Noisy Euclidean Distance Matrix Completion with a Single Missing Node

Stefan Sremac * Fei Wang† Henry Wolkowicz‡ Lucas Pettersson §

Thursday 28th March, 2019, 13:03

Abstract

We present several solution techniques for the noisy single source localization problem, i.e., the Euclidean distance matrix completion problem with a single missing node to locate under noisy data. For the case that the sensor locations are fixed, we show that this problem is implicitly convex, and we provide a purification algorithm along with the SDP relaxation to solve it *efficiently* and *accurately*. For the case that the sensor locations are relaxed, we study a model based on facial reduction. We present several approaches to solve this problem efficiently, and we compare their performance with existing techniques in the literature. Our tools are *semidefinite programming*, *Euclidean distance matrices*, *facial reduction*, and the *generalized trust region subproblem*. We include extensive numerical tests.

Keywords: single source localization, noise, Euclidean distance matrix completion, semidefinite programming, wireless communication, facial reduction, generalized trust region subproblem.

AMS subject classifications: 90C22, 15A83, 90C20, 62P30

Contents

*Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1. Research supported by the Natural Sciences and Engineering Research Council of Canada.

†Department of Mathematics, Royal Institute of Technology, Sweden.

‡Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1; Research supported by The Natural Sciences and Engineering Research Council of Canada.

§Undergraduate at Department of Physics, Royal Institute of Technology, Sweden. Research supported by an Undergraduate Student Research Awards Program.

List of Tables

List of Algorithms

1 Introduction

In this paper we consider the noisy, *single source localization problem*. The objective is to locate the source of a signal that is detected by a set of sensors with exactly known locations. Distances between sensors and source are given, but contaminated with noise. For instance, in an application to cellular networks, the source of the signal is a cellular phone and the cellular towers are the sensors. Our data is the, possibly noisy, distance measurements from each sensor to the source.

The single source localization problem has applications in e.g., navigation, structural engineering, and emergency response, [2, 3, 7, 25, 27, 38]. In general, it is related to distance geometry problems where the input consists of Euclidean distance measurements and a set of points in Euclidean space. The *sensor network localization problem* is a generalization of our single source

problem, where there are multiple sources and only some of the distance estimates are known. The general *Euclidean distance matrix completion problem* is yet a further generalization, where sensors do not have specified locations and only partial, possibly noisy, distance information is available, e.g., [1, 12, 14]. We refer the readers to the books [4, 8, 9] for applications, and to the paper [18] for algorithmic comparisons. We also refer the readers for the related *nearest Euclidean distance matrix (**NEDM**)* problem to the papers [30,31] where a semismooth Newton approach and a rank majorization approach is presented. The more general weighted **NEDM** is a much harder problem though. For theory that relates **NEDM** to semidefinite programming, see e.g., [11, 26].

A common approach to solving an instance of the single source localization problem is a modification of the least squares problem, referred to as the *squared least squares (**SLS**)* problem. We consider two equivalent formulations of **SLS**: the *generalized trust region subproblem (**GTRS**)* formulation; and the *nearest Euclidean distance matrix with fixed sensors (**NEDMF**)* formulation. We show that every extreme point of the semidefinite relaxation of **GTRS** may be easily transformed into a solution of **GTRS** and thus a solution of the **SLS** problem.

We also introduce and analyze several relaxations of the **NEDMF** formulation. These utilize semidefinite programming, *facial reduction*, and *parametric optimization*. We provide theoretical evidence that, generally, the solutions to these relaxations may be easily transformed into solutions of **SLS**. We also provide empirical evidence that the solutions to these relaxations may give better prediction for the location of the source.

1.1 Outline

In Section 1.2 we establish our notation and introduce background concepts. In Section 2.1 we prove strong duality for the **GTRS** formulation of **SLS** and in Section 2.2 we derive the semidefinite relaxation (**SDR**), and prove that it is tight. We also show that the extreme points of the optimal set of **SDR** correspond exactly to the optimizers of **SLS**. A *purification* algorithm for obtaining the extreme points is presented in Section 2.2.1. In Section 3 we introduce the **NEDM** formulation as well as several relaxations. We analyze the theoretical properties of the relaxations and present algorithms for solving them. The results of numerical comparisons of the algorithms are presented in Section 4.

1.2 Preliminaries

We denote by $\mathcal{S}^n$ the space of $n \times n$ real symmetric matrices endowed with the *trace inner product* and corresponding *Frobenius norm*,

$$\langle X, Y \rangle := \operatorname{trace}(XY) = \sum_{ij} X_{ij} Y_{ij}, \qquad \|X\|_F := \sqrt{\operatorname{trace}(XX)} = \sqrt{\sum_{ij} X_{ij}^2}.$$

Unless otherwise specified, the norm of a matrix is the Frobenius norm, and we may drop the subscript F. For a convex set C, the convex subset $f \subseteq C$ is a *face of* C if for all $x, y \in C, z \in f$ with $z \in (x, y)$, (the open line segment between x and y) we have $z \in f$.

The *cone of positive semidefinite matrices* is denoted by $\mathcal{S}^n_+$ and its interior is the *cone of positive definite matrices*, $\mathcal{S}^n_{++}$. The positive semidefinite cone is pointed, closed and convex. Moreover, the cone $\mathcal{S}^n_+$ induces a partial order on $\mathcal{S}^n$, that is $Y \succeq X$ if $Y - X \in \mathcal{S}^n_+$ and $Y \succ X$ if $Y - X \in \mathcal{S}^n_{++}$. Every face of $\mathcal{S}^n_+$ is characterized by the range or nullspace of matrices in its relative

interior, equivalently, by matrices of maximum rank. For $S \subseteq \mathcal{S}^n_+$, we denote the *minimal face of* S, face(S), the smallest face of $\mathcal{S}^n_+$ that contains S. Let $X \in \mathcal{S}^n_+$ have rank r with orthogonal spectral decomposition.

$$X = \begin{bmatrix} P & Q \end{bmatrix} \begin{bmatrix} D_+ & 0 \\ 0 & 0 \end{bmatrix}^T \begin{bmatrix} P & Q \end{bmatrix}^T, \; D_+ \in \mathcal{S}^r_{++}.$$

Then the range and nullspace characterizations of face(X) are,

$$\text{face}(X) = P\mathcal{S}^r_{++}P^T = \mathcal{S}^n_+ \cap \{QQ^T\}^\perp.$$

We say that the matrix QQ^T is an *exposing vector* for face(X).

Sometimes it is helpful to vectorize a symmetric matrix. Let svec : $\mathcal{S}^n \to \mathbb{R}^{n(n+1)/2}$ map the upper triangular elements of a symmetric matrix to a vector, and let sMat = svec^{-1}.

The *centered subspace of* $\mathcal{S}^n$*, denoted* $\mathcal{S}_C$, is defined as

$$\mathcal{S}_C := \{X \in \mathcal{S}^n : Xe = 0\},$$

where e is *the vector of all ones*. The *hollow subspace of* $\mathcal{S}^n$*, denoted* $\mathcal{S}_H$, is

$$\mathcal{S}_H := \{X \in \mathcal{S}^n : \text{diag}(X) = 0\},$$

where diag : $\mathcal{S}^n \to \mathbb{R}^n$, diag$(X)$:= $(X_{11}, \ldots, X_{nn})^T$. A matrix $D \in \mathcal{S}_H$ is said to be a *Euclidean distance matrix,* ***EDM*** if there exists an integer r and points $x^1, \ldots, x^n \in \mathbb{R}^r$ such that

$$\|x^i - x^j\|_2^2 = D_{ij}, \quad \text{for all } ij,$$

where $\|\cdot\|_2$ denotes the Euclidean norm. As for the Frobenius norm, we assume the norm of a vector to be the Euclidean norm when the subscript is omitted. The set of all $n \times n$ **EDM**s, denoted $\mathcal{E}^n$, forms a closed, convex cone with $\mathcal{E}^n \subset \mathcal{S}_H$.

The classical result of Schoenberg [32] states that **EDM**s are characterized by a face of the positive semidefinite cone. We state the result in terms of the Lindenstrauss mapping, $\mathcal{K} : \mathcal{S}^n \to \mathcal{S}^n$,

$$\mathcal{K}(X)_{ij} := X_{ii} + X_{jj} - 2X_{ij}.$$

with *adjoint* and *Moore-Penrose pseudoinverse*,

$$\mathcal{K}^*(D) = 2(\text{Diag}(De) - D), \quad \mathcal{K}^\dagger(D) = -\frac{1}{2}J_n \cdot \text{offDiag}(D) \cdot J_n,$$

respectively. Here Diag is the adjoint of diag, the matrix $J_n := I - \frac{1}{n}ee^T$ is the orthogonal projection onto $\mathcal{S}_C$, and offDiag(D) refers to zeroing out the diagonal of D, i.e., the orthogonal projection onto $\mathcal{S}_H$. The range of $\mathcal{K}$ is exactly $\mathcal{S}_H$ and the range of $\mathcal{K}^\dagger$ is the subspace $\mathcal{S}_C$. Moreover, $\mathcal{K}(\mathcal{S}^n_+) = \mathcal{E}^n$ and $\mathcal{K}$ is an isomorphism between $\mathcal{S}_C$ and $\mathcal{S}_H$.

The Schoenberg characterization states that $\mathcal{K}$ is an isomorphism between $\mathcal{S}^n_+ \cap \mathcal{S}_C$ and $\mathcal{E}^n$, see [1] for instance. Specifically,

$$\mathcal{K}(\mathcal{S}^n_+ \cap \mathcal{S}_C) = \mathcal{E}^n, \quad \mathcal{K}^\dagger(\mathcal{E}^n) = \mathcal{S}^n_+ \cap \mathcal{S}_C.$$

Moreover, if $D \in \mathcal{E}^n$ and $\mathcal{K}^\dagger(D) = PP^T$ has rank r with full column rank factorization PP^T, then the rows of P correspond to the points in $\mathbb{R}^r$ with pairwise distances corresponding to the elements of D. For more details, see e.g., [1, 10, 11, 21, 22].

2 SDP Formulation

We begin this section by formulating the **SLS** problem using the model and notation of [2]. We let n denote the number of sensors, $p^1, \dots, p^n \in \mathbb{R}^r$ denotes their locations, and r is the *embedding dimension*.

Assumption 2.1. *The following holds throughout:*

1. $n \geq r + 1$;
2. int $\operatorname{conv}(p^1, \dots, p^n) \neq \emptyset$;
3. $\sum_{i=1}^n p^i = 0$.

The first two items in Assumption 2.1 ensure that a signal can be uniquely recovered if we have accurate distance measurements. If the towers are positioned in a proper affine subspace of $\mathbb{R}^r$, and the signal is not contained within this affine subspace, then there are multiple possible locations for the signal with the given distance measurements. We assume that such poor designs are avoided in our applications. The third assumption is made so that the sources are *centered* about the origin. This property leads to a cleaner exposition in the **NEDM** relaxations of Section 3.

We let $d = \bar{d} + \varepsilon \in \mathbb{R}^n$ denote the vector of noisy distances from the source to the ith sensor,

$$d_i := \bar{d}_i + \varepsilon_i,\ i = 1, \dots, n,$$

where $\bar{d}_i$ is the true distance and ε_i is a perturbation, or noise. When the noise $\varepsilon_1, \dots, \varepsilon_n$ is not too large, then a satisfactory approximation of the location of the source can be obtained as a nearest distance problem to the sensors. Using the Euclidean norm as a metric, we obtain the least squares problem

$$p^*_{\mathbf{LS}} := \min_{x \in \mathbb{R}^r} \sum_{i=1}^n \left(\|x - p^i\| - d_i \right)^2. \tag{2.1}$$

This problem has the desirable property that its solution is the maximum likelihood estimator when the noise is assumed to be normal and the covariance matrix a multiple of the identity, e.g., [7]. However, it is a non-convex problem with an objective function that is not differentiable. Motivated by the success in [2], the main problem we consider instead is the optimization problem with squared distances

$$\boxed{(\mathbf{SLS}) \quad p^*_{\boldsymbol{SLS}} := \min_{x \in \mathbb{R}^r} \sum_{i=1}^n \left(\|x - p^i\|^2 - d_i^2 \right)^2.} \tag{2.2}$$

Though still a non-convex problem, in the subsequent sections we show that a solution of **SLS** can be obtained by solving at most $k \leq r + 1$ convex problems, see Theorem 2.7 below.

2.1 GTRS

The **GTRS** is an optimization problem where the objective is a quadratic and there is a single two-sided quadratic constraint, [29, 33]. Note that this class of problems also includes equality

constraints. If we expand the squared norm term in **SLS** and substitute using $\|x\|^2 = \alpha$ as in [2], we get the equivalent problem

$$p^*_{\boldsymbol{SLS}} = \min_{x,\alpha} \left\{ \sum_{i=1}^{n} \left(\alpha - 2x^T p^i + \|p^i\|^2 - d_i^2\right)^2 : \|x\|^2 - \alpha = 0,\ x \in \mathbb{R}^r \right\}. \tag{2.3}$$

In this formulation, we have a convex quadratic objective that is minimized over a level curve of a convex quadratic function. It follows that (2.3) is an instance of the standard trust region subproblem. Strong duality is proved in [29, 33]. For the sake of completeness, we include a proof of strong duality for our particular class of **GTRS**.

