Convergence Rate Estimates for Consensus over Random Graphs

Matthew T. Hale and Magnus Egerstedt[†]

Abstract— **Multi-agent coordination algorithms with randomized interactions have seen use in a variety of settings in the multi-agent systems literature. In some cases, these algorithms can be random by design, as in a gossip algorithm, and in other cases they are random due to external factors, as in the case of intermittent communications. Targeting both of these scenarios, we present novel convergence rate estimates for consensus problems solved over random graphs. Established results provide asymptotic convergence in this setting, and we provide estimates of the rate of convergence in two forms. First, we estimate decreases in a quadratic Lyapunov function over time to bound how quickly the agents' disagreement decays, and then we bound the probability of being at least a given distance from the point of agreement. Both estimates rely on (approximately) computing eigenvalues of the expected matrix exponential of a random graph's Laplacian, which we do explicitly in terms of the network's size and edge probability without assuming that any relationship between them holds. Simulation results are provided to support the theoretical developments made.**

I. Introduction

Distributed agreement, often broadly referred to as the *consensus* problem, is a canonical problem in distributed coordination and has received attention in diverse fields such as physics [23], signal processing [20], robotics [19], power systems [16], and communications [14]. The goal in such problems is to drive all agents in a network to a common final state. A key feature of consensus problems is their distributed nature; consensus is typically carried out across a network of agents in which each agent communicates with some other agents, though generally not all of them. The wide range of fields which study distributed agreement has given rise to corresponding diversity among consensus problem formulations, and a number of variants of consensus have been studied in the literature.

In this paper, we derive convergence rates for consensus over random graphs, studied previously in [10] where asymptotic convergence was shown. In some cases, the motivation for representing a communication network using random graphs comes from agents using an interaction protocol that is randomized by design, such as in a gossip-like algorithm [3]. In other cases, unreliable communications due to poor channel quality, interference, and other factors can be effectively represented by a random communication graph [15], and the work here applies to each of these scenarios. This problem formulation has each agent communicating with a random collection of other agents determined by a random graph. Each agent moves toward the average of its neighbors' states, then a new graph is generated, and each agent moves toward the average state of its new neighbors, with this process repeating until convergence is achieved.

We consider networks of a fixed size, and we examine consensus over random graphs generated by the Erdős-Rényi model [6], in which each possible edge in a graph is present with a fixed probability and is independent of all other edges. The Erdős-Rényi model is used because it accurately captures the behavior both of networks with intermittent and unreliable communications [10] and the behavior of some synchronous gossip algorithms [3]. Our approach consists of estimating the expected matrix exponential of the Laplacian associated with a random graph and then computing its eigenvalues in terms of the graph's size and edge probability. It is shown that the second-largest eigenvalue of this expected matrix exponential is key in estimating convergence rates, and the main contribution of this paper lies in explicitly computing this eigenvalue and using it to derive two novel rates of convergence. The first estimates the rate of decrease in a quadratic Lyapunov function to bound the rate at which agents approach agreement in their state values. The second bounds the probability of the agents' states being at least a given distance away from agreeing on a common state.

Both estimates rely in some form on computing eigenvalues of random matrices, and there is an established literature dedicated to doing so [5], [21], including for eigenvalues of random symmetric matrices [1], [8], [25], and eigenvalues of random graphs' Laplacians specifically [4]. A common approach to estimating or computing the eigenvalues of a random symmetric matrix is to let the size of the matrix get arbitrarily large [5], [8], [12], [25]. In graph theory, this approach corresponds to letting the number of nodes in a graph grow arbitrarily large and has seen use in spectral graph theory because it allows one to rigorously state results that hold for almost all graphs [2].

In the study of multi-agent systems, one is often interested in networks of a fixed, small size, such as in [14], [16], [24], and this makes results for asymptotically large networks less applicable in some cases. As a result, we are motivated to derive convergence results in terms of a network's size without taking it to grow asymptotically large. In addition, there are a number of graph theoretic results that estimate eigenvalues of random graphs' Laplacians when edge probabilities bear some known relationship to the size of the network, e.g., [13], [22]. In cases where a random graph is used to model unreliable communications, there is no guarantee that such a relationship will hold as the quality of communication channels can depend upon a wide variety

[†]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA. Email: {`matthale`, `magnus`}`@gatech.edu`.

of external factors. Therefore, we allow edge probabilities to take values independent of the network's size and state our results in terms of both a network's size and its edge probability without making assumptions about either.

The rest of the paper is organized as follows. Section II reviews the necessary background on consensus and graph theory, including that on random graphs. Next, Section III establishes asymptotic convergence of consensus over random graphs. Then Section IV presents our main results on the rate of convergence of consensus over random graphs. Section V then presents simulation results to verify the convergence rates we present. Finally, Section VI concludes the paper.

II. Review of Graph Theory and Consensus

In this section, we review the basic elements of graph theory required for the remainder of the paper. We introduce unweighted, undirected graphs, then review the consensus problem, and finally review random graphs.

A. Basic Graph Theory

All graphs in this paper are assumed to be unweighted and undirected. Such graphs are defined by pairwise relationships over a finite set of nodes or vertices. Suppose that a graph has a set V of n vertices, with $n \in \mathbb{N}$, and index these vertices over the set $\{1, \ldots, n\}$. We define the *edge set*

$$E \subset V \times V,$$

and say there is an edge between nodes i and j if $(i, j) \in E$. A graph G is then formally defined as the 2-tuple $G = (V, E)$. Throughout this paper, all edges are undirected and an edge $(i, j) \in E$ is not distinguished from the edge $(j, i) \in E$. We do not allow self loops and therefore $(i, i) \notin E$ for all i and all graphs G.

The degree associated with node i is defined as the total number of edges that connect node i to some other node. Using $|\cdot|$ to denote the cardinality of a set and using d_i to denote the degree of node i, we have

$$d_i = \big|\{j \mid (i, j) \in E\}\big|.$$

The $n \times n$ degree matrix associated with a graph G is then defined as

$$D(G) = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix},$$

which will be written simply as D when the graph G is understood.

