Noname manuscript No.
(will be inserted by the editor)

Perturbation Analysis of the Euclidean Distance Matrix Optimization Problem and Its Numerical Implications

Shaoyan Guo · Hou-Duo Qi · Liwei Zhang

Received / Accepted: date

Abstract Euclidean distance matrices have lately received increasing attention in applications such as multidimensional scaling and molecular conformation from nuclear magnetic resonance data in computational chemistry. In this paper, we focus on the perturbation analysis of the Euclidean distance matrix optimization problem (EDMOP). Under Robinson's constraint qualification, we establish a number of equivalent characterizations of strong regularity and strong stability at a locally optimal solution of EDMOP. Those results extend the corresponding characterizations in Semidefinite Programming and are tailored to the special structure in EDMOP. As an application, we demonstrate a numerical implication of the established results on an alternating direction method of multipliers (ADMM) to a stress minimization problem, which is an important instance of EDMOP. The implication is that the ADMM method converges to a strongly stable solution under reasonable assumptions.

Keywords Euclidean distance matrices · strong second order optimality condition · constraint nondegeneracy · strong regularity

Mathematics Subject Classification (2020) 90C33 · 90C25 · 90C20

Dedicated to Prof Asen L. Dontchev – a leader, a friend and a mentor.

S. Guo
School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China. E-mail: syguo@dlut.edu.cn.

H.-D Qi
Department of Applied Mathematics and CAS AMSS-PolyU Joint Lab, The Hong Kong Polytechnic University, Hong Kong. E-mail: houduo.qi@polyu.edu.hk.

L. Zhang
School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China. E-mail: lwzhang@dlut.edu.cn.

1 Introduction

This research has been strongly motivated by two significant advances in matrix optimization over the past decade. On one hand, much has been known about the perturbation analysis for semi-define programming (SDP) in the classical sense of Robinson [31] and Kojima [22]. Important results can be found in Bonnans and Shapiro [4], Sun [35], Chan and Sun [7], and Ding et. al. [10]. On the other hand, Euclidean Distance Matrix (EDM) optimization has appeared to be a powerful model for many practical problems, see the surveys [12, 24]. In particular, EDM optimization has a close link to SDP, see e.g., [1–3, 8, 14]. Moreover, error bounds have proved essential in establishing complexity results for numerical algorithms for EDM optimization [9, 40, 41]. An interesting question is whether we can establish perturbation results in the sense of Robinson [31] and Kojima [22] for general EDM optimization. As reported in this paper, well-formulated results can be obtained thanks to the advances on perturbation analysis in SDP. For general references on sensitivity and stability analysis in optimization, see, e.g., Bonnans and Shapiro [4], Dontchev and Rockafellar [13], Facchinei and Pang [15], Klatte and Kummer [21], Mordukhovich [26], and Rockafellar and Wets [32].

In the following, we first state the general EDM optimization problem, followed by two examples to show its tremendous capability in modeling. The first example is the well-known maximum-variance unfolding (MVU) problem [37, 39]. This example will show an advantage of EDM optimization over SDP. Moreover, MVU in the form of EDM is strongly stable. The second example is from calibrating correlation matrices in [30, Chp. 22] and it is well formulated in the general form of EDM optimization. We finish this section by briefly outlining the main contributions of the paper.

1.1 EDM Optimization: Formulation

The main concept in our formulation is the *conditionally positive semidefinite cone* $\mathcal{K}^n_+$ [25]. Let $\mathcal{S}^n$ and $\mathcal{S}^n_+$ denote respectively the space of $n \times n$ symmetric matrices and the cone of positive semidefinite matrices. $\mathcal{K}^n_+$ is the collection of all matrices in $\mathcal{S}^n$ that are positive semidefinite on the subspace $\mathbf{1}_n^\perp$: the subspace in $\mathbb{R}^n$ orthogonal to the vector of all ones $\mathbf{1}_n$. That is,

$$\mathcal{K}^n_+ := \left\{ A \in \mathcal{S}^n : \mathbf{x}^T A \mathbf{x} \ge 0, \ \mathbf{x} \in \mathbf{1}_n^\perp \right\},$$

where ":=" means "define". Furthermore, if $A \in \mathcal{K}^n_+$ has zero diagonal [33], then there must exist a set of points $\mathbf{x}_i \in \mathbb{R}^r$, $i = 1, \ldots, n$ such that the (i, j)th element of $(-A)$ is given by the squared Euclidean distance between $\mathbf{x}_i$ and $\mathbf{x}_j$, i.e.,

$$(-A)_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|^2, \quad i, j = 1, \ldots, n,$$

where r is called the embedding dimension of A and $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^r$. Such matrix $(\|\mathbf{x}_i - \mathbf{x}_j\|^2)^n_{i,j=1}$ is called Euclidean Distance Matrix

(EDM). The collection of all EDMs of the size $n \times n$ is the EDM (convex) cone $\mathcal{D}^n$, see [8]. Hence, $\mathcal{K}^n_+$ is intrinsically related to Euclidean geometry. The optimization problem below defined on $\mathcal{K}^n_+$ is hence referred to as the EDM optimization:

$$\text{(EDMOP)} \qquad \min_{\mathbf{x}\in\mathcal{X}} \ f(\mathbf{x}) \quad \text{s.t.} \quad \mathrm{diag}\,(G(\mathbf{x})) = 0, \quad G(\mathbf{x}) \in \mathcal{K}^n_+, \qquad (1.1)$$

where $\mathcal{X}$ is a finite-dimensional real Hilbert space, $f : \mathcal{X} \to \mathbb{R}$, $G : \mathcal{X} \to \mathcal{S}^n$ are twice continuously differentiable functions, and $\mathrm{diag}\,(\cdot)$ is a vector containing the diagonal elements of a matrix. We make following comments on the model (1.1).

(i) The constraints specify that $-G(\mathbf{x})$ is EDM. We could add more constraints such as $h(\mathbf{x}) = 0$ with $h(\cdot) : \mathcal{X} \mapsto \mathbb{R}^m$ being twice continuously differentiable. For simplicity of describing our results, we omit such constraints.

(ii) Problem (1.1) can be cast as a SDP by noting that

$$G(\mathbf{x}) \in \mathcal{K}^n_+ \quad \Longleftrightarrow \quad JG(\mathbf{x})J \in \mathcal{S}^n_+,$$

where $J := I - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$ is the orthogonal projection to the subspace $\mathbf{1}_n^\perp$. Perturbation analysis may be conducted on the resulting problem as done in [35] for nonlinear SDP. However, the composition between J and $G(\mathbf{x})$ may make analysis more complicated. For example, it would be hard to check the constraint nondegeneracy property [4] for the nearest EDM problem stated below, while it is straightforward from the EDM perspective. This is all because variational analysis on $\mathcal{K}^n_+$ can be thoroughly done. Note that constraint nondegeneracy is an important property for stability analysis in conic programming.

(iii) Many Euclidean embedding problems including the first example below take G as the identity mapping with $(-\mathbf{x})$ being EDM [42]. Our second example makes a case where $G(\mathbf{x})$ may take a nonlinear form.

1.2 Two Motivating Examples

We use two examples to demonstrate the modeling power of EDMOP. The first example is from nonlinear dimensionality reduction in machine learning and the second is from finance.

(a) Maximum Variance Unfolding (MVU). It was first proposed in [39] as a dimensionality reduction method and it has a close link to the Markovian mixing over graph [37]. The purpose is to embed n points in a low-dimensional Euclidean space while preserving the local structure among data.

Suppose there are n items, which are collected from a high-dimensional space. A certain type of dissimilarity (e.g., distance) may be computed for some pairs of items. Let δ_{ij} denote such a dissimilarity measurement between item i and item j. We let $\mathcal{N}$ be the collection of such pairs. The purpose is to embed those items as n points $\{\mathbf{x}_i\}_{i=1}^n$ in a low-dimensional space $\mathbb{R}^r$ such that the Euclidean distance between $\mathbf{x}_i$ and $\mathbf{x}_j$ approximates δ_{ij}:

$$\|\mathbf{x}_i - \mathbf{x}_j\| \approx \delta_{ij} \quad \forall \quad (i,j) \in \mathcal{N}. \tag{1.2}$$

A key principle in MVU is that it favours the embedding with high variance making the quantity

$$\sigma_X^2 := \frac{1}{2n} \sum_{i,j=1}^n \|\mathbf{x}_i - \mathbf{x}_j\|^2$$

as big as possible. Under the assumption that the embedding points are centered, i.e., $\mathbf{x}_1 + \cdots + \mathbf{x}_n = 0$ (centralization condition), the variance becomes $\sigma_X^2 = \sum_{i=1}^n \|\mathbf{x}_i\|^2$. Therefore, MVU aims to achieve (i) preserving the local distances in (1.2), and (ii) maximizing the variance σ_X^2:

$$\max_{\mathbf{x}_i} \sum_{i=1}^n \|\mathbf{x}_i\|^2 - \nu \sum_{(i,j)\in\mathcal{N}} \Big(\|\mathbf{x}_i - \mathbf{x}_j\|^2 - \delta_{ij}^2 \Big)^2, \tag{1.3}$$

where $\nu > 0$ is a balance parameter between the two aims.

Problem (1.3) is highly nonlinear in the space of embedding points $\mathbf{x}_i$. Fortunately, it has a nice SDP relaxation. Note that the squared distance $\|\mathbf{x}_i - \mathbf{x}_j\|^2$ has a linear representation in terms of a kernel matrix:

$$\|\mathbf{x}_i - \mathbf{x}_j\|^2 = \|\mathbf{x}_i\|^2 + \|\mathbf{x}_j\|^2 - 2\langle \mathbf{x}_i, \mathbf{x}_j \rangle = K_{ii} + K_{jj} - 2K_{ij},$$

where the kernel matrix K is defined by $K_{ij} = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$ with $\langle \cdot, \cdot \rangle$ being the standard dot product in $\mathbb{R}^r$. Consequently, (1.3) can be represented as a convex quadratic SDP (dropping the hidden rank constraint $\mathrm{rank}\,(K) = r$):

$$\text{(MVU)} \quad \begin{array}{l} \max_K \ \mathrm{Trace}\,(K) - \nu \sum_{(i,j)\in\mathcal{N}} \Big(K_{ii} + K_{jj} - 2K_{ij} - \delta_{ij}^2 \Big)^2, \\ \text{s.t. } K\mathbf{1}_n = 0, \quad K \succeq 0, \end{array} \tag{1.4}$$

where $K \succeq 0$ means that K is positive semidefinite (i.e., being a kernel), and the constraint $K\mathbf{1}_n = 0$ is the centralization condition.

Now, let us reformulate problem (1.3) as EDM optimization in terms of the EDM $D_{ij} := \|\mathbf{x}_i - \mathbf{x}_j\|^2$:

$$\begin{array}{ll} \min_D & \nu \sum_{(i,j)\in\mathcal{N}} \Big(D_{ij} - \delta_{ij}^2 \Big)^2 - \frac{1}{2n} \sum_{i<j}^n D_{ij} \\ \text{s.t.} & \mathrm{diag}\,(D) = 0, \quad (-D) \in \mathcal{K}_+^n. \end{array} \tag{1.5}$$

This problem is actually an instance of weighted nearest EDM problem, see e.g., [20, 27, 29]:

$$\text{(NEDM)} \quad \min_D \ \frac{1}{2} \|W \circ (D - \Delta)\|^2 \quad \text{s.t.} \quad \mathrm{diag}\,(D) = 0, \quad -D \in \mathcal{K}_+^n, \tag{1.6}$$

where $W \in \mathcal{S}^n$ is a nonnegative weight matrix, $\circ$ means componentwise multiplication between two matrices (we omit the specification of W and Δ corresponding to (1.5)).

When comparing with SDP (1.4), the nearest EDM formulation (1.6) has obvious numerical advantage mainly due to it being a projection problem. For example, when the weight matrix W is uniform (i.e., $W_{ij} = 1$ for all (i,j)), alternating projection method [17] and semismooth Newton method [27] can be used. However, standard SDP solvers for (1.4) were severely limited due to the quadratic term in the objective which depends on the size of $\mathcal{N}$ used in (1.2). When W is not uniform, it may be approximated by a rank-one matrix $\mathbf{w}\mathbf{w}^T$ where $\mathbf{w} \in \mathbb{R}^n$ is a positive weight vector. This results in a sequential rank-one weighted (1.6), which can be efficiently solved [29].

(b) Calibrating Correlations. This example comes from finance [30, Chp. 22]. Suppose we have n forwards with expiry dates denoted by T_i, $i = 1, \ldots, n$. A correlation function that models correlations between those forwards takes the following form:

$$\rho(T_i, T_j) := \exp(-\lambda(T_i, T_j)|T_i - T_j|) \text{ with } \lambda(T_i, T_j) := \beta \exp\{-\alpha \max\{i, j\}\},$$

where $\beta > 0$ and $\alpha \in \mathbb{R}$ are empirically chosen. It is hypothesized that those correlations form a correlation matrix. However, a numerical example given in [28] shows that such constructed correlation matrix may have negative eigenvalues. Hence it is not a legitimate correlation matrix, which should be positive semidefinite. A necessary and sufficient condition for $\rho(T_i, T_j)$ to be true correlations is that the matrix

$$G(\alpha, \beta) := \Big(\beta \exp\{-\alpha \max\{i, j\}\}|T_i - T_j|\Big)_{i,j=1}^n$$

is EDM [25]. A natural calibration strategy is to seek the best α_i and β_j from a target pair $\widehat{\alpha}_i$ and $\widehat{\beta}_i$ such that the matrix

$$G(\boldsymbol{\alpha}, \boldsymbol{\beta}) := \Big(\beta_j \exp\{-\alpha_i \max\{i, j\}\}|T_i - T_j|\Big)_{i,j=1}^n$$

is EDM, yielding the following EDM optimization:

$$\begin{array}{ll} \min_{\boldsymbol{\alpha},\boldsymbol{\beta}} & \frac{1}{2}\|\boldsymbol{\alpha} - \widehat{\boldsymbol{\alpha}}\|^2 + \frac{1}{2}\|\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}\|^2 \\ \text{s.t.} & \mathrm{diag}\,(G(\boldsymbol{\alpha}, \boldsymbol{\beta})) = 0, \quad -G(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{K}^n_+, \end{array} \tag{1.7}$$

where $\widehat{\boldsymbol{\alpha}}$ and $\widehat{\boldsymbol{\beta}}$ are respectively the target of α and β. This is an example that $G(\mathbf{x})$ in (1.1) may be nonlinear.

1.3 Contribution

Some geometric properties of $\mathcal{K}^n_+$ have been known in various places, see [17,18, 27]. However, there lacks a more detailed and constructive analysis. Our first contribution is to conduct a unified variational analysis on $\mathcal{K}^n_+$. In particular, we derive the following facts.

(i) $\mathcal{K}^n_+$ is $\mathcal{C}^2$-cone reducible.

(ii) We derive a formula for the second-order tangent set of $\mathcal{K}^n_+$. This forms a key part for stating the second-order necessary and sufficient conditions for EDMOP (1.1).

(iii) The orthogonal projection operator $\Pi_{\mathcal{K}^n_+}(\cdot)$ onto $\mathcal{K}^n_+$ plays an important role in numerical algorithms for (1.1). We derive a formula for the generalized Jacobian $\partial \Pi_{\mathcal{K}^n_+}(\cdot)$.

Our second contribution is to characterize the solution behaviour when the problem data is perturbed in the canonical fashion: $f(\mathbf{x})$ is replaced by $f(\mathbf{x}) + \langle \mathbf{u}_1, \mathbf{x} \rangle$ and $G(\mathbf{x})$ is perturbed to $G(\mathbf{x}) + \mathbf{u}_2$ in (1.1), where $\mathbf{u}_1 \in \mathcal{X}$ and $\mathbf{u}_2 \in \mathcal{S}^n$ are small perturbations. Two important solution concepts are the strong regularity of Robinson [31] and the strong stability of Kojima [22]. We propose computationally verifiable characterizations for the solution of (1.1) to have those two properties. We will see that the variational analysis on $\mathcal{K}^n_+$ is essential in such characterizations, which in turn were motivated by Sun [35]. It is important to note that more general results may be stated in terms of conic programming (see, e.g., [11]). Our results are tailored to the EDM optimization and hence are more specific. The benefit is that those characterizations may be directly applied as seen in our third contribution.

