Expected performance bounds for estimation on graphs from random relative measurements

Nicolas Boumal* Xiuyuan Cheng†

Compiled on: July 25, 2013.

Abstract

We consider estimation problems on graphs known as synchronization. Each node in the graph has a state to be estimated. An edge between two nodes is present if a (noisy) measurement of the relative state of those two nodes is available. It is known from other work that, under some assumptions including that noise be i.i.d., the variance of unbiased estimators for the states is lower-bounded by some constant times the trace of the pseudoinverse of the Laplacian of the measurement graph. In this work, we show that for Erdős-Rényi random graphs the latter bound (which is now a random variable) concentrates around its expectation, and we give a formula for the expectation. For synchronization, the results imply that acquiring measurements uniformly at random is a good strategy to drive the lower-bound to zero. This holds even for vanishing edge density, provided the graph remains sufficiently connected.

Keywords: Estimation on graphs, synchronization, Cramér-Rao bounds, graph Laplacian, random matrices.

1 Introduction

Synchronization is the general problem of estimating n elements $g_1, \ldots, g_n$ in a group $\mathcal{G}$ based on certain relative measurements $h_{ij} \in \mathcal{G}$ which bear information about $g_i \cdot g_j^{-1}$. The set of measurements defines an undirected graph G on n nodes, with a link between nodes i and j if a measurement h_{ij} is available. Synchronization is an important class of estimation problems on graphs and it occurs frequently in applications [14, 8, 17, 15, 1, 11, 16].

Synchronization proves useful for both discrete groups (e.g., $\mathbb{Z}_2 = \{+1, -1\}$ [7] and the group of permutations) and Lie groups (e.g., the group of translations [1, 7] and the group of rotations [3]). For example, consider synchronization of translations (the group is $\mathbb{R}^d$ and the group operation $\cdot$ is the sum). Let $x_1, \ldots, x_n \in \mathbb{R}^d$ be the n state vectors to estimate. In a sensor network localization context, they could represent the positions of n agents in some coordinate system. For each edge (i, j) in the (fixed) graph G, a measurement h_{ij} bears information about the relative position $x_i - x_j$.

For some Lie groups, namely the group of translations and the group of rotations, Cramér-Rao bounds (CRB) were established that put a lower-bound on the variance of any unbiased estimator for these estimation problems [2, 4, 11]. In the case of isotropic, i.i.d. noise on the measurements h_{ij}, these bounds are proportional to $\mathrm{trace}(\mathbf{L}_{(n)}^{\dagger})$, the

*Department of mathematical engineering, ICTEAM Institute, Université catholique de Louvain, Belgium.

†Program in Applied and Computational Mathematics, Princeton University, New Jersey, USA.

trace of the (Moore-Penrose) pseudoinverse[1] of the Laplacian $\mathbf{L}_{(n)} = \mathbf{D}_{(n)} - \mathbf{A}_{(n)}$ of the measurement graph G, where $\mathbf{D}_{(n)}$ is the diagonal degree matrix and $\mathbf{A}_{(n)}$ is the adjacency matrix of G. There is ground to believe that a similar statement applies to more Lie groups.

Continuing the sensor network localization example, assume the measurements are given by $h_{ij} = -h_{ji} = x_i - x_j + n_{ij}$, with the $n_{ij} \sim \mathcal{N}(0, \Sigma)$ i.i.d. normal random variables. Let $\hat{x}_1, \ldots, \hat{x}_n$ be any unbiased estimator of the state vectors. Assuming the x_i's and the $\hat{x}_i$'s are centered (since the estimation can only be resolved up to a global translation), then the CRB for this synchronization problem lower-bounds the variance as [2]:

$$\mathbb{E}\Big\{ \sum_{i=1}^{n} \|x_i - \hat{x}_i\|^2 \Big\} \geq \operatorname{trace}(\Sigma) \operatorname{trace}(\mathbf{L}_{(n)}^{\dagger}).$$

The expectation is taken w.r.t. the noise n_{ij}. The maximum likelihood estimator achieves this bound [1]. In such applications, for a growing number of nodes n it is desirable to drive the lower-bound on the average variance per node $\operatorname{trace}(\mathbf{L}_{(n)}^{\dagger})/n$ to zero, if possible.

It is well-known that the spectrum of $\mathbf{L}_{(n)}$ captures the connectivity properties of the graph, and it is hence not surprising to see it appear in bounds for estimation problems on graphs. Expander graphs for example, which are both sparse and well connected, have eigenvalues bounded away from zero (except for one) [10]. In turn, this translates in small eigenvalues for $\mathbf{L}_{(n)}^{\dagger}$ and a small value of $\operatorname{trace}(\mathbf{L}_{(n)}^{\dagger})$, meaning that CRB's allow for accurate synchronization on expanders.

A natural endeavor then is to assess the synchronizability of certain classes of random graphs by studying the random variable $\operatorname{trace}(\mathbf{L}_{(n)}^{\dagger})$. The purpose may be, for example, to study existing graphs which form randomly based on local interactions between nodes. Alternatively, the goal may be to devise decentralized experiment plans which can lead to accurate estimation, where each node (or sensor) i chooses which measurements h_{ij} to acquire, based on local rules (such as a budget constraint) and randomness.

