Rotation Averaging with the Chordal Distance: Global Minimizers and Strong Duality

Anders Eriksson, Carl Olsson, Fredrik Kahl, and Tat-Jun Chin

Abstract—In this paper we explore the role of duality principles within the problem of rotation averaging, a fundamental task in a wide range of applications. In its conventional form, rotation averaging is stated as a minimization over multiple rotation constraints. As these constraints are non-convex, this problem is generally considered challenging to solve globally. We show how to circumvent this difficulty through the use of Lagrangian duality. While such an approach is well-known it is normally not guaranteed to provide a tight relaxation. Based on spectral graph theory, we analytically prove that in many cases there is no duality gap unless the noise levels are severe. This allows us to obtain certifiably global solutions to a class of important non-convex problems in polynomial time.

We also propose an efficient, scalable algorithm that outperforms general purpose numerical solvers by a large margin and compares favourably to current state-of-the-art. Further, our approach is able to handle the large problem instances commonly occurring in structure from motion settings and it is trivially parallelizable. Experiments are presented for a number of different instances of both synthetic and real-world data.

Index Terms—Rotation averaging, Structure from Motion, Lagrangian duality, graph Laplacian, chordal distance.

1 INTRODUCTION

Rotation averaging appears as a subproblem in many important applications in computer vision, robotics, sensor networks and related areas. Given a number of relative rotation estimates between pairs of poses, the goal is to compute absolute camera orientations with respect to some common coordinate system. In computer vision, for instance, non-sequential structure from motion systems such as [1], [2], [3] rely on rotation averaging to initialize bundle adjustment. The overall idea is to consider as much data as possible in each step to avoid suboptimal reconstructions. In the context of rotation averaging this amounts to using as many camera pairs as possible.

The problem can be thought of as inference on a graph. An edge $(i, j)$ in this undirected graph represents a relative rotation measurement $\tilde{R}_{ij}$ and the objective is to find the absolute orientation $R_i$ for each vertex $i$ such that $R_i \tilde{R}_{ij} = R_j$ holds (approximately in the presence of noise) for all edges. The problem is generally considered difficult due to the need to enforce non-convex rotation constraints. Indeed, both $L_1$ and $L_2$ formulations of rotation averaging can have local minima, see Figure 1. Wilson et al. [4] studied local convexity of the problem and showed that instances with large loosely connected graphs are hard to solve with local, iterative optimization methods.

In contrast, our focus is on global optimality. In this paper we show that convex relaxation methods can in fact overcome the difficulties with local minima in rotation averaging. We utilize Lagrangian duality to handle the quadratic non-convex rotation constraints. While such an approach is normally not guaranteed to provide a tight relaxation, we give analytical error bounds that guarantee there will be no duality gap. For instance, it is sufficient that each angular residual is less than 42.9° to ensure optimality for complete graphs. Additionally, we develop a scalable and efficient algorithm, based on block coordinate descent, that outperforms standard semidefinite program (SDP) solvers for this problem. Further, we experimentally compare our approach to specialized state-of-the-art solvers.

1.1 Related Work

Rotation averaging has been under intense study in recent years, see [1], [5], [6], [7], [8], [9]. Despite progress in practical algorithms, they largely come without guarantees. One of the earliest averaging methods was due to Govindu [10], who showed that when representing the rotations with quaternions the problem can be viewed as a linear homogeneous least squares problem. There is however a sign ambiguity in the quaternion representation that has to be resolved before the formulation can be applied. Additionally, since a quaternion only represents a rotation if its
length is one, a norm constraint has to be incorporated for each rotation. It was observed by Fredriksson and Olsson in [11] that since both the objective and the constraints are quadratic, the Lagrange dual can be computed in closed form. The resulting SDP was experimentally shown to have no duality gap for moderate noise levels. In order to avoid solving a costly SDP for large scale problems, [11] proposed a verification scheme that given a candidate primal solution, computes a dual solution and evaluates the duality gap. In this way it is possible to test and verify optimality of a candidate solution obtained using any feasible algorithm.

A more straightforward rotation representation is using $3 \times 3$ matrices. Martinec and Pajdla [1] approximately solve the problem by ignoring the orthogonality and determinant constraints. This allows simple optimization by computing eigenvectors of the resulting homogeneous system of equations. A similar relaxation was derived by Arie-Nachimson et al. in [12]. In addition, an SDP formulation was presented which is equivalent to the one we address here, but with no performance guarantees. The tightness of SDP relaxations for 2D rotation averaging is studied in [13].

A number of robust approaches have been developed to handle outlier measurements. A sampling scheme over spanning trees of the graph is developed by Govindu in [14]. Enqvist et al. [2] also start from a spanning tree and add relative rotations that are consistent with the solution. In [15] the Weiszfeld algorithm is applied to single rotation averaging with the $L_1$ norm. In [16] convexity properties of the single rotation averaging problem are given. To our knowledge these results do not generalize to the case of multiple rotations. In [17] a robust formulation is solved using IRLS and in [18] Cramér-Rao lower bounds are computed for maximum likelihood estimators, but neither with any optimality guarantees.

A closely related problem is that of pose graph estimation, where camera orientations and positions are jointly optimized. In this context Lagrangian duality has been applied [19], [20], both for optimization and for verification similar to [11]. In [21] a consensus algorithm that allows for efficient distributed computations is presented. A fast verification technique for pose graph estimation was given in [22]. In a recent paper [23] an SDP relaxation for pose graph estimation with performance guarantees is analyzed. It is therefore easily seen that the chordal distance is related to the residual rotation angle by

$$d(R, S) = \sqrt{\det(R^T R - I)} = \sqrt{\det(R - S)}.$$  

For further details on rotation parametrizations and distance measures, see [5].

The optimality guarantees developed in this paper relies on some basic concepts from spectral graph theory. Let $G = (V, E)$ denote an undirected graph with vertex set $V$ and edge set $E$ and let $n = |V|$. The adjacency matrix $A$ is by definition the $n \times n$ matrix with elements

$$a_{ij} = \begin{cases} 0 & (i, j) \notin E \\ w_{ij} & (i, j) \in E \end{cases}$$

where $w_{ij}$ is a positive weight. In this paper we mostly consider unweighted graphs, where $w_{ij} = 1$, but our results extend to the weighted case where $w_{ij}$ can be interpreted as theoretical findings and which compare favourably to previous state-of-the-art.

This paper is an extension of the conference paper [24].

1.2 Preliminaries and Notation

The group of all rotations about the origin in three-dimensional Euclidean space is the Special Orthogonal Group, denoted $SO(3)$. This group is commonly represented by rotation matrices, orthogonal $3 \times 3$ real-valued matrices with positive determinant, i.e.,

$$SO(3) \in \{ R \in \mathbb{R}^{3 \times 3} \mid R^T R = I, \det(R) = 1 \}. \quad (1)$$

If we omit $\det(R) = 1$, we get the Orthogonal Group, $O(3)$.

Rotations can also be parametrized using the axis-angle representation predicated by Euler’s rotation theorem. This theorem states that any rotation in three-dimensional space is equivalent to a single rotation about a fixed axis $v$ by a magnitude of $\alpha$ degrees. This allows us to represent $R$ using the explicit expression

$$R = V \begin{bmatrix} \cos \alpha & - \sin \alpha & 0 \\
\sin \alpha & \cos \alpha & 0 \\
0 & 0 & 1 \end{bmatrix} V^T,$$  

where $V$ is an orthogonal matrix that maps the rotation axis $v$ onto the z-axis. Note that the orientation angle can easily be computed from $\text{tr}(R) = 1 + 2 \cos \alpha$. Both of the above representations of rotations in three dimensions will be used interchangeably.

