Robust Semidefinite Programming Approaches for Sensor Network Localization with Anchors

Nathan Krislock [*] Veronica Piccialli [†] Henry Wolkowicz [‡]

May 20, 2006

University of Waterloo
Department of Combinatorics and Optimization
Waterloo, Ontario N2L 3G1, Canada

Research Report CORR 2006-12

Key Words: Sensor Network Localization, Anchors, Graph Realization, Semidefinite Programming, Euclidean Distance Matrix Completions, Gauss-Newton Method

AMS Subject Classification:

Contents

[*]Research supported by Natural Sciences Engineering Research Council Canada. E-mail ngbkrislock@uwaterloo.ca
[†]Department of Computer and Systems Science "Antonio Ruberti" University of Rome "La Sapienza" Via Buonarroti 12, 00185 Rome, Italy. Email Veronica.Piccialli@dis.uniroma1.it
[‡]Research supported by Natural Sciences Engineering Research Council Canada. E-mail hwolkowicz@uwaterloo.ca

List of Tables

List of Figures

Abstract

We derive a robust primal-dual interior-point algorithm for a semidefinite programming, $\boldsymbol{SDP}$, relaxation for sensor localization with anchors and with noisy distance information. The relaxation is based on finding a *Euclidean Distance Matrix,* $\boldsymbol{EDM}$ *,* that is nearest in the Frobenius norm for the known noisy distances and that satisfies given upper and lower bounds on the unknown distances.

We show that the $\boldsymbol{SDP}$ relaxation for this nearest $\boldsymbol{EDM}$ problem is usually underdetermined and is an ill-posed problem. Our interior-point algorithm exploits the structure and maintains exact feasibility at each iteration. High accuracy solutions can be obtained despite the ill-conditioning of the optimality conditions.

Included are discussions on the strength and stability of the $\boldsymbol{SDP}$ relaxations, as well as results on invariant cones related to the operators that map between the cones of semidefinite and Euclidean distance matrices.

1 Introduction

Many applications use ad hoc wireless sensor networks for monitoring information, e.g. for earthquake detection, ocean current flows, weather, etc... Typical networks include a large number of sensor nodes which gather data and communicate among themselves. The location of a subset of the sensors is known; these sensor nodes are called *anchors*. From intercommunication among sensor nodes within a given (radio) range, we are able to establish approximate distances between a subset of the sensors and anchors. The sensor localization problem is to determine/estimate the location of all the sensors from this partial information on the distances. For more details on the various applications, see e.g. [7, 4, 14].

We model the problem by treating it as a nearest Euclidean Distance Matrix, $\boldsymbol{EDM}$, problem with lower bounds formed using the radio range between pairs of nodes for which no distance exists. We also allow for existing upper bounds. When solving for the sensor locations, the problem is nonconvex and hard to solve exactly. We study the semidefinite programming, $\boldsymbol{SDP}$, relaxation of the sensor localization problem. We project the feasible set onto the *minimal face* of the cone of semidefinite matrices, $\boldsymbol{SDP}$, and show that strict feasibility holds for the primal problem. However, we illustrate that this $\boldsymbol{SDP}$ relaxation has inherent instability, is usually underdetermined and ill-posed. This is due to both the nearest matrix model as well as the $\boldsymbol{SDP}$ relaxation. In addition, the relaxation can be strengthened using Lagrangian relaxation. Nevertheless, we derive a robust interior-point algorithm for the projected $\boldsymbol{SDP}$ relaxation. The algorithm exploits the special structure of the linear transformations for both efficiency and robustness. We eliminate, in advance, the primal and dual linear feasibility equations. We work with an overdetermined bilinear system, rather than using a square symmetrized approach, that is common for $\boldsymbol{SDP}$. This uses the Gauss-Newton approach in [15] and involves a crossover step after which we use the affine scaling direction with steplength one and without any backtracking to preserve positive definiteness. The search direction is found using a preconditioned conjugate gradient method, `LSQR` [18]. The algorithm maintains exact primal and dual feasibility throughout the iterations and attains high accurate solutions. Our numerical tests show that, despite the illconditioning, we are able to obtain high accuracy in the solution of the $\boldsymbol{SDP}$ problem. In addition, we obtain surprisingly highly accurate approximations of the original sensor localization problem.

1.1 Outline

The formulation of the sensor localization problem is outlined in Section 1.2. We continue in Section 2 with background and notation, including information on the linear transformation and adjoints used in the model. In Section 2.3 we present explicit representations of the Perron root and corresponding eigenvector for the important linear transformation that maps between the cones of $\mathcal{EDM}$ and $\mathcal{SDP}$.

The $\mathcal{SDP}$ relaxation is presented in Section 3. We use the Frobenius norm rather than the ℓ_1 norm that is used in the literature. We then project the problem onto the *minimal face* in order to obtain Slater's constraint qualification (strict feasibility) and guarantee strong duality. The dual is presented in Section 4.

The primal-dual interior-point (p-d i-p) algorithm is derived in Section 5. We include a heuristic for obtaining a strictly feasible starting point and the details for the diagonal preconditioning using the structure of the linear transformations. The algorithm uses a crossover technique, i.e. we use the affine scaling step without backtracking once we get sufficient decrease in the duality gap.

The instability in the $\mathcal{SDP}$ relaxation, and different approaches on dealing with it, is discussed in Section 6. We then continue with numerical tests and concluding remarks in Sections 7 and 8. We include an appendix with the details of the composition and adjoints of the linear transformations, as well as the details of the diagonal preconditioning.

1.2 Problem Formulation

Let the n unknown (sensor) points be $p^1, p^2, \ldots, p^n \in \mathbb{R}^r$, r the embedding dimension; and let the m known (anchor) points be $a^1, a^2, \ldots, a^m \in \mathbb{R}^r$. Let $A^T = [a^1, a^2, \ldots, a^m]$, $X^T = [p^1, p^2, \ldots, p^n]$, and define

$$P^T := \left(p^1, p^2, \ldots, p^n, a^1, a^2, \ldots, a^m\right) = \left(X^T \; A^T\right). \tag{1.1}$$

Note that we can always shift all the sensors and anchors so that the anchors are centered at the origin, $A^T e = 0$. We can then shift them all back at the end. To avoid some special trivial cases, we assume the following.

Assumption 1.1 *The number of sensors and anchors, and the embedding dimension satisfy*

$$n > m > r, \;\; A^T e = 0, \;\; \text{and } A \text{ is full column rank.}$$

Now define $(\mathcal{N}_e, \mathcal{N}_u, \mathcal{N}_l)$, respectively, to be the index sets of specified (distance values, upper bounds, lower bounds), respectively, of the distances d_{ij} between pairs of nodes from $\{p^i\}_1^n$ (sensors); and let $(\mathcal{M}_e, \mathcal{M}_u, \mathcal{M}_l)$, denote the same for distances between a node from $\{p^i\}_1^n$ (sensor) and a node from $\{a^k\}_1^m$ (anchor). Define (the partial Euclidean Distance Matrix) E with elements

$$E_{ij} = \begin{cases} d_{ij}^2 & \text{if} \quad ij \in \mathcal{N}_e \cup \mathcal{M}_e \\ 0 & \text{otherwise.} \end{cases}$$

The underlying graph is

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}), \tag{1.2}$$

with node set $\{1, \ldots, m+n\}$ and edge set $\mathcal{N}_e \cup \mathcal{M}_e$. Similarly, we define the matrix of (squared) upper distance bounds U and the matrix of (squared) lower distance bounds L for $ij \in \mathcal{N}_u \cup \mathcal{M}_u$ and $\mathcal{N}_l \cup \mathcal{M}_l$, respectively.

Our first formulation for finding the sensor locations p^j is the feasibility question for the constraints:

$$\begin{array}{lll} \|p^i - p^j\|^2 = E_{ij} & \forall (i,j) \in \mathcal{N}_e & \left(n_e = \frac{|\mathcal{N}_e|}{2}\right) \\ \|p^i - a^k\|^2 = E_{ik} & \forall (i,k) \in \mathcal{M}_e & \left(m_e = \frac{|\mathcal{M}_e|}{2}\right) \\ \|p^i - p^j\|^2 \leq U_{ij} & \forall (i,j) \in \mathcal{N}_u & \left(n_u = \frac{|\mathcal{N}_u|}{2}\right) \\ \|p^i - a^k\|^2 \leq U_{ik} & \forall (i,k) \in \mathcal{M}_u & \left(m_u = \frac{|\mathcal{M}_u|}{2}\right) \\ \|p^i - p^j\|^2 \geq L_{ij} & \forall (i,j) \in \mathcal{N}_l & \left(n_l = \frac{|\mathcal{N}_l|}{2}\right) \\ \|p^i - a^k\|^2 \geq L_{ik} & \forall (i,k) \in \mathcal{M}_l & \left(m_l = \frac{|\mathcal{M}_l|}{2}\right) \end{array} \tag{1.3}$$

Note that the first two and the last two sets of constraints are quadratic, nonconvex, constraints.

Let W_p, W_a be weight matrices. For example, they simply could be $0,1$ matrices that indicate when an exact distance is unknown or known. Or a weight could be used to indicate the confidence in the value of the distance. If there is noise in the data, the exact model (1.3) can be infeasible. Therefore, we can minimize the weighted least squares error.

$$(\boldsymbol{EDMC}) \qquad \begin{array}{rl} \min f_1(P) := & \frac{1}{2} \sum\limits_{(i,j)\in\mathcal{N}_e} (W_p)_{ij}(\|p^i - p^j\|^2 - E_{ij})^2 \\ & + \frac{1}{2} \sum\limits_{(i,k)\in\mathcal{M}_e} (W_a)_{ik}(\|p^i - a^k\|^2 - E_{ik})^2 \\ \text{subject to} & \|p^i - p^j\|^2 \leq U_{ij} \quad \forall (i,j) \in \mathcal{N}_u \quad \left(n_u = \frac{|\mathcal{N}_u|}{2}\right) \\ & \|p^i - a^k\|^2 \leq U_{ik} \quad \forall (i,k) \in \mathcal{M}_u \quad \left(m_u = \frac{|\mathcal{M}_u|}{2}\right) \\ & \|p^i - p^j\|^2 \geq L_{ij} \quad \forall (i,j) \in \mathcal{N}_l \quad \left(n_l = \frac{|\mathcal{N}_l|}{2}\right) \\ & \|p^i - a^k\|^2 \geq L_{ik} \quad \forall (i,k) \in \mathcal{M}_l \quad \left(m_l = \frac{|\mathcal{M}_l|}{2}\right). \end{array} \tag{1.4}$$

This is a *hard* problem to solve due to the nonconvex objective and constraints. One could apply global optimization techniques. In this paper, we use a convex relaxation approach, i.e. we use semidefinite relaxation. This results in a surprisingly robust, strong relaxation.

1.2.1 Related Approaches

Several recent papers have developed algorithms for the semidefinite relaxation of the sensor localization problem. Recent work on this area includes e.g. [14, 6, 4, 19, 5]. These relaxations use the ℓ_1 norm rather than the ℓ_2 norm that we use. We assume the noise in the radio signal is from a multivariate normal distribution with mean 0 and variance-covariance matrix $\sigma^2 I$, i.e. from a *spherical normal distribution* so that the least squares estimates are the maximum likelihood estimates. Our approach follows that in [2] for $\boldsymbol{EDM}$ completion *without* anchors.

2 Background and Preliminary Theory

2.1 Linear Transformations and Adjoints Related to EDM

We work in spaces of real matrices, $\mathcal{M}^{s\times t}$, equipped with the trace inner-product $\langle A, B\rangle = \operatorname{trace} A^T B$ and induced Frobenius norm $\|A\|^2 = \operatorname{trace} A^T A$. The space of $n \times n$ real symmetric matrices is denoted $\mathcal{S}^n$. It's dimension is $t(n) = n(n+1)/2$. For a given $B \in \mathcal{S}^n$, we denote

$b = \mathrm{svec}\,(B) \in \mathbb{R}^{t(n)}$ to be the vector obtained columnwise from the upper-triangular part of B with the strict upper-triangular part multiplied by $\sqrt{2}$. Thus svec is an isometry mapping $\mathcal{S}^n \to \mathbb{R}^{t(n)}$. The inverse and adjoint mapping is $B = \mathrm{sMat}\,(b)$. We now define several linear operators on $\mathcal{S}^n$. (A collection of linear transformations, adjoints and properties are given in the appendices.)

$$\mathcal{D}_e(B) \;:=\; \mathrm{diag}\,(B)\,e^T + e\,\mathrm{diag}\,(B)^T, \qquad \mathcal{K}(B) \;:=\; \mathcal{D}_e(B) - 2B, \tag{2.5}$$

where e is the vector of ones. The adjoint linear operators are

$$\mathcal{D}_e^*(D) = 2\mathrm{Diag}\,(De), \qquad \mathcal{K}^*(D) = 2(\mathrm{Diag}\,(De) - D). \tag{2.6}$$

By abuse of notation we allow $\mathcal{D}_e$ to act on $\mathbb{R}^n$:

$$\mathcal{D}_e(v) = ve^T + ev^T,\ v \in \mathbb{R}^n.$$

The linear operator $\mathcal{K}$ maps the cone of positive semidefinite matrices (denoted $\boldsymbol{SDP}$) onto the cone of Euclidean distance matrices (denoted $\boldsymbol{EDM}$), i.e. $\mathcal{K}(\boldsymbol{SDP}) = \boldsymbol{EDM}$. This allows us to change problem $\boldsymbol{EDMC}$ into a $\boldsymbol{SDP}$ problem.

We define the linear transformation $\mathrm{sblk}_i(S) = S_i \in \mathcal{S}^t$, on $S \in \mathcal{S}^n$, which pulls out the i-th diagonal block of the matrix S of dimension t. (The values of t and n can change and will be clear from the context.) The adjoint $\mathrm{sblk}_i^*(T) = \mathrm{sBlk}_i(T)$, where $T \in \mathcal{S}^t$, constructs a symmetric matrix of suitable dimensions with all elements zero expect for the i-th diagonal block given by T.

Similarly, we define the linear transformation $\mathrm{sblk}_{ij}(G) = G_{ij}$, on $G \in \mathcal{S}^n$, which pulls out the ij block of the matrix G of dimension $k \times l$ and multiplies it by $\sqrt{2}$. (The values of k, l, and n can change and will be clear from the context.) The adjoint $\mathrm{sblk}_{ij}^*(J) = \mathrm{sBlk}_{ij}(J)$, where $J \in \mathcal{M}^{k\times l} \cong \mathbb{R}^{kl}$, constructs a symmetric matrix which has all elements zero expect for the block ij which is given by J multiplied by $\frac{1}{\sqrt{2}}$, and for the block ji which is given by J^Tmultiplied by $\frac{1}{\sqrt{2}}$. We consider $J \in \mathcal{M}^{k\times l}$ to be a $k \times l$ matrix and equivalently $J \in \mathbb{R}^{kl}$ is a vector of length kl with the positions known. The multiplication by $\sqrt{2}$ (or $\frac{1}{\sqrt{2}}$) guarantees that the mapping is an isometry.

2.2 Properties of Transformations

Lemma 2.1 *([1])*

- *The nullspace $\mathcal{N}(\mathcal{K})$ equals the range $\mathcal{R}(\mathcal{D}_e)$.*
- *The range $\mathcal{R}(\mathcal{K})$ equals the hollow subspace of $\mathcal{S}^n$, denoted $S_H := \{D \in \mathcal{S}^n : \mathrm{diag}\,(D) = 0\}$.*
- *The range $\mathcal{R}(\mathcal{K}^*)$ equals the centered subspace of $\mathcal{S}^n$, denoted $S_c := \{B \in \mathcal{S}^n : Be = 0\}$.* ■

Corollary 2.1 *1. Let S_D denote the cone of diagonal matrices in $\mathcal{S}^n$. Then*

$$\begin{array}{c} S_c = \mathcal{N}(D_e^*) = \mathcal{R}(\mathcal{K}^*) \quad \perp \quad \mathcal{N}(K) = \mathcal{R}(\mathcal{D}_e) \\ S_H = \mathcal{R}(K) = \mathcal{N}(\mathcal{D}_e) \quad \perp \quad S_D = \mathcal{N}(K^*) = \mathcal{R}(\mathcal{D}_e^*). \end{array}$$

2. Let $\begin{pmatrix} V & \frac{1}{\sqrt{n}}e \end{pmatrix}$ *be an* $n \times n$ *orthogonal matrix. Then*

$$Y \succeq 0 \iff Y = V\hat{Y}V^T + \mathcal{D}_e(v) \succeq 0, \text{ for some } \hat{Y} \in \mathcal{S}^{n-1}, v \in \mathbb{R}^n.$$

Proof. Item (1) follows immediately from the definitions of the range $\mathcal{R}(\mathcal{K}^*) = S_c$ and the nullspace $\mathcal{N}(\mathcal{K}) = \mathcal{R}(\mathcal{D}_e)$. Item (2) now follows from the definition of V, i.e. the subspace of centered matrices $S_c = V\mathcal{S}^{n-1}V^T$. ■

Let $B = PP^T$. Then

$$D_{ij} = \|p^i - p^j\|^2 = (\mathrm{diag}\,(B)e^T + e\mathrm{diag}\,(B)^T - 2B)_{ij} = (\mathcal{K}(B))_{ij}\,, \tag{2.7}$$

i.e. the ***EDM*** $D = (D_{ij})$ and the points p_i in P are related by $D = \mathcal{K}(B)$, see (2.5).

Lemma 2.2 *Suppose that* $0 \preceq B \in \mathcal{S}^n$*. Then* $D = \mathcal{K}(B)$ *is EDM.*

Proof. Suppose that $\mathrm{rank}\,(B) = r$. Since $B \succeq 0$, there exists $P \in \mathcal{M}^{r \times n}$ such that $B = PP^T$. Thus

$$\begin{aligned} d_{ij} = (\mathcal{K}(B))_{ij} &= (\mathcal{K}(PP^T))_{ij} = (\mathrm{diag}\,(PP^T)\,e^T)_{ij} + (e\,\mathrm{diag}\,(PP^T))_{ij} - 2(PP^T)_{ij} \\ &= (p^i)^T p^i + (p^j)^T p^j - 2(p^i)^T p^j = \|p^i - p^j\|^2, \end{aligned}$$

where p^i is the i-th column of P. Thus $D = \mathcal{K}(B)$ is EDM. ■

The following shows that we can assume B is centered.

Lemma 2.3 *Let* $0 \preceq B \in \mathcal{S}^n$ *and define*

$$v := -\frac{1}{n}Be + \frac{e^T Be}{2n^2}e, \qquad C := B + ve^T + ev^T.$$

Then $C \succeq 0, D = \mathcal{K}(B) = \mathcal{K}(C)$ *and, moreover,* $Ce = 0$.

Proof. First, note that $\mathcal{K}(B) = \mathcal{K}(C) \iff \mathcal{K}(ve^T + ev^T) = 0$. But $(ve^T + ev^T) \in \mathcal{N}(\mathcal{K})$, by Lemma 2.1. Hence $\mathcal{K}(B) = \mathcal{K}(C)$. Moreover

$$\begin{aligned} Ce &= Be - \frac{1}{n}Bee^T e + \frac{e^T Be}{2n^2}ee^T e - \frac{1}{n}ee^T Be + \frac{e^T Be}{2n^2}ee^T e \\ &= Be - Be + \frac{e^T Be}{2n}e - \frac{e^T Be}{n}e + \frac{e^T Be}{2n}e = 0. \end{aligned}$$

Finally, let $P = \begin{pmatrix} e & V \end{pmatrix}$ be orthogonal. Then

$$P^T CP = \begin{pmatrix} 0 & 0 \\ 0 & V^T CV \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & V^T BV \end{pmatrix} \succeq 0.$$

Therefore, by Sylvester's Theorem on congruence, $C \succeq 0$. ■

2.3 Invariant Cones and Perron Roots

We digress in order to illustrate some of the relationships between the $\boldsymbol{SDP}$ and $\boldsymbol{EDM}$ cones. These relationships give rise to elegant results including generalizations of the Perron-Frobenius Theorem.

Lemma 2.4 *Suppose that $0 \leq H \in \mathcal{S}^n$, i.e. H is symmetric and nonnegative elementwise. Define the linear operator on $\mathcal{S}^n$*

$$\mathcal{U} := (H \circ \mathcal{K})^*(H \circ \mathcal{K}).$$

Then

$$\mathcal{U}(S^n_+) \subset \mathcal{F}_{\mathcal{E}}, \tag{2.8}$$

where the face of $\mathcal{S}^n_+$

$$\mathcal{F}_{\mathcal{E}} := \mathcal{S}^n_+ \cap S_C.$$

Proof. Let $X \in \mathcal{S}^n_+$ be given. We first show that $\mathcal{U}(X) \in \mathcal{S}^n_+$; equivalently, we show that $\operatorname{trace} Q\mathcal{U}(X) \geq 0,\ \forall Q \succeq 0$.

Let $Q \succeq 0$. We have

$$\operatorname{trace} Q\mathcal{U}(X) = \langle Q, (H \circ \mathcal{K})^*(H \circ \mathcal{K})(X)\rangle = \langle H \circ \mathcal{K}(Q), H \circ \mathcal{K}(X)\rangle = \langle H \circ D1, H \circ D2\rangle \geq 0,$$

since $D_1, D_2 \in \boldsymbol{EDM}$ and $H \geq 0$. Therefore $\mathcal{U}(X) = (H \circ \mathcal{K})^*(H \circ \mathcal{K})(X) \in \mathcal{S}^n_+$. The result (2.8) now follows since Lemma 2.1 implies $\mathcal{U}(X) \in \mathcal{R}(\mathcal{K}^*) = S_c$. ■

Corollary 2.2 *Suppose that $0 \leq H \in \mathcal{S}^n$ is defined as in Lemma 2.4. And suppose that* $\operatorname{diag} H = 0$ *and H has no zero row. Then the face $\mathcal{F}_{\mathcal{E}}$ (and its relative interior* $\operatorname{relint} \mathcal{F}_{\mathcal{E}}$*) are invariant under the operator $\mathcal{U} = (H \circ \mathcal{K})^*(H \circ \mathcal{K})$, i.e.*

$$\mathcal{U}(\mathcal{F}_{\mathcal{E}}) \subset \mathcal{F}_{\mathcal{E}}, \qquad \mathcal{U}(\operatorname{relint} \mathcal{F}_{\mathcal{E}}) \subset \operatorname{relint} \mathcal{F}_{\mathcal{E}}.$$

Proof. That the face $\mathcal{F}_{\mathcal{E}}$ is invariant follows immediately from Lemma 2.4.

