

# Performance Guarantees for Spectral Initialization in Rotation Averaging and Pose-Graph SLAM

Kevin J. Doherty<sup>\*1</sup>, David M. Rosen<sup>2</sup>, and John J. Leonard<sup>1</sup>

<sup>1</sup>*Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

<sup>2</sup>*Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

## Abstract

In this work, we describe an initialization method based on *spectral relaxation* for rotation averaging and pose-graph optimization in the setting of simultaneous localization and mapping (SLAM). These optimization problems are typically both large-scale and nonconvex. In consequence, efficient methods for computing reasonable initial estimates are often used to enhance the speed and performance of the algorithms underlying these robot perception problems. To this end, we present an estimator based on spectral relaxation tailored to the rotation averaging and SLAM problems. Critically, we also present a theoretical analysis that controls both the error of this estimate with respect to the latent ground truth signal as well as its distance to the corresponding *global minimizer* of the problem. This is, to our knowledge, *the first theoretical analysis of its kind*. The form of our bounds reveals the spectral properties of the measurement graphs to be central objects of interest in controlling these quantities. Moreover, our analysis controlling the distance of an initialization from the corresponding global minimizer is achieved by deriving new error bounds for global minimizers, which are likely to be of independent interest. Finally, our empirical results suggest that spectral estimates typically perform far better than the worst-case analysis suggests, producing solutions competitive with state-of-the-art initialization techniques in solution quality and computation time.

## 1 Introduction

Simultaneous localization and mapping (SLAM), the process by which a robot jointly infers its pose and the location of environmental landmarks, is a fundamental capability of mobile robots, supporting navigation, planning, and control [25]. Modern procedures by which SLAM solutions are recovered typically rely on nonconvex optimization. This nonconvexity manifests due to the constraint that robot rotation estimates lie in the nonconvex special orthogonal group  $\text{SO}(d)$ . In consequence, the quality of a recovered SLAM estimate depends upon the quality of the initial guess, and a great deal of research has been dedicated to initialization techniques (see Carlone et al. [9] for a review). While these techniques have been shown to work well in practice, the reasons for their empirical success are currently poorly understood. As a result, it is difficult to say when, or under what conditions, these techniques can be reliably deployed.

---

<sup>\*</sup>Corresponding author. Email: [kjd@csail.mit.edu](mailto:kjd@csail.mit.edu)

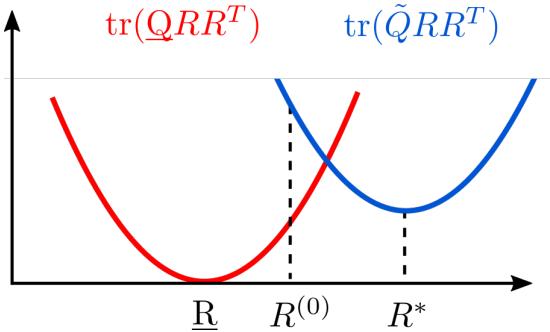


Figure 1: **Comparing true, optimal, and initial rotation estimates.** We are interested in bounds on the deviation of an initial estimate  $R^{(0)}$  from the (latent) ground truth  $\underline{R}$  and the globally optimal solution  $R^*$ .

This work provides *the first formal guarantees* on the error of an initialization method. Our analysis gives direct control over the error of a *spectral initialization* in terms of the spectral properties of the measurement graph<sup>1</sup>. In turn, this allows us to control the distance of the spectral estimate to the corresponding global minimizer, which is critical for ensuring that the initial guess is in the correct basin of attraction (see Figure 1 for an overview). Our proof of the latter result relies on new estimation error bounds for global optima, which are likely to be of independent interest. Our empirical results on both synthetic data and standard pose-graph SLAM benchmarks demonstrate that the spectral estimator typically performs far better than our worst-case analysis suggests, achieving solution quality and computation times competitive with state-of-the-art approaches. We also provide the first empirical results demonstrating that in many practical cases the spectral initialization method allows us to recover (verifiable) *globally optimal* solutions by straightforward nonlinear optimization (i.e. without the need to perform the semidefinite relaxation of [24]).<sup>2</sup> Beyond the immediate interpretation of these results as an analysis of an *initialization method* for optimization, our theoretical results reveal the spectral relaxation to be a cheap method for computing rotation averaging and pose-graph optimization solutions on its own (without the need to perform nonconvex optimization or semidefinite relaxation) that attains similar worst-case performance guarantees to the globally optimal estimate.

The remainder of the paper proceeds as follows: In Section 2, we discuss related literature on robot perception and rotation synchronization. In 3 we provide general notation and preliminaries, and Section 4 describes the spectral initialization procedure. In Section 5 we present our main theoretical results: an analysis controlling the estimation error of both the spectral initialization method and that of globally optimal estimates for the rotation averaging and pose-graph SLAM problems, as well as a corresponding bound on the distance between the spectral initialization and the corresponding globally optimal solution. Section 6 demonstrates empirical performance of the spectral initialization technique on benchmark SLAM datasets as well as the theoretical performance bound evaluated on synthetic data and shows, in particular, that the spectral estimator is competitive with state-of-the-art techniques for initialization.

<sup>1</sup>The spectral properties of measurement graphs have consistently emerged as key quantities controlling the performance of estimators for these problems, though this connection (particularly in the context of SLAM) remains under-explored (cf. [25] for a recent review).

<sup>2</sup>While our empirical results show that spectral initialization enables the recovery of globally optimally optimal solutions, the *verification* procedure used to determine global optimality is based on the properties of the corresponding semidefinite relaxation [24].

## 2 Related work

Simultaneous localization and mapping and rotation averaging problems are often formulated as high-dimensional, nonconvex optimization problems. Consequently, solving these problems typically requires efficient algorithms for producing an “initial guess.” Historically, research on this topic has focused on developing cheap, but typically inexact, convex or linear relaxations of the SLAM (resp. rotation averaging) problems (cf. [9, 20]). While these techniques often work well in practice, the *reasons* for their empirical success remain poorly understood, and it is difficult to assess *under which conditions* these techniques can be reliably deployed.

A related line of research is concerned with the development of Cramér-Rao bounds, which provide lower bounds on the expected mean-squared error of unbiased estimators for these problems. Our analysis complements previous work in this area [6, 11]. Such bounds control the *best* possible error of an estimator *in expectation*. In contrast, the bounds we present control the *worst-case* error on a *per instance* basis. Interestingly, our analysis, consistent with prior work, reveals the spectral properties of the graph parameterizing the estimation problem under consideration to be central objects of interest in controlling the estimation error (cf. [6, 11, 17]).

The spectral relaxation approach to initialization that we consider has previously appeared in other problem settings, particularly in the area of phase synchronization problems (cf. [4, 5, 18, 26]). In particular, Ling [18] describe theoretical results that are qualitatively similar to those described in this paper, though theirs are concerned specifically with *orthogonal* group synchronization problems. Liu et al. [19] take a similar approach to ours in order to derive error bounds for spectral estimators of synchronization problems defined over subgroups of the orthogonal group (including  $\text{SO}(d)$ ), though the form of the bounds we develop for  $\text{SO}(d)$  synchronization (rotation averaging) makes the connection between estimation error and the spectral properties of the measurement graphs that arise in robot perception problems more explicit.

Recently, Moreira et al. [21] proposed a computationally-efficient Krylov-Schur decomposition approach for computing pose-graph SLAM estimates. Their method is formally equivalent to the *rotation-only* variant of the spectral initialization procedure considered here in a way that can be made precise (see Appendix C for a full derivation), but our construction arises more directly and simply from convex relaxation. Arrigoni et al. [1] describe a spectral method for  $\text{SE}(d)$ -synchronization. A similar theoretical analysis to ours can be carried out for their approach, but the form of the relaxation they consider leads to more complicated bounds due to a dependence on the absolute translation scale. Finally, Boots and Gordon [3] consider spectral techniques for the range-only SLAM problem. Though their problem setting differs from the one considered here, extension of the techniques presented in this work to scenarios with different types of measurement models is an interesting area for future work.

Recently, *certifiably-correct* machine perception has emerged as a key area of interest to the robotics community, resulting in the development of algorithms capable of recovering globally optimal solutions in certain noise regimes [7, 8, 10, 12, 13, 24, 28]. Our analysis gives direct control over the estimation error for these solutions in terms of the magnitude of the noise corrupting the measurements. Moreover, the bounds we present suggest that these estimators, which are often based on large-scale semidefinite relaxations, admit similar *worst-case* performance guarantees to an estimate computed using the proposed spectral method. While it may be difficult to compute (or impossible to verify) a solution using the former methods, the spectral estimator is always available, and as we show, can typically be computed inexpensively.

### 3 Preliminaries and formulation

#### 3.1 Notation and preliminaries

**Lie groups:** We will make use of the matrix realizations of several Lie groups, most prominently the  $d$ -dimensional special Euclidean and special orthogonal groups, denoted  $\text{SE}(d)$  and  $\text{SO}(d)$ , respectively. The realization of  $\text{SE}(d)$  as a matrix group can be defined as follows:

$$\text{SE}(d) \triangleq \left\{ \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{(d+1) \times (d+1)} \mid R \in \text{SO}(d), t \in \mathbb{R}^d \right\}, \quad (1)$$

and the group  $\text{SO}(d)$  can be realized as:

$$\text{SO}(d) \triangleq \left\{ R \in \mathbb{R}^{d \times d} \mid R^T R = I_d, \det(R) = 1 \right\}. \quad (2)$$

**Linear algebra:** For a symmetric matrix  $S$ ,  $S \succeq 0$  denotes that  $S$  is positive-semidefinite. The eigenvalues of a symmetric matrix  $S \in \mathbb{R}^{n \times n}$  are denoted  $\lambda_1(S) \leq \lambda_2(S) \leq \dots \leq \lambda_n(S)$ . We will also consider several block-structured matrices and make use of a few special operators acting on block-structured matrices. Following the notation of Rosen et al. [24], given square matrices  $A_i \in \mathbb{R}^{d \times d}$ ,  $i = 1, \dots, n$ , we let  $\text{Diag}(A_1, \dots, A_n)$  denote the matrix with each  $A_i$  along the diagonal (i.e. the matrix direct sum). Furthermore, given a block-structured matrix  $B$ , let  $\text{BlockDiag}_d(B)$  denote the operator extracting the  $d \times d$  block-diagonal entries of  $B$ . Finally, let  $\text{SBD}(d, n)$  denote the set of  $dn \times dn$  symmetric block-diagonal matrices with diagonal blocks of size  $d \times d$ , and  $\text{SymBlockDiag}_d(A)$  be the operator extracting the symmetric part of the  $d \times d$  block-diagonal elements of  $A$ .

