# CONVERGENCE SPEED OF BINARY INTERVAL CONSENSUS

MOEZ DRAIEF* AND MILAN VOJNOVIĆ†

**Abstract.** We consider the convergence time for solving the binary consensus problem using the interval consensus algorithm proposed by Bénézit, Thiran and Vetterli (2009). In the binary consensus problem, each node initially holds one of two states and the goal for each node is to correctly decide which one of these two states was initially held by a majority of nodes.

We derive an upper bound on the expected convergence time that holds for arbitrary connected graphs, which is based on the location of eigenvalues of some contact rate matrices. We instantiate our bound for particular networks of interest, including complete graphs, paths, cycles, star-shaped networks, and Erdös-Rényi random graphs; for these graphs,we compare our bound with alternative computations. We find that for all these examples our bound is tight, yielding the exact order with respect to the number of nodes.

We pinpoint the fact that the expected convergence time critically depends on the voting margin defined as the difference between the fraction of nodes that initially held the majority and the minority states, respectively. The characterization of the expected convergence time yields exact relation between the expected convergence time and the voting margin, for some of these graphs, which reveals how the expected convergence time goes to infinity as the voting margin approaches zero.

Our results provide insights into how the expected convergence time depends on the network topology which can be used for performance evaluation and network design. The results are of interest in the context of networked systems, in particular, peer-to-peer networks, sensor networks and distributed databases.

**Key words.** Consensus, decentralized computation, distributed hypothesis testing

**AMS subject classifications.**

**1. Introduction.** Algorithms for distributed computation in networks have recently attracted considerable interest because of their wide-range of applications in networked systems such as peer-to-peer networks, sensor networks, distributed databases, and on-line social networks. A specific algorithmic problem of interest is the so called *binary consensus* [1, 2, 3, 4] where, initially, each node in the network holds one of two states and the goal for each node is to correctly decide which one of the two states was initially held by a majority of nodes. This is to be achieved by a decentralized algorithm where each node maintains its state based on the information exchanged at contacts with other nodes, where the contacts are restricted by the network topology. It is desired to reach a final decision by all nodes that is correct and within small convergence time.

A typical application scenario of the binary consensus corresponds to a set of agents who want to reach consensus on whether a given event has occurred based on their individual, one-off collected, information. Such cooperative decision-making settings arise in a number of applications such as environmental monitoring, surveillance and security, and target tracking [5], as well as voting in distributed systems [6]. Furthermore, it has been noted that one can use multiple binary consensus instances to solve multivalued consensuses; we refer to [7, 8] for an account on such algorithms.

We consider a decentralized algorithm known as *interval or quantized consensus* proposed by Bénézit, Thiran, and Vetterli [3]. The aim of this algorithm is to decide which one of $k \geq 2$ partitions of an interval contains the average of the initial values held by individual nodes. In this paper, we focus on *binary* interval consensus, i.e.

*Imperial College London

†Microsoft Research Cambridge

the case $k = 2$. An attractive feature of the interval consensus is its accuracy; it was shown in [3] that for any finite connected graph that describes the network topology, the interval consensus is guaranteed to converge to the correct state with probability 1. However, the following important question remained open: *How fast does the interval consensus converge to the final state?* We answer this question for the case of binary interval consensus.

The interval consensus could be considered a state-of-the-art algorithm for solving the binary consensus problem as it guarantees convergence to the correct consensus (i.e. has zero probability of error) for arbitrary finite connected graphs. Besides, it only requires a limited amount of memory and communication by individual nodes (only four states). Some alternative decentralized algorithms require fewer states of memory or communication but fail to reach the correct consensus with strictly positive probability. For instance, the traditional voter model requires only two states of memory and communication. It is however known that there are graphs for which the probability of error is a strictly positive constant, e.g. proportional to the number of nodes that initially held the minority state in the case of complete graphs. Another example is the ternary protocol proposed in [2] for which it was shown that for complete graphs, the probability of error diminishes to zero exponentially with the number of nodes, but provides no improvement over the voter model for some other graphs (e.g. a path).

In this paper, we provide an upper bound on the expected convergence time for solving the binary interval consensus on arbitrary connected graphs. This provides a unified approach for estimating the expected convergence time for particular graphs. The bound is tight in the sense that there exists a graph, namely the complete graph, for which the bound is achieved asymptotically for large number of nodes.

We demonstrate how the general upper bound can be instantiated for a range of particular graphs, including complete graphs, paths, cycles, star-shaped networks and Erdös-Rényi random graphs. Notice that the complete graph and the Erdös-Rényi random graph are good approximations of various unstructured and structured peer-to-peer networks and that star-shaped networks capture the scenarios where some node is a hub for other nodes.

Our results provide insights into how the expected convergence time depends on *the network structure* and *the voting margin*, where the latter is defined as the difference between the fraction of nodes initially holding the majority state and the fraction of nodes initially holding the minority state. For the network structure, we found that a key effect on the expected convergence time have the spectral properties of some matrices that dictate the contact rates between nodes. For the voting margin, we found that there exist graphs for which the voting margin significantly affects the expected convergence time. Specifically, we found that for some graphs, the expected convergence time goes to infinity as the voting margin approaches 0.

*Complete graph example.* For concreteness, we describe how the voting margin affects the expected convergence time for the complete graph of $n$ nodes. Let us denote with $\alpha > 1/2$ the fraction of nodes that initially held the majority state, and thus $\alpha - (1 - \alpha) = 2\alpha - 1$ is the voting margin. We found that that the convergence time $T$ satisfies

$$\mathbb{E}(T) = \frac{1}{2\alpha - 1} \log(n)(1 + o(1)).$$

Therefore, the expected convergence time is inversely proportional to the voting margin, and thus, goes to infinity as the voting margin goes to 0 (see Fig. 1.1 for an

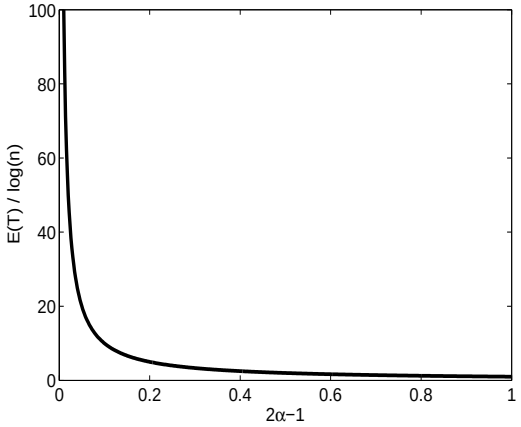


FIG. 1.1. *The expected convergence time vs. the voting margin* $2\alpha - 1$ *for the complete graph of* $n$ *nodes. The curve illustrates how the expected convergence time goes to infinity as the voting margin approaches zero.*

illustration). Hence, albeit the interval consensus guarantees convergence to the correct state, the expected convergence time can assume large values for small voting margins.

*Outline of the Paper.* In Section 2 we discuss the related work. Section 3 introduces the notation and the binary interval consensus algorithm considered in this paper. In Section 4, we consider the complete graph and establish exact mean-field limit as the number of nodes grows large. Section 5 presents our main result that consists of an upper bound on the expected convergence time that applies to arbitrary connected graphs (Theorem 5.2). Section 6 instantiates the upper bound for particular graphs, namely complete graphs, paths, cycles, star-shaped networks and Erdös-Rényi random graphs, and compares with alternative analysis. We conclude in Section 7. Some of the proofs are deferred to the appendix.

**2. Related Work.** In recent years there have been a large body of research on algorithms for decentralized computations over networks, under various constraints on the memory of individual nodes and communication between the nodes. For example, in the so called *quantized consensus* problem [1], the goal is to approximately compute the mean of the values that reside at individual nodes, in a decentralized fashion, where nodes communicate quantized information. In [4], the authors provided bounds on the convergence time in the context of averaging algorithms where nodes exchange quantized information.

The work that is most closely related to ours is [3] where the authors showed that the so called interval consensus algorithm guarantees correctness for arbitrary finite connected graphs. In particular, their work shows that for solving the binary consensus problem, it suffices to use *only two extra states* to guarantee convergence to the correct consensus in a finite time, for every finite connected graph. Our work advances this line of work by establishing the first tight characterizations of the expected convergence time for the binary interval consensus.

Previous work on the binary consensus problem considered algorithms under more stringent assumptions on the number of states stored and communicated by individual nodes. The standard *voter model* is an algorithm where each node stores and

communicates one of two states (0 or 1), where each instigator node switches to the state observed from the contacted node. The voter model has been studied in the context of various graph topologies [9, 10, 11] and the probability of reaching the correct consensus (i.e. corresponding to the initial majority state) is known in a closed-form for arbitrary connected graphs [12]. Specifically, the probability of reaching the correct consensus is proportional to the sum of degrees of the nodes that initially held the initial majority state. In particular, for the complete graphs, this means that the probability of reaching an incorrect consensus is proportional to the number of nodes that initially held the minority state. Moreover, for some network topologies, the convergence time of the voter model is known to be quadratic in the number of nodes, e.g. for a path [13].

In [2], the authors considered a ternary protocol for binary consensus problem where each node stores and communicates an extra state. It was shown that for the complete graph interactions, the probability of reaching the incorrect consensus is exponentially decreasing to 0 as the number of nodes $n$ grows large, with a rate that depends on the voting margin. Moreover, if the algorithm converges to the right consensus, then the time it takes to complete is logarithmic in the number of nodes $n$, and is independent of the voting margin. Notice that this is unlike to the binary interval consensus, for which we found that the expected convergence time, for the complete graph of $n$ nodes, is logarithmic in $n$, but with a factor that is dependent on the voting margin and going to infinity as the voting margin approaches zero. The main advantage of the binary interval consensus algorithm over the ternary protocol is the guaranteed convergence to the correct final state with probability 1, albeit this seems to be at some expense with respect to the convergence time for some graphs.

Finally, we would like to mention that in a bigger picture, our work relates to the cascading phenomena that arise in the context of social networks [14]; for example, in the viral marketing where an initial idea or behavior held by a portion of the population, spreads through the network, yielding a wide adoption across the whole population [15].

**3. Algorithm and Notation.** In this section, we introduce the interval consensus algorithm for the binary consensus problem. Each node is assumed to be in one of the following four states 0, $e_0$, $e_1$ and 1, at every time instant. It is assumed that the states satisfy the following order relations $0 < e_0 < e_1 < 1$. We can interpret the state $e_0$ as referring to the values in the interval $[0, 1/2)$ and $e_1$ as referring to the values in the interval $(1/2, 1]$. Since we assume that initial values held by the nodes are discrete, either 0 or 1, we have that both 0 and $e_0$ indicate value 0 and both $e_1$ and 1 indicate value 1. Without loss of generality, suppose that state 0 is initially held by a majority of nodes.

*State update rules.* The states held by the nodes are updated at pairwise contacts between nodes according to the following state update rules:

1. If a node in state 0 and a node in state 1 get in contact, they *switch* their states to state $e_1$ and state $e_0$, respectively.
2. If a node in state $e_0$ and a node in state 1 get in contact, they *switch* their states to state 1 and state $e_1$, respectively.
3. If a node in state $e_0$ and a node in state 0 get in contact, they *swap* their states to state 0 and $e_1$, respectively.
4. If a node in state $e_1$ and a node in state 1 get in contact, they *swap* their states to state 1 and state $e_1$, respectively.
5. If a node in state $e_0$ and a node in state $e_1$ get in contact, they *swap* their

states to state $e_1$ and state $e_0$, respectively.

6. For any other states of a pair of nodes that get in contact, their states remain unchanged.

*Temporal process of pairwise interactions.* We admit the standard asynchronous communication model [2, 16] where any pair of nodes $(i, j)$ interacts at instances of a Poisson process with rate $q_{i,j} \geq 0$. We denote with $V = \{1, 2, \ldots, n\}$ the set of nodes. The interaction rates are specified by the matrix $Q = (q_{i,j})_{i,j \in V}$ assumed to be symmetric. The transition matrix $Q$ induces an undirected graph $G = (V, E)$ where there is an edge $(i, j) \in E$ if and only if $q_{i,j} > 0$. We assume that graph $G$ is connected.

*Two convergence phases.* The state update rules ensure that in a finite time, a final state is reached at which all nodes are either in state 0 or state $e_0$ (state 0 is initial majority). We distinguish two phases in the convergence to the final state, which will be a key step for our analysis of the expected convergence time that relies on separately analyzing the two phases. The two convergence phases are defined as follows:

Phase 1 (depletion of state 1). This phase begins at the start of the execution of the algorithm and lasts until none of the nodes is in state 1. Whenever a node in state 0 and a node in state 1 get in contact, they switch to states $e_1$ and $e_0$, respectively. It is therefore clear that the number of nodes holding the minority state (state 1) decreases to 0 in a finite amount of time and from that time onwards, the number of nodes in state 0 remains equal to the difference of the (initial) number of nodes in state 0 and state 1.

