Approximate and Exact Completion Problems for Euclidean Distance Matrices using Semidefinite Programming

Suliman Al-Homidan[*]　　Henry Wolkowicz[†]

April 30, 2004

University of Waterloo
Department of Combinatorics & Optimization
Waterloo, Ontario N2L 3G1, Canada
Preliminary Research Report CORR 2004-15

Keywords: Euclidean Distance Matrix, Completion Problems, Nearest Matrix Approximation, Semidefinite Programming, Large Sparse Problems.

Abstract

A partial pre-distance matrix A is a matrix with zero diagonal and with certain elements *fixed* to given nonnegative values; the other elements are considered *free*. The Euclidean distance matrix completion problem chooses nonnegative values for the free elements in order to obtain a Euclidean distance matrix, EDM. The nearest (or approximate) Euclidean distance matrix problem is to find a Euclidean

---

[*] Department of Mathematical Sciences, King Fahd University of Petroleum and Minerals (KFUPM), Saudia Arabia. Email homidan@kfupm.edu.sa

[†] Research supported by The Natural Sciences and Engineering Research Council of Canada. Email hwolkowica@uwaterloo.ca

[0] URL for paper: http://orion.math.uwaterloo.ca/˜hwolkowi/henry/reports/ABSTRACTS.html

distance matrix, EDM, that is nearest in the Frobenius norm to the matrix A, when the free variables are discounted.

In this paper we introduce two algorithms: one for the exact completion problem and one for the approximate completion problem. Both use a reformulation of EDM into a semidefinite programming problem, SDP. The first algorithm is based on an implicit equation for the completion that for many instances provides an explicit solution. The other algorithm is based on primal-dual interior-point methods that exploit the structure and sparsity. Included are results on maps that arise that keep the EDM and SDP cones invariant.

We briefly discuss numerical tests.

Contents

1 Introduction

An $n \times n$ real symmetric matrix $D = (D_{ij})$ with nonnegative elements and zero diagonal is called a *pre-distance matrix*. In addition, if there exist points $x_1, x_2, \ldots, x_n$ in $\mathbb{R}^r$ such that

$$D_{ij} = \|x_i - x_j\|_2^2, \qquad i, j = 1, 2, \ldots, n, \tag{1.1}$$

then D is called a *Euclidean distance matrix*, denoted EDM. The smallest value of r is called the *embedding dimension* of D. Let A be a pre-distance matrix, H be an $n \times n$ symmetric (weight) matrix with nonnegative elements, and let $\|A\|_F = \sqrt{trace A^T A}$ denote the Frobenius norm of A. We consider the objective function

$$f_N(D) = \frac{1}{2}\|H \circ (A - D)\|_F^2, \tag{1.2}$$

where $\circ$ denotes Hadamard product. The *weighted, nearest (closest), Euclidean distance matrix problem* is

$$\text{(NEDM)} \qquad \begin{array}{rcc} \mu^* = & \min & f_N(D) \\ & \text{subject to} & D \in \mathcal{E} \end{array} \tag{1.3}$$

where $\mathcal{E}$ denotes the *convex cone of* EDM s. The (unknown) EDM is replaced, $D \leftarrow \mathcal{L}(X)$, where $X \in \mathcal{P}$, $\mathcal{P}$ denotes the cone of positive semidefinite matrices, and $\mathcal{L}$ is a linear transformation introduced below.

We also consider the exact completion problem, denoted EDMC (see (4.3) below for details)

$$\text{(EDMC)} \qquad \begin{array}{rcl} \mu^* := & \min & f_C(X) := \frac{1}{2}\|X\|^2 \\ & \text{subject to} & \mathcal{A}(X) = b \in \mathbb{R}^m \\ & & X \in \mathcal{P}. \end{array} \tag{1.4}$$

Here, the EDM problem is translated to a semidefinite programming (SDP) problem. The linear operator $\mathcal{A}$ forces the interpolation conditions corresponding to the fixed elements. It is formed using $\mathcal{L}$ mentioned above.

In this paper we solve NEDM in (1.3) using SDP . The mapping between EDM and SDP uses the linear transformation $\mathcal{L}$, see e.g. [2, 1]. In particular, we provide a stable algorithm that is particularly effective when the given matrix A is large and sparse. Our algorithm specifically exploits the equivalence between the EDM problem and optimization over the cone of positive semidefinite matrices. The algorithm uses the Gauss-Newton search direction with a preconditioned conjugate gradient method. The approach follows that in [26, 5].

We then provide an implicit solution for EDMC in (1.4), i.e. the optimal X is found from the solution of a ($m \times m$, positive definite) system of equations. In many instances this implicit solution becomes an explicit solution from a linear system of equations. This approach allows one to solve huge completion problems of order $n \cong 10^6$, as long as the number of fixed values is only moderately large, of order $m \cong 10^3$. Our empirical tests show that, generically, these large problems can be solved quickly and robustly.

A discussion of the complexity of EDMC is given in e.g. [20, 19, 18]. Special cases (e.g. chordal graphs) are shown to be completable in polynomial time. The complexity of other models of EDM are given in e.g. [22]. Previous SDP approaches appear in e.g. [4]. A general geometric description of EDM is at http://www.stanford.edu/˜dattorro/EDM.pdf.

1.1 Outline

We complete this Section 1 with notation. Section 2 describes the relations between the EDM and SDP cones. We introduce standard linear transformations that map between these two cones. In addition, we present known and new properties for these maps.

Section 3 presents several characterizations of EDM using SDP . These include using the linear operator $\mathcal{L}$, our main tool in our algorithms. Section 4 derives the quadratic SDP s that solve the two EDM problems. We include explicit expressions for the Perron root and vector for both $\mathcal{L}\mathcal{L}^*$ and $\mathcal{L}^*\mathcal{L}$. These linear operators hold the EDM and SDP cone invariant, respectively.

Section 5 presents the duality and optimality conditions for both quadratic programs. These are used to derive our two algorithms. We include details on the evaluations of the operators involved in the optimality conditions.

The primal-dual algorithm for NEDM is outlined in Section 6. We include explicit expressions for a diagonal preconditioner. A discussion on computational results for both algorithms appears in Section 7. Concluding remarks are given in Section 8.

1.2 Notation

We define several linear transformations between vector spaces. For a linear transformation $\mathcal{K} : \mathbb{R}^s \to \mathbb{R}^t$, the adjoint is denoted $\mathcal{K}^*$ and defined by $\langle \mathcal{K}^* v, w \rangle = \langle v, \mathcal{K} w \rangle$, $\forall v \in \mathbb{R}^t, w \in \mathbb{R}^s$. The unit vectors in $\mathbb{R}^t$ are denoted by e_i and e is the vector of ones.

For a general rectangular matrix $M \in \mathcal{M}^{m\times n}$, $v = \operatorname{vec}(M) \in \mathbb{R}^{mn}$ forms a vector from the columns of M. The inverse mapping, vec^{-1}, and the adjoint mapping, vec^*, are given by $\operatorname{Mat} = \operatorname{vec}^{-1} = \operatorname{vec}^*$. The linear transformation $\operatorname{diag}(X)$ is the vector formed from the diagonal of the (square) matrix X. Then the adjoint $\operatorname{Diag} := \operatorname{diag}^*$ forms a diagonal matrix from a given vector. We define $\operatorname{offDiag}(X) := X - \operatorname{Diag}(\operatorname{diag}(X))$.

We use the *trace inner product* $\langle M, N \rangle := \operatorname{trace} M^T N$, which induces the Frobenius norm. With this inner product, Mat (and vec) is an isometry. We use the same inner product on the space of $n \times n$ symmetric matrices, $\mathcal{S}^n$. The identity matrix is denoted by I_n or I when the meaning is clear. The basis elements (unit symmetric matrices) are $E_{ij} := \frac{1}{\sqrt{2}}\left(e_i e_j^T + e_j e_i^T\right)$, if $i \neq j$, and $E_{ii} := e_i e_i^T$. The cone of positive semidefinite matrices (SDP) in $\mathcal{S}^n$ is denoted by $\mathcal{P}_n$ (or by $\mathcal{P}$, when the meaning is clear). Positive semidefiniteness (resp. definiteness) is denoted by $\succeq$ (resp. $\succ$). The cone of EDM is denoted by $\mathcal{E}_n$ (or by $\mathcal{E}$ when the meaning is clear).

For $X \in \mathcal{S}^n$, let $x = \operatorname{svec} X \in \mathbb{R}^{t(n)}$, with $t(n) = n(n+1)/2$, be the vector obtained columnwise from the upper triangular part of X, where the strictly upper triangular part is multiplied by $\sqrt{2}$. The multiplication by $\sqrt{2}$ guarantees that the mapping is an isometry. Let $\operatorname{sMat} := \operatorname{svec}^{-1}$ denote the inverse mapping into $\mathcal{S}^n$. The adjoint operator $\operatorname{sMat}^* = \operatorname{svec}$, since

$$\langle \operatorname{sMat}(v), S \rangle = \operatorname{trace} \operatorname{sMat}(v) S \quad = \quad v^T \operatorname{svec}(S) = \langle \operatorname{svec}(S), v \rangle .$$

We similarly define usvec, usMat for the strictly upper-triangular part of a symmetric matrix. We also need the operator (and its adjoint)

$$\mathcal{D}_e(X) := \operatorname{diag}(X) e^T + e \operatorname{diag}(X)^T, \; \mathcal{D}_e^*(Y) = 2\operatorname{Diag}(Ye),$$

For $M \in \mathcal{M}^{m\times n}$, the space of $m \times n$ real matrices, we let $M^\dagger$ denote the *Moore-Penrose* generalized inverse, e.g. [6]. Then $P_{\mathcal{R}(M)} = MM^\dagger$ and $P_{\mathcal{R}(M^T)} = M^\dagger M$ are the orthogonal projections onto the ranges of M and M^T, respectively.

We collect the definition of various linear transformations and their adjoints in Appendix A.

2 Geometry of EDM and SDP

We list some known facts about the closed convex cones $\mathcal{E}, \mathcal{P}$, see e.g. [23, 13, 15, 24, 4, 27, 8]. We include new relationships between the two closed convex cones. In particular, the dimension of $\mathcal{E}$ is $n(n-1)/2$ and we can map (one-one and onto) $\mathcal{E}$ to any face of $\mathcal{P}$ with the same dimension, i.e. any face with matrices in the relative interior having rank $n-1$. We now provide some details.

It is well known that a pre-distance matrix D is a EDM if and only if D is negative semidefinite on M, the orthogonal complement of the vector of ones, e,

$$M := e^\perp = \left\{x \in \mathbb{R}^n : x^T e = 0\right\}.$$

Define the $n \times n$ orthogonal matrix

$$Q := \left[\frac{1}{\sqrt{n}} e \mid V\right], \qquad Q^T Q = I.$$

Thus

$$V^T e = 0, \quad V^T V = I, \quad V \in \mathcal{M}^{n,n-1}. \tag{2.1}$$

The subspace M can be represented as the range of V $(M = \mathcal{R}(V))$ and

$$J := P_M = VV^T = I - \frac{ee^T}{n} \tag{2.2}$$

is the orthogonal projection onto M. (Here $V^\dagger = V^T$.)

We also define the matrix

$$W \in \mathcal{M}^{n,n-1}, \quad W^T e = 0, \quad W \text{ full column rank.} \tag{2.3}$$

Then

$$W^\dagger W = I_{n-1}, \quad WW^\dagger = P_{\mathcal{R}(W)} = J,$$

where $P_{\mathcal{R}(W)} = P_{\mathcal{R}(V)}$ denotes the orthogonal projection onto the range of W. Thus V above can be considered as a special case of W with orthonormal columns.

Now define the *centered* and *hollow* subspaces of $\mathcal{S}^n$

$$\begin{array}{rcl} \mathcal{S}_C & := & \{B \in \mathcal{S}^n \, : Be = 0\}, \\ \mathcal{S}_H & := & \{D \in \mathcal{S}^n \, : \mathrm{diag}\,(D) = 0\}, \end{array} \tag{2.4}$$

and the linear transformations (with their adjoints which are easily verified)

$$\begin{array}{rcl} \mathcal{K}(B) & := & \mathrm{diag}(B)\, e^T + e\, \mathrm{diag}(B)^T - 2B \\ & := & \mathcal{D}_e(B) - 2(B); \\ \mathcal{K}^*(D) & = & 2\mathrm{Diag}\,(De) - 2D \\ & = & 2(\mathrm{Diag}\,(De) - D). \end{array} \tag{2.5}$$

And define the self-adjoint linear operator

$$\bar{\mathcal{T}}(D) \ := \ -\tfrac{1}{2} JDJ \quad (= \bar{\mathcal{T}}^*(D)). \tag{2.6}$$

The operator $-2\bar{\mathcal{T}}$ is an orthogonal projection onto $\mathcal{S}_C$; thus it is a self-adjoint idempotent. We denote this by $\bar{\mathcal{T}}$ rather than $\mathcal{T}$ (the latter notation is customary in the literature, e.g. [10]) since we modify it below.