Note that $nI - E \in \mathcal{F}_{\mathcal{E}}$ and has rank $n-1$; and $B \in \mathcal{F}_{\mathcal{E}} \Rightarrow \operatorname{trace} BE = 0$ implies that rank B is at most $n-1$. Therefore, the relative interior of the face $\mathcal{F}_{\mathcal{E}}$ consists of the centered, positive semidefinite matrices of rank exactly $n-1$.

Suppose that $B \in \operatorname{relint} \mathcal{F}_{\mathcal{E}}$, i.e. $B \succeq 0, \operatorname{trace} BE = 0$ and $\operatorname{rank} B = n-1$. Then we know that $\mathcal{U}(B) \in \mathcal{F}_{\mathcal{E}}$, by Lemma 2.4. It remains to show that $\mathcal{U}(B)$ is in the relative interior of the face, i.e. the rank of $\mathcal{U}(B)$ is $n-1$.

We show this by contradiction. Suppose that $\mathcal{U}(B)v = 0,\ 0 \neq v \neq \alpha e, \alpha \in \Re$. Then $0 = v^T\mathcal{U}(B)v = \operatorname{trace} vv^T\mathcal{U}(B) = \operatorname{trace}\mathcal{U}^*(vv^T)B$. Denote $C = \mathcal{U}^*(vv^T) = \mathcal{K}^*\left(H^{(2)} \circ \mathcal{K}(vv^T)\right)$. Since $B \in \operatorname{relint} \mathcal{F}_{\mathcal{E}}$, $\mathcal{U}(vv^T) \succeq 0$, and $\mathcal{U}^*(vv^T)B = \mathcal{U}(vv^T)B = 0$, we conclude C is in the conjugate face of $\mathcal{F}_{\mathcal{E}}$, i.e. $C = \alpha E$, for some $\alpha \geq 0$. By Lemma 2.1, we get C is centered. This implies that $C = 0$. Lemma 2.1 also implies that $\mathcal{N}(\mathcal{K}^*)$ is the set of diagonal matrices. Since H has no zero row, Lemma 2.1 implies that $vv^T \in \mathcal{N}(\mathcal{K}) = \mathcal{R}(\mathcal{D}_e)$. We now have $vv^T = we^T + ew^T$, for some w, i.e. $v = .5e$, the desired contradiction. ■

The above Lemma means that we can apply the generalized Perron-Frobenius Theorem, e.g. [20] to the operator $\mathcal{U}$, i.e. the spectral radius corresponds to a positive real eigenvalue with a corresponding eigenvector in the relative interior of the face $\mathcal{F}_{\mathcal{E}}$. In particular, we can get the following expression for the Perron eigenpair.

Theorem 2.1 *Suppose that* $0 \leq H \in \mathcal{S}^n$ *is defined as in Lemma 2.4 and that* $H_{ij} \in \{0,1\}$. *Suppose further that* $\operatorname{diag} H = 0$ *and* H *has no zero row. Denote the maximum eigenvalue*

$$\alpha := \max_i \left\{ \lambda_i \left(H + \operatorname{Diag}(He) \right) \right\},$$

and let x *be the corresponding eigenvector. Then the Perron root and eigenvector for the linear operator* $\mathcal{U}_H = \mathcal{K}^*(H^{(2)} \circ \mathcal{K}) = \mathcal{K}^*(H \circ \mathcal{K})$ *are, respectively,*

$$\lambda = 2\alpha - 2, \quad X = \operatorname{Diag}(x) + \operatorname{offDiag}\left(-\frac{1}{\alpha} H_{ij}(x_i + x_j) \right),$$

where $\operatorname{offDiag}(S) := S - \operatorname{Diag}(\operatorname{diag}(S))$.

Proof. By the the expression for $\mathcal{K}^*$, (2.6), we see that the graph of the eigenvector X is the same as for matrices in the range of $H \circ \mathcal{K}$, i.e. the graph of X is a subgraph of the graph of H. Moreover, since $X \in \mathcal{R}(\mathcal{K}^*)$, we have $Xe = 0$.

Now, the eigenvalue-eigenvector equation gives

$$\begin{aligned} \lambda X &= \mathcal{K}^*(H \circ \mathcal{K}(X)) = 2\operatorname{Diag}\left(\{H \circ \mathcal{K}(X)\}e\right) - 2H \circ \mathcal{K}(X) \\ &= 2\operatorname{Diag}\left(\{H \circ (\operatorname{diag}(X)e^T)\}e\right) + 2\operatorname{Diag}\left(\{H \circ (e\operatorname{diag}(X)^T)\}e\right) - 4\operatorname{Diag}\left(\{H \circ X\}e\right) \\ &\quad -2H \circ (\operatorname{diag}(X)e^T) - 2H \circ (e\operatorname{diag}(X)^T) + 4H \circ X. \end{aligned}$$

Since $H_{ij} \in \{0,1\}$, we have

$$H \circ X = X - \operatorname{Diag}(\operatorname{diag}(X))$$

Hence, we get

$$\begin{aligned} \lambda X &= 2\operatorname{Diag}\left(\{H \circ (\operatorname{diag}(X)e^T)\}e\right) + 2\operatorname{Diag}\left(\{H \circ (e\operatorname{diag}(X)^T)\}e\right) - 4\operatorname{Diag}(Xe) \\ &\quad -4\operatorname{Diag}(\operatorname{Diag}(\operatorname{diag}(X))e) - 2H \circ (\operatorname{diag}(X)e^T) - 2H \circ (e\operatorname{diag}(X)^T) + 4X + 4\operatorname{Diag}(\operatorname{diag}(X)). \end{aligned}$$

Combining with $Xe = 0$, this can be rewritten as

$$\begin{aligned} \left(\frac{\lambda}{2} - 2\right) X &= \operatorname{Diag}\left(\{H \circ (\operatorname{diag}(X)e^T)\}e\right) + \operatorname{Diag}\left(\{H \circ (e\operatorname{diag}(X)^T)\}e\right) \\ &\quad -H \circ (\operatorname{diag}(X)e^T) - H \circ (e\operatorname{diag}(X)^T). \end{aligned}$$

By the above equation we get for the elements on the diagonal

$$\alpha x_{ii} = (\sum_{k=1}^{n} H_{ik}) x_{ii} + \sum_{i=i}^{k} H_{ik} x_{kk}, \quad i = 1, \ldots, n,$$

i.e.

$$\alpha x = (H + \operatorname{Diag}(He))x.$$

As for the off-diagonal elements, we have

$$\alpha x_{ij} = -H_{ij}(x_{ii} + x_{jj}), \quad i \neq j.$$

■

Corollary 2.3 *The Perron root and eigenvector for the linear operator*

$$\mathcal{U}_{\mathcal{E}}(\cdot) := \mathcal{K}^*\mathcal{K}(\cdot)$$

are

$$\lambda = 4n > 0, \quad X = \beta I + \gamma E \succeq 0,$$

where $\gamma = -\frac{\beta}{n}$*,* $\beta > 0$*, and* $X \in \operatorname{relint} \mathcal{F}_{\mathcal{E}}$*.*

Proof. By Corollary 2.2 and the invariant cone results in [20], we conclude that the spectral radius of $\mathcal{U}_{\mathcal{E}}$ is a positive eigenvalue and the corresponding eigenvector $X \in \operatorname{relint} \mathcal{F}_{\mathcal{E}}$. We set $H = E - I$ and we verify that λ and X satisfy Theorem 2.1. By setting $H = E - I$, we get

$$\begin{aligned} \alpha &= \max\{\, \lambda_i\,(H + \operatorname{Diag}(He))\,\} = \max\{\, \lambda_i\,(E - I + \operatorname{Diag}((n-1)e))\,\} \\ &= \max\{\, \lambda_i\,(E + (n-2)I)\,\} = 2n - 2 \end{aligned}$$

that implies $\lambda = 4n$. On the other hand we know that X has the same graph of H, and hence $X = \beta I + \gamma E$. Now we have that

$$\operatorname{diag}(X) = x = (\gamma + \beta)e = (-\frac{\beta}{n} + \beta)e = (\frac{(n-1)\beta}{n})e$$

and we can see by substitution that x is the eigenvector of the matrix $(H + \operatorname{Diag}(He))$ corresponding to the eigenvalue $\alpha = 2n - 2$. Moreover we have that

$$X_{ij} = -\frac{\beta}{n} = -\frac{1}{\alpha}(x_i + x_j).$$

Finally,

$$X = (\alpha E + \beta I) = \beta \left(I - \frac{1}{n} E \right) \in \operatorname{relint} \mathcal{F}_{\mathcal{E}}.$$

.

■

The eigenvector matrix can be written as $X = \beta I - \frac{\beta}{n}E$, where the spectral radius $\rho(\frac{\beta}{n}E) \leq \beta$, and E is nonnegative, i.e. X is an M-matrix, e.g. [3].

From the above proof we note that the corresponding ***EDM*** for the eigenvector X is $D = \mathcal{K}(X) = 2\beta(E - I)$.

Corollary 2.4 *Suppose that* $0 \leq H \in \mathcal{S}^n$ *is the banded matrix* $H = E - I$ *but with* $H(1,n) = H(n,1) = 0$*. Then the Perron root and eigenvector for the linear operator*

$$\mathcal{U}(\cdot) := \mathcal{K}^*(H \circ \mathcal{K}(\cdot))$$

are given by

$$\lambda = 3n-2+\sqrt{n^2+4n-12} > 0, \quad X = \begin{pmatrix} X_{11} & -\beta & -\beta & \dots & \dots & -\beta & 0 \\ -\beta & X_{22} & -\alpha & \dots & \dots & -\alpha & -\beta \\ -\beta & -\alpha & X_{33} & -\alpha & \dots & -\alpha & -\beta \\ & \dots & & & \dots & & \\ -\beta & -\alpha & -\alpha & -\alpha & \dots & X_{n-1,n-1} & -\beta \\ 0 & -\beta & -\beta & \dots & \dots & -\beta & X_{11} \end{pmatrix}, \tag{2.9}$$

where the diagonal of X is defined by $Xe = 0$, and the Perron root λ and vector $v = \begin{pmatrix} \beta \\ \alpha \end{pmatrix}$ are found from the spectral radius and corresponding eigenvector of the matrix

$$A = 2\begin{pmatrix} n+2 & n-3 \\ 4 & 2(n-4) \end{pmatrix}, \tag{2.10}$$

i.e. $Av = \lambda v$, with

$$\beta = \left(\frac{3}{4} - \frac{1}{8}n + \frac{1}{8}\sqrt{n^2+4n-12}\right)\alpha, \quad \alpha > 0.$$

Proof. With X defined in (2.9), we find that $H \circ \mathcal{K}(X)$ has the same structure as X but with zero diagonal and $-\beta$ replaced by $(n+2)\beta + (n-3)\alpha$ and $-\alpha$ replaced by $4\beta + (2n-4)\alpha$. Then $\mathcal{K}^*(H \circ \mathcal{K}(X))$ doubles and negates these values. Therefore, the eigenpair λ, X can be found by solving the eigenpair equation given by $Av = \lambda v$, where $v = \begin{pmatrix} \beta \\ \alpha \end{pmatrix}$ and A is given by (2.10). ■

3 SDP Relaxation

3.1 EDMC Problem Reformulation using Matrices

Now let $\bar{Y} := PP^T = \begin{pmatrix} XX^T & XA^T \\ AX^T & AA^T \end{pmatrix}$ and $Z = \begin{pmatrix} I \\ P \end{pmatrix}\begin{pmatrix} I \\ P \end{pmatrix}^T = \begin{pmatrix} I & P^T \\ P & \bar{Y} \end{pmatrix}$. Therefore we have the dimensions:

X $n \times r$; A $m \times r$; P $m+n \times r$; $\bar{Y} = \bar{Y}^T$ $m+n \times m+n$; $Z = Z^T$ $m+n+r \times m+n+r$.

We can reformulate ***EDMC*** using matrix notation to get the equivalent problem

$$(\boldsymbol{EDMC}) \qquad \begin{array}{rrcl} \min & f_2(\bar{Y}) := \frac{1}{2}\|W \circ (\mathcal{K}(\bar{Y}) - E)\|_F^2 & & \\ \text{subject to} & g_u(\bar{Y}) := H_u \circ (\mathcal{K}(\bar{Y}) - U) & \leq & 0 \\ & g_l(\bar{Y}) := H_l \circ (\mathcal{K}(\bar{Y}) - L) & \geq & 0 \\ & \bar{Y} - PP^T & = & 0, \end{array} \tag{3.11}$$

where W is the (n+m) $\times$ (n+m) weight matrix having a positive ij-element if $(i,j) \in \mathcal{N}_e \cup \mathcal{M}_e$, 0 otherwise. H_u, H_l are $0,1$-matrices where the ij-th element equals 1 if $(i,j) \in \mathcal{N}_u \cup \mathcal{M}_u$ ($(i,j) \in \mathcal{N}_l \cup \mathcal{M}_l$), 0 otherwise. By abuse of notation, we consider the functions g_u, g_l as implicitly acting on only the nonzero components in the upper triangular parts of the matrices that result from the Hadamard products with H_u, H_l, respectively.

Remark 3.1 *The function $f_2(\bar{Y}) = f_2(PP^T)$, and it is clear that $f_2(PP^T) = f_1(P)$ in (1.4). Note that the functions f_2, g_u, g_l act only on $\bar{Y}$ and the locations of the anchors and sensors is completely hiding in the* hard, *nonconvex quadratic constraint $\bar{Y} = PP^T = \begin{pmatrix} XX^T & XA^T \\ AX^T & AA^T \end{pmatrix}$.*

*The problem **EDMC** is a linear least squares problem with nonlinear constraints. As we show in Section 6, the objective function can be either overdetermined or underdetermined, depending on the weights W. This can result in ill-conditioning problems, e.g. [12].*

3.2 Relaxation of the Hard Quadratic Constraint

We now consider the hard quadratic constraint $\bar{Y} = PP^T = \begin{pmatrix} XX^T & XA^T \\ AX^T & AA^T \end{pmatrix}$, P is defined in (1.1). We study the standard semidefinite relaxation. We include details on problems and weaknesses with the relaxation.

The constraint in (3.11) has the (blocked) property

$$\bar{Y} = PP^T = \begin{pmatrix} XX^T & XA^T \\ AX^T & AA^T \end{pmatrix} \iff \bar{Y}_{11} = XX^T \text{ and } \bar{Y}_{21} = AX^T, \bar{Y}_{22} = AA^T. \tag{3.12}$$

With P defined in (1.1) and $\bar{Y} = \begin{pmatrix} \bar{Y}_{11} & \bar{Y}_{21}^T \\ \bar{Y}_{21} & AA^T \end{pmatrix}$, it is common practice to relax (3.12) to the semidefinite constraint

$$G(P, \bar{Y}) := PP^T - \bar{Y} \preceq 0, \quad \mathcal{F}_G := \{(P, \bar{Y}) : G(P, \bar{Y}) \preceq 0\}. \tag{3.13}$$

The function G is convex in the Löwner (semidefinite) partial order. Therefore, the feasible set $\mathcal{F}_G$ is a convex set. By a Schur complement argument, the constraint is equivalent to the semidefinite constraint $Z = \begin{pmatrix} I \\ P \end{pmatrix} \begin{pmatrix} I \\ P \end{pmatrix}^T \succeq 0$, where Z has the following block structure.

$$Z := \begin{pmatrix} I & P^T \\ P & \bar{Y} \end{pmatrix}, \quad P = \begin{pmatrix} X \\ A \end{pmatrix}, \quad \bar{Y} := \begin{pmatrix} Y & Y_{21}^T \\ Y_{21} & Y_2 \end{pmatrix}, \tag{3.14}$$

and we equate $Y = \bar{Y}_{11}, Y_{21} = \bar{Y}_{21}, Y_2 = \bar{Y}_{22}$. When we fix the $2,2$ block of $\bar{Y}$ to be AA^T (which is equivalent to fixing the anchors), we get the following semidefinite constraint:

$$0 \preceq Z = \begin{pmatrix} I & P^T \\ P & \bar{Y} \end{pmatrix}, \quad P = \begin{pmatrix} X \\ A \end{pmatrix}, \quad \bar{Y} = \begin{pmatrix} Y & Y_{21}^T \\ Y_{21} & AA^T \end{pmatrix}, \quad Y_{21} = AX^T \tag{3.15}$$

Equivalently,

$$Z = \begin{pmatrix} I & X^T & A^T \\ X & Y & Y_{21}^T \\ A & Y_{21} & AA^T \end{pmatrix} \succeq 0, \quad Y_{21} = AX^T. \tag{3.16}$$

We denote the corresponding (convex) feasible set for the relaxation

$$\mathcal{F}_Z := \{(X, Y, Y_{21}) : (3.16) \text{ holds}\}, \qquad P, \bar{Y} \text{ formed using (3.15)}. \tag{3.17}$$

Remark 3.2 *We emphasize two points in the above relaxation.*

First, we relax from the hard nonlinear, nonconvex quadratic equality constraint in (3.12) *to the convex quadratic constraint in* (3.13)*, i.e. the latter is a convex constraint in the Löwner (semidefinite) partial order. And, using the identifications for* $P, \bar{Y}$ *in* (3.15) *above, and by abuse of notation, the feasible sets are equal, i.e.*

$$\{(\bar{Y}, P) : \bar{Y} = PP^T\} \subset \mathcal{F}_G = \mathcal{F}_Z.$$

However, the rationale for relaxing from $=$ *to* $\preceq$ *is based on obtaining a simple convexification, rather than obtaining a strong convexification. We see below, in Section 6.4.2, that the Lagrangian relaxation of* $\bar{Y} = PP^T$ *provides a stronger semidefinite relaxation.*

Second, the Schur complement is used to change the quadratic function to a linear one. The fact that the two feasible sets are equal, $\mathcal{F}_Z = \mathcal{F}_G$*, does* not *mean that the two relaxations are numerically equivalent. In fact, the Jacobian of the convex function* G *acting on* H_P, H_Y *is*

$$G'(P, \bar{Y})(H_P, H_Y) = PH_P^T + H_P P^T - H_Y.$$

This immediately implies that G' *is onto. In fact, this is a Lyapunov type equation with known stability properties, e.g. [13, Chap. 15]. In addition, if* $G'(P, \bar{Y})(0, H_Y) = 0$*, then* $H_Y = 0$*. And, if* $H_Y = 0$*, then one can take the singular value decomposition of the full rank* $P = U\Sigma V^T$*, where* Σ *contains the* $r \times r$ *nonsingular diagonal block* Σ_r*. Upon taking the orthogonal congruence* $U^T \cdot U$*, we see immediately that* $H_P = 0$ *as well. Therefore, the Jacobian* G' *is full rank and the columns corresponding to the variables* P, Y *are separately linearly independent.*

However, the constants in the matrix $Z = Z(X, Y, Y_{21})$ *imply that the Jacobian is* never onto*. Therefore, though this constraint has been linearized/simplified, the problem has become numerically unstable. This is evident in our numerical tests below.*

3.3 Facial Reduction

As seen in [2], the cone of ***EDM*** has a one-one correspondence to a proper face of the cone of ***SDP***. Therefore, to apply interior-point methods on an equivalent ***SDP***, one needs to reduce the problem to a smaller dimensional cone of ***SDP*** where Slater's constraint qualification holds. This guarantees both strictly feasible points for interior-point methods as well as the existence of dual variables. We now find the minimal face containing Z defined above in (3.15), and then we reduce the problem in order to obtain strictly feasible points. Strict feasibility (Slater's constraint qualification) also guarantees strong duality.

Define the (compact) singular value decomposition

$$A = U\Sigma V^T, \quad \Sigma = \mathrm{Diag}\,(\sigma) \succ 0, \quad U \in \mathcal{M}^{m\times r}, U^T U = I, \quad V \in \mathcal{M}^{r \times n}, V^T V = I, \tag{3.18}$$

and define the following matrix:

$$Z_s := \begin{pmatrix} I & X^T \\ X & Y \end{pmatrix} \qquad \text{(the } 2 \times 2 \text{ block from (3.16))}$$

We now see that checking semidefiniteness of the big 3×3 block matrix $Z \succeq 0$ in (3.15) is equivalent to just checking semidefiniteness of the smaller 2×2 block matrix Z_s as long as we set $Y_{21} = AX^T$, i.e. we can ignore the last row and column of blocks in Z.

Theorem 3.1 *Suppose that Z is defined by* (3.15) *and the corresponding feasible set $\mathcal{F}_Z$ is defined in* (3.17)*. Then*

$$\{Z \succeq 0\} \Leftrightarrow \left\{Z_s \succeq 0, \text{ and } Y_{21} = AX^T\right\},$$

and the feasible set

$$\mathcal{F}_Z = \mathcal{F}_S := \left\{(X,Y,Y_{21}) : \mathcal{Z}_s(X,Y) \succeq 0, \text{ and } Y_{21} = AX^T\right\}.$$

Proof. By substituting the singular value decomposition of A into (3.15) for Z, we get $Z \succeq 0$ if and only if

$$Z_1 := Z_1(X,Y,Y_{21}) = \begin{pmatrix} I & X^T & V\Sigma U^T \\ X & Y & Y_{21}^T \\ U\Sigma V^T & Y_{21} & U\Sigma^2 U^T \end{pmatrix} \succeq 0. \tag{3.19}$$

We have $I = I_r \in \mathcal{S}^r, Y \in \mathcal{S}^n$, and $U\Sigma^2 U^T \in \mathcal{S}^m$ with rank r. Therefore,

$$\max_{X,Y,Y_{21},Z_1 \succeq 0} \operatorname{rank} Z_1 = r + n + r = 2r + n, \qquad \text{(attained if } Y \in \mathcal{S}^n_+\text{)}. \tag{3.20}$$

The corresponding feasible set is

$$\mathcal{F}_{Z_1} := \{(X,Y,Y_{21}) : Z_1(X,Y,Y_{21}) \succeq 0\}, \tag{3.21}$$

and it is clear that $\mathcal{F}_{Z_1} = \mathcal{F}_Z$.