To measure distances between rotations we will in this work restrict ourselves to the chordal distance. This is the most commonly used metric when analyzing Lagrangian duality in rotation averaging. It is a convenient choice as it is quadratic in its entries leading to a particularly simple derivation and form of the associated dual problem. The chordal distance between two rotations $R$ and $S$ is defined as their Euclidean distance in the embedding space, that is,

$$d(R, S) = \| R - S \|_F.$$  

Since $\|R\|_F^2 = \|S\|_F^2 = 3$ we have

$$d(R, S)^2 = \| R \|_F^2 - 2 \text{tr} (R^T S) \| S \|_F^2 = 4(1 - \cos(\alpha)), \quad (4)$$

where $\alpha$ is the rotation angle of $R^T S$. It is therefore easily been that the chordal distance is related to the residual rotation angle by

$$d(R, S) = \sqrt{4(1 - \cos(\alpha))} = \sqrt{8 \sin^2 \frac{\alpha}{2}} = 2 \sqrt{2} \sin \frac{|\alpha|}{2}. \quad (5)$$

For further details on rotation parametrizations and distance measures, see [5].
a measurement uncertainty. The degree $d_i$ is the number of edges incident to vertex $i$, and the degree matrix $D$ is the diagonal matrix $D = \text{diag}(d_1, \ldots, d_n)$. The Laplacian $L_G$ of $G$ is defined by

$$L_G = D - A. \quad (7)$$

It is easily seen that for any vector $x \in \mathbb{R}^n$ we have

$$x^T L_G x = \sum_{(i,j) \in E} w_{ij} (x_i - x_j)^2, \quad (8)$$

and therefore $L_G \succeq 0$. Furthermore, since by construction the row sums over $L_G$ are zero, the vector $(1 \ 1 \ \ldots \ 1)^T$ is an eigenvector with eigenvalue zero. The second smallest eigenvalue $\lambda_2$ of $L_G$, also known as the algebraic connectivity or the Fiedler value, reflects the connectivity of $G$. For a connected graph $G$, which is the only case of interest to us, we always have $\lambda_2 > 0$. In general the magnitude of $\lambda_2$ reflects how well connected the overall graph is, see [25] for further analysis.

## 2 Problem Statement

The problem of rotation averaging is defined as the task of determining a set of $n$ absolute rotations $R_1, \ldots, R_n$ given distinct estimated relative rotations $\tilde{R}_{ij}$. Available relative rotations are represented by the edge set $E$ of the graph $G = (V, E)$. Under ideal conditions this amounts to finding the $n$ rotations compatible with the linear relations,

$$R_i \tilde{R}_{ij} = R_j, \quad (9)$$

for all $(i, j) \in E$. However, in the presence of noise, a solution to (9) is not guaranteed to exist. Instead, it is typically solved in a least-mean-square sense,

$$\min_{R_1, \ldots, R_n} \sum_{(i,j) \in E} d(R_i \tilde{R}_{ij}, R_j)^p, \quad (10)$$

where $p \geq 1$ and $d(\cdot, \cdot)$ is a distance function on the space of rotations. With the chordal distance, we define the rotation averaging problem as

$$\arg \min_{R_1, \ldots, R_n \in \text{SO}(3)} \sum_{(i,j) \in E} \|R_i \tilde{R}_{ij} - R_j\|_F^2, \quad (11)$$

which, with trace notation, can be simplified to

$$\arg \min_{R_1, \ldots, R_n \in \text{SO}(3)} - \sum_{(i,j) \in E} \text{tr} \left( R_i \tilde{R}_{ij} R_j^T \right), \quad (12)$$

which constitutes our primal problem.

It will prove convenient to next introduce a compact matrix formulation. Let

$$\tilde{R} = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 \end{bmatrix}, \quad (13)$$

where $\tilde{R}_{ij} = \tilde{R}_{ji}^T$ and $a_{ij}$ are the elements of the adjacency matrix $A$ of the graph $G$ and let

$$R = \begin{bmatrix} R_1 & R_2 & \ldots & R_n \end{bmatrix}. \quad (14)$$

We may now write the primal problem as

$$(P) \quad \min_{R \in \text{SO}(3)^n} -\text{tr} \left( \tilde{R} RR^T \right) \quad \text{s.t.} \quad R \in \text{SO}(3)^n. \quad (15)$$

### 3 Optimality Conditions

Since (P) is a constrained program it is natural to introduce Lagrange multipliers to enforce the constraints $R \in \text{SO}(3)^n$. Using the Lagrangian function there are two ways of trying to solve the original primal problem. Either by computing all stationary (KKT) points and evaluating the objective values of these or by maximizing the dual function. While the former approach is guaranteed to find an optimal point (assuming the existence of one) it is usually not feasible for large scale problems due to the non-linearity of the involved equations. If the dual function can be computed (which we will see is the case for our application) the latter approach yields a convex problem (specifically, a concave maximization) that is solvable for large scale instances. The resulting maximum yields a lower bound on the primal objective. In the following sections we will use properties of the KKT points to derive conditions that ensure that the primal and dual values are the same, enabling us to find a solution via a convex program. In this section we present the KKT conditions and derive the dual program that we will use.

#### 3.1 Necessary Local Optimality Conditions

The constraint set $R \in \text{SO}(3)^n$ consists of two types of constraints; the orthogonality constraints $R_i R_i^T = I$ and the determinant constraints $\det(R_i) = 1, i = 1, \ldots, n$.

Consider relaxing the rotation averaging problem by removing the determinant constraint,

$$(P') \quad \min_{R \in \text{O}(3)^n} -\text{tr} \left( RR^T \right) \quad \text{s.t.} \quad R \in \text{O}(3)^n. \quad (16)$$

The constraint $R \in \text{O}(3)^n$ still requires the $R_i$’s to be orthogonal. The orthogonal matrices consist of two disjoint, non-connected sets, with determinants 1 and −1 respectively. Hence, any local minimizer to the problem (P) also has to be a local minimizer, and therefore a KKT point, to (P'). We note that orthogonality can be enforced by restricting the $3 \times 3$ diagonal blocks of the symmetric matrix $R_i R_i^T$ to be identity matrices. That is, for any block diagonal matrix of the form

$$\Lambda = \begin{bmatrix} \Lambda_1 & 0 & 0 & \ldots \\ 0 & \Lambda_2 & 0 & \ldots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \Lambda_n \end{bmatrix}, \quad (17)$$

we have

$$\Lambda(I - R R^T) = 0. \quad (18)$$

Hence the Lagrangian can be written

$$L(R, \Lambda) = -\text{tr} \left( RR^T \right) - \text{tr} (\Lambda(I - R R^T)) = \text{tr} \left( R(\Lambda - \tilde{R})R^T \right) - \text{tr} (\Lambda). \quad (19)$$

Differentiating (19) we obtain the stationarity condition for optimality as

$$\left( \frac{1}{2} \Lambda + \Lambda^T - \tilde{R} \right) R^T = 0. \quad (20)$$
Noting that this permits us to, without loss of generality, assume that \( \Lambda \) is symmetric, we can now state the KKT conditions as

\[
\begin{align*}
\text{(Stationarity)} & \quad (\Lambda^* - \tilde{R}) R^T x = 0 \quad (21a) \\
\text{(Primal feasibility)} & \quad R^* \in \text{SO}(3)^n. \quad (21b)
\end{align*}
\]

Equation (21a) states that the rows of a local minimizer \( R^* \) will be eigenvectors of the matrix \( \Lambda^* - \tilde{R} \) with eigenvalue zero. This allows us to compute the optimal Lagrange multiplier \( \Lambda^* \) from a given minimizer \( R^* \) in \( \text{SO}(3)^n \). By (21a) we see that

\[
\Lambda^*_i R^*_iT = \sum_{j \neq i} a_{ij} \tilde{R}_{ij} R^*_jT \iff \Lambda^*_i = \sum_{j \neq i} a_{ij} \tilde{R}_{ij} R^*_jR^*_iT \quad (22)
\]

for \( i = 1, \ldots, n \). We therefore have:

**Lemma 3.1.** For a stationary point \( R^* \) to the primal problem (P), we can compute the corresponding Lagrangian multiplier \( \Lambda^* \) in closed form via (22).

