A CHEEGER INEQUALITY FOR THE GRAPH CONNECTION LAPLACIAN[*]

AFONSO S. BANDEIRA[†], AMIT SINGER[‡], AND DANIEL A. SPIELMAN[§]

Abstract. The $O(d)$ synchronization problem consists of estimating a set of n unknown orthogonal $d \times d$ matrices $O_1, \ldots, O_n$ from noisy measurements of a subset of the pairwise ratios $O_i O_j^{-1}$. We formulate and prove a Cheeger-type inequality that relates a measure of how well it is possible to solve the $O(d)$ synchronization problem with the spectra of an operator, the graph connection Laplacian. We also show how this inequality provides a worst-case performance guarantee for a spectral method to solve this problem.

Key words. $O(d)$ synchronization, Cheeger inequality, graph connection Laplacian, vector diffusion maps

AMS subject classifications. 05C50, 90C20, 90C35, 49M20

DOI. 10.1137/120875338

1. Introduction. While the graph Laplacian is used to encode similarities between connected vertices, the graph connection Laplacian endows the edges with transformations that describe the nature of the similarity. For example, consider a collection of two-dimensional photos of a three-dimensional object that are taken from many angles as it spins in mid-air. We could form a graph with one vertex for each photo by connecting photos that are similar. By applying simple transformations, such as in-plane rotations, we will discover similarities between photos that are not apparent when we merely treat them as vectors of pixels. In this case, we may also wish to keep track of the transformation under which a pair of photos is similar. Singer and Wu [27] defined the connection Laplacian to encode this additional information. The problem of using this information to assign a viewpoint to each picture is an instance of a synchronization problem on the graph.

More formally, the input to a synchronization problem over a group $\mathcal{G}$ is an undirected graph $G = (V, E)$ and a group element $\rho_{ij} \in \mathcal{G}$ for each edge $(i, j) \in E$, such that $\rho_{ji} = \rho_{ij}^{-1}$. We say that an assignment of group elements to vertices, also called a *group potential*, $g : V \to \mathcal{G}$, *satisfies* an edge ρ_{ij} if $g_i = \rho_{ij}g_j$. The objective in a synchronization problem is to find a group potential that satisfies the edges as much as possible. When there is a g that satisfies all of the edges, it is easy to find: one can arbitrarily fix the value of g at one vertex and then iteratively set g at neighbors of vertices whose values have already been set. When there is no group potential that satisfies all of the edges, we must specify a measure of how well a group potential

satisfies the edges. This is achieved by the notion of frustration which will be defined in section 1.3.

We focus on the group $O(d)$ of $d \times d$ orthogonal matrices (rotations, reflections, and compositions of both on $\mathbb{R}^d$). *Please keep in mind that the "O" has nothing to do with asymptotic notation.* This group is of particular interest in several applications. For example, when $d = 1$, i.e., $\mathcal{G} = O(1) \cong \mathbb{Z}/2\mathbb{Z}$, the solution to the synchronization problem can be used to determine whether a manifold is orientable [26]. The $\mathcal{G} = O(1)$ case is also a generalization of the MAX-CUT problem [13]: the group potential defines a partition of the vertices into two parts, and the group elements on edges specify whether the vertices they connect should be on the same or opposite sides of the partition. In fact, our inequality for *partial synchronization* can be understood as a generalization of Trevisan's inequality relating eigenvalues of the graph Laplacian and the maximum cut [32].

Synchronization over $O(d)$ plays a major role in an algorithm for the sensor network localization problem [10]. The similar problem of synchronization over $SO(d)$, the group of rotations in $\mathbb{R}^d$, also has several applications. The problem over $SO(3)$ can be used for global alignment of three-dimensional scans [33], and the problem over $SO(2)$ plays a major role in new algorithms for the cryo-electron microscopy problem (see [25, 28]). Other applications of $SO(2)$ synchronization may be found in [29, 35, 17].

Singer [29] proposed solving the $SO(2)$ synchronization problem by constructing a matrix, the connection Laplacian, whose eigenvectors associated with the smallest eigenvalues would provide the group potential if it were possible to satisfy all the edges. He then showed that under a model of random noise that prevents such a solution, a good solution can still be obtained by rounding the smallest eigenvectors. A similar algorithm was proposed in [11, 25] for $SO(3)$. These algorithms are analogous to spectral graph partitioning—the use of the smallest eigenvector of the Laplacian matrix to partition a graph [15]. The analysis of the $SO(2)$ algorithm under random noise can be viewed as an analog of McSherry's [21] analysis of spectral partitioning.

Our paper contains three theorems. We begin by considering the simpler problem of finding an assignment of unit vectors to each vertex that agrees with the transformations on the edges. Motivated by Trevisan [32], we first consider a variant of the problem in which we are allowed to find a *partial* assignment in which we assign the zero vector to some of the vertices. In Theorem 2.2, we prove a quadratic relationship between the smallest eigenvalue of the connection Laplacian and the minimum frustration of a partial assignment. To analyze the case in which we must assign a unit vector to every vertex, we observe that partial assignments are really only required when the underlying graph has poor connectivity. In Theorem 2.4, we prove a relation between the minimum frustration of a full assignment and the smallest eigenvalue of the connection Laplacian *and* the second-smallest eigenvalue of the underlying graph Laplacian. We see that when the second-smallest eigenvalue of the underlying graph Laplacian is large (i.e., when it is a good expander), the minimum frustration is well approximated by the smallest eigenvalue of the connection Laplacian. Our main result is Theorem 2.6, in which we relate the minimum frustration of a group potential to the sum of the smallest d eigenvalues of the connection Laplacian *and* the second-smallest eigenvalue of the underlying graph Laplacian.

In the same way that the classical Cheeger's inequality provides a worst-case performance guarantee for spectral clustering, the three Theorems described above provide worst-case performance guarantees for a spectral method to solve each of the described synchronization problems. It is worth noting that, in a different setup,

Trevisan [32] showed a particular case of this inequality, when the group is $O(1) \cong \mathbb{Z}/2\mathbb{Z}$ and all the offsets are -1. In that case the problem is equivalent to the MAX-CUT problem. Therefore, this $O(1)$ inequality gives a performance guarantee for a spectral method to solve MAX-CUT [32].

It is worth mentioning that there are several other adaptations of Cheeger's inequality. Recent progress in multiway partitioning problem gives a Cheeger inequality for the partitioning problem where one wants to partition the graph into more than two subsets (see [12, 19]). There is also a generalization of the Cheeger inequality for simplicial complexes, instead of graphs (see [24, 23, 30]).

The rest of this section includes both mathematical preliminaries that are needed in later sections and the formulation of the problem. Section 2 consists of our main contributions, algorithms to solve different forms of the synchronization problem and Cheeger-type inequalities that provide guarantees for these methods, as well as a brief overview of the proofs. In section 3 we provide the rigorous proofs for the core results described in section 2. We discuss an ℓ_1 version of the synchronization problem in section 4 and show a few tightness results in section 5. We end with some concluding remarks and a few open problems in section 6.

1.1. Notation. Throughout the paper we use the notation $[n]$ to refer to $\{1, \ldots, n\}$. Also, we make use of several matrix and vector notations. Given a matrix A we denote by $\|A\|_F$ its Frobenius norm. If A is symmetric we denote by $\lambda_1(A), \lambda_2(A), \ldots$ its eigenvalues in increasing order. Assuming further that A is positive semidefinite, we define the A-inner product of vectors x and y as $\langle x, y \rangle_A = x^T A y$ (and say that two vectors are A-orthogonal if this inner product is zero). Also, we define the A-norm of x as $\|x\|_A = \sqrt{\langle x, x \rangle_A}$, and we represent the ℓ_2 norm of x by $\|x\|$. Given $x \in \left(\mathbb{R}^d\right)^n$, we denote by x_i the ith $d \times 1$ block of x (that will correspond to the value of x on the vertex i) and, for any $u > 0$, we define x^u as

$$(1.1) \qquad x_i^u = \begin{cases} \frac{x_i}{\|x_i\|} & \text{if} \quad \|x_i\|^2 \geq u, \\ 0 & \text{if} \quad \|x_i\|^2 < u. \end{cases}$$

Furthermore, for $u = 0$ we denote x^0 by $\tilde{x}$, that is,

$$(1.2) \qquad \tilde{x}_i = \begin{cases} \frac{x_i}{\|x_i\|} & \text{if} \quad x_i \neq 0, \\ 0 & \text{if} \quad x_i = 0. \end{cases}$$

Finally, $\mathbb{S}^{d-1}$ denotes the unit sphere in $\mathbb{R}^d$.

1.2. Cheeger's inequality and the graph Laplacian. Before considering the synchronization problem we will briefly present the classical graph Cheeger's inequality in the context of spectral partitioning. The material presented in this section is well known, but it will help motivate the ideas that follow in the next sections.

Let $G = (V, E)$ be an undirected weighted graph with n vertices. In this section we discuss the problem of partitioning the vertices in two similarly sized sets in a way that minimizes the cut: the volume of edges across the subsets (of the partition).

There are several ways to measure the performance of a particular partition of the graph; we will consider the one known as the Cheeger constant. Given a partition (S, S^c) of V, let $h_S := \frac{\text{cut}(S)}{\min\{\text{vol}(S), \text{vol}(S^c)\}}$, where the value of the cut associated with S is $cut(S) = \sum_{i \in S} \sum_{j \in S^c} w_{ij}$, its volume is $vol(S) = \sum_{i \in S} d_i$, and $d_i = \sum_{j \in V} w_{ij}$ is the weighted degree of vertex i. We want to partition the graph so that h_S is minimized, and the minimum value is referred to as the Cheeger number of the graph, denoted

$h_G = \min_{S \subset V} h_S$. Finding the optimal S is known to be NP-hard, as it seems to require searching over an exponential number of possible partitions.

