrm{bd}\,(\mathcal{Q}) \setminus \{0\}, \\ 0 & \text{otherwise}, \end{cases} \tag{3.23}$$

where $x \in \Gamma$, $\lambda = (\lambda_0, \lambda_r) \in \mathbb{R} \times \mathbb{R}^m$, and $\nabla\widehat{\Phi}(x) = (-\nabla\Phi_0(x), \nabla\Phi_r(x))$. This form is a simplification of the one used in [5], reflects a nonzero *curvature* of the second-order cone $\mathcal{Q}$ at boundary points, and thus is not needed for polyhedra. Recall that $\nabla\Phi(x)$ is an $(m+1)\times n$ matrix and hence $\nabla\widehat{\Phi}(x)^*\nabla\Phi(x)$ is an $n \times n$ matrix in (3.23).

In our derivation of the formula for calculating the graphical derivative of the normal cone mapping N_Γ in Section 5, we appeal to the *linear conic optimization problem*

$$\min_{\lambda\in\mathbb{R}^{m+1}} \big\{ -\big\langle v, \big(\nabla^2\langle\lambda,\Phi\rangle(\bar{x}) + \mathcal{H}(x;\lambda)\big)v\big\rangle \;\big|\; \nabla\Phi(\bar{x})^*\lambda = \bar{x}^* \;\text{ and }\; \lambda \in N_{\mathcal{Q}}\big(\Phi(\bar{x})\big)\big\} \tag{3.24}$$

generated by the second-order cone $\mathcal{Q}$, where $(\bar{x}, \bar{x}^*) \in \mathrm{gph}\, N_\Gamma$ and $v \in K(\bar{x}, \bar{x}^*)$. Denote by $\Lambda(\bar{x}, \bar{x}^*; v)$ the set of optimal solutions to (3.24). The following result shows that if the primal problem (3.24) has an optimal solution, then its dual problem has an *approximate* feasible solution for which the optimal values of the primal and dual problems are “almost the same." This is one of the *principal differences* between the polyhedral case with the exact duality therein and the nonpolyhedral ice-cream setting. The duality result obtained below is known in case (LMS1) of Proposition 3.1 (actually in this setting we have the exact duality; see, e.g., [61, Theorem 4.14]), but even in this case our proof is new.

Theorem 3.7. (approximate duality in linear second-order cone optimization). *Taking* $(\bar{x},\bar{x}^*)\in\operatorname{gph}N_\Gamma$ *and* $v\in K(\bar{x},\bar{x}^*)$*, suppose that* $\Lambda(\bar{x},\bar{x}^*)\neq\emptyset$ *and* $\Phi(\bar{x})=0$*. Then for every* $\widetilde{\lambda}\in\Lambda(\bar{x},\bar{x}^*;v)$ *and any small* $\varepsilon>0$ *there exists* $z_\varepsilon\in\mathbb{R}^n$ *for which we have the relationships*

$$\operatorname{dist}\big(\nabla\Phi(\bar{x})z_\varepsilon+\langle v,\nabla^2\Phi(\bar{x})v\rangle;\mathcal{Q}\big)\le\varepsilon\quad\textit{and}\quad\langle\bar{x}^*,z_\varepsilon\rangle+\big\langle v,\nabla^2\langle\widetilde{\lambda},\Phi\rangle(\bar{x})v\big\rangle\ge-\varepsilon.\tag{3.25}$$

Proof. It follows from (3.23) that under $\Phi(\bar{x})=0$ the optimization problem (3.24) reduces to

$$\min_{\lambda\in\mathbb{R}^{m+1}}\big\{-\big\langle v,\nabla^2\langle\lambda,\Phi\rangle(\bar{x})v\big\rangle\,\big|\,\nabla\Phi(\bar{x})^*\lambda=\bar{x}^*\ \text{ and }\ \lambda\in-\mathcal{Q}\big\}.\tag{3.26}$$

The *dual problem* of (3.26) can be calculated via [6, page 125] and [60, Example 11.41] as

$$\max_{z\in\mathbb{R}^n}\big\{\langle\bar{x}^*,z\rangle\big|\ \nabla\Phi(\bar{x})z+\big\langle v,\nabla^2\Phi(\bar{x})v\big\rangle\in T_{\mathcal{Q}}\big(\Phi(\bar{x})\big)\big\}.\tag{3.27}$$

Employing Proposition 3.1, we examine all the three possible cases for the set of Lagrange multipliers $\Lambda(\bar{x},\bar{x}^*)$. Picking any $v\in K(\bar{x},\bar{x}^*)$ and $\varepsilon>0$ sufficiently small, consider first case (LMS1) in Proposition 3.1 and use the error bound estimate (3.4). This estimate allows us to use the intersection rule from [26, Proposition 3.2] for the normal cone to $\Lambda(\bar{x},\bar{x}^*)$ and thus to deduce for any $\widetilde{\lambda}\in\Lambda(\bar{x},\bar{x}^*;v)$ that

$$0\in-\langle v,\nabla^2\Phi(\bar{x})v\rangle+N_{\Lambda(\bar{x},\bar{x}^*)}(\widetilde{\lambda})\subset-\langle v,\nabla^2\Phi(\bar{x})v\rangle+N_{-\mathcal{Q}}(\widetilde{\lambda})+\operatorname{rge}\nabla\Phi(\bar{x}).$$

This allows us to find some $z \in \mathbb{R}^n$ for which we get

$$\nabla\Phi(\bar{x})z + \langle v, \nabla^2\Phi(\bar{x})v\rangle \in N_{-\mathcal{Q}}(\widetilde{\lambda}) \subset \mathcal{Q} = T_{\mathcal{Q}}\big(\Phi(\bar{x})\big).$$

Since $-\mathcal{Q}$ is a convex cone, this inclusion leads us to $\big\langle\widetilde{\lambda}, \nabla\Phi(\bar{x})z + \langle v, \nabla^2\Phi(\bar{x})v\big\rangle = 0$. Hence

$$\langle\bar{x}^*, z\rangle = \langle\widetilde{\lambda}, \nabla\Phi(\bar{x})z\rangle = -\big\langle v, \nabla^2\langle\widetilde{\lambda}, \Phi\rangle(\bar{x})v\big\rangle,$$

which in turns implies that z is an optimal solution for the dual problem (3.27) and that the optimal values of the primal and dual problems agree. Letting $z_\varepsilon := z$ justifies the validity of both relationships in (3.25) in case (LMS1).

In case (LMS2) of Proposition 3.1, the set of Lagrange multipliers is a singleton and so is bounded. Using [60, Proposition 11.39] tells us that the optimal values of the primal problem (3.26) and the dual problem (3.27) agree. Therefore we arrive at

$$\sup_{z\in\mathbb{R}^n}\big\{\langle\bar{x}^*, z\rangle\big|\ \nabla\Phi(\bar{x})z + \big\langle v, \nabla^2\Phi(\bar{x})v\big\rangle \in T_{\mathcal{Q}}\big(\Phi(\bar{x})\big)\big\} = -\,\big\langle v, \nabla^2\langle\lambda, \Phi\rangle(\bar{x})v\big\rangle$$

that allows us for any $\varepsilon > 0$ to find z_ε satisfying the second condition in (3.25) together with

$$\nabla\Phi(\bar{x})z_\varepsilon + \big\langle v, \nabla^2\Phi(\bar{x})v\big\rangle \in T_{\mathcal{Q}}\big(\Phi(\bar{x})\big) = \mathcal{Q}.$$

Thus z_ε satisfies the first condition in (3.25) as well, which completes the proof in case (LMS2).

Consider finally case (LMS3) in Proposition 3.1 where there is $\bar{\lambda} \in \operatorname{bd}(-\mathcal{Q})$ such that

$$\Lambda(\bar{x}, \bar{x}^*) = \ker \nabla\Phi(\bar{x})^* \cap (-\mathcal{Q}) = \big\{t\bar{\lambda} \,\big|\, t \geq 0\big\}.$$

In this case the primal problem (3.26) can be equivalently written as

$$\min_{\lambda \in \mathbb{R}^{m+1}} \big\{-\big\langle v, \nabla^2\langle \lambda, \Phi\rangle(\bar{x})v\big\rangle \;\big|\; \lambda = \alpha\bar{\lambda},\ \alpha \geq 0\big\}. \tag{3.28}$$

Since $\Lambda(\bar{x}, \bar{x}^*; v) \neq \emptyset$, we arrive at $\big\langle v,\ \nabla^2\langle\bar{\lambda},\ \Phi\rangle(\bar{x})v\big\rangle \leq 0$. Examine the two possible situations:

(1) $\big\langle v, \nabla^2\langle\bar{\lambda}, \Phi\rangle(\bar{x})v\big\rangle < 0$. In this setting problem (3.28) has a unique optimal solution $\lambda = 0$. Using the arguments similar to the case (LMS2) and applying again [60, Proposition 11.39], we can find some z_ε satisfying both relationships in (3.25).

(2) $\big\langle v, \nabla^2\langle\bar{\lambda}, \Phi\rangle(\bar{x})v\big\rangle = 0$. In this setting the set of optimal solutions to problem (3.28) is the entire ray $\{t\bar{\lambda} \,|\, t \geq 0\}$. Consider now a modified version of (3.26) defined by

$$\min_{\lambda = (\lambda_0, \lambda_r) \in \mathbb{R}\times\mathbb{R}^m} \big\{-\big\langle v, \nabla^2\langle\lambda, \Phi\rangle(\bar{x})v\big\rangle \;\big|\; \nabla\Phi(\bar{x})^*\lambda = 0,\ \lambda \in -Q,\ -\lambda_0 \leq 1\big\}. \tag{3.29}$$

Since $\lambda \in -\mathcal{Q}$, we get $\|\lambda_r\| \leq -\lambda_0$. This implies that the feasible region of problem (3.29) is nonempty and bounded, and so is the set of its optimal solutions. Moreover, its optimal value is zero due to $\big\langle v, \nabla^2\langle\bar{\lambda}, \Phi\rangle(\bar{x})v\big\rangle = 0$. It follows from [60, Theorem 11.39(a)] that the optimal value of the dual problem of (3.29) given by

$$\max_{(z,\alpha) \in \mathbb{R}^n \times \mathbb{R}} \big\{\langle 0, z\rangle - \alpha \big|\ \nabla\Phi(\bar{x})z + (\alpha, 0, \ldots, 0) + \big\langle v, \nabla^2\Phi(\bar{x})v\big\rangle \in \mathcal{Q},\ \alpha \geq 0\big\} \tag{3.30}$$

is zero as well. Thus we arrive at the equality

$$\sup_{(z,\alpha)\in\mathbb{R}^n\times\mathbb{R}} \left\{-\alpha \,\middle|\, \nabla\Phi(\bar{x})z + (\alpha,0,\ldots,0) + \left\langle v, \nabla^2\Phi(\bar{x})v\right\rangle \in \mathcal{Q},\ \alpha \geq 0\right\} = 0.$$

This tells us that for any $\varepsilon > 0$ there exists a feasible solution $(z_\varepsilon, \alpha_\varepsilon) \in \mathbb{R}^n \times \mathbb{R}$ to (3.30) such that $-\alpha_\varepsilon > -\varepsilon$. Therefore we have the estimates

$$\text{dist}\left(\nabla\Phi(\bar{x})z_\varepsilon + \langle v, \nabla^2\Phi(\bar{x})v\rangle; \mathcal{Q}\right) \leq \|(\alpha_\varepsilon, 0, \ldots, 0)\| = \alpha_\varepsilon < \varepsilon,$$

which verify the first condition in (3.25). Since $\bar{x}^* = 0$ and $\left\langle v, \nabla^2\langle\bar{\lambda}, \Phi\rangle(\bar{x})v\right\rangle = 0$, we get the second condition in (3.25) and thus complete the proof of the theorem. △

We conclude this section by deriving a second-order sufficient condition for strict local minima in SOCPs needed in what follows. Consider the problem

$$\min\ \varphi_0(x) \qquad \text{subject to} \quad x \in \Gamma, \tag{3.31}$$

where $\varphi_0\colon \mathbb{R}^n \to \mathbb{R}$ is twice differentiable, and where Γ is taken from (1.3). Such a second-order sufficient condition was established in [6, Theorem 3.86] under the validity of the Robinson constraint qualification (3.6) that is equivalent to the metric regularity of the mapping $x \mapsto \Phi(x) - \mathcal{Q}$. It occurs that the same result holds under weaker assumptions on the latter mapping including the validity of MSCQ that guarantees the existence of Lagrange multipliers.

Proposition 3.8. (second-order sufficient condition for strict local minimizers in SOCP). *Let $\bar{x} \in \Gamma$ be a feasible solution to* (3.31) *with $\Phi(\bar{x}) = 0$, and let $\Lambda(\bar{x}, \bar{x}^*) \neq \emptyset$ for $\bar{x}^* :=$*

$-\nabla\varphi_0(\bar{x})$. *Taking any* $\bar{\lambda} \in \Lambda(\bar{x}, \bar{x}^*)$*, impose the so-called second-order sufficient condition* (*SOSC*) *for optimality:*

$$\langle \nabla^2_{xx} L(\bar{x}, \bar{\lambda})u, u\rangle > 0 \qquad \text{for all } \; 0 \neq u \in \big\{u \in \mathbb{R}^n \big|\; \nabla\Phi(\bar{x})u \in \mathcal{Q} \cap \{\bar{\lambda}\}^{\perp}\big\}, \tag{3.32}$$

where $L(x, \lambda) := \varphi_0(x) + \langle \lambda, \Phi(x)\rangle$*. Then* $\bar{x}$ *is indeed a strict local minimizer for problem* (3.31).

Proof. Suppose that $\bar{x}$ is not a strict local minimizer for (3.31) and thus find a sequence $x^k \to \bar{x}$ as $k \to \infty$ with $\Phi(x^k) \in \mathcal{Q}$ and $\varphi_0(x^k) < \varphi_0(\bar{x})$; hence $x^k \neq \bar{x}$. Define $u_k := \frac{x^k - \bar{x}}{\|x^k - \bar{x}\|}$ and assume without loss of generality that $u_k \to \bar{u}$ for some $0 \neq \bar{u} \in \mathbb{R}^n$. It tells us that

$$\nabla\Phi(\bar{x})\bar{u} \in \mathcal{Q} \quad \text{and} \quad \langle \nabla\varphi_0(\bar{x}), \bar{u}\rangle \leq 0.$$

Combining this with $\bar{\lambda} \in \Lambda(\bar{x}, -\nabla\varphi_0(\bar{x}))$ yields $\nabla\Phi(\bar{x})\bar{u} \in \mathcal{Q} \cap \{\bar{\lambda}\}^{\perp}$. It is not hard to see that

$$\varphi_0(x^k) - \varphi_0(\bar{x}) + \langle \bar{\lambda}, \Phi(x^k)\rangle \leq 0,$$

which implies by the twice differentiability of φ_0 and Φ at $\bar{x}$ that

$$\langle \nabla^2_{xx} L(\bar{x}, \bar{\lambda})\bar{u}, \bar{u}\rangle \leq 0 \qquad \text{with} \quad \bar{u} \neq 0.$$

This contradicts (3.32) and hence completes the proof of the proposition. $\triangle$

3.2 Graphical Derivative of the Normal Cone Mapping

Here we present the main result of the paper on calculating the graphical derivative of the normal cone mapping generated by the constraint system (1.3) under imposing *merely* the *MSCQ* condition. Great progress in this direction was recently made by Gfrerer and Outrata [19] (preprint of 2014) who calculated this second-order object for polyhedral/NLP constraint systems under MSCQ and a certain additional condition instead of the standard nondegeneracy and Mangasarian-Fromovitz constraint qualifications. Then the additional condition to MSCQ was relaxed in [16] and fully dropped subsequently by Chieu and Hien [7] in the NLP setting. Various calculating formulas for the graphical derivative of the normal cone mappings to nonpolyhedral (including ice-cream) constraints were derived in [20, 40, 41]. However, all these results were obtained under the nondegeneracy condition (a conic extension of the classical linear independence of constraint gradients in NLPs). Thus the graphical derivative formula for the second-order cone constraints given in the next theorem is new even under the Robinson constraint qualification. Furthermore, our proof of this result is significantly different in the major part from that in [19] and the subsequent developments for polyhedral systems; see Remark 3.12 for more discussions.

Theorem 3.9. (graphical derivative of the normal cone mapping for the second-order cone constraint systems). *Let* $(\bar{x},\bar{x}^*)\in\operatorname{gph}N_\Gamma$*, and let MSCQ from Definition* 3.2 *hold at* $\bar{x}$ *with modulus* κ*. Then the tangent cone to* $\operatorname{gph}N_\Gamma$ *is represented by*

$$T_{\operatorname{gph}N_\Gamma}(\bar{x},\bar{x}^*)=\Big\{(v,v^*)\;\Big|\;\begin{array}{l}v^*\in\big(\nabla^2\langle\lambda,\Phi\rangle(\bar{x})+\mathcal{H}(\bar{x};\lambda)\big)v+N_{K(\bar{x},\bar{x}^*)}(v)\\ \text{for some}\;\;\lambda\in\Lambda(\bar{x},\bar{x}^*;v)\Big\},\end{array}\tag{3.33}$$

where $\Lambda(\bar{x},\bar{x}^*;v)$ *is the set of optimal solutions to* (3.24) *with* $\mathcal{H}$ *defined in* (3.23)*. Consequently, for all* $v\in\mathbb{R}^n$ *we have the graphical derivative formula*

$$(DN_\Gamma)(\bar{x},\bar{x}^*)(v)=\Big\{\big(\nabla^2\langle\lambda,\Phi\rangle(\bar{x})+\mathcal{H}(\bar{x};\lambda)\big)v\;\Big|\;\lambda\in\Lambda(\bar{x},\bar{x}^*;v)\Big\}+N_{K(\bar{x},\bar{x}^*)}(v).\tag{3.34}$$

Proof. It is sufficient to justify the tangent cone formula (3.33), which immediately yields the graphical derivative one (3.34) by definition (2.5). We split the proof of (3.33) into three different cases depending on the position of $\Phi(\bar{x})$ in $\mathcal{Q}$. First assume that $\Phi(\bar{x}) \in$ int$(\mathcal{Q})$ and thus get

$$\bar{x}^* \in N_\Gamma(\bar{x}) = \nabla\Phi(\bar{x})^* N_{\mathcal{Q}}(\Phi(\bar{x})) = \{0\}, \qquad T_\Gamma(\bar{x}) = \mathbb{R}^n, \quad \text{and} \quad K(\bar{x}, \bar{x}^*) = \mathbb{R}^n.$$

By the continuity of Φ around $\bar{x}$ we find a neighborhood U of $\bar{x}$ such that $\Phi(x) \in$ int$(\mathcal{Q})$ and $N_\Gamma(x) = \{0\}$ whenever $x \in U$. This tells us that

$$\text{gph}\, N_\Gamma \cap [U \times \mathbb{R}^n] = U \times \{0\},$$

which obviously provides the tangent cone representation

$$T_{\text{gph}\, N_\Gamma}(\bar{x}, 0) = \mathbb{R}^n \times \{0\}. \tag{3.35}$$

On the other hand, it follows from $\Lambda(\bar{x}, \bar{x}^*) = \{0\}$ that $\Lambda(\bar{x}, \bar{x}^*; v) = \{0\}$ for all $v \in K(\bar{x}, \bar{x}^*)$. This shows that the right-hand side of (3.33) amounts to $\mathbb{R}^n \times \{0\}$. Combining it with (3.35) verifies the tangent cone formula (3.33) in this case.

Next we consider the case where $\Phi(\bar{x}) \in \text{bd}\,(\mathcal{Q}) \setminus \{0\}$. As argued above, Γ can be described in this case by (3.9) via the mapping ψ from (2.17). Using Lemma 3.3 confirms that the mapping $x \mapsto \psi \circ \Phi(x) - \mathbb{R}^2_-$ is metrically subregular at $(\bar{x}, 0)$. Thus it follows from

[19, Theorem 1] that

$$T_{\mathrm{gph}\,N_\Gamma}(\bar{x},\bar{x}^*)=\big\{(v,v^*)\,\big|\,v^*\in\nabla^2\langle\widetilde{\lambda},\,\psi\circ\Phi\rangle(\bar{x})v+N_{K(\bar{x},\bar{x}^*)}(v)\text{ for some }\widetilde{\lambda}\in\widetilde{\Lambda}(\bar{x},\bar{x}^*;v)\big\},\tag{3.36}$$

where $\widetilde{\Lambda}(\bar{x},\bar{x}^*;v)$ is the set of optimal solutions to the linear program

$$\min_{\widetilde{\lambda}\in\mathbb{R}^2}\big\{-\big\langle v,\nabla^2\langle\widetilde{\lambda},\psi\circ\Phi\rangle(\bar{x})v\big\rangle\,\big|\,\nabla(\psi\circ\Phi)(\bar{x})^*\widetilde{\lambda}=\bar{x}^*,\ \widetilde{\lambda}\in N_{\mathbb{R}^2_-}\big(\psi\circ\Phi(\bar{x})\big)\big\}.$$

Define the set of Lagrange multipliers for the modified constraint system (3.9) by

$$\widetilde{\Lambda}(\bar{x},\bar{x}^*)=\big\{\widetilde{\lambda}\in\mathbb{R}^2_-\,\big|\,\nabla(\psi\circ\Phi)(\bar{x})^*\widetilde{\lambda}=\bar{x}^*,\ \widetilde{\lambda}\in N_{\mathbb{R}^2_-}\big(\psi\circ\Phi(\bar{x})\big)\big\}.$$

It is not hard to observe the implication

$$\widetilde{\lambda}\in\widetilde{\Lambda}(\bar{x},\bar{x}^*)\Longrightarrow\lambda:=\nabla\psi\big(\Phi(\bar{x})\big)^*\widetilde{\lambda}\in\Lambda(\bar{x},\bar{x}^*),\tag{3.37}$$

where $\Lambda(\bar{x},\bar{x}^*)$ is taken from (3.1). Conversely, we claim that

$$\lambda=(\lambda_0,\lambda_r)\in\Lambda(\bar{x},\bar{x}^*)\Longrightarrow\widetilde{\lambda}:=\Big(-\frac{\lambda_0}{2\Phi_0(\bar{x})},0\Big)\in\widetilde{\Lambda}(\bar{x},\bar{x}^*).\tag{3.38}$$

To verify (3.38), we need to show that any $\lambda=(\lambda_0,\lambda_r)\in\Lambda(\bar{x},\bar{x}^*)$ can be represented as $\lambda=\nabla\psi(\Phi(\bar{x}))^*\widetilde{\lambda}$ with some $\widetilde{\lambda}\in N_{\mathbb{R}^2_-}(\psi\circ\Phi(\bar{x}))$. Since $\Phi(\bar{x})=(\Phi_0(\bar{x}),\Phi_r(\bar{x}))\in\mathrm{bd}\,(\mathcal{Q})\setminus\{0\}$, it follows that $(\psi\circ\Phi)(\bar{x})=(0,-\Phi_0(\bar{x}))$ and $\Phi_0(\bar{x})>0$, which lead us to $N_{\mathbb{R}^2_-}((\psi\circ\Phi)(\bar{x}))=\mathbb{R}_+\times\{0\}$. Thus we need to find some $\alpha\geq 0$ such that the pair $\widetilde{\lambda}=(\alpha,0)$ satisfies the equation

$$\lambda = \nabla\psi\big(\Phi(\bar{x})\big)^*\widetilde{\lambda} = \begin{bmatrix} -2\Phi_0(\bar{x}) & -1 \\ 2\Phi_r(\bar{x}) & 0 \end{bmatrix} \begin{pmatrix} \alpha \\ 0 \end{pmatrix} = 2\alpha\widehat{\Phi}(\bar{x}),$$

which is clearly fulfilled for $\alpha = -\dfrac{\lambda_0}{2\Phi_0(\bar{x})}$ and hence justifies the claimed implication (3.38). Using these observations brings us to the following relationships:

$$\begin{aligned} \nabla^2\langle\widetilde{\lambda}, \psi\circ\Phi\rangle(\bar{x}) &= \nabla^2\left(\alpha(-\Phi_0^2(\cdot) + \|\Phi_r(\cdot)\|^2)\right)(\bar{x}) = 2\alpha\nabla\left[\widehat{\Phi}(\cdot)^*\nabla\Phi(\cdot)\right](\bar{x}) \\ &= 2\alpha\left[\nabla\widehat{\Phi}(\bar{x})^*\nabla\Phi(\bar{x}) + \langle\widehat{\Phi}(\bar{x}), \nabla^2\Phi(\bar{x})\rangle\right] \\ &= 2\alpha\nabla\widehat{\Phi}(\bar{x})^*\nabla\Phi(\bar{x}) + \langle 2\alpha\widehat{\Phi}(\bar{x}), \nabla^2\Phi(\bar{x})\rangle \\ &= -\frac{\lambda_0}{\Phi_0(\bar{x})}\nabla\widehat{\Phi}(\bar{x})^*\nabla\Phi(\bar{x}) + \langle\lambda, \nabla^2\Phi(\bar{x})\rangle = \mathcal{H}(\bar{x};\lambda) + \nabla^2\langle\lambda,\Phi\rangle(\bar{x}). \end{aligned}$$

Combining it with (3.37) and (3.38) confirms that (3.36) reduces to (3.33) in this case.

It remains to consider the most difficult nonpolyhedral case where $\Phi(\bar{x}) = 0$. We begin with verifying the inclusion "$\subset$" in (3.33). Picking any $(v, v^*) \in T_{\text{gph}\, N_\Gamma}(\bar{x}, \bar{x}^*)$, observe that it suffices to show the validity of the following two inclusions:

$$v \in K(\bar{x}, \bar{x}^*) \quad \text{and} \quad v^* - \nabla^2\langle\bar{\lambda}, \Phi\rangle(\bar{x})v \in N_{K(\bar{x},\bar{x}^*)}(v) \quad \text{for some } \bar{\lambda} \in \Lambda(\bar{x}, \bar{x}^*; v). \qquad (3.39)$$

To proceed, we get from the tangent cone definition (2.1) that for $(v, v^*) \in T_{\text{gph}\, N_\Gamma}(\bar{x}, \bar{x}^*)$ there are sequences $t_k \downarrow 0$ and $(v^k, v^{k,*}) \to (v, v^*)$ as $k \to \infty$ such that

$$(x^k, x^{k,*}) := (\bar{x} + t_k v^k, \bar{x}^* + t_k v^{k,*}) \in \text{gph}\, N_\Gamma, \quad k \in \mathbb{N}.$$

Let us split the subsequent proof of the inclusion "$\subset$" in (3.33) into the four steps.

Step 1: *There exists a sequence* $\{\lambda^k \in \Lambda(x^k, x^{k,*})\}$ *with* $\lambda^k \to \bar\lambda$ *as* $k \to \infty$ *for some* $\bar\lambda \in \Lambda(\bar x, \bar x^*)$. To verify this statement, we deduce first directly from [18, Lemma 2.1] and the robustness of MSCQ that there is a positive number δ such that $x^k \in \Gamma \cap \mathbb{B}_\delta(\bar x)$ and that

$$\Lambda(x^k, x^{k,*}) \cap \kappa\|x^{k,*}\|\mathbb{B} \neq \emptyset \quad \text{for all} \;\; k \in \mathbb{N},$$

where $\kappa > 0$ is the constant taken from Definition 3.2. This allows us to find $\lambda^k \in \Lambda(x^k, x^{k,*})$ so that $\|\lambda^k\| \le \kappa\|x^{k,*}\|$ for all $k \in \mathbb{N}$. Thus the boundedness of $\{x^{k,*}\}$ yields the one for $\{\lambda^k\}$, and therefore $\lambda^k \to \bar\lambda$ for some $\bar\lambda \in \mathbb{R}^{m+1}$ along a subsequence. In this way we conclude that $\bar\lambda \in \Lambda(\bar x, \bar x^*)$, where the latter set is represented by (3.3) due to $\Phi(\bar x) = 0$.