The $n \times n$ adjacency matrix associated with G, denoted $A(G)$, is defined element-wise as

$$a_{i,j} = \begin{cases} 1 & (i, j) \in E \\ 0 & \text{otherwise} \end{cases}$$

where $a_{i,j}$ is the (i, j)th entry in $A(G)$. Note that, because $(i, j) \in E$ implies $(j, i) \in E$, $A(G)$ is a symmetric matrix. In addition, the absence of self loops results in $A(G)$ having zeroes on its main diagonal for all graphs G. We will simply write A when G is clear from context.

Using $D(G)$ and $A(G)$, we define the Laplacian associated with a graph G as

$$L(G) = D(G) - A(G),$$

and note that the Laplacian of any undirected, unweighted graph is a symmetric, positive semi-definite matrix [9]. In particular, the Laplacians of such graphs have all non-negative eigenvalues. Letting $\lambda_i(M)$ denote the i^{th} smallest eigenvalue of a matrix M, for any graph G it is known that $\lambda_1\big(L(G)\big) = 0$ and that $\mathbb{1} = (1, 1, \ldots, 1)^T$ is an eigenvector associated with λ_1 [15], i.e., that $\mathbb{1}$ is in the nullspace of the Laplacian of any undirected, unweighted graph. For any graph Laplacian L of size $n \times n$, we find

$$0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n. \tag{1}$$

A graph G is said to be *connected* if, for any two nodes i and j in the graph, there exists a sequence of edges one can traverse to travel from node i to node j. A seminal result in graph theory provides that G is connected if and only if $\lambda_2 > 0$ [7], and we use this fact in the next section when studying the consensus problem.

B. The Consensus Protocol

A canonical problem in multi-agent control is that of consensus. Consensus consists of having a collection of agents, e.g., robots or mobile sensor nodes, agree on a common value in a distributed way. The term distributed refers to the fact that each agent in a network can only communicate with some other agents in the network, but in general not all other agents. Suppose each agent has a scalar state[1], with agent i's state denoted x_i, and assemble these into the ensemble state vector

$$x = (x_1, \ \ldots, \ x_n)^T \in \mathbb{R}^n.$$

If one represents the agents' communications using a graph G (where edges connect those agents that exchange information), then the continuous-time consensus protocol [18] takes the form

$$\dot{x} = -Lx, \tag{2}$$

where L is the Laplacian of the graph G.

The following well-studied theorem from [17] establishes convergence of the consensus protocol in continuous time when G is connected. In it, we use

$$\bar{x}(0) := \frac{1}{n}\sum_{i=1}^{n} x_i(0)$$

to denote the centroid of the agents' states (which is a scalar) and $\bar{x}(0)\mathbb{1}$ to denote the vector in $\mathbb{R}^n$ whose entries are all equal to $\bar{x}(0)$.

[1]All results in this paper are easily extended to the case of agents having vectors of states by replacing L in the consensus protocol in Equation (2) with $I \otimes L$, where $\otimes$ denotes the Kronecker product of two matrices. We focus on the scalar case to simplify issues of dimensionality.

Theorem 1: Let G be a connected graph and let an initial ensemble state $x(0)$ be given. Then the consensus protocol

$$\dot{x} = -Lx \tag{3}$$

asymptotically converges element-wise to the centroid of the agents' initial states, i.e., $x(t) \to \bar{x}(0)\mathbb{1}$. In addition, the rate at which $x(t)$ approaches $\bar{x}(0)\mathbb{1}$ is governed by λ_2.

Proof: See [17]. ■

Because the goal in consensus is for a team of nodes to reach a common value, one often refers to the agreement subspace of the agents, defined below as was done in [10].

Definition 1: ([10]) The agreement subspace is defined as the set of all points at which all agents have the same state value, i.e., $x_i = x_j$ for all i and j. Formally, it is defined as

$$\mathcal{A} = \text{span}\{\mathbb{1}\},$$

where $\mathbb{1}$ is the vector of all ones in $\mathbb{R}^n$. △

In Section III, we will show asymptotic convergence to $\mathcal{A}$ for consensus over random graphs by showing that the agents' disagreement goes to zero. Next, we introduce the random graph model used in the remainder of the paper.

C. Random Graphs

A common model for random graphs is the Erdős-Rényi model, originally published in [6], and we use it here because it accurately captures the behavior of two cases of interest. First, some algorithms are randomized by design, such as gossip algorithms [3], and Erdős-Rényi graphs can model the behavior of such algorithms in some cases. Second, members of a network sometimes communicate over communication channels which are intermittently lost and regained, and this behavior is well-modeled by Erdős-Rényi graphs [10]. This model takes two parameters to generate random graphs: a number of nodes $n \in \mathbb{N}$ and an edge probability[2] $p \in (0, 1)$. The Erdős-Rényi model generates graphs on n nodes whose edge sets contain each possible edge with probability p, independent of all other edges. Formally, for each admissible i and j, we have

$$\mathbb{P}[(i, j) \in E] = p.$$

An alternative characterization that we use later can be stated in terms of the elements of the adjacency matrix of a random graph: for n nodes and edge probability p, we find

$$\mathbb{P}[a_{i,j} = 1] = \mathbb{P}[a_{j,i} = 1] = p \\ \text{and } \mathbb{P}[a_{i,j} = 0] = \mathbb{P}[a_{j,i} = 0] = 1 - p.$$

Thus each $a_{i,j}$ is a Bernoulli random variable.