There is another way to measure the performance of a partition (S, S^c) known as the normalized cut (Ncut):

$$\text{Ncut}(S) = \text{cut}(S) \left(\frac{1}{\text{vol}(S)} + \frac{1}{\text{vol}(S^c)} \right).$$

As before, we want to find a subset with as small of an Ncut as possible. Note that the Ncut and the Cheeger constant are closely related:

$$\frac{1}{2} \text{Ncut}(S) \le h_S \le \text{Ncut}(S).$$

Let us introduce a few important definitions. Let W_0 be the weighted adjacency matrix of G and D_0 the degree matrix, a diagonal matrix with elements d_i. If we consider a vector $f \in \mathbb{R}^n$ whose entries take only two possible values, one associated with vertices in S and another in S^c, then the quadratic form $Q_f = \frac{1}{2} \sum_{ij} w_{ij} (f_i - f_j)^2$ is of fundamental importance as a measure of the cut between the sets. The symmetric positive semidefinite matrix that corresponds to this quadratic form, L_0, is known as the graph Laplacian of G. It is defined as $L_0 = D_0 - W_0$ and satisfies $v^T L_0 v = Q_v$ for any $v \in \mathbb{R}^n$. It is also useful to consider the normalized graph Laplacian $\mathcal{L}_0 = D_0^{-1/2} L_0 D_0^{-1/2} = I - D_0^{-1/2} W_0 D_0^{-1/2}$, which is also a symmetric positive semidefinite matrix.

Let us represent a partition (S, S^c) by a cut-function $f_S : V \to \mathbb{R}$ given by

$$f_S(i) = \begin{cases} \sqrt{\frac{\text{vol}(S^c)}{\text{vol}(S)\text{vol}(G)}} & \text{if} \quad i \in S, \\ -\sqrt{\frac{\text{vol}(S^c)}{\text{vol}(S)\text{vol}(G)}} & \text{if} \quad i \in S^c. \end{cases}$$

It is straightforward to show that $Q_{f_S} = f_S^T L_0 f_S = \text{Ncut}(S)$, $f_S^T D_0 f_S = 1$, and $f_S^T D_0 \mathbf{1} = 0$, where $\mathbf{1}$ is the all-ones vector in $\mathbb{R}^n$. This is the motivation for a spectral method to approximate the minimum Ncut problem. If we drop the constraint that f needs to be a cut-function and simply enforce the properties established above, then one would formulate the following relaxed problem:

$$(1.3) \qquad \min_{f:V \to \mathbb{R}, f^T D_0 f = 1, f^T D_0 \mathbf{1} = 0} f^T L_0 f.$$

Since $\mathbf{1}^T L_0 \mathbf{1} = 0$, we know by the Courant–Fisher formula that (1.3) corresponds to an eigenvector problem whose minimum is $\lambda_2(\mathcal{L}_0)$ and whose minimizer can be obtained by the corresponding eigenvector.

Since problem (1.3) is a relaxation of the minimum Ncut problem we automatically have $\frac{1}{2}\lambda_2(\mathcal{L}_0) \le \frac{1}{2}\min_{S \subset V} \text{Ncut} \le h_G$. Remarkably, one can show that the relaxation is not far from the partitioning problem. In fact, one can round the solution of (1.3) so that it corresponds to a partition (S, S^c) of G, whose h_S we can control. This is made precise by the following classical result in spectral graph theory. (Several different proofs for this inequality can be found in [9].)

THEOREM 1.1 (Cheeger inequality [3, 2]). *Let $G = (V, E)$ be a graph and $\mathcal{L}_0$ its normalized graph Laplacian. Then*

$$\frac{1}{2}\lambda_2(\mathcal{L}_0) \le h_G \le \sqrt{2\lambda_2(\mathcal{L}_0)},$$

where h_G is the Cheeger constant of G. Furthermore, the bound is constructive: using the solution of the eigenvector problem one can produce partition (S, S^c) that achieves the upper bound $\sqrt{2\lambda_2(\mathcal{L}_0)}$.

An alternative way to interpret Theorem 1.1 is through random walks on graphs. Note that the matrix $D_0^{-1}W_0$ is the transition probability matrix of a random walk in G, whose transition probabilities are proportional to the edge weights. It is known that the eigenvalues of $\mathcal{L}_0$ encode important information about the random walk. In fact, the second smallest eigenvalue[1] is a good measure of how well the random walk mixes. More specifically, the smaller $\lambda_2(\mathcal{L}_0)$, the slower the convergence to the limiting stationary distribution. It is clear that clusters will constitute obstacles to rapid mixing of the random walk, since the probability mass might be trapped inside such a set for a while. Cheeger's inequality (Theorem 1.1) shows that, in some sense, these sets are the only obstacles to rapid mixing.

1.3. Frustration, vector-valued walks, and the connection Laplacian.
If, in addition to a graph, we are given an orthogonal transformation $\rho_{ij} \in O(d)$ for each edge $(i, j) \in E$, we can consider a random walk that takes the transformations into account. One way of doing this is by defining a random walk that, instead of moving point masses, moves a vector from vertex to vertex and transforms it via the orthogonal transformation associated with the edge. One can similarly define a random walk that moves group elements on vertices. The connection Laplacian was defined by Singer and Wu [27] to measure the convergence of such random walks. The construction requires that $\rho_{ji} = \rho_{ij}^{-1} = \rho_{ij}^T$. Define the symmetric matrix $W_1 \in \mathbb{R}^{dn \times dn}$ so that the (i, j)th $d \times d$ block is given by $(W_1)_{ij} = w_{ij}\rho_{ij}$, where w_{ij} is the weight of the edge (i, j). Also, let $D_1 \in \mathbb{R}^{dn \times dn}$ be the diagonal matrix such that $(D_1)_{ii} = d_i I_{d \times d}$. We assume $d_i > 0$ for every i. The graph connection Laplacian L_1 is defined to be $L_1 = D_1 - W_1$, and the normalized graph connection Laplacian is

$$\mathcal{L}_1 = \mathrm{I} - D_1^{-1/2}W_1 D_1^{-1/2}.$$

If $v : V \to \mathbb{S}^{d-1}$ assigns a unit vector in $\mathbb{R}^d$ to each vertex, we may think of v as a vector in dn dimensions. In this case the quadratic form

$$v^T L_1 v = \sum_{(i,j) \in E} w_{ij} \|v_i - \rho_{ij} v_j\|^2 = \frac{1}{2} \sum_{i,j} w_{ij} \|v_i - \rho_{ij} v_j\|^2$$

is a measure of how well v satisfies the edges. This will be zero if $v_i = \rho_{ij} v_j$ for all edges (i, j). As $w_{ij} = 0$ when $(i, j) \notin E$, we can sum over all pairs of vertices without loss of generality. An assignment satisfying all edges will correspond to a stationary distribution in the vector-valued random walk.

Following our analogy with Cheeger's inequality for the normalized graph Laplacian, we normalize this measure by defining the *frustration* of v as

$$(1.4) \qquad \eta(v) = \frac{v^T L_1 v}{v^T D_1 v} = \frac{1}{2}\frac{\sum_{i,j} w_{ij} \|v_i - \rho_{ij} v_j\|^2}{\sum_i d_i \|v_i\|^2}.$$

We then define the $\mathbb{S}^{d-1}$ frustration constant of G as

$$(1.5) \qquad \eta_G = \min_{v:V \to \mathbb{S}^{d-1}} \eta(v).$$

---

[1] We note that the smallest eigenvalue is always 0.

The smallest eigenvalue of $\mathcal{L}_1$ provides a relaxation of η_G, as

$$\lambda_1(\mathcal{L}_1) = \min_{z \in \mathbb{R}^{dn}} \frac{z^T \mathcal{L}_1 z}{z^T z} = \min_{x \in \mathbb{R}^{dn}} \frac{(D_1^{\frac{1}{2}} x)^T \mathcal{L}_1 (D_1^{\frac{1}{2}} x)}{(D_1^{\frac{1}{2}} x)^T (D_1^{\frac{1}{2}} x)} = \min_{x \in \mathbb{R}^{dn}} \frac{x^T L_1 x}{x^T D_1 x} = \min_{x : V \to \mathbb{R}^d} \eta(x).$$

If there is a group potential $g : V \to O(d)$ that satisfies all the edges (which would again correspond to a stationary distribution for the $O(d)$-valued random walk), then we can obtain d orthogonal vectors on which the quadratic form defined by $\mathcal{L}_1$ is zero. For each $1 \le k \le d$, we obtain one of these vectors by setting $v(i)$ to the kth column of $g(i)$ for all $i \in V$. In particular, this means that the columns of the matrices of the group potential that satisfies all of the edges lie in the nullspace of $\mathcal{L}_1$. Since $g(i) \in O(d)$, these vectors are orthogonal. If G is connected, one can show that these are the only vectors in the nullspace of $\mathcal{L}_1$. This observation is the motivation for the use of a spectral algorithm for synchronization.

We define the frustration of a group potential $g : V \to O(d)$ to be

(1.6) $$\nu(g) = \frac{1}{2d} \frac{1}{\operatorname{vol}(G)} \sum_{i,j} w_{ij} \| g_i - \rho_{ij} g_j \|_F^2.$$

We then define the $O(d)$ frustration constant of G to be

$$\nu_G = \min_{g : V \to O(d)} \nu(g).$$

In Theorem 2.6, we prove that this frustration constant is small if and only if the sum of the first d eigenvalues of $\mathcal{L}_1$ is small as well.

2. Cheeger's type inequalities for the synchronization problem. In this section we present our main results. We present three spectral algorithms to solve three different formulations of synchronization problems and obtain for each a guarantee of performance in the form of a Cheeger's type inequality. We will briefly summarize both the results and the ideas to obtain them, leaving the rigorous proofs to section 3.

We start by considering the $\mathbb{S}^{d-1}$ synchronization problem. This corresponds to finding, for each vertex i of the graph, a vector $v_i \in \mathbb{S}^{d-1}$ in way that for each edge (i,j) the vectors agree with the edges, meaning $v_i = \rho_{ij} v_j$. Since this might not always be possible, we look for a function $v : V \to \mathbb{S}^{d-1}$ for which the frustration $\eta(v)$ is minimum (see (1.4)). Motivated by an algorithm to solve MAX-CUT by Trevisan [32], we first consider a version of the problem for which we allows ourselves to synchronize only a subset of the vertices, corresponding to the partial synchronization in $\mathbb{S}^{d-1}$. We then move on to consider the full synchronization problem in $\mathbb{S}^{d-1}$.

Finally, we will present our main result, an algorithm for $O(d)$ synchronization and a Cheeger-like inequality that equips it with a worst-case guarantee. Recall that the $O(d)$ synchronization corresponds to finding an assignment of an element $g_i \in O(d)$ to each vertex i in a way that minimizes the discrepancy with the pairwise measurements $\rho_{ij} \sim g_i g_j^{-1}$ obtained for each edge. This corresponds to minimizing the $O(d)$ frustration, $\nu(g)$ (see (1.6)).