Phase 2 (depletion of state $e_1$). This phase follows the end of phase 1 and lasts until none of the nodes is in state $e_1$. In this phase, the number of nodes in state $e_1$ decreases following each contact between a node in state $e_1$ and a node in state 0. Since, in this phase, no interaction between a pair of nodes results in increasing the number of nodes in state $e_1$, there are eventually no nodes in state $e_1$.

The duration of each of the two phases is ensured to be finite for arbitrary finite connected graphs by the definition of the state update rules where swapping of the states enables that state 1 nodes get in contact with state 0 nodes and similarly, enables that state $e_1$ nodes get in contact with state 0 nodes.

*Additional notation.* We denote by $S_i(t)$ the set of nodes in state $i \in \{0, e_0, e_1, 1\}$ at time $t$. With a slight abuse of notation, in some cases we will use the compact notation $|S_i| \equiv |S_i(0)|$, $i = 0, 1$, which should be clear from the context. We define $\alpha \in (1/2, 1]$ as the fraction of nodes that initially hold state 0, assumed to be the initial majority. Therefore, $|S_0| = \alpha n$ and $|S_1| = (1 - \alpha)n$.

**4. Mean-field Approximation for Complete Graphs.** In this section we study the dynamics of the algorithm described in Section 3 in the context of the *homogeneous mixing assumption*, i.e. for complete graph interactions. This is a special case that is commonly considered in the literature. The analysis in this section will already yield interesting results that will be later used for comparison with our main result in Section 5, where we establish an upper bound on the expected convergence time for arbitrary connected graphs.

In this case, we will establish that the expected convergence time is asymptotically equal to the expected duration of the first phase, as the number of nodes $n$ tends to infinity. This follows from the analysis of the limit system that we present in the following. The analysis yields the dynamics of the fraction of nodes in each of the

states over time when the number of nodes grows large.

We consider the complete graph of $n$ nodes, where the contact rate matrix is given by $q_{i,j} = 1/(n-1)$, for every $i, j \in V$ and $i \neq i$. From the definition of the node interaction process in Section 3, it is not difficult to note that the system state evolves according to a continuous-time Markov process $(\mathbf{X}(t))_{t\geq 0}$,

$$\mathbf{X}(t) = (|S_0(t)|, |S_{e_0}(t)|, |S_{e_1}(t)|, |S_1(t)|)$$

with the transition rates as follows

$$\mathbf{X} \to \begin{cases} (|S_0|-1, |S_{e_0}|+1, |S_{e_1}|+1, |S_1|-1) & \text{with rate } \frac{|S_0||S_1|}{n-1} \\ (|S_0|, |S_{e_0}|-1, |S_{e_1}|+1, |S_1|) & \text{with rate } \frac{|S_{e_0}||S_1|}{n-1} \\ (|S_0|, |S_{e_0}|+1, |S_{e_1}|-1, |S_1|) & \text{with rate } \frac{|S_0||S_{e_1}|}{n-1}. \end{cases}$$

In the remainder of this section, we will consider the limit dynamics as the number of nodes $n$ grows large.

**4.1. The Limit System.** We consider the limit dynamics as the number of nodes grows large, and to this end, consider the scaled system state $\mathbf{x}^{(n)}(t)$, defined for $t \geq 0$ as follows

$$\mathbf{x}^{(n)}(t) = \left(s_0^{(n)}(t), s_{e_0}^{(n)}(t), s_{e_1}^{(n)}(t), s_1^{(n)}(t)\right) = \frac{1}{n}\left(|S_0(t)|, |S_{e_0}(t)|, |S_{e_1}(t)|, |S_1(t)|\right).$$

Assume that the initial state is such that $\lim_{n\to\infty} \mathbf{x}^{(n)}(0) = \mathbf{x}(0)$, for some fixed initial values $\mathbf{x}(0) = (s_0(0), s_{e_0}(0), s_{e_1}(0), s_1(0))$ and let $\mathbf{x}(t) = (s_0(t), s_{e_0}(t), s_{e_1}(t), s_1(t))$ be the solution of the following system of ordinary differential equations, for $t \geq 0$,

$$\frac{d}{dt}s_0(t) = -s_1(t)s_0(t) \tag{4.1}$$

$$\frac{d}{dt}s_1(t) = -s_0(t)s_1(t) \tag{4.2}$$

$$\frac{d}{dt}s_{e_1}(t) = s_1(t)(1-s_1(t)) - (s_0(t)+s_1(t))s_{e_1}(t) \tag{4.3}$$

with $s_{e_0}(t) = 1 - s_0(t) - s_{e_1}(t) - s_1(t)$, $t \geq 0$.

From Kurtz's convergence theorem [17, Chapter 5], we have that $\mathbf{x}^{(n)}(t)$ converges to $\mathbf{x}(t)$ uniformly over compact intervals, i.e. for every finite $\tau > 0$,

$$\lim_{n\to\infty} \sup_{t\in[0,\tau]} ||\mathbf{x}^{(n)}(t) - \mathbf{x}(t)|| = 0, \text{ with probability 1.}$$

The system of differential equations (4.1)-(4.3) admits a closed-form solution that we present in the following proposition.

PROPOSITION 4.1. *The limit system dynamics, specified by (4.1)-(4.3), for $\alpha \in (1/2, 1]$, admits the following closed-form solution, for $t \geq 0$,*

$$s_0(t) = s_0(0)\frac{2\alpha-1}{s_0(0) - s_1(0)e^{-(2\alpha-1)t}} \tag{4.4}$$

$$s_1(t) = s_1(0)\frac{2\alpha-1}{s_0(0) - s_1(0)e^{-(2\alpha-1)t}}e^{-(2\alpha-1)t} \tag{4.5}$$

$$\begin{aligned} s_{e_1}(t) = {} & s_{e_1}(0)\left(\frac{2\alpha-1}{s_0(0) - s_1(0)e^{-(2\alpha-1)t}}\right)^2 e^{-(2\alpha-1)t} \\ & + s_1(0)\frac{s_0(0)(2\alpha-1)t - s_1(0)2\alpha(1-e^{-(2\alpha-1)t})}{(s_0(0) - s_1(0)e^{-(2\alpha-1)t})^2}e^{-(2\alpha-1)t}. \end{aligned} \tag{4.6}$$

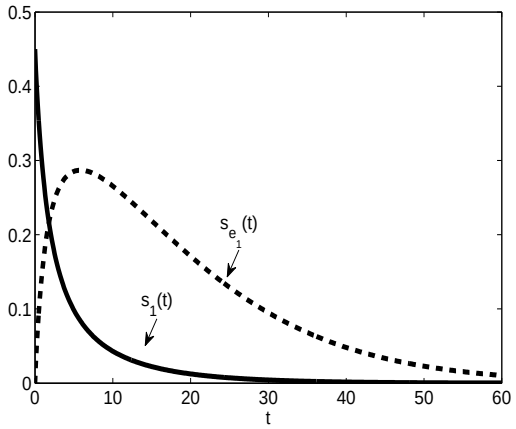


FIG. 4.1. *Depletion of state $e_1$ and state 1 over time, for large complete graphs.*

The proof is provided in Appendix B.

*Numerical example.* We present a numerical example that shows how the states $e_1$ and 1 deplete over time for $\alpha = 0.55$, i.e. the voting margin $2\alpha - 1 = 0.1$, according to the limit system dynamics in Proposition 4.1. Figure 4.1 shows the fraction of nodes in state 1 and state $e_1$ versus time. This numerical example suggests that the expected convergence time is of the same order as the expected time to deplete either state 1 or state $e_1$ node, which we will establish formally in the following theorem.

Let us define $t_{n,\alpha}$ to be the smallest time $t \geq 0$ such that $s_{e_1}(t) + s_1(t) = 1/n$. Notice that $t_{n,\alpha}$ may be interpreted as the time at which all but one node is in either state 0 or state $e_0$. Similarly, let us define $t^x_{n,\alpha}$ as the smallest time $t \geq 0$ at which $s_x(t)$ is equal to $1/n$, for $x = 1$ and $e_1$.

THEOREM 4.2. *The limit dynamics (4.4)-(4.6) with the initial value $(s_0(0), s_1(0)) = (\alpha, 1-\alpha)$, for arbitrarily fixed $\alpha \in (1/2, 1]$, has the following properties:*

1. *The fractions of nodes in state 1 and state $e_1$ satisfy, for large $t$,*

$$s_1(t) = (2\alpha - 1)\frac{1-\alpha}{\alpha} e^{-(2\alpha-1)t}(1 + o(1))$$
$$s_{e_1}(t) = (2\alpha - 1)\frac{1-\alpha}{\alpha} t e^{-(2\alpha-1)t}(1 + o(1)).$$

2. *The fraction of nodes that are either in state 1 or $e_1$ satisfies $s_1(t) + s_{e_1}(t) = s_{e_1}(t)(1 + o(1))$, for large $t$, and*

$$t_{n,\alpha} = \frac{1}{2\alpha - 1}\log(n)(1 + o(1)).$$

3. *The convergence times $t^1_{n,\alpha}$ and $t^{e_1}_{n,\alpha}$ satisfy*

$$t^1_{n,\alpha} = \frac{1}{2\alpha - 1}\log(n) + O(1)$$
$$t^{e_1}_{n,\alpha} = \frac{1}{2\alpha - 1}\left[\log(n) + \log(\log(n))\right] + O(1).$$

The proof is provided in Appendix A.

The theorem tells us that both state 1 and state $e_1$ deplete in time that is logarithmic in the network size $n$. Furthermore, we note that the extra time needed for depletion of state $e_1$, after the depletion of state 1, is small, of the order

$$t^{e_1}_{n,\alpha} - t^1_{n,\alpha} = \frac{1}{2\alpha - 1}\log(\log(n)) + O(1).$$

This tells us that for the complete graph, the expected duration of convergence phase 2 is negligible in comparison with the expected duration of convergence phase 1. This statement will be further supported with empirical results in Figure 6.1.

**5. General Bound for the Expected Convergence Time.** In this section we present our main result that consists of an upper bound on the expected convergence time for arbitrary connected graphs.

The bound is in terms of eigenvalues of a set of matrices $Q_S$ that is defined using the transition matrix $Q$ as follows. Let $S$ be a non-empty subset of the set of vertices $V$ of size smaller than $n$ and let $S^c = V \setminus S$. We consider the matrix $Q_S = (q^S_{i,j})_{i,j\in V}$ that is derived from the contact rate matrix $Q$ as follows

$$q^S_{i,j} = \begin{cases} -\sum_{l\in V} q_{i,l}, & i = j \\ q_{i,j}, & i \in S^c, j \neq i \\ 0, & i \in S, j \neq i. \end{cases} \tag{5.1}$$

We first establish that eigenvalues of the matrices $Q_S$, for $S \subset V$ such that $0 < |S| < n$, are strictly negative. This will be a key property that ensures finiteness of our bound which we present later in this section.

LEMMA 5.1. *For every finite graph $G$, there exists $\delta(Q,\alpha) > 0$ such that for every non-empty subset of vertices $S$, for which it holds $|S| < n$, if $\lambda$ is an eigenvalue of the the matrix $Q_S$, defined in (5.1), then it satisfies*

$$\lambda \leq -\delta(Q,\alpha) < 0.$$

We next present our main result that establishes an upper bound on the expected convergence time that holds for arbitrary connected graphs. Before stating the result, notice that at the end of phase 1 none of the nodes are in state 1, $(2\alpha-1)n$ nodes are in state 0, and the remaining $2(1-\alpha)n$ nodes are in either state $e_0$ or state $e_1$. At the end of phase 2, there are exactly $(2\alpha-1)n$ nodes in state 0 and $2(1-\alpha)n$ nodes in state $e_0$. The following theorem establishes a general bound for the expected duration of each convergence phases in terms of the number of nodes $n$ and the parameter $\delta(Q,\alpha)$, which we introduced in Lemma 5.1.

THEOREM 5.2. *Let $T_1$ be the smallest time at which all the nodes in state 1 are depleted. Then,*

$$\mathbb{E}(T_1) \leq \frac{1}{\delta(Q,\alpha)}(\log n + 1).$$

*Furthermore, letting $T_2$ be the time for all the nodes in state $e_1$ to be depleted, starting from an initial state with no nodes in state 1, we have*

$$\mathbb{E}(T_2) \leq \frac{1}{\delta(Q,\alpha)}(\log n + 1).$$

*In particular, if $T$ is the smallest time at which none of the nodes is in either state $e_1$ or state 1, then*

$$\mathbb{E}(T) \leq \frac{2}{\delta(Q,\alpha)}(\log n + 1).$$

It is worth noting that the above theorem holds for every positive integer $n$ and not just asymptotically in $n$.

The proof of the theorem is presented in Section 5.1 and here we outline the main ideas. The proof proceeds by first separately considering the two convergence phases. For phase 1, we characterize the evolution over time of the probability that a node is in state 1, for every node $i \in V$. This amounts to a "piecewise" linear dynamical system. Similarly, for phase 2, we characterize the evolution over time of the probability that a node is in state $e_1$, for every given node $i \in V$, and show that this also amounts to a "piecewise" linear dynamical system. The proof is then completed by using a spectral bound on the expected number of nodes in state 1, for phase 1, and in state $e_1$, for phase 2, which is then used to establish the asserted results.