Theorem 2.2. *Let*

$$\begin{aligned} P_T := \begin{bmatrix} p^1 & p^2 & \dots & p^n \end{bmatrix}^T,\ A := \begin{bmatrix} -2P_T & e \end{bmatrix},\ \tilde{I} := \begin{bmatrix} I_r & 0_{r\times 1} \\ 0_{1\times r} & 0 \end{bmatrix}, \\ b := \begin{pmatrix} d_1^2 - \|p^1\|^2 \\ \vdots \\ d_n^2 - \|p^n\|^2 \end{pmatrix},\ \tilde{b} := \begin{pmatrix} 0 \\ -\frac{1}{2} \end{pmatrix}. \end{aligned} \tag{2.4}$$

Consider ***SLS*** *in* (2.2) *and the equivalent form given in* (2.3)*. Then:*

1. *The problem* ***SLS*** *is equivalent to*

$$(\boldsymbol{GTRS}) \qquad p^*_{\boldsymbol{SLS}} = \min\{\|Ay - b\|^2 : y^T \tilde{I} y + 2\tilde{b}^T y = 0,\ y \in \mathbb{R}^{r+1}\}. \tag{2.5}$$

2. *The rank of A is $r+1$ and the optimal value of* ***GTRS*** *is finite and attained.*

3. *Strong duality holds for* ***GTRS****, i.e.,* ***GTRS*** *and its Lagrangian dual have a zero duality gap and the dual value is attained:*

$$p^*_{\boldsymbol{SLS}} = d^*_{\boldsymbol{SLS}} := \max_{\lambda} \min_{y} \{\|Ay - b\|^2 + \lambda(y^T \tilde{I} y + 2\tilde{b}^T y)\}. \tag{2.6}$$

Proof. The first claim that **SLS** can be rewritten as **GTRS** follows immediately using the substitution $y = (x^T, \alpha)^T$. For the second claim, note that by Assumption 2.1, Item 2, $\operatorname{rank}(P_T) = r$. Therefore, $(P_T)^T e = 0$ implies that

$$\operatorname{rank}(A) = \operatorname{rank}(P_T) + 1 = r + 1.$$

Now, since A has full column rank, we conclude that $A^T A$ is positive definite, and therefore the objective of **GTRS** is strictly convex and coercive. Moreover, the constraint set is closed and thus the optimal value of **GTRS** is finite and attained, as desired.

That we have a zero duality gap follows from coercivity in that there is no common recession direction with the objective function and constraint set. (For the convex case see e.g., [16, 17, 24].) We now prove this for our special case. Note that

$$A^T A = \begin{bmatrix} 4P_T^T P_T & 0 \\ 0 & n \end{bmatrix}.$$

Let $\gamma = \lambda_{\min}(4P_T^T P_T)$ be the (positive) smallest eigenvalue of $4P_T^T P_T$ so that we have $A^T A - \gamma \tilde{I} \succeq 0$, but singular. We note that the convex constraint $y^T \tilde{I} y + 2\tilde{b}^T y \leq 0$ satisfies the Slater condition, i.e., strict feasibility. Therefore, the following holds, with justification to follow.

$$\begin{aligned}
p^*_{\mathbf{SLS}} &= \min_y \{\|Ay - b\|^2 : y^T \tilde{I} y + 2\tilde{b}^T y = 0\} \\
&= \min_y \{\|Ay - b\|^2 - \gamma(y^T \tilde{I} y + 2\tilde{b}^T y) : y^T \tilde{I} y + 2\tilde{b}^T y = 0\} && (2.7) \\
&= \min_y \{\|Ay - b\|^2 - \gamma(y^T \tilde{I} y + 2\tilde{b}^T y) : y^T \tilde{I} y + 2\tilde{b}^T y \leq 0\} && (2.8) \\
&= \max_{\lambda \geq 0} \min_y \{\|Ay - b\|^2 - \gamma(y^T \tilde{I} y + 2\tilde{b}^T y) + \lambda(y^T \tilde{I} y + 2\tilde{b}^T y)\} && (2.9) \\
&= \max_{(\lambda - \gamma)} \min_y \{\|Ay - b\|^2 + (\lambda - \gamma)(y^T \tilde{I} y + 2\tilde{b}^T y)\} && (2.10) \\
&= d^*_{\mathbf{SLS}} \\
&\leq p^*_{\mathbf{SLS}}.
\end{aligned}$$

The first equality follows from Item 1 and the second equality holds since $\gamma(y^T \tilde{I} y + 2\tilde{b}^T y)$ is identically 0 for any feasible y.

For the third equality, let the objective and constraint, respectively, be denoted by

$$f(y) := \|Ay - b\|^2 - \gamma(y^T \tilde{I} y + 2\tilde{b}^T y), \quad g(y) := y^T \tilde{I} y + 2\tilde{b}^T y = 0.$$

The optimal value of (2.8) is a lower bound for $p^*_{\mathbf{SLS}}$ since the feasible set of (2.8) is a superset of the feasible set of (2.7) and the objectives are the same. Now suppose, for the sake of contradiction, that the optimal value of (2.8) is strictly less than $p^*_{\mathbf{SLS}}$. Then there exists $\bar{y}$ satisfying,

$$g(\bar{y}) < 0, \quad f(\bar{y}) < p^*_{\mathbf{SLS}}.$$

Let $0 \neq h \in \text{Null}(\nabla^2 f(\bar{y}))$. Then by the structure of $A^T A$ and construction of γ we see that $h = \begin{pmatrix} \bar{h}^T & 0 \end{pmatrix}^T$ with $\bar{h} \neq 0$. Moreover, we have,

$$\lim_{\alpha \to +\infty} g(\bar{y} \pm \alpha h) = \lim_{\alpha \to +\infty} g(\bar{y}) \pm 2\alpha \bar{y}^T h + \alpha^2 \|h\|^2 = +\infty. \tag{2.11}$$

Now we choose $\eta \in \{\pm 1\}$ such that,

$$f(\bar{y} + \eta \alpha h) \leq f(\bar{y}), \, \forall \alpha \geq 0.$$

These observations imply that that there exists $\bar{\alpha} > 0$ such that

$$g(\bar{y} + \bar{\alpha} h) = 0, \; f(\bar{y} + \bar{\alpha} h) \leq f(\bar{y}) < p^*_{\mathbf{SLS}},$$

a contradiction.

We have confirmed the third equality. Now (2.8) is a convex quadratic optimization problem where the Slater constraint qualification holds. This implies that strong duality holds, i.e., we get (2.9) with attainment for some $\lambda \geq 0$. Now if $\lambda < 0$ in (2.9) then the Hessian of the objective is indefinite (by construction of γ) and the optimal value of the inner minimization problem is $-\infty$. Thus since (2.9) is maximized with respect to λ in the outer optimization problem, we may remove the non-negativity constraint and obtain (2.10). The remaining lines are due to the definition of the Lagrangian dual and weak duality. Strong duality follows immediately. □

The above Theorem 2.2 shows that even though **SLS** is a non-convex problem, it can be formulated as an instance of **GTRS** and satisfies strong duality. Therefore it can be solved efficiently using, for instance, the algorithm of [29]. Moreover, in the subsequent results we show that **SLS** is equivalent to its *semidefinite programming (**SDP**)* relaxation in (2.15), a convex optimization problem.

We compare our **SDP** approach with the approach used by Beck et al [2]. In their approach they have to solve the following system obtained from the optimality conditions of **GTRS**:

$$\begin{array}{rcl} (AA^T + \lambda\tilde{I})y & = & A^Tb - \lambda\tilde{b}, \\ y^T\tilde{I}y + 2\tilde{b}^Ty & = & 0, \\ A^TA + \lambda\tilde{I} & \succeq & 0. \end{array} \tag{2.12}$$

The so-called *hard case* results in $A^TA + \lambda^*\tilde{I}$ being singular for the optimal λ^* and this can cause numerical difficulties. We note that in our **SDP** relaxation, we need not differentiate between the *'hard case'* and *'easy case'*.

2.2 The Semidefinite Relaxation, SDR

We now study the convex equivalent of **SLS**. We analyze the dual and the **SDP** relaxation of **GTRS**. By homogenizing the quadratic objective and constraint and using the fact that strong duality holds for the standard trust region subproblem [34], we obtain an equivalent formulation of the Lagrangian dual of **GTRS** as an **SDP**. We first define

$$\bar{A} := \begin{bmatrix} A^TA & -A^Tb \\ -b^TA & b^Tb \end{bmatrix}, \ \bar{B} := \begin{bmatrix} \tilde{I} & \tilde{b} \\ \tilde{b}^T & 0 \end{bmatrix}. \tag{2.13}$$

The Lagrangian dual of **GTRS** may be obtained as follows:

$$\begin{array}{rcl} d^*_{\mathbf{SLS}} & = & \max_\lambda \min_y \|Ay-b\|^2 - \lambda(y^T\tilde{I}y + 2\tilde{b}^Ty) \\ & = & \max_{\lambda,s}\{s : \|Ay-b\|^2 - \lambda(y^T\tilde{I}y + 2\tilde{b}^Ty) - s \geq 0, \forall y \in \mathbb{R}^{r+1}\} \\ & = & \max_{\lambda,s}\left\{s \ : \ \left[y^T,1\right]\left[\bar{A} - \lambda\bar{B} - se_{r+2}e_{r+2}^T\right]\left[y^T,1\right]^T \geq 0, \forall y \in \mathbb{R}^{r+1}\right\} \\ & = & \max_{\lambda,s}\left\{s \ : \ \lambda\bar{B} + se_{r+2}e_{r+2}^T \preceq \bar{A}\right\}. \end{array} \tag{2.14}$$

Here the first equality follows from the definition of the dual. The second and third equalities are just equivalent reformulations of the first one. For the last equality, let $\tilde{y} = [y^T, y_0]^T$ and $M = \bar{A} - \lambda\bar{B} - se_{r+2}e_{r+2}^T$ and suppose $\tilde{y}^TM\tilde{y} < 0$ for some $\tilde{y}$. If y_0 is nonzero, we can get a contradition by scaling $\tilde{y}$. If y_0 is zero, by the continutity of $\tilde{y} \to \tilde{y}^TM\tilde{y}$, we can perturb $\tilde{y}$ by a small enough amount so that the last element of $\tilde{y}$ is nonzero. This is a contradiction as in the previous case.

We observe that (2.14) is a dual-form **SDP** with the well known Lagrangian dual,

$$(\mathbf{SDR}) \qquad \begin{array}{rl} p^*_{\mathbf{SDR}} := \min & \langle\bar{A}, X\rangle \\ \text{s.t.} & \langle\bar{B}, X\rangle = 0 \\ & X_{r+2,r+2} = 1 \\ & X \in \mathcal{S}^{r+2}_+. \end{array} \tag{2.15}$$

Now let $\mathcal{F}$ and Ω, respectively, denote the feasible and optimal sets of solutions of **SDR**. We define the map $\rho : \mathbb{R}^{r+1} \to \mathcal{S}^{r+2}$ as,

$$\rho(y) = \begin{pmatrix} y \\ 1 \end{pmatrix} \begin{pmatrix} y \\ 1 \end{pmatrix}^T. \tag{2.16}$$

Note that ρ is an isomorphism between $\mathbb{R}^{r+1}$ and rank 1 matrices of $\mathcal{S}^{r+2}_+$, where the $(r+2, r+2)$ element is 1.

Lemma 2.3. *The map ρ is an isomorphism between the feasible sets of **GTRS** and **SDR**. Moreover, the objective value is preserved under ρ, i.e., $\|Ay - b\|^2 = \langle \bar{A}, \rho(y) \rangle$.*

Theorem 2.4. *The following holds:*

1. *The optimal values of **GTRS**, **SDR**, and (2.14) are all equal, finite, and attained.*
2. *The matrix X^* is an extreme point of Ω if, and only if, $X^* = \rho^{-1}(y^*)$ for some minimizer, y^*, of **GTRS**.*

Proof. From Theorem 2.2 and weak duality, we have that

$$p^*_{\mathbf{SLS}} = d^*_{\mathbf{SLS}} \leq p^*_{\mathbf{SDR}}. \tag{2.17}$$

Moreover, since **SDR** is a relaxation of **GTRS** we get,

$$p^*_{\mathbf{SDR}} \leq p^*_{\mathbf{SLS}} \implies p^*_{\mathbf{SLS}} = d^*_{\mathbf{SLS}} = p^*_{\mathbf{SDR}}.$$

Furthermore, from Theorem 2.2 the above values are all finite and the optimal values of **GTRS** and (2.14) are attained. To see that the optimal value of **SDR** is attained it suffices to show that (2.14) has a Slater point. Indeed, the feasible set of (2.14) consists of all $\mu, s \in \mathbb{R}$ such that,

$$\begin{bmatrix} A^TA + \mu\tilde{I} & -A^Tb + \mu\tilde{b} \\ -b^TA + \mu\tilde{b}^T & b^Tb - s \end{bmatrix} \succeq 0.$$

Setting $\mu = 0$ and applying the Schur complement condition, we have

$$\begin{bmatrix} A^TA & -A^Tb \\ -b^TA & b^Tb - s \end{bmatrix} \succ 0 \iff A^TA - \frac{1}{b^Tb - s} A^Tb(A^Tb)^T \succ 0,\ A^TA \succ 0.$$

By Theorem 2.2, A^TA is positive definite and a Slater point may be obtained by choosing s so that $b^Tb - s$ is sufficiently large.