2 Contribution

We address the simplest case: Erdős-Rényi random graphs with n nodes, where each edge is present with probability p_n, independently from the other edges. Consider a sequence of increasingly large Erdős-Rényi graphs. For each $n \geq 2$, let $\{A_{ij}^{(n)}\}$, $1 \leq i < j \leq n$, be a collection of independent Bernoulli random variables with

$$A_{ij}^{(n)} = \begin{cases} 1 & \text{with probability } p_n, \\ 0 & \text{with probability } 1 - p_n, \end{cases}$$

for some edge probability $0 < p_n < 1$ which may depend on n. Because we consider simple, undirected graphs, define $A_{ji}^{(n)} = A_{ij}^{(n)}$ and $A_{ii}^{(n)} = 0$. We also require the $A_{ij}^{(n)}$ for various n to be independent. For $i \neq j$,

$$\begin{aligned} &\mathbb{E}\{A_{ij}^{(n)}\} = p_n > 0, \\ &\mathbb{E}\{(A_{ij}^{(n)} - p_n)^2\} = \sigma_n^2 \triangleq p_n(1 - p_n). \end{aligned}$$

[1]For a symmetric matrix A which admits an eigenvalue decomposition $A = UDU^{\top}$, $U^{\top}U = I$ and $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, the pseudoinverse of A is given by $A^{\dagger} = UD^{\dagger}U^{\top}$, with $D^{\dagger} = \operatorname{diag}(\lambda_1^{\dagger}, \ldots, \lambda_n^{\dagger})$. The pseudoinverse of a scalar λ_i is $\lambda_i^{\dagger} = 1/\lambda_i$ if $\lambda_i \neq 0$ and $\lambda_i^{\dagger} = 0$ if $\lambda_i = 0$.

Define the $n \times n$ adjacency matrix $\mathbf{A}_{(n)}$ as

$$\mathbf{A}_{(n)} = \begin{pmatrix} 0 & A_{12}^{(n)} & \cdots & A_{1n}^{(n)} \\ A_{21}^{(n)} & 0 & \cdots & A_{2n}^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1}^{(n)} & A_{n2}^{(n)} & \cdots & 0 \end{pmatrix}.$$

Let $\mathbb{1}_n$ denote the all-ones vector of length n. The diagonal degree matrix is given by $\mathbf{D}_{(n)} = \operatorname{diag}(\mathbf{A}_{(n)}\mathbb{1}_n)$. Then, the $n \times n$ Laplacian matrix is defined as $\mathbf{L}_{(n)} = \mathbf{D}_{(n)} - \mathbf{A}_{(n)}$:

$$\mathbf{L}_{(n)} = \begin{pmatrix} \sum_{i\neq 1} A_{1i}^{(n)} & -A_{12}^{(n)} & \cdots & -A_{1n}^{(n)} \\ -A_{21}^{(n)} & \sum_{i\neq 2} A_{2i}^{(n)} & \cdots & -A_{2n}^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ -A_{n1}^{(n)} & -A_{n2}^{(n)} & \cdots & \sum_{i\neq n} A_{ni}^{(n)} \end{pmatrix}.$$

The Laplacian is a symmetric, positive semidefinite matrix since for all $u \in \mathbb{R}^n$,

$$u^\top \mathbf{L}_{(n)} u = \sum_{i<j} A_{ij}^{(n)} (u_i - u_j)^2 \geq 0.$$

Its eigenvalues $\lambda_i^{(n)}$, $1 \leq i \leq n$, are thus real and nonnegative, with the smallest eigenvalue necessarily equal to zero, since $\mathbf{L}_{(n)}\mathbb{1}_n = 0$:

$$0 = \lambda_1^{(n)} \leq \lambda_2^{(n)} \leq \cdots \leq \lambda_n^{(n)}.$$

The scalar random variable

$$X_n = p_n \operatorname{trace}(\mathbf{L}_{(n)}^\dagger) \tag{1}$$

is the variable of interest in this paper. We may gain a quick insight into this variable by considering the expected value of $\mathbf{L}_{(n)}$:

$$\mathbf{L}_{(n),0} \triangleq \mathbb{E}\{\mathbf{L}_{(n)}\} = p_n(n\mathbf{I}_{(n)} - \mathbb{1}_{n\times n}),$$

where $\mathbf{I}_{(n)}$ denotes the $n \times n$ identity matrix and $\mathbb{1}_{n\times n}$ is the all-ones matrix of size $n \times n$. The eigenvalues of $\mathbf{L}_{(n),0}$ are $0, np_n, np_n, \ldots, np_n$. Thus, $p_n \operatorname{trace}(\mathbf{L}_{(n),0}^\dagger) = (n-1)/n$, which goes to 1 as $n \to \infty$. Indeed, we will see that as n goes to infinity, $\mathbf{L}_{(n)}$ "behaves more and more like $\mathbf{L}_{(n),0}$", so that $X_n \to 1$. We aim at a more precise statement that will already be useful for moderate n.

In the limit, the positive eigenvalues of $\mathbf{L}_{(n)}$ tend to be distributed symmetrically around np_n [9]. Thus, the positive eigenvalues of $\mathbf{L}_{(n)}^\dagger$ will be distributed asymmetrically around $1/np_n$ and the expected value of X_n will be biased away from 1. In this communication, we establish a formula for the expected value of X_n for large n and bound its fluctuation around its mean, confirming concentration.

To this end, we look at the spectrum of the centered Laplacian. Define these centered random variables:

$$X_{ij}^{(n)} \triangleq A_{ij}^{(n)} - p_n.$$

Then, $\mathbb{E}\{X_{ij}^{(n)}\} = 0$ and $\mathbb{E}\{(X_{ij}^{(n)})^2\} = \sigma_n^2$. Define the $n \times n$ matrix $\mathbf{L}_{(n),1}$ such that $\mathbf{L}_{(n)} = \mathbf{L}_{(n),0} + \mathbf{L}_{(n),1}$:

$$\mathbf{L}_{(n),1} = \begin{pmatrix} \sum_{i\neq 1} X_{1i}^{(n)} & -X_{12}^{(n)} & \cdots & -X_{1n}^{(n)} \\ -X_{21}^{(n)} & \sum_{i\neq 2} X_{2i}^{(n)} & \cdots & -X_{2n}^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ -X_{n1}^{(n)} & -X_{n2}^{(n)} & \cdots & \sum_{i\neq n} X_{ni}^{(n)} \end{pmatrix}.$$

We make two working assumptions on the edge density p_n, which is the only parameter in the present setting. The first assumption ensures almost sure connectivity of the graph. When the graph is connected, λ_2 is known to be positive. Adding the second assumption further ensures that the eigenvalues of $\mathbf{L}_{(n),1}/(np_n)$ concentrate around 0, which will limit the fluctuation of X_n around its mean.