Theorem 2.1 *The linear operators satisfy*

$$\begin{array}{c} \mathcal{K}(\mathcal{S}_C) = \mathcal{S}_H, \\ \bar{\mathcal{T}}(\mathcal{S}_H) = \mathcal{S}_C, \end{array}$$

and $\mathcal{K}_{|\mathcal{S}_C}$ and $\bar{\mathcal{T}}_{|\mathcal{S}_H}$ are inverses of each other.

Proof. See e.g. [13, 16]. ■

To get a proper relationship for $\mathcal{K}$ and $\mathcal{K}^\dagger$, we modify $\bar{\mathcal{T}}$ and use the linear operator

$$\mathcal{T}(D) := \bar{\mathcal{T}}(\mathrm{offDiag}\,(D)) = -\frac{1}{2} J \mathrm{offDiag}\,(D) J,$$

where offDiag denotes the orthogonal projection onto the hollow matrices, i.e. zeroing out the diagonal. We now have the following relationships.

Proposition 2.2 *The generalized inverse*

$$\mathcal{K}^\dagger = \mathcal{T}. \tag{2.7}$$

Moreover,

$$\mathcal{R}(\mathcal{K}) = \mathcal{S}_H, \qquad \mathcal{N}(\mathcal{K}) = \mathcal{R}(\mathcal{D}_e) \tag{2.8}$$

$$\mathcal{R}(\mathcal{K}^*) = \mathcal{R}(\mathcal{T}) = \mathcal{S}_C, \qquad \mathcal{N}(\mathcal{K}^*) = \mathcal{N}(\mathcal{T}) = \operatorname{Diag}(\mathbb{R}^n) \tag{2.9}$$

$$\mathcal{S}^n = \mathcal{S}_H \oplus \operatorname{Diag}(\mathbb{R}^n) = \mathcal{S}_C \oplus \mathcal{R}(\mathcal{D}_e). \tag{2.10}$$

Proof. Let $X \in \mathcal{S}^n$, $x = \operatorname{diag}(X)$ and $B = \mathcal{D}_e(X)$. Then $B = xe^T + ex^T$, $\operatorname{diag}(B) = 2x$, and $\mathcal{K}(B) = 2xe^T + 2ex^T - 2B = 2\mathcal{D}_e(X) - 2\mathcal{D}_e(X) = 0$, i.e. this proves that $\mathcal{N}(\mathcal{K}) \supset \mathcal{R}(\mathcal{D}_e)$. The equations in (2.8) now follow from Theorem 2.1 and a dimension argument, i.e. using the fact that the dimensions of $\mathcal{S}_H, \mathcal{R}(\mathcal{D}_e)$ are $n(n-1)/2, n$, respectively.

The equations in (2.9) follow similarly. The equations in (2.10) now follow from the orthogonality of the nullspace and range space of the adjoint for any linear transformation.

The Moore-Penrose generalized inverse $\mathcal{K}^\dagger$ can be characterized by: $\mathcal{K}^\dagger$ equals the inverse of $\mathcal{K}$ when restricted to the range of $\mathcal{K}$, $\mathcal{S}_H$; the null space of $\mathcal{K}^\dagger$ is the orthogonal complement of the range of $\mathcal{K}$, i.e. the orthogonal complement of $\mathcal{S}_H$ which is the set of diagonal matrices. This follows from Theorem 2.1 and the above proof. ■

Details on the relationships between the EDM and SDP cones are discussed in the literature, e.g. see [16, 18]. We include the following.

Theorem 2.3 *The linear operators $\mathcal{T}, \mathcal{K}$ are* **one-one and onto** *mappings between the cone $\mathcal{E}$ in S_H and the face of the semidefinite cone*

$$\mathcal{F}_{\mathcal{E}} := \mathcal{P} \cap \mathcal{S}_C,$$

i.e.

$$\mathcal{T}(\mathcal{E}) = \mathcal{F}_{\mathcal{E}}, \qquad \mathcal{K}(\mathcal{F}_{\mathcal{E}}) = \mathcal{E}.$$

Proof. Note that $\mathcal{F}_{\mathcal{E}} = \{B \succeq 0 : Be = 0\} = \{B \succeq 0 : e^T Be = 0\} = \{B \succeq 0 : \operatorname{trace} ee^T B = 0\}$, i.e. $\mathcal{F}_{\mathcal{E}}$ is the face of $\mathcal{P}$ that is complementary to the face (ray) through ee^T in $\mathcal{P}$. The dimension of $\mathcal{F}_{\mathcal{E}}$ is $n(n-1)/2$, which is the same as the dimension of the subspace Sc. And $\mathcal{F}_{\mathcal{E}} = V\mathcal{P}_{n-1}V^T = W\mathcal{P}_{n-1}W^T$, where V, W are defined as in (2.1) and (2.3). The result now follows from Theorem 2.1. ■

Remark 2.4 *From the above Theorem 2.3, we see that we can transform the* EDM *problem to a* SDP *problem using the above two linear operators* $\mathcal{T}, \mathcal{K}$. *The cone* $\mathcal{E}$ *is mapped to the face* $\mathcal{F}_{\mathcal{E}}$ *so that we can optimize using this face. However, it can be more advantageous to rotate (using e.g. a congruence) and use another face of* $\mathcal{P}$, *e.g. to exploit the rank deficiency of the face and get a row and column of zeros. We explore these possibilities below.*

We now introduce the composite operators

$$\mathcal{K}_W(X) \quad := \quad \mathcal{K}(WXW^T), \tag{2.11}$$

and

$$\mathcal{T}_W(D) \quad := \quad W^{\dagger}\mathcal{T}(D)(W^T)^{\dagger} = -\tfrac{1}{2}W^{\dagger}J\mathrm{offDiag}\,(D)J(W^T)^{\dagger}, \tag{2.12}$$

where W is defined in (2.3).

Lemma 2.5 *Suppose that* W *satisfies the definition in (2.3). Then*

$$\begin{aligned} \mathcal{K}_W(\mathcal{S}^{n-1}) &= \mathcal{S}_H, \\ \mathcal{T}_W(\mathcal{S}_H) &= \mathcal{S}^{n-1}, \end{aligned}$$

and $\mathcal{K}_W = \mathcal{T}_W^{\dagger}$.

Proof. This immediately follows from Theorem 2.1 and the definition of W. Note that $\mathcal{N}(\mathcal{T}_W) = \mathrm{Diag}\,(\mathbb{R}^n) = \mathcal{S}_H^{\perp}\mathcal{R}(\mathcal{T}_W^*)$. ■

From (2.3) and (2.5) we get that

$$\mathcal{K}_W^*(D) \quad = \quad W^T\mathcal{K}^*(D)W \tag{2.13}$$

is the adjoint operator of $\mathcal{K}_W$. The following corollary summarizes useful relationships between $\mathcal{E}$, the cone of Euclidean distance matrices of order n, and $\mathcal{P}$, the cone of positive semidefinite matrices of order $n-1$.

Corollary 2.6 *Suppose that* W *is defined as in (2.3). Then:*

$$\begin{array}{rcl} \mathcal{K}_W(\mathcal{P}\,) & = & \mathcal{E}, \\ \mathcal{T}_W(\mathcal{E}) & = & \mathcal{P}\,. \end{array}$$

Proof. We saw earlier that D is EDM if and only if $D = \mathcal{K}(B)$ with $Be = 0$ and $B \succeq 0$. Let $X = W^T B W$, then since $Be = 0$ we have $B = WXW^T$. Therefore, $WXW^T \succeq 0$ if and only if $X \succeq 0$; and the result follows using the definitions (2.11) and (2.12) and Lemma 2.5. ■

Remark 2.7 *Note that the* $n \times (n-1)$ *matrix* V *as defined in (2.1) is not unique. One example is*

$$V := \begin{bmatrix} y & y & \dots & y \\ 1+x & x & \dots & x \\ x & 1+x & \dots & x \\ \vdots & \vdots & \ddots & \vdots \\ x & x & \dots & 1+x \end{bmatrix}, \tag{2.14}$$

where $x = \frac{-1}{n+\sqrt{n}}$ *and* $y = \frac{-1}{\sqrt{n}}$*. With this choice, it can be easily verified that* $V^T e = 0$*,* $V^T V = I$*, and* $VV^T = J$ *as required by (2.1). A sparse choice for* W *is*

$$W_s = \begin{pmatrix} e^T \\ -I_{n-1} \end{pmatrix}. \tag{2.15}$$

Note that if $B = WXW^T$*,* $X \succeq 0$*,* W *is defined in* (2.3)*, and* H *is the Householder matrix (orthogonal, symmetric, reflection)*

$$H = I - \frac{2}{v^T v} v v^T, \quad \text{with } v = \begin{pmatrix} \frac{1}{\sqrt{n}-1} e_{n-1} \\ -1 \end{pmatrix}, \tag{2.16}$$

then $HBH \succeq 0$ *and* $(HBH)_{nn} = 0$*, which implies that the last row (and column) are zero. We denote the face*

$$\mathcal{Z}_{\mathcal{E}} := \{B \succeq 0 : B_{nn} = 0\}. \tag{2.17}$$

3 SDP **Characterizations for** EDM s

We can exploit the geometry of the EDM and SDP cones and get several different characterizations for EDM . These can have numerical advantages if chosen properly. Many different characterizations appear in e.g. [4, 2, 1]. In this paper we concentrate on a characterization used in [2, 1], see Item 4 in Theorem 3.2 below.

We first prove the following.

Lemma 3.1 *Let* $X \in \mathcal{S}^{n-1}$ *and partition*

$$\mathcal{L}(X) := \begin{pmatrix} 0 & \operatorname{diag}(X)^T \\ \operatorname{diag}(X) & \mathcal{D}_e(X) - 2X \end{pmatrix} = \begin{pmatrix} 0 & d^T \\ d & \bar{D} \end{pmatrix} := D. \tag{3.1}$$

Then

$$\mathcal{L}^*(D) = 2\{\operatorname{Diag}(d) + \operatorname{Diag}(\bar{D}e) - \bar{D}\}, \quad \mathcal{L}^\dagger(D) = \frac{1}{2}(de^T + ed^T - \bar{D}).$$

Proof. The adjoint $\mathcal{L}^*$ is easily verified.

We show that

$$\mathcal{L}^\dagger \mathcal{L}(X) = X,$$

i.e.

$$\begin{aligned} \mathcal{L}^\dagger \mathcal{L}(X) &= \mathcal{L}^\dagger \begin{pmatrix} 0 & \operatorname{diag}(X)^T \\ \operatorname{diag}(X) & \mathcal{D}_e(X) - 2X \end{pmatrix} \\ &= \frac{1}{2}(\operatorname{diag}(X)e^T + e\operatorname{diag}(X)^T - (\mathcal{D}_e(X) - 2X)) \\ &= X \end{aligned}$$

This shows that $\mathcal{L}^\dagger$ is a left inverse of $\mathcal{L}$ and so must be the Moore-Penrose generalized inverse. (We note that $\mathcal{L}\mathcal{L}^\dagger = \operatorname{offDiag}$, i.e. it is the orthogonal projection onto the hollow matrices.) ■

Theorem 3.2 *Suppose that the matrices* W, H *are defined by* (2.3),(2.16), *respectively. Then the following are equivalent:*

1. $D \in \mathcal{E}$
2. $D = \mathcal{K}_W(X)$, *for some* $X \succeq 0$, $X \in \mathcal{S}^{n-1}$

3. *$D = \begin{pmatrix} 0 & \operatorname{diag}(X)^T + (s_X e^T - 2x_r^T) \\ \operatorname{diag}(X) + (s_X e - 2x_r) & \mathcal{D}_e(X) - 2X \end{pmatrix}$, for some $X \in \mathcal{P}_{n-1}$, where $s_X := e_{n-1}^T X e_{n-1},\ x_r := X e_{n-1}$*

4. *$D = \mathcal{L}(X) := \begin{pmatrix} 0 & \operatorname{diag}(X)^T \\ \operatorname{diag}(X) & \mathcal{D}_e(X) - 2X \end{pmatrix}$, for some $X \in \mathcal{P}_{n-1}$*

5. *$D = \mathcal{K}(B)$, for some $B \succeq 0$, with $Be = 0$, $B \in \mathcal{S}^n$*

6. *$D = \mathcal{K}(HBH)$, for some $B \succeq 0$, with $B_{nn} = 0$, $B \in \mathcal{S}^n$*

Proof. Note that $B \succeq 0, B_{nn} = 0$ implies that the last (n-th) row and column of B are zero. We show that Item 1 is equivalent to the other Items.

