École polytechnique fédérale de Lausanne

Institute of Mathematics

Master's semester project thesis

Spring Semester 2023

Relaxation in distance-based localization problem in the noiseless case

Carried out in the Chair of Continuous Optimization

Under the supervision and direction of
Andrew Ducan Mc Rae and Nicolas Boumal

Authored by
Kieran Vaudaux

Contents

1 Introduction

Distance-based problems is present in numerous domains, including computer vision, robotics, molecular biology, and wireless sensor networks. The underlying challenge in these problems lies in determining the spatial configuration of points given only some of their pairwise distances, possibly distorted by noise.

In many situations, rank relaxation techniques have emerged as a powerful tool to relax the constraints imposed by distance-based problems and in the same time "smooth" the optimization landscape. By relaxing the rank constraint, these methods open up the possibility of exploring a broader space of configurations, allowing for enhanced accuracy. However, a good understanding of the theoretical aspects surrounding rank relaxation is still lacking.

Distance-based problem could be formulated as a optimization problem on a Riemannian manifold. By considering the Riemannian manifold of possible configurations and the structure of Riemannian optimization problem behind the distance-based problem , we can develop good insights of the problem and its limitation. Additionally, the notion of global rigidity serves as a fundamental concept for investigating the uniqueness of solutions in distance-based problems. Global rigidity provides a mathematical framework to assess the extent to which a configuration can be uniquely determined from pairwise distances. By exploring the conditions under which a structure is globally rigid, we can obtain valuable information about the uniqueness and non-uniqueness of solutions, shedding light on the feasibility of rank relaxation techniques.

In this paper, we aims to provide some basic insights/understanding of the theoretical aspects related to rank relaxation in distance-based problems. By a brief look to the concepts of Riemannian geometry and global rigidity apply to the distance-based distance, we strive to advance our knowledge and and bring together concepts that may seem far apart but are in fact closely linked. After that, we will aims at exploring the behaviour of the rank relaxation through simulations.

In the subsequent sections, we will discuss the fundamental concepts of Riemannian geometry of our problem, elaborate on the notion of global rigidity, and explore how these ideas can be applied to gain insights into the uniqueness of solutions in distance-based problems. We will also highlight potential directions for future research.

2 Distance-based localization problem

Even if the focus of this report will be in the noiseless case, the following sections will present different theoretical aspect of the distance-based problem in the noisy case for generalisation purpose.

2.1 Notations

Let us define some notions before to dive in the subject of the project.

$\mathcal{OB}(d,m) = \{Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_m^T \end{bmatrix} \in \mathbb{R}^{m\times d} : y_i \in \mathbb{S}^{(d-1)},\ i = 1,\cdots,m\}$ is called the oblique manifold, it could be seen as the product of m spheres in $\mathbb{R}^d$. For $Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_m^T \end{bmatrix} \in \mathbb{R}^{m\times d}$, we defined the function normalize : $\mathbb{R}^{m\times d} \to \mathcal{OB}(d,m)$ such that normalize$(Y) = \begin{bmatrix} \frac{y_1^T}{\|y_1\|_2} \\ \vdots \\ \frac{y_m^T}{\|y_m\|_2} \end{bmatrix}$. We could also defined the two operators diag : $\mathbb{R}^{n\times n} \to \mathbb{R}^n$ such that diag$(X) = (X_{11},\cdots,X_{nn})^T \in \mathbb{R}^n$ and diag : $\mathbb{R}^n \to \mathbb{R}^{n\times n}$ such that diag(x) is the diagonal matrix with $x_1,\cdots,x_n$ on its diagonal. For a positive definite matrix W, we could defined the following matrix norm on $\mathbb{R}^{n\times n}$: $\|X\|_W^2 = Tr\left(X^T W X\right)$. Finally, let $\mathcal{G} = (\mathcal{V},\mathcal{E})$

be a graph consisting in n vertices and m edges, we could defined the incidence matrix $C(\mathcal{G}) \in \mathbb{R}^{m\times n}$ such that $C_{ki} = -1$ and $C_{kj} = 1$ for each edge $e_k = (i, j)$ for $k = 1, \cdots, m$.

2.2 Formulation of the problem

Given a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ and a set of measured edge distances $\tilde{d}_{ij}$ for all $(i, j) \in \mathcal{E}$, the distance-based localization problem aims at finding the vertex positions, $x_1, \cdots, x_n \in \mathbb{R}^d$ from which were the estimated the edges distances are as close as possible to the measured edge distances:

Problem 1 (MLE for distance-based localization)**.**

$$p^*_{MLE} = \min_{X\in\mathbb{R}^{d\times n}} \sum_{(i,j)\in\mathcal{E}} \frac{w_{ij}}{2} \left(\|x_i - x_j\|_2 - \tilde{d}_{ij}\right)^2 \tag{1}$$

The solution of Problem 1 is referred as the maximum likelihood estimate (MLE) under the assumptions that the measurement noise arises under the following model:

$$\tilde{d_{ij}} = \|x_j - x_i\|_2 + v_{ij},\ v_{ij} \sim \mathcal{N}\left(0, \sigma^2_{ij}\right) \tag{2}$$

where σ_{ij} is the variance of the noise The weights w_{ij} are proportional to inverse to inverse variance of the corresponding distance measurements.

There is many distanced-only localization methods which address the similar but globally differentiable problem of EDM completion:

Problem 2.

$$p^*_{EDM} = \min_{X\in\mathbb{R}^{d\times n}} \sum_{(i,j)\in\mathcal{E}} \frac{w_{ij}}{2} \left(\|x_i - x_j\|_2^2 - \tilde{d}^2_{ij}\right)^2 \tag{3}$$

Problem 2 can be seen as the MLE under the assumption that the Gaussian noise apply on the squared distances rather than the raw distances as in Eq (2):

$$\tilde{d_{ij}}^2 = \|x_j - x_i\|_2^2 + v_{ij},\ v_{ij} \sim \mathcal{N}\left(0, \sigma^2_{ij}\right) \tag{4}$$

However, this supposition that a measuring instrument measure directly the squared distances is not acceptable in most applications. That why, even if in the noiseless case Problems 1 and 2 have the same objective value, it's remain interesting to address this optimization problem since in general case the solution of both Problems will be different.

In ([1]), Halsted and Schwager proposed a new formulation of Problem 1 by introducing an additional variable $y_{ij} \in \mathcal{S}^{d-1}$ corresponding to the direction of each edge $(i, j) \in \mathcal{E}$:

Problem 3 (MLE over edge directions)**.**

$$p^*_{MLE+} = \min_{\substack{Y\in\mathcal{OB}(d,m)\\ X\in\mathbb{R}^{d\times n}}} \sum_{(i,j)\in\mathcal{E}} \frac{w_{ij}}{2} \left\|x_i - x_j - \tilde{d}_{ij} y_{ij}\right\|_2^2 \tag{5}$$

The inclusion of edge direction as additional variables makes the objective function convex at the price of introducing additional non-convex constraints $y_{ij}^T y_{ij} = 1, \forall (i, j) \in \mathcal{E}$. Furthermore, as shown in [1], Problem 3 is a **tight** relaxation of Problem 1, this means that the positions $X^* \in \mathbb{R}^{d\times n}$ which minimize the objective Eq. (5) also minimize the objective in (1).

It turns out that Problem 3 can be recast as an optimization problem over edge direction only, as stated in the following proposition.

Proposition 2.1. *Problem 3 can be reformulated as an optimization problem over the edge direction, $Y \in \mathcal{OB}(d, m)$, only, as:*

$$p^*_{MLE+} = \min_{Y\in\mathcal{OB}(d,m)} Tr\left(QYY^T\right) = \min_{Y\in\mathcal{OB}(d,m)} f(YY^T) = \min_{Y\in\mathcal{OB}(d,m)} g(Y), \tag{6}$$

where $Q = \tilde{D}\left(W - WC\left(C^TWC\right)^\dagger C^TW\right)\tilde{D}$, $f : \mathcal{C} \to \mathbb{R}$ with $f(Z) = Tr(QZ)$ and $g : \mathcal{OB}(d, m) \to \mathbb{R}$ with $g(Y) = f(YY^T)$.

Proof. Eq. (5) can be rewrite as

$$p^*_{MLE+} = \min_{\substack{Y\in\mathcal{OB}(d,m)\\ X\in\mathbb{R}^{d\times n}}} \left\|CX-\tilde{D}Y\right\|_W^2 \tag{7}$$

where $C = C(\mathcal{G}) \in \mathbb{R}^{m\times n}$ is the incidence matrix of the underlying graph $\mathcal{G}$, $\tilde{D} \in Diag\left(\mathbb{R}^m\right)$ is the diagonal matrix consisting of the measured distances along each edge and $W \in Diag\left(\mathbb{R}^m\right)$ contains the weights of each measurement. But let remark that for a fix $Y \in \mathcal{OB}(d,m)$, we have an unconstrained minimization problem over $X \in \mathbb{R}^{n\times d}$ with the objective:

$$\begin{aligned}\left\|CX-\tilde{D}Y\right\|_W^2 &= Tr\left[\left(CX-\tilde{D}Y\right)^T W\left(CX-\tilde{D}Y\right)\right]\\ &= Tr\left[X^TC^TWCX-\left(X^TC^TW\tilde{D}Y-Y^T\tilde{D}WCX\right)\right]+const(Y) = f(X)\end{aligned} \tag{8}$$

Thus, for a fixed $Y \in \mathcal{OB}(d,m)$ the objective in Eq.8 is a quadratic form in X which attains it's global minimum at X^* when $\nabla_X f(X^*) = 0$ (if $W \succ 0$ then $C^TWC \succeq 0$):

$$\begin{aligned}&\nabla_X f(X^*) = 2C^TWCX^* - 2C^TW\tilde{D}Y\\ \Rightarrow&\nabla_X f(X^*) = 0 \Leftrightarrow (C^TWC)X^* = C^TW\tilde{D}Y\\ \Rightarrow&X^* = (C^TWC)^\dagger C^TW\tilde{D}Y\end{aligned} \tag{9}$$

Thus, replacing this expression of X in (7) give us:

$$\begin{aligned}p^*_{MLE+} &= \min_{Y\in\mathcal{OB}(d,m)} \left\|\left(C(C^TWC)^\dagger C^TW\tilde{D}-\tilde{D}\right)Y\right\|_W^2\\ &= \min_{Y\in\mathcal{OB}(d,m)} Tr\left(Y^TQY\right)\\ &= \min_{Y\in\mathcal{OB}(d,m)} Tr\left(QYY^T\right)\end{aligned} \tag{10}$$

where $Q = \tilde{D}\left(W - WC\left(C^TWC\right)^\dagger C^TW\right)\tilde{D}$ and where the last equality comes from the cyclicity of the trace. □

With this new formulation of Problem 3, the following proposition is straightforward:

Proposition 2.2. *$Y^* \in \mathcal{OB}(d,m)$ is a global minimum of Problem 3 if and only if $Y^*Q \in \mathcal{OB}(d,m)$ is a global minimum of Problem 3 for any $Q \in \mathcal{O}(d)$.*

Our goal in this project, will be to study the following relaxation of Problem 3, in which we optimize over r-dimensional unit vector for some integer $r \geq d$:

Problem 4.

$$p^*_r = \min_{Y_r\in\mathcal{OB}(r,m)} Tr\left(QY_rY_r^T\right) = \min_{Y_r\in\mathcal{OB}(r,m)} g(Y) \tag{11}$$

with g as in Proposition 2.1.

