

Noisy Euclidean Distance Matrix Completion with a Single Missing Node

Stefan Sremac^{*} Fei Wang[†] Henry Wolkowicz[‡] Lucas Pettersson[§]

Tuesday 23rd July, 2019, 20:50

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

1	Introduction	2
1.1	Outline	3
1.2	Preliminaries	3
2	SDP Formulation	5
2.1	GTRS	5
2.2	The Semidefinite Relaxation, SDR	8
2.2.1	A Purification Algorithm	10

^{*}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.

26	3 EDM Formulation	12
27	3.1 The Relaxed NEDM Problem	14
28	3.1.1 Nearest Euclidean Distance Matrix Formulation	14
29	3.1.2 Weighted, Facially Reduced NEDM	16
30	3.1.3 Analysis of FNEDM	18
31	3.1.4 Solving FNEDM	21
32	3.1.5 Identifying Outliers using l_1 Minimization and Facial Reduction	25
33	3.2 Recovering Source Position from Gram Matrix	26
34	4 Numerical Results	27
35	5 Conclusion	30
36	Index	33
37	Bibliography	35

38 List of Tables

39	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 $r = 3$	28
40	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 $r = 3$	28
41	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 $\theta \sim U(5, 10)$	30
42	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 $\theta \sim U(3, 6)$	30

47 List of Algorithms

48	2.1 Purification Algorithm	11
49	3.1 Majorization Algorithm	25
50	3.2 Removing Outliers	27

51 1 Introduction

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

57 The single source localization problem has applications in e.g., navigation, structural engineer-
58 ing, and emergency response, [3, 4, 8, 24, 26, 38]. In general, it is related to distance geometry
59 problems where the input consists of Euclidean distance measurements and a set of points in Eu-
60 clidean space. The *sensor network localization problem* is a generalization of our single source

61 problem, where there are multiple sources and only some of the distance estimates are known. The
 62 general *Euclidean distance matrix completion problem* is yet a further generalization, where sensors
 63 do not have specified locations and only partial, possibly noisy, distance information is available,
 64 e.g., [2, 13, 15]. We refer the readers to the books [1, 5, 9, 10] and survey article [27] for background
 65 and applications, and to the paper [18] for algorithmic comparisons. We also refer the readers for
 66 the related *nearest Euclidean distance matrix (NEDM)* problem to the papers [30, 31] where a
 67 semismooth Newton approach and a rank majorization approach is presented. The more general
 68 weighted **NEDM** is a much harder problem though. For theory that relates **NEDM** to semidefinite
 69 programming, see e.g., [12, 25].

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

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

81 1.1 Outline

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

90 1.2 Preliminaries

We now present some preliminaries and background on **SDP** and the facial geometry, see e.g., [17].
 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*,

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

91 Unless otherwise specified, the norm of a matrix is the Frobenius norm, and we may drop the
 92 subscript F . For a convex set C , the convex subset $f \subseteq C$ is a *face of C* if for all $x, y \in C$, $x, y \in f$
 93 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* , $\text{face}(S)$, the smallest face of \mathcal{S}_+^n that contains S . Let $X \in \mathcal{S}_+^n$ have rank r with orthogonal spectral decomposition.

$$X = [P \ Q] \begin{bmatrix} D_+ & 0 \\ 0 & 0 \end{bmatrix}^T [P \ Q]^T, \quad D_+ \in \mathcal{S}_{++}^r.$$

Then the range and nullspace characterizations of $\text{face}(X)$ are,

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

94 We say that the matrix QQ^T is an *exposing vector* for $\text{face}(X)$.

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

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

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

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

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

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

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

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

100 respectively. Here Diag is the adjoint of diag , the matrix $J_n := I - \frac{1}{n}ee^T$ is the orthogonal projection
101 onto \mathcal{S}_C , and $\text{offDiag}(D)$ refers to zeroing out the diagonal of D , i.e., the orthogonal projection onto
102 \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, $\mathcal{K}(\mathcal{S}_+^n) = \mathcal{E}^n$
103 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 [2] for instance. Specifically,

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

104 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
105 the rows of P correspond to the points in \mathbb{R}^r with pairwise distances corresponding to the elements
106 of D . For more details, see e.g., [2, 11, 12, 21, 22].

2 SDP Formulation

We begin this section by formulating the **SLS** problem using the model and notation of [3]. 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. $\text{int conv}(p^1, \dots, p^n) \neq \emptyset$;

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 i th sensor,

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

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

$$p_{\text{LS}}^* := \min_{x \in \mathbb{R}^r} \sum_{i=1}^n (\|x - p^i\| - d_i)^2. \quad (2.1)$$

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., [8]. However, it is a non-convex problem with an objective function that is not differentiable. Motivated by the success in [3], the main problem we consider instead is the optimization problem with squared distances

$$\boxed{(\text{SLS}) \quad p_{\text{SLS}}^* := \min_{x \in \mathbb{R}^r} \sum_{i=1}^n (\|x - p^i\|^2 - d_i^2)^2.} \quad (2.2)$$

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 [3], we get the equivalent problem