Assumption 1 (connectivity). *$np_n \gg \log n$, that is, $\lim_{n\to\infty} \frac{\log n}{np_n} = 0$.*

Assumption 2 (concentration). *For all $\epsilon_0 > 0$, almost surely as $n \to \infty$,*

$$\|\mathbf{L}_{(n),1}\|_{\text{op}} \leq (2 + \epsilon_0)\sqrt{np_n \log n}.$$

For a symmetric matrix $M \in \mathbb{R}^{n\times n}$, $\|M\|_{\text{op}}$ denotes the operator norm of M:

$$\|M\|_{\text{op}} \triangleq \max_{u\in\mathbb{R}^n, u^\top u=1} \sqrt{u^\top M^\top M u} = \max_{1\leq i\leq n} |\lambda_i(M)|.$$

Under both assumptions, the following two theorems about the expectation and the fluctuation of X_n are established:

Theorem 1 (Expectation). *Under assumptions 1 and 2, almost surely as $n \to \infty$:*

$$\mathbb{E}X_n = 1 + \left(2\frac{1-p_n}{p_n} - 1\right)\frac{1}{n} + \mathcal{O}\left(\frac{\log^2 n}{(np_n)^2}\right), \tag{2}$$

with $X_n = p_n \text{trace}(\mathbf{L}_{(n)}^\dagger)$ (1).

Theorem 2 (Fluctuation). *Under assumptions 1 and 2, as $n \to \infty$ and for all $0 < \epsilon \leq 1/2$, it holds with probability at least $1 - 2\epsilon$ that:*

$$|X_n - \mathbb{E}X_n| \leq 1.01^2 \frac{\sqrt{\log(1/\epsilon)}}{np_n}.$$

The constant 1.01 can be made arbitrarily close to 1 for large enough n.

We show in a third theorem that a stronger version of Assumption 1 also implies Assumption 2:

Assumption 3 (stronger Assumption 1). *$np_n \gg \log^6 n$, that is, $\lim_{n\to\infty} \frac{\log^6 n}{np_n} = 0$.*

Theorem 3. *Assumption 3 implies assumptions 1 and 2.*

Theorem 3 rests essentially upon results from Chung et al. [6]. The latter paper treats a more general class of random graphs with given expected degrees. For the special case of Erdős-Rényi graphs, the condition could be replaced by the slightly weaker $np_n \gg \log^4 n$ (details omitted).

We spell out two simple examples. In both cases, Assumption 3 is readily checked so that Theorem 3 authorizes the use of theorems 1 and 2.

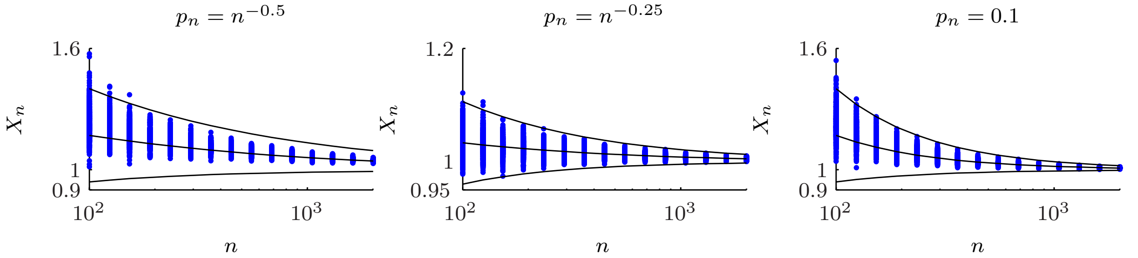

Figure 1: For a number of nodes n ranging from 100 to 2000 on a log-scale and for three different scenarios of edge densities p_n (decaying as $1/\sqrt{n}$, decaying as $1/\sqrt[4]{n}$ and constant), we computed 500 realizations of $X_n = p_n \mathrm{trace}(\mathbf{L}_{(n)}^\dagger)$ (blue dots). The middle black curve is the established formula for $\mathbb{E}X_n$ (3). The black curves above and below are separated from the middle one by the bound on the fluctuation (4), such that as $n \to \infty$, almost surely, all blue dots are within the two black lines with probability at least 99%. These bounds appear to make sense already for reasonable values of n.

Example 1 (p_n bounded away from 0 and 1)**.** *Assuming the edge presence probability p_n remains bounded away from zero and one, i.e., there exist constants c_ℓ, c_u such that $0 < c_\ell \leq p_n \leq c_u < 1$ for all n, then Assumption 3 holds and, almost surely as $n \to \infty$:*

$$\mathbb{E}X_n = 1 + \left(2\frac{1-p_n}{p_n} - 1\right)\frac{1}{n} + \mathcal{O}\left(\frac{\log^2 n}{n^2}\right)$$

Furthermore, for all $0 < \epsilon \leq 1/2$, as $n \to \infty$, with probability at least $1 - 2\epsilon$,

$$|X_n - \mathbb{E}X_n| \leq 1.01^2 \frac{\sqrt{\log(1/\epsilon)}}{np_n} = \mathcal{O}\left(\frac{1}{n}\right).$$