### 3.2 Sufficient Global Optimality Conditions

Next we present some simple constraints that ensure that the maximal value of the dual problem is the same as the minimum of (P). For simplicity we first consider the Lagrange dual of \( (P') \) which is a semidefinite program that we will use for optimization in later sections. The dual problem is defined by

\[
\max_{\Lambda} \min_{R} L(R, \Lambda).
\]

Since the (unrestricted) optimum of \( \min_{R} L(R, \Lambda) \) is either \(-\text{tr}(\Lambda)\), when \( \Lambda = \tilde{R} \succeq 0 \), or \(-\infty\) otherwise, we get

\[
(D) \quad \max_{\Lambda \succeq \tilde{R} \geq 0} -\text{tr}(\Lambda).\quad (24)
\]

It is clear (through standard duality arguments) that (D) gives a lower bound on (P'). Furthermore, if \( R^* \) is a stationary point fulfilling (21a) and (21b) with a corresponding Lagrangian multiplier \( \Lambda^* \) that satisfies \( \Lambda^* - \tilde{R} \succeq 0 \) then \( \Lambda^* \) is feasible in (D). It also follows that for the stationary point \( R^* \) we have \( \text{tr}(R^* \Lambda^* R^T) = \text{tr}(R^* \tilde{R} R^T) \) due to (21a). This together with (18) gives us

\[
-\text{tr}(\Lambda^*) = -\text{tr}(\Lambda^* R^T R^*) = -\text{tr}(R^* \tilde{R} R^T). \quad (25)
\]

This shows that the dual problem attains the same objective value as the primal one, that is, there is no duality gap between (P') and (D). Thus, the convex program (D) provides a way of solving the non-convex (P') when \( \Lambda^* - \tilde{R} \succeq 0 \).

We further note that if \( \Lambda^* - \tilde{R} \succeq 0 \) then by definition it is true that

\[
x^T (\Lambda^* - \tilde{R}) x \geq 0, \quad (26)
\]

for any \( 3n \)-vector \( x \). In particular, for any \( R \in \text{SO}(3)^n \),

\[
0 \leq \text{tr}(R(\Lambda^* - \tilde{R}) R^T) = \text{tr}(\Lambda^*) - \text{tr}(R \tilde{R} R^T) = \text{tr}(R^* \Lambda^* R^T) - \text{tr}(R \tilde{R} R^T),
\]

which shows that \(-\text{tr}(R^* \tilde{R} R^T) \leq -\text{tr}(R \tilde{R} R^T)\) for all \( R \in \text{SO}(3)^n \), that is, \( R^* \) is the global optimum of (P').

Finally we note that since \( R^* \) is also feasible in (P) and the global minimum of (P) is clearly at least as large as that of (P'), \( R^* \) must also be optimal in (P) if \( \Lambda^* - \tilde{R} \succeq 0 \). We summarize the above discussion in the following lemma:

**Lemma 3.2.** If a stationary point \( R^* \) (fulfilling (21a) and (21b)) with corresponding Lagrangian multiplier \( \Lambda^* \) fulfills \( \Lambda^* - \tilde{R} \succeq 0 \) then:

1. There is no duality gap between (P) and (D).
2. \( R^* \) is a global minimum for (P).

In the remainder of this paper we will study under which conditions \( \Lambda^* - \tilde{R} \succeq 0 \) holds and derive an efficient implementation for solving (D).

We note that it is possible to encode the determinant constraint with quadratic expressions by considering vector products of the rows of the rotation matrix (see e.g. [22]). It would therefore be easy to derive a Lagrangian for (P) instead of (P'). This approach would introduce a set of additional dual variables. This is likely to strengthen the dual formulation for problem instances where there is a duality gap. However under the above assumptions (which we will see in many cases hold under quite generous circumstances) it makes no difference as these would simply be set to zero.

## 4 MAIN RESULT

In this section, we will state our main result which gives error bounds that guarantee that strong duality holds for our primal and dual problems. From a practical point of view, the result means that it is possible to solve a convex semidefinite program and obtain the globally optimal solution to our non-convex problem, which is quite remarkable.

### 4.1 Strong Duality Theorem

Returning to our initial, primal rotation averaging problem (11). The goal is to find rotations \( R_i \) and \( R_j \) such that the sum of the residuals \( \|R_i R_j - R_j R_i\|^2 \) is minimized. For strong duality to hold, we need to show that there is a stationary point \( R^* \) for which the dual variable fulfills \( \Lambda^* - \tilde{R} \succeq 0 \).

Our main result shows that this can be done by bounding the angular error residuals.

**Theorem 4.1** (Strong Duality). Let \( R^*_i, i = 1, \ldots, n \) denote a stationary point to the primal problem (P) for a connected graph G with Laplacian \( L_G \). Let \( \alpha_{ij} \) denote the angular residuals, i.e., \( \alpha_{ij} = \angle(R^*_i R_{ij}, R^*_j) \). Then \( R^*_i, i = 1, \ldots, n \) will be globally optimal and strong duality will hold for (P) if

\[
|\alpha_{ij}| \leq \alpha_{\text{max}} \quad \forall (i,j) \in E,
\]

where

\[
\alpha_{\text{max}} = 2 \arcsin \left( \frac{1}{4} + \frac{\lambda_2(L_G)}{2d_{\text{max}}} - \frac{1}{2} \right),
\]

and \( d_{\text{max}} \) is the maximal vertex degree.

Note that the assumptions of the theorem are only stationarity of \( R^* \) and a bound on the angular error residuals. Thus an immediate consequence of the above result is the following converse:
Corollary 4.1 (Global/Local Minimizers). Any local minimizer of \((P)\) that is not global must have at least one angular error that is larger than (29).

It is clear that (29) will give a positive bound \(\alpha_{\text{max}}\) for any graph. Thus for any given problem instance, \(\alpha_{\text{max}}\) gives an explicit bound on the error residuals for which strong duality is guaranteed to hold. The strength of the bound will depend on the particular graph connectivity encapsulated by the Fiedler value \(\lambda_2(L_G)\) and the maximal vertex degree \(d_{\text{max}}\). We will see that for tightly connected graphs the bound ensures strong duality under surprisingly generous noise levels. In [4] it was observed that local convexity at a point holds under similar circumstances.