*Tightness of the bounds.* The bounds for individual convergence phases, asserted in Theorem 5.2, are tight in the sense that there exist graphs for which the asymptotically dominant terms of the expected convergence time and the corresponding bound are either equal or equal up to a constant factor. These tightness results will follow from the results in Section 6.

### 5.1. Proof of the Main Result.

**5.1.1. Proof of Lemma 5.1.** Let $S$ be a non-empty subset of $V$ of size $xn$, $x \in (0,1)$. Note that every eigenvalue $\lambda$ and the associated eigenvector $\vec{x}$ of the matrix $Q_S$ satisfy the following equations

$$\begin{array}{ll} \lambda x_i = -q_i x_i, & \text{for } i \in S \\ \lambda x_i = -q_i x_i + \sum_{l \in V} q_{i,l} x_l, & \text{for } i \in S^c \end{array} \tag{5.2}$$

where $q_i := \sum_{l \in V} q_{i,l}$, for every $i \in V$.

On the one hand, it is clear from the form of the matrices $Q_S$, given by (5.1), that for every $i \in S$, $\lambda = -q_i$ is an eigenvalue of $Q_S$. Since by assumption, the transition matrix $Q$ induces a connected graph $G$, we have that for every $i \in V$ there exists a $j \in V$ such that $q_{i,j} > 0$. Hence, it follows that $\lambda < 0$.

On the other hand, if $\lambda \neq -q_i$ for any $i \in S$, it is clear from (5.2) that $x_l = 0$ for every $l \in S$. In fact, since $Q$ is symmetric, the remaining eigenvalues of $Q_S$ are the eigenvalues of the symmetric matrix $M_S = (m^S_{i,j})_{i,j \in S^c}$ defined by

$$m^S_{i,j} = \begin{cases} -\sum_{l \in V} q_{i,l}, & i = j \in S^c \\ q_{i,j}, & i,j \in S^c, j \neq i \end{cases}$$

Let $\lambda$ be such an eigenvalue of $Q_S$ and let $\vec{x}$ be the corresponding eigenvector and, without loss of generality, assume that $||\vec{x}||_2^2 = \sum_{i \in V} x_i^2 = \sum_{i \in S^c} x_i^2 = 1$. Since $Q$ is symmetric, it is readily seen that

$$\begin{aligned} \lambda &= -\sum_{i \in S^c, j \in V} q_{i,j} x_i^2 + \sum_{i,j \in S^c} q_{i,j} x_i x_j \\ &= -\sum_{i \in S^c, j \in S} q_{i,j} x_i^2 - \sum_{i,j \in S^c} q_{i,j} x_i (x_i - x_j) \end{aligned}$$

$$= - \sum_{i \in S^c, j \in S} q_{i,j} x_i^2 - \frac{1}{2} \sum_{i,j \in S^c} q_{i,j} (x_i - x_j)^2. \tag{5.3}$$

Therefore, it is clear that $\lambda \leq 0$ with $\lambda = 0$ only if

$$\sum_{i \in S^c, j \in S} q_{i,j} x_i^2 + \frac{1}{2} \sum_{i,j \in S^c} q_{i,j} (x_i - x_j)^2 = 0.$$

Let $W \subset S^c$ be such that $x_i \neq 0$, for $i \in W$, and $x_i = 0$, for $i \in S^c \setminus W$. Since $\vec{x}$ is an eigenvector, then $W$ is non empty. If $\lambda = 0$, then

$$\sum_{i \in W, j \in S} q_{i,j} x_i^2 + \sum_{i \in W, j \in S^c \setminus W} q_{i,j} x_i^2 + \frac{1}{2} \sum_{i,j \in W} q_{i,j} (x_i - x_j)^2 = 0.$$

The above implies that there are no edges between $S$ and $W$, and that there are no edges between $W$ and $S^c \setminus W$, i.e. $W$ is an isolated component, which is a contradiction since $Q$ corresponds to a connected graph. Therefore, $\lambda < 0$.

We showed that for every $S \subset V$ such that $0 < |S| < n$, $Q_S$ has strictly negative eigenvalues. For every finite $n$, there is a finite number of subsets $S$. Therefore, there exists $\delta(Q, \alpha) > 0$ such that for every non-empty set $S \subset V$ of size $xn$, $x \in (0, 1)$, every eigenvalue $\lambda$ of the matrix $Q_S$ satisfies $\lambda \geq -\delta(Q, \alpha) > 0$, which proves the lemma.

**5.1.2. Proof of Theorem 5.2.** We first separately consider the two convergence phases and then complete with a step that applies to both phases.

*Phase 1: Depletion of nodes in state* 1. We describe the dynamics of the first phase through the following indicators of node states. Let $Z_i(t)$ and $A_i(t)$ be the indicators that node $i$ is in state 0 and 1 at time $t$, respectively. The indicator of being in either state $e_0$ or state $e_1$ at time $t$ is encoded by $A_i(t) = Z_i(t) = 0$. The system state evolves according to a continuous-time Markov process $(Z(t), A(t))_{t \geq 0}$, where $A(t) = (A_i(t))_{i \in V}$ and $Z(t) = (Z_i(t))_{i \in V}$, with the transition rates given as follows

$$(Z, A) \to \begin{cases} (Z - e_i, A - e_j) & \text{with rate} \quad q_{i,j} Z_i A_j \\ (Z - e_i + e_j, A) & \text{with rate} \quad q_{i,j} Z_i (1 - A_j - Z_j) \\ (Z, A - e_i + e_j) & \text{with rate} \quad q_{i,j} A_i (1 - A_j - Z_j) \end{cases}$$

where $i, j \in V$ and $e_i$ is the $n$-dimensional vector whose elements are all equal to 0 but the $i$-th element that is equal to 1.

Since $Q$ is a symmetric matrix, we have for every $i \in V$ and $t \geq 0$,

$$\begin{aligned} \frac{d}{dt} \mathbb{E}(A_i(t)) = &- \sum_{j \in V} q_{i,j} \mathbb{E}(A_i(t) Z_j(t)) - \sum_{j \in V} q_{i,j} \mathbb{E}\left(A_i(t)(1 - A_j(t) - Z_j(t))\right) \\ &+ \sum_{j \in V} q_{i,j} \mathbb{E}\left(A_j(t)(1 - A_i(t) - Z_i(t))\right) \end{aligned}$$

or, equivalently,

$$\frac{d}{dt} \mathbb{E}(A_i(t)) = - \left( \sum_{l \in V} q_{i,l} \right) \mathbb{E}(A_i(t)) + \sum_{j \in V} q_{i,j} \mathbb{E}\left(A_j(t)(1 - Z_i(t))\right).$$

Let us now consider the behavior of the set $S_0(t)$ of nodes in state 0, i.e. $S_0(t) = \{i \in V : Z_i(t) = 1\}$. From the above dynamics, we see that there are intervals $[t_k, t_{k+1})$ during which the set $S_0(t)$ does not change (the instants $t_k$ are stopping times of the Markov chain describing the evolution of the algorithm). Let $S_k \subset V$ be the set of nodes in state 0, for $t \in [t_k, t_{k+1})$, and let $S_k^c = V \setminus S_k$, i.e. $S_0(t) = S_k$ and $V \setminus S_0(t) = S_k^c$, for $t \in [t_k, t_{k+1})$. We then can write, for $t \in [t_k, t_{k+1})$,

$$\frac{d}{dt}\mathbb{E}_k(A_i(t)) = -\left(\sum_{l\in V} q_{i,l}\right)\mathbb{E}_k(A_i(t)) + \begin{cases} \sum_{j\in V} q_{i,j}\mathbb{E}_k\left(A_j(t)\right), & i \in S_k^c \\ 0, & i \in S_k \end{cases} \tag{5.4}$$

where $\mathbb{E}_k$ is the expectation conditional on the event $\{S_0(t) = S_k\}$. In a matrix form, this gives

$$\frac{d}{dt}\mathbb{E}_k(A(t)) = Q_{S_k}\mathbb{E}_k(A(t)), \text{ for } t_k \le t < t_{k+1},$$

where $Q_{S_k}$ is given by (5.1).

Solving the above differential equation, we have

$$\mathbb{E}_k(A(t)) = e^{Q_{S_k}(t-t_k)}\mathbb{E}_k(A(t_k)), \text{ for } t_k \le t < t_{k+1}.$$

Using the strong Markov property [18], it is not difficult to see that

$$\mathbb{E}(A(t)) = \mathbb{E}\left[e^{\lambda(t)}A(0)\right], \text{ for } t \ge 0,$$

where

$$\lambda(t) = Q_{S_k}(t - t_k) + \sum_{l=0}^{k-1} Q_{S_l}(t_{l+1} - t_l), \text{ for } t_k \le t < t_{k+1}.$$

Note that $\lambda(t)$ is a random matrix that depends of the stopping times $t_k$.

*Phase 2: Depletion of nodes in state $e_1$.* To describe the dynamics in the second phase, let $B_i(t)$ be the indicator that a node $i \in V$ is in state $e_1$ at time $t$. The notation $Z_i(t)$ has the same meaning as in phase 1, thus $Z_i(t)$ is the indicator that node $i \in V$ is in state 0 at time $t$. The indicator that a node $i \in V$ is in state $e_0$ at time $t$ is encoded by $B_i(t) = Z_i(t) = 0$.

The dynamics in this phase reduces to a continuous-time Markov process $(Z(t), B(t))_{t\ge 0}$, where $Z(t) = (Z_i(t))_{i\in V}$ and $B(t) = (B_i(t))_{i\in V}$, with the transition rates given as follows, for $i, j \in V$,

$$(Z, B) \to \begin{cases} (Z - e_i + e_j, B - e_j) & \text{with rate} \quad q_{i,j}Z_iB_j \\ (Z - e_i + e_j, B) & \text{with rate} \quad q_{i,j}Z_i(1 - B_j - Z_j) \\ (Z, B - e_i + e_j) & \text{with rate} \quad q_{i,j}B_i(1 - B_j - Z_j). \end{cases}$$

From this, we have for every $i \in V$ and $t \ge 0$,

$$\begin{aligned}\frac{d}{dt}\mathbb{E}(B_i(t)) = &-\sum_{i\in V} q_{i,j}\mathbb{E}(B_i(t)Z_j(t)) - \sum_{j\in V} q_{i,j}\mathbb{E}\left(B_i(t)(1 - Z_j(t) - B_j(t))\right) \\ &+ \sum_{j\in V} q_{i,j}\mathbb{E}\left(B_j(t)(1 - Z_i(t) - B_i(t))\right).\end{aligned}$$

Therefore, for every $i \in V$ and $t \geq 0$,

$$\frac{d}{dt}\mathbb{E}(B_i(t)) = -\left(\sum_{l\in V} q_{i,l}\right)\mathbb{E}(B_i(t)) + \sum_{j\in V} q_{i,j}\mathbb{E}\left(B_j(t)(1-Z_i(t))\right).$$

Similar to the first phase, we see that there are intervals $[t'_k, t'_{k+1})$ during which the set $S_0(t)$ does not change (the instants $t'_k$ are stopping times). Let $S'_k$ be such that $S_0(t) = S'_k$, for $t \in [t'_k, t'_{k+1})$. Similarly to the first phase, we have

$$\mathbb{E}(B(t)) = \mathbb{E}\left[e^{\lambda'(t)}B(t'_0)\right], \text{ for } t \geq 0,$$

where $\lambda'(t)$ is a random matrix given by

$$\lambda'(t) = Q_{S'_k}(t - t'_k) + \sum_{l=0}^{k-1} Q_{S'_l}(t'_{l+1} - t'_l), \text{ for } t'_k \leq t < t'_{k+1}.$$

Note that $t'_0 = T_1$ is the instant at which phase 2 starts (phase 1 ends).

*Duration of a phase.* In both phases, the process of interest is of the form

$$\mathbb{E}(Y(t)) = \mathbb{E}\left[e^{\lambda(t)}Y(0)\right], \text{ for } t \geq 0,$$

where for a (random) positive integer $m > 0$ and a sequence $0 = t_0 \leq t_1 \leq \cdots \leq t_m$, we have

$$\lambda(t) = Q_{S_k}(t - t_k) + \sum_{l=0}^{k-1} Q_{S_l}(t_{l+1} - t_l), \text{ for } t_k \leq t < t_{k+1}, \ k = 0, 1, \ldots, m-1.$$

For phase 1, $Y(t) \equiv A(t)$ while for phase 2, $Y(t) \equiv B(t)$. For every $t \geq 0$, we have

$$\begin{aligned}
||\mathbb{E}(Y(t))||_2 \leq \mathbb{E}\left[\left|\left|e^{\lambda(t)}Y(0)\right|\right|_2\right] &\leq \mathbb{E}\left[\left|\left|e^{\lambda(t)}\right|\right| \ ||Y(0)||_2\right] \\
&\leq \mathbb{E}\left[||e^{Q_{S_k}(t-t_k)}|| \prod_{l=0}^{k-1} ||e^{Q_{S_l}(t_{l+1}-t_l)}|| \ ||Y(0)||_2\right] \\
&\leq e^{-\delta(Q,\alpha)t}\mathbb{E}\left(||Y(0)||_2\right) \leq \sqrt{n}\, e^{-\delta(Q,\alpha)t}
\end{aligned}$$

where $||\cdot||$ denotes the matrix norm associated to the Euclidean norm $||\cdot||_2$. In the above, we used Jensen's inequality in the first inequality, followed by the property of matrix norms for the second and third inequalities, then Lemma 5.1 and finally the fact that $Y$ is a $n$-dimensional vector with elements taking values in $\{0, 1\}$.

