Noisy Euclidean distance matrix completion with a single missing node

Stefan Sremac · Fei Wang · Henry Wolkowicz · Lucas Pettersson

Received: 28 March 2019 / Accepted: 21 August 2019 / Published online: 26 August 2019 © The Author(s) 2019

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

Mathematical Subject Classification 90C22 · 15A83 · 90C20 · 62P30

1 Introduction

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

The single source localization problem has applications in e.g. navigation, structural engineering, and emergency response [3,4,8,24,26,38]. In general, it is related to distance geometry problems where the input consists of Euclidean distance measurements and a set of points in Euclidean space. The sensor network localization problem is a generalization of our single source problem, where there are multiple sources and only some of the distance estimates are known. The general Euclidean distance matrix completion problem is yet a further generalization, where sensors do not have specified locations and only partial, possibly noisy, distance information is available, e.g. [2,13,15]. We refer the readers to the books [1,5,9,10] and survey article [27] for background and applications, and to the paper [18] for algorithmic comparisons. We also refer the readers for the related nearest Euclidean distance matrix (NEDM) problem to the papers [30,31] where a semismooth Newton approach and a rank majorization approach is presented. The more general weighted NEDM is a much harder problem though. For theory that relates NEDM to semidefinite programming, see e.g. [12,25].

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

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

1.1 Outline

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

1.2 Preliminaries

We now present some preliminaries and background on SDP and the facial geometry, see e.g. [17]. We denote by $S^n$ the space of $n \times n$ real symmetric matrices endowed with the trace inner product and corresponding Frobenius norm,

$$(X, Y) := \text{trace}(XY) = \sum_{ij} X_{ij} Y_{ij}, \quad \|X\|_F := \sqrt{\text{trace}(XX)} = \sqrt{\sum_{ij} X_{ij}^2}.$$
Unless otherwise specified, the norm of a matrix is the Frobenius norm, and we may drop the subscript \( F \). For a convex set \( C \), the convex subset \( f \subseteq C \) is a face of \( C \) if for all \( x, y \in C, x, y \in f \) with \( z \in (x, y) \), (the open line segment between \( x \) and \( y \)) we have \( z \in f \).

The cone of positive semidefinite matrices is denoted by \( S^n_+ \) and its interior is the cone of positive definite matrices, \( S^n_{++} \). The positive semidefinite cone is pointed, closed and convex. Moreover, the cone \( S^n_+ \) induces a partial order on \( S^n \), that is \( Y \succeq X \) if \( Y - X \in S^n_+ \) and \( Y \succ X \) if \( Y - X \in S^n_{++} \). Every face of \( S^n_+ \) is characterized by the range or nullspace of matrices in its relative interior, equivalently, by matrices of maximum rank. For \( S \subseteq S^n_+ \), we denote the minimal face of \( S \), face\( (S) \), the smallest face of \( S^n_+ \) that contains \( S \). Let \( X \in S^n_+ \) have rank \( r \) with orthogonal spectral decomposition.

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

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

\[
\text{face}(X) = PS^n_{++}P^T = S^n_+ \cap \{QQ^T\}^\perp.
\]

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

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

The centered subspace of \( S^n \), denoted \( S_C \), is defined as

\[
S_C := \{ X \in S^n : Xe = 0 \},
\]

where \( e \) is the vector of all ones. The hollow subspace of \( S^n \), denoted \( S_H \), is

\[
S_H := \{ X \in S^n : \text{diag}(X) = 0 \},
\]

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

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

where \( \| \cdot \|_2 \) denotes the Euclidean norm. As for the Frobenius norm, we assume the norm of a vector to be the Euclidean norm when the subscript is omitted. The set of all \( n \times n \) EDMs, denoted \( E^n \), forms a closed, convex cone with \( E^n \subset S_H \).

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, \( \mathcal{K} : S^n \to S^n \),

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

with adjoint and Moore-Penrose pseudoinverse,

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

respectively. Here Diag is the adjoint of diag, the matrix \( J_n := I - \frac{1}{n} ee^T \) is the orthogonal projection onto \( S_C \), and offDiag\( (D) \) refers to zeroing out the diagonal of \( D \), i.e. the orthogonal projection onto \( S_H \). The range of \( \mathcal{K} \) is exactly \( S_H \) and the range of \( \mathcal{K}^\dagger \) is the subspace \( S_C \). Moreover, \( \mathcal{K}(S^n_+) = E^n \) and \( \mathcal{K} \) is an isomorphism between \( S_C \) and \( S_H \).
The Schoenberg characterization states that $\mathcal{K}$ is an isomorphism between $S^n_+ \cap S_C$ and $E^n$, see [2] for instance. Specifically,

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

Moreover, if $D \in E^n$ and $\mathcal{K}^*(D) = PP^T$ has rank $r$ with full column rank factorization $PP^T$, then the rows of $P$ correspond to the points in $\mathbb{R}^r$ with pairwise distances corresponding to the elements of $D$. For more details, see e.g. [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, $p^1, \ldots, p^n \in \mathbb{R}^r$ denotes their locations, and $r$ is the embedding dimension.

Assumption 2.1 The following holds throughout:

1. $n \geq r + 1$;
2. int conv($p^1, \ldots, p^n$) $\neq \emptyset$;
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 Sect. 3.

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

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

where $\bar{d}_i$ is the true distance and $\epsilon_i$ is a perturbation, or noise. When the noise $\epsilon_1, \ldots, \epsilon_n$ 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

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

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

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

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

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

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

**Theorem 2.2** Let

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

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

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

1. The problem SLS is equivalent to

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

2. The rank of \(A\) is \(r + 1\) and the optimal value of GTRS is finite and attained.

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

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

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

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

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

That we have a zero duality gap 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.

\[
p_{\text{SLS}}^* = \min_y \{ \|Ay - b\|^2 : y^T \bar{I} y + 2\bar{b}^T y = 0 \}
\]

\[
= \min_y \{ \|Ay - b\|^2 - \gamma(y^T \bar{I} y + 2\bar{b}^T y) : y^T \bar{I} y + 2\bar{b}^T y = 0 \}
\]

\[
= \min_y \{ \|Ay - b\|^2 - \gamma(y^T \bar{I} y + 2\bar{b}^T y) : y^T \bar{I} y + 2\bar{b}^T y \leq 0 \} \quad (2.7)
\]

\[
= \max_{\lambda \geq 0} \min_y \{ \|Ay - b\|^2 - \gamma(y^T \bar{I} y + 2\bar{b}^T y) + \lambda(y^T \bar{I} y + 2\bar{b}^T y) \} \quad (2.8)
\]

\[
= \max_{(\lambda - \gamma)} \min_y \{ \|Ay - b\|^2 + (\lambda - \gamma)(y^T \bar{I} y + 2\bar{b}^T y) \}
\]

\[
= d_{\text{SLS}}^* \leq p_{\text{SLS}}^*. \quad (2.10)
\]