(2.) The equivalence follows from Corollary 2.6.

(3.) The equivalence follows from Item 2 by setting $W = W_s$ defined in (2.15). Denote

$$B_s := W_s X W_s^T = \begin{pmatrix} e_{n-1}^T \\ -I_{n-1} \end{pmatrix} X \begin{pmatrix} e_{n-1} & -I_{n-1} \end{pmatrix} = \begin{pmatrix} s_X & -x_r^T \\ -x_r & X \end{pmatrix}.$$

For $X \in \mathcal{P}_{n-1}$, we now evaluate

$$\begin{array}{rcl} \mathcal{K}_{W_s}(X) & = & \mathcal{D}_e(W_s X W_s^T) - 2 W_s X W_s^T \\ & = & \begin{pmatrix} s_X \\ \operatorname{diag}(X) \end{pmatrix} e^T + \begin{pmatrix} s_X & \operatorname{diag}(X)^T \end{pmatrix} - 2B_s \\ & = & \begin{pmatrix} 2s_X & s_X e^T + \operatorname{diag}(X)^T \\ s_X e + \operatorname{diag}(X) & \mathcal{D}_e(X) \end{pmatrix} - 2B_s. \end{array}$$

(4.) Let $D \in \mathcal{E}$ be partitioned in the usual fashion as in (3.1), and set $W = W_s$ as in (2.15). Since J is the projection on the orthogonal complement of e, we get $JW_s = W_s$. Then

$$X = \mathcal{L}^\dagger(D) = \frac{1}{2}(de^T + ed^T - \bar{D}) = -\frac{1}{2} W_s^T D W_s = W_s^T \left(-\frac{1}{2} JDJ \right) W_s,$$

i.e. X is a congruence of a positive semidefinite matrix and so $X \succeq 0$. From Lemma 3.1, $\mathcal{L}\mathcal{L}^\dagger$ is the orthogonal projection on the hollow matrices, i.e. we conclude that

$$D = \mathcal{L}(X), \quad \text{where } X = \mathcal{L}^\dagger(D) \succeq 0.$$

Conversely, let $X \in \mathcal{P}^{n-1}$. Then $X = ZZ^T$ with $Z \in \mathcal{M}^{n-1 \times r}$. Let

$$p_1 = 0, \quad p_i = Z^T e_{i-1} \quad i = 2, \ldots, n.$$

Now for the first column and row of $D = \mathcal{L}(X)$ (diag (X))

$$\|p_1 - p_i\|_2^2 = \|p_i\|_2^2 = e_{i-1}^T X e_{i-1} = X_{i-1i-1} = D_{1i}$$

for $\bar{D}$

$$\begin{aligned} \|p_i - p_j\|_2^2 &= (e_{i-1} - e_{j-1})^T X (e_{i-1} - e_{j-1}) \\ &= X_{i-1i-1} + X_{j-1j-1} - 2X_{i-1j-1} \\ &= D_{1i} + D_{1j} - 2[\frac{1}{2}(D_{1i} + D_{1j} - D_{ij})] = D_{ij}, \end{aligned} \tag{3.2}$$

which shows that $p_1, p_2, \ldots, p_n$ are the n points satisfying the condition in (1.1). Equation (3.2) is equivalent to the part $\mathcal{D}_e(X) - 2X$ in $\mathcal{L}(X)$, and so $\mathcal{L}(X) = D$ is EDM.

(5.) The equivalence follows from Theorem 2.3.

(6.) The equivalence follows from the definition of the Householder matrix H, i.e. the mapping $H \cdot H$ is one-one and onto between the faces $\mathcal{F}_{\mathcal{E}}$ and $\mathcal{Z}_{\mathcal{E}}$.

■

4 Formulations into Quadratic SDP s

4.1 Nearest EDM Problem, NEDM

Since A is a given pre-distance matrix, diag$(A) =$ diag$(D) = 0$. Therefore, we can assume without loss of generality that diag$(H) = 0$. Note that $H_{ij} = 0$ means that D_{ij} is free, while $H_{ij} > 0$ forces D_{ij} to approximate A_{ij}. If we want $D_{ij} = A_{ij}$ exactly, then we can add a linear constraint to the program. (This is done in Section 4.2, in EDMC .) Recall that the graph of H is *connected* if for all indices $i \neq j$ there is a *path* of indices $i_1, i_2, ..., i_k$ such that $H_{i,i_1} \neq 0, H_{i_1,i_2} \neq 0, \ldots, H_{i_{k-1},i_k} \neq 0, H_{i_k,j} \neq 0$, see e.g. [9]. Thus, we can assume that the graph of H is connected or the problem can be solved more simply as two smaller problems. It is shown in [4] that Slater's condition holds for NEDM if the graph is connected.

By abuse of notation, define

$$f_N(X) := \|H \circ (A - \mathcal{L}(X))\|_F^2.$$

We now apply Theorem 3.2 and get the following problem, equivalent to NEDM.

$$(\text{NEDM}) \quad \begin{array}{rcc} \mu^* := & \min & f_N(X) \\ & \text{subject to} & X \succeq 0. \end{array}$$

Also, note that $X \in \mathcal{S}^{n-1}$. It is in this lower dimensional space that we solve the problem. We can recover the optimal distance matrix using the optimal X and the relation

$$D = \mathcal{L}(X).$$

Using finite precision, we can never solve the approximation problem exactly. In addition, we need to calculate the embedding dimension. The following lemma, from [4], shows we lose little in the objective function if we choose a small embedding dimension using a numerical rank approach, i.e. if we only discard *small* eigenvalues, then the change in the objective function is small.

Lemma 4.1 *([4]) Suppose that X^* solves (NEDM). Let $\bar{X}$ be the closest symmetric matrix to X^* with rank k, i.e. we set the smallest $n-k$ eigenvalues of X^* to 0, $\lambda_{k+1} = \ldots \lambda_n = 0$. Then,*

$$\sqrt{f(\bar{X})} \leq \sqrt{f(X^*)} + 2\gamma\left(\sqrt{n}+1\right)\sqrt{\sum_{i=k+1}^{n} \lambda_i^2}, \tag{4.1}$$

where $\gamma := \max_{ij} H_{ij}$. ■

4.2 EDM Completion Problem, EDMC

We now consider the exact completion problem, i.e. we are given certain fixed elements of a EDM matrix A, while the other elements are unknown (free). We want to complete this matrix to an EDM. We model this as a quadratic programming problem. As above, we assume that we are given a pre-distance matrix A, but with the proviso that it is completable to a EDM, i.e. free elements can be found to obtain a EDM. Therefore, for given $b \in \mathbb{R}^m$, and a set of indices (columnwise, $k \cong ij$)

$$\mathcal{S} = \{(i,j) : A_{i,j} = \frac{1}{\sqrt{2}} b_k \text{ is known, fixed, } i < j\}, \ |\mathcal{S}| = m, \tag{4.2}$$

define the quadratic program

$$\text{(EDMC)} \qquad \begin{array}{rcc} \mu^* := & \min & f(X) := \frac{1}{2}\|X\|_F^2 \\ & \text{subject to} & \mathcal{A}(X) = b \\ & & X \succeq 0, \end{array} \tag{4.3}$$

where $b \in \mathbb{R}^{|\mathcal{S}|}$ has components $b_k = \sqrt{2} A_{ij}$ and the constraint $\mathcal{A} = \mathcal{I} \cdot \mathcal{L} : \mathcal{S}^{n-1} \to \mathbb{R}^{|\mathcal{S}|}$ yields the interpolation conditions

$$\mathcal{A}(X)_{ij} = \operatorname{trace} E_{ij} \mathcal{L}(X) = b_k, \quad \forall k \cong (ij) \in \mathcal{S},$$

thus defining the interpolation operator $\mathcal{I}$. The transformation $\mathcal{I}$ is equivalent to usvec, though it only yields a subset of the upper triangular elements of the matrix. Therefore, for $s \neq t$, we assume that the interpolation conditions arise from

$$\langle E_{st}, D \rangle = 2 \frac{1}{\sqrt{2}} D_{st} = \sqrt{2} D_{st} = b_{st} \cong b_k.$$

4.3 Invariant Cones and Perron Roots

4.3.1 Invariant for NEDM

Lemma 4.2 *Let H be an $n \times n$ symmetric matrix with nonnegative elements and 0 diagonal and with no zero row (or column). Then*

$$X \succeq 0 \ \ (\textit{resp.} \ \succ 0) \Rightarrow \mathcal{L}^*(H^{(2)} \circ \mathcal{L}(X)) \succeq 0 \ \ (\textit{resp.} \ \succ 0),$$

i.e. the cone $\mathcal{P}$ (and its interior) is invariant under the operator $\mathcal{W} = \mathcal{L}^*(H^{(2)} \circ \mathcal{L}(\cdot))$.

Proof. Note that $\mathcal{W}(\cdot) = \mathcal{L}^*(H^{(2)} \circ \mathcal{L}(\cdot)) = (H \circ \mathcal{L})^*(H \circ \mathcal{L})(\cdot)$. Let $X \succeq 0$ be given and $Y \succeq 0$ be any other positive semidefinite matrix. Then

$$\left\langle Y, \mathcal{L}^*(H^{(2)} \circ \mathcal{L}(X)) \right\rangle = \langle H \circ \mathcal{L}(Y), H \circ \mathcal{L}(X) \rangle = \langle H \circ D_1, H \circ D_2 \rangle,$$

for some $D_1, D_2 \in \mathrm{EDM}$. This shows that $\left\langle Y, \mathcal{L}^*(H^{(2)} \circ \mathcal{L}(X))\right\rangle \geq 0$, by the nonnegativity of H and EDM s. Therefore,

$$\mathcal{L}^*(H^{(2)} \circ \mathcal{L}(X)) \succeq 0, \quad \forall X \succeq 0. \tag{4.4}$$

Note that $Z \succeq 0$ is singular if and only if there exists $0 \neq Y \succeq 0$ such that $\operatorname{trace} ZY = 0$. Therefore, to show that positive definite holds in (4.4), suppose that $X \succ 0, Y \succeq 0$ and

$$\left\langle Y, \mathcal{L}^*(H^{(2)} \circ \mathcal{L}(X))\right\rangle = 0.$$

Therefore, $\left\langle X, \mathcal{L}^*(H^{(2)} \circ \mathcal{L}(Y))\right\rangle = 0$. Since the nullspace of $\mathcal{L}^*$ is the set of diagonal matrices, we conclude that $H^{(2)} \circ \mathcal{L}(Y) = 0$. This implies that $Y = 0$. ■

4.3.2 Invariance for EDMC

The above Lemma 4.2 means that we can apply the generalized Perron-Frobenius Theorem to the operator $\mathcal{W}$, [25, 12, 7], i.e. the spectral radius corresponds to a positive real eigenvalue with a corresponding eigenvector in the (relative) interior of the cone. In particular, we can get the following explicit expressions for Perron eigenpairs.

Corollary 4.3 *The Perron root and eigenvector for*

$$\mathcal{W}_E(\cdot) := \mathcal{L}^*\mathcal{L}(\cdot)$$

are

$$\lambda = (2n+1) + \frac{\sqrt{(4n+2)^2 - 32}}{2} > 0, \qquad X = \alpha ee^T + \beta I \succ 0,$$

where $\alpha = -\frac{4\beta}{\lambda},\ \beta > 0$.

