

Convergence Speed of Binary Interval Consensus

Moez Draief* and Milan Vojnović†

* Imperial College London † Microsoft Research Cambridge

August 2009

Technical Report
MSR-TR-2009-86

Microsoft Research
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052

<http://www.research.microsoft.com>

Abstract – We consider the speed of convergence of an instance of the binary interval consensus, a distributed and decentralized algorithm for computing the quantized average value. With binary consensus problem, each node initially 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 the majority of nodes.

We derive an upper bound on the expected convergence time that holds for arbitrary connected graphs; it is based on the location of the eigenvalues of some contact rate matrices. We instantiate our bound for particular networks of interest, including complete graphs, star-shaped networks, and Erdős-Rényi random graphs, and in the former two cases compare with alternative computations. We find that for all these examples our bound is of the exact order with respect to the number of nodes and in some cases yields the exact multiplicative constant. We pinpoint the fact that the expected convergence time critically depends on the voting margin defined as the difference between the proportions of nodes that initially held the majority and the minority states, respectively. We derive an exact relation between the expected convergence time and the voting margin, for some of these graphs, that reveals how the expected convergence time tends to infinity as the voting margin approaches zero.

Our results provide insights on 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 peer-to-peer systems; in particular, for sensor networks and distributed databases.

I. INTRODUCTION

Algorithms for distributed computation in networks have recently attracted considerable interest because of their wide-range of applications in a number of contexts such as 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 the majority of nodes. This is to be achieved by a distributed 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 a short period of 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 consensus; we refer to [7], [8] for an account on such algorithms.

In this paper, we consider the *interval or quantized consensus*, a distributed algorithm for deciding in which one of $k \geq 2$ non-overlapping intervals, the average of the values held by the nodes resides. We focus on the *binary interval consensus*, i.e. the case $k = 2$. An attractive feature of the interval consensus is its accuracy; it was showed 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. What was unknown, however, is its speed of convergence.

In what follows, we provide an upper bound on the expected convergence time of the binary interval consensus for 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.

We demonstrate how the general upper bound can be instantiated for a range of examples, including complete graphs, star-shaped networks, and Erdős-Rényi random graphs. The complete graph and the Erdős-Rényi random graph, where contacts from each node are drawn uniformly at random across other nodes, are reasonable approximations of unstructured and structured peer-to-peer networks. The star-shaped network captures scenarios where some node is a hub for other nodes; for example, an information aggregator is a hub node.

Our results provide insights on how the expected convergence time depends on (1) the network structure and (2) the *voting margin*, defined as the difference between the proportion of nodes that initially hold the majority state and the proportion of nodes that initially hold the minority state. The network structure plays a role through the spectral properties of some matrices that dictate the contact rates between the nodes. We find that the voting margin has a significant effect on the expected convergence time.

Denoting by $\alpha > 1/2$ the proportion of nodes that initially hold the majority state, the voting margin is equal to $\alpha - (1 - \alpha) = 2\alpha - 1$. For example, we find that the expected convergence time for the complete graph on n nodes is $\log(n)/(2\alpha - 1)$ for large n , thus it decays as a power-law with the voting margin $2\alpha - 1$ (see Fig. 1 for an illustration). Therefore, albeit the interval consensus guarantees convergence to the correct state, the expected convergence time can assume large values as the voting margin approaches zero.

The contributions of this paper can be summarized in the following points:

- We provide an upper bound on the expected convergence time of the binary interval consensus that applies to arbitrary, connected network topologies. The bound is based on the location of the eigenvalues of some contact rate matrices. It provides a unified approach to estimate the convergence time for particular graphs by either deriving bounds on these eigenvalues analytically or undertaking efficient numerical computations.
- We instantiate our upper bound for several particular network topologies; namely, for complete graphs, stars, and

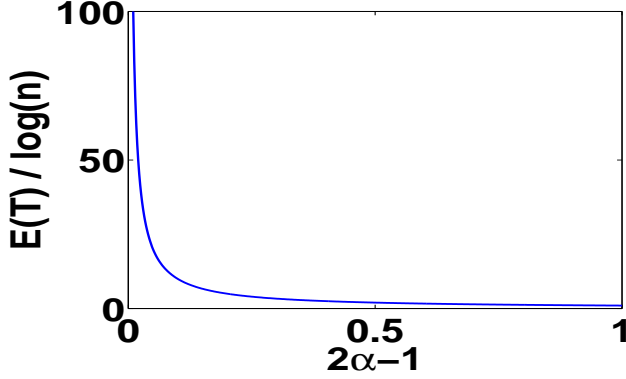


Fig. 1. Normalized expected convergence time for the complete graph on n nodes (for large n) versus the voting margin $2\alpha - 1$. The expected convergence time goes to infinity as the voting margin approaches zero. For example, the expected convergence time for the voting margin 0.02 (i.e. $\alpha = 0.51$) is factor 50 of that for the voting margin near to 1 (i.e. $\alpha = 0.99$).

Erdős-Rényi random graphs. For each of these cases, our upper bound is of the exact order with respect to the number of nodes n and in some cases yields the exact multiplicative constant.

- Our results provide insights into the convergence time for a given network topology. In addition, it supplies valuable clues for the network design problem where the goal is to optimize the network topology with respect to the convergence time of the algorithm. For instance, for Erdős-Rényi random graphs, we found a sufficient condition on the expected number of neighbours, for each node in the network, to ensure a bound on the expected convergence time.