Now we consider Item 2. By the existence of a Slater point for (2.14) we know that Ω is compact and convex. Now we show that Ω is actually a face of $\mathcal{F}$. To see this, let $\theta \in (0,1)$ and let $Z = \theta X + (1-\theta)Y \in \Omega$ for some $X, Y \in \mathcal{F}$. Since Z is optimal for **SDR** and X and Y are feasible for **SDR**, we have

$$\langle \bar{A}, Z \rangle = \theta \langle \bar{A}, Z \rangle + (1-\theta)\langle \bar{A}, Z \rangle \leq \theta \langle \bar{A}, X \rangle + (1-\theta)\langle \bar{A}, Y \rangle = \langle \bar{A}, Z \rangle.$$

Now equality holds throughout and we have $\langle \bar{A}, X \rangle = \langle \bar{A}, Y \rangle = \langle \bar{A}, Z \rangle$. Therefore $X, Y \in \Omega$ and by the definition of face, we conclude that Ω is a face of $\mathcal{F}$.

Since Ω is a compact convex set it has an extreme point, say X^*. Now X^* is also an extreme point of $\mathcal{F}$, as the relation 'face of' is transitive. Moreover, since there are exactly two equality

constraints in **SDR**, by Theorem 2.1 of [28], we have $\text{rank}(X^*)(1+\text{rank}(X^*))/2 \leq 2$. This equation is satisfied if, and only if, $\text{rank}(X^*) = 1$. Equivalently, $X^* = \rho(y^*)$ for some $y^* \in \mathbb{R}^{r+1}$. Now, by Lemma 2.3 and the first part of this proof we have that y^* is a minimizer of **GTRS**.

For the converse in Item 2, let y^* be a minimizer of **GTRS**. Then by Lemma 2.3, $X^* := \rho(y^*)$ is optimal for **SDR**. To see that X^* is an extreme point of Ω, let $Y, Z \in \Omega$ such that

$$\frac{1}{2}Y + \frac{1}{2}Z = X^*.$$

Since X^* has rank 1 and $Y, Z \succeq 0$, it follows that Y and Z are non-negative multiples of X^*. But by feasibility, $X^*_{r+2,r+2} = Y_{r+2,r+2} = Z_{r+2,r+2}$ and thus $Y = Z = X^*$. So, by definition, X^* is an extreme point of Ω, as desired. □

We have shown that the optimal value of **SLS** may be obtained by solving the *nice* convex problem **SDR**. Moreover, every extreme point of the optimal face of **SDR** can easily be transformed into an optimal solution of **SLS**. However, **SDR** is usually solved using an interior point method that is guaranteed to converge to a relative interior solution of Ω. In general, such a solution may not have rank 1. In the following corollary of Theorem 2.4 we address those instances for which the solution of **SDR** is readily transformed into a solution of **SLS**. For other instances, we present an algorithmic approach in Section 2.2.1.

Corollary 2.5. *The following hold.*

1. *If **GTRS** has a unique minimizer, say y^*, then the optimal set of **SDR** is the singleton $\rho(y^*)$.*
2. *If the optimal set of **SDR** is a singleton, say X^*, then $\text{rank}(X^*) = 1$ and $\rho^{-1}(X^*)$ is the unique minimizer of **GTRS**.*

Proof. Let y^* be the unique minimizer of **GTRS**. By Theorem 2.4 we know that $\rho(y^*)$ is an extreme point of Ω. Now suppose, for the sake of contradiction, that there exists $X \neq \rho(y^*)$ in Ω. Since Ω is a compact convex set it is the convex hull of its extreme points. Thus there exists an extreme point of Ω, say Y, that is distinct from $\rho(y^*)$. By Theorem 2.4, we know that $\rho^{-1}(Y)$ is a minimizer of **GTRS** and by Lemma 2.3, $\rho^{-1}(Y) \neq y^*$, contradicting the uniqueness of y^*.

For the converse, let X^* be the unique minimizer of **SDR**. Then X^* is the only extreme point of Ω and consequently $\rho^{-1}(X^*)$ is the unique minimizer of **GTRS**, as desired. □

2.2.1 A Purification Algorithm

Suppose the optimal solution of (2.15) is $\bar{X}$ with optimal value $p^*_{\mathbf{SDR}} = \langle \bar{A}, \bar{X} \rangle$ and $\text{rank}(\bar{X}) = \bar{r}$ where $\bar{r} > 1$. Note that we can not obtain an optimal solution of **GTRS** from $\bar{X}$ since the rank is too large. However, in this section we construct an algorithm that returns an extreme point of Ω which, by Theorem 2.4, is easily transformed into an optimal solution of **GTRS**. We note that this does *not* require the extreme point to be an *exposed* extreme point.

Let the compact spectral decomposition of $\bar{X}$ be $\bar{X} := UDU^T$ with $D \in \mathcal{S}^{\bar{r}}_{++}$. We use the substitution $X = USU^T$ and solve the problem (2.20), below, to obtain an optimal solution with lower rank. Note that $D \succ 0$ is a strictly feasible solution for (2.20). We choose the objective matrix $C \in \mathcal{S}^{\bar{r}}_{++}$ to be random and positive definite. To simplify the subsequent exposition, by abuse of notation, we redefine

$$\bar{B} \leftarrow U^T \bar{B} U,\ \bar{A} \leftarrow U^T \bar{A} U,\ \bar{E} \leftarrow U^T \bar{E} U, \tag{2.18}$$

where $\bar{E} := e_{r+2}e_{r+2}^T$. We define the linear map $\mathcal{A} : \mathbb{S}^{\bar{r}} \to \mathbb{R}^3$ and the vector $b \in \mathbb{R}^3$ as,

$$\mathcal{A}_S(S) := \begin{pmatrix} \langle \bar{B}, S \rangle \\ \langle \bar{A}, S \rangle \\ \langle \bar{E}, S \rangle \end{pmatrix}, \quad b_S := \begin{pmatrix} 0 \\ p^*_{\mathbf{SDR}} \\ 1 \end{pmatrix}, \tag{2.19}$$

respectively. The rank reducing program is

$$\begin{array}{ll} \min & \langle C, S \rangle \\ \text{s.t.} & \mathcal{A}_S(S) = b_S \\ & S \in \mathcal{S}^{\bar{r}}_+. \end{array} \tag{2.20}$$

In Algorithm 2.1 we extend the idea of the rank reducing program and in the subsequent results we prove that the output of the algorithm is a rank 1 optimal solution of **SDR**.

Algorithm 2.1 Purification Algorithm

INPUT: $\mathcal{A}_S$ as in (2.19) and $\bar{X} \in \Omega$.
initialize: $k = 1$, $\mathcal{A}^1_S := \mathcal{A}_S$, $S^1 := \bar{X}$, $U^0 = I$.
while $\text{rank}(S^k) \geq 2$ **do**
  Compute compact spectral decomposition, $S^k = U^k D^k (U^k)^T$, with $D^k \in \mathcal{S}^{r_k}_{++}$.
  Redefine $\mathcal{A}^k_S$ and b^k_S using U^k as in (2.18) and ensure that it is full rank.
  Choose $C^k \in \text{Null}(\mathcal{A}^k_S) \setminus \{0\}$.
  Obtain $S^{k+1} \in \arg\min\{\langle C^k, S \rangle : \mathcal{A}^k_S(S) = b^k_S,\ S \succeq 0\}$.
  Update $k \leftarrow k + 1$.
end while
OUTPUT: $X^* := U^0 \cdots U^{k-1} S^k (U^0 \cdots U^{k-1})^T$.

Lemma 2.6. *Let $k \geq 1$ be an integer and suppose that C^k, $\mathcal{A}^k_S$, and b^k_S are as in Algorithm 2.1. Then*

$$S^{k+1} \succ 0 \iff \mathcal{F}^k := \left\{ S \succeq 0 : \mathcal{A}^k_S(S) = b^k_S \right\} = \left\{ S^{k+1} \right\}.$$

Proof. By construction, $D^k \in \mathcal{F}^k$. Therefore,

$$\mathcal{F}^k = \left\{ S^{k+1} \right\} \implies S^{k+1} = D^k \succ 0.$$

For the forward direction, assume that $S^{k+1} \succ 0$ and, for the sake of contradiction, suppose that, S^{k+1} is not the only element of $\mathcal{F}^k$. Then $S^{k+1} \in \text{relint}(\mathcal{F}^k)$ and for any $T \in \text{Null}(\mathcal{A}^k_S)$ there exists $\varepsilon > 0$ such that,

$$\left\{ S^{k+1} + \varepsilon T, S^{k+1} - \varepsilon T \right\} \subset \mathcal{F}^k.$$

By the choice of C^k, there exists $T \in \text{Null}(\mathcal{A}^k_S)$ such that $\langle C^k, T \rangle \neq 0$ and we may assume, without loss of generality, that this inner product is in fact negative. Then,

$$\langle C^k, S^{k+1} + \varepsilon T \rangle < \langle C^k, S^{k+1} \rangle,$$

contradicting the optimality of S^{k+1}. □

Theorem 2.7. *Let $\bar{X} \in \mathcal{S}_+^{r+2}$ be an optimal solution to* ***SDR*** *. If $\bar{X}$ is an input to Algorithm 2.1, then the algorithm terminates with at most* $\operatorname{rank}(\bar{X}) - 1 \leq r + 1$ *calls to the while loop and the output, X^*, is a rank* 1 *optimal solution of* ***SDR****.*

Proof. We proceed by considering the trivial case, $\operatorname{rank}(\bar{X}) = 1$. Clearly $X^* = \bar{X}$ in this case, and we have the desired result. Thus we may assume that the while loop is called at least once. We show that for every S^k generated by Algorithm 2.1 with $k \geq 1$, we have,

$$X^k := U^0 \cdots U^{k-1} S^k (U^0 \cdots U^{k-1})^T \in \Omega. \tag{2.21}$$

To this end, let us consider the constraint $\langle \bar{B}, S^k \rangle = 0$. By the update formula, (2.18), we have,

$$0 = \langle (U^0 \cdots U^{k-1})^T \bar{B} U^0 \cdots U^{k-1}, S^k \rangle = \langle \bar{B}, U^0 \cdots U^{k-1} S^k (U^0 \cdots U^{k-1})^T = \langle \bar{B}, X^k \rangle.$$

Similarly the other two constraints comprising $\mathcal{A}_S^{k-1}$ are satisfied by X^k and therefore $X^k \in \Omega$.

Now we show that the sequence of ranks, $r_1, r_2, \ldots,$ generated by Algorithm 2.1 is strictly decreasing. It immediately follows that the algorithm terminates in at most $\operatorname{rank}(\bar{X}) - 1$ calls to the while loop and that the output matrix X^* has rank 1. Suppose, to the contrary, that there exists an integer $k \geq 2$ such that $r_k = r_{k-1}$. Then by construction, we have that $\operatorname{rank}(S^k) = r_k = r_{k-1}$ and S^k is a Slater point of the optimization problem,

$$\min\{\langle C^{k-1}, S \rangle : \mathcal{A}_S^{k-1}(S) = b_S^{k-1},\ S \succeq 0\}. \tag{2.22}$$

Therefore, by Lemma 2.6 we have that S^k is the only feasible solution of (2.22). Now we claim that X^k as defined above is an extreme point of Ω. To see this, let $Y^k, Z^k \in \Omega$ such that $X^k = \frac{1}{2}Y^k + \frac{1}{2}Z^k$. Since Y^k and Z^k are both positive semidefinite we have that

$$\operatorname{range}(X^k) \supseteq \left\{ \operatorname{range}(Y^k), \operatorname{range}(Z^k) \right\}.$$

Thus there exist $V^k, W^k \in \mathcal{S}_+^{r_{k-1}}$ such that,

$$Y^k = U^0 \cdots U^{k-1} V^k (U^0 \cdots U^{k-1})^T,\ Z^k = U^0 \cdots U^{k-1} W^k (U^0 \cdots U^{k-1})^T,$$

and it follows that V^k and W^k are feasible for (2.22). By uniqueness of S^k we have that $Y^k = Z^k = X^k$ and X^k is an extreme point of Ω. Then by Theorem 2.4, $\operatorname{rank}(S^k) = 1$ and Algorithm 2.1 terminates before generating r_k, a contradiction. □

We remark that in many of our numerical tests the rank of $\bar{X}$ was 2 or 3. Consequently, the purification process did not require many iterations.

3 EDM Formulation

In this section we use the Lindenstrauss operator, $\mathcal{K}$, and the Schoenberg characterization to formulate **SLS** as an **EDM** completion problem. Recall that the exact locations of the sensors (towers) are known, and that the tower-source distances are noisy. The corresponding **EDM** restricted to the towers is denoted D_T and is defined by

$$(D_T)_{ij} := \|p^i - p^j\|^2, \quad \forall 1 \leq i, j \leq n.$$

Then the approximate **EDM** for the sensors and the source is

$$D_{T_c} := \begin{bmatrix} D_T & d \circ d \\ (d \circ d)^T & 0 \end{bmatrix} \in \mathbb{S}^{n+1}.$$

Recall that

$$P_T = \begin{bmatrix} p^1 & p^2 & \dots & p^n \end{bmatrix}^T \in \mathbb{R}^{n \times r}.$$

From Assumption 2.1 the towers are centered, i.e. $e^T P_T = 0$. This property is desirable due to the Schoenberg characterization which states that $\mathcal{K}$ is an isomorphism between $\mathcal{S}^n_+ \cap \mathcal{S}_C$ and $\mathcal{E}^n$. Moreover, it allows for easy recovery of the towers in the last step of our algorithm by solving a *Procrustes* problem.