Proof. We confirm the Perron eigenvector matrix X and Perron root λ, i.e.

$$\mathcal{W}_E(X) := \mathcal{L}^*\left(\mathcal{L}(X)\right) = \lambda X,$$

where

$$X = \alpha ee^T + \beta I,$$

and

$$\lambda X = \lambda \alpha e e^T + \lambda \beta I. \tag{4.5}$$

Now $\operatorname{diag}(X) = (\alpha+\beta)e$ and

$$\mathcal{D}_e(X) = \operatorname{diag}(X)e^T + e\operatorname{diag}(X)^T = 2(\alpha+\beta)ee^T.$$

Therefore

$$\begin{array}{rcl}
\mathcal{L}(X) & = & \begin{pmatrix} 0 & \operatorname{diag}(X)^T \\ \operatorname{diag}(X) & \mathcal{D}_e(X) - 2(X) \end{pmatrix} \\
 & = & \begin{pmatrix} 0 & (\alpha+\beta)e^T \\ (\alpha+\beta)e & 2(\alpha+\beta)ee^T - 2(\alpha ee^T + \beta I) \end{pmatrix} \\
 & = & \begin{pmatrix} 0 & (\alpha+\beta)e^T \\ (\alpha+\beta)e & 2\beta(ee^T - I) \end{pmatrix} = \begin{pmatrix} 0 & d^T \\ d & \bar{D} \end{pmatrix} = D,
\end{array}$$

thus defining D, d, and $\bar{D}$. We now get

$$\begin{array}{rcl}
\mathcal{W}_E(X) & = & \mathcal{L}^*\left(\mathcal{L}(X)\right) = \mathcal{L}^*(D) \\
 & = & 2\{\operatorname{Diag}(d) + \operatorname{Diag}(\bar{D}e) - \bar{D}\} \\
 & = & 2\{\operatorname{Diag}((\alpha+\beta)e) + \operatorname{Diag}(2\beta(ee^T - I)e) - 2\beta(ee^T - I)\} \\
 & = & 2\{\operatorname{Diag}\{(\alpha+\beta+2\beta(n-1))e\} - 2\beta(ee^T - I)\} \\
 & = & 2\{\operatorname{Diag}\{(\alpha+\beta(2n-1))e\} - 2\beta(ee^T - I)\}
\end{array}$$

The eigenvalue-eigenvector equation now yields

$$\mathcal{W}_E(X) = 2\{\operatorname{Diag}\{(\alpha+\beta(2n-1))e\} - 2\beta(ee^T - I)\} = \lambda\alpha ee^T + \lambda\beta I = \lambda X,$$

i.e.

$$\lambda\alpha(E - I) + \lambda(\alpha+\beta)I \quad = \quad 2\{(\alpha+\beta(2n-1))I - 2\beta(E - I)\}.$$

Equating the diagonal terms yields

$$\lambda(\alpha+\beta) = 2(\alpha+\beta(2n-1)),$$

while the off-diagonal terms yield

$$\lambda\alpha = -4\beta.$$

We now solve and get $\alpha = \frac{-4\beta}{\lambda}$. We substitute and get

$$\lambda \frac{-4\beta}{\lambda} + \lambda\beta = \frac{-8\beta}{\lambda} + 4n\beta - 2\beta.$$

We can cancel λ in the first term and β on both sides to get

$$-4 + \lambda = \frac{-8}{\lambda} + 4n - 2.$$

This is equivalent to the quadratic equation

$$\lambda^2 - (4n+2)\lambda + 8 = 0,$$

i.e.

$$\lambda = (2n+1) + \frac{\sqrt{(4n+2)^2 - 32}}{2},$$

which is clearly positive. Moreover, using induction, we can show $\beta > n\alpha$ so that $X \succ 0$. ■

We get similar results for $\mathcal{LL}^*$.

Lemma 4.4 *The cone $\mathcal{E}$ is invariant under the linear operator $\mathcal{V} = \mathcal{LL}^*$. In fact, D a pre-distance matrix implies that $\mathcal{V}(D) \in \mathcal{E}$.*

Proof. Suppose that $D = \begin{pmatrix} 0 & d^T \\ d & \bar{D} \end{pmatrix}$ is a pre-distance matrix. Since D is nonnegative elementwise with zero diagonal, the matrix $X = \mathcal{L}^*(D) = 2\left(\text{Diag}\,(d) + \text{Diag}\,(\bar{D}e) - \bar{D}\right)$ is positive semidefinite, since it is diagonally dominant with nonnegative diagonal, i.e. $X \succeq 0$ by Gersgorin's disk theorem. Therefore

$$\mathcal{L}(X) \;=\; \begin{pmatrix} 0 & 2(d+\bar{D}e)^T \\ 2(d+\bar{D}e) & \mathcal{D}_e(X) - 2X \end{pmatrix}$$

is EDM by Theorem 3.2 Part 4. ■

Corollary 4.5 *The Perron root and vector of $\mathcal{LL}^*$ are*

$$\lambda = (2n-1) + \sqrt{(2n-3)^2 + 8(n-2)} > 0$$

and

$$D = \begin{pmatrix} 0 & e^T \\ e & \alpha(E-I) \end{pmatrix}, \quad \alpha > 0,$$

where

$$\alpha = \frac{1}{2(n-2)} \left\{ (2n-3) + \sqrt{(2n-3)^2 + 8(n-2)} \right\},$$

D is nonsingular and $\mathcal{L}^\dagger(D) \succ 0$.

Proof. Assume that the eigenvector in $\mathcal{E}$ is of the form

$$D = \begin{pmatrix} 0 & e^T \\ e & \alpha(E-I) \end{pmatrix}, \quad \alpha > 0.$$

Then

$$\begin{array}{rcl} X : & = & \mathcal{L}^*(D) \\ & = & 2(\operatorname{Diag}(e) + \operatorname{Diag}(\alpha(E-I)e) - \alpha(E-I)) \\ & = & 2(I + (n-2)\alpha I - \alpha E + \alpha I) \\ & = & 2\left([(n-1)\alpha + 1]I - \alpha E\right). \end{array}$$

Let $\beta := 2[(n-2)\alpha + 1]$. Then

$$\begin{array}{rcl} \mathcal{L}(X) & = & \begin{pmatrix} 0 & \operatorname{diag}(X)^T \\ \operatorname{diag}(X) & \mathcal{D}_e(X) - 2X \end{pmatrix} \\ & = & \begin{pmatrix} 0 & \beta e^T \\ \beta e & \beta(ee^T + ee^T) - 4\left([(n-1)\alpha+1]I - \alpha E\right) \end{pmatrix} \\ & = & \begin{pmatrix} 0 & \beta e^T \\ \beta e & 2\beta E - 4\left([(n-1)\alpha+1]I - \alpha E\right) \end{pmatrix} \\ & = & \begin{pmatrix} 0 & \beta e^T \\ \beta e & 4[\alpha(n-1)+1]E - 4\left([(n-1)\alpha+1]I\right) \end{pmatrix} \\ & = & \begin{pmatrix} 0 & 2[(n-2)\alpha+1]e^T \\ 2[(n-2)\alpha+1]e & 4[(n-1)\alpha+1](E-I) \end{pmatrix}. \end{array}$$

We now get the eigenvector-eigenvalue equation

$$\lambda \begin{pmatrix} 0 & e^T \\ e & \alpha(E-I) \end{pmatrix} = \begin{pmatrix} 0 & 2[(n-2)\alpha+1]e^T \\ 2[(n-2)\alpha+1]e & 4[(n-1)\alpha+1](E-I) \end{pmatrix}.$$

This yields the two equations

$$\lambda = 2[(n-2)\alpha + 1], \quad \lambda\alpha = 4[(n-1)\alpha + 1].$$

We can eliminate λ and get $2[(n-2)\alpha+1]\alpha = 4[(n-1)\alpha+1]$. The quadratic equation is

$$\begin{array}{rcl} 0 & = & 2(n-2)\alpha^2 + 2\alpha - 4(n-1)\alpha - 4 \\ & = & (n-2)\alpha^2 + \alpha - 2(n-1)\alpha - 2 \\ & = & (n-2)\alpha^2 + (3-2n)\alpha - 2, \end{array}$$

i.e.

$$\alpha = \frac{1}{2(n-2)}\left\{(2n-3) \pm \sqrt{(2n-3)^2 + 8(n-2)}\right\}$$

Therefore

$$\begin{array}{rcl} \lambda & = & 2 + 2(n-2)\left[\frac{1}{2(n-2)}\left\{(2n-3) \pm \sqrt{(2n-3)^2 + 8(n-2)}\right\}\right] \\ & = & (2n-1) \pm \sqrt{(2n-3)^2 + 8(n-2)}. \end{array}$$

That D is nonsingular follows by looking at the matrix vector product $D\begin{pmatrix} t \\ x \end{pmatrix} = 0$, i.e. this implies that $x^T e = 0$ which further implies x is a multiple of e, which forces $x = 0$. This also shows that $\mathcal{L}^\dagger(D) \succ 0$. ■

5 Duality and Optimality Conditions

5.1 Duality and Optimality Conditions for NEDM

We work on the equivalent problem to (1.3) stated above, i.e.

$$\mu^* := \min \frac{1}{2}\|H \circ (A - \mathcal{L}(X))\|_F^2 \quad \text{subject to} \quad X \succeq 0,\, X \in \mathcal{S}^{n-1}, \qquad (5.1)$$

where $\mathcal{L} : \mathcal{S}^{n-1} \to \mathcal{S}^n$ is given in the simple characterization in Item 4 in Theorem 3.2,

$$\mathcal{L}(X) = \begin{pmatrix} 0 & \operatorname{diag}(X)^T \\ \operatorname{diag}(X) & \mathcal{D}_e(X) - 2X \end{pmatrix}.$$

This is again a quadratic cone minimization problem. However, the EDM cone constraint is replaced by the SDP constraint and the dimension is reduced by 1.

The Lagrangian dual is

$$\mu^* \geq \nu^* := \max_{\Lambda \succeq 0} \min_X \frac{1}{2}\|H \circ (A - \mathcal{L}(X))\|_F^2 - \operatorname{trace} \Lambda X. \qquad (5.2)$$

We change the dual to the Wolfe dual by noting that the inner problem is a convex unconstrained minimization. If we let

$$C := \mathcal{L}^*(H^{(2)} \circ A),$$

then the unconstrained inner minimization in (5.2) has optimal solution characterized by the stationary conditions:

$$\begin{array}{rcl} 0 & = & -\left[(H\circ)\mathcal{L}\right]^* \{H \circ (A - \mathcal{L}(X))\} - \Lambda \\ & = & -\mathcal{L}^* \left\{H^{(2)} \circ (A - \mathcal{L}(X))\right\} - \Lambda \\ & = & \mathcal{L}^* \left\{H^{(2)} \circ (\mathcal{L}(X))\right\} - C - \Lambda, \end{array} \tag{5.3}$$

where we define $H^{(2)} := H \circ H$.

Thus we obtain the equivalent dual problem:

$$\begin{array}{cl} \max & \frac{1}{2}\|H \circ (A - \mathcal{L}(X))\|_F^2 - \operatorname{trace} \Lambda X \\ \text{subject to} & \Lambda = \mathcal{L}^* \left\{H^{(2)} \circ (\mathcal{L}(X))\right\} - C \\ & \Lambda \succeq 0. \end{array} \tag{5.4}$$

Without loss of generality, we can assume that H has no zero rows (or columns). We get the following optimality conditions.

Theorem 5.1 *Assume that H has no zero rows (or columns). The primal-dual optimal values satisfy $\mu^* = \nu^*$ and the primal-dual pair X, Λ are optimal for* (5.1) *and* (5.4) *if and only if*

$$\begin{array}{rcll} X & := & \operatorname{sMat}(x) \succeq 0 & \textit{(primal feasibility)} \\ \Lambda & := & \mathcal{L}^* \left\{H^{(2)} \circ (\mathcal{L}(X))\right\} - C, \quad \Lambda \succeq 0 & \textit{(dual feasibility)} \\ \Lambda X & := & 0 & \textit{(complementary slackness)} \end{array}$$

Proof. That Slater's condition holds for the primal is trivial. That it holds for the dual follows from Lemma 4.2 above. ■

For our p-d i-e-p algorithm we use

$$\Lambda X = \mu I \quad \text{perturbed complementary slackness.}$$

We can substitute the primal and dual feasibility equations into the perturbed complementary slackness equation and obtain a *single bilinear equa-*

tion in $x = \operatorname{svec}(X)$ that characterizes optimality for the perturbed log-barrier problem:

$$\begin{aligned} F_\mu(x) : \mathbb{R}^{t(n-1)} &\rightarrow \mathbb{R}^{(n-1)^2}. \\ F_\mu(x) : \;=\; & \operatorname{vec}\left[\mathcal{L}^*\left\{H^{(2)} \circ (\mathcal{L}(\operatorname{sMat}(x)))\right\} - C\right]\; \operatorname{sMat}(x) - \mu \operatorname{vec} I = 0. \end{aligned} \tag{5.5}$$

Viewing (5.5) as an overdetermined system of nonlinear equations, we solve it using an inexact Gauss-Newton method. Linearizing, we obtain a linear system for the search direction Δx

$$\begin{aligned} F_\mu(x+\Delta x) &= F_\mu(x) + F'_\mu(\Delta x) + o(\|\Delta x\|) \\ &= F_\mu(x) + \operatorname{vec}\{\Lambda \operatorname{sMat}(\Delta x)\} \;+ \\ &\qquad +\operatorname{vec}\{[\mathcal{L}^*(H^{(2)} \circ (\mathcal{L}(\operatorname{sMat}(\Delta x)))]\operatorname{sMat}(x)\} \\ &\qquad\qquad +o(\|\Delta x\|). \end{aligned} \tag{5.6}$$

Therefore,

$$F'_\mu(x)(\Delta x) = \operatorname{vec}\left\{\Lambda\left[\operatorname{sMat}(\Delta x)\right] + \left[\mathcal{L}^*\left(H^{(2)} \circ (\mathcal{L}(\operatorname{sMat}(\Delta x))\right)\right] X\right\}. \tag{5.7}$$

Note that

$$\mathcal{L}^*(D) \;=\; 2(\operatorname{Diag}(d) + \operatorname{Diag}(\bar{D}e) - \bar{D}).$$