Our third contribution is to demonstrate numerical implication of the obtained results to an alternating direction method of multipliers (ADMM) applied to a stress minimization problem, an instance of EDMOP studied in [43]. The established theoretical results allow us to make a stronger claim for the ADMM to the nondifferentiable stress problem that the sequence generated converges to a strongly stable solution, see Prop. 10. We delay the detailed description of the stress minimization problem to Section 5.

The paper is organized as follows. In Section 2, we give preliminaries needed in the sequential discussions, including the variational geometry of $\mathcal{K}^n_+$ and the differential properties of $\Pi_{\mathcal{K}^n_+}(\cdot)$. In Section 3 , we characterize second-order optimality conditions of EDMOP under various constraint qualifications. The equivalent characterizations of strong regularity of Robinson [31] and strong stability of Kojima [22] are derived in Section 4. In Section 5, we demonstrate numerical implication of the obtained results on a method of ADMM to the stress minimization problem. We conclude the paper in Section 6.

Before we move on to the main part of the paper, we would like to comment on a very important and yet a distinguishing approach that addresses the degeneracy issue of the EDM cone $\mathcal{D}^n$, whose interior is empty. The approach is known as the Facial Reduction (FR) thoroughly studied by Krislock and Wolkowicz [23] for EDM completion problems. We also refer to [14] for further development. FR roughly works as follows. Suppose some elements of EDM D are known. By exploring the clique structures among the known distances, one can move to a smaller dimensional space that contains D. This D often resides in a transformed face of a smaller positive SDP cone $\mathcal{S}^t_{++}$ $(t \ll n)$. Efficient iterative procedures can be designed to reduce t. This approach is extremely efficient in completing D when the known distances are exact or highly accurate to their true values [23]. In contrast, the approach introduced in this paper is to relax the EDM cone $\mathcal{D}^n$ to $\mathcal{K}^n_+$ ($\mathcal{K}^n_+$ has interior) and studies the optimization theory for the resulting EDM optimization problems. We refer the interested reader to the paper [42], where the FR-based algorithm `EepVecEDM` from [14] and a EDM optimization based algorithm `PREEEDM` are extensively compared for a large set of test problems.

Notation: Bold-faced lowercase letters denote vectors such as $\mathbf{x} \in \mathbb{R}^n$ and the uppercases are for matrices such as $A \in \mathcal{S}^n$. In particular, $\mathbf{e}_i$ is the ith unit vector in $\mathbb{R}^n$ with the ith element being one and zero elsewhere. The calligraphic letters denote sets such as $\mathcal{K}^n_+$. Its polar (or negative dual [4, Sect. 2.1.4]) cone is defined by

$$(\mathcal{K}^n_+)^\circ := \left\{ B \in \mathcal{S}^n \mid \langle A, B\rangle \le 0 \text{ for all } A \in \mathcal{K}^n_+ \right\},$$

where $\langle A, B\rangle$ is the standard trace product in $\mathcal{S}^n$. We let $\mathcal{S}^n_h := \{A \in \mathcal{S}^n \mid \mathrm{diag}\,(A) = 0\}$ known as the hollow subspace. Let $\mathcal{J}_x f(\mathbf{x})$ and $\mathcal{J}_{xx} f(\mathbf{x})$ denote the derivative and the second derivative of f with respect to $\mathbf{x} \in \mathcal{X}$, respectively. $\mathcal{T}_{\mathcal{K}^n_+}(\overline{A})$ and $\mathcal{N}_{\mathcal{K}^n_+}(\overline{A})$ are respectively the tangent cone and the normal cone of $\mathcal{K}^n_+$ at $\overline{A} \in \mathcal{K}^n_+$. $\mathrm{lin}(\mathcal{T}_{\mathcal{K}^n_+}(\overline{A}))$ is the largest linear space contained in $\mathcal{T}_{\mathcal{K}^n_+}(\overline{A})$. $\mathcal{T}^2_{\mathcal{K}^n_+}(\overline{A}, H)$ is the second order tangent set of $\mathcal{K}^n_+$ at $\overline{A} \in \mathcal{K}^n_+$ and along the direction $H \in \mathcal{T}_{\mathcal{K}^n_+}(\overline{A})$. Denote by $A \circ B := [A_{ij} B_{ij}]$ the Hadamard product between two matrices A and B of same size. For subsets α, β of $\{1, \ldots, n-1\}$, denote $B_{\alpha\beta}$ as the submatrix of B indexed by α and β. B_α denotes the submatrix consisting of columns of B indexed by α, and $|\alpha|$ is the cardinality of α. For a linear operator $\mathcal{A} : \mathcal{X} \mapsto \mathcal{Y}$, $\mathcal{A}^*$ denotes its conjugate.

2 Variational Analysis on $\mathcal{K}^n_+$

This section provides basic variational analysis on the convex cone $\mathcal{K}^n_+$ mostly through a study on the orthogonal projection onto $\mathcal{K}^n_+$:

$$\Pi_{\mathcal{K}^n_+}(A) := \mathop{\mathrm{argmin}}_{X \in \mathcal{K}^n_+} \|X - A\|,$$

where the norm is the Frobenius norm in $\mathcal{S}^n$ induced by the standard trace product. In particular, we will review two computational formulae for the projection, and characterize its directional derivatives and its generalized Jacobian. We also derive formulae for the tangent cone and the second-order tangent set of $\mathcal{K}^n_+$. These are needed for stating the second-order conditions for problem (1.1).

2.1 Two Formulae for $\Pi_{\mathcal{K}^n_+}(A)$

Recall that the matrix $J = I - \frac{1}{n}\mathbf{1}_n \mathbf{1}_n^T$ is the orthogonal projection to $\mathbf{1}_n^\perp$. The first formula is due to Gaffke and Mathar [18]:

$$\Pi_{\mathcal{K}^n_+}(A) = A + \Pi_{\mathcal{S}^n_+}(-JAJ), \quad \forall A \in \mathcal{S}^n. \tag{2.1}$$

This formula follows directly from the fact that $A \in \mathcal{K}^n_+$ if and only if $JAJ \succeq 0$ (i.e., $JAJ \in \mathcal{S}^n_+$).

A more structure-revealing formula was given by Hayden and Wells [19] using the Householder matrix Q:

$$Q := I - 2\mathbf{v}\mathbf{v}^T/\mathbf{v}^T\mathbf{v}, \tag{2.2}$$

where $\mathbf{v} := (1, \ldots, 1, \sqrt{n}+1)^T \in \mathbb{R}^n$. It has a close relationship with the matrix J:

$$J = Q \begin{bmatrix} I_{n-1} & 0 \\ 0 & 0 \end{bmatrix} Q. \tag{2.3}$$

We write the matrix QAQ in a block-structure:

$$QAQ =: \begin{bmatrix} A_1 & \mathbf{a} \\ \mathbf{a}^T & a_0 \end{bmatrix} \quad \text{with} \quad A_1 \in \mathcal{S}^{n-1}. \tag{2.4}$$

Substituting (2.3) into (2.1) and using the structure of QAQ in (2.4), we obtain the formula of Hayden and Wells [19]:

$$\Pi_{\mathcal{K}^n_+}(A) = Q \begin{bmatrix} \Pi_{\mathcal{S}^{n-1}_+}(A_1) & \mathbf{a} \\ \mathbf{a}^T & a_0 \end{bmatrix} Q, \quad \forall A \in \mathcal{S}^n.$$

An immediate observation from this formula is that $A \in \mathcal{K}^n_+$ if and only if $A_1 \in \mathcal{S}^{n-1}_+$ in (2.4). Hence the cone of $\mathcal{K}^n_+$ can be described as follows:

$$\mathcal{K}^n_+ = \left\{ Q \begin{bmatrix} Z_1 & \mathbf{z} \\ \mathbf{z}^T & z_0 \end{bmatrix} Q : Z_1 \in \mathcal{S}^{n-1}_+, \mathbf{z} \in \mathbb{R}^{n-1}, z_0 \in \mathbb{R} \right\}$$

Its polar cone $\left(\mathcal{K}^n_+\right)^\circ$ is then given by

$$\left(\mathcal{K}^n_+\right)^\circ = \left\{ Q \begin{bmatrix} Z & 0 \\ 0 & 0 \end{bmatrix} Q : -Z \in \mathcal{S}^{n-1}_+ \right\}.$$

From now on, we use A as a general matrix not necessarily belonging to $\mathcal{K}^n_+$ and $\overline{A}$ as a matrix in $\mathcal{K}^n_+$. Moreover, we often take $\overline{A}$ as the projection of A onto $\mathcal{K}^n_+$, i.e., $\overline{A} = \Pi_{\mathcal{K}^n_+}(A)$. In particular, we rewrite (2.4) as

$$A = Q \begin{bmatrix} A_1 & \mathbf{a} \\ \mathbf{a}^T & a_0 \end{bmatrix} Q \qquad \text{and} \qquad \overline{A} = Q \begin{bmatrix} \overline{A}_1 & \mathbf{a} \\ \mathbf{a}^T & a_0 \end{bmatrix} Q. \tag{2.5}$$

In the case $\overline{A} = \Pi_{\mathcal{K}^n_+}(A)$, we obviously have $\overline{A}_1 = \Pi_{\mathcal{S}^{n-1}_+}(A_1)$.

2.2 Directional Derivatives of $\Pi_{\mathcal{K}^n_+}(\cdot)$ and Tangent Cones of $\mathcal{K}^n_+$

The main task in this subsection is to derive the formulae for directional derivatives of $\Pi_{\mathcal{K}^n_+}(\cdot)$. The formulae will allow us to characterize the tangent cone and the second-order tangent set of $\mathcal{K}^n_+$. Our starting point is that the directional derivatives can be computed through the corresponding directional derivatives of the projection onto the positive semidefinite cone $\mathcal{S}^{n-1}_+$.

Let $H, W \in \mathcal{S}^n$ represent directions to be used. We again write H and W in the form of block structure:

$$H = Q \begin{bmatrix} H_1 & \mathbf{h} \\ \mathbf{h}^T & h_0 \end{bmatrix} Q, \qquad W = Q \begin{bmatrix} W_1 & \mathbf{w} \\ \mathbf{w}^T & w_0 \end{bmatrix} Q.$$

It is easy to see that the directional derivative of $\Pi_{\mathcal{K}^n_+}(\cdot)$ at A along H is given by

$$\Pi'_{\mathcal{K}^n_+}(A, H) = \lim_{t \to 0} \frac{\Pi_{\mathcal{K}^n_+}(A + tH) - \Pi_{\mathcal{K}^n_+}(A)}{t} = Q \begin{bmatrix} \Pi'_{\mathcal{S}^{n-1}_+}(A_1, H_1) & \mathbf{h} \\ \mathbf{h}^T & h_0 \end{bmatrix} Q, \tag{2.6}$$

and the second-order derivative is given by

$$\Pi''_{\mathcal{K}^n_+}(A; H, W) = Q \begin{bmatrix} \Pi''_{\mathcal{S}^{n-1}_+}(A_1; H_1, W_1) & \mathbf{w} \\ \mathbf{w}^T & w_0 \end{bmatrix} Q,$$

Computational formulae for the directional derivatives of the projection on the positive semidefinite cone have been known [34, 36]. We briefly state them. Let $\lambda_1(A_1) \geq \ldots \geq \lambda_{n-1}(A_1)$ be the eigenvalues of A_1 arranged in non-increasing order. Let $\Lambda := \mathrm{Diag}(\lambda_1(A_1), \ldots, \lambda_{n-1}(A_1))$ be the diagonal matrix with eigenvalues being its diagonal. Suppose A_1 has the following spectral decomposition:

$$A_1 = U \Lambda U^T, \tag{2.7}$$

where U is the matrix of orthonormal eigenvectors. Define three index sets of positive, zero, and negative eigenvalues of A_1, respectively, as

$$\alpha := \{i : \lambda_i > 0\}, \quad \beta := \{i : \lambda_i = 0\}, \quad \gamma := \{i : \lambda_i < 0\}.$$

The matrices Λ and U in (2.7) can be arranged as follows:

$$\Lambda = \begin{bmatrix} \Lambda_\alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Lambda_\gamma \end{bmatrix} \quad \text{and} \quad \mathrm{U} = \begin{bmatrix} U_\alpha & U_\beta & U_\gamma \end{bmatrix}$$

with $U_\alpha \in \mathbb{R}^{(n-1)\times|\alpha|}$, $U_\beta \in \mathbb{R}^{(n-1)\times|\beta|}$, and $U_\gamma \in \mathbb{R}^{(n-1)\times|\gamma|}$. Define the matrix $\Omega \in \mathcal{S}^{n-1}$ with entries

$$\Omega_{ij} := \frac{\max(\lambda_i, 0) + \max(\lambda_j, 0)}{|\lambda_i| + |\lambda_j|}, \quad i, j = 1, \ldots, n-1, \tag{2.8}$$

where 0/0 is defined to be 1. Sun and Sun [36] gave an explicit formula for the directional derivative of $\Pi'_{\mathcal{S}^{n-1}_+}(A_1, H_1)$:

$$\Pi'_{\mathcal{S}^{n-1}_+}(A_1; H_1) = U \begin{bmatrix} (\widetilde{H_1})_{\alpha\alpha} & (\widetilde{H_1})_{\alpha\beta} & \Omega_{\alpha\gamma} \circ (\widetilde{H_1})_{\alpha\gamma} \\ (\widetilde{H_1})^T_{\alpha\beta} & \Pi_{\mathcal{S}^{|\beta|}_+}((\widetilde{H_1})_{\beta\beta}) & 0 \\ (\widetilde{H_1})^T_{\alpha\gamma} \circ \Omega^T_{\alpha\gamma} & 0 & 0 \end{bmatrix} U^T \tag{2.9}$$

where $\widetilde{H_1} := U^T H_1 U$. The formulae (2.6) and (2.9) yield the characterization of the tangent cone of $\mathcal{K}^n_+$ at $\overline{A} = \Pi_{\mathcal{K}^n_+}(A)$:

$$\begin{aligned} \mathcal{T}_{\mathcal{K}^n_+}(\overline{A}) &= \left\{ H \in \mathcal{S}^n : \Pi'_{\mathcal{K}^n_+}(\overline{A}, H) = H \right\} \\ &= \left\{ Q \begin{bmatrix} H_1 & h \\ h^T & h_0 \end{bmatrix} Q \in \mathcal{S}^n : \Pi'_{\mathcal{S}^{n-1}_+}(\overline{A}_1; H_1) = H_1 \right\} \\ &= \left\{ Q \begin{bmatrix} H_1 & h \\ h^T & h_0 \end{bmatrix} Q \in \mathcal{S}^n : U^T_{\overline{\alpha}} H_1 U_{\overline{\alpha}} \succeq 0 \right\}, \end{aligned} \tag{2.10}$$

where $\overline{\alpha} := \beta \cup \gamma$ and $U_{\overline{\alpha}} := [U_\beta \, , U_\gamma]$. Note that the characterization of $\mathcal{T}_{\mathcal{K}^n_+}(\overline{A})$ was first obtained by [17] without using the directional derivative $\Pi'_{\mathcal{K}^n_+}(A, H)$. It follows from (2.10) that the largest linear subspace contained in the tangent cone is

$$\mathrm{lin}(\mathcal{T}_{\mathcal{K}^n_+}(\overline{A})) = \left\{ Q \begin{bmatrix} H_1 & \mathbf{h} \\ \mathbf{h}^T & h_0 \end{bmatrix} Q \in \mathcal{S}^n : U^T_{\overline{\alpha}} H_1 U_{\overline{\alpha}} = 0 \right\} \tag{2.11}$$

and the normal cone is given by

$$\mathcal{N}_{\mathcal{K}^n_+}(\overline{A}) = \left\{ Q \begin{bmatrix} U \begin{bmatrix} 0 & 0 \\ 0 & M \end{bmatrix} U^T & 0 \\ 0 & 0 \end{bmatrix} Q \in \mathcal{S}^n \;\Big|\; -M \in \mathcal{S}^{|\beta|+|\gamma|}_+ \right\}. \tag{2.12}$$

The critical cone of $\mathcal{K}^n_+$ at $\overline{A} \in \mathcal{K}^n_+$ is defined as

$$\mathcal{C}(A, \mathcal{K}^n_+) := \mathcal{T}_{\mathcal{K}^n_+}(\overline{A}) \cap (\overline{A} - A)^\perp,$$

which can be completely described:

$$\mathcal{C}(A, \mathcal{K}^n_+) = \left\{ Q \begin{bmatrix} H_1 & \mathbf{h} \\ \mathbf{h}^T & h_0 \end{bmatrix} Q \in \mathcal{S}^n \;\Big|\; \begin{array}{l} U^T_\beta H_1 U_\beta \succeq 0 \\ U^T_\beta H_1 U_\gamma = 0, \ U^T_\gamma H_1 U_\gamma = 0 \end{array} \right\}. \tag{2.13}$$

