

A Unified Approach to Synchronization Problems over Subgroups of the Orthogonal Group

Huikang Liu^{*} Man-Chung Yue[†] Anthony Man-Cho So[‡]

May 20, 2022

Abstract

The problem of synchronization over a group \mathcal{G} aims to estimate a collection of group elements $G_1^*, \dots, G_n^* \in \mathcal{G}$ based on noisy observations of a subset of all pairwise ratios of the form $G_i^* G_j^{*-1}$. Such a problem has gained much attention recently and finds many applications across a wide range of scientific and engineering areas. In this paper, we consider the class of synchronization problems in which the group is a closed subgroup of the orthogonal group. This class covers many group synchronization problems that arise in practice. Our contribution is fivefold. First, we propose a unified approach for solving this class of group synchronization problems, which consists of a suitable initialization step and an iterative refinement step based on the generalized power method, and show that it enjoys a strong theoretical guarantee on the estimation error under certain assumptions on the group, measurement graph, noise, and initialization. Second, we formulate two geometric conditions that are required by our approach and show that they hold for various practically relevant subgroups of the orthogonal group. The conditions are closely related to the error-bound geometry of the subgroup — an important notion in optimization. Third, we verify the assumptions on the measurement graph and noise for standard random graph and random matrix models. Fourth, based on the classic notion of metric entropy, we develop and analyze a novel spectral-type estimator. Finally, we show via extensive numerical experiments that our proposed non-convex approach outperforms existing approaches in terms of computational speed, scalability, and/or estimation error.

Keywords: Group Synchronization, (Special) Orthogonal Matrix, Permutation Matrix, Cyclic Group, Generalized Power Method, Estimation Error, Metric Entropy, Error Bound

^{*}Shanghai University of Finance and Economics. E-mail: liuhuikang@sufe.edu.cn

[†]The University of Hong Kong. E-mail: mc Yue@hku.hk

[‡]The Chinese University of Hong Kong. E-mail: mancho so@se.cuhk.edu.hk

1 Introduction

In many real-world estimation problems, the signals of interest, which are commonly referred to as *ground truths*, are constrained to lie in a *group*.¹ One such example is the *group synchronization problem* (or simply *synchronization problem*), in which the ground truth takes the form $G^* = (G_1^*, \dots, G_n^*)$ with G_1^*, \dots, G_n^* being elements of a group \mathcal{G} and is to be estimated from noisy measurements of a subset of all pairwise ratios of the form $G_i^* G_j^{*-1}$ that are computed using the binary operation on \mathcal{G} . Synchronization problems have found applications across a wide range of areas, such as social science, distributed network, signal processing, computer vision, robotics, structural biology, computational genomics, and machine learning, and hence have gained much attention over the past decade. Indeed, group synchronization problems often appear as sub-tasks in many important problems from the above areas, including community detection [1, 23] where \mathcal{G} is the Boolean group; ranking [24] and distributed clock synchronization [30] where \mathcal{G} is the group of 2D rotations; sensor network localization and cryo-electron microscopy where \mathcal{G} is the group of orthogonal matrices or special orthogonal matrices [25, 60]; the pose graph estimation problem [57] where \mathcal{G} is the group of Euclidean motions; the haplotype phasing problem [20, 61] where \mathcal{G} is the cyclic group (integers with modulo arithmetics); the multi-graph matching problem [37, 53] where \mathcal{G} is the group of permutations. In most applications, after solving the synchronization problem, the estimated group elements will in turn be used to estimate another underlying signal, which is the ultimate target. In principle, one could estimate the underlying signal directly or jointly with the group elements. However, as pointed out in [8], estimating the signal given the group elements is a considerably easier task and can be addressed using well-developed techniques for inverse problems. This motivates us to focus on the synchronization problem.

Numerical approaches to synchronization problems are roughly divided into three categories: Spectral-type estimators, semidefinite programming (SDP) relaxations, and non-convex approaches. In the context of synchronization problems, a spectral-type estimator was first introduced in [62] for phase synchronization (*i.e.*, \mathcal{G} is the group $\mathcal{SO}(2)$ of 2D rotations). It has later been generalized to synchronization problems over other subgroups of the orthogonal group (see [2, 9, 45, 53]) and even general compact groups [56]. The main cost of computing a spectral-type estimator comes in two parts. First, the eigenvectors corresponding to the first few eigenvalues of the *graph connection Laplacian* [9] or a data matrix defined using the noisy observations are computed (see Sections 2 and 5.4 for details). Second, a certain rounding procedure is invoked to ensure that the returned estimator lies in the feasible set \mathcal{G}^n . The major advantages of spectral-type estimators are

¹Recall that a *group* is a pair $(\mathcal{G}, *)$, where \mathcal{G} is a set and $*$ is a binary operation on \mathcal{G} , such that (i) $*$ is associative (*i.e.*, $g_1 * (g_2 * g_3) = (g_1 * g_2) * g_3$ for all $g_1, g_2, g_3 \in \mathcal{G}$); (ii) there exists an identity element $\text{id} \in \mathcal{G}$ (*i.e.*, $g * \text{id} = \text{id} * g = g$ for all $g \in \mathcal{G}$); (iii) for each $g \in \mathcal{G}$, there exists an inverse element g^{-1} of g (*i.e.*, $g * g^{-1} = g^{-1} * g = \text{id}$). For simplicity, we shall write $g_1 g_2$ for $g_1 * g_2$, where $g_1, g_2 \in \mathcal{G}$. Also, we shall abuse terminology and refer to \mathcal{G} as the group.

their low computational cost and ease of implementation.

The SDP relaxation approach is mainly based on convex relaxations of either the least squares formulation or the least unsquared deviation formulation of the synchronization problem, both of which are non-convex. These convex relaxations are SDP problems (hence the name of the approach) that are obtained via the by-now classic *lifting* technique (see, e.g., [50]) and are shown to be tight — the optimal solutions to the relaxed problem are also optimal for its non-relaxed counterpart — under certain assumptions on the observation and noise models [5, 6, 33, 38, 57, 71]. Recently, using results from harmonic analysis and representation theory, an SDP relaxation for general compact groups and general objective functions (as opposed to just the squared or unsquared deviation) has been derived and studied in [8]. A major drawback of the SDP relaxation approach is that it does not scale well with the number n of group elements to be estimated. Indeed, the approach requires solving an SDP problem with an $M \times M$ matrix decision variable, where M is linear in n . As such, the computational cost is on the order of $n^{3.5}$ [11, Section 6.6.3], which can be prohibitive when n is large. One remedy of this is to devise specialized algorithms for solving the large SDP problem. For example, an alternating direction augmented Lagrangian method was developed in [71] to solve an SDP relaxation of the least unsquared deviation formulation of the synchronization problem over the special orthogonal group. However, the computational cost is still not satisfactory when n is beyond a few hundreds. Moreover, the performance of the method is rather sensitive to how various parameters (e.g., step size, penalty parameter) are tuned. Another possible remedy is to solve the large SDP problem using the Burer-Monteiro heuristic [17, 18], which amounts to replacing the large semidefinite matrix variable $Z \in \mathbb{R}^{M \times M}$ by some factorization YY^\top with $Y \in \mathbb{R}^{M \times f}$ and solving the resulting optimization problem with the smaller matrix variable Y . However, the Burer-Monteiro heuristic, being a non-convex approach, is prone to sub-optimality due to the presence of local minima, unless the factorization dimension f is sufficiently large [7, 15, 16, 43].

In recent years, there have been attempts to solve the non-convex least squares formulation of the synchronization problem directly without relaxing it. Such a non-convex approach typically has two stages. In the first stage, a carefully designed initialization procedure is used to produce a point that is close enough to the ground truth. Then, in the second stage, an iterative procedure is used to refine the initial point. One procedure that is particularly suitable for the second stage is the *generalized power method* (GPM) [39, 48, 51]. The idea of applying GPM to synchronization problems was first introduced in [14] for phase synchronization and later further developed in [20] for the joint alignment problem and in [73] for the community detection problem. In addition, the analysis of GPM in [14] for phase synchronization was sharpened in [47, 79]. The advantage of the GPM-based non-convex approach is that it is much faster and more scalable than the SDP relaxation approach, as each iteration of GPM involves only matrix-vector multiplications and projections onto the group \mathcal{G} . As we shall see later, these projections can be computed efficiently for many concrete groups \mathcal{G} of practical relevance. Also, the

GPM-based approach is easy to implement and requires no parameter tuning. It is worth noting that although spectral-type estimators have been shown to be already qualitatively optimal in certain settings [14, 45], we have observed in our experiments that by refining a spectral-type estimator using GPM, the quality of the resulting estimator can be substantially better (see Section 6). In other words, we can greatly improve the performance of spectral-type estimators by paying a small amount of extra computational cost.

Interestingly, the procedures used to compute estimators of the ground truth can be viewed as *approximation algorithms* for the non-convex formulation of the synchronization problem at hand. For instance, it is known that the least squares formulation of synchronization over the Boolean group is equivalent to the MAX-CUT problem [71], while that of phase synchronization is equivalent to a complex quadratic maximization problem with unit modulus constraint [6]. As such, the *approximation accuracy* of various least squares estimators (*i.e.*, the gap between the objective value attained by the estimator in question and the optimal value of the formulation) can be determined, see, *e.g.*, [31, 64, 67]. However, since an optimal solution to the non-convex formulation is in general different from the ground truth, a more relevant and faithful measure of the quality of an estimator for a synchronization problem is the *estimation error*, which is defined as the deviation of the estimator from the ground truth. For synchronization over the special orthogonal group, the estimation error of various estimators has been studied in [6, 14, 47, 71].

Finally, we mention two relatively new yet effective approaches to general group synchronization problems, both of which fall into the category of *message passing* algorithms. First, an *approximate message passing* algorithm was derived in [54] for solving synchronization over general compact groups. Based on ideas from statistical physics, the work [54] provides a non-rigorous analysis on the asymptotic statistical guarantee of their algorithm in the large n limit. Second, a powerful framework has recently been developed in [42] for removing unreliable observations in the input data to general synchronization problems by leveraging a notion called *cycle-edge consistency*. It would be interesting to investigate both the theoretical and practical performance of our non-convex approach when combined with the framework in [42]. We leave this as a future work.

1.1 Contributions

In this paper, we consider synchronization problems over *closed subgroups of the orthogonal group* $\mathcal{O}(d)$, which include, *e.g.*, the orthogonal group $\mathcal{O}(d)$ itself, the special orthogonal group $\mathcal{SO}(d)$, the permutation group $\mathcal{P}(d)$, and the cyclic group \mathcal{Z}_m (see Section 2.1 for their definitions). We propose and analyze a non-convex approach for tackling this class of synchronization problems. Our main contribution is fivefold.

- First, we derive a *master theorem* for the proposed non-convex approach. Under four assumptions that are respectively related to the subgroup, noise, measurement graph, and initialization, our master theorem establishes an upper bound on the estimation error of the iterates of GPM and hence provides performance guarantees for the

proposed non-convex approach (see Theorem 1). The master theorem is applicable to general closed subgroups of the orthogonal group, measurement graphs, and noise. It also clearly reveals the roles played by these main components of a synchronization problem.

- Second, we formulate two key geometric conditions on the subgroup that can be used to verify the assumptions in the master theorem. These conditions are closely related to the *error-bound geometry* of the subgroup, which is a classic notion in optimization and plays an important role in the analysis of various iterative methods. We also prove that the two geometric conditions hold for the orthogonal group, the special orthogonal group, the permutation group, and the cyclic group, which are all practically relevant in the context of group synchronization problems.
- Third, we study random models of the measurement graph and noise. In particular, we show that if the measurement graph is the *Erdős-Rényi random graph* and the noise matrix is a random matrix with independent *sub-Gaussian* entries, then the assumptions on the measurement graph and noise in the master theorem will hold with high probability when the number n of target group elements is sufficiently large. This result is useful since many well-known random variables are sub-Gaussian, including the Gaussian, uniform, Bernoulli, and any bounded random variables.
- Fourth, we develop a novel spectral-type estimator, named the *entropic spectral estimator*, for our target class of synchronization problems. The entropic spectral estimator has an intimate connection to the classic geometric concept of *metric entropy*. We prove that under the above-mentioned geometric conditions and random models of the measurement graph and noise, the entropic spectral estimator will satisfy the assumption on the initialization of the non-convex approach with high probability.
- Finally, through extensive numerical experiments, we study the empirical performance of our proposed non-convex approach on synchronization problems over several subgroups of the orthogonal group. The experiment results show that the proposed approach outperforms existing ones in terms of computational speed, scalability, and/or estimation error.

Although the idea of applying GPM to solve synchronization problems over closed subgroups of the orthogonal group $\mathcal{O}(d)$ is natural in view of our earlier discussion, due to the non-commutativity of $\mathcal{O}(d)$, many of the key steps in [47] that rely on the commutativity of $\mathcal{SO}(2)$ break down. Hence, extending the theoretical results in [47] to synchronization problems over general subgroups of $\mathcal{O}(d)$ is highly non-trivial. Furthermore, the measurement and noise settings in this paper are significantly more general than those considered in [47]. It should also be pointed out that for cyclic synchronization \mathcal{Z}_m -sync, another non-convex approach was developed in [20]. However, unlike the approach in [20], the

dimension of the iterates and the computational cost of our approach do not increase with m . Therefore, our approach is arguably more efficient.

1.2 Organization

The rest of the paper is organized as follows. We formally introduce the group synchronization problem and some definitions related to it in Section 2. In Section 3, we propose a unified non-convex approach for solving synchronization problems over closed subgroups of the orthogonal group. We then prove a master theorem on the performance guarantee of the proposed approach in Section 4. In Section 5, we verify the conditions of the master theorem for various closed subgroups of the orthogonal group and standard random measurement graph and noise models. Lastly, we present results on the numerical performance of the proposed approach in Section 6 and conclude the paper in Section 7.

1.3 Notation

We use the following notation throughout the paper. For any $nd \times nd$ (resp. $nd \times d$) block matrix Y , we denote by $[Y]_{ij}$ (resp. $[Y]_i$) its (ij) -th (resp. i -th) $d \times d$ block. For any two matrices X and Y , we denote their Kronecker product by $X \otimes Y$ and, if they have conformable dimensions, their inner product by $\langle X, Y \rangle = \text{Tr}(X^\top Y)$. We use $\|X\|$ and $\|X\|_F$ to denote the operator norm and Frobenius norm of X , respectively. For any integer $k \geq 1$, we denote the $k \times k$ identity matrix by I_k . We use c, c_0, c_1, \dots to denote numerical constants in mathematical statements and proofs, whose values may change from appearance to appearance. For a graph with node set $[n] := \{1, \dots, n\}$, we denote by (i, j) the edge between nodes i and j .

2 Group Synchronization

2.1 Basic Setup

Let $d \geq 1$ be an integer. The basic objects in our study are the d -dimensional orthogonal group — *i.e.*, the set of $d \times d$ orthogonal matrices

$$\mathcal{O}(d) := \left\{ Q \in \mathbb{R}^{d \times d} : QQ^\top = Q^\top Q = I_d \right\}$$

with matrix multiplication as the binary operation — and its closed subgroups — *i.e.*, closed subsets of $\mathcal{O}(d)$ that form a group under matrix multiplication. These include many of the groups mentioned in Section 1, such as the orthogonal group $\mathcal{O}(d)$ itself (the Boolean group corresponds to $\mathcal{O}(1) = \{-1, +1\}$), the special orthogonal group

$$\mathcal{SO}(d) := \{Q \in \mathcal{O}(d) : \det(Q) = 1\},$$

the permutation group

$$\mathcal{P}(d) := \mathcal{O}(d) \cap \{0, 1\}^{d \times d},$$

and the cyclic group of order m

$$\mathcal{Z}_m := \left\{ \begin{bmatrix} \cos \frac{2k\pi}{m} & -\sin \frac{2k\pi}{m} \\ \sin \frac{2k\pi}{m} & \cos \frac{2k\pi}{m} \end{bmatrix} : k = 0, \dots, m-1 \right\}$$

(which is a closed subgroup of $\mathcal{O}(2)$).

Given a closed subgroup \mathcal{G} of the orthogonal group, the problem of *group synchronization over \mathcal{G}* , denoted by \mathcal{G} -sync, is to estimate the ground truth $G^* = (G_1^*, \dots, G_n^*) \in \mathcal{G}^n$ based on noisy observations of the pairwise ratios

$$C_{ij} \approx G_i^* G_j^{*-1}, \quad (i, j) \in E,$$

where $E \subseteq \{(i, j) \in [n]^2 : i < j\}$ is a collection of index pairs. We view \mathcal{G}^n as a subset of $\mathbb{R}^{nd \times d}$, so that $[G^*]_i = G_i^*$ for $i = 1, \dots, n$. Note that if we multiply the target elements G_1^*, \dots, G_n^* by a common group element $Q \in \mathcal{G}$ from the right, then the resulting elements $G_1^* Q, \dots, G_n^* Q$ would yield precisely the same set of measurements since

$$G_i^* Q (G_j^* Q)^{-1} = G_i^* Q Q^{-1} G_j^{*-1} = G_i^* G_j^{*-1}. \quad (1)$$

Therefore, we can at best recover the target elements up to some unknown common transformation $Q \in \mathcal{G}$ from the right. This motivates us to define the *estimation error* $\varepsilon(G)$ of an estimator $G \in \mathcal{G}^n$ as

$$\varepsilon(G) := \min_{Q \in \mathcal{G}} \|G - G^* Q\|_F. \quad (2)$$

It is immediate from the definition that $\varepsilon(G) = \varepsilon(GQ')$ for any $Q' \in \mathcal{G}$.

The pair $([n], E)$ forms a graph, called the *measurement graph* of the synchronization problem \mathcal{G} -sync. We assume that the measurement graph is connected. This is without loss of generality, since otherwise each connected component of the measurement graph defines a synchronization problem that can be solved separately from the others. There are various ways to model the noisy observations of the pairwise ratios. One way is to adopt the *additive noise model*

$$C_{ij} = G_i^* G_j^{*-1} + \Theta_{ij}, \quad (i, j) \in E,$$

where $\{\Theta_{ij} : (i, j) \in E\}$ are the *noise matrices*. Such a model appears frequently in the group synchronization literature, see, e.g., [6, 14, 38, 47, 53, 55, 62]. Another way is to adopt the *multiplicative noise model*

$$C_{ij} = G_i^* G_j^{*-1} \Theta_{ij}, \quad (i, j) \in E,$$

see, e.g., [20, 56, 71]. A special case of this is the so-called *outlier noise model*, where Θ_{ij} ($(i, j) \in E$) is either a random element distributed uniformly (with respect to the *Haar*

measure) over \mathcal{G} or the identity element of \mathcal{G} . We will report numerical results of our proposed non-convex approach for both the additive and multiplicative noise models in Section 6.

Most existing works on computational approaches to synchronization problems assume that the noise matrices $\{\Theta_{ij} : (i, j) \in E\}$ are independent. We shall also make this assumption in our subsequent development. However, it is worth noting that such an assumption may not be ideal for recovery. Indeed, in applications such as cryo-electron microscopy [63], the noisy observations $\{C_{ij} : (i, j) \in E\}$ are usually estimated as maximizers of the cross-correlation, in which case the noise matrices are *dependent*. For approaches that address the case of dependent noise matrices, we refer the reader to, *e.g.*, [12] and the references therein.

2.2 Least Squares Formulation

For any closed subgroup \mathcal{G} of the orthogonal group $\mathcal{O}(d)$, since $Q^{-1} = Q^\top$ for any $Q \in \mathcal{G}$, we can formulate the least squares estimation problem associated with \mathcal{G} -sync as

$$\begin{aligned} \min_{G \in \mathbb{R}^{nd \times d}} \quad & \sum_{(i,j) \in E} \|[G]_i [G]_j^\top - C_{ij}\|_F^2 \\ \text{subject to} \quad & [G]_i \in \mathcal{G}, \quad i = 1, \dots, n. \end{aligned}$$

Any optimal solution to this problem is said to be a least squares estimator for the problem \mathcal{G} -sync.² By the orthogonality of the blocks $[G]_1, \dots, [G]_n$, we can rewrite the above problem as

$$\begin{aligned} \max_{G \in \mathbb{R}^{nd \times d}} \quad & \text{Tr} \left(G^\top C G \right) \\ \text{subject to} \quad & G \in \mathcal{G}^n, \end{aligned} \tag{3}$$

where $C \in \mathbb{R}^{nd \times nd}$ is the block matrix defined by

$$[C]_{ij} := \begin{cases} C_{ij}, & \text{if } (i, j) \in E, \\ C_{ji}^\top, & \text{if } (j, i) \in E, \\ I_d, & \text{if } i = j, \\ \mathbf{0}, & \text{otherwise,} \end{cases} \quad i, j = 1, \dots, n. \tag{4}$$

Problem (3) is non-convex and in general NP-hard, as it is equivalent to the MAX-CUT problem when \mathcal{G} is the Boolean group $\mathcal{O}(1) = \{+1, -1\}$. A standard idea for tackling problem (3) is to consider its SDP relaxations [2, 6, 38, 57, 71], which can be solved by off-the-shelf solvers in polynomial time. However, this approach does not scale well with n since it requires solving an SDP problem with an $nd \times nd$ matrix variable.

²By the same argument as in (1), it can be seen that the least squares estimator is not unique.

3 A Non-Convex Approach

Instead of convexification, we propose to tackle the non-convex problem (3) directly by a two-stage approach. An important subroutine in our approach is the projection of a $d \times d$ matrix onto the group \mathcal{G} . This projection, denoted by $\Pi_{\mathcal{G}}$, is defined by

$$\Pi_{\mathcal{G}}(X) := \operatorname{argmin}_{Q \in \mathcal{G}} \|X - Q\|_F = \operatorname{argmax}_{Q \in \mathcal{G}} \langle X, Q \rangle, \quad X \in \mathbb{R}^{d \times d}, \quad (5)$$

so that

$$\operatorname{dist}(X, \mathcal{G}) := \min_{Q \in \mathcal{G}} \|X - Q\|_F = \|X - \Pi_{\mathcal{G}}(X)\|_F.$$

Using the projection $\Pi_{\mathcal{G}}$, we can map any $nd \times d$ matrix to the feasible region \mathcal{G}^n of the problem \mathcal{G} -sync using the *block-wise projection* $\Pi_{\mathcal{G}}^n : \mathbb{R}^{nd \times d} \rightarrow \mathcal{G}^n$ given by³

$$[\Pi_{\mathcal{G}}^n(Y)]_i = \Pi_{\mathcal{G}}([Y]_i), \quad i = 1, \dots, n.$$

When the group \mathcal{G} is clear from the context, we omit the subscript \mathcal{G} from the symbols $\Pi_{\mathcal{G}}^n$ and $\Pi_{\mathcal{G}}$. We record a useful property of the projection Π before moving on.