We use $\mathcal{G}(n, p)$ to denote the sample space of all possible random graphs generated by the Erdős-Rényi model on n nodes with edge probability p, and we use $\mathcal{L}(n, p)$ to denote the set of Laplacians of all such graphs. One approach to analyzing random graphs that has seen use in the graph theory literature consists of taking the limit as n goes to infinity, with the benefit of this approach being the ability to rigorously determine which properties hold for almost all graphs. However, motivated by the study of multi-agent systems, we are interested in networks of fixed size and therefore develop our results in terms of a fixed value of n. Toward doing so, we show asymptotic convergence of consensus over random graphs in the next section.

[2]The cases $p = 0$ and $p = 1$ provide edgeless graphs and complete graphs, respectively, and are omitted because their behavior is deterministic.

III. Consensus over Random Graphs

In this section we examine consensus where the agents' communication graph at each timestep is a random graph. We then show asymptotic convergence of this update law in order to help establish the role of convergence rates in its analysis. The theoretical results of this section are not new, but are presented to enable the development of convergence rates later in the paper. This section closely follows the approach of [10], which originally presented the results of this section.

A. Time-Sampled Consensus

We assume that all communication graphs hold constant for some positive amount of time $\delta > 0$ and we seek to examine the evolution of the system's state under this condition. We follow the problem setup of [10] in which states generated by the consensus protocol in Equation (3) are sampled in time by defining the state $z(k)$ as

$$z(k) = x(k\delta),$$

where the communication graph among the agents is assumed to hold constant over the interval $[k\delta, (k+1)\delta)$. Note that this problem is distinct from consensus performed in discrete time; the problem we consider analyzes samples of the continuous time state $x(t)$ rather than having states actually evolve in discrete time as in [11]. All agents still execute the protocol $\dot{x} = -Lx$ and, as in [10], $z(k)$ is introduced only as a theoretical tool for analyzing the behavior of $x(t)$ over time.

At each timestep k, the system will generally have a different graph Laplacian than it had at time $k - 1$. We denote the communication graph active at time k by G_k and denote its Laplacian by L_k. The solution to Equation (3) is $x(t) = e^{-Lt}x(0)$, and for any times t_1 and t_2 with $t_1 < t_2$ we find

$$x(t_2) = e^{-L(t_2-t_1)}x(t_1).$$

Setting $t_2 = (k + 1)\delta$ and $t_1 = k\delta$ then gives

$$x\big((k + 1)\delta\big) = e^{-\delta L_k}x(k\delta) \tag{4}$$

because the graph G_k is constant over the interval $[k\delta, (k + 1)\delta)$. Equation (4) then gives

$$z(k + 1) = e^{-\delta L_k}z(k) \tag{5}$$

and this is the protocol that we analyze in the remainder of the paper.

We emphasize that the agents' states still evolve in continuous time and that we analyze the continuous time signal $x(t)$ by analyzing samples taken every δ seconds. In

Section IV we provide convergence rates for Equation (5). Toward doing so, we next show that Equation (5) converges asymptotically to $\mathcal{A}$.

B. Establishing Asymptotic Convergence

We will assess convergence of consensus over random graphs by showing convergence to the agreement set $\mathcal{A}$, defined in Definition 1. Define the distance from a point $y \in \mathbb{R}^n$ to $\mathcal{A}$ as

$$\text{dist}(y, \mathcal{A}) = \inf_{z \in \mathcal{A}} \|z - y\|_2.$$

As in [10], define the orthogonal complement of the agreement subspace as

$$\mathcal{A}^\perp = \{x \mid x^T a = 0 \text{ for all } a \in \mathcal{A}\}.$$

In particular $x^T \mathbb{1} = 0$ for all $x \in \mathcal{A}^\perp$. Then define the Euclidean projection of $z(k)$ onto $\mathcal{A}^\perp$ as

$$\hat{z}(k) = \Pi_{\mathcal{A}^\perp}[z(k)] = \left(I - \frac{1}{n}J\right) z(k),$$

where $\hat{z}(k)$ captures the disagreement among agents by excluding the part of $z(k)$ that lies in $\mathcal{A}$.

It was noted in Theorem 1 that the consensus protocol over static graphs converges element-wise to centroid of the agents' initial states. The consensus protocol in Equation (5) also converges to the centroid of the agents' initial states, despite being run over random graphs. To see why, first observe that the consensus protocol in Equation (5) has converged when $z(k+1) - z(k) = 0$, i.e., when

$$e^{-\delta L_k} z(k) - z(k) = (e^{-\delta L_k} - I) z(k) = 0. \tag{6}$$

Because $\mathcal{A}$ is a subspace, we can decompose $z(k)$ into two parts according to

$$z(k) = \Pi_{\mathcal{A}}[z(k)] + \Pi_{\mathcal{A}^\perp}[z(k)] = \Pi_{\mathcal{A}}[z(k)] + \hat{z}(k).$$

Substituting this decomposition into Equation (6) we find

$$\begin{aligned}
&(e^{-\delta L_k} - I) z(k) = \\
&\Big(-\delta L_k + \frac{1}{2}\delta^2 L_k^2 - \frac{1}{3!}\delta^3 L_k^3 + \cdots\Big)\big(\Pi_{\mathcal{A}}[z(k)] + \hat{z}(k)\big) \\
&\quad = \left(-\delta L_k + \frac{1}{2}\delta^2 L_k^2 - \frac{1}{3!}\delta^3 L_k^3 + \cdots\right) \hat{z}(k) = 0,
\end{aligned}$$

where we have used the fact that $\Pi_{\mathcal{A}}[z(k)]$ is in the nullspace of all $L_k \in \mathcal{L}(n,p)$ by definition. For Equation (5) to have converged, one must therefore have $\hat{z}(k)$ in the nullspace of all $L_k \in \mathcal{L}(n,p)$. Since $\hat{z}(k) \perp \mathcal{A}$, we find $\hat{z}(k) = 0$. Then if Equation (5) has converged at time k, it has converged to $\Pi_{\mathcal{A}}[z(k)]$.