The affine hull of $\mathcal{C}(A, \mathcal{K}^n_+)$ denoted by $\mathrm{aff}(\mathcal{C}(A, \mathcal{K}^n_+))$ can thus be written as

$$\mathrm{aff}(\mathcal{C}(A, \mathcal{K}^n_+)) = \left\{ Q \begin{bmatrix} H_1 & \mathbf{h} \\ \mathbf{h}^T & h_0 \end{bmatrix} Q \in \mathcal{S}^n \;\Big|\; \begin{array}{l} U^T_\beta H_1 U_\gamma = 0 \\ U^T_\gamma H_1 U_\gamma = 0 \end{array} \right\}. \tag{2.14}$$

We now consider the second-order tangent set of $\mathcal{K}^n_+$ at $\overline{A}$ along a direction $H \in \mathcal{T}_{\mathcal{K}^n_+}(\overline{A})$. It can be characterized through the second derivative of $\Pi_{\mathcal{K}^n}$:

$$\begin{aligned}\mathcal{T}^2_{\mathcal{K}^n_+}(\overline{A}, H) &= \{W \in \mathcal{S}^n \;\big|\; \Pi''_{\mathcal{K}^n_+}(A, H, W) = W\} \\ &= \left\{ Q \begin{bmatrix} W_1 & \mathbf{w} \\ \mathbf{w}^T & w_0 \end{bmatrix} Q \in \mathcal{S}^n \;\Big|\; \Pi''_{\mathcal{S}^{n-1}_+}(A_1, H_1, W_1) = W_1 \right\} \\ &= \left\{ Q \begin{bmatrix} W_1 & \mathbf{w} \\ \mathbf{w}^T & w_0 \end{bmatrix} Q \in \mathcal{S}^n \;\Big|\; W_1 \in \mathcal{T}^2_{\mathcal{S}^{n-1}_+}(\overline{A}_1, H_1) \right\}\end{aligned}$$

The formula for the second-order tangent set of the semidefinite cone is known. Note that $\overline{A}_1 \in \mathcal{S}^{n-1}_+$. We consider the nontrivial case that its smallest eigenvalue is 0 (i.e., $\lambda_{n-1}(\overline{A}_1) = 0$). Otherwise $\mathcal{T}^2_{\mathcal{S}^{n-1}_+}(\overline{A}_1, H_1) = \mathcal{S}^{n-1}$. Let s be the multiplicity of the smallest eigenvalue of the matrix $U^T_{\overline{\alpha}} H_1 U^T_{\overline{\alpha}}$ and let $\mathbf{v}_1, \ldots, \mathbf{v}_s$ be a corresponding set of orthonormal eigenvectors and $V_s := [\mathbf{v}_1, \ldots, \mathbf{v}_s]$ be the corresponding matrix. Then, we know from [34, Eq. 35] that the second order tangent set of $\mathcal{S}^{n-1}_+$ at $\overline{A}_1 \in \mathcal{S}^{n-1}_+$ and along the direction $H_1 \in \mathcal{T}_{\mathcal{S}^{n-1}_+}(\overline{A}_1)$ can be written in the form,

$$\mathcal{T}^2_{\mathcal{S}^{n-1}_+}(\overline{A}_1, H_1) = \left\{ W_1 \in \mathcal{S}^{n-1} \;\Big|\; V_s^T U^T_{\overline{\alpha}} W_1 U_{\overline{\alpha}} V_s \succeq 2 V_s^T U^T_{\overline{\alpha}} H_1 A_1^{\dagger} H_1 U_{\overline{\alpha}} V_s \right\},$$

where $A_1^{\dagger}$ is the (generalized) Moore-Penrose inverse of A_1. Consequently,

$$\mathcal{T}^2_{\mathcal{K}^n_+}(\overline{A}, H) = \left\{ Q \begin{bmatrix} W_1 & \mathbf{w} \\ \mathbf{w}^T & w_0 \end{bmatrix} Q \;\Big|\; V_s^T U^T_{\overline{\alpha}} W_1 U_{\overline{\alpha}} V_s \succeq 2 V_s^T U^T_{\overline{\alpha}} H_1 A_1^{\dagger} H_1 U_{\overline{\alpha}} V_s \right\}. \tag{2.15}$$

As an application, we derive an expression for the support function of the second order tangent set $\mathcal{T}^2_{\mathcal{K}^n_+}(\overline{A}, H)$. To simplify the notation, we use $\mathcal{T}^2$ instead of $\mathcal{T}^2_{\mathcal{K}^n_+}(\overline{A}, H)$.

Proposition 1 *Let $\overline{A} \in \mathcal{K}^n_+$ with $\overline{A}_1$ being given in (2.5). Let $H \in \mathcal{T}_{\mathcal{K}^n_+}(\overline{A})$ and $Z \in \mathcal{N}_{\mathcal{K}^n_+}(\overline{A})$. We decompose H and Z as*

$$H = Q \begin{bmatrix} H_1 & \mathbf{h} \\ \mathbf{h}^T & h_0 \end{bmatrix} Q, \qquad Z = Q \begin{bmatrix} Z_1 & 0 \\ 0 & 0 \end{bmatrix} Q \qquad \textit{(by (2.12))}$$

The support function over $\mathcal{T}^2$ at Z is given by

$$\sigma(Z, \mathcal{T}^2) = \begin{cases} 2\langle Z_1,\ H_1 \overline{A}_1^{\dagger} H_1 \rangle & \textit{if } \langle Z,\ H \rangle = 0, \\ \infty & \textit{otherwise.} \end{cases}$$

Proof. We first consider the case that $\langle Z,\ H \rangle \neq 0$, i.e. Z is not contained in the polar of $\mathcal{T}_{\mathcal{T}_{\mathcal{K}^n_+}(\overline{A}_1)}(H)$. From [4, Proposition 3.34], we get $\sigma(Z, \mathcal{T}^2) = +\infty$.

Now, we suppose that $\langle Z,\ H\rangle = 0$. It follows from the structures of $\mathcal{T}_{\mathcal{K}_+^n}(\overline{A})$ and $\mathcal{N}_{\mathcal{K}_+^n}(\overline{A})$ that

$$0 = \langle Z,\ H\rangle = \langle Z_1,\ H_1\rangle = \langle U_{\overline{\alpha}}^T Z_1 U_{\overline{\alpha}},\ U_{\overline{\alpha}}^T H_1 U_{\overline{\alpha}}\rangle$$

with $U_{\overline{\alpha}}^T Z_1 U_{\overline{\alpha}} \preceq 0$ and $U_{\overline{\alpha}}^T H_1 U_{\overline{\alpha}} \succeq 0$. Therefore, both matrices share a same set of eigenvectors as they are complementary to each other. In particular, $U_{\overline{\alpha}}^T Z_1 U_{\overline{\alpha}}$ can be written in terms of eigenvectors in V_s corresponding to the smallest eigenvalue of $U_{\overline{\alpha}}^T H_1 U_{\overline{\alpha}} \succeq 0$:

$$U_{\overline{\alpha}}^T Z_1 U_{\overline{\alpha}} = V_s \Psi V_s^T \qquad \text{for some } \Psi \preceq 0.$$

Moreover, we have

$$Z_1 = U_{\overline{\alpha}} U_{\overline{\alpha}}^T Z_1 U_{\overline{\alpha}} U_{\overline{\alpha}}^T = U_{\overline{\alpha}} V_s \Psi V_s^T U_{\overline{\alpha}}^T.$$

We are ready to calculate the support function using the characterization (2.15):

$$\begin{aligned}
&\sigma(Z, \mathcal{T}^2) = \sup\{\langle Z,\ W\rangle : W \in \mathcal{T}^2\} \\
&= \sup\left\{\langle Q \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} Q, Q \begin{bmatrix} W_1 & w \\ w^T & w_0 \end{bmatrix} Q\rangle : V_s^T U_{\overline{\alpha}}^T W_1 U_{\overline{\alpha}} V_s \succeq 2 V_s^T U_{\overline{\alpha}}^T H_1 \overline{A}_1^{\dagger} H_1 U_{\overline{\alpha}} V_s \right\} \\
&= \sup\left\{\langle Z_1,\ W_1\rangle : V_s^T U_{\overline{\alpha}}^T W_1 U_{\overline{\alpha}} V_s \succeq 2 V_s^T U_{\overline{\alpha}}^T H_1 \overline{A}_1^{\dagger} H_1 U_{\overline{\alpha}} V_s \right\} \\
&= \sup\left\{\langle \Psi,\ V_s^T U_{\overline{\alpha}}^T W_1 U_{\overline{\alpha}} V_s\rangle : V_s^T U_{\overline{\alpha}}^T W_1 U_{\overline{\alpha}} V_s \succeq 2 V_s^T U_{\overline{\alpha}}^T H_1 \overline{A}_1^{\dagger} H_1 U_{\overline{\alpha}} V_s \right\} \\
&= 2\langle \Psi,\ V_s^T U_{\overline{\alpha}}^T H_1 \overline{A}_1^{\dagger} H_1 U_{\overline{\alpha}} V_s\rangle \\
&= \langle Z_1,\ H_1 \overline{A}_1^{\dagger} H_1\rangle
\end{aligned}$$

□

2.3 $\mathcal{K}_+^n$ is $\mathcal{C}^2$-cone Reducible

We recall the formal definition when a set is $\mathcal{C}^2$-cone reducible.

Definition 1 [4, Definition 3.135] Let $\mathcal{Y}$ and $\mathcal{Z}$ be two finite dimensional Euclidean spaces. Let $K \subset \mathcal{Y}$ and $C \subset \mathcal{Z}$ be convex closed sets. We say that the set K is $\mathcal{C}^2$-reducible to the set C, at a point $\overline{\mathbf{y}} \in K$, if there exist a neighbourhood $\mathcal{U}$ of $\overline{\mathbf{y}}$ and twice continuously differentiable mapping $\Xi : \mathcal{U} \to \mathcal{Z}$ such that (i) $\Xi'(\overline{\mathbf{y}}) : \mathcal{Y} \to \mathcal{Z}$ is onto, and (ii) $K \cap \mathcal{U} = \{\mathbf{y} \in \mathcal{U} : \Xi(\mathbf{y}) \in C\}$. We say that the reduction is pointed if the tangent cone $\mathcal{T}_C(\Xi(\overline{\mathbf{y}}))$ is a pointed cone. If, in addition, the set $C - \Xi(\overline{\mathbf{y}})$ is a pointed convex closed cone, we say that K is $\mathcal{C}^2$-cone reducible at $\overline{\mathbf{y}}$. We can assume without loss of generality that $\Xi(\overline{\mathbf{y}}) = 0$.

We now prove that $\mathcal{K}_+^n$ is $\mathcal{C}^2$-cone reducible at every point $\overline{A} \in \mathcal{K}_+^n$ and the proof is patterned after [4, Example 3.140]. For the sake of completeness, we include a short proof. Hence, $\mathcal{K}_+^n$ is second order regular ([4, Definition 3.85]) at every point.

Proposition 2 *Let $\overline{A}$ be an arbitrary point in $\mathcal{K}_+^n$ and let r be the rank of the matrix $J\overline{A}J$, i.e.,* $\mathrm{rank}\,(J\overline{A}J) = r$*. Then $\mathcal{K}_+^n$ is $\mathcal{C}^2$-cone reducible to $\mathcal{S}_+^{n-1-r}$ at $\overline{A} \in \mathcal{K}_+^n$.*

Proof. Let $\overline{A} \in \mathcal{K}_+^n$, i.e. $\overline{A} = Q \begin{bmatrix} \overline{A}_1 & \bar{\mathbf{a}} \\ \bar{\mathbf{a}}^T & \bar{a}_0 \end{bmatrix} Q$, and $\overline{A}_1 \in \mathcal{S}_+^{n-1}$. It follows from the property (2.3) between Q and J that $\mathrm{rank}(\overline{A}_1) = r$. For any $A = Q \begin{bmatrix} A_1 & \mathbf{a} \\ \mathbf{a}^T & a_0 \end{bmatrix} Q \in \mathcal{S}^n$, denote by $\lambda_1(A_1) \geq \ldots \geq \lambda_{n-1}(A_1)$ the eigenvalues of $(n-1) \times (n-1)$ symmetric matrix A_1 and $\mathbf{p}_1(A_1), \ldots, \mathbf{p}_{n-1}(A_1)$ an orthonormal set of corresponding eigenvectors. Let $E(A_1)$ be the $(n-1) \times (n-1-r)$ matrix whose columns are formed from vectors $\mathbf{p}_{r+1}(A_1), \ldots, \mathbf{p}_{n-1}(A_1)$.

Denote by $L(A_1)$ the eigenspace corresponding to the $(n-1-r)$ smallest eigenvalues of A_1, and $P(A_1)$ be the orthonormal projection matrix onto $L(A_1)$. Also let E_0 be a (fixed) $(n-1) \times (n-1-r)$ matrix whose columns are orthonormal and span the space $L(\overline{A}_1)$, i.e. $E_0 = E(\overline{A}_1)$. It is known that $P(A_1)$ is a continuously differentiable function of A_1 in a sufficiently small neighbourhood of $\overline{A}_1$. Consequently, $F(A_1) := P(A_1)E_0$ is also a continuously differentiable function of A_1 in a sufficiently small neighbourhood of $\overline{A}_1$ and $F(\overline{A}_1) = E_0$. It follows that for all A_1 sufficiently close to $\overline{A}_1$, the rank of $F(A_1)$ is $(n-1-r)$, i.e., its column vectors are linearly independent.

Let $U(A_1)$ be the matrix whose columns are obtained by applying the Gram-Schmidt orthonormalization procedure to the columns of $F(A_1)$. The matrix $U(A_1)$ is well defined and continuously differentiable near $\overline{A}_1$. Moreover, it satisfies the following conditions: $U(\overline{A}_1) = E_0$, the column space of $U(A_1)$ coincides with the column space of $E(A_1)$, and $U(A_1)^T U(A_1) = I_{n-1-r}$. We obtain that in a neighbourhood $\mathcal{N}_1$ of $\overline{A}_1$, the cone $\mathcal{S}_+^{n-1}$ can be defined in the form $\{A_1 : U(A_1)^T A_1 U(A_1) \succeq 0\}$. Furthermore, in a neighbourhood $\mathcal{N}$ of $\overline{A}$, the cone $\mathcal{K}_+^n$ can be represented as

$$\left\{ A \in \mathcal{S}^n : A = Q \begin{bmatrix} A_1 & \mathbf{a} \\ \mathbf{a}^T & a_0 \end{bmatrix} Q, U(A_1)^T A_1 U(A_1) \succeq 0 \right\}.$$

Consider the mapping $E : A \to U(A_1)^T A_1 U(A_1)$ from $\mathcal{N}$ into $\mathcal{S}_+^{n-1-r}$. The mapping E is continuously differentiable. Since the expression of $\mathcal{J}E(\overline{A})A$ is

$$(\mathcal{J}U(\overline{A}_1)^T A_1)\overline{A}_1 U(\overline{A}_1) + U(\overline{A}_1)^T A_1 U(\overline{A}_1) + U(\overline{A}_1)^T A_1 (\mathcal{J}U(\overline{A}_1)^T A_1)$$

and $\overline{A}_1 U(\overline{A}_1) = 0$, we have $\mathcal{J}E(\overline{A})A = E_0^T A E_0$. It follows that $\mathcal{J}E(\overline{A})$ is onto. By Def. 1, we proved $\mathcal{K}_+^n$ is $\mathcal{C}^2$-reducible to $\mathcal{S}_+^{n-1-r}$ at $\overline{A} \in \mathcal{K}_+^n$ with $r = \mathrm{rank}\,(J\overline{A}J)$. □

2.4 Generalized Jacobian of $\Pi_{\mathcal{K}_+^n}(\cdot)$

In this section, we state some differential properties of $\Pi_{\mathcal{K}_+^n}(\cdot)$. Those properties can be proved similarly as in [35] for the positive semidefinite cone $\mathcal{S}_+^n$. We will omit those proofs for the results in Prop. 3 and Prop. 5 below. Since $\Pi_{\mathcal{K}_+^n}(\cdot)$ is a Lipschitz continuous function on $\mathcal{S}^n$, by the well known Rademacher's theorem, we know that $\Pi_{\mathcal{K}_+^n}$ is F (Fréchet)-differentiable almost everywhere. Denote by $\mathrm{D}_{\Pi_{\mathcal{K}_+}^n}$ the set of all points where $\Pi_{\mathcal{K}_+^n}$ is F-differentiable, then Clarke's generalized Jacobian of $\Pi_{\mathcal{K}_+^n}$ at $A \in \mathcal{S}^n$ is defined as follows:

$$\partial \Pi_{\mathcal{K}_+^n}(A) := \operatorname{conv}\left\{\partial_B \Pi_{\mathcal{K}_+^n}(A)\right\}$$

where "conv" denotes the convex hull and the B-subdifferential $\partial_B \Pi_{\mathcal{K}_+^n}(A)$ is defined by

$$\partial_B \Pi_{\mathcal{K}_+^n}(A) := \left\{V : V = \lim_{k\to\infty} \Pi'_{\mathcal{K}_+^n}\left(A^k\right), A^k \to A, A^k \in \mathrm{D}_{\Pi^n_{\mathcal{K}+}}\right\}.$$

Proposition 3 *The following three statements are true.*

(i) $\Pi_{\mathcal{K}_+^n}(\cdot)$ *is F-differentiable at* $A \in \mathcal{K}_+^n$, *where* $A = Q\begin{bmatrix} A_1 & \mathbf{a} \\ \mathbf{a}^T & a_0 \end{bmatrix} Q$ *if and only if* A_1 *is nonsingular.*

(ii) For any $A \in \mathcal{K}_+^n$, *the directional derivative* $\Pi'_{\mathcal{K}_+^n}(A;\cdot)$ *is F-differentiable at* $H \in \mathcal{K}_+^n$ *if and only if* $U_\beta^T H_1 U_\beta$ *is nonsingular.*

(iii) Let $A \in \mathcal{K}_+^n$ *be arbitrary and* $\Theta(\cdot) := \Pi'_{\mathcal{K}_+^n}(A;\cdot)$. *It holds that*

$$\partial_B \Pi_{\mathcal{K}_+^n}(A) = \partial_B \Theta(0).$$

Proposition 3 allow us to prove the useful result on $\partial \Pi_{\mathcal{K}_+^n}(\cdot)$. First, we give the formula of describing $\partial \Pi_{\mathcal{S}_+^{n-1}}(A_1)$.