Step 2: *We have* $v \in T_\Gamma(\bar x) \cap \{\bar x^*\}^\perp = K(\bar x, \bar x^*)$. The equality here is by the definition of the critical cone (3.2); so getting the first one in (3.39) requires only the verification of the claimed inclusion. Recall from (3.8) that $T_\Gamma(\bar x) = \big\{w \in \mathbb{R}^n \,\big|\, \nabla\Phi(\bar x)w \in \mathcal{Q}\big\}$. It follows from $x^k \in \Gamma$ and $\Phi(\bar x) = 0$ that

$$\Phi(x^k) = t_k \nabla\Phi(\bar x)v^k + o(t_k) \in \mathcal{Q} \;\; \text{for all} \;\; k \in \mathbb{N}.$$

Dividing the latter by t_k and passing to the limit as $k \to \infty$ yield $\nabla\Phi(\bar x)v \in \mathcal{Q}$, and so $v \in T_\Gamma(\bar x)$. Since $\bar\lambda \in \Lambda(\bar x, \bar x^*)$ and $\langle \lambda^k, \Phi(x^k)\rangle = 0$ for all $k \in \mathbb{N}$, we get

$$\begin{aligned}
\langle \bar x^*, v\rangle &= \langle \nabla\Phi(\bar x)^*\bar\lambda, v\rangle = \langle \bar\lambda, \nabla\Phi(\bar x)v\rangle = \lim_{k\to\infty}\langle \lambda^k, \nabla\Phi(\bar x)v^k\rangle \\
&= \lim_{k\to\infty}\frac{\langle \lambda^k, \Phi(x^k) + o(t_k)\rangle}{t_k} = \lim_{k\to\infty}\Big\langle \lambda^k, \frac{o(t_k)}{t_k}\Big\rangle = 0
\end{aligned}$$

and thus finish the proof of the statement in Step 2.

Step 3: *We have the inclusion* $v^*-\nabla^2\langle\bar\lambda,\Phi\rangle(\bar x)v\in\big(K(\bar x,\bar x^*)\big)^*$ *for the multiplier* $\bar\lambda\in\Lambda(\bar x,\bar x^*)$ *constructed in Step* 1. Indeed, by the definition of $x^{k,*}$ we get

$$v^{k,*}=\frac{x^{k,*}-\bar x^*}{t_k}=\frac{\nabla\Phi(x^k)^*\lambda^k-\bar x^*}{t_k}=\frac{\nabla\Phi(\bar x)^*\lambda^k+t_k\nabla^2\langle\lambda^k,\Phi\rangle(\bar x)v^k+o(t_k)-\bar x^*}{t_k},$$

which in turn leads us to the equality

$$v^{k,*}-\nabla^2\langle\lambda^k,\Phi\rangle(\bar x)v^k+\frac{o(t_k)}{t_k}=\nabla\Phi(\bar x)^*\frac{\lambda^k}{t_k}-\frac{\bar x^*}{t_k}. \tag{3.40}$$

Using $\lambda^k\in-\mathcal{Q}=N_{\mathcal{Q}}(\Phi(\bar x))$ and (3.14) yields $v^{k,*}-\nabla^2\langle\lambda^k,\Phi\rangle(\bar x)v^k+\dfrac{o(t_k)}{t_k}\in(K(\bar x,\bar x^*))^*$. Since $(K(\bar x,\bar x^*))^*$ is closed, the passage to the limit as $k\to\infty$ gives us the desired inclusion.

Step 4: *We have* $\bar\lambda\in\Lambda(\bar x,\bar x^*;v)$ *and* $\big\langle v,v^*-\nabla^2\langle\bar\lambda,\Phi\rangle(\bar x)v\big\rangle=0$ *for the multiplier* $\bar\lambda$ *constructed above.* We first show that

$$\big\langle v,\nabla^2\langle\lambda,\Phi\rangle(\bar x)v\big\rangle\le\big\langle v,\nabla^2\langle\bar\lambda,\Phi\rangle(\bar x)v\big\rangle\quad\text{for any}\quad\lambda\in\Lambda(\bar x,\bar x^*), \tag{3.41}$$

which verifies the inclusion $\bar\lambda\in\Lambda(\bar x,\bar x^*;v)$. Picking $\lambda\in\Lambda(\bar x,\bar x^*)$ gives us $\lambda\in-\mathcal{Q}$ by (3.3). Using this together with $\Phi(x^k)\in\mathcal{Q}$ and $\langle\lambda^k,\Phi(x^k)\rangle=0$, we get the relationships

$$\begin{aligned}0&\le-\langle\lambda,\Phi(x^k)\rangle=\langle\lambda^k-\lambda,\Phi(x^k)\rangle\\&=t_k\langle\lambda^k-\lambda,\nabla\Phi(\bar x)v^k\rangle+\frac{1}{2}t_k^2\big\langle v^k,\nabla^2\langle\lambda^k-\lambda,\Phi\rangle(\bar x)v^k\big\rangle+o(t_k^2)\\&=t_k\langle\nabla\Phi(\bar x)^*\lambda^k-\bar x^*,v^k\rangle+\frac{1}{2}t_k^2\big\langle v^k,\nabla^2\langle\lambda^k-\lambda,\Phi\rangle(\bar x)v^k\big\rangle+o(t_k^2).\end{aligned}$$

Dividing by t_k^2 and employing (3.40) bring us to

$$0 \le \left\langle v^k, v^{k,*} - \nabla^2\langle\lambda^k, \Phi\rangle(\bar{x})v^k + \frac{o(t_k)}{t_k}\right\rangle + \frac{1}{2}\left\langle v^k, \nabla^2\langle\lambda^k - \lambda, \Phi\rangle(\bar{x})v^k\right\rangle + \frac{o(t_k^2)}{t_k^2},$$

which implies by passing to the limit as $k \to \infty$ that

$$0 \le \left\langle v, v^* - \nabla^2\langle\bar{\lambda}, \Phi\rangle(\bar{x})v\right\rangle + \frac{1}{2}\left\langle v, \nabla^2\langle\bar{\lambda} - \lambda, \Phi\rangle(\bar{x})v\right\rangle. \tag{3.42}$$

It follows from the relationships proved in Steps 2 and 3 that

$$\left\langle v, v^* - \nabla^2\langle\bar{\lambda}, \Phi\rangle(\bar{x})v\right\rangle \le 0. \tag{3.43}$$

which together with (3.42) yields (3.41). Finally, since (3.42) holds for any $\lambda \in \Lambda(\bar{x}, \bar{x}^*)$, letting $\lambda = \bar{\lambda}$ therein results in the inequality

$$\left\langle v, v^* - \nabla^2\langle\bar{\lambda}, \Phi\rangle(\bar{x})v\right\rangle \ge 0.$$

Combining it with (3.43) justifies Step 4, and thus we arrive at the inclusion “$\subset$" in (3.33).

Now we give a detailed proof of the opposite inclusion in (3.33), which occurs to be more involved. Pick (v, v^*) from the right-hand side of (3.33), which satisfies (3.39) in the case of $\Phi(\bar{x}) = 0$ under consideration. We proceed by showing that there are sequences $t_k \downarrow 0$ and $x^k \to \bar{x}$ as $k \to \infty$ satisfying the conditions

$$\bar{x} + t_k v - x^k = o(t_k) \quad \text{and} \quad \operatorname{dist}\big(\bar{x}^* + t_k v^*; N_\Gamma(x^k)\big) = o(t_k), \quad k \in \mathbb{N}. \tag{3.44}$$

These guarantee the existence of $x^{k,*} \in N_\Gamma(x^k)$ such that

$$(x^k, x^{k,*}) = \Big(\bar{x} + t_k\big(v + \frac{o(t_k)}{t_k}\big), \bar{x}^* + t_k\big(v^* + \frac{o(t_k)}{t_k}\big)\Big) \in \operatorname{gph} N_\Gamma,$$

and thus we arrive at $(v, v^*) \in T_{\operatorname{gph} N_\Gamma}(\bar{x}, \bar{x}^*)$, which is the goal.

To begin with, we conclude by the choice of (v, v^*) and the usage of Theorem 3.4 that there are $\lambda \in \Lambda(\bar{x}, \bar{x}^*; v)$ and $\mu \in T_{-\mathcal{Q}}(\lambda)$ satisfying the equalities

$$v^* = \nabla^2\langle\lambda, \Phi\rangle(\bar{x})v + \nabla\Phi(\bar{x})^*\mu \quad \text{and} \quad \langle\mu, \nabla\Phi(\bar{x})v\rangle = 0. \tag{3.45}$$

It comes from $\mu \in T_{-\mathcal{Q}}(\lambda)$ that there are sequences $t_i \downarrow 0$ and $\mu_i \to \mu$ as $i \to \infty$ with $\lambda + t_i\mu_i \in -\mathcal{Q}$. Choose $\alpha > 0$ so small that $\alpha\|\nabla^2\langle\lambda, \Phi\rangle(\bar{x})\| \le \dfrac{1}{2}$ holds. This ensures that the matrix $I + \alpha\nabla^2\langle\lambda, \Phi\rangle(\bar{x})$ is positive-definite, where I is the $n \times n$ identity matrix. Proposition 3.8 tells us that there exists $r > 0$ such that $\bar{x}$ is the strict global minimizer for the problem

$$\min_{x\in\mathbb{R}^n} \left\{\|\bar{x} + \alpha\bar{x}^* - x\|^2 \,\middle|\, x \in \Gamma \cap \mathbb{B}_r(\bar{x})\right\}. \tag{3.46}$$

For any fixed $k \in \mathbb{N}$ we select a positive number $\varepsilon_k < \big(16\alpha k^2(\kappa\|\bar{x}^*\| + 1)\big)^{-1}$. Since λ solves the linear optimization problem (3.26), Theorem 3.7 ensures the existence of $z^k \in \mathbb{R}^n$ with

$$\operatorname{dist}\big(\nabla\Phi(\bar{x})z^k + \big\langle v, \nabla^2\Phi(\bar{x})v\big\rangle; \mathcal{Q}\big) \le \varepsilon_k \quad \text{and} \quad \langle\bar{x}^*, z^k\rangle + \big\langle v, \nabla^2\langle\lambda, \Phi\rangle(\bar{x})v\big\rangle \ge -\varepsilon_k. \tag{3.47}$$

Picking next $i \in \mathbb{N}$, consider yet another optimization problem

$$\min_{x\in\mathbb{R}^n} \big\{ \|\bar{x} + t_i v + \frac{1}{2}t_i^2 z^k + \alpha(\bar{x}^* + t_i v^*) - x\|^2 \,\big|\, x \in \Gamma \cap \mathbb{B}_r(\bar{x}) \big\}, \tag{3.48}$$

which admits an optimal solution due to the classical Weierstrass theorem. It is not hard to check that $x^i \to \bar{x}$ as $i \to \infty$. Indeed, suppose that $x^i \to \widetilde{x}$ for some $\widetilde{x}$ along a subsequence, we see that

$$\|\bar{x} + \alpha\bar{x}^* - \widetilde{x}\|^2 \le \|\bar{x} + \alpha\bar{x}^* - x\|^2 \quad \text{for all } \ x \in \Gamma \cap \mathbb{B}_r(\bar{x}),$$

which yields $\widetilde{x} = \bar{x}$ since $\bar{x}$ is the strict global minimizer for (3.46). Assume now without loss of generality that $x^i \in \text{int}\mathbb{B}_r(\bar{x})$ for $i \in \mathbb{N}$ sufficiently large and utilize the first-order necessary optimality condition from [38, Proposition 5.1] at x^i for problem (3.48) to get the following inclusion:

$$\alpha(\bar{x}^* + t_i v^*) + t_i\Big(\frac{\bar{x} + t_i v - x^i}{t_i} + \frac{1}{2}t_i z^k\Big) \in N_\Gamma(x^i). \tag{3.49}$$

It follows from $\Phi(\bar{x}) = 0$ and the twice differentiability of Φ around $\bar{x}$ that

$$\Phi(\bar{x} + t_i v + \frac{1}{2}t_i^2 z^k) = t_i \nabla\Phi(\bar{x})v + \frac{1}{2}t_i^2\Big((\nabla\Phi(\bar{x})z^k + \big\langle v, \nabla^2\Phi(\bar{x})v\big\rangle\Big) + o(t_i^2).$$

Since v satisfies (3.39), we get $\nabla\Phi(\bar{x})v \in T_{\mathcal{Q}}(\Phi(\bar{x})) = \mathcal{Q}$. Taking this into account along with the first inequality in (3.47), we obtain the estimate

$$\text{dist}\Big(\Phi(\bar{x} + t_i v + \frac{1}{2}t_i^2 z^k); \mathcal{Q}\Big) \le \frac{\varepsilon_k}{2}t_i^2 + o(t_i^2),$$

which together with the assumed MSCQ at $\bar{x}$ results in

$$\operatorname{dist}\Big(\bar{x}+t_iv+\frac{1}{2}t_i^2z^k;\Gamma\Big)\le\frac{\kappa\varepsilon_k}{2}t_i^2+o(t_i^2).$$

This guarantees that for any $i\in\mathbb{N}$ there exists $\widetilde{x}^i\in\Gamma$ such that

$$\|\bar{x}+t_iv+\frac{1}{2}t_i^2z^k-\widetilde{x}^i\|\le\frac{\kappa\varepsilon_k}{2}t_i^2+o(t_i^2), \tag{3.50}$$

and so we verify that $\widetilde{x}^i\to\bar{x}$ as $i\to\infty$. This tells us that $\widetilde{x}^i\in\Gamma\cap\mathbb{B}_r(\bar{x})$ for all i sufficiently large. Since x^i is a global minimizer for (3.48), we get

$$\big\|\bar{x}+t_iv+\frac{1}{2}t_i^2z^k+\alpha(\bar{x}^*+t_iv^*)-x^i\big\|^2\le\big\|\bar{x}+t_iv+\frac{1}{2}t_i^2z+\alpha(\bar{x}^*+t_iv^*)-\widetilde{x}^i\big\|^2$$

for all large i, which together with (3.50) leads us to the estimates

$$\begin{aligned}
&\big\|\bar{x}+t_iv+\frac{1}{2}t_i^2z^k-x^i\big\|^2+2\alpha\big\langle\bar{x}^*+t_iv^*,\bar{x}+t_iv+\frac{1}{2}t_i^2z^k-x^i\big\rangle\\
\le\;&\big\|\bar{x}+t_iv+\frac{1}{2}t_i^2z^k-\widetilde{x}^i\big\|^2+2\alpha\big\langle\bar{x}^*+t_iv^*,\bar{x}+t_iv+\frac{1}{2}t_i^2z^k-\widetilde{x}^i\big\rangle\\
\le\;&\big\|\bar{x}+t_iv+\frac{1}{2}t_i^2z^k-\widetilde{x}^i\big\|^2+2\alpha\left(\|\bar{x}^*\|+t_i\|v^*\|\right)\big\|\bar{x}+t_iv+\frac{1}{2}t_i^2z^k-\widetilde{x}^i\big\|\\
\le\;&\alpha\kappa\|\bar{x}^*\|\varepsilon_kt_i^2+o(t_i^2).
\end{aligned}$$

These yield in turn the relationships

$$\begin{aligned}\left\|\bar{x}+t_iv+\frac{1}{2}t_i^2z^k-x^i\right\|^2 &\leq -2\alpha\left\langle\bar{x}^*+t_iv^*,\bar{x}+t_iv+\frac{1}{2}t_i^2z^k-x^i\right\rangle+\alpha\kappa\|\bar{x}^*\|\varepsilon_kt_i^2+o(t_i^2)\\ &= 2\alpha\left[\langle\bar{x}^*+t_iv^*,x^i-\bar{x}\rangle-t_i\langle\bar{x}^*,v\rangle-t_i^2\langle v^*,v\rangle-\frac{1}{2}t_i^2\langle\bar{x}^*,z^k\rangle\right]\\ &\quad+\alpha\kappa\|\bar{x}^*\|\varepsilon_kt_i^2+o(t_i^2). \qquad (3.51)\end{aligned}$$

Recall further from the first inclusion in (3.39) that $v\in K(\bar{x},\bar{x}^*)$ and hence $\langle\bar{x}^*,v\rangle=0$. It follows from (3.45) and (3.47), respectively, that

$$\langle v^*,v\rangle=\left\langle v,\nabla^2\langle\lambda,\Phi\rangle(\bar{x})v\right\rangle \quad\text{and}\quad -\langle\bar{x}^*,z^k\rangle\leq\left\langle v,\nabla^2\langle\lambda,\Phi\rangle(\bar{x})v\right\rangle+\varepsilon_k. \qquad (3.52)$$

Next we are going to find an upper estimate for the first term on the right-hand side of the equality in (3.51). It follows from both equalities in (3.45) that

$$\begin{aligned}\langle\bar{x}^*+t_iv^*,x^i-\bar{x}\rangle &= \left\langle\nabla\Phi(\bar{x})^*\lambda+t_i\left(\nabla^2\langle\lambda,\Phi\rangle(\bar{x})v+\nabla\Phi(\bar{x})^*\mu\right),x^i-\bar{x}\right\rangle\\ &= \left\langle\lambda+t_i\mu,\nabla\Phi(\bar{x})(x^i-\bar{x})\right\rangle+t_i\left\langle v,\nabla^2\langle\lambda,\Phi\rangle(\bar{x})(x^i-\bar{x})\right\rangle\\ &= \left\langle\lambda+t_i\mu_i,\nabla\Phi(\bar{x})(x^i-\bar{x})\right\rangle+t_i\left\langle\mu-\mu_i,\nabla\Phi(\bar{x})(x^i-\bar{x})\right\rangle\\ &\quad+t_i\left\langle v,\nabla^2\langle\lambda,\Phi\rangle(\bar{x})(x^i-\bar{x})\right\rangle\\ &= \left\langle\lambda+t_i\mu_i,\Phi(x^i)-\frac{1}{2}\langle x^i-\bar{x},\nabla^2\Phi(\bar{x})(x^i-\bar{x})\rangle\right\rangle\\ &\quad+t_i\left\langle v,\nabla^2\langle\lambda,\Phi\rangle(\bar{x})(x^i-\bar{x})\right\rangle+o(t_i\|x^i-\bar{x}\|)+o(\|x^i-\bar{x}\|^2)\\ &= \langle\lambda+t_i\mu_i,\Phi(x^i)\rangle-\frac{1}{2}\left\langle x^i-\bar{x},\nabla^2\langle\lambda,\Phi\rangle(\bar{x})(x^i-\bar{x})\right\rangle\\ &\quad+t_i\left\langle v,\nabla^2\langle\lambda,\Phi\rangle(\bar{x})(x^i-\bar{x})\right\rangle+o(t_i\|x^i-\bar{x}\|)+o(\|x^i-\bar{x}\|^2).\end{aligned}$$

Using these together with $\lambda + t_i\mu_i \in -\mathcal{Q}$ and $\Phi(x^i) \in \mathcal{Q}$ brings us to the estimate

$$\begin{aligned}\langle \bar{x}^* + t_i v^*, x^i - \bar{x}\rangle \;\leq\; & -\frac{1}{2}\left\langle \bar{x} - x^i, \nabla^2\langle\lambda,\Phi\rangle(\bar{x})(\bar{x} - x^i)\right\rangle - t_i\big\langle v, \nabla^2\langle\lambda,\Phi\rangle(\bar{x})(\bar{x} - x^i)\big\rangle \\ & + o(t_i\|x^i - \bar{x}\|) + o(\|x^i - \bar{x}\|^2). \qquad (3.53)\end{aligned}$$

Combining now the conditions in (3.51)–(3.53), we arrive at the following relationships:

$$\begin{aligned}\Big\|\bar{x} + t_i v + \frac{1}{2}t_i^2 z^k - x^i\Big\|^2 \;\leq\; & 2\alpha\Big[-\frac{1}{2}\left\langle \bar{x} - x^i, \nabla^2\langle\lambda,\Phi\rangle(\bar{x})(\bar{x} - x^i)\right\rangle - t_i\left\langle v, \nabla^2\langle\lambda,\Phi\rangle(\bar{x})(\bar{x} - x^i)\right\rangle \\ & - t_i^2\left\langle v, \nabla^2\langle\lambda,\Phi\rangle(\bar{x})v\right\rangle + \frac{1}{2}t_i^2\left(\left\langle v, \nabla^2\langle\lambda,\Phi\rangle(\bar{x})v\right\rangle + \varepsilon_k\right)\Big] \\ & + o(t_i\|x^i - \bar{x}\|) + o(\|x^i - \bar{x}\|^2) + \alpha\kappa\|\bar{x}^*\|\varepsilon_k t_i^2 + o(t_i^2) \\ = \;& -\alpha\Big[\left\langle \bar{x} - x^i, \nabla^2\langle\lambda,\Phi\rangle(\bar{x})(\bar{x} - x^i)\right\rangle + 2t_i\left\langle v, \nabla^2\langle\lambda,\Phi\rangle(\bar{x})(\bar{x} - x^i)\right\rangle \\ & + t_i^2\left\langle v, \nabla^2\langle\lambda,\Phi\rangle(\bar{x})v\right\rangle\Big] + \alpha\varepsilon_k t_i^2 + \alpha\kappa\|\bar{x}^*\|\varepsilon_k t_i^2 \\ & + o(t_i\|x^i - \bar{x}\|) + o(\|x^i - \bar{x}\|^2) + o(t_i^2) \\ = \;& -\alpha\left\langle \bar{x} + t_i v - x^i, \nabla^2\langle\lambda,\Phi\rangle(\bar{x})(\bar{x} + t_i v - x^i)\right\rangle + \alpha(\kappa\|\bar{x}^*\| + 1)\varepsilon_k t_i^2 \\ & + o(t_i\|x^i - \bar{x}\|) + o(\|x^i - \bar{x}\|^2) + o(t_i^2) \\ \leq \;& \frac{1}{2}\|\bar{x} + t_i v - x^i\|^2 + \alpha(\kappa\|\bar{x}^*\| + 1)\varepsilon_k t_i^2 \\ & + o(t_i\|x^i - \bar{x}\|) + o(\|x^i - \bar{x}\|^2) + o(t_i^2),\end{aligned}$$

where the last inequality comes from the fact that the matrix $\frac{1}{2}I + \alpha\nabla^2\langle\lambda,\Phi\rangle(\bar{x})$ is positive-semidefinite. This allows us to conclude that

$$\Big\|\bar{x}+t_i v+\frac{1}{2}t_i^2 z^k-x^i\Big\|^2-\frac{1}{2}\|\bar{x}+t_i v-x^i\|^2 \leq \alpha(\kappa\|\bar{x}^*\|+1)\varepsilon_k t_i^2+o(t_i\|x^i-\bar{x}\|)+o(\|x^i-\bar{x}\|^2)+o(t_i^2),$$

which verifies the inequality

$$\frac{1}{2}\|\bar{x}+t_iv-x^i\|^2+t_i^2\langle z^k,\bar{x}-x^i\rangle+t_i^3\langle z^k,v\rangle+\frac{1}{4}t_i^4\|z^k\|^2 \le \alpha(\kappa\|\bar{x}^*\|+1)\varepsilon_k t_i^2+o(t_i\|x^i-\bar{x}\|) +o(\|x^i-\bar{x}\|^2)+o(t_i^2).$$

Since $\varepsilon_k < \dfrac{1}{16\alpha(\kappa\|\bar{x}^*\|+1)}$, the latter inequality can be simplified as

$$\begin{aligned}\|\bar{x}+t_iv-x^i\|^2 &\le 2\alpha(\kappa\|\bar{x}^*\|+1)\varepsilon_k t_i^2+o(t_i\|x^i-\bar{x}\|)+o(\|x^i-\bar{x}\|^2)+o(t_i^2) \quad (3.54)\\ &\le \frac{1}{8}t_i^2+o(t_i^2)+o(t_i\|x^i-\bar{x}\|)+o(\|x^i-\bar{x}\|^2),\end{aligned}$$

and therefore we get for all i sufficiently large that

$$\|\bar{x}+t_iv-x^i\| \le \frac{1}{2}\left(t_i+\|x^i-\bar{x}\|\right).$$

In this way we arrive at the estimates

$$\|x^i-\bar{x}\| \le \|\bar{x}+t_iv-x^i\|+t_i\|v\| \le \frac{1}{2}t_i+\frac{1}{2}\|x^i-\bar{x}\|+t_i\|v\|,$$

which in turn imply that $\|x^i-\bar{x}\|=O(t_i)$ and so $o(t_i\|x^i-\bar{x}\|)=o(\|x^i-\bar{x}\|^2)=o(t_i^2)$. Using these relationships together with (3.54) gives us

$$\|\bar{x}+t_iv-x^i\|^2 \le 2\alpha(\kappa\|\bar{x}^*\|+1)\varepsilon_k t_i^2+o(t_i^2),$$

and so we come by passing to the limit as $i \to \infty$ to the inequalities

$$\lim_{i\to\infty} \frac{\|\bar{x} + t_i v - x^i\|^2}{t_i^2} \le 2\alpha(\kappa\|\bar{x}^*\| + 1)\varepsilon_k \le \frac{1}{8k^2}.$$

Remember that $k \in \mathbb{N}$ has been fixed through the above proof of the inclusion "$\supset$" in (3.33). This allows us to find an index i_k for which we have the estimates

$$\frac{\|\bar{x} + t_{i_k} v - x^{i_k}\|}{t_{i_k}} \le \frac{1}{2k} \qquad \text{and} \qquad t_{i_k}\|z^k\| \le \frac{1}{k}. \tag{3.55}$$

Repeating this process for any $k \in \mathbb{N}$, we construct sequences t_{i_k} and x^{i_k} that satisfy (3.55) and such that $t_{i_k} \downarrow 0$ and $x^{i_k} \to \bar{x}$ as $k \to \infty$. Combining finally (3.55) and (3.49) leads us to

$$\frac{\operatorname{dist}\big(\bar{x}^* + t_{i_k} v^*; N_\Gamma(x^{i_k})\big)}{t_{i_k}} \le \frac{1}{k}.$$

It yields (3.44) with $t_k := t_{i_k}$ and $x^k := x^{i_k}$ and so completes the proof of the theorem. $\square$

It is worth mentioning an equivalent version of the pointbased formula (3.33) in Theorem 3.9, which is an ice-cream counterpart of the polyhedral result established recently by Gfrerer and Ye [21, Theorem 4].

Corollary 3.10. (representation of the tangent cone to the normal cone graph for ice-cream constraint systems with bounded Lagrange multipliers). *Under the assumptions of Theorem* 3.9 *there is* $\delta > 0$ *such that for all* $x \in \Gamma \cap \mathbb{B}_\delta(\bar{x})$ *and all* $x^* \in N_\Gamma(x)$ *we have the* *representations*

$$T_{\mathrm{gph}\,N_\Gamma}(x,x^*) \;=\; \big\{(v,v^*)\,\big|\;\; v^*\in\big(\nabla^2\langle\lambda,\Phi\rangle(x)+\mathcal{H}(x;\lambda)\big)v+N_{K(x,x^*)}(v) \;\; \textit{for some}\;\; \lambda\in\Lambda(x,x^*;v)\cap\kappa\|x^*\|\mathbb{B}\big\}, \tag{3.56}$$

$$(DN_\Gamma)(x,x^*)(v)=\big\{\big(\nabla^2\langle\lambda,\Phi\rangle(x)+\mathcal{H}(x;\lambda)\big)v\,\big|\;\lambda\in\Lambda(x,x^*;v)\cap\kappa\|x^*\|\mathbb{B}\big\}+N_{K(x,x^*)}(v). \tag{3.57}$$

Proof. We first observe in Step 1 of the proof of Theorem 3.9 that the limit of $\{\lambda^k\}$ actually belongs to the set $\Lambda(\bar{x},\bar{x}^*)\cap\kappa\|\bar{x}^*\|\mathbb{B}$. Thus the claimed representations for $x=\bar{x}$ follow immediately. The robustness of MSCQ allows us to select $\delta>0$ so that this condition holds at any $x\in\Gamma\cap\mathbb{B}_\delta(\bar{x})$ with the same modulus κ. It implies therefore that both (3.56) and (3.57) are satisfied for all such x. □

The next consequence of Theorem 3.9 concerns an important case of the tangent cone formula in the case where $\bar{x}^*=0$, which is used in what follows.