In addition, we note that the operator $\mathcal{L}^*\left(H^{(2)} \circ \mathcal{L}(\cdot)\right)$ is self-adjoint; and the adjoint of $\Lambda\cdot$ is $\frac{1}{2}\left(\Lambda \cdot + \cdot^T \Lambda\right)$; while the adjoint of $\cdot X$ is $\frac{1}{2}\left(X \cdot^T + \cdot X\right)$. Therefore the adjoint $(F'_\mu)^*$ (see (5.7)) for $w \in \mathbb{R}^{(n-1)^2}$ is

$$\begin{aligned} (F'_\mu)^*(w) \;=\; & \tfrac{1}{2}\operatorname{svec}\left\{\left[\Lambda \operatorname{Mat}(w) + \operatorname{Mat}^T(w)\Lambda\right]\right. \\ & \quad \left. + \left[\mathcal{L}^*\left(H^{(2)} \circ (\mathcal{L}\left(X\operatorname{Mat}(w)^T + \operatorname{Mat}(w)X\right)\right)\right]\right\}. \end{aligned} \tag{5.8}$$

To simplify notation, we define the self-adjoint operator

$$\mathcal{W}(\cdot) := \mathcal{L}^*\left(H^{(2)} \circ \mathcal{L}(\operatorname{sMat}\cdot)\right), \tag{5.9}$$

and the transformations with their adjoints, by abuse of notation,

$$\mathcal{X}(s) = \operatorname{vec}\left(\operatorname{sMat}(s)X\right); \quad \mathcal{X}^*(w) = \frac{1}{2}\operatorname{svec}(X\operatorname{Mat}(w)^T + \operatorname{Mat}(w)X); \tag{5.10}$$

with adjoints

$$\Lambda(s) = \operatorname{vec}\left(\Lambda\operatorname{sMat}(s)\right); \quad \Lambda^*(w) = \frac{1}{2}\operatorname{svec}(\Lambda\operatorname{Mat}(w) + \operatorname{Mat}(w)^T\Lambda). \tag{5.11}$$

Therefore,

$$F'_\mu(s) = \Lambda(s) + \mathcal{X}\left(\mathcal{W}(s)\right), \quad s \in \mathbb{R}^{t(n-1)}; \tag{5.12}$$

$$(F'_\mu)^*(w) = \Lambda^*(w) + \mathcal{W}^*\left(\mathcal{X}^*(w)\right), \quad s \in \mathbb{R}^{(n-1)^2}. \tag{5.13}$$

5.2 Sparse Implementation

For an efficient implementation of a conjugate gradient method we must be able to evaluate $F'_\mu(x)(\Delta x)$ and $F'_\mu(x)^*(w)$ efficiently. We discuss the evaluation of $\mathcal{W}(y)$. We define

$$Y := \mathrm{sMat}\,(y); \quad y_d := \mathrm{diag}\,(Y); \quad y_u := \mathrm{usvec}\,(Y); \quad y_{\mathcal{D}_e u} = \mathrm{usvec}\,(\mathcal{D}_e(y_d)); \tag{5.14}$$

and, with the partition

$$H^{(2)} =: \begin{pmatrix} 0 & h^T \\ h & \bar{H} \end{pmatrix}; \quad \bar{H}_u := \mathrm{usvec}\,(\bar{H}); \tag{5.15}$$

and

$$y_{hd} := h \circ y_d; \quad y_{hu} := \bar{H}_u \circ y_u; \quad y_{H\mathcal{D}_e u} := \bar{H}_u \circ y_{\mathcal{D}_e u}.$$

Then

$$\begin{array}{rcl} H^{(2)} \circ \mathcal{L}(Y) & = & \begin{pmatrix} 0 & h^T \circ y_d^T \\ h \circ y_d & \bar{H} \circ \mathcal{D}_e(y_d) - \bar{H} \circ 2\mathrm{usMat}\,(y_u) \end{pmatrix} \\ & = & \begin{pmatrix} 0 & y_{hd}^T \\ y_{hd} & \bar{H} \circ \mathcal{D}_e(y_d) - 2\mathrm{usMat}\,(y_{hu}) \end{pmatrix} \\ & = & \begin{pmatrix} 0 & y_{hd}^T \\ y_{hd} & \mathrm{usMat}\,(y_{H\mathcal{D}_e u} - 2y_{hu}) \end{pmatrix} \\ & = & \begin{pmatrix} 0 & y_{hd}^T \\ y_{hd} & \bar{D} \end{pmatrix} \end{array}$$

and, by abuse of notation,

$$\begin{array}{rcl} \mathcal{W}(y) & = & \mathcal{L}^*(H^{(2)} \circ \mathcal{L}(\mathrm{sMat}\,(y))) \\ & = & 2\left((y_{hd} + \bar{D}e) - (y_{H\mathcal{D}_e u} - 2y_{hu})\right), \end{array} \tag{5.16}$$

i.e. we assume that these vectors are extended to vectors in $\mathbb{R}^{t(n-1)}$ by adding zeros appropriately. Therefore, to evaluate $\mathcal{W}(y)$, where y has the relationships in (5.14), we need only evaluate $y_{H\mathcal{D}_e u}, y_{hu}, y_{hd}, \bar{D}e$.

We note that the sparsity pattern of $\mathcal{W}(y)$ is essentially the same as that of H (except for the diagonal and the first row) and that the evaluation is a sparse operation on the sparse vector y_h and the $(n-1)$-vector y_d. Therefore, all the operations in the evaluations of both $F'_\mu(x)(\Delta x)$ and $F'_\mu(x)^*(w)$ are sparse operations, though the end result can be (and usually is) a dense vector.

Finally, we remark that the i,j-element

$$\left(\bar{H} \circ \mathcal{D}_e(v)\right)_{ij} = \begin{cases} v_i + v_j & \text{if } \bar{H}_{ij} \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

i.e. this is a sparse evaluation if $\bar{H}$ is sparse.

5.3 Duality and Optimality Conditions for EDMC

The objective function for EDMC is strictly convex and coercive. This implies that $\{X \succeq 0 : \mathcal{A}(X) = b,\ f(X) \leq \alpha\}$ is a convex, compact level set for each $\alpha \in \mathbb{R}_+$. Therefore, the primal problem EDMC is attained and there is no duality gap. (This can also be seen from Slater's condition, which holds for the dual feasibility equation in (5.19), with e.g. $\Lambda = \alpha I$, and $\alpha > 0$ sufficiently large.) In particular, Corollary 5.3 below implies that strong duality holds, i.e. the dual is attained as well.

The Lagrangian dual is

$$\mu^* = \nu^* := \max_{\Lambda \succeq 0, y \in \mathbb{R}^{|S|}} \min_X \frac{1}{2}\|X\|_F^2 + y^T(b - \mathcal{A}(X)) - \operatorname{trace} \Lambda X. \tag{5.17}$$

We get the following characterization of optimality.

Theorem 5.2 *Suppose that the feasible set of* EDMC *is not the empty set. Then the optimal solution of* EDMC *is* $D = \mathcal{L}([\mathcal{A}^*(y)]_+)$*, where* y *is the unique solution of the single equation*

$$\mathcal{A}([\mathcal{A}^*(y)]_+) = b, \tag{5.18}$$

and B_+ *denotes the projection of the symmetric matrix* $B \in \mathcal{S}^{n-1}$ *onto the cone* $\mathcal{P}_{n-1}$.

Proof. The optimality conditions obtained after differentiation are

$$\begin{array}{rlc} X = \mathcal{A}^*(y) + \Lambda \succeq 0, & \Lambda \succeq 0, & \text{dual feasibility} \\ \mathcal{A}(X) = b & & \text{primal feasibility} \\ \Lambda X = 0 & & \text{complementary slackness} \end{array} \tag{5.19}$$

This means that $\mathcal{A}^*(y) = X - \Lambda$, where both $X \succeq 0, \Lambda \succeq 0$, and $\Lambda X = 0$. Therefore the three symmetric matrices $W = \mathcal{A}^*(y), X, \Lambda$ are mutually diagonalizable. We write $X = PD_XP^T$, $\Lambda = PD_\Lambda P^T$, i.e. we conclude that $W = \mathcal{A}^*(y) = P\left(D_X - D_\Lambda\right)P^T$, $D_XD_\Lambda = 0$. Therefore $[\mathcal{A}^*(y)]_+ = PD_XP^T = X$. ∎

The following corollary provides an explicit solution for EDMC under the assumption that $y \geq 0$. This shows that a large class of completion problems can be efficiently solved.

Corollary 5.3 *The linear transformation $\mathcal{A}$ is onto and $\mathcal{A}\mathcal{A}^*$ is nonsingular. Suppose that $y = (A\mathcal{A}^*)^{-1}b \in \mathbb{R}^m_+$. Then*

$$D = \mathcal{L}\left(\mathcal{A}^*(y)\right) \tag{5.20}$$

is the unique solution of EDMC *.*

Proof. That $\mathcal{A}$ is onto follows from the definitions.
The proof continues as in the proof of Lemma 4.4. ∎

5.3.1 Matrix Representation A of $\mathcal{A}\mathcal{A}^*$

Under certain conditions, we can find an explicit solution from a single linear equation in Corollary 5.3. Therefore we need a matrix representation of the operator $\mathcal{A}\mathcal{A}^*$. We can do this by columns.

For each unit vector e_k, we get the k-th column of the matrix representation A from applying $\mathcal{A}\mathcal{A}^*(e_k)$. We identify $k \cong i_kj_k, i_k < j_k$, i.e. with the k-th element in the set S.

1. **Case 1**, $k \cong 1j_k$
   $D = \mathcal{I}^*(e_k) = \frac{1}{\sqrt{2}}E_k$ is a symmetric matrix with zero diagonal and with exactly two nonzero elements equal to $1/\sqrt{2}$ corresponding to the $k \cong 1j_k$ position, i.e. in the first row and column. The usual partition of D yields $d = e_{j_k-1}, \bar{D} = 0$. Therefore, $\mathcal{A}^*(e_k) = \mathcal{L}^*\mathcal{I}^*(e_k) = \sqrt{2}\mathrm{Diag}\,(e_{j_k-1})$. This implies that

$$\mathcal{L}\mathcal{A}^*(e_k) = \sqrt{2}\begin{pmatrix} 0 & e_{j_k-1}^T \\ e_{j_k-1} & e_{j_k-1}e^T + ee_{j_k-1}^T - 2\mathrm{Diag}\,(e_{j_k-1}) \end{pmatrix}.$$

The k-th column is now obtained by applying $\mathcal{I}$, i.e. the l-th element of the k-th column $A_{:k} \in \mathbb{R}^m$ is

$$A_{lk} = \begin{cases} 2 & \text{if } i_l = j_k \text{ or } j_l = j_k \\ 0 & \text{otherwise} \end{cases}$$

Note that A is symmetric, $A_{lk} = A_{kl}$, and the diagonal elements $A_{kk} = 2$, $\forall k$ in Case 1.

2. **Case 2**, $k \cong i_k j_k,\ i_k > 1$
(Note that E_k is in $\mathcal{S}^n$ or $\mathcal{S}^{n-1}$ depending on the context.) $D = \mathcal{I}^*(e_k) = \frac{1}{\sqrt{2}} E_k$ is a symmetric matrix with zero diagonal and with exactly two nonzero elements equal to $1/\sqrt{2}$ corresponding to the $k \cong i_k j_k$ position. The usual partition of D yields $d = 0, \bar{D} = \frac{1}{\sqrt{2}} E_{k-1}$, $d_e := d + \bar{D} = \sqrt{2}\,(e_{i_k-1} + e_{j_k-1})$, where $k - 1 \cong (i_k - 1, j_k - 1)$. Therefore, $X = \mathcal{A}^*(e_k) = \mathcal{L}^*\mathcal{I}^*(e_k) = \sqrt{2}\text{Diag}\,(e_{i_k-1} + e_{j_k-1}) - E_{i_k-1, j_k-1}$. Now

$$\mathcal{D}_e(X) = \sqrt{2}\left[(e_{i_k-1} + e_{j_k-1})\, e^T + e\,(e_{i_k-1} + e_{j_k-1})^T\right],$$

i.e. two rows and columns are all ones but with two at the diagonal intersection points. This implies that

$$\begin{aligned} \mathcal{L}\mathcal{A}^*(e_k) &= \mathcal{L}\mathcal{L}^*(D) \\ &= 2\begin{pmatrix} 0 & d_e^T \\ d_e & \text{offDiag}\,(d_e e^T + e d_e^T) + 2\bar{D} \end{pmatrix} \\ &= 2\begin{pmatrix} 0 & d_e^T \\ d_e & \text{offDiag}\,(d_e e^T + e d_e^T) + \sqrt{2}E_{k-1} \end{pmatrix} \end{aligned}$$

The k-th column is now obtained by applying $\mathcal{I}$, i.e. the l-th element of the k-th column is

$$A_{lk} = \begin{cases} 2 & \text{if } i_l = i_k \text{ or } i_l = j_k \\ 2 & \text{if } j_l = i_k \text{ or } j_l = j_k \\ 8 & \text{if } i_l = i_k \text{ and } j_l = j_k \\ 0 & \text{otherwise} \end{cases}$$

Note, as in Case 1, that A is symmetric, $A_{lk} = A_{kl}$, and the diagonal elements $A_{kk} = 8$, $\forall k$.