Furthermore, combining with Cauchy-Schwartz's inequality, we have

$$\sum_{i\in V} \mathbb{E}(Y_i(t)) \leq ||\mathbb{E}(Y(t))||_2 \ ||1||_2 \leq n\, e^{-\delta(Q,\alpha)t}, \text{ for every } t \geq 0.$$

Therefore, we have

$$\mathbb{P}(Y(t) \neq \mathbf{0}) \leq \sum_{i\in V} \mathbb{E}(Y_i(t)) \leq n\, e^{-\delta(Q,\alpha)t}, \quad \text{for every } t \geq 0.$$

Let $T_0$ be the time at which $Y(t)$ hits $\mathbf{0} = (0, \ldots, 0)^T$, which corresponds to $T_1$ for the process $A(t)$ and $T_2$ for the process $B(t)$. Then, we have

$$\begin{aligned}\mathbb{E}(T_0) = \int_0^\infty \mathbb{P}(T_0 > t)dt &= \int_0^\infty \mathbb{P}(Y(t) \neq \mathbf{0})dt \\ &\leq \frac{\log(n)}{\delta(Q,\alpha)} + n \int_{\frac{\log(n)}{\delta(Q,\alpha)}}^\infty e^{-\delta(Q,\alpha)t}dt \\ &= \frac{\log(n)+1}{\delta(Q,\alpha)}\end{aligned}$$

which completes the proof of the theorem.

**6. Application to Particular Graphs.** In this section we instantiate the bound of Theorem 5.2 for particular networks including complete graphs, paths, cycles, star-shaped networks and Erdös-Rényi random graphs. For all these cases, we compare with alternative computations and find that our bound is of exactly the same order as the expected convergence time with respect to the number of nodes. For the complete graph, we also examine the expected convergence time as the voting margin goes to zero.

**6.1. Complete Graphs.** We consider the complete graph of $n > 1$ nodes where each edge $e \in E$ is activated at instances of a Poisson process with rate $1/(n-1)$, i.e. we have $q_{i,j} = 1/(n-1)$ for all $i, j \in V$ such that $i \neq j$.

LEMMA 6.1. *For the complete graph of $n > 1$ nodes and every fixed $\alpha \in (1/2, 1]$, we have*

$$\delta(Q, \alpha) = (2\alpha - 1)\frac{n}{n-1}.$$

*Proof.* For the complete graph, the matrix $Q_S$ is as follows

$$q_{i,j}^S = \begin{cases} -1, & i = j \\ \frac{1}{n-1}, & i \in S^c, j \neq i \\ 0, & i \in S, j \neq i. \end{cases}$$

It is not difficult to see that the vector $\vec{x}$ such that $x_i = 0$ for $i \in S$ and $x_i = 1$ for $i \in S^c$ is an eigenvector of matrix $Q_S$ with the eigenvalue $-\frac{|S|}{n-1}$. Since in each of the two convergence phases, the matrices $Q_{S_k}$ are such that $|S_k| \geq (2\alpha - 1)n$, we have $\frac{|S_k|}{n-1} \geq (2\alpha - 1)\frac{n}{n-1} \geq (2\alpha - 1)$, which establishes the assertion. ∎

Combining the last lemma with Theorem 5.2, we have the following corollary.

COROLLARY 6.2. *For the complete graph of $n > 1$ nodes, the expected duration of phase $i = 1$ and 2 satisfies*

$$\mathbb{E}(T_i) \leq \frac{1}{2\alpha - 1}(\log(n) + 1).$$

In the following, we will show that the latter bound is asymptotically tight, for large number of nodes $n$, for convergence phase 1.

*Comparison with an alternative analysis.* For complete graphs, the convergence time can be studied by an analysis of the underlying stochastic system that we describe in the following.

We first consider the convergence phase 1. Let $0 = \tau_0 \leq \tau_1 \leq \cdots \leq \tau_{|S_1|}$ denote the time instances at which a node in state 0 and a node in state 1 get in contact. Recall that $|S_0|$ and $|S_1|$ denote the initial number of nodes in state 0 and state 1, respectively. It is readily observed that $|S_0(t)| = |S_0| - i$ and $|S_1(t)| = |S_1| - i$, for $\tau_i \leq t < \tau_{i+1}$ and $1 \leq i < |S_1|$.

It is not difficult to observe that $\tau_{i+1} - \tau_i$, $i = 0, 1, \ldots, |S_1| - 1$, is a sequence of independent random variables such that for each $0 \leq i < |S_1|$, $\tau_{i+1} - \tau_i$ is a minimum of a sequence of $(|S_0|-i)(|S_1|-i)$ i.i.d. random variables with exponential distribution with mean $n - 1$. Therefore, the distribution of $\tau_{i+1} - \tau_i$ is exponential with mean $1/\mu_i$, for $0 \leq i < |S_1|$, where $\mu_i = (|S_0| - i)(|S_1| - i)/(n - 1)$. In particular, we have $\mathbb{E}(T_1) = \sum_{i=0}^{|S_1|-1} \mu_i^{-1}$, i.e.

$$\mathbb{E}(T_1) = (n-1) \sum_{i=0}^{|S_1|-1} \frac{1}{(|S_0| - i)(|S_1| - 1)}. \tag{6.1}$$

The above considerations result in the following proposition.

Proposition 6.3. *For every complete graph with $n > 1$ nodes and initial state such that $|S_0| > |S_1| > 0$, the duration of the first convergence phase is a random variable $T_1$ with the following expected value*

$$\mathbb{E}(T_1) = \frac{n-1}{|S_0| - |S_1|} \left( H_{|S_1|} + H_{|S_0|-|S_1|} - H_{|S_0|} \right) \tag{6.2}$$

*where $H_k = \sum_{i=1}^{k} \frac{1}{i}$. Furthermore, for every fixed $\alpha \in (1/2, 1]$, we have*

$$\mathbb{E}(T_1) = \frac{1}{2\alpha - 1} \log(n) + O(1).$$

From the result of the proposition, we observe that the expected duration of the first phase is $\log(n)/\delta(Q, \alpha)$, asymptotically for large $n$, where $\delta(Q, \alpha)$ is given in Lemma 6.1, thus matching the upper bound of Theorem 5.2.

For the prevailing case of a complete graph, we can characterize the expected convergence time of the first phase as $\alpha$ approaches $1/2$, i.e. as the voting margin $2\alpha - 1$ approaches 0 from above. We first consider the limit case where initially there is an equal number of nodes in state 0 and state 1, i.e. $|S_0| = |S_1|$. From (6.1), it is straightforward to note

$$\mathbb{E}(T_1) = \frac{\pi^2}{6} n(1 + o(1)).$$

Therefore, we observe that in case of an initial draw, i.e. equal number of state 0 and state 1 nodes, the expected duration of the first phase scales linearly with the network size $n$. Note that in this case, the second phase starts with nodes in state $e_0$ and state $e_1$ and obviously no majority can follow.

We now discuss the case where $|S_0| - |S_1|$ is strictly positive but small. To this end, let $\mu_n$ denote the voting margin, i.e. $\mu_n = (|S_0| - |S_1|)/n$. From (6.2), is easy to

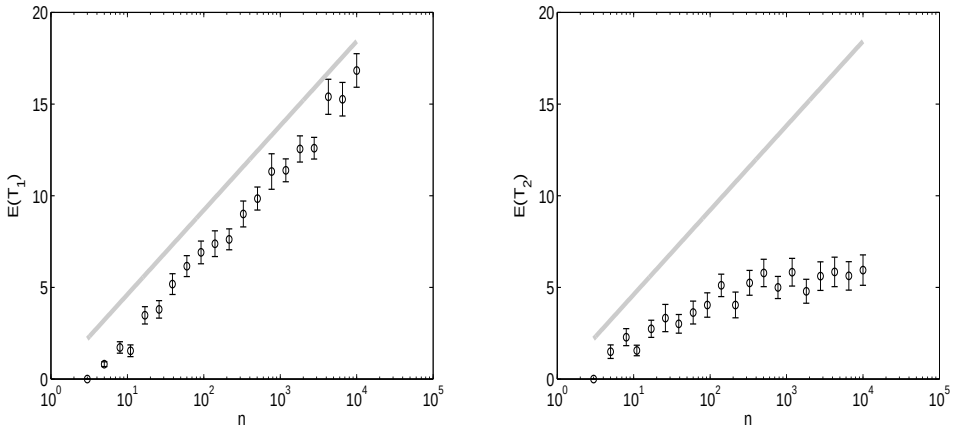


FIG. 6.1. *Complete graph: the expected duration of convergence phases vs. the number of nodes $n$. The initial majority state is held by $\lceil \alpha n \rceil$ nodes, where $\alpha = 3/4$. The solid line is the asymptote $\log(n)/(2\alpha - 1)$ while the bars are simulation results with 95% confidence intervals.*

observe that

$$\mathbb{E}(T_1) = \frac{1}{\mu_n} \log(n\mu_n) + O(1).$$

Therefore, we note that for the voting margin $\mu_n = O(1/n)$, $\mathbb{E}(T_1) = \Theta(n)$ while for $\mu_n > 0$ a fixed constant, we have $\mathbb{E}(T_1) = \Theta(\log(n))$. For the intermediate values of the voting margin, say for $\mu_n = 1/n^a$, for $0 < a < 1$, we have $\mathbb{E}(T_1) = \frac{1-a}{2} n^a \log(n)$.

We conclude the consideration of the convergence of phase 1 by showing that $T_1$ concentrates around $\log(n)/(2\alpha - 1)$ as $n$ grows large, in the following result whose proof is provided in Appendix C.

PROPOSITION 6.4. *For the duration of phase 1 given by $T_1$, we have that for every $\epsilon > 0$, we can find $\beta \in (0, \epsilon)$ such that*

$$T_1 \leq (1 + \epsilon) \frac{1}{2\alpha - 1} \log(n)$$

*with probability at least $1 - 1/n^\beta$, for large enough $n$.*

*Furthermore, for every $\epsilon > 0$,*

$$T_1 \geq (1 - \epsilon) \frac{1}{2\alpha - 1} \log(n)$$

*with probability at least $1 - 1/n$, i.e. with high probability.*

We now turn our attention to the convergence of phase 2. At time $T_1$, there are $|S_0| - |S_1|$ nodes in state 0 and the remaining $n - (|S_0| - |S_1|)$ nodes are either in state $e_0$ or state $e_1$. It is easy to observe that the expected duration of phase 2 is largest for an initial value for phase 2 in which none of the nodes is in state $e_1$, i.e. we have $|S_0| - |S_1|$ nodes in state 0 and the remaining $n - (|S_0| - |S_1|)$ nodes in state $e_1$. We denote with $\tau'_0 \leq \tau'_1 \leq \cdots \leq \tau'_{n-(|S_0|-|S_1|)}$ the time instances at which a node in state 0 and a node in state $e_1$ get in contact resulting in the decrease of the number of nodes

in state $e_1$. Notice that $\tau'_0 = T_1$ and that $\tau'_{n-(|S_0|-|S_1|)} - \tau'_0 = T_2$. Similarly as for phase 1, we have that $\tau'_1 - \tau'_0, \tau'_2 - \tau'_1, \ldots, \tau'_{n-(|S_0|-|S_1|)} - \tau'_{n-(|S_0|-|S_1|)-1}$ is a sequence of independent random variables where the distribution of $\tau'_{i+1} - \tau'_i$ is exponential with parameter $\mu'_i = (|S_0| - |S_1|)(n - |S_0| + |S_1| - i)/(n-1)$. Therefore,

$$\mathbb{E}(T_2) \leq \frac{1}{2\alpha - 1} \log(n) + O(1)$$

which establishes a matching bound to that in Theorem 5.2.

Finally, using similar steps as in the proof of Proposition 6.4, we can show the following.