Now let $G_T := P_T P_T^T$ be the Gram matrix restricted to the towers, and note that

$$\mathcal{K}(G_T) = D_T, \quad \mathcal{K}^\dagger(D_T) = G_T.$$

The nearest **EDM** problem with fixed sensors is

$$\min_{x \in \mathbb{R}^r} \frac{1}{2} \left\| \mathcal{K}\left(\begin{bmatrix} P_T \\ x^T \end{bmatrix} \begin{bmatrix} P_T \\ x^T \end{bmatrix}^T \right) - D_{Tc} \right\|^2. \tag{3.1}$$

For any $x \in \mathbb{R}^n$ let

$$d_x := \begin{pmatrix} \|x - p^1\|^2 \\ \vdots \\ \|x - p^n\|^2 \end{pmatrix}.$$

By simplifying the objective, we see that the **NEDMP** problem in (3.1) is indeed equivalent to **SLS**, i.e.,

$$\begin{aligned} \frac{1}{2} \left\| \mathcal{K}\left(\begin{bmatrix} P_T \\ x^T \end{bmatrix} \begin{bmatrix} P_T \\ x^T \end{bmatrix}^T \right) - D_{Tc} \right\|^2 &= \frac{1}{2} \left\| \mathcal{K}\left(\begin{bmatrix} G_T & P_T x \\ (P_T x)^T & 0 \end{bmatrix} \right) - D_{Tc} \right\|^2 \\ &= \frac{1}{2} \left\| \begin{bmatrix} D_T & d_x \\ d_x^T & 0 \end{bmatrix} - \begin{bmatrix} D_T & d \circ d \\ (d \circ d)^T & 0 \end{bmatrix} \right\|^2 \\ &= \sum_{i=1}^n \left(\|x - p^i\|^2 - d_i^2 \right)^2. \end{aligned}$$

The approach of [12] for the related sensor network localization problem is to replace the matrix $\begin{bmatrix} P_T \\ x^T \end{bmatrix} \begin{bmatrix} P_T \\ x^T \end{bmatrix}^T$ in (3.1) with the positive semidefinite matrix variable $X \in \mathcal{S}^{n+1}$, and then introduce a constraint on the block of X corresponding to the sensors. Taking this approach, we obtain *nearest Euclidean distance matrix with fixed sensors (**NEDMF**)* problem,

$$(\textbf{NEDMF}) \qquad \begin{aligned} V_S := \min \quad & \tfrac{1}{2} \|H_c \circ (\mathcal{K}(X) - D_{Tc})\|^2, \\ \text{s.t.} \quad & H_T \circ (\mathcal{K}(X) - D_{T_c}) = 0, \\ & \operatorname{rank}(X) \le r, \\ & X \succeq 0, \end{aligned} \tag{3.2}$$

where

$$H_T := \begin{bmatrix} ee^T - I & 0 \\ 0 & 0 \end{bmatrix}, \quad H_c := \begin{bmatrix} 0 & e \\ e^T & 0 \end{bmatrix}.$$

The objective of this (3.2) is exactly the objective of **SLS** (acting on the matrix variable) and the affine constraint restricts X to those Gram matrices for which the block corresponding to the sensors has exactly the same distances as $P_T P_T^T$. That is, if

$$X =: \begin{bmatrix} X_T \\ x_c^T \end{bmatrix} \begin{bmatrix} X_T \\ x_c^T \end{bmatrix}^T,$$

is feasible for (3.2), with $X_T \in \mathbb{R}^{n\times r}$ and $x_c \in \mathbb{R}^r$, then X_T differs from P_T only by translation and rotation. Since neither translation nor rotation affect the distances between the rows of X_T and x_c we translate the points in $\mathbb{R}^r$ so that X_T is centered. This corresponds to the assumption that P_T is centered. Then we solve the Procrustes problem

$$\min \{\|X_T Q - P_T\|^2 : Q^T Q = QQ^T = I,\ Q \in \mathbb{R}^{r\times r}\}, \tag{3.3}$$

to obtain the rotation and thus have a complete description of the transformation from X_T to P_T. Applying the transformation to x_c yields a vector feasible for **SLS**. Thus every feasible solution of (3.2), corresponds to a feasible solution of **SLS**. The converse is trivially true and we conclude that (3.2) due to the rank constraint, we show in the subsequent sections that the relaxation where the rank and the linear constraints are dropped, may be used to solve the problem accurately in a large number of instances.

3.1 The Relaxed NEDM Problem

3.1.1 Nearest Euclidean Distance Matrix Formulation

One relaxation of (3.2) is obtained by removing the affine constraint and modifying the objective as follows:

$$(\textit{NEDM}) \qquad \begin{aligned} \min\ & \frac{1}{2}||\,\mathcal{K}(X) - D_{T_c}||^2 \\ \text{s.t. }& \operatorname{rank}(X) \le r \\ & X \succeq 0. \end{aligned} \tag{3.4}$$

Due to the semidefinite characterization of $\mathcal{E}^{n+1}$ this problem is the projection of D_{T_c} onto the set of **EDM**s with embedding dimension at most r. The motivation behind this relaxation is the assumption that the distance measurements corresponding to the sensors are very accurate. Therefore, any minimizer of **NEDM** will likely have the first n points very near the sensors. As we show in the subsequent sections by introducing weights, we can obtain a solution arbitrarily close to that of (3.2).

The challenge in problem **NEDM** is the rank constraint. A simpler problem is to first solve the unconstrained least squares problem and then to project the solution onto the set of positive semidefinite matrices with rank at most r. This is equivalent to solving the inverse nearest

EDM problem:

$$(\textbf{\textit{NEDMinv}}) \qquad \begin{aligned} \min\ & \frac{1}{2}||X - \mathcal{K}^\dagger(D_{T_c})||^2 \\ \text{s.t. }& \operatorname{rank}(X) \le r \\ & X \succeq 0. \end{aligned} \tag{3.5}$$

Note that if the positive semidefinite constraint is removed, this problem is just the projection onto the matrices with rank at most r. By the Eckart-Young theorem, this projection is a rank r matrix obtained by setting the $n-r$ smallest eigenvalues (in magnitude) of $\mathcal{K}^\dagger(D_{T_c})$ to zero. In the following lemma we show that for sufficiently small noise, the negative eigenvalue is of small magnitude and hence the Eckart-Young rank r projection is positive semidefinite. We denote by $\bar{D} \in \mathcal{S}^{n+1}$ the true **EDM** of the sensors and the source, that is

$$\bar{D} := \begin{bmatrix} D_T & \bar{d} \circ \bar{d} \\ (\bar{d} \circ \bar{d})^T & 0 \end{bmatrix}. \tag{3.6}$$

It is easy to see from the definitions of $\bar{d}$ and ε that,

$$D_{T_c} = \bar{D} + e_{n+1}\xi^T + \xi e_{n+1}^T, \quad \text{where } \xi := \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix}.$$

Theorem 3.1. *The rank of $\mathcal{K}^\dagger(D_{T_c})$ is at most $r+2$. Moreover, $\mathcal{K}^\dagger(D_{T_c})$ has at most 1 negative eigenvalue with magnitude bounded above by $\frac{\sqrt{2}}{2}\|J_{n+1}\|^2\|\varepsilon\|$.*

Proof. First we note that the norm of $e_{n+1}\xi^T + \xi e_{n+1}^T$ is bounded above by the magnitude of the noise:

$$\|e_{n+1}\xi^T + \xi e_{n+1}^T\| = \sqrt{2\sum_{i=1}^{n} \varepsilon_i^2} = \sqrt{2}\|\varepsilon\|.$$

Next we observe that the matrix $e_{n+1}\xi^T + \xi e_{n+1}^T$ has trace 0 and rank 2. Thus $e_{n+1}\xi^T + \xi e_{n+1}^T$ has exactly one negative and one positive eigenvalue. By the Moreau decomposition theorem, e.g. [23], $e_{n+1}\xi^T + \xi e_{n+1}^T$ may be expressed as the sum of two rank one matrices, say $P \succeq 0$ and $Q \preceq 0$, that are the projections of $e_{n+1}\xi^T + \xi e_{n+1}^T$ onto $\mathcal{S}_+^{n+1}$ and $-\mathcal{S}_+^{n+1}$, respectively. Now we have,

$$\mathcal{K}^\dagger(D_{T_c}) = -\frac{1}{2}J_{n+1}D_{T_c}J_{n+1} = -\frac{1}{2}(J_{n+1}\bar{D}J_{n+1} + J_{n+1}QJ_{n+1}) + (-\frac{1}{2}J_{n+1}PJ_{n+1}),$$

where the first term is positive semidefinite with at least r and at most $r+1$ positive eigenvalues and the second term has one negative eigenvalue. Using the Cauchy-Schwartz inequality it can be shown that for $X, Y \in \mathcal{S}^n$,

$$\|XY\| \le \|X\|\|Y\|. \tag{3.7}$$

By (3.7) and the fact that P is a projection of $e_{n+1}\xi^T + \xi e_{n+1}^T$ onto $-\mathcal{S}_+^{n+1}$, we have

$$\left\| -\frac{1}{2}J_{n+1}PJ_{n+1} \right\| \le \frac{1}{2}\|J_{n+1}\|^2\|P\| \le \frac{1}{2}\|J_{n+1}\|^2\|e_{n+1}\xi^T + \xi e_{n+1}^T\| = \frac{\sqrt{2}}{2}\|J_{n+1}\|^2\|\varepsilon\|.$$

It follows that $\mathcal{K}^\dagger(D_{T_c})$ has rank at most $r+2$ and by the Courant-Fischer-Weyl theorem, e.g. [37], it has at most one negative eigenvalue whose magnitude is bounded above by $\frac{\sqrt{2}}{2}\|J_{n+1}\|^2\|\varepsilon\|$, as desired. □

The following corollary follows immediately.

Corollary 3.2. *If $\|\varepsilon\|$ is sufficiently small, the optimal solution of **NEDMinv** is the rank r Eckart-Young projection of $\mathcal{K}^\dagger(D_{T_c})$.*

3.1.2 Weighted, Facially Reduced NEDM

While we have discarded the information pertaining to the locations of the sensors in relaxing the problem (3.2) to the problem **NEDM**, we still make use of the distances between the sensors. Thus, to some extent the locations of the sensors have an implicit effect on the optimal solution of **NEDM** and the approximation **NEDMinv** from the previous section. In this section we take greater advantage of the known distances between the sensors by restricting **NEDM** to a face of $\mathcal{S}_+^{n+1}$ by *facial reduction.*

The true Gram matrix, $\mathcal{K}^\dagger(\bar{D})$, belongs to the set,

$$F_T := \{X \in \mathcal{S}_{c,+}^{n+1} : \mathcal{K}(X)_{1:n,1:n} = D_T\}. \tag{3.8}$$

Now the constraint $X \succeq 0$ in **NEDM**, may actually be refined to say $X \in \text{face}(F_T, \mathcal{S}_+^{n+1})$ which is the following:

$$(\textbf{\textit{NEDMP}}) \qquad \begin{array}{ll} \min & \frac{1}{2}||\,\mathcal{K}(X) - D_{T_c}||^2 \\ \text{s.t.} & \text{rank}(X) \le r \\ & X \in \text{face}(F_T, \mathcal{S}_+^{n+1}). \end{array} \tag{3.9}$$

Moreover, we may obtain a closed form expression for $\text{face}(F_T, \mathcal{S}_+^{n+1})$ in the form of an exposing vector. To see this, consider the spectral decomposition of the sensor Gram matrix,

$$G_T =: \begin{bmatrix} U & \frac{1}{\sqrt{n}}e & W_T \end{bmatrix} \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U & \frac{1}{\sqrt{n}}e & W_T \end{bmatrix}^T, \quad U^TU = I_r,\ U^Te = 0,\ \Lambda \in \mathcal{S}_{++}^r.$$

Note that $W_TW_T^T$ is an exposing vector for $\text{face}(G_T, \mathcal{S}_{c,+}^n)$ since the following two conditions hold:

$$\langle G_T, W_TW_T^T \rangle = 0,\ \text{rank}(G_T + W_TW_T^T) = n - 1 = \max_{X \in \mathcal{S}_{c,+}^n} \text{rank}(X).$$

We now extend $W_TW_T^T$ to an exposing vector for $\text{face}(F_T, \mathcal{S}_+^{n+1})$.