Now, choose $\bar{U}$ so that $\begin{pmatrix} U & \bar{U} \end{pmatrix}$ is orthogonal; and, consider the nonsingular congruence

$$\begin{aligned} 0 \preceq Z_2 := T^T Z T \quad &= \begin{pmatrix} V^T & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & \begin{pmatrix} U & \bar{U} \end{pmatrix}^T \end{pmatrix} Z \begin{pmatrix} V & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & \begin{pmatrix} U & \bar{U} \end{pmatrix} \end{pmatrix} \\ &= \begin{pmatrix} I & V^T X^T & \begin{pmatrix} \Sigma & 0 \end{pmatrix} \\ XV & Y & \begin{pmatrix} Y_{21}^T U & Y_{21}^T \bar{U} \end{pmatrix} \\ \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} & \begin{pmatrix} U^T Y_{21} \\ \bar{U}^T Y_{21} \end{pmatrix} & \begin{pmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{pmatrix} \end{pmatrix}. \end{aligned} \tag{3.22}$$

This shows that $Z \succeq 0$ is equivalent to

$$0 \preceq Z_3 := Z_3(X,Y,Y_{21}) = \begin{pmatrix} I & V^T X^T & \Sigma \\ XV & Y & Y_{21}^T U \\ \Sigma & U^T Y_{21} & \Sigma^2 \end{pmatrix}, \text{ and } Y_{21}^T \bar{U} = 0. \tag{3.23}$$

The corresponding feasible set is

$$\mathcal{F}_{Z_3} := \{(X,Y,Y_{21}) : Z_3(X,Y,Y_{21}) \succeq 0 \text{ and } Y_{21}^T \bar{U} = 0\}. \tag{3.24}$$

And, it is clear that $\mathcal{F}_{Z_3} = \mathcal{F}_{Z_1} = \mathcal{F}_Z$.

Note that with Z_1 in (3.19), we have

$$Z_1 Q = 0, \text{ where } Q = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \bar{U}\bar{U}^T \end{pmatrix} \succeq 0, \text{ rank } Q = m - r,$$

i.e. Q is in the relative interior of the conjugate face to the minimal face containing $\mathcal{F}_Z$, and the minimal face (the generated cone) is

$$\mathcal{F}_Z = \boldsymbol{SDP} \cap Q^\perp = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & U \end{pmatrix} \mathcal{S}_+^{2r+n} \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & U \end{pmatrix}^T . \tag{3.25}$$

We now apply a Schur complement argument to Z_3, i.e. (3.23) holds if and only if $Y_{21}^T \bar{U} = 0$ and

$$\begin{aligned} 0 &\preceq \begin{pmatrix} I & V^T X^T \\ XV & Y \end{pmatrix} - \begin{pmatrix} \Sigma \\ Y_{21}^T U \end{pmatrix} \Sigma^{-2} \begin{pmatrix} \Sigma & U^T Y_{21} \end{pmatrix} \\ &= \begin{pmatrix} 0 & V^T X^T - \Sigma^{-1} U^T Y_{21} \\ XV - Y_{21}^T U \Sigma^{-1} & Y - Y_{21}^T U \Sigma^{-2} U^T Y_{21} \end{pmatrix}. \end{aligned} \tag{3.26}$$

And this is equivalent to

$$Y - Y_{21}^T U \Sigma^{-2} U^T Y_{21} \succeq 0, \quad \Sigma V^T X^T - U^T Y_{21} = 0, \quad Y_{21}^T \bar{U} = 0. \tag{3.27}$$

We define the corresponding feasible set

$$\mathcal{F}_S := \{(X, Y, Y_{21}) : (3.27) \text{ holds}\}. \tag{3.28}$$

We conclude that $Z \succeq 0$ if and only if (3.27) holds and the feasible sets $\mathcal{F}_S = \mathcal{F}_{Z_3} = \mathcal{F}_{Z_1} = \mathcal{F}_Z$.

Now, (3.27) is equivalent to

$$Y - XVV^T X^T = Y - XX^T \succeq 0, \text{ and } V^T X^T = \Sigma^{-1} U^T Y_{21}, \; Y_{21}^T \bar{U} = 0, \tag{3.29}$$

which is equivalent to

$$\begin{pmatrix} I & X^T \\ X & Y \end{pmatrix} \succeq 0, \text{ and } V^T X^T = \Sigma^{-1} U^T Y_{21} \; Y_{21}^T \bar{U} = 0. \tag{3.30}$$

Therefore

$$\mathcal{F}_S = \{(X, Y, Y_{21}) : (3.30) \text{ holds}\}. \tag{3.31}$$

Define the full column rank matrix

$$K := \begin{pmatrix} V^T & 0 \\ 0 & I \\ \Sigma V^T & 0 \end{pmatrix} . \tag{3.32}$$

Then (3.30) is equivalent to

$$0 \preceq \begin{pmatrix} I & V^T X^T & \Sigma \\ XV & Y & XV\Sigma \\ \Sigma & \Sigma V^T X^T & \Sigma^2 \end{pmatrix} = K \begin{pmatrix} I & X^T \\ X & Y \end{pmatrix} K^T, \text{ and } \Sigma V^T X^T = U^T Y_{21}, \tag{3.33}$$

Since this implies that $Y_{21}^T \bar{U} = 0$. Or equivalently

$$\begin{pmatrix} I & V^T X^T & \Sigma \\ XV & Y & XV\Sigma \\ \Sigma & \Sigma V^T X^T & \Sigma^2 \end{pmatrix} = K \begin{pmatrix} I & X^T \\ X & Y \end{pmatrix} K^T, \begin{pmatrix} I & X^T \\ X & Y \end{pmatrix} \succeq 0, \Sigma V^T X^T = U^T Y_{21}. \tag{3.34}$$

The constraint in (3.34) is equivalent to $\begin{pmatrix} I & X^T \\ X & Y \end{pmatrix} \succeq 0, U\Sigma V^T X^T = Y_{21}$. The result now follows since $A = U\Sigma V^T$. ■

Corollary 3.1 *The minimal face containing the feasible set $\mathcal{F}_Z$ has dimension $2r+n$ and is given in the cone in* (3.25)*. In particular, each given Z in* (3.16)*,*

$$Z = \begin{pmatrix} I & X^T & A^T \\ X & Y & Y_{21}^T \\ A & Y_{21} & AA^T \end{pmatrix} \succeq 0,$$

can be expressed as

$$Z = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & U \end{pmatrix} \begin{pmatrix} I & X^T & V\Sigma \\ X & Y & XV\Sigma \\ \Sigma V^T & \Sigma V^T X^T & \Sigma^2 \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & U \end{pmatrix}^T.$$

■

3.4 Further Model Development with only X, Y

The above reduction from Z to Y allows us to replace the constraint $Z \succeq 0$ with the smaller dimensional constraint

$$Z_s = \begin{pmatrix} I & X^T \\ X & Y \end{pmatrix} \succeq 0, \quad Y_{21} = AX^T.$$

To develop the model, we introduce the following notation.

$$x := \mathrm{vec}\left(\mathrm{sblk}_{21}\begin{pmatrix} 0 & X^T \\ X & 0 \end{pmatrix}\right) = \sqrt{2}\mathrm{vec}\,(X), \quad y := \mathrm{svec}\,(Y),$$

where we add $\sqrt{2}$ to the definition of x since X appears together with X^T in Z_s and implicitly in $\bar{Y}$, with $Y_{21} = AX^T$. We define the following matrices and linear transformations:

$$\begin{aligned} \mathcal{Z}_s^x(x) &:= \mathrm{sBlk}_{21}(\mathrm{Mat}\,(x)), \quad \mathcal{Z}_s^y(y) := \mathrm{sBlk}_2(\mathrm{sMat}\,(y)), \\ \mathcal{Z}_s(x,y) &:= \mathcal{Z}_s^x(x) + \mathcal{Z}_s^y(y), \quad Z_s := \mathrm{sBlk}_1(I) + \mathcal{Z}_s(x,y), \end{aligned}$$

$$\begin{aligned} \mathcal{Y}^x(x) &:= \mathrm{sBlk}_{21}(A\mathrm{Mat}\,(x)^T), \quad \mathcal{Y}^y(y) := \mathrm{sBlk}_1(\mathrm{sMat}\,(y)) \\ \mathcal{Y}(x,y) &:= \mathcal{Y}^x(x) + \mathcal{Y}^y(y), \quad \bar{Y} := \mathrm{sBlk}_2(AA^T) + \mathcal{Y}(x,y). \end{aligned}$$

$$\begin{aligned} \bar{E} &:= W \circ \left[E - \mathcal{K}(\mathrm{sBlk}_2(AA^T))\right], \\ \bar{U} &:= H_u \circ \left[\mathcal{K}(\mathrm{sBlk}_2(AA^T)) - U\right], \\ \bar{L} &:= H_l \circ \left[L - \mathcal{K}(\mathrm{sBlk}_2(AA^T))\right]. \end{aligned}$$

The unknown matrix $\bar{Y}$ in (3.11) is equal to $\mathcal{Y}(x,y)$ with the additional constant in the 2, 2 block, i.e. our unknowns are the vectors x, y which are used to build $\bar{Y}$ and Z_s. Using this notation we can introduce the following relaxation of $\boldsymbol{EDMC}$ in (3.11).

$$
(\boldsymbol{EDMC-R}) \qquad \begin{array}{ll} \min & f_3(x,y) := \frac{1}{2}\|W\circ(\mathcal{K}(\mathcal{Y}(x,y))) - \bar{E}\|_F^2 \\ \text{subject to} & \begin{array}{rcl} g_u(x,y) := H_u \circ \mathcal{K}(\mathcal{Y}(x,y)) - \bar{U} & \leq & 0 \\ g_l(x,y) := \bar{L} - H_l \circ \mathcal{K}(\mathcal{Y}(x,y)) & \leq & 0 \\ \text{sBlk}_1(I) + \mathcal{Z}_s(x,y) & \succeq & 0. \end{array} \end{array} \tag{3.35}
$$

As above, we consider the functions g_u, g_l as implicitly acting only on the nonzero parts of the upper triangular part of the matrix that results from the Hadamard products with H_u, H_l, respectively.

Remark 3.3 *The constraint $Z_s \succeq 0$ in $\boldsymbol{EDMC-R}$ is equivalent to that used in e.g. [14, 4] and the references therein. However, the objective function is different, i.e. we use an ℓ_2 nearest matrix problem and a quadratic objective $\boldsymbol{SDP}$; whereas the papers in the literature generally use the ℓ_1 nearest matrix problem in order to obtain a linear $\boldsymbol{SDP}$.*

4 Duality for EDMC-R

We begin with the Lagrangian of $\boldsymbol{EDMC-R}$, problem (3.35),

$$
\begin{array}{rcl} L(x,y,\Lambda_u,\Lambda_l,\Lambda) & = & \frac{1}{2}\|W\circ\mathcal{K}(\mathcal{Y}(x,y)) - \bar{E}\|_F^2 + \left\langle \Lambda_u, H_u\circ\mathcal{K}(\mathcal{Y}(x,y)) - \bar{U}\right\rangle \\ & & + \left\langle \Lambda_l, \bar{L} - H_l\circ\mathcal{K}(\mathcal{Y}(x,y))\right\rangle - \left\langle \Lambda, \text{sBlk}_1(I) + \mathcal{Z}_s(x,y)\right\rangle, \end{array} \tag{4.36}
$$

where $0 \leq \Lambda_u, 0 \leq \Lambda_l \in \mathcal{S}^{m+n}$, and $0 \preceq \Lambda \in \mathcal{S}^{m+n}$. We partition the multiplier Λ as

$$
\Lambda = \begin{pmatrix} \Lambda_1 & \Lambda_{21}^T \\ \Lambda_{21} & \Lambda_2 \end{pmatrix}, \tag{4.37}
$$

corresponding to the blocks of Z_s. Recall that $x = \sqrt{2}\text{vec}\,(X),\ y := \text{svec}\,(Y)$. In addition, we denote

$$
\begin{array}{c} \lambda_u := \text{svec}\,(\Lambda_u), \quad \lambda_l := \text{svec}\,(\Lambda_l), \quad h_u := \text{svec}\,(H_u), \quad h_l := \text{svec}\,(H_l), \\ \lambda := \text{svec}\,(\Lambda), \quad \lambda_1 := \text{svec}\,(\Lambda_1), \quad \lambda_2 := \text{svec}\,(\Lambda_2), \quad \lambda_{21} := \text{vec}\,\text{sblk}_{21}(\Lambda). \end{array}
$$

And, for numerical implementation, we define the linear transformations

$$
h_u^{nz} = \text{svec}_u(H_u) \in \mathbb{R}^{nz_u}, \quad h_l^{nz} = \text{svec}_l(H_l) \in \mathbb{R}^{nz_l}, \tag{4.38}
$$

where h_u^{nz} is obtained from h_u by removing the zeros; thus, nz_u is the number of nonzeros in the upper-triangular part of H_u. Thus the indices are fixed from the given matrix H_u. Similarly, for h_l^{nz} with indices fixed from H_l. We then get the vectors

$$
\lambda_u^{nz} = \text{svec}_u(\Lambda_u) \in \mathbb{R}^{nz_u}, \quad \lambda_l^{nz} = \text{svec}_l(\Lambda_l) \in \mathbb{R}^{nz_l}.
$$

The adjoints are $\text{sMat}_u, \text{sMat}_l$; and, for any matrix M we get

$$
H_u \circ M = \text{sMat}_u\text{svec}_u(H_u \circ M).
$$

This holds similarly for $H_l \circ M$. Therefore, we could rewrite the Lagrangian as

$$
\begin{array}{rcl} L(x,y,\Lambda_u,\Lambda_l,\Lambda) & = & L(x,y,\lambda_u^{nz},\lambda_l^{nz},\Lambda) \\ & = & \frac{1}{2}\|W\circ\mathcal{K}(\mathcal{Y}(x,y)) - \bar{E}\|_F^2 + \left\langle \Lambda_u, H_u\circ\mathcal{K}(\mathcal{Y}(x,y)) - \bar{U}\right\rangle \\ & & \quad + \left\langle \Lambda_l, \bar{L} - H_l\circ\mathcal{K}(\mathcal{Y}(x,y))\right\rangle - \left\langle \Lambda, \text{sBlk}_1(I) + \mathcal{Z}_s(x,y)\right\rangle \\ & = & \frac{1}{2}\|W\circ\mathcal{K}(\mathcal{Y}(x,y)) - \bar{E}\|_F^2 + \left\langle \text{svec}_u(\Lambda_u), \text{svec}_u\left(H_u\circ\mathcal{K}(\mathcal{Y}(x,y)) - \bar{U}\right)\right\rangle \\ & & \quad + \left\langle \text{svec}_l(\Lambda_l), \text{svec}_l\left(\bar{L} - H_l\circ\mathcal{K}(\mathcal{Y}(x,y))\right)\right\rangle - \left\langle \Lambda, \text{sBlk}_1(I) + \mathcal{Z}_s(x,y)\right\rangle. \end{array} \tag{4.39}
$$

To simplify the dual of $\boldsymbol{EDMC-R}$, i.e. the max-min of the Lagrangian, we now find the stationarity conditions of the inner minimization problem, i.e. we take the derivatives of L with respect to x and y. We get

$$\begin{array}{rcl} 0 & = & \nabla_x L(x,y,\Lambda_u,\Lambda_l,\Lambda) \\ & = & [W\circ(\mathcal{K}\mathcal{Y}^x)]^*\left(W\circ\mathcal{K}(\mathcal{Y}(x,y))-\bar{E}\right)+[H_u\circ(\mathcal{K}\mathcal{Y}^x)]^*(\Lambda_u) \\ & & \qquad\qquad -[H_l\circ(\mathcal{K}\mathcal{Y}^x)]^*(\Lambda_l)-(\mathcal{Z}_s^x)^*(\Lambda). \end{array} \tag{4.40}$$

Note that both the Hadamard product and the transpose, $\cdot\circ\cdot, \cdot^T$, are self-adjoint linear transformations. Let H denote a weight matrix. Then

$$[H\circ(\mathcal{K}\mathcal{Y}^x)]^*(S)=(\mathcal{Y}^x)^*\mathcal{K}^*(H\circ(S))=\operatorname{vec}\left[\left\{\operatorname{sblk}_{21}\mathcal{K}^*(H\circ(S))\right\}^T A\right]. \tag{4.41}$$

Similarly,

$$\begin{array}{rcl} 0 & = & \nabla_y L(x,y,\Lambda_u,\Lambda_l,\Lambda) \\ & = & [W\circ(\mathcal{K}\mathcal{Y}^y)]^*\left(W\circ\mathcal{K}(\mathcal{Y}(x,y))-\bar{E}\right)+[H_u\circ(\mathcal{K}\mathcal{Y}^y)]^*(\Lambda_u) \\ & & \qquad\qquad -[H_l\circ(\mathcal{K}\mathcal{Y}^y)]^*(\Lambda_l)-(\mathcal{Z}_s^y)^*(\Lambda). \end{array} \tag{4.42}$$

And

$$[H\circ(\mathcal{K}\mathcal{Y}^y)]^*(S)=(\mathcal{Y}^y)^*\mathcal{K}^*(H\circ(S))=\operatorname{svec}\left\{\operatorname{sblk}_1\mathcal{K}^*(H\circ(S))\right\}. \tag{4.43}$$

Then $(\mathcal{Z}_s^x)^*(\Lambda)=\operatorname{vec}\operatorname{sblk}_{21}(\Lambda)=\lambda_{21}$. The first stationarity equation (4.40) yields

$$\begin{array}{rcl} (\mathcal{Z}_s^x)^*(\Lambda) & = & \lambda_{21} \\ & = & [W\circ(\mathcal{K}\mathcal{Y}^x)]^*\left(W\circ\mathcal{K}(\mathcal{Y}(x,y))-\bar{E}\right)+[H_u\circ(\mathcal{K}\mathcal{Y}^x)]^*(\Lambda_u)-[H_l\circ(\mathcal{K}\mathcal{Y}^x)]^*(\Lambda_l). \end{array} \tag{4.44}$$

Similarly, the second stationarity equation (4.42) yields

$$\begin{array}{rcl} (\mathcal{Z}_s^y)^*(\Lambda) & = & \lambda_2 \\ & = & [W\circ(\mathcal{K}\mathcal{Y}^y)]^*\left(W\circ\mathcal{K}(\mathcal{Y}(x,y))-\bar{E}\right)+[H_u\circ(\mathcal{K}\mathcal{Y}^y)]^*(\Lambda_u)-[H_l\circ(\mathcal{K}\mathcal{Y}^y)]^*(\Lambda_l). \end{array} \tag{4.45}$$

The Wolfe dual is obtained from applying the stationarity conditions to the inner minimization of the Lagrangian dual (max-min of the Lagrangian), i.e. we get the (dual $\boldsymbol{EDMC}$) problem:

$$(\boldsymbol{EDMC-D}) \quad \begin{array}{cl} \max & L(x,y,\lambda_u,\lambda_l,\lambda_1,\lambda_2,\lambda_{21}) \\ \text{subject to} & (4.44),(4.45) \\ & \operatorname{sMat}(\lambda_u)\geq 0, \operatorname{sMat}(\lambda_l)\geq 0 \\ & \operatorname{sBlk}_1\operatorname{sMat}(\lambda_1)+\operatorname{sBlk}_2\operatorname{sMat}(\lambda_2)+\operatorname{sBlk}_{21}\operatorname{Mat}(\lambda_{21})\succeq 0. \end{array} \tag{4.46}$$

We denote

$$\begin{array}{rcl} S_u & := & \bar{U}-H_u\circ(\mathcal{K}(\mathcal{Y}(x,y))), \quad s_u=\operatorname{svec}S_u \\ S_l & := & H_l\circ(\mathcal{K}(\mathcal{Y}(x,y)))-\bar{L}, \quad s_l=\operatorname{svec}S_l. \end{array} \tag{4.47}$$

We can now present the primal-dual characterization of optimality.

Theorem 4.1 *The primal-dual variables $x,y,\Lambda,\lambda_u,\lambda_l$ are optimal for $\boldsymbol{EDMC-R}$ if and only if:*

1. *Primal Feasibility:*
$$s_u \geq 0, \quad s_l \geq 0, \text{ in } (4.47),$$

$$Z_s = \mathrm{sBlk}_{\,1}(I) + \mathrm{sBlk}_{\,2}\mathrm{sMat}\,(y) + \mathrm{sBlk}_{\,21}\mathrm{Mat}\,(x) \succeq 0. \tag{4.48}$$

2. *Dual Feasibility: Stationarity equations* (4.44),(4.45) *hold and*
$$\Lambda = \mathrm{sBlk}_{\,1}\mathrm{sMat}\,(\lambda_1) + \mathrm{sBlk}_{\,2}\mathrm{sMat}\,(\lambda_2) + \mathrm{sBlk}_{\,21}\mathrm{Mat}\,(\lambda_{21}) \succeq 0; \lambda_u \geq 0; \lambda_l \geq 0. \tag{4.49}$$

3. *Complementary Slackness:*
$$\begin{aligned} \lambda_u \circ s_u &= 0 \\ \lambda_l \circ s_l &= 0 \\ \Lambda Z_s &= 0. \end{aligned}$$

■

We can use the structure of the optimality conditions to eliminate the linear dual equations and obtain a characterization of optimality based solely on a bilinear equation and nonnegativity/semidefiniteness.