To see that $\Pi_{\mathcal{A}}[z(k)] = \bar{z}(0)$, consider $\Pi_{\mathcal{A}}[z(1)]$. We have

$$\Pi_{\mathcal{A}}[z(1)] = z(1) - \hat{z}(1) = z(1) - \left(I - \frac{1}{n}J\right) z(1) = \bar{z}(1)\mathbb{1}.$$

Examining $\bar{z}(1)$, we see that

$$\begin{aligned}
\bar{z}(1) &= \frac{1}{n}\mathbb{1}^T z(1) = \frac{1}{n}\mathbb{1}^T e^{-\delta L_0} z(0) \\
&= \frac{1}{n}\mathbb{1}^T \left(I - \delta L_0 + \frac{\delta^2 L_0^2}{2} - \cdots\right) z(0) \\
&= \frac{1}{n}\mathbb{1}^T z(0) = \bar{z}(0),
\end{aligned}$$

where we have used $\mathbb{1}^T L_0 = (L_0^T \mathbb{1})^T = (L_0 \mathbb{1})^T = 0$ because L_0 is symmetric and $\mathbb{1}$ is in the nullspace of L_0. A simple inductive argument shows that $\bar{z}(k) = \bar{z}(0)$ for all k.

Therefore, the distance from $z(k)$ to $\mathcal{A}$ is equal to that from $z(k)$ to $\bar{z}(0)\mathbb{1}$. Noting that $\bar{z}(0)\mathbb{1}$, whose entries are all $\bar{z}(0)$, is equal to $\frac{1}{n}\mathbb{1}^T z(0)\mathbb{1}$, we have

$$\begin{aligned}
\text{dist}\big(z(k), \mathcal{A}\big)^2 &= \left\| z(k) - \frac{1}{n}\mathbb{1}^T z(0)\mathbb{1} \right\|^2 \\
&= z(k)^T z(k) - n\bar{z}(0)^2 = \frac{1}{n} z(k)^T \hat{L} z(k),
\end{aligned}$$

where we have used $\hat{L} := nI - J$, with I the $n \times n$ identity matrix and J the $n \times n$ matrix of ones.

In light of the form of $\text{dist}\big(z(k), \mathcal{A}\big)$, we will show convergence of the protocol in Equation (5) using the quadratic Lyapunov function

$$V\big(z(k)\big) = \frac{1}{n} z(k)^T \hat{L} z(k). \tag{7}$$

We also state the following definition which will be used to characterize stochastic convergence.

Definition 2: A random sequence $\{y(k)\}$ in $\mathbb{R}^n$ converges to $y \in \mathbb{R}^n$ with probability 1 if, for every $\epsilon > 0$,

$$\mathbb{P}\left[\sup_{N \le k < \infty} \|y(k) - y\|_2 \ge \epsilon\right] \to 0 \text{ as } N \to \infty.$$

△

We now have the following theorem that proves asymptotic convergence of the consensus protocol in Equation (5), due originally to [10].

Theorem 2: Fix a number of nodes $n \in \mathbb{N}$ and an edge probability $p \in (0,1)$, and let $z(0)$ be given. The consensus protocol

$$z(k+1) = e^{-\delta L_k} z(k),$$

where $L_k \in \mathcal{L}(n,p)$ for all k, converges to $\mathcal{A}$ with probability 1. In addition,

$$\hat{z}(k)^T \mathbb{E}[e^{-2\delta L_k} - I]\hat{z}(k) \to 0$$

with probability 1 and

$$\mathbb{P}\left[\sup_{N \le k < \infty} \|\hat{z}(k)\|_2^2 \ge \gamma\right] \le \frac{\hat{z}(0)^T \hat{z}(0)}{\gamma} \lambda_{n-1}\big(\mathbb{E}[e^{-2\delta L_k}]\big)^N$$

where $\lambda_{n-1}(\cdot)$ denotes the second-largest eigenvalue of a matrix.

Proof: See [10]. ■

More details on this theorem can be found in [10], where its proof originally appeared, and in [15]. Having established asymptotic convergence, we derive rates of convergence for consensus over random graphs in the next section.

IV. Convergence Rates for Consensus over Random Graphs

Section III showed that consensus over random graphs converges asymptotically, and in this section we derive our main results on two rates of convergence for consensus over random graphs. Section IV-A is based on work in [10], though the form of convergence rate we derive is different from the one derived there. The remainder of the section then presents our main results on novel, explicit convergence rates.

A. Convergence Rates

We now highlight the utility of convergence rates in the context of Theorem 2. To do so, we examine the evolution of the disagreement among agents. Using the same Lyapunov function as in Equation (7), we find

$$\mathbb{E}[V\big(\hat{z}(k+1)\big) - V\big(\hat{z}(k)\big) \mid \hat{z}(k)] = \hat{z}(k)^T \mathbb{E}[e^{-2\delta L_k} - I]\hat{z}(k)$$

$$= \hat{z}(k)^T \mathbb{E}[e^{-2\delta L_k}]\hat{z}(k) - \hat{z}(k)^T\hat{z}(k).$$

By definition, $\hat{z}(k)$ is orthogonal to $\mathbb{1}$. $\mathbb{1}$ is also the eigenvector associated with the largest eigenvalue of $e^{-2\delta L_k}$ for all $L_k \in \mathcal{L}(n,p)$ (because the eigenvalues of $e^{-2\delta L_k}$ are $e^{-2\delta\lambda_i}$ for each λ_i in Equation (1)). Then we find that

$$\hat{z}(k)^T \mathbb{E}\big[e^{-2\delta L_k}\big]\hat{z}(k) \leq \lambda_{n-1}\left(\mathbb{E}\big[e^{-2\delta L_k}\big]\right) \|\hat{z}(k)\|_2^2,$$

where $\lambda_{n-1}(M)$ denotes the second largest eigenvalue of an $n \times n$ matrix M. Consequently, we have

$$\mathbb{E}[V\big(\hat{z}(k+1)\big) - V\big(\hat{z}(k)\big) \mid \hat{z}(k)] \leq \big(\lambda_{n-1}(\mathbb{E}[e^{-2\delta L_k}]) - 1\big)\|\hat{z}(k)\|_2^2, \quad (8)$$

and by Theorem 2 we have

$$\mathbb{P}\left[\sup_{N \leq k < \infty} \|\hat{z}(k)\|_2^2 \geq \gamma\right] \leq \frac{\hat{z}(0)^T\hat{z}(0)}{\gamma}\lambda_{n-1}\big(\mathbb{E}[e^{-2\delta L_k}]\big)^N. \quad (9)$$

Therefore, both convergence rates depend upon $\lambda_{n-1}(\mathbb{E}[e^{-2\delta L_k}])$, and we compute it next.