Lemma 1. *For any $X \in \mathbb{R}^{d \times d}$ and any $\eta > 0$, we have*

$$\Pi(\eta \cdot X) = \Pi(X).$$

The proof of Lemma 1 is straightforward and thus omitted.

We now describe the details of the two stages of our approach. The first stage aims to find a feasible point $G^0 \in \mathcal{G}^n$ that has a sufficiently small estimation error $\varepsilon(G^0)$. Theorem 1 below provides an explicit upper bound on the estimation error $\varepsilon(G^0)$ that our non-convex approach requires. In Section 5.4, we propose a novel spectral-type estimator, called the *entropic spectral estimator*, for general group synchronization problems and show that under certain random models of the measurements and noise, the entropic spectral estimator satisfies the said upper bound on the estimation error with overwhelming probability. Moreover, the entropic spectral estimator can be computed efficiently. Hence, it can be used as G^0 in the first stage.

In the second stage, starting with the initial point G^0 obtained from the first stage, we iteratively refine the estimates using GPM, which is described below.

Algorithm 1 Generalized Power Method (GPM) for \mathcal{G} -sync

- 1: **Input:** the matrix C and an initial point G^0
 - 2: **for** $t = 0, 1, 2, \dots$ **do**
 - 3: $G^{t+1} = \Pi^n(CG^t)$
 - 4: **end for**
-

³Strictly speaking, the maps $\Pi_{\mathcal{G}}$ and $\Pi_{\mathcal{G}}^n$ are set-valued since in general problem (5) can have multiple minimizers. Nevertheless, all the results in this paper hold for any of the minimizers.

The name “generalized power method” of Algorithm 1 comes from its close resemblance to the classic power method [58] for computing the dominant eigenvector of a matrix. In fact, problem (3) can be seen as a constrained eigenvalue problem. More precisely, each column v of G is not only normalized to a fixed length (indeed, $\|v\|_2 = \sqrt{n}$) and orthogonal to all other columns as in the usual eigenvalue problem, but also subject to the extra constraint that the d -dimensional sub-vectors $v_{1:d}, v_{d+1:2d}, \dots, v_{(n-1)d+1:nd}$ are all of unit length (here, we use the MATLAB notation $v_{k:\ell}$ to denote the sub-vector $(v_k, \dots, v_\ell)^\top$ of v). Despite this resemblance, the analysis of GPM for solving group synchronization problems is highly non-trivial due to the said extra constraint.

Overall, the proposed non-convex approach enjoys several computational advantages and is very efficient. For the first stage, the main cost of computing the entropic spectral estimator lies in the computation of the first d eigenvectors of the matrix C and the generation of independent copies of uniformly random orthogonal matrices, which can be done efficiently by a host of modern eigen-solvers [46, 58] and random orthogonal matrix samplers [29, 52, 65], respectively. For the second stage, we can see from Algorithm 1 that GPM does not require the tuning of any parameter and can be implemented extremely easily. The computational cost in each iteration consists of two parts: (i) d matrix-vector multiplications for forming the product CG^t and (ii) the block-wise projection Π^n , which can be decomposed into n projections Π onto the group \mathcal{G} . As we will see in Section 5.1, for the orthogonal group $\mathcal{O}(d)$, the special orthogonal group $\mathcal{SO}(d)$, and the permutation group $\mathcal{P}(d)$, the projection Π reduces to a singular value decomposition (SVD) or a d -dimensional linear programming problem; for the cyclic group \mathcal{Z}_m , the projection Π can be computed using a simple, explicit formula involving trigonometric functions. Thus, the two parts mentioned above are well-suited for parallelization and can be implemented efficiently for many closed subgroups of the orthogonal group.

4 A Master Theorem

In this section, we present a master theorem on the estimation performance of our proposed non-convex approach. We begin with some definitions. Let $\bar{E} = E \cup \{(1, 1), \dots, (n, n)\}$ and consider the *extended measurement graph* $([n], \bar{E})$. Every node in this graph has precisely one self-loop. For $i, j \in [n]$, let $w_{ij} = 1$ if either $(i, j) \in \bar{E}$ or $(j, i) \in \bar{E}$; otherwise let $w_{ij} = 0$. The degree r_i of node $i \in [n]$ of the extended measurement graph is then given by $r_i = \sum_{j=1}^n w_{ij}$. Note that for any $i \in [n]$, we have $r_i \geq 1$ since $w_{ii} = 1$. Let $\Delta \in \mathbb{R}^{nd \times nd}$ be the block matrix defined by

$$[\Delta]_{ij} := w_{ij} \left([C]_{ij} - [G^* G^{*\top}]_{ij} \right) = \begin{cases} C_{ij} - G_i^* G_j^{*\top}, & \text{if } (i, j) \in E, \\ C_{ji}^\top - G_i^* G_j^{*\top}, & \text{if } (j, i) \in E, \\ \mathbf{0}, & \text{otherwise,} \end{cases} \quad i, j = 1, \dots, n, \quad (6)$$

where $C \in \mathbb{R}^{nd \times nd}$ is the matrix defined in (4) and $G^* \in \mathcal{G}^n$ is the ground truth. Furthermore, let $\bar{D} := \text{Diag}(r_1, \dots, r_n) \in \mathbb{R}^{n \times n}$, $\bar{W} \in \mathbb{R}^{n \times n}$ be the matrix defined by $\bar{W}_{ij} = w_{ij}$, $i, j = 1, \dots, n$, and $e \in \mathbb{R}^n$ be the all-one vector. Set $D = \bar{D} \otimes I_d \in \mathbb{R}^{nd \times nd}$, $W = \bar{W} \otimes I_d \in \mathbb{R}^{nd \times nd}$, and $F = ee^\top \otimes I_d \in \mathbb{R}^{nd \times nd}$. We define the following parameter for the measurement graph:

$$\kappa := \left\| D^{-1}W - \frac{1}{n}F \right\|.$$

The parameter κ is related to the connectedness of the measurement graph $([n], E)$, see the discussion after Theorem 1. For any $G \in \mathcal{G}^n$, we define

$$Q_G := \underset{Q \in \mathcal{G}}{\operatorname{argmin}} \|G - G^*Q\|_F = \underset{Q \in \mathcal{G}}{\operatorname{argmax}} \left\langle G^{*\top}G, Q \right\rangle = \Pi \left(G^{*\top}G \right), \quad (7)$$

where the last equality follows from the definition of Π in (5). In particular, we have $\varepsilon(G) = \|G - G^*Q_G\|_F$. For simplicity, we write $Q^t = Q_{G^t}$ for all $t \geq 0$.

The following theorem is the first main theoretical result of this paper, which provides a bound on the rate at which the estimation error of the iterates generated by the proposed non-convex approach decays under certain assumptions on the subgroup, measurement graph, noise, and initialization.

Theorem 1 (Master Theorem for GPM). *Suppose that*

(i) *there exists a constant $\alpha \geq 1$ such that for any $t \geq 0$,*

$$\|G^{*\top}G^t - n \cdot Q^t\|_F \leq \frac{\alpha}{2} \|G^t - G^*Q^t\|_F^2;$$

(ii) $\kappa \leq \frac{1}{32}$;

(iii) $\|D^{-1}\Delta\| \leq \frac{1}{32}$ and $\|D^{-1}\Delta G^*\|_F \leq \frac{\sqrt{n}}{32\alpha}$;

(iv) $\varepsilon(G^0) \leq \frac{\sqrt{n}}{2\alpha}$.

Then, the iterates $\{G^t\}_{t \geq 1}$ generated by GPM satisfy

$$\varepsilon(G^t) \leq \left(\frac{5}{8}\right)^{t+1} \varepsilon(G^0) + \frac{16}{3} \|D^{-1}\Delta G^*\|_F, \quad \forall t \geq 1.$$

Let us elaborate on conditions (i)–(iv) in and the implications of Theorem 1 before giving its proof. Condition (i) reflects the geometry of the subgroup \mathcal{G} through the projection map $\Pi_{\mathcal{G}}$. Indeed, as we shall see in Section 5.1, it can be interpreted as an *error bound condition*, which provides a measure of proximity of points in the convex hull of \mathcal{G} to \mathcal{G} itself. We will also show in Section 5.1 that condition (i) holds with $\alpha = 1$ when \mathcal{G} is the orthogonal group $\mathcal{O}(d)$, the special orthogonal group $\mathcal{SO}(d)$, or the permutation group $\mathcal{P}(d)$. We

conjecture that condition (i) actually holds for arbitrary closed subgroups of the orthogonal group, see Conjecture 1. Condition (ii) is related to the measurement graph $([n], E)$ and represents the requirement that the measurement graph needs to be sufficiently connected. In particular, we have $\kappa = 0$ when the measurement graph is complete (*i.e.*, $w_{ij} = 1$ for all $1 \leq i < j \leq n$). Condition (iii) depends on normalized deviation matrix $D^{-1}\Delta$ and, roughly speaking, requires that the noise in the measurements cannot be too large. This condition suggests that when we quantify the information contained in the observations $\{C_{ij} : (i, j) \in E\}$, we should normalized the deviations matrices $\{\Delta_{ij} : (i, j) \in E\}$ by the degrees $\{r_i : i \in [n]\}$ of the nodes in the (extended) measurement graph. This is consistent with our intuition that under the same level of noise, more information is available at node i if there are more observations related to G_i^* . For the special case of $\mathcal{SO}(2)$ -sync with a complete measurement graph, a similar condition is used in the analysis of the SDP relaxation approach, see [6, Definition 3.1]. Condition (iv) captures the requirement that GPM needs to be initialized with a point G^0 of sufficiently small estimation error.

Under the setting of Theorem 1, we see that any accumulation point G^∞ of the sequence of iterates $\{G^t\}_{t \geq 0}$ generated by GPM satisfy

$$\varepsilon(G^\infty) \leq \frac{16}{3} \|D^{-1}\Delta G^*\|_F. \quad (8)$$

In particular, in the absence of noise (*i.e.*, $\Delta = \mathbf{0}$), every accumulation point of the sequence $\{G^t\}_{t \geq 0}$ is, up to some common transformation, equal to the ground truth G^* . For the general case, we have $\varepsilon(G^\infty) \leq \frac{\sqrt{n}}{6\alpha}$ by inequality (8) and condition (iii). Although such a bound does not seem to improve by much the bound $\frac{\sqrt{n}}{2\alpha}$ on the estimation error of the initial point G^0 , the numerical constants $\frac{1}{32}$, $\frac{5}{8}$, and $\frac{16}{3}$ in the statement of Theorem 1 are not important, as they can be improved by bootstrapping the analysis. After all, the condition $\|D^{-1}\Delta G^*\|_F \leq \frac{\sqrt{n}}{32\alpha}$ used to obtain the bound is only a very weak requirement for the conclusion of Theorem 1 to hold and may not reflect the true magnitude of $\|D^{-1}\Delta G^*\|_F$. The key message conveyed by Theorem 1 is that the estimation error of the iterates produced by the suitably initialized GPM decreases at least geometrically to some parameter characterized by the measurement graph and noise. It is also instructive to specialize Theorem 1 to the case of $\mathcal{SO}(2)$ -sync with a complete measurement graph. A simple calculation yields $D = nI_{2n}$ and $\|D^{-1}\Delta G^*\|_F \leq \sqrt{\frac{2}{n}}\|\Delta\|$, which implies that $\varepsilon(G^\infty)$ is on the order of $\frac{\|\Delta\|}{\sqrt{n}}$. Under the Gaussian noise model, this matches the order of the Cramér-Rao lower bound on the estimation error of any unbiased estimator of G^* , see the discussion in [47, Section 1].

Note that Theorem 1 does not guarantee the convergence of the sequence $\{G^t\}_{t \geq 0}$ generated by GPM. In the recent paper [44], which appeared after this paper was posted on arXiv, Ling considered the problem $\mathcal{O}(d)$ -sync under the setting of additive Gaussian noise and complete measurement graph and showed that if the standard deviation of the noise is on the order of $\frac{\sqrt{n}}{\sqrt{d}(\sqrt{d} + \sqrt{\log n})}$, then GPM, when initialized by a spectral estimator,

will generate iterates that converge linearly to a maximum likelihood estimator (which coincides with an optimal solution to the non-convex least squares formulation) with high probability. However, whether such a result can be extended to settings involving other closed subgroups of $\mathcal{O}(d)$ or more general measurement graphs or deterministic noise models remains open.

4.1 Proof of Theorem 1

The proof of Theorem 1 requires Lemma 2 below, which asserts that the projection map Π satisfies certain contraction-type property and is a generalization of [47, Proposition 3.3] from $\mathcal{SO}(2)$ to arbitrary closed subgroups of the orthogonal group. Since [47, Proposition 3.3] has already been applied in a number of works to study phase synchronization problems [79] or even other estimation problems [28, 69], we believe that Lemma 2 will find further applications in synchronization problems or other estimation/optimization problems over general subgroups of the orthogonal group. Furthermore, the proof of Lemma 2 is different from and much simpler than that of [47, Proposition 3.3].

Lemma 2. *For any $r > 0$, $X \in \mathbb{R}^{d \times d}$, and $Q \in \mathcal{G}$,*

$$\|\Pi(X) - Q\|_F \leq 2 \|r^{-1}X - Q\|_F.$$

Proof. Using Lemma 1 and the definition of Π , we have

$$\begin{aligned} \|\Pi(X) - Q\|_F &= \|\Pi(r^{-1}X) - Q\|_F \\ &\leq \|\Pi(r^{-1}X) - r^{-1}X\|_F + \|r^{-1}X - Q\|_F \\ &\leq \|Q - r^{-1}X\|_F + \|r^{-1}X - Q\|_F \\ &= 2 \|r^{-1}X - Q\|_F, \end{aligned}$$

which completes the proof. \square

The proof of Theorem 1 also relies on the following technical lemma:

Lemma 3. *Let $\alpha \geq 1$ be given. If for some $t \geq 0$,*

$$\left\| G^{*\top} G^t - n \cdot Q^t \right\|_F \leq \frac{\alpha}{2} \|G^t - G^* Q^t\|_F^2,$$

then

$$\varepsilon(G^{t+1}) \leq 2 \left(\left(\frac{\alpha \cdot \varepsilon(G^t)}{2\sqrt{n}} + \kappa + \|D^{-1}\Delta\| \right) \varepsilon(G^t) + \|D^{-1}\Delta G^*\|_F \right).$$

Proof. Using Lemma 2 and the fact that $G_i^* Q^t \in \mathcal{G}$, we have, for any $t \geq 1$ and $i \in [n]$, that

$$\|\Pi([CG^t]_i) - G_i^* Q^t\|_F \leq 2 \|r_i^{-1} [CG^t]_i - G_i^* Q^t\|_F. \quad (9)$$

Hence,

$$\begin{aligned}
\varepsilon(G^{t+1}) &= \min_{Q \in \mathcal{G}} \|G^{t+1} - G^*Q\|_F \leq \|G^{t+1} - G^*Q^t\|_F = \|\Pi^n(CG^t) - G^*Q^t\|_F \\
&\leq 2 \|D^{-1}CG^t - G^*Q^t\|_F \leq 2 (\|D^{-1}(C - \Delta)G^t - G^*Q^t\|_F + \|D^{-1}\Delta G^t\|_F) \\
&\leq 2 (\|D^{-1}(C - \Delta)G^t - G^*Q^t\|_F + \|D^{-1}\Delta(G^t - G^*Q^t)\|_F + \|D^{-1}\Delta G^*Q^t\|_F) \\
&\leq 2 (\|D^{-1}(C - \Delta)G^t - G^*Q^t\|_F + \|D^{-1}\Delta\| \cdot \varepsilon(G^t) + \|D^{-1}\Delta G^*\|_F),
\end{aligned} \tag{10}$$

where the second inequality follows from (9). Using the definitions of C in (4) and Δ in (6), it is easy to verify that $[C - \Delta]_{ij} = w_{ij}G_i^*G_j^{*\top}$ for $i, j \in [n]$. It follows that

$$\begin{aligned}
&\|D^{-1}(C - \Delta)G^t - G^*Q^t\|_F \\
&= \|D^{-1} \cdot \text{BlkDiag}(G_1^*, \dots, G_n^*) \cdot W \cdot \text{BlkDiag}(G_1^{*\top}, \dots, G_n^{*\top}) \cdot G^t - G^*Q^t\|_F \\
&= \|\text{BlkDiag}(G_1^*, \dots, G_n^*) \left((D^{-1}W) \cdot \text{BlkDiag}(G_1^{*\top}, \dots, G_n^{*\top}) \cdot G^t - e \otimes Q^t \right)\|_F \\
&= \|D^{-1}W \left(\text{BlkDiag}(G_1^{*\top}, \dots, G_n^{*\top}) \cdot G^t - e \otimes Q^t \right)\|_F \\
&\leq \frac{1}{n} \|F \left(\text{BlkDiag}(G_1^{*\top}, \dots, G_n^{*\top}) \cdot G^t - e \otimes Q^t \right)\|_F \\
&\quad + \left\| \left(D^{-1}W - \frac{1}{n}F \right) \left(\text{BlkDiag}(G_1^{*\top}, \dots, G_n^{*\top}) \cdot G^t - e \otimes Q^t \right) \right\|_F \\
&\leq \frac{1}{\sqrt{n}} \|G^{*\top}G^t - n \cdot Q^t\|_F + \kappa \cdot \|G^t - G^*Q^t\|_F \\
&\leq \left(\frac{\alpha \cdot \varepsilon(G^t)}{2\sqrt{n}} + \kappa \right) \varepsilon(G^t),
\end{aligned} \tag{11}$$

where the second equality is due to the fact that D^{-1} and $\text{BlkDiag}(G_1^*, \dots, G_n^*)$ commute and the identity

$$G^*Q^t = \text{BlkDiag}(G_1^*, \dots, G_n^*) \cdot (e \otimes Q^t),$$

the third equality is due to the fact that $D^{-1}W(e \otimes Q^t) = e \otimes Q^t$, the second-to-last inequality is due to the identities

$$\begin{aligned}
F \left(\text{BlkDiag}(G_1^{*\top}, \dots, G_n^{*\top}) \cdot G^t - e \otimes Q^t \right) &= e \otimes (G^{*\top}G^t - n \cdot Q^t), \\
\left\| \text{BlkDiag}(G_1^{*\top}, \dots, G_n^{*\top}) \cdot G^t - e \otimes Q^t \right\|_F &= \|G^t - \text{BlkDiag}(G_1^*, \dots, G_n^*) \cdot (e \otimes Q^t)\|_F \\
&= \|G^t - G^*Q^t\|_F,
\end{aligned}$$

and the last inequality is due to the assumption of the lemma and the definition of Q^t . The desired result now follows by putting (10) and (11) together. \square

Armed with Lemma 3, we can now prove the master theorem.

Proof of Theorem 1. We first show by induction that

$$\varepsilon(G^t) \leq \frac{\sqrt{n}}{2\alpha}, \quad \forall t \geq 0. \quad (12)$$

For $t = 0$, it follows from condition (iv) that $\varepsilon(G^0) \leq \frac{\sqrt{n}}{2\alpha}$. Next, we assume that $\varepsilon(G^t) \leq \frac{\sqrt{n}}{2\alpha}$ for some $t \geq 0$. By Lemma 3, conditions (ii)–(iii), and the inductive hypothesis, we have

$$\begin{aligned} \varepsilon(G^{t+1}) &\leq 2 \left(\left(\frac{\alpha \cdot \varepsilon(G^t)}{2\sqrt{n}} + \kappa + \|D^{-1}\Delta\| \right) \varepsilon(G^t) + \|D^{-1}\Delta G^*\|_F \right) \\ &\leq 2 \left(\frac{1}{4} + \frac{1}{32} + \frac{1}{32} \right) \frac{\sqrt{n}}{2\alpha} + \frac{\sqrt{n}}{16\alpha} < \frac{\sqrt{n}}{2\alpha}, \end{aligned}$$

which yields (12). Using Lemma 3, conditions (ii)–(iii), and inequality (12), we have

$$\begin{aligned} \varepsilon(G^{t+1}) &\leq 2 \left(\left(\frac{\alpha \cdot \varepsilon(G^t)}{2\sqrt{n}} + \kappa + \|D^{-1}\Delta\| \right) \varepsilon(G^t) + \|D^{-1}\Delta G^*\|_F \right) \\ &\leq \frac{5}{8} \varepsilon(G^t) + 2 \|D^{-1}\Delta G^*\|_F. \end{aligned}$$

Unrolling this recurrence gives

$$\varepsilon(G^{t+1}) \leq \left(\frac{5}{8} \right)^{t+1} \varepsilon(G^0) + \frac{16}{3} \|D^{-1}\Delta G^*\|_F,$$

which completes the proof. \square

In the next section, we show that condition (i) in Theorem 1 is satisfied by various closed subgroups of the orthogonal group, while conditions (ii) and (iii) are satisfied by certain random measurement graph and random noise models. Moreover, we propose a novel spectral-type estimator and show that for the said subgroups and under the said random measurement graph and noise models, condition (iv) in Theorem 1 is satisfied by the proposed estimator. These results demonstrate the utility and power of Theorem 1.

5 Verifying the Conditions of the Master Theorem

5.1 Error-Bound Geometry of the Subgroup \mathcal{G}

In order to apply the master theorem in the last section, the subgroup \mathcal{G} has to satisfy certain geometric conditions. In this section, we formulate these conditions and verify them for four specific subgroups, namely the orthogonal group $\mathcal{O}(d)$, the special orthogonal group $\mathcal{SO}(d)$, the permutation group $\mathcal{P}(d)$, and the cyclic group \mathcal{Z}_m . Along the way, we show that the projection maps associated with these subgroups can be computed in a tractable manner.

To begin, let $\text{conv}(\mathcal{G})$ denote the convex hull of the subgroup \mathcal{G} . It is easy to check that a point $X \in \text{conv}(\mathcal{G})$ lies in the subgroup \mathcal{G} if and only if any (and hence all) of the following three quantities vanishes:

$$\text{dist}(X, \mathcal{G}), \quad \text{Tr}(I_d - X^\top \Pi_{\mathcal{G}}(X)), \quad d - \|X\|_F^2.$$

Therefore, these quantities can serve as proximity measures to \mathcal{G} for points in $\text{conv}(\mathcal{G})$. We now investigate the connection between our non-convex approach to the geometry of \mathcal{G} , as reflected through the ratios of these proximity measures. The precise geometric conditions are given as follows.