Corollary 4.1 *The dual linear equality constraints* (4.44),(4.45) *in Theorem 4.1 can be eliminated after using them to substitute for* λ_2, λ_{21} *in* (4.49)*. The complementarity conditions now yield a bilinear system of equations*
$$F(x, y, \lambda_u, \lambda_l, \lambda_1) = 0,$$
with nonnegativity and semidefinite conditions that characterize optimality of $\boldsymbol{EDMC - R}$. ■

In addition, we can guarantee a strictly feasible primal starting point from the facial reduction in Theorem 3.1. This, along with a strictly feasible dual starting point are found using the heuristic in Section 5.2.

5 A Robust Primal-Dual Interior-Point Method

To solve $\boldsymbol{EDMC - R}$ we use the Gauss-Newton method on the perturbed complementary slackness conditions (written with the block vector notation):
$$F_\mu(x, y, \lambda_u, \lambda_l, \lambda_1) := \begin{pmatrix} \lambda_u \circ s_u - \mu_u e \\ \lambda_l \circ s_l - \mu_l e \\ \Lambda Z_s - \mu_c I \end{pmatrix} = 0, \tag{5.50}$$
where $s_u = s_u(x, y)$, $s_l = s_l(x, y)$, $\Lambda = \Lambda(\lambda_1, x, y, \lambda_u, \lambda_l)$, and $Z_s = Z_s(x, y)$. This is an overdetermined system with
$$(m_u + n_u) + (m_l + n_l) + (n + r)^2 \text{ equations}; \quad nr + t(n) + (m_u + n_u) + (m_l + n_l) + t(r) \text{ variables}.$$

5.1 Linearization

We denote the Gauss-Newton search direction for (5.50) by

$$\Delta s := \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta\lambda_u \\ \Delta\lambda_l \\ \Delta\lambda_1 \end{pmatrix}.$$

The linearized system for the search direction Δs is:

$$F'_\mu(\Delta s) \cong F'_\mu(x, y, \lambda_u, \lambda_l, \lambda_1)(\Delta s) = -F_\mu(x, y, \lambda_u, \lambda_l, \lambda_1).$$

To further simplify notation, we use the following composition of linear transformations. Let H be symmetric. Then

$$\begin{array}{rcl} \mathcal{K}_H^x(x) & := & H \circ (\mathcal{K}(\mathcal{Y}^x(x))), \\ \mathcal{K}_H^y(y) & := & H \circ (\mathcal{K}(\mathcal{Y}^y(y))), \\ \mathcal{K}_H(x,y) & := & H \circ (\mathcal{K}(\mathcal{Y}(x,y))). \end{array}$$

The linearization of the complementary slackness conditions results in three blocks of equations

1. $$\lambda_u \circ \operatorname{svec} \mathcal{K}_{H_u}(\Delta x, \Delta y) + s_u \circ \Delta\lambda_u = \mu_u e - \lambda_u \circ s_u$$

2. $$\lambda_l \circ \operatorname{svec} \mathcal{K}_{H_l}(\Delta x, \Delta y) + s_l \circ \Delta\lambda_l = \mu_l e - \lambda_l \circ s_l$$

3. $$\begin{array}{rl} & \Lambda \mathcal{Z}_s(\Delta x, \Delta y) + \left[\operatorname{sBlk}_1(\operatorname{sMat}(\Delta\lambda_1)) + \operatorname{sBlk}_2(\operatorname{sMat}(\Delta\lambda_2)) + \operatorname{sBlk}_{21}(\operatorname{sMat}(\Delta\lambda_{21}))\right] Z_s \\ = & \Lambda \mathcal{Z}_s(\Delta x, \Delta y) + [\operatorname{sBlk}_1(\operatorname{sMat}(\Delta\lambda_1)) \\ & + \operatorname{sBlk}_2\left(\operatorname{sMat}\left\{(\mathcal{K}_W^y)^* \mathcal{K}_W(\Delta x, \Delta y) + (\mathcal{K}_{H_u}^y)^*(\operatorname{sMat}(\Delta\lambda_u)) - (\mathcal{K}_{H_l}^y)^*(\operatorname{sMat}(\Delta\lambda_l))\right\}\right) \\ & + \operatorname{sBlk}_{21}\left(\operatorname{Mat}\left\{(\mathcal{K}_W^x)^* \mathcal{K}_W(\Delta x, \Delta y) + (\mathcal{K}_{H_u}^x)^*(\operatorname{sMat}(\Delta\lambda_u)) - (\mathcal{K}_{H_l}^x)^*(\operatorname{sMat}(\Delta\lambda_l))\right\}\right)\Big] Z_s \\ = & \mu_c I - \Lambda Z_s \end{array}$$

and hence

$$F'_\mu(\Delta s) = \begin{pmatrix} \lambda_u \circ \operatorname{svec} \mathcal{K}_{H_u}(\Delta x, \Delta y) + s_u \circ \Delta\lambda_u \\ \lambda_l \circ \operatorname{svec} \mathcal{K}_{H_l}(\Delta x, \Delta y) + s_l \circ \Delta\lambda_l \\ \Lambda \mathcal{Z}_s(\Delta x, \Delta y) + [\operatorname{sBlk}_1(\operatorname{sMat}(\Delta\lambda_1)) + \\ \operatorname{sBlk}_2\left(\operatorname{sMat}\left\{(\mathcal{K}_W^y)^* \mathcal{K}_W(\Delta x, \Delta y) + (\mathcal{K}_{H_u}^y)^*(\operatorname{sMat}(\Delta\lambda_u)) - (\mathcal{K}_{H_l}^y)^*(\operatorname{sMat}(\Delta\lambda_l))\right\}\right) + \\ \operatorname{sBlk}_{21}\left(\operatorname{Mat}\left\{(\mathcal{K}_W^x)^* \mathcal{K}_W(\Delta x, \Delta y) + (\mathcal{K}_{H_u}^x)^*(\operatorname{sMat}(\Delta\lambda_u)) - (\mathcal{K}_{H_l}^x)^*(\operatorname{sMat}(\Delta\lambda_l))\right\}\right)\Big] Z_s \end{pmatrix}$$

where $F'_\mu : \mathcal{M}^{r\times n} \times \Re^{t(n)} \times \Re^{t(m+n)} \times \Re^{t(m+n)} \times \Re^{t(m+n)} \to \Re^{t(m+n)} \times \Re^{t(m+n)} \times \Re^{(r+n)^2}$, i.e. the linear system is overdetermined.

We need to calculate the adjoint $(F'_\mu)^*$. We first find $\mathcal{Z}_s^*$, $(\mathcal{K}_H^x)^*$, $(\mathcal{K}_H^y)^*$, and $(\mathcal{K}_H)^*$. By the expression of $\mathcal{Z}_s$ we get

$$\mathcal{Z}_s^*(S) = \begin{pmatrix} (\mathcal{Z}_s^x)^*(S) \\ (\mathcal{Z}_s^y)^*(S) \end{pmatrix} = \begin{pmatrix} \mathrm{Mat}^*(\mathrm{sblk}_{21}(S)) \\ \mathrm{sMat}^*(\mathrm{sblk}_2(S)) \end{pmatrix} = \begin{pmatrix} \mathrm{vec}\,(\mathrm{sblk}_{21}(S)) \\ \mathrm{svec}\,(\mathrm{sblk}_2(S)) \end{pmatrix}. \tag{5.51}$$

By the expression of $\mathcal{Y}(\Delta x, \Delta y)$, we get

$$\mathcal{Y}^*(S) = \begin{pmatrix} (\mathcal{Y}^x)^*(S) \\ (\mathcal{Y}^y)^*(S) \end{pmatrix} = \begin{pmatrix} \mathrm{Mat}^*(\mathrm{sblk}_{21}(S)^T A) \\ \mathrm{sMat}^*(\mathrm{sblk}_1(S)) \end{pmatrix} = \begin{pmatrix} \mathrm{vec}\,(\mathrm{sblk}_{21}(S)^T A) \\ \mathrm{svec}\,(\mathrm{sblk}_1(S)) \end{pmatrix}. \tag{5.52}$$

By the expression of $\mathcal{K}_H(\Delta x, \Delta y)$, we get

$$\mathcal{K}_H^*(S) = \begin{pmatrix} (\mathcal{K}_H^x)^*(S) \\ (\mathcal{K}_H^y)^*(S) \end{pmatrix} = \begin{pmatrix} (\mathcal{Y}^x)^*(\mathcal{K}^*(H \circ S)) \\ (\mathcal{Y}^y)^*(\mathcal{K}^*(H \circ S)) \end{pmatrix}. \tag{5.53}$$

We have:

$$\begin{aligned} \langle \Lambda \mathcal{Z}_s(\Delta x, \Delta y), M \rangle &= \mathrm{trace}\,(M^T \Lambda \mathcal{Z}_s(\Delta x, \Delta y)) \\ &= \left\langle \tfrac{1}{2}\left(M^T\Lambda + \Lambda M\right), \mathcal{Z}_s(\Delta x, \Delta y)) \right\rangle \\ &= \left\langle \tfrac{1}{2}\mathcal{Z}_s^*((M^T\Lambda + \Lambda M)), \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right\rangle. \end{aligned} \tag{5.54}$$

The left half of the inner-product in the final row of (5.54) defines the required adjoint. Moreover,

$$\begin{aligned} \langle \mathrm{sBlk}_1\,(\mathrm{sMat}\,(\Delta\lambda_1))\, Z_s, M \rangle &= \mathrm{trace}\,\left(Z_s M^T\right) \mathrm{sBlk}_1\,(\mathrm{sMat}\,(\Delta\lambda_1)) \\ &= \left\langle \frac{1}{2}\left(Z_s M^T + M Z_s\right), \mathrm{sBlk}_1\,(\mathrm{sMat}\,(\Delta\lambda_1)) \right\rangle \\ &= \left\langle \frac{1}{2}\mathrm{svec}\,\left(\mathrm{sblk}_1\left(Z_s M^T + M Z_s\right)\right), \Delta\lambda_1 \right\rangle \end{aligned}$$

Similarly,

$$\begin{aligned} \left\langle \mathrm{sBlk}_2\left(\mathrm{sMat}\,\left\{(\mathcal{K}_H^y)^*\mathcal{K}_H(\Delta x, \Delta y)\right\}\right) Z_s, M \right\rangle &= \mathrm{trace}\, Z_s M^T \mathrm{sBlk}_2\left(\mathrm{sMat}\,\left\{(\mathcal{K}_H^y)^*\mathcal{K}_H(\Delta x, \Delta y)\right\}\right) \\ &= \left\langle \frac{1}{2}(Z_s M^T + M Z_s), \mathrm{sBlk}_2\left(\mathrm{sMat}\,((\mathcal{K}_H^y)^*\mathcal{K}_H(\Delta x, \Delta y))\right) \right\rangle \\ &= \left\langle \frac{1}{2}(\mathcal{K}_H)^*\mathcal{K}_H^y\left(\mathrm{svec}\,\left(\mathrm{sblk}_2\left(Z_s M^T + M Z_s\right)\right)\right), \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right\rangle \end{aligned}$$

and

$$\begin{aligned} \langle \mathrm{sBlk}_{21}\,(\mathrm{Mat}\,((\mathcal{K}_H^x)^*\mathcal{K}_H(\Delta x, \Delta y)))\, Z_s, M \rangle &= \mathrm{trace}\,\left(Z_s M^T\right) \mathrm{sBlk}_{21}\,(\mathrm{Mat}\,((\mathcal{K}_H^x)^*\mathcal{K}_H(\Delta x, \Delta y)))) \\ &= \left\langle \frac{1}{2}\left(Z_s M^T + M Z_s\right), \mathrm{sBlk}_{21}\,(\mathrm{Mat}\,((\mathcal{K}_H^x)^*\mathcal{K}_H(\Delta x, \Delta y))) \right\rangle \\ &= \left\langle \frac{1}{2}(\mathcal{K}_H)^*\mathcal{K}_H^x\left(\mathrm{vec}\,\left(\mathrm{sblk}_{21}(Z_s M^T + M Z_s)\right)\right), \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right\rangle \end{aligned}$$

again yielding the desired adjoints on the left of the last inner-products. Now we can evaluate $(F'_\mu)^*(w_1, w_2, W_3)$. This consists of three columns of blocks with five rows per column. We list this by columns C_1, C_2, C_3.

$$C_1 = \begin{pmatrix} (\mathcal{K}^x_{H_u})^* \left(\mathrm{sMat}\left(\lambda_u \circ w_1\right)\right) \\ (\mathcal{K}^y_{H_u})^* \left(\mathrm{sMat}\left(\lambda_u \circ w_1\right)\right) \\ w_1 \circ s_u \\ 0 \\ 0 \end{pmatrix}$$

$$C_2 = \begin{pmatrix} (\mathcal{K}^x_{H_l})^* \left(\mathrm{sMat}\left(\lambda_l \circ w_2\right)\right) \\ (\mathcal{K}^y_{H_l})^* \left(\mathrm{sMat}\left(\lambda_l \circ w_2\right)\right) \\ 0 \\ w_2 \circ s_l \\ 0 \end{pmatrix}$$

$$C_3 = \begin{pmatrix} C_{13} \\ C_{23} \\ C_{33} \\ C_{43} \\ C_{53} \end{pmatrix}$$

with elements of C_3

$$\begin{aligned}
C_{13} &= \frac{1}{2}(\mathcal{Z}^x_s)^*(W_3^T\Lambda + \Lambda^T W_3) + \frac{1}{2}(\mathcal{K}^x_W)^*\mathcal{K}^x_W \left(\mathrm{vec}\left(\mathrm{sblk}_{21}(Z_s W_3^T + W_3 Z_s)\right)\right) \\
&\quad + \frac{1}{2}(\mathcal{K}^x_W)^*\mathcal{K}^y_W \left(\mathrm{svec}\left(\mathrm{sblk}_2\left(Z_s W_3^T + W_3 Z_s\right)\right)\right) \\
C_{23} &= \frac{1}{2}(\mathcal{Z}^y_s)^*(W_3^T\Lambda + \Lambda^T W_3) + \frac{1}{2}(\mathcal{K}^y_W)^*\mathcal{K}^x_W \left(\mathrm{vec}\left(\mathrm{sblk}_{21}(Z_s W_3^T + W_3 Z_s)\right)\right) \\
&\quad + \frac{1}{2}(\mathcal{K}^y_W)^*\mathcal{K}^y_W \left(\mathrm{svec}\left(\mathrm{sblk}_2\left(Z_s W_3^T + W_3 Z_s\right)\right)\right) \\
C_{33} &= \frac{1}{2}\mathrm{svec}\left(\mathcal{K}^y_{H_u}\left(\mathrm{svec}\left(\mathrm{sblk}_2(Z_s W_3^T + W_3 Z_s)\right)\right)\right) + \frac{1}{2}\mathrm{svec}\left(\mathcal{K}^x_{H_u}\left(\mathrm{vec}\left(\mathrm{sblk}_{21}(Z_s W_3^T + W_3 Z_s)\right)\right)\right) \\
C_{43} &= -\frac{1}{2}\mathrm{svec}\left(\mathcal{K}^y_{H_l}\left(\mathrm{svec}\left(\mathrm{sblk}_2(Z_s W_3^T + W_3 Z_s)\right)\right)\right) - \frac{1}{2}\mathrm{svec}\left(\mathcal{K}^x_{H_l}\left(\mathrm{vec}\left(\mathrm{sblk}_{21}(Z_s W_3^T + W_3 Z_s)\right)\right)\right) \\
C_{53} &= \frac{1}{2}\mathrm{svec}\left(\mathrm{sblk}_1\left(Z_s W_3^T + W_3 Z_s\right)\right),
\end{aligned}$$

where $w_1 \in \Re^{t(m+n)}, w_2 \in \Re^{t(m+n)}, W_3 \in \mathcal{M}^{r+n}$. Thus the desired adjoint is given by $(F'_\mu)^* = C_1 + C_2 + C_3$.

5.2 Heuristic for an Initial Strictly Feasible Point

To compute an initial strictly feasible point, we begin by considering $x = 0$ and $y = \mathrm{svec}\,(Y)$, where $Y = \alpha I$. Then for any $\alpha > 0$ we have

$$Z_s = \begin{pmatrix} I & 0 \\ 0 & \alpha I \end{pmatrix} \succ 0 \text{ and } \mathcal{K}(\mathcal{Y}(x,y)) = \begin{pmatrix} 2\alpha(E-I) & \alpha E^T \\ \alpha E & 0 \end{pmatrix},$$

where E is a matrix of ones of appropriate size. Since we require for strict primal feasibility that $s_l^{nz} := \mathrm{svec}_l(S_l) = \mathrm{svec}_l(H_l \circ \mathcal{K}(\mathcal{Y}(x,y)) - \bar{L}) > 0$, it suffices to choose $\alpha > \bar{L}_{ij}$, for all ij such that $(H_l)_{ij} = 1$. In practice, we choose

$$\alpha = (2.1 + \epsilon) \max \left\{ 1, \max_{ij} |\bar{L}_{ij}| \right\},$$

where $0 \leq \epsilon \leq 1$ is a random number. For the initial strictly feasible dual variables, we choose $\Lambda_l = \frac{1}{100} S_l$ so that $\lambda_l^{nz} = \mathrm{svec}_l \Lambda_l > 0$. Defining λ_{21} and λ_2 according to equations (4.44) and (4.45), respectively, we find that $\Lambda_2 = \mathrm{sMat}\,(\lambda_2) \succ 0$. We would now like to choose $\Lambda_1 = \beta I$, for some $\beta > 0$, so that

$$\Lambda = \begin{pmatrix} \beta I & \Lambda_{21}^T \\ \Lambda_{21} & \Lambda_2 \end{pmatrix} \succ 0.$$

The last inequality holds if and only if $\Lambda_2 - \frac{1}{\beta} \Lambda_{21} \Lambda_{21}^T \succ 0$. We can therefore guarantee $\Lambda \succ 0$ by choosing

$$\beta = \frac{\mathrm{trace}\,(\Lambda_{21} \Lambda_{21}^T)}{\lambda_{\min}(\Lambda_2)}.$$

5.3 Diagonal Preconditioning

It has been shown in [11] that for a full rank matrix $A \in \mathcal{M}^{m \times n}$, $m \geq n$, and using the condition number of a positive definite matrix K, given by $\omega(K) := \frac{\mathrm{trace}\,(K)/n}{\det(K)^{1/n}}$, we conclude that the optimal diagonal scaling, i.e the solution of

$$\min \omega \left((AD)^T (AD) \right), \quad D \text{ positive diagonal matrix},$$

is given by $D = \mathrm{diag}\,(d_{ii} = 1/\|A_{:,i}\|)$. Therefore, we need to evaluate the columns of the linear transformation $F'_\mu(\cdot)$. The details are presented in Appendix D, page 41.

6 Instability in the Model

The use of the matrix Z_s is equivalent to the matrix of unknown variables used in the literature, e.g. in [4]. We include the facial reduction in Theorem 3.1 to emphasize that we are finding the smallest face containing the feasible set $\mathcal{F}_Z$ and obtaining an equivalent problem. Such a reduction allows for Slater's constraint qualification to hold, i.e. for a strictly feasible solution. It also guarantees strong duality, i.e. the existence of a nonempty (compact) set of Lagrange multipliers. This should result in stabilizing the program. However, as mentioned above, we still have instability due in part to the constant I in Z_s.

We now see that we have a conundrum; a stronger $\boldsymbol{SDP}$ relaxation, as seen by $Y - XX^T \cong 0$, results in increased ill-conditioning.

6.1 Case of an Accurate Relaxation

We first suppose that at optimality the relaxation $Z_s \succeq 0$ is a good approximation of the original constraint in (3.12), i.e. $\|\bar{Y} - PP^T\|_F < \epsilon$, ϵ small. Therefore, the norm of the first leading block is small as well $\|Y - XX^T\|_F < \epsilon$. Now

$$Z_s := \begin{pmatrix} I & X^T \\ X & Y \end{pmatrix} = \begin{pmatrix} I & 0 \\ -X & I \end{pmatrix}^{-1} \begin{pmatrix} I & 0 \\ 0 & Y - XX^T \end{pmatrix} \begin{pmatrix} I & -X^T \\ 0 & I \end{pmatrix}^{-1}.$$

Therefore, if $0 < \epsilon << 1$, then Z_s has r eigenvalues order 1, $O(1)$, and the rest are approximately 0, i.e. Z_s has numerical rank r. Therefore, for strict complementarity to hold, the dual variable Λ to the constraint $\mathrm{sBlk}_1(I) + \mathcal{Z}_s(x, y) \succeq 0$ in (3.35) must have numerical rank $n - r$. This also follows from Cauchy's interlacing property extended to Schur complements, see [17, Theorem 2.1, Page 49]

But, our numerical tests show that this is *not* the case, i.e. strict complementarity fails. In fact, when the noise factor is low and the approximation is good, we lose strict complementarity and have an ill-conditioned problem. Paradoxically, a good ***SDP*** approximation theoretically results in a poor approximation numerically and/or slow convergence to the optimum.

Note that if the approximation is good, i.e. $\|Y - XX^T\|_F$ is small, then, by interlacing of eigenvalues, this means that Z_s has many small eigenvalues and so has vectors that are numerically in the nullspace of Z_s, i.e. numerically $Z_s N = 0$, where $N = \begin{pmatrix} N_X \\ N_Y \end{pmatrix} \in \mathbb{R}^{n+r \times t}$ is a full column rank t matrix of (numerical) orthonormal eigenvectors in the nullspace. This implies that the linear transformation $\mathcal{N}_Z(X, Y) : \mathbb{R}^{nr+t(n)} \to \mathbb{R}^{nt}$

$$\mathcal{N}_Z(X, Y) = \begin{pmatrix} X & Y \end{pmatrix} \begin{pmatrix} N_X \\ N_Y \end{pmatrix} = XN_X + YN_Y = 0. \tag{6.55}$$

The adjoint is

$$\mathcal{N}_Z^*(M) = M \begin{pmatrix} N_X^T & N_Y^T \end{pmatrix}.$$

Since the nullity of $\mathcal{N}_Z^*$ is 0, we conclude that $\mathcal{N}_Z$ is onto and so we can eliminate nt variables using (6.55), i.e. nt is an estimate on the (numerical) nullity of the optimality conditions for the ***SDP*** relaxation.