Any p^*_{MLE} must also be feasible for this relaxed problem, meaning that $p^*_r \leq p^*_{MLE}$ for all $r \geq d$, and that $p^*_{r_2} \leq p^*_{r_1}$ for any $r_1 < r_2$. In the noiseless case, the objective value for each relaxation will be 0 since the graph is embedded in $\mathbb{R}^d$ and also realise the distance $\tilde{D}$ when there is no noise.
In the extreme case of $r = m$, Problem 4 is even equivalent to the semi-definite positive (SDP) problem:

Problem 5.

$$p^*_{SDP} = \min_{Z\in\mathcal{C}} Tr\left(QZ\right) = \min_{Z\in\mathcal{C}} f(Z) \tag{12}$$

with f as in Proposition 2.1.

Algorithm 1: Projection from $\mathcal{OB}(r,m)$ to $\mathcal{OB}(d,m)$

Input: $Y_r \in \mathcal{OB}(r,m)$

1 Compute the SVD decomposition of Y_r, $Y_r = U\Sigma V^T$, with $\Sigma = diag(\sigma_1, \cdots, \sigma_r)$, $\sigma_1 \geq \cdots \geq \sigma_r \geq 0$.

2 Set $\tilde{Y}_d = U \begin{bmatrix} diag(\sigma_1, \cdots, \sigma_d) \\ \mathbf{O}_{(r-d)\times d} \end{bmatrix}$

3 Project $\tilde{Y}_d$ on $\mathcal{OB}(d,m)$ by normalizing each row: $Y_d = normlaize(\tilde{Y}_d)$.

The relaxed problems 4 and 5 will, respectively, give solutions Y_r^* and Y_m^* (s.t. $Z^* = Y_m^*(Y_m^*)^T$) embedded in $\mathbb{R}^r$ and $\mathbb{R}^m$. But, as the original problem 3 is on $\mathcal{OB}(d,m)$, the solution need to be project back to the original space. For that, one can follow this algorithm:

The goal of the relaxation of the original problem 3 is that, it has often been observed that increasing the dimension of the search space often allows to prevent convergence to spurious local minima, and in the extreme case where $r = m$ Problem 5 even becomes convex and thus the optimization problem is not anymore subject to spurious local minima. However, this comes at two prices, firstly the increase of the dimension of the search space may considerably slow down any optimization algorithm and secondly , even if we obtain a optimal solution in $\mathcal{OB}(r,m)$, we don't have any guaranty that the solution obtain after projection in the original search space will still be a optimal solution. Even if our experiments will tend to show that the solution obtained after projection remain often a optimal solution of Problem 3 or at least a good "first guess" for a initialization point, we later in Section **??** that relaxing to much the dimension may create new global optimal solution which does not live in a d-dimensional subspace anymore.

It's obvious that in the noiseless case, there always exists a low-rank solution embedded in $\mathbb{R}^d$, but in the general case it will often not be the case since the noise make that the graph $\mathcal{G}$ and the distances $\tilde{D}$ are certainly not realisable in the original space $\mathbb{R}^d$. However, it as been show ([2]) that if a graph is realizable in some dimension than it must be realisable in dimension $r = \left\lfloor \frac{-1+\sqrt{8m+1}}{2} \right\rfloor$:

Theorem 2.1 (Existence of a low-rank solution)**.** *Problem 5 is a tight relaxation of Problem 4 (i.e. $p^*_{SDP} = p^*_r$) for all r satisfying*

$$r \geq \left\lfloor \frac{-1+\sqrt{8m+1}}{2} \right\rfloor. \tag{13}$$

2.3 Theoretical consideration of the problems

This section will be based on the article [3] in which N. Boumal address a more general problem, we gonna to adapt some of it's result to our problems in order to get a better comprehension and intuition of our problems. It will first focus on the geometrical characterization of the problems 4 and 5 to gain a good basic understanding of the geometry of each space and then state the different optimization tools which will be needed in the second part to understand the characterization of optimal points in the different problems and the potentials links between them.

2.3.1 Geometry of the problems

The ambient space $\mathbb{R}^{m\times r}$ is endowed with the classical Euclidean metric $\langle U, V\rangle = Tr\left(U^T V\right)$, which correspond to the Frobenius norm. The search space of Problem 4 is see as a submanifold of $\mathbb{R}^{m\times r}$. As a remainder the oblique manifold can be simply express via the following equality constraint:

$$\mathcal{OB}(r,m) \left\{ Y \in \mathbb{R}^{m\times r} : \mathrm{diag}(YY^T) = \mathbf{1} \right\}. \tag{14}$$

The space is always connected since it can be seen as the product of m sphere $\mathbb{S}^{r-1}$. Furthermore, the tangent space is easily obtained with the differential of the equality constraint:

$$T_Y\mathcal{OB}(r,m) = \left\{ \dot{Y} \in \mathbb{R}^{m\times r} : \mathrm{diag}\left(\dot{Y}Y^T + Y\dot{Y}^T\right) = \mathbf{0} \right\} \tag{15}$$

Each of this tangent space is equipped with the restriction of the ambient metric, making of $\mathcal{OB}(r,m)$ a Riemannian submanifold of $\mathcal{E}$. The orthogonal projector from the embedding space $\mathbb{R}^{m\times r}$ to the tangent space at Y is given by:

$$\mathrm{Proj}_Y(Z) = Z - \mathrm{diag}(ZY^T)Y \tag{16}$$

And one can define the natural retraction from the tangent space to the manifold which is just the retraction of the sphere manifold apply to each row of $Y \in \mathcal{OB}(r,m)$:

$$\mathrm{Retraction}_Y(\dot{Y}) = \mathrm{normalize}(Y + \dot{Y}). \tag{17}$$

Let $\nabla g(Y) = \nabla\left(f(YY^T)\right) = 2\left(\nabla f(YY^T)\right)Y = 2QY$ be the classical Euclidean gradient of g defined in (2.1), its Riemannian gradient is then naturally given by the projection of the classical gradient to the tangent space:

$$\mathrm{grad}\ g(Y) = \mathrm{Proj}_Y\left(\nabla g(Y)\right) = 2\mathrm{Proj}_Y\left(QY\right). \tag{18}$$

Furthermore, if we denote by $\nabla^2 g(Y)$ the classical Hessian of g at Y, which is such that $\nabla^2 g(Y)\left[\dot{Y}\right] = 2Q\dot{Y}$ for $\dot{Y} \in T_Y\mathcal{OB}(r,m)$. The Riemannian Hessian of g at Y is then the symmetric operator on the tangent space $T_Y\mathcal{OB}(r,m)$ obtained as the projection of the derivative of the gradient vector field:

$$\begin{aligned}\mathrm{Hess}\ g(Y)\left[\dot{Y}\right] &= \mathrm{Proj}_Y\left(\nabla^2 g(Y)\left[\dot{Y}\right] - \mathrm{diag}\left(\nabla g(Y)Y^T\right)\dot{Y}\right)\\ &= 2\mathrm{Proj}_Y\left(Q\dot{Y} - \mathrm{diag}\left(QYY^T\right)\dot{Y}\right)\end{aligned} \tag{19}$$

For his part $\mathcal{C}$ is a compact and convex set, which could be seen as the set of correlation matrices. Moreover, there is some useful way of see the space $\mathcal{C}$, indeed it can be show that the relative interior point of the non-empty faces of $\mathcal{C}$ forms partition of $\mathcal{C}$.

Definition 2.1 (Faces)**.** *A face of $\mathcal{C}$ is a convex subset $\mathcal{F}$ of $\mathcal{C}$ such that every closed line segment in $\mathcal{C}$with a relative point in $\mathcal{F}$ has both end points in $\mathcal{F}$.*

Furthermore, the special case where $\dim(\mathcal{F}) = 0$ is of special interest and are called extreme points.

Definition 2.2 (Extreme points)**.** *$X \in \mathcal{C}$ is an extreme point of $\mathcal{C}$ if there does not exist $X', X'' \in \mathcal{C}\setminus\{X\}$ and $0 < \lambda < 1$ such that $X = \lambda X' + (1-\lambda)X''$.*

Extreme points are of special interest because they often arise as the solution of optimization problems, especially when f is linear, which is our case, since f attains its minimum on $\mathcal{C}$ at one of its extreme points.

2.3.2 Characterization of optimality in Problem 3, 4 and 5

Lemma 2.1 (Necessary optimality conditions for Problem 5)**.** *$X \in \mathcal{C}$ is called a KKT point for Problem 5 if there exist a symmetric matrix $\hat{S} \in \mathbb{S}^{m\times m}$ and diagonal matrix $\hat{\Lambda} \in \mathbb{S}^{m\times m}$ such that*

$$\hat{S}X = 0, \quad \hat{S} = \nabla f(X) + \hat{\lambda} \quad \textit{and} \quad \hat{S} \succeq 0. \tag{20}$$

If X is a local optimizer for Problem 5, then X is a KKT point. Moreover, as f is convex, all KKT points are global optimizer.

Lemma 2.2 (Necessary optimality conditions for Problem 4)**.** *Let $Y \in \mathcal{OB}(r,m)$ and $X = YY^T$. A critical point of Problem 4 satisfies grad $g(Y) = 0$, that is,*

$$\left(\nabla f(X) - diag\left(\nabla f(X)X\right)\right)Y = 0. \tag{21}$$

A second-order critical point is a critical point which satisfies Hess $g(Y) \succeq 0$, that is, for all $\dot{Y} \in T_Y\mathcal{OB}(r,m)$

$$\left\langle \dot{Y}, \nabla^2 g(Y)\left[\dot{Y}\right] - diag\left(\nabla g(Y)Y^T\right)\dot{Y}\right\rangle \geq 0. \tag{22}$$

If Y a local optimizer for Problem 4, then it is a second-order critical point.

Theorem 2.2 (S is the right certificate)**.** *$X \in \mathcal{C}$ is a KKT point for Problem 5 if and only if $S(X) = \nabla f(X) - diag(\nabla f(X)X)$ is positive semi-definite. If so, $S(X)$ is the unique dual certificate for Lemma 2.1.*

Theorem 2.3. *If $X \in \mathcal{C}$ is an extreme point for Problem 5, and $S(X) \succeq 0$, and $rank(X) + rank(S(X)) = m$ (strict complementarity), then X is the unique global optimizer of Problem 5.*

Corollary 2.1 (Global optimality conditions)**.** *Let $Y \in \mathcal{OB}(r, m)$, $X = YY^T$ is globally optimal for Problem 4 if and only if $S(X) \succeq 0$. If so, then Y is a global optimizer for the nonconvex problem 5. Furthermore, if X is extreme and if $rank(X) + rank(S(X)) = m$, then X is the unique global optimizer of Problem 5 and Y is the unique global optimizer of Problem 4, up to orthogonal action YQ for $Q \in \mathcal{O}(r)$.*

The two followings results give a interesting comprehension of how to escape local minima of Problem 3 and 4. Based [3] proposed a interesting algorithm which allows to escape local minima by increasing incrementally the search space until convergence, since the procedure will must converge at least when $r = m$, $\mathcal{C}$ being convex.