**Probability and statistics:** We denote the multivariate Gaussian distribution with mean  $\mu \in \mathbb{R}^d$  and covariance  $\Sigma \in \mathbb{S}^d$  as  $\mathcal{N}(\mu, \Sigma)$ . We denote the isotropic Langevin distribution on  $\text{SO}(d)$  with mode  $M \in \text{SO}(d)$  and concentration parameter  $\kappa \geq 0$  as  $\text{Langevin}(M, \kappa)$ . The probability density function for the isotropic Langevin distribution is given as:

$$p(R; M, \kappa) = \frac{1}{c_d(\kappa)} \exp(\kappa \text{tr}(M^T R)), \quad (3)$$

where  $c_d(\kappa)$  is a normalization constant.

Finally, for an unknown variable  $Z$  we aim to infer, we denote its true latent value by  $Z$  and  $\tilde{Z}$  denote a noisy measurement of  $Z$ .

**Gauge-invariant distance metrics:** A key property of the problems under consideration is *gauge symmetry*. Synchronization problems over  $\text{SO}(d)$  and  $\text{SE}(d)$  with  $d \geq 2$  admit infinitely many solutions. In particular, we will see that if  $R^* \in \text{SO}(d)^n$  is an optimal estimate of the rotational states, then  $GR^*$  is also optimal for any  $G \in \text{SO}(d)$ . We therefore define the following *orbit distances* in order to compare solutions to the problems at hand:

$$d_{\mathcal{S}}(X, Y) \triangleq \min_{G \in \text{SO}(d)} \|X - GY\|_F, \quad X, Y \in \text{SO}(d)^n \quad (4a)$$

$$d_{\mathcal{O}}(X, Y) \triangleq \min_{G \in \text{O}(d)} \|X - GY\|_F, \quad X, Y \in \text{O}(d)^n. \quad (4b)$$

It will be convenient to “overload” the  $\text{O}(d)$  orbit distance to act on elements of the Stiefel manifold  $\text{St}(d, dn)$ . That is, for  $X, Y \in \text{St}(d, dn)$ :

$$d_{\mathcal{O}}(X, Y) \triangleq \min_{G \in \text{O}(d)} \|X - GY\|_F. \quad (5)$$

Each of these distances can be computed in closed form (cf. Rosen et al. [24]).

### 3.2 Problem formulation

We consider the problem of synchronization over the  $\text{SO}(d)$  group: this is the problem of estimating  $n$  unknown values  $R_1, \dots, R_n \in \text{SO}(d)$  given a set of noisy measurements  $\tilde{R}_{ij}$  of a subset of their pairwise relative rotations. The problem of  $\text{SO}(d)$ -synchronization captures, in particular, the problems of rotation averaging and, under common modeling assumptions, pose graph optimization, where the variables of interest are the orientations of a robot (or more generally, a rigid body) at different points in time (see, for example Grisetti et al. [14]).

This problem possesses a natural graphical structure  $\mathcal{G} \triangleq (\mathcal{V}, \vec{\mathcal{E}})$ , where nodes  $\mathcal{V}$  correspond to latent variables  $R_i \in \text{SO}(d)$  and edges  $(i, j) \in \vec{\mathcal{E}}$  correspond to (noisy) measured relative rotations  $\tilde{R}_{ij}$  between  $R_i$  and  $R_j$ . In particular, for the problem of *rotation averaging*, we adopt the following generative model for rotation measurements for each edge  $(i, j) \in \vec{\mathcal{E}}$ :

$$\tilde{R}_{ij} = R_{ij} R_{ij}^\epsilon, \quad R_{ij}^\epsilon \sim \text{Langevin}(I_d, \kappa_{ij}). \quad (6)$$

Given a set of noisy pairwise relative rotations  $\tilde{R}_{ij}$  sampled according to the generative model (6), a maximum-likelihood estimate  $R^* \in \text{SO}(d)^n$  for the latent rotational states  $R_1, \dots, R_n$  is obtained as a minimizer of the following problem (cf. [12, 24]):

**Problem 1** (Maximum-likelihood estimation for rotation averaging).

$$p_{MLE}^* = \min_{R_i \in \text{SO}(d)} \sum_{(i, j) \in \vec{\mathcal{E}}} \kappa_{ij} \|R_j - R_i \tilde{R}_{ij}\|_F^2. \quad (7)$$

For pose-graph SLAM ( $\text{SE}(d)$ -synchronization), we adopt the following generative model for rotation and translation measurements: For each edge  $(i, j) \in \vec{\mathcal{E}}$ :

$$\tilde{R}_{ij} = R_{ij} R_{ij}^\epsilon, \quad R_{ij}^\epsilon \sim \text{Langevin}(I_d, \kappa_{ij}) \quad (8a)$$

$$\tilde{t}_{ij} = \underline{t}_{ij} + t_{ij}^\epsilon, \quad t_{ij}^\epsilon \sim \mathcal{N}(0, \tau_{ij}^{-1} I_d), \quad (8b)$$

where  $x_{ij} = (t_{ij}, R_{ij})$  is the true value of  $x_{ij}$ . That is,  $x_{ij} = x_i^{-1} x_j$  given the true values of poses  $x_i$  and  $x_j$ . Under this noise model, the typical nonlinear least-squares formulation of  $\text{SE}(d)$  synchronization is written as follows:

**Problem 2** (Maximum-likelihood estimation for  $\text{SE}(d)$  synchronization).

$$p_{MLE}^* = \min_{\substack{t_i \in \mathbb{R}^d \\ R_i \in \text{SO}(d)}} \sum_{(i, j) \in \vec{\mathcal{E}}} \kappa_{ij} \|R_j - R_i \tilde{R}_{ij}\|_F^2 + \tau_{ij} \|t_j - t_i - R_i \tilde{t}_{ij}\|_2^2, \quad (9)$$

where  $\tilde{t}_{ij}$  are (noisy) measured translations between vehicle poses  $i$  and  $j$  (having corresponding measurement precision  $\tau_{ij}$ ), and  $\tilde{R}_{ij}$  are (noisy) measured rotations between vehicle poses  $i$  and  $j$  (having corresponding measurement precision  $\kappa_{ij}$ ).

Under these modeling assumptions, both pose-graph optimization and rotation averaging can be written in the following form:

**Problem 3** (Rotation Synchronization).

$$p^* = \min_{R \in \text{SO}(d)^n} \text{tr}(\tilde{Q} R^T R), \quad (10)$$

where  $\tilde{Q} \in \text{Sym}(dn)$ ,  $\tilde{Q} \succeq 0$ . In particular, we have:

$$\tilde{Q} = L(\tilde{G}^\rho) \quad (\text{RA})$$

$$\tilde{Q} = L(\tilde{G}^\rho) + \underbrace{\tilde{\Sigma} - \tilde{V}^\top L(W^\tau)^\dagger \tilde{V}}_{\tilde{Q}^\tau}. \quad (\text{PGO})$$

For our theoretical results, the specific structure of  $\tilde{Q}$  is not important, and indeed they would apply so long as in the *noiseless* case where  $\tilde{Q} = Q$ , we have that  $R^\top \in \ker(Q)$  where  $R$  is the latent ground-truth assignment to the rotation estimates. Nonetheless, the complete definition of  $\tilde{Q}$  follows.

Here  $L(W^\tau)$  denotes the Laplacian of the translation weight graph  $W^\tau \triangleq (\mathcal{V}, \mathcal{E}, \{\tau_{ij}\})$  with *undirected edges*  $\{i, j\} \in \mathcal{E}$ , which is an  $n \times n$  matrix with  $i, j$ -entry:

$$L(W^\tau)_{ij} = \begin{cases} \sum_{e \in \delta(i)} \tau_e, & i = j, \\ -\tau_{ij}, & \{i, j\} \in \mathcal{E}, \\ 0, & \{i, j\} \notin \mathcal{E} \end{cases} \quad (12)$$

Similarly,  $L(\tilde{G}^\rho)$  denotes the *connection Laplacian* for the rotational measurements, which is a  $dn \times dn$  symmetric block-diagonal matrix with  $d \times d$  blocks determined by:

$$L(\tilde{G}^\rho)_{ij} \triangleq \begin{cases} d_i^\rho I_d, & i = j, \\ -\kappa_{ij} \tilde{R}_{ij}, & \{i, j\} \in \mathcal{E}, \\ 0_{d \times d}, & \{i, j\} \notin \mathcal{E}, \end{cases} \quad (13a)$$

$$d_i^\rho \triangleq \sum_{e \in \delta(i)} \kappa_e, \quad (13b)$$

where  $\delta(i)$  denotes the set of edges *incident to* node  $i$ .  $\tilde{V} \in \mathbb{R}^{n \times dn}$  denotes the  $(1 \times d)$ -block-structured matrix with  $(i, j)$  block given by:

$$\tilde{V}_{ij} \triangleq \begin{cases} \sum_{e \in \delta^-(j)} \tau_e \tilde{t}_e^\top, & i = j, \\ -\tau_{ji} \tilde{t}_{ji}^\top, & (j, i) \in \vec{\mathcal{E}}, \\ 0_{1 \times d}, & \text{otherwise.} \end{cases} \quad (14)$$

Finally, let  $\tilde{\Sigma} \in \text{SBD}(d, n)$  denote the symmetric block-structured diagonal matrix given by:

$$\begin{aligned} \tilde{\Sigma} &\triangleq \text{Diag}(\tilde{\Sigma}_1, \dots, \tilde{\Sigma}_n) \in \text{SBD}(d, n) \\ \tilde{\Sigma}_i &\triangleq \sum_{e \in \delta^-(i)} \tau_e \tilde{t}_e \tilde{t}_e^\top, \end{aligned} \quad (15)$$

where  $\delta^-(i)$  denotes the set of edges *leaving* node  $i$ .