Corollary 3.11. (simplification of the graphical derivative formula for $\bar{x}^*=0$). *Let $\bar{x}^*=0$ in the framework of Theorem 3.9. Then we have*

$$T_{\mathrm{gph}\,N_\Gamma}(\bar{x},0)=\big\{(v,v^*)\,\big|\,v^*\in N_{K(\bar{x},0)}(v)\big\}=\mathrm{gph}\,N_{K(\bar{x},0)} \tag{3.58}$$

and correspondingly the graphical derivative formula

$$(DN_\Gamma)(\bar{x},0)(v)=N_{K(\bar{x},0)}(v)=\nabla\Phi(\bar{x})^*\big[N_\mathcal{Q}\big(\Phi(\bar{x})\big)\cap\{\nabla\Phi(\bar{x})v\}^\perp\big]. \tag{3.59}$$

Proof. If $\bar{x}^*=0$, we deduce from (3.56) that $\lambda=0$. Using this together with $\mathcal{H}(\bar{x};\lambda)=0$ for $\lambda=0$, we arrive at (3.58) and hence at (3.59). □

Remark 3.12. (discussions on the graphical derivative formulas).

(i) First we highlight some important differences between our proof of Theorem 3.9 for *nonpolyhedral* second-order constraint systems and its *polyhedral* counterpart for NLPs in [19, Theorem 1] and in the similar devices from [7, 16]. Unlike the latter proof that exploits the *Hoffman lemma* to verify the inclusion "$\subset$" in (3.33), we *do not appeal* to any error bound estimate; this is new even for polyhedral systems. Our approach is applicable to other cone-constrained frameworks; however, we believe that some *error bound* estimate *is needed* for the general setting. The reason for avoiding error bounds in the proof of Theorem 3.9 is that in the ice-cream case we have the inclusion $N_{\mathcal{Q}}(x) \subset N_{\mathcal{Q}}(0)$ for any $x \in \mathbb{R}^{m+1}$. Another difference between our proof and that in [19] lies in the justification of the inclusion "$\supset$" in the tangent cone formula. Indeed, the proof in [19] employs the *exact duality*, which holds in the polyhedral setting. In contrast, our proof relies on the *approximate duality* established in Theorem 3.7.

(ii) The first result on the tangent cone and the graphical derivative of normal cone mapping to the general conic constraint system

$$\Gamma := \big\{x \in \mathbb{R}^n \big|\ \Phi(x) \in \Theta\big\}, \tag{3.60}$$

where $\Theta \subset \mathbb{R}^m$ is a closed and convex, was established by Mordukhovich, Outrata and Ramírez [40, Theorem 3.3] under the *nondegeneracy condition* from [6] and the rather restrictive assumption on the convexity of Γ. This result was derived not in the form of (3.33) but in terms of the directional derivative of the *projection mapping* associated with Θ. Later the same authors improved this result in [41, Theorem 5.2] by dropping the convexity of Γ under the *projection derivation condition* discussed in Sect. 3.1, which enabled them to write the main result for (3.60) in the form of (3.33). However, as proved in Corollary 2.7, this PDC does not hold at nonzero boundary points of $\mathcal{Q}$ and so [41, Theorem 5.2]—obtained also under the nondegeneracy condition —cannot be utilized in the ice-cream framework when $\Phi(\bar{x}) \in \operatorname{bd}(\mathcal{Q}) \setminus \{0\}$.

(iii) Quite recently, Gfrerer and Outrata [20, Theorem 2] calculated the graphical derivative of the normal mapping to (3.60) under *nondegeneracy condition* when Θ is not necessarily convex. Combining their result with Corollary 2.4 above in the ice-cream framework, we see that it agrees with Theorem 3.33 *provided that* the nondegeneracy condition is satisfied. However, our results can be applied to much broader settings since it only demands the fulfillment of MSCQ. As mentioned above, our results seem to be new for SOCPs even under RCQ (3.6), which is equivalent to the metric regularity of $x \mapsto \Phi(x) - \mathcal{Q}$ around $(\bar{x}, 0)$. Note that in the latter case the Lagrange multiplier set $\Lambda(\bar{x}, \bar{x}^*)$ admits either the (LMS1) or the (LMS2) representation from its description in Proposition 3.3.

Next we illustrate the applicability of the main result in Theorem 3.9 to the ice-cream constraint systems at points where neither nondegeneracy nor Robinson constraint qualification is satisfied.

Example 3.13. (calculation of graphical derivative for ice-cream normal cone systems). Define the mapping $\Phi : \mathbb{R}^2 \to \mathbb{R}^3$ by

$$\Phi(x) := \big(\sqrt{2}x_1^2 + x_2, x_1^2 + \frac{1}{\sqrt{2}}x_2, x_1^2 - \frac{1}{\sqrt{2}}x_2\big) \quad \text{for} \quad x = (x_1, x_2) \in \mathbb{R}^2$$

and consider the constraint system associated with the three-dimensional ice-cream cone $\mathcal{Q}_3$:

$$\Gamma = \big\{x \in \mathbb{R}^2 \,\big|\, \Phi(\bar{x}) \in \mathcal{Q}_3\big\} = \big\{(x_1, x_2) \in \mathbb{R}^2 \,\big|\, x_2 \geq 0\big\}.$$

Given any $x \in \Gamma$, we claim that the mapping $x \mapsto \Phi(x) - \mathcal{Q}_3$ is metrically subregular at $(x, 0)$, i.e., MSCQ holds at x. To begin with, observe by (3.12) and direct calculations that

$$\operatorname{dist}\big((x_1, x_2); \Gamma\big) = \begin{cases} 0 & \text{if } x_2 \geq 0, \\ -x_2 & \text{if } x_2 < 0; \end{cases}$$

$$\operatorname{dist}\big(\Phi(x_1, x_2); \mathcal{Q}_3\big) = \begin{cases} 0 & \text{if } x_2 \geq 0, \\ -\sqrt{2}x_2 & \text{if } x_1 = 0,\ x_2 < 0, \\ \dfrac{\sqrt{2}}{2}\Big(- x_2 + \sqrt{2x_1^4 + x_2^2} - \sqrt{2}x_1^2\Big) & \text{otherwise,} \end{cases}$$

which gives us $\operatorname{dist}((x_1, x_2); \Gamma) \leq \sqrt{2}\operatorname{dist}(\Phi(x_1, x_2); \mathcal{Q}_3)$ for all $(x_1, x_2) \in \mathbb{R}^2$ and thus verifies the validity of MSCQ at any $x \in \Gamma$. It is not hard to check that

$$N_\Gamma(x) = \begin{cases} \{(0,0)\} & \text{if } x_2 > 0, \\ \{0\} \times \mathbb{R}_- & \text{if } x_2 = 0, \\ \varnothing & \text{if } x_2 < 0 \end{cases} \quad \text{and} \quad T_\Gamma(x) = \begin{cases} \mathbb{R}^2 & \text{if } x_2 > 0, \\ \mathbb{R} \times \mathbb{R}_+ & \text{if } x_2 = 0, \\ \varnothing & \text{if } x_2 < 0. \end{cases}$$

On the other hand, the direct calculation tells us that

$$T_{\text{gph}\, N_\Gamma}(\bar{x},\bar{x}^*) = \begin{cases} [\mathbb{R}\times(0,\infty)\times\{(0,0)\}]\cup[\mathbb{R}\times\{0\}\times\{0\}\times\mathbb{R}_-] & \text{if } x_2=0, \bar{x}^*=0,\\ \mathbb{R}\times\{0\}\times\{0\}\times\mathbb{R} & \text{if } x_2=0, \bar{x}^*\neq 0,\\ \mathbb{R}^2\times\{(0,0)\} & \text{if } x_2>0, \bar{x}^*=0. \end{cases} \tag{3.61}$$

Let us now apply Theorem 3.9 to calculate the tangent cone to $\text{gph}\, N_\Gamma$ and the graphical derivative of the normal cone mapping. For $\lambda=(\lambda_0,\lambda_1,\lambda_2)\in\mathbb{R}^3$ we have

$$\nabla\Phi(x)^* = \begin{bmatrix} 2\sqrt{2}x_1 & 2x_1 & 2x_1\\ 1 & \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \end{bmatrix}, \qquad \nabla^2\langle\lambda,\Phi\rangle(x) = \begin{bmatrix} 2\sqrt{2}\lambda_0+2\lambda_1+2\lambda_2 & 0\\ 0 & 0 \end{bmatrix}.$$

Consider further the following five characteristic cases:

Case 1: $\bar{x}=(0,0)$ and $\bar{x}^*=(0,0)\in N_\Gamma(\bar{x})$. In this case we have $\Phi(\bar{x})=0$, $\mathcal{H}(\bar{x};\lambda)=0$, and $K(\bar{x},\bar{x}^*)=T_\Gamma(\bar{x})=\mathbb{R}\times\mathbb{R}_+$. Applying Corollary 3.11 tells us that

$$T_{\text{gph}\, N_\Gamma}(\bar{x},\bar{x}^*) = \text{gph}\, N_{K(\bar{x},\bar{x}^*)} = \big[\mathbb{R}\times(0,\infty)\times\{(0,0)\}\big]\cup\big[\mathbb{R}\times\{0\}\times\{0\}\times\mathbb{R}_-\big],$$

$$(DN_\Gamma)(\bar{x},\bar{x}^*)\big((v_1,v_2)\big) = N_{K(\bar{x},\bar{x}^*)}\big((v_1,v_2)\big) = \begin{cases} \big\{(0,0)\big\} & \text{if } v_2>0,\\ \big\{0\big\}\times\mathbb{R}_- & \text{if } v_2=0 \end{cases}$$

for $v=(v_1,v_2)$, which agrees with the calculation in (3.61).

Case 2: $\bar{x}=(0,0)$ and $\bar{x}^*=(0,-1)$ with $K(\bar{x},\bar{x}^*)=\mathbb{R}\times\{0\}$. Take $\big((v_1,v_2),(v_1^*,v_2^*)\big)$

from the right-hand side of (3.33) and observe that for any $v := (v_1, v_2) \in K(\bar{x}, \bar{x}^*)$ it holds

$$N_{K(\bar{x},\bar{x}^*)}(v) = \{0\}\times\mathbb{R} \quad \text{and} \quad \Lambda(\bar{x}, \bar{x}^*; v) = \begin{cases} \Big\{(-1, \dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}})\Big\} & \text{if } v_1 \neq 0, \\ \big\{\lambda \in -\mathcal{Q}_3 \,\big|\, \sqrt{2}\lambda_0 + \lambda_1 - \lambda_2 = -\sqrt{2}\big\} & \text{if } v_1 = 0. \end{cases}$$

Thus Theorem 3.9 gives us the following inclusions:

(i) if $v_1 \neq 0$ and $v_2 = 0$, then

$$v^* \in \begin{bmatrix} -2\sqrt{2} + \sqrt{2} + \sqrt{2} & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} v_1 \\ 0 \end{pmatrix} + \{0\} \times \mathbb{R} = \big\{0\big\} \times \mathbb{R};$$

(ii) if $v_1 = v_2 = 0$, then there exists $\lambda \in \Lambda(\bar{x}, \bar{x}^*; v)$ such that

$$v^* \in \begin{bmatrix} 2\sqrt{2}\lambda_1 + 2\lambda_2 + 2\lambda_3 & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \{0\} \times \mathbb{R} = \big\{0\big\} \times \mathbb{R}.$$

We therefore arrive at the tangent cone formula

$$T_{\operatorname{gph} N_\Gamma}(\bar{x}, \bar{x}^*) = \big\{(v, v^*) \,\big|\, v_2 = 0 \quad \text{and} \quad v_1^* = 0\,\big\},$$

which yields for $v = (v_1, v_2)$ with $v_2 = 0$ the graphical derivative one

$$(DN_\Gamma)(\bar{x}, \bar{x}^*)\big((v_1, v_2)\big) = \big\{0\big\} \times \mathbb{R}.$$

Thus in this case we again agree with the calculation in (3.61).

Case 3: $\bar{x} = (1,0)$ and $\bar{x}^* = (0,0) \in N_\Gamma(\bar{x})$. Observe that in this case we have $\Phi(\bar{x}) \in \text{bd}\,(\mathcal{Q}_3) \setminus \{0\}$, $K(\bar{x}, \bar{x}^*) = \mathbb{R} \times \mathbb{R}_+$, and it follows from (3.23) that

$$\mathcal{H}(x;\lambda) = -\frac{\lambda_0}{\sqrt{2}} \begin{bmatrix} 0 & -2\sqrt{2} \\ -2\sqrt{2} & 0 \end{bmatrix}.$$

Applying Corollary 3.11 gives us the same formulas for $T_{\text{gph}\, N_\Gamma}$ and DN_Γ as in Case 1.

Case 4: $\bar{x} = (1,0)$ and $\bar{x}^* = (0,-1) \in N_\Gamma(\bar{x})$ with $K(\bar{x}, \bar{x}^*) = \mathbb{R} \times \{0\}$. Taking $\big((v_1, v_2), (v_1^*, v_2^*)\big)$ from the right-hand side of (3.33), observe that for all $v = (v_1, v_2) \in K(\bar{x}, \bar{x}^*)$ we get $N_{K(\bar{x},\bar{x}^*)}(v) = \{0\} \times \mathbb{R}$. It is easy to check that

$$\Lambda(\bar{x}, \bar{x}^*) = \Lambda(\bar{x}, \bar{x}^*; v) = \Big\{\Big(-1, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\Big)\Big\},$$

which implies that for any $\lambda \in \Lambda(\bar{x}, \bar{x}^*; v)$ we have

$$\nabla^2\langle \lambda, \Phi\rangle(\bar{x}) + \mathcal{H}(\bar{x};\lambda) = \begin{bmatrix} -2\sqrt{2} + \sqrt{2} + \sqrt{2} & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -2\sqrt{2} \\ -2\sqrt{2} & 0 \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ -2 & 0 \end{bmatrix}.$$

Appealing to Theorem 3.9 tells us that

$$\begin{aligned} T_{\text{gph}\, N_\Gamma}(\bar{x}, \bar{x}^*) &= \Big\{\big((v_1, v_2), (v_1^*, v_2^*)\big) \,\Big|\, v_2 = 0,\ v^* \in (0, -2v_1) + \{0\} \times \mathbb{R}\Big\} \\ &= \Big\{\big((v_1, v_2), (v_1^*, v_2^*)\big) \,\Big|\, v_2 = 0,\ v_1^* = 0\Big\}, \end{aligned}$$

which readily implies that for any $v = (v_1, v_2)$ with $v_2 = 0$ we get

$$(DN_\Gamma)(\bar{x}, \bar{x}^*)\big((v_1, v_2)\big) = \Big\{0\Big\} \times \mathbb{R}.$$

Case 5: $\bar{x} = (0,1)$ and $\bar{x}^* = (0,0)$. In this case we have $K(\bar{x},\bar{x}^*) = \mathbb{R}^2$ and so $N_{K(\bar{x},\bar{x}^*)}(v) = \{(0,0)\}$ for all $v \in \mathbb{R}^2$. It is easy to see that $\Lambda(\bar{x},\bar{x}^*) = (\sqrt{2},-1,1)\mathbb{R}_-$, which tells us that the Lagrange multipliers set has the representation in (LMS3) of Proposition 3.1. Employing again Corollary 3.11 ensures the validity of the relationships

$$T_{\text{gph}\, N_\Gamma}(\bar{x},\bar{x}^*) = \text{gph}\, N_{K(\bar{x},\bar{x}^*)} = \mathbb{R}^2 \times \big\{(0,0)\big\},$$

and therefore we arrive at the graphical derivative formula

$$(DN_\Gamma)(\bar{x},\bar{x}^*)((v_1,v_2)) = \big\{(0,0)\big\}, \quad v \in \mathbb{R}^2,$$

which illustrates the applicability of Theorem 3.9 under the imposed MSCQ condition. Since the set of Lagrange multipliers is *unbounded* in some cases above, both *metric regularity* (which equivalent to the Robinson constraint qualification characterizing the boundedness of Lagrange multipliers) and nondegeneracy conditions *fail* in this example. This completes our considerations in this example.

3.3 Application to Isolated Calmness

In this section, we provide an application of Theorem 3.9 to an important stability property well recognized in variational analysis and optimization; see, e.g., [9, 11, 40] and the references therein. Recall that a mapping $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is said to be *isolatedly calm* at $(\bar{x},\bar{y}) \in \text{gph}\, F$ if there exist a constant $\ell \geq 0$ and neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

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

In what follows we apply the graphical derivative formula established above to characterize the isolated calmness property of the *parametric variational system*

$$S(p)=\big\{x\in\mathbb{R}^n\,\big|\,p\in f(x)+N_\Gamma(x)\big\} \tag{3.62}$$

generated by the the ice-cream cone $\mathcal{Q}\subset\mathbb{R}^{m+1}$ via (1.3), where $f\colon\mathbb{R}^n\to\mathbb{R}^n$ is a differentiable mapping. The following theorem provides a complete characterization of the isolated calmness of the variational system (3.62) entirely via its given data.

Theorem 3.14. (characterization of isolated calmness for ice-cream variational systems). *Let* $(\bar p,\bar x)\in\operatorname{gph}S$ *with* S *taken from* (3.62). *In addition to the standing assumptions on* Γ *from* (1.3) *and the MSCQ condition of Theorem* 3.9*, suppose that* f *is Fréchet differentiable at* $\bar x\in\Gamma$*. Then* S *enjoys the isolated calmness property at* $(\bar p,\bar x)$ *if and only if*

$$\left\{\begin{array}{l} 0\in\nabla f(\bar x)v+\big(\nabla^2\langle\lambda,\Phi\rangle(\bar x)+\mathcal{H}(\bar x;\lambda)\big)v+N_{K(\bar x,\bar p-f(\bar x))}(v)\\ \lambda\in\Lambda\big(\bar x,\bar p-f(\bar x);v\big)\cap\kappa\,\|\bar p-f(\bar x)\|\mathbb{B}\end{array}\right.\Longrightarrow v=0, \tag{3.63}$$

where $\kappa>0$ *is the metric subregularity constant of the mapping* $x\mapsto\Phi(x)-\mathcal{Q}$ *at* $(\bar x,0)$.

Proof. We invoke a graphical derivative characterization of the isolated calmness property (3.3) for arbitrary closed-graph multifunctions written as

$$DF(\bar x,\bar y)(0)=\{0\}. \tag{3.64}$$

This result goes back to Rockafellar [56] although it was not explicitly formulated in [56]; see [11, Theorem 4C.1] with the commentaries. It easily follows from the Fréchet differentiability of f at $\bar x$ and the structure of S in (3.62) that $v\in DS(\bar p,\bar x)(u)$ if and only if

$u \in \nabla f(\bar{x})v + (DN_\Gamma)(\bar{x}, \bar{p} - f(\bar{x}))(v)$. Using now the calmness criterion (3.64) and substituting there the graphical derivative formula from Corollary 3.10, we arrive at the claimed characterization (3.63). $\triangle$

Finally in this section, we present a numerical example of the ice-cream variational system (3.62) where the application of Theorem 3.14 allows us to reveal that the isolated calmness property holds at some feasible points while failing at other ones.

Example 3.15. (verification of isolated calmness). Consider the variational system (3.62) with the mapping $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$f(x) := \big(x_1, x_2^2\big) \quad \text{for} \quad x = (x_1, x_2)$$

and the constraint set Γ taken from Example 3.13. We examine the following cases:

Case 1: $\bar{x} = (0,0)$ and $\bar{p} = f(\bar{x}) = (0,0)$. In this case we have

$$\nabla f(\bar{x})v + DN_\Gamma\big(\bar{x}, \bar{p} - f(\bar{x})\big)\big((v_1, v_2)\big) = \begin{pmatrix} v_1 \\ 0 \end{pmatrix} + \begin{cases} \{(0,0)\} & \text{if } v_2 > 0, \\ \{0\} \times \mathbb{R}_- & \text{if } \ v_2 = 0. \end{cases}$$

Invoking the corresponding calculations from Example 3.13 shows implication (3.63) does not hold. Thus the isolated calmness of (3.62) fails at this point $(\bar{p}, \bar{x})$.

Case 2: $\bar{x} = (0,0)$ and $\bar{p} = (0,-1)$. In this case we have $\bar{p} - f(\bar{x}) = (0,-1)$ and

$$\nabla f(\bar{x})v + DN_\Gamma\big(\bar{x}, \bar{p} - f(\bar{x})\big)\big((v_1, v_2)\big) = \begin{pmatrix} v_1 \\ 0 \end{pmatrix} + \{0\} \times \mathbb{R} \quad \text{if } \ v_2 = 0.$$

It is clear that implication (3.63) holds for this case, and so does the isolated calmness at $(\bar{p}, \bar{x})$.

Case 3: $\bar{x} = (1, 0)$ and $\bar{p} = f(\bar{x}) = (1, 0)$. The right-hand side of the inclusion in (3.63) for this case is the same as that in Case 1. Therefore we come up with the same conclusion that isolated calmness does not hold at this point.

Case 4: $\bar{x} = (1, 0)$ and $\bar{p} = (1, -1)$. We get the validity of the same implication (3.63) as that in Case 2 and therefore justify the isolated calmness of (3.62) at the point under consideration.

Case 5: $\bar{x} = (0, 1)$ and $\bar{p} = f(\bar{x}) = (0, 1)$. Then the right-hand side of the inclusion in (3.63) reduces to $(v_1, 2v_2) + \{(0, 0)\}$. It is easy to see that implication (3.63) holds, which therefore justifies the isolated calmness of (3.62) in this case.

CHAPTER 4 SECOND-ORDER VARIATIONAL ANALYSIS OF AUGMENTED LAGRANGIANS

This chapter aims at providing characterizations of the second-order growth condition for the penalized problem (1.8), and thus a second-order sufficient condition for strict local minimizers of this problem. Our main device to obtain such characterizations is the *second subderivative* (2.11). As observed by Rockafellar [58, Theorem 2.2], the second-order growth condition for a proper extended-real-valued function can be characterized via its second subderivative. Using this rather simple albeit powerful result for the penalized problem (1.8) requires the *calculation* of the second subderivative of the augmented Lagrangian (1.7).

4.1 Twice Epi-Differentiability of Augmented Lagrangians

We begin with the following assertion that calculates the second subderivative of the Moreau envelope of a convex function. Given $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\rho > 0$, recall that the *Moreau envelope* of φ relative to ρ is defined by the infimal convolution

$$(e_{1/\rho}\varphi)(x) := \inf_{w} \Big\{\varphi(w) + \tfrac{\rho}{2}\|w - x\|^2\Big\}, \quad x \in \mathbb{R}^n. \tag{4.1}$$

Proposition 4.1 (second subderivatives of Moreau envelopes). *Let $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper, lower semicontinuous (l.s.c.), and convex function, and let $\bar{v} \in \partial\varphi(\bar{x})$. If φ is twice epi-differentiable at $\bar{x}$ for $\bar{v}$, then for any $\rho > 0$ the Moreau envelope $e_{1/\rho}\varphi$ is properly twice epi-differentiable at $\bar{x} + \rho^{-1}\bar{v}$ for $\bar{v}$ and its second subderivative at this point is calculated by*

$$\mathrm{d}^2(e_{1/\rho}\varphi)(\bar{x} + \rho^{-1}\bar{v}, \bar{v})(w) = e_{1/2\rho}\big(\mathrm{d}^2\varphi(\bar{x}, \bar{v})\big)(w) \ \textit{for all}\ \ w \in \mathbb{R}^n. \tag{4.2}$$

Proof. Fix $\rho > 0$. It follows from [60, Theorem 11.23] that

$$(e_{1/\rho}\varphi)^*(z) = \varphi^*(z) + \tfrac{1}{2}\rho^{-1}\|z\|^2 \ \text{ for all } \ z \in \mathbb{R}^n. \tag{4.3}$$

Because φ is proper, convex, and twice epi-differentiable at $\bar{x}$ for $\bar{v}$, we deduce from [60, Proposition 13.20] that $\mathrm{d}^2\varphi(\bar{x}, \bar{v})$ is proper, l.s.c., and convex as well. Furthermore, it follows from [60, Theorem 13.21] that the proper twice epi-differentiability of φ at $\bar{x}$ for $\bar{v}$ yields this property for the conjugate function φ^*. Employing [60, Proposition 12.19] tells us that the inclusion $\bar{v} \in \partial\varphi(\bar{x})$ ensures that $\nabla(e_{1/\rho}\varphi)(\bar{x} + \rho^{-1}\bar{v}) = \bar{v}$. Combining these facts with (4.3) and the sum rule for twice epi-differentiability from [60, Exercise 13.18] implies that $(e_{1/\rho}\varphi)^*$ is properly twice epi-differentiable at $\bar{v}$ for $\bar{x} + \rho^{-1}\bar{v}$ and that its second subderivative is given by

$$\mathrm{d}^2(e_{1/\rho}\varphi)^*(\bar{x} + \rho^{-1}\bar{v}, \bar{v})(w) = \mathrm{d}^2\varphi^*(\bar{v}, \bar{x})(w) + \rho^{-1}\|w\|^2 \ \text{ for all } w \in \mathbb{R}^n. \tag{4.4}$$

This together with [60, Theorem 13.21] yields the proper twice epi-differentiability of $(e_{1/\rho}\varphi)^*$ at $\bar{x} + \rho^{-1}\bar{v}$ for $\bar{v}$. Thus the second subderivative of the latter function can be calculated by

$$\begin{aligned}
\tfrac{1}{2}\mathrm{d}^2(e_{1/\rho}\varphi)(\bar{x} + \rho^{-1}\bar{v}, \bar{v})(w) &= \Big(\tfrac{1}{2}\mathrm{d}^2(e_{1/\rho}\varphi)^*(\bar{x} + \rho^{-1}\bar{v}, \bar{v})\Big)^*(w) \\
&= \inf_{u \in \mathbb{R}^n} \Big\{\big(\tfrac{1}{2}\mathrm{d}^2\varphi^*(\bar{v}, \bar{x})\big)^*(u) + \tfrac{1}{2}\rho\|u - w\|^2\Big\} \\
&= \inf_{u \in \mathbb{R}^n} \Big\{\tfrac{1}{2}\mathrm{d}^2\varphi(\bar{x}, \bar{v})(u) + \tfrac{1}{2}\rho\|u - w\|^2\Big\},
\end{aligned}$$

where the first equality comes from [60, Theorem 13.21], the second one is due to (4.4) and [3, Proposition 14.1(i)], and the last equality follows from [60, Theorem 13.21]. This readily justifies the claimed formula for the second subderivative of $e_{1/\rho}\varphi$ at $\bar{x}+\rho^{-1}\bar{v}$ for $\bar{v}$. □

The second subderivative of the Moreau envelope for general prox-regular functions was established in [60, Exercise 13.45]. However, there are several differences between the latter result and Proposition 4.1. Firstly, the result of [60] was obtained for $\bar{v}=0$ and $\rho>0$ sufficiently large. Our result does not demand neither of these requirements. Secondly, there is the coefficient $1/2$ in [60, Exercise 13.45], which does not appear in (4.2). The price for a nicer formula, however, is confining ourselves to the framework to convex functions.