Condition 1. *There exists a constant $\alpha \geq 1$ such that*

$$\text{dist}(X, \mathcal{G}) \leq \alpha \text{Tr}(I_d - X^\top \Pi_{\mathcal{G}}(X)), \quad \forall X \in \text{conv}(\mathcal{G}).$$

Condition 2. *There exists a constant $\beta \in (0, 1]$ such that*

$$\beta \text{Tr}(I_d - X^\top \Pi_{\mathcal{G}}(X)) \leq d - \|X\|_F^2, \quad \forall X \in \text{conv}(\mathcal{G}).$$

The above two conditions play an important role in our development. Indeed, if we take $X = \frac{1}{n} G^{*\top} G$, then Condition 1 immediately implies that

$$\left\| G^{*\top} G - n \cdot Q_G \right\|_F \leq \alpha \text{Tr} \left(n \cdot I_d - Q_G^\top G^{*\top} G \right) = \frac{\alpha}{2} \|G - G^* Q_G\|_F^2,$$

which is precisely condition (i) in the master theorem. The usefulness of Condition 2 will become clear when we study the initialization for GPM in Section 5.4.

It is worth noting that Conditions 1 and 2 are reminiscent of *error bound conditions*, which have been extensively studied in the optimization literature and applied to analyze the convergence rates of various iterative methods, see, *e.g.*, [46, 72, 77, 78, 80] and the references therein. Roughly speaking, an error bound condition postulates that the distance function associated with some target set (often difficult to characterize theoretically and not computable in practice) is bounded above by a continuous surrogate function (often easier to characterize theoretically and compute in practice) that vanishes on the target set. In the context of synchronization problems, such a condition was first introduced in [47] to study the optimization performance of GPM.

The main result of this subsection is summarized in the following theorem, which will be proved in a case-by-case manner.

Theorem 2. *For $\mathcal{G} = \mathcal{O}(d)$ or $\mathcal{G} = \mathcal{SO}(d)$, the projection $\Pi_{\mathcal{G}}$ can be computed in closed form via the SVD of a $d \times d$ matrix; for $\mathcal{G} = \mathcal{P}(d)$, the projection $\Pi_{\mathcal{G}}$ can be computed by solving a d -dimensional linear programming problem; for $\mathcal{G} = \mathcal{Z}_m$, the projection $\Pi_{\mathcal{G}}$ can*

be computed in closed form via the formula in Proposition 2. Moreover, Conditions 1 and 2 hold for the groups $\mathcal{O}(d)$, $\mathcal{SO}(d)$, $\mathcal{P}(d)$, and \mathcal{Z}_m with

$$\alpha = \begin{cases} 1, & \text{if } \mathcal{G} = \mathcal{O}(d), \mathcal{SO}(d), \text{ or } \mathcal{P}(d), \\ 1, & \text{if } \mathcal{G} = \mathcal{Z}_m \text{ with } m = 1, 2, \\ (\sqrt{2} \sin \frac{\pi}{m})^{-1}, & \text{if } \mathcal{G} = \mathcal{Z}_m \text{ with } m \geq 3 \end{cases}$$

and

$$\beta = \begin{cases} 1, & \text{if } \mathcal{G} = \mathcal{O}(d), \\ \frac{1}{2}, & \text{if } \mathcal{G} = \mathcal{SO}(d), \\ \frac{2}{d}, & \text{if } \mathcal{G} = \mathcal{P}(d), \\ 1, & \text{if } \mathcal{G} = \mathcal{Z}_m \text{ with } m = 1, \\ \sin^2 \frac{\pi}{m}, & \text{if } \mathcal{G} = \mathcal{Z}_m \text{ with } m \geq 2. \end{cases}$$

Motivated by Theorem 2, we formulate the following conjecture:

Conjecture 1. *Let \mathcal{G} be any closed subgroup of the orthogonal group $\mathcal{O}(d)$. Then, there exist constants $\alpha \geq 1$ and $\beta \in (0, 1]$ such that*

$$\frac{\text{dist}(X, \mathcal{G})}{\alpha} \leq \text{Tr}(I_d - X^\top \Pi_{\mathcal{G}}(X)) \leq \frac{d - \|X\|_F^2}{\beta}, \quad \forall X \in \text{conv}(\mathcal{G}).$$

The validity of Conjecture 1 would imply that condition (i) in the master theorem is superfluous.

5.1.1 Proof of Theorem 2: Orthogonal Group $\mathcal{O}(d)$

The projection $\Pi_{\mathcal{O}(d)}$ is given by the solution to the well-known *orthogonal Procrustes problem* [32], i.e., for any $X \in \mathbb{R}^{d \times d}$ with SVD $U_X \Sigma_X V_X^\top$, we have

$$\Pi_{\mathcal{O}(d)}(X) = U_X V_X^\top.$$

Now, let $X \in \text{conv}(\mathcal{O}(d))$ be arbitrary. By Carathéodory's theorem, there exist $Q_1, \dots, Q_L \in \mathcal{O}(d)$ and $\omega_1, \dots, \omega_L \geq 0$ such that

$$\sum_{\ell=1}^L \omega_\ell = 1 \quad \text{and} \quad X = \sum_{\ell=1}^L \omega_\ell Q_\ell.$$

Then, for any $j \in [d]$,

$$\sigma_j(X) = \sigma_j \left(\sum_{\ell=1}^L \omega_\ell Q_\ell \right) \leq \sum_{\ell=1}^L \omega_\ell \sigma_j(Q_\ell) = \sum_{\ell=1}^L \omega_\ell = 1, \quad (13)$$

where $\sigma_j(\cdot)$ denotes the j -th largest singular value. This implies that $I_d - \Sigma_X$ is positive semidefinite and

$$\begin{aligned} \text{dist}(X, \mathcal{O}(d)) &= \|X - \Pi_{\mathcal{O}(d)}(X)\|_F = \|U_X \Sigma_X V_X^\top - U_X V_X^\top\|_F = \|I_d - \Sigma_X\|_F \\ &\leq \text{Tr}(I_d - \Sigma_X) = \text{Tr}(I_d - V_X \Sigma_X V_X^\top) = \text{Tr}(I_d - X^\top \Pi_{\mathcal{O}(d)}(X)). \end{aligned}$$

Thus, Condition 1 holds for $\mathcal{O}(d)$ with $\alpha = 1$. Furthermore, we have

$$\begin{aligned} \text{Tr}(I_d - X^\top \Pi_{\mathcal{O}(d)}(X)) &= \text{Tr}(I_d - \Sigma_X) = \sum_{j=1}^d (1 - \sigma_j(X)) \\ &\leq \sum_{j=1}^d (1 - \sigma_j(X))(1 + \sigma_j(X)) = \sum_{j=1}^d (1 - \sigma_j^2(X)) = d - \|X\|_F^2, \end{aligned} \tag{14}$$

which shows that Condition 2 holds for $\mathcal{O}(d)$ with $\beta = 1$.

5.1.2 Proof of Theorem 2: Special Orthogonal Group $\mathcal{SO}(d)$

We adopt a similar strategy to the one used in Section 5.1.1 to establish Condition 1 for $\mathcal{SO}(d)$. First, we have the following result, which gives an explicit formula for the projection $\Pi_{\mathcal{SO}(d)}$ and is known as the *Kabsch algorithm* [40].

Lemma 4. *Let $X \in \mathbb{R}^{d \times d}$ be a matrix with SVD $X = U_X \Sigma_X V_X^\top$. Let $I_X \in \mathbb{R}^{d \times d}$ be the diagonal matrix defined by $I_X = \text{Diag}(1, \dots, 1, \det(U_X V_X^\top))$. Then, we have*

$$\Pi_{\mathcal{SO}(d)}(X) = U_X I_X V_X^\top.$$

Now, let $X \in \text{conv}(\mathcal{SO}(d))$ be arbitrary. Using Lemma 4, we have

$$\begin{aligned} \text{dist}(X, \mathcal{SO}(d)) &= \|X - \Pi_{\mathcal{SO}(d)}(X)\|_F = \|U_X \Sigma_X V_X^\top - U_X I_X V_X^\top\|_F = \|I_d - \Sigma_X I_X\|_F \\ &\leq \text{Tr}(I_d - \Sigma_X I_X) = \text{Tr}(I_d - V_X \Sigma_X I_X V_X^\top) = \text{Tr}(I_d - X^\top \Pi_{\mathcal{SO}(d)}(X)), \end{aligned}$$

where the inequality follows from (13). This shows that Condition 1 holds for $\mathcal{SO}(d)$ with $\alpha = 1$.

To establish Condition 2 for $\mathcal{SO}(d)$, let $X \in \text{conv}(\mathcal{SO}(d))$ be arbitrary and consider first the case where $\det(U_X V_X) = 1$. Then, we have $I_X = I_d$ and $\Pi_{\mathcal{SO}(d)}(X) = \Pi_{\mathcal{O}(d)}(X)$. It follows from (14) that

$$\text{Tr}(I_d - X^\top \Pi_{\mathcal{SO}(d)}(X)) = \text{Tr}(I_d - X^\top \Pi_{\mathcal{O}(d)}(X)) \leq d - \|X\|_F^2.$$

Next, consider the case where $\det(U_X V_X) = -1$, *i.e.*, $\Pi_{\mathcal{O}(d)}(X) \in \mathcal{O}(d) \setminus \mathcal{SO}(d)$. Since $X \in \text{conv}(\mathcal{SO}(d))$, by Carathéodory's theorem, there exist $Q_1, \dots, Q_L \in \mathcal{SO}(d)$ and $\omega_1, \dots, \omega_L \geq 0$ such that

$$\sum_{\ell=1}^L \omega_\ell = 1 \quad \text{and} \quad X = \sum_{\ell=1}^L \omega_\ell Q_\ell.$$

Hence, we have

$$d - \|X\|_F^2 \geq \text{Tr}(I_d - X^\top \Pi_{\mathcal{O}(d)}(X)) = \sum_{\ell=1}^L \omega_\ell \sum_{j=1}^d (1 - \lambda_j(Q_\ell^\top \Pi_{\mathcal{O}(d)}(X))) \geq \sum_{\ell=1}^L \omega_\ell \cdot 2 = 2, \quad (15)$$

where the first inequality follows from (14) and the second inequality follows from the fact that $Q_\ell^\top \Pi_{\mathcal{O}(d)}(X) \in \mathcal{O}(d) \setminus \mathcal{SO}(d)$ and hence the eigenvalues of $Q_\ell^\top \Pi_{\mathcal{O}(d)}(X)$ are either $+1$ or -1 with at least one being -1 . This gives

$$\begin{aligned} \text{Tr}(I_d - X^\top \Pi_{\mathcal{SO}(d)}(X)) &= \text{Tr}(I_d - \Sigma_X I_X) = 1 + \sigma_d(X) + \sum_{j=1}^{d-1} (1 - \sigma_j(X)) \\ &\leq 2 + \sum_{j=1}^{d-1} (1 - \sigma_j^2(X)) \leq 2 + \sum_{j=1}^d (1 - \sigma_j^2(X)) \leq 2 \left(d - \|X\|_F^2 \right), \end{aligned}$$

where the first equality follows from Lemma 4, the second equality follows from the assumption that $\det(U_X V_X) = -1$, the first inequality follows from the fact that $\sigma_j(X) \leq 1$ for $j = 1, \dots, d$, and the last inequality follows from (15). This shows that Condition 2 holds for $\mathcal{SO}(d)$ with $\beta = \frac{1}{2}$.

5.1.3 Proof of Theorem 2: Permutation Group $\mathcal{P}(d)$

It is easy to verify that the projection $\Pi_{\mathcal{P}(d)}$ is given by an optimal solution to the *assignment problem* [19], *i.e.*, for any $X \in \mathbb{R}^{d \times d}$, we have

$$\Pi_{\mathcal{P}(d)}(X) = \underset{Q \in \mathcal{P}(d)}{\text{argmax}} \langle Q, X \rangle.$$

The above problem can be solved in polynomial time by linear programming or the Hungarian (also known as the Kuhn–Munkres) algorithm.

To establish Condition 1 for $\mathcal{P}(d)$, we first note that for any $Q \in \mathcal{P}(d)$, we either have

$d = \text{Tr}(Q)$ or $d - \text{Tr}(Q) \geq 2$. It follows that

$$\begin{aligned} \|I_d - Q\|_F^2 &= \sum_{i=1}^d \sum_{j=1}^d (I_d - Q)_{ij}^2 = \sum_{i=1}^d \sum_{j=1}^d |(I_d - Q)_{ij}| \\ &= \sum_{i \neq j} Q_{ij} + \sum_{i=1}^d (1 - Q_{ii}) = 2d - 2\text{Tr}(Q) \leq (d - \text{Tr}(Q))^2. \end{aligned} \quad (16)$$

For any $Y \in \text{conv}(\mathcal{P}(d))$, by Carathéodory's theorem, there exist $Q_1, \dots, Q_L \in \mathcal{P}(d)$ and $\omega_1, \dots, \omega_L \geq 0$ such that

$$\sum_{\ell=1}^L \omega_\ell = 1 \quad \text{and} \quad Y = \sum_{\ell=1}^L \omega_\ell Q_\ell.$$

Using (16), we get

$$\|I_d - Y\|_F \leq \sum_{\ell=1}^L \omega_\ell \|I_d - Q_\ell\|_F \leq \sum_{\ell=1}^L \omega_\ell (d - \text{Tr}(Q_\ell)) = d - \text{Tr}(Y). \quad (17)$$

Now, let $X \in \text{conv}(\mathcal{P}(d))$ be arbitrary. Since $\Pi_{\mathcal{P}(d)}(X)^\top X \in \text{conv}(\mathcal{P}(d))$, by invoking (17) with $Y = \Pi_{\mathcal{P}(d)}(X)^\top X$, we obtain

$$\begin{aligned} \text{dist}(X, \mathcal{P}(d)) &= \|X - \Pi_{\mathcal{P}(d)}(X)\|_F = \|\Pi_{\mathcal{P}(d)}(X)^\top X - I_d\|_F \\ &\leq d - \text{Tr}(\Pi_{\mathcal{P}(d)}(X)^\top X) = \text{Tr}(I_d - X^\top \Pi_{\mathcal{P}(d)}(X)). \end{aligned}$$

This shows that Condition 1 holds for $\mathcal{P}(d)$ with $\alpha = 1$.

To establish Condition 2 for $\mathcal{P}(d)$, we prove the following stronger result, which states that Condition 2 actually holds for any discrete subgroup of $\mathcal{O}(d)$.

Proposition 1. *Let \mathcal{G} be a discrete subgroup of $\mathcal{O}(d)$ with at least two elements. Then, Condition 2 holds for \mathcal{G} with*

$$\beta = \frac{d - \max_{Q, Q' \in \mathcal{G}, Q \neq Q'} \text{Tr}(Q^\top Q')}{d - \min_{Q, Q' \in \mathcal{G}, Q \neq Q'} \text{Tr}(Q^\top Q')}.$$

Proof. The compactness of $\mathcal{O}(d)$ implies that \mathcal{G} must be finite. Hence, we may let $\mathcal{G} = \{Q_1, \dots, Q_L\}$ with $Q_1, \dots, Q_L \in \mathcal{O}(d)$. For any $X \in \text{conv}(\mathcal{G})$, there exist $\omega_1, \dots, \omega_L \geq 0$ such that

$$\sum_{\ell=1}^L \omega_\ell = 1 \quad \text{and} \quad X = \sum_{\ell=1}^L \omega_\ell Q_\ell.$$

Without loss of generality, suppose that $\omega_1 = \max_{\ell \in [L]} \omega_\ell$. Then, we have

$$\max_{Q \in \mathcal{G}} \text{Tr}(X^\top Q) \geq \text{Tr}(X^\top Q_1) = \text{Tr}\left(\sum_{\ell=1}^L \omega_\ell Q_\ell^\top Q_1\right) \geq \omega_1 d + (1 - \omega_1) \min_{i \neq j} \text{Tr}(Q_i^\top Q_j),$$

which implies that

$$\text{Tr}(I_d - X^\top \Pi_{\mathcal{G}}(X)) = d - \max_{Q \in \mathcal{G}} \text{Tr}(X^\top Q) \leq (1 - \omega_1) \left(d - \min_{i \neq j} \text{Tr}(Q_i^\top Q_j)\right). \quad (18)$$

On the other hand, we have

$$\begin{aligned} \|X\|_F^2 &= \sum_{\ell=1}^L \omega_\ell^2 \|Q_\ell\|_F^2 + \sum_{\ell' \neq \ell} \omega_\ell \omega_{\ell'} \text{Tr}(Q_{\ell'}^\top Q_\ell) \leq \sum_{\ell=1}^L \omega_\ell^2 d + \sum_{\ell' \neq \ell} \omega_\ell \omega_{\ell'} \max_{i \neq j} \text{Tr}(Q_i^\top Q_j) \\ &\leq \sum_{\ell=1}^L \omega_\ell^2 d + \left(1 - \sum_{\ell=1}^L \omega_\ell^2\right) \max_{i \neq j} \text{Tr}(Q_i^\top Q_j) = d - \left(d - \max_{i \neq j} \text{Tr}(Q_i^\top Q_j)\right) \left(1 - \sum_{\ell=1}^L \omega_\ell^2\right) \\ &\leq d - \left(d - \max_{i \neq j} \text{Tr}(Q_i^\top Q_j)\right) (1 - \omega_1), \end{aligned}$$

which, upon substitution into (18), shows that Condition 2 holds for \mathcal{G} with

$$\beta = \frac{d - \max_{i \neq j} \text{Tr}(Q_i^\top Q_j)}{d - \min_{i \neq j} \text{Tr}(Q_i^\top Q_j)}.$$

This completes the proof. □

When $\mathcal{G} = \mathcal{P}(d)$, a simple calculation shows that

$$\max_{Q, Q' \in \mathcal{P}(d), Q \neq Q'} \text{Tr}(Q^\top Q') = d - 2, \quad \min_{Q, Q' \in \mathcal{P}(d), Q \neq Q'} \text{Tr}(Q^\top Q') = 0.$$

Hence, by Proposition 1, Condition 2 holds for $\mathcal{P}(d)$ with $\beta = \frac{2}{d}$.

5.1.4 Proof of Theorem 2: The Cyclic Group \mathcal{Z}_m

We first prove that Condition 1 holds for \mathcal{Z}_m for any integer $m \geq 1$. In particular, we show that

$$\text{dist}(X, \mathcal{Z}_m) \leq \begin{cases} \text{Tr}(I_d - X^\top \Pi_{\mathcal{Z}_m}(X)), & \text{if } m = 1, 2, \\ \frac{1}{\sqrt{2} \sin(\frac{\pi}{m})} \text{Tr}(I_d - X^\top \Pi_{\mathcal{Z}_m}(X)), & \text{if } m \geq 3. \end{cases}$$

To do so, we note that every $X \in \text{conv}(\mathcal{Z}_m)$ is of the form

$$X = \begin{bmatrix} x & -y \\ y & x \end{bmatrix}$$

and gives rise to a complex number $z_X \in \mathbb{C}$ defined by $z_X = x + \mathbf{i}y$, where \mathbf{i} is the imaginary unit. Let $\Re(z)$ denote the real part of any complex number $z \in \mathbb{C}$,

$$Q_k = \begin{bmatrix} \cos \frac{2k\pi}{m} & -\sin \frac{2k\pi}{m} \\ \sin \frac{2k\pi}{m} & \cos \frac{2k\pi}{m} \end{bmatrix}, \quad k = 0, \dots, m-1, \quad (19)$$

and

$$\hat{k}(X) \in \underset{k \in \{0, \dots, m-1\}}{\operatorname{argmin}} \|X - Q_k\|_F^2.$$

For $m = 1$, we have $\hat{k}(X) = 0$ for any $X \in \operatorname{conv}(\mathcal{Z}_1)$ and hence $\Pi_{\mathcal{Z}_1}(X) = Q_0 = I_2$. It follows that

$$\operatorname{dist}(X, \mathcal{Z}_1) = \|X - I_2\|_F \leq \operatorname{Tr}(I_2 - X) = \operatorname{Tr}(I_2 - X^\top \Pi_{\mathcal{Z}_1}(X)),$$

i.e., Condition 1 holds for \mathcal{Z}_1 with $\alpha = 1$. For $m = 2$, if $\Re(z_X) = x \geq 0$, then $\hat{k}(X) = 0$ and $\Pi_{\mathcal{Z}_2}(X) = Q_0 = I_2$; if $\Re(z_X) = x < 0$, then $\hat{k}(X) = 1$ and $\Pi_{\mathcal{Z}_2}(X) = Q_1 = -I_2$. Therefore, we have

$$\operatorname{dist}(X, \mathcal{Z}_2) = \begin{cases} \|X - I_2\|_F & \leq \begin{cases} \operatorname{Tr}(I_2 - X) \\ \operatorname{Tr}(I_2 + X) \end{cases} \leq \operatorname{Tr}(I_2 - X^\top \Pi_{\mathcal{Z}_2}(X)), & \text{if } x \geq 0, \\ \|X + I_2\|_F & \leq \operatorname{Tr}(I_2 + X) \leq \operatorname{Tr}(I_2 - X^\top \Pi_{\mathcal{Z}_2}(X)), & \text{if } x < 0, \end{cases}$$

which shows that Condition 1 holds for \mathcal{Z}_2 with $\alpha = 1$.