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

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

Theorem 2.2. *Let*

$$P_T := [p^1 \ p^2 \ \dots \ p^n]^T, \quad A := [-2P_T \ e], \quad \tilde{I} := \begin{bmatrix} I_r & 0_{r \times 1} \\ 0_{1 \times r} & 0 \end{bmatrix},$$

$$b := \begin{pmatrix} d_1^2 - \|p^1\|^2 \\ \vdots \\ d_n^2 - \|p^n\|^2 \end{pmatrix}, \quad \tilde{b} := \begin{pmatrix} 0 \\ -\frac{1}{2} \end{pmatrix}. \quad (2.4)$$

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

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

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

129 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:

$$p_{SLS}^* = d_{SLS}^* := \max_{\lambda} \min_y \{ \|Ay - b\|^2 + \lambda(y^T \tilde{I}y + 2\tilde{b}^T y) \}. \quad (2.6)$$

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, $\text{rank}(P_T) = r$. Therefore, $(P_T)^T e = 0$ implies that

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

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

That we have a zero duality gap for **GTRS** follows from [29], since this is a *generalized trust region subproblem*. 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.

$$\begin{aligned} p_{\text{SLS}}^* &= \min_y \{ \|Ay - b\|^2 : y^T \tilde{I} y + 2\tilde{b}^T y = 0 \} \\ &= \min_y \{ \|Ay - b\|^2 - \gamma(y^T \tilde{I} y + 2\tilde{b}^T y) : y^T \tilde{I} y + 2\tilde{b}^T y = 0 \} \end{aligned} \quad (2.7)$$

$$= \min_y \{ \|Ay - b\|^2 - \gamma(y^T \tilde{I} y + 2\tilde{b}^T y) : y^T \tilde{I} y + 2\tilde{b}^T y \leq 0 \} \quad (2.8)$$

$$= \max_{\lambda \geq 0} \min_y \{ \|Ay - b\|^2 - \gamma(y^T \tilde{I} y + 2\tilde{b}^T y) + \lambda(y^T \tilde{I} y + 2\tilde{b}^T y) \} \quad (2.9)$$

$$= \max_{(\lambda - \gamma)} \min_y \{ \|Ay - b\|^2 + (\lambda - \gamma)(y^T \tilde{I} y + 2\tilde{b}^T y) \} \quad (2.10)$$

$$\begin{aligned} &= d_{\text{SLS}}^* \\ &\leq p_{\text{SLS}}^*. \end{aligned}$$

133 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
134 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_{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_{SLS}^* . Then there exists \bar{y} satisfying,

$$g(\bar{y}) < 0, \quad f(\bar{y}) < p_{\text{SLS}}^*.$$

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 $h = (\bar{h}^T \ 0)^T$ with $\bar{h} \neq 0$. Moreover, we have,

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

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

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

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

$$g(\bar{y} + \bar{\alpha} h) = 0, \quad f(\bar{y} + \bar{\alpha} h) \leq f(\bar{y}) < p_{\text{SLS}}^*,$$

135 a contradiction.

136 We have confirmed the third equality. Now (2.8) is a convex quadratic optimization problem
137 where the Slater constraint qualification holds. This implies that strong duality holds, i.e., we get
138 (2.9) with attainment for some $\lambda \geq 0$. Now if $\lambda < 0$ in (2.9) then the Hessian of the objective is
139 indefinite (by construction of γ) and the optimal value of the inner minimization problem is $-\infty$.
140 Thus since (2.9) is maximized with respect to λ in the outer optimization problem, we may remove
141 the non-negativity constraint and obtain (2.10). The remaining lines are due to the definition of
142 the Lagrangian dual and weak duality. Strong duality follows immediately. \square

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

We compare our **SDP** approach with the approach used by Beck et al [3]. In their approach they have to solve the following system obtained from the optimality conditions of **GTRS**:

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

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

151 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

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

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

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

152 Here the first equality follows from the definition of the dual. The second and third equalities
 153 are just equivalent reformulations of the first one. For the last equality, let $\tilde{y} = [y^T, y_0]^T$ and
 154 $M = \bar{A} - \lambda\bar{B} - se_{r+2}e_{r+2}^T$ and suppose $\tilde{y}^TM\tilde{y} < 0$ for some \tilde{y} . If y_0 is nonzero, we can get a
 155 contradiction by scaling \tilde{y} . If y_0 is zero, by the continuity of $\tilde{y} \rightarrow \tilde{y}^TM\tilde{y}$, we can perturb \tilde{y} by a
 156 small enough amount so that the last element of \tilde{y} is nonzero. This is a contradiction as in the
 157 previous case.

We observe that (2.14) is a dual-form **SDP** corresponding to the primal SDP problem, e.g., [39],

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

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

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

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

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

162 **Theorem 2.4.** *The following holds:*

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

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

$$p_{\text{SLS}}^* = d_{\text{SLS}}^* \leq p_{\text{SDR}}^*. \quad (2.17)$$

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

$$p_{\text{SDR}}^* \leq p_{\text{SLS}}^* \implies p_{\text{SLS}}^* = d_{\text{SLS}}^* = p_{\text{SDR}}^*.$$

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,

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

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

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

166 By Theorem 2.2, $A^T A$ is positive definite and a Slater point may be obtained by choosing s so that
167 $b^T b - s$ 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

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

168 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
169 by the definition of face, we conclude that Ω is a face of \mathcal{F} .