The above arguments emphasize the fact that as we get closer to the optimum and Z_s gets closer to being singular, the ***SDP*** optimality conditions become more ill-conditioned. To correct this we use the strengthened relaxation in Section 6.4.2. Alternatively, we can deflate. Suppose that we detect that the smallest eigenvalue $0 < \lambda_{\min}(Z_s) << 1, Z_s v = \lambda_{\min}(Z_s)v, \|v\| = 1$ is close to zero. Let $P = \begin{pmatrix} v & V \end{pmatrix}$ be an orthogonal matrix. Then we can replace the complementary slackness ΛZ with $\bar{\Lambda}\bar{Z} = 0$, where $\bar{\Lambda} = V^T \Lambda V, \bar{Z} = V^T Z V$. This reduces the dimension of the symmetric matrix space by one and removes the singularity. We can recover X, Y at each iteration using P. This avoids the ill-conditioning as we can repeat this for each small eigenvalue of Z.

6.2 Linear Independence of Constraints and Redundancy in Objective Function

Suppose that the distances in the data matrix E are exact. Then our problem is a feasibility question, i.e. an ***EDM*** completion problem. We could add the constraints $W \circ (\mathcal{K}(\mathcal{Y}(x, y))) - \bar{E} = 0$ to our model problem. Since there are $rn + t(n)$ variables, we would get a unique solution if we had the same number of linear independent constraints. (This is pointed out in [4, Proposition 1].) But, as we now see, this can rarely happen. To determine which constraints are linearly independent we could use the following Lemma. Recall that $A^T e = 0$ and $\mathrm{rank}\,(A) = r < m < n$.

Lemma 6.1 *Let* $\begin{pmatrix} V & \frac{1}{\sqrt{n}}e \end{pmatrix}$ *be an* $n \times n$ *orthogonal matrix. Under the assumption that* $A^T e = 0$ *and* $\operatorname{rank}(A) = r < m < n$*, we have:*

1. *The nullspace* $\mathcal{N}(\mathcal{K}\mathcal{Y}) = 0$*; the range space* $\mathcal{R}(\mathcal{Y}^*\mathcal{K}^*) = \mathbb{R}^{rn+t(n)}$*;* $\operatorname{rank}(\mathcal{Y}^*\mathcal{K}^*) = rn + t(n)$*; and* $\dim \mathcal{R}(\mathcal{K}\mathcal{Y}) - \dim \mathcal{R}(\mathcal{Y}^*\mathcal{K}^*) = n(m-r-1) \geq 0$*, i.e. the corresponding linear system is either square or overdetermined and always full column rank.*

2. *Suppose that* $W \geq 0$ *is a given weight matrix with corresponding index sets* $\mathcal{N}_e, \mathcal{M}_e$ *for the positive elements. We let* $y = (y_k) = (y_{ij}) = svec(Y) \in \mathbb{R}^{t(n)}$*, where the indices* $k = ij$ *refer to both the position in* $\mathbb{R}^{t(n)}$ *as well as the position in the matrix* Y*. Similarly,* $x = (x_k) = (x_{ij}) = \sqrt{2}\operatorname{vec} X$*, where the indices refer to both the position in* $\mathbb{R}^{rn}$ *as well as the position in the matrix* X*.*

   (a) *The nullspace of vectors (x,y) is given by the span of the orthogonal unit vectors,*

   $$\begin{aligned} \mathcal{N}(W \circ \mathcal{K}\mathcal{Y}) \quad = \quad & \operatorname{span}\left[\{(0, e_k) : e_k = e_{ij}, i < j, ij \notin \mathcal{N}_e\}\right. \\ & \cup \{(0, e_k) : e_k = e_{ii}, it \notin \mathcal{M}_e, \forall t\} \\ & \left.\cup \{(e_k, 0) : e_k = e_{ij}, 1 \leq j \leq r, it \notin \mathcal{M}_e, \forall t\}\right]. \end{aligned}$$

   Therefore,

   $$\dim \mathcal{N}(W \circ \mathcal{K}\mathcal{Y}) = t(n-1) - |\mathcal{N}_e| + (r+1)|\{i : it \notin \mathcal{M}_e, \forall t\}|$$

   and

   $$\begin{aligned} \operatorname{rank}(W \circ \mathcal{K}\mathcal{Y}) &= [t(m+n-1) - t(m-1)] - t(n-1) + |\mathcal{N}_e| - (r+1)|\{i : it \notin \mathcal{M}_e, \forall t\}| \\ &= mn + |\mathcal{N}_e| - (r+1)|\{i : it \notin \mathcal{M}_e, \forall t\}|. \end{aligned}$$

   In the nontrivial case, $|\{i : it \notin \mathcal{M}_e, \forall t\}| = 0$*, we conclude that* $W \circ \mathcal{K}\mathcal{Y}$ *is full column rank if and only if*

   $$mn + |\mathcal{N}_e| \geq rn + t(n). \tag{6.56}$$

   (b) *The range space*

   $$\begin{aligned} \mathcal{R}(W \circ \mathcal{K}\mathcal{Y}) \quad = \quad & \operatorname{span}\left[\{\mathcal{K}\mathcal{Y}(0, e_k) : e_k = e_{ij}, i < j, ij \in \mathcal{N}_e\}\right. \\ & \cup \{\mathcal{K}\mathcal{Y}(0, e_k) : e_k = e_{ii}, it \in \mathcal{M}_e, \textit{ for some } t\} \\ & \left.\cup \{\mathcal{K}\mathcal{Y}(e_k, 0) : e_k = e_{ij}, 1 \leq j \leq r, it \in \mathcal{M}_e, \textit{ for some } t\}\right]. \end{aligned}$$

3. *Consider the blocked matrix* $W = \begin{pmatrix} W_1 & W_{21}^T \\ W_{21} & W_2 \end{pmatrix}$*. The nullspace*

   $$\mathcal{N}(\mathcal{Y}^*\mathcal{K}^*) = \left\{ W = \begin{pmatrix} W_1 & W_{21}^T \\ W_{21} & W_2 \end{pmatrix} : W_1 = \operatorname{Diag}(w),\ W_{21}^T \begin{pmatrix} A & e \end{pmatrix} = 0 \right\}.$$

Proof.

1. The nullspace of $\mathcal{K}\mathcal{Y}$ is determined using the details of the composite linear transformation given in the Appendix (B.74), i.e. with notation $\mathcal{K}(\mathcal{Y}(x,y)) = 0, Y = \operatorname{sMat}(y), \bar{y} = \operatorname{diag}(Y), X = \frac{1}{\sqrt{2}}\operatorname{vec} x$, we get $\mathcal{D}_e(\bar{y}) - 2Y = 0 \Rightarrow Y = 0 \Rightarrow AX^T = 0 \Rightarrow X = 0$. This implies that $\mathcal{R}(\mathcal{Y}^*\mathcal{K}^*) = \mathbb{R}^{rn+t(n)}$. The difference in the dimensions is calculated in the Appendix B.1.

2a. The unit vectors are found using the expression in (B.74). The nullspace is given by the span of these unit vectors since $-2XA^T + \bar{y}e^T = 0$ and $A^Te = 0$ implies that $\bar{y} = 0$.

2b. The range is found using the expression in (B.74).

3. Suppose that $\mathcal{Y}^*\mathcal{K}^*(W) = 0$. From (B.75), we get $W_{21}^T = 0$ and

$$\text{svec}\left(\text{Diag}\left(\begin{pmatrix} W_1 & W_{21}^T \end{pmatrix} e\right) - W_1\right) = 0.$$

This implies that $W_1 = \text{Diag}(w)$ is a diagonal matrix. We get

$$\text{svec}\left(\text{Diag}\left(w + W_{21}^T e\right) - \text{Diag}(w)\right) = 0,$$

and so $W_{21}^T e = 0$. ■

For given distances in E, we can now determine when a unique solution x, y exists.

Corollary 6.1 *Suppose that the assumption $A^Te = 0$, $\text{rank}(A) = r$ holds.*

1. *If the number of specified distances $n_e + m_e < nr + t(n)$, then $W \circ \mathcal{K}(\mathcal{Y}(x, y))$ has a nontrivial nullspace, i.e. the least squares objective function $f_3(x, y)$ is underdetermined.*

2. *If the number of anchors is one more than the embedding dimension, $m = r+1$, then, there is a unique solution x, y that satisfies $W \circ (\mathcal{K}(\mathcal{Y}(x, y) - \bar{E}) = 0$ if and only if all the exact distances in E are specified (known). If some of the distances are not specified, then $W \circ \mathcal{K}(\mathcal{Y}(x, y))$ has a nontrivial nullspace.*

■

Remark 6.1 *The above Corollary 6.1 states that if $m = r + 1$ and for* an arbitrary *number of sensors $n > 0$, we need to know all the distances in E exactly to guarantee that we have a unique solution x, y. Moreover, the least squares problem is* underdetermined *if some of the distances are not specified. This is also the case if the number of specified distances is less than $nr + t(n)$. This means that the least squares problem is* ill-conditioned. *We should replace the objective function with finding the least squares solution of minimum norm, i.e. minimize $\|(x, y)\|$ and use the current objective as a linear constraint, after projecting $\bar{E}$ appropriately. See e.g. [16, 12, 10] for related work on underdetermined least squares problems.*

6.3 Instability from Linearization

The original relaxation replaced $\bar{Y} = PP^T$ by $\bar{Y} \succeq PP^T$ or equivalently $XX^T - Y = 0$ is replaced by

$$G(x, y) = G(X, Y) = XX^T - Y \preceq 0,$$

where $x = \sqrt{2}\text{vec}\, X, y = \text{svec}(Y)$. In the formulation used in the literature, we replaced this by the equivalent linearized form $\mathcal{Z}_s(x, y) = \begin{pmatrix} I & X^T \\ X & Y \end{pmatrix} \succeq 0$. Though this is equivalent in theory, it is not equivalent numerically, since the dimensions of the range of G and that of range of $\mathcal{Z}_s$ are different, while the domains are the same.

Consider the simple case where W is the matrix of all ones and there are no upper or lower bound constraints. The Lagrangian is now

$$L(x,y,\Lambda) = \frac{1}{2}\|\mathcal{K}(\mathcal{Y}(x,y) - \bar{E}\|_F^2 + \langle \Lambda, G(x,y)\rangle\,, \tag{6.57}$$

where $0 \preceq \Lambda = \operatorname{svec}(\lambda) \in \mathcal{S}^n$. Note that

$$\begin{array}{rcl} \langle \Lambda, G(x,0)\rangle & = & \frac{1}{2}\langle \operatorname{Mat}(x), \Lambda \operatorname{Mat}(x)\rangle = \frac{1}{2}x^T \operatorname{vec}(\Lambda \operatorname{Mat}(x))\,; \\ \langle \Lambda, G(0,y)\rangle & = & \langle \operatorname{sMat}(y), \Lambda\rangle = y^T \operatorname{svec}(\Lambda). \end{array}$$

We now find the stationarity conditions of the inner minimization problem of the Lagrangian dual, i.e. we take the derivatives of L with respect to x and y. We get

$$\begin{array}{rcl} 0 & = & \nabla_x L(x,y,\Lambda) \\ & = & [(\mathcal{K}\mathcal{Y}^x)]^* \left((\mathcal{K}\mathcal{Y})(x,y)) - \bar{E}\right) + \operatorname{vec}(\Lambda \operatorname{Mat}(x))\,. \end{array} \tag{6.58}$$

Similarly,

$$\begin{array}{rcl} 0 & = & \nabla_y L(x,y,\Lambda) \\ & = & [(\mathcal{K}\mathcal{Y}^y)]^* \left((\mathcal{K}\mathcal{Y})(x,y)) - \bar{E}\right) - \operatorname{svec}(\Lambda). \end{array} \tag{6.59}$$

The Wolfe dual is obtained from applying the stationarity conditions to the inner minimization of the Lagrangian dual (max-min of the Lagrangian), i.e. we get the dual problem:

$$(\boldsymbol{EDMC} - \mathit{nonlin}) \quad \begin{array}{rl} \max & L(x,y,\lambda) \\ \text{subject to} & (6.58), (6.59) \\ & \operatorname{sMat}(\lambda) \succeq 0. \end{array} \tag{6.60}$$

Equivalently,

$$(\boldsymbol{EDMC} - \mathit{nonlin}) \quad \begin{array}{rl} \max & L(x,y,\lambda) \\ \text{subject to} & [(\mathcal{K}\mathcal{Y}^x)]^* \left((\mathcal{K}\mathcal{Y})(x,y)) - \bar{E}\right) + \operatorname{vec}(\Lambda \operatorname{Mat}(x)) = 0 \\ & \operatorname{sMat}\left([(\mathcal{K}\mathcal{Y}^y)]^* \left((\mathcal{K}\mathcal{Y})(x,y)) - \bar{E}\right)\right) \succeq 0. \end{array} \tag{6.61}$$

6.4 Alternative Models for Stability

Following are two approaches that avoid the ill-conditioning problems discussed above. (These are studied in a forthcoming research report.)

6.4.1 Minimum Norm Solutions for Underdetermined Systems

Underdetermined least squares problems suffer from ill-conditioning (as our numerics confirm for our problems). The usual approach, see e.g. [8], is to use the underdetermined system as a constraint and to find the solution of minimum norm. Therefore, we can follow the approach in [1] and change the nearest $\boldsymbol{EDM}$ problem into a minimum norm interpolation problem. For given data $\bar{E}$, and weight matrix W, we add the constraint $W \circ (\mathcal{K}\mathcal{Y}(x,y)) = W \circ \bar{E} = 0$.

$$(\boldsymbol{EDMC} - \boldsymbol{RC}) \quad \begin{array}{rrcl} \min & \multicolumn{3}{c}{\frac{1}{2}\|(x,y)\|^2} \\ \text{subject to} & W \circ (\mathcal{K}(\mathcal{Y}(x,y))) & = & W \circ \bar{E} \\ & \operatorname{sBlk}_1(I) + \mathcal{Z}_s(x,y) & \succeq & 0. \end{array} \tag{6.62}$$

For simplicity we have ignored the upper and lower bound constraints.

6.4.2 Strengthened Relaxation using Lagrangian Dual

A stronger relaxation is obtained if we take the Lagrangian relaxation of (3.35) but with the constraint $\bar{Y} = PP^T$ rather than the weaker $\bar{Y} \succeq PP^T$.

We partition

$$\bar{Y} = \begin{pmatrix} Y & Y_{21} \\ Y_{21}^T & Y_2 \end{pmatrix}.$$

$$(\boldsymbol{EDMC-R}) \qquad \begin{array}{rl} \min & \frac{1}{2}\|W \circ \mathcal{K}(Y) - \bar{E}\|_F^2 \\ \text{subject to} & H_u \circ (\mathcal{K}(Y)) - \bar{U} \leq 0 \\ & \bar{L} - H_l \circ (\mathcal{K}(Y)) \leq 0 \\ & Y = \begin{pmatrix} XX^T & XA^T \\ AX^T & AA^T \end{pmatrix}. \end{array} \tag{6.63}$$

The Lagrangian is quadratic and convex in Y but not convex in X.

$$\begin{array}{rcl} L(X, Y, \Lambda_u, \Lambda_l, \Lambda) & = & \frac{1}{2}\|W \circ \mathcal{K}(Y) - \bar{E}\|_F^2 \\ & & + \left\langle \Lambda_u, H_u \circ \mathcal{K}(Y) - \bar{U} \right\rangle \\ & & + \left\langle \Lambda_l, \bar{L} - H_l \circ \mathcal{K}(Y) \right\rangle \\ & & - \left\langle \Lambda, Y - \begin{pmatrix} XX^T & XA^T \\ AX^T & AA^T \end{pmatrix} \right\rangle. \end{array}$$

After homogenizing the Lagrangian, the optimality conditions for the Lagrangian dual include stationarity as well as a semidefiniteness condition, i.e. we get a semidefinite form for the Lagrangian dual.

As an illustration of the relative strength of the different relaxations, suppose that we have the simplified problem

$$\begin{array}{rll} \mu^* := & \min & f(P) \\ & \text{subject to} & PP^T = Y \\ & & P \in \mathcal{M}^{m \times n}. \end{array} \tag{6.64}$$

The two different relaxations for (6.64) are:

$$\begin{array}{rll} \mu_I^* := & \min & f(P) \\ & \text{subject to} & PP^T \preceq Y \\ & & P \in \mathcal{M}^{m \times n}; \end{array} \tag{6.65}$$

$$\mu_E^* := \max_{S \in \mathcal{S}^n} \min_P f(P) + \operatorname{trace}\left(S(PP^T - Y)\right). \tag{6.66}$$

We now show that if the relaxation (6.65) does not solve the original problem (6.64) exactly, then the Lagrangian relaxation (6.66) is a stronger relaxation.

Theorem 6.1 *Suppose that $Y \succeq 0$ is given and that f is a strictly convex coercive function of $P \in \mathcal{M}^{mn}$. Then μ_I^* in (6.65) is finite and attained at a single matrix P^*. Moreover, if*

$$Y - P^* P^{*T} \neq 0, \tag{6.67}$$

then $\mu^* \geq \mu_E^* > \mu_I^*$.

Proof. First, note that $g(P) := \operatorname{trace} S(PP^T - Y)$. Then the Hessian $\nabla^2 g(P) = I \otimes S \succeq 0$, for all $S \succeq 0$. (Here $\otimes$ denotes Kronecker product.) Therefore, $g(P)$ is a Löwner convex function, i.e. it is convex with respect to the Löwner partial order, $g(\lambda M_1 + (1-\lambda)M_2) \preceq \lambda g(M_1) + (1-\lambda) g(M_2)$, for all $0 \leq \lambda \leq 1$. Therefore, strong duality holds, i.e. This implies that the feasible set of (6.65) is a closed convex set.

Now, since $Y \succeq 0$, we get (6.65) is feasible, i.e. $\mu_I^* < \infty$. In addition, by coercivity of f and the fact that the feasible set in (6.65) is closed and convex, we conclude that $\mu_I^* > -\infty$ and μ_I^* is attained at a single matrix P^* since the problem is strictly convex.

Now let $Y = Q \begin{pmatrix} 0 & 0 \\ 0 & D_Y \end{pmatrix} Q^T$ be an orthogonal block diagonalization of Y, with $D_Y \succ 0$. Then $PP^T \preceq Y$ if and only if

$$(Q^T P)(Q^T P)^T \preceq \begin{pmatrix} 0 & 0 \\ 0 & D_Y \end{pmatrix}. \tag{6.68}$$

If we partition $W = Q^T P = \begin{pmatrix} W_1 & W_2 \\ W_3 & W_4 \end{pmatrix}$, then

$$(Q^T P)(Q^T P)^T = \begin{pmatrix} W_1 W_1^T + W_2 W_2^T & \cdots \\ \cdots & \cdots \end{pmatrix}$$

and (6.68) implies that both $W_1 = 0, W_2 = 0$, i.e. without loss of generality we can assume that $P = \begin{pmatrix} 0 & 0 \\ P_3 & P_4 \end{pmatrix}$. After substitution into the objective function f, we can discard the zeros in P and we can assume, without loss of generality, that Slater's condition holds in (6.65) respect to the Löwner partial order, i.e.

$$\mu_E^* = \max_{S \succeq 0} \min_P f(P) + \operatorname{trace}(S(PP^T - Y)) = \min_P f(P) + \operatorname{trace}(S^*(PP^T - Y)),$$

for some $S^* \succeq 0$.

Now, by our assumption, $0 \preceq Q := Y - P^* P^{*T} \neq 0$ and

$$\begin{array}{rcl} \mu_I^* & = & f(P^*) \\ & = & \max_{S \succeq 0} \min_P f(P) + \operatorname{trace}(S(PP^T - Y)) \\ & = & \max_{S \succeq 0} \phi(S) \\ & = & \phi(S^*) = \min_P f(P) + \operatorname{trace}(S^*(PP^T - Y)) \\ & = & f(P^*) + \operatorname{trace}(S^*(P^* P^{*T} - Y)). \end{array}$$

Now consider the concave function $\phi(S) = \min_P f(P) + \operatorname{trace}(S(PP^T - Y))$. We have $\phi(S^*) = \mu_I^*$. Denote by P_S the set of matrices such that $P_S = \arg\min f(P) + \operatorname{trace}(S^*(PP^T - Y)$ The directional derivative of $\phi(S)$ at any matrix S along a direction T is given by [9, Page 6]

$$\begin{array}{rcl} D_T \phi(S) & = & \min_{P_S} \left\langle T, \frac{\partial}{\partial S}(f(P_S) + \operatorname{trace}(S(P_S {P_S}^T - Y)) \right\rangle \\ & = & \min_{P_S} \left\langle T, (P_S {P_S}^T - Y) \right\rangle. \end{array}$$

Thus if we consider $T = -Q$ and $S = S^*$, then $P_{S^*} = P^*$, and

$$D_{-Q}\phi(S^*) =< -Q, -Q >$$

and hence the directional derivative along the direction $-Q$ is positive. Thus if we consider the matrix $\hat{S} = S^* - \epsilon Q$, for $\epsilon > 0$ sufficiently small, we get $\phi(\hat{S}) > \phi(S^*) = \mu_I^*$, which implies $\mu_E^* > \mu_I^*$. ■

A numerical study for the strengthened relaxation is in progress.

7 Numerical Tests

We now present results on randomly generated problems with connected underlying graphs. The tests were done using MATLAB 7.1. These tests illustrate the robustness of the algorithm and the relaxation, i.e. high accuracy solutions can be obtained.