Remark 5.4 *Suppose we restrict ourselves to Case 2. (We can add a dummy vertex, or equivalently a column/row of zeros to the matrix* $D = \mathcal{A}^*(b)$*.) Finding the matrix representation* A *can be explained using the graph,* $G = (V_G, E_G)$*, of the matrix* $D = \mathcal{A}^*(b)$*. The nodes* V_G *correspond to the columns of* D*. The edge* $e_{ij} \in E_G$ *if* $D_{ij} \neq 0$*. Therefore,* $m = |E_G| = |S|$*. The* k*-column of* A *corresponds to* b_k *which corresponds to the* $k \cong i_k j_k$ *element* $D_{i_k j_k} \neq 0$*. The* A_{lk} *element is nonzero for each arc emanating from the two nodes* $i_k, j_k \in V_G$*, i.e.*

$$A_{lk} = \begin{cases} 2 & \textit{if } (i_k, j_l) \in E_G \\ 2 & \textit{if } (i_l, j_k) \in E_G \\ 8 & \textit{if } (i_l, j_l) \in E_G \\ 0 & \quad \textit{otherwise} \end{cases}$$

Proposition 5.5 *Let* order (j) *denote the* order *of the node* j *in* G*. Suppose that for each pair of nodes* j_1, j_2 *in* V

$$(j_1, j_2) \in E_G \Rightarrow \operatorname{order}(j_1) + \operatorname{order}(j_2)(<) \leq 6.$$

Then the matrix representation A *is (strictly) diagonally dominant.*

Proof. The $k \cong j_1 j_2$-th row of the matrix representation A consists of a 2 in each position corresponding to an arc leaving either node j_1, j_2. The result follows since the sum of the orders counts the arc joining the two nodes twice and this latter position is a diagonal element 8. ■

Remark 5.6 *For a given* $y > 0$ *with rational numbers, we conclude that* $b = Ay > 0$ *and there is a neighbourhood around* b *where the completion exists and can be found by solving the simple positive definite system defined by* A*. Therefore, for the given data* b *found this way,* EDMC *is a polytime problem, i.e. one can decide if a completion exists and find the exact completion in polynomial time.*

Lemma 5.7 *Suppose that* $k(x), x \in \mathbb{R}^n$ *is a norm with (compact, convex) unit ball* $B = \{x \in \mathbb{R}^n : k(x) \leq 1\}$*. And* $k^*(y) = \max_{x \in B} x^T y$ *is the* dual norm. *Let* A *be the matrix representation of* $\mathcal{A}\mathcal{A}^*$*,* $y > 0$ *be given, and* $b = Ay$*. Define*

$$r := \min_i \left\{ \frac{y_i}{k^*(e_i^T A^{-1})} \right\}.$$

Then $b > 0$ and

$$k(\Delta b) \leq r \Rightarrow A^{-1}(b + \Delta b) \geq 0.$$

In particular, if $Av = \lambda v, \lambda > 0, v > 0$ yields the Perron root and vector (eigenvalue-eigenvector pair, see Lemma 4.4) for A, then we can choose $y = \frac{1}{\lambda}v$.

Proof. We know that $Ay = b > 0$. (Since $\mathcal{A}^*(y) \succeq 0$.) We want to guarantee that $A^{-1}(b+\Delta b) \geq 0$ or equivalently that $-A^{-1}\Delta b \leq y$, Therefore, we want to find the maximum r such that

$$\max_{k(\Delta b) \leq 1} r(-e_i^T A^{-1})\Delta b \leq y_i, \quad \forall i.$$

The latter is equivalent to finding the largest r for which

$$rk^*(e_i^T A^{-1}) \leq y_i, \quad \forall i.$$

■

We now look at the converse perturbation, i.e. given $b > 0$ but $A^{-1}b$ is not nonnegative. We look for results on perturbations $b + \Delta b$ so that $A^{-1}(b + \Delta b) \geq 0$.

Lemma 5.8 *Suppose that $b \in \mathbb{R}^m_+$ and $\mathcal{A}^\dagger(b) \succeq 0$. Then $\mathcal{L}(\mathcal{A}\mathcal{A}^*)^{-1}(b)$ is the EDM completion of b.*

Proof. Note that $\mathcal{A}\mathcal{A}^*(\mathcal{A}\mathcal{A}^*)^{-1}(b) = b$ implies that $\mathcal{A}^\dagger = \mathcal{A}^*(\mathcal{A}\mathcal{A}^*)^{-1}$. ■

We would like to study the cone

$$\mathcal{E}_b := \{b \geq 0 : \mathcal{A}^\dagger(b) \succeq 0\},$$

since this is exactly the set of vectors b for which the EDMC problem can be solved quickly using $(\mathcal{A}\mathcal{A}^*)^{-1}$. We note that

$$\begin{array}{rcl} b \in \mathcal{E}_b & \Leftrightarrow & b \geq 0 \text{ and trace}\, S\mathcal{A}^\dagger(b) \geq 0, \ \forall S \in \mathcal{P} \\ & \Leftrightarrow & b \geq 0 \text{ and } \sum_{k=1}^m b_k \text{trace}\,(SA_k) \geq 0, \ \forall S \in \mathcal{P} \\ & \Leftrightarrow & b \geq 0 \text{ and } b^T a_S \geq 0, \ \forall S \in \mathcal{P}, \end{array}$$

where $a_S = (\,\text{trace}\,(SA_k)\,) \in \mathbb{R}^m$. (The matrices A_k are implicity defined from the above.)

6 Primal-Dual Interior-Exterior-Point Algorithm for NEDM

We can use equation (5.5) to develop a primal-dual interior-exterior-point (p-d i-e-p) algorithm, i.e., we linearize to find the search direction using a linear least squares problem. We can assume that we start strictly feasible, see Lemma 4.2.

6.1 Framework

The p-d i-e-p framework that we use is different in several ways from the common framework for both linear and semidefinite programming, see e.g. [28, 21]. We have eliminated, in advance, the primal and dual linear feasibility equations. We work with an overdetermined nonlinear system rather than a square symmetrized system; thus we use an (inexact) Gauss-Newton approach [17]. We include a centering parameter σ_k (instead of the customary predictor-corrector approach). We enforce positive semidefiniteness rather than definiteness in the steplengths. In addition, once we are *close enough* to the optimum we set the centering parameter σ to zero (crossover step) and we no longer enforce interiority, i.e. we allow negative eigenvalues. This allows for (fast) asymptotic quadratic convergence.

At each iteration, we have available the iterate x and we find a new iterate by taking a step in the (inexact) Gauss-Newton search direction Δx. Up until the crossover, we ensure that the new iterate $x + \alpha\Delta x$ results in both X, Λ sufficiently positive definite; then, we take $\alpha = 1$ after the crossover. By our construction, the iterates maintain exact primal and dual feasibility.

We let $\mathcal{F}^0$ denote the set of strictly feasible primal-dual points; F' denotes the derivative of the function of optimality conditions.

Algorithm 6.1 *(Primal-Dual Gauss-Newton via PCG for NEDM)*
Input: *Objective: pre-distance and weight matrices* $A, H \in \mathcal{S}^n$
Tolerances: δ_1 *(gap),* δ_2 *(crossover)*
Initialization:

$$X^0, \Lambda^0 := \mathcal{L}^*\left\{H^{(2)} \circ (\mathcal{L}(X^0))\right\} - C \ \succ \ 0$$

gap $= \operatorname{trace} \Lambda^0 X^0$; $\mu = \text{gap}/(n-1)$; $\sigma = 1$; objval $= \dfrac{1}{2}\|H \circ (A - \mathcal{L}(X))\|_F^2$.

```
while $\min\{\frac{\text{gap}}{\text{objval}+1}, \text{objval}\} > \delta_1$
    if  $\min\{\frac{\text{gap}}{\text{objval}+1}, \text{objval}\} < \delta_2$  then
        $\sigma = 0$
    else
        update $\sigma$
    end if
    Find diagonal preconditioner $p$ see Section 6.2.
    Find LSS of $F'_{\sigma\mu}(x)\,(\Delta x) = -F_{\sigma\mu}(x)$  (using LSQR)
    update $X := X + \alpha\ \mathrm{sMat}\,(\Delta x), \alpha > 0,\ \ \Lambda := \mathcal{L}^*\left\{H^{(2)} \circ (\mathcal{L}(X))\right\} - C,$
            $(X, \Lambda\ \succ\ 0)$
            $\text{gap} = \operatorname{trace} \Lambda^0 X^0;\quad \mu = \text{gap}/(n-1);$
            $\text{objval} = \frac{1}{2}\|H \circ (A - \mathcal{L}(X))\|_F^2$
endwhile
```

6.2 Preconditioning

Preconditioning is essential for efficient solution of the least squares problem (5.7). We find an operator P and find Δx,the least squares solution of

$$(\Lambda + \mathcal{X}\mathcal{W})\, P^{-1}(\widehat{\Delta x}) = -F_\mu(x),$$

where

$$\widehat{\Delta x} = P(\Delta x).$$

The inverse is not found explicitly. The operator P has simple structure so that the linear system can be solved efficiently.

6.2.1 Diagonal Preconditioning

Optimal diagonal scaling has been studied in, e.g., [14, Sect. 10.5], and [11, Prop. 2.1(v)]. In the latter reference, it was shown that for a full rank matrix $A \in \mathcal{M}^{m\times n}$, $m \geq n$, and using the condition number $\omega(K) := n^{-1}\operatorname{trace}(K)/\det(K)^{1/n}$, the optimal scaling, i.e. the solution of the optimization problem

$$\min \omega((AD)^T(AD)) \quad \text{subject to } D \text{ positive diagonal matrix,} \tag{6.1}$$

is given by $d_{ii} = 1/\|A_{:i}\|_2, i = 1, \ldots, n$.

Therefore, the operator P is diagonal and is evaluated using the columns of the matrix representation of the operator

$$F'_\mu(\cdot) = \Lambda(\cdot) + \mathcal{X}\left(\mathcal{W}(\cdot)\right).$$

The columns are ordered using $k = 1, 2, \ldots$ where k represents (i, j), $1 \leq i \leq j \leq n-1$ for the upper triangular part of the symmetric matrix ΔX, taken columnwise. As above, we let $X = \mathrm{sMat}\,(x)$ and $E_{ij} = \frac{1}{\sqrt{2}}(e_i e_j^T + e_j e_i^T)$ if $i < j$ while $E_{ii} = e_i e_i^T$.