170 Since Ω is a compact convex set it has an extreme point, say X^* . Now X^* is also an extreme
171 point of \mathcal{F} , as the relation *face of* is transitive, i.e., *a face of a face is a face*. Moreover, since

172 there are exactly two equality constraints in **SDR**, by Theorem 2.1 of [28], we have $\text{rank}(X^*)(1 +$
173 $\text{rank}(X^*))/2 \leq 2$. This equation is satisfied if, and only if, $\text{rank}(X^*) = 1$. Equivalently, $X^* = \rho(y^*)$
174 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
175 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

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

176 Since X^* has rank 1 and $Y, Z \succeq 0$, it follows that Y and Z are non-negative multiples of X^* . But
177 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
178 extreme point of Ω , as desired. \square

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

186 **Corollary 2.5.** *The following hold.*

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

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

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

197 2.2.1 A Purification Algorithm

198 Suppose the optimal solution of (2.15) is \bar{X} with optimal value $p_{\text{SDR}}^* = \langle \bar{A}, \bar{X} \rangle$ and $\text{rank}(\bar{X}) = \bar{r}$
199 where $\bar{r} > 1$. Note that we can not obtain an optimal solution of **GTRS** from \bar{X} since the rank
200 is too large. However, in this section we construct an algorithm that returns an extreme point of
201 Ω which, by Theorem 2.4, is easily transformed into an optimal solution of **GTRS**. We note that
202 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

$$\bar{B} \leftarrow U^T \bar{B} U, \bar{A} \leftarrow U^T \bar{A} U, \bar{E} \leftarrow U^T \bar{E} U, \quad (2.18)$$

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

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

respectively. The rank reducing program is

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

203 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

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

204

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

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

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 $T \in \text{Null}(\mathcal{A}_S^k)$ 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 $T \in \text{Null}(\mathcal{A}_S^k)$ 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,

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

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

206 **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,
 207 then the algorithm terminates with at most $\text{rank}(\bar{X}) - 1 \leq r + 1$ calls to the while loop and the
 208 output, X^* , is a rank 1 optimal solution of **SDR**.

Proof. We proceed by considering the trivial case, $\text{rank}(\bar{X}) = 1$. 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,

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

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

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

209 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, \dots , generated by Algorithm 2.1 is strictly decreasing. It immediately follows that the algorithm terminates in at most $\text{rank}(\bar{X}) - 1$ 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 $\text{rank}(S^k) = r_k = r_{k-1}$ and S^k is a Slater point of the optimization problem,

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

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

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

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

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

210 and it follows that V^k and W^k are feasible for (2.22). By uniqueness of S^k we have that $Y^k =$
 211 $Z^k = X^k$ and X^k is an extreme point of Ω . Then by Theorem 2.4, $\text{rank}(S^k) = 1$ and Algorithm 2.1
 212 terminates before generating r_k , a contradiction. \square

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

215 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

$$P_T = [p^1 \quad p^2 \quad \dots \quad p^n]^T \in \mathbb{R}^{n \times r}.$$

216 From Assumption 2.1 the towers are centered, i.e. $e^T P_T = 0$. This property is desirable due to
 217 the Schoenberg characterization which states that \mathcal{K} is an isomorphism between $\mathcal{S}_+^n \cap \mathcal{S}_C$ and \mathcal{E}^n .
 218 Moreover, it allows for easy recovery of the towers in the last step of our algorithm by solving a
 219 *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

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

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.,

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

The approach of [13] 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,

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

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

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

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

226 3.1 The Relaxed NEDM Problem

227 3.1.1 Nearest Euclidean Distance Matrix Formulation

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

$$\begin{aligned} & \min \frac{1}{2} \| \mathcal{K}(X) - D_{T_c} \|^2 \\ (\mathbf{NEDM}) \quad & \text{s.t. } \text{rank}(X) \leq r \\ & X \succeq 0. \end{aligned} \quad (3.4)$$

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

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

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

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,

$$\begin{aligned}
\mathcal{K}^\dagger(D_{T_c}) &= -\frac{1}{2} J_{n+1} D_{T_c} J_{n+1} \\
&= -\frac{1}{2} (J_{n+1} \bar{D} J_{n+1} + J_{n+1} Q J_{n+1}) + (-\frac{1}{2} J_{n+1} P J_{n+1}),
\end{aligned}$$

where the first term in the last line is positive semidefinite with at least r and at most $r + 1$ positive eigenvalues ($-J_{n+1} \bar{D} J_{n+1}$ is a positive semidefinite matrix with rank r and $-J_{n+1} Q J_{n+1}$ is positive semidefinite with rank at most 1); and the second term is negative semidefinite with at most one negative eigenvalue. Using the Cauchy-Schwartz inequality it can be shown that for $X, Y \in \mathcal{S}^n$,