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

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

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

The optimal value of (2.8) is a lower bound for \(p_{\text{SLS}}^*\) 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 \(p_{\text{SLS}}^*\). Then there exists \(\tilde{y}\) satisfying,

\[
g(\tilde{y}) < 0, \quad f(\tilde{y}) < p_{\text{SLS}}^*.
\]

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

\[
\lim_{a \to +\infty} g(\tilde{y} \pm \alpha h) = \lim_{a \to +\infty} g(\tilde{y}) \pm 2\alpha \tilde{y}^T h + \alpha^2 \|h\|^2 = +\infty. \quad (2.11)
\]

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

\[
f(\tilde{y} + \eta \alpha h) \leq f(\tilde{y}), \quad \forall \alpha \geq 0.
\]

These observations imply that that there exists \(\tilde{\alpha} > 0\) such that

\[
g(\tilde{y} + \tilde{\alpha} h) = 0, \quad f(\tilde{y} + \tilde{\alpha} h) \leq f(\tilde{y}) < p_{\text{SLS}}^*.
\]

a contradiction.

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

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

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

\[
\begin{align*}
(AA^T + \lambda \tilde{I})y &= A^T b - \lambda \tilde{b}, \\
y^T \tilde{I}y + 2\tilde{b}^Ty &= 0, \\
A^T A + \lambda \tilde{I} &\succeq 0.
\end{align*}
\] (2.12)

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

### 2.2 The semidefinite relaxation, SDR

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

\[
\bar{A} := \begin{bmatrix} A^T A & -A^T b \\ -b^T A & b^T b \end{bmatrix}, \quad \bar{B} := \begin{bmatrix} \tilde{I} & \tilde{b} \\ \tilde{b}^T & 0 \end{bmatrix}.
\] (2.13)

The Lagrangian dual of GTRS may be obtained as follows:

\[
d^*_{\text{SLS}} = \max_{\lambda} \min_y \|Ay - b\|^2 - \lambda(y^T \tilde{I}y + 2\tilde{b}^Ty)
\]

\[
= \max_{\lambda,s} \left\{ s : \|Ay - b\|^2 - \lambda(y^T \tilde{I}y + 2\tilde{b}^Ty) - s \geq 0, \forall y \in \mathbb{R}^{r+1} \right\}
\]

\[
= \max_{\lambda,s} \left\{ s : \begin{bmatrix} y^T, 1 \end{bmatrix} \begin{bmatrix} \bar{A} - \lambda \bar{B} - se_{r+2}e_{r+2}^T \\ y^T, 1 \end{bmatrix}^T \geq 0, \forall y \in \mathbb{R}^{r+1} \right\}
\] (2.14)

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

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

\[
p_{\text{SDR}}^* := \min (\bar{A}, X)
\]

**(SDR)**

\[
s. t. (\bar{B}, X) = 0
\]

\[
X_{r+2, r+2} = 1
\]

\[
X \in S_+^{r+2}
\] (2.15)

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

\[
\rho(y) = \begin{pmatrix} y \\ 1 \end{pmatrix} \begin{pmatrix} y^T \\ 1 \end{pmatrix}.
\] (2.16)
Now equality holds throughout and we have
\[ \langle \bar{A}, \bar{A}, Z \rangle = \theta \langle \bar{A}, Z \rangle + (1 - \theta) \langle \bar{A}, Z \rangle \leq \theta \langle \bar{A}, X \rangle + (1 - \theta) \langle \bar{A}, Y \rangle = \langle \bar{A}, Z \rangle. \]

Note that \( \rho \) is an isomorphism between \( \mathbb{R}^{r+1} \) and rank 1 matrices of \( S_{r+2}^+ \), where the \((r + 2, r + 2)\) element is 1.

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

**Theorem 2.4** The following holds:

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

**Proof** From Theorem 2.2 and weak duality, we have that
\[ p^\ast_{\text{SLS}} \leq p^\ast_{\text{SDR}} \leq p^\ast_{\text{SLS}} \] (2.17)

Moreover, since \( \text{SDR} \) is a relaxation of \( \text{GTRS} \) we get,
\[ p^\ast_{\text{SDR}} \leq p^\ast_{\text{SLS}} \implies p^\ast_{\text{SLS}} = d^\ast_{\text{SLS}} = p^\ast_{\text{SDR}}. \]

Furthermore, from Theorem 2.2 the above values are all finite and the optimal values of \( \text{GTRS} \) and (2.14) are attained. To see that the optimal value of \( \text{SDR} \) 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,
\[ \begin{bmatrix} A^T A + \mu I - A^T b + \mu \bar{b} \\ -b^T A + \mu \bar{b}^T \\ b^T b - s \end{bmatrix} \succeq 0. \]

Setting \( \mu = 0 \) and applying the Schur complement condition, we have
\[ \begin{bmatrix} A^T A & -A^T b \\ -b^T A & b^T b - s \end{bmatrix} > 0 \iff A^T A - \frac{1}{b^T b - s} A^T b (A^T b)^T > 0, A^T A > 0. \]

By Theorem 2.2, \( A^T A \) is positive definite and a Slater point may be obtained by choosing \( s \) so that \( b^T b - s \) is sufficiently large.

Now we consider Item 2.4. By the existence of a Slater point for (2.14) we know that \( \Omega \) is compact and convex. Now we show that \( \Omega \) 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 \( \text{SDR} \) and \( X \) and \( Y \) are feasible for \( \text{SDR} \), we have
\[ \langle \bar{A}, Z \rangle = \theta \langle \bar{A}, Z \rangle + (1 - \theta) \langle \bar{A}, Z \rangle \leq \theta \langle \bar{A}, X \rangle + (1 - \theta) \langle \bar{A}, Y \rangle = \langle \bar{A}, Z \rangle. \]

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

Since \( \Omega \) is a compact convex set it has an extreme point, say \( X^\ast \). Now \( X^\ast \) is also an extreme point of \( \mathcal{F} \), as the relation face of is transitive, i.e. a face of a face is a face. Moreover, since there are exactly two equality constraints in \( \text{SDR} \), by Theorem 2.1 of [28], we have
\[ \text{rank}(X^\ast)(1 + \text{rank}(X^\ast))/2 \leq 2. \]

Equivalently, \( X^\ast = \rho(y^\ast) \) for some \( y^\ast \in \mathbb{R}^{r+1} \). Now, by Lemma 2.3 and the first part of this proof we have that \( y^\ast \) is a minimizer of \( \text{GTRS} \).