## 4 Spectral methods for initialization

The nonconvexity of the  $\text{SO}(d)$  constraint renders Problem 3 computationally hard to solve in general. Spectral methods for synchronization simplify the problem by relaxing the  $\text{SO}(d)$  constraint to the constraint  $YY^\top = nI_d$ ,  $Y \in \mathbb{R}^{d \times dn}$ :

---

**Algorithm 1** Spectral Initialization procedure

---

**Input:** The data matrix  $\tilde{Q}$  from (RA) or (PGO)  
**Output:** A spectral initialization  $R^{(0)}$

- 1: **function** SPECTRALINITIALIZATION( $\tilde{Q}$ )
- 2:   Compute the  $d$  eigenvectors  $\Phi$  of  $\tilde{Q}$  with smallest eigenvalues                       $\triangleright$  Solve Problem 4.
- 3:   **for**  $i = 1, \dots, n$  **do**
- 4:     Set  $R_i^{(0)} \leftarrow \Pi_S(\Phi_i)$ , where  $\Phi_i$  is the  $i$ -th ( $d \times d$ ) block of  $\Phi$ .                       $\triangleright$  Definition 1
- 5:     **end for**
- 6:   **return**  $R^{(0)}$
- 7: **end function**

---

**Problem 4** (Spectral Relaxation of the Rotation Synchronization Problem).

$$\begin{aligned} p_S^* &= \min_{Y \in \mathbb{R}^{d \times dn}} \text{tr}(\tilde{Q} Y^\top Y) \\ \text{s.t. } &YY^\top = nI_d. \end{aligned} \tag{16}$$

These methods are termed *spectral* relaxations due to their connection to the eigenvectors of the data matrix  $\tilde{Q}$ . In particular, it is straightforward to verify that an estimate  $Y^*$  is a minimizer for Problem 3 if and only if  $Y^*$  consists of the  $d$  eigenvectors corresponding to the minimum  $d$  eigenvalues of  $\tilde{Q}$ .

For the noiseless problem parameterized by  $\underline{Q}$ , the relaxation in Problem 4 is exact in the sense that  $\underline{R} = Y^*$ . This follows from the fact that by construction, the ground truth rotations  $\underline{R}^\top$  lie in  $\ker(\underline{Q})$ ,<sup>3</sup> and  $\underline{R}\underline{R}^\top = nI_d$  since  $\underline{R} \in \text{SO}(d)^n$ . In general, however, we do not expect such a straightforward relationship between a minimizer of the spectral relaxation in Problem 4 and a corresponding minimizer of Problem 3. In such cases where exactness does not hold, we can *round* the estimate provided by the spectral relaxation to an approximate solution  $R^{(0)} \in \text{SO}(d)^n$  in the feasible set of Problem 3. The following definition makes this precise.

**Definition 1** (Projection onto  $\text{SO}(d)$ ). Let  $X \in \mathbb{R}^{d \times d}$ , the projection  $\Pi_S(X)$  of  $X$  onto  $\text{SO}(d)$  is by definition a minimizer of the following:

$$\Pi_S(X) \triangleq \underset{G \in \text{SO}(d)}{\operatorname{argmin}} \|X - G\|_F. \tag{17}$$

A minimizer for this problem is given in closed-form as (cf. [15, 29]):

$$\Pi_S(X) = U\Sigma V^\top. \tag{18}$$

where  $U\Sigma V^\top$  is a singular value decomposition of  $X$ , and  $\Sigma$  is the matrix:

$$\Sigma = \text{Diag}(1, 1, \det(UV^\top)) \tag{19}$$

In the context of subsequent derivations, it will be convenient to “overload” this rounding operation to  $Y \in \mathbb{R}^{d \times dn}$  as follows:

$$\Pi_S(Y) = (\Pi_S(Y_1), \dots, \Pi_S(Y_n)), \tag{20}$$

where  $Y_i \in \mathbb{R}^{d \times d}$  are the  $n$  blocks of  $Y$ .

Therefore, we can obtain an approximate solution to Problem 3 from a minimizer  $Y^*$  of the relaxation in Problem 4 as  $R^{(0)} \triangleq \Pi_S(Y^*)$ . This entire procedure is summarized in Algorithm 1.

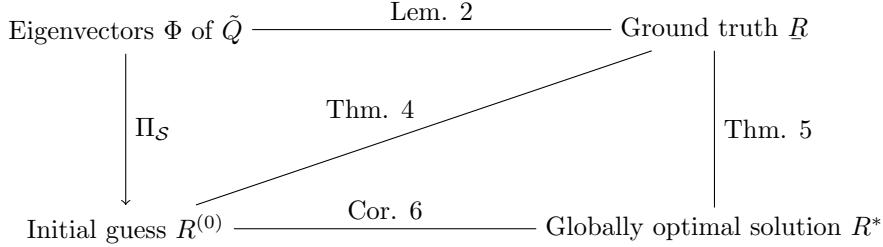


Figure 2: **Guide to the theoretical bounds.** This figure presents a diagrammatic guide to the bounds presented in Section 5. In particular, Lemma 2 gives a bound on the orbit distance between the eigenvectors of the data matrix and the ground truth. We then use this result in Theorem 4, giving a bound on the deviation of the spectral initialization from the ground truth. In Theorem 5 we bound the deviation of a *globally optimal solution* to the maximum likelihood estimation problems for rotation averaging and pose-graph SLAM from the ground truth. Finally, relating these bounds we obtain Corollary 6, which bounds the distance between a spectral initialization and the globally optimal solution.

## 5 Main results

This section presents our main theoretical results, which are three-fold: First, we provide a bound on the error of an initial guess  $R^{(0)}$  with respect to the (latent) ground-truth rotations  $R$ . Second, we give a new bound on the error of *globally optimal* solutions  $R^*$  with respect to  $R$ . The latter bound differs from prior work (e.g. Preskitt [22], Rosen et al. [24]) in that it is defined with respect to the orbit distance on  $\text{SO}(d)^n$ . Previous work used the orbit distance on  $\text{O}(d)^n$  due to mathematical convenience. However, the estimation error one considers in application is over  $\text{SO}(d)^n$ , since this is the domain on which the estimation problem is *actually* defined. Combining these results, we obtain an upper bound on the  $\text{SO}(d)$  orbit distance between an initial guess  $R^{(0)}$  and a globally optimal solution  $R^*$ . Our analysis gives direct control over the mutual deviation between the three quantities of interest:  $R^{(0)}$ ,  $R^*$ , and  $R$  as a function of the noise magnitude. We conclude with additional remarks about computing these bounds for practical SLAM scenarios and a few straightforward adaptations of the main results. Figure 2 gives an overview of the main theoretical results and the quantities they relate.

Recall from Problem 4 that an estimate  $Y \in \mathbb{R}^{d \times dn}$  is a minimizer of Problem 4 if and only if it is comprised of a set of the minimum  $d$  eigenvectors<sup>4</sup> of  $\tilde{Q}$ , and that in the noiseless case this minimizer is given by  $R$ , the minimum  $d$  eigenvectors of  $\underline{Q}$ . Since a spectral initialization  $R^{(0)}$  is obtained as the block-wise projection of the solution  $\Phi$  to Problem 4 onto  $\text{SO}(d)$ , we can bound the estimation error  $d_S(R, R^{(0)})$  by first bounding the deviation of a  $\Phi$  from  $R$ , and subsequently bounding the additional error incurred by the projection of each estimate onto  $\text{SO}(d)^n$ .

We can address the problem of bounding the deviation of a solution  $\Phi$  to Problem 4 from the solution  $R$  in the noiseless case, via the Davis-Kahan Theorem [30], a classical result relating the perturbation of a matrix's eigenvectors under a symmetric perturbation to the magnitude of the perturbation itself. Here, we take  $\underline{Q}$  to be the matrix under consideration, and define the perturbation  $\Delta Q \triangleq \tilde{Q} - \underline{Q}$ . The following lemma gives the desired characterization, a detailed proof for which is provided in Appendix A:

**Lemma 2.** *Let  $\Phi$  be a minimizer of Problem 4 and  $R$  be the corresponding ground truth rotations.*

<sup>3</sup>We refer the reader to [24, Appendix C.3] for detailed analysis of the noiseless case.

<sup>4</sup>By “minimum  $d$  eigenvectors” we mean the eigenvectors corresponding to the minimum  $d$  eigenvalues.

Then the estimation error of  $\Phi$  satisfies:

$$d_{\mathcal{O}}(\underline{R}, \Phi) \leq \frac{2\sqrt{2dn}\|\Delta Q\|_2}{\lambda_{d+1}(Q)}. \quad (21)$$

Lemma 2 provides control over the deviation of an “unrounded” solution  $\Phi$  from the  $\underline{R}$ . The second technical ingredient we require is the following simple bound controlling the maximum distance between a matrix  $X$  and its projection  $\Pi_{\mathcal{S}}(X)$  onto  $\text{SO}(d)$ :

**Lemma 3.** *Let  $X \in \mathbb{R}^{d \times d}$  and  $R \in \text{SO}(d)$ :*

$$\|\Pi_{\mathcal{S}}(X) - R\|_F \leq 2\|X - R\|_F. \quad (22)$$

*Proof.*

$$\|\Pi_{\mathcal{S}}(X) - R\|_F = \|\Pi_{\mathcal{S}}(X) - X + X - R\|_F \quad (23)$$

$$\leq \|\Pi_{\mathcal{S}}(X) - X\|_F + \|X - R\|_F \quad (24)$$

$$\leq 2\|X - R\|_F, \quad (25)$$

where the last inequality follows from the fact that  $\Pi_{\mathcal{S}}(X)$  is a minimizer over  $\text{SO}(d)$  of the distance to  $X$  with respect to the Frobenius norm, and that, by hypothesis,  $R \in \text{SO}(d)$ .  $\square$

Lemma 3 provides a straightforward approach for converting a bound of the form (21) on the error of  $\Phi$  into the desired bound on  $\Pi_{\mathcal{S}}(\Phi)$ . In turn, we obtain the following theorem (a detailed proof of which is given in Appendix B.1)

**Theorem 4.** *Let  $\Phi$  be a minimizer of Problem 4 and  $R^{(0)} = \Pi_{\mathcal{S}}(\Phi) \in \text{SO}(d)^n$  be the corresponding spectral initialization. Finally, let  $R \in \text{SO}(d)^n$  be set of ground truth rotations in Problem 3. Then the estimation error of  $R^{(0)}$  satisfies:*

$$d_{\mathcal{S}}(\underline{R}, R^{(0)}) \leq \frac{4\sqrt{2dn}\|\Delta Q\|_2}{\lambda_{d+1}(Q)}. \quad (26)$$

The bound (26) gives a direct (linear) relationship between the magnitude of the perturbation  $\Delta Q$  and the worst-case error of a spectral estimate. Moreover, Theorem 4 suggests that  $\Delta Q \rightarrow 0$  implies  $d_{\mathcal{S}}(\underline{R}, R^{(0)}) \rightarrow 0$ . That is to say, as the measurements approach their “noiseless” counterparts, we recover the latent ground-truth signal.