Proposition 4.1 allows us to obtain the required calculation of the second subderivative of the augmented Lagrangian (1.7).

Theorem 4.2 (second epi-derivatives of augmented Lagrangians). *Let* $(\bar{x},\bar{\lambda})$ *be a solution to the KKT system* (1.4). *Then for any* $\rho>0$ *the function* $x\mapsto\mathcal{L}(x,\bar{\lambda},\rho)$ *defined via the augmented Lagrangian* (1.7) *is twice epi-differentiable at* $\bar{x}$ *for* 0 *and its second subderivative is given by*

$$d_x^2\mathcal{L}\big((\bar{x},\bar{\lambda},\rho),0\big)(w)=\big\langle w,\nabla_{xx}^2L(\bar{x},\bar{\lambda})w\big\rangle+Q_{\bar{x},\bar{\lambda},\rho}(\nabla\Phi(\bar{x})w)+\rho\,\mathrm{dist}^2\big(\nabla\Phi(\bar{x})w;K_{\mathcal{Q}}(\Phi(\bar{x}),\bar{\lambda})\big),\tag{4.5}$$

for $w \in \mathbb{R}^n$, where the quadratic function $Q_{\bar{x},\bar{\lambda},\rho}: \mathbb{R}^{m+1} \to \mathbb{R}$ is defined by

$$Q_{\bar{x},\bar{\lambda},\rho}(v) := \begin{cases} 0 & \textit{if } \Phi(\bar{x}) \in (\operatorname{int} \mathcal{Q}) \cup \{0\} \textit{ or } \bar{\lambda} = 0, \\ \dfrac{\rho\|\bar{\lambda}\|}{\rho\|\Phi(\bar{x})\| + \|\bar{\lambda}\|}\Big(\|v_r\|^2 - \dfrac{\langle \bar{\lambda}_r, v_r\rangle^2}{\|\bar{\lambda}_r\|^2}\Big) & \textit{if } \Phi(\bar{x}) \in (\operatorname{bd} \mathcal{Q}) \setminus \{0\} \textit{ and } \bar{\lambda} \neq 0. \end{cases} \tag{4.6}$$

Proof. Since $(\bar{x}, \bar{\lambda})$ is a solution to the KKT system (1.4), we have $\nabla_x \mathcal{L}(\bar{x}, \bar{\lambda}, \rho) = 0$, where $\nabla_x \mathcal{L}$ is calculated in (1.9). The twice epi-differentiability of the function $x \mapsto \mathcal{L}(x, \bar{\lambda}, \rho)$ at $\bar{x}$ for $\bar{v} = 0$ follows from [35, Theorem 8.3(i)]. Let us proceed with the second subderivative calculation for the latter function. If either $\Phi(\bar{x}) \in (\operatorname{int} \mathcal{Q}) \cup \{0\}$ or $\bar{\lambda} = 0$, then by (2.13) we get

$$\mathrm{d}^2 \delta_{\mathcal{Q}}\big(\Phi(\bar{x}), \bar{\lambda}\big)(w) = \delta_{K_{\mathcal{Q}}(\Phi(\bar{x}),\bar{\lambda})}(w) \quad \text{whenever} \quad w \in \mathbb{R}^{m+1}.$$

Employing again [35, Theorem 8.3(i,iii)] and the second subderivative calculation (4.2) from Proposition 4.1 for the Moreau envelope (4.1) of $\varphi = \delta_{\mathcal{Q}}$ tells us that

$$\begin{aligned} \mathrm{d}_x^2 \mathcal{L}\big((\bar{x}, \bar{\lambda}, \rho), 0\big)(w) &= \big\langle w, \nabla_{xx}^2 L(\bar{x}, \bar{\lambda}) w\big\rangle + e_{1/2\rho}\big(\mathrm{d}^2 \delta_{\mathcal{Q}}(\Phi(\bar{x}), \bar{\lambda})\big)(w) \\ &= \big\langle w, \nabla_{xx}^2 L(\bar{x}, \bar{\lambda}) w\big\rangle + \inf_{u \in \mathbb{R}^{m+1}} \big\{\delta_{K_{\mathcal{Q}}(\Phi(\bar{x}),\bar{\lambda})}(u) + \rho\|u - \nabla\Phi(\bar{x})w\|^2\big\} \\ &= \big\langle w, \nabla_{xx}^2 L(\bar{x}, \bar{\lambda}) w\big\rangle + \rho \operatorname{dist}^2\big(\nabla\Phi(\bar{x})w; K_{\mathcal{Q}}(\Phi(\bar{x}), \bar{\lambda})\big), \end{aligned}$$

which verifies formula (4.5) with $Q_{\bar{x},\bar{\lambda},\rho}(w)$ from (4.6) in this case. Assuming next that $\Phi(\bar{x}) \in (\operatorname{bd} \mathcal{Q}) \setminus \{0\}$ and $\bar{\lambda} \neq 0$, define the function $\theta(y) := \frac{1}{2}\operatorname{dist}^2(y; \mathcal{Q})$ for $y \in \mathbb{R}^{m+1}$. It is well known that θ is continuously differentiable on $\mathbb{R}^{m+1}$ and its gradient is given by

$$\nabla\theta(y) = \Pi_{-\mathcal{Q}}(y) \quad \text{whenever} \quad y \in \mathbb{R}^{m+1}.$$

Since $\Phi(\bar{x}) \in (\operatorname{bd}\mathcal{Q}) \setminus \{0\}$ and $0 \neq \bar{\lambda} \in N_{\mathcal{Q}}(\Phi(\bar{x}))$, we get $\bar{y} \notin \mathcal{Q} \cup (-\mathcal{Q})$ with $\bar{y} = (\bar{y}_0, \bar{y}_r) := \Phi(\bar{x}) + \rho^{-1}\bar{\lambda}$. This clearly yields $\|\bar{y}_r\| > 0$, and so we arrive at

$$\nabla\theta(y) = \Pi_{-\mathcal{Q}}(y) = \frac{1}{2}\Big(1 - \frac{y_0}{\|y_r\|}\Big)\big(-\|y_r\|, y_r\big) = \frac{1}{2}\Big(y_0 - \|y_r\|, y_r - y_0\frac{y_r}{\|y_r\|}\Big)$$

for all y close to $\bar{y}$. This confirms, in particular, that θ is $\mathcal{C}^2$-smooth around $\bar{y}$ with

$$\nabla^2\theta(\bar{y}) = \nabla\Pi_{-\mathcal{Q}}(\bar{y}) = \frac{1}{2}\begin{pmatrix} 1 & -\dfrac{\bar{y}_r^*}{\|\bar{y}_r\|} \\ -\dfrac{\bar{y}_r}{\|\bar{y}_r\|} & I_m - \dfrac{\bar{y}_0}{\|\bar{y}_r\|}I_m + \dfrac{\bar{y}_0}{\|\bar{y}_r\|}\dfrac{\bar{y}_r\bar{y}_r^*}{\|\bar{y}_r\|^2} \end{pmatrix}, \tag{4.7}$$

where I_m the $m \times m$ identity matrix, and where $\bar{y}_r^*$ stands for the corresponding vector row. Since $\Phi(\bar{x}) \in (\operatorname{bd} Q) \setminus \{0\}$ and $\bar{\lambda} \in N_{\mathcal{Q}}(\Phi(\bar{x})) \setminus \{0\}$, it follows that $\bar{\lambda} = t\widetilde{\Phi}(\bar{x}) = t\,(-\Phi_0(\bar{x}), \Phi_r(\bar{x}))$ for some $t > 0$ and $\bar{\lambda}_0 = -\|\bar{\lambda}_r\|$. Thus we have

$$\bar{y} = \Phi(\bar{x}) + \rho^{-1}\bar{\lambda} = \frac{1}{t}\big(-\bar{\lambda}_0, \bar{\lambda}_r\big) + \frac{1}{\rho}\big(\bar{\lambda}_0, \bar{\lambda}_r\big) = \Big(\frac{t-\rho}{t\rho}\bar{\lambda}_0, \frac{t+\rho}{t\rho}\bar{\lambda}_r\Big) = \Big(\frac{\rho-t}{t\rho}\|\bar{\lambda}_r\|, \frac{\rho+t}{t\rho}\bar{\lambda}_r\Big).$$

Plugging the latter into (4.7) gives us the gradient formula

$$\nabla\Pi_{-\mathcal{Q}}(\bar{y}) = \frac{1}{2}\begin{pmatrix} 1 & -\dfrac{\bar{\lambda}_r^*}{\|\bar{\lambda}_r\|} \\ -\dfrac{\bar{\lambda}_r}{\|\bar{\lambda}_r\|} & \dfrac{2t}{\rho+t}I_m + \dfrac{\rho-t}{\rho+t}\dfrac{\bar{\lambda}_r\bar{\lambda}_r^*}{\|\bar{\lambda}_r\|^2} \end{pmatrix},$$

which being combined with (4.7) and $\bar\lambda_0 = -\|\bar\lambda_r\|$ results in

$$
\begin{aligned}
\big\langle \nabla^2\theta(\bar y)v, v\big\rangle &= \frac{1}{2}\left(v_0^2 - \frac{2v_0}{\|\bar\lambda_r\|}\langle\bar\lambda_r, v_r\rangle + \frac{2t}{\rho+t}\|v_r\|^2 + \frac{\rho-t}{\rho+t}\frac{\langle\bar\lambda_r, v_r\rangle^2}{\|\bar\lambda_r\|^2}\right)\\
&= \frac{1}{2}\left[v_0^2 - 2v_0\frac{\langle\bar\lambda_r, v_r\rangle}{\|\bar\lambda_r\|} + \left(\frac{\langle\bar\lambda_r, v_r\rangle}{\|\bar\lambda_r\|}\right)^2\right] + \frac{t}{\rho+t}\left(\|v_r\|^2 - \frac{\langle\bar\lambda_r, v_r\rangle^2}{\|\bar\lambda_r\|^2}\right)\\
&= \frac{(\bar\lambda_0 v_0)^2 + 2\widehat{\bar\lambda_0} v_0\langle\bar\lambda_r, v_r\rangle + \langle\bar\lambda_r, v_r\rangle^2}{2\|\bar\lambda_r\|^2} + \frac{t}{\rho+t}\left(\|v_r\|^2 - \frac{\langle\bar\lambda_r, v_r\rangle^2}{\|\bar\lambda_r\|^2}\right)\\
&= \frac{\langle\bar\lambda, v\rangle^2}{\|\bar\lambda\|^2} + \frac{t}{\rho+t}\left(\|v_r\|^2 - \frac{\langle\bar\lambda_r, v_r\rangle^2}{\|\bar\lambda_r\|^2}\right)\\
&= \operatorname{dist}^2\big(v; K_{\mathcal Q}(\Phi(\bar x), \bar\lambda)\big) + \frac{\|\bar\lambda\|}{\rho\|\Phi(\bar x)\| + \|\bar\lambda\|}\left(\|v_r\|^2 - \frac{\langle\bar\lambda_r, v_r\rangle^2}{\|\bar\lambda_r\|^2}\right), \qquad (4.8)
\end{aligned}
$$

for all $v = (v_0, v_r) \in \mathbb{R}^{m+1}$. In the last equality we use the facts that $K_{\mathcal Q}(\Phi(\bar x), \bar\lambda) = \{\bar\lambda\}^\perp$ and $\|\bar\lambda\| = t\|\Phi(\bar x)\|$. It follows from the twice differentiability of θ at $\bar y$ that the function $x \mapsto \mathcal{L}(x, \bar\lambda, \rho)$ is twice differentiable at $\bar x$ with its second subderivative computed by

$$
\mathrm{d}_x^2\mathcal{L}((\bar x, \bar\lambda, \rho), 0)(w) = \big\langle \nabla_{xx}^2\mathcal{L}(\bar x, \bar\lambda, \rho)w, w\big\rangle = \big\langle w, \nabla_{xx}^2 L(\bar x, \bar\lambda)w\big\rangle + \rho\big\langle\nabla^2\theta(\bar y)v, v\big\rangle
$$

with $v = \nabla\Phi(\bar x)w$. Combining this and (4.8) gives us the claimed second subderivative formula in this case and thus finishes the proof of the theorem. □

4.2 Second-Order Growth Conditions of Augmented Lagrangians

We recall first in this section the following result from [35, Proposition 7.3] justifying the ability of the second subderivertive (2.11) to characterize the second-order growth condition for SOCPs.

Proposition 4.3 (SOSC yields second-order growth). *Let* $(\bar x, \bar\lambda) \in \mathbb{R}^n \times \mathbb{R}^{m+1}$ *be a solution* *to the KKT system* (1.4)*, and let the second-order sufficient condition*

$$\big\langle \nabla^2_{xx} L(\bar{x},\bar{\lambda})w, w\big\rangle + \mathrm{d}^2\delta_{\mathcal{Q}}\big(\Phi(\bar{x}),\bar{\lambda}\big)\big(\nabla\Phi(\bar{x})w\big) > 0 \quad \textit{for all } \ w \in \mathbb{R}^n \setminus \{0\}. \tag{4.9}$$

hold. Then there exist positive numbers ℓ, γ *such that the second-order growth condition*

$$f(x) \geq f(\bar{x}) + \tfrac{\ell}{2}\|x-\bar{x}\|^2 \quad \textit{for all } \ x \in \mathbb{B}_\gamma(\bar{x}) \textit{ with } \ \Phi(x) \in \mathcal{Q} \tag{4.10}$$

is satisfied for the second-order cone program (1.1).

Observe that the presented SOSC (4.9) is equivalent to the second-order conditions used for SOCPs in Proposition 3.8 (for the case of $\Phi(\bar{x}) = 0$) and in other publications [5, 30]. This indeed follows from the second subderivative formula (2.13). Note also that SOSC (4.9) is stronger than the conventional second-order sufficient condition for (1.1), the latter requires the supremum of the quadratic term in (4.9) over all the Lagrange multipliers from (1.5) be positive. This stronger condition is in fact equivalent to the second-order growth (4.10) under an appropriate constraint qualification; see [35, Theorem 7.2]. Let us now provide an equivalent version of SOSC (4.9) that is often used in what follows.

Remark 4.4 (equivalent version of SOSC). It is not hard to check that the formulated SOSC (4.9) amounts to saying that there exists a number $\bar{\ell} > 0$ such that we have

$$\big\langle \nabla^2_{xx} L(\bar{x},\bar{\lambda})w, w\big\rangle + \mathrm{d}^2\delta_{\mathcal{Q}}\big(\Phi(\bar{x}),\bar{\lambda}\big)\big(\nabla\Phi(\bar{x})w\big) \geq \bar{\ell}\,\|w\|^2 \quad \text{for all } \ w \in \mathbb{R}^n. \tag{4.11}$$

Conversely, the fulfillment of (4.11) at $(\bar{x},\bar{\lambda})$ ensures that for any $\ell \in (0,\bar{\ell})$ there exists a positive number γ such that the second-order growth condition (4.10) is satisfied at $\bar{x}$.

Now we are ready to establish complete pointwise characterizations of the second-order growth condition for the penalized problem (1.8) in terms of SOSC (4.9) and the second subderivative of the augmented Lagrangian (1.7).

Theorem 4.5 (characterizations of second-order growth condition for augmented Lagrangians). *Let* $(\bar{x}, \bar{\lambda})$ *be a solution to the KKT system* (1.4) *for SOCP* (1.1). *Then the following assertions are equivalent:*

(i) *The second-order sufficient condition* (4.9) *holds at* $(\bar{x}, \bar{\lambda})$.

(ii) *There exists a constant* $\rho_{\bar{\lambda}} > 0$ *such that for any* $\rho \geq \rho_{\bar{\lambda}}$ *we have*

$$\mathrm{d}_x^2 \mathcal{L}((\bar{x}, \bar{\lambda}, \rho), 0)(w) > 0 \ \textit{ whenever } \ w \in \mathbb{R}^n \setminus \{0\}. \tag{4.12}$$

(iii) *There exist positive constants* $\rho_{\bar{\lambda}}, \gamma_{\bar{\lambda}}$, *and* $\ell_{\bar{\lambda}}$ *such that for any* $\rho \geq \rho_{\bar{\lambda}}$ *we have*

$$\mathcal{L}(x, \bar{\lambda}, \rho) \geq f(\bar{x}) + \ell_{\bar{\lambda}} \|x - \bar{x}\|^2 \ \textit{ for all } \ x \in \mathbb{B}_{\gamma_{\bar{\lambda}}}(\bar{x}). \tag{4.13}$$

Proof. Since $(\bar{x}, \bar{\lambda})$ is a solution to the KKT system (1.4), for all $\rho > 0$ we have $\mathcal{L}(\bar{x}, \bar{\lambda}, \rho) = f(\bar{x})$ and $\nabla_x \mathcal{L}(\bar{x}, \bar{\lambda}, \rho) = 0$. Assuming that (ii) holds, deduce from [60, Theorem 13.24] that the second-order growth condition (4.13) for $\rho = \rho_{\bar{\lambda}}$ follows from (4.12) with the same constant ρ. Appealing now to Proposition 1.1(i) tells us that

$$\mathcal{L}(x, \lambda, \rho) \geq \mathcal{L}(x, \lambda, \rho_{\bar{\lambda}}) \ \text{ whenever } \ \rho \geq \rho_{\bar{\lambda}}.$$

This combined with (4.13) for $\rho = \rho_{\bar{\lambda}}$ justifies the second-order growth condition for any $\rho \geq \rho_{\bar{\lambda}}$ and thus verifies (iii). The opposite implication (iii) $\Longrightarrow$ (ii) follows directly from the definition of the second subderivative.

Assume now that (ii) holds and let $\rho \geq \rho_{\bar{\lambda}}$. To justify (i), pick $w \in \mathbb{R}^n \setminus \{0\}$ with $v := \nabla\Phi(\bar{x})w \in K_{\mathcal{Q}}(\Phi(\bar{x}), \bar{\lambda})$. We now show that

$$\mathrm{d}^2\delta_{\mathcal{Q}}\big(\Phi(\bar{x}), \bar{\lambda}\big)\big(v\big) \geq Q_{\bar{x},\bar{\lambda},\rho}(v) + \rho\, \mathrm{dist}^2\big(v; K_{\mathcal{Q}}(\Phi(\bar{x}), \bar{\lambda})\big) \tag{4.14}$$

for all $\rho > 0$. If either $\Phi(\bar{x}) \in (\mathrm{int}\, \mathcal{Q}) \cup \{0\}$ or $\bar{\lambda} = 0$, we get from (2.13) and (4.6) that

$$\mathrm{d}^2\delta_{\mathcal{Q}}\big(\Phi(\bar{x}), \bar{\lambda}\big)\big(v\big) = Q_{\bar{x},\bar{\lambda},\rho}(v) + \rho\, \mathrm{dist}^2\big(v; K_{\mathcal{Q}}(\Phi(\bar{x}), \bar{\lambda})\big) = 0 \tag{4.15}$$

where the last equality comes from $\nabla\Phi(\bar{x})w \in K_{\mathcal{Q}}(\Phi(\bar{x}), \bar{\lambda})$. Otherwise, if $\Phi(\bar{x}) \in (\mathrm{bd}\, \mathcal{Q}) \setminus \{0\}$ and $\bar{\lambda} \neq 0$, then we get that $K_{\mathcal{Q}}(\Phi(\bar{x}), \bar{\lambda}) = \{\bar{\lambda}\}^{\perp}$. It follows from $v \in K_{\mathcal{Q}}(\Phi(\bar{x}), \bar{\lambda})$ and $\bar{\lambda} \in \mathrm{bd}\,(-\mathcal{Q}) \setminus \{0\}$ that

$$\langle v_r, \bar{\lambda}_r \rangle^2 = v_0^2 \bar{\lambda}_0^2 = v_0^2 \|\bar{\lambda}_r\|^2 \quad \text{and} \quad \|v_r\|^2 \geq v_0^2. \tag{4.16}$$

We then deduce from (2.13) and (4.6) that

$$\begin{aligned}
& \mathrm{d}^2\delta_{\mathcal{Q}}\big(\Phi(\bar{x}),\bar{\lambda}\big)\big(v\big) - Q_{\bar{x},\bar{\lambda},\rho}(v) - \rho\,\mathrm{dist}^2\big(v; K_{\mathcal{Q}}(\Phi(\bar{x}),\bar{\lambda})\big) \\
= \;& \frac{\|\bar{\lambda}\|}{\|\Phi(\bar{x})\|}\big(\|v_r\|^2 - v_0^2\big) - \frac{\rho\|\bar{\lambda}\|}{\rho\|\Phi(\bar{x})\| + \|\bar{\lambda}\|}\Big(\|v_r\|^2 - \frac{\langle \bar{\lambda}_r, v_r\rangle^2}{\|\bar{\lambda}_r\|^2}\Big) \\
= \;& \frac{\|\bar{\lambda}\|^2}{\|\Phi(\bar{x})\|(\rho\|\Phi(\bar{x})\| + \|\bar{\lambda}\|)}\big(\|v_r\|^2 - v_0^2\big) - \frac{\rho\|\bar{\lambda}\|}{\rho\|\Phi(\bar{x})\| + \|\bar{\lambda}\|}\Big(v_0^2 - \frac{\langle \bar{\lambda}_r, v_r\rangle^2}{\|\bar{\lambda}_{i,r}\|^2}\Big) \\
\geq \;& 0, \qquad (4.17)
\end{aligned}$$

where the last inequality is due to estimates in (4.16). Thus, we justify (4.14) for $v \in K_{\mathcal{Q}}(\Phi(\bar{x}),\bar{\lambda})$. Note that (4.14) is obvious if $v \notin K_{\mathcal{Q}}(\Phi(\bar{x}),\bar{\lambda}) = \mathrm{dom}\,\mathrm{d}^2\delta_{\mathcal{Q}}\big(\Phi(\bar{x}),\bar{\lambda}\big)$. Referring to (4.5), we get SOSC (4.9) from (4.12) and (4.14). Thus we are done with (ii) $\Longrightarrow$ (i).

To complete the proof of the theorem, it remains to verify implication (i) $\Longrightarrow$ (ii). Since the second subderivative is positive homogenous of degree 2, to prove (4.12) it is neccesary and sufficient to verify the condition: for all $\rho > 0$ sufficiently large we get

$$d^2\mathcal{L}\big((\bar{x},\bar{\lambda};\rho)|0\big)(w) > 0 \quad \text{whenever} \;\; w \in \mathbb{S}. \qquad (4.18)$$

Assuming that (i) holds, we first justify the claim that (4.18) holds for all $w \in \mathbb{S}$ with $v := \nabla\Phi(\bar{x})w \in K_{\mathcal{Q}}(\Phi(\bar{x}),\bar{\lambda})$. It is worth mentioning that the quadratic function (in w) on the left-hand side of SOSC (4.9) must attain its minimum value on the compact set $\mathbb{S}$. Let ℓ_0 denote such a value, then by (4.9) we have $\ell_0 > 0$. We now show that

$$Q_{\bar{x},\bar{\lambda},\rho}(v) + \rho\,\mathrm{dist}^2\big(v; K_{\mathcal{Q}}(\Phi(\bar{x}),\bar{\lambda})\big) \geq \mathrm{d}^2\delta_{\mathcal{Q}}\big(\Phi(\bar{x}),\bar{\lambda}\big)\big(v\big) - \frac{\ell_0}{2} \qquad (4.19)$$

for all $\rho > 0$ sufficiently large. In the above proof of the implication (ii) $\Longrightarrow$ (i), it is proved that the latter holds for all $\rho > 0$ whenever $\Phi(\bar{x}) \in (\operatorname{int} \mathcal{Q}) \cup \{0\}$ or $\bar{\lambda} = 0$, see (4.15). Turning now to the remaining case with $\Phi(\bar{x}) \in (\operatorname{bd} \mathcal{Q}) \setminus \{0\}$ and $\bar{\lambda} \neq 0$. Recall from (4.16) and (4.17) that

$$\begin{aligned} & Q_{\bar{x},\bar{\lambda},\rho}(v) + \rho \operatorname{dist}^2\big(v; K_{\mathcal{Q}}(\Phi(\bar{x}),\bar{\lambda})\big) - \mathrm{d}^2\delta_{\mathcal{Q}}\big(\Phi(\bar{x}),\bar{\lambda}\big)(v) \\ = \ & -\frac{\|\bar{\lambda}\|^2}{\|\Phi(\bar{x})\|(\rho\|\Phi(\bar{x})\| + \|\bar{\lambda}\|)}\big(\|v_r\|^2 - v_0^2\big) \geq -\frac{\|\bar{\lambda}\|^2\|v\|^2}{\|\Phi(\bar{x})\|(\rho\|\Phi(\bar{x})\| + \|\bar{\lambda}\|)} \\ \geq \ & -\frac{\|\bar{\lambda}\|^2\|\nabla\Phi(\bar{x})\|^2}{\|\Phi(\bar{x})\|(\rho\|\Phi(\bar{x})\| + \|\bar{\lambda}\|)}, \end{aligned}$$

where, in the last equality, we use the fact that $v = \nabla\Phi(\bar{x})w$ with $\|w\| = 1$. Pick $\varrho_0 > 0$ such that the condition

$$\frac{\big(\|\nabla\Phi(\bar{x})\| \cdot \|\bar{\lambda}\|\big)^2}{\|\Phi(\bar{x})\|(\rho\|\Phi(\bar{x})\| + \|\bar{\lambda}\|)} \leq \frac{\ell_0}{2} \quad \text{for all } \rho \geq \varrho_0 \tag{4.20}$$

is fulfilled. Then (4.19) is satisfied for the case with $\Phi(\bar{x}) \in (\operatorname{bd} \mathcal{Q}) \setminus \{0\}$ and $\bar{\lambda} \neq 0$, and therefore, for all possible position of $\Phi(\bar{x}) \in \mathcal{Q}$ and $\bar{\lambda} \in N_{\mathcal{Q}}(\Phi(\bar{x}))$ whenever $\rho \geq \varrho_0$. Referring to (4.5) and SOSC (4.9), we get by (4.19) that

$$\mathrm{d}^2\mathcal{L}\big((\bar{x},\bar{\lambda};\rho)|0\big)(w) \geq \big\langle w, \nabla^2_{xx}L(\bar{x},\bar{\lambda})w\big\rangle + \mathrm{d}^2\delta_{\mathcal{Q}}\big(\Phi(\bar{x}),\bar{\lambda}\big)(v) - \frac{\ell_0}{2} \geq \frac{\ell_0}{2} \tag{4.21}$$

for all $w \in \mathbb{S}$ with $v = \nabla\Phi_i(\bar{x})w \in K_{\mathcal{Q}}(\Phi(\bar{x}),\bar{\lambda})$ and for all $\rho \geq \varrho_0$, which just completes the verification of (4.18) for such w.