For the first operator in F'_μ, we get (recall, by abuse of notation, $\Lambda(s) = \Lambda \mathrm{sMat}\,(s)$):

$$\begin{array}{rcl}
\Lambda(e_k) & = & \Lambda \mathrm{sMat}\,(e_k) = \Lambda\ E_{ij} \\
 & = & \left\{ \begin{array}{cl} \frac{1}{\sqrt{2}} \Lambda \left(e_i e_j^T + e_j e_i^T \right), & \text{if } i < j \\ \Lambda \left(e_i e_i^T \right), & \text{if } i = j \end{array} \right. \\
 & = & \left\{ \begin{array}{cl} \frac{1}{\sqrt{2}} \left(\Lambda_{:i} e_j^T + \Lambda_{:j} e_i^T \right), & \text{if } i < j \\ \left(\Lambda_{:i} e_i^T \right), & \text{if } i = j. \end{array} \right.
\end{array}$$

By abuse of notation we use operators such as $\mathcal{L}$ on matrices. For the second operator we get

$$\begin{array}{rcl}
\mathcal{L}(E_{ij}) & = & \left\{ \begin{array}{ll} \begin{pmatrix} 0 & \mathrm{diag}\,(E_{ij})^T \\ \mathrm{diag}\,(E_{ij}) & \mathcal{D}_e(E_{ij}) - 2\sqrt{2} E_{ij} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & -2\sqrt{2} E_{ij} \end{pmatrix}, & \text{if } i < j \\ \begin{pmatrix} 0 & \mathrm{diag}\,(E_{ii})^T \\ \mathrm{diag}\,(E_{ii}) & \mathcal{D}_e(E_{ii}) - 2 E_{ii} \end{pmatrix} = \begin{pmatrix} 0 & e_i^T \\ e_i & \mathcal{D}_e(E_{ii}) - 2 E_{ii} \end{pmatrix}, & \text{if } i = j. \end{array} \right. \\
 & = & \left\{ \begin{array}{ll} \begin{pmatrix} 0 & 0 \\ 0 & -2\sqrt{2} E_{ij} \end{pmatrix}, & \text{if } i < j \\ \begin{pmatrix} 0 & e_i^T \\ e_i & e_i (e - e_i)^T + (e - e_i) e_i^T \end{pmatrix}, & \text{if } i = j. \end{array} \right.
\end{array}$$

Therefore, with the partition in (5.15), we define

$$\bar{H}_{ij}^E := \bar{H} \circ E_{ij}, \quad h_i^e := h \circ e_i,$$

and get

$$H^{(2)} \circ \mathcal{L}(E_{ij}) = \begin{cases} \begin{pmatrix} 0 & 0 \\ 0 & -2\sqrt{2}\bar{H}^E_{ij} \end{pmatrix}, & \text{if } i < j \\ \begin{pmatrix} 0 & (h^e_i)^T \\ h^e_i & \bar{H} \circ \mathcal{D}_e(E_{ii}) \end{pmatrix}, & \text{if } i = j. \end{cases}$$

$$= \begin{cases} \begin{pmatrix} 0 & 0 \\ 0 & -2\sqrt{2}\bar{H}^E_{ij} \end{pmatrix}, & \text{if } i < j \\ \begin{pmatrix} 0 & (h^e_i)^T \\ h^e_i & \bar{H}_{i:}e_i^T + e_i\bar{H}_{:i} \end{pmatrix}, & \text{if } i = j. \end{cases}$$

Finally, with $e_k = E_{ij}$,

$$\begin{aligned} \mathcal{W}(e_k) &= \begin{cases} \frac{1}{\sqrt{2}}2\left[\operatorname{Diag}(-2\sqrt{2}\bar{H}^E_{ij}e) - (-2\sqrt{2}\bar{H}^E_{ij})\right] & \text{if } i < j \\ 2\left[\operatorname{Diag}(h^e_i) + \operatorname{Diag}((\bar{H} \circ \mathcal{D}_e(E_{ii}))e) - (\bar{H} \circ \mathcal{D}_e(E_{ii})\right] & \text{if } i = j. \end{cases} \\ &= \begin{cases} 2\sqrt{2}\bar{H}_{ij}\left(\sqrt{2}E_{ij} - E_{ii} - E_{jj}\right) & \text{if } i < j \\ 2\left[\operatorname{Diag}(h^e_i) + \operatorname{Diag}(\bar{H} \circ \mathcal{D}_e(E_{ii}))e - (\bar{H} \circ \mathcal{D}_e(E_{ii})\right] & \text{if } i = j. \end{cases} \\ &= \begin{cases} 2\sqrt{2}\bar{H}_{ij}\left(\sqrt{2}E_{ij} - E_{ii} - E_{jj}\right) & \text{if } i < j \\ 2\left[\operatorname{Diag}\left(\bar{H}_{:i} + (e^T\bar{H}_{:i} + h_i)e_i\right) - \left(\bar{H}_{i:}e_i^T + e_i\bar{H}_{:i}\right)\right] & \text{if } i = j. \end{cases} \end{aligned} \tag{6.2}$$

Now, for the case $k \cong (ij), i < j$**,** $\mathcal{W}(e_k)$ with at most 4 nonzero elements positioned at the intersection of the i, j rows and columns, i.e.

$$2\sqrt{2}\bar{H}_{ij}\left(\sqrt{2}E_{ij} - E_{ii} - E_{jj}\right) = 2\sqrt{2}\bar{H}_{ij} \begin{array}{c} \left(\begin{array}{ccccc} \ddots & \vdots & \vdots & \vdots & \ddots \\ \dots & -1 & \dots & 1 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \dots & 1 & \dots & -1 & \dots \\ \ddots & \vdots & \vdots & \vdots & \ddots \end{array}\right) \end{array} \begin{array}{l} \\ \dots i \\ \\ \dots j \\ \\ \end{array}$$

i.e. it consists of the four equal elements except for the sign. Therefore,

$$\mathcal{X}\mathcal{W}(e_k) = (X(\mathcal{W}(e_k)))^T = 2\sqrt{2}\bar{H}_{ij}\left(X(\sqrt{2}E_{ij} - E_{ii} - E_{jj})\right)^T,$$

i.e. we need only evaluate it if $\bar{H}_{ij} \neq 0$! And, we get

$$\text{for } i < j: \qquad \mathcal{X}\mathcal{W}(e_k) = \begin{cases} 2\sqrt{2}\bar{H}_{ij}(X_{j:} - X_{i:}) & \text{in row } i \\ 2\sqrt{2}\bar{H}_{ij}(X_{i:} - X_{j:}) & \text{in row } j \end{cases} \tag{6.3}$$

$$\begin{array}{rcl}\|(\Lambda+\mathcal{X}\mathcal{W})(e_k)\|_F^2 & = & \|(\Lambda+\mathcal{X}\mathcal{W})(E_{ij})\|_F^2 \\ & = & \|(\Lambda+\mathcal{X}\mathcal{W})(E_{ij})\|_F^2 \\ & = & \|\Lambda(E_{ij})\|_F^2+\|\mathcal{X}\mathcal{W}(E_{ij})\|_F^2+2\left\langle\Lambda(E_{ij}),\mathcal{X}\mathcal{W}(E_{ij})\right\rangle \\ & = & \|\Lambda(e_k)\|_F^2+\|(\mathcal{W}(e_k))X\|_F^2+2\left\langle\Lambda(E_{ij}),(\mathcal{W}(E_{ij}))X\right\rangle\end{array} \tag{6.4}$$

We need to find

$$\begin{array}{rcl}\left\langle\Lambda(E_{ij}),(\mathcal{W}(E_{ij}))X\right\rangle & = & \operatorname{trace}\Lambda(E_{ij})^T(\mathcal{W}(E_{ij}))X \\ & = & \operatorname{trace}X\Lambda(E_{ij})^T(\mathcal{W}(E_{ij})) \\ & = & \frac{1}{\sqrt{2}}\operatorname{trace}X(\Lambda_{:i}e_j^T+\Lambda(:j)e_i^T)^T4\bar{H}_{ij}\left(E_{ij}-E_{ii}-E_{jj}\right) \\ & = & \frac{1}{\sqrt{2}}4\bar{H}_{ij}\operatorname{trace}\left(X_{:j}\Lambda_{:i}+X_{:i}\Lambda_{:j}\right)\left(E_{ij}-E_{ii}-E_{jj}\right) \\ & = & 2\bar{H}_{ij}\left\{X_{ij}\Lambda_{ji}+X_{ii}\Lambda_{jj}+X_{jj}\Lambda_{ii}+X_{ji}\Lambda_{ij}\right. \\ & & \qquad \left.-X_{ij}\Lambda_{ii}-X_{ii}\Lambda_{ij}-X_{jj}\Lambda_{ji}-X_{ji}\Lambda_{jj}\right\}.\end{array} \tag{6.5}$$

Thus we see that here as well, this need only be evaluated if $\bar{H}_{ij} \neq 0$!

To summarize the evaluation of the diagonal preconditioner for $i < j$, we continue from (6.4),

$$\begin{array}{rcl}\|(\Lambda+\mathcal{X}\mathcal{W})(e_k)\|_F^2 & = & \|\Lambda(e_k)\|_F^2+\|(\mathcal{W}(e_k))X\|_F^2+\left\langle\Lambda(E_{ij}),(\mathcal{W}(E_{ij}))X\right\rangle \\ & = & \frac{1}{2}\left(\|\Lambda_{:i}\|^2+\|\Lambda_{:j}\|^2\right)+ \\ & & \quad 16\bar{H}_{ij}^2\left(\|X_{j:}\|^2+\|X_{i:}\|^2-2X_{i:}^TX_{j:}\right)+ \\ & & \qquad 4\bar{H}_{ij}\left\{X_{ij}\Lambda_{ji}+X_{ii}\Lambda_{jj}+X_{jj}\Lambda_{ii}+X_{ji}\Lambda_{ij}\right. \\ & & \qquad\quad \left.-X_{ij}\Lambda_{ii}-X_{ii}\Lambda_{ij}-X_{jj}\Lambda_{ji}-X_{ji}\Lambda_{jj}\right\}.\end{array} \tag{6.6}$$

Now, for $k \cong (ij), i = j$**,** recall $\mathcal{W}(e_k)$ from (6.2) with $\bar{H}_{:i}$ in the diagonal and minus it in the ith row and column and $e^T\bar{H}_{:i}+h_i$ in the ii position

$$\begin{array}{l} 2\left[\operatorname{Diag}\left(\bar{H}_{:i}+(e^T\bar{H}_{:i}+h_i)e_i\right)-\left(\bar{H}_{i:}e_i^T+e_i\bar{H}_{:i}\right)\right] \\ = \begin{pmatrix} \bar{H}_{1i} & 0 & \dots & -\bar{H}_{1i} & \dots & 0 \\ 0 & \bar{H}_{2i} & \dots & -\bar{H}_{2i} & \dots & 0 \\ \dots & \dots & \ddots & \vdots & \dots & 0 \\ -\bar{H}_{1i} & -\bar{H}_{2i} & \dots & e^T\bar{H}_{:i}+h_i & \dots & -\bar{H}_{n-1i} \\ \dots & \dots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & -\bar{H}_{n-1i} & \dots & -\bar{H}_{n-1i} \end{pmatrix}. \end{array}$$

Therefore,

$$\mathcal{X}\mathcal{W}(e_k) = (X(\mathcal{W}(e_k)))^T = 2\bar{H}_{:i}e^T \circ (X - eX_{i:}),$$

This is true for all the rows of $\mathcal{X}\mathcal{W}$ excepts for the ith row. (Note: that $\bar{H}_{ii} = 0$ then the ith row in the above expression is all zero). The ith row then is given by

$$(X(\mathcal{W}(e_k)))^T_{i:} = 2\{-\bar{H}_{i:}X + (h_i + e^T\bar{H}_{:i})X_{i:}\},$$

$$\begin{array}{rcl} \|(\Lambda + \mathcal{X}\mathcal{W})(e_k)\|_F^2 & = & \|(\Lambda + \mathcal{X}\mathcal{W})(E_{ii})\|_F^2 \\ & = & \|\Lambda(e_k)\|_F^2 + \|(\mathcal{W}(e_k))X\|_F^2 + 2\,\langle \Lambda(e_k), (\mathcal{W}(e_k))X\rangle \end{array} \tag{6.7}$$

Let $\bar{H}^v = \bar{H} + \operatorname{Diag}(v),\ v = \bar{H}e + h$. We need to find

$$\begin{array}{rcl} \langle \Lambda(E_{ii}), (\mathcal{W}(E_{ii}))X\rangle & = & \operatorname{trace}\Lambda(E_{ii})^T(\mathcal{W}(E_{ii}))X \\ & = & 2\operatorname{trace}\Lambda_{:i}\left(\operatorname{Diag}(\bar{H}_{:i} + (e^T\bar{H}\colon i + h_i)e_i) - \bar{H}_{:i}e^T + e\bar{H}_{i:}\right)Xe_i \\ & = & 2\left(\operatorname{diag}(\Lambda(\bar{H}^v \circ X)) - \operatorname{diag}(\Lambda\bar{H}) \circ \operatorname{diag}(X)\right. \\ & & \qquad \left. -\operatorname{diag}(\Lambda) \circ \operatorname{diag}(\bar{H}X)\right)e_i \end{array} \tag{6.8}$$

To summarize the evaluation of the diagonal preconditioner for $i = j$, we continue from (6.7),

$$\begin{array}{rcl} \|(\Lambda + \mathcal{X}\mathcal{W})(e_k)\|_F^2 & = & \|\Lambda(e_k)\|_F^2 + \|(\mathcal{W}(e_k))X\|_F^2 + 2\,\langle \Lambda(e_k), (\mathcal{W}(e_k))X\rangle \\ & = & \|\Lambda_{:i}\|^2 + \|2\bar{H}_{:i}e^T \circ (X - eX_{i:})\|_F^2 + \\ & & \quad \|2\{-\bar{H}_{i:}X + (h_i + e^T\bar{H}_{:i})X_{i:}\}\|^2 + \\ & & \qquad 4\left(\operatorname{diag}(\Lambda(\bar{H}^v \circ X)) - \operatorname{diag}(\Lambda\bar{H}) \circ \operatorname{diag}(X)\right. \\ & & \qquad \left. -\operatorname{diag}(\Lambda) \circ \operatorname{diag}(\bar{H}X)\right)e_i \end{array} \tag{6.9}$$

The diagonal preconditioners are inexpensive to calculate. However, in general, they are not strong enough, e.g. [14].

7 Computational Results

We now present some preliminary computational results. More extensive tests are being done in [3].