Lemma 3.3. *Let $\overline{W}_T := [W_T^T \quad 0]^T$ and let $W := \overline{W}_T\overline{W}_T^T + ee^T$. Then,*

1. *$\overline{W}_T\overline{W}_T^T$ exposes $\text{face}(F_T, \mathcal{S}_{c,+}^{n+1})$,*
2. *W exposes $\text{face}(F_T, \mathcal{S}_+^{n+1})$.*

Proof. This statement is a special case of Theorem 4.13 of [15]. □

Note that $\text{face}(\mathcal{K}^\dagger(\bar{D}), \mathcal{S}_+^{n+1}) \subsetneq \text{face}(F_T, \mathcal{S}_+^{n+1})$ since,

$$\text{rank}(\mathcal{K}^\dagger(\bar{D})) + \text{rank}(W) = r + n - r < n + 1.$$

Through W we have a 'nullspace' characterization of face$(F_T, \mathcal{S}_+^{n+1})$. However, the 'range space' characterization is more useful in the context of semidefinite optimizaiton as it leads to dimension reduction, numerical stability, and strong duality. To this end, we consider any $(n+1)\times(r+1)$ matrix such that its columns form a basis for null(W). One such choice is,

$$V = J_{n+1}\begin{bmatrix} P_T & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} P_T & -\frac{1}{n+1}e \\ 0 & 1-\frac{1}{n+1} \end{bmatrix}. \tag{3.10}$$

To verify that the columns of V indeed form a basis for null(W), we first observe that rank$(V) = r+1$ and secondly we have,

$$\begin{aligned} WV &= \left(\begin{bmatrix} W_T W_T^T & 0 \\ 0 & 0 \end{bmatrix} + ee^T\right)\begin{bmatrix} P_T & -\frac{1}{n+1}e \\ 0 & 1-\frac{1}{n+1} \end{bmatrix} \\ &= \begin{bmatrix} W_T W_T^T P_T - \frac{1}{n+1}W_T W_T^T e \\ 0 \end{bmatrix} + ee^T\begin{bmatrix} P_T & -\frac{1}{n+1}e \\ 0 & 1-\frac{1}{n+1} \end{bmatrix} \\ &= 0. \end{aligned}$$

It follows that

$$\text{face}(F_T, \mathcal{S}_+^{n+1}) = \mathcal{S}_+^{n+1} \cap W^\perp = V\mathcal{S}_+^{r+1}V^T. \tag{3.11}$$

Thus we may replace the variable X in **NEDMP** by VRV^T for $R \in \mathcal{S}_+^{r+1}$. To simplify the notation, we define the composite map $\mathcal{K}_V := \mathcal{K}(V \cdot V^T)$. Moreover, we introduce a weight matrix to the objective and obtain the *weighted facially reduced problem,* ***FNEDM***,

$$(\textbf{\textit{FNEDM}}) \qquad \begin{aligned} V_\alpha := \min\ & \frac{1}{2}||H_\alpha \circ (\mathcal{K}_V(R) - D_{T_c})||^2, \qquad (=: f(R,\alpha)) \\ \text{s.t. }& \operatorname{rank} R \le r, \\ & R \succeq 0. \end{aligned} \tag{3.12}$$

Here $H_\alpha := \alpha H_T + H_c$ and α is positive. Let us make a few comments regarding this problem. When $\alpha = 1$ the weight matrix has no effect and **FNEDM** reduces to **NEDMP** . On the other hand, when α is very large, the solution has to satisfy the distance constraints for the sensors more accurately and in this case **FNEDM** approximates (3.2). In fact, in Theorem 3.9 we prove that the solution to **FNEDM** approaches that of (3.2) as α increases.

We begin our analysis by proving that V_α is attained.

Lemma 3.4. *Let $\alpha > 0$. Then*

1. null$(H_\alpha \circ \mathcal{K}_V) = \{0\}$,
2. *$f(R,\alpha)$ is strictly convex and coercive,*
3. *the problem* ***FNEDM*** *admits a minimizer.*

Proof. For Item 1, under the assumption that $\alpha > 0$, we have $H_\alpha \circ \mathcal{K}_V(R) = 0$ if, and only if, $\mathcal{K}_V(R) = 0$. Recall that $\mathcal{K}$ is one-to-one between the centered and hollow subspaces and $\mathcal{K}(0) = 0$. By construction, range$(V \cdot V^T)$ is a subset of the centered matrices. Hence $H \circ \mathcal{K}_V(R) = 0$ if, and only if, $VRV^T = 0$. Since V is full column rank, $VRV^T = 0$ if, and only if, $R = 0$, as desired.

Now we turn to Item 2. The function $f(R, \alpha)$ is quadratic with a positive semidefinite second derivative. Moreover, by Item 1, the second derivative is positive definite. Therefore $f(R, \alpha)$ is strictly convex and coercive.

Finally, the feasible set of **FNEDM** is closed. Combining this observation with coercivity of the objective, from Item 2, we obtain Item 3. □

We conclude this subsection by deriving the optimality conditions for the convex relaxation of **FNEDM** , which is obtained by dropping the rank constraint.

Lemma 3.5. *The matrix $R \in \mathcal{S}_+^{r+1}$ is optimal for the relaxation of* (3.12) *obtained by ignoring the rank constraint if, and only if,*

$$0 \preceq \nabla f(R) = V^T \left(H_\alpha^* \circ \mathcal{K}^* \left[(H_\alpha \circ \mathcal{K})(VRV^T) - H_\alpha \circ D_{T_c} \right] \right) V, \qquad \langle \nabla f(R), R \rangle = 0.$$

In addition, R is optimal for (3.12) *if* $\operatorname{rank} R \leq r$.

Proof. From the Pshenichnyi-Rockafellar conditions, R is optimal if, and only if, $\nabla f(R) \in (\mathcal{S}_+^{r+1} - R)^+$, the nonnegative polar cone. This condition holds if, and only if, for all $X \in \mathcal{S}_+^{r+1}$ and $\alpha > 0$, we have

$$0 \leq \langle \nabla f(R), \alpha X - R \rangle = \alpha \langle \nabla f(R), X \rangle - \langle \nabla f(R), R \rangle.$$

which implies that $\alpha \langle \nabla f(R), X \rangle \geq \langle \nabla f(R), R \rangle$ for every $\alpha > 0$. Since α may be arbitrarily large we get that $\langle \nabla f(R), X \rangle \geq 0$ for all $X \in \mathcal{S}_+^{r+1}$. Therefore, we conclude that $\nabla f(R) \in (\mathcal{S}_+^{r+1})^+ = \mathcal{S}_+^{r+1}$. Moreover, setting $X = 0$, we get,

$$0 \leq \langle \nabla f(R), 0 - R \rangle = -\langle \nabla f(R), R \rangle \leq 0,$$

hence orthogonality holds. □

3.1.3 Analysis of FNEDM

In this section we show that the optimal value of **FNEDM** is a lower bound for the optimal value of **SLS**. Moreover, the this lower bound becomes exact as α is increased to $+\infty$.

In the **SLS** model, the distances between the towers are fixed, while in the **NEDM** model (3.4), the distances between towers are free. The facial reduction model allows the distances between the towers to change but the towers can still be transformed back to their original positions by a square matrix $Q \in \mathbb{R}^{r \times r}$. Note that Q does not have to be orthonormal, so it is possible that $QQ^T \neq I$.

Theorem 3.6. *Let P_T be as above, V as in* (3.10)*, and let P be a centered matrix with,*

$$P = \begin{bmatrix} T \\ c^T \end{bmatrix}, \ T \in \mathbb{R}^{n \times r}, \ c \in \mathbb{R}^r.$$

Then there exists a matrix $Q \in \mathbb{R}^{r \times r}$ such that $P_T Q = J_n T$ if, and only if,

$$PP^T \in V \mathcal{S}_+^{r+1} V^T.$$

Proof. Since P is centered,

$$0 = P^T e = T^T e + c.$$

Substituting into the equation $P_T Q = J_n T$ we get,

$$P_T Q = J_n T = T - \frac{1}{n} e e^T T = T - \frac{1}{n} e c^T,$$

which yields the following expression for P,

$$P = \begin{bmatrix} P_T Q + \frac{1}{n} e c^T \\ c^T \end{bmatrix}. \tag{3.13}$$

Now by (3.11) we have,

$$PP^T \in V \mathcal{S}_+^{r+1} V^T \iff PP^T \in \mathcal{S}_+^{n+1} \cap W^\perp \iff P^T W = 0.$$

Applying (3.13) we verify that the last statement in the equivalence holds,

$$\begin{aligned} P^T W &= P^T (\overline{W}_T \overline{W}_T^T + e e^T) \\ &= P^T \overline{W}_T \overline{W}_T^T \\ &= \begin{bmatrix} Q^T P_T^T + \frac{1}{n} c e^T & c \end{bmatrix} \begin{bmatrix} W_T \\ 0 \end{bmatrix} \overline{W}_T^T \\ &= \left(Q^T P_T^T W_T + \frac{1}{n} c e^T W_T \right) \overline{W}_T^T \\ &= 0, \end{aligned}$$

as desired.

For the other direction, let

$$V_1 := \begin{bmatrix} P_T & 0 \\ 0 & 1 \end{bmatrix},$$

and recall that $V = J_{n+1} V_1$. Suppose PP^T belongs to the face $V \mathcal{S}_+^{r+1} V^T$. Then $P = J_{n+1} V_1 M$ for some $M \in \mathbb{R}^{(r+1)\times r}$. We show that if $Q \in \mathbb{R}^{r\times r}$ denotes the first r rows of M, then $P_T Q = J_n T$. To this end, let $\bar{J} = [J_n \quad 0]$ and observe that $\bar{J} P = J_n T$. Moreover, since $\bar{J}$ is centered, $\bar{J} J_{n+1} = \bar{J}$. Then,

$$J_n T = \bar{J} P = \bar{J} J_{n+1} V_1 M = \bar{J} V_1 M = J_n P_T Q = P_T Q,$$

as desired. □

Theorem 3.6 indicates that when using the facial reduction model **FNEDM** we can use a least square approach to exactly get back the original positions of the sensors. This approach will be discussed in section 3.2 along with the Procrustes approach.

In the following, we show that the optimal value of the problem in (3.12) is not greater than the optimal value of the **SLS** estimates (2.2) or (3.2). We also prove that the solution to **FNEDM** approaches that of (3.2) as α increases.

Lemma 3.7. *Consider the problem,*

$$\begin{aligned} V_T := \min \ & \frac{1}{2} ||H_c \circ (\mathcal{K}_V(R) - D_{T_c})||^2 \quad (=: h(R)) \\ \text{s.t. } & H_T \circ (\mathcal{K}_V(R) - D_{T_c}) = 0 \\ & \operatorname{rank} R \leq r \\ & R \in \mathcal{S}_+^{r+1}. \end{aligned} \tag{3.14}$$

Then V_T is finite and satisfies $V_T = V_S$.

Proof. That V_T is finite, follows from arguments analogous to those used in Lemma 3.4.

For the equality claim, it is clear that $V_S \leq V_T$. To show that $V_S \geq V_T$, consider X that is feasible for (3.2). First we show that X may be assumed to be centered. To see this, consider $\hat{X} = J_n X J_n$. Note that $\hat{X}$ is the orthogonal projection of X onto $\mathcal{S}_c$ and it can be verified that $\hat{X} = \mathcal{K}^\dagger \mathcal{K}(X)$. Now it is clear that $\hat{X} \succeq 0$ and that $\mathcal{K}(\hat{X}) = \mathcal{K}(X)$. Moreover, since J_n is singular we have, $\text{rank}(\hat{X}) \leq \text{rank}(X)$. Therefore, $\hat{X}$ is also feasible for (3.2) and provides the same objective value as X.

Now there exists $T \in \mathbb{R}^{n \times r}$ and $c \in \mathbb{R}^r$ such that,

$$X = \begin{bmatrix} T \\ c^T \end{bmatrix} \begin{bmatrix} T \\ c^T \end{bmatrix}^T.$$

Then, from the tower constraint of (3.2) we get the implications,

$$\mathcal{K}(TT^T) = D_T \implies \mathcal{K}^\dagger \mathcal{K}(TT^T) = \mathcal{K}^\dagger(D_T) \implies J_n T T^T J_n = P_T P_T^T.$$

Thus, there exists an orthogonal Q such that $J_n T = P_T Q$. By Theorem 3.6 we have $X \in V \mathcal{S}_+^{r+1} V^T$ and it follows that $V_S \geq V_T$. □

Lemma 3.8. *Let $0 < \alpha_1 < \alpha_2$. Then,*

$$V_{\alpha_1} < V_{\alpha_2} \leq V_T.$$

*Moreover, the first inequality is strict if D_{T_c} is not an **EDM** with embedding dimension r.*

Proof. For the first inequality, let $R \succeq 0$ be such that $\text{rank}(R) \leq r$. Then,

$$\begin{aligned} f(R, \alpha_1) &= \frac{1}{2} \|H_{\alpha_1} \circ (\mathcal{K}_V(R) - D_{T_c})\|^2 \\ &= \frac{1}{2} \|(\alpha_1 H_T + H_c) \circ (\mathcal{K}_V(R) - D_{T_c})\|^2 \\ &= \frac{1}{2} \alpha_1^2 \|H_T \circ (\mathcal{K}_V(R) - D_{T_c})\|^2 + \frac{1}{2} \|H_c \circ (\mathcal{K}_V(R) - D_{T_c})\|^2 \\ &\leq \frac{1}{2} \alpha_2^2 \|H_T \circ (\mathcal{K}_V(R) - D_{T_c})\|^2 + \frac{1}{2} \|H_c \circ (\mathcal{K}_V(R) - D_{T_c})\|^2 \\ &= f(R, \alpha_2). \end{aligned}$$

Note that the inequality is strict if, and only if, $\|H_T \circ (\mathcal{K}_V(R) - D_{T_c})\|$ is positive. This holds for all R, if D_{T_c} is not an **EDM** with embedding dimension r.

For the second inequality, we first observe that,

$$\begin{aligned} V_T = \min\ & f(R, \alpha_2) \\ \text{s.t. } & H_T \circ (\mathcal{K}_V(R) - D_{T_c}) = 0 \\ & \text{rank}(R) \leq r \\ & R \succeq 0. \end{aligned}$$

Now V_T and V_{α_2} are both optimal values of $f(R, \alpha_2)$ over their respective domains, but the domain for V_T is smaller than that of V_{α_2}. Hence, the second inequality holds. □

Theorem 3.9. *For any $\alpha > 0$, let R_α denote the minimizer of **FNEDM**. Let $\{\alpha_\ell\}_{\ell\in\mathbb{N}} \subset \mathbb{R}_{++}$ be a sequence of increasing numbers such that $R_{\alpha_\ell} \to \bar{R}$ for some $\bar{R} \in \mathcal{S}^{r+1}$. Then $V_\alpha \uparrow V_T$ and $\bar{R}$ is a minimizer of* (3.14).