Table 1: Problems with increasing noise factor are presented in Table 1, page 31. The densities for the weight matrices in the objective and lower bound matrix in the constraint are: $W.75, L.8$, respectively. The dimensions of the problem are: $n = 15, m = 5, r = 2$, respectively. The results illustrate the robustness of the algorithm, i.e. though both linear transformations corresponding to $\mathcal{Z}_s$ and the Jacobian F'_μ have many zero singular values (last two columns of the table), the optimal value and norm of the error in the relaxation, $\frac{\|PP^T - \bar{Y}\|}{\|PP^T}$, are the same order as the noise factor. Our tests showed that strict complementarity fails. In fact, the dual multiplier Λ is often 0. Though failure of strict complementarity indicates singularity in the Jacobian of the optimality conditions, the shadow price interpretation of a small dual multiplier Λ indicates stability in the objective value under small perturbations in the data. This appears to be one of the good effects from using the ℓ_2 norm.

Table 2: In Table 2, page 32, we illustrate the typical convergence rate as well as the high accuracy in the solution. We include the steplength and the centering parameter σ. The values for n, m, r are $20, 6, 2$, respectively. The system of optimality conditions is underdetermined, i.e. the number of known distances is less than the number of variables: $182 < 250$. The relative error in the relaxation is small: $\frac{\|Ybar - PP'\|}{\|PP'\|} = 2.2985e-007$, $\frac{\|Y - XX'\|}{\|XX'\|} = 2.8256e-007$. Strict complementarity fails, but we do get high accuracy in the solution, i.e. we get the 10 decimals accuracy in the duality gap that is set in the algorithm. The crossover to the affine scaling direction with step length 1 and no backtracking is started after 2 decimal accuracy is obtained. However, we do not see quadratic convergence after the crossover, most likely due to the lack of strict complementarity.

Table 3: In Table 3, page 33, we start with density .999 for both W, L and decrease to density .14495. We can see that the quality of the relaxation decreases once the density decreases below .5. The decrease is from order .00001 to order 1.5. The number of singular values in F'_μ increases, while the number of zero eigenvalues in Z_s decreases.

The progress of one random problem is illustrated in Figures 1 to Figure 7.

- **Figure 1**, page 34, shows the original positions of the anchors as stars and the sensors as dots labelled with R#. The anchors and sensors are generated randomly. Once they are fixed, then the distances within radio range are calculated. Only the distances and positions of the anchors are used by the program.

- **Figure 2**, page 35, shows the positions found by the algorithm after one iteration, i.e. the points marked N# indicate the corresponding sensor that is searching for R#. The original positions for N# are all clustered together, since $x = 0$ is used as the initial solution. The value $-log_{10}$(relative duality gap) is $-4.69e-01$.

nf	optvalue	relaxation	cond.number F'_μ	sv($\mathcal{Z}_s$)	sv(F'_μ)
0.0000e+000	3.9909e-009	1.1248e-005	3.8547e+006	15	19
5.0000e-002	7.5156e-004	4.4637e-002	1.0244e+011	6	27
1.0000e-001	3.7103e-003	1.1286e-001	1.9989e+010	5	25
1.5000e-001	6.2623e-003	1.3125e-001	1.0065e+010	6	14
2.0000e-001	1.3735e-002	1.3073e-001	6.8833e+009	7	12
2.5000e-001	2.3426e-002	2.4828e-001	2.4823e+010	8	6
3.0000e-001	6.0509e-002	2.3677e-001	3.4795e+010	7	7
3.5000e-001	5.5367e-002	3.7260e-001	2.3340e+008	6	4
4.0000e-001	7.6703e-002	3.6343e-001	8.9745e+010	8	3
4.5000e-001	1.2493e-001	6.9625e-001	3.2590e+010	6	9
5.0000e-001	1.3913e-001	3.9052e-001	2.2870e+005	8	0
5.5000e-001	8.8552e-002	3.8742e-001	5.8879e+007	8	2
6.0000e-001	4.2425e-001	4.1399e-001	4.9251e+012	8	4
6.5000e-001	2.0414e-001	6.6054e-001	2.4221e+010	7	4
7.0000e-001	1.2028e-001	3.4328e-001	1.9402e+010	7	6
7.5000e-001	2.6590e-001	7.9316e-001	1.3643e+011	7	4
8.0000e-001	4.7155e-001	3.7822e-001	6.6910e+009	8	2
8.5000e-001	1.8951e-001	5.8652e-001	1.4185e+011	6	7
9.0000e-001	2.1741e-001	9.8757e-001	2.9077e+005	8	0
9.5000e-001	4.4698e-001	4.6648e-001	2.7013e+006	9	2

Table 1: Increasing Noise: density$W = .75$, density$L = .8, n = 15, m = 5, r = 2$

noiter	-log10(relgapp)	step	σ	optval
1	-4.15e-001	7.6000e-001	1.0e+000	1.9185e+003
2	-4.38e-001	9.5000e-001	7.7e-001	1.0141e+003
3	-4.52e-001	7.6000e-001	5.7e-001	4.3623e+002
4	-4.69e-001	9.5000e-001	7.7e-001	2.2990e+002
5	-4.52e-001	7.6000e-001	5.7e-001	1.0327e+002
6	-4.70e-001	9.5000e-001	7.7e-001	5.3836e+001
7	-4.58e-001	6.0800e-001	5.7e-001	2.8919e+001
8	-4.74e-001	9.5000e-001	8.2e-001	1.5349e+001
9	-4.29e-001	7.6000e-001	5.7e-001	6.6988e+000
10	-4.11e-001	9.5000e-001	7.7e-001	3.3713e+000
11	-2.02e-001	9.5000e-001	5.7e-001	1.1047e+000
12	1.11e-001	9.5000e-001	5.7e-001	3.5139e-001
13	5.53e-001	9.8085e-001	5.7e-001	1.0648e-001
14	1.06e+000	9.9503e-001	5.6e-001	3.1561e-002
15	1.60e+000	9.9875e-001	5.5e-001	9.3173e-003
16	2.13e+000	9.9967e-001	5.5e-001	2.7479e-003
	CROSSOVER			
17	2.66e+000	1.0000e+000	0.0e+000	8.0958e-004
18	3.39e+000	1.0000e+000	0.0e+000	2.0202e-004
19	4.00e+000	1.0000e+000	0.0e+000	5.0503e-005
20	4.60e+000	1.0000e+000	0.0e+000	1.2626e-005
21	5.20e+000	1.0000e+000	0.0e+000	3.1564e-006
22	5.80e+000	1.0000e+000	0.0e+000	7.8911e-007
23	6.40e+000	1.0000e+000	0.0e+000	1.9728e-007
24	7.01e+000	1.0000e+000	0.0e+000	4.9319e-008
25	7.61e+000	1.0000e+000	0.0e+000	1.2330e-008
26	8.21e+000	1.0000e+000	0.0e+000	3.0825e-009
27	8.81e+000	1.0000e+000	0.0e+000	7.7062e-010
28	9.41e+000	1.0000e+000	0.0e+000	1.9265e-010
29	1.00e+001	1.0000e+000	0.0e+000	4.8163e-011
30	1.06e+001	1.0000e+000	0.0e+000	1.2041e-011

Table 2: Convergence Rate: density$W = .80$, density$L = .8, n = 20, m = 6, r = 2$

densityW	optvalue	$\|Y - XX\| / \|XX^T\|$	$\|\bar{Y} - PP^T\| / \|PP^T\|$	sv(Z_s)	sv(F'_μ)
9.9900e-001	9.4954e-009	2.8991e-005	2.3742e-005	15	6
9.5405e-001	5.1134e-009	1.5913e-005	1.3441e-005	15	10
9.0910e-001	5.9131e-009	2.0394e-005	1.8625e-005	15	12
8.6415e-001	3.9076e-009	1.4313e-005	1.0464e-005	15	18
8.1920e-001	4.0578e-009	2.7828e-005	1.7884e-005	15	18
7.7425e-001	6.8738e-009	4.0661e-005	2.7265e-005	15	15
7.2930e-001	3.6859e-009	8.4506e-006	6.1864e-006	15	21
6.8435e-001	4.6248e-009	1.9904e-005	1.3379e-005	15	22
6.3940e-001	4.6809e-009	2.0534e-005	1.5896e-005	15	29
5.9445e-001	9.6305e-009	3.3416e-005	2.6535e-005	15	18
5.4950e-001	8.2301e-010	1.0282e-005	8.8234e-006	15	91
5.0455e-001	6.3663e-009	1.2188e-004	1.0635e-004	13	51
4.5960e-001	6.4292e-010	4.6191e-004	3.7895e-004	12	95
4.1465e-001	3.0968e-009	7.4456e-005	6.1611e-005	14	54
3.6970e-001	9.7092e-010	3.5303e-004	2.8009e-004	14	98
3.2475e-001	5.5651e-011	2.7590e-002	2.2310e-002	7	101
2.7980e-001	1.3733e-015	4.7867e-001	3.6324e-001	5	104
2.3485e-001	9.4798e-012	2.0052e+000	1.1612e+000	6	110
1.8990e-001	2.3922e-010	7.5653e-001	6.2430e-001	4	119
1.4495e-001	3.2174e-029	2.5984e+000	1.7237e+000	0	122

Table 3: Decreasing density W and L: $n = 15, m = 5, r = 2$

- **Figures 3,4,5,6,7**, on pages 36,37,38,39, 40, respectively, show the positions found after 2,8,12,22,29 iterations, respectively, with values for $-log_{10}$(relative duality gap): $-.391, -.405, 3.58, 9.02$, respectively. We see that the sensors are almost found after 8 iterations and essentially found exactly after 12 iterations.

The crossover occurred after iteration 20. Though the density of W was low (approx. .3), the quality of the relaxation was excellent, i.e. correct to approximately 7 digits.

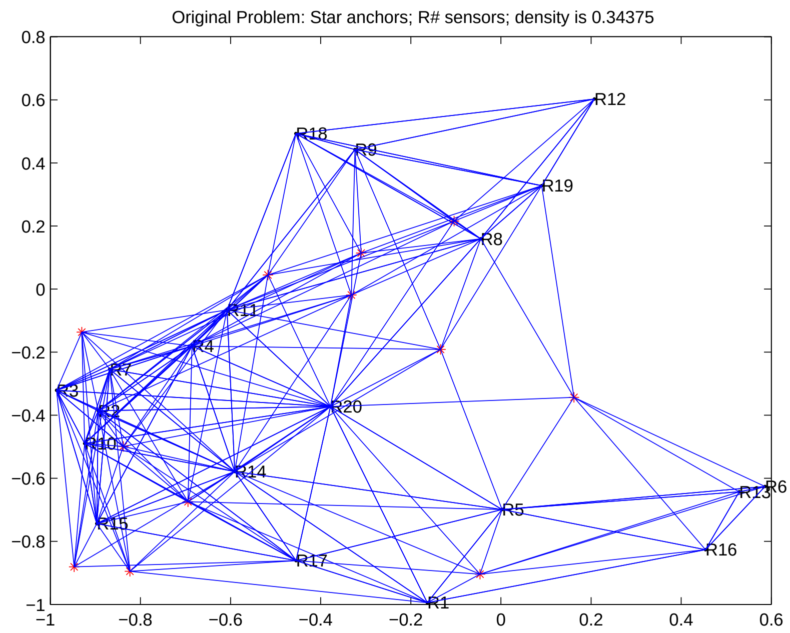

Figure 1: Sample random problem with 20 sensors, 12 anchors

8 Conclusion

In this paper we studied a robust algorithm for a semidefinite relaxation of the sensor localization problem with anchors. We showed that the linearization step used in the literature on the ***SDP*** relaxation has several problems with instability. But, our algorithm was able to obtain high accuracy solutions. In addition, contrary to other approaches in the literature, we solve a nearest Euclidean Matrix problem with the Frobenius norm rather than the ℓ_1 norm.

Our numerical tests show that the robust algorithm obtains high accuracy solutions for the ***SDP*** relaxation; and the relaxation yields surprisingly strong approximations to the original sensor localization problem. In fact, when the underlying graph is connected, it is difficult to find examples where the original true sensor positions are not found. These tests are preliminary and the algorithm is currently being modified to handle large instances.

We also discussed stronger and more robust relaxations. Numerical tests with these are in progress.

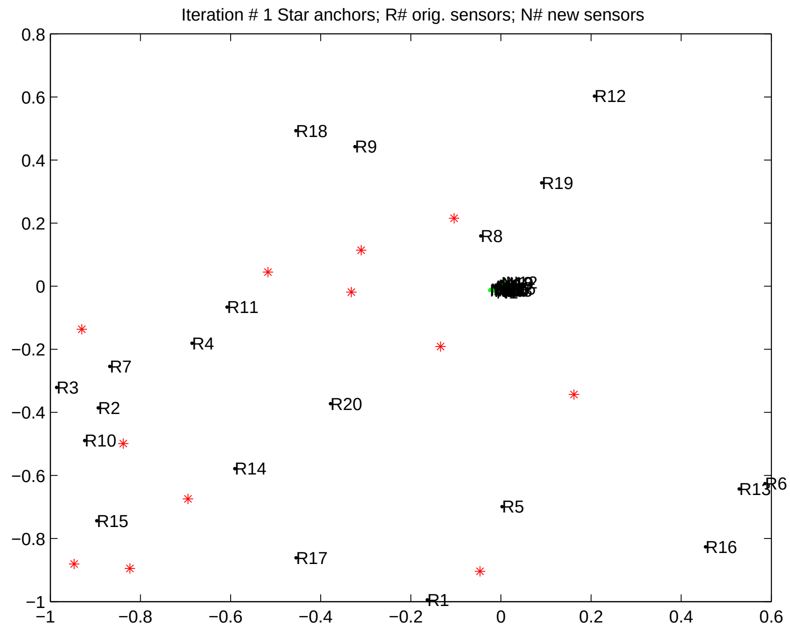

Figure 2: After one iteration for randomly generated problem

A Linear Transformations and Adjoints

$$\text{From (3.14)}: \quad Z := \begin{pmatrix} I & P^T \\ P & \bar{Y} \end{pmatrix}, \quad P = \begin{pmatrix} X \\ A \end{pmatrix}, \quad \bar{Y} := \begin{pmatrix} Y & Y_{21}^T \\ Y_{21} & Y_2 \end{pmatrix},$$

$$Z_s = \begin{pmatrix} I & X^T \\ X & Y \end{pmatrix} \succeq 0, \quad Y_{21} = AX^T, \quad Y_2 = AA^T.$$

$$x := \sqrt{2}\mathrm{vec}\,(X) \quad y := \mathrm{svec}\,(Y),$$

$$\mathcal{D}_e(B) \;:=\; \mathrm{diag}\,(B)\,e^T + e\,\mathrm{diag}\,(B)^T, \qquad \mathcal{K}(B) \;:=\; \mathcal{D}_e(B) - 2B, \tag{A.69}$$

$$\mathcal{D}_e^*(D) = 2\mathrm{Diag}\,(De), \qquad \mathcal{K}^*(D) = 2(\mathrm{Diag}\,(De) - D) \tag{A.70}$$

$$\begin{array}{rr} \mathcal{Z}_s^x(x) := \mathrm{sBlk}_{\,21}(\mathrm{Mat}\,(x)), & \mathcal{Z}_s^y(y) := \mathrm{sBlk}_{\,2}(\mathrm{sMat}\,(y)), \\ \mathcal{Z}_s(x,y) := \mathcal{Z}_s^x(x) + \mathcal{Z}_s^y(y), & Z_s := \mathrm{sBlk}_{\,1}(I) + \mathcal{Z}_s(x,y), \end{array}$$

$$\begin{array}{rr} \mathcal{Y}^x(x) := \mathrm{sBlk}_{\,21}(A\mathrm{Mat}\,(x)^T), & \mathcal{Y}^y(y) := \mathrm{sBlk}_{\,1}(\mathrm{sMat}\,(y)) \\ \mathcal{Y}(x,y) := \mathcal{Y}^x(x) + \mathcal{Y}^y(y), & \bar{Y} := \mathrm{sBlk}_{\,2}(AA^T) + \mathcal{Y}(x,y). \end{array}$$

$$\mathcal{Z}_s^*(S) = \begin{pmatrix} (\mathcal{Z}_s^x)^*(S) \\ (\mathcal{Z}_s^y)^*(S) \end{pmatrix} \begin{pmatrix} \mathrm{Mat}^*(\mathrm{sblk}_{\,21}(S)) \\ \mathrm{sMat}^*(\mathrm{sblk}_{\,2}(S)) \end{pmatrix} = \begin{pmatrix} \mathrm{vec}\,(S_{21}) \\ \mathrm{svec}\,(S_2) \end{pmatrix} \tag{A.71}$$

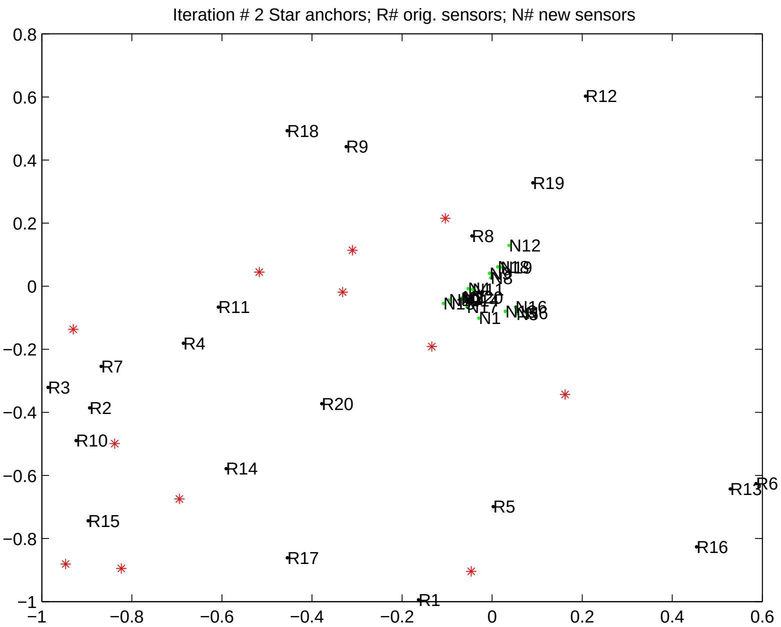

Figure 3: After two iterations for randomly generated problem

$$\mathcal{Y}^*(S) = \begin{pmatrix} (\mathcal{Y}^x)^*(S) \\ (\mathcal{Y}^y)^*(S) \end{pmatrix} = \begin{pmatrix} \mathrm{Mat}^*(\sqrt{2}\mathrm{sblk}_{21}(S)^T A) \\ \mathrm{sMat}^*(\mathrm{sblk}_1(S)) \end{pmatrix} = \begin{pmatrix} \mathrm{vec}\,(\sqrt{2} S_{21}^T A) \\ \mathrm{svec}\,(S_1) \end{pmatrix} \tag{A.72}$$

$$\begin{array}{rcl} \mathcal{K}_H^x(x) & := & H \circ (\mathcal{K}(\mathcal{Y}^x(x))), \\ \mathcal{K}_H^y(y) & := & H \circ (\mathcal{K}(\mathcal{Y}^y(y))), \\ \mathcal{K}_H(x,y) & := & H \circ (\mathcal{K}(\mathcal{Y}(x,y))). \end{array}$$

$$\mathcal{K}_H^*(S) = \begin{pmatrix} (\mathcal{K}_H^x)^*(S) \\ (\mathcal{K}_H^y)^*(S) \end{pmatrix} = \begin{pmatrix} (\mathcal{Y}^x)^*(\mathcal{K}^*(H \circ S)) \\ (\mathcal{Y}^y)^*(\mathcal{K}^*(H \circ S)) \end{pmatrix} \tag{A.73}$$

$$\begin{array}{c} \bar{E} := W \circ \left[E - \mathcal{K}(\mathrm{sBlk}_2(AA^T)) \right], \\ \bar{U} := H_u \circ \left[\mathcal{K}(\mathrm{sBlk}_2(AA^T)) - U \right], \\ \bar{L} := H_l \circ \left[L - \mathcal{K}(\mathrm{sBlk}_2(AA^T)) \right]. \end{array}$$

$$\Lambda = \begin{pmatrix} \Lambda_1 & \Lambda_{21}^T \\ \Lambda_{21} & \Lambda_2 \end{pmatrix},$$

$$\begin{array}{c} \lambda_u := \mathrm{svec}\,(\Lambda_u), \quad \lambda_l := \mathrm{svec}\,(\Lambda_l), \quad h_u := \mathrm{svec}\,(H_u), \quad h_l := \mathrm{svec}\,(H_l), \\ \lambda := \mathrm{svec}\,(\Lambda), \quad \lambda_1 := \mathrm{svec}\,(\Lambda_1), \quad \lambda_2 := \mathrm{svec}\,(\Lambda_2), \quad \lambda_{21} := \mathrm{vec}\,\mathrm{sblk}_{21}(\Lambda). \end{array}$$

For a given symmetric matrix H:

$$\begin{array}{rcl} \mathcal{K}_H^x(x) & := & H \circ (\mathcal{K}(\mathcal{Y}^x(x))), \\ \mathcal{K}_H^y(y) & := & H \circ (\mathcal{K}(\mathcal{Y}^y(y))), \\ \mathcal{K}_H(x,y) & := & H \circ (\mathcal{K}(\mathcal{Y}(x,y))). \end{array}$$