2.1. Partial synchronization in $\mathbb{S}^{d-1}$. The motivation for considering a spectral relaxation for the synchronization problem in $\mathbb{S}^{d-1}$ is the observation that $\lambda_1(\mathcal{L}_1) = \min_{x : V \to \mathbb{R}^d} \eta(x)$. In order to understand how tight the relaxation is we need to relate $\lambda_1(\mathcal{L}_1)$ with $\eta_G = \min_{x : V \to \mathbb{S}^{d-1}} \eta(x)$.

Consider, however, the following example: a graph consisting of two disjoint components, one whose ρ_{ij} measurements are perfectly compatible and another one on which they are not. Its graph connection Laplacian would have a nonzero vector in its null space, corresponding to synchronizations on the compatible component and zero on the incompatible part (thus $\lambda_1(\mathcal{L}_1) = 0$). On the other hand, the constraint that v has to take values on $\mathbb{S}^{d-1}$ will force it to try to synchronize the incompatible part, thereby bounding η_G away from zero. This example motivates a different formulation of the $\mathbb{S}^{d-1}$ synchronization problem where vertices are allowed not to be labeled (labeled with 0). We thus define the partial $\mathbb{S}^{d-1}$ frustration constant of G as the minimum possible frustration value for such an assignment,

$$(2.1) \qquad \eta_G^* = \min_{v:V \to \mathbb{S}^{d-1} \cup \{0\}} \eta(v).$$

We propose the following algorithm to solve the partial $\mathbb{S}^{d-1}$ synchronization problem.

ALGORITHM 2.1. *Given a graph $G = (V, E)$ and a function $\rho : E \to O(d)$, construct the normalized connection Laplacian $\mathcal{L}_1$ and the degree matrix D_1. Compute z, the eigenvector corresponding to the smallest eigenvalue of $\mathcal{L}_1$. Let $x = D_1^{-\frac{1}{2}} z$. For each vertex index i, let $u_i = \|x_i\|$ and set $v^i : V \to \mathbb{S}^{d-1} \cup \{0\}$ as $v^i = x^{u_i}$, according to (1.1). Output v equal to the v^i that minimizes $\eta(v^i)$.*

Lemma 3.1 guarantees that the solution v given by Algorithm 2.1 satisfies $\eta(v) \leq \sqrt{10\eta(x)}$. Since x was computed so that $\eta(x) = \lambda_1(\mathcal{L}_1)$, Algorithm 2.1 is guaranteed to output a solution v such that

$$\eta(v) \leq \sqrt{10\lambda_1(\mathcal{L}_1)}.$$

Note that $\lambda_1(\mathcal{L}_1) \leq \eta_G^*$, which is the optimum value for the partial $\mathbb{S}^{d-1}$ synchronization problem (see (2.1)). The proof for Lemma 3.1 will appear below. The idea to show that the rounding, from x, to the solution v done by Algorithm 2.1 produces a solution with $\eta(v) \leq \sqrt{10\eta(x)}$ is to use the probabilistic method. One considers a random rounding scheme by rounding x as in Algorithm 2.1 and (1.1) but thresholding at a random value u, drawn from a well-chosen distribution. One then shows that, in expectation, the frustration of the rounded vector is bounded by $\sqrt{10\eta(x)}$. This automatically ensures that there must exist a value u that produces a solution with frustration bounded by $\sqrt{10\eta(x)}$. The rounding described in Algorithm 2.1 runs through all possible such roundings and is thus guaranteed to produce a solution satisfying the bound. An $O(1)$ version of this algorithm and analysis appeared in [32], when ρ is the constant function equal to -1, in the context of the MAX-CUT problem. In fact, if $d = 1$, the factor 10 can be substituted by 8 and the stronger inequality holds $\eta(v) \leq \sqrt{8\lambda_1(\mathcal{L}_1)}$.

The above performance guarantee for Algorithm 2.1 automatically implies the following Cheeger-like inequality.

THEOREM 2.2. *Let $G = (V, E)$ be a weighted graph. Given a function $\rho : E \to O(d)$, let η_G^* be the partial $\mathbb{S}^{d-1}$ frustration constant of G and $\lambda_1(\mathcal{L}_1)$ the smallest eigenvalue of the normalized graph connection Laplacian. Then*

$$(2.2) \qquad \lambda_1(\mathcal{L}_1) \leq \eta_G^* \leq \sqrt{10\lambda_1(\mathcal{L}_1)}.$$

Furthermore, if $d = 1$, the stronger inequality holds, $\eta_G^* \leq \sqrt{8\lambda_1(\mathcal{L}_1)}$.

We note that Trevisan [32], in the context of MAX-CUT, iteratively performs this partial synchronization procedure in the subgraph composed of the vertices left

unlabeled by the previous iteration, in order to label the entire graph. We, however, consider only one iteration.

2.2. Full synchronization in $\mathbb{S}^{d-1}$. In this section we adapt Algorithm 2.1 to solve (full) synchronization in $\mathbb{S}^{d-1}$ and show performance guarantees, under reasonable conditions, by obtaining bounds for η_G, the frustration constant for synchronization in $\mathbb{S}^{d-1}$. The intuition given to justify the relaxation to partial $\mathbb{S}^{d-1}$ synchronization was based on the possibility of poor connectivity of the graph (small spectral gap). In this section we show that poor connectivity, as measured by a small spectral gap in the normalized graph Laplacian, is the only condition under which one can have large discrepancy between the frustration constants and the spectra of the graph connection Laplacian. We will show that, as long as the spectral gap is bounded away from zero, one can in fact control the full frustration constants.

ALGORITHM 2.3. *Given a weighted graph $G = (V, E)$ and a function $\rho : E \to O(d)$, construct the normalized connection Laplacian $\mathcal{L}_1$ and the degree matrix D_1. Compute z, the eigenvector corresponding to the smallest eigenvalue of $\mathcal{L}_1$. Let $x = D_1^{-\frac{1}{2}} z$. Output the solution $v : V \to \mathbb{S}^{d-1} \cup \{0\}$, where each v_i is defined as*

$$v_i = \frac{x_i}{\|x_i\|}.$$

If $x_i = 0$, have v_i be any vector in $\mathbb{S}^{d-1}$.

Similarly to Algorithm 2.1, Lemma 3.6 guarantees that the solution v given by Algorithm 2.3 satisfies $\eta(v) \leq 44 \frac{1}{\lambda_2(\mathcal{L}_0)} \eta(x)$. Again, since x was computed so that $\eta(x) = \lambda_1(\mathcal{L}_1)$, Algorithm 2.1 is guaranteed to output a solution v such that

$$\eta(v) \leq 44 \frac{\lambda_1(\mathcal{L}_1)}{\lambda_2(\mathcal{L}_0)}.$$

Recall that, trivially, $\lambda_1(\mathcal{L}_1) \leq \eta_G$, which is the optimum value for the (full) $\mathbb{S}^{d-1}$ synchronization problem (see (1.5)). The proof for Lemma 3.6 is also deferred until section 3. The idea here is to look at the vector of the local norms of x: $n_x \in \mathbb{R}^n$, where $n_x(i) = \|x_i\|$. It is not hard to show that $\frac{n_x^T L_0 n_x}{n_x^T D_0 n_x} \leq \eta(x)$, which means that if $\eta(x)$ is small, then n_x cannot vary much between two vertices that share an edge. Since $\lambda_2(\mathcal{L}_0)$ is large, one can show that such a vector needs to be close to constant, which means that the norms of x across the vertices are similar. If the norms were all the same, then the rounding $v_i = \frac{x_i}{\|v_i\|}$ would not affect the value of $\eta(\cdot)$, we take this slightly further by showing that if the norms are similar, then we can control how much the rounding affects the penalty function.

The above performance guarantee for Algorithm 2.3 automatically implies another Cheeger-like inequality.

THEOREM 2.4. *Let $G = (V, E)$ be a graph. Given a function $\rho : E \to O(d)$, let η_G be the $\mathbb{S}^{d-1}$ frustration constants of G, $\lambda_1(\mathcal{L}_1)$ the smallest eigenvalue of the normalized graph connection Laplacian, and $\lambda_2(\mathcal{L}_0)$ the second smallest eigenvalue of the normalized graph Laplacian. Then,*

$$\lambda_1(\mathcal{L}_1) \leq \eta_G \leq 44 \frac{\lambda_1(\mathcal{L}_1)}{\lambda_2(\mathcal{L}_0)}.$$

2.3. The $O(d)$ synchronization problem. We now present our main contribution, a spectral algorithm for $O(d)$ synchronization together with a Cheeger-type inequality that provides a worst-case performance guarantee for the algorithm.

Before presenting the algorithm and the results, let us note the differences between this problem and the $\mathbb{S}^{d-1}$ synchronization problem presented above. For the $\mathbb{S}^{d-1}$ case, the main difficulty that we faced in trying to obtain candidate solutions from eigenvectors was the local unit norm constraint. This is due to the fact that the synchronization problem requires its solution to be a function from V to $\mathbb{S}^{d-1}$, corresponding to a vector in $\mathbb{R}^{dn}$ whose vertex subvectors have unit norm, while the eigenvector, in general, does not satisfy such a constraint. Nevertheless, the results in the previous section show that, by simply rounding the eigenvector, one does not lose more than a linear term, given that the graph Laplacian has a spectral gap bounded away from zero.

However, the $O(d)$ synchronization setting is more involved. The reason being that, besides the local normalization constraint, there is also a local orthogonality constraint. (At each vertex, the d vectors have to be orthogonal so that they can be the columns of an orthogonal matrix.) For $\mathbb{S}^{d-1}$ we locally normalized the vectors, by choosing for each vertex the unit vector closest to x_i. For $O(d)$ synchronization we will pick, for each vertex, the orthogonal matrix closest (in the Frobenius norm) to the matrix $\left[x_i^1 \cdots x_i^d\right]$, where x_i^j corresponds to the d-dimensional vector assigned to vertex i by the jth eigenvector. This rounding can be achieved by the polar decomposition. Given a $d \times d$ matrix X, the matrix $U(X)$, solution of $\min_{U \in O(d)} \|U(X) - X\|_F$, is one of the components of the polar decomposition of X (see [16, 20] and references therein). We note that $U(X)$ can be computed efficiently through the SVD decomposition of X. In fact, given the SVD of X, $X = U\Sigma V^T$, the closest orthogonal matrix to X is given by $U(X) = UV^T$ (see [16]). This approach is made precise in the following spectral algorithm for $O(d)$-synchronization.