Example 2 (vanishing p_n)**.** *Assuming the edge presence probability p_n decays as $p_n = \gamma n^{-1+\alpha}$, for some constant $\gamma > 0$ and $0 < \alpha \leq 1$ (and $\gamma < 1$ if $\alpha = 1$), then Assumption 3 holds and, almost surely as $n \to \infty$:*

$$\mathbb{E}X_n = 1 + \frac{2}{\gamma n^\alpha} - \frac{3}{n} + \mathcal{O}\left(\frac{\log^2 n}{n^{2\alpha}}\right). \tag{3}$$

Furthermore, for all $0 < \epsilon \leq 1/2$, as $n \to \infty$, with probability at least $1 - 2\epsilon$,

$$|X_n - \mathbb{E}X_n| \leq 1.01^2 \frac{\sqrt{\log(1/\epsilon)}}{\gamma n^\alpha} = \mathcal{O}\left(\frac{1}{n^\alpha}\right). \tag{4}$$

All the aforementioned results have an "as n goes to infinity" proviso, which in theory limits their usability. Through numerical tests, Figure 1 illustrates the fact that already for small values of n (in the hundreds), the predictions hold quite well.

With respect to synchronization, the take home message of this second example is the following. If we could afford a complete graph of measurements, the Cramér-Rao lower-bound on the average variance per node would decay as $\mathrm{trace}(\mathbf{L}_{(n)}^\dagger)/n = \mathcal{O}(1/n)$. If on the other hand we can only afford to obtain a vanishing fraction of, say, $p_n \sim 1/\sqrt{n}$ of the $\mathcal{O}(n^2)$ measurements following an Erdős-Rényi graph—which is easy to set up in a decentralized manner—then we would have that $\mathbb{E}\{\mathrm{trace}(\mathbf{L}_{(n)}^\dagger)\}/n = \mathbb{E}X_n/np_n = \mathcal{O}(1/\sqrt{n})$ and that this quantity concentrates around its mean as $\mathcal{O}(1/n)$. Hence, although we only acquire a vanishing fraction of the measurements, acquiring them uniformly at random guarantees sufficient connectivity in the graph that the average CRB still goes to zero as $n \to \infty$, with a fairly predictable price to pay in accuracy.

Of course, in practical applications it may be impossible or overly expensive to acquire measurements between arbitrary pairs of nodes. Physical sensors may indeed be constrained by range limitations, so that sensors (embedded in $\mathbb{R}^3$) can only sense close-by sensors. This calls for the study of the synchronizability of so-called disk graphs, where nodes are placed at random in a region of space and are connected if the distance separating them is smaller than some visibility radius. Other random connectivity models of interest in this respect (to the social sciences for example) include small-world graphs and scale-free networks [12].

The remainder of the paper presents proofs for theorems 1–3.

Further notation. We denote the (random) eigenvalues of $\mathbf{L}_{(n),1}$ as

$$t_1^{(n)} = 0, \quad t_2^{(n)} \le \cdots \le t_n^{(n)}.$$

These can be either positive or negative. The eigenvalues of $\mathbf{L}_{(n)}$ obey

$$\lambda_1^{(n)} = 0, \quad \lambda_{i>1}^{(n)} = np_n + t_i^{(n)}.$$

3 Proof of Theorem 1 about the expectation

In this proof of Theorem 1, we establish an expression for the expectation of X_n (1). Consider the sequence

$$c_n = 2.01\sqrt{\frac{\log n}{np_n}}.$$

Assumption 2 implies that $\frac{1}{np_n}\|\mathbf{L}_{(n),1}\|_{\mathrm{op}} \le c_n$ almost surely as $n \to \infty$. Furthermore, Assumption 1 guarantees that $c_n \to 0$ as $n \to \infty$. In particular, for all n larger than some threshold, $c_n < 1$. For n large enough then, almost surely $\lambda_2^{(n)} = np_n + t_2^{(n)} > 0$, meaning that the graph defined by the adjacency matrix $\mathbf{A}_{(n)}$ is connected. When such is the case, X_n obeys

$$X_n = \sum_{i=2}^{n} \frac{p_n}{\lambda_i^{(n)}} = \sum_{i=2}^{n} \frac{p_n}{np_n + t_i^{(n)}} = \frac{1}{n}\sum_{i=2}^{n} \frac{1}{1 + \frac{1}{np_n} t_i^{(n)}}.$$

Then, using the series expansion $\frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - \cdots$, which is convergent for $|x| < 1$, we get:

$$X_n = \frac{1}{n}\sum_{i=2}^{n}\left[1 + \sum_{k=1}^{\infty}\left(-\frac{1}{np_n} t_i^{(n)}\right)^k\right] = \frac{1}{n}\left[(n-1) + \sum_{k=1}^{\infty}\sum_{i=2}^{n}\left(-\frac{1}{np_n} t_i^{(n)}\right)^k\right].$$

The summations commute because the series is convergent. Observe that

$$\sum_{i=2}^{n} (t_i^{(n)})^k = \mathrm{trace}(\mathbf{L}_{(n),1}^k).$$

Thus,

$$X_n = \frac{1}{n}\left[(n-1) + \sum_{k=1}^{\infty}\left(-\frac{1}{np_n}\right)^k \operatorname{trace}(\mathbf{L}_{(n),1}^k)\right].$$

In order to compute the expectation of $\operatorname{trace}(\mathbf{L}_{(n)}^{\dagger})$ then, we must understand the expectations $\mathbb{E}\Big\{\operatorname{trace}(\mathbf{L}_{(n),1}^k)\Big\}$ for $k = 1, 2, \dots$ The first few terms are given by:

$$\begin{aligned}
\mathbb{E}\Big\{\operatorname{trace}(\mathbf{L}_{(n),1}^1)\Big\} &= \mathbb{E}\Big\{\sum_{i\neq j} X_{ij}^{(n)}\Big\} = 0, \\
\mathbb{E}\Big\{\operatorname{trace}(\mathbf{L}_{(n),1}^2)\Big\} &= \mathbb{E}\Big\{\big\|\mathbf{L}_{(n),1}\big\|_{\mathrm{F}}^2\Big\} = 2n(n-1)\sigma_n^2, \\
\mathbb{E}\Big\{\operatorname{trace}(\mathbf{L}_{(n),1}^3)\Big\} &= 0.
\end{aligned}$$