B. Computing $\lambda_{n-1}\big(\mathbb{E}[e^{-2\delta L_k}]\big)$

We have the following lemma concerning eigenvalues of a matrix of the form $aI + b(J-I)$, which we will use below.

Lemma 1: A matrix M of the form $aI + b(J-I)$, namely

$$M = \begin{pmatrix} a & b & \cdots & b \\ b & a & \cdots & b \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \cdots & a \end{pmatrix} \in \mathbb{R}^{n\times n},$$

has $a + (n-1)b$ as an eigenvalue with multiplicity one and $a - b$ as an eigenvalue with multiplicity $n-1$.

Proof: We proceed using a series of row operations that will preserve the characteristic polynomial of M. We see that

$$|M - \lambda I| = \left|\begin{pmatrix} a-\lambda & b & \cdots & b \\ b & a-\lambda & \cdots & b \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \cdots & a-\lambda \end{pmatrix}\right|.$$

Next, we add rows 2 through n to row 1, giving

$$|M - \lambda I| = \big(a + (n-1)b - \lambda\big)\left|\begin{pmatrix} 1 & 1 & \cdots & 1 \\ b & a-\lambda & \cdots & b \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \cdots & a-\lambda \end{pmatrix}\right|.$$

Subtracting b times row 1 from each other row, we find

$$|M - \lambda I| = \big(a + (n-1)b - \lambda\big)\cdot \left|\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & a-b-\lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a-b-\lambda \end{pmatrix}\right|,$$

where the matrix on the right-hand side is upper-triangular. The determinant on the right-hand side is then the product of the diagonal entries of that matrix, resulting in

$$|M - \lambda I| = (a + (n-1)b - \lambda)(a - b - \lambda)^{n-1},$$

whose roots are indeed $a + (n-1)b$, with multiplicity 1, and $a - b$ with multiplicity $n-1$. ■

We now derive the expected value of the first four powers of a random graph's Laplacian, and below we will use these results to approximate the Laplacian's expected matrix exponential. As above, we use the notation $\hat{L} = nI - J$.

Lemma 2: Let a number of nodes $n \in \mathbb{N}$ and edge probability $p \in (0,1)$ be given. The Laplacian $L \in \mathcal{L}(n,p)$ of a graph $G \in \mathcal{G}(n,p)$ satisfies

$$\begin{aligned} \mathbb{E}[L] &= p\hat{L} \\ \mathbb{E}[L^2] &= \big[(n-2)p^2 + 2p\big]\,\hat{L} \\ \mathbb{E}[L^3] &= \big[(n-2)(n-4)p^3 + 6(n-2)p^2 + 4p\big]\,\hat{L} \\ \mathbb{E}[L^4] &= \big[(n-7)(n-3)(n-2)p^4 \\ &\quad + 6(2n-7)(n-2)p^3 + 25(n-2)p^2 + 8p\big]\hat{L}. \end{aligned}$$

Proof sketch: We sketch the proof to avoid exposition on many tedious computations and instead elaborate on the core arguments used to derive the above results.

The general form of graph Laplacian for $G \in \mathcal{G}(n,p)$ is

$$\begin{pmatrix} \sum_{\substack{j=1\\j\neq 1}}^n a_{1,j} & -a_{1,2} & \cdots & -a_{1,n} \\ -a_{1,2} & \sum_{\substack{j=1\\j\neq 2}}^n a_{2,j} & \cdots & -a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1,n} & -a_{2,n} & \cdots & \sum_{\substack{j=1\\j\neq n}}^n a_{n,j} \end{pmatrix},$$

where each term $a_{i,j}$ is a Bernoulli random variable with expectation equal to p. The off-diagonal entries of L have $\mathbb{E}[L_{ij}] = \mathbb{E}[-a_{i,j}] = -p$, while linearity of $\mathbb{E}[\cdot]$ gives $\mathbb{E}[L_{ii}] = (n-1)p$ for diagonal entries of L. From this we find $\mathbb{E}[L] = (n-1)pI - p(J-I) = p\hat{L}$.

Computing the expectation of the entries of L^2 requires one to consider two cases. The diagonal entries of L^2 are formed by the product of row i of L with column i of L (and by symmetry of L these are identical), while the off-diagonal entries result from the product of row i and column j of L (which are not identical when $i \neq j$). It is important to note that $a_{i,j}^2 = a_{i,j}$ because $a_{i,j}$ is a Bernoulli random variable. As a result, when computing expectations one finds that $\mathbb{E}[a_{i,j}^k] = p$ for all $k \in \mathbb{N}$, while products of ℓ distinct off-diagonal entries of A have expectation equal to p^ℓ. In the case of $\mathbb{E}[L^2]$, a diagonal entry takes the form

$$\begin{aligned}\mathbb{E}\left[(L^2)_{ii}\right] &= \mathbb{E}\left[\sum_{\substack{p=1\\p\neq i}}^{n} a_{i,p}^2 + \left(\sum_{\substack{q=1\\q\neq i}}^{n} a_{i,q}\right)^2\right] \\ &= (n-1)(n-2)p^2 + 2(n-1)p,\end{aligned}$$

while an off-diagonal entry takes the form

$$\begin{aligned}\mathbb{E}[(L^2)_{ij}] = \mathbb{E}\left[\sum_{\substack{k=1\\k\neq i,j}}^{n} a_{i,k}a_{k,j}\right] - \mathbb{E}\left[a_{i,j}\sum_{\substack{p=1\\p\neq i}}^{n} a_{i,p}\right] \\ - \mathbb{E}\left[a_{i,j}\sum_{\substack{q=1\\q\neq i}}^{n} a_{q,j}\right] = -(n-2)p^2 - 2p.\end{aligned}$$

Then we find

$$\begin{aligned}\mathbb{E}[L^2] &= \big((n-1)(n-2)p^2 + 2(n-1)p\big)I \\ &+ \big(-(n-2)p^2 - 2p\big)(J-I) = \big((n-2)p^2 + 2p\big)\hat{L},\end{aligned}$$

as above.