4.1.1 Complete Graphs

Consider a graph with \(n = 3\) vertices that are connected, and all degrees are equal, \(d_{\text{max}} = 2\), denoted \(K_3\). The Laplacian matrix \(L_{K_3}\) has the Fiedler value \(\lambda_2(L_{K_3}) = 3\). This gives \(\alpha_{\text{max}} = \frac{\pi}{3}\) rad = 60°. So, any local minimizer which has angular residuals less than 60° is also a global solution. For a complete graph with \(n\) vertices (Figure 2),

![Fig. 2. A complete graph (left) and a cycle graph (right), both with 6 vertices.](image)

denoted \(K_n\), it follows that \(d_{\text{max}} = n - 1\) since every pair of nodes is connected. Further, it is well-known (and easy to show) that \(\lambda_2(L_{K_n}) = n\), see [25]. As \(n\) becomes larger than 3, we get a decreasing series of upper bounds which in the limit tends to \(2 \arcsin \left( \frac{\sqrt{n-1}}{n} \right) \approx 0.749\) rad = 42.9°. Hence, as long as the residual angular errors are less than 42.9° - which is quite generous from a practical point of view - we can compute the optimal solution via a convex program. Also note that this bound holds independently of \(n\).

Corollary 4.2. For a complete graph \(K_n\) with \(n\) vertices, the residual upper bound \(\alpha_{\text{max}} = 2 \arcsin \left( \frac{\sqrt{n-1}}{n} \right) \approx 0.749\) rad = 42.9° ensures global optimality and strong duality for any \(n\).

4.1.2 Circular Graphs

Next we consider a graph that arises during circular camera motions. Such problems are frequently occurring when using turn tables for 3D reconstruction. The simplest example is a cycle graph, see Figure 2. Here each node is connected to its two closest neighbors.

The Laplacian for the graph in Figure 3 is

\[
L_G = \begin{pmatrix}
4 & -1 & -1 & 0 & 0 & 0 & -1 & -1 \\
-1 & 4 & -1 & -1 & 0 & 0 & 0 & -1 \\
-1 & -1 & 4 & -1 & -1 & 0 & 0 & 0 \\
0 & -1 & -1 & 4 & -1 & -1 & 0 & 0 \\
0 & 0 & -1 & -1 & 4 & -1 & -1 & 0 \\
0 & 0 & 0 & -1 & -1 & 4 & -1 & -1 \\
-1 & 0 & 0 & 0 & -1 & -1 & 4 & -1 \\
-1 & -1 & 0 & 0 & 0 & -1 & -1 & 4
\end{pmatrix}
\]

![Fig. 3. A circular graph with connections to its four closest neighbors.](image)

In general a graph with \(n\) vertices of this type has a Laplacian that is a circulant matrix, which corresponds to a discrete periodic convolution. It is diagonalized by a discrete Fourier matrix \(F_n\) with elements \(f_{kl} = e^{\frac{2\pi i (k-1)(l-1)}{n}}\), where \(k,l = 1,\ldots, n\) and its eigenvalues can be computed from \(F_n c^T\), where \(c\) is the first row of \(L_G\) [26]. Note that the eigenvalues are real when the elements of \(c\) are real (and by combining conjugate columns of \(F_n\), the eigenvectors can also be selected real). For a general circular graph with \(n\) vertices where each vertex is connected to its \(2b\) closest neighbors we get the eigenvalues

\[
\eta_k = \begin{cases} 
2b + 1 - \frac{\sin \left( \frac{\pi}{n} \left( 2b + 1 \right) \right)}{\sin \left( \frac{\pi k}{n} \right)} & k \neq 0 \\
0 & k = 0.
\end{cases}
\]

Note that \(L_G\) is essentially \((2b + 1) I - M\), where \(M\) is a sliding average filter. The eigenvalues of \(M\) are the Fourier coefficients of a step function, which is the sampled sinc-function. The Fiedler value is thus

\[
\lambda_2(L_G) = \eta_2 = \eta_n = 2b + 1 - \frac{\sin \left( \frac{\pi}{n} \left( 2b + 1 \right) \right)}{\sin \left( \frac{\pi}{n} \right)}.
\]

Since this graph has \(d_{\text{max}} = 2b\) we have

\[
\frac{\lambda_2(L_G)}{2d_{\text{max}}} \approx \frac{1}{2} - \frac{n}{4\pi b} \sin \left( \frac{\pi}{n} 2b \right),
\]

for large \(n\) and \(b\). This expression only depends on the quotient \(\frac{2b}{n}\) which corresponds to the fraction of available measurements. (Note also that when \(2b + 1 < \frac{n}{2}\) the right hand side is a proper lower bound on \(\lambda_2(L_G)\) since \(\sin \left( \frac{\pi}{n} \right) < \frac{\pi}{n}\) and \(\sin \left( \frac{\pi}{n} \left( 2b + 1 \right) \right) > \sin \left( \frac{\pi}{n} (2b) \right)\). Inserting (32) into (29) gives the function plotted in Figure 4. For \(\frac{2b}{n} = 1\)

![Fig. 4. The maximal angle residual \(\alpha_{\text{max}}\) (y-axis) versus the level of available data \(\frac{2b}{n}\) (x-axis) for circular graphs.](image)

the estimate (32) yields \(\alpha_{\text{max}} = 42.9^\circ\) which coincides with the estimate for large complete graphs. Additionally, it is strictly positive as long as \(\frac{2b}{n} > 0\). This analysis suggests that
the case \( b = 1 \) which corresponds to a single cycle graph, see Figure 2, is most likely to have multiple local minima. On the other hand, with less data the residual errors are usually smaller and therefore strong duality may still hold. In fact, we show below that for a cycle graph strong duality always holds. In Figure 1, we have a real example of an orbital camera motion which is close to a cycle. It may seem hard to determine if the camera motion consists of one or more loops around the object - we give three different local minima for this example.

For a cycle graph (31) with \( b = 1 \) reduces to \( \lambda_2(L_G) = (1 - \cos(\frac{\pi}{n})) \). Inserting into (29) and simplifying, we get \( \alpha_{\text{max}} = 2 \arcsin \left( \sqrt{\frac{1}{4} + \sin^2(\frac{\pi}{n})} - \frac{1}{2} \right) \). For large values of \( n \), the upper bound decreases rapidly. In fact, the upper bound is quite conservative and it is possible to show a stronger one using a different analysis.

**Theorem 4.2.** Let \( R^*_{ij} i = 1, \ldots, n \) denote a stationary point to the primal problem (P) for a cycle graph with \( n \) vertices. Let \( \alpha_{ij} \) denote the angular residuals, i.e., \( \alpha_{ij} = \angle(R^*_{ij} R_{ij}, R^*_{ij}) \). Then, \( R^*_{ij} i = 1, \ldots, n \) will be globally optimal and strong duality will hold for (P) if \( |\alpha_{ij}| \leq \frac{\pi}{n} \) for all \((i, j) \in E\).

**Proof.** A sufficient condition for strong duality to hold is that \( \Lambda^* - \tilde{R} \succeq 0 \) (Lemma 3.2). We now return to the proof of our main result. Recall that we have shown that if \( \alpha_{ij} = \alpha = \frac{\pi}{n} \) (which is always possible, see Theorem 23 in [27]). In conclusion, the angular residuals \( |\alpha_{ij}| \) of the globally optimal solution for a cycle is always less than or equal to \( \frac{\pi}{n} \), which means that strong duality holds. Conversely, if the angular residual is larger than \( \frac{\pi}{n} \) for a local minimizer, then it does not correspond to the global solution.