PROPOSITION 6.5. *For the duration of convergence phase 2, $T_2$, we have that for every $\epsilon > 0$, we can fix $\beta \in (0, \epsilon)$ such that*

$$T_2 \leq (1+\epsilon)\frac{1}{2\alpha - 1} \log(n)$$

*with probability at least $1 - 1/n^\beta$, for large enough $n$.*

Finally, we compare the bound $\log(n)/(2\alpha - 1)$ with simulation results, in Figure 6.1. We observe that the bound is tight for phase 1 and not tight for phase 2, due to our choice of the initial condition in phase 2 . Note also that Figure 6.1 indicates that the expected duration of convergence phase 2 scales as $\Theta(\log(\log n))$, which is consistent with the result in Theorem 4.2.

**6.2. Paths.** We consider a path of $n > 1$ nodes where each edge is activated at instances of a Poisson process of rate 1. Therefore, the contact rate matrix $Q$ is given by $q_{i,i+1} = 1$, for $i = 1, \ldots, n-1$, $q_{i-1,i} = 1$, for $i = 2, 3, \ldots, n$, and all other elements equal to 0.

LEMMA 6.6. *For a path of $n > 1$ nodes, we have, for $\alpha \in (1/2, 1]$,*

$$\begin{aligned}\delta(Q, \alpha) &= 2\left(1 - \cos\left(\frac{\pi}{4(1-\alpha)n + 1}\right)\right) \\ &= \frac{\pi^2}{16(1-\alpha)^2 n^2}(1 + o(1)).\end{aligned}$$

The proof is provided in Appendix D.1.The previous lemma, together with Theorem 5.2, yields the following result.

COROLLARY 6.7. *For a path of $n > 1$ nodes and $\alpha \in (1/2, 1)$, we have for phase $i = 1$ and 2,*

$$\mathbb{E}(T_i) \leq \frac{16(1-\alpha)^2}{\pi^2} n^2 \log(n) + O(1).$$

Finally, we compare the asymptotic bound with simulation results in Figure 6.2. The results indicate that the bound is rather tight for phase 1 and is not tight for phase 2.

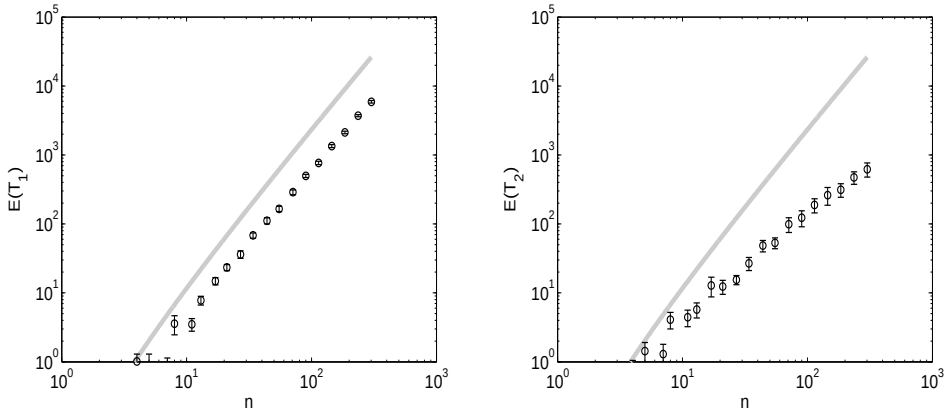


FIG. 6.2. *Path: the expected duration of convergence phases vs. the number of nodes $n$, for $\alpha = 3/4$. The solid line is the asymptote in Corollary 6.7 while the bars are simulation results with 95% confidence intervals.*

**6.3. Cycles.** We consider a cycle of $n > 1$ nodes where each edge is activated at instances of a Poisson process with rate 1. Therefore, the contact rate matrix $Q$ is given by $q_{i,i+1} = 1$, for $i = 1, \ldots, n-1$, $q_{1,n} = q_{n,1} = 1$, and all other elements equal to 0.

LEMMA 6.8. *For a cycle network of $n > 1$ nodes, we have, for $\alpha \in (1/2, 1]$,*

$$\begin{aligned}\delta(Q, \alpha) &= 2\left(1 - \cos\left(\frac{\pi}{2(1-\alpha)n + 1}\right)\right) \\ &= \frac{\pi^2}{4(1-\alpha)^2 n^2}(1 + o(1)).\end{aligned}$$

The proof is provided in Appendix D.2.

COROLLARY 6.9. *For the cycle with $\alpha \in (1/2, 1)$, we have for phase $i = 1$ and 2,*

$$\mathbb{E}(T_i) \leq \frac{4(1-\alpha)^2}{\pi^2} n^2 \log(n) + O(1).$$

Finally, we compare the last bound with simulation results in Figure 6.3. Similar as in other cases, we observe that the bound has the same scaling with the number of nodes as the expected duration of convergence phase 1, and is not tight for convergence phase 2.

**6.4. Star-Shaped Networks.** We consider a star-shaped network that consists of a hub node and $n - 1$ leaf nodes. Without loss of generality, let the hub node be node 1 and let $i = 2, 3, \ldots, n$ be the leaf nodes. The contacts between a leaf node and the hub are assumed to occur at instances of a Poisson process of rate $1/(n-1)$.

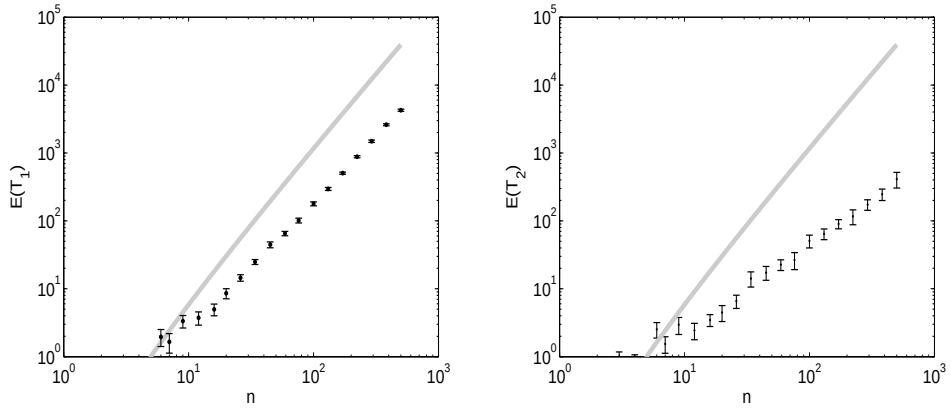


FIG. 6.3. *Cycle: the expected durations of convergence phases vs. the number of the nodes $n$. The initial state is such that state 0 is held by a set of $\lceil \alpha n \rceil$ consecutive nodes along the cycle with $\alpha = 3/4$.*

This setting is motivated in practice by networks where a designated node assumes the role of an information aggregator to which other nodes are connected and this aggregator node has access capacity of rate 1. The elements of matrix $Q$ are given by $q_{1,i} = q_{i,1} = 1/(n-1)$, for $i = 2, 3, \ldots, n$ and other elements equal to 0.

We have the following lemma for the star-shaped network that we defined above.

LEMMA 6.10. *For the star network of $n > 1$ nodes, we have*

$$\begin{aligned}\delta(Q,\alpha) &= \frac{n}{2(n-1)}\left(1 - \sqrt{1 - \frac{4(2\alpha-1)}{n}}\right) \\ &\geq \frac{2\alpha - 1}{n}\end{aligned}$$

*where the inequality is tight for large $n$.*

The proof is provided in Appendix D.3. The previous lemma yields the following corollary.

COROLLARY 6.11. *For the star network with $n > 1$ nodes and every fixed $\alpha \in (1/2, 1]$, the expected duration of phase $i = 1$ and 2 satisfies*

$$\mathbb{E}(T_i) \leq \frac{1}{2\alpha - 1} n(\log(n) + 1).$$

*Comparison with an alternative analysis for phase 1.* For the star-shaped network of $n$ nodes, we can compute the exact asymptotically dominant term of the expected duration of phase 1, for large $n$, which is presented in the following proposition.

PROPOSITION 6.12. *For the star-shaped network of $n$ nodes, the expected time to deplete nodes in state 1 satisfies*

$$\mathbb{E}(T_1) = \frac{1}{(2\alpha - 1)(3 - 2\alpha)} n \log(n) + O(n). \tag{6.3}$$

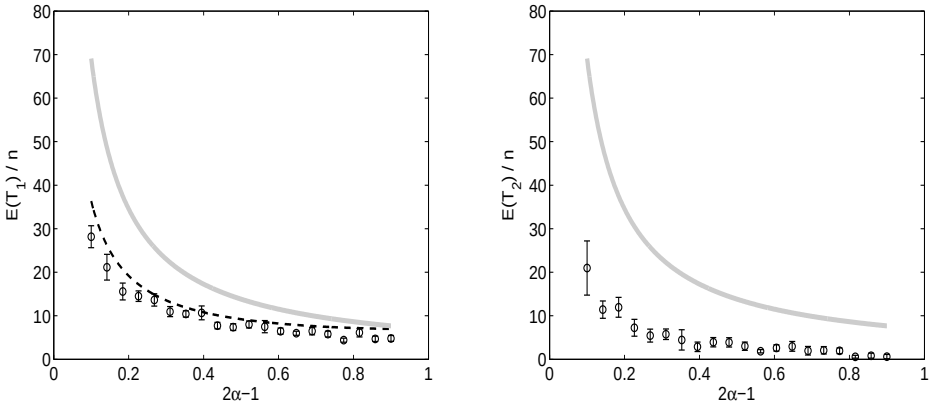


FIG. 6.4. *Star-shaped network: expected duration of convergence phases versus the voting margin* $2\alpha - 1$*, for* $n = 1000$*. The solid curves indicate* $\log(n)/(2\alpha - 1)$*; the dashed line indicates* $\log(n)/[(2\alpha - 1)(3 - 2\alpha)]$*; the bars indicate* 95%*-confidence intervals of estimates obtained by simulations.*

The proof is provided in Appendix E.

Notice that the dominant term in Proposition 6.12 is smaller than the upper bound in Corollary 6.11 for the factor $1/(3 - 2\alpha)$.

*Remark.* We only consider the expected convergence time for phase 1. Similar analysis could be pursued for phase 2 but is more complicated, because the lumping of the states as done in the proof for phase 1 cannot be made.

Finally, we compare our bound with simulation results in Figure 6.4. The results indicate that the bound of Corollary 6.11 is not tight. We also observe that the asymptote in Proposition 6.12 conforms well with simulation results.

**6.5. Erdös-Rényi Random Graphs.** We consider random graphs for which the matrix of contact rates $Q$ is defined as follows. Given a parameter $p_n \in (0, 1)$ that corresponds to the probability that a pair of nodes interact with a strictly positive rate, we define the contact rate of a pair of nodes $i, j \in V$, $i \neq j$, as follows

$$q_{i,j} = \frac{1}{np_n} X_{i,j}$$

where $X_{i,j}$ is a sequence of i.i.d. random variables such that $\mathbb{P}(X_{i,j} = 1) = 1 - \mathbb{P}(X_{i,j} = 0) = p_n$, for every $i, j \in V$, $j \neq i$. The rates are normalized with the factor $1/(np_n)$, so that for each node, the interaction rate with other nodes is 1, asymptotically for large $n$.

Furthermore, we assume that $p_n$ is chosen such that, for a constant $c > 1$,

$$p_n = c\frac{\log(n)}{n}$$

which ensures that the induced random graph is connected with high probability.

We have the following lemma.

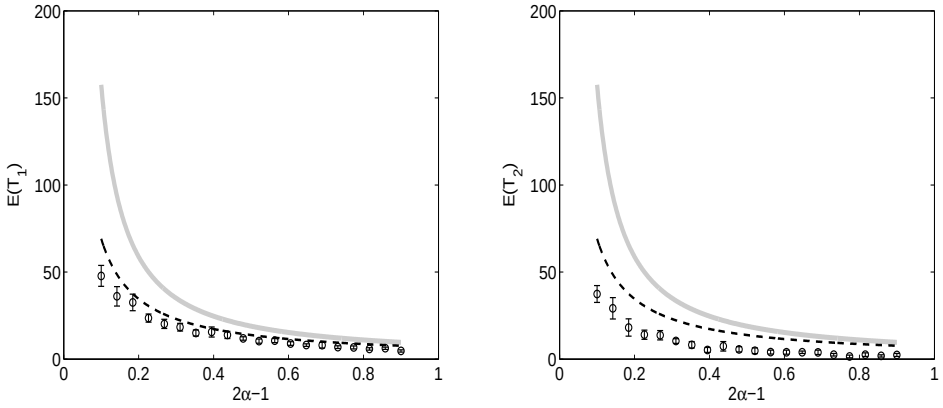


FIG. 6.5. *Erdös-Rényi random graphs: the expected duration of convergence phases vs. the voting margin* $2\alpha-1$*, for* $n = 1000$ *and* $c = 100$*. The solid curves indicate the bound of Corollary 6.5; the dashed lines indicate* $\log(n)/(2\alpha-1)$*; the bars indicate* 95%*-confidence estimates.*

LEMMA 6.13. *Suppose* $c > \frac{2}{2\alpha-1}$ *and* $\alpha \in (1/2, 1]$*. We then have*

$$\frac{1}{\delta(Q,\alpha)} \leq \frac{1}{(2\alpha-1)\varphi^{-1}\left(\frac{2}{c(2\alpha-1)}\right)} + o(1), \text{ with high probability,} \tag{6.4}$$

*where* $\varphi^{-1}(\cdot)$ *is the inverse function of* $\varphi(x) := x\log(x) + 1 - x$*, for* $x \in [0, 1]$*.*

The proof is provided in Appendix D.4.