Now, consider the case where $m \geq 3$. On one hand, we have

$$\operatorname{dist}(X, \mathcal{Z}_m) = \|X - Q_{\hat{k}(X)}\|_F = \sqrt{2} \left| z_X - z_{Q_{\hat{k}(X)}} \right| = \sqrt{2} \left| z_X - e^{\frac{2\hat{k}(X)\pi\mathbf{i}}{m}} \right|, \quad (20)$$

where $|\cdot|$ denotes the modulus of a complex number. On the other hand,

$$\operatorname{Tr}(I_2 - X^\top \Pi_{\mathcal{Z}_m}(X)) = 2 - \operatorname{Tr}(X^\top Q_{\hat{k}(X)}) = 2 \left(1 - \Re \left(z_X \cdot e^{-\frac{2\hat{k}(X)\pi\mathbf{i}}{m}} \right) \right). \quad (21)$$

Therefore, in order to establish Condition 1 for \mathcal{Z}_m , we need to bound the ratio

$$\frac{\left| z - e^{\frac{2k\pi\mathbf{i}}{m}} \right|}{1 - \Re \left(z \cdot e^{-\frac{2k\pi\mathbf{i}}{m}} \right)} \quad (22)$$

subject to the constraint that $z = z_X$ for some $X \in \operatorname{conv}(\mathcal{G})$ with $\hat{k}(X) = k$. By symmetry, we can assume without loss of generality that $\hat{k}(X) = 0$ and consider the shaded region \mathcal{R}_0 shown in Figure 1(a), which is defined by

$$\mathcal{R}_0 := \left\{ z \in \mathbb{C} : z \in \operatorname{conv}(\{z_{Q_0}, \dots, z_{Q_{m-1}}\}), 0 \in \underset{k \in \{0, \dots, m-1\}}{\operatorname{argmin}} \left| z - e^{\frac{2k\pi\mathbf{i}}{m}} \right| \right\}.$$

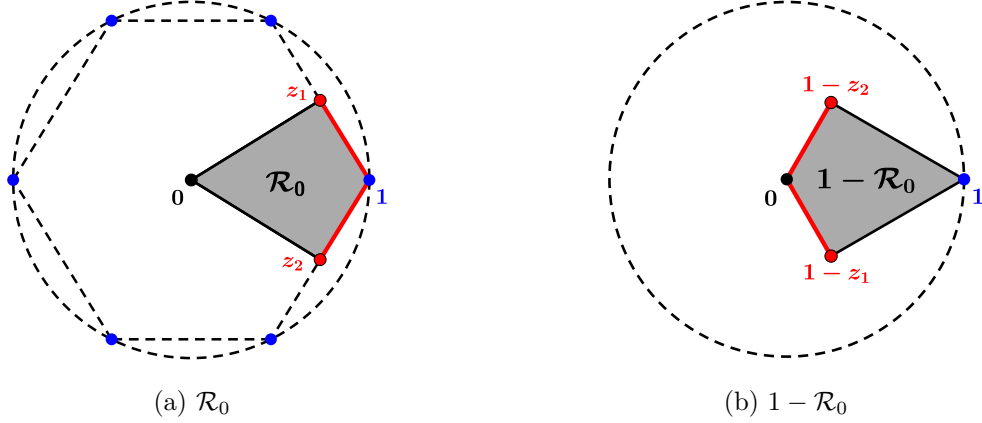


Figure 1: The regions \mathcal{R}_0 and $1 - \mathcal{R}_0$ associated with the cyclic group \mathcal{Z}_6 . On panel (a), the blue points are the group elements of \mathcal{Z}_6 and the points on the two red line segments are the maximizers of the ratio (23).

Over the region \mathcal{R}_0 , the ratio (22) reduces to

$$\frac{|1 - z|}{1 - \Re(z)}. \quad (23)$$

Let z_1 and z_2 be the midpoints between 1 and its two neighboring group elements $e^{2\pi i/m}$ and $e^{-2\pi i/m}$, respectively. We claim that any point on the line segments $[1, z_1] \cup [1, z_2]$ (see Figure 1(a)) is a maximizer of the ratio (23) over \mathcal{R}_0 . To see this, we consider the polar representation $z' = |z'| \cdot e^{i\phi}$ of the transformed variable $z' = 1 - z$, which lies in the region $1 - \mathcal{R}_0$ (see Figure 1(b)). By the angle sum formula for the regular m -gon, we have $\phi \in [-\pi(\frac{1}{2} - \frac{1}{m}), \pi(\frac{1}{2} - \frac{1}{m})]$. It follows that

$$\begin{aligned} \max_{z \in \mathcal{R}_0} \frac{|1 - z|}{1 - \Re(z)} &= \max_{z' \in 1 - \mathcal{R}_0} \frac{|z'|}{\Re(z')} = \max_{\phi \in [-\pi(\frac{1}{2} - \frac{1}{m}), \pi(\frac{1}{2} - \frac{1}{m})]} \frac{1}{\Re(e^{i\phi})} \\ &= \max_{\phi \in [-\pi(\frac{1}{2} - \frac{1}{m}), \pi(\frac{1}{2} - \frac{1}{m})]} \frac{1}{\cos \phi} \leq \frac{1}{\cos(\frac{1}{2} - \frac{1}{m})} = \frac{1}{\sin \frac{\pi}{m}}. \end{aligned} \quad (24)$$

This proves the claim since

$$\frac{|1 - z|}{1 - \Re(z)} = \frac{1}{\sin \frac{\pi}{m}}, \quad \forall z \in [1, z_1] \cup [1, z_2].$$

Now, using (20), (21), and (24), we get

$$\text{dist}(X, \mathcal{Z}_m) \leq \frac{1}{\sqrt{2} \sin \frac{\pi}{m}} \text{Tr} \left(I_d - X^\top \Pi_{\mathcal{Z}_m}(X) \right).$$

This shows that Condition 1 holds for \mathcal{Z}_m (where $m \geq 3$) with $\alpha = \frac{1}{\sqrt{2} \sin \frac{\pi}{m}}$.

Next, we establish Condition 2 for \mathcal{Z}_m with $m \geq 1$. It is trivial to show that Condition 2 holds for \mathcal{Z}_1 with $\beta = 1$. For $m \geq 2$, since \mathcal{Z}_m is a discrete subgroup of $\mathcal{O}(2)$, we compute

$$\max_{Q, Q' \in \mathcal{Z}_m, Q \neq Q'} \text{Tr}(Q^\top Q') = 2 \cos \frac{2\pi}{m}, \quad \min_{Q, Q' \in \mathcal{Z}_m, Q \neq Q'} \text{Tr}(Q^\top Q') = -2$$

and apply Proposition 1 to conclude that Condition 2 holds for \mathcal{Z}_m with

$$\beta = \frac{2 - 2 \cos \frac{2\pi}{m}}{4} = \sin^2 \frac{\pi}{m}.$$

Finally, let us derive an explicit formula for the projection $\Pi_{\mathcal{Z}_m}$, where $m \geq 1$.

Proposition 2. *For any*

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \in \mathbb{R}^{2 \times 2},$$

define

$$\theta = \begin{cases} \arccos \frac{x_{21} - x_{12}}{\sqrt{(x_{11} + x_{22})^2 + (x_{21} - x_{12})^2}}, & \text{if } x_{11} + x_{22} \geq 0, \\ 2\pi - \arccos \frac{x_{21} - x_{12}}{\sqrt{(x_{11} + x_{22})^2 + (x_{21} - x_{12})^2}}, & \text{if } x_{11} + x_{22} < 0, \end{cases}$$

which always lies in $[0, 2\pi)$. Then, for any $m \geq 1$,

$$\Pi_{\mathcal{Z}_m}(X) = \begin{cases} Q_{\text{round}(\frac{m}{4} - \frac{m\theta}{2\pi})}, & \text{if } 0 \leq \theta \leq \frac{\pi}{2} + \frac{\pi}{m}, \\ Q_{\text{round}(\frac{5m}{4} - \frac{m\theta}{2\pi})}, & \text{if } \frac{\pi}{2} + \frac{\pi}{m} < \theta < 2\pi, \end{cases}$$

where $\text{round}(\cdot)$ rounds a number to its closest integer and Q_k is defined in (19).

Proof. The result is trivial when $m = 1$. Thus, we assume that $m \geq 2$. Then, we have

$$\begin{aligned} \argmin_{k \in \{0, \dots, m-1\}} \|X - Q_k\|_F^2 &= \argmin_{k \in \{0, \dots, m-1\}} \langle X, Q_k \rangle \\ &= \argmin_{k \in \{0, \dots, m-1\}} \left\{ (x_{11} + x_{22}) \cos \frac{2k\pi}{m} + (x_{21} - x_{12}) \sin \frac{2k\pi}{m} \right\} \\ &= \argmin_{k \in \{0, \dots, m-1\}} \sin \left(\theta + \frac{2k\pi}{m} \right). \end{aligned}$$

Since θ can take any value in $[0, 2\pi)$ and $\frac{2k\pi}{m}$ lies in $[0, \frac{2(m-1)\pi}{m}]$, we have

$$0 \leq \theta + \frac{2k\pi}{m} \leq 2\pi + \frac{2(m-1)\pi}{m} < 4\pi.$$

Moreover, the function $\varphi \mapsto \sin \varphi$ has two peaks in $\left[0, 2\pi + \frac{2(m-1)\pi}{m}\right)$, namely at $\varphi = \frac{\pi}{2}$ and $\varphi = \frac{5\pi}{2}$. It follows that

$$\operatorname{argmin}_{k \in \{0, \dots, m-1\}} \|X - Q_k\|_F^2 = \operatorname{argmin}_{k \in \{0, \dots, m-1\}} \min \left\{ \left| \theta + \frac{2k\pi}{m} - \frac{\pi}{2} \right|, \left| \theta + \frac{2k\pi}{m} - \frac{5\pi}{2} \right| \right\}. \quad (25)$$

We first consider the case where $0 \leq \theta \leq \frac{\pi}{2} + \frac{\pi}{m}$. Note that

$$\theta + \frac{2k\pi}{m} \leq \frac{\pi}{2} + \frac{\pi}{m} + \frac{2(m-1)\pi}{m} = \frac{5\pi}{2} - \frac{\pi}{m},$$

which implies that

$$\left| \theta + \frac{2k\pi}{m} - \frac{5\pi}{2} \right| \geq \frac{\pi}{m}.$$

Also, since

$$\frac{m}{4} - \frac{m\theta}{2\pi} \geq \frac{m}{4} - \frac{m}{2\pi} \left(\frac{\pi}{2} + \frac{\pi}{m} \right) = -\frac{1}{2},$$

we have $\operatorname{round} \left(\frac{m}{4} - \frac{m\theta}{2\pi} \right) \in \{0, \dots, m-1\}$. Setting $k = \operatorname{round} \left(\frac{m}{4} - \frac{m\theta}{2\pi} \right)$, we get

$$\left| k - \left(\frac{m}{4} - \frac{m\theta}{2\pi} \right) \right| \leq \frac{1}{2}$$

and hence

$$\left| \theta + \frac{2k\pi}{m} - \frac{\pi}{2} \right| \leq \frac{\pi}{m}.$$

These, together with (25), yield

$$\begin{aligned} \operatorname{argmin}_{k \in \{0, \dots, m-1\}} \|X - Q_k\|_F^2 &= \operatorname{argmin}_{k \in \{0, \dots, m-1\}} \left| \theta + \frac{2k\pi}{m} - \frac{\pi}{2} \right| \\ &= \operatorname{argmin}_{k \in \{0, \dots, m-1\}} \left| k - \left(\frac{m}{4} - \frac{m\theta}{2\pi} \right) \right| = \operatorname{round} \left(\frac{m}{4} - \frac{m\theta}{2\pi} \right). \end{aligned}$$

Next, we consider the case where $\frac{\pi}{2} + \frac{\pi}{m} < \theta < 2\pi$. Note that

$$\theta + \frac{2k\pi}{m} - \frac{\pi}{2} > \frac{\pi}{2} + \frac{\pi}{m} - \frac{\pi}{2} = \frac{\pi}{m},$$

which implies that

$$\left| \theta + \frac{2k\pi}{m} - \frac{\pi}{2} \right| \geq \frac{\pi}{m}.$$

Also, since

$$\frac{5m}{4} - \frac{m\theta}{2\pi} < \frac{5m}{4} - \frac{m}{2\pi} \left(\frac{\pi}{2} + \frac{\pi}{m} \right) = m - \frac{1}{2},$$

we have $\text{round}\left(\frac{5m}{4} - \frac{m\theta}{2\pi}\right) \in \{0, \dots, m-1\}$. Setting $k = \text{round}\left(\frac{5m}{4} - \frac{m\theta}{2\pi}\right)$, we get

$$\left|k - \left(\frac{5m}{4} - \frac{m\theta}{2\pi}\right)\right| \leq \frac{1}{2}$$

and hence

$$\left|\theta + \frac{2k\pi}{m} - \frac{5\pi}{2}\right| \leq \frac{\pi}{m}.$$

These, together with (25), yield

$$\begin{aligned} \operatorname{argmin}_{k \in \{0, \dots, m-1\}} \|X - Q_k\|_F^2 &= \operatorname{argmin}_{k \in \{0, \dots, m-1\}} \left|\theta + \frac{2k\pi}{m} - \frac{5\pi}{2}\right| \\ &= \operatorname{argmin}_{k \in \{0, \dots, m-1\}} \left|k - \left(\frac{5m}{4} - \frac{m\theta}{2\pi}\right)\right| = \text{round}\left(\frac{5m}{4} - \frac{m\theta}{2\pi}\right). \end{aligned}$$

This completes the proof. \square

5.2 Erdős-Rényi Measurement Graphs

Recall from our discussion in Section 4 that the parameter κ can be viewed as a measure of the connectivity of the measurement graph $([n], E)$. In particular, when the measurement graph is complete, we have $\kappa = 0$, which shows that condition (ii) in the master theorem is satisfied. As it turns out, the condition can be satisfied by measurement graphs that are much sparser. In this subsection, we show that if the measurement graph is an *Erdős-Rényi* random graph with *observation rate* $p \geq \frac{c \log n}{n}$ for some constant $c > 0$ — *i.e.*, the edge weights $\{w_{ij} : 1 \leq i < j \leq n\}$ are independent and identically distributed (i.i.d.) Bernoulli random variables with

$$w_{ij} = \begin{cases} 1, & \text{with probability } p, \\ 0, & \text{with probability } 1 - p, \end{cases} \quad 1 \leq i < j \leq n$$

— then condition (ii) in the master theorem will be satisfied with high probability. More precisely, we have the following result:

Theorem 3. *Suppose that the measurement graph is an Erdős-Rényi random graph with observation rate $p \in (0, 1]$. Then, there exist constants $c_1, c_2 > 0$ such that whenever $p \geq \frac{c_1 \log n}{n}$, we will have $\kappa \leq \frac{1}{32}$ with probability at least $1 - \frac{1}{n^{c_2}}$.*

Proof. Using the definitions of D , W , F , and κ in Section 4, we compute

$$\begin{aligned} \kappa &= \left\| D^{-1}W - \frac{1}{n}F \right\| = \left\| (\bar{D} \otimes I_d)^{-1}(\bar{W} \otimes I_d) - \frac{1}{n}(ee^\top \otimes I_d) \right\| \\ &= \left\| (\bar{D}^{-1} \otimes I_d)(\bar{W} \otimes I_d) - \frac{1}{n}(ee^\top \otimes I_d) \right\| \end{aligned}$$

$$\begin{aligned}
&= \left\| \left(\bar{D}^{-1} \bar{W} \right) \otimes I_d - \frac{1}{n} (ee^\top \otimes I_d) \right\| \\
&= \left\| \left(\bar{D}^{-1} \bar{W} - \frac{1}{n} ee^\top \right) \otimes I_d \right\| \\
&= \left\| \bar{D}^{-1} \bar{W} - \frac{1}{n} ee^\top \right\|,
\end{aligned}$$

where the second line follows from the fact that $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ for any invertible matrices A and B , the third line follows from the fact that $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ for any matrices A, B, C, D with conformable dimensions, the fourth line follows from the bilinearity of the Kronecker product, and the last line follows from the fact that $\|A \otimes B\| = \|A\| \cdot \|B\|$ (these properties of the Kronecker product can be found in, *e.g.*, [35, Chapter 4.2]). Now, we bound

$$\begin{aligned}
\left\| \bar{D}^{-1} \bar{W} - \frac{1}{n} ee^\top \right\| &\leq \left\| I_n - \frac{1}{p(n-1)+1} \bar{D} \right\| \cdot \|\bar{D}^{-1} \bar{W}\| + \frac{1}{p(n-1)+1} \|\bar{W} - \mathbb{E}[\bar{W}]\| \\
&\quad + \left\| \frac{1}{p(n-1)+1} \mathbb{E}[\bar{W}] - \frac{1}{n} ee^\top \right\|.
\end{aligned}$$

Since $\bar{D}^{-1} \bar{W}$ has non-negative entries and each of its rows sums to 1, we have $\|\bar{D}^{-1} \bar{W}\| \leq 1$ by [34, Corollary 6.1.5]. Moreover, observe that

$$\left\| I_n - \frac{1}{p(n-1)+1} \bar{D} \right\| = \max_{i \in [n]} \left| 1 - \frac{r_i}{p(n-1)+1} \right|,$$

where $r_i = \sum_{j=1}^n w_{ij} = 1 + \sum_{j \neq i} w_{ij}$ for $i = 1, \dots, n$. By Chernoff's inequality (cf. [70, Exercise 2.3.5]), for any $t \in (0, 1]$, we have

$$\Pr \left(\left| \sum_{j \neq i} w_{ij} - p(n-1) \right| \geq tp(n-1) \right) \leq 2 \exp \left(-\frac{p(n-1)t^2}{3} \right), \quad i = 1, \dots, n.$$

It follows that for any $t \in (0, 1]$,

$$\Pr \left(\left\| I_n - \frac{1}{p(n-1)+1} \bar{D} \right\| \geq \frac{tp(n-1)}{p(n-1)+1} \right) \leq 2n \cdot \exp \left(-\frac{p(n-1)t^2}{3} \right).$$

Next, by adapting the results in [13, Examples 3.14 and 6.8] and [10, Corollary 3.6], we have, for any $p \geq \frac{\log n}{n}$, that

$$\Pr (\|\bar{W} - \mathbb{E}[\bar{W}]\| \geq 172\sqrt{pn}) \leq \exp \left(-\frac{pn}{8} \right), \tag{26}$$

see the proof of [49, Lemma 2]. Lastly, since $\mathbb{E}[\bar{W}] = (1-p)I_n + p \cdot ee^\top$, we have

$$\left\| \frac{1}{p(n-1)+1} \mathbb{E}[\bar{W}] - \frac{1}{n} ee^\top \right\| = \left\| \frac{1-p}{p(n-1)+1} I_n + \left(\frac{p}{p(n-1)+1} - \frac{1}{n} \right) ee^\top \right\|$$

$$= \frac{1-p}{p(n-1)+1}.$$

Upon setting $t = \frac{1}{66}$ and assuming that $p \geq \frac{(172 \times 66)^2 \log n}{n}$, we conclude that

$$\kappa \leq \frac{tp(n-1)}{p(n-1)+1} + \frac{172\sqrt{pn}}{p(n-1)+1} + \frac{1-p}{p(n-1)+1} < \frac{1}{32}$$

with probability at least $1 - \frac{1}{n^{9000}}$. \square

It is worth noting that Theorem 3 holds even when the order of p is as small as $\frac{\log n}{n}$. Such an order is optimal in the sense that it matches the order of the connectivity threshold of an Erdős-Rényi random graph [27] (cf. [68, Chapter 5.3]).

5.3 Additive Sub-Gaussian Noise Model

Recall that under the additive noise model, the measurements are given by

$$C_{ij} = G_i^* G_j^{*\top} + \Theta_{ij}, \quad (i, j) \in E. \quad (27)$$

The purpose of this subsection is to show that condition (iii) in the master theorem holds for a large class of noise matrices $\{\Theta_{ij} : (i, j) \in E\}$. Specifically, we focus on the case where $\{\Theta_{ij} : (i, j) \in E\}$ is a collection of independent random matrices whose entries are i.i.d. *sub-Gaussian* random variables.

Definition 1 (Sub-Gaussian Random Variable). *A random variable ξ is said to be sub-Gaussian with parameter $\sigma > 0$ if*

$$\mathbb{E}[\exp(\lambda\xi)] \leq \exp\left(\frac{\lambda^2\sigma^2}{2}\right), \quad \lambda \in \mathbb{R}.$$

It can be shown that if ξ is sub-Gaussian in the above sense, then it necessarily satisfies $\mathbb{E}[\xi] = 0$. The class of sub-Gaussian random variables is rich. It contains, for example, the Gaussian, uniform, Bernoulli, and any bounded random variables, see [70, Example 2.5.8]. In particular, a Gaussian random variable with mean zero and standard deviation $\sigma > 0$ is sub-Gaussian with parameter σ .

We first establish a tail inequality for the operator norm of the block matrix Δ (see (6) for the definition), which will be useful for verifying condition (iii) in the master theorem.

Proposition 3. *Suppose that the measurement graph is an Erdős-Rényi random graph with observation rate $p \in (0, 1]$. Let $\{\Theta_{ij} : 1 \leq i < j \leq n\}$ be independent noise matrices that are independent of the measurement graph and whose entries are i.i.d. sub-Gaussian random variables with parameter $\sigma > 0$. Then, the following hold:*

(a) For $p = 1$, there exists a constant $c_0 > 0$ such that

$$\Pr \left(\|\Delta\| \geq c_0 \sigma \sqrt{nd} \right) \leq \frac{1}{n}.$$

(b) For $p < 1$, there exist constants $c_1, c_2, c_3 > 0$ such that whenever $p \geq \frac{c_1 \log n}{n}$, we have

$$\Pr \left(\|\Delta\| \geq c_2 \sigma \sqrt{pnd \log(nd)} \right) \leq \frac{c_3}{n}.$$

Proof. Using (6) and (27), we can write

$$[\Delta]_{ij} = w_{ij} \left([C]_{ij} - \left[G^* G^{*\top} \right]_{ij} \right) = w_{ij} \left(C_{ij} - G_i^* G_j^{*\top} \right) = w_{ij} \Theta_{ij}, \quad 1 \leq i < j \leq n.$$

For $p = 1$, we have $w_{ij} = 1$ for $1 \leq i < j \leq n$. In particular, we see that $\Delta \in \mathbb{R}^{nd \times nd}$ is a symmetric matrix whose entries on or above the diagonal are independent sub-Gaussian random variables with parameters bounded above by $\sigma > 0$. In view of Definition 1 and [70, equation (2.16)], the desired bound follows from [70, Corollary 4.4.8].