Proposition 4 [35, Proposition 2.2] Suppose that $A_1 \in \mathcal{S}^{n-1}$ has the decomposition (2.7). Then for any $V \in \partial \Pi_{\mathcal{S}_+^{n-1}}(A_1)$ (respectively, $V \in \partial_B \Pi_{\mathcal{S}_+^{n-1}}(A_1)$), there exists a $V_1 \in \partial \Pi_{\mathcal{S}_+^{|\beta|}}(0)$ (respectively, $V_1 \in \partial_B \Pi_{\mathcal{S}_+^{|\beta|}}(0)$) such that

$$V(H) = U\begin{bmatrix} (\widetilde{H_1})_{\alpha\alpha} & (\widetilde{H_1})_{\alpha\beta} & \Omega_{\alpha\gamma} \circ (\widetilde{H_1})_{\alpha\gamma} \\ (\widetilde{H_1})_{\alpha\beta}^T & V_1((\widetilde{H_1})_{\beta\beta}) & 0 \\ (\widetilde{H_1})_{\alpha\gamma}^T \circ \Omega_{\alpha\gamma}^T & 0 & 0 \end{bmatrix} U^T,$$

with $\widetilde{H_1} := U^T H_1 U$ and the matrix Ω is defined in (2.8).

Proposition 5 *Suppose that* $A \in \mathcal{S}^n$ *has the decomposition (2.5). Then for any* $\Psi \in \partial \Pi_{\mathcal{K}_+^n}(A)$ *(respectively,* $\Psi \in \partial_B \Pi_{\mathcal{K}_+^n}(A)$*), there exists a* $V \in \partial \Pi_{\mathcal{S}_+^{n-1}}(A_1)$ *(respectively,* $V_1 \in \partial_B \Pi_{\mathcal{S}_+^{n-1}}(A_1)$*) such that*

$$\Psi(H) = Q\begin{bmatrix} V(H_1) & \mathbf{h} \\ \mathbf{h}^T & h_0 \end{bmatrix} Q.$$

For any given $B \in \mathcal{S}^n$ we below introduce a linear-quadratic function $\Upsilon_B : \mathcal{S}^n \times \mathcal{S}^n \to \mathbb{R}$.

Definition 2 For any given $B \in \mathcal{S}^n$, define the linear-quadratic function $\Upsilon_B : \mathcal{S}^n \times \mathcal{S}^n \to \mathbb{R}$ by

$$\Upsilon_B(\Gamma, A) := 2\langle \Gamma_1, B_1 A_1^\dagger B_1\rangle, \quad (\Gamma, A) \in \mathcal{S}^n \times \mathcal{S}^n,$$

where the submatrices Γ_1, B_1, and A_1 are defined as follows:

$$\Gamma = Q \begin{bmatrix} \Gamma_1 & \xi \\ \xi^T & \xi_0 \end{bmatrix} Q, \quad B = Q \begin{bmatrix} B_1 & \mathbf{b} \\ \mathbf{b}^T & b_0 \end{bmatrix} Q, \quad \text{and} \quad A = Q \begin{bmatrix} A_1 & \mathbf{a} \\ \mathbf{a}^T & a_0 \end{bmatrix} Q.$$

The following result plays an important role in our subsequent analysis.

Proposition 6 *Suppose that $B \in \mathcal{K}_+^n$ and $\Gamma \in \mathcal{N}_{\mathcal{K}_+^n}(B)$, then for any $\Psi \in \partial \Pi_{\mathcal{K}_+^n}(B+\Gamma)$ and $\triangle B, \triangle \Gamma \in \mathcal{S}^n$ such that $\Psi(\triangle B + \triangle \Gamma) = \triangle B$, it holds that*

$$\langle \triangle B, \triangle \Gamma\rangle \geq -\Upsilon_B(\Gamma, \triangle B).$$

Proof. Let $A := B + \Gamma$, Then, we know that

$$B = \Pi_{\mathcal{K}_+^n}(B+\Gamma) = \Pi_{\mathcal{K}_+^n}(A).$$

Thus, we can assume that A has the decomposition as in (2.5), and

$$B = Q \begin{bmatrix} B_1 & \mathbf{b} \\ \mathbf{b}^T & b_0 \end{bmatrix} = Q \begin{bmatrix} U \begin{bmatrix} \Lambda_\alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} U^T & \mathbf{a} \\ \mathbf{a}^T & a_0 \end{bmatrix} Q \,, \Gamma = Q \begin{bmatrix} U \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Lambda_\gamma \end{bmatrix} U^T & 0 \\ 0 & 0 \end{bmatrix} Q.$$

By Proposition 5, there exists a $V \in \partial \Pi_{\mathcal{S}_+^{n-1}}(B_1 + \Gamma_1)$ such that

$$\Psi(\triangle B + \triangle \Gamma) = Q \begin{bmatrix} V(\triangle B_1 + \triangle \Gamma_1) & \triangle b + \triangle \xi \\ (\triangle b + \triangle \xi)^T & \triangle b_0 + \triangle \xi_0 \end{bmatrix} Q,$$

where $\triangle \Gamma = Q \begin{bmatrix} \triangle \Gamma_1 & \triangle \xi \\ \triangle \xi^T & \triangle \xi_0 \end{bmatrix} Q$ and $\triangle B = Q \begin{bmatrix} \triangle B_1 & \triangle b \\ \triangle b^T & \triangle b_0 \end{bmatrix} Q$.

It follows from $\Psi(\triangle B + \triangle \Gamma) = \triangle B$ that

$$V(\triangle B_1 + \triangle \Gamma_1) = \triangle B_1, \triangle \xi = 0, \triangle \xi_0 = 0.$$

Since $B_1 \in \mathcal{S}_+^{n-1}, \Gamma_1 \in \mathcal{N}_{\mathcal{S}_+^{n-1}}(B_1)$, then by [35, Proposition 2.3], we have

$$\langle \triangle B_1, \triangle \Gamma_1\rangle \geq -\langle \Gamma_1, \triangle B_1 B_1^\dagger \triangle B\rangle.$$

Hence,

$$\begin{aligned} \langle \triangle B, \triangle \Gamma\rangle &= \left\langle Q \begin{bmatrix} \triangle B_1 & \triangle b \\ \triangle b^T & \triangle b_0 \end{bmatrix} Q, Q \begin{bmatrix} \triangle \Gamma_1 & 0 \\ 0 & 0 \end{bmatrix} Q \right\rangle \\ &= \langle \triangle B_1, \triangle \Gamma_1\rangle \\ &\geq -2\langle \Gamma_1, \triangle B_1 B_1^\dagger \triangle B\rangle \\ &= -\Upsilon_B(\Gamma, \triangle B). \end{aligned}$$

□

3 Optimality Conditions and Constraint Qualifications

First, we rewrite EDMOP (1.1) in the following form:

$$\text{(EDMOP)} \qquad \min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}) \qquad \text{s.t.} \quad \mathcal{A}(G(\mathbf{x}))=0, \quad G(\mathbf{x})\in\mathcal{K}^n_+, \tag{3.1}$$

with $\mathcal{A}(\cdot)$ being the diagonal operator. The first order optimality condition, namely the Karush-Kuhn-Tucker (KKT) condition, for EDMOP takes the following form:

$$\mathcal{J}_x L(\mathbf{x},\zeta,\Gamma)=0,\ \mathcal{A}(G(\mathbf{x}))=0 \text{ and } \Gamma\in\mathcal{N}_{\mathcal{K}^n_+}(G(\mathbf{x})), \tag{3.2}$$

where the Lagrangian function $L:\mathcal{X}\times\mathbb{R}^n\times\mathcal{S}^n\to\mathbb{R}$ is defined by

$$L(\mathbf{x},\zeta,\Gamma):=f(\mathbf{x})+\langle\zeta,\mathcal{A}(G(\mathbf{x}))\rangle+\langle\Gamma,G(\mathbf{x})\rangle,\ (\mathbf{x},\zeta,\Gamma)\in\mathcal{X}\times\mathbb{R}^n\times\mathcal{S}^n.$$

If $(\mathbf{x},\zeta,\Gamma)$ satisfies (3.2), then it is called a KKT point of EDMOP, $\mathbf{x}$ a stationary point, and (ζ,Γ) a Lagrange multiplier at $\mathbf{x}$. We denote by $\Lambda(\mathbf{x})$ the set of all Lagrange multipliers at $\mathbf{x}$.

Let $\overline{\mathbf{x}}$ be a feasible solution to EDMOP. The critical cone $\mathcal{C}(\overline{\mathbf{x}})$ of EDMOP at $\overline{\mathbf{x}}$ is defined by

$$\mathcal{C}(\overline{\mathbf{x}}):=\{\mathbf{d}:\mathcal{A}(\mathcal{J}_xG(\overline{\mathbf{x}})\mathbf{d})=0,\mathcal{J}_xG(\overline{\mathbf{x}})\mathbf{d}\in\mathcal{T}_{\mathcal{K}^n_+}(G(\overline{\mathbf{x}})),\mathcal{J}_xf(\overline{\mathbf{x}})\mathbf{d}\le 0\}.$$

If $\overline{\mathbf{x}}$ is a stationary point of EDMOP with $\Lambda(\overline{\mathbf{x}})$ nonempty, then

$$\mathcal{C}(\overline{\mathbf{x}})=\{\mathbf{d}:\ \mathcal{A}(\mathcal{J}_xG(\overline{\mathbf{x}})\mathbf{d})=0,\mathcal{J}_xG(\overline{\mathbf{x}})\mathbf{d}\in\mathcal{T}_{\mathcal{K}^n_+}(G(\overline{\mathbf{x}})),\mathcal{J}_xf(\overline{\mathbf{x}})\mathbf{d}=0\},$$

and there exists $(\overline{\zeta},\overline{\Gamma})\in\Lambda(\overline{\mathbf{x}})$ such that

$$\mathcal{J}_xL(\overline{\mathbf{x}},\overline{\zeta},\overline{\Gamma})=0,\ \mathcal{A}(G(\overline{\mathbf{x}}))=0 \text{ and } \overline{\Gamma}\in\mathcal{N}_{\mathcal{K}^n_+}(G(\overline{\mathbf{x}})).$$

By using the fact

$$\overline{\Gamma}\in\mathcal{N}_{\mathcal{K}^n_+}(G(\overline{\mathbf{x}})) \quad\Leftrightarrow\quad G(\overline{\mathbf{x}})=\Pi_{\mathcal{K}^n_+}(G(\overline{\mathbf{x}})+\overline{\Gamma}),$$

we may assume that $A:=G(\overline{x})+\overline{\Gamma}$ has the following decomposition as in (2.5) and (2.7):

$$G(\overline{\mathbf{x}})=Q\begin{bmatrix} U\begin{bmatrix}\Lambda_\alpha & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}U^T & \mathbf{a}\\ \mathbf{a}^T & a_0\end{bmatrix}Q,\quad \overline{\Gamma}=Q\begin{bmatrix} U\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & \Lambda_\gamma\end{bmatrix}U^T & 0\\ 0 & 0\end{bmatrix}Q. \tag{3.3}$$

Then, by (2.10) and (2.11), we have

$$\mathcal{T}_{\mathcal{K}^n_+}(G(\overline{\mathbf{x}}))=\left\{Q\begin{bmatrix}H_1 & \mathbf{h}\\ \mathbf{h}^T & h_0\end{bmatrix}Q\in\mathcal{S}^n: U_{\overline{\alpha}}^TH_1U_{\overline{\alpha}}\succeq 0\right\},$$

$$\mathrm{lin}(\mathcal{T}_{\mathcal{K}^n_+}(G(\overline{\mathbf{x}})))=\left\{Q\begin{bmatrix}H_1 & \mathbf{h}\\ \mathbf{h}^T & h_0\end{bmatrix}Q\in\mathcal{S}^n: U_{\overline{\alpha}}^TH_1U_{\overline{\alpha}}=0\right\}. \tag{3.4}$$

Furthermore, since $\Lambda(\overline{\mathbf{x}})$ is nonempty,

$$\mathcal{C}(\overline{\mathbf{x}}) = \{\mathbf{d} :\ \mathcal{A}(\mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d}) = 0,\ \mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d} \in \mathcal{C}(A; \mathcal{K}^n_+)\},$$

where $\mathcal{C}(A; \mathcal{K}^n_+)$ is defined in (2.13). However, it is not easy to obtain an explicit formula for $\mathrm{aff}(\mathcal{C}(\overline{x}))$. Instead, we define the following outer approximation set to $\mathrm{aff}(\mathcal{C}(\overline{x}))$ with respect to $(\overline{\zeta}, \overline{\Gamma})$ by

$$\mathrm{app}(\overline{\zeta}, \overline{\Gamma}) := \left\{\mathbf{d} : \mathcal{A}(\mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d}) = 0,\ \mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d} \in \mathrm{aff}(\mathcal{C}(A; \mathcal{K}^n_+))\right\}.$$

Suppose $\mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d} = Q \begin{bmatrix} \overline{H}_1 & \bar{\mathbf{h}} \\ \bar{\mathbf{h}}^T & \bar{h}_0 \end{bmatrix} Q$, by (2.14), it holds that

$$\mathrm{app}(\overline{\zeta}, \overline{\Gamma}) := \left\{\mathbf{d} : \mathcal{A}\left(\mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d}\right) = 0,\ U_\beta^T \overline{H}_1 U_\gamma = 0,\ U_\gamma^T \overline{H}_1 U_\gamma = 0\right\}. \quad (3.5)$$

Having stated the first-order optimality condition and the tangent-cone related sets for (3.1), we are ready to state two constraint qualifications. The Robinson constraint qualification (CQ) is said to hold at $\overline{\mathbf{x}}$ if

$$\begin{pmatrix} \mathcal{J}_x(\mathcal{A} \circ G)(\overline{\mathbf{x}}) \\ \mathcal{J}_x G(\overline{\mathbf{x}}) \end{pmatrix} \mathcal{X} + \begin{pmatrix} 0 \\ \mathcal{T}_{\mathcal{K}^n_+}(G(\overline{\mathbf{x}})) \end{pmatrix} = \begin{pmatrix} \mathbb{R}^n \\ \mathcal{S}^n \end{pmatrix},$$

and the constraint nondegeneracy, which is stronger, is said to hold at $\overline{\mathbf{x}}$ if

$$\begin{pmatrix} \mathcal{J}_x(\mathcal{A} \circ G)(\overline{\mathbf{x}}) \\ \mathcal{J}_x G(\overline{\mathbf{x}}) \end{pmatrix} \mathcal{X} + \begin{pmatrix} 0 \\ \mathrm{lin}(\mathcal{T}_{\mathcal{K}^n_+}(G(\overline{\mathbf{x}}))) \end{pmatrix} = \begin{pmatrix} \mathbb{R}^n \\ \mathcal{S}^n \end{pmatrix}. \quad (3.6)$$

The following results provides a criterion for checking the constraint nondegeneracy.