For the converse in Item 2.4, let \( y^\ast \) be a minimizer of \( \text{GTRS} \). Then by Lemma 2.3, \( X^\ast := \rho(y^\ast) \) is optimal for \( \text{SDR} \). To see that \( X^\ast \) is an extreme point of \( \Omega \), let \( Y, Z \in \Omega \) such that
\[ \frac{1}{2} Y + \frac{1}{2} Z = X^\ast. \]
Since \( X^* \) has rank 1 and \( Y, Z \geq 0 \), it follows that \( Y \) and \( Z \) are non-negative multiples of \( X^* \). But by feasibility, \( X_{r+2,r+2}^* = Y_{r+2,r+2} = Z_{r+2,r+2} \) and thus \( Y = Z = X^* \). So, by definition, \( X^* \) is an extreme point of \( \Omega \), as desired. \( \square \)

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

**Corollary 2.5**  
The following hold.

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

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

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

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

**2.2.1 A purification algorithm**

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

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

\[
\tilde{B} \leftarrow U^T \tilde{B} U, \quad \tilde{A} \leftarrow U^T \tilde{A} U, \quad \tilde{E} \leftarrow U^T \tilde{E} U, \quad \text{(2.18)}
\]

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

\[
A_S(S) := \begin{pmatrix} \langle \tilde{B}, S \rangle \\ \langle \tilde{A}, S \rangle \\ \langle \tilde{E}, S \rangle \end{pmatrix}, \quad b_S := \begin{pmatrix} 0 \\ p_{\text{SDR}}^* \\ 1 \end{pmatrix}, \quad \text{(2.19)}
\]

respectively. The rank reducing program is

\[
\begin{align*}
\min & \quad \langle C, S \rangle \\
\text{s.t.} & \quad A_S(S) = b_S \\
& \quad S \in \mathcal{S}^r_{++}.
\end{align*} \quad \text{(2.20)}
\]
In Algorithm 2.1 we extend the idea of the rank reducing program and in the subsequent results we prove that the output of the algorithm is a rank 1 optimal solution of SDR.

**Algorithm 2.1 Purification Algorithm**

1. **INPUT**: \( A_S \) as in (2.19) and \( X \in \Omega \).
2. **initialize**: \( k = 1, A_S^1 := A_S, S^1 := X, U^0 = I \).
3. **while** \( \text{rank}(S^k) \geq 2 \) **do**
   4. Compute compact spectral decomposition, \( S^k = U^k D^k (U^k)^T \), with \( D^k \in S^k_{++} \).
   5. Redefine \( A_S^k \) and \( b^k_S \) using \( U^k \) as in (2.18) and ensure that it is full rank.
   6. Choose \( C^k \in \text{Null}(A_S^k) \setminus \{0\} \).
   7. Obtain \( s^{k+1} \in \arg \min \{(C^k, S) : A_S^k(S) = b^k_S, S \geq 0\} \).
   8. Update \( k \leftarrow k + 1 \).
9. **end while**
10. **OUTPUT**: \( X^* := U^0 \cdots U^{k-1} S^k (U^0 \cdots U^{k-1})^T \).

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

\[
S^{k+1} > 0 \iff S^k := \left\{ S \geq 0 : A_S^k(S) = b^k_S \right\} = \left\{ S^{k+1} \right\}.
\]

**Proof** By construction, \( D^k \in S^k \). Therefore,

\[
S^k = \left\{ S^{k+1} \right\} \implies S^{k+1} = D^k > 0.
\]

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

\[
\left\{ S^{k+1} + \varepsilon T, S^{k+1} - \varepsilon T \right\} \subset S^k.
\]

By the choice of \( C^k \), there exists \( T \in \text{Null}(A_S^k) \) 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,

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

contradicting the optimality of \( S^{k+1} \).

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

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

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

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

\[
0 = \langle (U^0 \cdots U^{k-1})^T \tilde{B} U^0 \cdots U^{k-1}, S^k \rangle = \langle \tilde{B}, U^0 \cdots U^{k-1} S^k (U^0 \cdots U^{k-1})^T \rangle = \langle \tilde{B}, X^k \rangle.
\]
Similarly the other two constraints comprising \( A_k^{k-1} \) are satisfied by \( X^k \) and therefore \( X^k \in \Omega \).

Now we show that the sequence of ranks, \( r_1, r_2, \ldots \), generated by Algorithm 2.1 is strictly decreasing. It immediately follows that the algorithm terminates in at most \( \text{rank}(\bar{X}) - 1 \) 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,

\[
\min \langle C^{k-1}, S \rangle : A_k^{k-1}S = b^{k-1}_S, \quad S \succeq 0.
\] (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 \( \Omega \). 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

\[
\text{range}(X^k) \supseteq \{ \text{range}(Y^k), \text{range}(Z^k) \}.
\]

Thus there exist \( V^k, W^k \in S^{r_k-1}_+ \) such that,

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

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

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

### 3 EDM Formulation

In this section we use the Lindenstrauss operator, \( K \), 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

\[D_{Tc} := \begin{bmatrix} D_T & d \circ d \\ (d \circ d)^T & 0 \end{bmatrix} \in S^{n+1}_+.\]

Recall that

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

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

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

\[K(G_T) = D_T, \quad K^\dagger(D_T) = G_T.\]
The nearest EDM problem with fixed sensors is
\[
\min_{x \in \mathbb{R}^r} \frac{1}{2} \| \mathcal{K} \left( \begin{bmatrix} P_T & P_T \\ x^T & x^T \end{bmatrix} \right) - D_{Tc} \|^2 .
\] (3.1)

For any \( x \in \mathbb{R}^n \) let
\[
d_x := \left( \begin{array}{c} \| x - p^1 \|^2 \\ \vdots \\ \| x - p^n \|^2 \end{array} \right).
\]

By simplifying the objective, we see that the NEDMP problem in (3.1) is indeed equivalent to SLS, i.e.
\[
\frac{1}{2} \| \mathcal{K} \left( \begin{bmatrix} P_T & P_T \\ x^T & x^T \end{bmatrix} \right) - D_{Tc} \|^2 = \frac{1}{2} \| \mathcal{K} \left( \begin{bmatrix} G_T & P_T x \\ (P_T x)^T & 0 \end{bmatrix} \right) - D_{Tc} \|^2
\]
\[
= \frac{1}{2} \left\| \begin{bmatrix} D_T & d_x \\ d_x^T & 0 \end{bmatrix} - \begin{bmatrix} D_T & d \circ d \\ (d \circ d)^T & 0 \end{bmatrix} \right\|^2
\]
\[
= \sum_{i=1}^n \left( \| x - p^i \|^2 - d_i^2 \right)^2.
\]