For $p < 1$, consider, for $k, \ell \in [d]$, the matrix $E_{k\ell} \in \mathbb{R}^{d \times d}$ defined by

$$(E_{k\ell})_{uv} := \begin{cases} 1, & \text{if } (u, v) = (k, \ell), \\ 0, & \text{otherwise,} \end{cases} \quad u, v = 1, \dots, d.$$

Furthermore, consider, for $1 \leq i < j \leq n$ and $k, \ell \in [d]$, the symmetric matrices $B_{ijkl} \in \mathbb{R}^{nd \times nd}$ and $S_{ij} \in \mathbb{R}^{nd \times nd}$ defined by

$$B_{ijkl} := \begin{bmatrix} & & \underbrace{w_{ij} E_{k\ell}}_{(i,j)\text{-th block}} \\ & & \\ \underbrace{w_{ij} E_{k\ell}^\top}_{(j,i)\text{-th block}} & & \end{bmatrix}$$

and

$$S_{ij} := \begin{bmatrix} & \underbrace{[\Delta]_{ij}}_{(i,j)\text{-th block}} \\ & \\ \underbrace{[\Delta]_{ij}^\top}_{(j,i)\text{-th block}} & \end{bmatrix} = w_{ij} \begin{bmatrix} & \Theta_{ij} \\ \Theta_{ij}^\top & \end{bmatrix} = \sum_{k, \ell \in [d]} (\Theta_{ij})_{k\ell} B_{ijkl},$$

respectively. Then, conditioning on the measurement graph (*i.e.*, on the values of the random variables $\{w_{ij} : 1 \leq i < j \leq n\}$), we see that

$$\Delta = \sum_{1 \leq i < j \leq n} S_{ij} = \sum_{1 \leq i < j \leq n} \sum_{k, \ell \in [d]} (\Theta_{ij})_{k\ell} B_{ijk\ell}$$

is a *matrix sub-Gaussian series*; cf. [68, Chapter 4]. Upon noting that

$$\sum_{k, \ell \in [d]} B_{ijk\ell}^2 = w_{ij}^2 \sum_{k, \ell \in [d]} \begin{bmatrix} \underbrace{E_{k\ell} E_{k\ell}^\top}_{(i,i)\text{-th block}} & & \\ & \underbrace{E_{k\ell}^\top E_{k\ell}}_{(j,j)\text{-th block}} & \\ & & \end{bmatrix} = w_{ij} \begin{bmatrix} d \cdot I_d & & \\ & & \\ & & d \cdot I_d \end{bmatrix},$$

we have

$$\rho^2 := \left\| \sum_{1 \leq i < j \leq n} \sum_{k, \ell \in [d]} B_{ijk\ell}^2 \right\| = \|\text{BlkDiag}(w_1 d \cdot I_d, \dots, w_n d \cdot I_d)\| = d \cdot \max_{i \in [n]} w_i,$$

where

$$w_i = \sum_{j=i+1}^n w_{ij} + \sum_{j=1}^{i-1} w_{ji}, \quad i = 1, \dots, n.$$

It follows from [21, Theorem 1] and the symmetry of sub-Gaussian random variables (cf. [68, Chapters 4.6.1 and 4.6.2]) that for any $s \geq 0$,

$$\Pr(\|\Delta\| \geq s \mid \{w_{ij} : 1 \leq i < j \leq n\}) \leq 2nd \cdot \exp\left(-\frac{s^2}{2\sigma^2 \rho^2}\right). \quad (28)$$

Now, since $\{w_{ij} : 1 \leq i < j \leq n\}$ are i.i.d. Bernoulli random variables with parameter p , by the union bound and Chernoff's inequality [70, Theorem 2.3.1], for any $t > p(n-1)$, we have

$$\Pr\left(\max_{i \in [n]} w_i \geq t\right) \leq \sum_{i=1}^n \Pr(w_i \geq t) \leq n \cdot \exp(-p(n-1)) \cdot \left(\frac{ep(n-1)}{t}\right)^t. \quad (29)$$

Upon taking $t = e^2 p(n-1)$ in (29) and $s = 2e\sigma\sqrt{pd(n-1)\log(nd)}$ in (28), we have

$$\Pr\left(\|\Delta\| \geq 2e\sigma\sqrt{pd(n-1)\log(nd)}\right) \leq \frac{2}{nd} + n \cdot \exp(-(e^2 + 1)p(n-1)).$$

In particular, for $p \geq \frac{2}{e^2+1} \cdot \frac{\log n}{n-1}$, the above probability is at most $\frac{3}{n}$. \square

We note that the bound on $\|\Delta\|$ for the case where $p < 1$ has an extra logarithmic factor when compared with that for the case where $p = 1$. This is due to the application of the norm bound developed in [21] for general matrix sub-Gaussian series. It would be interesting to see whether such a factor can be removed. We leave this as an open question.

The following theorem, which is a substantial generalization of [6, Proposition 3.3], shows that under the setting of Proposition 3, condition (iii) in Theorem 1 will be satisfied with high probability.

Theorem 4. *Consider the setting of Proposition 3 and let $\alpha \geq 1$ be arbitrary. Then, there exist constants $c_0, c_1, c_2, c_3 > 0$ such that when either (a) $p = 1$ and $\sigma \leq \frac{c_0\sqrt{n}}{\alpha d}$ or (b) $p \in \left[\frac{c_1 \log n}{n}, 1\right)$ and $\sigma \leq \frac{c_2\sqrt{pn}}{\alpha d\sqrt{\log(nd)}}$, we will have*

$$\|D^{-1}\Delta\| \leq \frac{1}{32} \quad \text{and} \quad \|D^{-1}\Delta G^*\|_F \leq \frac{\sqrt{n}}{32\alpha}$$

with probability at least $1 - \frac{c_3}{n}$.

Proof. We first bound $\|D^{-1}\Delta\|$. For $p = 1$, the definition of r_i ($i = 1, \dots, n$) in Section 4 yields $r_1 = \dots = r_n = n$. Since $D = \text{BlkDiag}(r_1 \cdot I_d, \dots, r_n \cdot I_d)$, we have $\|D^{-1}\| = \frac{1}{n}$. This, together with Proposition 3, implies the existence of a constant $c_0 > 0$ such that whenever $\sigma \leq \frac{c_0\sqrt{n}}{\alpha d}$, we will have

$$\|D^{-1}\Delta\| \leq \|D^{-1}\| \cdot \|\Delta\| = \frac{\|\Delta\|}{n} \leq \frac{1}{32\alpha\sqrt{d}} \leq \frac{1}{32} \quad (30)$$

with probability at least $1 - \frac{1}{n}$. For $p < 1$, upon applying Chernoff's inequality [70, Exercise 2.3.2], we have

$$\Pr\left(r_i \leq 1 + \frac{p(n-1)}{2}\right) \leq \left(\frac{2}{e}\right)^{p(n-1)/2}, \quad i = 1, \dots, n.$$

The above inequality and the union bound lead to

$$\Pr\left(\|D^{-1}\| \geq \left(1 + \frac{p(n-1)}{2}\right)^{-1}\right) \leq \sum_{i=1}^n \Pr\left(r_i \leq 1 + \frac{p(n-1)}{2}\right) \leq n \cdot \left(\frac{2}{e}\right)^{p(n-1)/2}.$$

This, together with Proposition 3, implies the existence of constants $c_1, c_2, c_3 > 0$ such that whenever $p \geq \frac{c_1 \log n}{n}$ and $\sigma \leq \frac{c_2\sqrt{pn}}{\alpha d\sqrt{\log(nd)}}$, we will have

$$\|D^{-1}\Delta\| \leq \|D^{-1}\| \cdot \|\Delta\| \leq \left(1 + \frac{p(n-1)}{2}\right)^{-1} \|\Delta\| \leq \frac{1}{32\alpha\sqrt{d}} \leq \frac{1}{32} \quad (31)$$

with probability at least $1 - \frac{c_3}{n}$.

Next, we bound $\|D^{-1}\Delta G^*\|_F$. Since G^* has n blocks, each of which is a $d \times d$ orthogonal matrix, we have $\|G^*\|_F = \sqrt{nd}$. Using the inequality $\|D^{-1}\Delta G^*\|_F \leq \|D^{-1}\Delta\| \cdot \|G^*\|_F$, the desired bound then follows from (30) (for the case where $p = 1$) or (31) (for the case where $p < 1$). \square

5.4 Entropic Spectral Initialization

In order for our non-convex approach to enjoy the theoretical guarantee offered by the master theorem, we need to initialize GPM by a point that has a sufficiently small estimation error. As it turns out, the geometry of the closed subgroup \mathcal{G} contains much information that can be used to guide our construction of such a point. Specifically, by considering the error-bound geometry of \mathcal{G} as encapsulated in Condition 2 and the classic notion of *metric entropy* of the quotient $\mathcal{O}(d)/\mathcal{G}$ of the orthogonal group $\mathcal{O}(d)$ by the closed subgroup \mathcal{G} ,⁴ we design a novel initialization procedure for GPM that produces a point satisfying condition (iv) in the master theorem. Before we present our proposed procedure, let us introduce some basic results concerning the metric entropy of the quotient $\mathcal{O}(d)/\mathcal{G}$.

5.4.1 The Quotient $\mathcal{O}(d)/\mathcal{G}$ and Its Metric Entropy

Given any closed subgroup \mathcal{G} of $\mathcal{O}(d)$ and any orthogonal matrix $Q \in \mathcal{O}(d)$, the (left-) *coset* $[Q]$ of \mathcal{G} in $\mathcal{O}(d)$ is defined by

$$[Q] := \{O \in \mathcal{O}(d) : O = QQ' \text{ for some } Q' \in \mathcal{G}\}.$$

The quotient $\mathcal{O}(d)/\mathcal{G}$ is then defined as the set of all cosets of \mathcal{G} in $\mathcal{O}(d)$. We can define a natural distance on $\mathcal{O}(d)/\mathcal{G}$ by

$$\text{dist}([Q_1], [Q_2]) = \min_{Q' \in \mathcal{G}} \|Q_1 - Q_2 Q'\|_F.$$

It can be easily seen that this distance is independent of the choice of the class representatives Q_1 and Q_2 of the cosets $[Q_1]$ and $[Q_2]$, respectively. Moreover, it turns $\mathcal{O}(d)/\mathcal{G}$ into a compact (under the quotient topology) metric space.

We next introduce the concepts of *net*, *covering number*, and *metric entropy*, see, e.g., [41].

Definition 2 (Net, Covering Number, and Metric Entropy). *Let (\mathcal{S}, ν) be a compact metric space and $\epsilon > 0$ be a parameter. A subset $\mathcal{N} \subseteq \mathcal{S}$ is said to be an ϵ -net of \mathcal{S} if for any point $x \in \mathcal{S}$, there exists a point $y \in \mathcal{N}$ such that $\nu(x, y) \leq \epsilon$. The cardinality of the smallest ϵ -net is called the ϵ -covering number, denoted by $N(\mathcal{S}, \epsilon)$. The quantity $\log N(\mathcal{S}, \epsilon)$ is called the ϵ -metric entropy of \mathcal{S} .*

The compactness of \mathcal{S} implies that $N(\mathcal{S}, \epsilon)$ is finite for any $\epsilon > 0$. The following proposition offers an explicit and efficient construction of an ϵ -net of the quotient $\mathcal{O}(d)/\mathcal{G}$.

Proposition 4. *Let $\epsilon > 0$ be given and H_ϵ be the $(\epsilon/2)$ -metric entropy of $\mathcal{O}(d)/\mathcal{G}$. For any $K \geq 1$, let Q_1, \dots, Q_K be random orthogonal matrices that are independently and uniformly distributed on $\mathcal{O}(d)$. Then, the set $\mathcal{Q} = \{[Q_1], \dots, [Q_K]\}$ is an ϵ -net of $\mathcal{O}(d)/\mathcal{G}$ with probability at least $1 - (1 - e^{-H_\epsilon})^K$.*

⁴Note that the quotient $\mathcal{O}(d)/\mathcal{G}$ may not be a group in general as we do not require the subgroup \mathcal{G} to be normal.

Proof. Consider any $(\epsilon/2)$ -net $\mathcal{N} = \{[\bar{Q}_1], \dots, [\bar{Q}_{|\mathcal{N}|}]\}$ of the quotient $\mathcal{O}(d)/\mathcal{G}$ and fix an arbitrary $O \in \mathcal{O}(d)$. Since the matrices Q_1, \dots, Q_K are independently and uniformly distributed on $\mathcal{O}(d)$, the cosets $[Q_1], \dots, [Q_K]$ are independently and uniformly distributed on $\mathcal{O}(d)/\mathcal{G}$. Therefore, for any $k \in [K]$ and $\ell \in [|\mathcal{N}|]$,

$$1 = \Pr \left([Q_k] \in \bigcup_{\ell=1}^{|\mathcal{N}|} \mathcal{B}([\bar{Q}_\ell], \frac{\epsilon}{2}) \right) \leq \sum_{\ell=1}^{|\mathcal{N}|} \Pr([Q_k] \in \mathcal{B}([\bar{Q}_\ell], \frac{\epsilon}{2})) = |\mathcal{N}| \cdot \Pr([Q_k] \in \mathcal{B}([\bar{Q}_\ell], \frac{\epsilon}{2})),$$

where $\mathcal{B}([\bar{Q}_\ell], \frac{\epsilon}{2}) := \{[Q] : \text{dist}([\bar{Q}_\ell], [Q]) \leq \frac{\epsilon}{2}\}$ and the last equality follows from the fact that the distribution of Q_k is invariant under multiplication by any fixed orthogonal matrix. This gives

$$\Pr([Q_k] \in \mathcal{B}([\bar{Q}_\ell], \frac{\epsilon}{2})) \geq |\mathcal{N}|^{-1}, \quad k = 1, \dots, K; \ell = 1, \dots, |\mathcal{N}|. \quad (32)$$

Next, there exists some $\bar{\ell} \in [|\mathcal{N}|]$ such that $[O] \in \mathcal{B}([\bar{Q}_{\bar{\ell}}], \frac{\epsilon}{2})$. Hence, by (32), we have

$$\begin{aligned} \Pr([O] \in \mathcal{B}([Q_k], \epsilon) \text{ for some } k \in [K]) &= 1 - \prod_{k=1}^K (1 - \Pr([O] \in \mathcal{B}([Q_k], \epsilon))) \\ &\geq 1 - \prod_{k=1}^K (1 - \Pr([Q_k] \in \mathcal{B}([\bar{Q}_{\bar{\ell}}], \frac{\epsilon}{2}))) \geq 1 - (1 - |\mathcal{N}|^{-1})^K. \end{aligned}$$

Optimizing the lower bound over all possible $(\epsilon/2)$ -nets \mathcal{N} of $\mathcal{O}(d)/\mathcal{G}$ completes the proof. \square

In view of Proposition 4, we are naturally interested in determining the metric entropy of the quotient $\mathcal{O}(d)/\mathcal{G}$, particularly when \mathcal{G} is one of the four subgroups (*i.e.*, $\mathcal{O}(d)$, $\mathcal{SO}(d)$, $\mathcal{P}(d)$, and \mathcal{Z}_m) we considered earlier. For $\mathcal{G} = \mathcal{O}(d)$, the quotient $\mathcal{O}(d)/\mathcal{O}(d)$ is the trivial group that contains only one element. Therefore, its ϵ -metric entropy is 0 for any $\epsilon > 0$. For $\mathcal{G} = \mathcal{SO}(d)$, the quotient $\mathcal{O}(d)/\mathcal{SO}(d)$ is isomorphic to the Boolean group, which implies that its ϵ -metric entropy is at most $\log 2$ for any $\epsilon > 0$. In what follows, we provide an estimate of the metric entropy of the quotient $\mathcal{O}(d)/\mathcal{G}$ when \mathcal{G} is a discrete subgroup of $\mathcal{O}(d)$. In particular, such an estimate applies to the cases where $\mathcal{G} = \mathcal{P}(d)$ and $\mathcal{G} = \mathcal{Z}_m$.

5.4.2 Metric Entropy of the Quotient $\mathcal{O}(d)/\mathcal{G}$ for Discrete \mathcal{G}

To begin, let us introduce the concepts of *packing* and *packing number*, which are closely related to the concepts of net and covering number, respectively.

Definition 3 (Packing and Packing Number). *Let (\mathcal{S}, ν) be a compact metric space and $\epsilon > 0$ be a parameter. A subset $\mathcal{P} \subseteq \mathcal{S}$ is said to be an ϵ -packing of \mathcal{S} if any two points $x, y \in \mathcal{P}$ satisfy $\nu(x, y) > \epsilon$. The cardinality of the largest ϵ -packing is called the ϵ -packing number, denoted by $P(\mathcal{S}, \epsilon)$.*

Given a compact metric space (\mathcal{S}, ν) and a parameter $\epsilon > 0$, we have the following relationship between the covering and packing numbers, see, *e.g.*, [66, Inequality (3)]:

$$N(\mathcal{S}, \epsilon) \leq P(\mathcal{S}, \epsilon) \leq N(\mathcal{S}, \frac{\epsilon}{2}). \quad (33)$$

This can then be used to establish the following result:

Proposition 5. *Let \mathcal{G} be a discrete subgroup of $\mathcal{O}(d)$ and $\epsilon > 0$ be a given parameter. Then, we have*

$$N(\mathcal{O}(d)/\mathcal{G}, \epsilon) \leq |\mathcal{G}|^{-1} \left(\frac{c}{\epsilon} \right)^{\frac{d(d-1)}{2}}$$

for some constant $c > 0$. Consequently, the ϵ -metric entropy of the quotient $\mathcal{O}(d)/\mathcal{G}$ is at most

$$\frac{d(d-1)}{2} \log \frac{c}{\epsilon} - \log |\mathcal{G}|.$$

Proof. It suffices to consider the case where ϵ is smaller than the minimum separation of the discrete subgroup \mathcal{G} , *i.e.*,

$$\epsilon < \min_{G'_1, G'_2 \in \mathcal{G}, G'_1 \neq G'_2} \|G'_1 - G'_2\|_F. \quad (34)$$

Let $\tilde{\mathcal{P}}$ be a maximal ϵ -packing of the quotient $\mathcal{O}(d)/\mathcal{G}$, *i.e.*, $|\tilde{\mathcal{P}}| = P(\mathcal{O}(d)/\mathcal{G}, \epsilon)$. Then, for any distinct cosets $[Q_1], [Q_2] \in \tilde{\mathcal{P}}$, we have

$$\text{dist}([Q_1], [Q_2]) = \min_{Q' \in \mathcal{G}} \|Q_1 - Q_2 Q'\|_F > \epsilon. \quad (35)$$

Consider the set $\mathcal{P} := \{Q \in \mathcal{O}(d) : [Q] \in \tilde{\mathcal{P}}\}$. We claim that \mathcal{P} is an ϵ -packing of $\mathcal{O}(d)$. To prove this, let $Q_1, Q_2 \in \mathcal{P}$ be two distinct points. If $[Q_1] \neq [Q_2]$, then we have

$$\|Q_1 - Q_2\|_F \geq \min_{Q' \in \mathcal{G}} \|Q_1 - Q_2 Q'\|_F > \epsilon$$

by (35). If $[Q_1] = [Q_2]$, then $Q_1 = Q_2 Q'$ for some $Q' \in \mathcal{G} \setminus \{I_d\}$. It follows from (34) that

$$\|Q_1 - Q_2\|_F = \|I_d - Q'\|_F > \epsilon.$$

Thus, the claim is established. Now, it is elementary to show that (i) for any $Q \in \mathcal{O}(d)$, we have $|[Q]| = |\mathcal{G}|$; (ii) for any two cosets $[Q_1], [Q_2] \in \mathcal{O}(d)/\mathcal{G}$, we either have $[Q_1] = [Q_2]$ or $[Q_1] \cap [Q_2] = \emptyset$. In particular, we see that $|\mathcal{P}| = |\tilde{\mathcal{P}}| \cdot |\mathcal{G}| = P(\mathcal{O}(d)/\mathcal{G}, \epsilon) \cdot |\mathcal{G}|$. Consequently, there exists a constant $c > 0$ such that

$$N(\mathcal{O}(d)/\mathcal{G}, \epsilon) \cdot |\mathcal{G}| \leq P(\mathcal{O}(d)/\mathcal{G}, \epsilon) \cdot |\mathcal{G}| = |\mathcal{P}| \leq P(\mathcal{O}(d), \epsilon) \leq N(\mathcal{O}(d), \frac{\epsilon}{2}) \leq \left(\frac{c}{\epsilon} \right)^{\frac{d(d-1)}{2}},$$

where the first and third inequalities follow from (33), the second inequality follows the fact that \mathcal{P} is an ϵ -packing, and the last inequality follows from [66, Theorem 7]. This completes the proof. \square

5.4.3 Entropic Spectral Estimator

We are now ready to develop our advertised initialization procedure for GPM. We will make use of the results in the previous subsection and the following variant of the Davis-Kahan Theorem [76].

Theorem 5. *Let $Z, Z^* \in \mathbb{R}^{M \times M}$ be symmetric matrices with eigenvalues $\lambda_1 \geq \dots \geq \lambda_M$ and $\lambda_1^* \geq \dots \geq \lambda_M^*$, respectively. For any integers k, ℓ such that $1 \leq k \leq \ell \leq M$, let $V, V^* \in \mathbb{R}^{M \times (\ell - k + 1)}$ be the matrices defined by*

$$V = \begin{bmatrix} | & & | \\ v_k & \cdots & v_\ell \\ | & & | \end{bmatrix} \quad \text{and} \quad V^* = \begin{bmatrix} | & & | \\ v_k^* & \cdots & v_\ell^* \\ | & & | \end{bmatrix},$$

where v_i and v_i^* are the eigenvectors of Z and Z^* corresponding to the eigenvalues λ_i and λ_i^* , respectively, for $i \in [M]$. Suppose that $\min\{\lambda_{k-1}^* - \lambda_k^*, \lambda_\ell^* - \lambda_{\ell+1}^*\} > 0$, where $\lambda_0 = +\infty$ and $\lambda_{M+1} = -\infty$ by convention. Then, there exists an orthogonal matrix $Q^* \in \mathcal{O}(\ell - k + 1)$ such that

$$\|VQ^* - V^*\|_F \leq \frac{2\sqrt{2} \min\{\sqrt{\ell - k + 1} \|Z - Z^*\|, \|Z - Z^*\|_F\}}{\min\{\lambda_{k-1}^* - \lambda_k^*, \lambda_\ell^* - \lambda_{\ell+1}^*\}}.$$

To begin, let $\mathcal{G} \subseteq \mathcal{O}(d)$ be the closed subgroup of interest and $\eta > 0$ be a given parameter. Consider taking $Z = C$, $Z^* = \eta \cdot G^* G^{*\top}$, $k = 1$, $\ell = d$ in Theorem 5, where, as before, $C \in \mathbb{R}^{nd \times nd}$ is the matrix defined in (4) and $G^* \in \mathcal{G}^n$ is the ground truth. Note that each of the d columns of $\frac{1}{\sqrt{n}}G^*$ is an eigenvector of $\eta \cdot G^* G^{*\top}$ with eigenvalue ηn . The remaining $(n - 1)d$ eigenvalues of $\eta \cdot G^* G^{*\top}$ are all equal to 0. Hence, by letting $V_C \in \mathbb{R}^{nd \times d}$ to be the matrix whose j -th column is the eigenvector associated with the j -th largest eigenvalue of C for $j = 1, \dots, d$, we deduce from Theorem 5 the existence of a $Q^* \in \mathcal{O}(d)$ that satisfies

$$\begin{aligned} \left\| V_C Q^* - \frac{1}{\sqrt{n}} G^* \right\|_F &\leq \frac{2\sqrt{2} \min\left\{ \sqrt{d} \left\| C - \eta \cdot G^* G^{*\top} \right\|, \left\| C - \eta \cdot G^* G^{*\top} \right\|_F \right\}}{\min\{\lambda_{k-1}^* - \lambda_k^*, \lambda_\ell^* - \lambda_{\ell+1}^*\}} \\ &\leq \frac{2\sqrt{2d} \left\| C - \eta \cdot G^* G^{*\top} \right\|}{\eta n}. \end{aligned} \tag{36}$$

This, together with the definition of ε in (2), Lemmas 1 and 2, and inequality (36), implies that $\Pi^n(V_C Q^*)$ enjoys the estimation error bound

$$\begin{aligned} \varepsilon(\Pi^n(V_C Q^*)) &\leq \|\Pi^n(V_C Q^*) - G^*\|_F = \|\Pi^n(\sqrt{n} \cdot V_C Q^*) - G^*\|_F \\ &\leq 2 \left\| \sqrt{n} \cdot V_C Q^* - G^* \right\|_F \leq \frac{4\sqrt{2d}}{\eta\sqrt{n}} \left\| C - \eta \cdot G^* G^{*\top} \right\|. \end{aligned}$$

Unfortunately, the estimator $\Pi^n(V_C Q^*)$ is not implementable as we do not know Q^* in general. To work around this, let us construct an approximation \tilde{Q} of the unknown Q^* as follows. Consider an ϵ -net \mathcal{Q} of the quotient $\mathcal{O}(d)/\mathcal{G}$. By Definition 2, we have

$$\min_{[Q] \in \mathcal{Q}} \text{dist}([Q], [Q^*]) = \min_{[Q] \in \mathcal{Q}} \min_{Q' \in \mathcal{G}} \|Q - Q^* Q'\|_F \leq \epsilon.$$

Therefore, there must exist $[\tilde{Q}] \in \mathcal{Q}$ and $Q' \in \mathcal{G}$ such that

$$\|\tilde{Q} - Q^* Q'\|_F \leq \epsilon. \quad (37)$$

This motivates us to define the estimator

$$\tilde{G} = \Pi^n(V_C \tilde{Q}). \quad (38)$$

Let us now study the estimation performance of \tilde{G} .