7.1 Explicit Completions

7.1.1 Small Problems

We first look at applying Theorem 5.2 without the projection, i.e. we test empirically how often we can find the completion explicitly using

$$D = \mathcal{L}\left(\mathcal{A}^*(\mathcal{A}\mathcal{A}^*)^{-1}b)\right) = \mathcal{L}\left(\mathcal{A}^\dagger b\right).$$

We generate a random sparsity pattern for D but ensure that the graph of the pattern is connected . We then generate a random $X \succeq 0$ and set the original distance matrix to $D = \mathcal{L}(X)$ and the original data b using the generated sparsity pattern.

We start with low dimensional tests since generating the data is time consuming. We see that though most of the randomly generated problems do not yield $y = \mathcal{A}^{\dagger}b$ nonnegative, they still generally yield a distance matrix D, i.e. $\mathcal{A}^*(y) \succeq 0$. The tests were done with n increasing in steps of 10 and the density increasing in steps of .1, i.e. $n = 10 : 10 : 100$ with density $.1 : .1 : .8$. Each element of the following matrix Results (dimension versus density) contains the number of failures in 100 tests.

$$\text{Results} = \begin{pmatrix} 19 & 27 & 29 & 25 & 32 & 27 & 20 & 38 \\ 6 & 20 & 23 & 22 & 27 & 21 & 28 & 28 \\ 8 & 8 & 9 & 9 & 11 & 16 & 17 & 24 \\ 2 & 2 & 6 & 5 & 14 & 17 & 20 & 17 \\ 2 & 0 & 2 & 8 & 7 & 8 & 15 & 12 \\ 1 & 1 & 1 & 1 & 3 & 8 & 15 & 11 \\ 2 & 0 & 3 & 1 & 5 & 7 & 6 & 15 \\ 1 & 0 & 0 & 4 & 2 & 4 & 9 & 9 \\ 1 & 0 & 0 & 1 & 3 & 2 & 5 & 6 \\ 0 & 0 & 0 & 0 & 1 & 6 & 5 & 5 \end{pmatrix}.$$

7.1.2 Large Problems

We solved many large/huge problems of with n several million and m approximately 10^5. The time taken for these problems was small as generating the matrix is quick as is solving a positive definite system.

7.2 Nearest EDM Problem

We applied the primal-dual algorithm outline above in Section 6. We used MATLAB and solved many small problems. The algorithm is robust and an optimal solution (verified) was found in general. However, the algorithm was slow and had difficulties because the Jacobian of the optimality conditions became singular in many sparse instances. This will be studied further in [3].

8 Conclusion

We have presented an implicit solution for EDMC, the EDM completion problem. For many completable problems, this algorithm provides an explicit solution technique that can be applied to huge problems.

We also derived a p-d-i-e-p algorithm for finding the nearest Euclidean Distance matrix to a given matrix, NEDM. The algorithm uses an inexact Gauss-Newton method with preconditioned conjugate gradients. The preliminary numerical tests show promise. However, the cost for finding the search direction using a least squares approach is still too high. New tests are in progress that take advantage of the high singularity of the Jacobian of the optimality conditions.

References

[1] S. AL-HOMIDAN. *Hybrid methods for optimization problems with positive semidefinite matrix constraints*. PhD thesis, University of Dundee, 1993.

[2] S. AL-HOMIDAN and R. FLETCHER. Hybrid methods for finding the nearest Euclidean distance matrix. In *Recent advances in nonsmooth optimization*, pages 1–17. World Sci. Publishing, River Edge, NJ, 1995.

[3] S. AL-HOMIDAN, V. PICCIALLI, and H. WOLKOWICZ. Relations between approximate and exact completion problems for Euclidean distance matrices. Technical Report CORR 2004-???, University of Waterloo, Waterloo, Ontario, 2004.

[4] A. ALFAKIH, A. KHANDANI, and H. WOLKOWICZ. Solving Euclidean distance matrix completion problems via semidefinite programming. *Comput. Optim. Appl.*, 12(1-3):13–30, 1999. Computational optimization—a tribute to Olvi Mangasarian, Part I.

[5] M.F. ANJOS, N. HIGHAM, M. TAKOUDA, and H. WOLKOWICZ. A semidefinite programming approach for the closest correlation matrix problem. Technical Report CORR 2001-in progress, University of Waterloo, Waterloo, Ontario, 2001.

[6] A. BEN-ISRAEL and T.N.E. GREVILLE. *Generalized Inverses: Theory and Applications*. Wiley-Interscience, 1974.

[7] A. BERMAN and R. PLEMMONS. *Nonnegative matrices in the mathematical sciences*. Academic Press, New York, NY, 1979.

[8] F. BOHNENBLUST. Joint positiveness of matrices. Technical report, UCLA, 1948. Manuscript.

[9] R. A. BRUALDI and H. J. RYSER. *Combinatorial Matrix Theory*. Cambridge University Press, New York, 1991.

[10] F. CRITCHLEY. On certain linear mappings between inner-product and squared distance matrices. *Linear Algebra Appl.*, 105:91–107, 1988.

[11] J.E. DENNIS Jr. and H. WOLKOWICZ. Sizing and least-change secant methods. *SIAM J. Numer. Anal.*, 30(5):1291–1314, 1993.

[12] L. ELSNER. Monotonie und randspektrum bei vollstetigen operatoren. *Archive for Rational Mechanics and Analysis*, 36:356–365, 1970.

[13] J. C. GOWER. Properties of Euclidean and non-Euclidean distance matrices. *Linear Algebra Appl.*, 67:81–97, 1985.

[14] A. GREENBAUM. *Iterative methods for solving linear systems*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[15] T.L. HAYDEN, J. WELLS, W-M. LIU, and P. TARAZAGA. The cone of distance matrices. *Linear Algebra Appl.*, 144:153–169, 1991.

[16] C.R. JOHNSON and P. TARAZAGA. Connections between the real positive semidefinite and distance matrix completion problems. *Linear Algebra Appl.*, 223/224:375–391, 1995. Special issue honoring Miroslav Fiedler and Vlastimil Pták.

[17] S. KRUK, M. MURAMATSU, F. RENDL, R.J. VANDERBEI, and H. WOLKOWICZ. The Gauss-Newton direction in linear and semidefinite programming. *Optimization Methods and Software*, 15(1):1–27, 2001.

[18] M. LAURENT. A connection between positive semidefinite and Euclidean distance matrix completion problems. *Linear Algebra Appl.*, 273:9–22, 1998.

[19] M. LAURENT. A tour d'horizon on positive semidefinite and Euclidean distance matrix completion problems. In *Topics in Semidefinite and Interior-Point Methods*, volume 18 of *The Fields Institute for Research in Mathematical Sciences, Communications Series*, pages 51–76, Providence, Rhode Island, 1998. American Mathematical Society.

[20] M. LAURENT. Polynomial instances of the positive semidefinite and Euclidean distance matrix completion problems. *SIAM J. Matrix Anal. Appl.*, 22:874–894, 2000.

[21] R.D.C. MONTEIRO and M.J. TODD. Path-following methods. In *Handbook of Semidefinite Programming*, pages 267–306. Kluwer Acad. Publ., Boston, MA, 2000.

[22] J.B. SAXE. Embeddability of weighted graphs in k-space is strongly NP-hard. In *Seventeenth Annual Allerton Conference on Communication, Control, and Computing, Proceedings of the Conference held in Monticello, Ill., October 10–12, 1979*, pages xiv+1036, Urbana, 1979. University of Illinois Department of Electrical Engineering. Proceedings of the International School of Physics "Enrico Fermi", LXX*.

[23] I.J. SCHOENBERG. Remarks to Maurice Frechet's article: Sur la definition axiomatique d'une classe d'espaces vectoriels distancies applicables vectoriellement sur l'espace de Hilbert. *Ann. Math.*, 36:724–732, 1935.

[24] W.S. TORGERSON. Multidimensional scaling. I. Theory and method. *Psychometrika*, 17:401–419, 1952.

[25] J. VANDERGRAFT. Spectral properties of matrices which have invariant cones. *SIAM J. Applied Math.*, 16:1208–1222, 1968.

[26] H. WOLKOWICZ. Solving semidefinite programs using preconditioned conjugate graditents. *Optim. Methods Softw.*, 19(1):?–?, 2004. Research report CORR 2001-49.

[27] H. WOLKOWICZ, R. SAIGAL, and L. VANDENBERGHE, editors. *HANDBOOK OF SEMIDEFINITE PROGRAMMING: Theory, Algorithms, and Applications.* Kluwer Academic Publishers, Boston, MA, 2000. xxvi+654 pages.

[28] S. WRIGHT. *Primal-Dual Interior-Point Methods.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa, 1996.

A Transformations and Adjoints

We collect various definitions of linear transformations and their adjoints here. More details are given in Section 1.2.

Let $X \in \mathcal{S}^{n-1}$ and $D = \begin{pmatrix} 0 & d^T \\ d & \bar{D} \end{pmatrix} \in \mathcal{S}^n$ with $d_e = d + \bar{D}e$. The interpolation operator $\mathcal{I} : \mathcal{S}^n \to \mathbb{R}^m$.

1.
$$\mathcal{I}(S) = \sum_{k=1}^{m} \sqrt{2} S_{i_k j_k} e_k, \quad \mathcal{I}^*(b) = \sum_{k=1}^{m} b_k E_{i_k j_k}, \quad \mathcal{I}^\dagger = \mathcal{I}^*.$$

2.
$$\text{offDiag}\,(X) = X - \text{Diag}\,(\text{diag}\,(X)), \quad \text{zeros out the diagonal.}$$

3.
$$\mathcal{D}_e(X) = \text{diag}\,(X) e^T + e\,\text{diag}\,(X)^T, \quad \mathcal{D}_e^*(D) = 2\text{Diag}\,(De).$$

4. (a)
$$\mathcal{L}(X) = \begin{pmatrix} 0 & \text{diag}\,(X)^T \\ \text{diag}\,(X) & \mathcal{D}_e(X) - 2X \end{pmatrix};$$

   $\text{diag}\,(\mathcal{L}(X)) = 0$; and $X \succeq 0 \Rightarrow \mathcal{L}(X) \in \mathcal{E}$;

   (b)
$$\mathcal{L}^*(D) = 2\{\text{Diag}\,(d) + \text{Diag}\,(\bar{D}e) - \bar{D}\} = 2\{\text{Diag}\,(d_e) - \bar{D}\};$$

   D pre-distance matrix $\Rightarrow \mathcal{L}^*(D) \succeq 0$; $\mathcal{D}_e(\mathcal{L}^*(D)) = 2(d_e e^T + e d_e^T)$ (if $\text{diag}\,(D) = 0$).

(c) Assume that $\operatorname{diag}(D) = 0$. Then

$$\begin{array}{rcl} \mathcal{L}\mathcal{L}^*(D) & = & 2\mathcal{L}\{\operatorname{Diag}(d) + \operatorname{Diag}(\bar{D}e) - \bar{D}\} \\ & = & \begin{pmatrix} 0 & 2d_e^T \\ 2d_e & \mathcal{D}_e(\mathcal{L}^*(D)) - 2\mathcal{L}^*(D) \end{pmatrix} \\ & = & \begin{pmatrix} 0 & 2d_e^T \\ 2d_e & 2(d_e e^T + e d_e^T) - 2(2(\operatorname{Diag}(d_e) - \bar{D})) \end{pmatrix} \\ & = & 2\begin{pmatrix} 0 & d_e^T \\ d_e & \operatorname{offDiag}(d_e e^T + e d_e^T) + \bar{D} \end{pmatrix}; \end{array}$$

D pre-distance matrix $\Rightarrow \mathcal{L}\mathcal{L}^*(D) \in \mathcal{E}$;

(d)

$$\mathcal{L}^\dagger(D) = \frac{1}{2}\left(de^T + ed^T - \bar{D}\right), \quad \mathcal{L}^\dagger\mathcal{L} = I \text{ and } \mathcal{L}\mathcal{L}^\dagger = \operatorname{offDiag}.$$

See Lemma 3.1.

5.

$$\mathcal{A}^\dagger(b) = \mathcal{L}^\dagger\mathcal{I}^*(b) = \sum_{k=1}^{m} \mathcal{L}^\dagger b_k E_k,$$

6.

$$\begin{array}{rcl} \mathcal{K}(B) & := & \operatorname{diag}(B)\, e^T + e \operatorname{diag}(B)^T - 2B \\ & := & \mathcal{D}_e(B) - 2(B); \\ \mathcal{K}^*(D) & = & 2\operatorname{Diag}(De) - 2D \\ & = & 2(\operatorname{Diag}(De) - D). \end{array}$$

7.

$$\bar{\mathcal{T}}(D) \;:=\; -\tfrac{1}{2}JDJ \quad (= \bar{\mathcal{T}}^*(D)).$$