Theorem 2.4 (Escape direction for rank-deficient critical point)**.** *Let Y be a rank-deficient critical point for Problem 4 such that $X = YY^T$ is not a KKT point of Problem 5. Then, for all nonzero vectors $z \in \mathbb{R}^r$ and $u \in \mathbb{R}^m$ such that $Yz = 0$ and $u^T S(X)u < 0$, $\dot{Y} = uz^T \in T_Y\mathcal{OB}(r, m)$ is a descent direction for g from Y.*

Corollary 2.2 (Escape direction in the full rank case)**.** *Let Y be a full-rank critical point for Problem 4 such that $X = YY^T$ is not a KKT point of Problem 5. Let $Y_+ = \begin{pmatrix} Y & O_{n\times(p_+-p)} \end{pmatrix} \in \mathcal{OB}(r, m)$. Then,*

- *(a) $g(Y_+) = g(Y)$;*
- *(b) Y_+ is a critical point for Problem 4 for $r = p_+$;*
- *(c) for all $u \in \mathbb{R}^n$ such that $u^T S(X)u < 0$, $\dot{Y} = ue_{r+1}^T \in T_{Y_+}\mathcal{OB}(r, m)$ is a descent direction for g from Y_+, where $e_{r+1} \in \mathbb{R}^{p_+}$ is a zero vector except for its $(r + 1)^{st}$ entry, equal to 1.*

3 Graph Rigidity

The landscape of the optimization problem 4 can be quite wild and have numerous spurious local minima, even in the noiseless case. As we place ourselves in the noiseless case, we know that there exists at least one graph embedded in $\mathbb{R}^d$ which match the measured distances $\tilde{D}$ and then have a zero objective value in Problem 3 and thus, by embedding of $\mathbb{R}^d$ in $\mathbb{R}^r$, it must hold also for Problem 4. However, we would be able to understand or at least have a good intuition under which conditions on the graph $\mathcal{G}$ Problem 3 and his relaxation Problem 4 have a unique optimal solution. And, when it will not be the case, how could we obtain a new optimal solution from the old one.

Unfortunately, there are no easy answers to these questions. We will therefore first look at the notion of graph rigidity, which will enable us to provide an answer to our questions but which will be difficult to use in our simulations. We will then present necessary but not sufficient conditions for guaranteeing uniqueness, as well as methods for obtaining new solutions in certain cases.

One can already observe that if $X^* \in \mathbb{R}^{n\times d}$ is a optimal solution of Problem 1, if $R : \mathbb{R}^{m\times d} \to \mathbb{R}^{m\times d}$ is a rigid transform of the Euclidean space $\mathbb{R}^{m\times d}$, i.e. a rotation, symmetry or translations than $R(X)$ is also a optimal solution of Problem 1. Analogously, as show in 2.2, if Y^* is a optimal solution for Problem 3 or 4, then $R(Y)$ will also be a optimal solution for R a rotation or a symmetry. Thus, we have to keep in mind that the solutions found will always be up to the action on the right of $\mathcal{O}(r)$.

Corollary 2.1 give a way to check if a global optimizer, $Y \in \mathcal{OB}(r, m)$ is the unique global optimizer of Problem 4, up to the action of $\mathcal{O}(r)$, by checking if $X = YY^T$ is a extreme point of $\mathcal{C}$, if $S(X) \succeq 0$ and if $\text{rank}(X) + \text{rank}(S(X)) = m$. However, we must already have a solution to use this result to check if the global optimizer is unique or not. Even it's could still be useful in the case where we generate a graph and we want to check if there exist another graph which match its structure and the distances $\tilde{D}$, it's in practice difficult to apply because of the numerical error which arise when we try to compute the rank of X and $S(X)$.

3.1 Uniqueness of the solution

To try to get better understanding of this question, this section will be placed in the context of a fundamental problem in distance geometry, the Graph Realization Problem (GRP), which aims is to determine when the distance between certain pairs of vertices of a finite configuration in Euclidean space $\mathbb{E}^d$ determine it up to rigid motion. To make this more precise, let define some useful notions.

Definition 3.1 (Configuration)**.** *A configuration is a finite collection of n labeled points,* $\mathbf{p} = (p_1, \cdots, p_n)$, $p_i \in \mathbb{E}^d$ *for* $1 \leq i \leq n$.

Definition 3.2 (Framework)**.** *A framework in* $\mathbb{E}^d$ *is a graph* $\mathcal{G}$ *with n vertices together with a corresponding configuration* $\mathbf{p} = (p_1, \cdots, p_n)$ *in* $\mathbb{E}^d$*, and is denoted by* $\mathcal{G}(\mathbf{p})$.

Definition 3.3 (Equivalent framework and congruent configuration)**.** *Let consider the two following definitions:*

- *Two framework* $\mathcal{G}(\mathbf{p})$ *and* $\mathcal{G}(\mathbf{q})$ *are equivalent,* $\mathcal{G}(\mathbf{p}) \equiv \mathcal{G}(\mathbf{q})$*, if when* $\{i, j\}$ *forms an edge of* $\mathcal{G}$*, then* $\|p_i - p_j\| = \|q_i - q_j\|$.
- *A configuration* $\mathbf{p}$ *is congruent to* $\mathbf{q}$*,* $\mathbf{p} \equiv \mathbf{q}$*, if for* $\{i, j\} \in \{1, \cdots, n\}$ *we have* $\|p_i - p_j\| = \|q_i - q_j\|$.

Definition 3.4 (Globally rigid framework)**.** *A framework* $\mathcal{G}(\mathbf{p})$ *is called globally rigid in* $\mathbb{E}^d$ *if* $\mathcal{G}(\mathbf{p}) \equiv \mathcal{G}(\mathbf{p})$ *implies* $\mathbf{p} \equiv \mathbf{q}$.

Definition 3.5 (Rigid framework)**.** *A framework* $\mathcal{G}(\mathbf{p})$ *in* $\mathbb{E}^d$ *is said to be rigid if there is an* $\epsilon > 0$ *such that for any other configuration* $\mathbf{q}$ *in* $\mathbb{E}^d$*, where* $\|\mathbf{p} - \mathbf{q}\| < \epsilon$ *and* $\mathcal{G}(\mathbf{p}) \equiv \mathcal{G}(\mathbf{q})$*, then* $\mathbf{p} \equiv \mathbf{q}$.

Definition 3.6. *(Vertex k-connected) A connected graph G is said to be vertex k-connected (or k-connected) if it has more than k vertices and remains connected whenever fewer than k vertices are removed.*

However, even rigid graphs can have multiple realizations, indeed there is some discontinuous transformations which allows to transform a realization into a another. A situation in which this can happen is when there is some vertices about which a portion of the graph can be reflected. These vertices form a kind of "mirror". For that, there must be no edges between the two halves of the graph separated by this mirror. For our problem in $\mathbb{E}^d$, the mirror vertices must lie in a $(d-1)-$dimensional subspace. We will say that a framework in a d-dimensional space allows a *partial reflection* if a separating set of vertices lies in a $(d-1)-$dimensional subspace. This bring us to the following result of [4]:

Theorem 3.1. *A rigid graph positioned generically in* $\mathbb{E}^d$ *will have a partial reflection if and only if it is not vertex* $(d+1)-$*connected.*

What interests us now is, for a given framework $\mathcal{G}(\mathbf{p})$, whether it is globally rigid or not. Nevertheless, it has been show that this problem is strongly NP-hard, even in $\mathbb{E}^1$. However, in [5], Robert Connelly show that there is an algebraic set of configurations such that when the configuration $\mathbf{p}$ is outside that set, $\mathcal{G}(\mathbf{p})$ is globally rigid in $\mathbb{E}^2$. So, we are led to consider the question of whether "most" configurations $\mathbf{p}$ for a given graph $\mathcal{G}$ are globally rigid.

Definition 3.7 (Algebraic dependent and generic configuration)**.** *A set* $(\alpha_1, \cdots, \alpha_m)$ *of distinct real numbers is said to be algebraically dependent if there exist a non-zero polynomial* $f(x_1, \cdots, x_m)$ *with integer coefficients such that* $h(\alpha_1, \cdots, \alpha_m) = 0$*. If the set* $(\alpha_1, \cdots, \alpha_m)$ *is not algebraically dependent, it is called generic.*
If a configuration $\mathbf{p} = (p_1, \cdots, p_n)$ *in* $\mathbb{E}$ *is such that its dn coordinates are generic, we say* $\mathbf{p}$ *is generic.*

To handle a more tractable problem, we then focus on this new notion of global rigidity.

Definition 3.8 (Generic global rigidity)**.** *For a given graph* $\mathcal{G}$*, when* $\mathcal{G}(\mathbf{p})$ *is globally rigid for all generic configurations* $\mathbf{p}$ *in* $\mathbb{E}$*, we say that* $\mathcal{G}$ *is generically globally rigid in* $\mathbb{E}^d$.

Definition 3.9 (Vertex m-connected and redundantly rigid)**.** *A graph is vertex m-connected if it takes the removal of at least m vertices of $\mathcal{G}$ to make the graph disconnected.*
A framework $\mathcal{G}(\mathbf{p})$ is said to be redundantly rigid if $\mathcal{G}(\mathbf{p})$ is rigid in $\mathbb{E}^d$ even after the removal of any edge of $\mathcal{G}$.

All of this leads initially to the following result of [4]:

Theorem 3.2. *Let $\mathcal{G}(\mathbf{p})$ be a framework in $\mathbb{E}^d$ such that the configuration $\mathbf{p}$ is generic, and $\mathcal{G}(\mathbf{p})$ is globally rigid with at least $d+1$ vertices. Then the following conditions must hold:*

- *(i) The graph $\mathcal{G}$ is vertex $(d+1)-$connected.*
- *(ii) The framework $\mathcal{G}(\mathbf{p})$is redundantly rigid in $\mathbb{E}^d$.*

One could remark that condition (i) is straightforward by Theorem 3.1. However, for $d \geq 3$ it is shown that the conjecture, that conditions (i) and (ii) are sufficient for generic global rigidity, is false. Furthermore, for $d \geq 3$ it is still not known whether global rigidity is a generic property, which is the case for $d \in \{1, 2\}$. On the other hand, it is known that rigidity is a generic property, and thus rigidity in $\mathbb{E}^d$ is entirely a combinatorial property of the graph $\mathcal{G}$.

In order to state the main result of [5], which will give a theoretical way to check that a framework $\mathcal{G}(\mathbf{p})$ is globally rigid in $\mathbb{E}^d$, we only need some new notions.

Definition 3.10 (Equilibrium stress vector and stress matrix)**.** *Let $\omega = (\cdots, \omega_{ij}, \cdots)$, for all $\{i, j\} \in \{1, \cdots, n\}$, be a set of scalar such that $\omega_{ij} = \omega_{ji}$, $\omega_{ij} = 0$ if $\{i, j\}$ is not an edge of $\mathcal{G}$. We say that ω is an equilibrium stress for the framework $\mathcal{G}(\mathbf{p})$ if, for each vertex i of $\mathcal{G}$, the following equation holds:*

$$\sum_j \omega_{ij}(p_i - p_j) = 0. \tag{23}$$

To each stress vector ω for a graph $\mathcal{G}$ on n vertices, there is an n-by-n symmetric matrix Ω, the associated stress matrix, such that for $i \neq j$, $\{i, j\} \subset \{1, \cdots, n\}$, we have $\Omega_{ij} = -\omega_{ij}$, and the diagonal entries are such that the row and the columns sums of the entries of Ω are zero.