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

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

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

242 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],
 243 it has at most one negative eigenvalue whose magnitude is bounded above by $\frac{\sqrt{2}}{2}\|J_{n+1}\|^2\|\varepsilon\|$, as
 244 desired. \square

245 The following corollary follows immediately.

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

248 3.1.2 Weighted, Facially Reduced NEDM

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

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

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

255 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
 256 is the following:

$$\begin{aligned} \min \quad & \frac{1}{2} \|\mathcal{K}(X) - D_{T_c}\|^2 \\ \text{(NEDMP)} \quad & \text{s.t. } \text{rank}(X) \leq r \\ & X \in \text{face}(F_T, \mathcal{S}_+^{n+1}). \end{aligned} \quad (3.9)$$

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,

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

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

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

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

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

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

260 2. W exposes $\text{face}(F_T, \mathcal{S}_+^{n+1})$.

261 *Proof.* This statement is a special case of Theorem 4.13 of [16]. \square

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 $\text{face}(F_T, \mathcal{S}_+^{n+1})$. However, the ‘range space’ characterization is more useful in the context of semidefinite optimization 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 $\text{null}(W)$. One such choice is,

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

262 To verify that the columns of V indeed form a basis for $\text{null}(W)$, we first observe that $\text{rank}(V) = r + 1$
 263 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

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

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

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

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

272 We begin our analysis by proving that V_α is attained.

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

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

277 *Proof.* For Item 1, under the assumption that $\alpha > 0$, we have $H_\alpha \circ \mathcal{K}_V(R) = 0$ if, and only if,
 278 $\mathcal{K}_V(R) = 0$. Recall that \mathcal{K} is one-to-one between the centered and hollow subspaces and $\mathcal{K}(0) = 0$.
 279 By construction, $\text{range}(V \cdot V^T)$ is a subset of the centered matrices. Hence $H \circ \mathcal{K}_V(R) = 0$ if, and
 280 only if, $VRV^T = 0$. Since V is full column rank, $VRV^T = 0$ if, and only if, $R = 0$, as desired.

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

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

286 We conclude this subsection by deriving the optimality conditions for the convex relaxation of
 287 **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,*

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

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

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,

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

289 hence orthogonality holds. \square

290 3.1.3 Analysis of FNEDM

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

293 In the **SLS** model, the distances between the towers are fixed, while in the **NEDM** model (3.4),
 294 the distances between towers are free. The facial reduction model allows the distances between the
 295 towers to change but the towers can still be transformed back to their original positions by a square
 296 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,*

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

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 ,

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

Now by (3.11) we have,

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

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

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

297 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 $P P^T$ 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,

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

298 as desired. □

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

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

305 **Lemma 3.7.** *Consider the problem,*

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

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

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

308 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
309 feasible for (3.2). First we show that X may be assumed to be centered. To see this, consider
310 $\hat{X} = J_n X J_n$. Note that \hat{X} is the orthogonal projection of X onto \mathcal{S}_c and it can be verified that
311 $\hat{X} = \mathcal{K}^\dagger \mathcal{K}(X)$. Now it is clear that $\hat{X} \succeq 0$ and that $\mathcal{K}(\hat{X}) = \mathcal{K}(X)$. Moreover, since J_n is singular
312 we have, $\text{rank}(\hat{X}) \leq \text{rank}(X)$. Therefore, \hat{X} is also feasible for (3.2) and provides the same objective
313 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,

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

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

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

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

316 *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,

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

317 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
318 all R , if D_{T_c} is not an **EDM** with embedding dimension r .

For the second inequality, we first observe that,

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

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

321 **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
322 a sequence of increasing numbers such that $R_{\alpha_\ell} \rightarrow \bar{R}$ for some $\bar{R} \in \mathcal{S}^{r+1}$. Then $V_\alpha \uparrow V_T$ and \bar{R} is
323 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,

$$V_\alpha \uparrow V^* \leq V_T. \quad (3.15)$$

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,

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

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

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

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,

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

324 Now equality holds throughout (3.18) and the desired results are immediate. \square

325 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,

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

326 In this subsection we explore the relationship between the optimal solution of **FNEDM** and the
327 eigenvalues of R_{LS} . In general the Moore-Penrose inverse may be difficult to obtain, however, the
328 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*

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

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

333 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
 334 many applications, the inverse of $(H_\alpha \circ \mathcal{K}_V)^*(H_\alpha \circ \mathcal{K}_V)$, and consequently R_{LS} , may be obtained
 335 efficiently.