Proposition 7 *Let $\mathcal{X} = \mathbb{R}^m$. Let $\overline{\mathbf{x}}$ be a feasible point of EDMOP and $G(\overline{\mathbf{x}})$ has the decomposition as in* (3.5)*. Then, $\overline{\mathbf{x}}$ is constraint nondegenerate if and only if the following m-dimensional vectors are linearly independent:*

$$\mathbf{v}_\ell := \begin{pmatrix} \left(\frac{\partial G(\overline{\mathbf{x}})}{\partial x_1}\right)_{\ell\ell} \\ \vdots \\ \left(\frac{\partial G(\overline{\mathbf{x}})}{\partial x_m}\right)_{\ell\ell} \end{pmatrix} \quad \text{and} \quad \mathbf{p}_{ij} := \begin{pmatrix} \bar{\mathbf{u}}_i^T Q \frac{\partial G(\overline{\mathbf{x}})}{\partial x_1} Q \bar{\mathbf{u}}_j \\ \vdots \\ \bar{\mathbf{u}}_i^T Q \frac{\partial G(\overline{\mathbf{x}})}{\partial x_m} Q \bar{\mathbf{u}}_j \end{pmatrix}, \quad \begin{array}{l} \ell = 1, \ldots, n \\ 1 \le i \le j \le |\overline{\alpha}|, \end{array}$$

where $\overline{U}_{\overline{\alpha}} = \begin{bmatrix} U_{\overline{\alpha}} \\ 0 \end{bmatrix} := \left[\bar{\mathbf{u}}_1, \ldots, \bar{\mathbf{u}}_{|\overline{\alpha}|}\right]$ and $U_{\overline{\alpha}}$ consists of the eigenvectors in U corresponding to the non-positive eigenvalues (those not in Λ_α) in (3.3).

Proof. By taking the orthogonal complements of both sides in (3.6), we obtain that $\overline{\mathbf{x}}$ is constraint nondegenerate if and only if the following condition holds:

$$\left[\begin{pmatrix} \mathcal{J}_x(\mathcal{A} \circ G)(\overline{\mathbf{x}}) \\ \mathcal{J}_x G(\overline{\mathbf{x}}) \end{pmatrix} \mathbb{R}^m\right]^\perp \cap \left[\begin{pmatrix} \{0\} \\ \mathrm{lin}\left(\mathcal{T}_{\mathcal{K}_n^+}(G(\overline{\mathbf{x}}))\right) \end{pmatrix}\right]^\perp = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (3.7)$$

It follows from (3.4) that

$$
\begin{aligned}
&\operatorname{lin}\left(\mathcal{T}_{\mathcal{K}^n_+}(G(\overline{\mathbf{x}}))\right) = \left\{H \in \mathcal{S}^n : \overline{U}^T_{\overline{\alpha}} QHQ\overline{U}_{\overline{\alpha}} = 0\right\} \\
&= \left\{H \in \mathcal{S}^n : \bar{\mathbf{u}}_i^T QHQ\bar{\mathbf{u}}_j = 0, 1 \le i \le j \le |\overline{\alpha}|\right\} \\
&= \left\{H \in \mathcal{S}^n : \left\langle Q\left(\frac{\bar{\mathbf{u}}_i\bar{\mathbf{u}}_j^T + \bar{\mathbf{u}}_j\bar{\mathbf{u}}_i^T}{2}\right)Q, H\right\rangle = 0, 1 \le i \le j \le |\overline{\alpha}|\right\}.
\end{aligned}
$$

Since

$$
\begin{aligned}
&\left[\begin{pmatrix} \mathcal{J}_x(\mathcal{A}\circ G)(\overline{\mathbf{x}}) \\ \mathcal{J}_x G(\overline{\mathbf{x}}) \end{pmatrix}\mathbb{R}^m\right]^\perp \\
&= \{(\mathbf{s}, Y) \in \mathbb{R}^n \times \mathcal{S}^n : \langle \mathbf{s}, \mathcal{A}\left(\mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d}\right)\rangle + \langle Y, \mathcal{J}_x G(\bar{x})\mathbf{d}, \forall \mathbf{d} \in \mathbb{R}^m\} \\
&= \left\{(\mathbf{s}, Y) \in \mathbb{R}^n \times \mathcal{S}^n : \left\langle \mathcal{A}^*\mathbf{s} + Y, \sum_{k=1}^m d_k \frac{\partial G(\overline{\mathbf{x}})}{\partial x_k}\right\rangle = 0, \forall \mathbf{d} \in \mathbb{R}^m\right\} \\
&= \left\{(\mathbf{s}, Y) \in \mathbb{R}^n \times \mathcal{S}^n : \left\langle \mathcal{A}^*\mathbf{s} + Y, \frac{\partial G(\overline{\mathbf{x}})}{\partial x_k}\right\rangle = 0, k = 1, \ldots, m\right\}.
\end{aligned}
$$

We obtain that (3.7) holds if and only if the following system of linear equations with unknown $a_i, 1 \le i \le n$ and $b_{ij}, 1 \le i \le j \le |\overline{\alpha}|$ have only the zero solution:

$$
\sum_{l=1}^n a_i \left\langle \frac{\partial G(\overline{\mathbf{x}})}{\partial x_k}, \mathbf{e}_i\mathbf{e}_i^T\right\rangle + \sum_{1\le i\le j\le|\overline{\alpha}|} b_{ij}\left\langle Q\left(\frac{\bar{\mathbf{u}}_i\bar{\mathbf{u}}_j^T + \bar{\mathbf{u}}_j\bar{\mathbf{u}}_i^T}{2}\right)Q, \frac{\partial G(\overline{\mathbf{x}})}{\partial x_k}\right\rangle = 0,\ k \le m,
$$

which is equivalent to saying that the vectors $\mathbf{v}_i, 1 \le i \le n$ and $\mathbf{p}_{ij}, 1 \le i \le j \le |\overline{\alpha}|$ are linearly independent. This completes the proof. □

Example 1 (Constraint nondegeneracy for the nearest EDM problem (1.6)) We use the nearest EDM (1.6) to demonstrate the application of Prop. 7. In this case. $\mathbf{x} = X \in \mathcal{S}^n$ and $G(X) = -X$. The vectors $\mathbf{v}_i$ and $\mathbf{p}_{ij}$ become matrices and are respectively denoted by $V_i \in \mathcal{S}^n$ and $P_{(ij)} \in \mathcal{S}^n$, which are given by

$$
V_i = \mathbf{e}_i\mathbf{e}_i^T, \qquad P_{(ij)} = (Q\bar{\mathbf{u}}_i)(Q\bar{\mathbf{u}}_j)^T, \quad 1 \le i \le j \le \overline{\alpha}.
$$

We prove the linear independence of those matrices. It is equivalent to prove the linear independence of the following matrices:

$$
\widehat{V}_i := QV_iQ = Q\mathbf{e}_i\mathbf{e}_i^TQ, \qquad \widehat{P}_{(ij)} := QP_{(ij)}Q = \bar{\mathbf{u}}_i\bar{\mathbf{u}}_j^T.
$$

Consider the linear combination of those matrices:

$$
M := \sum_{i=1}^n \xi_i\widehat{V}_i + \sum_{1\le i\le j}^{|\overline{\alpha}|} \xi_{ij}\widehat{P}_{(ij)} = 0. \tag{3.8}
$$

Using the facts $\bar{\mathbf{u}}_i^T\mathbf{e}_n = 0$ for all $i = 1, \ldots, |\overline{\alpha}|$ and $Q\mathbf{e}_n = -(1/\sqrt{n})\mathbf{1}_n$, we derive from (3.8) that

$$0 = M\mathbf{e}_n = \sum_{i=1}^{n} \xi_i \widehat{V}_i \mathbf{e}_n = -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i Q\mathbf{e}_i. \tag{3.9}$$

Combining with the fact

$$Q\mathbf{e}_i = -\frac{1}{n+\sqrt{n}}\mathbf{1}_n - \frac{1}{\sqrt{n}+1}\mathbf{e}_n + \mathbf{e}_i, \qquad i = 1, \ldots, n-1,$$

the equation (3.9) reduces to the following system:

$$\begin{cases} (\xi_i + \cdots + \xi_{n-1})\mathbf{1}_n + \sqrt{n}(\xi_i + \cdots + \xi_{n-1})\mathbf{e}_n - \\ (n+\sqrt{n})\sum_{i=1}^{n-1} \xi_i \mathbf{e}_i + (\sqrt{n}+1)\xi_n \mathbf{1}_n = 0 \\ \xi_1 + \cdots + \xi_n = 0. \end{cases}$$

The only solution of the system is $\xi_1 = \xi_2 = \cdots = \xi_n = 0$. Consequently, the equation (3.8) becomes

$$\sum_{1 \le i \le j}^{|\overline{\alpha}|} \xi_{ij} \widehat{P}_{(ij)} = \sum_{1 \le i \le j}^{|\overline{\alpha}|} \xi_{ij} \bar{\mathbf{u}}_i \bar{\mathbf{u}}_j^T = 0.$$

For an arbitrary pair (i_0, j_0), computing the inner product of this equation with $\mathbf{u}_{i_0}\mathbf{u}_{j_0}^T$ and using the fact that $\bar{\mathbf{u}}_i^T\bar{\mathbf{u}}_j = 0$ for all $i \neq j$ lead to

$$0 = \xi_{i_0 j_0} \|\mathbf{u}_{i_0}\|^2 \|\mathbf{u}_{j_0}\|^2 = \xi_{i_0 j_0}.$$

Hence, $\xi_{ij} = 0$ for all i, j. This proves that (3.8) has zero as its solution. By Prop. 7, constraint nondegeneracy holds at every feasible point X for the nearest EDM problem (1.6).

Under the Robinson CQ, we can establish the second-order necessary condition at a local optimum. Conversely, for a first-order stationary point with Robinson's CQ, the stationary point becomes a local minimum under a second-order sufficient condition. Those results are included in the following theorem, whose proof is a combination of Proposition 2 and [4, Theorem 3.86].

Theorem 1 *Suppose that $\overline{\mathbf{x}}$ is a locally optimal solution to EDMOP and Robinson's CQ holds at $\overline{\mathbf{x}}$. Then, the following inequality holds:*

$$\sup_{(\zeta,\Gamma)\in\Lambda(\overline{\mathbf{x}})} \langle \mathbf{d}, \mathcal{J}^2_{xx} L(\overline{\mathbf{x}}, \zeta, \Gamma)\mathbf{d}\rangle - \sigma(\Gamma, \mathcal{T}^2_{\mathcal{K}^n_+}(G(\overline{\mathbf{x}}), \mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d})) \ge 0, \quad \forall \mathbf{d} \in \mathcal{C}(\overline{\mathbf{x}}).$$

Conversely, let $\overline{\mathbf{x}}$ be a feasible solution to EDMOP satisfying the first order optimality condition. Suppose that Robinson's CQ holds at $\overline{\mathbf{x}}$. Then the condition

$$\sup_{(\zeta,\Gamma)\in\Lambda(\overline{\mathbf{x}})} \langle \mathbf{d}, \mathcal{J}^2_{xx} L(\overline{\mathbf{x}}, \zeta, \Gamma)\mathbf{d}\rangle - \sigma(\Gamma, \mathcal{T}^2_{\mathcal{K}^n_+}(G(\overline{\mathbf{x}}), \mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d})) > 0, \forall \mathbf{d} \in \mathcal{C}(\overline{\mathbf{x}})\backslash\{0\}.$$

is necessary and sufficient for the quadratic growth condition at the point $\overline{\mathbf{x}}$*:*

$$f(\mathbf{x}) \geq f(\overline{\mathbf{x}}) + c\|\mathbf{x} - \overline{\mathbf{x}}\|^2 \quad \forall \mathbf{x} \in \mathcal{N} \textit{ such that } G(\mathbf{x}) \in \mathcal{K}_+^n \cap \mathcal{S}_h^n$$

for some constant $c > 0$ *and a neighbourhood* $\mathcal{N}$ *of* $\overline{\mathbf{x}}$ *in* $\mathcal{X}$*.*

We end this section with the definition of a strong second order sufficient condition for EDMOP, which will be needed in the sensitivity analysis in next section.

Definition 3 Let $\overline{\mathbf{x}}$ be a stationary point of EDMOP. We say that the strong second order sufficient condition holds at $\overline{\mathbf{x}}$ if for any $\mathbf{d} \in \widehat{\mathcal{C}}(\overline{\mathbf{x}}) \backslash \{0\}$, we have

$$\sup_{(\zeta,\Gamma)\in\Lambda(\overline{\mathbf{x}})} \{\langle \mathbf{d}, \mathcal{J}^2_{xx} L(\overline{\mathbf{x}}, \zeta, \Gamma)\mathbf{d}\rangle - \Upsilon_{G(\overline{\mathbf{x}})}(\Gamma, \mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d})\} > 0, \tag{3.10}$$

where

$$\widehat{\mathcal{C}}(\overline{\mathbf{x}}) := \bigcap_{(\zeta,\Gamma)\in\Lambda(\overline{\mathbf{x}})} \operatorname{app}(\zeta, \Gamma).$$

4 Sensitivity analysis

It is well known that the first-order optimality condition can be reformulated as a system of nonlinear equations. Hence, a KKT point can be regarded as a solution of an equation, which provides an approach to study the behaviour of the KKT point. One such behaviour can be studied via the strong regularity of Robinson [31]. The purpose of this section is to provide a complete characterization of strong regularity of the KKT points of EDMOP.

4.1 Strongly regularity

In this subsection we characterize the strong regularity of a KKT point of EDMOP by using the strong second order optimality condition. First , we note that the KKT condition (3.2) can be equivalently expressed as

$$F(\mathbf{x}, \zeta, \Gamma) = \begin{bmatrix} \mathcal{J}_x L(\mathbf{x}, \zeta, \Gamma) \\ \mathcal{A}(G(\mathbf{x})) \\ -G(\mathbf{x}) + \Pi_{\mathcal{K}_+^n}(G(\mathbf{x}) + \Gamma) \end{bmatrix} \tag{4.1}$$

$$= \begin{bmatrix} \mathcal{J}_x L(\mathbf{x}, \zeta, \Gamma) \\ \mathcal{A}(G(\mathbf{x})) \\ \Gamma - \Pi_{(\mathcal{K}_+^n)^\circ}(G(\mathbf{x}) + \Gamma) \end{bmatrix} = 0.$$

System (4.1) is also equivalent to

$$0 \in \begin{bmatrix} \mathcal{J}_x L(\mathbf{x}, \zeta, \Gamma) \\ \mathcal{A}(G(\mathbf{x})) \\ -G(\mathbf{x}) \end{bmatrix} + \begin{bmatrix} \mathcal{N}_{\mathcal{X}}(\mathbf{x}) \\ \mathcal{N}_{\mathbb{R}^n}(\zeta) \\ \mathcal{N}_{(\mathcal{K}_+^n)^\circ}(\Gamma) \end{bmatrix}. \tag{4.2}$$

Denote $\mathcal{Z} := \mathcal{X} \times \mathbb{R}^n \times \mathcal{S}^n, \mathcal{D} := \mathcal{X} \times \mathbb{R}^n \times (\mathcal{K}^n_+)^\circ$. For any $\mathbf{z} = (\mathbf{x}, \zeta, \Gamma) \in \mathcal{Z}$. Define

$$\phi(\mathbf{z}) := \begin{bmatrix} \mathcal{J}_x L(\mathbf{x}, \zeta, \Gamma) \\ \mathcal{A}(G(\mathbf{x})) \\ -G(\mathbf{x}) \end{bmatrix},$$

then system (4.2) is expressed as a compact form

$$0 \in \phi(\mathbf{z}) + \mathcal{N}_{\mathcal{D}}(\mathbf{z}). \tag{4.3}$$

Let

$$\Pi_{\mathcal{D}}(\mathbf{z}) := (\mathbf{x}, \zeta, \Pi_{(\mathcal{K}^n_+)^\circ}(Y)), \quad \forall\, \mathbf{z} = (\mathbf{x}, \zeta, Y).$$

The normal mapping of the generalized equation (4.3) is defined as follows:

$$\begin{aligned} F(\mathbf{z}) &:= \phi(\Pi_{\mathcal{D}}(\mathbf{z})) + \mathbf{z} - \Pi_{\mathcal{D}}(\mathbf{z}) \\ &= \begin{bmatrix} \mathcal{J}_x L(\mathbf{x}, \zeta, Y - \Pi_{\mathcal{K}^n_+}(Y)) \\ \mathcal{A}(G(\mathbf{x})) \\ -G(\mathbf{x}) + \Pi_{\mathcal{K}^n_+}(Y) \end{bmatrix}. \end{aligned} \tag{4.4}$$

Then $(\overline{\mathbf{x}}, \overline{\zeta}, \overline{\Gamma})$ is a solution of the generalized equation (4.3) if and only if $F(\overline{\mathbf{x}}, \overline{\zeta}, \overline{Y}) = 0$, with $\overline{Y} = \overline{\Gamma} + G(\overline{\mathbf{x}})$.