Next, we address the issue of furnishing a bound on  $d_{\mathcal{S}}(\underline{R}, R^*)$ . The following theorem gives the desired bound, which we prove in Appendix B.2:

**Theorem 5** (Bounding the estimation error for  $R^*$ ). *Let  $R^*$  be a minimizer of Problem 3 and  $R$  be the set of ground-truth rotations. Then the estimation error of  $R^*$  satisfies:*

$$d_{\mathcal{S}}(\underline{R}, R^*) \leq \frac{8\sqrt{dn}\|\Delta Q\|_2}{\lambda_{d+1}(Q)}. \quad (27)$$

We are not aware of any prior bounds that control the estimation error of a maximum-likelihood estimate  $R^*$  over  $\text{SO}(d)^n$  specifically. Prior work considered the estimation error over  $\text{O}(d)^n$  [2, 18, 24]. In our application, however, we are specifically concerned with the estimation error over  $\text{SO}(d)^n$ ; as one can see from inspection, this is the domain on which Problem 3 is defined.

While Theorem 4 establishes error bounds for the spectral estimator, when viewed as an *initialization method*, the distance between the initial guess  $R^{(0)}$  and the globally optimal solution is of central importance. A corollary to Theorems 4 and 5, which we prove in Appendix B.3, allows us to control  $d_{\mathcal{S}}(R^{(0)}, R^*)$  in terms of the noise matrix  $\Delta Q$ . We have:

**Corollary 6.** *The distance between the initialization  $R^{(0)}$  and the globally optimal solution  $R^*$  satisfies:*

$$d_S(R^{(0)}, R^*) \leq \frac{(8 + 4\sqrt{2})\sqrt{dn}\|\Delta Q\|_2}{\lambda_{d+1}(Q)}. \quad (28)$$

These bounds define a clear relationship between the spectral properties of  $\underline{Q}$  and  $\Delta Q$  and the deviation between a spectral estimator  $R^{(0)}$ , maximum-likelihood estimator  $R^*$ , and the ground-truth  $R$ . An important consequence of these bounds is that as  $\Delta Q \rightarrow 0$ , we have (at least<sup>5</sup>) linear convergence of the estimation error for *both* the spectral estimator and the maximum-likelihood estimator to zero. This, in turn, guarantees that  $\Delta Q \rightarrow 0$  implies  $R^*, R^{(0)} \rightarrow R$  (up to symmetry), which is what we would expect.

In practice, however, we do not have access to  $\underline{Q}$ . This presents some difficulty in the computation of  $\Delta Q$  and  $\lambda_{d+1}(Q)$ . Fortunately, the noiseless rotation matrices admit a description in terms of quantities that *are* typically assumed to be known. In particular, we have (cf. [24, Lemma 8]):

$$L(W^\rho) \otimes I_d = SL(G^\rho)S^{-1}, \text{ for } S = \text{Diag}(\underline{R}_1, \dots, \underline{R}_n) \in \mathbb{R}^{d \times dn} \quad (29a)$$

$$\lambda_{d+1}(L(G^\rho)) = \lambda_2(L(W^\rho)). \quad (29b)$$

Now,  $L(W^\rho)$  depends only on the concentration parameters  $\kappa_{ij}$  attached to each edge, which are generally assumed to be known *a priori* from the noise model. In the rotation averaging case, we have  $\underline{Q} = L(G^\rho)$ , and therefore the denominator  $\lambda_{d+1}(Q)$  is readily available as  $\lambda_2(L(W^\rho))$ , the algebraic connectivity of the rotational weight Laplacian.

In the case of pose-graph SLAM, where the matrix  $\underline{Q}$  contains the translational terms  $Q^\tau$ , we can make use of the fact that  $\underline{Q} = L(G^\rho) + Q^\tau$  is the sum of positive-semidefinite matrices, so that  $\lambda_{d+1}(L(G^\rho)) \leq \lambda_{d+1}(L(G^\rho) + Q^\tau) = \lambda_{d+1}(\underline{Q})$ . In particular, the (weaker) bounds obtained by substituting  $\lambda_{d+1}(\underline{Q})$  with  $\lambda_{d+1}(L(G^\rho))$  in (26) and (27) hold.

Moreover, a common SLAM initialization technique is to make use of *rotation only initialization* – i.e., compute the initializer  $R^{(0)}$  using *only* the relative rotation measurements [9]. This can have computational advantages in practice since  $L(\tilde{G}^\rho)$  is generally *sparse*; the same cannot be said for the pose-graph SLAM data matrix  $\tilde{Q}$ , as it arises via analytic elimination of the translational states, in which the resulting data matrix  $\tilde{Q}$  is formed as a (dense) generalized Schur complement (cf. [24]). An interesting fact is that for pose-graph SLAM, a spectral initialization  $R^{(0)}$  computed using the eigenvectors of  $L(\tilde{G}^\rho)$  (i.e. ignoring  $Q^\tau$ ) attains the bound:

$$d_S(R, R^{(0)}) \leq \frac{4\sqrt{2dn}\|\Delta L(\tilde{G}^\rho)\|_2}{\lambda_{d+1}(L(G^\rho))}. \quad (30)$$

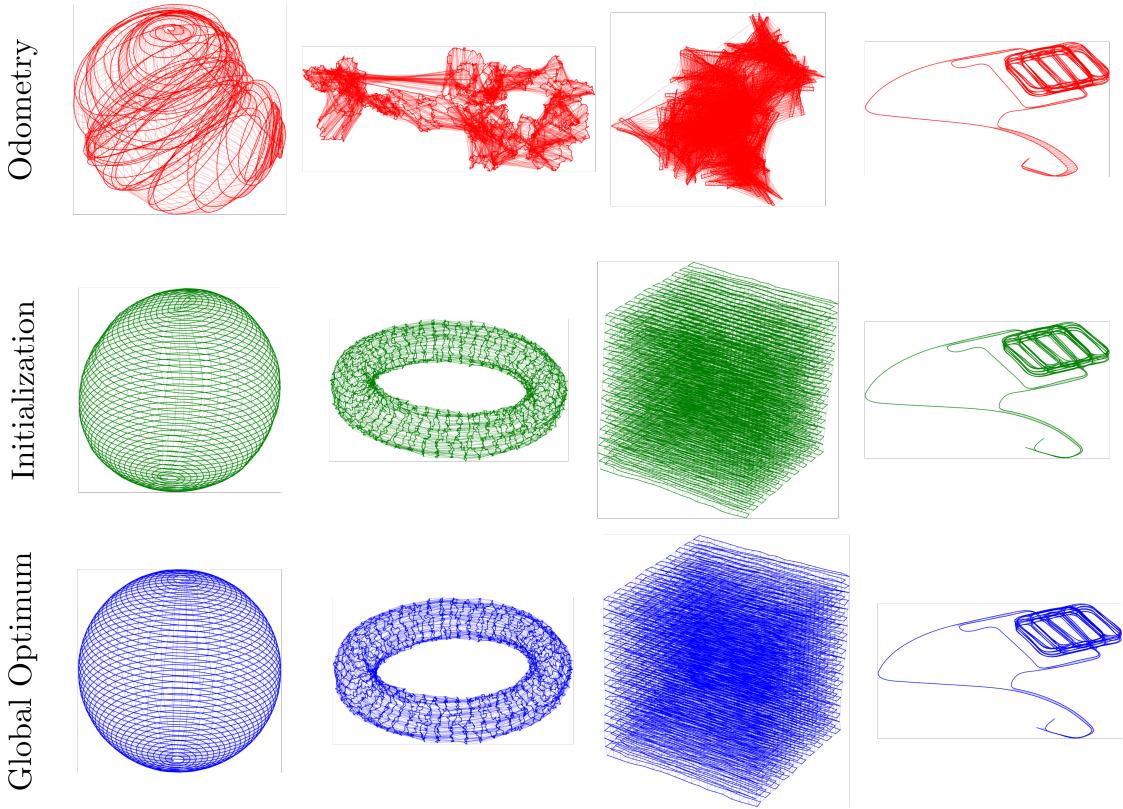
This bound holds by the same reasoning as Theorem 4, but with the consideration that  $R^\top \in \ker(L(G^\rho))$ .

As a final consideration, one may observe that in typical applications we also do not have access to  $\Delta Q$ . Indeed, if we had access to  $\Delta Q$ , we could simply recover the (latent) true data matrix  $\underline{Q}$  as  $\tilde{Q} - \Delta Q$ . The best we can typically do in practice is to simulate measurements from the generative model to empirically estimate statistics of  $\|\Delta Q\|_2$ .<sup>6</sup> This, in turn, can give a

---

<sup>5</sup>It is of course possible that sharper bounds will be derived that improve upon this result.

<sup>6</sup>Simulating measurements in the case of pose-graph SLAM requires knowledge of the ground-truth translation measurement scale, which is typically also unavailable in practice. However, the *rotation-only* initialization bound (30) applies in general and depends only upon the rotation measurements, which can be simulated to produce an empirical bound.



**Figure 3: Spectral relaxation produces high-quality initializations.** Qualitative comparison with the globally optimal solution suggests that the spectral relaxation produces estimates that are very close to the globally optimal solution for a variety of SLAM benchmark datasets.

*distribution* over the estimation error bound for an initializer  $d_S(R, R^{(0)})$  (likewise for the error of the maximum likelihood estimate, and the distance between the maximum likelihood estimate and the initializer).

## 6 Experimental results

In this section, we compare the theoretical bounds in Theorem 4 to the actual realized estimation error of the spectral initialization and globally optimal pose-graph SLAM solutions. In Section 6.1 we construct synthetic pose-graph SLAM scenarios for which the ground-truth poses are known. Since the bounds presented in our analysis depend upon knowledge of the noise magnitude  $\|\Delta Q\|_2$  and the spectral gap of the *true* data matrix  $Q$ , which are unknown in practice, our first set of empirical results shed light on the behavior of these worst-case bounds (as well as the *actual* error realized by different estimators) as we vary the noise parameters controlling the generative model (8).