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

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

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

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

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

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

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

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

355 In **Case III** we are motivated by the primal-dual approach of [30,31] and the penalty approach
 356 of [19, 30, 31]. Let $h = [1, \dots, \alpha]^T$, we notice that $H_\alpha \circ Y = hh^T \circ Y = \text{Diag}(h)Y \text{Diag}(h)$ if
 357 $\text{diag}(Y) = 0$. Let $T = \text{Diag}(h)$, it is easy to see that (3.12) is equivalent to the problem:

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

358 As in [30, 31], we define $\mathcal{K}_+^{n+1}(r) := \{Y \in \mathcal{S}^{n+1} \mid -J_{n+1}YJ_{n+1} \succeq 0, \text{rank}(J_{n+1}YJ_{n+1}) \leq r\}$, and let
 359 $\mathcal{B}(Y) = 0$ represent the linear constraints $\text{diag}(Y) = 0$ and $\langle J_{n+1}YJ_{n+1}, W \rangle = 0$. Then (3.20) may
 360 be written as,

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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 $\Pi(X)$ is the nearest point to X from this ray. Now we can use orthogonality: $\Pi(X) - X$ is orthogonal to $\Pi(X) - 0$. Then the triangle inequality follows:

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

The Lagrangian dual problem is then defined by,

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

367 as desired. □

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

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

To penalize rank, we consider the concave penalty function,

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

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

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

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

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

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

385 The majorization approach is outlined below in Algorithm 3.1. Central to the approach is the
 386 observation that p is majorized by its linear approximation, since it is concave. In the algorithm,
 387 $\partial p(R)$ denotes the subdifferential of p at R . Thus at every iterate, the convex subproblem (3.31)
 388 is solved to obtain the next iterate.

Algorithm 3.1 Majorization Algorithm

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

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

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

389 **Theorem 3.13.** *Suppose Algorithm 3.1 converges to a stationary point \bar{R} , and that $\operatorname{rank}(\bar{R}) = r$.*
 390 *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:

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

Under the assumption $\operatorname{rank}(\bar{R}) = r$, we have $\bar{R} = V \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} V^T$ where $\Lambda = \operatorname{Diag}(\lambda_1, \dots, \lambda_r)$ with $\lambda_1 \geq \dots \geq \lambda_r > 0$ 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 $\lambda_1, \dots, \lambda_r$. We have

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

and

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

391 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 $\hat{R} \in \operatorname{face}(\bar{R})$, we have

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

392 Hence our claim is proved. □

393 3.1.5 Identifying Outliers using l_1 Minimization and Facial Reduction

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

399 Problem (3.12) is equivalent to minimizing the residual of an overdetermined linear system in
400 the domain of an **SDP** cone. Let $z := \text{svec}(R)$ for $R \in \mathcal{S}^{r+1}$. Abusing our previous notation, let
401 $b := \text{svec}(H_\alpha \circ D_{T_c})$ and let A denote the matrix representation of $H_\alpha \circ \mathcal{K}_V$. Then $z \in \mathbb{R}^{(r+1)(r+2)/2}$
402 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.
403 In other words, we have an overdetermined system. Under this new notation, problem (3.12) is
404 equivalent to,

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

405 To motivate the compressed sensing approach, suppose that only the outlier measurements are
406 noisy and that the remaining measurements are accurate. If \bar{z} denotes the true solution, then
407 $A\bar{z} - b$ is sparse and we consider the popular l_1 norm minimization problem,

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

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

412 The problem (3.34) differs from the classical compressed sensing model in the positive semidefi-
413 nite constraint. However, in our numerical tests, we have found that adding the positive semidefinite
414 constraint greatly increases the success rate in identifying outliers. In compressed sensing, If the
415 matrix N satisfy the so-called *restricted isometry property*, then the sparse signal can be recovered
416 exactly [7, Theorem 1.1]. However, there are no practical algorithms available right now to check
417 if a given matrix satisfies the restricted isometry property. If δ_0 is the solution to (3.34) and most
418 of the elements of δ_0 are 0, then the non-zero elements indicate the outlier measurements.

419 Thus far, we have assumed that most of the measurements in b are exact and a few have large
420 error. Now let us revert to the original assumption of this section: that most elements of b are
421 slightly inaccurate and few elements are very inaccurate. If the positive semidefinite constraint is
422 ignored, then the identification of outliers is guaranteed to be accurate assuming that N satisfies
423 the restricted isometry property. To be specific, if $\delta^\#$ represents the optimal solution of (3.34)
424 without the positive semidefinite constraint, then $\|\delta^\# - \delta_0\|_{l_2} \leq C_S \cdot \epsilon$ where C_S and ϵ are small
425 constants, [6, 7]. The specifics for our outlier-detection algorithm are stated in Algorithm 3.2.

426 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,

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

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

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

- 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.
-

427 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]

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

428 The optimal solution can be found explicitly from the singular value decomposition of $\bar{P}_f^T P_T$. If
 429 $\bar{P}_f^T P_T =: U_f \Sigma_f V_f^T$, then the optimal solution to (3.36) is $Q^* := U_f V_f^T$. The recovered position of

430 the source is then $\boxed{p_c^T = p_f^T Q^*}$.

The second approach is to solve the least square problem

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

431 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

432 \bar{P}_f . The recovered position of the source is then $\boxed{p_c^T = p_f^T \bar{Q}}$.

433 4 Numerical Results

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

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

436 where D is the generated **EDM** and $\varepsilon \in U(-\eta, \eta)$. The outliers are obtained by multiplying (4.1)
 437 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

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

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

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

445 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 $r = 3$.