### 4.2 Proof of Theorem 4.1

We now return to the proof of our main result. Recall that a sufficient condition for strong duality to hold is that \( \Lambda^* - \tilde{R} \succeq 0 \) (Lemma 3.2). To prove Theorem 4.1 we will show that this is true under the conditions of the theorem.

To simplify the presentation we denote the residual rotations \( \tilde{R} = R^*_{ij} R_{ij} R^*_{ij}^T \) and define

\[
D_{R^*} = \begin{bmatrix}
R^*_{ij} & 0 & 0 \\
0 & R^*_{ij} & 0 \\
0 & 0 & R^*_{ij}
\end{bmatrix}.
\]

Then \( D_{R^*}(\Lambda^* - \tilde{R})D_{R^*}^T = \sum_{j \neq i} a_{ij} E_{ij} = \sum_{j \neq i} a_{ij} (\tilde{E}_{ij}^T + \tilde{E}_{ij}) \) and by induction, the matrix \( D_{R^*}(\Lambda^* - \tilde{R})D_{R^*}^T \) has the following tridiagonal (Laplacian-like) structure

\[
\begin{bmatrix}
E_{+T} & -E & \ldots & -E \\
-E & E_{+T} & \ldots & -E \\
\vdots & \ddots & \ddots & \vdots \\
-E & \ldots & -E & E_{+T}
\end{bmatrix}.
\]

Note that this means that the total error is equally distributed in an optimal solution among all the residuals, in particular, \( \alpha_{ij} = \alpha \) for all \((i, j) \in E\), where \( \alpha \) is the residual rotation angle of \( \tilde{E} \).

Let \( v \) denote the rotation axis of \( \tilde{E} \) and let \( u \) and \( w \) be an orthogonal base which is orthogonal to \( v \). Then, define the two vectors \( v_{\pm} = (v_{\pm,1} \ v_{\pm,2} \ \ldots \ v_{\pm,n})^T \), where \( v_{\pm,i} = \cos(\frac{\pi}{n})u \pm \sin(\frac{\pi}{n})w \) for \( i = 1, \ldots, n \). Now it is straightforward to check that \( v_{\pm} \) are eigenvectors to (34) with eigenvalues \( 4 \sin(\frac{\pi}{n} \pm \alpha) \sin(\frac{\pi}{n}) \). The sign of the smallest of these two eigenvalues determines the positive definiteness of the matrix in (34). In other words, we have shown that if \( |\alpha| \leq \frac{\pi}{n} \) then \( D_{R^*}(\Lambda^* - \tilde{R})D_{R^*}^T \succeq 0 \).

Requiring that the angular residuals \( |\alpha_{ij}| \) must be less than \( \pi/n \) for the global solution may seem like a restriction, but it is actually not. To see this, note that a non-optimal solution to the rotation averaging problem can be obtained by choosing \( R^* \) such that the first residual \( \alpha_{12} \) is zero, and then continuing in the same fashion such that all but the last residual \( \alpha_{1n} \) in the cycle is zero. In the worst case, \( \alpha_{1n} = \pi \).

However, this is (obviously) non-optimal. A better solution is obtained if we distribute the angular residual error evenly so that \( \alpha_{ij} = \frac{\pi}{n} \) (which is always possible, see Theorem 23 in [27]). In conclusion, the angular residuals \( |\alpha_{ij}| \) of the globally optimal solution for a cycle is always less than or equal to \( \frac{\pi}{n} \), which means that strong duality holds. Conversely, if the angular residual is larger than \( \frac{\pi}{n} \) for a local minimizer, then it does not correspond to the global solution.
Lemma 4.2. Let \( \alpha_{ij} \) be the residual angle of \( E_{ij} \). Then
\[
\| \Delta_{ii} \| \leq 2 \sum_{j \neq i} a_{ij} \sin^2 \left( \frac{\alpha_{ij}}{2} \right). \tag{41}
\]
If in addition \( 0 \leq \alpha_{\max} \leq \frac{\pi}{2} \), where \( \alpha_{\max} = \max_{i \neq j} |\alpha_{ij}|, \) then
\[
\| \Delta_{ii} \| \leq 2d_i \sin^2 \left( \frac{\alpha_{\max}}{2} \right). \tag{42}
\]
for all \( i = 1, \ldots, n \), where \( d_i \) is the degree of vertex \( i \).

Proof. It is easy to see that by applying a change of coordinates \( E_{ij} \) can be written
\[
E_{ij} = V_i \left[ \begin{array}{ccc}
\cos(\alpha_{ij}) & -\sin(\alpha_{ij}) & 0 \\
\sin(\alpha_{ij}) & \cos(\alpha_{ij}) & 0 \\
0 & 0 & 1
\end{array} \right] V_i^T,
\]
and hence
\[
\frac{1}{2}(E_{ij} + E_{ji}^T) = V_i \left[ \begin{array}{ccc}
\cos(\alpha_{ij}) & 0 & 0 \\
0 & \cos(\alpha_{ij}) & 0 \\
0 & 0 & 1
\end{array} \right] V_i^T.
\]
Therefore
\[
(\cos(\alpha_{ij}) - 1)I \leq \frac{1}{2}(E_{ij} + E_{ji}^T) - I \preceq 0,
\]
and since \( \Delta_{ii} = \sum_{j \neq i} a_{ij} \left( \frac{1}{2}(E_{ij} + E_{ji}^T) - I \right) \) we get
\[
\sum_{j \neq i} a_{ij} (\cos(\alpha_{ij}) - 1)I \preceq \Delta_{ii} \preceq 0.
\]
Thus \( \| \Delta_{ii} \| \leq \sum_{j \neq i} a_{ij} (1 - \cos(\alpha_{ij})) \). Using the identity \( 1 - \cos(\alpha_{ij}) = 2\sin^2 \left( \frac{\alpha_{ij}}{2} \right) \) and the fact that the sin-function is non-decreasing on \( [-\frac{\pi}{2}, \frac{\pi}{2}] \) gives the desired result.

Lemma 4.3. If \( i \neq j \) then
\[
\| \Delta_{ij} \| \leq 2a_{ij} \sin \left( \frac{|\alpha_{ij}|}{2} \right). \tag{47}
\]
If in addition \( 0 \leq \alpha_{\max} \leq \frac{\pi}{2} \) then
\[
\| \Delta_{ij} \| \leq 2a_{ij} \sin \left( \frac{\alpha_{\max}}{2} \right). \tag{48}
\]
Proof. To estimate the off-diagonal blocks \( \| \Delta_{ij} \| = a_{ij}\|I - E_{ij}\| \) we note that for a unit vector \( v \) we have
\[
\sqrt{\| v - E_{ij}v \|^2} = \sqrt{\| v \|^2 - 2 \cos \langle v, E_{ij}v \rangle + \| E_{ij}v \|^2} \leq \sqrt{2(1 - \cos(\alpha_{ij}))},
\]
where \( \langle v, E_{ij}v \rangle \) is the angle between \( v \) and \( E_{ij}v \). Furthermore, we will have equality if \( v \) is perpendicular to the rotation axis of \( E_{ij} \). Therefore
\[
\| \Delta_{ij} \| = a_{ij} \sqrt{2(1 - \cos(\alpha_{ij}))} \leq 2a_{ij} \sin \left( \frac{\alpha_{\max}}{2} \right). \tag{50}
\]

Summarizing the results in Lemmas 4.1-4.3 we get that the eigenvalues \( \lambda \) of \( \Delta \) fulfill
\[
|\lambda(\Delta)| \leq 2d_i \sin^2 \left( \frac{\alpha_{\max}}{2} \right) + \sum_{j \neq i} 2a_{ij} \sin \left( \frac{\alpha_{\max}}{2} \right)
\]
\[
\leq 2d_{\max} \sin \left( \frac{\alpha_{\max}}{2} \right) \left( 1 + \sin \left( \frac{\alpha_{\max}}{2} \right) \right), \tag{51}
\]
where \( d_{\max} \) is the maximal vertex degree. Note that the same bound holds for all eigenvalues of \( \Delta \), in particular, the one with the largest magnitude \( \lambda_{\max}(\Delta) \).