ALGORITHM 2.5. *Given a weighted graph $G = (V, E)$ and a function $\rho : E \to O(d)$, construct the normalized connection Laplacian $\mathcal{L}_1$ and the degree matrix D_1. Compute $z^1, \ldots, z^d$, the first d eigenvectors corresponding to the d smallest eigenvalues of $\mathcal{L}_1$. Let $x^j = D_1^{-\frac{1}{2}} z^j$ for each $j = 1, \ldots, d$. Output the solution $g : V \to O(d)$, where each g_i is defined as*

$$g_i = U(X_i),$$

where $X_i = \left[x_i^1 \cdots x_i^d\right]$ and $U(X_i)$ is the closest orthogonal matrix of X_i, which can be computed via the SVD of X_i; if $X_i = U_i \Sigma_i V_i^T$, then $U(X_i) = U_i V_i^T$. If X_i is singular,[2] simply set $U(X_i)$ to be I_d.

Similarly to how the performance of the $\mathbb{S}^{d-1}$ synchronization algorithms was obtained, Lemma 3.11 bounds the effect of the rounding step in Algorithm 2.5. Before rounding, the frustration of the solution $[x^1 \cdots x^d]$ is $\frac{1}{d} \sum_{i=1}^{d} \eta(x^i)$. Lemma 3.11 guarantees that the solution g obtained by the rounding in Algorithm 2.5 satisfies $\nu(g) \leq 1026 d^3 \frac{1}{\lambda_2(\mathcal{L}_0)} \sum_{i=1}^{d} \eta(x^i)$. Because of how the vectors $x^1, \ldots, x^d$ were built, $\sum_{i=1}^{d} \eta(x^i) = \sum_{i=1}^{d} \lambda_i(\mathcal{L}_1)$, and this means that the solution g computed by Algorithm 2.5 satisfies

$$\nu(g) \leq 1026 d^3 \frac{1}{\lambda_2(\mathcal{L}_0)} \sum_{i=1}^{d} \lambda_i(\mathcal{L}_1).$$

This performance guarantee automatically implies our main result, a Cheeger inequality for the connection Laplacian.

---

[2]In this case the uniqueness of $U(X_i)$ is not guaranteed, and thus the map is not well defined.

THEOREM 2.6. *Let $\lambda_i(\mathcal{L}_1)$ and $\lambda_i(\mathcal{L}_0)$ denote the ith smallest eigenvalues of the normalized connection Laplacian $\mathcal{L}_1$ and the normalized graph Laplacian $\mathcal{L}_0$, respectively. Let ν_G denote the frustration constant for $O(d)$ synchronization. Then,*

$$\frac{1}{d}\sum_{i=1}^{d}\lambda_i(\mathcal{L}_1) \leq \nu_G \leq 1026d^3\frac{1}{\lambda_2(\mathcal{L}_0)}\sum_{i=1}^{d}\lambda_i(\mathcal{L}_1).$$

Note that, once again, the lower bound is trivially obtained by noting that the eigenvector problem is a relaxation of the original synchronization problem.

Although the rigorous statement and proof of Lemma 3.11 will be presented in section 3, we give a brief intuitive explanation of how the result is obtained.

As discussed above, the performance guarantee for Algorithm 2.3 relies on a proper understanding of the effect of the rounding step. In particular we showed that if $\lambda_2(\mathcal{L}_0)$ is small, then locally normalizing the candidate solution (which corresponds to the rounding step) has an effect over the penalty function that we can control. The case of $O(d)$ synchronization is dealt with similarly. Instead of local normalization, the rounding step for Algorithm 2.5 is based on the polar decomposition. We start by understanding when the polar decomposition is stable (in the sense of changing the penalty function on a given edge) and see that this is the case when the candidate solution $X_i \in \mathbb{R}^{d \times d}$ is not close to being singular. The idea then is to show that only a small portion (which will depend on $\sum_{i=1}^{d}\lambda_i(\mathcal{L}_1)$ and $\lambda_2(\mathcal{L}_0)$) of the graph can have candidate solutions X_i close to singular and use that to show that, overall, we can bound the harmful contribution potentially caused by the rounding procedure on the penalty function.

3. Proof of the main results. In this section we prove the results described above.

3.1. Proofs for synchronization in $\mathbb{S}^{d-1}$. We start with the main lemma regarding partial $\mathbb{S}^{d-1}$ synchronization.

LEMMA 3.1. *Given $x \in \mathbb{R}^{dn}$, there exists $u > 0$ such that*

$$\eta(x^u) \leq \sqrt{10\eta(x)}.$$

Moreover, if $d = 1$, the right-hand side can be replaced by $\sqrt{8\eta(x)}$.

Proof. This lemma immediately follows from Lemma 3.2 as

$$\eta(x^u) = \frac{1}{2}\frac{\sum_{ij} w_{ij}\|x_i^u - \rho_{ij}x_j^u\|^2}{\sum_i d_i\|x_i^u\|^2} \leq \left(\frac{1}{2}\max_{i,j}\|x_i^u - \rho_{ij}x_j^u\|\right)\frac{\sum_{ij} w_{ij}\|x_i^u - \rho_{ij}x_j^u\|}{\sum_i d_i\|x_i^u\|}$$
$$\leq \frac{\sum_{ij} w_{ij}\|x_i^u - \rho_{ij}x_j^u\|}{\sum_i d_i\|x_i^u\|},$$

where the last inequality was obtained by noting that $\|x_i^u - \rho_{ij}x_j^u\| \leq \|x_i^u\| + \|x_j^u\| \leq 2$ ☐

LEMMA 3.2. *Given $x \in \mathbb{R}^{dn}$, there exists $u > 0$ such that*

$$\frac{\sum_{ij} w_{ij}\|x_i^u - \rho_{ij}x_j^u\|}{\sum_i d_i\|x_i^u\|} \leq \sqrt{10\eta(x)}.$$

Moreover, if $d = 1$, the right-hand side can be replaced by $\sqrt{8\eta(x)}$.

Proof. Let us suppose, without loss of generality, that x is normalized so that $\max_i \|x_i\| = 1$. We will use a probabilistic argument. Let us consider the random

variable u drawn uniformly from $[0,1]$ and recall that x^u is defined by $x_i^u = \frac{x_i}{\|x_i\|}$ if $\|x_i\|^2 > u$ or $x_i^u = 0$ if $\|x_i\|^2 \le u$. We will show that $\frac{1}{2}\frac{\mathbb{E}\sum_{ij} w_{ij}\|x_i^u - \rho_{ij}x_j^u\|}{\mathbb{E}\sum_i d_i \|x_i^u\|} \le \sqrt{\frac{5}{2}}\eta(x)$, which implies that at least one of the realizations of u must satisfy the inequality, and proves the lemma.

We start by showing that, for each edge (i,j),

$$(3.1) \qquad \mathbb{E}\|x_i^u - \rho_{ij}x_j^u\| \le \frac{\sqrt{5}}{2}\|x_i - \rho_{ij}x_j\| \left(\|x_i\| + \|x_j\|\right).$$

Without loss of generality we can consider $\rho_{ij} = I$ and $\|x_j\| \le \|x_i\|$ and get

$$\mathbb{E}\|x_i^u - x_j^u\| = \|x_j\|^2 \left\| \frac{x_i}{\|x_i\|} - \frac{x_j}{\|x_j\|} \right\| + \left(\|x_i\|^2 - \|x_j\|^2\right).$$

Thus, it suffices to show

$$\|x_j\|^2 \left\| \frac{x_i}{\|x_i\|} - \frac{x_j}{\|x_j\|} \right\| + \left(\|x_i\|^2 - \|x_j\|^2\right) \le \frac{\sqrt{5}}{2}\|x_i - x_j\| \left(\|x_i\| + \|x_j\|\right),$$

which is a consequence of Proposition A.1 for $y = \frac{x_j}{\|x_j\|}$, $z = \frac{x_i}{\|x_i\|}$, and $\alpha = \frac{\|x_i\|}{\|x_j\|}$. Now, using (3.1), the linearity of expectation, and the Cauchy–Schwarz inequality we have

$$\mathbb{E}\sum_{ij} w_{ij}\|x_i^u - \rho_{ij}x_j^u\| \le \frac{\sqrt{5}}{2}\sum_{ij} w_{ij}\|x_i - \rho_{ij}x_j\|(\|x_i\| + \|x_j\|)$$

$$\le \frac{\sqrt{5}}{2}\sqrt{\sum_{ij} w_{ij}\|x_i - \rho_{ij}x_j\|^2}\sqrt{\sum_{ij} w_{ij}(\|x_i\| + \|x_j\|)^2}.$$

Since $\sum_{ij} w_{ij}\|x_i - \rho_{ij}x_j\|^2 = 2\eta(x)\sum_i d_i\|x_i\|^2$ and

$$\sum_{ij} w_{ij}(\|x_i\| + \|x_j\|)^2 \le 2\sum_{ij} w_{ij}(\|x_i\|^2 + \|x_j\|^2) = 4\sum_i d_i\|x_i\|^2,$$

we have

$$\mathbb{E}\sum_{ij} w_{ij}\|x_i^u - \rho_{ij}x_j^u\| \le \frac{\sqrt{5}}{2}\sqrt{8\eta(x)\sum_i d_i\|x_i\|^2}$$

$$= \frac{\sqrt{5}}{2}\sqrt{8\eta(x)}\mathbb{E}\sum_i d_i\|x_i^u\| = 2\sqrt{\frac{5}{2}\eta(x)}\mathbb{E}\sum_i d_i\|x_i^u\|,$$

which completes the proof. When $d = 1$, the sharper result can be obtained by noting that (3.1) holds even without the $\frac{\sqrt{5}}{2}$ factor. $\qquad \square$

In the context of full $\mathbb{S}^{d-1}$ synchronization, the intuition given to justify a requirement in connectivity of the graph is that it forces the solution of the relaxed problem to have balanced norm across the vertices of the graph. The following lemma makes this thought precise.