The approach of [13] for the related sensor network localization problem is to replace the matrix \( \begin{bmatrix} P_T & P_T \\ x^T & x^T \end{bmatrix} \) in (3.1) with the positive semidefinite matrix variable \( X \in S^{n+1} \), 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,
\[
V_S := \min \frac{1}{2} \| \mathcal{H}_c \circ (\mathcal{K}(X) - D_{Tc}) \|^2 \\
\text{s.t.} \quad H_T \circ (\mathcal{K}(X) - D_{Tc}) = 0, \\
\quad \text{rank}(X) \leq r, \\
\quad X \succeq 0,
\] (NEDMF) (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
\[
X = \begin{bmatrix} X_T & X_c \\ x_c^T & x_c^T \end{bmatrix}^T,
\]
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 affects 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
\[
\min \{ \| X_T Q - P_T \|^2 : Q^T Q = Q Q^T = I, \quad Q \in \mathbb{R}^{r \times r} \},
\] (3.3)
to obtain the rotation and thus have a complete description of the transformation from $X_T$ to $P_T$. Applying the transformation to $x_c$ yields a vector feasible for SLS. Thus every feasible solution of (3.2), corresponds to a feasible solution of SLS. The converse is trivially true and we conclude that (3.2) is equivalent to SLS due to the rank constraint. We show in the subsequent sections that the relaxation where the rank and the linear constraints are dropped, may be used to solve the problem accurately in a large number of instances.

3.1 The relaxed NEDM problem

3.1.1 Nearest Euclidean distance matrix formulation

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

$$\begin{align*}
(NEDM) \quad & \min \frac{1}{2} \| K(X) - D_T \|_2^2 \\
& \text{s.t. rank}(X) \leq r \\
& X \succeq 0.
\end{align*}$$

(3.4)

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

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

$$\begin{align*}
(NEDMinv) \quad & \min \frac{1}{2} \| X - K^\dagger(D_T) \|_2^2 \\
& \text{s.t. rank}(X) \leq r \\
& X \succeq 0.
\end{align*}$$

(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 $K^\dagger(D_T)$ 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 $\overline{D} \in S^{n+1}$ the true EDM of the sensors and the source, that is

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

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

$$D_T = \overline{D} + e_{n+1} \xi^T + \xi e_{n+1}^T, \quad \text{where} \; \xi := \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix}.$$
Theorem 3.1 The rank of $\mathcal{K}^\dagger(D_T)$ is at most $r + 2$. Moreover, $\mathcal{K}^\dagger(D_T)$ has at most 1 negative eigenvalue with magnitude bounded above by $\frac{\sqrt{2}}{2} \|J_{n+1}\| \|\varepsilon\|$.

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 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 $P_{n+1}$ and $Q_{n+1}$, respectively. Now we have,

$$\mathcal{K}^\dagger(D_T) = -\frac{1}{2}J_{n+1}D_TJ_{n+1} = -\frac{1}{2}(J_{n+1}D_jJ_{n+1} + J_{n+1}Q_{n+1}) + \left(-\frac{1}{2}J_{n+1}P_{n+1}\right),$$

where the first term in the last line is positive semidefinite with at least $r$ and at most $r + 1$ positive eigenvalues ($-J_{n+1}D_jJ_{n+1}$ is a positive semidefinite matrix with rank $r$ and $-J_{n+1}Q_{n+1}$ 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 $X, Y \in \mathbb{S}^n$,

$$\|XY\| \leq \|X\| \|Y\|.$$  (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 $-S_{+}^{n+1}$, we have

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

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

The following corollary follows immediately.

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

3.1.2 Weighted, facially reduced NEDM

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

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

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

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

$\begin{align*}
\min \frac{1}{2} \| \mathcal{K}(X) - DT_c \|^2 \\
\text{s.t. } \text{rank}(X) \leq r \\
X \in \text{face}(F_T, S_+^{n+1}).
\end{align*}$

(\text{NEDMP})

(3.9)

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

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

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

\begin{align*}
\langle G_T, W_T W_T^T \rangle &= 0, \\
\text{rank}(G_T + W_T W_T^T) &= n - 1 = \max_{X \in S_{c,+}^n} \text{rank}(X).
\end{align*}

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

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

1. $W_T W_T^T$ exposes $\text{face}(F_T, S_+^{n+1})$.
2. $W$ exposes $\text{face}(F_T, S_+^{n+1})$.

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

Note that $\text{face}(\mathcal{K}^+(D), S_+^{n+1}) \subseteq \text{face}(F_T, S_+^{n+1})$ since,

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

Through $W$ we have a ‘nullspace’ characterization of $\text{face}(F_T, S_+^{n+1})$. 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 null($W$). One such choice is,

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

(3.10)

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

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

It follows that

$\text{face}(F_T, S_+^{n+1}) = S_+^{n+1} \cap W^\perp = VS_+^{r+1} V^T.$

(3.11)
Thus we may replace the variable $X$ in NEDMP by $VRV^T$ for $R \in S_{+}^{r+1}$. To simplify the notation, we define the composite map $K_{V} := K(V \cdot V^T)$. Moreover, we introduce a weight matrix to the objective and obtain the weighted facially reduced problem, FNEDM,

$$V_{\alpha} := \min \frac{1}{2} ||H_{\alpha} \circ (K_{V}(R) - DT_{c})||^2, \quad (\text{FNEDM})$$

s.t. $\text{rank } R \leq r,$

$$R \succeq 0.$$  

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

We begin our analysis by proving that $V_{\alpha}$ is attained.

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

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

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

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

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

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

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

$$0 \preceq \nabla f(R) = V^T \left(H_{\alpha}^* \circ K^* \left( (H_{\alpha} \circ K)(VRV^T) - H_{\alpha} \circ DT_{c} \right) \right) V, \quad \langle \nabla f(R), R \rangle = 0.$$  

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

**Proof** From the Pshenichnyi–Rockafellar conditions, $R$ is optimal if, and only if, $\nabla f(R) \in (S_{+}^{r+1} - R)^+$, the nonnegative polar cone. This condition holds if, and only if, for all $X \in S_{+}^{r+1}$ 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 $\alpha$ may be arbitrarily large we get that $\langle \nabla f(R), X \rangle \geq 0$ for all $X \in S_{+}^{r+1}$. Therefore, we conclude that $\nabla f(R) \in (S_{+}^{r+1})^+ = S_{+}^{r+1}$. Moreover, setting $X = 0$, we get,

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

hence orthogonality holds. 