Now returning to our goal of showing that \( D_{R^*}(\Lambda^* - \tilde{R})D_{R^*} \preceq 0 \). Let \( N = [I \quad I \ldots I]^T \). The columns of \( N \) will be in the null space of \( D_{R^*}(\Lambda^* - \tilde{R})D_{R^*} \). Therefore \( D_{R^*}(\Lambda^* - \tilde{R})D_{R^*} \) is positive semidefinite if \( D_{R^*}(\Lambda^* - \tilde{R})D_{R^*} + \mu NN^T \) is, and hence it is enough to show that
\[
\lambda_1 \left( D_{R^*}(\Lambda^* - \tilde{R})D_{R^*} + \mu NN^T \right) \geq 0 \tag{52}
\]
for sufficiently large \( \mu \). The Laplacian \( L_G \) is positive semidefinite with smallest eigenvalue \( \lambda_1 = 0 \) and corresponding eigenvector \( v = (1, 1, \ldots, 1)^T \). Furthermore, as \( N = v \otimes I_3 \), it is clear that for sufficiently large \( \mu \) we have
\[
\lambda_1(L_G \otimes I_3 + \mu NN^T) = \lambda_1(L_G + \mu vv^T) = \lambda_2(L_G).
\]
Since \( D_{R^*}(\Lambda^* - \tilde{R})D_{R^*} + \mu NN^T = L_G \otimes I_3 + \mu NN^T + \Delta \), (53) we therefore get
\[
\lambda_2(L_G - |\lambda_{\max}(\Delta)|).
\]
If the right-hand side is positive, then so is the left-hand side. Using (51) for \( \lambda_{\max}(\Delta) \) yields the following result.

Lemma 4.4. The matrix \( \Lambda^* - \tilde{R} \) is positive semidefinite if \( 0 \leq \alpha_{\max} \leq \frac{\pi}{2} \) and
\[
\lambda_2(L_G - 2d_{\max} \sin \left( \frac{\alpha_{\max}}{2} \right) \left( 1 + \sin \left( \frac{\alpha_{\max}}{2} \right) \right) \geq 0. \tag{55}
\]

By completing squares, one obtains the equivalent condition
\[
\left( \sin \left( \frac{\alpha_{\max}}{2} \right) + \frac{1}{2} \right)^2 \leq \frac{\lambda_2(L_G)}{2d_{\max}^2} + \frac{1}{4}, \tag{56}
\]
which proves Theorem 4.1.

What these results show, is that if there is a KKT point in \((P)\) then it is also a KKT point to \((P')\). If this KKT point fulfills the prescribed error conditions it will be globally optimal in \((P')\) and strong duality holds. But a solution that is globally optimal in \((P')\) and feasible in \((P)\) will also be globally optimal in \((P)\) since the objective functions are the same. Thus, as long as there is a solution to \((P)\) with small enough errors, the programs \((P), (P')\) and \((D)\) will all yield the same objective value.

5 Problems with Unknown Graph Topology

The bound (29) on \( \alpha_{\max} \) depends on the algebraic connectivity of the underlying graph \( \lambda_2(L_G) \). It may also be of interest to have an estimate that is independent of the particular graph and only depends on the amount of missing data. In this section we show how to derive such a bound.

Again we consider a matrix \( \Delta \) defined as in (38), but with respect to the complete graph \( K_n \), that is,
\[
\Delta = D_{R^*}(\Lambda^* - \tilde{R})D_{R^*} - L_{K_n} \otimes I_3. \tag{57}
\]
The \( 3 \times 3 \) sub-blocks of \( \Delta \) can be written
\[
\Delta_{ij} = \sum_{j \neq i} a_{ij} (E_{ij} + E_{ji}^T) - (n - 1)I
\]
\[
= \sum_{j \neq i} a_{ij} (E_{ij} + E_{ji}^T - I) - (n - 1 - d_i)I \tag{58}
\]
and
\[ \Delta_{ij} = -a_{ij}x_{ij} - I. \] (59)

From Lemmas 4.2 and 4.3 we now get that
\[ \|\Delta_{ii}\| \leq 2d_i \sin^2 \left( \frac{\alpha_{\max}}{2} \right) + (n - 1 - d_i) \] (60)
and
\[ \|\Delta_{ij}\| \leq \begin{cases} \frac{1}{2} \sin \left( \frac{\alpha_{\max}}{2} \right) & a_{ij} = 0 \\ a_{ij} = 1. \end{cases} \] (61)

Lemma 4.1 shows that any eigenvalue of \( \Delta \) has a property such that
\[ |\lambda| \leq 2d_i \left( \sin^2 \left( \frac{\alpha_{\max}}{2} \right) + \sin \left( \frac{\alpha_{\max}}{2} \right) \right) + 2(n - 1 - d_i), \] (62)
for some \( i \). We get that the sufficient condition \( \lambda_2(L_{K_n}) - |\lambda_{\max}(\Delta)| \geq 0 \) for strong duality is true if \( \lambda_2(L_{K_n}) \) is greater or equal to the right-hand side of (62). As the complete graph \( K_n \) has \( \lambda_2(L_{K_n}) = n \), a sufficient condition for strong duality can be stated as follows,
\[ 1 - \frac{n}{2d_i} + \frac{1}{d_i} \geq \sin^2 \left( \frac{\alpha_{\max}}{2} \right) + \sin \left( \frac{\alpha_{\max}}{2} \right). \] (63)
Solving for \( \alpha_{\max} \) gives
\[ \alpha_{\max} \leq 2 \arcsin \left( \frac{\sqrt{\frac{5}{4} - \frac{n}{2d_i} + \frac{1}{d_i}} - 1}{2} \right). \] (64)

Figure 5 shows the right hand side as a function of the vertex-connectivity (normalized with \( n - 2 \)). As long as \( d_i/(n - 2) > 0.5 \) we get a positive bound on \( \alpha_{\max} \) that is, as long as no more than roughly half of the measurements are missing for each rotation this bound ensures that \( \Lambda^* - \tilde{R} \geq 0 \). Since this has to hold for all \( i \) and the bound is increasing with \( d_i \) we get the following result.

**Theorem 5.1** (Strong Duality with Unknown Topology). The stationary point \( R_i \), \( i = 1, \ldots, n \) will be globally optimal and strong duality will hold for (P) if
\[ |\alpha_{ij}| \leq \alpha_{\max} \quad \forall (i, j) \in E, \] (65)
where
\[ \alpha_{\max} \leq 2 \arcsin \left( \frac{\sqrt{\frac{5}{4} - \frac{n - 2}{2d_{\min}} - \frac{1}{2}}}{2} \right), \] (66)
and \( d_{\min} \) is the minimal vertex degree.