So,

$$\begin{aligned}
\mathbb{E}X_n &= \frac{1}{n}\Big[(n-1) + 2(n-1)\frac{\sigma_n^2}{np_n^2} + \sum_{k=4}^{\infty}\left(-\frac{1}{np_n}\right)^k \mathbb{E}\Big\{\operatorname{trace}(\mathbf{L}_{(n),1}^k)\Big\}\Big] \\
&= 1 + \left(2\frac{\sigma_n^2}{p_n^2} - 1\right)\frac{1}{n} - 2\frac{\sigma_n^2}{p_n^2}\frac{1}{n^2} + \frac{1}{n}\sum_{k=4}^{\infty}\left(-\frac{1}{np_n}\right)^k \mathbb{E}\Big\{\operatorname{trace}(\mathbf{L}_{(n),1}^k)\Big\}. \qquad (5)
\end{aligned}$$

Let us bound the series:

$$\begin{aligned}
\left|\frac{1}{n}\sum_{k=4}^{\infty}\left(-\frac{1}{np_n}\right)^k \mathbb{E}\Big\{\operatorname{trace}(\mathbf{L}_{(n),1}^k)\Big\}\right| &= \left|\frac{1}{n}\mathbb{E}\left\{\sum_{i=1}^{n}\sum_{k=4}^{\infty}\left(-\frac{1}{np_n}t_i^{(n)}\right)^k\right\}\right| \\
&\le \frac{1}{n}\mathbb{E}\left\{\sum_{i=1}^{n}\left|\sum_{k=4}^{\infty}\left(-\frac{1}{np_n}t_i^{(n)}\right)^k\right|\right\} \\
&= \frac{1}{n}\mathbb{E}\left\{\sum_{i=1}^{n}\left|\frac{1}{np_n}t_i^{(n)}\right|^4\left|\sum_{k=0}^{\infty}\left(-\frac{1}{np_n}t_i^{(n)}\right)^k\right|\right\} \\
&= \frac{1}{n}\mathbb{E}\left\{\sum_{i=1}^{n}\left|\frac{1}{np_n}t_i^{(n)}\right|^4\left|\frac{1}{1+\frac{1}{np_n}t_i^{(n)}}\right|\right\}.
\end{aligned}$$

Notice that $|x| \le \varepsilon < 1 \Rightarrow \left|\frac{1}{1+x}\right| \le \frac{1}{1-\varepsilon}$. Since under assumptions 1 and 2, almost surely as $n \to \infty$, we have that $\left|\frac{1}{np_n}t_i^{(n)}\right| \le c_n < 1$, it holds that:

$$\left|\frac{1}{n}\sum_{k=4}^{\infty}\left(-\frac{1}{np_n}\right)^k \mathbb{E}\Big\{\operatorname{trace}(\mathbf{L}_{(n),1}^k)\Big\}\right| \le \frac{1}{n}\mathbb{E}\left\{\sum_{i=1}^{n}\frac{c_n^4}{1-c_n}\right\} = \frac{c_n^4}{1-c_n}.$$

Plugging the latter in (5) yields (2).

4 Proof of Theorem 2 about the fluctuation

We now show that X_n concentrates around its expected value—Theorem 2. The proof rests on McDiarmid's inequality [13, Thm. 3.1], which can be stated as follows:

Lemma 1 (McDiarmid's inequality)**.** *Let $Y_1, \ldots, Y_m$ be independent random variables taking values in the set $\mathcal{Y}$. Let $f\colon \mathcal{Y}^m \to \mathbb{R}$ be a function of these random variables such that for all realizations $y_1, \ldots, y_m \in \mathcal{Y}$ and for all $1 \le i \le m$, $y_i' \in \mathcal{Y}$, it holds that*

$$\left|f(y_1, \ldots, y_i, \ldots, y_m) - f(y_1, \ldots, y_i', \ldots, y_m)\right| \le c_i,$$

where c_i only depends on i. Then, for all $\alpha > 0$,

$$\mathbf{Pr}\left[|f - \mathbb{E}f| \ge \alpha\right] \le 2\exp\left(\frac{-2\alpha^2}{\sum_{i=1}^m c_i^2}\right).$$

Define the following matrices for all $n \ge 2$:

$$\mathbf{M}_{(n)} = \frac{1}{\sqrt{np_n}}\mathbf{L}_{(n),1}, \text{ and } \mathbf{R}_{(n)} = (\mathbf{M}_{(n)} + \sqrt{np_n}\mathbf{I}_{(n)})^{-1}.$$

Under Assumption 1, the random variable of interest X_n is linked to these matrices as follows:

$$\begin{aligned}
X_n = p_n \mathrm{trace}(\mathbf{L}_{(n)}^\dagger) &= p_n \sum_{i=2}^n \frac{1}{\lambda_i^{(n)}} = p_n \sum_{i=2}^n \frac{1}{np_n + t_i^{(n)}} \frac{1/\sqrt{np_n}}{1/\sqrt{np_n}} \\
&= \sqrt{\frac{p_n}{n}} \left[\sum_{i=1}^n \frac{1}{\sqrt{np_n} + \frac{t_i^{(n)}}{\sqrt{np_n}}} - \frac{1}{\sqrt{np_n}}\right] = \sqrt{\frac{p_n}{n}} \mathrm{trace}(\mathbf{R}_{(n)}) - \frac{1}{n}.
\end{aligned} \tag{6}$$