LEMMA 3.3. *Given $x \in \mathbb{R}^{dn}$, there exists $\alpha_x \ge 0$ such that $r_x = x - \alpha_x \tilde{x}$ satisfies $\|r_x\|_{D_1}^2 \le \frac{\eta(x)}{\lambda_2(\mathcal{L}_0)}\|x\|_{D_1}^2$.*

Proof. Let us define $n_x \in \mathbb{R}^n$ by $(n_x)_i = \|x_i\|$ and recall $\tilde{x}$ defined in (1.2). We now set $\alpha_x = \operatorname{argmin}_\alpha \|n_x - \alpha \mathbf{1}\|_{D_0}$. A simple calculation reveals that this gives

$$\alpha_x = \frac{\mathbf{1}^T D_0 n_x}{\mathbf{1}^T D_0 \mathbf{1}}.$$

Since n_x is a nonnegative vector, α_x is nonnegative as well. Let us also define $u_x \in \mathbb{R}^n$ so that $(r_x)_i = (u_x)_i \tilde{x}_i$. This implies that $u_x = n_x - (\frac{\mathbf{1}^T D_0 n_x}{\mathbf{1}^T D_0 \mathbf{1}})\mathbf{1}$. Thus,

$$u_x^T L_0 u_x = n_x^T L_0 n_x = \frac{1}{2} \sum_{ij} w_{ij} (\|x_i\| - \|x_j\|)^2 \le \frac{1}{2} \sum_{ij} w_{ij} \|x_i - \rho_{ij} x_j\|^2 = \eta(x) \|x\|_{D_1}^2.$$

Since $u_x^T D_0 \mathbf{1} = 0$, we have $\frac{(u_x)^T L_0 u_x}{\|u_x\|_{D_0}^2} \ge \lambda_2(\mathcal{L}_0)$. This shows that $\|r_x\|_{D_1}^2 = \|u_x\|_{D_0}^2 \le \frac{1}{\lambda_2(\mathcal{L}_0)} \eta(x) \|x\|_{D_1}^2$.   □

This allows one to bound the volume of the subset of vertices where the norm of x is not typical. Let us first define this set.

DEFINITION 3.4. *Given $x \in \mathbb{R}^{dn}$, normalized so that $\|x\|_{D_1}^2 = \operatorname{vol}(G)$, and a positive number δ, we define the ill-balanced vertex subset of the graph G as $\mathcal{I}b_x(\delta) = \{i \in V : |\|x_i\| - 1| \ge \delta\}$.*

The volume of $\mathcal{I}b_x(\delta)$ is controlled by the following lemma.

LEMMA 3.5. *Let $x \in \mathbb{R}^{dn}$ satisfy $\|x\|_{D_1}^2 = \operatorname{vol}(G)$. Then,*

$$\frac{\operatorname{vol}(\mathcal{I}b_x(\delta))}{\operatorname{vol}(G)} \le \frac{4}{\delta^2} \frac{\eta(x)}{\lambda_2(\mathcal{L}_0)}.$$

Proof. Lemma 3.3 guarantees the existence of $\alpha_x \in \mathbb{R}^+$ such that $r_x = x - \alpha_x \tilde{x}$ satisfies $\|r_x\|_{D_1}^2 \le \frac{\eta(x)}{\lambda_2(\mathcal{L}_0)} \|x\|_{D_1}^2$.

Let us start by bounding α_x; by the triangle inequality,

$$(1 - \alpha_x)^2 \operatorname{vol}(G) = (\|x\|_{D_1} - \alpha_x \|\tilde{x}\|_{D_1})^2 \le \|r_x\|_{D_1}^2 \le \frac{\eta(x)}{\lambda_2(\mathcal{L}_0)} \operatorname{vol}(G),$$

which implies $(1 - \alpha_x)^2 \le \frac{\eta(x)}{\lambda_2(\mathcal{L}_0)}$.

If $i \in \mathcal{I}b_x(\delta)$, then $|\|x_i\| - 1| \ge \delta$, which implies $\|(r_x)_i\| = |\|x_i\| - \alpha_x| \ge |\|x_i\| - 1| - |1 - \alpha_x| \ge \delta - \sqrt{\frac{\eta(x)}{\lambda_2(\mathcal{L}_0)}}$. Squaring both sides of the inequality and summing over all $i \in \mathcal{I}b_x(\delta)$ gives

$$(3.2) \quad \frac{\eta(x)}{\lambda_2(\mathcal{L}_0)} \operatorname{vol}(G) \ge \|r_x\|_{D_1}^2 \ge \sum_{i \in \mathcal{I}b_k} d_i \|(r_x)_i\|_2^2 \ge \operatorname{vol}(\mathcal{I}b_k) \left(\delta - \sqrt{\frac{\eta(x)}{\lambda_2(\mathcal{L}_0)}} \right)^2,$$

as long as $\delta > \sqrt{\frac{\eta(x)}{\lambda_2(\mathcal{L}_0)}}$. Let us separate two cases:

If $\frac{\delta}{2} > \sqrt{\frac{\eta(x)}{\lambda_2(\mathcal{L}_0)}}$, then, using (3.2), we have

$$\frac{\operatorname{vol}(\mathcal{I}b_k)}{\operatorname{vol}(G)} \le \frac{\eta(x)}{\lambda_2(\mathcal{L}_0)} \left(\delta - \sqrt{\frac{\eta(x)}{\lambda_2(\mathcal{L}_0)}} \right)^{-2} \le \frac{\eta(x)}{\lambda_2(\mathcal{L}_0)} \left(\delta - \frac{\delta}{2} \right)^{-2} = \frac{4}{\delta^2} \frac{\eta(x)}{\lambda_2(\mathcal{L}_0)}.$$

If, on the other hand, $\frac{\delta}{2} \leq \sqrt{\frac{\eta(x)}{\lambda_2(\mathcal{L}_0)}}$, then, since $\frac{\text{vol}(\mathcal{I}b_k)}{\text{vol}(G)} \leq 1$,

$$\frac{\text{vol}(\mathcal{I}b_k)}{\text{vol}(G)} \leq 1 \leq \frac{\eta(x)}{\lambda_2(\mathcal{L}_0)} \left(\frac{\delta}{2}\right)^{-2} = \frac{4}{\delta^2} \frac{\eta(x)}{\lambda_2(\mathcal{L}_0)}. \qquad \square$$

By placing an upper bound on the number of ill-balanced vertices (Lemma 3.5), we can control how much $\eta(x)$ is affected when we locally normalize x. This is achieved in the following lemma, which contains the central technical result regarding full $\mathbb{S}^{d-1}$ synchronization.

LEMMA 3.6. *For every* $x \in \mathbb{R}^{dn}$, $\eta(\tilde{x}) \leq \frac{44}{\lambda_2(\mathcal{L}_0)}\eta(x)$.

Proof. We want to bound $\eta(\tilde{x}) = \frac{1}{2\,\text{vol}(G)} \sum_{ij} w_{ij}\|\tilde{x}_i - \rho_{ij}\tilde{x}_j\|^2$. Without loss of generality we can assume $\|x\|_{D_1}^2 = \text{vol}(G)$. Let $0 < \gamma < 1$, and then

$$\eta(\tilde{x}) \leq \frac{1}{2\,\text{vol}(G)} \left(\sum_{i \in \mathcal{I}b_x(\gamma)} \sum_j w_{ij}\|\tilde{x}_i - \rho_{ij}\tilde{x}_j\|^2 + \sum_{j \in \mathcal{I}b_x(\gamma)} \sum_i w_{ij}\|\tilde{x}_i - \rho_{ij}\tilde{x}_j\|^2 \right.$$

$$\left. + \sum_{i,j \notin \mathcal{I}b_x(\gamma)} w_{ij}\|\tilde{x}_i - \rho_{ij}\tilde{x}_j\|^2 \right)$$

$$\leq 4\frac{\text{vol}(\mathcal{I}b_x(\gamma))}{\text{vol}(G)} + \frac{1}{2\,\text{vol}(G)} \sum_{i,j \notin \mathcal{I}b_x(\gamma)} w_{ij}\|\tilde{x}_i - \rho_{ij}\tilde{x}_j\|^2.$$

By Lemma 3.5 we have $4\frac{\text{vol}(\mathcal{I}b_x(\gamma))}{\text{vol}(G)} \leq \frac{16}{\gamma^2} \frac{\eta(x)}{\lambda_2(\mathcal{L}_0)}$.

Note that, for any $y, z \in \mathbb{R}^d$, $\|\frac{y}{\|y\|} - \frac{z}{\|z\|}\| \leq \frac{\|y-z\|}{\min\{\|y\|, \|z\|\}}$. By setting $y = x_i$ and $z = \rho_{ij}x_j$, we get $\|\tilde{x}_i - \rho_{ij}\tilde{x}_j\| \leq \frac{\|x_i - \rho_{ij}x_j\|}{\min\{\|x_i\|, \|x_j\|\}}$. This implies that

$$\frac{1}{2\,\text{vol}(G)} \sum_{i,j \notin \mathcal{I}b_x(\gamma)} w_{ij}\|\tilde{x}_i - \rho_{ij}\tilde{x}_j\|^2 \leq \frac{1}{2\,\text{vol}(G)} \sum_{i,j \notin \mathcal{I}b_x(\gamma)} w_{ij} \left(\frac{\|x_i - \rho_{ij}x_j\|}{\min\{\|x_i\|, \|x_j\|\}}\right)^2$$

$$\leq \frac{1}{2\,\text{vol}(G)} \frac{1}{(1-\gamma)^2} \sum_{i,j \notin \mathcal{I}b_x(\gamma)} w_{ij} \|x_i - \rho_{ij}x_j\|^2.$$

This means that

$$\eta(\tilde{x}) \leq \frac{16}{\gamma^2} \frac{\eta(x)}{\lambda_2(\mathcal{L}_0)} + \frac{1}{2\,\text{vol}(G)} \frac{1}{(1-\gamma)^2} \sum_{i,j \notin \mathcal{I}b_x(\gamma)} w_{ij} \|x_i - \rho_{ij}x_j\|^2$$

$$\leq \left(\frac{16}{\gamma^2} \frac{1}{\lambda_2(\mathcal{L}_0)} + \frac{1}{(1-\gamma)^2}\right) \eta(x).$$

Since $\lambda_2(\mathcal{L}_0) \leq 1$ (see, e.g., [8]), it is possible to pick γ (e.g., 0.7) such that $\frac{16}{\gamma^2} \frac{1}{\lambda_2(\mathcal{L}_0)} + \frac{1}{(1-\gamma)^2} \leq \frac{44}{\lambda_2(\mathcal{L}_0)}$. $\square$

3.2. Proofs for synchronization in $O(d)$. As described, the rounding procedure for $O(d)$ synchronization is based on the polar decomposition. We need to understand how much the polar decomposition can potentially affect the penalty on each edge. With this purpose we use the following result of Li [20].