\(\square\)
3.1.3 Analysis of FNEDM

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

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

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

$$P = \begin{bmatrix} T \\ c^T \end{bmatrix}, \quad T \in \mathbb{R}^{n \times r}, \quad c \in \mathbb{R}^r.$$  \hspace{1cm} (3.13)

Then there exists a matrix $Q \in \mathbb{R}^{r \times r}$ such that $P_T Q = J_n T$ if, and only if, $PP_T \in VS_{r+1}^{r+1} V^T$.

**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,

$$P_T Q = J_n T = T - \frac{1}{n} ee^T T = T - \frac{1}{n} ec^T,$$

which yields the following expression for $P$,

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

Now by (3.11) we have,

$$PP_T \in VS_{r+1}^{r+1} V^T \iff PP_T \in S_n^{r+1} \cap W \iff PP_T W = 0.$$

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

$$P^T W = P^T (W_T W_T + ee^T)$$

$$= P^T W_T W_T$$

$$= \begin{bmatrix} Q^T P_T^T + \frac{1}{n} c e^T \\ c \end{bmatrix} \begin{bmatrix} W_T \\ 0 \end{bmatrix} W_T^T$$

$$= \left( Q^T P_T^T W_T + \frac{1}{n} c e^T W_T \right) W_T^T$$

$$= 0,$$

as desired.

For the other direction, let

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

and recall that $V = J_{n+1} V_1$. Suppose $PP_T^T$ belongs to the face $VS_{r+1}^{r+1} V^T$. 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 $PTQ = JnT$. To this end, let $\bar{J} = [Jn 0]$ and observe that $\bar{J}P = JnT$. Moreover, since $\bar{J}$ is centered, $\bar{J}J_{n+1} = \bar{J}$. Then,

$$JnT = \bar{J}P = \bar{J}J_{n+1}V_1M = \bar{J}V_1M = JnPTQ = PTQ,$$

as desired. \(\square\)

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

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

**Lemma 3.7** Consider the problem,

$$V_T := \min \frac{1}{2} ||H_c \circ (K_V(R) - D_{Tc})||^2 =: h(R)$$

s.t. $H_T \circ (K_V(R) - D_{Tc}) = 0$

$$\text{rank } R \leq r$$

$$R \in S_{r+1}^r.$$

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

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

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

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

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

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

$$\mathcal{K}(TT^T) = D_T \implies K^\dagger K(TT^T) = K^\dagger (D_T) \implies JnTT^TJ_n = PTPT^T.$$

Thus, there exists an orthogonal $Q$ such that $JnT = PTQ$. By Theorem 3.6 we have $X \in VSS_{r+1}^rV^T$ and it follows that $V_S \geq V_T$. \(\square\)

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

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

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

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

$$f(R, \alpha_1) = \frac{1}{2} ||H_{\alpha_1} \circ (K_V(R) - D_{Tc})||^2$$

$$= \frac{1}{2} \| (\alpha_1 H_T + H_c) \circ (K_V(R) - D_{Tc}) \|^2$$

\(\square\) Springer
\[
\frac{1}{2} \alpha_2^2 \|H_T \circ (\mathcal{K}_V (R) - D_T)\|^2 + \frac{1}{2} \|H_c \circ (\mathcal{K}_V (R) - D_T)\|^2 \leq \frac{1}{2} \alpha_2^2 \|H_T \circ (\mathcal{K}_V (R) - D_T)\|^2 + \frac{1}{2} \|H_c \circ (\mathcal{K}_V (R) - D_T)\|^2 = f(R, \alpha_2).
\]

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

For the second inequality, we first observe that,
\[
V_T = \min f(R, \alpha_2)
\]
\[
s.t. H_T \circ (\mathcal{K}_V (R) - D_T) = 0
\]
\[
\text{rank}(R) \leq r
\]
\[
R \succeq 0.
\]

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

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

**Proof** First we note that \(R_\alpha\) is well defined by Lemma 3.4. Now, from Lemma 3.8, we have that \(V_\alpha\) is monotonically increasing and bounded above by \(V_T\). Hence there exists \(V^*\) such that,
\[
V_\alpha \uparrow V^* \leq V_T. \tag{3.15}
\]

Next, we show that \(R\) is feasible for (3.14). Since \(S^{r+1}_+\) is closed and the rank function is lower semicontinuous, we have \(\text{rank}(R) \leq r\) and \(R \succeq 0\). Moreover, for every \(\ell \in \mathbb{N}\),
\[
V_{\alpha_\ell} = f(R_{\alpha_\ell}, \alpha_\ell) = \frac{1}{2} \alpha_\ell^2 \|H_T \circ (\mathcal{K}_V (R_{\alpha_\ell}) - D_T)\|^2 + h(R_{\alpha_\ell})
\]
Rearranging and taking the limit we get,
\[
0 \leq \lim_{\ell \to +\infty} \frac{1}{2} \alpha_\ell^2 \|H_T \circ (\mathcal{K}_V (R_{\alpha_\ell}) - D_T)\|^2 = \lim_{\ell \to +\infty} V_{\alpha_\ell} - h(R_{\alpha_\ell}) = V^* - h(R). \tag{3.16}
\]

The last equality follows from the continuity of \(h\). Since the limit in (3.16) exists we get,
\[
0 = \lim_{\ell \to +\infty} \|H_T \circ (\mathcal{K}_V (R_{\alpha_\ell}) - D_T)\| = \|H_T \circ (\mathcal{K}_V (R) - D_T)\|, \tag{3.17}
\]
by continuity. Thus \(R\) is feasible for (3.14) and we have \(h(R) \geq V_T\). On the other hand, from (3.16) we have \(h(R) \leq V^*\). Combining these observations with (3.15) we get,
\[
h(R) \leq V^* \leq V_T \leq h(R). \tag{3.18}
\]

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

### 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,
\[
R_{LS} := (H_\alpha \circ \mathcal{K}_V)^\dagger (H_\alpha \circ D_T) \in \arg \min f(R). \tag{3.19}
\]
In this subsection we explore the relationship between the optimal solution of FNEDM and the eigenvalues of \( R_{LS} \). In general the Moore–Penrose inverse may be difficult to obtain, however, the following result implies that \( R_{LS} \) may be derived efficiently and it is the unique minimizer of \( f \).

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

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

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

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

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

- **Case I** \( R_{LS} \geq 0 \) and \( \text{rank}(R_{LS}) \leq r \).
- **Case II** \( R_{LS} \notin S_r^{++} \).
- **Case III** \( R_{LS} > 0 \).

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

**Theorem 3.11** Let \( R^* \) denote the minimizer of the relaxation of FNEDM where the rank constraint is removed. If \( R_{LS} \notin S_r^{++} \), then \( R^* \) is a minimizer of FNEDM.