We now recall the definition of strongly regular solution of the generalized equation (4.3) introduced by Robinson [31].

Definition 4 Let $\overline{\mathbf{z}}$ be a solution of the generalized equation (4.3). It is said that $\overline{\mathbf{z}}$ is strongly regular if there exist neighbourhoods $\mathcal{B}$ of $0 \in \mathcal{Z}$ and $\mathcal{V}$ of $\overline{\mathbf{z}} \in \mathcal{Z}$ such that for every $\omega \in \mathcal{B}$, the following linearized generalized equation

$$\omega \in \phi(\overline{\mathbf{z}}) + \mathcal{J}_z\phi(\overline{\mathbf{z}})(\mathbf{z} - \overline{\mathbf{z}}) + \mathcal{N}_{\mathcal{D}}(\mathbf{z})$$

has a unique solution in $\mathcal{V}$, denoted by $\mathbf{z}_{\mathcal{V}}(\omega)$, and the mapping $\mathbf{z}_{\mathcal{V}} : \mathcal{B} \to \mathcal{V}$ is Lipschitz continuous.

The strong regularity of a solution to the generalized equation is closely related to the Lipschitz homeomorphism of the normal mapping.

Lemma 1 *[15] $(\overline{\mathbf{x}}, \overline{\zeta}, \overline{\Gamma})$ is a strongly regular solution of the generalized equation (4.3) if and only if $F(\cdot)$ in (4.4) is locally Lipschitz homeomorphism near $(\overline{\mathbf{x}}, \overline{\zeta}, \overline{Y})$.*

The next proposition relates the strong second order sufficient condition and constraint nondegeneracy to the nonsingularity of Clarke's Jacobian of the mapping $F(\cdot)$ and the strong regularity of a solution to the generalized equation. The result below was motivated by and patterned after [35, Prop. 3.2].

Theorem 2 *Let $\overline{\mathbf{x}}$ be a locally optimal solution of EDMOP and $(\overline{\zeta}, \overline{\Gamma})$ be the Lagrange multiplier at $\overline{\mathbf{x}}$. Denote $\overline{Y} = \overline{\Gamma} + G(\overline{\mathbf{x}})$. Consider the following three statements :*

(a) The strong second order sufficient condition holds at $\overline{\mathbf{x}}$ *and* $\overline{\mathbf{x}}$ *is constraint nondegenerate.*
(b) Any element in $\partial F(\overline{\mathbf{x}}, \overline{\zeta}, \overline{Y})$ *is nonsingular.*
(c) The KKT point $(\overline{\mathbf{x}}, \overline{\zeta}, \overline{\Gamma})$ *is a strongly regular solution of the generalized equation (4.3).*

It holds that $(a) \Rightarrow (b) \Rightarrow (c)$.

Proof. For $(a) \Rightarrow (b)$, since the constraint nondegeneracy condition holds at $\overline{\mathbf{x}}$, it follows from [4, Proposition 4.50] that $\Lambda(\overline{\mathbf{x}}) = \{(\overline{\zeta}, \overline{\Gamma})\}$ is a singleton. The strong second order sufficient condition takes the form

$$\langle \mathbf{d}, \mathcal{J}_{xx} L(\overline{\mathbf{x}}, \overline{\zeta}, \overline{\Gamma})\mathbf{d}\rangle - \Upsilon_{G(\overline{\mathbf{x}})}(\overline{\Gamma}, \mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d}) > 0, \quad \forall \mathbf{d} \in \mathrm{app}(\overline{\zeta}, \overline{\Gamma})\backslash\{0\}.$$

Let W be any arbitrary element in $\partial F(\overline{\mathbf{x}}, \overline{\zeta}, \overline{Y})$. We shall show that W is nonsingular. Let $(\triangle\mathbf{x}, \triangle\zeta, \triangle Y) \in \mathcal{X} \times \mathbb{R}^n \times \mathcal{S}^n$ be such that

$$W(\triangle\mathbf{x}, \triangle\zeta, \triangle Y) = 0.$$

Then there exists an element $\Psi \in \partial\Pi_{\mathcal{K}_+^n}(\overline{Y})$ such that

$$\begin{aligned} &W(\triangle\mathbf{x}, \triangle\zeta, \triangle Y) \\ &= \begin{bmatrix} \mathcal{J}_{xx} L(\overline{\mathbf{x}}, \overline{\zeta}, \overline{\Gamma})\triangle\mathbf{x} + \mathcal{J}_x\mathcal{A}(G(\overline{\mathbf{x}}))^*\triangle\zeta + \mathcal{J}_x G(\overline{\mathbf{x}})^*(\triangle Y - \Psi(\triangle Y)) \\ -\mathcal{J}_x\mathcal{A}(G(\overline{\mathbf{x}}))\triangle\mathbf{x} \\ -\mathcal{J}_x G(\overline{\mathbf{x}})\triangle\mathbf{x} + \Psi(\triangle Y) \end{bmatrix} = 0. \end{aligned} \tag{4.5}$$

By Proposition 5 and the last two equations of (4.5), we obtain that

$$\triangle\mathbf{x} \in \mathrm{app}(\overline{\zeta}, \overline{\Gamma}). \tag{4.6}$$

Denote $\triangle\Gamma := \triangle Y - \Psi(\triangle Y)$ and from the third equation of (4.5), we have

$$\Psi\left(\triangle\Gamma + \mathcal{J}_x G(\overline{\mathbf{x}})\triangle x\right) = \mathcal{J}_{\mathbf{x}} G(\overline{\mathbf{x}})\Delta\mathbf{x}.$$

Then, by Proposition 6, we obtain

$$\langle \triangle\Gamma, \mathcal{J}_x G(\overline{\mathbf{x}})\triangle\mathbf{x}\rangle \geq -\Upsilon_{G(\overline{\mathbf{x}})}(\overline{\Gamma}, \mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d}).$$

It follows from the first equation of (4.5) that

$$\begin{aligned} 0 &= \langle \triangle\mathbf{x}, \mathcal{J}_{xx} L(\overline{\mathbf{x}}, \overline{\zeta}, \overline{\Gamma})\triangle\mathbf{x} + \mathcal{J}_x\mathcal{A}(G(\overline{\mathbf{x}}))^*\triangle\zeta + \mathcal{J}_x G(\overline{\mathbf{x}})^*\triangle\Gamma\rangle \\ &= \langle \triangle\mathbf{x}, \mathcal{J}_{xx} L(\overline{\mathbf{x}}, \overline{\zeta}, \overline{\Gamma})\triangle\mathbf{x}\rangle + \langle \mathcal{J}_x G(\overline{\mathbf{x}})\triangle\mathbf{x}, \triangle\Gamma\rangle \\ &\geq \langle \triangle\mathbf{x}, \mathcal{J}_{xx} L(\overline{\mathbf{x}}, \overline{\zeta}, \overline{\Gamma})\triangle x\rangle - \Upsilon_{G(\overline{\mathbf{x}})}(\overline{\Gamma}, \mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d}). \end{aligned}$$

Thus, we conclude from (4.6) and the strong second order sufficient condition that

$$\triangle\mathbf{x} = 0, \quad \triangle Y = \triangle\Gamma.$$

Consequently, (4.5) reduces to

$$\begin{bmatrix} \mathcal{J}_x \mathcal{A}(G(\overline{\mathbf{x}}))^* \triangle\zeta + \mathcal{J}_x G(\overline{\mathbf{x}})^* \triangle\Gamma \\ \Psi(\triangle\Gamma) \end{bmatrix} = 0.$$

By the constraint nondegeneracy condition, there exists a vector $\mathbf{d} \in \mathcal{X}$ and $S \in \mathrm{lin}(\mathcal{T}_{\mathcal{K}^n_+}(G(\overline{\mathbf{x}}))$ with $S = Q \begin{bmatrix} S_1 & \mathbf{s} \\ \mathbf{s}^T & s_0 \end{bmatrix} Q$ such that

$$\mathcal{A}(\mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d}) = \triangle\zeta, \;\; \mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d} + S = \triangle\Gamma.$$

From Proposition 5 and $\Psi(\triangle\Gamma) = 0$, where $\triangle\Gamma = Q \begin{bmatrix} \triangle\Gamma_1 & \Delta\xi \\ \triangle\xi^T & \Delta\xi_0 \end{bmatrix} Q$, we obtain

$$\triangle\xi = 0, \; \triangle\xi_0 = 0, \; U_\alpha^T \triangle\Gamma_1 U_\alpha = 0, \; U_\alpha^T \triangle\Gamma_1 U_\beta = 0, \; U_\alpha^T \triangle\Gamma_1 U_\gamma = 0. \quad (4.7)$$

Then one has the following chain of equalities:

$$\begin{aligned} &\langle \triangle\zeta, \triangle\zeta \rangle + \langle \triangle\Gamma, \triangle\Gamma \rangle \\ &= \langle \mathcal{A}\left(\mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d}\right), \Delta\zeta \rangle + \langle \Delta\Gamma, \mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d} \rangle + \langle \Delta\Gamma, S \rangle \\ &= \langle \mathcal{J}_x(\mathcal{A} \circ G)(\overline{\mathbf{x}})^* \triangle\zeta + \mathcal{J}_x G(\overline{\mathbf{x}})^* \triangle\Gamma, \mathbf{d} \rangle + \langle \Delta\Gamma, S \rangle \\ &= \langle \triangle\Gamma, S \rangle \\ &= \langle U^T \triangle\Gamma U, U^T S U \rangle, \end{aligned}$$

which, together with (4.7) and (3.4), implies that

$$\langle \triangle\zeta, \triangle\zeta \rangle + \langle \triangle\Gamma, \triangle\Gamma \rangle = \langle U^T \triangle\Gamma U, U^T S U \rangle = 0.$$

Thus, $\triangle\zeta = 0$, $\triangle\Gamma = 0$, which implies $\triangle Y = 0$. Together with $\triangle\mathbf{x} = 0$, this shows that W is nonsingular.

$(b) \Rightarrow (c)$. By Clarke's inverse function theorem [13, Thm. 4G.1], if every element in $\partial F(\overline{\mathbf{x}}, \overline{\zeta}, \overline{Y})$ is nonsingular, then F is locally Lipschitz homeomorphism near $(\overline{\mathbf{x}}, \overline{\zeta}, \overline{Y})$. Thus, by Lemma 1, $(\overline{\mathbf{x}}, \overline{\zeta}, \overline{\Gamma})$ is strongly regular solution of the generalized equation (4.3). □

4.2 Equivalent Characterizations

In this part, we derive six equivalent characterizations of the strong regularity of a KKT point of EDMOP. We recall from Thm.1 that the quadratic growth condition is equivalent to a second-order sufficient condition under the Robinson CQ. For strong regularity, we need the uniform second order growth condition defined in [4, Definition 5.16].

Let $\mathcal{U}$ be a finite-dimensional space. We say that $f(\mathbf{x}, \mathbf{u})$ and $G(\mathbf{x}, \mathbf{u})$ with $\mathbf{u} \in \mathcal{U}$ is $\mathcal{C}^2$-smooth parametrizations if $f : \mathcal{X} \times \mathcal{U} \to \mathbb{R}$ and $G : \mathcal{X} \times \mathcal{U} \to \mathcal{S}^n$ are twice continuous differentiable and there exists $\overline{\mathbf{u}} \in \mathcal{U}$ such that $f(\cdot, \overline{\mathbf{u}}) = f(\cdot)$

and $G(\cdot,\overline{\mathbf{u}}) = G(\cdot)$. The corresponding parametrized problem of EDMOP takes the form

$$\min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x},\mathbf{u}) \quad \text{s.t.} \quad G(\mathbf{x},\mathbf{u}) \in \mathcal{K}_+^n \cap \mathcal{S}_h^n. \tag{4.8}$$

We say that a parametrization is canonical if $U = \mathcal{X} \times \mathcal{S}^n, \overline{\mathbf{u}} := (0,0) \in \mathcal{X}^* \times \mathcal{S}^n$, and

$$(f(\mathbf{x},\mathbf{u}), G(\mathbf{x},\mathbf{u})) := (f(\mathbf{x}) - \langle \mathbf{u}_1, \mathbf{x}\rangle, G(\mathbf{x}) + \mathbf{u}_2), \mathbf{x} \in \mathcal{X}, \mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2) \in \mathcal{X} \times \mathcal{S}^n.$$

Definition 5 Let $\overline{\mathbf{x}}$ be a stationary point of EDMOP. We say that the uniform second order growth condition holds at $\overline{\mathbf{x}}$ with respect to a $\mathcal{C}^2$-smooth parametrization $(f(\mathbf{x},\mathbf{u}), G(\mathbf{x},\mathbf{u}))$ if there exist $c > 0$ and neighbourhoods $\mathcal{N}_{\overline{\mathbf{x}}}$ of $\overline{\mathbf{x}}$ and $\mathcal{N}_{\overline{\mathbf{u}}}$ of $\overline{\mathbf{u}}$ such that for any $\mathbf{u} \in \mathcal{N}_{\overline{\mathbf{u}}}$ and any stationary point $\mathbf{x}(\mathbf{u}) \in \mathcal{N}_{\overline{\mathbf{x}}}$ of the corresponding parametrized problem (4.8), the following holds:

$$f(\mathbf{x},\mathbf{u}) \geq f(\mathbf{x}(\mathbf{u}),\mathbf{u}) + c\|\mathbf{x} - \mathbf{x}(\mathbf{u})\|^2, \forall \mathbf{x} \in \mathcal{N}_{\overline{\mathbf{x}}},\ G(\mathbf{x},\mathbf{u}) \in \mathcal{K}_+^n \cap \mathcal{S}_h^n. \tag{4.9}$$

The next theorem shows that for EDMOP, the uniform second order growth condition with respect to the canonical parametrization implies the strong second order sufficient condition in Def. 3.

Theorem 3 *Let $\overline{\mathbf{x}}$ be a stationary point of EDMOP. Suppose that Robinson's CQ holds at $\overline{\mathbf{x}}$. If the uniform second order growth condition holds at $\overline{\mathbf{x}}$ with respect to the canonical parametrization, then the strong second order sufficient condition holds at $\overline{\mathbf{x}}$.*

Proof. Let $(\overline{\zeta}, \overline{\Gamma}) \in \Lambda(\overline{\mathbf{x}})$. We may assume that $A := G(\overline{\mathbf{x}}) + \overline{\Gamma}$ has the decomposition as in (2.5), and $G(\overline{\mathbf{x}})$ and $\overline{\Gamma}$ satisfy (3.3). We first prove that there exist $\mathbf{s} \in \mathbb{R}^{n-1}$ and $s_0 \in \mathbb{R}$ such that

$$\operatorname{diag}(QSQ) = 0 \quad \text{with} \quad S := \begin{bmatrix} U_\beta U_\beta^T & \mathbf{s} \\ \mathbf{s}^T & s_0 \end{bmatrix}. \tag{4.10}$$

In fact,

$$S = \underbrace{\begin{bmatrix} 0 & \mathbf{s} \\ \mathbf{s}^T & s_0 \end{bmatrix}}_{:=S_0} + \underbrace{\begin{bmatrix} U_\beta U_\beta^T & 0 \\ 0 & 0 \end{bmatrix}}_{:=B}.$$

Let $\mathbf{b} := -\operatorname{diag}(QBQ)$. Then the requirement $\operatorname{diag}(QSQ) = 0$ means $\operatorname{diag}(QS_0Q) = \mathbf{b}$. By using the properties of Q matrix (2.2), it is easy to verify that the choice

$$\begin{bmatrix} 2\mathbf{s} \\ a_0 \end{bmatrix} = \mathbf{b}$$

satisfies $\operatorname{diag}(QS_0Q) = \mathbf{b}$. Thus, we can claim the following three facts:

(i) There exist $\mathbf{s} \in \mathbb{R}^{n-1}$ and $s_0 \in \mathbb{R}$ such that (4.10)) holds.