In Section 6.2, we evaluate the performance of spectral relaxation as a practical initialization method in the context of 3D pose-graph SLAM applications. We show that, consistent with our results on synthetic data, the spectral initialization method offers high-quality initial solutions for pose-graph optimization, and, in particular, that the inclusion of translation data when computing a spectral estimator improves the objective value of the initial estimate compared to

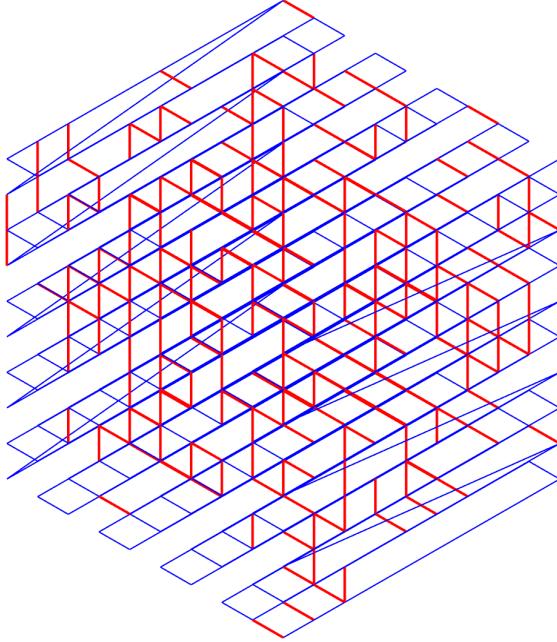


Figure 4: **Cube experiments.** Ground truth poses for the synthetic Cube dataset [8, 24] with  $s = 10$  vertices per side and  $p_{LC} = 0.1$ . The robot’s trajectory is shown in blue with loop closures shown in red.

the common approach of using exclusively rotation data.

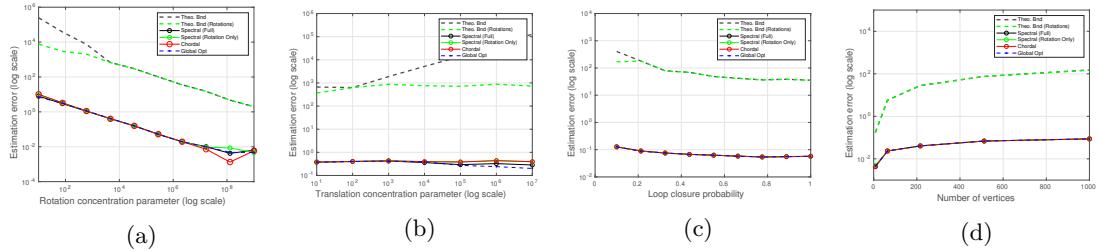
For all of our experiments, we also provide results for the well-known *chordal* initialization method [20], which relaxes the feasible set of Problem 3 to  $\mathbb{R}^{d \times dn}$ , with the constraint that  $R_1^{(0)} = I_d$ , for which the solution can be obtained by solving a linear system.

The spectral initialization method was implemented in C++ using Spectra to efficiently solve large-scale eigenvalue problems [23]. Computation of the theoretical bounds in Section 6.1 was performed in MATLAB using `eigs`. All experiments were performed on a laptop with a 2.2 GHz Intel i7 CPU. Where (verified) globally optimal solutions were needed, we used the C++ implementation of SE-Sync [24].

## 6.1 Evaluation on synthetic data

The bounds presented in our analysis depend upon knowledge of the noise magnitude  $\|\Delta Q\|_2$ , which is unknown in practice. In light of this fact, we examine empirically the behavior of the bounds as function of the noise parameters specifically. Beyond providing access to the ground-truth rotations, this experimental setup allows us to compare the worst-case analytic bounds we provide with empirical performance in noise regimes well outside the range typically encountered in real SLAM scenarios.

In this set of experiments, we examine the performance bounds for spectral initialization across a variety of noise parameters. We make use of the Cube dataset [8, 24]. The dataset consists of a set of vertices (poses) organized in a three-dimensional cube, with  $s$  vertices per dimension. Consecutive poses have a simulated “odometry” edge between them, and loop closures are sampled randomly from the remaining possible edges with probability  $p_{LC}$ . Each synthetic baseline is generated by randomly sampling problem instances from the generative model (8) with fixed noise parameters  $\kappa$  and  $\tau$  for all measurements. A sample configuration for the Cube



**Figure 5: Influence of dataset parameters on the performance bounds for the Cube experiments.** We examine empirically the change in the theoretical bounds (26) and (30) as well as the estimation error of several pose-graph optimization estimates while varying (a) the rotation concentration parameter  $\kappa$ , (b) the translation concentration parameter  $\tau$ , (c) the probability of a loop closure  $p_{LC}$ , (d) the number of vertices  $s^3$ .

dataset is provided in Figure 4.

**Influence of noise parameters on performance bounds:** In Figure 5, we study the performance of the spectral initialization approach across a variety of typical noise configurations. In each case, we provide the theoretical worst-case bounds (26) and (30) along with the empirical error of the different estimators under consideration. In Figure 5a, we sample Cube problem instances with logarithmically spaced values of  $\kappa$  while fixing the other parameters:  $\tau = 150$ ,  $p_{LC} = 0.2$ , and  $s = 10$ . In Fig. 5b, we fix  $\kappa = 10^5$ ,  $p_{LC} = 0.25$  and  $s = 10$  and sample problem instances with logarithmically-spaced translation concentration parameter  $\tau$ . In Fig. 5c, we fix  $\kappa = 10^5$ ,  $\tau = 150$ ,  $s = 10$  and vary  $p_{LC}$  from 0 to 1.

We observe that across a wide range of concentration parameters, the spectral relaxation produces estimates that attain very similar estimation error to the globally optimal estimate.<sup>7</sup> In particular, we find that the error of these estimates often improves upon the worst-case bounds (26) and (30) by orders of magnitude. This is consistent with earlier observations of qualitatively similar bounds for phase synchronization [22]. Moreover, in practical applications of rotation averaging and pose-graph optimization, previous work has shown that the maximum-likelihood estimator often attains expected error close to the Cramér-Rao *lower bound* (see [6] for rotation averaging and [11] for pose-graph optimization). The specific case varying the translation concentration parameter in Figure 5b is interesting, since while the spectral estimator improves with increasing  $\tau$ , the bound suggests the opposite worst-case behavior. This seems to stem from the fact that errors are scaled by increasing  $\tau$ , but without a commensurate increase in  $\lambda_{d+1}(Q)$  in the denominator of 26. This warrants further investigation, and also impacts similar prior bounds for pose-graph optimization, e.g. in [24]. With this exception, the behavior of the bounds with varying noise parameters seems to accurately capture the behavior of the actual estimation error.

**Dependence on problem dimensionality:** Due to the explicit appearance of the problem dimension  $n$  in the bounds of Theorems 5 and 4, as well as Corollary 6, it is interesting to consider how the number of rotations to be estimated affects these bounds. In Figure 5d, we fix  $\kappa = 10^5$ ,  $\tau = 150$ ,  $p_{LC} = 0.2$  and vary the number of vertices in the Cube dataset. Indeed, we find that the behavior of the worst-case bounds suggests an unfavorable scaling in the problem dimension: at  $s^3 = 8$  vertices, the worst-case bound overestimates the true error by approximately an order of magnitude; at  $s^3 = 1000$ , it overestimates the true error by approximately 3 orders

<sup>7</sup>The global minimizer  $R^*$  is optimal specifically with respect to the cost function defined in Problem 3. That is, a particular estimate which happens to be *suboptimal* with respect to this cost function may (by happenstance) be closer to  $R$  than  $R^*$ , a phenomenon observed in Fig. 5a.

Dataset		Odometry	Chordal	<b>Spectral (Rotation Only)</b>	<b>Spectral</b>	Global Opt.
Sphere	Iter	65	6	8	4	
	Cost	$1.14 \times 10^9$	1971.17	5594.19	1742.75	1687
	Time (s)	-	0.707	0.602	0.779	
Torus	Iter	32	5	5	4	
	Cost	$3.87 \times 10^8$	24669.2	25833.2	24272.7	24227
	Time (s)	-	1.316	1.501	1.199	
Grid	Iter	30	6	6	4	
	Cost	$1.97 \times 10^{10}$	87252	86966.1	84486.4	84320
	Time (s)	-	8.747	18.806	0.25	
Garage	Iter	1028	3	4	4	
	Cost	$2.31 \times 10^9$	1.42	3.215	2.7	1.26
	Time (s)	-	0.201	0.136	25.7	

Table 1: **Standard SLAM benchmarks** Objective value (cost), time, and number of RTR iterations (Iter.) required to compute SLAM solutions for several standard benchmarks. Proposed approaches are **bold**.

of magnitude. It is unclear, at present, however, whether it is possible to reduce or remove this dependence on the problem dimension, or whether there are particular configurations of  $\tilde{Q}$  and  $Q$  of the structure encountered in rotation averaging and pose-graph SLAM that produce estimates attaining (something close to) the theoretical bounds. A more sophisticated analysis taking into consideration the specific structure of these matrices may yield more refined bounds.

## 6.2 Evaluation on standard SLAM benchmark datasets

In these experiments, we consider evaluation of the spectral initialization method on several standard SLAM benchmark datasets. Figure 3 provides a qualitative comparison of three techniques for initialization: odometry only (i.e. composing measurements between consecutive poses), the proposed spectral initialization approach, and the globally optimal solution. We observe that spectral initialization provides solutions that visually resemble the globally optimal solution. Table 1 gives the corresponding quantitative comparison. We provide: the number of iterations required for a Riemannian trust-region (RTR) optimization method to converge to the globally optimal solution. With the exception of odometry-only initialization, all of the methods considered enabled the recovery of rank  $d$  (verifiably) globally optimal solutions using SE-Sync: in these cases, initialization methods coupled with standard optimization techniques recovered globally optimal solutions *without* the need to explicitly solve a large-scale semidefinite program.

Both of the spectral methods (using the “full” pose-graph optimization data matrix  $\tilde{Q}$  and the “rotation only” version using only  $L(\tilde{G}^{\rho})$ ) provide estimates competitive with the state-of-the-art chordal initialization method, generally attaining near-optimal objective values. An interesting observation is that, in their work, Moreira et al. [21] found that the rotation-only spectral estimator attains a higher cost on the Sphere dataset than alternative methods. We observe the same phenomenon, but interestingly, when we include the translation measurements (i.e. using the full data matrix  $\tilde{Q}$ ), we find that this discrepancy disappears. Similarly, the chordal estimator also performs well on this dataset, despite the fact that, like the rotation-only spectral initialization, it does not make use of translation information.