Computing the general form of L^3 is done by multiplying the general form of L^2 by that of L and the general form of L^4 is found by squaring the general form of L^2. Having found these two general forms, one follows the above strategy for computing their expected values: first replace $a_{i,j}^k$ with $a_{i,j}$ for all $k > 1$ and then compute the expectation of the products of ℓ distinct off-diagonal entries of A as p^ℓ, resulting in the above. ■

We use the first four powers of L because that is all that is required for accurate convergence rate estimates, as will be shown in Section V. One way in which this accuracy is attained is through the choice of δ. Many choices of δ are possible and we choose $\delta = 1/n$, for several reasons. First, it was assumed in Section III-A that the communication graph of the system is constant between samples, i.e., that G_k does not change over the interval $[k\delta, (k+1)\delta)$. As a network grows, the number of possible edges does too and thus a larger network has more ways in which its communication topology may change. As a result, a larger network should use shorter sampling times and δ should decrease as n increases. The choice of $\delta = 1/n$ provides a simple means of enforcing this condition.

Second, the use of $\delta = 1/n$ is also partly inspired by the same choice made in [11] for discrete-time consensus where it is a necessary condition for stability; though not necessary for stability in the current paper, we make the same choice to help retain the same broad applicability of the results in [11]. Using this choice of δ, we now give the approximate value of $\lambda_{n-1}(\mathbb{E}[e^{-2\delta L_k}])$ in terms of n and p.

Theorem 3: Suppose $\delta = 1/n$ and define

$$\begin{aligned}\kappa_1 &:= p \\ \kappa_2 &:= (n-2)p^2 + 2p \\ \kappa_3 &:= (n-2)(n-4)p^3 + 6(n-2)p^2 + 4p \\ \kappa_4 &:= (n-7)(n-3)(n-2)p^4 \\ &\quad + 6(2n-7)(n-2)p^3 + 25(n-2)p^2 + 8p \\ \mu(n,p) &:= -2\frac{\kappa_1}{n} + 2\frac{\kappa_2}{n^2} - \frac{4}{3}\frac{\kappa_3}{n^3} + \frac{2}{3}\frac{\kappa_4}{n^4}.\end{aligned}$$

Then

$$E[e^{-2\delta L_k}] \approx I + \mu(n,p)\hat{L},$$

and consequently we have

$$\lambda_{n-1}(\mathbb{E}[e^{-2\delta L_k}]) \approx 1 + n\mu(n,p).$$

Proof: Taylor expanding the matrix exponential, we find

$$\mathbb{E}[e^{-2\delta L_k}] \approx I - 2\delta\mathbb{E}[L_k] + 2\delta^2\mathbb{E}[L_k^2] - \frac{4}{3}\delta^3\mathbb{E}[L_k^3] + \frac{2}{3}\delta^4\mathbb{E}[L_k^4], \tag{10}$$

where the exact matrix exponential is well-approximated by this truncation in part due to the aforementioned choice of $\delta = 1/n$. Substituting $\delta = 1/n$ and the results of Lemma 2 into Equation (10) gives

$$\begin{aligned}\mathbb{E}[e^{-2\delta L_k}] &\approx I + \left(-2\frac{\kappa_1}{n} + 2\frac{\kappa_2}{n^2} - \frac{4}{3}\frac{\kappa_3}{n^3} + \frac{2}{3}\frac{\kappa_4}{n^4}\right)\hat{L} \\ &= I + \mu(n,p)\hat{L},\end{aligned}$$

which establishes the first part of the theorem.

Next, we use the definition of $\hat{L}$ as $nI - J$ to find

$$I + \mu(n,p)\hat{L} = \big(1 + (n-1)\mu(n,p)\big)I - \mu(n,p)\big(J - I\big).$$

Using Lemma 1 with $a = 1 + (n-1)\mu(n,p)$ and $b = -\mu(n,p)$, we find that the largest eigenvalue of $\mathbb{E}[e^{-2\delta L_k}]$ is 1, and the second largest through smallest eigenvalues of $\mathbb{E}[e^{-2\delta L_k}]$ are all approximately $1 + n\mu(n,p)$. In particular, $\lambda_{n-1}\big(\mathbb{E}[e^{-2\delta L_k}]\big) \approx 1 + n\mu(n,p)$, as desired. ■

C. Explicit Convergence Rates for Consensus over Random Graphs

We now present our unified main convergence rates for consensus over random graphs, stated in terms of network size n and edge probability p.