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

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

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

\[
\begin{align*}
\min & \quad \frac{1}{2}||T(Y - D_{T_c})T||^2 \\
\text{s.t.} & \quad \text{diag}(Y) = 0, \\
& \quad \langle J_{n+1}YJ_{n+1}, W \rangle = 0, \\
& \quad -J_{n+1}YJ_{n+1} \succeq 0, \\
& \quad \text{rank}(J_{n+1}YJ_{n+1}) \leq r.
\end{align*}
\]
As in [30,31], we define $\mathcal{K}^{n+1}_+(r) := \{ Y \in S^{n+1} | - J_{n+1} Y J_{n+1} \succeq 0, \ \text{rank}(J_{n+1} Y J_{n+1}) \leq r \}$, and let $\mathcal{B}(Y) = 0$ represent the linear constraints $\text{diag}(Y) = 0$ and $\langle J_{n+1} Y J_{n+1}, W \rangle = 0$. Then (3.20) may be written as,

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

(3.21)

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

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

$$- \min_Y \frac{1}{2} \sum_{\kappa^{n+1}_T(r)} (TD_{Tc} T + B^* (y)) \| 2 - \frac{1}{2} \| T D_{Tc} T \|^2,$$

(3.22)

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

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

$$L(Y, y) = \frac{1}{2} \| T(Y - D_{Tc})T \|^2 - \langle \mathcal{B}(Y), y \rangle$$

$$= \frac{1}{2} \left( \| T(Y - (D_{Tc} + T^{-1} B^* (y) T^{-1})) T \|^2 + \| T D_{Tc} T \|^2 - \| T(D_{Tc} + T^{-1} B^* (y) T^{-1}) T \|^2 \right).$$

(3.23)

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

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

(3.24)

$$= \frac{1}{2} \sum_{\kappa^{n+1}_T(r)} (TD_{Tc} T + B^* (y)) - (TD_{Tc} T + B^* (y)) \| 2$$

(3.25)

$$- \frac{1}{2} \| TD_{Tc} T + B^* (y) \|^2 + \frac{1}{2} \| TD_{Tc} T \|^2$$

(3.26)

$$= - \frac{1}{2} \sum_{\kappa^{n+1}_T(r)} (TD_{Tc} T + B^* (y)) \| 2 + \frac{1}{2} \| TD_{Tc} T \|^2$$

(3.27)

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

$$\left\| \sum_{\kappa^{n+1}_T(r)} (TD_{Tc} T + B^* (y)) \right\|^2 + \left\| \sum_{\kappa^{n+1}_T(r)} (TD_{Tc} T + B^* (y)) \right\|^2$$

$$= \left\| TD_{Tc} T + B^* (y) \right\|^2.$$

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

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

The Lagrangian dual problem is then defined by,

\[
V_d := \max_{\theta} \left( -\min_y \frac{1}{2} \left\| T \prod K \n + 1 \right\| \right) + \frac{1}{2} \left\| T D_{Tc} T + B^*(y) \right\|^2 - \frac{1}{2} \left\| T D_{Tc} T \right\|^2,
\]

as desired. \( \square \)

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

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

To penalize rank, we consider the concave penalty function,

\[
p : S_{r+1} \rightarrow \mathbb{R}, \quad p(R) := \langle I, R \rangle - \sum_{i=1}^{r} \lambda_i(R).
\]

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

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

Hence, \( p \) is an appropriate penalty function for the rank constraint of \textbf{FNEDM}. Now we consider the penalized version of \textbf{FNEDM},

\[
\begin{aligned}
\text{(PNEDM)} \quad & \min_{R} \frac{1}{2} \| H_{\alpha} \circ (\mathcal{K}_V(R) - D_{Tc}) \|^2 + \gamma p(R), \\
\text{s.t.} \quad & R \succeq 0.
\end{aligned}
\]

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

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

\begin{theorem}
Suppose Algorithm 3.1 converges to a stationary point \( \bar{R} \), and that \( \text{rank}(\bar{R}) = r \). Then \( \bar{R} \) is a global minimizer of \textbf{FNEDM} restricted to face(\( \bar{R} \)).
\end{theorem}
Algorithm 3.1: Majorization Algorithm

1: INPUT: \( R_0 \succeq 0, \gamma > 0, 1 > \epsilon > 0 \)
2: initialize: \( k = 0, err = 1 \)
3: while \( err > \epsilon \) do
4: Choose \( U^k \in \partial p(R^k) \)
5: Obtain \( R^{k+1}, R^{k+1} \in \arg\min_{R \succeq 0} \left[ \frac{1}{2} ||H_\alpha \circ (K_V(R)) - D_Tc||^2 + \gamma (\rho p(R^k) + \langle U^k, R - R^k \rangle) \right] \) \( (3.31) \)
6: Update \( err \leftarrow ||R^{k+1} - R^k||, k \leftarrow k + 1 \)
7: end while

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

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

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

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

and

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

Therefore we have

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

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

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

Hence our claim is proved. \( \square \)

3.1.5 Identifying outliers using \( l_1 \) minimization and facial reduction

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

Problem (3.12) is equivalent to minimizing the residual of an overdetermined linear system in the domain of an SDP cone. Let \( z := \text{svec}(R) \) for \( R \in S^{r+1} \). Abusing our previous notation, let \( b := \text{svec}(H_\alpha \circ D_{T_\alpha}) \) and let \( A \) denote the matrix representation of \( H_\alpha \circ K_V \). Then \( z \in \mathbb{R}^{(r+1)(r+2)/2} \) and \( b \in \mathbb{R}^{n(n+1)/2} \). In practice, \( n \) is much larger than \( r+1 \), so \( A \) will have more rows than columns. In other words, we have an overdetermined system. Under this new notation, problem (3.12) is equivalent to,
min \|\delta\| \\
\text{s.t} \ A z - b = \delta \\
\text{sMat}(z) \succeq 0 \quad (3.33)

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

\[
\min \|\delta\|_1 \\
\text{s.t} \ A z - b = \delta \\
\text{sMat}(z) \succeq 0. \quad (3.34)
\]

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

The problem (3.34) differs from the classical compressed sensing model in the positive semidefinite constraint. However, in our numerical tests, we have found that adding the positive semidefinite constraint greatly increases the success rate in identifying outliers. In compressed sensing, if the matrix \( N \) satisfy the so-called restricted isometry property, then the sparse signal can be recovered exactly [7, Theorem 1.1]. However, there are no practical algorithms available right now to check if a given matrix satisfies the restricted isometry property. If \( \delta_0 \) is the solution to (3.34) and most of the elements of \( \delta_0 \) are 0, then the non-zero elements indicate the outlier measurements.