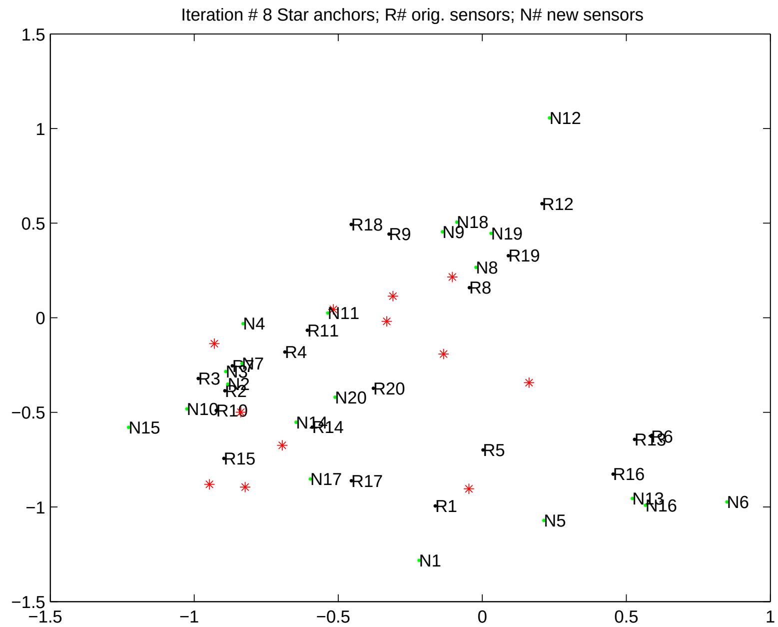

Figure 4: After eight iterations for randomly generated problem

B Composition of Transformations and Properties

B.1 $\mathcal{KY} : \mathbb{R}^{rn+t(n)} \to \mathbb{R}^{t(m+n)-n-t(m)}$

Note that

$$
\begin{aligned}
2\left(rn + t(n) - (t(m+n) - n - t(m))\right) &= 2rn + n^2 + n - (m+n)(m+n+1) + 2n + m(m+1) \\
&= 2rn + 2n - 2mn \\
&= 2n(r+1-m),
\end{aligned}
$$

i.e. $\mathcal{KY}$ has nontrivial nullspace if $r+1-m>0$. Now, recall that $x = \sqrt{2}\mathrm{vec}\,(X), y = \mathrm{svec}\,(Y)$. Let $\bar{y} = \mathrm{diag}\,(Y)$.

$$
\begin{aligned}
\mathcal{KY}(x,y) &= \mathcal{K}\left(\mathrm{sBlk}_{21}(A\mathrm{Mat}\,(x)^T) + \mathrm{sBlk}_1(\mathrm{sMat}\,(y))\right) \\
&= \mathcal{K}\begin{pmatrix} Y & XA^T \\ AX^T & 0 \end{pmatrix} \\
&= \mathcal{D}_e\begin{pmatrix} \bar{y} \\ 0 \end{pmatrix} - 2\begin{pmatrix} Y & XA^T \\ AX^T & 0 \end{pmatrix} \\
&= \begin{pmatrix} \mathcal{D}_e(\bar{y}) - 2Y & -2XA^T + \bar{y}e^T \\ -2AX^T + e\bar{y}^T & 0 \end{pmatrix}.
\end{aligned}
\tag{B.74}
$$

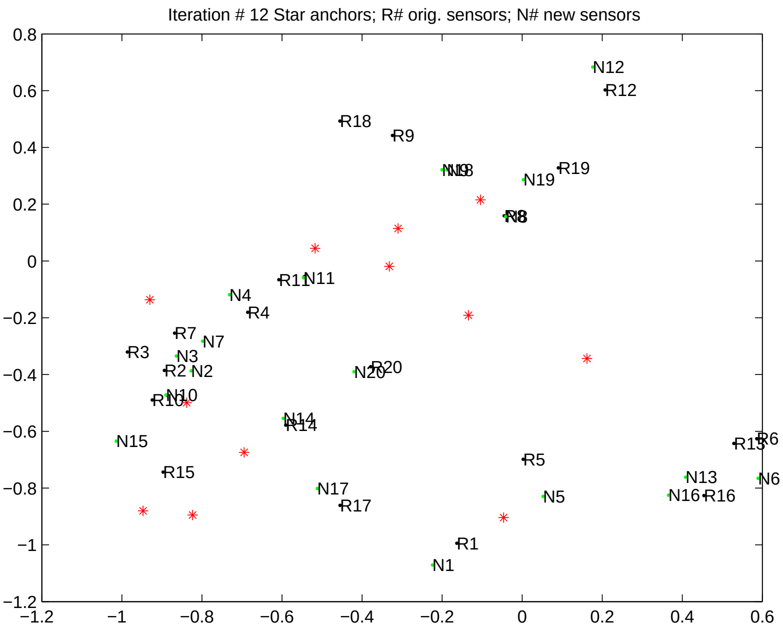

Figure 5: After twelve iterations for randomly generated problem

B.2 $\mathcal{Y}^*\mathcal{K}^*$

Consider the blocked matrix $W = \begin{pmatrix} W_1 & W_{21}^T \\ W_{21} & W_2 \end{pmatrix}$. Let $w = We = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$. Then

$$
\begin{aligned}
\mathcal{Y}^*\mathcal{K}^*(W) &= \mathcal{Y}^* 2 \left(\operatorname{Diag}(We) - W\right) \\
&= \mathcal{Y}^* 2 \left(\operatorname{Diag}\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} - W\right) \\
&= \mathcal{Y}^* 2 \begin{pmatrix} \operatorname{Diag}(w_1) - W_1 & W_{21}^T \\ W_{21} & \operatorname{Diag}(w_2) - W_2 \end{pmatrix} \\
&= 2 \begin{pmatrix} \operatorname{vec}\left(\sqrt{2} W_{21}^T A\right) \\ \operatorname{svec}\left(\operatorname{Diag}(w_1) - W_1\right) \end{pmatrix}.
\end{aligned} \tag{B.75}
$$

B.3 $(\mathcal{K}_W)^*\mathcal{K}_W$

We consider the linear transformation

$$
(\mathcal{K}_W)^*\mathcal{K}_W(x,y) = \begin{pmatrix} (\mathcal{K}_W^x)^*\mathcal{K}_W(x,y) \\ (\mathcal{K}_W^y)^*\mathcal{K}_W(x,y) \end{pmatrix}.
$$

We have that

$$
\mathcal{K}_W(x,y) = W \circ \begin{pmatrix} \mathcal{K}(Y) & \begin{pmatrix} \bar{y}_1 e_m^T \\ \vdots \\ \bar{y}_n e_m^T \end{pmatrix} - 2XA^T \\ \begin{pmatrix} \bar{y}_1 e_m & \cdots & \bar{y}_n e_m \end{pmatrix} - 2AX^T & 0_{m\times m} \end{pmatrix}
$$

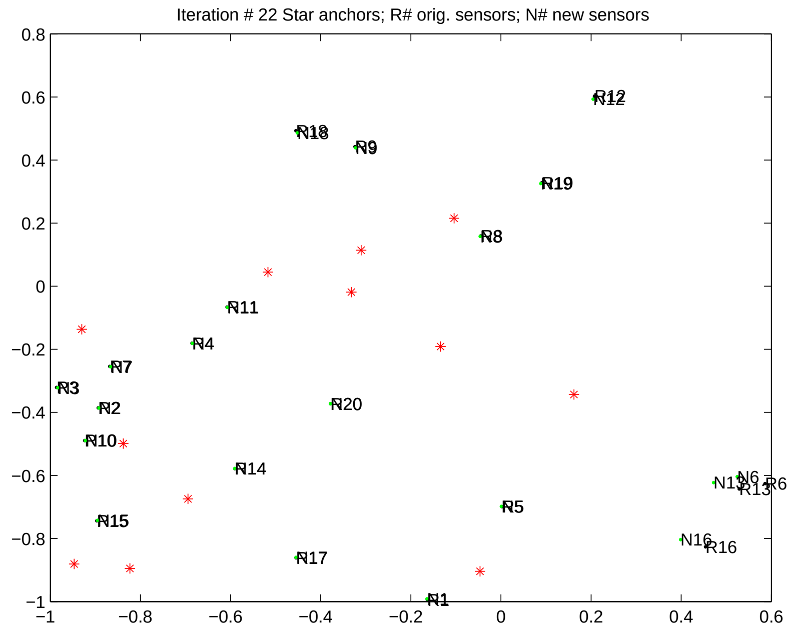

Figure 6: After 22 iterations for randomly generated problem

where $Y = \mathrm{sMat}\,(y)$ and $\bar{y} = \mathrm{diag}\,(Y)$. Let $W^{(2)} := W \circ W$. Then

$$\begin{aligned}
\mathcal{K}^*(W \circ \mathcal{K}_W(x,y)) &= \mathcal{K}^* \left(W^{(2)} \circ \begin{pmatrix} \mathcal{K}(Y) & \begin{pmatrix} \bar{y}_1 e_m^T \\ \vdots \\ \bar{y}_n e_m^T \end{pmatrix} - 2XA^T \\ \begin{pmatrix} \bar{y}_1 e_m & \cdots & \bar{y}_n e_m \end{pmatrix} - 2AX^T & 0_{m\times m} \end{pmatrix} \right) \\
&= 2\mathrm{Diag}\left(\left[W^{(2)} \circ \begin{pmatrix} \mathcal{K}(Y) & \begin{pmatrix} \bar{y}_1 e_m^T \\ \vdots \\ \bar{y}_n e_m^T \end{pmatrix} - 2XA^T \\ \begin{pmatrix} \bar{y}_1 e_m & \cdots & \bar{y}_n e_m \end{pmatrix} - 2AX^T & 0_{m\times m} \end{pmatrix} \right] e \right) \\
&\quad - 2W^{(2)} \circ \begin{pmatrix} \mathcal{K}(Y) & \begin{pmatrix} \bar{y}_1 e_m^T \\ \vdots \\ \bar{y}_n e_m^T \end{pmatrix} - 2XA^T \\ \begin{pmatrix} \bar{y}_1 e_m & \cdots & \bar{y}_n e_m \end{pmatrix} - 2AX^T & 0_{m\times m} \end{pmatrix}.
\end{aligned}$$

Thus

$$(\mathcal{K}_W^x)^* \mathcal{K}_W(x,y) = -2\mathrm{vec}\left(\left[W^{(2)}(1:n, n+1:n+m) \circ \left(\begin{pmatrix} \bar{y}_1 e_m^T \\ \vdots \\ \bar{y}_n e_m^T \end{pmatrix} - 2XA^T \right) \right] A \right).$$

Moreover,

$$(\mathcal{K}_W^y)^* \mathcal{K}_W(x,y) = \mathrm{svec}\,\mathrm{sblk}_1\left(\mathcal{K}^*(W \circ \mathcal{K}_W(x,y))\right).$$

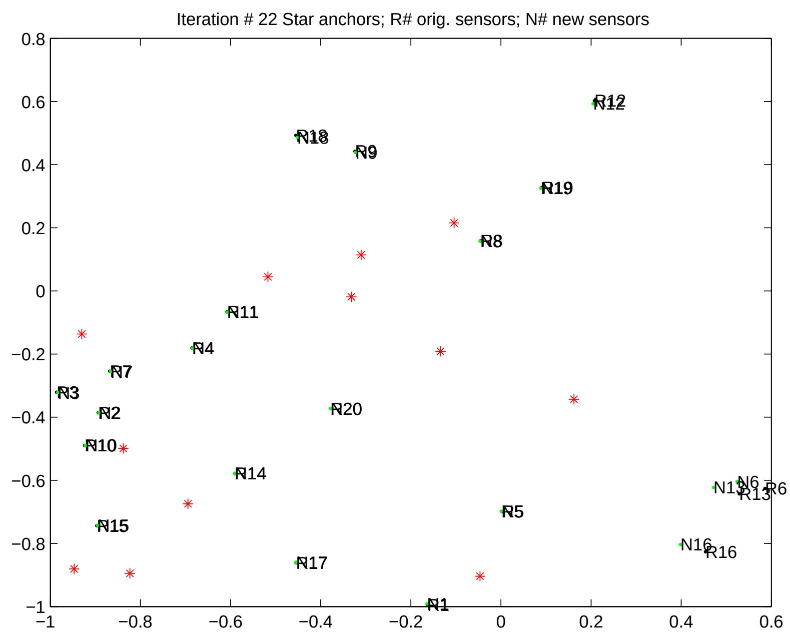

Figure 7: After 29 iterations for randomly generated problem

To evaluate this quantity, note that

$$\begin{aligned}\operatorname{sblk}_1\left(\mathcal{K}^*(W\circ\mathcal{K}_W(x,y))\right) &= 2\mathrm{Diag}\left(\left[W^{(2)}(1:n,1:n+m)\circ\left(\mathcal{K}(Y)\quad\begin{pmatrix}\bar{y}_1e_m^T\\ \vdots\\ \bar{y}_ne_m^T\end{pmatrix}-2XA^T\right)\right]e\right)\\ &\quad -2W^{(2)}(1:n,1:n)\circ\mathcal{K}(Y).\end{aligned}$$

By the expression of $\mathcal{K}(Y)$, we get

$$\begin{aligned}&\left(\left[W^{(2)}(1:n,1:n+m)\circ\left(\mathcal{K}(Y)\quad\begin{pmatrix}\bar{y}_1e_m^T\\ \vdots\\ \bar{y}_ne_m^T\end{pmatrix}-2XA^T\right)\right]e\right)_i\\ &=\sum_{j=1}^n W_{ij}^2(\bar{y}_i+\bar{y}_j-2Y_{ij})+\sum_{j=1}^m W_{i,j+n}^2(\bar{y}_i+X(:,i)A(j,:)).\end{aligned}$$

C Elements of λ_2

From (4.45),

$$\lambda_2=\mathcal{L}_2(x,y,\lambda_u,\lambda_l)-\left[W\circ(\mathcal{K}\mathcal{Y}^y)\right]^*(\bar{E}),$$

where the linear transformation

$$\mathcal{L}_2(x,y,\lambda_u,\lambda_l)=\left[W\circ(\mathcal{K}\mathcal{Y}^y)\right]^*(W\circ\mathcal{K}(\mathcal{Y}(x,y)))+\left[H_u\circ(\mathcal{K}\mathcal{Y}^y)\right]^*(\Lambda_u)-\left[H_l\circ(\mathcal{K}\mathcal{Y}^y)\right]^*(\Lambda_l).$$

We now evaluate the columns of the matrix representation of $\mathcal{L}_2$. For $x = e_i$, $i = 1, \ldots, nr$, we get $X = \frac{1}{\sqrt{2}} E_{st}$. By the expression of $(\mathcal{K}_W^y)\mathcal{K}_W(x,y)$, we get

$$\begin{array}{rcl}\mathcal{L}_2(e_i, 0_{t(n)+2t(m+n)}) & = & \operatorname{svec} \operatorname{sblk}_1 \left[\mathcal{K}^* \left(W^{(2)} \circ \mathcal{K} \begin{pmatrix} 0_{n\times n} & \frac{1}{\sqrt{2}} E_{st} A^T \\ \frac{1}{\sqrt{2}} A E_{st}^T & 0_{m\times m} \end{pmatrix} \right) \right] \\ & = & -2\sqrt{2} \operatorname{svec} \left(\operatorname{Diag} \begin{pmatrix} 0_{s-1} \\ \sum_{j=1}^m W_{s,j+n}^2 A_{jt} \\ 0_{n-s} \end{pmatrix} \right),\end{array}$$

i.e. the first nr columns of the matrix representation of $\mathcal{L}_2$ have only one element that can be different from zero. Moreover if the s-th sensor is not related to any anchor, in the sense that there is no information on the distances between this sensor and any of the anchors, then the r corresponding columns of $\mathcal{L}_2$ are zero.

Now, suppose that $x = 0_{nr}$ and $y = e_i$ with i such that $\operatorname{sMat}(Y) = \operatorname{Diag}(e_s)$. Then, by the expression of $(\mathcal{K}_W^y)^* \mathcal{K}_W$, we get

$$\begin{array}{rcl}\mathcal{L}_2(0_{nr}, e_i, 0_{2t(m+n)}) & = & \operatorname{svec} \operatorname{sblk}_1 \left[\mathcal{K}^* \left(W^{(2)} \circ \mathcal{K} \begin{pmatrix} \operatorname{Diag}(e_s) & 0_{n\times m} \\ 0_{m\times n} & 0_{m\times m} \end{pmatrix} \right) \right] \\ & = & \operatorname{svec} \left(2\operatorname{Diag} \left(\left[W^{(2)}(1:n, 1:n+m) \circ \begin{pmatrix} \mathcal{K}(\operatorname{Diag}(e_s)) & \begin{matrix} 0_{(s-1)\times m} \\ e_m^T \\ 0_{(n-s)\times m} \end{matrix} \end{pmatrix} \right] e \right) \right. \\ & & \left. -2W^{(2)}(1:n, 1:n) \circ \mathcal{K}(\operatorname{Diag}(e_s)) \right).\end{array}$$

Moreover,

$$\mathcal{K}(\operatorname{Diag}(e_s)) = \begin{pmatrix} 0_{(s-1)\times(s-1)} & e & 0_{(s-1)\times(n-s)} \\ e^T & 0 & e^T \\ 0_{(n-s)\times(s-1)} & e & 0_{(n-s)\times(n-s)} \end{pmatrix},$$

where e is the vector of all ones of the right dimension. Thus

$$\begin{array}{rcl}\mathcal{L}_2(0_{nr}, e_i, 0_{2t(m+n)}) & = & \operatorname{svec} \left(2\operatorname{Diag} \begin{pmatrix} W^{(2)}(1:s-1, s) \\ \sum_{k=1,k\neq s}^{n+m} W_{s,k}^{(2)} \\ W^{(2)}(s+1:n, s) \end{pmatrix} \right. \\ & & \left. -2 \begin{pmatrix} 0_{(s-1)\times(s-1)} & W^{(2)}(1:s-1, s) & 0_{(s-1)\times(n-s)} \\ W^{(2)}(s, 1:s-1) & 0 & W^{(2)}(s, s+1:n) \\ 0_{(n-s)\times(s-1)} & W^{(2)}(s+1:n, s) & 0_{(n-s)\times(n-s)} \end{pmatrix} \right) \\ & = & 2 \begin{pmatrix} \operatorname{Diag}(W^{(2)}(1:s-1, s)) & -W^{(2)}(1:s-1, s) & 0_{(s-1)\times(n-s)} \\ -W^{(2)}(s, 1:s-1) & \sum_{k=1,k\neq s}^{n+m} W_{s,k}^{(2)} & -W^{(2)}(s, s+1:n) \\ 0_{(n-s)\times(s-1)} & -W^{(2)}(s+1:n, s) & \operatorname{Diag}(W^{(2)}(s+1:n, s)) \end{pmatrix}.\end{array}$$

D Diagonal Preconditioning Details

To clarify some of the calculations, we recall the notation for the nonzeros in a $0,1$ symmetric matrix H, see (4.38); i.e. $h_H^{nz} = \operatorname{svec}_H(H) \in \mathbb{R}^{nz_H}$, where h_H^{nz} is obtained from $\operatorname{svec}(H)$ by removing the zeros. Therefore, for a symmetric matrix S,

$$H \circ S = \operatorname{sMat}_H \operatorname{svec}_H(H) \circ \operatorname{sMat}_H \operatorname{svec}_H(S) = \operatorname{sMat}_H(s_H^{nz}). \tag{D.76}$$

To evaluate the first nr columns, we need to evaluate $F'_\mu(e_i, 0_{t(n)+t(m+n)+t(r)})$, $i = 1, \ldots, nr$. We note that if $\Delta x = e_i$, then $\mathrm{Mat}\,(x) = E_{st}$, where we denote by E_{st} the $n \times r$ matrix of all zeros except for the (s,t) element which is equal to 1. Thus we get

$$F'_\mu(e_i, 0_{t(n)+nzu+nzl+t(r)}) = \begin{pmatrix} \lambda_u \circ \mathrm{svec}_{\,H_u} \mathcal{K}_{H_u}(e_i, 0_{t(n)}) \\ \lambda_l \circ \mathrm{svec}_{\,H_l} \mathcal{K}_{H_l}(e_i, 0_{t(n)}) \\ \Lambda \mathcal{Z}_s(e_i, 0_{t(n)}) + [\mathrm{sBlk}_{\,2} \left(\mathrm{sMat}\,\left\{(\mathcal{K}_W^y)^* \mathcal{K}_W(e_i, 0_{t(n)})\right\}\right) + \\ \mathrm{sBlk}_{\,21} \left(\mathrm{Mat}\,\left\{(\mathcal{K}_W^x)^* \mathcal{K}_W(e_i, 0_{t(n)})\right\}\right)] \, Z_s \end{pmatrix}$$

First, we note that, given a matrix H,

$$\begin{array}{rcl} \mathcal{K}_H(e_i, 0_{t(n)}) & = & H \circ \mathcal{K}\left(\mathcal{Y}(e_i, 0_{t(n)})\right) = H \circ \mathcal{K} \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} E_{st} A^T \\ \frac{1}{\sqrt{2}} A E_{ts} & 0 \end{pmatrix} \\ & = & -2H \circ \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} E_{st} A^T \\ \frac{1}{\sqrt{2}} A E_{ts} & 0 \end{pmatrix} \\ & = & \begin{pmatrix} & 0 & & \begin{matrix} 0_{(s-1)\times m} \\ -\frac{2}{\sqrt{2}} \bar{h}_s^T \circ A(:,t)^T \\ 0_{(n-s)\times m} \end{matrix} \\ 0_{m\times(s-1)} & -\frac{2}{\sqrt{2}} \bar{h}_s \circ A(:,t) & 0_{m\times(n-s)} & 0 \end{pmatrix}, \end{array} \tag{D.76}$$

where $\bar{h}_s = H(n+1:n+m, s)$. By this relation we get

$$\lambda_u \circ \mathrm{svec}\, \mathcal{K}_{H_u}(e_i, 0_{t(n)}) = \lambda_u \circ \mathrm{svec}_{\,H_u} \begin{pmatrix} & 0 & & \begin{matrix} 0_{(s-1)\times m} \\ -\frac{2}{\sqrt{2}} (\bar{h}_s^u)^T \circ A(:,t)^T \\ 0_{(n-s)\times m} \end{matrix} \\ 0_{m\times(s-1)} & -\frac{2}{\sqrt{2}} \bar{h}_s^u \circ A(:,t) & 0_{m\times(n-s)} & 0 \end{pmatrix}$$

$$\lambda_l \circ \mathrm{svec}\, \mathcal{K}_{H_l}(e_i, 0_{t(n)}) = \lambda_l \circ \mathrm{svec}_{\,H_l} \begin{pmatrix} & 0 & & \begin{matrix} 0_{(s-1)\times m} \\ -\frac{2}{\sqrt{2}} (\bar{h}_s^l)^T \circ A(:,t)^T \\ 0_{(n-s)\times m} \end{matrix} \\ 0_{m\times(s-1)} & -\frac{2}{\sqrt{2}} \bar{h}_s^l \circ A(:,t) & 0_{m\times(n-s)} & 0 \end{pmatrix}.$$