## 7 Conclusion

In this work, we presented the first performance guarantees for an initialization method in the context of robot perception problems (i.e. assuming incomplete graphs and typical noise models). In particular, the spectral relaxation of the original rotation averaging (resp. pose-graph SLAM) problem admits a solution in terms of the eigenvectors of the data matrix parameterizing the problem which has error upper-bounded as a function of the noise corrupting the data. These bounds also allow us to reason about the worst-case deviation of an initial guess from a maximum-likelihood estimate. Finally, we empirically examine the tightness of these bounds and compare the quality of two different spectral initializers in the context of 3D pose-graph SLAM applications using benchmark datasets.

The bounds presented here also emphasize an important connection between the spectral graph-theoretic properties of the graph parameterizing rotation averaging and SLAM problems and the error of estimators for those problems. This arises through the presence of the value  $\lambda_{d+1}(Q)$ , the algebraic connectivity of the rotation weight Laplacian in the case of rotation averaging, in the error bounds.

Moreover, the use of initialization techniques with performance guarantees raises an interesting line of reasoning. By virtue of the nonconvexity of Problem 3, the maximum-likelihood estimator attaining the performance bound in Theorem 5 may be difficult to obtain and indeed may be impossible to *verify*. Rosen et al. [24] observed problems in the high-noise regime where verification of solutions to the maximum-likelihood estimation problem for pose-graph SLAM could not be verified, but nonetheless where descent methods coupled with chordal initialization could obtain reasonable solutions. In contrast, for these problems the spectral estimator, which attains a similar worst-case error bound, can *always* be computed (and indeed may provide a very reasonable starting point for subsequent refinement using a descent method).

Finally, in practice one does not have access to the error matrix  $\Delta Q$ . In practical applications where we would like to understand the performance of estimators for these problems *a priori* (i.e. given only the noise statistics of a set of sensors and the topology of the graph), we may resort practically to empirical statistics about  $\|\Delta Q\|_2$  by simulation of noisy measurements using the generative model.

## A Symmetric perturbations of symmetric matrices

Recall that  $R$  and  $\Phi$  are the solutions of the noiseless and noisy versions of the spectral relaxation in Problem 4. It is straightforward to verify that these are in fact Stiefel manifold elements giving the  $d$  minimum eigenvectors for their corresponding data matrices. The Davis-Kahan Theorem is a classical result in linear algebra that measures the perturbation of a matrix's eigenvectors under a symmetric perturbation of that matrix [27]. Therefore, we make use of this theorem to derive a bound on the estimation error of a spectral estimator as a function of the noise in the data matrix. In particular, the proof of Lemma 2 (and consequently Theorem 4) relies on a particular variant of the Davis-Kahan  $\sin \theta$  Theorem [30, Theorem 2]. Here, we briefly restate the main result of [30] and give a proof of Lemma 2.

**Theorem 7** (Yu et al. [30], Theorem 2). *Let  $\Sigma, \hat{\Sigma} \in \mathbb{R}^{p \times p}$  be symmetric, with eigenvalues  $\lambda_p \geq \dots \geq \lambda_1$  and  $\hat{\lambda}_p \geq \dots \geq \hat{\lambda}_1$  respectively. Fix  $1 \leq s \leq r \leq p$  and assume that  $\min(\lambda_{s-1} - \lambda_s, \lambda_r - \lambda_{r+1}) > 0$ , where  $\lambda_{p+1} \triangleq \infty$  and  $\lambda_0 \triangleq -\infty$ . Let  $d \triangleq r - s + 1$ , and let  $V = (v_s, v_{s+1}, \dots, v_r) \in \mathbb{R}^{p \times d}$  and  $\hat{V} = (\hat{v}_s, \hat{v}_{s+1}, \dots, \hat{v}_r) \in \mathbb{R}^{p \times d}$  have orthonormal columns satisfying  $\Sigma v_j = \lambda_j v_j$  and*

$\hat{\Sigma}\hat{v}_j = \hat{\lambda}_j\hat{v}_j$  for  $j = s, s+1, \dots, r$ . Then there exists an orthogonal matrix  $G \in O(d)$  such that

$$\|\hat{V}G - V\|_F \leq \frac{2^{3/2} \min(d^{1/2} \|\hat{\Sigma} - \Sigma\|_{op}, \|\hat{\Sigma} - \Sigma\|_F)}{\min(\lambda_{s-1} - \lambda_s, \lambda_r - \lambda_{r+1})}. \quad (31)$$

With this result in hand, we are ready to prove Lemma 2.

*Proof of Lemma 2.* The data matrices  $\tilde{Q}$  and  $Q$  are symmetric  $dn \times dn$  matrices with eigenvalues  $\lambda_1, \dots, \lambda_{dn}$  and  $\tilde{\lambda}_1, \dots, \tilde{\lambda}_{dn}$ , respectively. Recall from Lemma 2 that  $\underline{R}$  consists of the ground-truth rotations and  $\Phi$  is the estimate obtained from the spectral relaxation in Problem 4, parameterized by  $\tilde{Q}$ . It is straightforward to verify from the statement of Problem 4 that the  $d$  normalized eigenvectors corresponding to  $\lambda_1, \dots, \lambda_d$  of  $Q$  and  $\tilde{\lambda}_1, \dots, \tilde{\lambda}_d$  are exactly  $\frac{1}{\sqrt{n}}\underline{R}^\top$  and  $\frac{1}{\sqrt{n}}\Phi^\top$ , respectively. Then, letting  $s = 1$  and  $r = d$  and applying Theorem 7, there exists an orthogonal matrix  $G \in O(d)$  such that:

$$\frac{1}{\sqrt{n}}\|\Phi^\top G - \underline{R}^\top\|_F \leq \frac{2\sqrt{2d}\|\tilde{Q} - Q\|_2}{\lambda_{d+1}(Q) - \lambda_d(Q)}. \quad (32)$$

Multiplying both sides of this expression by  $\sqrt{n}$ , we have:

$$\|\Phi^\top G - \underline{R}^\top\|_F \leq \frac{2\sqrt{2dn}\|\tilde{Q} - Q\|_2}{\lambda_{d+1}(Q) - \lambda_d(Q)}. \quad (33)$$

Now, by definition  $\Delta Q = \tilde{Q} - Q$ , and from the fact that  $\underline{R}^\top \in \ker(Q)$  and that  $\lambda_{d+1}(Q) > 0$  (since  $\mathcal{G}$  is connected),<sup>8</sup> we have that  $\lambda_d = 0$  and the above expression simplifies to:

$$\|\Phi^\top G - \underline{R}^\top\|_F \leq \frac{2\sqrt{2dn}\|\Delta Q\|_2}{\lambda_{d+1}(Q)}. \quad (34)$$

Taking the transpose of the terms inside the norm gives the desired result.  $\square$

## B Proof of the main results

In this section, we prove the main results, i.e. Theorem 4, Theorem 5, and Corollary 6.

### B.1 An upper bound for the estimation error in Problem 4

*Proof.* To simplify the subsequent derivation, we will assume without loss of generality that  $\underline{R}$  and  $\Phi$  are the representatives of their orbits satisfying  $d_{\mathcal{O}}(\underline{R}, \Phi) = \|\underline{R} - \Phi\|_F$ . Recall from the definition of  $d_{\mathcal{S}}(\underline{R}, R^{(0)})$  that:

$$d_{\mathcal{S}}(\underline{R}, R^{(0)}) = \min_{G \in SO(d)} \|\underline{R} - GR^{(0)}\|_F. \quad (35)$$

Therefore, we have:

$$\begin{aligned} d_{\mathcal{S}}(\underline{R}, R^{(0)})^2 &= \min_{G \in SO(d)} \|\underline{R} - GR^{(0)}\|_F^2 \\ &\leq \|\underline{R} - R^{(0)}\|_F^2, \\ &= \sum_{i=1}^n \|\underline{R}_i - \Pi_{\mathcal{S}}(\Phi_i)\|_F^2, \end{aligned} \quad (36)$$

---

<sup>8</sup>It is not particularly restrictive to assume that  $\mathcal{G}$  is connected. In the case that  $\mathcal{G}$  is not connected, the estimation problem splits over the connected components of  $\mathcal{G}$ , and all of our results hold separately for each connected component.

where in the last line we have used the fact that  $R^{(0)}$  consists of the projections of individual  $(d \times d)$  blocks of  $\Phi$  onto  $\text{SO}(d)$ . From Lemma 3, we have that each of the  $n$  summands above satisfies:

$$\|R_i - \Pi_{\mathcal{S}}(\Phi_i)\|_F^2 \leq 4\|R_i - \Phi_i\|_F^2. \quad (37)$$

This, in turn, gives a corresponding bound on the summation:

$$\begin{aligned} \sum_{i=1}^n \|R_i - \Pi_{\mathcal{S}}(\Phi_i)\|_F^2 &\leq 4 \sum_{i=1}^n \|R_i - \Phi_i\|_F^2 \\ &= 4\|\underline{R} - \Phi\|_F^2. \end{aligned} \quad (38)$$

Since, by hypothesis,  $\Phi$  and  $\underline{R}$  are representatives of their orbits satisfying  $d_{\mathcal{O}}(\underline{R}, \Phi) = \|\underline{R} - \Phi\|_F$ , we have:

$$4\|\underline{R} - \Phi\|_F^2 = 4d_{\mathcal{O}}(\underline{R}, \Phi)^2. \quad (39)$$

Applying Lemma 2, we directly obtain:

$$4d_{\mathcal{O}}(\underline{R}, \Phi)^2 \leq 4(2\sqrt{2dn})^2 \frac{\|\Delta Q\|_2^2}{\lambda_{d+1}(\underline{Q})^2}. \quad (40)$$

In summary, we have:

$$d_{\mathcal{S}}(\underline{R}, R^{(0)})^2 \leq 4(2\sqrt{2dn})^2 \frac{\|\Delta Q\|_2^2}{\lambda_{d+1}(\underline{Q})^2}. \quad (41)$$

Taking the square root of both sides of the inequality in the last line gives:

$$d_{\mathcal{S}}(\underline{R}, R^{(0)}) \leq \frac{4\sqrt{2dn}\|\Delta Q\|_2}{\lambda_{d+1}(\underline{Q})}, \quad (42)$$

which concludes the proof.  $\square$

## B.2 An upper bound for the estimation error in Problem 3

We begin following the arguments of Preskitt [22, Appendix D.4]. From the optimality of  $R^*$  we have:

$$\begin{aligned} \text{tr}(\tilde{Q}\underline{R}^T\underline{R}) &= \text{tr}(Q\underline{R}^T\underline{R}) + \text{tr}(\Delta Q\underline{R}^T\underline{R}) \\ &\geq \text{tr}(Q\underline{R}^{*\top}R^*) + \text{tr}(\Delta Q\underline{R}^{*\top}R^*) = \text{tr}(\tilde{Q}\underline{R}^{*\top}R^*). \end{aligned} \quad (43)$$