Next, we decompose the unit sphere into the two pieces:

$$\mathbb{S}_+ := \big\{ w \in \mathbb{S} \,\big|\, \big\langle w, \nabla^2_{xx} L(\bar{x}, \bar{\lambda}) w \big\rangle + Q_{\bar{x},\bar{\lambda},\varrho_0}(v) > 0 \big\}$$

and

$$\mathbb{S}_- := \big\{ w \in \mathbb{S} \,\big|\, \big\langle w, \nabla^2_{xx} L(\bar{x}, \bar{\lambda}) w \big\rangle + Q_{\bar{x},\bar{\lambda},\varrho_0}(v) \le 0 \big\},$$

where ϱ_0 is taken from (4.20). We see from (4.6) that the function $\rho \mapsto Q_{\bar{x},\bar{\lambda},\rho}(v)$ is nondecreasing on $\mathbb{R}_+$, then by (4.5) the estimate (4.18) must be satisfied for any $w \in \mathbb{S}_+$ and any $\rho \ge \varrho_0$. Define the function $\vartheta \colon \mathbb{S}_- \to \mathbb{R}$ by

$$\vartheta(w) := -\frac{\big\langle \nabla^2_{xx} L(\bar{x}, \bar{\lambda}) w, w \big\rangle}{\mathrm{dist}^2\big(\nabla \Phi(\bar{x}) w; K_{\mathcal{Q}}(\Phi(\bar{x}), \bar{\lambda})\big)}, \quad w \in \mathbb{S}_-.$$

Picking an arbitrary vector $w \in \mathbb{S}_-$, we conclude from the just proved claim that $\nabla \Phi(\bar{x}) w \notin K_{\mathcal{Q}}(\Phi(\bar{x}), \bar{\lambda})$. This confirms that $\mathrm{dist}\big(\nabla \Phi(\bar{x}) w; K_{\mathcal{Q}}(\Phi(\bar{x}), \bar{\lambda})\big) > 0$. Also we get by (4.6) that $Q_{\bar{x},\bar{\lambda},\varrho_0}(v)$ is always positive, then $w \in \mathbb{S}_-$ implies that $\big\langle \nabla^2_{xx} L(\bar{x}, \bar{\lambda}) w, w \big\rangle \le 0$. Thus the function ϑ is continuous and nonnegative on the compact set $\mathbb{S}_-$, and hence its maximum value over this set, denoted by ϱ_1, is finite and nonnegative. This demonstrates that for any $\rho > \varrho_1$ we have the estimate

$$\big\langle \nabla^2_{xx} L(\bar{x}, \bar{\lambda}) w, w \big\rangle + \rho \, \mathrm{dist}^2\big(\nabla \Phi(\bar{x}) w; K_{\mathcal{Q}}(\Phi(\bar{x}), \bar{\lambda})\big) > 0 \;\text{ whenever }\; w \in \mathbb{S}_-.$$

This together with the above estimate for the case of $w \in \mathbb{S}_+$ and $\rho > \varrho_0$ verifies (4.12) for all $w \in \mathbb{R}^n \setminus \{0\}$ and $\rho \geq \rho_{\bar{\lambda}} > \max\{\varrho_0, \varrho_1\}$ and thus completes the proof of the theorem.

□

Implication (i) $\Longrightarrow$ (iii) in Theorem 4.5 was established by Rockafellar in [59, Theorem 7.4] for nonlinear programming problems. His proof strongly exploits the geometry of NLPs and does not appeal to the second subderivative as in our proof. For the second-order cone programming problem (1.1), the aforementioned implication, not the established equivalencies in Theorem 4.5, was obtained in [33, Proposition 10], where in addiction the strict complementarity and nondegeneracy conditions were imposed.

To proceed further, observe that both constants $\ell_{\bar{\lambda}}$ and $\gamma_{\bar{\lambda}}$ in (4.13) depend on $\bar{\lambda}$. Now we are going to find additional assumptions that allow us to justify the second-order growth condition (4.13) for *all* $\lambda \in \Lambda(\bar{x})$ sufficiently close to $\bar{\lambda}$, where the aforementioned constants do not depend on λ. This is crucial for the convergence analysis of the ALM in the case of *nonunique* Lagrange multipliers. The rest of this section is mainly focusing on achieving such a *uniform second-order growth condition* for the augmented Lagrangian (1.7).

We begin with the following lemma, which provides a common constant $\ell_{\bar{\lambda}}$ that works for all λ sufficiently close to $\bar{\lambda}$. Then we derive a similar result for $\gamma_{\bar{\lambda}}$ in the proof of the next theorem.

Lemma 4.6 (uniform estimate for second subderivatives of augmented Lagrangians). *Let $(\bar{x}, \bar{\lambda})$ be a solution to the KKT system* (1.4)*, and let SOSC* (4.9) *hold at $(\bar{x}, \bar{\lambda})$. Then there exist positive constants $\rho_{\bar{\lambda}}$, ℓ_1, ε_0 such that for all $\rho \geq \rho_{\bar{\lambda}}$ and $\lambda \in \Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_0}(\bar{\lambda})$ we have*

$$\mathrm{d}_x^2 \mathcal{L}((\bar{x}, \lambda, \rho), 0)(w) \geq \frac{\ell_1}{2}\|w\|^2 \ \textit{ whenever } \ w \in \mathbb{R}^n. \tag{4.22}$$

Proof. Theorem 4.5 gives us a constant $\rho_{\bar\lambda} > 0$ for which condition (4.12) holds when $\rho \geq \rho_{\bar\lambda}$. Recall that the second subderivative is l.s.c. and positive homogenous of degree 2. Owing to (4.5), condition (4.12) amounts to the existence of a constant $\ell_1 > 0$ such that

$$\mathrm{d}^2_x\mathcal{L}((\bar x, \bar\lambda, \rho_{\bar\lambda}), 0)(w) = \big\langle w, \nabla^2_{xx}L(\bar x, \bar\lambda)w\big\rangle + Q_{\bar x,\bar\lambda,\rho}(v) + \rho_{\bar\lambda}\operatorname{dist}^2\big(v; K_{\mathcal{Q}}(\Phi(\bar x), \bar\lambda)\big) \geq \ell_1 \quad (4.23)$$

for all w from the unit sphere $\mathbb{S} \subset \mathbb{R}^n$ and $v = \nabla\Phi(\bar x)w$, where the quadratic form $Q_{\bar x,\bar\lambda,\rho_{\bar\lambda}}(\cdot)$ is taken from (4.6). Let us now verify the existence of $\varepsilon_0 > 0$ so that for any $\lambda \in \Lambda(\bar x) \cap \mathbb{B}_{\varepsilon_0}(\bar\lambda)$ we have

$$\mathrm{d}^2_x\mathcal{L}((\bar x, \lambda, \rho_{\bar\lambda}), 0)(w) = \big\langle w, \nabla^2_{xx}L(\bar x, \lambda)w\big\rangle + Q_{\bar x,\lambda,\rho}(v) + \rho_{\bar\lambda}\operatorname{dist}^2\big(v; K_{\mathcal{Q}}(\Phi(\bar x), \bar\lambda)\big) \geq \frac{\ell_1}{2}, \quad w \in \mathbb{S}, \quad (4.24)$$

where $Q_{\bar x,\lambda,\rho_{\bar\lambda}}(\cdot)$ is taken from (4.6) with replacing $\bar\lambda$ by λ. We first observe that $\big\langle w, \nabla^2_{xx}L(\bar x, \lambda)w\big\rangle \to \big\langle w, \nabla^2_{xx}L(\bar x, \bar\lambda)w\big\rangle$ as $\lambda \to \bar\lambda$ with $\lambda \in \Lambda(\bar x)$ uniformly for all $w \in \mathbb{S}$ due to the following estimate

$$|\big\langle w, \nabla^2_{xx}L(\bar x, \lambda)w\big\rangle - \big\langle w, \nabla^2_{xx}L(\bar x, \bar\lambda)w\big\rangle| \leq \|\nabla^2\Phi(\bar x)\| \cdot \|\lambda - \bar\lambda\|.$$

We now prove the uniform convergence of $Q_{\bar x,\lambda,\rho_{\bar\lambda}}(v) \to Q_{\bar x,\bar\lambda,\rho_{\bar\lambda}}(v)$ as $\lambda \to \bar\lambda$ with $\lambda \in \Lambda(\bar x)$ for all $v \in \nabla\Phi(\bar x)(\mathbb{S})$. It is obvious for the case with $\Phi(\bar x) \in (\operatorname{int}\mathcal{Q}) \cup \{0\}$, since quadratic forms reduce to 0 by (4.6). Assume that $\Phi(\bar x) \in (\operatorname{bd}\mathcal{Q})\setminus\{0\}$. If $\bar\lambda = 0$, $\lambda \to \bar\lambda$ with $\lambda \in \Lambda(\bar x)$, then it follows from (4.6) that

$$\begin{array}{rcl}|Q_{\bar{x},\lambda,\rho_{\bar{\lambda}}}(w)-Q_{\bar{x},\bar{\lambda},\rho_{\bar{\lambda}}}(w)| & = & \dfrac{\rho_{\bar{\lambda}}\|\lambda\|}{\rho_{\bar{\lambda}}\|\Phi(\bar{x})\|+\|\lambda\|}\Big(\|v_r\|^2-\dfrac{\langle\lambda_r,v_r\rangle^2}{\|\lambda_r\|^2}\Big)\\ & \le & \dfrac{\|\lambda\|}{\|\Phi(\bar{x})\|}\|v\|^2\le\dfrac{\|\nabla\Phi(\bar{x})\|^2}{\|\Phi(\bar{x})\|}\|\lambda-\bar{\lambda}\|,\end{array}$$

which justifies the claimed uniform convergence in this case as well. Finally, assume that $\bar{\lambda}\neq 0$ and $\lambda\to\bar{\lambda}$ with $\lambda\in\Lambda(\bar{x})$ and suppose without loss of generality that $\lambda\neq 0$. Since $\lambda\in\Lambda(\bar{x})$, $\bar{\lambda}\in\Lambda(\bar{x})$, and $\Phi(\bar{x})\in(\mathrm{bd}\,\mathcal{Q})\setminus\{0\}$, it follows from (2.3) that there exist positive constants t and $\bar{t}$ such that $\lambda=t\widetilde{\Phi}(\bar{x})$ and $\bar{\lambda}=\bar{t}\widetilde{\Phi}(\bar{x})$. These relationships result in the equality

$$\frac{\langle\lambda_r,v_r\rangle^2}{\|\lambda_r\|^2}=\frac{\langle\bar{\lambda}_r,v_r\rangle^2}{\|\bar{\lambda}_r\|^2}.$$

Using this together with (4.6) brings us to the estimates

$$\begin{array}{rcl}|Q_{\bar{x},\lambda,\rho_{\bar{\lambda}}}(v)-Q_{\bar{x},\bar{\lambda},\rho_{\bar{\lambda}}}(v)| & = & \Big|\dfrac{\rho_{\bar{\lambda}}\|\lambda\|}{\rho_{\bar{\lambda}}\|\Phi(\bar{x})\|+\|\lambda\|}-\dfrac{\rho_{\bar{\lambda}}\|\bar{\lambda}\|}{\rho_{\bar{\lambda}}\|\Phi(\bar{x})\|+\|\bar{\lambda}\|}\Big|\Big(\|v_r\|^2-\dfrac{\langle\bar{\lambda}_r,v_r\rangle^2}{\|\bar{\lambda}_r\|^2}\Big)\\ & \le & \dfrac{\|\lambda-\bar{\lambda}\|}{\|\Phi(\bar{x})\|}\|v\|^2\le\dfrac{\|\nabla\Phi(\bar{x})\|^2}{\Phi(\bar{x})}\|\lambda-\bar{\lambda}\|,\end{array}$$

which again justify the claimed uniform convergence in this last case. Thus we find a number $\varepsilon_1>0$ ensuring the uniform condition

$$\big\langle w,\nabla^2_{xx}L(\bar{x},\lambda)w\big\rangle+Q_{\bar{x},\lambda,\rho_{\bar{\lambda}}}(v)\ge\big\langle w,\nabla^2_{xx}L(\bar{x},\bar{\lambda})w+Q_{\bar{x},\bar{\lambda},\rho_{\bar{\lambda}}}(v)-\frac{\ell_1}{4}\qquad(4.25)$$

whenever $w\in\mathbb{S}$ and $\lambda\in\Lambda(\bar{x})\cap\mathbb{B}_{\varepsilon_1}(\bar{\lambda})$. Next we intend to verify the existence of $\varepsilon_2>0$

such that

$$\begin{cases} \operatorname{dist}^2\big(v; K_{\mathcal{Q}}(\Phi(\bar{x}), \lambda)\big) \geq \operatorname{dist}^2\big(v; K_{\mathcal{Q}}(\Phi(\bar{x}), \bar{\lambda})\big) - \dfrac{\ell_1}{4\rho_{\bar{\lambda}}} \\ \text{for all } \ v \in \nabla\Phi(\bar{x})(\mathbb{S}) \ \text{ and all } \ \lambda \in \Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_2}(\bar{\lambda}). \end{cases} \tag{4.26}$$

To proceed, consider the following four possible locations of $\bar{\lambda}$ in $-\mathcal{Q}$:

(a) $\bar{\lambda} = 0$. In this case we have

$$K_{\mathcal{Q}}\big(\Phi(\bar{x}), \bar{\lambda}\big) = T_{\mathcal{Q}}\big(\Phi(\bar{x})\big) \supset K_{\mathcal{Q}}\big(\Phi(\bar{x}), \lambda\big)$$

for all $\lambda \in \Lambda(\bar{x})$, which verifies the fulfillment of (4.26).

(b) $\bar{\lambda} \in \operatorname{int}(-\mathcal{Q})$ with $\Phi(\bar{x}) = 0$. If λ is sufficiently close to $\bar{\lambda}$, then $\lambda \in \operatorname{int}(-\mathcal{Q})$. This yields

$$K_{\mathcal{Q}}\big(\Phi(\bar{x}), \lambda\big) = K_{\mathcal{Q}}\big(\Phi(\bar{x}), \bar{\lambda}\big) = \{0\},$$

which immediately ensures that (4.26) holds.

(c) $\bar{\lambda} \in \operatorname{bd}(-\mathcal{Q}) \setminus \{0\}$ with $\Phi(\bar{x}) \in \operatorname{bd}(\mathcal{Q}) \setminus \{0\}$. If $\lambda \to \bar{\lambda}$ with $\lambda \in \Lambda(\bar{x})$, we get $\lambda = t\bar{\lambda}$ for some $t > 0$, which confirms that

$$K_{\mathcal{Q}}(\Phi(\bar{x}), \lambda) = K_{\mathcal{Q}}\big(\Phi(\bar{x}), \bar{\lambda}\big).$$

This clearly justifies the claimed estimate (4.26).

(d) $\bar{\lambda} \in \mathrm{bd}\,(-\mathcal{Q}) \setminus \{0\}$ with $\Phi(\bar{x}) = 0$. In this case, we have for all $\lambda \in \Lambda(\bar{x}) \setminus \{0\}$ that

$$K_{\mathcal{Q}}\big(\Phi(\bar{x}), \lambda\big) = \begin{cases} \mathbb{R}_+ \widetilde{\lambda} & \text{if } \lambda \in \mathrm{bd}\,(-\mathcal{Q}) \setminus \{0\}, \\ \{0\} & \text{if } \lambda \in \mathrm{int}\,(-\mathcal{Q}), \end{cases}$$

where the tilde-notation for the ice-cream cone is defined at the end of Section 1. Then (4.26) is obviously satisfied when $\lambda \in \Lambda(\bar{x}) \cap \mathrm{int}\,(-\mathcal{Q})$. Assume now that $\lambda \in [\Lambda(\bar{x}) \cap \mathrm{bd}\,(-\mathcal{Q})] \setminus \{0\}$. It is not hard to verify that for $\lambda \in [\Lambda(\bar{x}) \cap \mathrm{bd}\,(-\mathcal{Q})] \setminus \{0\}$ we get

$$\mathrm{dist}^2\big(v; K_{\mathcal{Q}}(\Phi(\bar{x}), \lambda)\big) = \|v\|^2 - \frac{1}{\|\lambda\|^2}\big(\max\{0, \langle \widetilde{\lambda}, v\rangle\}\big)^2.$$

It is worth mentioning that the function $\big(\max(0,t)\big)^2$ is $\mathcal{C}^1$ on the whole real line. It follows from the latter formula that $\lambda \in [\Lambda(\bar{x}) \cap \mathrm{bd}\,(-\mathcal{Q})] \setminus \{0\} \mapsto \mathrm{dist}^2\big(v; K_{\mathcal{Q}}(\Phi(\bar{x}), \lambda)\big)$ is a $\mathcal{C}^1$ function relative to the set $[\Lambda(\bar{x}) \cap \mathrm{bd}\,(-\mathcal{Q})] \setminus \{0\}$. Taking this into account and choosing λ to be sufficiently close to $\bar{\lambda}$ ensure the existence of $\varepsilon_2 > 0$ for which the uniform estimate (4.26) is guaranteed. This completes the justification of (4.26) for all the possible cases.

Finally, denote $\varepsilon_0 := \min\{\varepsilon_1, \varepsilon_2\}$ with ε_1 and ε_2 taken from (4.25) and (4.26), respectively. Combining (4.23), (4.25), and (4.26) tells us that estimate (4.24) is satisfied for any $\lambda \in \Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_0}(\bar{\lambda})$. Thus for any such a multiplier λ we have

$$\mathrm{d}_x^2 \mathcal{L}\big((\bar{x}, \lambda, \rho_{\bar{\lambda}}), 0\big)(w) \geq \frac{\ell_1}{2}\|w\|^2 \ \text{ whenever } \ w \in \mathbb{R}^n.$$

This together with (4.5) and the fact that $\rho \mapsto Q_{\bar{x},\bar{\lambda},\rho}(v)$ is nondecreasing on $\mathbb{R}_+$ implies for any $\lambda \in \Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_0}(\bar{\lambda})$ that

$$\mathrm{d}_x^2\mathcal{L}\big((\bar{x},\lambda,\rho),0\big)(w) \geq \tfrac{1}{2}\ell_1\|w\|^2 \;\text{ for all }\; w \in \mathbb{R}^n \;\text{ and all }\; \rho \geq \rho_{\bar{\lambda}},$$

which therefore completes the proof of the lemma. □

Now we are ready to derive a uniform version of the second-order growth condition for (1.7).

Theorem 4.7 (uniform second-order growth condition for augmented Lagrangians). *Let $(\bar{x},\bar{\lambda})$ be a solution to the KKT system* (1.4)*, and let SOSC* (4.9) *hold at $(\bar{x},\bar{\lambda})$. Assume in addition that the Lagrange multiplier set $\Lambda(\bar{x})$ in* (1.5) *is either a polyhedron, or that the multiplier $\bar{\lambda}$ belongs to the interior of $-\mathcal{Q}$. Then there are positive constants $\rho_{\bar{\lambda}},\gamma_{\bar{\lambda}},\varepsilon_{\bar{\lambda}},\ell_{\bar{\lambda}}$ such that for all $\lambda \in \Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_{\bar{\lambda}}}(\bar{\lambda})$ and $\rho \geq \rho_{\bar{\lambda}}$ we have the uniform second-order growth condition*

$$\mathcal{L}(x,\lambda,\rho) \geq f(\bar{x}) + \ell_{\bar{\lambda}}\|x-\bar{x}\|^2 \;\textit{ whenever }\; x \in \mathbb{B}_{\gamma_{\bar{\lambda}}}(\bar{x}). \tag{4.27}$$

Proof. Take the positive constants ℓ_1, ε_0, and $\rho_{\bar{\lambda}}$ from Lemma 4.6 for which (4.22) holds whenever $\lambda \in \Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_0}(\bar{\lambda})$ and $\rho \geq \rho_{\bar{\lambda}}$. Using [60, Theorem 13.24] and remembering that $\mathcal{L}(\bar{x},\lambda,\rho_{\bar{\lambda}}) = f(\bar{x})$ for all $\lambda \in \Lambda(\bar{x})$, we deduce from (4.22) that for any $\lambda \in \Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_0}(\bar{\lambda})$ there exists $\gamma_\lambda > 0$ ensuring the estimate

$$\mathcal{L}(x,\lambda,\rho_{\bar{\lambda}}) \geq f(\bar{x}) + \frac{\ell_1}{4}\|x-\bar{x}\|^2 \;\text{ whenever }\; x \in \mathbb{B}_{\gamma_\lambda}(\bar{x}), \tag{4.28}$$

where the constant $\frac{\ell_1}{4}$ can be chosen the same for all the multipliers $\lambda \in \Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_0}(\bar{\lambda})$. This comes from (4.22) and the proof of [60, Theorem 13.24]; see also Remark 4.4 for a similar discussion. However, the radii of the balls centered at $\bar{x}$ in (4.28) depend on λ. It is shown below that we can find a common radius for all the multipliers $\lambda \in \Lambda(\bar{x})$ that are sufficiently close to $\bar{\lambda}$. To proceed, define the function $\varphi\colon \mathbb{R}^{m+1} \to \overline{\mathbb{R}}$ by

$$\varphi(\lambda) := \sup_{x \in \mathbb{B}_{\gamma_{\bar{\lambda}}}(\bar{x})} \frac{f(\bar{x}) - \mathcal{L}(x, \lambda, \rho_{\bar{\lambda}})}{\|x - \bar{x}\|^2} + \delta_{\Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_0}(\bar{\lambda})}(\lambda), \quad \lambda \in \mathbb{R}^{m+1}. \tag{4.29}$$

Proposition 1.1(ii) tells us that the function $\lambda \mapsto \mathcal{L}(x, \lambda, \rho_{\bar{\lambda}})$ is concave. This together with the convexity of the set $\Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_0}(\bar{\lambda})$ ensures that φ in (4.29) is a convex function. Let us now verify that for any $\lambda \in \Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_0}(\bar{\lambda})$ the value $\varphi(\lambda)$ is finite. To this end, pick such a multiplier λ and observe that for $\gamma_\lambda \geq \gamma_{\bar{\lambda}}$ we get by (4.28) the estimates

$$\varphi(\lambda) \leq \sup_{x \in \mathbb{B}_{\gamma_\lambda}(\bar{x})} \frac{f(\bar{x}) - \mathcal{L}(x, \lambda, \rho_{\bar{\lambda}})}{\|x - \bar{x}\|^2} \leq -\frac{\ell_1}{4}.$$

In particular, this implies that $\varphi(\bar{\lambda}) \leq -\frac{\ell_1}{4}$. If $\gamma_\lambda < \gamma_{\bar{\lambda}}$, then

$$\varphi(\lambda) \leq \max\left\{ \sup_{x \in \mathbb{B}_{\gamma_\lambda}(\bar{x})} \frac{f(\bar{x}) - \mathcal{L}(x, \lambda, \rho_{\bar{\lambda}})}{\|x - \bar{x}\|^2}, \max_{\gamma_\lambda \leq \|x - \bar{x}\| \leq \gamma_{\bar{\lambda}}} \frac{f(\bar{x}) - \mathcal{L}(x, \lambda, \rho_{\bar{\lambda}})}{\|x - \bar{x}\|^2} \right\} < \infty,$$

where the first term inside the maximum does not exceed $-\ell_1/4$ because of (4.28), and where the second term is finite since it is the maximum of a continuous function over a compact set. This implies that $\varphi(\lambda)$ is finite for all $\lambda \in \Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_0}(\bar{\lambda})$, which ensures that

$$\operatorname{dom} \varphi = \Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_0}(\bar{\lambda}).$$

If $\bar{\lambda} \in \operatorname{int}(-\mathcal{Q})$, we get $\bar{\lambda} \in \operatorname{ri} \Lambda(\bar{x})$, which clearly implies that $\bar{\lambda} \in \operatorname{ri}(\operatorname{dom} \varphi)$. Since φ is convex, it is continuous at $\bar{\lambda}$ relative to its domain. Hence we find $\varepsilon_{\bar{\lambda}} \in (0, \varepsilon_0]$ such that

$$\varphi(\lambda) \le \varphi(\bar{\lambda}) + \frac{\ell_1}{8} \le -\frac{\ell_1}{8} \text{ for all } \lambda \in \operatorname{dom}\varphi \cap \mathbb{B}_{\varepsilon_{\bar{\lambda}}}(\bar{\lambda}) = \Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_{\bar{\lambda}}}(\bar{\lambda}). \tag{4.30}$$

Next we proceed to achieve a similar result when $\Lambda(\bar{x})$ is a polyhedral convex set. In this case the collection of Lagrange multipliers $\Lambda(\bar{x})$ is either a ray on the boundary of $-\mathcal{Q}$, or a singleton. If the latter holds, we obtain $\Lambda(\bar{x}) = \{\bar{\lambda}\}$, and hence the uniform growth condition (4.27) follows directly from (4.13). If $\Lambda(\bar{x})$ is a ray on the boundary of $-\mathcal{Q}$, then $\Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_0}(\bar{\lambda})$ is a segment. If now $\bar{\lambda} \neq 0$, then we get $\bar{\lambda} \in \text{ri}\,[\Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_0}(\bar{\lambda})] = \text{ri}\,(\operatorname{dom}\varphi)$. Arguing as above leads us to (4.30) in this case. Otherwise, $\bar{\lambda}$ is an endpoint of the aforementioned segment, and thus $\bar{\lambda} = 0$. Let λ_e be the other endpoint. If $\varphi(\lambda_e) \le \varphi(\bar{\lambda}) + \ell_1/8$, then (4.30) holds for $\varepsilon_{\bar{\lambda}} := \varepsilon_0$, which follows from the convexity of φ. Otherwise, we have that $\varphi(\widehat{\lambda}_e) > \varphi(\bar{\lambda}) + \ell_1/8$. Denote

$$\bar{t} := \frac{\ell_1}{8\big(\varphi(\lambda_e) - \varphi(\bar{\lambda})\big)} \in (0,1) \text{ and } \lambda_{\bar{t}} := (1-\bar{t})\bar{\lambda} + \bar{t}\lambda_e.$$

Then using the convexity of φ tells us that

$$\varphi(\lambda_{\bar{t}}) \le (1-\bar{t})\varphi(\bar{\lambda}) + \bar{t}\varphi(\lambda_e) = \varphi(\bar{\lambda}) + \bar{t}\big(\varphi(\lambda_e) - \varphi(\bar{\lambda})\big) = \varphi(\bar{\lambda}) + \frac{\ell_1}{8} \le -\frac{\ell_1}{8},$$

which readily yields (4.30) with $\varepsilon_{\bar{\lambda}} := \|\lambda_{\bar{t}} - \bar{\lambda}\| \in (0, \varepsilon_0]$. This completes the verification of (4.30) with some constant $\varepsilon_{\bar{\lambda}} \in (0, \varepsilon_0]$ if either $\Lambda(\bar{x})$ is a polyhedral convex set, or

$\bar{\lambda} \in \text{int}\,(-\mathcal{Q})$. Consequently, it follows from (4.29) and (4.30) that

$$\mathcal{L}(x, \lambda, \rho_{\bar{\lambda}}) \geq f(\bar{x}) + \frac{\ell_1}{8}\|x - \bar{x}\|^2 \text{ for all } x \in \mathbb{B}_{\gamma_{\bar{\lambda}}}(\bar{x}) \text{ and } \lambda \in \Lambda(\bar{x}) \cap \mathbb{B}_{\varepsilon_{\bar{\lambda}}}(\bar{\lambda}). \tag{4.31}$$

Employing now Proposition 1.1(i) gives us the inequality

$$\mathcal{L}(x, \lambda, \rho) \geq \mathcal{L}(x, \lambda, \rho_{\bar{\lambda}}) \text{ for all } \rho \geq \rho_{\bar{\lambda}}.$$

Combining this with (4.31) and setting $\ell_{\bar{\lambda}} := \frac{\ell_1}{8}$ verify the uniform growth condition (4.27).

□

A similar result to Theorem 4.7 was derived in [14, Proposition 3.1] for NLPs. The given proof therein seems however to be rather sketchy in some details. We are not familiar with any previous results on the uniform second-order growth condition (4.27) for SOCPs. As shown in the next section, the second-order growth conditions obtained above are crucial for developing the augmented Lagrangian method for this class of optimization problems.

CHAPTER 5 CONVERGENCE ANALYSIS OF AUGMENTED LAGRANGIAN METHOD FOR SOCPS

5.1 Error Bounds for Perturbed KKT Systems of SOCPs

Here we derive an efficient error bound estimate for the KKT system of problem (1.1) under the validity of SOSC (4.9). This is highly important for the subsequent results of the paper.