Thus, the fluctuation of X_n is governed by the fluctuation of $\mathrm{trace}(\mathbf{R}_{(n)})$:

$$|X_n - \mathbb{E}X_n| = \sqrt{\frac{p_n}{n}} \left|\mathrm{trace}(\mathbf{R}_{(n)}) - \mathbb{E}\mathrm{trace}(\mathbf{R}_{(n)})\right|.$$

We set out to apply McDiarmid's inequality to the function

$$f(X_{12}, X_{13}, \ldots, X_{1n}, X_{23}, \ldots, X_{2n}, \ldots, X_{(n-1),n}) \triangleq \mathrm{trace}(\mathbf{R}_{(n)}).$$

The inputs X_{ij} are i.i.d. shifted Bernoulli's. For any fixed pair $(i, j), i < j$, let $\mathbf{M}_{(n)}'$ and $\mathbf{R}_{(n)}'$ be $\mathbf{M}_{(n)}$ and $\mathbf{R}_{(n)}$ respectively, with X_{ij} replaced by $X_{ij}' = 1 - X_{ij} + 2p_n$ (the other possible value for this random variable). We find a bound c, uniform over the choice of (i, j), such that for all realizations of the random variables,

$$\left|\mathrm{trace}(\mathbf{R}_{(n)}) - \mathrm{trace}(\mathbf{R}_{(n)}')\right| \le c.$$

Since $\mathbf{M}_{(n)}' - \mathbf{M}_{(n)} = (\mathbf{R}_{(n)}')^{-1} - \mathbf{R}_{(n)}^{-1}$, it holds that

$$\mathbf{R}_{(n)}(\mathbf{M}_{(n)}' - \mathbf{M}_{(n)})\mathbf{R}_{(n)}' = \mathbf{R}_{(n)}\left((\mathbf{R}_{(n)}')^{-1} - \mathbf{R}_{(n)}^{-1}\right)\mathbf{R}_{(n)}' = \mathbf{R}_{(n)} - \mathbf{R}_{(n)}'$$

so that

$$\left|\mathrm{trace}(\mathbf{R}_{(n)}) - \mathrm{trace}(\mathbf{R}_{(n)}')\right| = \left|\mathrm{trace}(\mathbf{R}_{(n)}(\mathbf{M}_{(n)}' - \mathbf{M}_{(n)})\mathbf{R}_{(n)}')\right|. \tag{7}$$

The matrix $\mathbf{M}_{(n)}' - \mathbf{M}_{(n)}$ is zero except for the entries affected by X_{ij}. More specifically, letting e_i denote the i^{th} canonical basis vector of length n, by definition,

$$\mathbf{M}_{(n)}' - \mathbf{M}_{(n)} = \frac{X_{ij}' - X_{ij}}{\sqrt{np_n}}(e_i - e_j)(e_i - e_j)^\top. \tag{8}$$

Since $X'_{ij} - X_{ij} = \pm 1$,

$$\begin{aligned}\operatorname{trace}(\mathbf{R}_{(n)}(\mathbf{M}'_{(n)} - \mathbf{M}_{(n)})\mathbf{R}'_{(n)}) &= \frac{\pm 1}{\sqrt{np_n}}\operatorname{trace}\left(\mathbf{R}_{(n)}(e_i - e_j)(e_i - e_j)^\top \mathbf{R}'_{(n)}\right) \\ &= \frac{\pm 1}{\sqrt{np_n}}(e_i - e_j)^\top \mathbf{R}'_{(n)}\mathbf{R}_{(n)}(e_i - e_j).\end{aligned}$$

Combine this with (7) and $\|e_i - e_j\| = \sqrt{2}$ to obtain the following inequality:

$$\left|\operatorname{trace}(\mathbf{R}_{(n)}) - \operatorname{trace}(\mathbf{R}'_{(n)})\right| \leq \frac{2}{\sqrt{np_n}}\|\mathbf{R}_{(n)}\|_{\text{op}}\|\mathbf{R}'_{(n)}\|_{\text{op}}. \tag{9}$$

Assumptions 1 and 2 guarantee that $\sqrt{np_n} \gg \sqrt{\log n}$ and that almost surely, as $n \to \infty$, $\|\mathbf{M}_{(n)}\|_{\text{op}} < 2.01\sqrt{\log n}$. Thus, it holds almost surely as $n \to \infty$ that

$$\|\mathbf{R}_{(n)}\|_{\text{op}} = \max_{1 \leq i \leq n} \frac{1}{|\lambda_i(\mathbf{M}_{(n)}) + \sqrt{np_n}|} < \frac{1}{\sqrt{np_n} - 2.01\sqrt{\log n}} < \frac{1.01}{\sqrt{np_n}}.$$

At the same time, $\|(e_i - e_j)(e_i - e_j)^\top\|_{\text{op}} = 2$ and equation (8) yield,

$$\|\mathbf{M}'_{(n)}\|_{\text{op}} \leq \|\mathbf{M}_{(n)}\|_{\text{op}} + \frac{2}{\sqrt{np_n}} < 2.01\sqrt{\log n} + \frac{2}{\sqrt{np_n}} < 2.02\sqrt{\log n},$$

so that, similarly to $\mathbf{R}_{(n)}$, we have $\|\mathbf{R}'_{(n)}\|_{\text{op}} < \frac{1.01}{\sqrt{np_n}}$ (almost surely as $n \to \infty$). As a result, going back to (9), we have that

$$\left|\operatorname{trace}(\mathbf{R}_{(n)}) - \operatorname{trace}(\mathbf{R}'_{(n)})\right| \leq \frac{2 \cdot 1.01^2}{(np_n)^{3/2}}. \tag{10}$$

Applying McDiarmid's inequality to X_n using (6) and (10), it is seen that almost surely as $n \to \infty$, for all $\alpha > 0$,

$$\mathbf{Pr}\left[|X_n - \mathbb{E}X_n| \geq \alpha\right] \leq 2\exp\left(\frac{-2\alpha^2}{\frac{n(n-1)}{2}\,\frac{p_n}{n}\,\frac{4\cdot 1.01^4}{(np_n)^3}}\right) < 2\exp\left(-\alpha^2(np_n)^2/1.01^4\right).$$

Let us fix $-\alpha^2(np_n)^2/1.01^4 = \log\epsilon$ for some $\epsilon > 0$. Then, almost surely as $n \to \infty$, with probability at least $1 - 2\epsilon$,

$$|X_n - \mathbb{E}X_n| \leq 1.01^2\frac{\sqrt{\log 1/\epsilon}}{np_n}.$$

This concludes the proof.