From the last lemma and Theorem 5.2, we have the following corollary.

COROLLARY 6.14. *Under* $c > \frac{2}{2\alpha-1}$ *and* $\alpha \in (1/2, 1]$*, we have for the duration of phase* $i = 1$ *and* 2*,*

$$T_i \leq \frac{1}{(2\alpha-1)\varphi^{-1}\left(\frac{2}{c(2\alpha-1)}\right)}\log(n) + O(1) \tag{6.5}$$

*with high probability.*

*Remark.* We note the following intuitive observation: the asserted bound for the expected convergence time for each of the phases boils down to that of the complete graph, for large expected degree of a node, i.e. large $c$. Indeed, this holds because for every fixed $\alpha \in (1/2, 1]$, the term $\varphi^{-1}(\frac{2}{c(2\alpha-1)})$ goes to 1 as $c$ grows large.

Finally, we compare the bound of Corollary 6.14 with estimates obtained by simulations in Fig. 6.5. The results confirm that the bound is indeed a bound and that it is not tight, which is because of a bounding technique that we used in the proof.

**7. Conclusion.** We established an upper bound on the expected convergence time of the binary interval consensus that applies to arbitrary connected graphs. We showed that for a range of particular graphs, the bound is of exactly the same order as the expected convergence time with respect to the network size. The bound provides insights into how the network topology and the voting margin affect the expected convergence time. In particular, we showed that there exist network graphs for which the expected convergence time becomes much larger when the voting margin approaches zero. The established bound provides a unifying approach to bound the expected convergence time of binary interval consensus on arbitrary finite and connected graphs.

An important direction of future work is to consider lower bounds on the convergence time. In particular, it would be of interest to better understand how to fine tune the interaction parameters $q_{ij}$ to achieve the best possible convergence time for a given connected graph under given memory and communication constraints.

REFERENCES


[1] A. Kashyap, T. Basar, and R. Srikant. Quantized consensus. *Automatica*, 43:1192–1203, 2007.

[2] E. Perron, D. Vasudevan, and M. Vojnović. Using Three States for Binary Consensus on Complete Graphs. In *Proc. of IEEE Infocom 2009*, April 2009.

[3] F. Bénézit, P. Thiran, and M. Vetterli. Interval Consensus: From Quantized Gossip to Voting. In *Proc. of IEEE ICASP (Int'l Conference on Acoustics, Speech, and Signal Processing)*, pages 3661–3664, 2009.

[4] A. Nedic, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis. Distributed Averaging Algorithms and Quantization Effects. In *47th IEEE Conference on Decision and Control*, pages 4825–4830, 2008.

[5] Y. Ruan and Y. Mostofi. Binary consensus with soft information processing in cooperative networks. In *Proc. of the 47th IEEE Conference on Decision and Control*, pages 3613–3619, 2008.

[6] B. Hardekopf, K. Kwiat, and S. Upadhyaya. Secure and fault-tolerant voting in distributed systems. In *IEEE Aerospace Conference*, pages 3/1117–3/1126, 2001.

[7] A. Mostéfaoui, M. Raynal, and F. Tronel. From Binary Consensus to Multivalued Consensus in asynchronous message-passing systems. *Information Processing Letters*, 73(5–6):207–212, 2000.

[8] J. Zhanga and W. Chenb. Bounded cost algorithms for multivalued consensus using binary consensus instances. *Information Processing Letters*, 109(17):1005–1009, 2009.

[9] P. Donnelly and D. Welsh. Finite particle systems and infection models. *Math. Proc. Camb. Phil. Soc.*, 94:167–182, 1993.

[10] T. M. Liggett. *Interacting particle systems*. Springer, 2 edition, 2006.

[11] V. Sood and S. Redner. Voter model on heterogenous graphs. *Physical Review Letters*, 94(17):178701, 2005.

[12] Y. Hassin and D. Peleg. Distributed probabilistic polling and applications to proportionate agreement. *Information and Computation*, 171(2):248–268, 2002.

[13] D. Aldous and J.A. Fill. *Reversible Markov Chains and Random Walks on Graphs*. in preparation, (http://www.stat.berkeley.edu/ aldous/RWG/book.html), 1999.

[14] J. Kleinberg. Cascading behavior in networks: Algorithmic and economic issues. In *In Algorithmic Game Theory (N. Nisan, T. Roughgarden, E. Tardos, V. Vazirani). Cambridge University Press*, 2007.

[15] H. Amini, M. Draief, and M. Lelarge. Marketing in Random Networks. In *Proc. of NetCoop '08 (Int'l Workshop on Network Control and Optimization), LNCS 5425*, pages 17–25, 2009.

[16] S. Boyd, A. Ghosh, and and D. Shah B. Prabhakar. Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52(6):2508–2530, 2006.

[17] M. Draief and L. Massoulié. *Epidemics and Rumours in Complex Networks*. Cambridge University Press, 2010.

[18] P. Brémaud. *Markov Chains: Gibbs Fields, Monte Carlo simulations and Queues*. Springer, 1999.

[19] W.-C. Yueh. Eigenvalues of Several Tridiagonal Matrices. *Applied Mathematics E-Notes*, 5:66–74, 2005.

**Appendix A. Proof of Theorem 4.2.** We prove separately each item of the theorem as follows:

*Item 1.* The large time asymptotes follow readily from (4.5) and (4.6) under the assumed initial state.

*Item 2.* By definition, $s_1(t_{n,\alpha})+s_{e_1}(t_{n,\alpha}) = 1/n$ and from item 1, $s_1(t)+s_{e_1}(t) = (2\alpha-1)\frac{1-\alpha}{\alpha}te^{-(2\alpha-1)t}(1+o(1))$. Combining the latter two equations, the asserted asymptote follows.

*Item 3.* . First, the asymptote for $t^1_{n,\alpha}$ follows from $s_1(t^1_{n,\alpha}) = 1/n$ and $s_1(t) = (2\alpha-1)\frac{1-\alpha}{\alpha}e^{-(2\alpha-1)t}(1+o(1))$ (item 1). Second, the asymptote for $t^{e_1}_{n,\alpha}$ is established as follows. By definition, $s_{e_1}(t^{e_1}_{n,\alpha}) = 1/n$ and $s_{e_1}(t) = (2\alpha-1)\frac{1-\alpha}{\alpha}te^{-(2\alpha-1)t}(1+o(1))$, we have

$$(2\alpha-1)t^{e_1}_{n,\alpha} - \log(t^{e_1}_{n,\alpha}) = \log(n) + O(1). \tag{A.1}$$

Let $x_n$ be such that $t^{e_1}_{n,\alpha} = \frac{1}{2\alpha-1}(\log(n)+x_n)$. Then, from (A.1),

$$x_n - \log\left(1+\frac{x_n}{\log(n)}\right) = \log(\log(n)) + O(1).$$

From this, it is readily seen that $x_n = \log(\log(n)) + O(1)$, which completes the proof.

**Appendix B. Proof of Proposition 4.1.** For brevity, we will denote the voting margin by $\gamma$, i.e. $\gamma = 2\alpha - 1$. From (4.1)-(4.2), notice that $s_0(t)$ and $s_1(t)$ evolve autonomously of $s_{e_0}(t)$ and $s_{e_1}(t)$, for every $t \geq 0$. Furthermore,

$$\frac{d}{dt}(s_0(t) - s_1(t)) = 0, \quad \text{for every } t \geq 0.$$

Therefore, $s_0(t) - s_1(t) = s_0(0) - s_1(0) = \gamma$, for every $t \geq 0$. Combining with (4.1), we obtain

$$\frac{d}{dt}s_0(t) = (\gamma - s_0(t))s_0(t), \ t \geq 0.$$

This yields (4.4) and (4.5). Note that (4.3) is a linear differential equation with time inhomogeneous parameters given as follows

$$\frac{d}{dt}s_{e_1}(t) + a(t)s_{e_1}(t) = b(t) \tag{B.1}$$

where $a(t) = s_0(t) + s_1(t)$ and $b(t) = s_1(t)(1-s_1(t))$, for $t \geq 0$.

From (4.4) and (4.5), we have

$$a(t) = \gamma\frac{s_0(0)+s_1(0)e^{-\gamma t}}{s_0(0)-s_1(0)e^{-\gamma t}} \text{ and } b(t) = s_0(0)\gamma\frac{s_0(0)-s_1(0)(1+\gamma)e^{-\gamma t}}{(s_0(0)-s_1(0)e^{-\gamma t})^2}e^{-\gamma t}.$$

Now, it is well known that the solution to the linear differential equation (B.1) is

$$s_{e_1}(t) = s_{e_1}(0)e^{-\int_0^t a(s)ds} + \int_0^t b(s)e^{-\int_s^t a(x)dx}ds. \tag{B.2}$$

Following some elementary integrations, we have

$$\exp\left(-\int_0^t a(s)ds\right) = \left(\frac{\gamma}{s_0(0) - s_1(0)e^{-\gamma t}}\right)^2 e^{-\gamma t}. \tag{B.3}$$

Using the latter identity,

$$\int_0^t b(s)e^{-\int_s^t a(x)dx}ds = \frac{e^{-\gamma t}}{(s_0(0) - s_1(0)e^{-\gamma t})^2}\phi(t) \tag{B.4}$$

where

$$\begin{aligned}\phi(t) &= \int_0^t b(s)(s_0(0) - s_1(0)(1+\gamma)e^{-\gamma s})^2 e^{\gamma s}ds \\ &= s_1(0)[s_0(0)\gamma t - s_1(0)(1+\gamma)(1-e^{-\gamma t})].\end{aligned}$$

Combined with (B.4), (B.3), and (B.2), we obtain (4.6), which completes the proof.

**Appendix C. Proof of Proposition 6.4.** We first note that

$$T_1 = \sum_{i=0}^{|S_1|-1} \sigma_i$$

where $\sigma_0, \sigma_1, \ldots, \sigma_{|S_1|-1}$ is a sequence of independent random variables and the distribution of $\sigma_i$ is exponential with parameter $\mu_i$ given as follows

$$\mu_i = \frac{(|S_0| - i)(|S_1| - i)}{n-1},\ i = 0, 1, \ldots, |S_1| - 1.$$

We will use the inequality stated in the following lemma. Recall that $|S_0| = \alpha n$ and $|S_1| = (1-\alpha)n$ for $\alpha \in (1/2, 1]$.

LEMMA C.1. *For every real value $s$ such that $-(2\alpha-1)^2 n < s < 2\alpha - 1$ and $s/(2\alpha-1)$ a non-positive integer, we have*

$$\mathbb{E}(e^{sT_1}) \le K_\alpha(s) n^{\frac{s}{2\alpha-1}}(1 + o(1))$$

*where*

$$K_\alpha(s) = \Gamma\left(1 - \frac{s}{2\alpha-1}\right)\left(\frac{1-\alpha}{\alpha}(2\alpha-1)\right)^{\frac{s}{2\alpha-1}}.$$

*Proof.* Since $\tau_0, \tau_1, \ldots, \tau_{|S_1|-1}$ is a sequence of independent random variables and $\tau_i$ is an exponential random variable with parameter $\mu_i$, we have, $\mathbb{E}(e^{s\tau_i}) = \frac{\mu_i}{\mu_i - s}$, for every $s < \mu_i$. Therefore, it is readily observed that

$$\mathbb{E}(e^{sT_1}) = \prod_{i=0}^{|S_1|-1}\left(1 - \frac{s(n-1)}{(|S_0| - i)(|S_1| - i)}\right)^{-1}, \quad \text{for every } s < 2\alpha - 1. \tag{C.1}$$

Now, we have for every $i = 0, 1, \ldots, |S_1| - 1$,

$$\left(1 - \frac{s(n-1)}{(|S_0| - i)(|S_1| - i)}\right)^{-1} \le \left(1 - \frac{sn}{(|S_0| - i)(|S_1| - i)}\right)^{-1}$$

$$
\begin{aligned}
&= \left(1 - \frac{sn/(|S_0| - |S_1|)}{|S_1| - i} + \frac{sn/(|S_0| - |S_1|)}{|S_0| - i}\right)^{-1} \\
&\leq \left(1 - \frac{sn/(|S_0| - |S_1|)}{|S_1| - i}\right)^{-1} \left(1 + \frac{sn/(|S_0| - |S_1|)}{|S_0| - i}\right)^{-1} \\
&= \left(1 - \frac{s/(2\alpha-1)}{(1-\alpha)n - i}\right)^{-1} \left(1 + \frac{s/(2\alpha-1)}{\alpha n - i}\right)^{-1}
\end{aligned}
$$

where the-right side is finite and positive under the assumption of the lemma $-(2\alpha - 1)^2 n < s < 2\alpha - 1$. Plugging the last above inequality in (C.1), we establish

$$
\text{(C.2)}\quad \mathbb{E}(e^{sT_1}) \leq \prod_{i=1}^{(1-\alpha)n} \left(1 - \frac{s/(2\alpha-1)}{i}\right)^{-1} \prod_{i=(2\alpha-1)n+1}^{\alpha n} \left(1 + \frac{s/(2\alpha-1)}{i}\right)^{-1}.
$$