Theorem 3.3. *Suppose that $\mathbf{p} = (p_1, \cdots, n)$, $n \geq d+2$, is a generic configuration in $\mathbb{E}^d$ such that there is an equilibrium stress for a framework $\mathcal{G}(\mathbf{p})$, where the rank of the associated stress matrix Ω is $n-d-1$. Then $\mathcal{G}(\mathbf{p})$ is globally rigid in $\mathbb{E}^d$.*

Let us remark that this theorem give us a way to numerically check if a framework is globally rigid or not. Indeed, if we have a generic configuration $\mathbf{p}$, it is possible to solve the equilibrium equation (23) of an appropriate equilibrium stress ω, and then calculate the rank of Ω. If the rank is equal to $n-d-1$, than by Theorem 3.3 we can be assured that $\mathcal{G}(\mathbf{p})$ is globally rigid in $\mathbb{E}^d$. Furthermore, one could ask if the converse holds, and it indeed the case as shown in [6]:

Theorem 3.4. *If a graph $\mathcal{G}$ with at least $d+2$ vertices does not have a stress matrix Ω of rank $n-d-1$ in $\mathbb{E}^d$, then any generic framework $\mathbf{p}$ is not globally rigid.*

3.2 Generation of congruent framework in the non globally rigid case

Given a framework $\mathcal{G}(\mathbf{p})$, even if Theorem 3.1 give us a way to obtain a new congruent framework $\mathcal{G}(\mathbf{q})$ in the case where $\mathcal{G}(\mathbf{p})$ is not vertex $(d+1)-$connected with a partial reflection of the framework $\mathcal{G}(\mathbf{p})$ around the "mirror" set of points, it's remain not the easier way to find congruent framework since the vertex $k-$connectivity is not so easy to check. But, we nonetheless mention that in [4], there is different algorithm which are mentioned to solve this problem of $k-$connectivity.

We instead propose another specific case, the non rigid case, where we can easily obtain congruent framework of a framework $\mathcal{G}(\mathbf{p})$ in $\mathbb{R}^d$. Suppose that a vertex v_i of $\mathcal{G}$ associated with the position $p_i \in \mathbb{R}^d$ has a degree of m and his neighbours are $v_{i_1}, \cdots, v_{i_m}$ have the position $p_{i_1}, \cdots, p_{i_m} \in \mathbb{R}^d$. If the affine hull of these m points has dimension $k \in \mathbb{N}$ and if $k \leq d$, than we have that there is two case:

- either p_i is the only points in $\mathbb{R}^d$ which match the distances $d_j = \|p_{i_j} - p_i\|_2$ for all $j = 1, \cdots, m$,

- either there is a $(n-k)$-sphere of points which match the distances $d_j = \|p_{i_j} - p_i\|_2$ for all $j = 1, \cdots, m$.

Indeed, this comes from the fact that for fixed $p_{i_1}, \cdots, p_{i_m}$, the only constraints that are on a potential position $\tilde{p}_i$ of the vertex v_i is to respect the distances $\|p_{i_j} - \tilde{p}_i\|_2 = \|p_{i_j} - p_i\|_2$ for all $j = 1, \cdots, m$. But, this is in fact equivalent to saying that the point $\tilde{p}_i$ lie on the intersection of the m spheres of centers $p_{i_1}, \cdots, p_{i_m}$ and respective radius $d_j = \|p_{i_j} - p_i\|_2$ for all $j = 1, \cdots, m$.

To show that we will then study the intersection of m sphere in $\mathbb{R}^d$. Let $a_1, \cdots, a_m$ be the centres of m spheres in $\mathbb{R}^d$, and $d_1, \cdots, d_m$ be their respective radius. Then, the points x at the intersection are given by the following equations:

$$\|x - a_j\|_2^2 = d_j^2, \ j = 1, \cdots, m. \tag{24}$$

Theorem 3.5 (Spheres intersection)**.** *Let $a_1, \cdots, a_m$ be the centres of m spheres in $\mathbb{R}^d$, and $d_1, \cdots, d_m$ be their respective radius. If the affine hull of these m centres has dimension $k \in \mathbb{N}$, then the possibilities for the sphere intersection are:*

- *(1) the empty set;*
- *(2) a single point;*
- *(3) an $(n-k)-$sphere.*

Proof. Let $\hat{A}$ denote the $d \times (m-1)$ matrix of shifted centres by a_m, that is,

$$\hat{A} = [a_1 - a_m, \cdots, a_{m-1} - a_m]. \tag{25}$$

This matrix $\hat{A}$ has rank k, since the affine hull of these m centres has dimension k, and thus $\hat{A}$ has k linearly independent columns, where $k \leq \min\{d, m-1\}$. If a point x is in the intersection of the spheres, then $\bar{x} = x - a_m$ is also at the intersection of the translated spheres. Then,

$$\|\bar{x}\|_2^2 = d_m^2 \tag{26}$$

and

$$\|\bar{x} - (a_i - a_m)\|_2^2 = d_i^2, \ i = 1, \cdots, m-1 \tag{27}$$

or equivalently

$$\|\bar{x}\|_2^2 - 2\bar{x}^T(a_i - a_m) + \|a_i - a_m\|_2^2 = d_i^2, \ i = 1, \cdots, m-1. \tag{28}$$

From Eq.26,

$$\begin{aligned} &d_m^2 - 2\bar{x}^T(a_i - a_m) + \|a_i - a_m\|_2^2 = d_i^2, \ i = 1, \cdots, m-1 \\ &(a_i - a_m)^T\bar{x} = -\frac{1}{2}\left(d_i^2 - d_m^2 - \|a_i - a_m\|_2^2\right), \ i = 1, \cdots, m-1. \end{aligned} \tag{29}$$

Thus, in matrix form, Eq.29 are given by $\hat{A}^T\bar{x} = c$, where $\frac{-1}{2}\left(d_i^2 - d_m^2 - \|a_i - a_m\|_2^2\right), i = 1, \cdots, m-1$. Now, consider the QR decomposition of $\hat{A}$:

$$\hat{A} = QR = Q\begin{bmatrix}\hat{R} \\ 0\end{bmatrix}, \tag{30}$$

where Q is an $d \times d$ orthogonal matrix and R is an $d \times (m-1)$ matrix, with the last $n-k$ rows null. Also, R has the same rank as $\hat{A}$. We obtain

$$\begin{bmatrix}\hat{R}^T & 0\end{bmatrix} Q^T\bar{x} = c \tag{31}$$

let write $Q^T\bar{x} = \begin{bmatrix}y \\ z\end{bmatrix}$, where $y \in \mathbb{R}^k$ and $z \in \mathbb{R}^{d-k}$, implying

$$\begin{bmatrix}\hat{R}^T & 0\end{bmatrix}\begin{bmatrix}y \\ z\end{bmatrix} = c \Leftrightarrow \hat{R}^T y = c. \tag{32}$$

On the other hand,

$$d_m^2 = \|\bar{x}\|_2^2 = \left\| Q \begin{bmatrix} y \\ z \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} y \\ z \end{bmatrix} \right\|_2^2 = \|y\|_2^2 + \|z\|_2^2, \tag{33}$$

which implies that

$$\|z\|_2^2 m = d_m^2 - \|y\|_2^2. \tag{34}$$

The intersection of sphere will be nonempty iff both the linear system (32) and equation (34) are consistent. Once the values of y and z are determined, the points in the intersection are given by

$$x = Q \begin{bmatrix} y \\ z \end{bmatrix} + a_m. \tag{35}$$

Since the linear system (32) is over-determined and $\hat{R}^T$ has full rank, either this system has none or a unique solution. If the system (32) is inconsistent, then the intersection of m spheres is empty. Whenever this linear system has a unique solution, Eq. (34) implies the following cases.

- If $d_m^2 < \|y\|_2$, then the intersection is empty.
- If $d_m^2 = \|y\|_2$, then there is a single point in the intersection.
- If $d_m^2 > \|y\|_2$, then the intersection is a $(d-k)$-sphere.

In the last case, as $z \in \mathbb{R}^{n-k}$ if $d_m^2 > \|y\|_2$, we have that all the points of an $(n-k)-$sphere satisfy equation (34). So, when we apply the rotation and translation given by equation (35), the solution set remains an $(n-k)-$sphere. □

The proof of Theorem (3.5) give a direct procedure to compute spheres intersections and thus a new point $\tilde{p}_i$ for which the framework $\mathcal{G}(\tilde{\mathbf{p}})$, $\tilde{\mathbf{p}} = (p_1, \cdots, \tilde{p}_i, \cdots, p_m)$, is equivalent to $\mathcal{G}(\mathbf{p})$.

Algorithm 2: Sphere intersection

Input: The centres $a_1, \cdots, a_m \in \mathbb{R}^d$ and the radius $d_1, \cdots, d_m \in \mathbb{R}_+$ of the m spheres.

Output: If the intersection is empty, the algorithm informs that. Otherwise, we have two cases:

- (a) the intersection is a point: return the solution x;
- (b) the intersection is an $(n-k)-$sphere: return y and Q.

Define the matrix $\hat{A} = [a_1 - a_m, \cdots, a_{m-1} - a_m]$;
Compute $\hat{A} = QR$ and set $k = \text{rank}(\hat{A})$;
Compute the vector $c \in \mathbb{R}^{m-1}$ defined by $c_i = -\frac{1}{2}(d_i^2 - d_m^2 - \|a_i - a_m\|_2^2)$;
Obtain the matrix $\hat{R}$ by removing the $n-k$ last rows of R; **if** $c \notin Range(\hat{R})$ **then**
    Stop, the intersection is empty;
Solve the system $\hat{R}^T y = c$;
if $d_m^2 - \|y\|_2^2 < 0$ **then**
    Stop, the intersection is empty;
if $d_m^2 - \|y\|_2^2 = 0$ **then**
    Set $z = 0$, and return $x = Q \begin{bmatrix} y \\ z \end{bmatrix} + a_m$;
if $d_m^2 - \|y\|_2^2 > 0$ **then**
    Return Q and y. Any point x in the intersection canbe obtained by $x = Q \begin{bmatrix} y \\ z \end{bmatrix} + a_m$, where z is any vector satisfying $\|z\|_2^2 = d_m^2 - \|y\|_2^2$.

4 Experiments

In this section, we gonna to try to use the comprehension/intuition on the Problem 3 that we have build in the previous sections to explore the effect of relaxation on the convergence to optimal solution in the noiseless case. The following experiments use the Riemannian geometry of Problem 3 and use the Manopt ([7]) Toolbox on MATLAB, which is design to handle Riemannian optimization problems. More specifically, we gonna to the Riemannian Trust Region algorithm to handle our optimization problem.

Before to begin any simulations, let define some structures of graphs that we gonna to use along this section:

- **Local structure**: For $m \in \mathbb{N}$ given, we will say that a graph have a local structure with the parameter m if we created this graph by connecting each vertex to its m nearest neighbours.
- **Non-local structure**: We will say that a graph have a non-local structure if he have a local structure with parameter m and we add a extra edge between each vertex and its furthest neighbour. Thus, each vertex have at least a degree of $m + 1$.
- **Circular graph in $\mathbb{R}^2$**: We create equally spaced points on the circle and connect them with edges following the local structure policy.
- **(Non-)Local structure on the ball in $\mathbb{R}^d$**: We generate n i.i.d points from a Gaussian distribution and the edges are created according to the (non-)local structure definition.
- **(Non-)Local structure on the sphere in $\mathbb{R}^d$**: We generate n i.i.d points from a Gaussian distribution and then we normalize each of them to obtain points on the unit sphere. The edges are created according to the (non-)local structure definition.
- **(Non-)Local structure on the Torus in $\mathbb{R}^3$**: We generate points uniformly on the Torus and then generate the edges according to the (non-)local structure policy.
- **Random graph** We generate n uniformly on the unit ball and then we create the edges completely at random with a given probability p, following the Erdós-Rényi model.