Theorem 4: Let a network size $n \in \mathbb{N}$ and an edge probability $p \in (0,1)$ be given, and let $\mu(n,p)$ be as defined in Theorem 3. For sampling constant $\delta = 1/n$ we have

$$\mathbb{P}\left[\sup_{N \leq k < \infty} \|\hat{z}(k)\|_2^2 \geq \gamma\right] \leq \frac{\hat{z}(0)^T\hat{z}(0)}{\gamma}\big(1 + n\mu(n,p)\big)^N \tag{11}$$

for all $N \in \mathbb{N}$. In addition, the expected decrease in the Lyapunov function $V(z) = \frac{1}{n}z^T\hat{L}z$ from time k to time $k+1$ is bounded according to

$$\mathbb{E}\big[V\big(\hat{z}(k+1)\big) - V\big(\hat{z}(k)\big) \mid \hat{z}(k)\big] \leq n\mu(n,p)\|\hat{z}(k)\|_2^2. \tag{12}$$

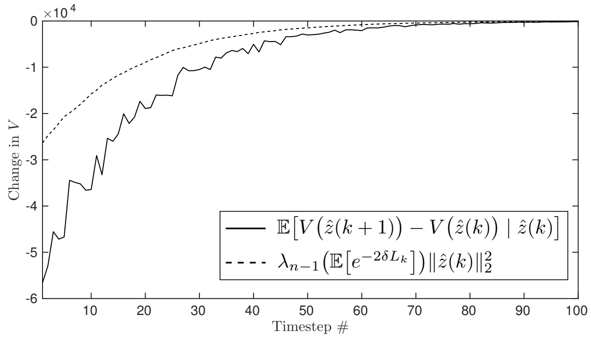

Figure 1. A simulation run showing the expected decrease in V (solid line) and its theoretical upper bound (dashed line) as in Equation (13). Here we see that the upper bound indeed holds across the whole time horizon, with the dashed line always above the solid line. In addition, the upper bound becomes more accurate across iterations, with it accurately predicting when the expected decrease in V is near zero.

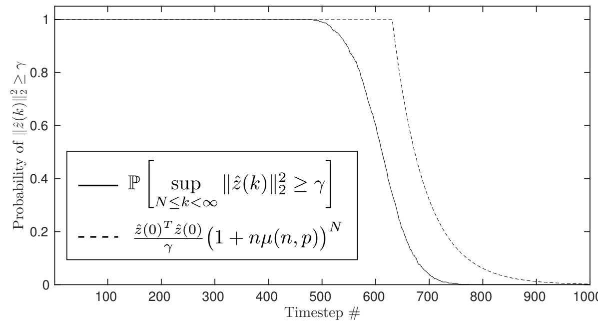

Figure 2. A simulation run showing the probability of having $\|\hat{z}(k)\|_2^2 \geq 3$. The empirical probability is plotted as the solid line, while the upper bound is shown as the dashed line. We see that the upper bound indeed holds and, despite relying only on expected values, provides a close enough approximation to the empirical probability to be useful in a variety of settings.

Proof: This follows from Equations (8) and (9) and Theorem 3. ■

The appeal of using these convergence rate estimates together is that one need only compute the constant $\mu(n,p)$ and then both estimates can be used. Both provide information at each timestep because they rely on the current iteration count and can therefore be applied in real time. One can also use Equation (11) for a range of values of γ to obtain the probability of being contained in each member of a family of super-level sets, letting one associate probabilities with all points in state space.

These two estimates also provide information about both individual consensus runs and families of consensus runs. Specifically, Equation (12) applies to single consensus runs and evolves the current expected decrease in V based upon $\hat{z}(k)$, providing a rate estimate specific to that run. On the other hand, Equation (11) applies to all runs starting from a given initial condition, giving information about how often we should expect one trajectory out of a family to be at least some distance from the point of convergence. Furthermore, they are complementary in that Equation (12) provides an upper bound on the rate of decrease in V and is optimistic in the sense that it over-estimates the expected decrease in disagreement in the system. On the other hand, Equation (11) over-estimates the probability of the agents' disagreement being a certain size, and is therefore pessimistic.

Taken together, these two convergence rate estimates enable one to probe the behavior of any consensus problem over random graphs, regardless of network size n, edge probability p, or initial condition $z(0)$, and provide quantitative data on such problems while requiring only the computation of $\mu(n,p)$. In the next section we present numerical results that verify both bounds presented in Theorem 4.

V. Simulation Results

In this section we present numerical results to support the results in Theorem 4. We simulate consensus over random graphs and first numerically examine the expected decrease in V, and second bound the probability of being at least some distance away from the point of agreement.

A. Estimating Decreases in Disagreement

We now present simulation results to verify the upper bound on the expected decrease in the Lyapunov function $V(z) = \frac{1}{n}z^T\hat{L}z$ as defined in Equation (7). The consensus problem we ran consists of $n = 50$ agents and edge probability $p = 0.03$. All agents had two states and were initialized to be evenly spaced along a circle of radius 100 centered on the origin. In this case, to estimate the rate of convergence using Equation (12), we compute $n\mu(n,p) = -0.0561$, from which we find

$$\mathbb{E}\big[V\big(\hat{z}(k+1)\big) - V\big(\hat{z}(k)\big) \mid \hat{z}(k)\big] \leq -0.0561\|\hat{z}(k)\|_2^2. \quad (13)$$

To validate Equation (13), a single consensus run was simulated. At each timestep k, the value of $\mathbb{E}\big[V\big(\hat{z}(k+1)\big) - V\big(\hat{z}(k)\big) \mid \hat{z}(k)\big]$ was computed numerically by fixing $\hat{z}(k)$, generating $1,000$ random graphs to compute $\tilde{z}(k+1) = e^{-2\delta L(G)}\hat{z}(k)$ for each graph G generated this way, and then computing $V\big(\tilde{z}(k+1)\big) - V\big(\hat{z}(k)\big) \mid \hat{z}(k)$ for each G. These values were then averaged to numerically determine $\mathbb{E}\big[V\big(\hat{z}(k+1)\big) - V\big(\hat{z}(k)\big) \mid \hat{z}(k)\big]$.

The results of this simulation run are shown in Figure 1, where the right-hand side of Equation (13) is shown as a dashed line and the left-hand side of Equation (13) is shown as a solid line; though $1,000$ timesteps of consensus were run, only the first 100 are shown because the two lines are indistinguishable and approximately zero beyond this point. Figure 1 shows that indeed the upper bound on the decreases in V from Theorem 4 holds because the dashed line is always above the solid line. Furthermore, as the iteration count increases, the upper bound becomes more accurate, meaning that we not only have an upper bound on the rate of decrease of V, but also that this upper bound can accurately predict when decreases in V go to zero, thereby accurately predicting when consensus is achieved.