Thus far, we have assumed that most of the measurements in \( b \) are exact and a few have large error. Now let us revert to the original assumption of this section: that most elements of \( b \) are slightly inaccurate and few elements are very inaccurate. If the positive semidefinite constraint is ignored, then the identification of outliers is guaranteed to be accurate assuming that \( N \) satisfies the restricted isometry property. To be specific, if \( \delta^\# \) represents the optimal solution of (3.34) without the positive semidefinite constraint, then \( ||\delta^\# - \delta_0||_2 \leq C_\xi \cdot \epsilon \) where \( C_\xi \) and \( \epsilon \) are small constants [6,7]. The specifics for our outlier-detection algorithm are stated in Algorithm 3.2.

\textbf{Algorithm 3.2 Removing Outliers}

1: \textbf{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

\[
\min \|K_V(R) - D_T_c\|_1, \\
\text{s.t} \ R \succeq 0. \quad (3.35)
\]

3: Obtain \( \delta := (K_V(R) - D_T_c)|_{1:n,n+1} \).
4: Normalize: \( \delta \leftarrow \frac{1}{||\delta||_2} \).
5: Remove \( p_i \) from \( P_T \) and \( d_i \) from \( d \) for all \( i \) satisfying \( \delta_i \geq \frac{1}{\sqrt{n}} \).
6: \textbf{OUTPUT:} Sensor matrix \( P_T \) and distance vector \( d \) with outliers removed.
3.2 Recovering source position from gram matrix

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

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

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

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

\[
\min_Q \| P_T - \bar{P}_f Q \|_F^2 \quad \text{s.t.} \quad Q^T Q = I_r.
\]

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

The second approach is to solve the least square problem

\[
\min_Q \| P_T - \bar{P}_f Q \|_F^2 \quad \text{s.t.} \quad Q \in \mathbb{R}^{r \times r}.
\]

The least square solution is \( \bar{Q} = \bar{P}_f \bar{P}_f \). Recall that \( \bar{P}_f \) is the Moore–Penrose generalized inverse of \( \bar{P}_f \). The recovered position of the source is then \( p_f^T = p_f^T \bar{Q} \).

4 Numerical results

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

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

where \( D \) is the generated EDM and \( \varepsilon \in U(-\eta, \eta) \). The outliers are obtained by multiplying (4.1) by another factor \( \theta \) 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, \( c_M^r \), between the true location of the source, \( c \), and the location obtained using method \( M \), denoted \( c_M \), is given by

\[ c_M^r = \frac{\|c_M - c\|}{\|c\|}. \]

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

The methods in the tables are labelled according to the models with some additional prefixes. To be specific, the \( L \) and \( P \) prefixes represent the different ways used to obtain the position of the source, \( c \). By \( L \) we denote the least square approach of (3.37) and \( P \) represents the Procrustes approach in (3.36). We choose \( \alpha = 1 \) in \text{FNEDM} and the constant \( \gamma \) for \text{PNEDM} in (3.30) is chosen to be 1000.
Table 1 The mean relative error \( c_{Mre} \) of 100 simulations for varying amount of sensors and error factors with no outliers for dimension \( r = 3 \)

| Error factor \( \eta \) | \( \eta = 0.002 \) | \( \eta = 0.02 \) | \( \eta = 0.2 \) |
|------------------------|------------------|------------------|------------------|
| \# 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 2 The mean relative error \( c_{Mre} \) of 100 simulations for varying amount of sensors and error factors with no outliers for dimension \( r = 3 \)

| Error factor \( \eta \) | \( \eta = 0.005 \) | \( \eta = 0.05 \) | \( \eta = 0.15 \) |
|------------------------|------------------|------------------|------------------|
| \# 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           |

We report some results in the following table.

From Tables 1 and 2 we can see generally P-FNEDM has the smallest error, and occasionally L-FNEDM is better. Also we can see that as the number of towers \( n \) increases, the relative error \( c_{Mre} \) decreases which is expected as we have more sensors, the location of the source should be more accurate.

To compare the overall performance of all the methods, we use the well known performance profiles [14]. The approach is outlined below.

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

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

We then compute the performance ratio,

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

and the function,

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

The performance profile is a plot of \( \psi_M(\tau) \) for \( \tau \in (1, +\infty) \) and all choices of \( M \in \mathcal{M} \). Note that \( r_{n,\eta,M} \geq 1 \) and equality holds if, and only if, the solution obtained by \( M \) is best for the pair \((n, \eta)\). In general, smaller values of \( r_{n,\eta,M} \) indicate better performance. The function \( \psi_M(\tau) \) measures how many pairs \((n, \eta)\) were solved with a performance ratio of \( \tau \) or better. The function is monotonically non-decreasing and larger values are better.
The performance profiles can be seen in Fig. 1a and b, the P-FNEDM approach has the best performance over all 5 methods. Also using the Procrustes approach (3.36) is better than using the least squares approach (3.37). Allowing the sensors to move in FNEDM model is better than fixing the sensors in SDR or making the sensors completely free in NEDM for recovering the location of the source.

We also generate the data with outliers. In FNEDM, the outliers are detected and removed using the $\ell_1$ norm approach described in Sect. 3.1.5. We report the results with outliers added in the following Tables 3 and 4.

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

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

Fig. 1 Performance profiles for $\psi_M(\tau)$ with $n = [5, 10, 15], r = 3$, no outliers
Table 3 The mean relative error $c_{re}^M$ of 100 simulations for varying amount of sensors and error factors with 1 outlier for dimension $r = 3$

| Error factor $\eta$ | $\eta = 0.001$ | $\eta = 0.01$ | $\eta = 0.1$ |
|---------------------|-----------------|---------------|---------------|
| # 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 |

Outlier factor $\theta \sim U(5, 10)$

Table 4 The mean relative error $c_{re}^M$ of 100 simulations for varying amount of sensors and error factors with 2 outliers for dimension $r = 3$

| Error factor $\eta$ | $\eta = 0.001$ | $\eta = 0.01$ | $\eta = 0.1$ |
|---------------------|-----------------|---------------|---------------|
| # 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 |

Outlier factor $\theta \sim U(3, 6)$

From Table 3 and 4 we can see clearly that when outliers are added, the FNEDM outperforms both SDR and NEDM with a big improvement, as the outliers can be removed. It is also consistent with our previous conclusion that using the Procrustes approach (3.36) is better than using the least squares approach (3.37).

5 Conclusion

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

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

Finally, we used the $\ell_1$ norm approach in Algorithm 3.2, to remove outlier measurements. In Tables 3 and 4 we demonstrate the effectiveness of this approach.
Acknowledgements

Open access funding provided by Royal Institute of Technology.