5 Proof of Theorem 3, about a stronger condition

Assumption 1 is certainly satisfied. Let us investigate Assumption 2. We prove that, for any $\epsilon_0 > 0$, almost surely as $n \to \infty$,

$$\|\mathbf{L}_{(n),1}\|_{\text{op}} \leq (2 + \epsilon_0)\sqrt{np_n \log n}.$$

Recall that $\mathbf{L}_{(n),1} = \mathbf{L}_{(n)} - \mathbb{E}\{\mathbf{L}_{(n)}\}$. Define similarly

$$\begin{aligned}\mathbf{D}_{(n),1} &= \mathbf{D}_{(n)} - \mathbb{E}\{\mathbf{D}_{(n)}\} = \mathbf{D}_{(n)} - (n-1)p_n\mathbf{I}_{(n)}, \qquad \text{and}\\ \mathbf{A}_{(n),1} &= \mathbf{A}_{(n)} - \mathbb{E}\{\mathbf{A}_{(n)}\} = \mathbf{A}_{(n)} - p_n(\mathbb{1}_{n\times n} - \mathbf{I}_{(n)})\end{aligned}$$

so that $\mathbf{L}_{(n),1} = \mathbf{D}_{(n),1} - \mathbf{A}_{(n),1}$. We show subsequently that for all $\epsilon_1, \epsilon_2 > 0$, almost surely as $n \to \infty$,

$$\|\mathbf{A}_{(n),1}\|_{\mathrm{op}} \leq (2+\epsilon_1)\sqrt{np_n}, \text{ and} \tag{11}$$

$$\|\mathbf{D}_{(n),1}\|_{\mathrm{op}} \leq (2+\epsilon_2)\sqrt{np_n \log n}. \tag{12}$$

Then, since $\|\mathbf{L}_{(n),1}\|_{\mathrm{op}} \leq \|\mathbf{A}_{(n),1}\|_{\mathrm{op}} + \|\mathbf{D}_{(n),1}\|_{\mathrm{op}}$, the proof will be complete. Proving (12) just requires Assumption 1. We require $np_n \gg \log^6 n$ to establish (11) using results from [6].

We first consider $\mathbf{A}_{(n),1}$ and resort to [6, Thm. 3.2]. In the latter paper, self loops in the graph are allowed. To take this into account, let $Y_i^{(n)}$, $n \geq 2$, $1 \leq i \leq n$, be i.i.d. Bernoulli random variables with the same distribution as, and independent from, the $A_{ij}^{(n)}$'s, $i < j$, and let $\mathbf{Y}_{(n)} = \mathrm{diag}(Y_1^{(n)}, \ldots, Y_n^{(n)})$. Then, in the present work's notation, reference [6] defines a matrix C such that

$$np_nC = \mathbf{A}_{(n),1} + (\mathbf{Y}_{(n)} - p_n\mathbf{I}_{(n)}). \tag{13}$$

Theorem 3.2 in that reference states that, provided $np_n \gg \log^6 n$, then almost surely

$$\|C\|_{\mathrm{op}} = (1+o(1))\frac{2}{\sqrt{np_n}}. \tag{14}$$

The contribution of $\mathbf{Y}_{(n)} - p_n\mathbf{I}_{(n)}$ is negligible:

$$\sup_n \|\mathbf{Y}_{(n)} - p_n\mathbf{I}_{(n)}\|_{\mathrm{op}} = \sup_n \max_{1\leq i\leq n} |Y_i^{(n)} - p_n| \leq 1. \tag{15}$$

Combining (13), (14) and (15), it is seen that for all $\epsilon_1 > 0$, almost surely as $n \to \infty$:

$$\begin{aligned}\|\mathbf{A}_{(n),1}\|_{\mathrm{op}} &\leq \|np_nC\|_{\mathrm{op}} + 1 = (2+o(1))\sqrt{np_n} + 1\\ &\leq (2+\epsilon_1)\sqrt{np_n}.\end{aligned}$$

This shows (11).

Now consider $\|\mathbf{D}_{(n),1}\|_{\mathrm{op}} = \max_{1\leq i\leq n} |(\mathbf{D}_{(n),1})_{ii}|$. Each diagonal entry

$$(\mathbf{D}_{(n),1})_{ii} = \sum_{j\neq i} X_{ij}^{(n)} = \Big(\sum_{j\neq i} A_{ij}^{(n)}\Big) - (n-1)p_n$$

is the difference between a sum of independent, nonnegative Bernoulli random variables and its mean. The Chernoff bound for nonnegative Bernoulli's [5, Thm. 2.4] controls such differences as follows, for positive λ:

$$\begin{aligned}\mathbf{Pr}\left[(\mathbf{D}_{(n),1})_{ii} \geq +\lambda\right] &\leq \exp\left(-\frac{\lambda^2}{2((n-1)p_n + \frac{\lambda}{3})}\right),\\ \mathbf{Pr}\left[(\mathbf{D}_{(n),1})_{ii} \leq -\lambda\right] &\leq \exp\left(-\frac{\lambda^2}{2(n-1)p_n}\right).\end{aligned}$$