B. Bounding the Probability of Being away from $\bar{z}(0)\mathbb{1}$

We now present a consensus problem to verify the upper bound on the probability that the disagreement among agents will be at least a certain amount after a fixed point in time, stated in Theorem 4. In particular, for $n = 10$ agents and edge probability $p = 0.01$, $1,000$ trials were run to find experimentally the value of $\mathbb{P}\big[\sup_{N \leq k < \infty} \|\hat{z}(k)\|_2^2 \geq \gamma\big]$ for each $N \in \{1, \ldots, 1,000\}$, where we set $\gamma = 3$. All trials were initialized with the agents spaced equally along a circle of radius 100 whose center was at the origin. The results of these numerical experiments are shown in Figure 2 in the solid line, while the theoretical upper bound $\frac{\hat{z}(0)^T \hat{z}(0)}{\gamma}\big(1 + n\mu(n,p)\big)^N$ is shown in Figure 2 as the dashed line.

Figure 2 shows that the probability bound in Theorem 4 indeed holds because the dashed line is always aligned with or above the solid line. In addition, we see that the upper bound's graph over time stays close to that of the empirical probability, indicating that, despite relying only on expected values, the bound on the probability of being at least some distance from consensus provides a useful estimate of the actual probability, enabling one to make predictions about the magnitude of disagreement in a network over time.

VI. Conclusion

Explicit convergence rate bounds were presented for consensus over random graphs. A key feature was that convergence rate estimates were given in terms of the network size and edge probability without making any assumptions about either. Eigenvalues of the expected exponential of random graphs' Laplacians were computed and used to derive approximate convergence rate bounds. Numerical results confirmed that these results accurately capture the behavior of consensus over random graphs.

Acknowledgments

The authors would like to thank Professor Daniel Spielman for his helpful comments on this work and the literature on spectral graph theory.

References

[1] Noga Alon, Michael Krivelevich, and Van H. Vu. On the concentration of eigenvalues of random symmetric matrices. *Israel Journal of Mathematics*, 131(1):259–267, 2002.

[2] Béla Bollobás. *Random Graphs:*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, Aug 2001.

[3] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE/ACM Trans. Netw.*, 14(SI):2508–2530, June 2006.

[4] Amin Coja-Oghlan. On the Laplacian eigenvalues of G(n, p). *Comb. Probab. Comput.*, 16(6):923–946, November 2007.

[5] Persi Diaconis and Mehrdad Shahshahani. On the eigenvalues of random matrices. *Journal of Applied Probability*, 31:49–62, 1994.

[6] Paul Erdős and Alfréd Rényi. On random graphs I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[7] Miroslav Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2):298–305, 1973.

[8] Z. Füredi and J. Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3):233–241, 1981.

[9] Chris Godsil and Gordon Royle. *Algebraic graph theory*, volume 207 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2001.

[10] Y. Hatano and M. Mesbahi. Agreement over random networks. *IEEE Transactions on Automatic Control*, 50(11):1867–1872, Nov 2005.

[11] A. Jadbabaie, Jie Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988–1001, June 2003.

[12] F. Juhász. On the asymptotic behaviour of the spectra of non-symmetric random (0, 1) matrices. *Disc. Math.*, 41(2):161–165, 1982.

[13] Michael Krivelevich and Benny Sudakov. The largest eigenvalue of sparse random graphs. *Comb. Probab. Comput.*, 12(1):61–72, 2003.

[14] M. Mehyar, D. Spanos, J. Pongsajapan, S. H. Low, and R. M. Murray. Asynchronous distributed averaging on communication networks. *IEEE/ACM Transactions on Networking*, 15(3):512–520, June 2007.

[15] Mehran Mesbahi and Magnus Egerstedt. *Graph theoretic methods in multiagent networks*. Princeton University Press, 2010.

[16] M. H. Nazari, Z. Costello, M. J. Feizollahi, S. Grijalva, and M. Egerstedt. Distributed frequency control of prosumer-based electric energy systems. *IEEE Trans. on Power Systems*, 29(6):2934–2942, Nov 2014.

[17] Reza Olfati-Saber and Richard M Murray. Consensus protocols for networks of dynamic agents. In *Proceedings of the 2003 American Controls Conference*, 2003.

[18] Reza Olfati-Saber and Richard M Murray. Consensus problems in networks of agents with switching topology and time-delays. *Automatic Control, IEEE Transactions on*, 49(9):1520–1533, 2004.

[19] Wei Ren and Randal W. Beard. *Distributed Consensus in Multi-vehicle Cooperative Control: Theory and Applications*. Springer Publishing Company, Incorporated, 1st edition, 2007.

[20] I. D. Schizas, G. Mateos, and G. B. Giannakis. Distributed LMS for consensus-based in-network adaptive processing. *IEEE Transactions on Signal Processing*, 57(6):2365–2382, June 2009.

[21] Terence Tao. *Topics in random matrix theory*. American Mathematical Society Providence, RI, 2012.

[22] Linh V. Tran, Van H. Vu, and Ke Wang. Sparse random graphs: Eigenvalues and eigenvectors. *Random Structures & Algorithms*, 42(1):110–134, 2013.

[23] Tamás Vicsek, András Czirók, Eshel Ben-Jacob, Inon Cohen, and Ofer Shochet. Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.*, 75:1226–1229, Aug 1995.

[24] G. Werner-Allen, K. Lorincz, M. Ruiz, O. Marcillo, J. Johnson, J. Lees, and M. Welsh. Deploying a wireless sensor network on an active volcano. *IEEE Internet Computing*, 10(2):18–25, March 2006.

[25] Eugene P. Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, 67(2):325–327, 1958.