A crucial role of error bounds in convergence analysis of major numerical algorithms has been well understood in optimization theory; see, e.g., the books [13, 29]. To the best of our knowledge, the first error bound estimate for KKT systems of NLPs under the classical second-order sufficient condition alone was derived in Hager and Gowda [22, Lemma 2] and then was improved by Izmailov [27] who replaced the conventional SOSC with the weaker *noncriticality* of Lagrange multipliers introduced therein. It has been recently observed by Mordukhovich and Sarabi [44] that similar results for nonpolyhedral conic programs require an additional assumption of the *calmness* of Lagrange multiplier mappings associated with canonically perturbed KKT systems. The latter assumption automatically holds for NLPs.

For any fixed $\bar{x} \in \mathbb{R}^n$ the *multiplier mapping* $M_{\bar{x}} \colon \mathbb{R}^n \times \mathbb{R}^{m+1} \rightrightarrows \mathbb{R}^{m+1}$, associated with the canonically perturbed KKT system (1.4) of (1.1), is defined by

$$M_{\bar{x}}(v,w) := \big\{\lambda \in \mathbb{R}^{m+1} \;\big|\; \nabla_x L(\bar{x},\lambda) = v,\ \lambda \in N_{\mathcal{Q}}\big(\Phi(\bar{x}) + w\big)\big\},\quad (v,w) \in \mathbb{R}^n \times \mathbb{R}^{m+1}. \tag{5.1}$$

It is easy to see that $M_{\bar{x}}(0,0)$ reduces to the set of Lagrange multipliers $\Lambda(\bar{x})$ of the unperturbed system (1.5). Given a solution $(\bar{x},\bar{\lambda})$ to the KKT system (1.4), the calmness condition (3.3) for $M_{\bar{x}}$ at $((0,0),\bar{\lambda})$ reads as the existence of positive constants τ and γ

such that

$$M_{\bar{x}}(v,w)\cap \mathbb{B}_\gamma(\bar{\lambda})\subset \Lambda(\bar{x})+\tau\big(\|w\|+\|v\|\big)\mathbb{B}\ \text{ whenever }\ (v,w)\in \mathbb{B}_\gamma(0,0).$$

This can be equivalently rewritten as the existence of $\tau,\gamma>0$ such that the estimate

$$\operatorname{dist}\big(\lambda;\Lambda(\bar{x})\big)\le \tau\Big(\|\nabla_x L(\bar{x},\lambda)\|+\operatorname{dist}\big(\Phi(\bar{x});N_{\mathcal{Q}}^{-1}(\lambda)\big)\Big) \tag{5.2}$$

holds for all $\lambda\in\mathbb{B}_\gamma(\bar{\lambda})$. We can easily check that for (polyhedral) NLPs the calmness of the multiplier mapping follows automatically from the classical Hoffman lemma. Efficient conditions for the calmness of (5.1) in the SOCP framework (1.1) are presented at the end of this section.

Now we are ready to derive the main result of this section ensuring the aforementioned error bound estimate. Define the *residual function* $\sigma\colon \mathbb{R}^n\times\mathbb{R}^{m+1}\to\mathbb{R}$ of the KKT system (1.4) by

$$\sigma(x,\lambda):=\|\nabla_x L(x,\lambda)\|+\|\Phi(x)-\Pi_{\mathcal{Q}}\big(\Phi(x)+\lambda\big)\|,\ \ (x,\lambda)\in\mathbb{R}^n\times\mathbb{R}^{m+1}. \tag{5.3}$$

It is easy to see that if $(\bar{x},\bar{\lambda})$ is a solution to the KKT system (1.4), then it follows from property (P4) of the projection mapping that $\sigma(\bar{x},\bar{\lambda})=0$. Using this and the Lipschitz continuity of σ with respect to both x and λ around $(\bar{x},\bar{\lambda})$, we can find constants $\gamma_2>0$ and $\kappa_2\ge 0$ such that

$$\sigma(x,\lambda)\le \kappa_2\big(\|x-\bar{x}\|+\operatorname{dist}(\lambda;\Lambda(\bar{x}))\big)\quad \text{for all }\ (x,\lambda)\in\mathbb{B}_{\gamma_2}(\bar{x},\bar{\lambda}). \tag{5.4}$$

Below we show that the opposite inequality in (5.4), which is crucial for our subsequent developments of the ALM, can be achieved if in addition both SOSC (4.9) and the calmness of the multiplier mapping are satisfied. The provided proof, being strongly based on the geometry of the second-order cone (1.2), is much simpler than the one given recently in [44, Theorem 5.9] for $\mathcal{C}^2$-cone reducible cone programs that is based on a highly involved reduction technique.

Theorem 5.1 (error bound for SOCPs under calmness and SOSC). *Let $(\bar{x}, \bar{\lambda})$ be a solution to the KKT system* (1.4)*, and let SOSC* (4.9) *hold at $(\bar{x}, \bar{\lambda})$. If the multiplier mapping $M_{\bar{x}}$ in* (5.1) *is calm at $((0,0), \bar{\lambda})$, then there exist constants $\gamma_1 > 0$ and $\kappa_1 \geq 0$ such that*

$$\|x - \bar{x}\| + \operatorname{dist}\big(\lambda; \Lambda(\bar{x})\big) \leq \kappa_1 \, \sigma(x, \lambda) \quad \textit{for all} \quad (x, \lambda) \in \mathbb{B}_{\gamma_1}(\bar{x}, \bar{\lambda}), \tag{5.5}$$

where the residual function σ is taken from (5.3).

Proof. Observe that if $x = \bar{x}$ and $\lambda \in \Lambda(\bar{x})$, then (5.5) holds since both sides are equal to 0. Let us now verify (5.5) while assuming that either $x \neq \bar{x}$ or $\lambda \notin \Lambda(\bar{x})$. We first show that

$$\|x - \bar{x}\| = O\big(\sigma(x, \lambda)\big) \quad \text{as} \quad (x, \lambda) \to (\bar{x}, \bar{\lambda}). \tag{5.6}$$

Arguing by contradiction, suppose that there exists a sequence $(x^k, \lambda^k) \to (\bar{x}, \bar{\lambda})$ with either $x^k \neq \bar{x}$ or $\lambda^k \notin \Lambda(\bar{x})$ satisfying the strict inequalities

$$\|x^k - \bar{x}\| > k \, \sigma(x^k, \lambda^k) > 0 \quad \text{for all} \;\; k \in \mathbb{N},$$

which imply that $\sigma(x^k,\lambda^k)=o(\|x^k-\bar{x}\|)$. By the definition of σ the latter means that

$$\nabla_x L(x^k,\lambda^k)=o(\|x^k-\bar{x}\|) \quad \text{and} \quad \alpha^k:=\Phi\big(x^k\big)-\Pi_{\mathcal{Q}}(\Phi(x^k)+\lambda^k\big)=o(\|x^k-\bar{x}\|). \tag{5.7}$$

Using the second equality in (5.7) combined with property (P1), we get the relationships

$$\Phi(x^k)-\alpha^k\in\mathcal{Q}, \quad \lambda^k+\alpha^k\in-\mathcal{Q}, \quad \text{and} \quad \big\langle\Phi(x^k)-\alpha^k,\lambda^k+\alpha^k\big\rangle=0, \tag{5.8}$$

which in turn bring us to the inclusion

$$\lambda^k+\alpha^k\in N_{\mathcal{Q}}\big(\Phi(x^k)-\alpha^k\big). \tag{5.9}$$

It follows from the calmness estimate (5.2) that

$$\text{dist}\big(\lambda^k+\alpha^k;\Lambda(\bar{x})\big)\le\tau\big(\|\nabla_x L(\bar{x},\lambda^k+\alpha^k)\|+\text{dist}(\Phi(\bar{x});N_{\mathcal{Q}}^{-1}(\lambda^k+\alpha^k))\big)$$

for all $k\in\mathbb{N}$ sufficiently large. Since the gradient ∇f and Jacobian $\nabla\Phi$ mappings are Lipschitz continuous around $\bar{x}$, we always have the estimate

$$\begin{aligned}\|\nabla_x L(\bar{x},\lambda^k+\alpha^k)\| \;&\le\; \|\nabla f(x^k)-\nabla f(\bar{x})\|+\|\nabla_x L(x^k,\lambda^k)\|+\|(\nabla\Phi(x^k)-\nabla\Phi(\bar{x}))^*\lambda^k\|\\ &\quad+\|\nabla\Phi(\bar{x})^*\alpha^k\|=O(\|x^k-\bar{x}\|).\end{aligned}$$

On the other hand, it follows from (5.9) that $\Phi(x^k)-\alpha^k\in N_{\mathcal{Q}}^{-1}(\lambda^k+\alpha^k)$, and hence

$$\mathrm{dist}\big(\Phi(\bar{x});N_{\mathcal{Q}}^{-1}(\lambda^k+\alpha^k)\big)\le\|\Phi(x^k)-\alpha^k-\Phi(\bar{x})\|=O(\|x^k-\bar{x}\|),$$

where the last equality comes from the Lipschitz continuity of Φ around $\bar{x}$ and the condition $\alpha^k=o(\|x^k-\bar{x}\|)$. This ensures in turn that $\lambda^k-\widehat{\lambda}^k=O(\|x^k-\bar{x}\|)$, where $\widehat{\lambda}^k:=\Pi_{\Lambda(\bar{x})}(\lambda^k)$. Passing to subsequences if necessary gives us

$$\frac{x^k-\bar{x}}{\|x^k-\bar{x}\|}\to\xi\neq 0\quad\text{and}\quad\frac{\lambda^k-\widehat{\lambda}^k}{\|x^k-\bar{x}\|}\to\eta\ \text{ as }\ k\to\infty. \tag{5.10}$$

Appealing now to the first estimate in (5.7), we arrive at the equalities

$$\begin{aligned}
o(\|x^k-\bar{x}\|) &= \nabla_x L(x^k,\lambda^k)=\nabla_x L(x^k,\bar{\lambda})+\nabla\Phi(x^k)^*(\lambda^k-\bar{\lambda})\\
&= \nabla_x L(\bar{x},\bar{\lambda})+\nabla^2_{xx}L(\bar{x},\bar{\lambda})(x^k-\bar{x})+o(\|x^k-\bar{x}\|)\\
&\quad+\nabla\Phi(x^k)^*(\lambda^k-\widehat{\lambda}^k)+\big(\nabla\Phi(x^k)-\nabla\Phi(\bar{x})\big)^*(\widehat{\lambda}^k-\bar{\lambda})\\
&= \nabla^2_{xx}L(\bar{x},\bar{\lambda})(x^k-\bar{x})+\nabla\Phi(x^k)^*(\lambda^k-\widehat{\lambda}^k)\\
&\quad+\big(\nabla^2\Phi(\bar{x})(x^k-\bar{x})+o(\|x^k-\bar{x}\|)\big)^*(\widehat{\lambda}^k-\bar{\lambda})+o(\|x^k-\bar{x}\|).
\end{aligned}$$

Dividing both sides by $\|x^k-\bar{x}\|$ and then passing to the limit as $k\to\infty$ show that

$$0=\nabla^2_{xx}L(\bar{x},\bar{\lambda})\xi+\nabla\Phi(\bar{x})^*\eta. \tag{5.11}$$

Let us now verify the inclusion $\nabla\Phi(\bar{x})\xi\in K_{\mathcal{Q}}(\Phi(\bar{x}),\bar{\lambda})=T_{\mathcal{Q}}(\Phi(\bar{x}))\cap\{\bar{\lambda}\}^{\perp}$. Indeed, using the first relation in (5.8) yields

$$\mathcal{Q} \ni \Phi(x^k) - \alpha^k = \Phi(\bar{x}) + \|x^k - \bar{x}\| \Big[\nabla\Phi(\bar{x}) \Big(\frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \Big) + \frac{o(\|x^k - \bar{x}\|)}{\|x^k - \bar{x}\|} \Big],$$

which tells us that $\nabla\Phi(\bar{x})\xi \in T_{\mathcal{Q}}(\Phi(\bar{x}))$. Combining this with $\bar{\lambda} \in N_{\mathcal{Q}}(\Phi(\bar{x}))$, we obtain $\langle \bar{\lambda}, \nabla\Phi(\bar{x})\xi \rangle \le 0$. To prove the equality therein, deduce from (5.9) that

$$0 \ge \big\langle \lambda^k + \alpha^k, \Phi(\bar{x}) - \Phi(x^k) + \alpha^k \big\rangle = -\Big\langle \lambda^k + \alpha^k, \|x^k - \bar{x}\| \Big[\nabla\Phi(\bar{x}) \Big(\frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \Big) + \frac{o(\|x^k - \bar{x}\|)}{\|x^k - \bar{x}\|} \Big] \Big\rangle.$$

Dividing both sides by $\|x^k - \bar{x}\|$ and then passing to the limit as $k \to \infty$ verify that $\langle \bar{\lambda}, \nabla\Phi(\bar{x})\xi \rangle \ge 0$. Thus we get $\langle \bar{\lambda}, \nabla\Phi(\bar{x})\xi \rangle = 0$ and hence arrive at $\nabla\Phi(\bar{x})\xi \in K_{\mathcal{Q}}(\Phi(\bar{x}), \bar{\lambda})$.

Our next step is to prove the following inequality involving the second subderivative (2.11):

$$\big\langle \nabla\Phi\big(\bar{x}\big)\xi, \eta \big\rangle \ge \mathrm{d}^2\delta_{\mathcal{Q}}\big(\Phi(\bar{x}), \bar{\lambda}\big)\big(\nabla\Phi(\bar{x})\xi\big). \tag{5.12}$$

To proceed, remember that $\widehat{\lambda}^k \in N_{\mathcal{Q}}(\Phi(\bar{x}))$. Using (5.9) and the monotonicity of the normal cone mapping to a convex set, we get

$$\begin{aligned} 0 &\le \big\langle \Phi(x^k) - \Phi(\bar{x}) - \alpha^k, \lambda^k - \widehat{\lambda}^k + \alpha^k \big\rangle \\ &= \big\langle \nabla\Phi(\bar{x})(x^k - \bar{x}) + o(\|x^k - \bar{x}\|), \lambda^k - \widehat{\lambda}^k + o(\|x^k - \bar{x}\| \big\rangle. \end{aligned}$$

Dividing both sides by $\|x^k - \bar{x}\|^2$ and passing to the limit as $k \to \infty$ give us

$$\big\langle \nabla\Phi(\bar{x})\xi, \eta \big\rangle \ge 0.$$

This combined with (2.13) verifies (5.12) if either $\Phi(\bar{x}) = 0$, $\Phi(x) \in \text{int}\,\mathcal{Q}$, or $\bar{\lambda} = 0$.

It remains to validate (5.12) in the case where $\Phi(\bar{x}) \in (\text{bd}\,\mathcal{Q}) \setminus \{0\}$ and $\bar{\lambda} \neq 0$. Then (5.9) and the normal cone representation (2.3) allow us to find $t_k \in \mathbb{R}_+$ and $\widehat{t}_k \in \mathbb{R}_+$ such that $\lambda^k + \alpha^k = t_k(\widetilde{\Phi}(x^k) - \widetilde{\alpha}^k)$ and $\widehat{\lambda}^k = \widehat{t}_k \widetilde{\Phi}(\bar{x})$ for large $k \in \mathbb{N}$. We clearly have $\lim_{k\to\infty} t_k = \lim_{k\to\infty} \widehat{t}_k = \|\bar{\lambda}\|/\|\Phi(\bar{x})\|$. Passing to a subsequence if necessary, assume without loss of generality that either $t_k \geq \widehat{t}_k$ or $t_k \leq \widehat{t}_k$ for all $k \in \mathbb{N}$. If the former holds, then

$$
\begin{aligned}
& \big\langle \Phi(x^k) - \Phi(\bar{x}) - \alpha^k, \lambda^k - \widehat{\lambda}^k + \alpha^k \big\rangle = \big\langle \Phi(x^k) - \Phi(\bar{x}) - \alpha^k, t_k\widetilde{\Phi}(x^k) - \widehat{t}_k\widetilde{\Phi}(\bar{x}) - t_k\widetilde{\alpha}^k \big\rangle \\
= \;& \widehat{t}_k \big\langle \Phi(x^k) - \Phi(\bar{x}) - \alpha^k, \widetilde{\Phi}(x^k) - \widetilde{\alpha}^k - \widetilde{\Phi}(\bar{x}) \big\rangle + (t_k - \widehat{t}_k)\big\langle \Phi(x^k) - \Phi(\bar{x}) - \alpha^k, \widetilde{\Phi}(x^k) - \widetilde{\alpha}^k \big\rangle \\
= \;& \widehat{t}_k \big\langle \nabla\Phi(\bar{x})(x^k - \bar{x}) + o(\|x^k - \bar{x}\|), \nabla\widetilde{\Phi}(\bar{x})(x^k - \bar{x}) + o(\|x^k - \bar{x}\|) \big\rangle - (t_k - \widehat{t}_k)\big\langle \Phi(\bar{x}), \widetilde{\Phi}(x^k) - \widetilde{\alpha}^k \big\rangle \\
\geq \;& \widehat{t}_k \big\langle \nabla\Phi(\bar{x})(x^k - \bar{x}) + o(\|x^k - \bar{x}\|), \nabla\widetilde{\Phi}(\bar{x})(x^k - \bar{x}) + o(\|x^k - \bar{x}\|) \big\rangle,
\end{aligned}
$$

where the second equality comes from $\Phi(x^k) - \alpha^k \in \text{bd}\,\mathcal{Q}$ and the last inequality is due to $\Phi(\bar{x}) \in \mathcal{Q}$ while $\widetilde{\Phi}(x^k) - \widetilde{\alpha}^k \in -\mathcal{Q}$. If the latter holds, a similar argument brings us to

$$
\begin{aligned}
& \big\langle \Phi(x^k) - \Phi(\bar{x}) - \alpha^k, \lambda^k - \widehat{\lambda}^k + \alpha^k \big\rangle = \big\langle \Phi(x^k) - \Phi(\bar{x}) - \alpha^k, t_k\widetilde{\Phi}(x^k) - \widehat{t}_k\widetilde{\Phi}(\bar{x}) - t_k\widetilde{\alpha}^k \big\rangle \\
= \;& t_k \big\langle \Phi(x^k) - \Phi(\bar{x}) - \alpha^k, \widetilde{\Phi}(x^k) - \widetilde{\alpha}^k - \widetilde{\Phi}(\bar{x}) \big\rangle + (t_k - \widehat{t}_k)\big\langle \Phi(x^k) - \Phi(\bar{x}) - \alpha^k, \widetilde{\Phi}(\bar{x}) \big\rangle \\
= \;& t_k \big\langle \nabla\Phi(\bar{x})(x^k - \bar{x}) + o(\|x^k - \bar{x}\|), \nabla\widetilde{\Phi}(\bar{x})(x^k - \bar{x}) + o(\|x^k - \bar{x}\|) \big\rangle + (t_k - \widehat{t}_k)\big\langle \Phi(x^k) - \alpha^k, \widetilde{\Phi}(\bar{x}) \big\rangle \\
\geq \;& t_k \big\langle \nabla\Phi(\bar{x})(x^k - \bar{x}) + o(\|x^k - \bar{x}\|), \nabla\widetilde{\Phi}(\bar{x})(x^k - \bar{x}) + o(\|x^k - \bar{x}\|) \big\rangle.
\end{aligned}
$$

Dividing these estimates by $\|x^k - \bar{x}\|^2$ and passing to the limit as $k \to \infty$ result in

$$\big\langle \nabla\Phi(\bar{x})\xi, \eta\big\rangle \geq \frac{\|\bar{\lambda}\|}{\|\Phi(\bar{x})\|}\big\langle \nabla\Phi(\bar{x})\xi, \nabla\widetilde{\Phi}(\bar{x})\xi\big\rangle = \mathrm{d}^2\delta_{\mathcal{Q}}(\Phi(\bar{x}), \bar{\lambda})(\nabla\Phi(\bar{x})\xi),$$

where the last equality is taken from (2.13). This fully justifies (5.12).

Combining now (5.12) with (5.11) implies that

$$\big\langle \xi, \nabla^2_{xx}L(\bar{x}, \bar{\lambda})\xi\big\rangle + \mathrm{d}^2\delta_{\mathcal{Q}}\big(\Phi(\bar{x}), \bar{\lambda}\big)\big(\nabla\Phi(\bar{x})\xi\big) \leq \big\langle \xi, \nabla^2_{xx}L(\bar{x}, \bar{\lambda})\xi\big\rangle + \big\langle \nabla\Phi(\bar{x})\xi, \eta\big\rangle = 0,$$

which contradicts the second-order sufficient condition (4.9) since $\Phi(\bar{x})\xi \in T_{\mathcal{Q}}(\Phi(\bar{x}))$ and $\xi \neq 0$, and thus verifies estimate (5.6).

To finish the proof of the claimed error bound (5.5), it remains to show that

$$\operatorname{dist}\big(\lambda; \Lambda(\bar{x})\big) = O\big(\sigma(x, \lambda)\big) \quad \text{as} \quad (x, \lambda) \to (\bar{x}, \bar{\lambda}). \tag{5.13}$$

To proceed, pick (x, λ) satisfying (5.6) and denote $y := \Pi_{\mathcal{Q}}(\Phi(x) + \lambda) - \Phi(x)$. Thus we get $\lambda - y \in N_{\mathcal{Q}}(\Phi(x) + y)$. Moreover, since $(x, \lambda) \to (\bar{x}, \bar{\lambda})$, we get $y \to 0$. Combining the latter with (5.2) readily yields the relationships

$$\begin{aligned} \operatorname{dist}\big(\lambda - y; \Lambda(\bar{x})\big) &= O\big(\|\nabla_x L(\bar{x}, \lambda - y)\| + \operatorname{dist}(\Phi(\bar{x}); N_{\mathcal{Q}}^{-1}(\lambda - y))\big) \\ &= O\big(\|\nabla_x L(x, \lambda)\| + \|y\| + \|x - \bar{x}\|\big) = O\big(\sigma(x, \lambda)\big), \end{aligned}$$

where the last equality comes from (5.6). Since

$$\operatorname{dist}\big(\lambda; \Lambda(\bar{x})\big) - \operatorname{dist}\big(\lambda - y; \Lambda(\bar{x})\big) = O(\|y\|) = O\big(\sigma(x, \lambda)\big),$$

we arrive at (5.13). The error bound (5.5) follows from the combination of (5.6) and (5.13), and hence completes the proof of the theorem. □

Next we present an example showing that the assumed calmness of the multiplier mapping in Theorem 5.1 is essential for the validity of the error bound (5.5). In fact, the following example demonstrates more: not only does the *primal-dual* error bound (5.5) fail without the calmness assumption on (5.1), but even the *primal estimate* (5.6) is violated in the absence of calmness. This illustrates a striking difference between NLPs and nonpolyhedral SOCPs.

Example 5.2 (failure of error bound in the absence of calmness of multiplier mappings). *Consider SOCP* (1.1) *with the data* $f\colon \mathbb{R}^2 \to \mathbb{R}$ *and* $\Phi\colon \mathbb{R}^2 \to \mathbb{R}^3$ *defined by*

$$f(x :) = x_2^2 \quad \textit{and} \quad \Phi(x) := (-x_1^2 + x_2, x_2, 0) \quad \textit{with} \;\; x = (x_1, x_2) \in \mathbb{R}^2.$$

Take $\bar{x} := (0,0)$ *and observe that* $\Phi(\bar{x}) = 0$ *and that*

$$\nabla f(\bar{x}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \nabla\Phi(\bar{x})^* = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}, \quad \Lambda(\bar{x}) = -\mathcal{Q} \cap \{(1,1,0)\}^{\perp} = \mathbb{R}_+(-1,1,0).$$

Letting $\bar{\lambda} := (-1,1,0) \in \Lambda(\bar{x})$*, we conclude that the pair* $(\bar{x}, \bar{\lambda})$ *satisfies the KKT system* (1.4).

It follows from the equality

$$\nabla^2_{xx}L(\bar{x},\bar{\lambda}) = \nabla^2 f(\bar{x}) + \nabla^2\langle\bar{\lambda},\Phi\rangle(\bar{x}) = 2I_2,$$

with I_2 standing for the 2×2 identity matrix, that SOSC (4.9) *holds at $(\bar{x},\bar{\lambda})$. To show now that the multiplier mapping $M_{\bar{x}}$ from* (5.1) *is not calm at $((0,0),\bar{\lambda})$, select $\lambda^k := \big(-1, t_k, \sqrt{1-t_k^2}\big)$ with $t_k \uparrow 1$ as $k\to\infty$, which yields $\lambda^k \to \bar{\lambda}$ as $k\to\infty$ and $\lambda^k \in -\mathcal{Q}$ for all $k\in\mathbb{N}$. Direct calculations give us the expressions*

$$\operatorname{dist}^2\big(\lambda^k;\Lambda(\bar{x})\big) = \Big\|\lambda^k - \frac{\langle\lambda^k,\bar{\lambda}\rangle}{\|\bar{\lambda}\|^2}\bar{\lambda}\Big\|^2 = \frac{3-2t_k-t_k^2}{2} \ \text{ and}$$

$$\|\nabla f(\bar{x}) + \nabla\Phi(\bar{x})^*\lambda^k\|^2 = (t_k-1)^2,$$

which lead us to the limit calculations

$$\lim_{k\to\infty}\frac{\operatorname{dist}^2\big(\lambda^k;\Lambda(\bar{x})\big)}{\|\nabla f(\bar{x}) + \nabla\Phi(\bar{x})^*\lambda^k\|^2} = \frac{1}{2}\lim_{k\to\infty}\frac{3-2t_k-t_k^2}{(t_k-1)^2} = \infty.$$

This tells us that the multiplier mapping $M_{\bar{x}}$ is not calm at $((0,0),\bar{\lambda})$.

Next we check that the primal estimate (5.6) *fails in this example. To proceed, take $x^k := (0,\alpha_k)$ with $\alpha_k := -(t_k-1)/2$ and observe that $(x^k,\lambda^k)\to(\bar{x},\bar{\lambda})$ as $k\to\infty$. This yields*

$$\nabla f(x^k) + \nabla\Phi(x^k)^*\lambda^k = \begin{pmatrix}0\\ 2\alpha_k\end{pmatrix} + \begin{pmatrix}0\\ t_k-1\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix} = o(\|x^k-\bar{x}\|). \tag{5.14}$$

On the other hand, since $\Phi(x^k)$ is a nonzero point on the boundary of $\mathcal{Q}$ and λ^k is a nonzero *point on the boundary of* $-\mathcal{Q}$*, it follows that*

$$\Phi(x^k)+\lambda^k=\alpha_k(1,1,0)+\Big(-1,t_k,\sqrt{1-t_k^2}\Big)\notin\mathcal{Q}\cup-\mathcal{Q}.$$

Letting $y^k:=\Phi(x^k)+\lambda^k$*, we calculate that*

$$\Pi_{\mathcal{Q}}(y^k)=\frac{1}{2}\Big(y_0^k+\|y_r^k\|\Big)\Big(1,\frac{y_r^k}{\|y_r^k\|}\Big)$$

and then easily check as $k\to\infty$ *that*

$$\Big(1,\frac{y_r^k}{\|y_r^k\|}\Big)\to(1,1,0)\quad\textit{and}\quad\lim_{k\to\infty}\frac{y_0^k+\|y_r^k\|}{\|x^k-\bar{x}\|}=\lim_{k\to\infty}\frac{\alpha_k-1+\sqrt{\alpha_k^2+2\alpha_k t_k+1}}{\alpha_k}=2.$$

This allows us to compute the limits

$$\begin{aligned}\lim_{k\to\infty}\frac{\|\Phi(x^k)-\Pi_{\mathcal{Q}}(\Phi(x^k)+\lambda^k)\|}{\|x^k-\bar{x}\|}&=\lim_{k\to\infty}\Big\|\frac{\Phi(x^k)}{\alpha_k}-\frac{y_0^k+\|y_r^k\|}{2\|x^k-\bar{x}\|}\Big(1,\frac{y_r^k}{\|y_r^k\|}\Big)\Big\|\\&=\|(1,1,0)-(1,1,0)\|=0.\end{aligned}$$

Combining the latter with (5.14) *demonstrates that the primal estimate* (5.6) *and hence the error bound* (5.5) *both fail in this simple example.*

Let us now turn our attention to efficient conditions that ensure the fulfillment of the imposed calmness of the multiplier mapping (5.1). First we provide an improvement of a result established recently in [44, Theorem 4.1], which gives a complete characterization of the calmness property of (5.1) together with the *uniqueness* of Lagrange multipliers in terms of the *dual qualification condition* that involves the graphical derivative (2.5) of the normal cone mapping for (1.2). To proceed, consider the *fully perturbed* set of Lagrange multipliers $M\colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{m+1} \rightrightarrows \mathbb{R}^{m+1}$, where—in contrast to $M_{\bar{x}}(v,w)$ in (5.1)—the decision variable x is also included in the perturbation procedure. We define this mapping by

$$M(x,v,w) := \big\{\lambda \in \mathbb{R}^{m+1} \;\big|\; \nabla_x L(x,\lambda) = v,\ \lambda \in N_{\mathcal{Q}}\big(\Phi(x)+w\big)\big\} \tag{5.15}$$

for $(x,v,w) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{m+1}$ and observe that $M(\bar{x},0,0) = M_{\bar{x}}(0,0) = \Lambda(\bar{x})$. The next proposition provides a full characterization of the *upper Lipschitzian* property of the fully perturbed multiplier mapping M via the dual qualification condition, which plays a key role in the convergent analysis of the ALM method for SOCP (1.1).