4.1 Pathological cases of local minima and relaxation improvement

Before to really diving into the study of the the effect of the relaxation, we present different situations in $\mathbb{R}^2$ for which Problem 3 could be quite wild and have numerous spurious local minima. And, we will begin to see what can be achieved by relaxing the Problem 3.

Figure (1) show the solution obtained for circular graph of reasonable size $n \in \{20, 50\}$ without relaxation, but one can already observe that the solutions obtained are quite far from the ground truth for the graphs (A), (C) and (D). For this ones the solution obtained looks like to the ground truth graph that has been twisted and folded in different directions. The graph (B) is for its parts quite close (at least visually) from the original one, but there is a vertex which has somehow been mirrored with respect to the segment generated by these two nearest neighbours which leaves it at a local minimum.

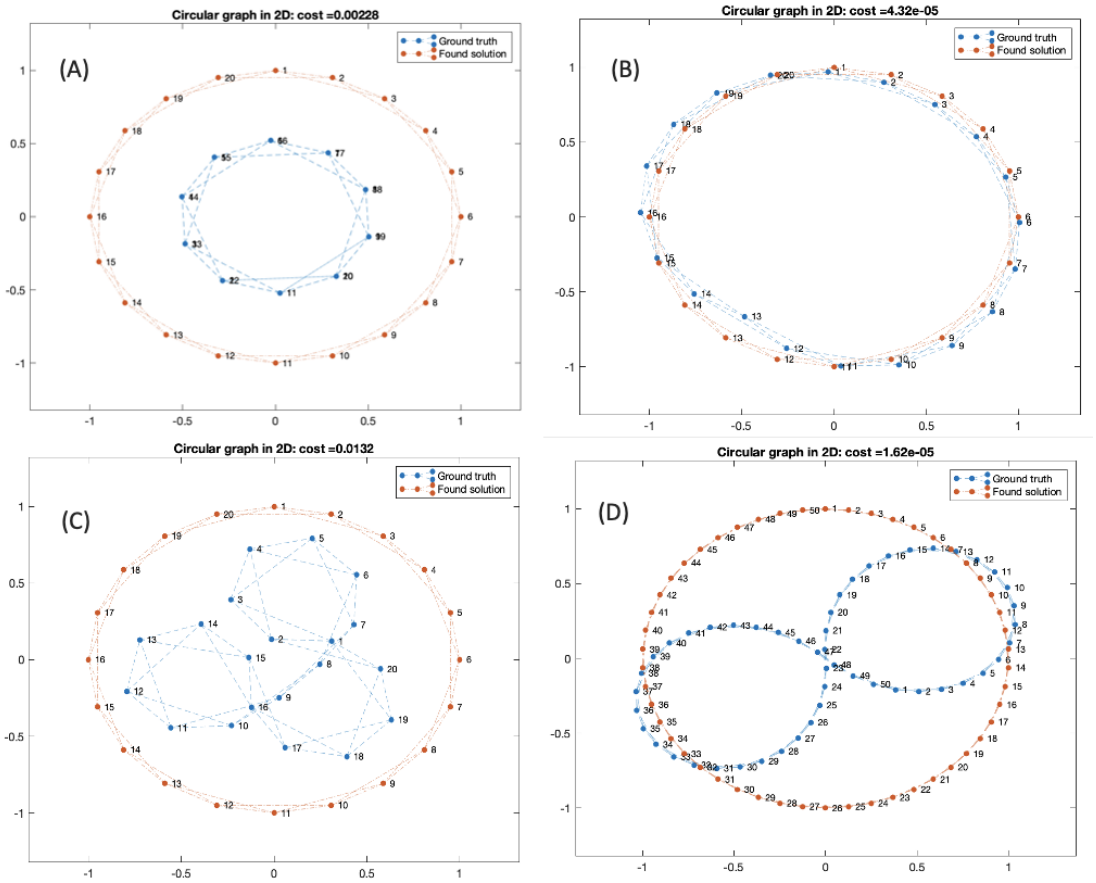

Figure 1: These figures are the ground truth and the solution found by the RTR algorithm for circular graph in $\mathbb{R}^2$; $n = 20,\ m = 4$ for (A), (B) and (C); $n = 50,\ m = 4$ for (D).

The set up of Figure (2) is similar to Fig. 1 but with more vertex and twice more edges. The graph (A) show us a solution which seems to be really close to the ground truth, however even if its cost is quite small $\approx 10^{-7}$ its still to big to be considered as a global optimum. What we can learn from this is that when the number of vertices increases and the number of edges remains low on this type of circular graph, it has a numerous local minima that are relatively close to the optimal solution. If we increase again the number of vertex, we observe the same kind of folding of the graph observed in Fig.(1), but with smaller associated cost.

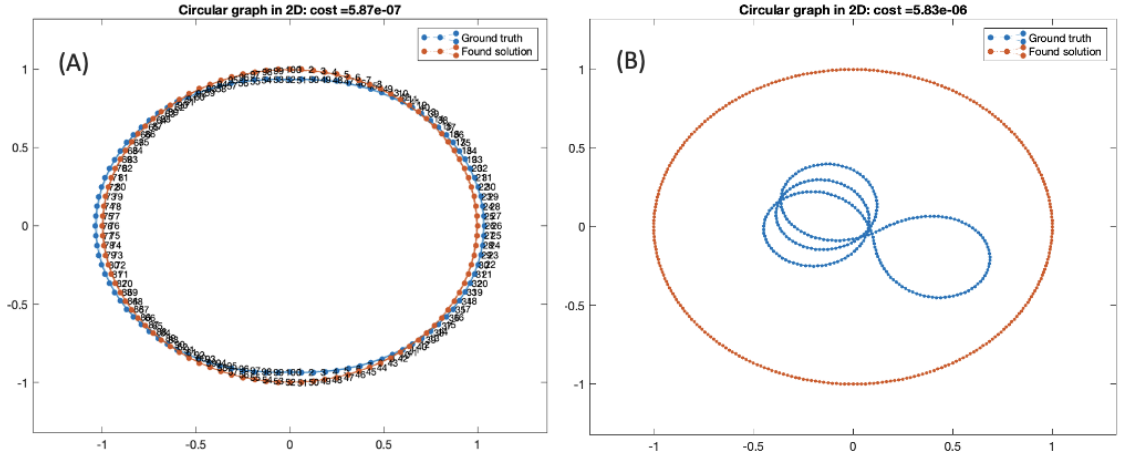

Figure 2: These figures are the ground truth and the solution found by the RTR algorithm for circular graph in $\mathbb{R}^2$; $n = 100,\ m = 8$ for (A) and $n = 300,\ m = 8$ for (B).

Graphs (3)-(A) and (B) are similar of the previous graphs at the exception that this time the ground truth graphs are not equally spaced on the circle but randomly distributed around this one. Even with more edges by vertex, the solution found is still far from the ground truth both visually and in terms of cost. The same behaviour is observed on graphs (C) and (D).

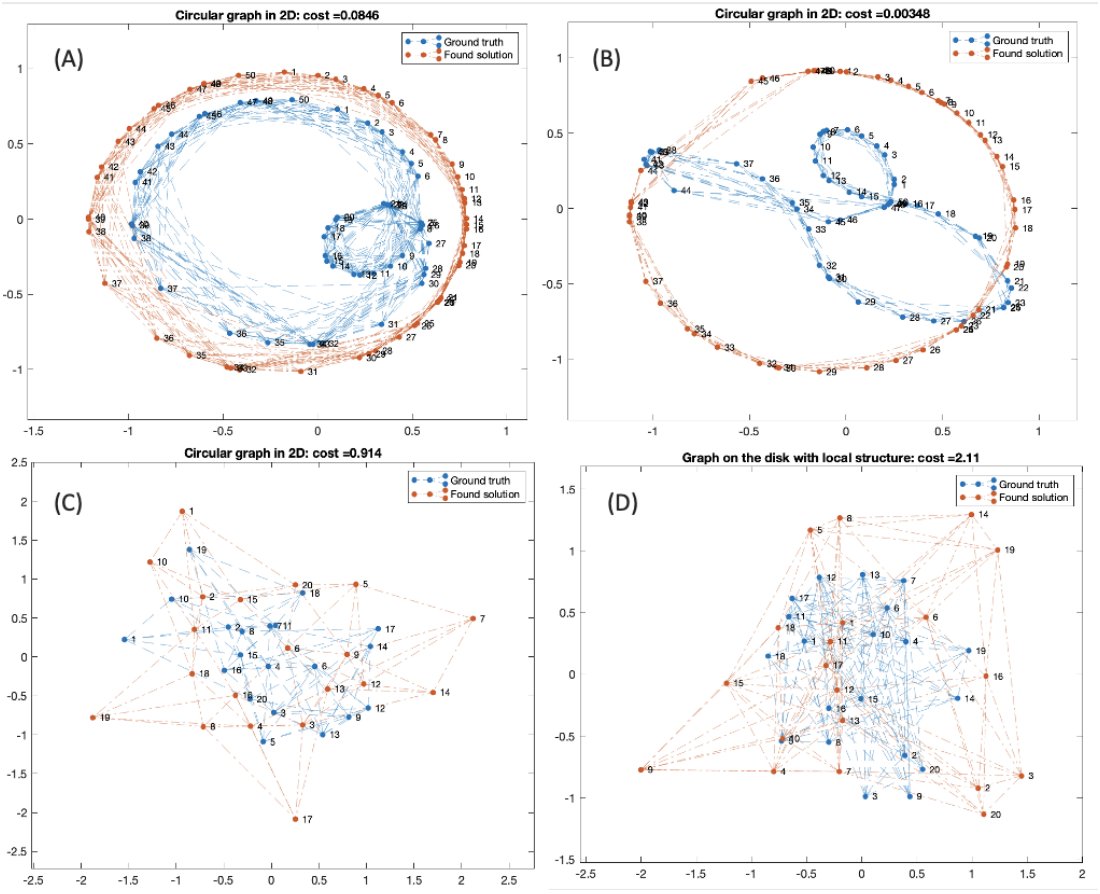

Figure 3: These figures are the ground truth and the solution found by the RTR algorithm for local structure on the circle graph for (A) and (B), and on the disk for (C) and (D); (A) $n = 50,\ m = 8$, (B) $n = 50,\ m = 14$, (C) $n = 20,\ m = 5$ and (D) $n = 20,\ m = 8$.