To get the last $(r+n)^2$ elements of $F'_\mu(e_i, 0_{t(n)+t(m+n)+t(r)})$, we need the following quantities:

$$\Lambda \mathcal{Z}_s(e_i, 0_{t(n)}) = \begin{pmatrix} 0_{(r+n)\times(t-1)} & \frac{1}{\sqrt{2}} \Lambda(:, r+s) & 0_{(r+n)\times(r+s-t-1)} & \frac{1}{\sqrt{2}} \Lambda(:,t) & 0_{(r+n)\times(n-s)} \end{pmatrix}$$

Moreover, by (D.76)

$$\begin{array}{rcl} (\mathcal{K}_W)^* \mathcal{K}_W(e_i, 0_{t(n)}) & = & -2\mathcal{Y}^* \left(\mathcal{K}^* \left(W^{(2)} \circ \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} E_{st} A^T \\ \frac{1}{\sqrt{2}} A E_{ts} & 0 \end{pmatrix}\right)\right) \\ & = & -4\mathcal{Y}^* \left(\mathrm{Diag} \begin{pmatrix} 0_{s-1} \\ \sum_{k=1}^m W(s, n+k)^2 A(k,t) \\ 0_{n-s} \\ W(s, n+1)^2 A(1,t) \\ \vdots \\ W(s, n+m)^2 A(m,t) \end{pmatrix} \right. - \end{array}$$

$$
- \left(\begin{array}{cccc} & 0_{n\times n} & & \begin{array}{c} 0_{(s-1)\times m} \\ \frac{1}{\sqrt{2}} \bar{w}_s^T \circ A(:,t)^T \\ 0_{(n-s)\times m} \end{array} \\ 0_{m\times(s-1)} & \frac{1}{\sqrt{2}} \bar{w}_s \circ A(:,t) & 0_{m\times(n-s)} & 0_{m\times m} \end{array} \right) \Bigg) \Bigg)
$$

$$
= \quad -4\mathcal{Y}^*(Q)
$$

where $\bar{w}_s = W(n+1:n+m,s)^2$. By the expression of $\mathcal{Y}^*$, we get

$$
\begin{aligned}
\mathrm{Mat}\ ((\mathcal{Y}^x)^*(Q)) &= \mathrm{sblk}_{21} \begin{pmatrix} 0_{(s-1)\times(s-1)} & 0_{(s-1)\times 1} & 0_{(s-1)\times(n-s)} & 0_{(s-1)\times m} \\ 0_{1\times(s-1)} & \bar{w}_s^T A(:,t) & 0_{(1)\times(n-s)} & -\frac{1}{\sqrt{2}} \bar{w}_s^T \circ A(:,t)^T \\ 0_{(n-s)\times(s-1)} & 0_{(n-s)\times 1} & 0_{(n-s)\times(n-s)} & 0_{(n-s)\times m} \\ 0_{m\times(s-1)} & -\frac{1}{\sqrt{2}} \bar{w}_s \circ A(:,t) & 0_{m\times(n-s)} & \mathrm{diag}\,(\bar{w}_s \circ A(:,t)) \end{pmatrix}^T A \\
&= \begin{pmatrix} 0_{(s-1)\times r} \\ -(\bar{w}_s \circ A(:,t))^T A \\ 0_{(n-s)\times r} \end{pmatrix}
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{sMat}\ ((\mathcal{Y}^y)^*(Q)) &= \mathrm{sblk}_{1} \begin{pmatrix} 0_{(s-1)\times(s-1)} & 0_{(s-1)\times 1} & 0_{(s-1)\times(n-s)} & 0_{(s-1)\times m} \\ 0_{1\times(s-1)} & \bar{w}_s^T A(:,t) & 0_{(1)\times(n-s)} & -\frac{1}{\sqrt{2}} \bar{w}_s^T \circ A(:,t)^T \\ 0_{(n-s)\times(s-1)} & 0_{(n-s)\times 1} & 0_{(n-s)\times(n-s)} & 0_{(n-s)\times m} \\ 0_{m\times(s-1)} & -\frac{1}{\sqrt{2}} \bar{w}_s \circ A(:,t) & 0_{m\times(n-s)} & \mathrm{diag}\,(\bar{w}_s \circ A(:,t)) \end{pmatrix} \\
&= \begin{pmatrix} 0_{(s-1)\times(s-1)} & 0_{(s-1)\times 1} & 0_{(s-1)\times(n-s)} \\ 0_{1\times(s-1)} & \bar{w}_s^T A(:,t) & 0_{(1)\times(n-s)} \\ 0_{(n-s)\times(s-1)} & 0_{(n-s)\times 1} & 0_{(n-s)\times(n-s)} \end{pmatrix}
\end{aligned}
$$

and hence

$$
\begin{aligned}
&\left[\mathrm{sBlk}_2 \left(\mathrm{sMat}\ \left\{(\mathcal{K}_W^y)^* \mathcal{K}_W(e_i, 0_{t(n)})\right\}\right) + \mathrm{sBlk}_{21} \left(\mathrm{Mat}\ \left\{(\mathcal{K}_W^x)^* \mathcal{K}_W(e_i, 0_{t(n)})\right\}\right)\right] Z_s \\
&= \begin{pmatrix} 0_{r\times r} & 0_{r\times(s-1)} & 4A^T(\bar{w}_s \circ A(:,t)) & 0_{r\times(n-s)} \\ 0_{(s-1)\times r} & 0_{(s-1)\times(s-1)} & 0_{(s-1)\times 1} & 0_{(s-1)\times(n-s)} \\ 4(\bar{w}_s \circ A(:,t))^T A & 0_{1\times(s-1)} & -4\bar{w}_s^T A(:,t) & 0_{(1)\times(n-s)} \\ 0_{(n-s)\times r} & 0_{(n-s)\times(s-1)} & 0_{(n-s)\times 1} & 0_{(n-s)\times(n-s)} \end{pmatrix} Z_s
\end{aligned}
$$

Now assume that $\Delta x = \Delta\lambda_l = \Delta\lambda_1 = 0$, and that $\Delta y = e_i$, with i such that $\mathrm{sMat}\,(y) = \mathrm{Diag}\,(e_s)$. We get

$$
F'_\mu(0_{nr}, e_i, 0_{nzl+nzu+t(r)}) = \begin{pmatrix} \lambda_u \circ \mathrm{svec}\, \mathcal{K}_{H_u}(0_{nr}, e_i) \\ \lambda_l \circ \mathrm{svec}\, \mathcal{K}_{H_l}(0_{nr}, e_i) \\ \Lambda \mathcal{Z}_s(0_{nr}, e_i) + \left[\mathrm{sBlk}_2 \left(\mathrm{sMat}\ \left\{(\mathcal{K}_W^y)^* \mathcal{K}_W(0_{nr}, e_i)\right\}\right) + \right. \\ \left. \mathrm{sBlk}_{21} \left(\mathrm{Mat}\ \left\{(\mathcal{K}_W^x)^* \mathcal{K}_W(0_{nr}, e_i)\right\}\right)\right] Z_s \end{pmatrix}
$$

Moreover, given a matrix H, we have

$$
\mathcal{K}_H(0_{nr}, e_i) \quad = \quad H \circ \begin{pmatrix} \mathcal{K}(\mathrm{Diag}\,(e_s)) & e_s & \dots & e_s \\ e_s^T & & & \\ \vdots & & 0_{m\times m} & \\ e_s^T & & & \end{pmatrix}
$$

and

$$\mathcal{K}(\mathrm{Diag}\,(e_s)) = \begin{pmatrix} 0_{(s-1)\times(s-1)} & e & 0_{(s-1)\times(n-s)} \\ e^T & 0 & e^T \\ 0_{(n-s)\times(s-1)} & e & 0_{(n-s)\times(n-s)} \end{pmatrix},$$

Thus the first rows of F'_μ are given by

$$\lambda_u \circ \mathrm{svec}_{H_u}\left(H_u \circ \begin{pmatrix} \begin{matrix} 0_{(s-1)\times(s-1)} & e & 0_{(s-1)\times(n-s)} \\ e^T & 0 & e^T \\ 0_{(n-s)\times(s-1)} & e & 0_{(n-s)\times(n-s)} \end{matrix} & \begin{matrix} e_s & \dots & e_s \end{matrix} \\ \begin{matrix} e_s^T \\ \vdots \\ e_s^T \end{matrix} & 0_{m\times m} \end{pmatrix}\right)$$

$$\lambda_l \circ \mathrm{svec}_{H_l}\left(H_l \circ \begin{pmatrix} \begin{matrix} 0_{(s-1)\times(s-1)} & e & 0_{(s-1)\times(n-s)} \\ e^T & 0 & e^T \\ 0_{(n-s)\times(s-1)} & e & 0_{(n-s)\times(n-s)} \end{matrix} & \begin{matrix} e_s & \dots & e_s \end{matrix} \\ \begin{matrix} e_s^T \\ \vdots \\ e_s^T \end{matrix} & 0_{m\times m} \end{pmatrix}\right)$$

As for the last $(n+r)^2$ rows, we get

$$\Lambda\mathcal{Z}(0_{nr}, e_i) = \begin{pmatrix} 0_{n\times(s-1)} & \Lambda(:,s) & 0_{n\times(n-s)} \end{pmatrix},$$

$$\mathrm{sBlk}_2\left(\mathrm{sMat}\left\{(\mathcal{K}_W^y)^*\mathcal{K}_W(0_{nr}, e_i)\right\}\right) =$$
$$2\begin{pmatrix} 0_{r\times r} & & 0_{r\times n} & \\ & \mathrm{Diag}\,(W^{(2)}(1:s-1,s)) & -W^{(2)}(1:s-1,s) & 0_{(s-1)\times(n-s)} \\ 0_{n\times r} & -W^{(2)}(s,1:s-1) & \sum_{k=1,k\neq s}^{n+m} W^{(2)}_{s,k} & -W^{(2)}(s,s+1:n) \\ & 0_{(n-s)\times(s-1)} & -W^{(2)}(s+1:n,s) & \mathrm{Diag}\,(W^{(2)}(s+1:n,s)) \end{pmatrix},$$

and

$$\mathrm{sBlk}_{21}\left(\mathrm{Mat}\left\{(\mathcal{K}_W^x)^*\mathcal{K}_W(0_{nr}, e_i)\right\}\right) =$$
$$2\begin{pmatrix} 0_{r\times r} & \begin{matrix} 0_{r\times s-1} \\ A^TW^{(2)}(n+1:n+m,s) \\ 0_{r\times n-s} \end{matrix} \\ \begin{matrix} 0_{s-1\times r} \\ W^{(2)}(s,n+1:n+m)A \\ 0_{n-s\times r} \end{matrix} & 0_{n\times n} \end{pmatrix}.$$

Now suppose that $\Delta x = \Delta\lambda_l = \Delta\lambda_1 = 0$ and $\Delta y = e_i$, where i is an index such that $\mathrm{sMat}\,(\Delta y) = \frac{1}{\sqrt{2}}E_{st}$. This in particular implies $s \neq t$. Then given a matrix H, we have

$$\mathcal{K}_H(0_{nr}, e_i) \quad = \quad H \circ \begin{pmatrix} -\frac{2}{\sqrt{2}}E_{st} & 0_{n\times m} \\ 0_{m\times n} & 0_{m\times m} \end{pmatrix}$$

Thus the first rows of F'_μ are given by

$$\begin{array}{l} \lambda_u \circ \operatorname{svec}_{H_u}\left(H_u \circ \begin{pmatrix} -\frac{2}{\sqrt{2}}E_{st} & 0_{n\times m} \\ 0_{m\times n} & 0_{m\times m} \end{pmatrix}\right) \\ \lambda_l \circ \operatorname{svec}_{H_l}\left(H_l \circ \begin{pmatrix} -\frac{2}{\sqrt{2}}E_{st} & 0_{n\times m} \\ 0_{m\times n} & 0_{m\times m} \end{pmatrix}\right) \end{array}$$

As for the last $(n+r)^2$ rows, we get

$$\begin{array}{l} \operatorname{sBlk}_2\left(\operatorname{sMat}\left\{(\mathcal{K}_W^y)^*\mathcal{K}_W(0_{nr}, e_i)\right\}\right) = \\ \operatorname{sBlk}_2\left(2\left[\operatorname{Diag}\left(W^{(2)}(1:n, 1:n+m) \circ \begin{pmatrix} -\frac{2}{\sqrt{2}}E_{st} & 0_{n\times m} \end{pmatrix}\right)\right] e\right), \end{array}$$

and

$$\operatorname{sBlk}_{21}\left(\operatorname{Mat}\left\{(\mathcal{K}_W^x)^*\mathcal{K}_W(0_{nr}, e_i)\right\}\right) = 0_{n+m\times n+m}.$$

Assume $\Delta x = \Delta y = \Delta\lambda_l = \Delta\lambda_1 = 0$ and $\Delta\lambda_u = e_i$, where i is an index such that $\operatorname{sMat}(\Delta\lambda_u) = \frac{1}{\sqrt{2}}E_{st}$, with (s,t) such that $H_u(s,t) \neq 0$. This in particular implies $s \neq t$. The first two blocks of rows of F'_μ are given by

$$F'_\mu(0_{nr+t(n)}, e_i, 0_{nzl+t(r)}) = \begin{pmatrix} s_u \circ e_i \\ 0_{nl} \end{pmatrix}$$

We note that $H_u \circ E_{st} = E_{st}$. Thus the last $(n+r)^2$ rows of F'_μ are given by

$$\left[\operatorname{sBlk}_2\left(\operatorname{sMat}\left((\mathcal{K}_{H_u}^y)^*(\frac{1}{\sqrt{2}}E_{st})\right)\right) + \operatorname{sBlk}_{21}\left(\operatorname{sMat}\left((\mathcal{K}_{H_u}^x)^*(\frac{1}{\sqrt{2}}E_{st})\right)\right)\right] Z_s$$

with

$$\operatorname{sMat}\left((\mathcal{K}_{H_u}^y)^*(\frac{1}{\sqrt{2}}E_{st})\right) = \frac{1}{\sqrt{2}}\operatorname{sblk}_1\left(\mathcal{K}^*(E_{st})\right)$$

and

$$\operatorname{sMat}\left((\mathcal{K}_{H_u}^x)^*(\frac{1}{\sqrt{2}}E_{st})\right) = \frac{1}{\sqrt{2}}\operatorname{sblk}_{21}\left(\mathcal{K}^*(E_{st})\right)$$

Moreover,

$$\mathcal{K}^*(E_{st}) = \begin{pmatrix} 0_{(s-1)\times(t-1)} & 0_{s-1} & 0_{(s-1)\times(s-t-1)} & 0_{s-1} & 0_{(s-1)\times(t-s-1)} & 0_{s-1} & 0_{(s-1)\times(n+m-t-s)} \\ 0_{t-1}^T & 0 & 0_{s-1-t}^T & 1 & 0_{t-s-1}^T & -1 & 0_{n+m-t-s}^T \\ 0_{(t-s-1)\times(t-1)} & 0 & 0_{(t-s-1)\times(s-t-1)} & 0_{t-s-1} & 0_{(t-s-1)\times(t-s-1)} & 0_{t-s-1} & 0_{(t-s-1)\times(n+m-t-s)} \\ 0_{t-1}^T & -1 & 0_{s-1-t}^T & 0_{t-s-1}^T & 0_{t-s-1}^T & 0 & 0_{n+m-t-s}^T \\ 0_{(n+m-s-t)\times(t-1)} & 0 & 0_{(n+m-s-t)\times(s-t-1)} & 0_{(n+m-s-t)} & 0_{(n+m-s-t)\times(t-s-1)} & 0_{n+m-s-t} & 0_{(n+m-s-t)\times(n+m-t-s)} \end{pmatrix}$$

Similar reasoning can be done for the columns corresponding to $\Delta\lambda_l = e_i$, where i is an index such that $\operatorname{sMat}(\Delta\lambda_l) = \frac{1}{\sqrt{2}}E_{st}$, with (s,t) such that $H_l(s,t) \neq 0$. Now, to evaluate the last $t(r)$ columns of F'_μ, we set $\Delta x = \Delta y = \Delta\lambda_u = \Delta\lambda_l = 0$, and $\Lambda_1 = e_i$, $i = 1, \ldots, t(r)$. We have two cases: $\operatorname{sMat}(\lambda_1) = \operatorname{Diag}(e_s)$ or $\operatorname{sMat}(\lambda_1) = 1/\sqrt{2}E_{st}$. Suppose $\operatorname{sMat}(\lambda_1) = \operatorname{Diag}(e_s)$. Then, recalling the expression of F'_μ and the expression of Z_s:

$$F'_\mu(0_{nr+t(n)+nzl+nzu}, e_i) \quad = \quad \begin{pmatrix} 0_{nzl+nzu} \\ \operatorname{sBlk}_1(\operatorname{Diag}(e_s))Z_s \end{pmatrix}$$

where

$$\mathrm{sBlk}_1(\mathrm{Diag}(e_s))Z_s = \begin{pmatrix} & 0_{(s-1)\times n} \\ \mathrm{Diag}(e_s) & X(:,s)^T \\ & 0_{(r-s)\times n} \\ 0_{r\times n} & 0_{n\times n} \end{pmatrix}$$

If $\mathrm{sMat}(\lambda_1) = 1/\sqrt{2}E_{st}$, then

$$F'_\mu(0_{nr+t(n)+2t(m+n)}, e_i) = \begin{pmatrix} 0_{nzl+nzu} \\ \mathrm{sBlk}_1(\frac{1}{\sqrt{2}}E_{st})Z_s \end{pmatrix}$$

where

$$\mathrm{sBlk}_1(\frac{1}{\sqrt{2}}E_{st})Z_s = \frac{1}{\sqrt{2}} \begin{pmatrix} & 0_{(s-1)\times n} \\ & X(:,s)^T \\ E_{st} & 0_{(r-s-t-1)\times n} \\ & X(:,t)^T \\ & 0_{(r-t)\times n} \\ 0_{r\times n} & 0_{n\times n} \end{pmatrix}.$$

References

[1] S. AL-HOMIDAN and H. WOLKOWICZ. Approximate and exact completion problems for Euclidean distance matrices using semidefinite programming. *Linear Algebra Appl.*, 406:109–141, 2005.

[2] 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.

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

[4] P. BISWAS, T.C. LIANG, K.C. TOH, T.C. WANG, and Y. YE. Semidefinite programming approaches for sensor network localization with noisy distance measurements. *IEEE Transactions on Automation Science and Engineering*, 2006. To appear.

[5] P. BISWAS, T.C. LIANG, K.C. TOH, and Y. YE. An SDP based approach for anchor-free 3D graph realization. Technical report, Operation Research, Stanford University, Stanford, CA, 2005.

[6] P. BISWAS and Y. YE. Semidefinite programming for ad hoc wireless sensor network localization. In *Information Processing In Sensor Networks, Proceedings of the third international symposium on Information processing in sensor networks*, pages 46–54, Berkeley, Calif., 2004.

[7] P. BISWAS and Y. YE. A distributed method for solving semidefinite programs arising from ad hoc wireless sensor network localization. In *Multiscale optimization methods and applications*, volume 82 of *Nonconvex Optim. Appl.*, pages 69–84. Springer, New York, 2006.

[8] Å. BJÖRCK. Least squares methods. In P. G. Ciarlet and J. L. Lions, editors, *Handbook of Numerical Analysis. I. Solution of Equations in R^n. Part 1*, pages 466–647. Elsevier/North-Holland, Amsterdam, 1990.

[9] J.M. DANSKIN. *The theory of max-min and its application to weapons allocation problems.* Econometrics and Operations Research, Vol. V. Springer-Verlag New York, Inc., New York, 1967.

[10] Y-B. DENG, X-Y. HU, and L. ZHANG. Least squares solution of $BXA^T = T$ over symmetric, skew-symmetric, and positive semidefinite X. *SIAM J. Matrix Anal. Appl.*, 25(2):486–494 (electronic), 2003.

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

[12] J. ERIKSSON and M.E. GULLIKSSON. Local results for the Gauss-Newton method on constrained rank-deficient nonlinear least squares. *Math. Comp.*, 73(248):1865–1883 (electronic), 2004.

[13] N.J. HIGHAM. *Accuracy and Stability of Numerical Algorithms.* SIAM, Philadelphia, 1995.

[14] H. JIN. *Scalable Sensor Localization Algorithms for Wireless Sensor Networks.* PhD thesis, Toronto University, Toronto, Ontario, Canada, 2005.

[15] 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.

[16] R-C LI and Q. YE. A Krylov subspace method for quadratic matrix polynomials with application to constrained least squares problems. *SIAM J. Matrix Anal. Appl.*, 25(2):405–428 (electronic), 2003.

[17] J. LIU. Eigenvalue and singular value inequalities of Schur complements. In F. ZHANG, editor, *The Schur complement and its applications*, volume 4 of *Numerical Methods and Algorithms*, pages xvi+295. Springer-Verlag, New York, 2005.

[18] C.C. PAIGE and M.A. SAUNDERS. LSQR: an algorithm for sparse linear equations and sparse least squares. *ACM Trans. Math. Software*, 8(1):43–71, 1982.

[19] A.M. SO and Y. YE. Theory of semidefinite programming for sensor network localization. *Math. Programming*, to appear, 2006.

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