(ii) By making use the structure of $\overline{\Gamma}$ in (3.3), we have

$$\langle \overline{\Gamma},\ QSQ\rangle = \langle U_\beta U_\beta^T,\ U_\gamma \Lambda_\gamma U_\gamma\rangle = 0. \tag{4.11}$$

(iii) By making use of (2.3), we have

$$JQSQJ = Q\begin{bmatrix} U_\beta U_\beta^T & 0 \\ 0 & 0\end{bmatrix} Q \succeq 0, \quad \text{implying} \quad QSQ \in \mathcal{K}_+^n. \tag{4.12}$$

Next, we consider the parametrized problem of EDMOP in the following form

$$\begin{array}{ll} \min\limits_{\mathbf{x}\in\mathcal{X}} & f(\mathbf{x}) \\ \text{s.t.} & \mathcal{A}(G(\mathbf{x})) = 0, \\ & G(\mathbf{x}) + \kappa QSQ \in \mathcal{K}_+^n. \end{array} \tag{4.13}$$

where $\kappa > 0$ is the perturbation parameter and S is constructed in (4.10). It follows from (4.12) that any feasible point of EDMOP is also a feasible point of (4.13). Moreover, the property in (4.11) implies that $(\overline{\zeta}, \overline{\Gamma})$ is also a Lagrange multiplier of the perturbed problem at $\overline{\mathbf{x}}$:

$$\mathcal{J}_x L_\kappa(\overline{\mathbf{x}}, \zeta, \Gamma) = 0, \mathcal{A}(G(\overline{\mathbf{x}})) = 0, \begin{cases} \Gamma \in (\mathcal{K}_+^n)^\circ, G(\overline{\mathbf{x}}) + \kappa QSQ \in \mathcal{K}_+^n \\ \langle \Gamma,\ G(\overline{\mathbf{x}}) + \kappa QSQ\rangle = 0, \end{cases} \tag{4.14}$$

where for each $\kappa \in \mathbb{R}$,

$$L_\kappa(\mathbf{x}, \zeta, \Gamma) := L(\mathbf{x}, \zeta, \Gamma) + \kappa\langle \Gamma,\ QSQ\rangle.$$

Since $QSQ \in \mathcal{K}_+^n$ and $G(\overline{\mathbf{x}}) \in \mathcal{K}_+^n$, the complementarity condition in (4.14) implies

$$\langle \Gamma,\ G(\overline{\mathbf{x}})\rangle = 0.$$

This means that any Lagrange multiplier of the perturbed problem is also a Lagrange multiplier of EDMOP, $\Lambda_\kappa(\overline{\mathbf{x}}) \subseteq \Lambda(\overline{\mathbf{x}})$, where $\Lambda_\kappa(\overline{\mathbf{x}})$ is the set of all (ζ, Γ) satisfying the KKT condition (4.14).

Denote $\widehat{A} := G(\overline{x}) + \overline{\Gamma} + \kappa QSQ$. We then have

$$Q\widehat{A}Q = QAQ + \kappa S = \begin{bmatrix} A_1 & \mathbf{a} \\ \mathbf{a}^T & a_0\end{bmatrix} + \kappa \begin{bmatrix} U_\beta U_\beta^T & \mathbf{s} \\ \mathbf{s}^T & s_0\end{bmatrix},$$

and the eigen-decomposition (3.3) implies

$$\widehat{A}_1 := A_1 + \kappa U_\beta U_\beta^T = [U_\alpha, U_\beta]\begin{bmatrix} \Lambda_\alpha & \\ & I_{|\beta|}\end{bmatrix}\begin{bmatrix} U_\alpha^T \\ U_\beta^T\end{bmatrix} + U_\gamma \Lambda_\gamma U_\gamma.$$

This means that $\widehat{A}_1$ is nonsingular and has only positive and negative eigenvalues. Therefore, the critical cone of $\mathcal{K}_+^n$ at $\widehat{A}$ is a subspace and is given by

$$\mathcal{C}(\widehat{A}, \mathcal{K}_+^n) = \left\{ Q\begin{bmatrix} H_1 & \mathbf{h} \\ \mathbf{h}^T & h_0\end{bmatrix} Q \;\middle|\; U_\gamma^T H_1 U_\gamma = 0\right\} \supseteq \mathcal{C}(A, \mathcal{K}_+^n).$$

Consequently, the critical cone $\mathcal{C}_\kappa(\overline{\mathbf{x}})$ of the perturbation problem at $\overline{\mathbf{x}}$ must contain $\mathrm{app}(\overline{\zeta}, \overline{\Gamma})$:

$$\mathcal{C}_\kappa(\overline{\mathbf{x}}) \supseteq \mathrm{app}(\overline{\zeta}, \overline{\Gamma}).$$

Since the uniform second order growth condition of EDMOP holds at $\overline{\mathbf{x}}$ with respect to the canonical parametrization, for any small enough $\kappa > 0$, one has from Thm. 1 that

$$\langle \mathbf{d},\ \mathcal{J}^2_{xx} L_\kappa(\overline{\mathbf{x}}, \overline{\zeta}, \overline{\Gamma})\mathbf{d}\rangle - \Upsilon_{G(\overline{\mathbf{x}})+\kappa\Sigma}(\overline{\Gamma},\ \mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d}) > 0, \quad \forall \mathbf{d} \in \mathcal{C}_\kappa(\overline{\mathbf{x}})\backslash\{0\}.$$

By noting that

$$\mathcal{J}^2_{xx} L_\kappa(\overline{\mathbf{x}}, \zeta, \Gamma) = \mathcal{J}^2_{xx} L(\overline{\mathbf{x}}, \zeta, \Gamma)$$

and

$$\Upsilon_{G(\overline{\mathbf{x}})}(\overline{\Gamma}, \mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d})) = \Upsilon_{G(\overline{\mathbf{x}})+\kappa QSQ}(\overline{\Gamma}, \mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d}), \quad \forall\, \mathbf{d} \in \mathrm{app}(\overline{\zeta}, \overline{\Gamma}),$$

we obtain

$$\sup_{(\zeta,\Gamma)\in\Lambda_\kappa(\overline{\mathbf{x}})} \langle \mathbf{d}, \mathcal{J}^2_{xx} L(\overline{\mathbf{x}}, \zeta, \Gamma)\mathbf{d}\rangle - \Upsilon_{G(\overline{\mathbf{x}})}(\overline{\Gamma}, \mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d}) > 0, \forall \mathbf{d} \in \mathrm{app}(\zeta, \Gamma)\backslash\{0\}.$$

Since $\Lambda_\kappa(\overline{\mathbf{x}}) \subseteq \Lambda(\overline{\mathbf{x}})$, we get

$$\sup_{(\zeta,\Gamma)\in\Lambda(\overline{\mathbf{x}})} \langle \mathbf{d}, \mathcal{J}^2_{xx} L(\overline{\mathbf{x}}, \zeta, \Gamma)\mathbf{d}\rangle - \Upsilon_{G(\overline{\mathbf{x}})}(\overline{\Gamma}, \mathcal{J}_x G(\overline{\mathbf{x}})\mathbf{d}) > 0, \forall \mathbf{d} \in \mathrm{app}(\zeta, \Gamma)\backslash\{0\}.$$

This is the strong order sufficient condition at $\overline{\mathbf{x}}$. □

Another characterization of the strong regularity is through the strong stability of a stationary point, introduced by Kojima in [22]. The following definition of strong stability is from Bonnans and Shapiro [4, Definition 5.33].

Definition 6 Let $\overline{\mathbf{x}}$ be a stationary point of Problem (1.1). We say that $\overline{\mathbf{x}}$ is strongly stable with respect to a $\mathcal{C}^2$-smooth parametrization $(f(\mathbf{x}, \mathbf{u}), G(\mathbf{x}, \mathbf{u}))$ if there exist neighbourhoods $\mathcal{N}_{\overline{\mathbf{x}}}$ of $\overline{\mathbf{x}}$ and $\mathcal{N}_{\overline{\mathbf{u}}}$ of $\overline{\mathbf{u}}$ such that for any $\mathbf{u} \in \mathcal{N}_{\overline{\mathbf{u}}}$, the corresponding parametrized problem (4.8) has a unique stationary point $\mathbf{x}(\mathbf{u}) \in \mathcal{N}_{\overline{\mathbf{x}}}$ and $\mathbf{x}(\cdot)$ is continuous on $\mathcal{N}_{\overline{\mathbf{u}}}$. If this holds for any $\mathcal{C}^2$-smooth parametrization, we say that is $\overline{\mathbf{x}}$ strongly stable.

Now we are ready to state the main result of this subsection.

Theorem 4 *Let $\overline{\mathbf{x}}$ be a locally optimal solution of EDMOP. Suppose that Robinson's CQ holds at $\overline{\mathbf{x}}$ and $\overline{Y} = \overline{\Gamma} + G(\overline{\mathbf{x}})$. Let $(\overline{\zeta}, \overline{\Gamma})$ be a Lagrange multiplier at $\overline{\mathbf{x}}$. Then, the following statements are equivalent:*

(a) The strong second order sufficient condition holds at $\overline{\mathbf{x}}$ and $\overline{\mathbf{x}}$ is constraint nondegenerate.
(b) Any element in $\partial F(\overline{\mathbf{x}}, \overline{\zeta}, \overline{Y})$ is nonsingular.
(c) The KKT point $(\overline{\mathbf{x}}, \overline{\zeta}, \overline{\Gamma})$ is a strongly regular solution to the generalized equation (4.3).

(d) The uniform second order growth condition holds at $\overline{\mathbf{x}}$ *and* $\overline{\mathbf{x}}$ *is constraint nondegenerate.*
(e) The point $\overline{\mathbf{x}}$ *is strongly stable and* $\overline{\mathbf{x}}$ *is constraint nondegenerate.*
(f) F *is locally Lipschitz homeomorphism near* $(\overline{\mathbf{x}}, \overline{\zeta}, \overline{Y})$.

Proof. We know from Theorem 2 that $(a) \Rightarrow (b) \Rightarrow (c)$ and by the Lemma 1 $(c) \Leftrightarrow (f)$. The relations $(c) \Leftrightarrow (d) \Leftrightarrow (e)$ follow from [4, Theorems 5.24, 5.35]. Finally, by Theorem 3, we have $(d) \Rightarrow (a)$. The proof is completed. □

5 Numerical Implication

In this part, we demonstrate the usefulness of our obtained theoretical results using two instances of the nearest EDM problem. The first one is (1.6), which is continuously differentiable. Our second example will be non-differentiable and we will show how the strong regularity results in Thm. 4 will ensure that the problem has a unique solution and guarantees a global convergence of an alternating direction method of multipliers to the unique solution.

5.1 Strong Regularity of NEDM (1.6)

We first consider the problem (1.6). We make the following assumption on the weight matrix W:

Assumption H1: We assume $W_{ij} > 0$ for all $i \neq j$.

Let $\overline{D} \in \mathcal{X}$ be an optimal solution of (1.6). In Example 1, we proved that the constraint nondegeneracy is satisfied at any feasible point of (1.6). It follows from [4, Prop. 4.50] that the Lagrange multipliers $\Lambda(\overline{D}) = \{(\overline{\zeta}, \overline{\Gamma})\}$ is a singleton. Moreover, for any $H \in \mathrm{app}(\overline{\zeta}, \overline{\Gamma})$, the first constraint in (3.5) becomes $\mathrm{diag}\,(H) = 0$. Hence,

$$\mathrm{app}(\overline{\zeta}, \overline{\Gamma}) \subseteq \mathcal{S}^{\mathrm{n}}_{\mathrm{h}}.$$

This means for any $0 \neq H \in \mathrm{app}(\overline{\zeta}, \overline{\Gamma})$, there must be a pair (i, j) with $i \neq j$ such that $H_{ij} \neq 0$. Then the strong second order sufficient condition becomes

$$\begin{aligned} &\langle H, \mathcal{J}^2_{XX} L(\overline{D}, \overline{\zeta}, \overline{\Gamma}) H\rangle - \Upsilon_{\overline{X}} \\ &= \sum_{i \neq j} W_{ij}^2 H_{ij}^2 - \Upsilon_{\overline{X}}(\overline{\Gamma}, H) > 0, \quad \forall\ H \in \mathrm{app}(\overline{\zeta}, \overline{\Gamma}) \backslash \{0\}, \end{aligned}$$

where we used the fact $-\Upsilon_{\overline{D}}(\overline{\Gamma}, H) \geq 0$. Therefore, the following is a direct consequence of Thm. 4.

Proposition 8 *Let* $\overline{D} \in \mathcal{X}$ *be an optimal solution of NEDM and suppose Assumption H1 is satisfied. Then the following hold:*

(a) The strong second order sufficient condition holds at $\overline{D}$.
(b) Any element in $\partial F(\overline{D}, \overline{\zeta}, \overline{\Gamma})$ *is nonsingular.*
(c) The KKT point $(\overline{X}, \overline{\zeta}, \overline{\Gamma})$ *is a strongly regular solution to the generalized equation (4.3).*

5.2 Strong Regularity of NEDM with Stress Loss

One critique of the NEDM model (1.6) is that it tries to approximate a given measurement δ_{ij} through the approximation to its squared measurement:

$$D_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|^2 \approx \delta_{ij}^2.$$

It then uses the least-squares criterion to get the best approximation, which favours large distances over short ones. An alternative and widely used approximation is

$$\sqrt{D_{ij}} = \|\mathbf{x}_i - \mathbf{x}_j\| \approx \delta_{ij}.$$

That is, we directly approximate δ_{ij} by Euclidean distance $\|\mathbf{x}_i - \mathbf{x}_j\|$ instead of approximating its squared value. Applying the least-squares criterion yields the following problem:

$$\min_D \ f(D) := \frac{1}{2}\|\sqrt{W} \circ (\sqrt{D} - \Delta)\|^2, \qquad \text{s.t.} \ \ \operatorname{diag}(D) = 0, \ \ -D \in \mathcal{K}^n_+, \quad (5.1)$$

where $\Delta_{ij} = \delta_{ij}$. We make the following comments on (5.1).

(i) In the field of Multi-Dimension Scaling (MDS) [6], the objective function in (5.1) is called the stress loss function while the objective function in (1.6) is called the squared-stress loss function. The embedding quality from the stress loss is often better than that from the squared-stress loss. The squared-stress function is quadratic in D and is strongly convex in $\text{app}(\overline{\zeta}, \overline{\Gamma})$. That is why we can claim that the solution of (1.6) is strongly regular in Prop. 8. However, the stress loss function in (5.1) is not quadratic in D and is not differentiable when some $D_{ij} = 0$, $i \neq j$.

(ii) The non-differentiability of the stress function not only poses a computational challenge, but also renders the theoretical results obtained not valid directly to (5.1). For example, one question is whether the optimal solution is still strongly regular under some reasonable assumptions. We answer this question below by studying an alternating direction method of multipliers (ADMM) for (5.1).

The ADMM is stated on an equivalent reformulation of (5.1). Let

$$f_1(D) := \frac{1}{2}\|\sqrt{W} \circ (\sqrt{D} - \Delta)\|^2 + \mathcal{I}_{\mathcal{S}^n_h}(D), \qquad f_2(Z) := \mathcal{I}_{\mathcal{K}^n_+}(Z),$$

where $\mathcal{I}_{\mathcal{C}}(\cdot)$ is the indicator function on the set $\mathcal{C}$. Then NEDM (5.1) is equivalent to

$$\min_{D,Z} \ f_1(D) + f_2(Z) \quad \text{s.t.} \quad D + Z = 0. \qquad (5.2)$$

The augmented Lagrange function for (5.2) is given by

$$\mathcal{L}(D, Z; Y) = f_1(D) + f_2(Z) + \langle Y,\ D + Z\rangle + \frac{\rho}{2}\|D + Z\|^2,$$

where $Y \in \mathcal{S}^n$ is the multiplier and $\rho > 0$ is the penalty parameter.

Algorithm 1 ADMM for (5.1)

1: Initialization: D^0, Z^0, Y^0, $\rho > 0$ and $\nu \in (0, (\sqrt{5}+1)/2)$. set the iteration index $k := 0$.