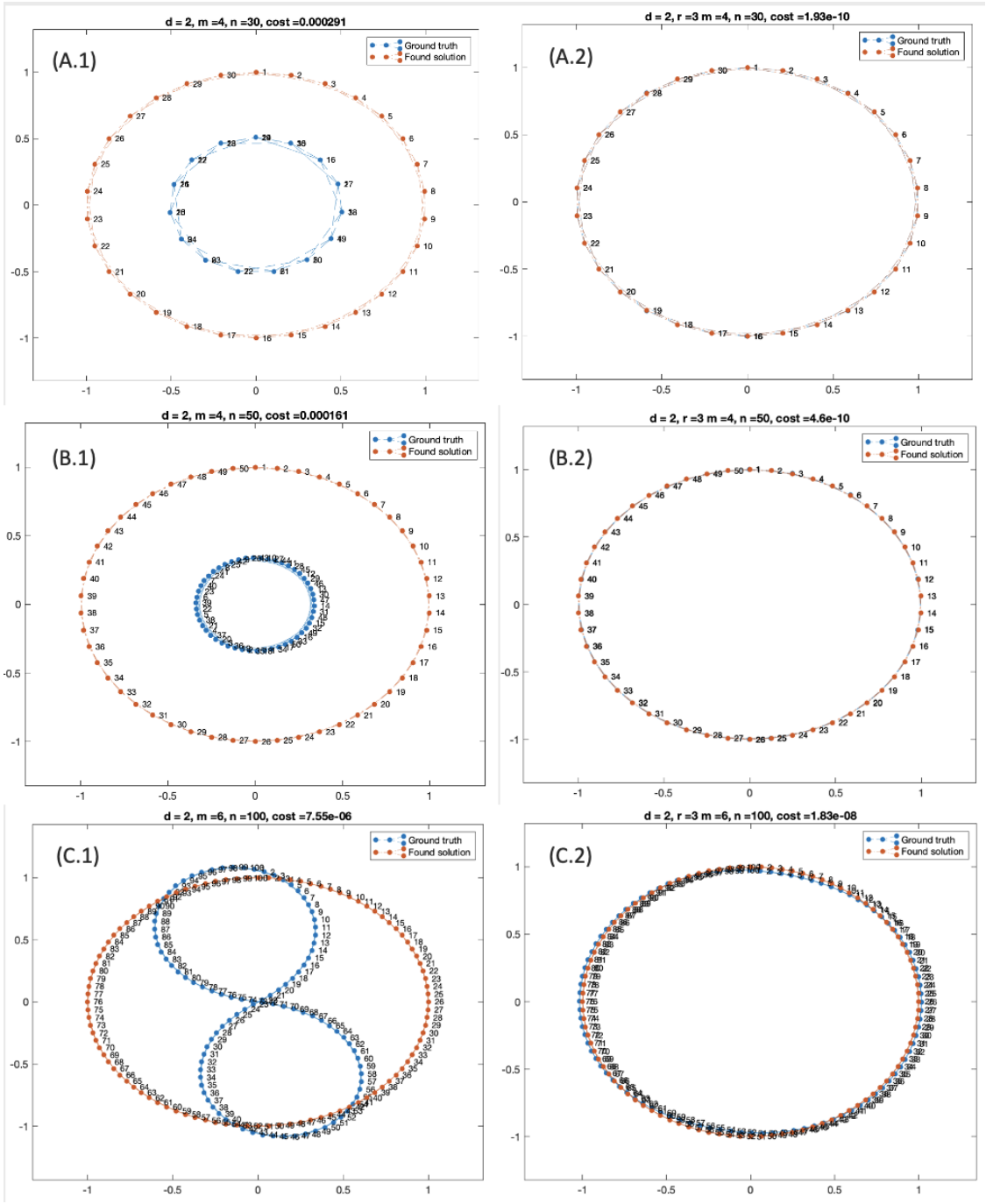

Figure 4: The ground truth graph on the left and the right are the same circular graph in $\mathbb{R}^2$, however the solution found on the left is the solution found without relaxation and the solution on the right is the solution found with relaxation of one dimension and then project back to $\mathbb{R}^2$. We have: $n = 30,\ m = 4$ for (A.1) and (A.2), $n = 50,\ m = 4$ for (B.1) and (B.2), and $n = 100,\ m = 6$ for (C.1) and (C.2).

The first benefice of relaxing Problem 3 are observable on Fig. (4) and on Fig. (5)-(A) and (B). Indeed, these graphs converges to the same kind of local minima observed on the previous graphs, however for the same ground truth graphs and with the relaxing of one dimension, these graphs manage to escape the previous local minima or others to converges to the optimal solution or at least really close to it. These examples "show" that in general the relaxation of Problem 3 often allows to escape from the convergences to spurious local minima. Nevertheless, as it can be see on Fig. (5-(C)), it also happen the converse and sometimes even if the solution in the relaxed space converge to 0 value objective function, and so a potential global optimizer, the solution obtain after projection in

the original space could be far from being a local or global solution.

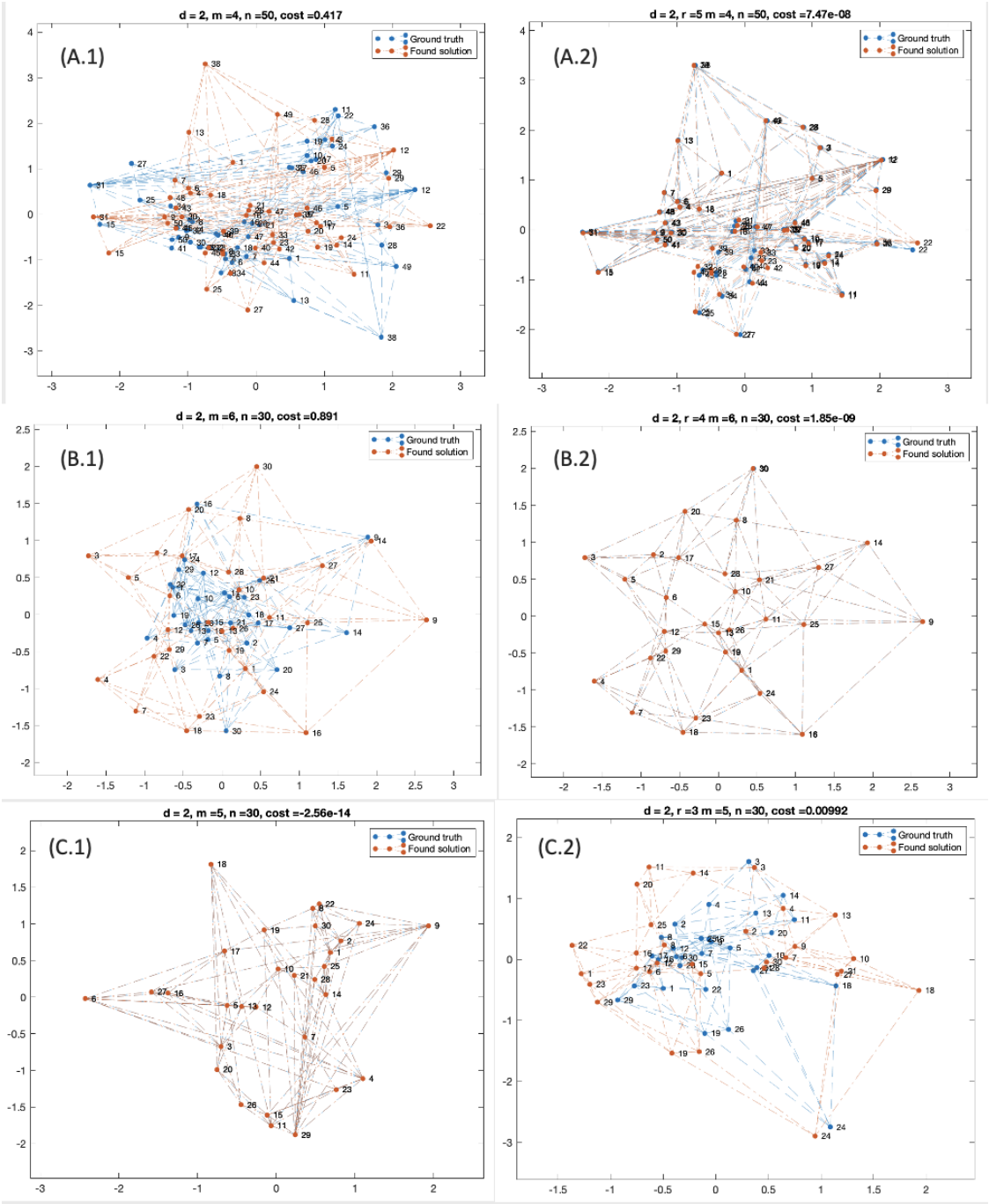

Figure 5: The ground truth graph on the left and the right are the same local structured graph in the disk in $\mathbb{R}^2$, however the solution found on the left is the solution found without relaxation and the solution on the right is the solution found with relaxation of one dimension and then project back to $\mathbb{R}^2$. We have: $n = 50,\ m = 4$ for (A.1) and (A.2), $n = 30,\ m = 6$ for (B.1) and (B.2), and $n = 30,\ m = 6$ for (C.1) and (C.2).

4.2 Behaviour of relaxation on different graphs

We observe that on graphs of reasonable size (number of vertices smaller than few hundreds), the result of relaxing Problem 3 of 1, 3, 10 or more dimension does not change the convergence to the optimal value in general, but more we increase the number of dimensions more the Problem 4 becomes

computationally demanding to be solved. So we prefer to focus on the case where we only consider relaxations of 3 dimensions.

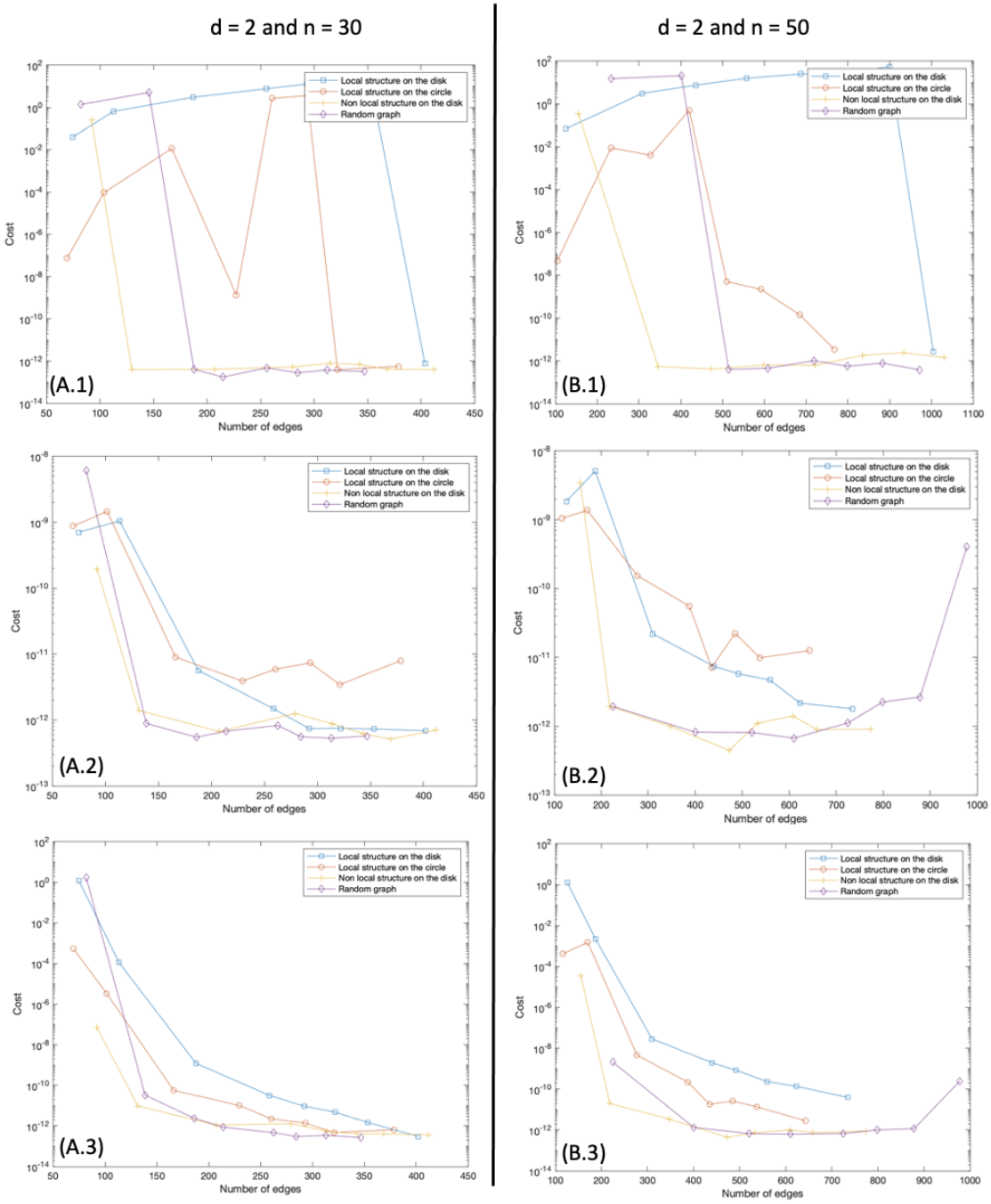

Figure 6: Convergence of the solution for different types of graphs in $\mathbb{R}^2$ and for increasing number of edge. The left column correspond to graphs on $n = 30$ vertices and the right one correspond to graphs with $n = 50$ vertices. The first row of plots show the convergence of the solution without relaxation, the second one show the convergence of the solution in $\mathbb{R}^5$, and the third one show the convergence of the projection of the relaxed solution in $\mathbb{R}^2$. We run 15 experiments for each case and we plot the median result of this experiments.