Proof. First we note that R_α is well defined by Lemma 3.4. Now, from Lemma 3.8, we have that V_α is monotonically increasing and bounded above by V_T. Hence there exists V^* such that,

$$V_\alpha \uparrow V^* \leq V_T. \tag{3.15}$$

Next, we show that $\bar{R}$ is feasible for (3.14). Since $\mathcal{S}_+^{r+1}$ is closed and the rank function is lower semicontinuous, we have $\text{rank}(\bar{R}) \leq r$ and $\bar{R} \succeq 0$. Moreover, for every $\ell \in \mathbb{N}$,

$$V_{\alpha_\ell} = f(R_{\alpha_\ell}, \alpha_\ell) = \frac{1}{2}\alpha_\ell^2 ||H_T \circ (\mathcal{K}_V(R_{\alpha_\ell}) - D_{T_c})||^2 + h(R_{\alpha_\ell})$$

Rearranging and taking the limit we get,

$$0 \leq \lim_{\ell\to+\infty} \frac{1}{2}\alpha_\ell^2 ||H_T \circ (\mathcal{K}_V(R_{\alpha_\ell} V^T) - D_{T_c})||^2 = \lim_{\ell\to+\infty} V_{\alpha_\ell} - h(R_{\alpha_\ell}) = V^* - h(\bar{R}). \tag{3.16}$$

The last equality follows from the continuity of h. Since the limit in (3.16) exists we get,

$$0 = \lim_{\ell\to+\infty} ||H_T \circ (\mathcal{K}_V(R_{\alpha_\ell}) - D_{T_c})|| = ||H_T \circ (\mathcal{K}_V(\bar{R}) - D_{T_c})||, \tag{3.17}$$

by continuity. Thus $\bar{R}$ is feasible for (3.14) and we have $h(\bar{R}) \geq V_T$. On the other hand, from (3.16) we have $h(\bar{R}) \leq V^*$. Combining these observations with (3.15) we get,

$$h(\bar{R}) \leq V^* \leq V_T \leq h(\bar{R}). \tag{3.18}$$

Now equality holds throughout (3.18) and the desired results are immediate. □

3.1.4 Solving FNEDM

The solution set of the unconstrained version of (3.12) can be stated in terms of the Moore-Penrose generalized inverse of $H_\alpha \circ \mathcal{K}_V$, denoted by $(H_\alpha \circ \mathcal{K}_V)^\dagger$. Indeed, the solution to the least squares problem is,

$$R_{LS} := (H_\alpha \circ \mathcal{K}_V)^\dagger (H_\alpha \circ D_{T_c}) \in \operatorname{argmin} f(R). \tag{3.19}$$

In this subsection we explore the relationship between the optimal solution of **FNEDM** and the eigenvalues of R_{LS}. In general the Moore-Penrose inverse may be difficult to obtain, however, the following result implies that R_{LS} may be derived efficiently and it is the *unique* minimizer of f.

Lemma 3.10. *Let R_{LS} be as in* (3.19). *Then, R_{LS} is the unique minimizer of f and*

$$R_{LS} = ((H_\alpha \circ \mathcal{K}_V)^*(H_\alpha \circ \mathcal{K}_V))^{-1}(H_\alpha \circ \mathcal{K}_V)^*(H_\alpha \circ D_{T_c}).$$

Proof. That R_{LS} is the unique minimizer of f follows from strict convexity as in Item 2 of Lemma 3.4. Moreover, by Item 1 of Lemma 3.4, we have $\text{null}(H_\alpha \circ \mathcal{K}_V) = \{0\}$ which implies that $(H_\alpha \circ \mathcal{K}_V)^\dagger$ is the left inverse. The desired expression for R_{LS} is obtained by substituting the left inverse into (3.19). □

Note that $(H_\alpha \circ \mathcal{K}_V)^*(H_\alpha \circ \mathcal{K}_V)$ admits an $r \times r$ matrix representation. Thus if r is small, as in many applications, the inverse of $(H_\alpha \circ \mathcal{K}_V)^*(H_\alpha \circ \mathcal{K}_V)$, and consequently R_{LS}, may be obtained efficiently.

We consider three cases regarding the eigenvalues of R_{LS}, each of which corresponds to a different approach to solving **FNEDM**.

Case I: $R_{LS} \succeq 0$ and $\text{rank}(R_{LS}) \leq r$.

Case II: $R_{LS} \notin \mathcal{S}_+^{r+1}$.

Case III: $R_{LS} \succ 0$.

In the best scenario, **Case I**, we have that R_{LS} is the unique minimizer of **FNEDM**. In this case **FNEDM** reduces to an unconstrained convex optimization problem. Moreover, we have a closed form solution for the minimizer, R_{LS}. In **Case II**, the minimizer of **FNEDM** may also be obtained through a convex relaxation as is indicated by the following result.

Theorem 3.11. *Let $R^\star$ denote the minimizer of the relaxation of **FNEDM** where the rank constraint is removed. If $R_{LS} \notin \mathcal{S}_+^{r+1}$, then $R^\star$ is a minimizer of **FNEDM**.*

Proof. Let $R^\star$ denote the optimal solution of **FNEDM** without the rank constraint. Note that $R^\star$ exists by arguments analogous to those in Lemma 3.4. If $\text{rank}(R^\star) \leq r$, then clearly $R^\star$ is a minimizer of **FNEDM**. Thus we may assume that $R^\star \succ 0$.

Since R_{LS} is the unique minimizer of f, we have $f(R_{LS}) < f(R^\star)$. Moreover, by strict convexity of f, every matrix R in the relative interior of the line segment $[R_{LS}, R^\star]$ satisfies $f(R) < f(R^\star)$. Now since $R^\star \succ 0$ there exists $\bar{R} \in \text{relint}[R_{LS}, R^\star] \cap \mathcal{S}_+^{r+1}$. Then, $\bar{R}$ is feasible for the relaxation of **FNEDM** where the rank constraint is removed. However, $f(\bar{R}) < f(R^\star)$, contradicting the optimality of $R^\star$. □

In **Case III** we are motivated by the primal-dual approach of [30,31] and the penalty approach of [19, 30, 31]. Let $h = [1, \cdots, \alpha]^T$, we notice that $H_\alpha \circ Y = hh^T \circ Y = \text{Diag}(h) Y \,\text{Diag}(h)$ if $\text{diag}(Y) = 0$. Let $T = \text{Diag}(h)$, it is easy to see that (3.12) is equivalent to the problem:

$$\begin{aligned} \min\ & \frac{1}{2}||T(Y - D_{T_c})T||^2 \\ \text{s.t. } & \text{diag}(Y) = 0, \\ & \langle J_{n+1} Y J_{n+1}, W \rangle = 0, \\ & -J_{n+1} Y J_{n+1} \succeq 0, \\ & \text{rank}(J_{n+1} Y J_{n+1}) \leq r. \end{aligned} \tag{3.20}$$

As in [30,31], we define $\mathcal{K}_+^{n+1}(r) := \{Y \in \mathcal{S}^{n+1} | -J_{n+1} Y J_{n+1} \succeq 0,\ \text{rank}(J_{n+1} Y J_{n+1}) \leq r\}$, and let $\mathcal{B}(Y) = 0$ represent the linear constraints $\text{diag}(Y) = 0$ and $\langle J_{n+1} Y J_{n+1}, W \rangle = 0$. Then (3.20) may be written as,

$$\begin{aligned} \min\ & \frac{1}{2}||T(Y - D_{T_c})T||^2 \\ \text{s.t. } & \mathcal{B}(Y) = 0 \\ & Y \in \mathcal{K}_+^{n+1}(r), \end{aligned} \tag{3.21}$$

If $\mathcal{B}$ consists only of the diagonal constraint and $T = I$, then (3.21) is exactly the problem considered in [30, 31], where a sufficient condition for strong duality was presented. In the subsequent results, we present an analogous dual problem for the general constraint $\mathcal{B}(Y) = 0$.

Lemma 3.12. *The Lagrangian dual of* (3.21) *is*

$$-\min_y \frac{1}{2}\| \prod_{\mathcal{K}_T^{n+1}(r)} (TD_{T_c}T + \mathcal{B}^*(y))\|^2 - \frac{1}{2}\|TD_{T_c}T\|^2, \tag{3.22}$$

where $\mathcal{K}_T^{n+1}(r) = \{Y \in \mathcal{S}^{n+1} | -Y \succeq 0 \text{ on } \{Te\}^\perp, \text{ rank}(J_{n+1}T^{-1}YT^{-1}J_{n+1}) \leq r\}$

Proof. The Lagrangian function $L : \mathcal{S}^{n+1} \times \mathbb{R}^{n+2} \to \mathbb{R}$ of (3.21) is,

$$\begin{aligned} L(Y, y) = &\ \frac{1}{2}\|T(Y - D_{T_c})T\|^2 - \langle \mathcal{B}(Y), y\rangle \\ = &\ \frac{1}{2}\left(\|T(Y - (D_{T_c} + T^{-1}\mathcal{B}^*(y)T^{-1}))T\|^2 + \|TD_{T_c}T\|^2 - \|T(D_{T_c} + T^{-1}\mathcal{B}^*(y)T^{-1})T\|^2\right). \end{aligned} \tag{3.23}$$

we then write the dual object function $\theta : \mathbb{R}^{n+2} \to \mathbb{R}$ as,

$$\theta(y) := \min_{Y \in \mathcal{K}_+^{n+1}(r)} L(Y, y) \tag{3.24}$$

$$= \frac{1}{2}\| \prod_{\mathcal{K}_T^{n+1}(r)} (TD_{T_c}T + \mathcal{B}^*(y)) - (TD_{T_c}T + \mathcal{B}^*(y))\|^2 \tag{3.25}$$

$$-\frac{1}{2}\|TD_{T_c}T + \mathcal{B}^*(y)\|^2 + \frac{1}{2}\|TD_{T_c}T\|^2 \tag{3.26}$$

$$= -\frac{1}{2}\| \prod_{\mathcal{K}_T^{n+1}(r)} (TD_{T_c}T + \mathcal{B}^*(y))\|^2 + \frac{1}{2}\|TD_{T_c}T\|^2 \tag{3.27}$$

From equation (3.26) to equation (3.27) we need to prove the triangle equality holds, i.e.,

$$\begin{aligned} \| \prod_{\mathcal{K}_T^{n+1}(r)} (TD_{T_c}T + \mathcal{B}^*(y)) - (TD_{T_c}T + \mathcal{B}^*(y))\|^2 + \| \prod_{\mathcal{K}_T^{n+1}(r)} (TD_{T_c}T + \mathcal{B}^*(y))\|^2 \\ = \|TD_{T_c}T + \mathcal{B}^*(y)\|^2. \end{aligned}$$

To this end, consider any matrix $X \in \mathcal{S}^{n+1}$ and let $\prod(X)$ be a nearest point in $\mathcal{K}_T^{n+1}(r)$ to X. Since $\mathcal{K}_T^{n+1}(r)$ is a cone, the ray $\theta \prod(X)$ for all $\theta \geq 0$ is contained in the set $\mathcal{K}_T^{n+1}(r)$. Moreover this ray is convex and $\prod(X)$ is the nearest point to X from this ray. Now we can use orthogonality: $\prod(X) - X$ is orthogonal to $\prod(X) - 0$. Then the triangle inequality follows:

$$\| \prod(X) - X\|^2 + \| \prod(X)\|^2 = \|X\|^2.$$

The Lagrangian dual problem is then defined by,

$$V_d := \max \theta(y) = -\min_y \frac{1}{2}\| \prod_{\mathcal{K}_T^{n+1}(r)} (TD_{T_c}T + \mathcal{B}^*(y))\|^2 - \frac{1}{2}\|TD_{T_c}T\|^2, \tag{3.28}$$

as desired. □

In [30,31] it is shown that the Lagrangian dual has compact level sets and therefore the optimal value is finite and attained. The dual problem (3.28) can be solved by the semi-smooth Newton approach proposed in [30].

In [30, 31], the authors proposed a rank majorization approach where strong duality is guaranteed if the penalty function goes to zero. The approach can be readily modified to replace the diagonal constraint by the linear constraint $\mathcal{B}$ and to include the diagonal weight matrix T. The strong duality result and global optimal condition can also be carried out to our problem (3.21). The drawback of this approach is the slow convergence when n is large. Therefore, in our facial reduction model we prefer to stay in $\mathcal{S}^{r+1}$ rather than $\mathcal{S}^{n+1}$ since the dimension is lower. Hence we develop a rank majorization approach in $\mathcal{S}^{r+1}$ in the following:

To penalize rank, we consider the concave penalty function,

$$p : \mathcal{S}^{r+1} \to \mathbb{R}, \quad p(R) := \langle I, R\rangle - \sum_{i=1}^{r} \lambda_i(R). \tag{3.29}$$

Note that p is non-negative over the positive semidefinite matrices and

$$R \succeq 0, \quad \operatorname{rank}(R) \leq r \iff p(R) = 0 \quad R \succeq 0.$$

Hence, p is an appropriate penalty function for the rank constraint of **FNEDM**. Now we consider the penalized version of **FNEDM** ,

$$(\boldsymbol{PNEDM}) \qquad \begin{array}{l} \min \ \frac{1}{2}||H_\alpha \circ (\mathcal{K}_V(R)) - D_{T_c})||^2 + \gamma p(R), \\ \text{s.t.} \quad R \succeq 0. \end{array} \tag{3.30}$$

where γ is a positive constant. The objective is a difference of convex functions and the feasible set is convex. The literature on this type of optimization problem is extensive and the theory well established. In particular, the well-known *majorization* approach guarantees convergence to a matrix satisfying the first order necessary conditions for **PNEDM**, i.e., a stationary point. See for instance [35, 36].