Proposition 6. *Let $\eta > 0$, $\epsilon \geq 0$ be given parameters and \mathcal{Q} be an ϵ -net of the quotient $\mathcal{O}(d)/\mathcal{G}$. Consider the estimator \tilde{G} defined in (38). Then,*

$$\varepsilon(\tilde{G}) \leq 2\sqrt{n}\epsilon + \frac{4\sqrt{2d}}{\eta\sqrt{n}} \|C - \eta \cdot G^* G^{*\top}\|.$$

Proof. We have

$$\begin{aligned} \varepsilon(\tilde{G}) &= \min_{Q \in \mathcal{G}} \|\tilde{G} - G^* Q\|_F \leq \|\Pi^n(V_C \tilde{Q}) - G^* Q'\|_F \leq 2\sqrt{n} \left\| V_C \tilde{Q} - \frac{1}{\sqrt{n}} G^* Q' \right\|_F \\ &\leq 2\sqrt{n} \left(\left\| V_C \tilde{Q} - V_C Q^* Q' \right\|_F + \left\| V_C Q^* Q' - \frac{1}{\sqrt{n}} G^* Q' \right\|_F \right) \\ &\leq 2\sqrt{n} \left(\|V_C\| \cdot \|\tilde{Q} - Q^* Q'\|_F + \left\| V_C Q^* - \frac{1}{\sqrt{n}} G^* \right\|_F \right) \\ &\leq 2\sqrt{n}\epsilon + \frac{4\sqrt{2d}}{\eta\sqrt{n}} \|C - \eta \cdot G^* G^{*\top}\|, \end{aligned}$$

where the first inequality follows from (38); the second inequality follows from Lemmas 1 and 2; the last inequality follows from (36), (37), and the fact that $\|V_C\| \leq 1$. This completes the proof. \square

Compared with the bound on $\varepsilon(\Pi^n(V_C Q^*))$, we see that the bound on $\varepsilon(\tilde{G})$ has an extra term that is on the order of $\sqrt{n}\epsilon$. This can be attributed to the error incurred when using an element of an ϵ -net of the quotient $\mathcal{O}(d)/\mathcal{G}$ to approximate the unknown element $[Q^*]$. However, since the estimator \tilde{G} is defined via inequality (37), which involves the unknown orthogonal matrix Q^* , it is still not implementable. Nevertheless, the above idea can be

further developed to construct another estimator that not only is implementable but also enjoys a good estimation error bound. Specifically, motivated by the objective function of the least squares formulation (3), let us consider the function $\psi : \mathcal{O}(d)/\mathcal{G} \rightarrow \mathbb{R}$ defined by

$$\psi([Q]) := \langle C\Pi^n(V_C Q), \Pi^n(V_C Q) \rangle.$$

Since $\Pi^n(V_C Q Q') = \Pi^n(V_C Q) Q'$ for any $Q' \in \mathcal{G}$, the function ψ is independent of the choice of the representative Q of the equivalent class $[Q]$. This shows that ψ is a well-defined function on $\mathcal{O}(d)/\mathcal{G}$. Partly inspired by the work [75], in which a randomized rounding scheme for approximating the optimal solution to certain robust non-convex quadratic optimization problem is analyzed using an ϵ -net of the sphere, we propose the following estimator:

Algorithm 2 Entropic Spectral Estimator

- 1: **Input:** the matrix C and an ϵ -net \mathcal{Q} of the quotient $\mathcal{O}(d)/\mathcal{G}$
- 2: find the maximizer $[\hat{Q}]$ of the problem

$$\max_{[Q] \in \mathcal{Q}} \psi([Q])$$

- 3: **Output:** the matrix $\hat{G} = \Pi^n(V_C \hat{Q})$
-

When $\mathcal{G} = \mathcal{O}(d)$, there is only one equivalent class, viz. the orthogonal group $\mathcal{O}(d)$ itself. In this case, there is no need to input an ϵ -net, and the entropic spectral estimator \hat{G} reduces to the spectral estimator in the works [14] and [53]. Therefore, the entropic spectral estimator can be seen as a generalization of these spectral estimators. When $\mathcal{G} = \mathcal{SO}(d)$, there are two equivalent classes: One formed by the set of orthogonal matrices with determinant $+1$ and the other with determinant -1 . Taking the representatives I_d and $\text{Diag}(-1, 1, \dots, 1)$ for these two equivalent classes, respectively, the entropic estimator is either $\Pi^n(V_C)$ or $\Pi^n(V'_C)$ with $V'_C = V_C \cdot \text{Diag}(-1, 1, \dots, 1)$, depending on whether $\langle C\Pi^n(V_C), \Pi^n(V_C) \rangle \geq \langle C\Pi^n(V'_C), \Pi^n(V'_C) \rangle$. For a general closed subgroup \mathcal{G} of $\mathcal{O}(d)$, in order to use the proposed estimator, we have to construct an ϵ -net of the quotient $\mathcal{O}(d)/\mathcal{G}$. Proposition 4 provides an explicit and efficient way to achieve this. All we have to do is to generate a set of independent, uniformly distributed random orthogonal matrices.

Interestingly, Proposition 4 reveals an intimate relation between our proposed estimator and the notion of metric entropy: The smaller the metric entropy of the quotient $\mathcal{O}(d)/\mathcal{G}$, the fewer independent copies of uniformly distributed random orthogonal matrices we need to construct the ϵ -net \mathcal{Q} . This explains why we name our estimator \hat{G} the *entropic spectral estimator*.

The main result of this subsection is the following theorem, which concerns the estimation performance of the entropic spectral estimator.

Theorem 6. Suppose that the group \mathcal{G} satisfies Condition 2 with parameter $\beta \in (0, 1]$. Let $\eta > 0$ and $\epsilon \geq 0$ be given parameters. Then, the entropic spectral estimator \hat{G} returned by Algorithm 2 satisfies

$$\varepsilon(\hat{G}) \leq 2\sqrt{\frac{2n}{\beta}}\epsilon + \frac{12\sqrt{d}}{\beta\eta\sqrt{n}} \|C - \eta \cdot G^* G^{*\top}\|.$$

Theorem 6 shows that even though we do not have access to the element $[\tilde{Q}] \in \mathcal{Q}$ defined in (37) and hence cannot construct the estimator \tilde{G} in (38), we can get hold of another element $[\hat{Q}] \in \mathcal{Q}$ by maximizing the least squares-based function ψ over \mathcal{Q} and use it to construct the estimator \hat{G} , whose estimation error bound is worse than that of the estimator \tilde{G} by roughly a factor of $\frac{1}{\sqrt{\beta}}$ (recall that the parameter $\beta \in (0, 1]$ is related to the geometry of the subgroup \mathcal{G} , see Condition 2). As shown in Theorem 2, for many groups of interest (such as the orthogonal group $\mathcal{O}(d)$, the special orthogonal group $\mathcal{SO}(d)$, and the cyclic group \mathcal{Z}_m), the parameter β is a constant, which implies that the bounds on $\varepsilon(\tilde{G})$ and $\varepsilon(\hat{G})$ differ by at most a constant factor.

Proof of Theorem 6. Let \tilde{G} be defined as in (38). By definition of the entropic spectral estimator \hat{G} ,

$$\text{Tr}(\hat{G}^\top C \hat{G}) \geq \text{Tr}(\tilde{G}^\top C \tilde{G}). \quad (39)$$

Upon letting $\Delta_\eta = C - \eta \cdot G^* G^{*\top}$, for any $Q \in \mathcal{G}$, we have

$$\begin{aligned} & \text{Tr}(\hat{G}^\top \Delta_\eta \hat{G}) - \text{Tr}(G^{*\top} \Delta_\eta G^*) = \text{Tr}((\hat{G}Q - G^*)^\top \Delta_\eta (\hat{G}Q + G^*)) \\ & \leq \|\hat{G}Q - G^*\|_F \cdot \|\Delta_\eta\| \cdot (\|\hat{G}Q\|_F + \|G^*\|_F) \leq 2\sqrt{nd} \cdot \|\Delta_\eta\| \cdot \|\hat{G}Q - G^*\|_F. \end{aligned}$$

This implies that

$$\begin{aligned} \text{Tr}(\hat{G}^\top C \hat{G}) &= \eta \cdot \|G^{*\top} \hat{G}\|_F^2 + \text{Tr}(\hat{G}^\top \Delta_\eta \hat{G}) \\ &\leq \eta \cdot \|G^{*\top} \hat{G}\|_F^2 + \text{Tr}(G^{*\top} \Delta_\eta G^*) + 2\sqrt{nd} \cdot \|\Delta_\eta\| \cdot \varepsilon(\hat{G}). \end{aligned} \quad (40)$$

Similarly, we can get

$$\begin{aligned} \text{Tr}(\tilde{G}^\top C \tilde{G}) &= \eta \cdot \|G^{*\top} \tilde{G}\|_F^2 + \text{Tr}(\tilde{G}^\top \Delta_\eta \tilde{G}) \\ &\geq \eta \cdot \|G^{*\top} \tilde{G}\|_F^2 + \text{Tr}(G^{*\top} \Delta_\eta G^*) - 2\sqrt{nd} \cdot \|\Delta_\eta\| \cdot \varepsilon(\tilde{G}). \end{aligned} \quad (41)$$

Now, for any $G \in \mathcal{G}^n$, we have the identity

$$\varepsilon(G)^2 = \|G - G^* Q_G\|_F^2 = 2 \text{Tr}(n \cdot I_d - Q_G^\top G^{*\top} G), \quad (42)$$

where $Q_G \in \mathcal{G} \subseteq \mathcal{O}(d)$ is defined in (7). This yields

$$\left\| G^{*\top} \tilde{G} \right\|_F^2 = \left\| Q_{\tilde{G}}^\top G^{*\top} \tilde{G} \right\|_F^2 \geq \frac{1}{d} \left(\text{Tr} \left(Q_{\tilde{G}}^\top G^{*\top} \tilde{G} \right) \right)^2 = \left(n\sqrt{d} - \frac{1}{2\sqrt{d}} \varepsilon(\tilde{G})^2 \right)^2,$$

where the inequality follows from the Cauchy-Schwarz inequality. Upon substituting the above into (41), we obtain

$$\begin{aligned} \text{Tr} \left(\tilde{G}^\top C \tilde{G} \right) &\geq \eta \left(n\sqrt{d} - \frac{1}{2\sqrt{d}} \varepsilon(\tilde{G})^2 \right)^2 + \text{Tr} \left(G^{*\top} \Delta_\eta G^* \right) - 2\sqrt{nd} \cdot \|\Delta_\eta\| \cdot \varepsilon(\tilde{G}) \\ &\geq \eta n^2 d - \eta n \cdot \varepsilon(\tilde{G})^2 + \text{Tr} \left(G^{*\top} \Delta_\eta G^* \right) - 2\sqrt{nd} \cdot \|\Delta_\eta\| \cdot \varepsilon(\tilde{G}). \end{aligned} \quad (43)$$

It follows from (39), (40), and (43) that

$$\begin{aligned} \eta \left\| G^{*\top} \hat{G} \right\|_F^2 &\geq \text{Tr} \left(\hat{G}^\top C \hat{G} \right) - \text{Tr} \left(G^{*\top} \Delta_\eta G^* \right) - 2\sqrt{nd} \cdot \|\Delta_\eta\| \cdot \varepsilon(\hat{G}) \\ &\geq \text{Tr} \left(\tilde{G}^\top C \tilde{G} \right) - \text{Tr} \left(G^{*\top} \Delta_\eta G^* \right) - 2\sqrt{nd} \cdot \|\Delta_\eta\| \cdot \varepsilon(\hat{G}) \\ &\geq \eta n^2 d - \eta n \cdot \varepsilon(\tilde{G})^2 - 2\sqrt{nd} \cdot \|\Delta_\eta\| \cdot \left(\varepsilon(\hat{G}) + \varepsilon(\tilde{G}) \right). \end{aligned} \quad (44)$$

Next, using the definition of $Q_{\hat{G}}$ (see (7)), identity (42), Condition 2 (with $X = \frac{1}{n} G^{*\top} \hat{G}$), and inequality (44), we have

$$\begin{aligned} \frac{\beta}{2} \varepsilon(\hat{G})^2 &= \beta \text{Tr} \left(n \cdot I_d - Q_{\hat{G}}^\top G^{*\top} \hat{G} \right) \leq nd - \frac{1}{n} \left\| G^{*\top} \hat{G} \right\|_F^2 \\ &\leq \varepsilon(\tilde{G})^2 + \frac{2\sqrt{d} \cdot \|\Delta_\eta\|}{\eta\sqrt{n}} \left(\varepsilon(\tilde{G}) + \varepsilon(\hat{G}) \right). \end{aligned}$$

This, together with the fact that $\beta \in (0, 1]$, gives

$$\begin{aligned} \left(\frac{\beta}{2} \varepsilon(\hat{G}) - \frac{\sqrt{d} \cdot \|\Delta_\eta\|}{\eta\sqrt{n}} \right)^2 &= \frac{\beta^2}{4} \varepsilon(\hat{G})^2 - \frac{\beta\sqrt{d} \cdot \|\Delta_\eta\|}{\eta\sqrt{n}} \varepsilon(\hat{G}) + \left(\frac{\sqrt{d} \cdot \|\Delta_\eta\|}{\eta\sqrt{n}} \right)^2 \\ &\leq \frac{\beta}{2} \varepsilon(\tilde{G})^2 + \frac{\beta\sqrt{d} \cdot \|\Delta_\eta\|}{\eta\sqrt{n}} \varepsilon(\tilde{G}) + \left(\frac{\sqrt{d} \cdot \|\Delta_\eta\|}{\eta\sqrt{n}} \right)^2 \\ &\leq \left(\sqrt{\frac{\beta}{2}} \varepsilon(\tilde{G}) + \frac{\sqrt{d} \cdot \|\Delta_\eta\|}{\eta\sqrt{n}} \right)^2. \end{aligned}$$

It follows that

$$\varepsilon(\hat{G}) \leq \sqrt{\frac{2}{\beta}} \varepsilon(\tilde{G}) + \frac{4\sqrt{d}}{\beta\eta\sqrt{n}} \|\Delta_\eta\| \leq \sqrt{\frac{2}{\beta}} \left(2\epsilon\sqrt{n} + \frac{4\sqrt{2d}}{\eta\sqrt{n}} \|\Delta_\eta\| \right) + \frac{4\sqrt{d}}{\beta\eta\sqrt{n}} \|\Delta_\eta\|$$

$$\leq 2\sqrt{\frac{2n}{\beta}}\epsilon + \frac{4\sqrt{d}}{\eta\sqrt{n}}\left(2\sqrt{\frac{1}{\beta}} + \frac{1}{\beta}\right)\|\Delta_\eta\| \leq 2\sqrt{\frac{2n}{\beta}}\epsilon + \frac{12\sqrt{d}}{\beta\eta\sqrt{n}}\|\Delta_\eta\|,$$

where the second inequality follows from Proposition 6. This completes the proof. \square

Armed with Theorem 6, we now show that when the measurement graph and additive noise follow the settings in Sections 5.2 and 5.3, respectively, condition (iv) in the master theorem will be satisfied with high probability.

Theorem 7. *Suppose that (i) the group \mathcal{G} satisfies Conditions 1 and 2 with parameters $\alpha \geq 1$ and $\beta \in (0, 1]$, respectively; (ii) the measurement graph is an Erdős-Rényi random graph with observation rate $p \in (0, 1]$; (iii) the noise matrices $\{\Theta_{ij} : 1 \leq i < j \leq n\}$ are independent of each other and of the measurement graph, and whose entries are i.i.d. sub-Gaussian random variables with parameter $\sigma > 0$. Then, there exist constants $c_0, c_1, c_2, c_3, c_4 > 0$ such that when either (a) $p = 1$ and $\sigma \leq \frac{c_0\beta\sqrt{n}}{\alpha d}$ or (b) $p \in \left[c_1 \cdot \max\left\{\frac{\alpha^2 d}{\beta^2 n}, \frac{\log n}{n}\right\}, 1\right)$ and $\sigma \leq \frac{c_2\beta\sqrt{pn}}{\alpha d\sqrt{\log(nd)}}$, the entropic spectral estimator \hat{G} generated by Algorithm 2 with $\epsilon \leq \frac{c_3\sqrt{\beta}}{\alpha}$ will satisfy*

$$\varepsilon(\hat{G}) \leq \frac{\sqrt{n}}{2\alpha}$$

with probability at least $1 - \frac{c_4}{n}$.

Let us briefly discuss how to construct an ϵ -net of the quotient $\mathcal{O}(d)/\mathcal{G}$ that satisfies the requirement of Theorem 7 for some specific groups \mathcal{G} . If \mathcal{G} is either the orthogonal group $\mathcal{O}(d)$ or the special orthogonal group $\mathcal{SO}(d)$, then we can simply take the 0-net of the quotient $\mathcal{O}(d)/\mathcal{G}$; see the remark following Algorithm 2. If \mathcal{G} is a discrete subgroup of $\mathcal{O}(d)$, then we can find a desired ϵ -net by invoking Propositions 4 and 5. Specifically, there exists a constant $c > 0$ such that given any $\delta \in (0, 1)$, if we generate

$$K \geq \frac{\log \delta}{\log \left(1 - |\mathcal{G}| \left(\frac{c\sqrt{\beta}}{\alpha}\right)^{\frac{d(d-1)}{2}}\right)} \quad (45)$$

random orthogonal matrices Q_1, \dots, Q_K that are independently and uniformly distributed on $\mathcal{O}(d)$, then with probability at least $1 - \delta$, the set $\mathcal{Q} = \{[Q_1], \dots, [Q_K]\}$ will be an ϵ -net of $\mathcal{O}(d)/\mathcal{G}$ satisfying the requirement of Theorem 7.

Proof of Theorem 7. Let $C^* \in \mathbb{R}^{nd \times nd}$ be the block matrix defined by $[C^*]_{ij} = w_{ij}G_i^*G_j^{*\top}$, $i, j = 1, \dots, n$ and $\Delta^* := C^* - p \cdot G^*G^{*\top}$. Using the definition of Δ in (6), we can write

$$C - p \cdot G^*G^{*\top} = \Delta^* + C - C^* = \Delta^* + \Delta.$$

Thus, by invoking Theorem 6 with $\eta = p$ and $\epsilon \leq \frac{\sqrt{\beta}}{8\sqrt{2}\alpha}$, we have

$$\varepsilon(\hat{G}) \leq 2\sqrt{\frac{2n}{\beta}}\epsilon + \frac{12\sqrt{d}}{\beta p\sqrt{n}} \left\| C - p \cdot G^* G^{*\top} \right\| \leq \frac{\sqrt{n}}{4\alpha} + \frac{12\sqrt{d}}{\beta p\sqrt{n}} (\|\Delta^*\| + \|\Delta\|). \quad (46)$$

Let us now bound $\|\Delta^*\|$ and $\|\Delta\|$ separately.

For $p = 1$, Proposition 3 implies the existence of a constant $c' > 0$ such that

$$\Pr \left(\|\Delta\| \geq c' \sigma \sqrt{nd} \right) \leq \frac{1}{n}.$$

Thus, there exists a constant $c_0 > 0$ such that when $\sigma \leq \frac{c_0 \beta \sqrt{n}}{\alpha d}$, we will have

$$\frac{12\sqrt{d}}{\beta\sqrt{n}} \|\Delta\| \leq \frac{\sqrt{n}}{4\alpha}$$

with probability at least $1 - \frac{1}{n}$. Since $\Delta^* = \mathbf{0}$ when $p = 1$, we conclude from (46) that $\varepsilon(\hat{G}) \leq \frac{\sqrt{n}}{2\alpha}$.