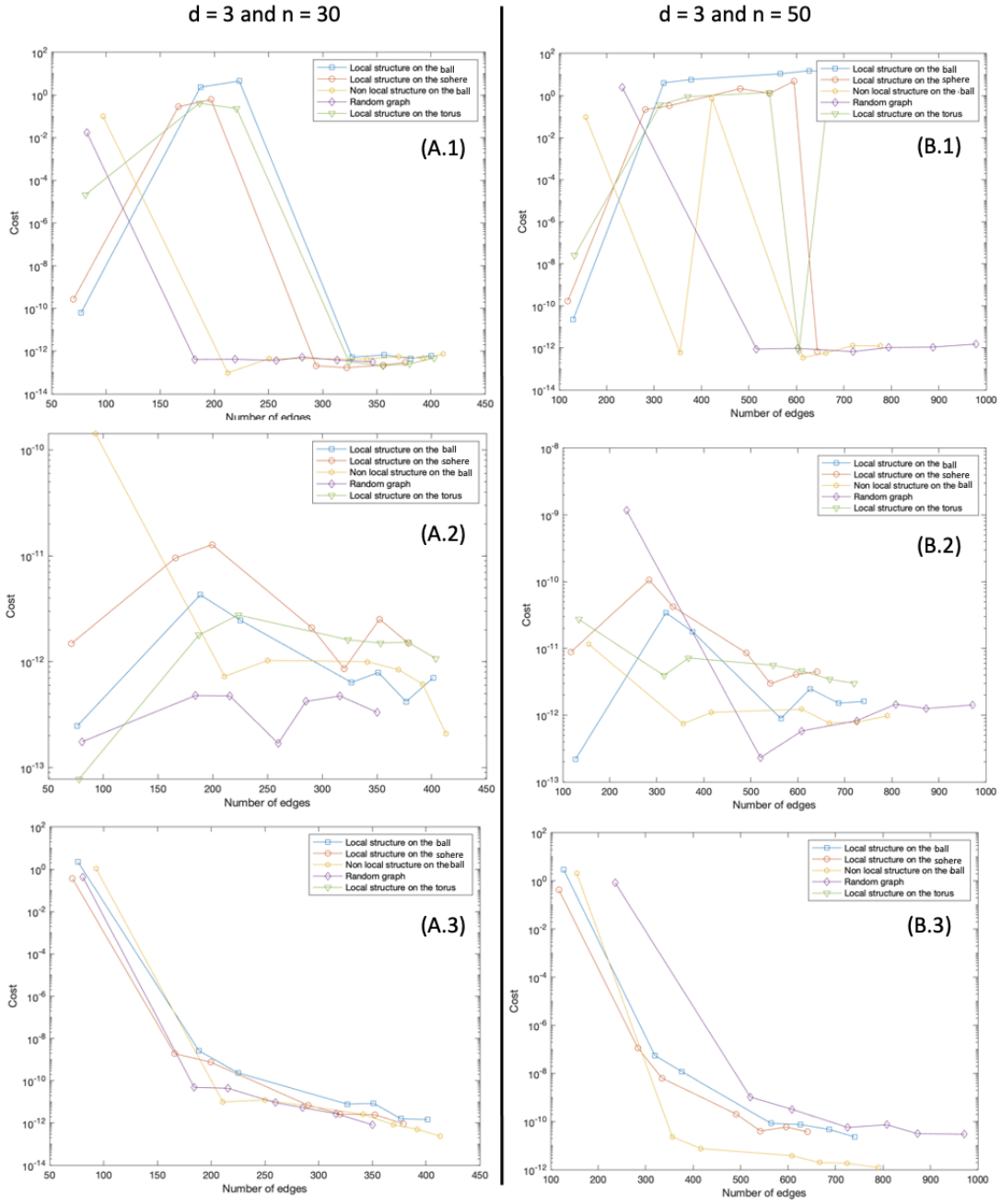

Figure 7: Convergence of the solution for different types of graphs in $\mathbb{R}^3$ and for increasing number of edge. The left column correspond to graphs on $n = 30$ vertices and the right one correspond to graphs with $n = 50$ vertices. The first row of plots show the convergence of the solution without relaxation, the second one show the convergence of the solution in $\mathbb{R}^6$, and the third one show the convergence of the projection of the relaxed solution in $\mathbb{R}^3$. We run 15 experiments for each case and we plot the median result of this experiments.

Globally, the behaviour of the convergence presented on Fig. (6,7,8) are quite similar . Without relaxation, the solution does not necessarily converge to a optimal value when the number of edges is to low. However, it tend to converge when the number of edge is high enough. It seems that the graphs with non-local structure and the random graphs needs less connections in the graph to converge to a optimal value then the graphs with a local-structure. In every case, the relaxed solution converge to a solution which is almost always numerically optimal ($\approx 10^{-10}$). Moreover, for the projection of the relaxed solution, we also observe that the different figures looks similar, since the cost of the

projection is sub-optimal for graphs which does not have many edges and the cost of the projection keep to decreasing when the connectivity of the graphs increase to quickly reach a numerically optimal value.

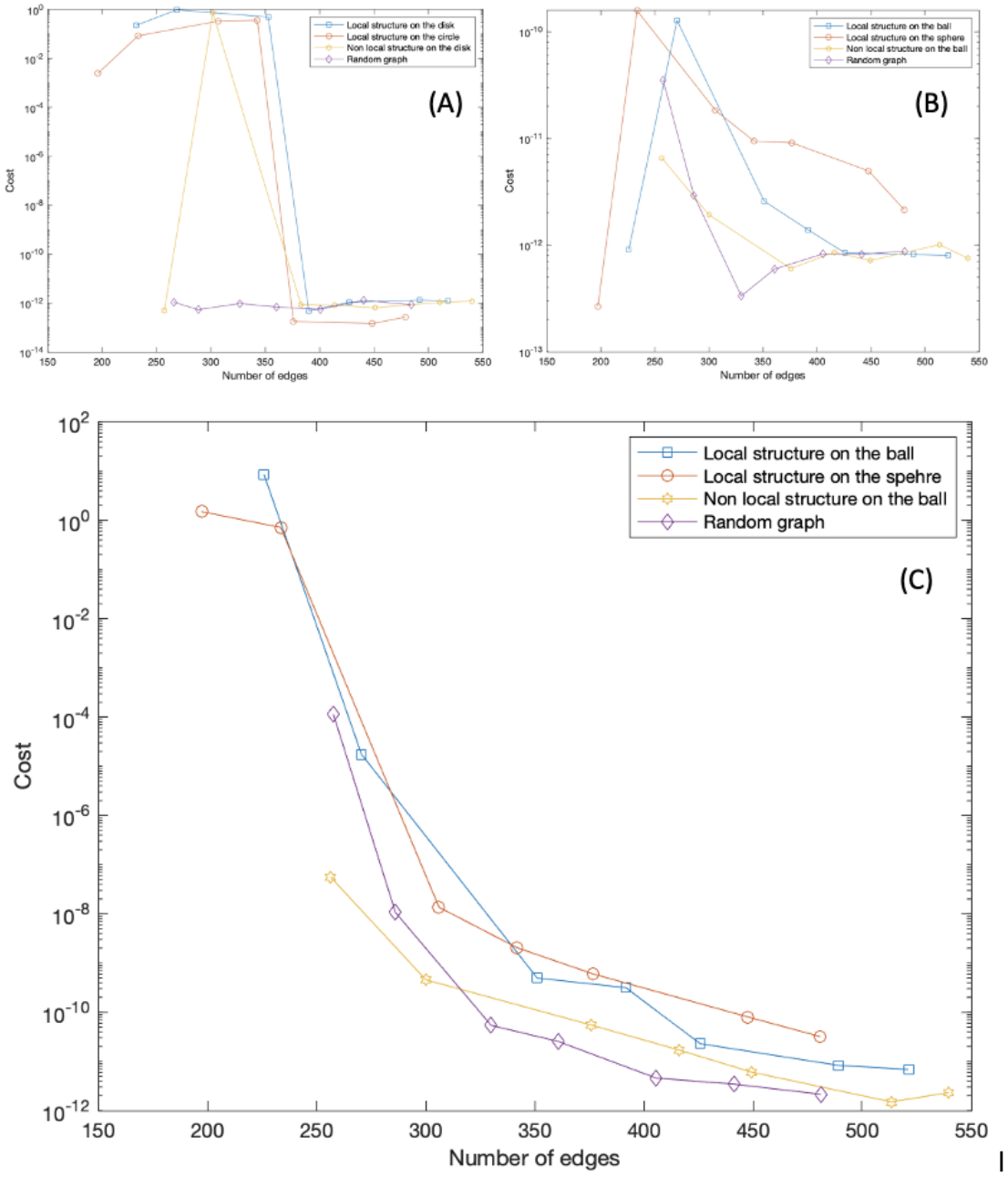

Figure 8: Convergence of the solution for different types of graphs ($n = 30$) in $\mathbb{R}^2$ and for increasing number of edge. (A) show the convergence of the solution without relaxation, (B) show the convergence of the solution in $\mathbb{R}^9$, and (C) show the convergence of the projection of the relaxed solution in $\mathbb{R}^6$. We run 15 experiments for each case and we plot the median result of this experiments.

4.3 Convergence of the last singular value during the optimization path with relaxation

In this section, we will present a really interesting pattern in the convergence of the last singular values of the solution along the optimization path. At each iteration of the RTR solvers we add the computation of the SVD decomposition to store the singular value at each iteration. Plotting the

evolution of these singular values along the optimization path allows us to observe the last singular values, here the last three, convergence exponentially to zero. Trying to characterize this convergence, we found, after some heuristics tests, that the convergences of these singular values behave as

$$i \mapsto \|Y_0\|_F \left(\left(\frac{1}{2} \right)^{\frac{\bar{\beta}}{n \log_2(d+1)}} \right)^i, \quad \text{for the } i^{th} \text{ iteration,} \tag{36}$$

where $\beta = \frac{1}{n} \sum_{j=1}^{n} degree(v_i)$ is the mean degree of the graph $\mathcal{G}$.

Until now, we are not able to give a explication plausible of this pattern of convergence, and if there is a real quantification of the convergence of the last singular value in this problem, we believe that it should certainly be different from the one that we only find "with the hands". Nevertheless, we are able to give a kind of interpretation of this coefficient. Indeed, more the graph is connected more the graph should be rigid and thus it could make sens that the convergence to a subspace of the relaxed space is faster. In the same way, more there is vertices and higher is the dimension of the original problem, the slowest should be the convergence to a subspace.