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

450 To compare the overall performance of all the methods, we use the well known *performance*
 451 *profiles*, [14]. 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

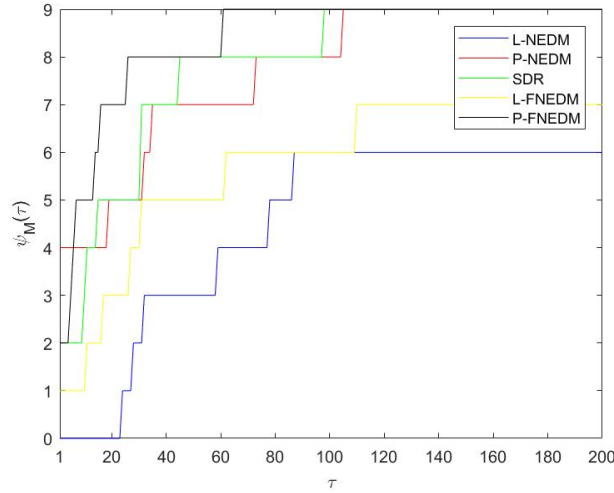
$$c_{n,\eta,M} = \text{mean over 100 instances, for } n \text{ towers, with error factor } \eta \text{ and method } M.$$

We then compute the *performance ratio*,

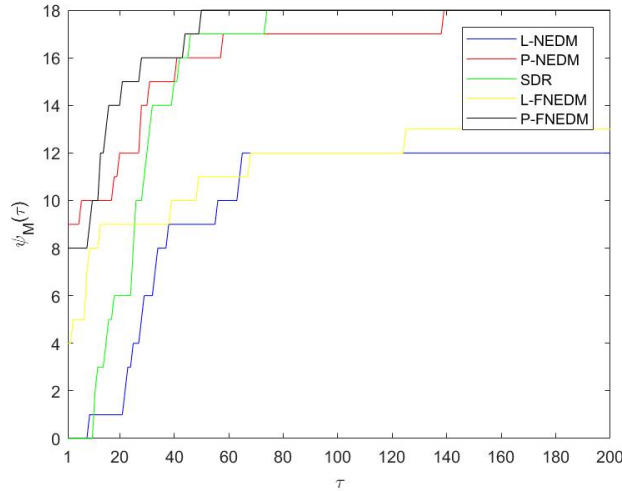
$$r_{n,\eta,M} = \frac{c_{n,\eta,M}}{\min\{c_{n,\eta,M} : M \in \mathcal{M}\}},$$

and the function,

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



(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.

452 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
 453 $r_{n,\eta,M} \geq 1$ and equality holds if, and only if, the solution obtained by M is best for the pair (n, η) .
 454 In general, smaller values of $r_{n,\eta,M}$ indicate better performance. The function $\psi_M(\tau)$ measures how
 455 many pairs (n, η) were solved with a performance ratio of τ or better. The function is monotonically
 456 non-decreasing and larger values are better.

457 The performance profiles can be seen in Figure 4.1a and 4.1b, the P-FNEDM approach has
 458 the best performance over all 5 methods. Also using the Procrustes approach (3.36) is better than
 459 using the least squares approach (3.37). Allowing the sensors to move in **FNEDM** model is better
 460 than fixing the sensors in **SDR** or making the sensors completely free in **NEDM** for recovering the
 461 location of the source.

462 We also generate the data with outliers. In **FNEDM**, the outliers are detected and removed
 463 using the ℓ_1 norm approach described in Section 3.1.5. We report the results with outliers added
 464 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 $\theta \sim U(3, 6)$

465 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
 466 consistent with our previous conclusion that using the Procrustes approach (3.36) is better than
 467 using the least squares approach (3.37).
 468