Open Access

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Alfakih, A.Y.: Euclidean Distance Matrices and Their Applications in Rigidity Theory. Springer, Cham (2018)
2. Alfakih, A.Y., Khandani, A., Wolkowicz, H.: Solving Euclidean distance matrix completion problems via semidefinite programming. Comput. Optim. Appl. 12(1–3), 13–30 (1999). A tribute to Olvi Mangasarian
3. Beck, A., Stoica, P., Li, J.: Exact and approximate solutions of source localization problems. IEEE Trans. Signal Process. 56(5), 1770–1778 (2008)
4. Beck, A., Teboulle, M., Chikishev, Z.: Iterative minimization schemes for solving the single source localization problem. SIAM J. Optim. 19(3), 1397–1416 (2008)
5. Borg, I., Groenen, P.: Modern multidimensional scaling: theory and applications. J. Educ. Meas. 40(3), 277–280 (2003)
6. Candès, E., Rudelson, M., Tao, T., Vershynin, R.: Error correction via linear programming. In: Proceedings of HITE 2005 46th Annual EIII Symposium on Foundations of Computer Science, (FOCS’05), pp. 1–14. IEEE, New York (2005)
7. Candès, E.J., Romberg, J.K., Tao, T.: Stable signal recovery from incomplete and inaccurate measurements. Commun. Pure Appl. Math. 59(8), 1207–1223 (2006)
8. Cheung, K.W., So, H.-C., Ma, W.-K., Chan, Y.-T.: Least squares algorithms for time-of-arrival-based mobile location. IEEE Trans. Signal Process. 52(4), 1121–1130 (2004)
9. Cox, T.F., Cox, M.A.: Multidimensional Scaling. Chapman and hall/CRC, Boca Raton (2000)
10. Critchley, F.: Dimensionality theorems in multidimensional scaling and hierarchical cluster analysis. In: Data Analysis and Informatics (Versailles, 1985), pp. 45–70. North-Holland, Amsterdam (1986)
11. Dattorro, J.: Convex optimization & Euclidean distance geometry. Lulu. com (2010)
12. Ding, Y., Krislock, N., Qian, J., Wolkowicz, H.: Sensor network localization, Euclidean distance matrix completions, and graph realization. Optim. Eng. 11(1), 45–66 (2010)
13. Dolan, E.D., Moré, J.J.: Benchmarking optimization software with performance profiles. Math. Program. 91(2, Ser. A), 201–213 (2002)
14. Drusvyatskiy, D., Krislock, N., Cheung Voronin, Y.-L., Wolkowicz, H.: Noisy Euclidean distance realization: robust facial reduction and the Pareto frontier. SIAM J. Optim. 27(4), 2301–2331 (2017)
15. Drusvyatskiy, D., Pataki, G., Wolkowicz, H.: Coordinate shadows of semidefinite and Euclidean distance matrices. SIAM J. Optim. 25(2), 1160–1178 (2015)
16. Drusvyatskiy, D., Wolkowicz, H.: The many faces of degeneracy in conic optimization. Found. Trends® Optim. 3(2), 77–170 (2017)
17. Fang, H., O’Leary, D.P.: Euclidean distance matrix completion problems. Optim. Methods Softw. 27(4–5), 695–717 (2012)
18. Gao, Y., Sun, D.: A majorized penalty approach for calibrating rank constrained correlation matrix problems. Technical Report (2010)
19. Golub, G.H., Van Loan, C.F.: Matrix Computations, 3rd edn. Johns Hopkins University Press, Baltimore (1996)
20. Gower, J.C.: Properties of Euclidean and non-Euclidean distance matrices. Linear Algebra Appl. 67, 81–97 (1985)
21. Hayden, T.L., Wells, J., Liu, W.M., Taragaza, P.: The cone of distance matrices. Linear Algebra Appl. 144, 153–169 (1991)
22. Hiriart-Urruty, J.-B., Lemaréchal, C.: Fundamentals of Convex Analysis, Grundlehren Text Editions. Springer, Berlin (2001). Abridged version of it Convex analysis and minimization algorithms. I [Springer, Berlin, 1993; MR1261420 (95m:90001)] and it II [ibid.; MR1295240 (95m:90002)]
23. Koshima, H., Hoshen, J.: Personal locator services emerge. IEEE Spectr. 37(2), 41–48 (2000)
25. Krislock, N., Wolkowicz, H.: Euclidean distance matrices and applications. In: Handbook on Semidefinite, Cone and Polynomial Optimization, Number 2009-06 in International Series in Operations Research & Management Science, pp. 879–914. Springer, Berlin (2011)
26. Kundu, T.: Acoustic source localization. Ultrasonics 54(1), 25–38 (2014)
27. Liberti, L., Lavor, C., Maculan, N., Mucherino, A.: Euclidean distance geometry and applications. SIAM Rev. 56(1), 3–69 (2014)
28. Pataki, G.: On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. Math. Oper. Res. 23(2), 339–358 (1998)
29. Pong, T.K., Wolkowicz, H.: The generalized trust region subproblem. Comput. Optim. Appl. 58(2), 273–322 (2014)
30. Qi, H.-D.: A semismooth Newton method for the nearest Euclidean distance matrix problem. SIAM J. Matrix Anal. Appl. 34(1), 67–93 (2013)
31. Qi, H.-D., Yuan, X.: Computing the nearest Euclidean distance matrix with low embedding dimensions. Math. Program. 147(1), 351–389 (2014)
32. Schoenberg, I.J.: Metric spaces and positive definite functions. Trans. Am. Math. Soc. 44(3), 522–536 (1938)
33. Stern, R., Wolkowicz, H.: Trust region problems and nonsymmetric eigenvalue perturbations. SIAM J. Matrix Anal. Appl. 15(3), 755–778 (1994)
34. Stern, R., Wolkowicz, H.: Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations. SIAM J. Optim. 5(2), 286–313 (1995)
35. Tao, P.D., An, L.T.H.: Convex analysis approach to dc programming: theory, algorithms and applications. Acta Math. Vietnam. 22(1), 289–355 (1997)
36. Tao, P.D., An, L.T.H.: The dc (difference of convex functions) programming and dca revisited with dc models of real world nonconvex optimization problems. Ann. Oper. Res. 133(1–4), 23–46 (2005)
37. Tuncel, L.: Polyhedral and Semidefinite Programming Methods in Combinatorial Optimization. Fields Institute Monographs, vol. 27. American Mathematical Society, Providence, RI (2010)
38. Warrior, J., McHenry, E., McGee, K.: They know where you are [location detection]. IEEE Spectr. 40(7), 20–25 (2003)
39. Wolkowicz, H., Saigal, R., Vandenberghe, L. (eds.): Handbook of Semidefinite Programming. International Series in Operations Research & Management Science, vol. 27. Kluwer Academic Publishers, Boston, MA (2000)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.