6 Chordal Error Bounds

The result in Lemma 4.4 gives a condition for strong duality that uses the worst case residual angle \( \alpha_{\max} \) and the largest degree \( d_{\max} \). While such an estimate provides an intuitive and simple sufficient condition it is clear that it could be violated, for example, in the case of outlier measurements. In this section we show that by considering all residuals, and not just the worst one, we can derive conditions that will allow measurement errors that are larger than those required by (29).

We show the following stronger version of Lemma 4.4.

**Lemma 6.1.** The matrix \( \Lambda^* - \tilde{R} \) is positive semidefinite if
\[ \sum_{j \neq i} 2a_{ij} \sin \left( \frac{\alpha_{ij}}{2} \right) \left( 1 + \sin \left( \frac{\alpha_{ij}}{2} \right) \right) \leq \lambda_2(L_G), \] (67)
for all \( i = 1, \ldots, n \).

Proof. With \( \Delta \) as in Section 4.1 we have from Lemmas 4.1, 4.2 and 4.3 that the eigenvalues of \( \Delta \) fulfill
\[ |\lambda(\Delta)| \leq \sum_{ij \neq i} 2a_{ij} \sin^2 \left( \frac{\alpha_{ij}}{2} \right) + \sum_{ij \neq i} 2a_{ij} \sin \left( \frac{\alpha_{ij}}{2} \right), \] (68)
for some \( i \). We therefore have that \( \lambda_2(L_G) \geq |\lambda_{\max}(\Delta)| \) if (67) holds. And since \( \Lambda^* - \tilde{R} \) \( \geq 0 \) when \( \lambda_2(L_G) - |\lambda_{\max}(\Delta)| \geq 0 \) this proves the result.

Since \( \|R_i \tilde{R}_{ij} - R_j\|_F = 2\sqrt{2} \sin \left( \frac{\alpha_{ij}}{2} \right) \) we see that \( \Lambda^* - \tilde{R} \) is positive semidefinite if for all \( i = 1, \ldots, n \) we have
\[ e_i \leq \lambda_2(L_G), \] (69)
where
\[ e_i = \sum_{ij \neq i} a_{ij} \left( \frac{\|R_i \tilde{R}_{ij} - R_j\|_F^2}{4} + \frac{\|R_i \tilde{R}_{ij} - R_j\|_F}{\sqrt{2}} \right). \] (70)

This bound allows individual residual errors to be larger than \( \alpha_{\max} \) as long as the sum (70) is sufficiently small.

7 Solving the Rotation Averaging Problem

The dual problem (D) is a convex semidefinite program, and although it is provably solvable in polynomial time by interior point methods [29], in practice such problems quickly become intractable as the dimension of the entering variables grow.

In this section we present a first-order method for solving semidefinite programs with constant block diagonals. Our approach solves the dual of (D) and consists of two simple matrix operations only, matrix multiplication and square roots of \( 3 \times 3 \) symmetric matrices, the latter which can be solved in closed form. Consequently, these two operations permit a simple and efficient implementation without the need for dedicated numerical libraries.

The dual of (D) is given by
\[ \min_{Y \succeq 0} \max_{\Lambda} - \text{tr}(\Lambda) + \text{tr} \left( Y(\Lambda - \tilde{R}) \right). \] (71)
Let the matrix $Y$ be partitioned as follows,

\[
Y = \begin{bmatrix}
Y_{11} & Y_{12} & \cdots & Y_{1n} \\
Y_{21} & Y_{22} & \cdots & Y_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
Y_{n1} & Y_{n2} & \cdots & Y_{nn}
\end{bmatrix}
\]  

(72)

where each block $Y_{ij} \in \mathbb{R}^{3 \times 3}$ for $i, j = 1, \ldots, n$. Since $\Lambda$ is block-diagonal (17) it is clear that the inner maximization is unbounded when $Y_{ii} - I_{3 \times 3} \neq 0$ and zero otherwise. We therefore get

\[
(DD) \min_Y -\operatorname{tr}\left( \tilde{R}Y \right) \\
\text{s.t.} \quad Y_{ii} = I_3, \quad i = 1, \ldots, n, \quad Y \succeq 0.
\]  

(73)

Since $Y \succeq 0$ it is clear that

\[-\operatorname{tr}(\Lambda) + \operatorname{tr}(Y(\Lambda - R^*)) \succeq -\operatorname{tr}(\Lambda),
\]

for all $\Lambda$ of the form (17). Therefore $(DD) \succeq (D)$ and assuming strong duality holds $(D) = (P)$. Furthermore if $R^*$ is the global optimum of $(P)$ then $Y = R^{\ast T} R^{\ast}$ is feasible in (73) which shows that $(DD) = (P)$.

Thus, when strong duality holds, recovering a primal solution to $(P)$ is then achieved by simply reading off the first three rows of $Y^*$ and choosing their signs to ensure positive determinants of the resulting rotation matrices.

### 7.1 Block Coordinate Descent

In this section we present a block coordinate descent method for solving semidefinite programs with block diagonal constraints on the form (73). This method is a generalization of the row-by-row algorithms derived in [30].

Consider the following semidefinite program,

\[
\min_{S \in \mathbb{R}^{n \times 3}} \operatorname{tr}(W^T S) \\
\text{s.t.} \quad \left[ \begin{array}{c}
I \\
\frac{1}{2} S^T_B
\end{array} \right] \succeq 0.
\]  

(74)

This is a subproblem that arises when attempting to solve (DD) in (73) using a block coordinate descent approach, i.e., by fixing all but one row and column of blocks in (72) and reordering as necessary. It turns out that this subproblem has a particularly simple, closed form solution, established by the following lemma.

**Lemma 7.1.** Let $B$ be a positive semidefinite matrix. Then, the solution to (74) is given by,

\[
S^* = -BW \left[ \left( W^T BW \right)^{\frac{1}{2}} \right]^{\dagger}.
\]  

(75)

Here $\dagger$ denotes the Moore–Penrose pseudoinverse.

**Proof.** From the Schur complement, we have that the $2 \times 2$ block matrix in (74) is positive semidefinite if and only if

\[
I - ST B^\dagger S \succeq 0, \quad (I - BB^\dagger) S = 0.
\]  

(76)

(77)

Hence the problem (74) is equivalent to

\[
\min_{S \in \mathbb{R}^{n \times 3}} \langle W, S \rangle \\
\text{s.t.} \quad I - ST B^\dagger S \succeq 0, \quad (I - BB^\dagger) S = 0.
\]  

(78a)

(78b)

(78c)

The KKT conditions for (78), with Lagrangian multipliers $\Gamma$ and $\Upsilon$, become

\[
W + 2B^\dagger ST + (I - BB^\dagger) \Upsilon = 0, \quad (I - BB^\dagger) \Upsilon = 0, \quad \Gamma \succeq 0, \quad (I - ST B^\dagger) \Gamma = 0.
\]  

(79)

(80)

(81)

(82)

(83)

Rewrite (79) and (83) as

\[
B^\dagger \Gamma = -\frac{1}{2} W - \frac{1}{2} (I - BB^\dagger) \Upsilon, \quad \Gamma^T \Gamma = \Gamma^T ST B^\dagger ST.
\]  

(84)

(85)

Since the pseudoinverse fulfills $B^\dagger B B^\dagger = B^\dagger$, combining (84) and (85) we obtain

\[
\Gamma^2 = \Gamma^T ST B^\dagger B B^\dagger ST = \frac{1}{4} \left( W + (I - BB^\dagger) \Upsilon \right)^T B \left( W + (I - BB^\dagger) \Upsilon \right) = \frac{1}{4} W^T BW.
\]  