For $p < 1$, the proof of Proposition 3 shows that for any $p \geq \frac{2}{e^2+1} \cdot \frac{\log n}{n-1}$, we have

$$\Pr \left(\|\Delta\| \geq 2e\sigma \sqrt{pd(n-1)\log(nd)} \right) \leq \frac{3}{n}.$$

On the other hand, observe that

$$\Delta^* = \left(\bar{W} - p \cdot ee^\top \right) \boxtimes \left(G^* G^{*\top} \right),$$

where $\bar{W} \in \mathbb{R}^{n \times n}$ is the matrix defined in Section 4 and \boxtimes denotes the block Kronecker product introduced in [36, Definition 2.5]. Since $G^* G^{*\top}$ is positive semidefinite with $\left[G^* G^{*\top} \right]_{ii} = I_d$ for $i = 1, \dots, n$, by [36, Corollary 4.5(b)] and the identity $\bar{W} - p \cdot ee^\top = \bar{W} - \mathbb{E}[\bar{W}] + (1-p)I_n$, we have

$$\|\Delta^*\| \leq \left\| \bar{W} - p \cdot ee^\top \right\| \leq 1 + \left\| \bar{W} - \mathbb{E}[\bar{W}] \right\|.$$

This, together with (26), implies that for any $p \geq \frac{\log n}{n}$,

$$\Pr \left(\|\Delta^*\| \geq 173\sqrt{pn} \right) \leq \exp \left(-\frac{pn}{8} \right).$$

The above calculations show that when

$$p \geq \max \left\{ \frac{(173 \times 96)^2 \alpha^2 d}{\beta^2 n}, \frac{8 \log n}{n} \right\}, \quad \sigma \leq \frac{\beta \sqrt{pn}}{192 e \alpha d \sqrt{\log(nd)}},$$

we will have

$$\frac{12\sqrt{d}}{\beta p\sqrt{n}}\|\Delta^*\| \leq \frac{\sqrt{n}}{8\alpha} \quad \text{and} \quad \frac{12\sqrt{d}}{\beta p\sqrt{n}}\|\Delta\| \leq \frac{\sqrt{n}}{8\alpha}$$

with probability at least $1 - \frac{4}{n}$. Upon substituting these bounds into (46), we obtain $\varepsilon(\hat{G}) \leq \frac{\sqrt{n}}{2\alpha}$. \square

We remark that there are works studying the estimation performance of various spectral estimators for $\mathcal{O}(d)$ -sync [45], $\mathcal{SO}(2)$ -sync [14], and $\mathcal{P}(d)$ -sync [4, 45, 53, 59], and it is worth comparing the results for these estimators with that for our entropic spectral estimator. For $\mathcal{O}(d)$ -sync, the work [45] establishes a bound on the estimation error, measured in the *operator norm*, of each block $[\hat{G}]_1, \dots, [\hat{G}]_n$ of its proposed spectral estimator \hat{G} . By contrast, our work establishes a bound on the estimation error, measured in the *Frobenius norm*, of the proposed entropic spectral estimator in its entirety. Although for $\mathcal{O}(d)$ -sync the blockwise error bound in [45, Theorem 3.1] is generally sharper than the error bound in Theorem 7 of our work, the former applies only to the setting of additive *Gaussian* noise and *complete* measurement graph, while the latter applies to the more general setting of additive *sub-Gaussian* noise and *Erdős-Rényi* measurement graph with observation rate p that can go down to the order of $\max\left\{\frac{d}{n}, \frac{\log n}{n}\right\}$. For $\mathcal{SO}(2)$ -sync, the work [14] studies the estimation error of a spectral estimator under the additive noise model with a *complete* measurement graph. Since such a setting is covered by that of Theorem 6, we can compare the corresponding estimation error bounds. Recall from our discussion immediately following Theorem 7 that we may take $\epsilon = 0$. Moreover, we have $\beta = \frac{1}{2}$ by Theorem 2. Thus, Theorem 6 (with $\epsilon = 0$, $d = 2$, $\beta = \frac{1}{2}$, $\eta = 1$) and the definition of Δ in (6) imply that the estimation error of the entropic spectral estimator is at most on the order of $\frac{\|\Delta\|}{\sqrt{n}}$, which is the same as that of the spectral estimator in [14]; see [14, Lemma 6]. Lastly, for $\mathcal{P}(d)$ -sync, the works [4, 45, 53, 59] consider an outlier noise model, which is different from the additive noise model considered in our work. As such, the corresponding estimation error bounds cannot be compared directly.

5.5 Putting the Pieces Together

Now, let us summarize our technical developments so far. By combining the results in Theorems 2, 3, 4, and 7, we see that under the setting of Theorem 7 and with \mathcal{G} being either the orthogonal group $\mathcal{O}(d)$, the special orthogonal group $\mathcal{SO}(d)$, the permutation group $\mathcal{P}(d)$, or the cyclic group \mathcal{Z}_m , all four conditions (i)–(iv) in the master theorem (Theorem 1) will be satisfied with high probability. Consequently, the iterates generated by GPM will enjoy the estimation error bound stated in the theorem. To the best of our knowledge, this is the first time an estimation performance guarantee of such generality is obtained for GPM.

As is evident from Theorems 2, 3, 4, and 7, the aforementioned guarantee requires the observation rate p to be at least on the order of $\max \left\{ \frac{\alpha^2 d}{\beta^2 n}, \frac{\log n}{n} \right\}$ and the noise level σ to be at most on the order of $\frac{\beta \sqrt{n}}{\alpha d}$ (resp. $\frac{\beta \sqrt{pn}}{\alpha d \sqrt{\log(nd)}}$) when $p = 1$ (resp. $p < 1$), where $\alpha \geq 1$ and $\beta \in (0, 1]$ are related to the error-bound geometry of the subgroup \mathcal{G} (see Conditions 1 and 2). When d is at most on the order of $\log n$, the order of the bound on p is optimal in the sense it matches the order of the connectivity threshold of an Erdős-Rényi random graph; see the remark after the proof of Theorem 3. Moreover, when $\mathcal{G} = \mathcal{O}(d)$ and $p = 1$, our bound on σ has a more favorable order than that in [44, Theorem 3.2]. Nevertheless, we should point out that our result pertains to the *estimation* performance of GPM, while that in [44] pertains to the *optimization* performance (*i.e.*, convergence behavior) of GPM. In particular, as detailed in the discussion following Theorem 1, our master theorem guarantees that the estimation error of the iterates generated by the suitably initialized GPM decreases at least geometrically to some threshold, while [44, Theorem 3.2] establishes the linear convergence of the iterates, albeit only for the setting of additive Gaussian noise and complete measurement graph.

6 Numerical Results

We have conducted numerical experiments to compare the computational speed, scalability, and estimation performance of our proposed entropic spectral estimator and the GPM-based non-convex approach with those of existing methods. In our experiments, we have considered noise models that go beyond the additive one covered by our theoretical development so as to test the viability of our proposed approach. All our codes are implemented using MATLAB and tested on a desktop with Intel Core i7-10700 CUP (2.90GHz \times 8). As will be seen from the results, our approach demonstrates superior performance in many different experiment settings.

6.1 Special Orthogonal Synchronization

We first present numerical results on $\mathcal{SO}(d)$ -sync.

6.1.1 Setting

We focus on the case where $d = 3$, which is most relevant to real-world applications. The experiment setting, which is the same as that in [71], is as follows. We take an Erdős-Rényi random graph with observation rate $p \in (0, 1]$ as the measurement graph $([n], E)$. We consider a multiplicative noise model with two layers of multiplicative noise. More precisely, the observations are given by

$$C_{ij} = G_i^* G_j^{*\top} \Theta_{ij}^{\text{out}} \Theta_{ij}^{\text{Lan}}, \quad (i, j) \in E,$$

where Θ_{ij}^{out} is the so-called outlier noise defined by

$$\Theta_{ij}^{\text{out}} = \begin{cases} I_3, & \text{with probability } q, \\ Q_{ij} \sim \text{Uniform}(\mathcal{SO}(3)), & \text{with probability } 1 - q \end{cases}$$

with $q \in (0, 1]$ being the non-corruption rate, $\text{Uniform}(\mathcal{SO}(3))$ being the uniform distribution on $\mathcal{SO}(3)$, and $\Theta_{ij}^{\text{Lan}} \in \mathcal{SO}(3)$ being generated according to the Langevin distribution (also called the von Mises-Fisher distribution) [22, 71] on $\mathcal{SO}(3)$ with mean I_3 and concentration parameter $\gamma \geq 0$, *i.e.*, the density function of each Θ_{ij}^{Lan} is given by

$$c(\gamma) \exp(\gamma \text{Tr}(\Theta_{ij}^{\text{Lan}}))$$

for some normalization constant $c(\gamma) > 0$. The parameter $\gamma \geq 0$ controls the concentration of the random matrix Θ_{ij}^{Lan} around the mean I_3 — the larger the parameter γ , the more concentrated around the mean I_3 the random matrix Θ_{ij}^{Lan} is. In particular, it is the uniform distribution $\text{Uniform}(\mathcal{SO}(3))$ when $\gamma = 0$. As $\gamma \rightarrow +\infty$, the distribution behaves like a Gaussian distribution with mean I_3 and variance $\frac{1}{\gamma}$.

6.1.2 Results

We compare our proposed entropic spectral estimator for $\mathcal{SO}(d)$ -sync and the GPM-based non-convex approach with the least unsquared deviation approach in [71] and the low-rank-sparse decomposition approach in [3]. We also include the standard spectral estimator in the comparison as baseline. The codes for the least unsquared deviation approach are provided by the authors of [71], while those for the low-rank-sparse decomposition approach are available online.⁵ In our experiments, we use the default choice for all the parameters in their codes. Since the two competing methods are known to perform well against outlier noise, we mainly study the recovery performance and computational time by varying the proportion of outliers $1 - q$. For ease of comparison, we measure the recovery performance using the *normalized estimation error* $\frac{\epsilon(G)}{\sqrt{2nd}}$, whose value always lies between 0 and 1. Figures 2 and 3 show the experiment results in the low noise ($\gamma = 1$) and high noise ($\gamma = 0.4$) regimes, respectively. The reported time for our non-convex approach includes the time for computing the entropic spectral estimator, as the former is initialized by the latter. All the points in the figures are obtained by averaging over 30 independent random instances.

From Figure 2, we see that the estimation performance of the proposed entropic spectral estimator is significantly better than that of the standard spectral estimator. Moreover, when GPM is initialized by the entropic spectral estimator, it yields an estimator whose performance further improves upon that of the entropic spectral estimator. We also note that in the low noise regime, the non-convex approach performs generally on par with

⁵<https://fusiello.github.io/demo/gmf/index.html>

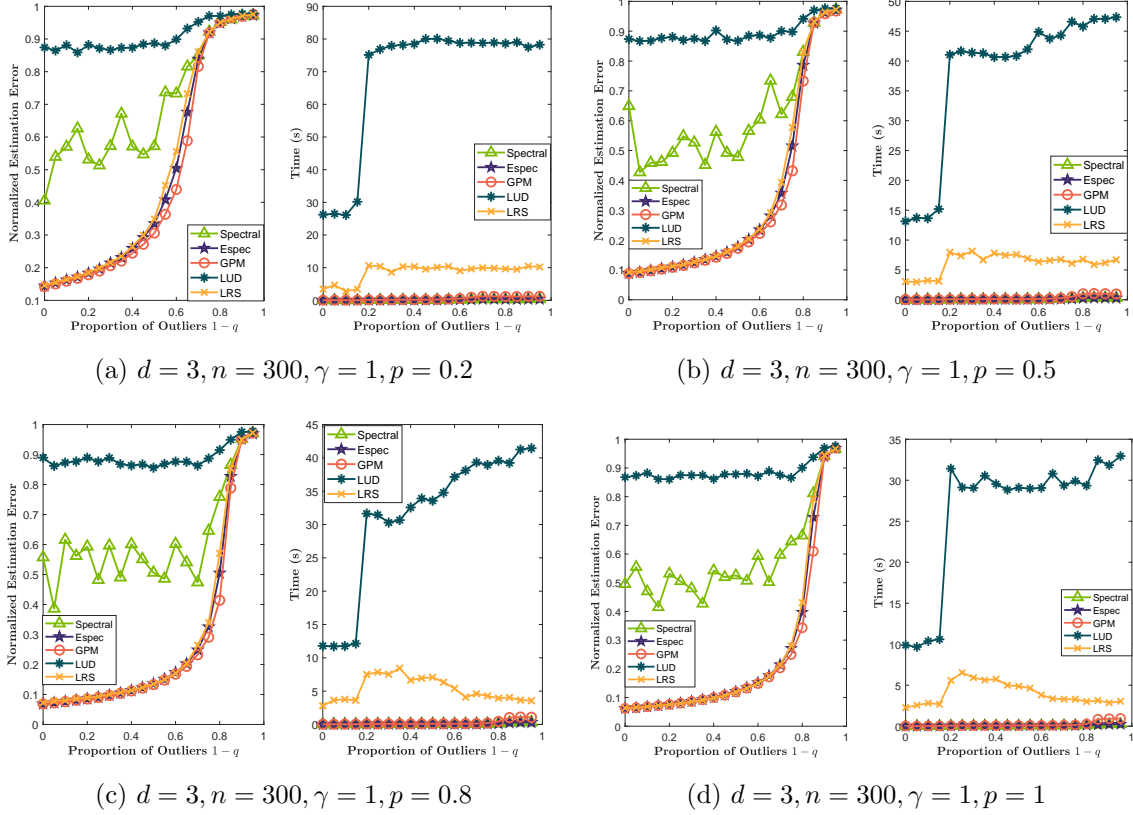


Figure 2: Estimation error and computational time of the standard spectral estimator (labeled as Spectral), the entropic spectral estimator (labeled as Espec), the non-convex approach initialized by the entropic spectral estimator (labeled as GPM), the least unsquared deviation approach (labeled as LUD), and the low-rank-sparse decomposition approach (labeled as LRS) under the setting $d = 3$, $n = 300$, $\gamma = 1$, and $p \in \{0.2, 0.5, 0.8, 1\}$.

the low-rank-sparse decomposition approach in terms of estimation error. This suggests that our approach is fairly robust to multiplicative and outlier noise, given that the low-rank-sparse decomposition approach was shown empirically to possess such a desirable property [3]. As for the computation time, both our entropic spectral estimator and the GPM-based non-convex approach are considerably faster than the low-rank-sparse decomposition approach. Quite surprisingly, the least unsquared deviation approach performs worse than the spectral estimator in terms of estimation error. The optimization problem associated with the least unsquared deviation approach is a nonlinear convex semidefinite program obtained via the relaxation technique and solved by the alternating direction augmented Lagrangian method [74]. We suspect that the unsatisfactory performance of the

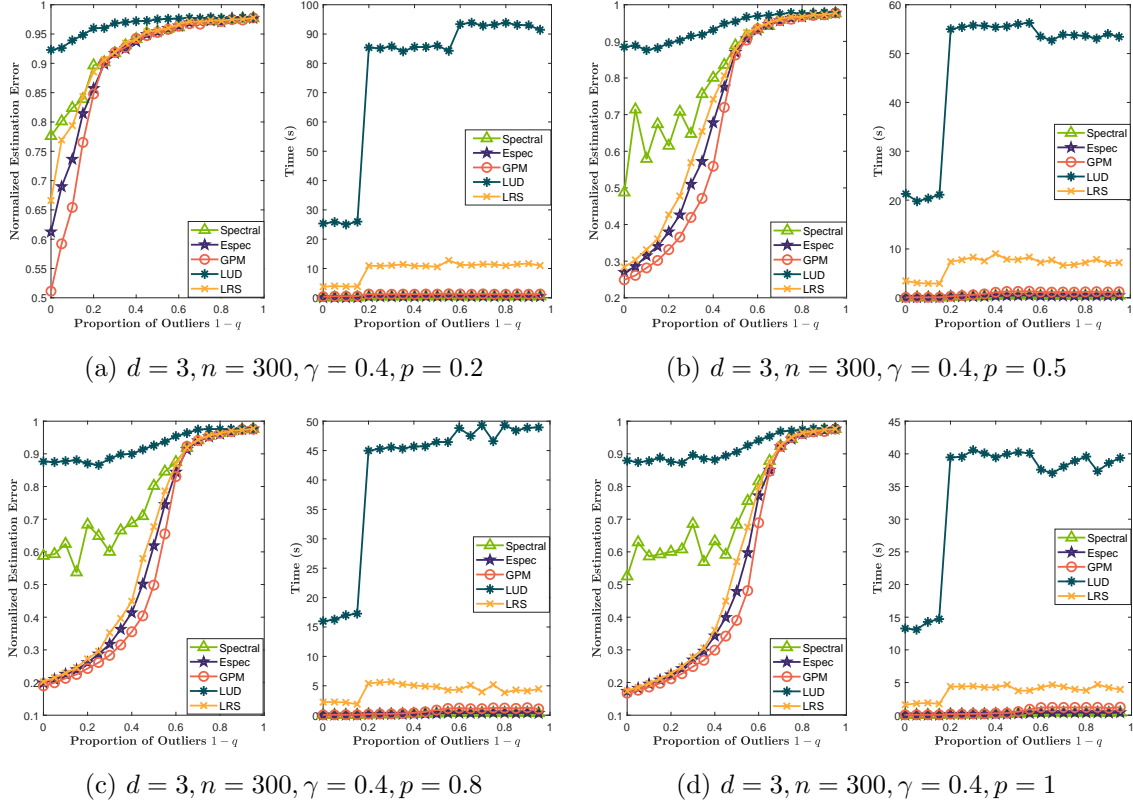


Figure 3: Estimation error and computational time of the standard spectral estimator (labeled as Spectral), the entropic spectral estimator (labeled as Espec), the non-convex approach initialized by the entropic spectral estimator (labeled as GPM), the least unsquared deviation approach (labeled as LUD), and the low-rank-sparse decomposition approach (labeled as LRS) under the setting $d = 3, n = 300, \gamma = 0.4$, and $p \in \{0.2, 0.5, 0.8, 1\}$.

least unsquared deviation approach might be due to the fact that the alternating direction augmented Lagrangian method can only achieve low to medium accuracy and/or is sensitive to the choice of the penalty parameter in the augmented Lagrangian term.

The experiment results in the high noise regime are shown in Figure 3. The behavior of the algorithms is mostly similar to that in the low noise case, except that the estimation performance of our GPM-based non-convex approach is fairly better than that of the low-rank-sparse decomposition approach.

6.2 Permutation Synchronization

Next, we present numerical results on $\mathcal{P}(d)$ -sync.

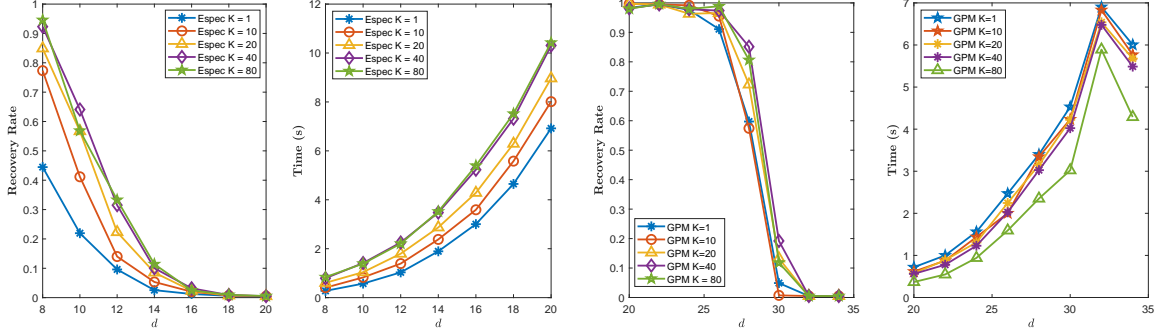


Figure 4: Recovery rate and computational time of the proposed entropic spectral estimator (labeled as Espec) and the non-convex approach initialized by the entropic spectral estimator (labeled as GPM) under the setting $d = 20$, $n = 200$, $p = 0.5$, $q = 0.8$, $\delta = 1$, and $K \in \{1, 10, 20, 40, 80\}$.

6.2.1 Setting

Again, we take an Erdős-Rényi random graph with observation rate $p \in (0, 1)$ as the measurement graph $([n], E)$. We consider an adversarial measurement model that contains both additive and multiplicative noise. Specifically, the observations are given by

$$C_{ij} = \Pi_{\mathcal{P}(d)}(G_i^* G_j^{*\top} \Xi_{ij}^{\text{out}} + \delta W_{ij}), \quad (i, j) \in E,$$

where Ξ_{ij}^{out} is the outlier noise defined by

$$\Xi_{ij}^{\text{out}} = \begin{cases} I_d, & \text{with probability } q, \\ P_{ij} \sim \text{Uniform}(\mathcal{P}(d)), & \text{with probability } 1 - q \end{cases}$$

with $q \in (0, 1]$ being the non-corruption rate, $\text{Uniform}(\mathcal{P}(d))$ being the uniform distribution on $\mathcal{P}(d)$, $\{W_{ij} : (i, j) \in E\}$ being independent random matrices with i.i.d. standard Gaussian entries, and $\delta \geq 0$ being a parameter controlling the magnitude of the additive noise. Due to the discrete nature of the signals (permutation matrices), instead of the estimation error $\varepsilon(\cdot)$, we quantify the estimation performance by the *recovery rate*, which is defined as

$$\text{Recovery Rate} = \frac{\text{number of correctly recovered group elements}}{n}. \quad (47)$$

6.2.2 Results

Recall that the entropic spectral estimator for $\mathcal{P}(d)$ -sync requires as input an ϵ -net of the quotient $\mathcal{O}(d)/\mathcal{P}(d)$ (see Algorithm 2). Following the approach discussed in the paragraph

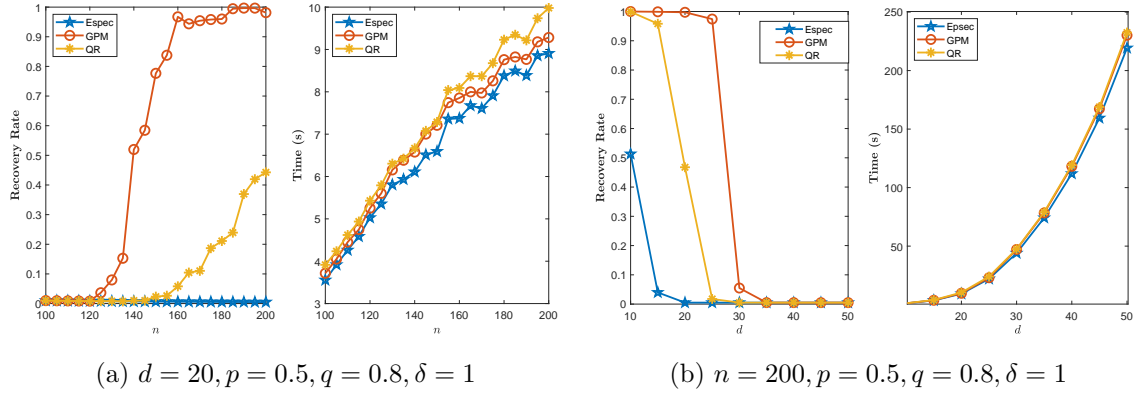


Figure 5: Recovery rate and computational time of the entropic spectral estimator with $K = 40$ (labeled as Espec), the non-convex approach initialized by the entropic spectral estimator (labeled as GPM), and the QR factorization-based approach initialized by the entropic spectral estimator (labeled as QR) as the parameters n and d vary.

immediately after Theorem 7, such an ϵ -net can be found with high probability by generating K independent, uniformly distributed random orthogonal matrices, where K is sufficiently large (see (45)). In our first experiment, we investigate the performance of the entropic spectral estimator and the GPM-based non-convex approach as K varies. Figure 4 shows the results for $K \in \{1, 10, 20, 40, 80\}$. The algorithm we used to generate uniformly distributed random orthogonal matrices is based on the paper [26], see also [52] for more details. Note that when $K = 1$, the entropic spectral estimator reduces to the standard spectral estimator. From the figure, we see that the recovery performance of both the entropic spectral estimator and the non-convex approach improves as K increases, but the marginal benefit becomes smaller and smaller. Furthermore, we see that the computational time does not increase significantly as K increases.