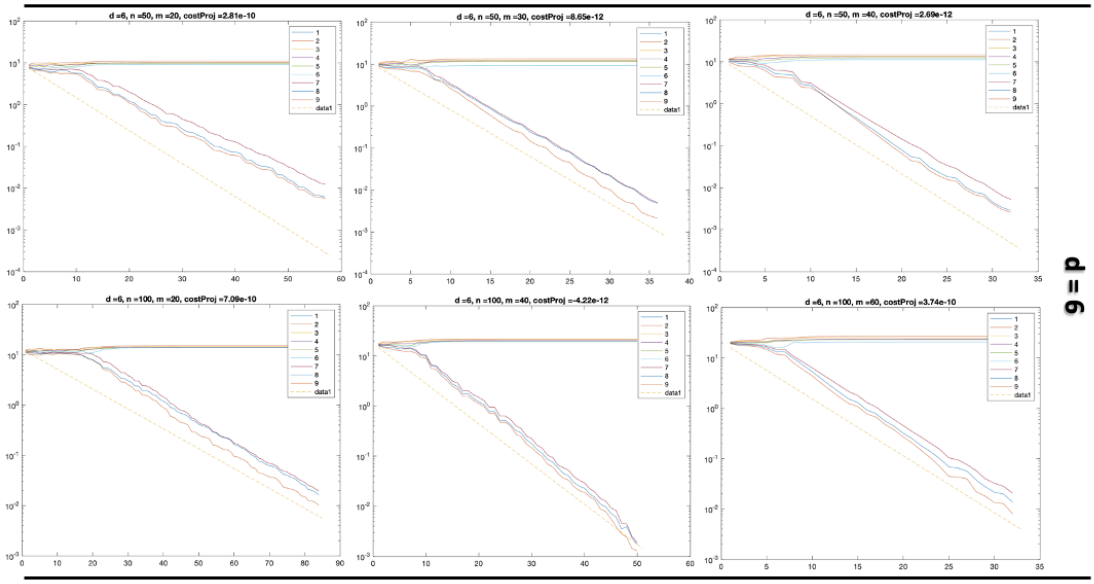

Figure 9: Convergence of the singular values in the optimization path on local structured graphs in the ball of different dimension, different number of vertices and different connectivity, with a relaxation of 3 dimensions.

This convergence pattern can be observed on Figures (10) and (10), the dotted line represent the convergence pattern given in Eq.(36). Whether in terms of dimension, number of vertices or edges, convergence seems to be broadly in line with the pattern we've outlined. Moreover, we observed the same pattern when we tested this on other types of graph. The correspondence between the two convergences was in some cases less "perfect", as with random graphs, on which we observed the same type of convergence but slightly faster for the latter. This can certainly be explained by a multitude of reasons, but we tend to think that it is partly due to the fact that the average degree of the graph's greens is used as a sort of measure of the graph's connectivity/rigidity, but this measure is certainly too basic to characterise this.

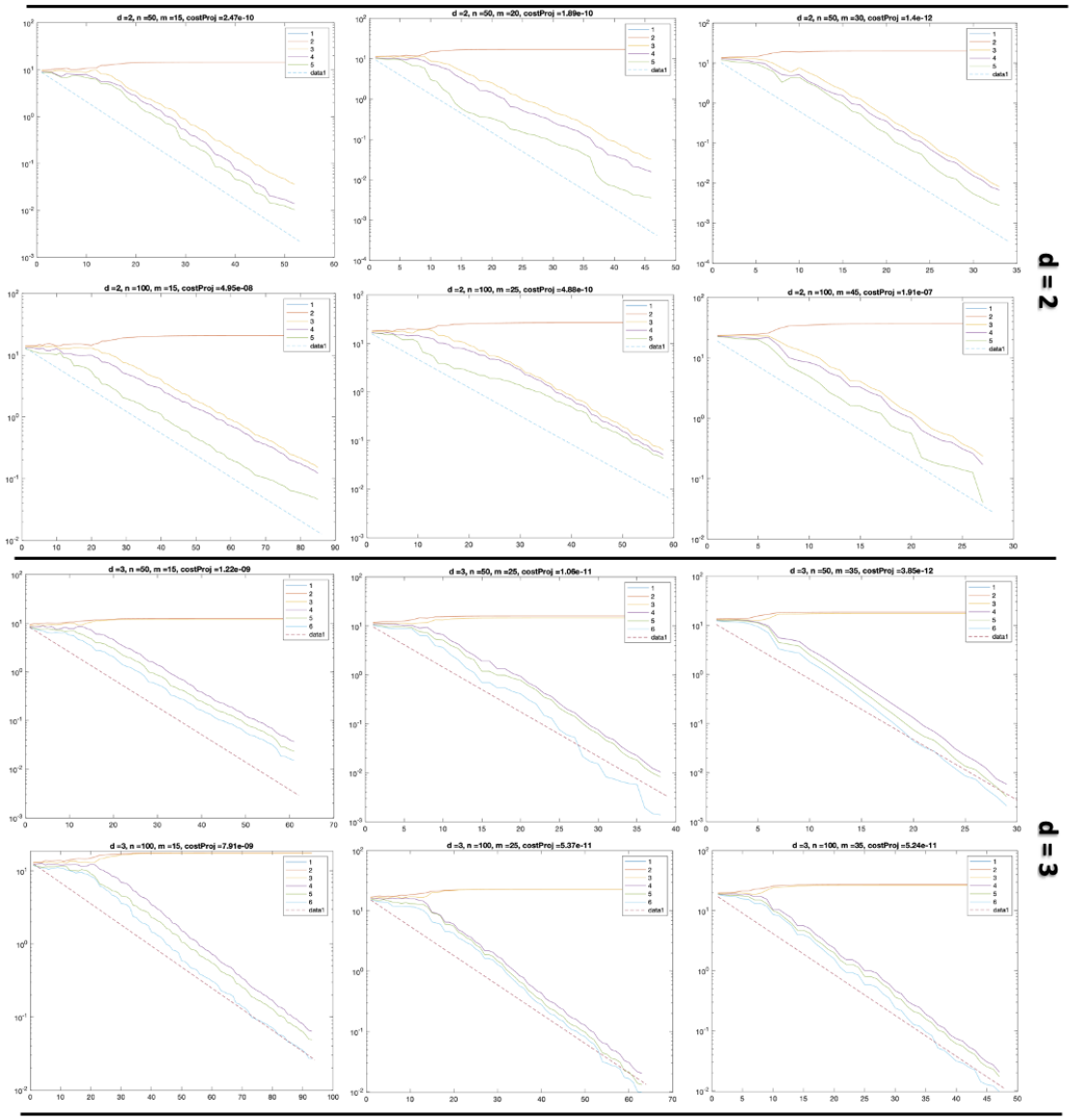

Figure 10: Figure (10) continued. Convergence of the singular values in the optimization path on local structured graphs in the ball of different dimension, different number of vertices and different connectivity, with a relaxation of 3 dimensions.

4.4 Ideas to be investigated

We algorithm which increase temporally the dimension to escape local minima and then project back when "far" from the local minima.

- **Convergence of the singular value**: The pattern observed in the previous section seems to us to be the most interesting thing to investigate in future work. Indeed, we could first try to see if this pattern also appears when using other optimisation methods, so that we would have an initial idea of whether this phenomenon is linked only to the geometry of the problem or whether it is also intrinsically linked to our optimisation method. Next, we should try to find a more representative measure of graph connectivity than the one used here. One idea we had was not to focus just on the convergence of singular values, but possibly to study the evolution of the main right-singular space of the relaxed solution $Y\mathbb{R}^r$.
- **Relaxed space with last singular values constrained :** From what we have been able to observe on our problem, the relaxation of our problem very often allows us to converge to a

global minimum in the relaxed space. However, in some cases, this solution is not contained in a space of dimension close to the dimension of the initial problem, d, with the consequence that the solution obtained after projection is no longer optimal. One idea we had, but didn't have time to implement, was to change our relaxed space $\mathcal{OB}(r,m)$ to the following space $\{Y = U\Sigma V^T \in \mathcal{OB}(r,m) : \Sigma_{ii} < \alpha, \text{ for } i = d+1, \cdots, r\}$ and where $\alpha\mathbb{R}_+$ is a constant. The idea is that we hope that by choosing α correctly we will obtain a space in which our solution always has the possibility of escaping from spurious local minima, but in which our graph is globally rigid and thus still have a unique global minimum.

- **Study of the projection with perturbation of the relaxed solution :** In the case where we obtain a solution in the relaxed space $\mathbb{R}^r$, but where the graph as at least a vertex of degree $l \leq r-1$, then we know by Theorem 3.5 that we could move this vertex on a sphere of dimension $r-l$ without changing the relative distances to its neighbours. For that we can use algorithm (2) to perturb the relaxed solution, however even if the solution was well perturb with method, the solution stay quite close to the original one and the projection stay a global solution in our experiments. We have some difficulty to well understand why its the case. It will certainly demand more time and investigation to understand what is happening exactly here.

5 Conclusion

We began this project by studying the geometry of the problem as well as the relaxations associated with it. This allowed us to gain a better understanding of the problem and build a strong knowledge foundation to develop our intuition and understanding of our objective, which was to investigate rank relaxation in this distance-based problem.

By questioning the existence and uniqueness of the solution, that led us to a topic that initially seemed far from the optimization context we were in: graph rigidity. Exploring this subject allowed us to develop a solid understanding of the complexity associated with determining whether a graph can have a unique realization in Euclidean space but also of the limitation of what we could hope to obtain as an answer with the information at our disposal. However, we gained numerous insights and intuition that proved valuable for the remainder of the project, including methods to obtain graphs that realized the measured distances from other realization graphs. This enabled us to develop several ideas that we further investigated.

In the various simulations we have carried out, we have been able to test this rank relaxation method in numerous situations and for various graph structures, but only in the noiseless case of the distance-bases problem. We were able to observe that rank relaxation made it possible to find the optimal solution almost every time, as long as the graph was sufficiently connected/rigid, and that in the opposite case the optimal solution found in the relax space did not result in an optimal solution. In addition, we found that relaxing the problem by many dimensions was often not beneficial and that a 'small' relaxation was often sufficient and less demanding from a computational point of view. During the relaxation of the problem, we also observed an interesting pattern of convergence of the last singular values. However, we have no conclusive explanation for this but it could be the subject of future investigation.

References

[1] Trevor Halsted and Mac Schwager. The Riemannian Elevator for certifiable distance-based localization, 2022. Under Review.

[2] A. I. Barvinok. Problems of distance geometry and convex properties of quadratic maps. *Discrete & Computational Geometry*, 13(2):189–202, 1995.

[3] Nicolas Boumal. A riemannian low-rank method for optimization over semidefinite matrices with block-diagonal constraints, 2016.

[4] Bruce Hendrickson. Conditions for unique graph realizations. *SIAM Journal on Computing*, 21(1):65–84, 1992.

[5] Robert Connelly. Generic global rigidity. *Discrete & Computational Geometry*, 33(4):549–563, 2005.

[6] Yuan Zhang, Shutang Liu, Xiuyang Zhao, and Zhongtian Jia. Theoretic analysis of unique localization for wireless sensor networks. *Ad Hoc Networks*, 10(3):623–634, may 2012.

[7] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre. Manopt, a Matlab toolbox for optimization on manifolds. *Journal of Machine Learning Research*, 15(42):1455–1459, 2014.