Now, use Euler's formula $\Gamma(1+z) = z\Gamma(z) = \lim_{n\to\infty} n^z \prod_{i=1}^{n} \left(1 + \frac{z}{i}\right)^{-1}$, for $z$ a real number except for negative integers. Then, we note that for all real values $s$ except for $s/(2\alpha - 1)$ a positive negative integer

$$
\text{(C.3)}\quad \prod_{i=1}^{(1-\alpha)n} \left(1 - \frac{s/(2\alpha-1)}{i}\right)^{-1} = \Gamma\left(1 - \frac{s}{2\alpha-1}\right) [(1-\alpha)n]^{\frac{s}{2\alpha-1}} (1 + o(1))
$$

and for all real values $s$ except for $s/(2\alpha - 1)$ a negative integer, we have

$$
\begin{aligned}
\prod_{i=(2\alpha-1)n+1}^{\alpha n} \left(1 + \frac{s/(2\alpha-1)}{i}\right)^{-1} &= \frac{\prod_{i=1}^{\alpha n} \left(1 + \frac{s/(2\alpha-1)}{i}\right)^{-1}}{\prod_{i=1}^{(2\alpha-1)n} \left(1 + \frac{s/(2\alpha-1)}{i}\right)^{-1}} \\
&= \frac{[\alpha n]^{-\frac{s}{2\alpha-1}} \Gamma\left(1 + \frac{s}{2\alpha-1}\right)(1 + o(1))}{[(2\alpha-1)n]^{-\frac{s}{2\alpha-1}} \Gamma\left(1 + \frac{s}{2\alpha-1}\right)(1 + o(1))} \\
&= \left(\frac{2\alpha-1}{\alpha}\right)^{\frac{s}{2\alpha-1}} (1 + o(1)). \qquad \text{(C.4)}
\end{aligned}
$$

Combing (C.2) with (C.3) and (C.4), the assertion of the lemma follows. □

In remainder of the proof, we separately consider the upper and lower bounds asserted in the theorem.

*Lower bound.* By Chernoff's bound, we have for every $\theta > 0$

$$
\mathbb{P}\left(T_1 < (1-\epsilon)\frac{1}{2\alpha-1}\log(n)\right) \leq n^{\frac{(1-\epsilon)\theta}{2\alpha-1}} \mathbb{E}(e^{-\theta T_1}).
$$

Combining with Lemma C.1, for every $0 < \theta < (2\alpha-1)^2 n$ and $\theta/(2\alpha-1)$ not a positive integer,

$$
\mathbb{P}\left(T_1 < (1-\epsilon)\frac{1}{2\alpha-1}\log(n)\right) \leq K_\alpha(-\theta) n^{-\frac{\epsilon\theta}{2\alpha-1}} (1 + o(1)).
$$

Therefore, taking $\theta \geq (2\alpha-1)/\epsilon$, we have $\mathbb{P}\left(T_1 > (1+\epsilon)\frac{1}{2\alpha-1}\log(n)\right) \leq 1/n$, for large enough $n$, which completes the proof for the lower bound.

*Upper bound.* Again, by Chernoff's bound, for every $\theta > 0$,

$$\mathbb{P}\left(T_1 > (1+\epsilon)\frac{1}{2\alpha-1}\log(n)\right) \leq n^{-\frac{(1+\epsilon)\theta}{2\alpha-1}}\mathbb{E}(e^{\theta T_1}).$$

Combining with Lemma C.1, we observe that for every $0 < \theta < 2\alpha - 1$ and $\theta/(2\alpha-1)$ not a positive integer,

$$\mathbb{P}\left(T_1 > (1+\epsilon)\frac{1}{2\alpha-1}\log(n)\right) \leq K_\alpha(\theta)n^{-\frac{\epsilon\theta}{2\alpha-1}}(1+o(1)).$$

Therefore, provided that $c < \epsilon$, $\mathbb{P}\left(T_1 > (1+\epsilon)\frac{1}{2\alpha-1}\log(n)\right) \leq 1/n^c$, for large enough $n$. This completes the proof.

## Appendix D. Characterization of $\delta$ for Particular Graphs.

**D.1. Proof of Lemma 6.6 (Path).** Recall that the matrix $Q$ is the tridiagonal matrix with $q_{i,i+1} = 1$, for $i = 1, 2, \ldots, n-1$, $q_{i,i-1} = 1$, for $i = 2, 3, \ldots, n$, and all other elements equal to 0.

We will separately consider four cases depending on whether the respective end-node 1 and $n$ is in $S$ or $S^c$. In the following, we denote with $\xi_A(\lambda)$, the characteristic polynomial of a matrix $A$.

*Case 1*: $1 \in S$ and $n \in S^c$. In this case, we repeatedly expand the matrix $\lambda I - Q^S$ along the rows $i \in S$, using Laplace's formula, to obtain

$$\xi_{Q^S}(\lambda) = (\lambda+1)(\lambda+2)^{|S|-1}\xi_B(\lambda) \tag{D.1}$$

where the matrix $B$ is a block-diagonal matrix with blocks $B_1$, $B_2$, ..., $B_b$, for $1 < b \leq |S|$ that are symmetric tridiagonal matrices of the form

$$\begin{pmatrix} -2+c_1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2+c_2 \end{pmatrix} \tag{D.2}$$

where $c_1 = c_2 = 0$, for $i = 1, 2, \ldots, b-1$ (type 1), and $c_1 = 0$ and $c_2 = 1$, for $i = b$ (type 2).

Since $B$ is a block-diagonal matrix, notice that $\xi_B(\lambda) = \prod_{i=1}^b \xi_{B_i}(\lambda)$. Hence, together with (D.1), we have that $Q^S$ the largest eigenvalue of a block that is either of type 1 or type 2.

The eigenvalues of tridiagonal matrices of the form (D.2) are well known for some values of the parameters $c_1$ and $c_2$, see e.g. [19]. In particular, for a $m \times m$ tridiagonal matrix of type 1, i.e. for $c_1 = c_2 = 0$, we have eigenvalues

$$\lambda_k = -2\left(1 - \cos\left(\frac{\pi k}{m+1}\right)\right),\ k = 1, 2, \ldots, m. \tag{D.3}$$

For a $m \times m$ tridiagonal matrix of type 2, i.e. for $c_1 = 0$ and $c_2 = 1$, we have eigenvalues

$$\kappa_k = -2\left(1 - \cos\left(\frac{(2k-1)\pi}{2m+1}\right)\right),\ k = 1, 2, \ldots, m.$$

It is readily checked that the largest eigenvalue is $\kappa_1$ with $m = |S^c|$. This corresponds to the case where the nodes in the set $S$ are $1, 2, \ldots, |S|$ (i.e. form a cluster).

*Case 2*: $1 \in S^c$ and $n \in S$. In this case, we have that (D.1) holds but the blocks of matrix $B$ redefined so that $c_1 = 1$ and $c_2 = 0$, for $B_1$, and $c_1 = c_2 = 0$, for $B_2$, $B_3$, ..., and $B_b$. It is readily observed that the eigenvalues of the block matrix $B_1$ are the same as for $c_1 = 0$ and $c_2 = 1$ (type 1 tridiagonal matrix in Case 1). Hence, the largest eigenvalue is same as under Case 1. Notice that in this case, it corresponds to taking nodes $n - |S| + 1$, $n - |S| + 2$, ..., and $n$ to be in the set $S$.

*Case 3*: $1 \in S^c$ and $n \in S^c$. In this case, by the same arguments as in Case 1, we have

$$\xi_{Q^S}(\lambda) = (\lambda + 2)^{|S|}\xi_B(\lambda)$$

where $B$ is a block-diagonal matrix with blocks $B_1$, $B_2$, ..., $B_b$, $1 < b \leq |S^c|$, which are of the form (D.2) such that $c_1 = 1$ and $c_2 = 0$ for $B_1$, $c_1 = c_2 = 0$, for $B_2$, $B_3$, ..., and $B_b$, and $c_1 = 0$ and $c_2 = 1$, for $B_b$. In this case, the largest eigenvalue is $\kappa_1$ with $m = |S^c| - 1$.

*Case 4*: $1 \in S$ and $n \in S$. In this case, we have

$$\xi_{Q^S}(\lambda) = (\lambda + 1)^2(\lambda + 2)^{|S|-2}\xi_B(\lambda)$$

where $B$ is a block-diagonal matrix with blocks $B_1$, $B_2$, ..., $B_b$, $1 < b \leq |S^c|$ that are all of the form (D.2) with $c_1 = c_2 = 0$. In this case, the largest eigenvalue is $\lambda_1$ with $m = |S^c|$.

Finally, we observe that for each of the four cases, since $|S^c| \leq 2(1 - \alpha)n$, we can take $\delta(Q, \alpha)$ as asserted in the lemma, which completes the proof.

**D.2. Proof of Lemma 6.8 (Cycle).** The proof is similar to that for a path in Section D.1. We will see that for the cycle, we will deal with blocks of tridiagonal matrices of the form (D.2) with $c_1 = c_2 = 0$. Recall that for the cycle, we have matrix $Q$ such that $q_{i,i+1} = 1$, for $i = 1, 2, \ldots, n-1$, $q_{i-1,i} = 1$, for $i = 2, 3, \ldots, n$, $q_{1,n} = q_{n,1} = 1$, and all other elements equal to 0.

Again, we separately consider the following four cases.

*Case 1*: $1 \in S$ and $n \in S^c$. By successive expansion along rows $i \in S$, we obtain

$$\xi_{Q^S}(\lambda) = (\lambda + 2)^{|S|}\xi_B(\lambda) \tag{D.4}$$

where $B$ is a block-diagonal matrix with blocks of the form (D.2) with $c_1 = c_2 = 0$ (referred to as type 1). Since $B$ is a block-diagonal matrix, we have $\xi_B(\lambda) = \prod_{i=1}^{b} \xi_{B_i}(\lambda)$, $1 < b \leq |S^c|$, where $B_i$ is a matrix of type 1. The largest eigenvalue is $\lambda_1$, given in (D.3), for $m = |S^c|$.

*Case 2*: $1 \in S^c$ and $n \in S$. In this case, the same arguments hold as in Case 1.

*Case 3*: $1 \in S^c$ and $n \in S^c$. In this case, (D.4) holds, with matrix $B$ with diagonal blocks and other elements as in Case 1, except that $b_{1,|S^c|} = b_{|S^c|,1} = 1$. This matrix

can be transformed into a block-diagonal matrix of the same form as in Case 1 by permuting the rows of the matrix, hence, it has the same spectral properties as the matrix $B$ under Case 1. Specifically, this can be done by moving the block of rows that correspond to $B_1$ to the bottom of the matrix $B$.

*Case 4*: $1 \in S$ and $n \in S$. In this case, the same arguments apply as in Case 1.

Finally, we observe that for each of the four cases, since $|S^c| \leq 2(1-\alpha)$, we can take $\delta(Q,\alpha)$ as asserted in the lemma, which completes the proof.

**D.3. Proof of Lemma 6.10 (Star).** We separately consider the two cases for which either the hub is in the set $S$ or not.

*Case 1.* Suppose that the hub is in the set $S$, i.e. $1 \in S$. It is easy to observe that in this case, the matrix $Q^S$ is a triangular matrix with all upper diagonal elements equal to 0, and the diagonal elements equal to $(-1, -\frac{1}{n-1}, \ldots, -\frac{1}{n-1})$. Hence, the largest eigenvalue is $-\frac{1}{n-1}$.