In our second experiment, we compare the performance of our entropic spectral estimator (with $K = 40$) and GPM-based non-convex approach with that of the QR factorization-based iterative approach developed in [59] as the parameters n and d vary. The results are summarized in Figure 5. All the points in the figure are obtained by averaging over 30 independent random instances. As we can see from Figure 5, the recovery performance of our GPM-based non-convex approach is significantly better than that of the entropic spectral estimator and the QR factorization-based approach. Furthermore, we see that our GPM-based non-convex approach is slightly faster than the QR factorization-based approach. Of course, the computational time of the entropic spectral estimator is the shortest among the three tested methods, since both the GPM-based non-convex approach and the QR factorization-based approach are initialized by the entropic spectral estimator.

6.3 Cyclic Synchronization (Joint Alignment Problem)

Finally, we present our numerical results on \mathcal{Z}_m -sync. We remark that \mathcal{Z}_m -sync is equivalent to the joint alignment problem considered in [20].

6.3.1 Setting

We adopt the same experiment setting as in [20]. Specifically, we take an Erdős-Rényi random graph with observation rate $p \in (0, 1)$ as the measurement graph $([n], E)$. We consider a multiplicative noise model, in which the observations are given by

$$C_{ij} = G_i^* G_j^{*\top} \Theta_{ij}^{\text{out}}, \quad (i, j) \in E.$$

Here,

$$\Theta_{ij}^{\text{out}} = \begin{cases} I_2, & \text{with probability } q, \\ Q_k, & \text{with probability } 1 - q \end{cases}$$

with $q \in (0, 1]$ being the non-corruption rate, Q_k being defined in (19), and $k \sim \text{Uniform}([m])$ being uniformly distributed on $[m]$. To quantify the estimation performance, we again use the recovery rate⁶ defined in (47).

6.3.2 Results

We compare the performance of the proposed entropic spectral estimator and GPM-based non-convex approach with that of another non-convex approach named *projected power method*, which is developed in [20]. The standard spectral estimator for \mathcal{Z}_m -sync is once again included in the experiments as a baseline. We focus on how the recovery rate and computational time of these approaches depend on the order m of the cyclic group. The results are plotted in Figure 6. All the points in the figure are obtained by averaging over 30 independent random instances.

As Figure 6 shows, our proposed entropic spectral estimator with $K = 10$ significantly outperforms the standard spectral estimator, and the non-convex approach can further improve the recovery rate, albeit by a small margin. The projected power method has the best recovery rate, especially when the group order m is large. Nevertheless, in terms of computational time, the proposed GPM-based non-convex approach is substantially faster than the projected power method. Therefore, for cyclic synchronization problems, the projected power method is the way to go if recovery performance is the main concern. However, if speed and scalability are of great concern, then our results suggest that the proposed GPM-based non-convex approach would be a better alternative.

⁶In [20], the *misclassification rate* is used to quantify the estimation performance, which is equal to $1 - \text{Recovery Rate}$.

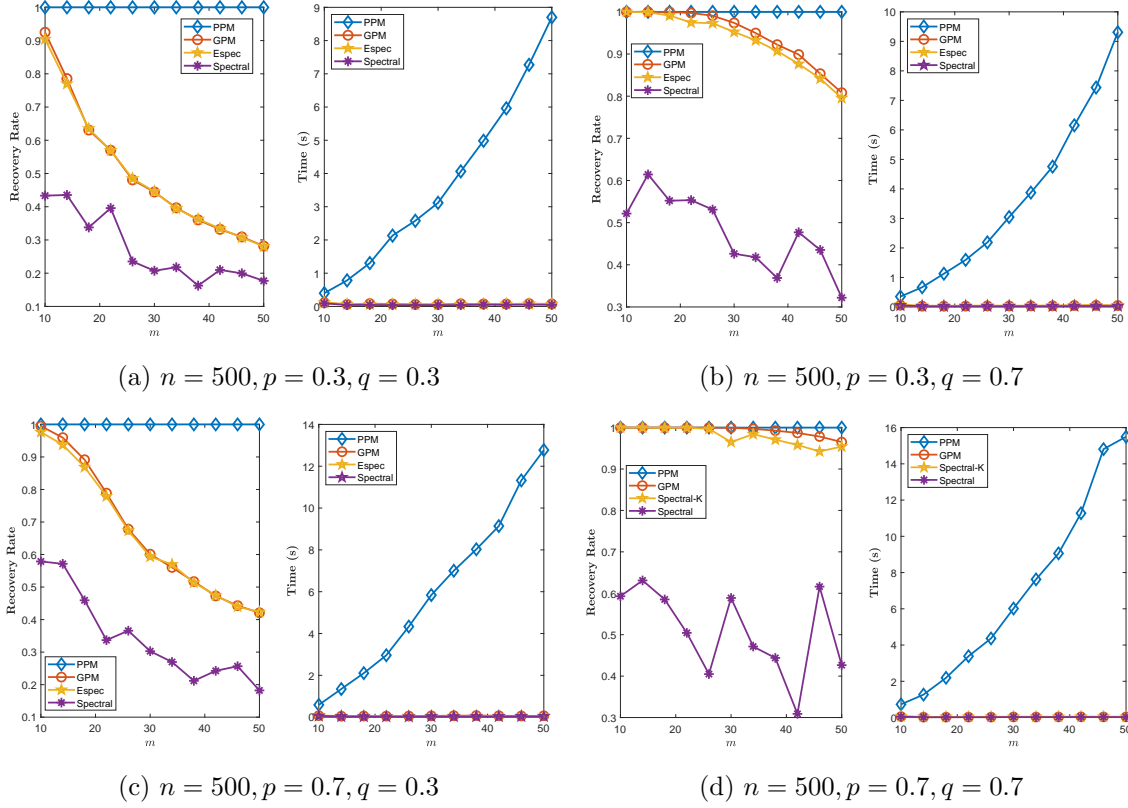


Figure 6: Recovery rate and computational time of the standard spectral estimator (labeled as Spectral), the proposed entropic spectral estimator with $K = 10$ (labeled as Espec), the non-convex approach initialized by the entropic spectral estimator (labeled as GPM), and the projected power method (labeled as PPM) under the setting $p \in \{0.3, 0.7\}$ and $q \in \{0.3, 0.7\}$.

7 Conclusion

In this paper, we proposed a unified approach for tackling a class of synchronization problems over closed subgroups of the orthogonal group. The approach consists of a suitable initialization step and an iterative refinement step based on GPM. We then proved a master theorem, which shows that the estimation error of the iterates produced by GPM decreases geometrically under certain assumptions on the subgroup, measurement graph, noise, and initialization. We verified these assumptions for various practically relevant subgroups under standard random measurement graph and noise models. In the process, we formulated two conditions concerning the geometry of subgroups of the orthogonal group and developed a novel spectral-type estimator called the entropic spectral estimator based

on the notion of metric entropy. These can be of independent interest. Our experiment results showed that the proposed approach outperforms existing approaches in terms of computational speed, scalability, and/or estimation error.

Besides Conjecture 1, there are two other research questions concerning non-convex approaches for solving group synchronization problems that are worth investigating. First, although the theory in Section 5.3 covers only additive noise models, our experiments showed that the proposed approach is also effective under various multiplicative noise models. Thus, it would be interesting to see whether our analysis can be extended to cover more general noise models. Second, an important example of group synchronization problems is $SE(d)$ -sync [57], where $SE(d)$ is the group of d -dimensional Euclidean motions. Such a problem arises in areas such as robotics and computer vision. The group $SE(d)$ is not a subgroup of the orthogonal group. It would be interesting to develop a non-convex approach similar to ours for solving $SE(d)$ -sync.

Acknowledgment

We thank Mihai Cucuringu, Yin-Tat Lee and Michael Kwok-Po Ng for helpful discussions and Amit Singer for kindly sharing with us their codes for the experiments. Man-Chung Yue is supported by the Hong Kong Research Grants Council (RGC) under the Early Career Scheme (ECS) project 25302420. The research of Anthony Man-Cho So is supported in part by the Hong Kong Research Grants Council (RGC) General Research Fund (GRF) Project CUHK 14205421.

References

- [1] E. Abbe, A. S. Bandeira, and G. Hall. Exact recovery in the stochastic block model. *IEEE Transactions on Information Theory*, 62(1):471–487, 2016.
- [2] M. Arie-Nachimson, S. Z. Kovalsky, I. Kemelmacher-Shlizerman, A. Singer, and R. Basri. Global motion estimation from point matches. In *2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission*, pages 81–88. IEEE, 2012.
- [3] F. Arrigoni, B. Rossi, P. Fragneto, and A. Fusiello. Robust synchronization in $SO(3)$ and $SE(3)$ via low-rank and sparse matrix decomposition. *Computer Vision and Image Understanding*, 174:95–113, 2018.
- [4] C. Bajaj, T. Gao, Z. He, Q. Huang, and Z. Liang. SMAC: Simultaneous mapping and clustering using spectral decompositions. In *International Conference on Machine Learning*, pages 324–333, 2018.

- [5] A. S. Bandeira. Random Laplacian matrices and convex relaxations. *Foundations of Computational Mathematics*, 18(2):345–379, 2018.
- [6] A. S. Bandeira, N. Boumal, and A. Singer. Tightness of the maximum likelihood semidefinite relaxation for angular synchronization. *Mathematical Programming*, 163(1-2):145–167, 2017.
- [7] A. S. Bandeira, N. Boumal, and V. Voroninski. On the low-rank approach for semidefinite programs arising in synchronization and community detection. In *Conference on Learning Theory*, pages 361–382, 2016.
- [8] A. S. Bandeira, Y. Chen, R. R. Lederman, and A. Singer. Non-unique games over compact groups and orientation estimation in cryo-EM. *Inverse Problems*, 36(6):064002, 2020.
- [9] A. S. Bandeira, A. Singer, and D. A. Spielman. A Cheeger inequality for the graph connection Laplacian. *SIAM Journal on Matrix Analysis and Applications*, 34(4):1611–1630, 2013.
- [10] A. S. Bandeira and R. van Handel. Sharp nonasymptotic bounds on the norm of random matrices with independent entries. *The Annals of Probability*, 44(4):2479–2506, 2016.
- [11] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. MPS-SIAM Series on Optimization. SIAM, 2001.
- [12] T. Bendory, N. Boumal, C. Ma, Z. Zhao, and A. Singer. Bispectrum inversion with application to multireference alignment. *IEEE Transactions on Signal Processing*, 66(4):1037–1050, 2018.
- [13] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, 2013.
- [14] N. Boumal. Nonconvex phase synchronization. *SIAM Journal on Optimization*, 26(4):2355–2377, 2016.
- [15] N. Boumal, V. Voroninski, and A. Bandeira. The non-convex Burer-Monteiro approach works on smooth semidefinite programs. In *Advances in Neural Information Processing Systems*, pages 2757–2765, 2016.
- [16] N. Boumal, V. Voroninski, and A. S. Bandeira. Deterministic guarantees for Burer-Monteiro factorizations of smooth semidefinite programs. *Communications on Pure and Applied Mathematics*, 73(3):581–608, 2020.

- [17] S. Burer and R. D. Monteiro. A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Mathematical Programming*, 95(2):329–357, 2003.
- [18] S. Burer and R. D. Monteiro. Local minima and convergence in low-rank semidefinite programming. *Mathematical Programming*, 103(3):427–444, 2005.
- [19] R. Burkard, M. Dell’Amico, and S. Martello. *Assignment Problems (Revised Reprint)*, volume 106. SIAM, 2012.
- [20] Y. Chen and E. J. Candès. The projected power method: An efficient algorithm for joint alignment from pairwise differences. *Communications on Pure and Applied Mathematics*, 71(8):1648–1714, 2018.
- [21] S.-S. Cheung, A. M.-C. So, and K. Wang. Linear matrix inequalities with stochastically dependent perturbations and applications to chance-constrained semidefinite optimization. *SIAM Journal on Optimization*, 22(4):1394–1430, 2012.
- [22] A. Chiuso, G. Picci, and S. Soatto. Wide-sense estimation on the special orthogonal group. *Communications in Information and Systems*, 8(3):185–200, 2008.
- [23] M. Cucuringu. Synchronization over Z_2 and community detection in signed multiplex networks with constraints. *Journal of Complex Networks*, 3(3):469–506, 2015.
- [24] M. Cucuringu. Sync-Rank: Robust ranking, constrained ranking and rank aggregation via eigenvector and SDP synchronization. *IEEE Transactions on Network Science and Engineering*, 3(1):58–79, 2016.
- [25] M. Cucuringu, A. Singer, and D. Cowburn. Eigenvector synchronization, graph rigidity and the molecule problem. *Information and Inference: A Journal of the IMA*, 1(1):21–67, 2012.
- [26] P. Diaconis and M. Shahshahani. The subgroup algorithm for generating uniform random variables. *Probability in the Engineering and Informational Sciences*, 1(1):15–32, 1987.
- [27] P. Erdős and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.
- [28] M. Fanuel and H. Tyagi. Denoising modulo samples: k-NN regression and tightness of SDP relaxation. *Information and Inference: A Journal of the IMA*, to appear, 2021.
- [29] A. Genz. Methods for generating random orthogonal matrices. In H. Niederreiter and J. Spanier, editors, *Monte-Carlo and Quasi-Monte Carlo Methods 1998*, pages 199–213. Springer, 2000.

- [30] A. Giridhar and P. R. Kumar. Distributed clock synchronization over wireless networks: Algorithms and analysis. In *Decision and Control, 2006 45th IEEE Conference on*, pages 4915–4920. IEEE, 2006.
- [31] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995.
- [32] J. C. Gower and G. B. Dijkstra. *Procrustes Problems*, volume 30. Oxford University Press, 2004.
- [33] B. Hajek, Y. Wu, and J. Xu. Achieving exact cluster recovery threshold via semidefinite programming. *IEEE Transactions on Information Theory*, 62(5):2788–2797, 2016.
- [34] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.
- [35] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.
- [36] R. A. Horn and R. Mathias. Block-matrix generalizations of Schur’s basic theorems on Hadamard products. *Linear Algebra and Its Applications*, 172:337–346, 1992.
- [37] Q.-X. Huang and L. Guibas. Consistent shape maps via semidefinite programming. In *Computer Graphics Forum*, volume 32, pages 177–186. Wiley Online Library, 2013.
- [38] A. Javanmard, A. Montanari, and F. Ricci-Tersenghi. Phase transitions in semidefinite relaxations. *Proceedings of the National Academy of Sciences*, 113(16):E2218–E2223, 2016.
- [39] M. Journée, Y. Nesterov, P. Richtárik, and R. Sepulchre. Generalized power method for sparse principal component analysis. *Journal of Machine Learning Research*, 11(Feb):517–553, 2010.
- [40] W. Kabsch. A discussion of the solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, 34(5):827–828, 1978.
- [41] A. N. Kolmogorov and V. M. Tikhomirov. ε -entropy and ε -capacity of sets in function spaces. *Uspekhi Matematicheskikh Nauk*, 14(2):3–86, 1959.
- [42] G. Lerman and Y. Shi. Robust group synchronization via cycle-edge message passing. *Foundations of Computational Mathematics*, to appear, 2021.
- [43] S. Ling. Solving orthogonal group synchronization via convex and low-rank optimization: Tightness and landscape analysis. *arXiv preprint arXiv:2006.00902*, 2020.

- [44] S. Ling. Improved performance guarantees for orthogonal group synchronization via generalized power method. Accepted for publication in *SIAM Journal on Optimization*, 2021.
- [45] S. Ling. Near-optimal performance bounds for orthogonal and permutation group synchronization via spectral methods. *Applied and Computational Harmonic Analysis*, 60:20–52, 2022.
- [46] H. Liu, A. M.-C. So, and W. Wu. Quadratic optimization with orthogonality constraint: Explicit Łojasiewicz exponent and linear convergence of retraction-based line-search and stochastic variance-reduced gradient methods. *Mathematical Programming*, 178(1–2):215–262, 2019.
- [47] H. Liu, M.-C. Yue, and A. M.-C. So. On the estimation performance and convergence rate of the generalized power method for phase synchronization. *SIAM Journal on Optimization*, 27(4):2426–2446, 2017.
- [48] H. Liu, M.-C. Yue, A. M.-C. So, and W.-K. Ma. A discrete first-order method for large-scale MIMO detection with provable guarantees. In *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5. IEEE, 2017.
- [49] G. Lugosi, S. Mendelson, and N. Zhivotovskiy. Concentration of the spectral norm of Erdős-Rényi random graphs. *Bernoulli*, 26(3):2253–2274, 2020.
- [50] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang. Semidefinite relaxation of quadratic optimization problems. *IEEE Signal Processing Magazine*, 27(3):20–34, 2010.
- [51] R. Luss and M. Teboulle. Conditional gradient algorithms for rank-one matrix approximations with a sparsity constraint. *SIAM Review*, 55(1):65–98, 2013.
- [52] F. Mezzadri. How to generate random matrices from the classical compact groups. *Notices of the American Mathematical Society*, 54(5):592–604, 2007.
- [53] D. Pachauri, R. Kondor, and V. Singh. Solving the multi-way matching problem by permutation synchronization. In *Advances in Neural Information Processing Systems*, pages 1860–1868, 2013.
- [54] A. Perry, A. S. Wein, A. S. Bandeira, and A. Moitra. Message-passing algorithms for synchronization problems over compact groups. *Communications on Pure and Applied Mathematics*, 71(11):2275–2322, 2018.
- [55] A. Perry, A. S. Wein, A. S. Bandeira, and A. Moitra. Optimality and sub-optimality of PCA I: Spiked random matrix models. *The Annals of Statistics*, 46(5):2416–2451, 2018.

- [56] E. Romanov and M. Gavish. The noise-sensitivity phase transition in spectral group synchronization over compact groups. *Applied and Computational Harmonic Analysis*, 2019.
- [57] D. M. Rosen, L. Carlone, A. S. Bandeira, and J. 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.
- [58] Y. Saad. *Numerical Methods for Large Eigenvalue Problems: Revised Edition*. SIAM, 2011.
- [59] Y. Shen, Q. Huang, N. Srebro, and S. Sanghavi. Normalized spectral map synchronization. In *Advances in Neural Information Processing Systems*, pages 4925–4933, 2016.
- [60] Y. Shkolnisky and A. Singer. Viewing direction estimation in Cryo-EM using synchronization. *SIAM Journal on Imaging Sciences*, 5(3):1088–1110, 2012.
- [61] H. Si, H. Vikalo, and S. Vishwanath. Haplotype assembly: An information theoretic view. In *Information Theory Workshop (ITW), 2014 IEEE*, pages 182–186. IEEE, 2014.
- [62] A. Singer. Angular synchronization by eigenvectors and semidefinite programming. *Applied and Computational Harmonic Analysis*, 30(1):20–36, 2011.
- [63] A. Singer and F. J. Sigworth. Computational methods for single-particle cryo-EM. *Annual Review of Biomedical Data Science*, 3:163–190, 2020.
- [64] A. M.-C. So, J. Zhang, and Y. Ye. On approximating complex quadratic optimization problems via semidefinite programming relaxations. *Mathematical Programming*, 110(1):93–110, 2007.
- [65] G. W. Stewart. The efficient generation of random orthogonal matrices with an application to condition estimators. *SIAM Journal on Numerical Analysis*, 17(3):403–409, 1980.
- [66] S. J. Szarek. Metric entropy of homogeneous spaces. In *Quantum Probability (Gdańsk, 1997)*, volume 43, pages 395–410. Banach Center Publications, Warsaw, 1998.
- [67] L. Trevisan. Max cut and the smallest eigenvalue. *SIAM Journal on Computing*, 41(6):1769–1786, 2012.
- [68] J. A. Tropp. An Introduction to Matrix Concentration Inequalities. *Foundations and Trends® in Machine Learning*, 8(1-2):1–230, 2015.

- [69] H. Tyagi. Error analysis for denoising smooth modulo signals on a graph. *arXiv preprint arXiv:2009.04859*, 2020.
- [70] R. Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*, volume 47 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, 2018.
- [71] L. Wang and A. Singer. Exact and stable recovery of rotations for robust synchronization. *Information and Inference: A Journal of the IMA*, 2(2):145–193, 2013.
- [72] P. Wang, H. Liu, and A. M.-C. So. Linear convergence of a proximal alternating minimization method with extrapolation for ℓ_1 -norm principal component analysis. *arXiv preprint arXiv:2107.07107*, 2021.
- [73] P. Wang, Z. Zhou, and A. M.-C. So. Non-convex exact community recovery in stochastic block model. *Mathematical Programming*, to appear, 2021.
- [74] Z. Wen, D. Goldfarb, and W. Yin. Alternating direction augmented Lagrangian methods for semidefinite programming. *Mathematical Programming Computation*, 2(3-4):203–230, 2010.
- [75] S. X. Wu, M.-C. Yue, A. M.-C. So, and W.-K. Ma. SDR approximation bounds for the robust multicast beamforming problem with interference temperature constraints. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4054–4058. IEEE, 2017.
- [76] Y. Yu, T. Wang, and R. J. Samworth. A useful variant of the Davis–Kahan theorem for statisticians. *Biometrika*, 102(2):315–323, 2015.
- [77] M.-C. Yue, Z. Zhou, and A. M.-C. So. A family of inexact SQA methods for non-smooth convex minimization with provable convergence guarantees based on the Luo–Tseng error bound property. *Mathematical Programming*, 174(1):327–358, 2019.
- [78] M.-C. Yue, Z. Zhou, and A. M.-C. So. On the quadratic convergence of the cubic regularization method under a local error bound condition. *SIAM Journal on Optimization*, 29(1):904–932, 2019.
- [79] Y. Zhong and N. Boumal. Near-optimal bounds for phase synchronization. *SIAM Journal on Optimization*, 28(2):989–1016, 2018.
- [80] Z. Zhou and A. M.-C. So. A unified approach to error bounds for structured convex optimization problems. *Mathematical Programming*, 165(2):689–728, 2017.