Proposition 5.3 (calmness and uniqueness of Lagrange multipliers). *Let $(\bar{x},\bar{\lambda})$ be a solution to the KKT system* (1.4). *Then the following assertions are equivalent:*

(i) *The multiplier mapping $M_{\bar{x}}$ is calm at $((0,0),\bar{\lambda})$, and $\Lambda(\bar{x}) = \{\bar{\lambda}\}$, i.e., the mapping $M_{\bar{x}}$ has the isolated calmness property at $(\bar{x},\bar{\lambda})$.*

(ii) *We have the dual qualification condition*

$$DN_{\mathcal{Q}}\big(\Phi(\bar{x}),\bar{\lambda}\big)(0) \cap \ker \nabla\Phi(\bar{x})^* = \{0\}. \tag{5.16}$$

(iii) *There exist positive numbers γ_3 and κ_3 such that the upper Lipschitzian estimate*

$$M(x,v,w) \subset \{\bar{\lambda}\} + \kappa_3(\|x-\bar{x}\| + \|v\| + \|w\|)\mathbb{B} \ \textit{for all}\ \ (x,v,w) \in \mathbb{B}_{\gamma_3}(\bar{x},0,0) \tag{5.17}$$

holds for the fully perturbed multiplier mapping (5.15).

Proof. The equivalence between (i) and (ii) was established in [44, Theorem 4.1]. Also it is not hard to see that (iii) implies (i) since $M(\bar{x}, 0, 0) = \Lambda(\bar{x})$. Thus it remains to verify the last implication (ii) $\Longrightarrow$ (iii). Observe to this end due to Corollary 2.4 that

$$DN_{\mathcal{Q}}\big(\Phi(\bar{x}), \bar{\lambda}\big)(0) = N_{K_{\mathcal{Q}}(\Phi(\bar{x}),\bar{\lambda})}(0) = K_{\mathcal{Q}}\big(\Phi(\bar{x}), \bar{\lambda}\big)^* = \big(T_{\mathcal{Q}}\big(\Phi(\bar{x})\big) \cap \{\bar{\lambda}\}^{\perp}\big)^*,$$

which in turn yields the inclusion

$$N_{\mathcal{Q}}\big(\Phi(\bar{x})\big) = T_{\mathcal{Q}}\big(\Phi(\bar{x})\big)^* \subset DN_{\mathcal{Q}}\big(\Phi(\bar{x}), \bar{\lambda}\big)(0).$$

Then the dual qualification (5.3) ensures the fulfillment of the basic constraint qualification

$$N_{\mathcal{Q}}\big(\Phi(\bar{x})\big) \cap \ker \nabla\Phi(\bar{x})^* = \{0\},$$

which implies that the Lagrange multiplier sets $M(x, v, w)$ are uniformly bounded for all (x, v, w) in some neighborhood $\mathcal{U}$ of the nominal triple $(\bar{x}, 0, 0)$.

Having this in hand and arguing by contraposition, suppose on the contrary that the upper Lipschitzian property (5.17) fails. The equivalence between (i) and (ii) readily implies that $M(\bar{x}, 0, 0) = \Lambda(\bar{x}) = \{\bar{\lambda}\}$. Thus it follows from the contraposition assumption that there exist sequences of $(x^k, v^k, w^k) \to (\bar{x}, 0, 0)$ as $k \to \infty$ and of the corresponding multipliers $\lambda^k \in M(x^k, v^k, w^k)$ satisfying the inequality

$$\|\lambda^k - \bar{\lambda}\| > k(\|x^k - \bar{x}\| + \|v^k\| + \|w^k\|) \;\text{ whenever }\; k \in \mathbb{N}. \tag{5.18}$$

Suppose without loss of generality that $(x^k, v^k, w^k) \in \mathcal{U}$ for all $k \in \mathbb{N}$. Hence the sequence

$\{\lambda^k\}$ is bounded, and so it has a limiting point $\widehat{\lambda}$. Taking into account the robustness (closed graph property) of the normal cone mapping $N_\mathcal{Q}$ with respect to perturbations of the initial point, the continuity of the mappings $\Phi, \nabla f$, and $\nabla\Phi$ as well as the convergence $(x^k, v^k, w^k) \to (\bar{x}, 0, 0)$, we arrive at $\widehat{\lambda} \in \Lambda(\bar{x}) = \{\bar{\lambda}\}$, which tells us that $\lambda^k \to \bar{\lambda}$ as $k \to \infty$. Letting now $t_k := \|\lambda^k - \bar{\lambda}\|$ ensures that $t_k \downarrow 0$ and allows us to conclude by (5.18) that

$$x^k - \bar{x} = o(t_k), v^k = o(t_k) \text{ and } w^k = o(t_k) \text{ as } k \to \infty. \tag{5.19}$$

Furthermore, the passage to a subsequence if necessary gives us a vector $\eta \in \mathbb{R}^{m+1} \setminus \{0\}$ such that $\dfrac{\lambda^k - \bar{\lambda}}{t_k} \to \eta$. Recalling that $\lambda^k \in M(x^k, v^k, w^k)$, we get

$$\begin{aligned} o(t_k) = v^k &= \nabla f(x^k) + \nabla\Phi(x^k)^*\lambda^k \\ &= \nabla f(x^k) - \nabla f(\bar{x}) + \nabla\Phi(x^k)^*\lambda^k - \nabla\Phi(\bar{x})^*\bar{\lambda} \\ &= \nabla f(x^k) - \nabla f(\bar{x}) + \big(\nabla\Phi(x^k) - \nabla\Phi(\bar{x})\big)^*\lambda^k + \nabla\Phi(\bar{x})(\lambda^k - \bar{\lambda}) \\ &= o(t_k) + \nabla\Phi(\bar{x})(\lambda^k - \bar{\lambda}), \end{aligned}$$

where the verification of the last equality uses the Lipschitz continuity of ∇f and $\nabla\Phi$ around $\bar{x}$, the boundedness of $\{\lambda^k\}$, and the first estimate in (5.19). Dividing both sides of the latter by t_k and passing to the limit as $k \to \infty$ result in $\eta \in \ker \nabla\Phi(\bar{x})^*$. On the other hand, we have

$$\Big(\frac{\Phi(x^k) + w^k - \Phi(\bar{x})}{t_k}, \frac{\lambda^k - \bar{\lambda}}{t_k}\Big) = \frac{\big(\Phi(x^k) + w^k, \lambda^k\big) - \big(\Phi(\bar{x}), \bar{\lambda}\big)}{t_k} \in \frac{\operatorname{gph} N_\mathcal{Q} - \big(\Phi(\bar{x}), \bar{\lambda}\big)}{t_k},$$

which yields $(0,\eta)\in T_{\mathrm{gph}\,N_{\mathcal Q}}(\Phi(\bar x),\bar\lambda)$ and hence verifies the condition

$$\eta\in DN_{\mathcal Q}\big(\Phi(\bar x,\bar\lambda)\big)(0)\cap\ker\nabla\Phi(\bar x)^*.$$

Since $\eta\neq 0$, the latter contradicts (5.16) and thus justifies the claimed estimate (5.17).

□

A different sufficient condition for the upper Lipschitzian property (5.17) was obtained in [6, Proposition 4.47] by using a condition called the “strict constraint qualification." This condition is strictly more restrictive than the dual qualification (5.16), which—as shown in Proposition 5.3—is indeed equivalent to the upper Lipschitzian estimate in (5.17).

Our next goal is to provide a more detailed analysis of the calmness of the multiplier mapping for (1.1) entirely via the given SOCP data at the fixed solution $(\bar x,\bar\lambda)$ to the KKT system (1.4). Consider all the possible cases. If $\Phi(\bar x)\in\mathrm{int}\,\mathcal Q$, then it follows from the normal cone representation (2.3) that $\Lambda(\bar x)=\{0\}$ for the set of Lagrange multipliers in (1.5). Since $M_{\bar x}(0,0)=\Lambda(\bar x)$ and since $M_{\bar x}(u,v)=\{0\}$ whenever the pair (u,v) is sufficiently close to $(0,0)$, we surely get the calmness of the multiplier mapping at $((0,0),\bar\lambda)$ with $\bar\lambda=0$ in this case. If further $\Phi(\bar x)\in(\mathrm{bd}\,\mathcal Q)\setminus\{0\}$, then it follows from (2.3) that $\Lambda(\bar x)$ is the intersection of two polyhedral convex sets. Employing the classical Hoffman lemma ensures that

$$\mathrm{dist}\big(\lambda;\Lambda(\bar x)\big)=O\big(\|\nabla_x L(\bar x,\lambda)\|+\mathrm{dist}(\lambda;N_{\mathcal Q}(\Phi(\bar x))\big)=O\big(\|\nabla_x L(\bar x,\lambda)\|+\mathrm{dist}(\Phi(\bar x);N_{\mathcal Q}^{-1}(\lambda)\big)$$

for all λ close enough to $\bar\lambda\in N_{\mathcal Q}(\Phi(\bar x))$, where the last equality comes from the fact that the mapping $N_{\mathcal{Q}}$ is clearly calm at $(\Phi(\bar{x}), \bar{\lambda})$ in this case. This again verifies the calmness property of the multiplier mapping (5.1) at $((0,0), \bar{\lambda})$.

Considering further the remaining case where $\Phi(\bar{x}) = 0$, we deduce from Proposition 3.1 that the set of Lagrange multipliers $\Lambda(\bar{x})$ admits one of the following representations:

(a) The strict complementarity holds for $\Lambda(\bar{x})$, i.e., $\Lambda(\bar{x})$ contains an interior point of $-\mathcal{Q}$.

(b) $\Lambda(\bar{x}) = \{0\}$.

(c) $\Lambda(\bar{x}) = \{\bar{\lambda}\}$ and $\bar{\lambda} \in \mathrm{bd}\,(-\mathcal{Q}) \setminus \{0\}$.

(d) $\Lambda(\bar{x}) = \mathbb{R}_+\bar{\lambda}$ and $\bar{\lambda} \in \mathrm{bd}\,(-\mathcal{Q}) \setminus \{0\}$.

The next proposition describes the calmness of multiplier mapping for (1.1) when $\Phi(\bar{x}) = 0$.

Proposition 5.4 (calmness of SOCP multipliers at vertex). *Let $(\bar{x}, \bar{\lambda})$ be a solution for the generalized KKT system* (1.4)*, and let $\Phi(\bar{x}) = 0$. The following hold:*

(i) *In cases* (a) *and* (b) *for $\Lambda(\bar{x})$ the multiplier mapping $M_{\bar{x}}$ is calm at $((0,0), \bar{\lambda})$.*

(ii) *In case* (c) *for $\Lambda(\bar{x})$ the calmness of $M_{\bar{x}}$ at $((0,0), \bar{\lambda})$ is equivalent to the full rank of $\nabla\Phi(\bar{x})$.*

Proof. In case (a) we get from Proposition 3.1 that estimate (5.2) is satisfied, which verifies the claimed calmness property of the multiplier mapping. In case (b) it follows from (1.5) that $\nabla f(\bar{x}) = 0$, which yields the equalities

$$-\mathcal{Q} \cap \ker \nabla\Phi(\bar{x})^* = N_{\mathcal{Q}}\big(\Phi(\bar{x})\big) \cap \ker \nabla\Phi(\bar{x})^* = \Lambda(\bar{x}) = \{0\}, \tag{5.20}$$

and so $\bar{\lambda} = 0$ and $K_{\mathcal{Q}}(\Phi(\bar{x}), \bar{\lambda}) = T_{\mathcal{Q}}(0) = \mathcal{Q}$. By Corollary 2.4 we have

$$DN_{\mathcal{Q}}\big(\Phi(\bar{x}), \bar{\lambda}\big)(0) = N_{K_{\mathcal{Q}}(\Phi(\bar{x}),\bar{\lambda})}(0) = -\mathcal{Q}.$$

This together with (5.20) tells us the dual qualification condition (5.16) holds in this case. Employing Proposition 5.3 confirms the calmness of the multiplier mapping $M_{\bar{x}}$ at $((0,0), \bar{\lambda})$.

Finally, consider case (c). If $\nabla\Phi(\bar{x})$ has full rank, then the dual qualification condition (5.16) is satisfied. Hence Proposition 5.3 ensures that the multiplier mapping $M_{\bar{x}}$ is calm at $((0,0), \bar{\lambda})$. Conversely, the validity of the calmness property for $M_{\bar{x}}$ in the framework of (c) implies by Proposition 5.3 that the dual qualification condition (5.16) holds. Combining this with the fact that $\bar{\lambda} \in \mathrm{bd}\,(-\mathcal{Q}) \setminus \{0\}$ in (c) confirms that the matrix $\nabla\Phi(\bar{x})$ has full rank; see Theorem 3.5 for the verification of this claim. This completes the proof of the proposition. $\square$

The above discussions paint a clear picture for the calmness of the multiplier mapping in all the possible cases but (d). It has not been clarified at this stage how to provide verifiable conditions ensuring the calmness property of $M_{\bar{x}}$ in case (d).

5.2 Well-Posedness and Convergence Analysis of ALM for SOCPs

In this concluding section of the dissertation we apply the suggested approach and results of second-order variational analysis (which are undoubtedly of their independent interest) to the convergence analysis of the *augmented Lagrangian method* for solving SOCPs (1.1).

The principal idea of the ALM for (1.1) is to solve a sequence of *unconstrained mini-*

mization problems for which the objective functions, at each iteration, are approximations of the augmented Lagrangian (1.7). Namely, given the current iteration (x^k, λ^k, ρ_k), the ALM solves the following unconstrained problem (called a *subproblem*):

$$\text{minimize } \ \mathcal{L}(x, \lambda^k, \rho_k) \ \text{ for } \ x \in \mathbb{R}^n \tag{5.21}$$

for next primal iterate x^{k+1} and then use it to construct the next dual iterate λ^{k+1}. More specifically, we aim at solving the *stationary equation*

$$\nabla_x \mathcal{L}(x, \lambda^k, \rho_k) = 0 \tag{5.22}$$

for x^{k+1} and then to update the corresponding multiplier by $\lambda^{k+1} := \Pi_{-\mathcal{Q}}(\rho_k \Phi(x^{k+1}) + \lambda^k)$.

Since solving (5.22) is not easy in practice, it is more convenient to choose an approximate solution x^{k+1} satisfying the *approximate stationary condition*

$$\|\nabla_x \mathcal{L}(x^{k+1}, \lambda^k, \rho_k)\| \leq \varepsilon_k \tag{5.23}$$

with a given accuracy/tolerance $\varepsilon_k \geq 0$. Following the conventional terminology of nonlinear programming, we say that the ALM is *exact* of $\varepsilon_k = 0$, i.e., the exact stationary equation (5.22) is used, and *inexact* if (5.23) with $\varepsilon_k > 0$ is under consideration. In this paper we deal with both exact and inexact versions of the ALM by choosing an arbitrary accuracy $\varepsilon_k \geq 0$ sufficiently small. The ALM algorithm for (1.1) is described as follows.

Algorithm 5.5 (augmented Lagrangian method for SOCPs). *Choose* $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^{m+1}$ *and* $\bar{\rho} > 0$. *Pick* $\varepsilon_k \to 0$ *as* $k \to \infty$ *and* ρ_k *with* $\rho_k \geq \bar{\rho}$ *for all* k *and set* $k := 0$. *Then:*

(1) *If* (x^k, λ^k) *satisfies a suitable termination criterion, stop.*

(2) *Otherwise, find* x^{k+1} *satisfying* (5.23) *and update the Lagrange multiplier by*

$$\lambda^{k+1} := \Pi_{-\mathcal{Q}}\big(\rho_k \Phi(x^{k+1}) + \lambda^k\big). \tag{5.24}$$

(3) *Set* $k \leftarrow k+1$ *and go to Step 1.*

To perform the well-posedness and convergence analysis of Algorithm 5.5, we need to make sure first of all that the ALM is *well-defined*, i.e., its subproblems constructed in (5.21) are *solvable*. The following theorem reveals that the optimal solution mappings to subproblems (5.21) enjoy the robust isolated calmness property uniformly in ρ. This confirms, in particular, that subproblems (5.21) always admit a local optimal solution. Note that the developed proof of the theorem requires only the second-order growth condition (4.13), which is based on SOSC (4.9), without any additional assumptions.

Theorem 5.6 (solvability and robust stability of subproblems in ALM). *Let* $\rho_{\bar\lambda}$*,* $\gamma_{\bar\lambda}$*, and* $\ell_{\bar\lambda}$ *be positive constants for which the second-order growth condition* (4.13) *holds whenever* $\rho \ge \rho_{\bar\lambda}$*. Then there exist constants* $\ell > 0$*,* $\widehat\gamma \in (0, \gamma_{\bar\lambda}]$*, and* $\varepsilon > 0$ *such that the local optimal solution mapping* $S_\rho \colon \mathbb{R}^{m+1} \to \mathbb{R}^n$ *defined by*

$$S_\rho(\lambda) := \operatorname{argmin}\big\{\mathcal{L}(x, \lambda, \rho) \;\big|\; x \in \mathbb{B}_{\widehat\gamma}(\bar x)\big\}, \quad \lambda \in \mathbb{R}^{m+1}, \tag{5.25}$$

satisfies, for all $\lambda \in \mathbb{B}_\varepsilon(\bar\lambda)$ *and all* $\rho \in [\rho_{\bar\lambda}, \infty)$*, the inclusions*

$$S_\rho(\lambda) \subset \{\bar x\} + \ell\|\lambda - \bar\lambda\|\mathbb{B} \;\; \textit{and} \;\; \emptyset \ne S_\rho(\lambda) \subset \operatorname{int}\mathbb{B}_{\widehat\gamma}(\bar x), \tag{5.26}$$

which, in particular, implies that the mapping S_ρ enjoys the isolated calmness property at $(\bar{x},\bar{\lambda})$ uniformly in ρ on the interval $[\rho_{\bar{\lambda}},\infty)$.

Proof. Since Φ is twice differentiable at $\bar{x}$, there are constants $\widehat{\gamma}\in(0,\gamma_{\bar{\lambda}}]$ and $\kappa>0$ with

$$\|\Phi(x)-\Phi(\bar{x})\|\le\kappa\|x-\bar{x}\|\ \text{ for all }\ x\in\mathbb{B}_{\widehat{\gamma}}(\bar{x}). \tag{5.27}$$

Employing the second-order growth condition (4.13) tells us that $S_\rho(\bar{\lambda})\cap\mathbb{B}_{\gamma_{\bar{\lambda}}}(\bar{x})=\{\bar{x}\}$ for all $\rho\ge\rho_{\bar{\lambda}}$. Define now the the positive constant

$$\ell:=\frac{\kappa}{\ell_{\bar{\lambda}}}+\sqrt{\frac{\kappa^2}{\ell_{\bar{\lambda}}^2}+\frac{1}{\ell_{\bar{\lambda}}\rho_{\bar{\lambda}}}}, \tag{5.28}$$

select a positive number $\varepsilon<\ell^{-1}\widehat{\gamma}$, and then pick any $\lambda\in\mathbb{B}_\varepsilon(\bar{\lambda})$ and $\rho\ge\rho_{\bar{\lambda}}$. Observe further that for all such λ and ρ we have $S_\rho(\lambda)\neq\emptyset$, since the optimization problem in (5.25) admits an optimal solution by the classical Weierstrass theorem. Fix any $u\in S_\rho(\lambda)$ and recall from Proposition 1.1(ii) that the function $\lambda\mapsto\mathcal{L}(u,\lambda,\rho)$ is concave. This together with (1.9) yields

$$\begin{aligned}\mathcal{L}(u,\lambda,\rho) &\ge \mathcal{L}(u,\bar{\lambda},\rho)-\langle\nabla_\lambda\mathcal{L}(u,\lambda,\rho),\bar{\lambda}-\lambda\rangle\\ &= \mathcal{L}(u,\bar{\lambda},\rho)-\rho^{-1}\big\langle\Pi_{-\mathcal{Q}}\big(\rho\Phi(u)+\lambda\big)-\lambda,\bar{\lambda}-\lambda\big\rangle\\ &\ge f(\bar{x})+\ell_{\bar{\lambda}}\|u-\bar{x}\|^2-\rho^{-1}\big\langle\Pi_{-\mathcal{Q}}\big(\rho\Phi(u)+\lambda\big)-\lambda,\bar{\lambda}-\lambda\big\rangle,\end{aligned} \tag{5.29}$$

where we use (4.13) for the last inequality. It follows from the optimality of u that

$$\mathcal{L}(u,\lambda,\rho)\le\mathcal{L}(\bar{x},\lambda,\rho)=f(\bar{x})+\frac{\rho}{2}\text{dist}^2\big(\Phi(\bar{x})+\rho^{-1}\lambda;\mathcal{Q}\big)-\frac{1}{2}\rho^{-1}\|\lambda\|^2\le f(\bar{x}),$$

which together with (5.29) brings us to the estimate

$$\|u-\bar{x}\|^2 \le \frac{1}{\rho\ell_{\bar{\lambda}}}\left\langle \Pi_{-\mathcal{Q}}\big(\rho\Phi(u)+\lambda\big)-\lambda, \bar{\lambda}-\lambda\right\rangle. \tag{5.30}$$

Employing the projection properties (P2) and (P4) from Sect. 1.3, we get

$$\begin{aligned}
\big\|\Pi_{-\mathcal{Q}}\big(\rho\Phi(u)+\lambda\big)-\lambda\big\| &= \big\|\rho\Phi(u)+\lambda-\Pi_{\mathcal{Q}}\big(\rho\Phi(u)+\lambda\big)-\lambda\big\| \\
&= \|\rho\Phi(u)-\Pi_{\mathcal{Q}}\big(\rho\Phi(u)+\lambda\big)\| \\
&= \|\rho\left(\Phi(u)-\Phi(\bar{x})\right)+\Pi_{\mathcal{Q}}\big(\rho\Phi(\bar{x})+\bar{\lambda}\big)-\Pi_{\mathcal{Q}}\big(\rho\Phi(u)+\lambda\big)\| \\
&\le \rho\|\Phi(u)-\Phi(\bar{x})\|+\rho\|\Phi(u)-\Phi(\bar{x})\|+\|\bar{\lambda}-\lambda\| \\
&\le 2\rho\kappa\|u-\bar{x}\|+\|\bar{\lambda}-\lambda\|,
\end{aligned}$$

where the last inequality comes from (5.27). Using this and (5.30) tells us that

$$\|u-\bar{x}\|^2 \le \frac{1}{\rho\ell_{\bar{\lambda}}}\Big(2\rho\kappa\|u-\bar{x}\|+\|\lambda-\bar{\lambda}\|\Big)\|\lambda-\bar{\lambda}\|,$$

which can be written in the equivalent form as

$$\ell_{\bar{\lambda}}\|u-\bar{x}\|^2-2\kappa\|\lambda-\bar{\lambda}\|\cdot\|u-\bar{x}\|-\frac{\|\lambda-\bar{\lambda}\|^2}{\rho}\le 0.$$

This in turn gives us the estimate

$$\|u-\bar{x}\| \le \left(\frac{\kappa}{\ell_{\bar{\lambda}}}+\sqrt{\frac{\kappa^2}{\ell_{\bar{\lambda}}^2}+\frac{1}{\ell_{\bar{\lambda}}\rho}}\right)\|\lambda-\bar{\lambda}\| \le \ell\|\lambda-\bar{\lambda}\| \le \ell\varepsilon < \widehat{\gamma},$$

which simultaneously verifies both inclusions in (5.26) and thus completes the proof. □

It follows from Theorem 5.6 that, at each iteration k, the condition $\lambda^k \in \mathbb{B}_\varepsilon(\bar{\lambda})$ on the current multiplier in Algorithm 5.5 allows us to find an exact local solution to the optimization problem (5.21) such that $\|u^k - \bar{x}\| \le \ell\|\lambda^k - \bar{\lambda}\|$. Then the Lipschitz continuity of $\nabla_x \mathcal{L}(\cdot, \lambda^k, \rho_k)$ around u^k ensures that for any $\varepsilon_k \ge 0$ we can get an ε_k-solution x^{k+1} satisfying both the approximate stationary condition (5.23) and the same estimate

$$\|x^{k+1} - \bar{x}\| \le \ell\|\lambda^k - \bar{\lambda}\| \tag{5.31}$$

as the exact solution u^k to the optimization problem (5.21) under consideration.

Now we are ready to proceed with local convergence analysis of Algorithm 5.5, which mainly exploits the two major ingredients and the corresponding results developed above: **(1)** *SOSC* (4.9) at $(\bar{x}, \bar{\lambda})$ and the associated *second-order growth* of the augmented Lagrangian, and **(2)** the *calmness* of the multiplier mapping. In addition, we assume that the set of Lagrange multipliers is a *singleton* in the *most interesting case* where $\Phi(\bar{x}) = 0$. The main reason for imposing this restriction is that the convergent analysis of the general case is conducted by using an iterative framework proposed by Fisher in [15, Theorem 1]. However, the latter result demands an error bound estimate the for consecutive terms of the ALM method. Deriving such an estimate for SOCPs with $\Phi(\bar{x}) = 0$ is our ongoing research project. When $\Phi(\bar{x}) \neq 0$, the desired estimate for the consecutive terms in the ALM algorithm can be established by using the *uniform growth condition* from Theorem 4.7 without the uniqueness requirement for Lagrange multipliers, while we omit this consideration in what follows by taking into account the size of the paper. Note that for NLPs such an analysis has been conducted by Fernández and Solodov [14].

The following theorem establishes the linear convergence of Algorithm 5.5 in both exact and inexact frameworks of the ALM with an arbitrarily chosen tolerance in (5.23) in the form $\varepsilon_k = o\big(\sigma(x^k,\lambda^k)\big)$, where $\sigma(x,\lambda)$ is the error bound from (5.3).