*Case 2.* Suppose now that the hub is not in the set $S$, i.e. $1 \in S^c$. If $\lambda$ is an eigenvalue of $Q_S$ with an eigenvector $\vec{x}$, then we have

$$\begin{aligned}
\lambda x_1 &= -x_1 + \frac{1}{n-1} \sum_{i \in S^c \setminus \{1\}} x_i \\
\lambda x_i &= -\frac{x_i}{n-1} + \frac{x_1}{n-1}, \quad \text{for } i \in S^c \setminus \{1\} \\
\lambda x_i &= -\frac{1}{n-1} x_i, \quad \text{for } i \in S.
\end{aligned}$$

This implies

$$\begin{aligned}
\lambda x_1 &= -x_1 + \frac{1}{n-1} \sum_{i \in S^c \setminus \{1\}} x_i \\
x_1 &= ((n-1)\lambda + 1) x_i, \quad \text{for } i \in S^c \setminus \{1\} \\
\lambda x_i &= -\frac{1}{n-1} x_i, \quad \text{for } i \in S.
\end{aligned}$$

Suppose that $\vec{x}$ is such that $x_i = 0$, for every $i \in S$. From the last above identities, it readily follows that $\lambda$ is a solution of the quadratic equation

$$\lambda^2 + \frac{n}{n-1}\lambda + \frac{1}{n-1}\left(1 - \frac{|S^c|-1}{n-1}\right) = 0.$$

It is straightforward to show that the two solutions are

$$\lambda_1 = -\frac{1}{2}\frac{n}{n-1}\left(1 - \sqrt{1 - \frac{4|S|}{n^2}}\right) \text{ and } \lambda_2 = -\frac{1}{2}\frac{n}{n-1}\left(1 + \sqrt{1 - \frac{4|S|}{n^2}}\right).$$

Clearly, the largest eigenvalue is $\lambda_1$ and since $|S| \geq (2\alpha - 1)n$, it is maximized for $|S| = (2\alpha - 1)n$.
Finally, we note that the largest eigenvalue is attained in Case 2, which establishes the first equality in the lemma, from which the asserted inequality and its tightness readily follow.

**D.4. Proof of Lemma 6.13 (Erdös-Rényi).** From (5.3), note that for every $S \subset V$ such that $0 < |S| < n$, if $\lambda$ is an eigenvalue of matrix $Q^S$, then

$$\lambda \leq -\min_{i \in S^c} \left\{ \sum_{j \in S} q_{i,j} \right\}.$$

In the following, we would like to find a value $x_n > 0$ such that $\min_{i \in S^c} \left\{ \sum_{j \in S} q_{i,j} \right\} > x_n$ holds with high probability. To this end, we consider the probability that the latter event does not hold, i.e., for $x_n > 0$, we consider

$$p_e := \mathbb{P}\left( \min_{i \in S^c} \left\{ \sum_{j \in S} q_{i,j} \right\} \leq x_n \right).$$

We first show that the following bound holds $p_e \leq \bar{p}_e$ with

$$\bar{p}_e = 2(1-\alpha)n \exp\left( -(2\alpha - 1) n p_n \varphi\left( \frac{x_n}{2\alpha - 1} \right) \right) \tag{D.5}$$

where we define $\varphi(x) = x\log(x) + 1 - x$, for $x \geq 0$.

To see this, note that for every fixed $\theta > 0$,

$$\begin{aligned} p_e = \mathbb{P}\left( \bigcup_{i \in S^c} \left\{ \sum_{j \in S} q_{i,j} < x_n \right\} \right) &\leq |S^c| \mathbb{P}\left( \sum_{j \in S} q_{i,j} < x_n \right) \\ &\leq |S^c| e^{\theta x_n} \mathbb{E}\left( e^{-\frac{\theta}{np_n} X_{i,j}} \right)^{|S|} \\ &= |S^c| e^{\theta x_n} \left( 1 + p_n \left( e^{-\frac{\theta}{np_n}} - 1 \right) \right)^{|S|} \end{aligned}$$

where the first inequality follows by the union bound, the second inequality by the Chernoff's inequality, and the third equality by the fact that $X_{i,j}$ is a Bernoulli random variable with mean $p_n$.

Since $|S^c| \leq 2(1-\alpha)n$ and $|S| \geq (2\alpha - 1)n$, we have

$$p_e \leq 2(1-\alpha) n e^{\theta x_n} \left( 1 + p_n \left( e^{-\frac{\theta}{np_n}} - 1 \right) \right)^{(2\alpha - 1)n}.$$

Furthermore, using the fact $1 + p_n(e^{-\frac{\theta}{np_n}} - 1) \leq \exp(p_n(e^{-\frac{\theta}{np_n}} - 1))$, it follows

$$p_e \leq 2(1-\alpha) n e^{\theta x_n + (2\alpha - 1) n p_n \left( e^{-\frac{\theta}{np_n}} - 1 \right)}.$$

It is straightforward to check that the right-hand side in the last inequality is minimized for $\theta = -np_n \log\left( \frac{x_n}{2\alpha - 1} \right)$ and for this value is equal to $\bar{p}_e$ given in (D.5).

Requiring $\bar{p}_e \leq \frac{1}{n}$ is equivalent to

$$\varphi\left( \frac{x_n}{2\alpha - 1} \right) \geq \frac{2\log(n) + \log(2(1-\alpha))}{(2\alpha - 1) n p_n}.$$

From this it follows that $\lambda \leq -x_n$ holds with high probability provided that $x_n \geq 0$ can be chosen such that

$$\varphi\left(\frac{x_n}{2\alpha-1}\right) \geq \frac{2}{c(2\alpha-1)}\left(1+\frac{\log(2(1-\alpha))}{2\log(n)}\right). \tag{D.6}$$

Such a value $x_n$ exists as $\varphi(x)$ is a decreasing function on $[0,1]$, with boundary values $\varphi(0)=0$ and $\varphi(1)=0$, and under our assumption, $c(2\alpha-1)>2$, the right-hand side in (D.6) is smaller than 1 for large enough $n$.

Finally, from (D.6), we note

$$x_n \geq (2\alpha-1)\varphi^{-1}\left(\frac{2}{c(2\alpha-1)}\right)+O\left(\frac{1}{\log(n)}\right)$$

from which the asserted result follows.

**Appendix E. Proof of Proposition 6.12.** Let $H(t)$ denote the state of the hub at time $t$. Due to the symmetry of the considered graph, it is not difficult to observe that the dynamics is fully described by a continuous-time Markov process $(H(t),|S_0(t)|,|S_1(t)|,|S_{e_0}(t)|,|S_{e_1}(t)|)_{t\geq 0}$. We need to compute the expected value of the smallest time $t$ such that $|S_1(t)|=0$, i.e. the time when the state 1 becomes depleted. To this end, it suffices to consider $(H(t),|S_0(t)|,|S_1(t)|,|S_e(t)|)$ where $S_e(t)=S_{e_0}(t)\cup S_{e_1}(t)$, i.e. the system states $e_0$ and $e_1$ are lumped into one state, which we denote with $e$. The system will be said to be in mode $i$ at time $t$, whenever the number of depleted state 1 nodes before time $t$ is equal to $i$, for $i=0,1,\ldots,|S_1(0)|-1$. Notice that if the system is in mode $i$ at time $t$, then $|S_0(t)|=|S_0(0)|-i$, $|S_1(t)|=|S_1(0)|-i$, and $|S_e(t)|=|S_e(0)|+2i$. For simplicity, we will use the following notation $x_0^i=|S_0(0)|-i$, $x_1^i=|S_1(0)|-i$, and $x_e^i=|S_e(0)|+2i$, for every given mode $i$.

We will compute the expected sojourn time in each of the modes by analyzing a discrete-time Markov chain $\phi^i=(\phi_k^i)_{k\geq 0}$, for given mode $i$, defined as follows. This Markov chain is embedded at time instances at which the hub node interacts with a leaf node. The state space of $\phi^i$ consists of the states 0, 1, $e$, $e^*$ with the transition probabilities given in Figure E.1. The state of the Markov chain $\phi^i$ indicates the state of the hub at contact instances of the hub with the leaf nodes, where we introduced an extra state $e^*$ to encode the event where the hub is in state $e$ and that this state was reached by the hub from either state 0 or 1, thus indicating a depletion of state 1. Note that the expected duration of mode 0 is equal to the mean hitting time of state $e^*$ for the Markov chain $\phi^0$ started at either state 0 or 1, while the expected duration of mode $i$, for $0<i<|S_1(0)|$, is equal to the mean hitting time of state $e^*$ for the Markov chain $\phi^i$ started at state $e$. We compute these mean hitting times in the following.

Fix an arbitrary mode $i$, and then let $\varphi_s(i)$ be the mean hitting time of state $e^*$ for the Markov chain $\phi^i$ started at state $s=0$, 1, and $e$. By the first-step analysis, we have that the latter mean hitting times are the solution of the following system of linear equations

$$\begin{aligned}
\varphi_0(i) &= \tfrac{x_0^i-1}{n-1}\varphi_0(i)+\tfrac{x_e^i}{n-1}\varphi_e(i)+1\\
\varphi_e(i) &= \tfrac{x_0^i}{n-1}\varphi_0(i)+\tfrac{x_e^i-1}{n-1}\varphi_e(i)+\tfrac{x_1^i}{n-1}\varphi_1(i)+1\\
\varphi_1(i) &= \tfrac{x_e^i}{n-1}\varphi_e(i)+\tfrac{x_1^i-1}{n-1}\varphi_1(i)+1.
\end{aligned} \tag{E.1}$$

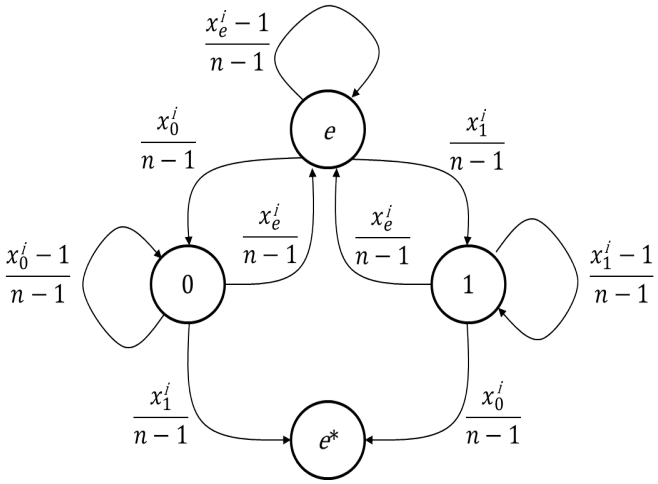


FIG. E.1. *Star-shaped network: the transition probabilities of* $\phi^i$.

From this, it is straightforward to derive

$$\varphi_0(i) = (n-1)\frac{n^2 x_e^i + x_0^i x_1^i}{x_0^i x_1^i (n-x_0^i)(n+x_e^i)}, \tag{E.2}$$

$$\varphi_1(i) = (n-1)\frac{n^2 x_e^i + x_0^i x_1^i}{x_0^i x_1^i (n-x_1^i)(n+x_e^i)}, \tag{E.3}$$

$$\varphi_e(i) = (n-1)\frac{n^2 - x_0^i x_1^i}{x_0^i x_1^i (n+x_e^i)}. \tag{E.4}$$

The expected sojourn time in each given mode is as follows. For mode $i = 0$, it holds $x_0^0 = \alpha n$, $x_1^0 = (1-\alpha)n$, $x_e^0 = 0$, and thus, from (E.2) and (E.3), the expected sojourn in mode 0 is $\varphi_0(0)$ and $\varphi_1(0)$, for the initial state of the hub equal to 0 and 1, respectively, where

$$\varphi_0(0) = \frac{n-1}{(1-\alpha)n} \text{ and } \varphi_1(0) = \frac{n-1}{\alpha n}. \tag{E.5}$$

On the other hand, for $0 < i < |S_1(0)|$, the expected sojourn time in mode $i$ is $\varphi_e(i)$, and from (E.4) and $x_0^i = |S_0(0)| - i$, $x_1^i = |S_1(0)| - i$, $x_e^i = |S_e(0)| + 2i$, we have

$$\varphi_e(i) = (n-1)\left(\frac{n^2}{(\alpha n - i)((1-\alpha)n - i)(n+2i)} - \frac{1}{n+2i}\right). \tag{E.6}$$

Finally, the expected duration of phase 1 is equal to $\varphi_s(0) + \sum_{i=1}^{(1-\alpha)n-1} \varphi_e(i)$, where $s$ denotes the initial state of the hub, either 0 or 1. On the one hand, from (E.5), we have that for every fixed $\alpha \in (1/2, 1]$, both $\varphi_0(0)$ and $\varphi_1(0)$ are asymptotically constants, as the number of nodes $n$ grows large, thus $\varphi_0(0) = \Theta(1)$ and $\varphi_1(0) = \Theta(1)$. On the other hand, using (E.6) and some elementary calculus, we obtain

$$\sum_{i=1}^{(1-\alpha)n-1} \varphi_e(i) = \frac{n-1}{(2\alpha-1)(3-2\alpha)} H_{(1-\alpha)n-1} - \frac{n-1}{(2\alpha-1)(1+2\alpha)}[H_{\alpha n-1} - H_{(2\alpha-1)n}]$$
$$+ \left(\frac{2}{(2\alpha-1)(3-2\alpha)} - \frac{2}{(2\alpha-1)(1+2\alpha)} - 1\right) \sum_{i=1}^{(1-\alpha)n-1} \frac{n-1}{n+2i}.$$

where, recall, $H_k = \sum_{i=1}^{k} \frac{1}{i}$. From this, it can be observed that

$$\sum_{i=1}^{(1-\alpha)n-1} \varphi_e(i) = \frac{1}{(2\alpha-1)(3-2\alpha)} n \log(n) + O(n). \tag{E.7}$$

This completes the proof of the proposition.