(86)

(87)

(88)

Here the last equality follows since $B(I - BB^\dagger) = 0$. This gives

\[
\Gamma = \frac{1}{2} \left( W^T BW \right)^{\frac{1}{2}}.
\]  

(89)

Inserting (89) in (84)

\[
B^\dagger S \left( W^T BW \right)^{\frac{1}{2}} = -W - (I - BB^\dagger) \Upsilon,
\]  

(90)

multiplying with $B$ from the left on both sides and using (81), $BB^\dagger S = S$, we arrive at

\[
S \left( W^T BW \right)^{\frac{1}{2}} = -BW,
\]  

(91)

and consequently

\[
S = -BW \left[ \left( W^T BW \right)^{\frac{1}{2}} \right]^{\dagger}.
\]  

(92)

Finally, since

\[
\Gamma = \frac{1}{2} \left( W^T BW \right)^{\frac{1}{2}} \succeq 0, \quad I - ST B^\dagger S = I - \left[ \left( W^T BW \right)^{\frac{1}{2}} \right]^{\dagger} W^T BW \left[ \left( W^T BW \right)^{\frac{1}{2}} \right]^{\dagger} \succeq 0,
\]  

(93)

(94)

the conditions (80) and (82) are satisfied then (75) must be a feasible and optimal solution to (78) and consequently also to (74).

The resulting algorithm is summarised in Algorithm 1 below.

### 8 Experimental Results

In this section we present an experimental study aimed at characterizing the performance and computational efficiency of the proposed algorithm compared to existing numerical solvers.
input: \( \tilde{R}, Y^{(0)} \succeq 0, \ t = 0 \).

repeat

- Select an integer \( k \in [1, \ldots, n] \).
- \( B_k \): the result of eliminating the \( k \)th row and column from \( Y^t \).
- \( W_k \): the result of eliminating the \( k \)th row and all but the \( k \)th column from \( \tilde{R} \).
- \( S^*_k = -B_k W_k \left( W_k^T B_k W_k \right)^\dagger \) as in (75).
- \( Y^{t} = \left[ I \ S^*_k \right] \), (succeeded by the appropriate reordering).
- \( t = t + 1 \)

until convergence

Algorithm 1: A block coordinate descent algorithm for the semidefinite relaxation (DD) in (73).

8.1 Synthetic Data

In our first set of experiments we compared the computational efficiency of the following algorithms:

- The Levenberg-Marquardt (LM) algorithm [31], a standard nonlinear optimization method.
- Algorithm 1, our approach developed in Section 7.
- The Staircase method of [32], a Riemannian optimization based approach with a very efficient and well-written implementation which is publicly available.
- SeDuMi [33], a well-known software package for conic optimization.

We constructed a large number of synthetic problem instances of increasing size, perturbed by varying levels of noise. Each absolute rotation was obtained by rotation about the z-axis by \( 2\pi/n \) rad and by construction, forming a cycle graph. The relative rotations were perturbed by noise in the form of a random rotation about an axis sampled from a uniform distribution on the unit sphere with angles normally distributed with mean 0 and variance \( \sigma \). The absolute rotations were initialized (if required) in a similar fashion but with the angles uniformly distributed over \([0, 2\pi] \) rad.

The results, averaged over 100 runs, can be seen in Table 1. As expected, the LM algorithm significantly outperforms our algorithm as well as SeDuMi and the Staircase method, but it only manages to obtain the global optima in about 30–70% of the time. As predicted by Theorem 4.2 and the discussion in Section 4.1 on cycle graphs, Algorithm 1, the Staircase method and SeDuMi produce globally optimal solutions at every single problem instance, independent of the noise level and independent on the number of cameras. As expected, Algorithm 1 does appear to outperform SeDuMi quite significantly with respect to computational efficiency. From this table we also observe that our proposed approach also appears competitive with the Staircase method of [32]. It should be noted that we here solved the problem only to moderate precision (3-4 decimal places). For a greater precision we would expect the Staircase method to perform better. However, for the applications considered here, moderate accuracies are generally sufficient.

8.2 Real-World Data

In our second set of experiments we compared the computational efficiency on a number of publicly available real-
world datasets\(^1\)\(^2\). The results are presented in Table 2. Here, as in the previous experiment, both methods correctly produce the global optima at each instance. Algorithm 1 again performs comparable to the Staircase method and significantly outperforms SeDuMi with respect to computational cost, providing further evidence of the efficiency of the proposed algorithm. It can further be seen that Theorem 4.1 provides bounds sufficiently large to guarantee strong duality, and hence global optimality, in all the real-world instances except for the pumpkin and buddah datasets. Although strong duality does indeed hold in both of these cases, the resulting certificate is insufficient with respect to the angle bound (29) for the pumpkin dataset and both the angle and chordal bounds for the buddah dataset. The visibility graphs of both of these datasets are comprised both of densely as well as sparsely connected cameras, resulting in a large value of \(d_{\text{max}}\) in combination with a small value of \(d_{\text{min}}\) (minimum degree). Since \(\lambda_2 \leq d_{\text{min}}\), a limited bound on \(\alpha_{\text{max}}\) follows directly from (29). These instances serve as a representative example of when the bounds of Theorem 4.1, although still valid and strictly positive, become too conservative in practice.

9 Conclusions

In this paper we have presented a theoretical analysis of Lagrangian duality in rotation averaging based on spectral graph theory. Our main result states that for this class of problems strong duality will provably hold between the primal and dual formulations if the noise levels are sufficiently restricted. In many cases the noise levels required for strong duality not to hold can be shown to be quite severe. To the best of our knowledge, this is the first time such practically useful sufficient conditions for strong duality have been established for optimization over multiple rotations.

A scalable first-order algorithm, a generalization of coordinate descent methods for semidefinite cone programming, was also presented. Our empirical validation demonstrates the potential of this proposed algorithm, significantly outperforming existing general purpose semidefinite numerical solvers and performing comparable, in many instances even better than, state-of-the-art methods. We found this result particularly encouraging as, owing to its simplicity, our proposed approach would be exceedingly straightforward to parallelise, with potentially considerable speed-ups as a consequence.

Acknowledgments

This work has been funded by the Australian Research Council through grant FT170100072, the Swedish Research Council (no. 2016-04445 and 2018-05375), the Swedish Foundation for Strategic Research (Semantic Mapping and Visual Navigation for Smart Robots) and Vinnova / FFI (Perception, no. 2017-01942).

References


\(^1\) Available from http://www.maths.lth.se/~calle

\(^2\) Theorem 4.1 provides bounds sufficiently large to guarantee strong duality, and hence global optimality, in all the real-world instances except for the pumpkin and buddah datasets. Although strong duality does indeed hold in both of these cases, the resulting certificate is insufficient with respect to the angle bound (29) for the pumpkin dataset and both the angle and chordal bounds for the buddah dataset. The visibility graphs of both of these datasets are comprised both of densely as well as sparsely connected cameras, resulting in a large value of \(d_{\text{max}}\) in combination with a small value of \(d_{\text{min}}\) (minimum degree). Since \(\lambda_2 \leq d_{\text{min}}\), a limited bound on \(\alpha_{\text{max}}\) follows directly from (29). These instances serve as a representative example of when the bounds of Theorem 4.1, although still valid and strictly positive, become too conservative in practice.

\[^{1}\] Available from http://www.maths.lth.se/~calle