LEMMA 3.7 (Theorem 1 in [20]). *Let* $A, B \in \mathbb{C}^{d \times d}$ *be nonsingular matrices with polar decompositions* $A = U(A)P$ *and* $B = U(B)P'$. *Then* $\|U(A) - U(B)\|_F \leq \frac{2}{\sigma_{\min}(A) + \sigma_{\min}(B)}\|A - B\|_F$, *where* $\sigma_{\min}(A)$ *is the smallest singular value of the matrix* A.

This means that in order to bound possible instabilities of the polar decomposition, one needs to control the size of the smallest singular value of $X_i = \left[x_i^1 \cdots x_i^d \right]$. To achieve this, we will introduce a notion similar to $\mathcal{I}b_x(\delta)$ but designed to take into account local orthogonality instead of local normalization.

DEFINITION 3.8. *Given $x, y \in \mathbb{R}^{dn}$, two D_1-orthogonal vectors, normalized so that $\|x\|_{D_1}^2 = \|y\|_{D_1}^2 = \mathrm{vol}(G)$, and a positive number δ, we define the following ill-balanced vertex subset of the graph G as*

$$\mathcal{I}b_{xy}(\delta) = \{i \in V : |\langle x_i, y_i \rangle| \geq \delta\}.$$

The next lemma shows that, for an edge whose incident vertices are not ill-balanced (see Definitions 3.8 and 3.4), the polar decomposition only slightly affects the penalty function.

LEMMA 3.9. *Let $x^1, \ldots, x^d \in \mathbb{R}^{dn}$ be D_1-orthogonal vectors, normalized so that $\|x^k\|_{D_1}^2 = \mathrm{vol}(G)$. Let us define the "balanced" set $\mathcal{B}$ as the complement of $\bigcup_{k \in [d]} (\mathcal{I}b_{x^k}(\frac{1}{8d}) \cup \bigcup_{m \in [d] \setminus \{k\}} \mathcal{I}b_{x^k x^m}(\frac{1}{2d}))$. For all $i, j \in \mathcal{B}$, we have*

$$\|U(X_i) - \rho_{ij} U(X_j)\|_F \leq \sqrt{2} \|X_i - \rho_{ij} X_j\|_F.$$

Proof. For $i \in \mathcal{B}$, consider the gram matrix $X_i^T X_i$. Its kth diagonal entry satisfies $\|x_i^k\|^2 \geq (1 - \frac{1}{8d})^2 \geq 1 - \frac{1}{4d}$. On the other hand, the nondiagonal entries are, in magnitude, smaller or equal to $\frac{1}{2d}$. By the Gershgorin circle theorem, the smallest eigenvalue of $X_i^T X_i$, which is equal to $\sigma_{\min}(X_i)^2$, satisfies $\sigma_{\min}(X_i)^2 \geq 1 - \frac{1}{4d} - (d-1)\frac{1}{2d}$. Hence, $\sigma_{\min}(X_i) \geq \frac{1}{\sqrt{2}}$. By observing that $U(\rho_{ij} X_j) = \rho_{ij} U(X_j)$ and using Lemma 3.7, we get $\|U(X_i) - \rho_{ij} U(X_j)\|_F \leq \sqrt{2} \|X_i - \rho_{ij} X_j\|_F$. $\square$

The last step is to control the size of the ill-balanced sets.

LEMMA 3.10. *Let $x, y \in \mathbb{R}^{dn}$ be D_1-orthogonal vectors such that $\|x\|_{D_1}^2 = \|y\|_{D_1}^2 = \mathrm{vol}(G)$. Then,*

$$\frac{\mathrm{vol}\left(\mathcal{I}b_{xy}\left(\frac{1}{2d}\right) \setminus \left(\mathcal{I}b_x\left(\frac{1}{8d}\right) \cup \mathcal{I}b_y\left(\frac{1}{8d}\right)\right)\right)}{\mathrm{vol}(G)} \leq 4(8d)^2 \frac{\eta(x) + \eta(y)}{\lambda_2(\mathcal{L}_0)}.$$

Proof. Let us consider the vector $u = \frac{1}{\sqrt{2}}(x + y)$. It satisfies $\|u\|_{D_1}^2 = \mathrm{vol}(G)$, and by the triangle inequality on the norm $\|\cdot\|_{L_1}$, $\eta(u) \leq \eta(x) + \eta(y)$. By Lemma 3.5 we get $\frac{\mathrm{vol}(\mathcal{I}b_u(\frac{1}{8d}))}{\mathrm{vol}(G)} \leq 4(8d)^2 \frac{\eta(x) + \eta(y)}{\lambda_2(\mathcal{L}_0)}$. We conclude the proof by noting that $\mathcal{I}b_{xy}(\frac{1}{2d}) \subset \mathcal{I}b_x(\frac{1}{8d}) \cup \mathcal{I}b_y(\frac{1}{8d}) \cup \mathcal{I}b_u(\frac{1}{8d})$. In fact, if $i \notin \mathcal{I}b_x(\frac{1}{8d}) \cup \mathcal{I}b_y(\frac{1}{8d}) \cup \mathcal{I}b_u(\frac{1}{8d})$, then $|\langle x_i, y_i \rangle| = |\|u_i\|^2 - \frac{\|x_i\|^2 + \|y_i\|^2}{2}| \leq (1 + \frac{1}{8d})^2 - (1 - \frac{1}{8d})^2 = \frac{1}{2d}$. $\square$

At this point we have built all the foundations needed for the proof of the central lemma regarding $O(d)$ synchronization.

LEMMA 3.11. *Given $x^1, \ldots, x^d \in \mathbb{R}^{dn}$ such that $\langle x^k, x^l \rangle_{D_1} = 0$ for all $k \neq l$, consider the potential $g : V \to O(d)$ given as $g_i = U(X_i)$, where $X_i = [x_i^1 \cdots x_i^d]$ and $U(X)$ is the closest (in the Frobenius norm) orthogonal matrix of X. If X_i is singular, $U(X_i)$ is simply set to be I_d. Then,*

$$\nu(g) \leq \left(2d^{-1} + 2^{10} d^3\right) \frac{1}{\lambda_2(\mathcal{L}_0)} \sum_{i=1}^{d} \eta\left(x^i\right).$$

Proof. Let us consider $\mathcal{B}$ as defined in Lemma 3.9, meaning $\mathcal{B}^c = \bigcup_{k \in [d]} (\mathcal{I}b_{x^k}(\frac{1}{8d}) \cup \bigcup_{m \in [d] \setminus \{k\}} \mathcal{I}b_{x^k x^m}(\frac{1}{2d}))$. We want to bound $\nu(g) = \frac{1}{2d \, \mathrm{vol}(G)} \sum_{ij} w_{ij} \|g_i - \rho_{ij} g_j\|_F^2$.

Since $(\mathcal{B} \times \mathcal{B})^c \subset (\mathcal{B}^c \times V) \cup (V \times \mathcal{B}^c)$,

$$\sum_{ij} w_{ij} \|g_i - \rho_{ij} g_j\|_F^2 \leq \sum_{(i,j) \in \mathcal{B} \times \mathcal{B}} w_{ij} \|U(X_i) - \rho_{ij} U(X_j)\|_F^2$$
$$+ 2 \sum_{i \in \mathcal{B}^c} \sum_{j \in V} w_{ij} \|g_i - \rho_{ij} g_j\|_F^2$$
$$\leq 2 \sum_{ij} w_{ij} \|X_i - \rho_{ij} X_j\|_F^2 + 8d \operatorname{vol}(\mathcal{B}^c),$$

where the second inequality was obtained by using Lemma 3.9 and noting that O_i is an orthogonal matrix. To bound $\operatorname{vol}(\mathcal{B}^c)$ we make use of Lemmas 3.5 and 3.10 and get that $\frac{\operatorname{vol}(\mathcal{B}^c)}{\operatorname{vol}(G)}$ is bounded above by

$$\sum_{k \in [d]} \frac{\operatorname{vol}\left(\mathcal{I}b_{x^k}\left(\frac{1}{8d}\right)\right)}{\operatorname{vol}(G)}$$
$$+ \frac{1}{2} \sum_{k \in [d]} \sum_{m \in [d] \setminus k} \frac{\operatorname{vol}\left(\mathcal{I}b_{x^k x^m}\left(\frac{1}{2d}\right) \setminus \left(\mathcal{I}b_x\left(\frac{1}{8d}\right) \cup \mathcal{I}b_y\left(\frac{1}{8d}\right)\right)\right)}{\operatorname{vol}(G)} \leq 2^8 d^3 \frac{\sum_{k \in [d]} \eta(x^k)}{\lambda_2(\mathcal{L}_0)}.$$

Since $\sum_{ij} w_{ij} \|X_i - \rho_{ij} X_j\|_F^2 = 2 \operatorname{vol}(G) \sum_{k=1}^d \eta(x^k)$, we get $\nu(g) \leq (2d^{-1} + 2^{10} d^3) \frac{1}{\lambda_2(\mathcal{L}_0)} \sum_{k=1}^d \eta(x^k)$. □

4. An unsquared version of the frustration constant. We formulated the synchronization problem as minimizing the square of the Frobenius norm of the incompatibilities (in some sense, an ℓ_2 penalty function). This penalty function is particularly nice because it is close in spirit to the Rayleigh quotient formulation of an eigenvector problem and thus more related to what the spectral method will try to minimize, as we indeed showed in the theorems above.

On the other hand, considering the sum of the Frobenius norms of the incompatibilities (in some sense, an ℓ_1 penalty function) will induce sparsity on the edge inconsistencies, meaning that it will favor candidate solutions for which some edges are perfectly correct even if there are some edges with large errors. This type of penalty function is often favorable under some noise models. In fact, the noise model analyzed in [29] consists of a few randomly chosen edges having a measurement that is randomly drawn with respect to the uniform distribution in the space of possible measurements (in our case $O(d)$; in [29] $SO(2)$), and therefore the original rotation potential will perfectly agree with some edges and have a large error on others. This motivates us to look at ℓ_1 versions of frustration constants. Let us define

$$\nu_1(O) = \frac{1}{\sqrt{d} \operatorname{vol}(G)} \sum_{ij} w_{ij} \|O_i - \rho_{ij} O_j\|_F,$$

and the $O(d)$ frustration ℓ_1 constant of G as $\nu_{G,1} = \min_{O:V \to O(d)} \vartheta(O)$. Similarly,

$$\eta_1(v) = \frac{\sum_{ij} w_{ij} \|v_i - \rho_{ij} v_j\|}{\sum_i d_i \|v_i\|}$$

and the ℓ_1 constant of G, as $\eta_{G,1} = \min_{v:V \to \mathbb{S}^{d-1}} \zeta(v)$. We also define a partial version of it, $\eta_{G,1}^* = \min_{v:V \to \mathbb{S}^{d-1} \cup \{0\}} \zeta(v)$.