The majorization approach is outlined below in Algorithm 3.1. Central to the approach is the observation that p is majorized by its linear approximation, since it is concave. In the algorithm, $\partial p(R)$ denotes the subdifferential of p at R. Thus at every iterate, the convex subproblem (3.31) is solved to obtain the next iterate.

Theorem 3.13. *Suppose Algorithm 3.1 converges to a stationary point* $\bar{R}$*, and that* $\operatorname{rank}(\bar{R}) = r$*. Then* $\bar{R}$ *is a global minimizer of* ***FNEDM*** *restricted to* $\operatorname{face}(\bar{R})$*.*

Proof. By [35, 36], the stationary point $\bar{R}$ satisfies the following condition:

$$(\nabla f(\bar{R}, \alpha) + \mathcal{N}_{\mathcal{S}_+^{r+1}}(\bar{R})) \cap (\gamma \partial p(\bar{R})) \neq \emptyset. \tag{3.32}$$

Under the assumption $\operatorname{rank}(\bar{R}) = r$, we have $\bar{R} = V \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} V^T$ where $\Lambda = \operatorname{Diag}(\lambda_1, ..., \lambda_r)$ with $\lambda_1 \geq ... \geq \lambda_r > 0$ being the eigenvalues of Z and $V^T V = I$. Let $V = [V_1, V_2]$ with the columns of V_1 being the eigenvectors corresponding to $\lambda_1, ..., \lambda_r$. We have

$$\mathcal{N}_{\mathcal{S}_+^{r+1}}(\bar{R}) = \{V \begin{bmatrix} 0 & 0 \\ 0 & t \end{bmatrix} V^T : t \geq 0\}$$

Algorithm 3.1 Majorization Algorithm

INPUT: $R_0 \succeq 0$, $\gamma >> 0$, $1 > \epsilon > 0$
initialize: $k = 0$, $err = 1$
while $err > \epsilon$ **do**
  Choose $U^k \in \partial p(R^k)$
  Obtain R^{k+1},

$$R^{k+1} \in \underset{R \succeq 0}{\operatorname{argmin}} \ \frac{1}{2}||H_\alpha \circ (\mathcal{K}_V(R)) - D_{T_c})||^2 + \gamma(p(R^k) + \langle U^k, R - R^k \rangle) \tag{3.31}$$

  Update $err \leftarrow \|R^{k+1} - R^k\|$, $k \leftarrow k + 1$
end while

and

$$\partial p(\bar{R}) = I - \begin{bmatrix} V_1 \\ 0 \end{bmatrix} \begin{bmatrix} V_1^T & 0 \end{bmatrix} = V \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} V^T.$$

Therefore we have $\nabla f(\bar{R}, \alpha) = V \begin{bmatrix} 0 & 0 \\ 0 & \gamma - t \end{bmatrix} V^T$.

Due to the convexity of $f(R, \alpha)$, for any $\hat{R} \in \text{face}(\bar{R})$, we have

$$f(\hat{R}, \alpha) \geq f(\bar{R}, \alpha) + \langle \nabla f(\bar{R}, \alpha), \hat{R} - \bar{R} \rangle = f(\bar{R}, \alpha) + 0.$$

Hence our claim is proved. □

3.1.5 Identifying Outliers using l_1 Minimization and Facial Reduction

In this section, we address the issue of unequal noise, where a few distance measurements are outliers, i.e., much more inaccurate than others. We use l_1 norm minimization to try and identify the outliers, and remove them to obtain a more stable problem. We assume that we have many more towers available than is necessary, so that removal of a few outliers leaves us with towers that still satisfy Assumption 2.1.

Problem (3.12) is equivalent to minimizing the residual of an overdetermined linear system in the domain of an **SDP** cone. Let $z := \text{svec}(R)$ for $R \in \mathcal{S}^{r+1}$. Abusing our previous notation, let $b := \text{svec}(H_\alpha \circ D_{T_c})$ and let A denote the matrix representation of $H_\alpha \circ \mathcal{K}_V$. Then $z \in \mathbb{R}^{(r+1)(r+2)/2}$ and $b \in \mathbb{R}^{n(n+1)/2}$. In practice, n is much larger than $r+1$, so A will have more rows than columns. In other words, we have an overdetermined system. Under this new notation, problem (3.12) is equivalent to,

$$\begin{aligned} &\min \ \|\delta\| \\ &\text{s.t } Az - b = \delta \\ &\quad \text{sMat}(z) \succeq 0 \end{aligned} \tag{3.33}$$

To motivate the compressed sensing approach, suppose that only the outlier measurements are noisy and that the remaining measurements are accurate. If $\bar{z}$ denotes the true solution, then

$A\bar{z} - b$ is sparse and we consider the popular l_1 norm minimization problem,

$$\begin{aligned} &\min \|\delta\|_1 \\ &\text{s.t } Az - b = \delta \\ &\quad \text{sMat}(z) \succeq 0. \end{aligned} \tag{3.34}$$

Aside from the positive semidefinite cone (3.34) is a compressed sensing problem. To see this, note than $\delta + b = Az$ if, and only if, $\delta + b \in \text{range}(A)$. Let N be a matrix such that $\text{range}(A) = \text{null}(N)$. Then $\delta + b = Az$ if, and only if, $\delta + b \in \text{null}(N)$. Therefore the constraint, $Az - b = \delta$ is equivalent to $N\delta = -Nb$ which is exactly the compressed sensing constraint.

The problem (3.34) differs from the classical compressed sensing model in the positive semidefinite constraint. However, in our numerical tests, we have found that adding the positive semidefinite constraint greatly increases the success rate in identifying outliers. In compressed sensing, If the matrix N satisfy the so-called *restricted isometry property*, then the sparse signal can be recovered exactly [6, Theorem 1.1]. However, there are no practical algorithms available right now to check if a given matrix satisfies the restricted isometry property. If δ_0 is the solution to (3.34) and most of the elements of δ_0 are 0, then the non-zero elements indicate the outlier measurements.

Thus far, we have assumed that most of the measurements in b are exact and a few have large error. Now let us revert to the original assumption of this section: that most elements of b are slightly inaccurate and few elements are very inaccurate. If the positive semidefinite constraint is ignored, then the identification of outliers is guaranteed to be accurate assuming that N satisfies the restricted isometry property. To be specific, if $\delta^{\#}$ represents the optimal solution of (3.34) without the positive semidefinite constraint, then $||\delta^{\#} - \delta_0||_{l_2} \leq C_S \cdot \epsilon$ where C_S and ϵ are small constants, [5, 6]. The specifics for our outlier-detection algorithm are stated in Algorithm 3.2.

Algorithm 3.2 Removing Outliers

1: **INPUT:** Matrix of sensor locations, P_T, and vector of noisy distances, d, from sensors to the source.

2: Solve the following l_1 norm minimization problem

$$\begin{aligned} &\min \| \mathcal{K}_V(R) - D_{T_c}\|_1, \\ &\text{s.t. } R \succeq 0. \end{aligned} \tag{3.35}$$

3: Obtain $\delta := (\mathcal{K}_V(R) - D_{T_c})_{1:n,n+1}$.

4: Normalize: $\delta \leftarrow \frac{1}{\|\delta\|_2}\delta$.

5: Remove p_i from P_T and d_i from d for all i satisfying $\delta_i \geq \frac{1}{\sqrt{n}}$.

6: **OUTPUT:** Sensor matrix P_T and distance vector d with outliers removed.

3.2 Recovering Source Position from Gram Matrix

After finding the **EDM** from our data, we need to rotate the sensors back to their original positions in order to recover the position of the source. This is done by solving a Procrustes problem. That is, suppose that the, appropriately partitioned, final **EDM**, corresponding Gram matrix and points

are,

$$D_f = \begin{bmatrix} \bar{D}_f & d_f \\ d_f^T & 0 \end{bmatrix}, \; G_f = P_f P_f^T \in \mathcal{S}^{n+1}, \quad P_f = \begin{bmatrix} \bar{P}_f \\ p_f^T \end{bmatrix} \in \mathbb{R}^{N+1,r}.$$

Assuming $\bar{P}_f$ and the original data P_T are both centered, we now have two approaches.

The first approach solves the following Procrustes problem using [20, Algorithm 12.4.1]

$$\begin{array}{ll} \min_Q & \|P_T - \bar{P}_f Q\|_F^2 \\ \text{s.t.} & Q^T Q = I_r. \end{array} \tag{3.36}$$

The optimal solution can be found explicitly from the singular value decomposition of $\bar{P}_f^T P_T$. If $\bar{P}_f^T P_T =: U_f \Sigma_f V_f^T$, then the optimal solution to (3.36) is $Q^* := U_f V_f^T$. The recovered position of the source is then $\boxed{p_c^T = p_f^T Q^*}$.

The second approach is to solve the least square problem

$$\begin{array}{ll} \min_Q & \|P_T - \bar{P}_f Q\|_F^2 \\ \text{s.t.} & Q \in \mathbb{R}^{r \times r}. \end{array} \tag{3.37}$$

The least square solution is $\bar{Q} = \bar{P}_f^\dagger P_T$. Recall that $\bar{P}_f^\dagger$ is the Moore-Penrose generalized inverse of $\bar{P}_f$. The recovered position of the source is then $\boxed{p_c^T = p_f^T \bar{Q}}$.

4 Numerical Results

To compare the different methods, we used randomly generated data with an error proportional to the distance to each tower. The proportionality is given by η. This gives

$$D_{n+1,i} = D_{i,n+1} = \left[\bar{d}_i \left(1 + \varepsilon_i\right)\right]^2, \tag{4.1}$$

where D is the generated **EDM** and $\varepsilon \in U(-\eta, \eta)$. The outliers are obtained by multiplying (4.1) by another factor θ for a small subset of the indices.

We let $\mathcal{M}$ denote the set of optimization methods to be tested. Then for $M \in \mathcal{M}$, the relative error, c_{re}^M, between the true location of the source, c, and the location obtained using method M, denoted c_M, is given by

$$c_{re}^M = \frac{\|c_M - c\|}{\|c\|}. \tag{4.2}$$

The data is then found by calculating this error for all the methods and varying the error η in equation (4.1) and the amount of sensors n. For each pair (n, η), one hundred instances are solved.

The methods in the tables are labelled according to the models with some additional prefixes. To be specific, the L and P prefixes represent the different ways used to obtain the position of the source, c. By L we denote the least square approach of (3.37) and P represents the Procrustes approach in (3.36). We choose $\alpha = 1$ in **FNEDM** and the constant γ for **PNEDM** in (3.30) is chosen to be 1000.

We report some results in the following table.

Error factor η	$\eta = 0.002$			$\eta = 0.02$			$\eta = 0.2$		
# Sensors	5	10	15	5	10	15	5	10	15
L-NEDM	0.0045	0.0014	0.0010	0.0408	0.0140	0.0120	0.3550	0.1466	0.1153
P-NEDM	0.0025	0.0013	0.0010	0.0231	0.0133	0.0117	0.2813	0.1385	0.1171
SDR	0.0024	0.0014	0.0010	0.0223	0.0137	0.0119	0.2739	0.1373	0.1164
L-FNEDM	0.0042	0.0013	0.0010	0.0356	0.0141	0.0119	0.2910	0.1395	0.1061
P-FNEDM	0.0024	0.0013	0.0010	0.0237	0.0134	0.0118	0.2623	0.1360	0.1088

Table 4.1: The mean relative error c_{re}^M of 100 simulations for varying amount of sensors and error factors with no outliers for dimension $r = 3$.

Error factor η	$\eta = 0.005$			$\eta = 0.05$			$\eta = 0.15$		
# Sensors	5	10	15	5	10	15	5	10	15
L-NEDM	0.0101	0.0033	0.0027	0.0970	0.0328	0.0262	0.2473	0.1037	0.0786
P-NEDM	0.0070	0.0031	0.0027	0.0610	0.0320	0.0262	0.1925	0.1041	0.0760
SDR	0.0071	0.0031	0.0027	0.0576	0.0322	0.0261	0.1933	0.1030	0.0779
L-FNEDM	0.0090	0.0032	0.0026	0.0800	0.0311	0.0255	0.2151	0.1001	0.0769
P-FNEDM	0.0069	0.0031	0.0027	0.0536	0.0310	0.0258	0.1914	0.1000	0.0772

Table 4.2: The mean relative error c_{re}^M of 100 simulations for varying amount of sensors and error factors with no outliers for dimension $r = 3$.

From Table 4.1 and 4.2 we can see generally P-FNEDM has the smallest error, and occasionally L-FNEDM is better. Also we can see that as the number of towers n increases, the relative error c_{re}^M decreases which is expected as we have more sensors, the location of the source should be more accurate.

To compare the overall performance of all the methods, we use the well known *performance profiles*, [13]. The approach is outlined below.