469 5 Conclusion

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

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

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

Index

- 484 D_T , 12
 485 D_{T_c} , 13
 486 $G_T := P_T P_T^T$, 13
 487 H_T , 14
 488 H_c , 14
 489 $H_\alpha := \alpha H_T + H_c$, 17
 490 $J_n := I - \frac{1}{n} e e^T$, 4
 491 J_n , orthogonal projection onto \mathcal{S}_C , 4
 492 P_T , 13
 493 X_T , 14
 494 Diag , 4
 495 \mathcal{E}^n , **EDM** cone, 4
 496 \mathcal{F} , feasible set of **SDR**, 8
 497 int , interior, 5
 498 \mathcal{K} , Lindenstrauss mapping, 4
 499 \mathcal{K}^* , adjoint, 4
 500 \mathcal{K}^\dagger , Moore-Penrose pseudoinverse, 4
 501 \mathcal{M} , set of methods used, 28
 502 Ω , optimal solutions of **SDR**, 8
 503 \mathcal{S}_C , centered subspace of \mathcal{S}^n , 4
 504 \mathcal{S}_H , hollow subspace of \mathcal{S}^n , 4
 505 conv , convex hull, 5
 506 $\text{diag}(X)$, 4
 507 η , 27
 508 $\text{face}(X)$, 4
 509 $\text{face}(C)$, minimal face of C , 4
 510 $\text{offDiag}(D)$, 4
 511 \bar{D} , 15
 512 ∂ , subdifferential, 24
 513 ρ , map to rank 1 matrices, 8
 514 sMat , svec^{-1} , 4
 515 svec , vectorize a symmetric matrix, 4
 516 svec^{-1} , sMat , 4
 517 **FNEDM**, 17
 518 **PNEDM**, 24
 519 c , cell/source, 5
 520 $c_{n,\eta,M}$, 28
 521 $d = \bar{d} + \varepsilon \in \mathbb{R}^n$, distances (noisy), 5
 522 d_{SLS}^* , 6
 523 e , vector of all ones, 4
 524 $f(R, \alpha)$, 17
 525 $h(R)$, 19
 526 n , 5
 527 n , number of sensors, 5
 528 p_{SLS}^* , 5, 6
 529 r , embedding dimension, 5
 530 $r_{n,\eta,M}$, performance ratio, 28
 531 x_c^T , 14
 532 **EDM** cone, \mathcal{E}^n , 4
 533 **EDM**, Euclidean distance matrix, 4
 534 **FNEDM**, weighted facially reduced problem, 17
 535 **GTRS**, generalized trust region subproblem, 5, 6
 536 **NEDMP**, nearest **EDM** problem with fixed sensors, 13
 537 **NEDMF**, nearest Euclidean distance matrix with fixed sensors, 2
 538 **NEDM**, nearest Euclidean distance matrix, 3
 539 **SDP**, semidefinite program, 8
 540 **SDR**, semidefinite programming relaxation, 8
 541 **SLS**, squared least squares, 2
 542 **NEDMP**, 16
 543 **NEDMinv**, 15
 544 **NEDM**, 14
 545 adjoint, \mathcal{K}^* , 4
 546 cell, c , 5
 547 centered, 5
 548 centered subspace of \mathcal{S}^n , denoted \mathcal{S}_C , 4
 549 cone of positive definite matrices, 3
 550 cone of positive semidefinite matrices, 3
 551 convex hull, conv , 5
 552 distances (noisy), $d = \bar{d} + \varepsilon \in \mathbb{R}^n$, 5
 553 embedding dimension, r , 5
 554 Euclidean distance matrix, **EDM**, 4
 555 exposing vector, 4
 556 face of C , 3
 557 feasible set of **SDR**, \mathcal{F} , 8
 558 Frobenius norm, 3
 559 generalized trust region subproblem, **GTRS**, 6
 560 hollow subspace of \mathcal{S}^n , \mathcal{S}_H , 4

565 interior, int , 5
566 Lindenstrauss mapping, \mathcal{K} , 4
567 minimal face of S , $\text{face}(S)$, 4
568 Moore-Penrose pseudoinverse, \mathcal{K}^\dagger , 4
569 nearest **EDM** problem with fixed sensors, **NEDMP**,
570 13
571 nearest Euclidean distance matrix with fixed sen-
572 sors, **NEDMF**, 2
573 nearest Euclidean distance matrix, **NEDM**, 3
574 number of sensors, n , 5

575 optimal solutions of **SDR**, Ω , 8
576 orthogonal projection onto \mathcal{S}_C , J_n , 4

577 performance ratio, $r_{n,\eta,M}$, 28

578 semidefinite program, **SDP**, 8
579 semidefinite programming relaxation, **SDR**, 8
580 set of methods used, \mathcal{M} , 28
581 single source localization problem, 2
582 source, c , 5
583 squared least squares, **SLS**, 2
584 subdifferential, ∂ , 24

585 trace inner product, 3

586 vector of all ones, e , 4
587 vectorize a symmetric matrix, svec , 4

588 weighted facially reduced problem, **FNEDM**,
589 17

References

- [1] Abdo Y. Alfakih. *Euclidean distance matrices and their applications in rigidity theory*. Springer, Cham, 2018. 3
- [2] 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. 3, 4
- [3] 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, 5, 6, 8
- [4] 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
- [5] I. Borg and P. Groenen. Modern multidimensional scaling: theory and applications. *Journal of Educational Measurement*, 40(3):277–280, 2003. 3
- [6] E. Candes, M. Rudelson, T. Tao, and R. Vershynin. Error correction via linear programming. In *Proceedings of the 2005 46th Annual EIII Symposium on Foundations of Computer Science, (FOCS’05)*, pages 1–14. IEEE, 2005. 26
- [7] 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. 26
- [8] 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
- [9] T.F. Cox and M.A. Cox. *Multidimensional scaling*. Chapman and hall/CRC, 2000. 3
- [10] G.M. Crippen and T.F. Havel. *Distance geometry and molecular conformation*, volume 74. Research Studies Press Taunton, 1988. 3
- [11] 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
- [12] J. Dattorro. *Convex optimization & Euclidean distance geometry*. Lulu. com, 2010. 3, 4
- [13] 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. 3, 13
- [14] E.D. Dolan and J.J. Moré. Benchmarking optimization software with performance profiles. *Math. Program.*, 91(2, Ser. A):201–213, 2002. 28
- [15] 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. 3
- [16] D. Drusvyatskiy, G. Pataki, and H. Wolkowicz. Coordinate shadows of semidefinite and Euclidean distance matrices. *SIAM J. Optim.*, 25(2):1160–1178, 2015. 16