Theorem 5.7 (primal-dual convergence of ALM). *Let $(\bar{x},\bar{\lambda})$ be a solution to the KKT system* (1.4)*, let SOSC* (4.9) *hold at $(\bar{x},\bar{\lambda})$, and let the multiplier mapping $M_{\bar{x}}$ from* (5.1) *be calm at $((0,0),\bar{\lambda})$ and $\Lambda(\bar{x}) = \{\bar{\lambda}\}$. Then there exist positive numbers $\bar{\gamma}$ and $\bar{\rho}$ ensuring the following: for any starting point $(x^0,\lambda^0) \in \mathbb{B}_{\bar{\gamma}}(\bar{x},\bar{\lambda})$ and any $\rho_k \geq \bar{\rho}$, Algorithm* 5.5 *generates a sequence of iterates (x^k,λ^k) with a tolerance in* (5.23) *arbitrary chosen as $\varepsilon_k = o(\sigma(x^k,\lambda^k))$ such that (x^k,λ^k) converges to $(\bar{x},\bar{\lambda})$ as $k \to \infty$, and the rate of this convergence is linear.*

Proof. Let $\rho_{\bar{\lambda}}, \gamma_{\bar{\lambda}}$, $\ell_{\bar{\lambda}}$ be the positive constants taken from Theorem 4.5(iii), and let κ_i and γ_i for $i = 1,2,3$ be positive constants taken from the Lipschitzian estimates (5.5), (5.4), and (5.17), respectively. Picking the positive constants κ and $\widehat{\gamma}$ from (5.27), ℓ from (5.28), and ε from Theorem 5.6, define the positive numbers

$$\widehat{\gamma}_1 := \min\big\{\gamma_1,\widehat{\gamma}\big\}, \quad \gamma_{1,2} := \max\big\{\gamma_1,\gamma_2\big\}, \quad \gamma := \min\Big\{\gamma_3,\frac{\widehat{\gamma}_1}{2\kappa_3}\Big\}, \tag{5.32}$$

$$\bar{\rho} := \max\big\{\rho_{\bar{\lambda}},2\kappa_1,8\kappa_1^2\kappa_2\big\}, \ \text{ and } \ \bar{\gamma} := \min\Big\{\widehat{\gamma}_1,\gamma_2,\varepsilon,\frac{\bar{\rho}\gamma}{2\sqrt{10}},\frac{\widehat{\gamma}_1}{2\ell},\frac{\gamma}{2\ell(\kappa+1)}\Big\}. \tag{5.33}$$

Assume also without loss of generality that

$$o\big(\sigma(x,\lambda)\big) \leq \min\Big\{\frac{1}{\kappa_2\bar{\rho}},\frac{1}{8\kappa_1\kappa_2}\Big\}\sigma(x,\lambda) \ \text{ whenever } \ (x,\lambda) \in \mathbb{B}_{\gamma_{1,2}}(\bar{x},\bar{\lambda}) \tag{5.34}$$

and then show that for any starting point $(x^0,\lambda^0) \in \mathbb{B}_{\bar{\gamma}}(\bar{x},\bar{\lambda})$ there exists a sequence

$\{(x^k, \lambda^k)\}$ generated by Algorithm 5.5 with any $\rho_k \geq \bar{\rho}$ such that

$$(x^k, \lambda^k) \in \mathbb{B}_{\bar{\gamma}}(\bar{x}, \bar{\lambda}) \text{ for all } \ k \in \mathbb{N} \cup \{0\}. \tag{5.35}$$

Arguing by induction, observe that (5.35) obviously holds for $k = 0$ and suppose that (5.35) is satisfied for some $k \in \mathbb{N}$ with $\rho_k \geq \bar{\rho}$. We are going to verify that (5.35) fulfills for $k+1$. To furnish this, deduce first from (5.33) that $\|\lambda^k - \bar{\lambda}\| \leq \varepsilon$. This together with the remark after the proof of Theorem 5.6 ensures the existence of an approximate solution x^{k+1} with

$$\|\nabla_x \mathcal{L}(x^{k+1}, \lambda^k, \rho_k)\| \leq \varepsilon_k = o\big(\sigma(x^k, \lambda^k)\big),$$

where $\varepsilon_k \geq 0$ can be chosen arbitrary in this form. It follows from (5.31) that the obtained ε_k-solution satisfies the estimates

$$\|x^{k+1} - \bar{x}\| \leq \ell \|\lambda^k - \bar{\lambda}\| \leq \ell \bar{\gamma} \leq \frac{\widehat{\gamma}_1}{2}, \tag{5.36}$$

where the last inequality comes from (5.33). We proceed now to establish a similar estimate for the dual iterate λ^{k+1}. Using (5.24) and the projection property (P4) yields $\lambda^{k+1} \in N_{\mathcal{Q}}(\Phi(x^{k+1}) + \rho_k^{-1}(\lambda^k - \lambda^{k+1}))$ and hence $\lambda^{k+1} \in M(x^{k+1}, v^{k+1}, w^{k+1})$ with $w^{k+1} := \dfrac{\lambda^k - \lambda^{k+1}}{\rho_k}$ and

$$v^{k+1} := \nabla_x L(x^{k+1}, \lambda^{k+1}) = \nabla_x \mathcal{L}(x^{k+1}, \lambda^k, \rho_k) = o\big(\sigma(x^k, \lambda^k)\big). \tag{5.37}$$

The inclusion $(x^k,\lambda^k)\in\mathbb{B}_{\gamma_2}(\bar{x},\bar{\lambda})$ allows us to deduce from (5.4), (5.34), and (5.37) that

$$\|v^{k+1}\|\le\frac{\sigma(x^k,\lambda^k)}{\kappa_2\bar{\rho}}\le\frac{\|x^k-\bar{x}\|+\|\lambda^k-\bar{\lambda}\|}{\bar{\rho}}. \tag{5.38}$$

Employing again the updating scheme (5.24), we arrive at the relationships

$$\begin{aligned}\|w^{k+1}\| &= \|\rho_k^{-1}\lambda^k-\Pi_{-\mathcal{Q}}\big(\Phi(x^{k+1})+\rho_k^{-1}\lambda^k\big)\|\\ &\le \rho_k^{-1}\|\lambda^k-\bar{\lambda}\|+\big\|\Pi_{-\mathcal{Q}}\big(\Phi(x^{k+1})+\rho_k^{-1}\lambda^k\big)-\Pi_{-\mathcal{Q}}\big(\Phi(\bar{x})+\rho_k^{-1}\bar{\lambda}\big)\big\|\\ &\le 2\rho_k^{-1}\|\lambda^k-\bar{\lambda}\|+\|\Phi(x^{k+1})-\Phi(\bar{x})\|\\ &\le 2\bar{\rho}^{-1}\|\lambda^k-\bar{\lambda}\|+\kappa\|x^{k+1}-\bar{x}\|\end{aligned}$$

with the last estimate coming from (5.27) and $x^{k+1}\in\mathbb{B}_{\widehat{\gamma}}(\bar{x})$. Thus (5.36) and (5.38) bring us to

$$\begin{aligned}\|x^{k+1}-\bar{x}\|+\|v^{k+1}\|+\|w^{k+1}\| &\le (\kappa+1)\|x^{k+1}-\bar{x}\|+\rho_{\bar{\lambda}}^{-1}\|x^k-\bar{x}\|+3\rho_{\bar{\lambda}}^{-1}\|\lambda^k-\bar{\lambda}\|\\ &\le \ell(\kappa+1)\|\lambda^k-\bar{\lambda}\|+\sqrt{10}\bar{\rho}^{-1}\|(x^k,\lambda^k)-(\bar{x},\bar{\lambda})\|\le\gamma,\end{aligned}$$

where the last inequality employs the induction assumption (5.35) together with (5.33). This along with $\gamma\le\gamma_3$ due to (5.32) ensures that $(x^{k+1},v^{k+1},w^{k+1})\in\mathbb{B}_{\gamma_3}(\bar{x},0,0)$. Hence we deduce from the upper Lipschitzian property in (5.17) and the definition of γ in (5.32) that

$$\|\lambda^{k+1}-\bar{\lambda}\|\le\kappa_3\big(\|x^{k+1}-\bar{x}\|+\|v^{k+1}\|+\|w^{k+1}\|\big)\le\kappa_3\gamma\le\frac{\widehat{\gamma}_1}{2}$$

verifying therefore the promised estimate for the dual iterate λ^{k+1}. This together with

(5.36) shows that $(x^{k+1},\lambda^{k+1})\in\mathbb{B}_{\widehat{\gamma}_1}(\bar{x},\bar{\lambda})$. Using the latter, the imposed SOSC (4.9), and the calmness of the multiplier mappings $M_{\bar{x}}$ from (5.1), we conclude from Theorem 5.1 that

$$\|x^{k+1}-\bar{x}\|+\|\lambda^{k+1}-\bar{\lambda}\|\le\kappa_1\sigma_{k+1}\ \text{ with}$$

$$\sigma_{k+1}=\|\nabla_x L(x^{k+1},\lambda^{k+1})\|+\|\Phi(x^{k+1})-\Pi_{\mathcal{Q}}\big(\Phi(x^{k+1})+\lambda^{k+1}\big)\|. \tag{5.39}$$

Define further the projection vector

$$p^{k+1}:=\Pi_{\mathcal{Q}}(\Phi(x^{k+1})+\rho_k^{-1}\lambda^k)$$

and deduce from the updating scheme (5.24) that

$$\Phi(x^{k+1})-p^{k+1}=\frac{\lambda^{k+1}-\lambda^k}{\rho_k}. \tag{5.40}$$

Employing the projection properties (P1) and (P2) results in $\langle p^{k+1},\lambda^{k+1}\rangle=0$ due to

$$\rho_k^{-1}\lambda^{k+1}=\Pi_{-\mathcal{Q}}\big(\Phi(x^{k+1})+\rho_k^{-1}\lambda^k\big)=\Phi(x^{k+1})+\rho_k^{-1}\lambda^k-p^{k+1},$$

which together with $p^{k+1}\in\mathcal{Q}$ yields $\lambda^{k+1}\in N_{\mathcal{Q}}(p^{k+1})$. Hence $p^{k+1}=\Pi_{\mathcal{Q}}(p^{k+1}+\lambda^{k+1})$ by property (P4). Since the mapping $y\mapsto y-\Pi_{\mathcal{Q}}(y+\lambda^{k+1})=\Pi_{-\mathcal{Q}}(y+\lambda^{k+1})-\lambda^{k+1}$ is clearly nonexpansive, we arrive at the relationships

$$\begin{aligned}
& \big\|\Phi(x^{k+1}) - \Pi_{\mathcal{Q}}\big(\Phi(x^{k+1}) + \lambda^{k+1}\big)\big\| \\
= \; & \big\|\Phi(x^{k+1}) - \Pi_{\mathcal{Q}}\big(\Phi(x^{k+1}) + \lambda^{k+1}\big)\big\| - \big\|p^{k+1} - \Pi_{\mathcal{Q}}(p^{k+1} + \lambda^{k+1})\big\| \\
\le \; & \big\|\Phi(x^{k+1}) - \Pi_{\mathcal{Q}}\big(\Phi(x^{k+1}) + \lambda^{k+1}\big) - \big(p^{k+1} - \Pi_{\mathcal{Q}}(p^{k+1} + \lambda^{k+1})\big)\big\| \\
\le \; & \|\Phi(x^{k+1}) - p^{k+1}\| \\
\le \; & \rho_k^{-1}\|\lambda^{k+1} - \lambda^k\| \qquad (\text{by } (5.40)) \\
\le \; & \rho_k^{-1}\big(\|\lambda^{k+1} - \bar{\lambda}\| + \|\lambda^k - \bar{\lambda}\|\big) \\
\le \; & \kappa_1\rho_k^{-1}(\sigma_{k+1} + \sigma_k).
\end{aligned}$$

Using this together with (5.37) and (5.39) leads us to the estimates

$$\sigma_{k+1} \le \varepsilon_k + \big\|\Phi(x^{k+1}) - \Pi_{\mathcal{Q}}\big(\Phi(x^{k+1}) + \lambda^{k+1}\big)\big\| \le \varepsilon_k + \frac{\kappa_1}{\rho_k}\Big(\sigma_{k+1} + \sigma_k\Big),$$

which can be equivalently rewritten as

$$\Big(1 - \frac{\kappa}{\rho_k}\Big)\sigma_{k+1} \le \varepsilon_k + \frac{\kappa_1}{\rho_k}\sigma_k.$$

Since $\rho_k \ge \bar{\rho}$, by (5.33), we get $1 - \dfrac{\kappa_1}{\rho_k} > \dfrac{1}{2}$, which ensures that

$$\sigma_{k+1} \le 2\sigma_k\left(\frac{\varepsilon_k}{\sigma_k} + \frac{\kappa_1}{\rho_k}\right).$$

Applying finally the error bounds (5.5) and (5.4) and then appealing to (5.33) and (5.34) yields

$$\begin{aligned}
\|x^{k+1}-\bar{x}\|+\|\lambda^{k+1}-\bar{\lambda}\| &\le \kappa_1\sigma_{k+1}\le 2\kappa_1\left(\frac{\varepsilon_k}{\sigma_k}+\frac{\kappa_1}{\rho_k}\right)\sigma_k\\
&\le 2\kappa_1\kappa_2\left(\frac{\varepsilon_k}{\sigma_k}+\frac{\kappa_1}{\rho_k}\right)\big(\|x^k-\bar{x}\|+\|\lambda^k-\bar{\lambda}\|\big)\\
&\le \frac{1}{2}\big(\|x^k-\bar{x}\|+\|\lambda^k-\bar{\lambda}\|\big),
\end{aligned} \tag{5.41}$$

which together with the induction assumption (5.35) brings us to

$$(x^{k+1},\lambda^{k+1})\in\mathbb{B}_{\bar{\gamma}}(\bar{x},\bar{\lambda}).$$

This finishes our induction argument to justify (5.35) for all $k\in\mathbb{N}$. Observe that the latter inclusion along with (5.33) implies that $\|\lambda^{k+1}-\bar{\lambda}\|\le\varepsilon$ while allowing us to use Theorem 5.6 to construct the next primal iterate x^{k+2}. Since (5.41) holds for all $k\in\mathbb{N}$, we clearly get that $(x^k,\lambda^k)\to(\bar{x},\bar{\lambda})$ as $k\to\infty$. Furthermore, the obtained estimate tells us that rate of convergence of (x^k,λ^k) to $(\bar{x},\bar{\lambda})$ is linear, which therefore completes the proof of the theorem. □

To conclude the chapter, let us compare the convergence analysis of Algorithm 5.5 given in Theorem 5.7 with the one provided recently by Kanzow and Steck [30, 31] for the class of $\mathcal{C}^2$-cone reducible conic programs that includes SOCPs. There are significant differences between Algorithm 5.5 and the ALM method developed in [30, 31]. First and foremost, the latter publications use instead of λ^k a certain vector w^k from a bounded set in the formation of subproblems (5.21). This is different from the classical ALM method for constrained optimization, including NLPs. It seems to us that the main reason for such a change is that the usage of λ^k from the updating scheme (5.24) is essentially more challenging to conduct an adequate convergence analysis of the ALM method, since it requires to prove the uniform boundedness of the sequence of multipliers. While the algorithm in [30, 31] uses a particular updating scheme for the penalty parameter ρ_k, our approach reveals that there is no need to confine the convergence analysis to a particular updating scheme for ρ_k as long as we keep it sufficiently large. Also, as mentioned in Sect. 1.3, the solvability of subproblems (5.21) was not addressed in [30, 31]. Let us finally emphasize that the progress achieved in this paper is largely based on the application and development of powerful tools of second-order variational analysis and generalized differentiation.

REFERENCES

[1] F. Alizadeh and D. Goldfarb, Second-order cone programming, *Math. Program.* **95** (2003), 3–51.

[2] 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), 135–160.

[3] H. H. Bauschke and P. L. Combettes: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York (2011).

[4] D. P. Bertsekas: *Constrained Optimization and Lagrangian Multiplier Method*. Academic Press, New York (1982).

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

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

[7] N. H. Chieu and L. V. Hien, Computation of graphical derivative for a class of normal cone mappings under a very weak condition, *SIAM J. Optim.* **27** (2017), 190–204.

[8] Y. Cui, D. Sun, and K.-C. Toh: On the R-superlinear convergence of the KKT residues generated by the augmented Lagrangian method for convex composite conic programming. *Math. Program.* **178** (2019), 381–415.

[9] 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), 67–90.

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

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

[12] D. Drusvyatskiy and A. S. Lewis: Error bounds, quadratic growth, and linear convergence of proximal methods. *Math. Oper. Res.* (to appear). arXiv:1602.06661.

[13] F. Facchinei and J.-C. Pang: *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer, New York (2003).

[14] D. Fernández and M. V. Solodov: Local convergence of exact and inexact augmented Lagrangian methods under the second-order sufficient optimality condition. *SIAM J. Optim.* **22** (2012), 384–407.

[15] A. Fischer: Local behavior of an iterative framework for generalized equations with nonisolated solutions. *Math. Program.* **94** (2002), 91–124.

[16] H. Gfrerer and B. S. Mordukhovich, Complete characterizations of tilt stability in nonlinear programming under weakest qualification conditions, *SIAM J. Optim.* **25** (2015), 2081–2119.

[17] H. Gfrerer: First-order and second-order characterizations of metric subregularity and calmness of constraint set mappings. *SIAM J. Optim.* **21** (2011), 1439–1474.

[18] H. Gfrerer and B. S. Mordukhovich, Robinson stability of parametric constraint systems via variational analysis, *SIAM J. Optim.* **27** (2017), 438–465.

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

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

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

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

[23] R. Henrion, A. Jourani, and J. V. Outrata, On the calmness of a class of multifunctions, *SIAM J. Optim.* **13** (2002), 603–618.

[24] R. Henrion and J. V. Outrata, Calmness of constraint systems with applications, *Math. Program* **104** (2005), 437–464.

[25] M. R. Hestenes: Multiplier and gradient methods. *J. Optim. Theory Appl.* **4** (1969), 303–320.

[26] A. D. Ioffe and J. V. Outrata, On metric and calmness qualification conditions in subdifferential calculus, *Set-Valued Anal.* **16** (2008), 199–227.

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

[28] A. F. Izmailov, A. S. Kurennoy, and M. S. Solodov: Local convergence of the method of multipliers for variational and optimization problems under the noncriticality assumption. *Comput. Optim. Appl.* **60** (2015), 111–140.

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

[30] C. Kanzow and D. Steck: On error bounds and multiplier methods for variational problems in Banach spaces. *SIAM J. Optim.* **56** (2018), 1716–1738.

[31] C. Kanzow and D. Steck: Improved local convergence results for augmented Lagrangian methods in $\mathcal{C}^2$-cone reducible constrained optimization. *Math. Program.*

177 (2019), 425–438.

[32] A. B. Levy, Sensitivity of solutions to variational inequalities on Banach spaces, *SIAM J. Control Optim.* **38** (1999), 50–60.

[33] Y. J. Liua and L. Zhang: Convergence analysis of the augmented Lagrangian method for nonlinear second-order cone optimization problems. *Nonlin. Anal.* **67** (2007), 1359–1373.

[34] A. Mohammadi, B. S. Mordukhovich, and M. E. Sarabi: Variational analysis of composite models with applications to continuous optimization. *Math. Oper. Res.* (to appear). arXiv:1905.08837v2.

[35] A. Mohammadi, B. S. Mordukhovich, and M. E. Sarabi: Parabolic regularity in geometric variational analysis *Trans. Amer. Math. Soc.* (to appear). arXiv:1909.00241.

[36] A. Mohammadi, B. S. Mordukhovich, and M. E. Sarabi: Stability of KKT systems and superlinear convergene of the SQP method under parabolic regularity. *J. Optim. Theory Appl.* (to appear). arXiv:1910.06894v1.

[37] A. Mohammadi and M. E. Sarabi: Twice epi-differentiability of extended-real-valued functions with applications in composite optimization. *SIAM J. Optim.* (to apear). arXiv:1911.05236.

[38] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic Theory; II: Applications*, Springer, Berlin, 2006.

[39] B. S. Mordukhovich: *Variational Analysis and Applications*. Springer, Cham, Switzerland, 2018.

[40] B. S. Mordukhovich, J. V. Outrata, and H. Ramírez C., Second-order variational analysis in conic programming with applications to optimality and stability, *SIAM J. Op-*

tim. **25** (2015), 76–101.

[41] B. S. Mordukhovich, J. V. Outrata, and H. Ramírez C., Graphical derivatives and stability analysis for parameterized equilibria with conic constraints, *Set-Valued Var. Anal.* **23** (2015), 687–704.

[42] B. S. Mordukhovich, J. V. Outrata, and M. E. Sarabi, Full stability of locally optimal solution in second-order cone programming, *SIAM J. Optim.* **24** (2014), 1581–1613.

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

[44] B. S. Mordukhovich and M. E. Sarabi: Criticality of Lagrange multipliers in variational systems. *SIAM J. Optim.* **29** (2019), 1524–1557.

[45] J. Nocedal and S. J. Wright: *Numerical Optimization*. Springer, New York, 2006.

[46] J. V. Outrata and H. Ramírez, C., On the Aubin property of critical points to perturbed second-order cone programs, *SIAM J. Optim.* **21** (2011), 798–823.

[47] J. V. Outrata and D. Sun, On the coderivative of the projection operator onto the second-order cone, *Set-Valued Anal.* **16** (2008), 999–1014.

[48] M. J. D. Powell: A method for nonlinear constraints in minimization problems. In: *R. Fletcher (ed.): Optimization*, Academic Press, New York (1969), 283–298.

[49] S. M. Robinson: Stability theory for systems of inequalities, II: differentiable nonlinear systems. *SIAM J. Numer. Anal.* **13** (1976), 497–513.

[50] S. M. Robinson, Some continuity properties of polyhedral multifunctions, *Math. Program. Stud.* **14** (1981), 206–214.

[51] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[52] R. T. Rockafellar: A dual approach to solving nonlinear programming problems by unconstrained optimization. *Math. Program.* **5** (1973), 354–373.

[53] R. T, Rockafellar: The multiplier method of Hestenes and Powell applied to convex programming. *J. Optim. Theory Appl.* **12** (1973), 555–562.

[54] R. T. Rockafellar: Augmented Lagrangian multiplier functions and duality in nonconvex programming. *SIAM J. Control* **12** (1974), 268–285.

[55] R. T. Rockafellar, Maximal monotone relations and the second derivatives of nonsmooth functions. *Ann. Inst. H. Pointcaré Analyse Non-Linéaire* **2** (1985), 167–184.

[56] R. T. Rockafellar, Proto-differentiability of set-valued mappings and its applications in optimization, *Ann. Inst. H. Poincaré: Analyse Non Linéaire* **6** (1987), 448–482.

[57] R. T. Rockafellar, First- and second-order epi-differentiability in nonlinear programming, *Trans. Amer. Math. Soc.* **307** (1988), 75–108.

[58] R. T. Rockafellar: Second-order optimality conditions in nonlinear programming obtained by way of epi-derivatives. *Math. Oper. Res.* **14** (1989), 462–484.

[59] R. T. Rockafellar: Lagrange multipliers and optimality. *SIAM Rev.* **35** (1993), 183–238.

[60] R. T. Rockafellar and R. J-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.

[61] A. P. Ruszczynski, *Nonlinear Optimization*, Princeton University Press, Priceton, NJ, 2006.

[62] A. Shapiro and A. S. Nemirovski, *Duality of Linear Conic Problems,* School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, 2003; available in www.optimization-online.org/DB_FILE/2003/12/793.pdf.

[63] A. Shapiro and J. Sun: Some properties of the augmented Lagrangian in cone constrained optimization. *Math. Oper. Res.* **29** (2004), 479–491.

[64] D. Sun, J. Sun, and L. Zhang: The rate of convergence of the augmented Lagrangian method for nonlinear semidefinite programming. *Math. Program.* **114** (2008), 349–391.

[65] J. J. Ye and X. Y. Ye: Necessary optimality conditions for optimization problems with variational inequality constraints. *Math. Oper. Res.* **22** (2007), 977–997.

[66] G. Wachsmuth, Strong stationarity for optimization problems with complementarity constraints in absence of polyhedricity, *Set-Valued Anal.* **25** (2017), 133–175.

ABSTRACT

VARIATIONAL ANALYSIS IN SECOND-ORDER CONE PROGRAMMING WITH APPLICATIONS

by

HANG THI VAN NGUYEN

December 2020

Advisor: Dr. Boris S. Mordukhovich

Major: Applied Mathematics

Degree: Doctor of Philosophy

This dissertation conducts a second-order variational analysis for an important class on nonpolyhedral conic programs generated by the so-called *second-order/Lorentz/ice-cream cone* $\mathcal{Q}$. These *second-order cone programs (SOCPs)* are mathematically challenging due to the nonpolyhedrality of the underlying second-order cone while being important for various applications. The two main devices in our study are *second epi-derivative* and *graphical derivative* of the *normal cone mapping* which are proved to accumulate vital second-order information of functions/constraint systems under investigation. Our main contribution is threefold:

- proving the *twice epi-differentiability* of the indicator function of $\mathcal{Q}$ and of the *augmented Lagrangian* associated with SOCPs, and deriving explicit formulae for the calculation of the second epi-derivatives of both functions;

- establishing a precise formula–entirely via the initial data– for *calculating the graphical derivative* of the *normal cone mapping* generated by the constraint set of SOCPs *without* imposing any *nondegeneracy* condition;

- conducting a complete convergence analysis of the *Augmented Lagrangian Method* (ALM) for SOCPs with *solvability*, *stability* and *local convergence analysis* of both *exact* and *inexact* versions of the ALM under fairly *mild assumptions*.

These results have strong potentials for applications to SOCPs and related problems. Among those presented in this dissertation we mention characterizations of the *uniqueness of Lagrange multipliers* together with an *error bound* estimate for second-order cone constraints; of the *isolated calmness* property for solutions maps of perturbed variational systems associated with SOCPs; and also of *(uniform) second-order growth condition* for the augmented Lagrangian associated with SOCPs.

AUTOBIOGRAPHICAL STATEMENT

HANG THI VAN NGUYEN

Education

- *Ph.D. (expected) in Applied Mathematics*, Wayne State University, MI, USA (2020)
- *M.S. in Applied Mathematics*, University of Limoges, Limoges, France (2011)
- *B.S. in Applied Mathematics*, Hanoi National University of Education, Hanoi, Vietnam (2009)

Awards

- *University Graduate Research Fellowship*, Wayne State University, August 2019–August 2020.
- *Karl W. and Helen L. Folley Endowed Scholarship*, Department of Mathematics, Wayne State University, 2018.
- *Karl W. and Helen L. Folley Endowed Scholarship*, Department of Mathematics, Wayne State University, 2017.
- *Scholarship for International Master Program in Sciences from Vietnamese Government*, Ministry of Education and Training, Vietnam, September 2009–July 2011.

Publications and Preprints

1. N. T. V. Hang, B. S. Mordukhovich, and M. E. Sarabi, *Augmented Lagrangian method for second-order cone programs under second-order sufficiency,* submitted.
2. N. T. V. Hang, B. S. Mordukhovich, and M. E. Sarabi, *Second-order variational analysis in second-order cone programming,* Math. Program. 180 (2020), 75–116.
3. N. T. V. Hang and N. D. Yen, *On the problem of minimizing a difference of polyhedral convex functions under linear constraints,* J. Optim. Theory Appl. 171(2) (2016), 617–642.
4. N. T. V. Hang and J.-C. Yao, *Sufficient conditions for error bounds of difference functions and applications,* J. Glob. Optim. 66(3) (2016), 439–456.
5. N. T. V. Hang and N. D. Yen, *Optimality conditions and stability analysis via the Mordukhovich subdifferential,* Numer. Func. Anal. Optim. 36(3) (2015), 364–386.
6. N. T. V. Hang, *The penalty functions method and multiplier rules based on the Mordukhovich subdifferential,* Set-Valued Var. Anal. 22(2) (2014), 299–312.