Combining these two inequalities and setting $\lambda = \sqrt{cnp_n \log n}$ for some positive c yet to determine, we obtain:

$$\begin{aligned}\mathbf{Pr}\left[\left|(\mathbf{D}_{(n),1})_{ii}\right| \geq \sqrt{cnp_n \log n}\right] &\leq \exp\left(-\frac{cnp_n \log n}{2\left((n-1)p_n + \frac{\sqrt{cnp_n \log n}}{3}\right)}\right) + \exp\left(-\frac{cnp_n \log n}{2(n-1)p_n}\right) \\ &\leq \exp\left(-\frac{c}{2}\frac{\log n}{\frac{n-1}{n} + \frac{1}{3}\sqrt{\frac{c \log n}{np_n}}}\right) + \exp\left(-\frac{c}{2}\log n\right).\end{aligned}$$

Given that $\sqrt{c \log n / np_n} \to 0$ as $n \to \infty$, for all $\epsilon' > 0$, there exists $n_{\epsilon'}$ such that for all $n \geq n_{\epsilon'}$, the denominator in the first exponential obeys $\frac{n-1}{n} + \frac{1}{3}\sqrt{\frac{c \log n}{np_n}} \leq 1 + \epsilon'$. Hence, we further obtain:

$$\begin{aligned}\mathbf{Pr}\left[\left|(\mathbf{D}_{(n),1})_{ii}\right| \geq \sqrt{cnp_n \log n}\right] &\leq \exp\left(-\frac{c}{2}\frac{\log n}{1+\epsilon'}\right) + \exp\left(-\frac{c}{2}\log n\right) \\ &\leq 2\exp\left(-\frac{c}{2}\frac{\log n}{1+\epsilon'}\right) = \frac{2}{n^{\frac{c}{2+2\epsilon'}}}.\end{aligned}$$

By the union bound, which states that the probability of at least one event among n events to occur is bounded by the sum of the probabilities of those n events, it follows that:

$$\mathbf{Pr}\left[\left|(\mathbf{D}_{(n),1})_{ii}\right| \leq \sqrt{cnp_n \log n}, \text{ for all } i\right] \geq 1 - \frac{2}{n^{\frac{c}{2+2\epsilon'}-1}}.$$

For all $\epsilon_2 > 0$, let $\sqrt{c} = 2 + \epsilon_2$. We may pick $\epsilon' > 0$ such that $c > 4(1+\epsilon')$. Thus, it holds that for all $\epsilon_2 > 0$ and for large n, with probability at least $1 - 2/n^\gamma$ for some $\gamma > 1$,

$$\|\mathbf{D}_{(n),1}\|_{\mathrm{op}} \leq (2+\epsilon_2)\sqrt{np_n \log n}.$$

By the Borel-Cantelli lemma, this shows (12) holds almost surely as $n \to \infty$.

Acknowledgment

We thank Amit Singer for suggesting this collaboration and for interesting discussions. This paper presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme initiated by the Belgian Science Policy Office. NB is an FNRS research fellow.

References

[1] P. Barooah and J.P. Hespanha. Estimation on graphs from relative measurements. *Control Systems Magazine, IEEE*, 27(4):57–74, 2007.

[2] N. Boumal. On intrinsic Cramér-Rao bounds for Riemannian submanifolds and quotient manifolds. *Signal Processing, IEEE Transactions on*, 61(7):1809–1821, 2013.

[3] N. Boumal, A. Singer, and P.-A. Absil. Robust estimation of rotations from relative measurements by maximum likelihood. In preparation, 2013.

[4] N. Boumal, A. Singer, P.-A. Absil, and V. D. Blondel. Cramér-Rao bounds for synchronization of rotations. *CoRR*, abs/1211.1621, 2012.

[5] F. Chung and L. Lu. *Complex graphs and networks*. Number 107. AMS Bookstore, 2006.

[6] F. Chung, L. Lu, and V. Vu. The spectra of random graphs with given expected degrees. *Internet Mathematics*, 1(3):257–275, 2004.

[7] M. Cucuringu, Y. Lipman, and A. Singer. Sensor network localization by eigenvector synchronization over the Euclidean group. *ACM Transactions on Sensor Networks*, 8(3):19:1–19:42, 2012.

[8] 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.

[9] X. Ding and T. Jiang. Spectral distributions of adjacency and Laplacian matrices of random graphs. *The Annals of Applied Probability*, 20(6):2086–2117, 2010.

[10] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–562, 2006.

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

[12] A. Jamakovic and S. Uhlig. On the relationship between the algebraic connectivity and graph's robustness to node and link failures. In *Next Generation Internet Networks, 3rd EuroNGI Conference on*, pages 96–102. IEEE, 2007.

[13] C. McDiarmid. Concentration. In M. Habib, C. McDiarmid, J. Ramirez-Alfonsin, and B. Reed, editors, *Probabilistic Methods for Algorithmic Discrete Mathematics*, volume 16 of *Algorithms and Combinatorics*, pages 195–248. Springer Berlin Heidelberg, 1998.

[14] A. Singer. Angular synchronization by eigenvectors and semidefinite programming. *Applied and Computational Harmonic Analysis*, 30(1):20–36, 2011.

[15] B. Sonday, A. Singer, and I.G. Kevrekidis. Noisy dynamic simulations in the presence of symmetry: Data alignment and model reduction. *Computers and Mathematics with Applications*, 65(10):1535–1557, 2013.

[16] R. Tron and R. Vidal. Distributed image-based 3D localization of camera sensor networks. In *Decision and Control, held jointly with the 28th Chinese Control Conference. Proceedings of the 48th IEEE Conference on*, pages 901–908. IEEE, 2009.

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