Since  $\text{tr}(Q\underline{R}^T\underline{R}) = 0$ , we can rearrange the above expression to obtain:

$$\text{tr}(Q\underline{R}^{*\top}R^*) \leq \text{tr}(\Delta Q\underline{R}^T\underline{R}) - \text{tr}(\Delta Q\underline{R}^{*\top}R^*). \quad (44)$$

Using the fact that  $\text{tr}(\Delta Q\underline{R}^T\underline{R}) = \text{vec}(\underline{R})^T(\Delta Q \otimes I_n) \text{vec}(\underline{R})$  (and likewise for  $\text{tr}(\Delta Q\underline{R}^{*\top}R^*)$ ), we have:

$$\begin{aligned} \text{tr}(Q\underline{R}^{*\top}R^*) &\leq \text{vec}(\underline{R} - R^*)^T(\Delta Q \otimes I_n) \text{vec}(\underline{R} + R^*) \\ &\leq \|\text{vec}(\underline{R} - R^*)\|_2 \|\Delta Q \otimes I_n\|_2 \|\text{vec}(\underline{R} + R^*)\|_2 \\ &= \|\underline{R} - R^*\|_F \|\Delta Q\|_2 \|\underline{R} + R^*\|_F \\ &\leq 2\sqrt{dn} \|\underline{R} - R^*\|_F \|\Delta Q\|_2. \end{aligned} \quad (45)$$

In order to lower-bound the right-hand side of (45) in terms of the estimation error  $d_{\mathcal{S}}(\underline{R}, R^*)$ , we will make use of the following technical lemma of Rosen et al. [24]:

**Lemma 8** (Lemma 11 of Rosen et al. [24]). Let  $R \in O(d)^n \subset \mathbb{R}^{d \times dn}$  and furthermore let  $M = \{WR \mid W \in \mathbb{R}^{d \times d}\} \subset \mathbb{R}^{d \times dn}$  be the subspace of matrices with rows contained in  $\text{image}(R^\top)$ . Then

$$\begin{aligned}\text{Proj}_V : \mathbb{R}^{dn} &\rightarrow \text{image}(R^\top) \\ \text{Proj}_V(x) &= \frac{1}{n} R^\top Rx\end{aligned}\tag{46}$$

is the orthogonal projection onto  $\text{image}(R^\top)$  with respect to the  $\ell_2$  inner product, and the map

$$\begin{aligned}\text{Proj}_M : \mathbb{R}^{d \times dn} &\rightarrow M \\ \text{Proj}_M(X) &= \frac{1}{n} X R^\top R\end{aligned}\tag{47}$$

which applies  $\text{Proj}_V$  to the rows of  $X$  is the orthogonal projection onto  $M$  with respect to the Frobenius inner product.

Since  $\ker(Q) = \text{image}(R^\top)$  and  $\dim(\text{image}(R^\top)) = d$ , from Lemma 8, we have:

$$\text{tr}(Q R^{*\top} R^*) \geq \lambda_{d+1}(Q) \|P\|_F^2,\tag{48}$$

where

$$\begin{aligned}R^* &= K + P \\ K &= \text{Proj}_M(R^*) = \frac{1}{n} R^* R^\top R \\ P &= R^* - \text{Proj}_M(R^*) = R^* - \frac{1}{n} R^* R^\top R\end{aligned}\tag{49}$$

is an orthogonal decomposition of  $R^*$  and the rows of  $P$  are contained in the orthogonal complement of  $\text{image}(R^\top)$

The following lemma provides a bound on  $d_S(R, R^*)^2$  in terms of  $\|P\|_F^2$ .

**Lemma 9.** Let  $R^*$  and  $R$  be representatives of their orbits such that  $d_S(R, R^*) = \|R - R^*\|_F$ , and  $P = R^* - \text{Proj}_M(R^*)$  as defined in (49). Then:

$$\frac{1}{4} d_S(R, R^*)^2 \leq \|P\|_F^2.\tag{50}$$

*Proof.* Let  $X = \frac{1}{n} R R^{*\top}$ , so that  $K = X^\top R$ . Expanding the left hand side, we have:

$$\begin{aligned}d_S(R, R^*)^2 &= \|R^* - R\|_F^2 \\ &\leq \|R^* - \Pi_S(X^\top)R\|_F^2,\end{aligned}\tag{51}$$

from the fact that the orbit distance is obtained as the minimum over  $G \in SO(d)$  of the quantity  $\|R^* - GR\|_F$ , and that by hypothesis this minimum is obtained as  $\|R^* - R\|_F$ . Breaking up the norm into its blockwise summands, and from the orthogonal invariance of the Frobenius norm, we can rearrange this expression as follows:

$$\begin{aligned}\|R^* - \Pi_S(X^\top)R\|_F^2 &= \sum_{i=1}^n \|R_i^* - \Pi_S(X^\top)R_i\|_F^2 \\ &= \sum_{i=1}^n \|R_i^* R_i^\top - \Pi_S(X^\top)\|_F^2.\end{aligned}\tag{52}$$

From Lemma 3, we know that each summand in the above expression satisfies

$$\|R_i^* R_i^\top - \Pi_S(X^\top)\|_F^2 \leq 4 \|R_i^* R_i^\top - X^\top\|_F^2.\tag{53}$$

Since this bound is satisfied for each summand, the total summation satisfies

$$\begin{aligned}
\sum_{i=1}^n \|R_i^* R_i^\top - \Pi_S(X^\top)\|_F^2 &\leq 4 \sum_{i=1}^n \|R_i^* R_i^\top - X^\top\|_F^2 \\
&= 4 \sum_{i=1}^n \|R_i^* - X^\top \underline{R}_i\|_F^2 \\
&= 4 \|R^* - X^\top \underline{R}\|_F^2.
\end{aligned} \tag{54}$$

Since  $K = X^\top \underline{R}$ , we have:

$$\begin{aligned}
4 \|R^* - X^\top \underline{R}\|_F^2 &= 4 \|R^* - K\|_F^2 \\
&= 4 \|P\|_F^2,
\end{aligned} \tag{55}$$

which gives the desired bound.  $\square$

With this result, we are ready to prove Theorem 5.

*Proof.* From (48) and (45), we have:

$$\lambda_{d+1}(Q) \|P\|_F^2 \leq 2\sqrt{dn} \|R - R^*\|_F \|\Delta Q\|_2. \tag{56}$$

Since, by hypothesis,  $R^*$  and  $\underline{R}$  are the representatives of their orbits satisfying  $d_S(\underline{R}, R^*) = \|\underline{R} - R^*\|_F$ , from Lemma 9 we have

$$d_S(\underline{R}, R^*)^2 \leq 4 \|P\|_F^2. \tag{57}$$

Combining (57) with (56), we obtain:

$$d_S(\underline{R}, R^*) \leq \frac{8\sqrt{dn} \|\Delta Q\|_2}{\lambda_{d+1}(Q)}, \tag{58}$$

which is what we intended to show.  $\square$

### B.3 An upper bound on the distance between a spectral estimator and the maximum-likelihood estimator

**Lemma 10** (Orbit distances are pseudometrics). *The orbit distances  $d_S$  and  $d_O$  are pseudometrics on  $\text{SO}(d)^n$  and  $\text{O}(d)^n$ , respectively. In particular, for all  $X, Y, Z \in \text{SO}(d)^n$ , we have:*

1.  $d_S(X, X) = 0$
2.  $d_S(X, Y) = d_S(Y, X)$
3.  $d_S(X, Z) \leq d_S(X, Y) + d_S(Y, Z),$

and likewise for  $d_O$ .

*Proof.* To simplify the subsequent derivation, we will continue with the proof for the case of  $d_S$  on  $\text{SO}(d)^n$ , but the same methodology applies directly to  $d_O$  on  $\text{O}(d)^n$ . A *pseudometric* on  $\text{SO}(d)^n$  (resp.  $\text{O}(d)^n$ ) is any nonnegative function  $\text{SO}(d)^n \times \text{SO}(d)^n \rightarrow \mathbb{R}_{\geq 0}$  satisfying the properties 1–3 [16]. To establish 1, we have:

$$d_S(X, X) = \min_{G \in \text{SO}(d)} \|X - GX\|_F = 0, \tag{59}$$

since  $d_S(X, Y) \geq 0$  for all  $X, Y \in \text{SO}(d)^n$  and taking  $G = I$  gives the desired result. For 2, we have:

$$\begin{aligned} d_S(X, Y) &= \min_{G \in \text{SO}(d)} \|X - GY\|_F \\ &= \min_{G \in \text{SO}(d)} \|Y - G^\top X\|_F, \end{aligned} \quad (60)$$

where the second line follows from the orthogonal invariance of the Frobenius norm. Finally, letting  $S = G^\top \in \text{SO}(d)$ , we have:

$$\begin{aligned} d_S(X, Y) &= \min_{G \in \text{SO}(d)} \|Y - G^\top X\|_F \\ &= \min_{S \in \text{SO}(d)} \|Y - SX\|_F \\ &= d_S(Y, X). \end{aligned} \quad (61)$$

Finally, to establish 3, we aim to prove that for any  $X, Y, Z \in \text{SO}(d)^n$ :

$$d_S(X, Z) \leq d_S(X, Y) + d_S(Y, Z). \quad (62)$$

Suppose the orbit distance  $d_S(X, Y)$  is attained with minimizer  $G_{XY}^* \in \text{SO}(d)$  and likewise the distance  $d_S(Y, Z)$  is attained with minimizer  $G_{YZ}^* \in \text{SO}(d)$ . Define:

$$G' \triangleq G_{XY}^* G_{YZ}^*. \quad (63)$$

Now, since  $G'$  is itself the product of two elements of  $\text{SO}(d)$ , we know  $G' \in \text{SO}(d)$ , and therefore:

$$d_S(X, Z) = \min_{G \in \text{SO}(d)} \|X - GZ\|_F \leq \|X - G'Z\|_F. \quad (64)$$

Examining the right-hand side of this expression, we have:

$$\begin{aligned} \|X - G'Z\|_F &= \|X - G_{XY}^* Y + G_{XY}^* Y - G'Z\|_F \\ &\leq \underbrace{\|X - G_{XY}^* Y\|_F}_{d_S(X, Y)} + \|G_{XY}^* Y - G'Z\|_F, \end{aligned} \quad (65)$$

where the last line follows from the triangle inequality for the Frobenius norm. Now, substitution of the definition (63) into the second term of (65) reveals:

$$\begin{aligned} \|G_{XY}^* Y - G'Z\|_F &= \|G_{XY}^* Y - G_{XY}^* G_{YZ}^* Z\|_F \\ &= \|Y - G_{YZ}^* Z\|_F \\ &= d_S(Y, Z), \end{aligned} \quad (66)$$

where the second line follows from the orthogonal invariance of the Frobenius norm. Taken together, these results give:

$$d_S(X, Z) \leq \|X - G'Z\|_F \leq d_S(X, Y) + d_S(Y, Z), \quad (67)$$

which is what we intended to show.  $\square$

Lemma 10 suggests a straightforward proof of Corollary 6.