From our results in the above section it is easy to obtain the following Cheeger type inequalities for these frustration constants.

THEOREM 4.1. *Let $\lambda_i(\mathcal{L}_1)$ and $\lambda_i(\mathcal{L}_0)$ denote the ith smallest eigenvalue of, respectively, the normalized connection Laplacian $\mathcal{L}_1$ and the normalized graph Laplacian $\mathcal{L}_0$. Let $\eta_{G,1}^*$, $\eta_{G,1}$, and $\nu_{G,1}$ denote the ℓ_1 frustration constants defined above. Then,*

$$\lambda_1(\mathcal{L}_1) \leq \eta_{G,1}^* \leq \sqrt{10\lambda_1(\mathcal{L}_1)}, \tag{4.1}$$

$$\lambda_1(\mathcal{L}_1) \leq \eta_{G,1} \leq 2\sqrt{\frac{22}{\lambda_2(\mathcal{L}_0)}\lambda_1(\mathcal{L}_1)}, \tag{4.2}$$

and,

$$\frac{1}{d}\sum_{i=1}^{d}\lambda_i(\mathcal{L}_1) \leq \nu_{G,1} \leq 6d\sqrt{\frac{57d}{\lambda_2(\mathcal{L}_0)}\sum_{i=1}^{d}\lambda_i(\mathcal{L}_1)}. \tag{4.3}$$

Proof. For any $v : V \to S^{d-1} \cup \{0\}$, we have

$$\eta(v) = \frac{1}{2}\frac{\sum_{ij} w_{ij}\|v_i - \rho_{ij}v_j\|^2}{\sum_i d_i\|v_i\|^2} \leq \frac{\left(\sum_{ij} w_{ij}\|v_i - \rho_{ij}v_j\|\right)\frac{\max_{ij}\|v_i - \rho_{ij}v_j\|}{2}}{\left(\sum_i d_i\|v_i\|\right)\min\|v_i\|}$$

$$\leq \frac{\sum_{ij} w_{ij}\|v_i - \rho_{ij}v_j\|}{\sum_i d_i\|v_i\|} = \eta_1(v),$$

which, together with Theorems 2.2 and 2.4, gives the lower bound on both (4.1) and (4.2).

Note that Lemma 3.2 actually guarantees that there exists $v : V \to S^{d-1} \cup \{0\}$ such that $\eta_1(v) \leq \sqrt{10\lambda_i(\mathcal{L}_1)}$, which concludes the proof of (4.1).

Let $v : V \to S^{d-1} \cup \{0\}$ be a solution that satisfies $\eta(v) \leq 44\frac{\lambda_1(\mathcal{L}_1)}{\lambda_2(\mathcal{L}_0)}$, guaranteed to exist by Theorem 2.4. We then have,

$$\eta_1(v) = \frac{\sum_{ij} w_{ij}\|v_i - \rho_{ij}v_j\|}{\sum_i d_i\|v_i\|} = \frac{1}{\mathrm{vol}(G)}\sum_{ij}\left(w_{ij}^{\frac{1}{2}}\|v_i - \rho_{ij}v_j\|\right)w_{ij}^{\frac{1}{2}}$$

$$\leq \frac{1}{\mathrm{vol}(G)}\left(\sum_{ij} w_{ij}\|v_i - \rho_{ij}v_j\|^2\right)^{\frac{1}{2}}\left(\sum_{ij} w_{ij}\right)^{\frac{1}{2}} = \sqrt{\frac{\sum_{ij} w_{ij}\|v_i - \rho_{ij}v_j\|^2}{\sum_i d_i\|v_i\|^2}}$$

$$= \sqrt{2\eta(v)}, \tag{4.4}$$

where the inequality is obtained using Cauchy–Schwarz. This completes the proof of (4.2).

Inequality (4.3) is shown in the same way as (4.2): since $g_i, g_j \in O(d)$, we have $\|g_i - \rho_{ij}g_j\|_F \leq 2\sqrt{d}$, which gives $\nu_1(g) \geq \nu(g)$. On the other hand, using Cauchy–Schwarz in the same way as in (4.4) gives $\nu_1(g) \leq \sqrt{2\nu(g)}$, which implies (4.3) and concludes the proof of the theorem. $\square$

In fact, Wang and Singer [34] recently showed that under the random outlier's noise model (described above), a semidefinite relaxation of the $O(d)$ synchronization problem formulated with the ℓ_1 penalty function was able to recover the ground truth solution with high probability, provided the underlying graph is drawn from the Erdős–Rényi random graph model and the ratio of outliers is below a certain threshold.

5. Tightness of results. Let us consider the ring graph on n vertices $G_n = (V_n, E_n)$ with $V_n = [n]$ and $E = \{(i, (i+1) \bmod n), i \in [n]\}$ with the edge weights all equal to 1 and $\rho : V \to O(d)$ as $\rho_{(n,1)} = -\mathrm{I}$ and $\rho = \mathrm{I}$ for all other edges. Define $x \in \mathbb{R}^{dn}$ by $x_k = [2\frac{k}{n} - 1, 0, \ldots, 0]^T$. It is easy to check that $\eta(x) = \mathcal{O}(n^{-2})$ and that, for any $u > 0$, if $x^u \not\equiv 0$, there will have to be at least one edge that is not compatible with x^u, implying $\eta(x^u) \geq \frac{1}{2n}$. This shows that the $1/2$ exponent in Lemma 3.1 is needed. In fact, by adding a few more edges to the graph G_n, one can also show the tightness of Theorem 2.2: Consider the "rainbow" graph H_n that is constructed by adding to G_n, for each nonnegative integer k smaller than $n/2$, an edge between vertex k and vertex $n - k$ with $\rho_{(k,n-k)} = -\mathrm{I}$. The vector x still satisfies $\eta(x) = \mathcal{O}(n^{-2})$, however, for any nonzero vector $v : V \to \mathbb{S}^{d-1} \cup \{0\}$, it is not hard to show that $\eta(v)$ has to be of order at least n^{-1}, meaning that η_G^* is $\Omega(\sqrt{\lambda_1(\mathcal{L}_1)})$. This also means that, even if considering η_G^*, one could not get a linear bound (as provided by Lemma 3.6) without the control on $\lambda_2(\mathcal{L}_0)$.

Theorem 2.6 provides a nontrivial bound only if $\lambda_2(\mathcal{L}_0)$ is sufficiently large. It is clear that if one wants to bound full frustration constants, a dependency on $\lambda_2(\mathcal{L}_0)$ is needed. It is, nevertheless, nonobvious that this dependency is still needed if we consider partial versions of $O(d)$ frustration constants, ϑ_G^* or ν_G^*. This can, however, be illustrated by a simple example in $O(2)$; consider a disconnected graph G with two sufficiently large complete components, $G^1 = (V^1, E^1)$ and $G^2 = (V^2, E^2)$. For each edge let $\rho_{i,j} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$. It is clear that the vectors x^1 and x^2 defined such that $x_i^1 = [0, 1_{V^1}(i)]^T$ and $x_i^2 = [0, 1_{V^2}(i)]^T$ are orthogonal to each other and lie in the null space of the graph connection Laplacian of G. This implies that $\lambda_2(\mathcal{L}_1) = 0$. On the other hand, it is straightforward to check that ν_G^* is not zero because it is impossible to perfectly synchronize the graph (or any of the components, for that matter).

6. Concluding remarks. Synchronization is a challenging problem. Recent discoveries suggest that spectral relaxations are promising as feasible methods to solve this problem. In fact, in [29], probability guarantees of performance are given for the performance of a spectral method to solve the $SO(2)$ synchronization problem under a certain random noise model. Nevertheless, to the best of our knowledge, Algorithm 2.5 is the first method for $O(d)$ synchronization having a (deterministic) worst-case performance guarantee. As one would expect, the worst-case performance is significantly weaker than the kind of probabilistic guarantees, given a specific noise model, e.g., as the one given in [29]. In fact, the guarantees in [29] and [34] are given in terms of distance between the candidate solution and the ground truth, while the one we provide is given in terms of the compatibility error. Recently, in the context of the phase retrieval problem, the guarantees from this paper were used to obtain guarantees on the distance between the candidate solution and the ground truth (see [1]).

In special applications, one knows, a priori, that every element in the potential has positive determinant. This corresponds to a synchronization problem in $SO(d)$. Although this can be viewed as a special case of the $O(d)$ problem, it is expected that the additional structure can be leveraged to improve the algorithm (and the analysis). In particular, the first $d - 1$ columns of a matrix in $SO(d)$ completely determine the matrix. This suggests that the $SO(d)$ synchronization problem is solvable by just the first $d - 1$ eigenvectors of the connection Laplacian, instead of the d first ones. In fact, the $SO(2)$ synchronization problem is equivalent to the $\mathbb{S}^1$ localization one, and the guarantees for the $\mathbb{S}^1$ localization problem were given solely in terms of the first eigenvalue of the graph connection Laplacian. We leave the improved $SO(d)$ analysis for future work.

The performance guarantee for this algorithm relies on the fact that the vectors obtained are D_1-orthogonal. However, in practice, due to possible errors in the calculations, this condition might be perturbed and the D_1 inner products of these vectors, although small, may no longer be exactly zero. It is easy to adapt the analysis to this setting and show that it is in fact robust to such perturbations.

The results in this paper also suggest an alternative to Algorithm 2.5 which corresponds to, instead of solves, the eigenvector problem, determining the d vectors sequentially, and, at each step, constraining on the vector being locally orthogonal to the previous ones. (This can still be done efficiently; see [14].) After the d vectors are obtained, one can simply locally normalize each one and output that as the synchronization solution candidate. The issue with this method is that its iterative nature[3] makes its analysis more difficult, as it is hard to guarantee that small errors in the first few vectors would not greatly affect the remaining ones. Also, numerical simulations suggest that the performance, in practice, of both methods is roughly the same. Independently of which method is used, the solution, although guaranteed to have a certain performance, is not guaranteed to be a local optimum. Recently, Boumal et al. [5] suggest that an iterative smooth optimization (in manifolds) method, when started in the solution given by the rounding procedure, can be used to locally search for a better solution.