For each pair (n, η) and one hundred solved instances, we calculate the mean of the relative error c_{re}^M for method M. We denote this

$$c_{n,\eta,M} = \text{ mean over 100 instances, for } n \text{ towers, with error factor } \eta \text{ and method } M.$$

We then compute the *performance ratio*,

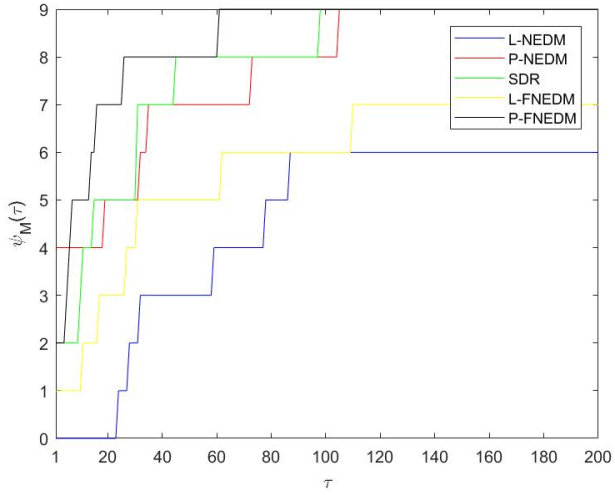
$$r_{n,\eta,M} = \frac{c_{n,\eta,M}}{\min\{c_{n,\eta,M} : M \in \mathcal{M}\}},$$

and the function,

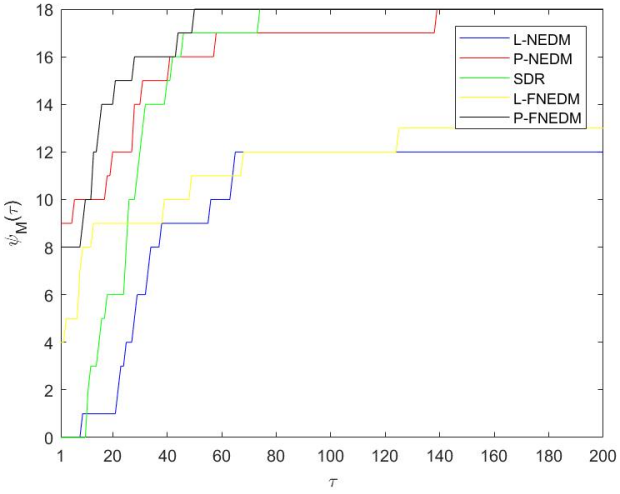
$$\psi_M(\tau) = \frac{|\{(n,\eta) : r_{n,\eta,M} \leq \tau\}|}{|\mathcal{M}|}.$$

The performance profile is a plot of $\psi_M(\tau)$ for $\tau \in (1, +\infty)$ and all choices of $M \in \mathcal{M}$. Note that $r_{n,\eta,M} \geq 1$ and equality holds if, and only if, the solution obtained by M is best for the pair (n, η). In general, smaller values of $r_{n,\eta,M}$ indicate better performance. The function $\psi_M(\tau)$ measures how many pairs (n, η) were solved with a performance ratio of τ or better. The function is monotonically non-decreasing and larger values are better.

The performance profiles can be seen in Figure 4.1a and 4.1b, the P-FNEDM approach has the best performance over all 5 methods. Also using the Procrustes approach (3.36) is better than using the least squares approach (3.37). Allowing the sensors to move in **FNEDM** model is better



(a) $\eta = [0.002, 0.02, 0.2]$



(b) $\eta = [0.0005, 0.001, 0.005, 0.01, 0.05, 0.15]$

Figure 4.1: Performance Profiles for $\psi_M(\tau)$ with $n = [5, 10, 15]$, $r = 3$, no outliers.

than fixing the sensors in **SDR** or making the sensors completely free in **NEDM** for recovering the location of the source.

We also generate the data with outliers. In **FNEDM**, the outliers are detected and removed using the ℓ_1 norm approach described in Section 3.1.5. We report the results with outliers added in the following Table 4.3 and Table 4.4.

Error factor η	$\eta = 0.001$			$\eta = 0.01$			$\eta = 0.1$		
# Sensors	7	12	16	7	12	16	7	12	16
L-RNEDM	0.8076	0.6189	0.4579	0.8695	0.6376	0.4738	0.8006	0.5935	0.4068
P-RNEDM	1.0319	0.6789	0.4755	1.0819	0.6869	0.4677	0.9939	0.6374	0.4312
SDR	1.0618	0.7150	0.5398	1.0825	0.6981	0.5343	0.9968	0.6732	0.4983
L-FNEDM	0.1358	0.0546	0.0388	0.1556	0.0525	0.0402	0.2308	0.0799	0.0710
P-FNEDM	0.1364	0.0546	0.0388	0.1588	0.0527	0.0401	0.2150	0.0799	0.0708

Table 4.3: The mean relative error c_{re}^M of 100 simulations for varying amount of sensors and error factors with 1 outlier for dimension $r = 3$. Outlier factor $\theta \sim U(5, 10)$

Error factor η	$\eta = 0.001$			$\eta = 0.01$			$\eta = 0.1$		
# Sensors	7	12	16	7	12	16	7	12	16
L-RNEDM	0.7035	0.5299	0.3909	0.7686	0.5186	0.3905	0.7219	0.5296	0.4271
P-RNEDM	0.9533	0.5838	0.4488	0.9160	0.5817	0.4371	0.9324	0.6183	0.4739
SDR	0.9337	0.5386	0.4623	0.8905	0.5600	0.4390	0.8927	0.5917	0.4663
L-FNEDM	0.5777	0.1032	0.0571	0.5637	0.0961	0.0560	0.5860	0.1409	0.0878
P-FNEDM	0.5740	0.1033	0.0561	0.5388	0.0925	0.0544	0.5619	0.1380	0.0864

Table 4.4: The mean relative error c_{re}^M of 100 simulations for varying amount of sensors and error factors with 2 outliers for dimension $r = 3$. Outlier factor $\theta \sim U(3, 6)$

From Table 4.3 and 4.4 we can see clearly that when outliers are added, the **FNEDM** outperforms both **SDR** and **NEDM** with a big improvement, as the outliers can be removed. It is also consistent with our previous conclusion that using the Procrustes approach (3.36) is better than using the least squares approach (3.37).

5 Conclusion

We showed that the **SLS** formulation of the single source localization problem is inherently convex, by considering the semidefinite relaxation, **SDR**, of the **GTRS** formulation. The extreme points of the optimal set of **SDR** correspond exactly to the optimal solutions of the **SLS** formulation and these extreme points can be obtained by solving no more than $r + 1$ convex optimization problems.

We also analyzed several **EDM** based relaxations of the **SLS** formulation and introduced the weighted facial reduction model **FNEDM**. The optimal value of **FNEDM** was shown to converge to the optimal value of **SLS** by increasing α. In our numerical tests, we showed that our newly proposed model **FNEDM** performs the best for recovering the location of the source. Without any outliers present, the performance of each method improves as the number of towers increases. This is expected since more information is available. All the methods tend to perform similarly as the number of towers increases but the facial reduction model, **FNEDM**, using the Procrustes approach performs the best.

Finally, we used the ℓ_1 norm approach in Algorithm 3.2, to remove outlier measurements. In Table 4.3 and Table 4.4 we demonstrate the effectiveness of this approach.

Index

References

[1] A.Y. Alfakih, A. Khandani, and H. Wolkowicz. Solving Euclidean distance matrix completion problems via semidefinite programming. *Comput. Optim. Appl.*, 12(1-3):13–30, 1999. A tribute to Olvi Mangasarian. 2, 4

[2] A. Beck, P. Stoica, and J. Li. Exact and approximate solutions of source localization problems. *IEEE Transactions and signal processing*, 56(5):1770–1778, 2008. 2, 4, 5, 7

[3] A. Beck, M. Teboulle, and Z. Chikishev. Iterative minimization schemes for solving the single source localization problem. *SIAM J. Optim.*, 19(3):1397–1416, 2008. 2

[4] I. Borg and P. Groenen. Modern multidimensional scaling: theory and applications. *Journal of Educational Measurement*, 40(3):277–280, 2003. 2

[5] E. Candes, M. Rudelson, T. Tao, and R. Vershynin. Error correction via linear programming. In *Proceedings of hte 2005 46th Annual EIII Symposium on Foundations of Computer Science, (FOCS'o5)*, pages 1–14. IEEE, 2005. 25

[6] E.J. Candès, J.K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math.*, 59(8):1207–1223, 2006. 25

[7] K.W. Cheung, H.-C. So, W-K Ma, and Y.-T. Chan. Least squares algorithms for time-of-arrival-based mobile location. *IEEE Transactions on Signal Processing*, 52(4):1121–1130, 2004. 2, 5

[8] T.F Cox and M.A. Cox. *Multidimensional scaling*. Chapman and hall/CRC, 2000. 2

[9] G.M. Crippen and T.F. Havel. *Distance geometry and molecular conformation*, volume 74. Research Studies Press Taunton, 1988. 2

[10] F. Critchley. Dimensionality theorems in multidimensional scaling and hierarchical cluster analysis. In *Data analysis and informatics (Versailles, 1985)*, pages 45–70. North-Holland, Amsterdam, 1986. 4

[11] J. Dattorro. *Convex optimization & Euclidean distance geometry*. Lulu. com, 2010. 2, 4

[12] Y. Ding, N. Krislock, J. Qian, and H. Wolkowicz. Sensor network localization, Euclidean distance matrix completions, and graph realization. *Optim. Eng.*, 11(1):45–66, 2010. 2, 12

[13] E.D. Dolan and J.J. Moré. Benchmarking optimization software with performance profiles. *Math. Program.*, 91(2, Ser. A):201–213, 2002. 27

[14] D. Drusvyatskiy, N. Krislock, Y-L. Cheung Voronin, and H. Wolkowicz. Noisy Euclidean distance realization: robust facial reduction and the Pareto frontier. *SIAM Journal on Optimization*, 27(4):2301–2331, 2017. 2

[15] D. Drusvyatskiy, G. Pataki, and H. Wolkowicz. Coordinate shadows of semidefinite and Euclidean distance matrices. *SIAM J. Optim.*, 25(2):1160–1178, 2015. 16

[16] R.J. Duffin. Clark's theorem on linear programs holds for convex programs. *Proc. Nat. Acad. Sci. U.S.A.*, 75(4):1624–1626, 1978. 6

[17] E. Ernst and M. Volle. Zero duality gap for convex programs: a generalization of the Clark-Duffin theorem. *J. Optim. Theory Appl.*, 158(3):668–686, 2013. 6

[18] H. Fang and D.P. O'Leary. Euclidean distance matrix completion problems. *Optimization Methods and Software*, 27(4-5):695–717, 2012. 2

[19] Y. Gao and D. Sun. A majorized penalty approach for calibrating rank constrained correlation matrix problems. *Technical Report*, 2010. 21

[20] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, Maryland, 3^{nd} edition, 1996. 26

[21] J.C. Gower. Properties of Euclidean and non-Euclidean distance matrices. *Linear Algebra Appl.*, 67:81–97, 1985. 4

[22] T.L. Hayden, J. Wells, W.M. Liu, and P. Tarazaga. The cone of distance matrices. *Linear Algebra Appl.*, 144:153–169, 1991. 4

[23] J.-B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of convex analysis*. Grundlehren Text Editions. Springer-Verlag, Berlin, 2001. Abridged version of ıt Convex analysis and minimization algorithms. I [Springer, Berlin, 1993; MR1261420 (95m:90001)] and ıt II [ibid.; MR1295240 (95m:90002)]. 14

[24] V. Jeyakumar and H. Wolkowicz. Zero duality gaps in infinite-dimensional programming. *J. Optim. Theory Appl.*, 67(1):87–108, 1990. 6

[25] H. Koshima and J. Hoshen. Personal locator services emerge. *IEEE spectrum*, 37(2):41–48, 2000. 2

[26] N. Krislock and H. Wolkowicz. Euclidean distance matrices and applications. In *Handbook on Semidefinite, Cone and Polynomial Optimization*, number 2009-06 in International Series in Operations Research & Management Science, pages 879–914. Springer-Verlag, 2011. 2

[27] T. Kundu. Acoustic source localization. *Ultrasonics*, 54(1):25–38, 2014. 2

[28] G. Pataki. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Math. Oper. Res.*, 23(2):339–358, 1998. 9

[29] T.K. Pong and H. Wolkowicz. The generalized trust region subproblem. *Comput. Optim. Appl.*, 58(2):273–322, 2014. 5, 7

[30] H.-D. Qi. A semismooth Newton method for the nearest Euclidean distance matrix problem. *SIAM Journal on Matrix Analysis and Applications*, 34(1):67–93, 2013. 2, 21, 22, 23

[31] H.-D. Qi and X. Yuan. Computing the nearest Euclidean distance matrix with low embedding dimensions. *Mathematical Programming*, 147(1):351–389, Oct 2014. 2, 21, 22, 23

[32] I.J. Schoenberg. Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.*, 44(3):522–536, 1938. 3

[33] R. Stern and H. Wolkowicz. Trust region problems and nonsymmetric eigenvalue perturbations. *SIAM J. Matrix Anal. Appl.*, 15(3):755–778, 1994. 5

[34] R. Stern and H. Wolkowicz. Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations. *SIAM J. Optim.*, 5(2):286–313, 1995. 7

[35] P.D. Tao and L.T.H. An. Convex analysis approach to dc programming: Theory, algorithms and applications. *Acta mathematica vietnamica*, 22(1):289–355, 1997. 23, 24

[36] P.D. Tao and L.T.H. An. The dc (difference of convex functions) programming and dca revisited with dc models of real world nonconvex optimization problems. *Annals of operations research*, 133(1-4):23–46, 2005. 23, 24

[37] L. Tunçel. *Polyhedral and Semidefinite Programming Methods in Combinatorial Optimization*, volume 27 of *Fields Institute Monographs*. American Mathematical Society, Providence, RI, 2010. 15

[38] J. Warrior, E. McHenry, and K. McGee. They know where you are [location detection]. *IEEE Spectrum*, 40(7):20–25, 2003. 2