*Proof.* From the triangle inequality for  $d_S$ , we have:

$$d_S(R^{(0)}, R^*) \leq d_S(R, R^{(0)}) + d_S(R, R^*). \quad (68)$$

Substitution of (26) and (27) into (68) gives the desired result.  $\square$

## C Relationship to the method of Moreira et al. [21]

In their recent work, Moreira et al. [21] construct an estimator for pose-graph SLAM problems based on eigenvector computations (and a computationally-efficient procedure for eigenvector recovery). In this section, we show that their approach is formally equivalent to the *rotation-only* variant of the spectral initialization we discuss in Section 5 and therefore has estimation error satisfying the bound (30). Moreira et al. [21] specifically consider *unweighted* rotation measurements, for which it suffices to consider the generative model (6) taking  $\kappa_{ij} = 1$  for all edges  $(i, j) \in \mathcal{E}$ .

Their construction begins by considering the matrix  $\tilde{M} \in \mathbb{R}^{dn \times dn}$  with  $d \times d$  block  $i, j$  given by:

$$\tilde{M}_{ij} = \begin{cases} \tilde{R}_{ij} & \text{if } (i, j) \in \mathcal{E} \\ I_d & \text{if } i = j \\ 0_d & \text{otherwise.} \end{cases} \quad (69)$$

They observe that for all stationary points  $\hat{R} \in \text{SO}(d)^n \subset \mathbb{R}^{d \times dn}$ , there is a corresponding matrix  $\Lambda \in \mathbb{R}^{dn \times dn}$  such that:

$$\underbrace{(\Lambda - \tilde{M})}_{S} \hat{R}^T = 0, \quad (70)$$

where  $\Lambda$  has the symmetric  $d \times d$  block diagonal structure:

$$\Lambda = \begin{bmatrix} \Lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Lambda_n \end{bmatrix}. \quad (71)$$

In the *noiseless* case where  $\tilde{M} = M$ ,<sup>9</sup> the matrix  $S = \Lambda - M$  is given by:

$$S = (\mathcal{L} \otimes J_d) \circ M, \quad (72)$$

where  $\mathcal{L}$  is the scalar (unweighted) rotation graph Laplacian,  $J_d \in \mathbb{R}^{d \times d}$  is an all-ones matrix, and  $\circ$  denotes the Hadamard product. Moreover,  $S \succeq 0$  and  $R^T \in \ker(S)$ , so the ground-truth rotations  $R$  can be recovered by computing the  $d$  eigenvectors of  $S$  corresponding to the smallest eigenvalues of  $S$ . In the case of noisy measurements, they compute the eigenvectors of  $S = (\mathcal{L} \otimes J_3) \circ \tilde{M}$ . Now, by comparing the definition, we observe that the quantity  $(\mathcal{L} \otimes J_3) \circ \tilde{M}$  is identical to  $L(\tilde{G}^\rho)$  (cf. equation (13a)) in the case where all  $\kappa_{ij} = 1$ . Consequently, their spectral estimator based on recovering eigenvectors of  $S$  is exactly identical to the procedure outlined in Section 4 in the case of unweighted measurements, and where translation measurements have been discarded.

## References

- [1] Federica Arrigoni, Beatrice Rossi, and Andrea Fusiello. Spectral synchronization of multiple views in  $\text{SE}(3)$ . *SIAM Journal on Imaging Sciences*, 9(4):1963–1990, 2016.
- [2] Afonso S Bandeira, Nicolas Boumal, and Amit Singer. Tightness of the maximum likelihood semidefinite relaxation for angular synchronization. *Mathematical Programming*, 163(1-2):145–167, 2017.

---

<sup>9</sup>In keeping with the notation in the rest of this manuscript, we use the notation  $M$  to denote the measurement matrix (69) constructed from the *ground-truth* rotations  $R_{ij}$ .

- [3] Byron Boots and Geoff Gordon. A spectral learning approach to range-only SLAM. In *International Conference on Machine Learning*, pages 19–26. PMLR, 2013.
- [4] Nicolas Boumal. Nonconvex phase synchronization. *SIAM Journal on Optimization*, 26(4):2355–2377, 2016.
- [5] Nicolas Boumal, Amit Singer, and P-A Absil. Robust estimation of rotations from relative measurements by maximum likelihood. In *52nd IEEE Conference on Decision and Control*, pages 1156–1161. IEEE, 2013.
- [6] Nicolas Boumal, Amit Singer, P-A Absil, and Vincent D Blondel. Cramér–Rao bounds for synchronization of rotations. *Information and Inference: A Journal of the IMA*, 3(1):1–39, 2014.
- [7] Jesus Briales and Javier Gonzalez-Jimenez. Cartan-Sync: Fast and global SE(d)-synchronization. *IEEE Robotics and Automation Letters*, 2(4):2127–2134, 2017.
- [8] Luca Carlone, David M Rosen, Giuseppe Calafiore, John J Leonard, and Frank Dellaert. Lagrangian duality in 3D SLAM: Verification techniques and optimal solutions. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 125–132. IEEE, 2015.
- [9] Luca Carlone, Roberto Tron, Kostas Daniilidis, and Frank Dellaert. Initialization techniques for 3D SLAM: a survey on rotation estimation and its use in pose graph optimization. In *2015 IEEE international conference on robotics and automation (ICRA)*, pages 4597–4604. IEEE, 2015.
- [10] Luca Carlone, Giuseppe C Calafiore, Carlo Tommolillo, and Frank Dellaert. Planar pose graph optimization: Duality, optimal solutions, and verification. *IEEE Transactions on Robotics*, 32(3):545–565, 2016.
- [11] Yongbo Chen, Shoudong Huang, Liang Zhao, and Gamini Dissanayake. Cramér–Rao bounds and optimal design metrics for pose-graph SLAM. *IEEE Transactions on Robotics*, 37(2):627–641, 2021.
- [12] Frank Dellaert, David M Rosen, Jing Wu, Robert Mahony, and Luca Carlone. Shanon rotation averaging: Global optimality by surfing  $\text{SO}(p)^N$ . In *European Conference on Computer Vision*, pages 292–308. Springer, 2020.
- [13] Taosha Fan, Hanlin Wang, Michael Rubenstein, and Todd Murphy. CPL-SLAM: Efficient and certifiably correct planar graph-based SLAM using the complex number representation. *IEEE Transactions on Robotics*, 36(6):1719–1737, 2020.
- [14] Giorgio Grisetti, Rainer Kummerle, Cyrill Stachniss, and Wolfram Burgard. A tutorial on graph-based SLAM. *IEEE Intelligent Transportation Systems Magazine*, 2(4):31–43, 2010.
- [15] Richard J Hanson and Michael J Norris. Analysis of measurements based on the singular value decomposition. *SIAM Journal on Scientific and Statistical Computing*, 2(3):363–373, 1981.
- [16] John L Kelley. *General Topology*. Van Nostrand, New York, 1955.
- [17] Kasra Khosoussi, Shoudong Huang, and Gamini Dissanayake. Novel insights into the impact of graph structure on SLAM. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2707–2714. IEEE, 2014.

- [18] Shuyang Ling. Near-optimal performance bounds for orthogonal and permutation group synchronization via spectral methods. *arXiv preprint arXiv:2008.05341*, 2020.
- [19] Huikang Liu, Man-Chung Yue, and Anthony Man-Cho So. A unified approach to synchronization problems over subgroups of the orthogonal group. *arXiv preprint arXiv:2009.07514*, 2020.
- [20] Daniel Martinec and Tomas Pajdla. Robust rotation and translation estimation in multiview reconstruction. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2007.
- [21] Gabriel Moreira, Manuel Marques, and Joao Paulo Costeira. Fast pose graph optimization via krylov-schur and cholesky factorization. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 1898–1906, 2021.
- [22] Brian Patrick Preskitt. *Phase retrieval from locally supported measurements*. PhD thesis, UC San Diego, 2018.
- [23] Yixuan Qiu. Spectra: C++ library for large scale eigenvalue problems. <https://spectralib.org>, 2015.
- [24] David M Rosen, Luca Carlone, Afonso S Bandeira, and John J Leonard. SE-Sync: A certifiably correct algorithm for synchronization over the special Euclidean group. *The International Journal of Robotics Research*, 38(2-3):95–125, 2019.
- [25] David M Rosen, Kevin J Doherty, Antonio Terán Espinoza, and John J Leonard. Advances in Inference and Representation for Simultaneous Localization and Mapping. *Annual Review of Control, Robotics, and Autonomous Systems*, 4, 2021.
- [26] Amit Singer. Angular synchronization by eigenvectors and semidefinite programming. *Applied and computational harmonic analysis*, 30(1):20–36, 2011.
- [27] G.W. Stewart, J.W. Stewart, J. Sun, Academic Press (Londyn)., and Harcourt Brace Jovanovich. *Matrix Perturbation Theory*. Computer Science and Scientific Computing. Elsevier Science, 1990. ISBN 9780126702309. URL <https://books.google.com/books?id=178PAQAAQAAJ>.
- [28] Yulun Tian, Kasra Khosoussi, David M Rosen, and Jonathan P How. Distributed certifiably correct pose-graph optimization. *arXiv preprint arXiv:1911.03721*, 2019.
- [29] Shinji Umeyama. Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 13(04):376–380, 1991.
- [30] Yi Yu, Tengyao Wang, and Richard J Samworth. A useful variant of the Davis–Kahan theorem for statisticians. *Biometrika*, 102(2):315–323, 2015.