One might argue that, in some applications, the weights on the edges of the graph do not have a clear meaning, the reason being that we may be given a few relative measurements ρ_{ij} and it is unclear how to give weights to such measurements. In such cases, since we have the freedom of choosing the weights of the edges and Theorems 2.4 and 2.6 suggest that our method will work better with a large $\lambda_2(\mathcal{L}_0)$, one could compute the weights of the edges in such a way that $\lambda_2(\mathcal{L}_0)$ is maximized. This problem is solved in [31]. The caveat is that the new weights will affect the way the compatibility error is measured, as well as the eigenvalues of the connection Laplacian. It is thus still unclear if such an approach would improve the method. Another interesting possible outcome of a procedure of this nature is a possible ranking of the edges; large weights would likely tend to be given to edges that are more important to ensure the connectivity of the graph.

The classical Cheeger inequality has an analogous result on smooth manifolds (actually, the first to be shown [7]). One interesting question is whether the theorems in this paper have an analogous smooth version. One difficulty is to understand what would correspond to the frustration constant in the smooth case. Recent work on vector diffusion maps [27] suggests that an analogous result on smooth manifolds would be related to the parallel transport, and its incompatibility and some results in differential geometry [4] suggest that the holonomy could be a geometric property that corresponds to the frustration constant. These suggestions are "coherent" because holonomy can, in some sense, be viewed as the incompatibility of the parallel transport (due to the curvature of the manifold).

In some cases the incompatibilities have some structure; the continuous setting described above may be such an example. Although, in this paper, we make no attempt to understand or take advantage of such structure, we believe this would be an interesting direction for future work, and we direct the reader to papers on which topological tools are used to understand (and leverage) the structure of inconsistencies in synchronization-like problems [6, 18, 22].

---

[3]This is due to the fact that the calculation of one of the vectors greatly depends on the ones already computed.

Appendix A. Some technical steps.

PROPOSITION A.1. *For any y and z unit vectors in $\mathbb{R}^d$, the following holds for any $\alpha \geq 1$:*

$$\|y - z\| + \alpha^2 - 1 \leq \frac{\sqrt{5}}{2}\|y - \alpha z\|(1 + \alpha).$$

Proof. Let $t = \|y - z\|$, which implies that $0 \leq t \leq 2$. Since y and z are unit vectors, it is straightforward to check that $\|y - \alpha z\| = \sqrt{1 + \alpha^2 - 2\alpha(1 - \frac{1}{2}t^2)}$. Thus, it suffices to show

$$(A.1) \qquad t + \alpha^2 - 1 \leq \frac{\sqrt{5}}{2}\sqrt{1 + \alpha^2 - 2\alpha\left(1 - \frac{1}{2}t^2\right)}(1 + \alpha)$$

for all $0 \leq t \leq 2$ and $\alpha \geq 1$. Since both sides of (A.1) are positive, it is enough to show the inequality with both sides squared. Squaring and rearranging yields

$$t^2 + 2t\alpha^2 - 2t + \left(\alpha^2 - 1\right)^2 \leq \frac{5}{4}\left((\alpha^2 - 1)^2 + \alpha t^2(1 + \alpha)^2\right).$$

This is equivalent to the nonnegativity, in the interval $[0, 2]$, of a certain quadratic function of t:

$$\left(\frac{5}{4}\alpha(1 + \alpha)^2 - 1\right)t^2 - \left(2\alpha^2 - 2\right)t + \frac{1}{4}\left(\alpha^2 - 1\right)^2 \geq 0.$$

Since, for $\alpha \geq 1$,

$$\left(2\alpha^2 - 2\right)^2 - 4\left(\frac{5}{4}\alpha(1 + \alpha)^2 - 1\right)\frac{1}{4}\left(\alpha^2 - 1\right)^2 = \left(\alpha^2 - 1\right)^2\left(4 - \frac{5}{4}\alpha(1 + \alpha)^2 + 1\right) \leq 0,$$

the quadratic is always nonnegative and thus nonnegative in $[0, 2]$. $\square$

REFERENCES

[1] B. ALEXEEV, A. S. BANDEIRA, M. FICKUS, AND D. G. MIXON, *Phase retrieval with polarization*, SIAM J. Imaging Sci., to appear.

[2] N. ALON AND V. MILMAN, *Isoperimetric inequalities for graphs, and superconcentrators*, J. Combin. Theory, 38 (1985), pp. 73–88.

[3] N. ALON, *Eigenvalues and expanders*, Combinatorica, 6 (1986), pp. 83–96.

[4] W. BALLMANN, J. BRÜNING, AND G. CARRON, *Eigenvalues and holonomy*, Int. Math. Res. Not., (2003), pp. 657–665.

[5] N. BOUMAL, A. SINGER, P.-A. ABSIL, AND V. D. BLONDEL, *Cramér-Rao bounds for synchronization of rotations*, CoRR, abs/1211.1621 (2012).

[6] O. CANDOGAN, I. MENACHE, A. OZDAGLAR, AND P. A. PARRILO, *Flows and decompositions of games: Harmonic and potential games*, arXiv:1005.2405v2, 2010.

[7] J. CHEEGER, *A lower bound for the smallest eigenvalue of the Laplacian*, in Problems in Analysis, Princeton University Press, Princeton, 1970, pp. 195–199.

[8] F. CHUNG AND L. LU, *Complex Graphs and Networks*, CBMS Regional Conference Series in Mathematics, AMS, Providence, RI, 2006.

[9] F. Chung, *Four proofs for the Cheeger inequality and graph partition algorithms*, Fourth International Congress of Chinese Mathematicians, 2010, pp. 331–349.

[10] M. Cucuringu, Y. Lipman, and A. Singer, *Sensor network localization by eigenvector synchronization over the Euclidean group*, ACM Trans. Sensor Networks, 8 (2012), pp. 19:1–19:42.

[11] M. Cucuringu, A. Singer, and D. Cowburn, *Eigenvector synchronization, graph rigidity and the molecule problem*, Information and Inference: A Journal of the IMA, 1 (2012), pp. 21–67.

[12] S. O. Gharan and L. Trevisan, *A higher–order Cheeger's inequality*, arXiv:1107.2686, 2011.

[13] M. X. Goemans and D. P. Williamson, *Improved apprximation algorithms for maximum cut and satisfiability problems using semidefine programming*, J. Assoc. Comput. Mach., 42 (1995), pp. 1115–1145.

[14] G. H. Golub, *Some modified matrix eigenvalue problems*, SIAM Rev., 15 (1973), pp. 318–334.

[15] Hagen and Kahng, *New spectral methods for ratio cut partitioning and clustering*, IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems, 11 (1992), pp. 1074–1085.

[16] N. J. Higham, *Computing the polar decomposition with applications*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 1160–1174.

[17] S. D. Howard, D. Cochran, W. Moran, and F. R. Cohen, *Estimation and registration on graphs*, arXiv:1010.2983, 2010.

[18] X. Jiang, L.-H. Lim, Y. Yao, and Y. Ye, *Statistical ranking and combinatorial hodge theory*, arXiv:0811.1067v2, 2009.

[19] J. R. Lee, S. O. Gharan, and L. Trevisan, *Multi-way spectral partitioning and higher–order Cheeger inequalities*, arXiv:1111.1055, 2011.

[20] R.-C. Li, *New perturbation bounds for the unitary polar factor*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 327–332.

[21] F. McSherry, *Spectral partitioning of random graphs*, in FOCS '01: Proceedings of the 42nd IEEE symposium on Foundations of Computer Science, 2001, pp. 529–537.

[22] P. Molavi and A. Jadbabaie, *A topological view of estimation from noisy relative measurements*, in Proceedings of the American Control Conference (ACC), 2011, pp. 3615–3620.

[23] O. Parzanchevski, R. Rosenthal, and R. J. Tessler, *Isoperimetric inequalities in simplicial complexes*, arXiv:1207.0638v2, 2012.

[24] O. Parzanchevski and R. Rosenthal, *Simplicial complexes: Spectrum, homology and random walks*, arXiv:1211.6775v1, 2012.

[25] A. Singer and Y. Shkolnisky, *Three-dimensional structure determination from common lines in cryo-em by eigenvectors and semidefinite programming*, SIAM J. Imaging Sci., 4 (2011), pp. 543–572.

[26] A. Singer and H.-T. Wu, *Orientability and diffusion maps*, Appl. Comput. Harmon. Anal., 31 (2011), pp. 44–58.

[27] A. Singer and H.-T. Wu, *Vector diffusion maps and the connection Laplacian*, Comm. Pure Appl. Math., 65 (2012), pp. 1067–1144.

[28] A. Singer, Z. Zhao, Y. Shkolnisky, and R. Hadani, *Viewing angle classification of cryo-electron microscopy images using eigenvectors*, SIAM J. Imaging Sci., 4 (2011), pp. 723–759.

[29] A. Singer, *Angular synchronization by eigenvectors and semidefinite programming*, Appl. Comput. Harmon. Anal., 30 (2011), pp. 20–36.

[30] J. Steenbergen, C. Klivans, and S. Mukherjee, *A Cheeger-type inequality on simplicial complexes*, arXiv:1209.5091v3, 2012.

[31] J. Sun, S. Boyd, L. Xiao, and P. Diaconis, *The fastest mixing Markov process on a graph and a connection to a maximum variance unfolding problem*, SIAM Rev., 48 (2006), pp. 681–699.

[32] L. Trevisan, *Max cut and the smallest eigenvalue*, in Proceedings of the 41st annual ACM symposium on Theory of computing, STOC '09, New York, ACM, 2009, pp. 263–272.

[33] T. Tzeneva, *Global Alignment of Multiple 3-D Scans Using Eigevector Synchronization*, Senior Thesis, Princeton University, Princeton., NJ, 2011.

[34] L. Wang and A. Singer, *Exact and stable recovery of rotations for robust synchronization*, arXiv:1211.2441v1, 2012.

[35] S. X. Yu, *Angular embedding: A robust quadratic criterion*, IEEE Trans. Pattern Anal. Mach. Intell., 34 (2012), pp. 158–173.