2: For $k = 0, 1, \dots,$ update

$$\begin{cases} D^{k+1} = \ \text{argmin} \ \ \mathcal{L}(D, Z^k; Y^k) \\ Z^{k+1} = \ \text{argmin} \ \ \mathcal{L}(D^{k+1}, Z; Y^k). \end{cases} \tag{5.3}$$

3: Update Y^k by

$$Y^{k+1} = Y^k + \nu\rho\Big(D^{k+1} + Z^{k+1}\Big).$$

We note that Algorithm 1 is a special case of the Proximal ADMM of [16] in its Appendix B by choosing $S = T = 0$ therein. We have proved the constraint nondegeneracy in Example 1 for the constraints in (5.1). Since the objective is convex in D and the solution set is nonempty and bounded under the assumption H2 below, there exists a Lagrange multiplier for (5.1). Consequently, the reformulated problem (5.2) also has a Lagrange multiplier. Moreover, the matrices

$$\Sigma_{f_1} + \rho I \quad \text{and} \quad \Sigma_{f_2} + \rho I \tag{5.4}$$

both are positive definite for $\rho > 0$, where Σ_{f_1} and Σ_{f_2} are the positive semidefinite operators associated with f_1 and f_2 respectively (for the definition of such Σ matrices, please refer to [16, Eq. (B.4), Eq. (B.5)]). The existence of Lagrange multipliers and the positive definiteness of the two matrices in (5.4) are the conditions used for the global convergence of the proximal ADMM.

Therefore, the sequence $\{D^k, Z^k, Y^k\}$ generated by Alg. 1 converges to a KKT point of (5.2), denoted as $(\overline{D}, \overline{Z}, \overline{Y})$ with $\overline{D}$ being an optimal solution of (5.1). We show that $\overline{D}$ is also strongly regular under the assumption H2 below.

Assumption H2: We assume $\delta_{ij} > 0$ for all $i \neq j$.

In the following, we prove that $\overline{D}_{ij} > 0$ for all $i \neq j$. We first see how D^{k+1} is calculated.

$$\begin{aligned} D^{k+1} &= \ \text{argmin} \ \ \frac{1}{2}\|\sqrt{W} \circ (\sqrt{D} - \Delta)\|^2 + \langle Y^k,\ D\rangle + \frac{\rho}{2}\|D + Z^k\|^2 \\ &= \ \text{argmin} \ \ \frac{\rho}{2}\|D + Z^k + (W + Y^k)/\rho\|^2 - \langle W \circ \Delta,\ \sqrt{D}\rangle \\ &= \ \text{argmin} \ \ \frac{1}{2}\|D + \widehat{Z}^k\|^2 - \frac{1}{\rho}\langle W \circ \Delta,\ \sqrt{D}\rangle, \qquad \text{s.t.} \ \ \text{diag}\,(D) = 0,\ D \geq 0 \end{aligned}$$

where $\widehat{Z}^k := Z^k + (W + Y^k)/\rho$. It is easy to see that the variables D_{ij} are separable and hence they can be computed individually.

$$\begin{cases} D_{ij}^{k+1} = \text{argmin}_{D_{ij}\geq 0} \ \frac{1}{2}\Big(D_{ij} + \widehat{Z}_{ij}^k\Big)^2 - \frac{1}{\rho} W_{ij}\delta_{ij}\sqrt{D_{ij}}, & \text{for } i \neq j \\ D_{ii}^{k+1} = 0, & \text{for } i = 1, \ldots, n. \end{cases} \tag{5.5}$$

For each D_{ij}^{k+1} with $i \neq j$, it is actually the solution of the following one-dimensional optimization problem with properly defined ω and η:

$$\texttt{dcroot}(\omega, \eta) := \arg\min_{x\geq 0} \ \frac{1}{2}(x-\omega)^2 - 2\eta\sqrt{x}.$$

The solution of this problem has been studied in [43] and it is closely related to the roots of so-called depressed cubic equation. This is the reason why we named the solution as `dcroot` depending on ω and η, which define the one-dimensional problem. We omit the computational formula for $\texttt{dcroot}(\omega, \eta)$. We only need its nice property that it is bounded away from 0.

Proposition 9 *[43, Prop. 3.4] Suppose $c_1 > 0$ and $c_2 > 0$ are given two constants. There exists a constant $c_0 > 0$ such that*

$$|\textit{dcroot}(\omega, \eta)| \geq c_0 \qquad \forall\, (\omega, \eta) \ \ \textit{satisfying} \ \ |\omega| \leq c_1, \quad \eta \geq c_2.$$

The update D_{ij}^{k+1} for $i \neq j$ is given by

$$D_{ij}^{k+1} = \texttt{dcroot}(\omega_{ij}^{(k)}, \eta_{ij}), \qquad \text{with} \quad \omega_{ij}^{(k)} := -\widehat{Z}_{ij}^k \quad \text{and} \quad \eta_{ij} := \frac{1}{2\rho} W_{ij}\delta_{ij}.$$

Since

$$\lim_{k\to\infty} \widehat{Z}^k = \overline{Z} + (W + \overline{Y})/\rho,$$

and under Assumptions H1 and H2, there exist constants c_1 and c_2 such that

$$|\omega_{ij}^{(k)}| = |\widehat{Z}_{ij}^k| \leq c_1 \quad \text{and} \quad \eta_{ij} \geq c_2 := \min_{s\neq t} \ W_{st}\delta_{st} \qquad \forall\, i \neq j \quad \text{and} \quad \forall\, k.$$

Prop.9 implies that there exists $c_0 > 0$ such $D_{ij}^{k+1} \geq c_0$ for all k and $i \neq j$. We further note that the objective $f(D)$ in (5.1) has the following form:

$$f(D) = \frac{1}{2}\|\sqrt{W} \circ (\sqrt{D} - \Delta)\|^2 = \frac{1}{2}\sum_{i\neq j} \Big(\sqrt{W_{ij}}(\sqrt{D_{ij}} - \delta_{ij})\Big)^2,$$

which does not include the term for $i = j$. This is because $\operatorname{diag}(D) = 0$.

What we have proved is that the objective function $f(D)$ is continuously differentiable at $\overline{D}$ and it is strictly convex. Therefore the strong second order sufficient condition holds

$$\begin{aligned} &\langle H, \mathcal{J}_{XX}\mathcal{L}(\overline{D}, \overline{\zeta}, \overline{\Gamma})H\rangle - \Upsilon_{\overline{X}}(\overline{\Gamma}, H) \\ &= \sum_{i\neq j} \frac{1}{4} W_{ij}\delta_{ij}\Big(\overline{D}_{ij}\Big)^{-3/2} H_{ij}^2 - \Upsilon_{\overline{X}}(\overline{\Gamma}, H) > 0, \quad \forall\, H \in \text{app}(\overline{\zeta}, \overline{\Gamma})\backslash\{0\}, \end{aligned}$$

where we used that facts that $-\Upsilon_{\overline{X}}(\overline{\Gamma}, H) \geq 0$, $W_{ij}\delta_{ij} > 0$ and $\overline{D}_{ij} > 0$ for all $i \neq j$. The constraint nondegeneracy of (5.1) has been proved in Example 1. Hence we can make a strong claim about problem (5.1) by Theorem 4. We state it below.

Proposition 10 *Under Assumptions H1 and H2, the sequence $\{D^k\}$ generated by ADMM Alg. 1 converges to the unique and strongly stable solution $\overline{D}$ of Problem (5.1). Moreover, the unique KKT point of (5.1) at $\overline{D}$ is strongly regular.*

6 Conclusion

In this paper, we studied perturbation analysis on the Euclidean Distance Matrix OPtimization (EDMOP), which have found many applications in distance related problems. We established various characterizations of strong regularity and strong stability at a locally optimal solution of EDMOP. Those results are related to the second-order information of EDMOP and are particularly useful in justifying when Newton-type methods are efficient. For example, Theorem 4 can be used to explain why a semimsooth Newton method studied in [27] is quadratically convergent for the nearest Euclidean distance matrix problem (1.6). We also demonstrated an implication of the obtained results to a first-order method (ADMM) for the stress minimization problem. Therefore, the theoretical results established have important numerical applications too.

This paper also motivates some interesting questions. The first one is whether the constraint nondegeneracy property in Example 1 still holds if some fixed-distance constraints are added. Such constraints appear in Euclidean Distance Matrix Completion, see, e.g., [1]. The second question is whether the constraint nondegeneracy holds for the correlation calibration problem described in (1.7). This would open the possibility of developing fast Newton's method for it. The third question is whether one can develop Newton's method with efficient implementation for the stress minimization problem (5.1). This question is valid because we showed in Prop. 10 that the optimal solution is strongly stable. Hence, Newton's method would enjoy quadratic convergence rate. However, the obstacle is that the objective function is not everywhere differentiable. Finally, one referee proposed whether an interior-point method can be developed for (5.1).

Acknowledgements We thank the editor and the two referees for their detailed comments that have improved the quality of the paper. In particular, one referee points us to the relevant papers [2, 3] on SDP approach for EDM optimization.

Data Availability Statement: We do not analyse or generate any datasets, because our work proceeds within a theoretical and mathematical approach.
Funding: The research of Guo was partially supported by NSFC 11801057 and Qi was partially supported by IEC/NSFC/191543 and P0045347.

Conflict of Interest Declaration: The authors have no relevant financial or non-financial interests to disclose. Furthermore, the authors have no competing or conflict of interests to declare that are relevant to the content of this article.

References

1. A. Y. Alfakih, A. Khandani, and H. Wolkowicz, *Solving Euclidean distance matrix completion problems via semidefinite programming*, Comput. Optim. Appl., 1999, 12: 13-30.
2. A. Y. Alfakih and H. Wolkowicz, *Matrix completion problems*, In *Handbook of semidefinite programming*, volume 27 of Internat. Ser. Oper. Res. Management Sci., 533-545, Kluwer Acad. Publ., Boston, MA, 2002.
3. S. Al-Homidan and H. Wolkowicz, *Approximate and exact completion problems for Euclidean distance matrices using semidefinite programming*, Linear Algebra Appl., 406 (2005), 109-141.
4. J. F. Bonnans, and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.
5. J. F. Bonnans, C. H. Ramirez, *Perturbation analysis of second-order cone programming problems*, Math. Program. Ser.B, 2005, 104:205-227.
6. I. Borg and P.J.F. Groenen, *Modern Multidimensional Scaling: Theory and Applications*, 2nd ed., Springer Ser. Statist., Springer, New York, 2005.
7. Z. X. Chan and D. F. Sun, *Constraint nondegeneracy, strong regularity and nonsingularity in semidefinite programming*, SIAM J. Optim., 19 (2008), 370-396.
8. J. Dattorro, *Convex Optimization and Euclidean Distance Geometry*, Meboo Publishing USA, Palo Alto, CA, 2005.
9. C. Ding and H.-D. Qi, *Convex optimization learning of faithful Euclidean distance representations in nonlinear dimensionality reduction*, Math. Program., 164 (2017), 341-381.
10. C. Ding, D.F. Sun, and K.C Toh, *An introduction to a class of matrix cone programming*, Maths. Prog., 144 (2014), 141-179.
11. C. Ding, D.F. Sun and L.W. Zhang, *Characterization of the robust isolated calmness for a class of conic programming problems*, SIAM J. Optim., 27 (2017), 67-90.
12. I. Dokmanic, R. Parhizkar, J. Ranieri and M. Vetterli, *Euclidean distance matrices: Essential theory, algorithms, and applications*, IEEE Signal Process. Mag., 32 (2015), 12-30.
13. A.L. Dontchev and R.T. Rockafellar, *Implicit Functions and Solution Mappings*, Springer 2009.
14. D. Drusvyatskiy, N. Krislock, T.-L. Voronin, H. Wolkowicz, *Noisy Euclidean distance realization: robust facial reduction and the Pareto frontier*. SIAM J. Optim. 27 (2017), 2301-2331.
15. F. Facchinei and J. S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems: Volume I, Volumn II*, Springer, New York, 2003.
16. M. Fazel, T.K. Pong, D. Sun, and P. Tseng, *Hankel matrix rank mininization with applications to system identification and realization*, SIAM J. Matrix Anal. Appl. 34 (2013), 946–977.
17. W. Glunt, T. L. Hayden, S. Hong, and J. Wells, *An alternating projection algorithm for computing the nearest Euclidean distance matrix*, SIAM J. Matrix Anal. Appl., 11 (1990) 589-600.
18. N. Gaffke and R. Mathar, *A cyclic projection algorithm via duality*, Metrika, 36 (1989), 29-54.
19. T. L. Hayden and J. Wells, *Approximation by matrices positive semidefinite on a subspace*, Linear Algebra Appl., 109 (1988), 115-130.
20. C.R. Johnson, B. Kroschel, and H. Wolkowicz, *An interior-point method for approximate positive semidefinite completion*, Comput. Optim. Appl., 9 (1998), 175-190.
21. D. Klatte and B. Kummer, *Nonsmooth Equations in Optimization:Regularity, Calculus, Methods and Applications*, Kluwer Academic Publishers, Boston, 2002.

22. M. Kojima, *Strongly stable stationary solutions in nonlinear programs*, Analysis and Computation of Fixed Points, Academic Press, New York, 1980, 93-138.
23. N. Krislock and H. Wolkowicz, *Explicit sensor network localization using semidefinite representations and facial reductions*, SIAM J. Optim., 20 (2010), 2679-2708.
24. L. Liberti, C. Lavor, N. Maculan, and A. Mucherino, *Euclidean distance geometry and applications*, SIAM Review, 56 (2014), 3-69.
25. C.A. Micchelli, *Interpolation of scattered data: distance matrices and conditionally positive definite functions*, Constr. Approx. 2 (1986), 11-22.
26. B.S. Mordukhovich, *Variational Analysis and Applications*, Springer, 2018.
27. H. D. Qi, *A semismooth Newton method for the nearest Euclidean distance matrix problem*, SIAM J. Matrix Anal. Appl., 34 (2013), 67–93.
28. H.-D. Qi, *Conditional quadratic semidefinite programming: examples and methods*, J. Oper. Res. Soc. China 2 (2014) 143-170.
29. H.-D Qi and X. Yuan, *Computing the nearest Euclidean distance matrix with low embedding dimensions*, Math. Prog., 147 (2014), 351-389.
30. R. Rebonato, *Volatility and Correlation* (2nd Ed.), John Wiley & Sons Ltd, 2004.
31. S.M. Robinson, *Strongly regular generalized equations*, Math. Oper. Res., 5 (1980), 43-62.
32. R.T. Rockafellar and R.-B. Wets, *Variational Analysis*, Springer Verlag, 1998.
33. I.J. Schoenberg *Remarks to Maurice Frechét's article "Sur la définition axiomatique d'une classe d'espace distancés vectoriellement applicable sur l'espace de Hilbert"*, Annals of Mathematics 36 (1935), 724-732.
34. A. Shapiro, *First and second order analysis of nonlinear semidefinite programs*, Math. Prog., 77 (1997), 301-320.
35. D.F. Sun, *The strong second order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications*, Math. Oper. Res., 31 (2006), 761-776.
36. D.F. Sun and J. Sun, *Semismooth matrix valued functions*, Math. Oper. Res., 27 (2002), 150-169.
37. J. Sun, S. Boyd and L. Xiao, and P. Diaconis, *The fastest mixing Markov process on a graph and a connection to a maximum variance unfolding problem*, SIAM Review, 48 (2006), 681-699.
38. Y. Wang and L.W. Zhang, *Nonsingularity in second-order cone programming via the smoothing metric projector*, Science in China Series A, 53 (2010), 1025-1038.
39. K.Q Weinberger and K.S. Lawrence, *An introduction to nonlinear dimensionality reduction by maximum variance unfolding*, In AAAI, 6 (2006), 1683-1686.
40. H. Zhang, Y. Liu, and H. Lei, *Localization from incomplete Euclidean distance matrix: Performance analysis for the svd–mds approach*, IEEE Trans. Sig. Processing, 67 (2019), 2196-2209.
41. Q. Zhang, X. Zhao, and C. Ding, *Matrix optimization based Euclidean embedding with outliers*, Comput. Optim. Appl., 79 (2021), 235-271.
42. S. Zhou, N. Xiu, and H.-D. Qi, *Robust Euclidean embedding via EDM optimization* Math. Prog. Comput., 12 (2020), 337-387.
43. S. Zhou, N. Xiu, and H.-D. Qi, *A fast matrix majorization-projection method for penalized stress minimization with box constraints*, IEEE Trans. Sig. Processing, 66 (2018), 4331-4346.