- 626 [17] D. Drusvyatskiy and H. Wolkowicz. The many faces of degeneracy in conic optimization.
627 *Foundations and Trends[®] in Optimization*, 3(2):77–170, 2017. 3
- 628 [18] H. Fang and D.P. O’Leary. Euclidean distance matrix completion problems. *Optimization*
629 *Methods and Software*, 27(4-5):695–717, 2012. 3
- 630 [19] Y. Gao and D. Sun. A majorized penalty approach for calibrating rank constrained correlation
631 matrix problems. *Technical Report*, 2010. 22
- 632 [20] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins University Press,
633 Baltimore, Maryland, 3rd edition, 1996. 27
- 634 [21] J.C. Gower. Properties of Euclidean and non-Euclidean distance matrices. *Linear Algebra*
635 *Appl.*, 67:81–97, 1985. 4
- 636 [22] T.L. Hayden, J. Wells, W.M. Liu, and P. Tarazaga. The cone of distance matrices. *Linear*
637 *Algebra Appl.*, 144:153–169, 1991. 4
- 638 [23] J.-B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of convex analysis*. Grundlehren Text
639 Editions. Springer-Verlag, Berlin, 2001. Abridged version of it Convex analysis and minimiza-
640 tion algorithms. I [Springer, Berlin, 1993; MR1261420 (95m:90001)] and it II [ibid.; MR1295240
641 (95m:90002)]. 15
- 642 [24] H. Koshima and J. Hoshen. Personal locator services emerge. *IEEE spectrum*, 37(2):41–48,
643 2000. 2
- 644 [25] N. Krislock and H. Wolkowicz. Euclidean distance matrices and applications. In *Handbook on*
645 *Semidefinite, Cone and Polynomial Optimization*, number 2009-06 in International Series in
646 Operations Research & Management Science, pages 879–914. Springer-Verlag, 2011. 3
- 647 [26] T. Kundu. Acoustic source localization. *Ultrasonics*, 54(1):25–38, 2014. 2
- 648 [27] L. Liberti, C. Lavor, N. Maculan, and A. Mucherino. Euclidean distance geometry and appli-
649 cations. *SIAM Rev.*, 56(1):3–69, 2014. 3
- 650 [28] G. Pataki. On the rank of extreme matrices in semidefinite programs and the multiplicity of
651 optimal eigenvalues. *Math. Oper. Res.*, 23(2):339–358, 1998. 10
- 652 [29] T.K. Pong and H. Wolkowicz. The generalized trust region subproblem. *Comput. Optim.*
653 *Appl.*, 58(2):273–322, 2014. 5, 6, 8
- 654 [30] H.-D. Qi. A semismooth Newton method for the nearest Euclidean distance matrix problem.
655 *SIAM Journal on Matrix Analysis and Applications*, 34(1):67–93, 2013. 3, 22, 23, 24
- 656 [31] H.-D. Qi and X. Yuan. Computing the nearest Euclidean distance matrix with low embedding
657 dimensions. *Mathematical Programming*, 147(1):351–389, Oct 2014. 3, 22, 23, 24
- 658 [32] I.J. Schoenberg. Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.*,
659 44(3):522–536, 1938. 4
- 660 [33] R. Stern and H. Wolkowicz. Trust region problems and nonsymmetric eigenvalue perturbations.
661 *SIAM J. Matrix Anal. Appl.*, 15(3):755–778, 1994. 5, 6

- 662 [34] R. Stern and H. Wolkowicz. Indefinite trust region subproblems and nonsymmetric eigenvalue
663 perturbations. *SIAM J. Optim.*, 5(2):286–313, 1995. 8
- 664 [35] P.D. Tao and L.T.H. An. Convex analysis approach to dc programming: Theory, algorithms
665 and applications. *Acta mathematica vietnamica*, 22(1):289–355, 1997. 24, 25
- 666 [36] P.D. Tao and L.T.H. An. The dc (difference of convex functions) programming and dca
667 revisited with dc models of real world nonconvex optimization problems. *Annals of operations
668 research*, 133(1-4):23–46, 2005. 24, 25
- 669 [37] L. Tunçel. *Polyhedral and Semidefinite Programming Methods in Combinatorial Optimization*,
670 volume 27 of *Fields Institute Monographs*. American Mathematical Society, Providence, RI,
671 2010. 16
- 672 [38] J. Warrior, E. McHenry, and K. McGee. They know where you are [location detection]. *IEEE
673 Spectrum*, 40(7):20–25, 2003. 2
- 674 [39] H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors. *Handbook of semidefinite programming*.
675 International Series in Operations Research & Management Science, 27. Kluwer Academic
676 Publishers, Boston, MA, 2000. Theory, algorithms, and applications. 8