- We find that the expected convergence time critically depends on the voting margin and that it can assume large values as the voting margin approaches zero. This motivates future work towards designing speedier consensus algorithms.

A. Related Work

In recent years, there have been a number of papers that analyzed the effect of the quantization of the values exchanged between nodes known as *quantized consensus* [1], wherein nodes exchange values drawn from a finite set. In [4], the authors provided bounds on the convergence time in the context of the averaging algorithm when agents have access to quantized values of the other agents.

In our case, we assume that each node holds an initial value which is either 0 or 1 and the aim is to reach an actual consensus consistent with the initial majority observation. Similar to the approach in [3], nodes are allowed to update their states to one of two additional intermediate states denoted by e_0 and e_1 . In [3], the authors proved that the addition of these states guarantees the convergence of the algorithm to the correct consensus. They however fell short from providing any analysis of the convergence time.

A related dynamics is the *voter model* whereby each agent adopts the opinion of a randomly chosen neighbour. The voter model has been studied in the context of various graph topologies [9], [10], [11], [12] and it has been proved

that the probability of incorrect consensus (one that is not consistent with the initial majority) is constant bounded away from zero [13]. In [2], the authors proposed and studied the binary consensus with a ternary state maintained and signaled by the nodes. This algorithm fails to converge to the correct consensus with a positive probability. [2] established results for complete graphs showing that (1) the probability of failing to reach the correct consensus is decreasing exponentially with the number of nodes n , with the rate that is decreasing with the voting margin; (2) provided that the convergence is to the correct consensus, the convergence time is $\log(n)$, independent of the voting margin $2\alpha - 1$.

For the complete graph, we find that the binary interval consensus with two intermediate states has the expected convergence time $\log(n)/(2\alpha - 1)$. We note that the convergence time of the ternary protocol [2] is faster by a factor $1/(2\alpha - 1)$, provided that it converges to the correct consensus (which fails to be the case with probability exponentially decaying with the number of nodes n).

Finally, we would like to mention that our work relates to cascading behaviours in on-line social networks [14]. In particular, the viral marketing problem whereby an initial idea or behaviour, held by a portion of the population, “percolates” through the network yielding wide adoption across the whole population [15].

B. Outline of the Paper

In Section II we present the binary interval consensus algorithm, which is the primary focus of this paper, and introduce some basic notation. Section III contains our upper bound on the expected convergence time that applies to arbitrary connected graphs (Theorem III.1) together with the proof of this result. Section IV provides results for particular graphs that we consider. The results are further discussed and compared with related work in Section V. We conclude in Section VI.

We deferred some of the proofs to Appendix.

II. ALGORITHM AND NOTATION

For the binary interval consensus algorithm, each node is in one of four states: 0, e_0 , e_1 , and 1. The states satisfy the following order $0 < e_0 < e_1 < 1$. At each contact of a pair of nodes, their respective states x and y (without loss of generality) ordered such that $x \leq y$, are updated according to the following mapping $(x, y) \mapsto (x', y')$ defined by

$$\begin{aligned} (0, e_0) &\rightarrow (e_0, 0) \\ (0, e_1) &\rightarrow (e_0, 0) \\ (0, 1) &\rightarrow (e_1, e_0) \\ (e_0, e_1) &\rightarrow (e_1, e_0) \\ (e_0, 1) &\rightarrow (1, e_1) \\ (e_1, 1) &\rightarrow (1, e_1) \\ (s, s) &\rightarrow (s, s), \text{ for } s = 0, e_0, e_1, 1. \end{aligned}$$

We assume that the nodes correspond to the vertices of a connected undirected graph $G = (V, E)$ where $V =$

$\{1, \dots, n\}$ corresponds to the set of vertices and E to the set of edges.

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 $q_{i,j}$. We suppose that $q_{i,j} = q_{j,i} \neq 0$ if $(i, j) \in E$.

The number of nodes in state 0 and the number of nodes in state 1 are decremented by 1 following the interaction of a node in state 0 and a node in state 1, and are otherwise unchanged. Note that the mapping is such that the output states x', y' satisfy $x' \geq y'$ whenever the input states are such that $x \leq y$. In particular, this is achieved by swapping the states at encounters of the type $(0, e_0)$, (e_0, e_1) , $(e_1, 1)$. For finite connected graphs, this ensures the existence of encounters in the network such that any state 0 held by a node can get in contact with a state 1 held by some other node through swapping in finite time.

In what follows, we denote by $S_i(t)$ the set of nodes in state $i = 0, e_0, e_1, 1$, at time t , and we use the abbreviate notation $|S_i| \equiv |S_i(0)|$, $i = 0, 1$.

For $\alpha \in (1/2, 1]$, let $|S_0| = \alpha n$ and $|S_1| = (1 - \alpha)n$ be the number of node initially in state 0 and state 1, respectively. By the properties of the algorithm described above, the number of nodes in the initial minority state 1 becomes equal to zero and the number of nodes in the initial majority state 0 becomes equal to $|S_0| - |S_1|$ in finite time. Once there are no more nodes in the initial minority state 1, the number of nodes in state e_1 decreases at encounters with nodes in state 0. This guarantees the convergence of the system to a set of states in which all nodes are either in state e_0 or state 0; this is a correct consensus, indicating 0 as the initial majority.

III. A BOUND ON THE EXPECTED CONVERGENCE TIME FOR ARBITRARY GRAPHS

In this section, we provide an upper bound on the expected convergence time. The bound is in terms of the eigenvalues of a family of matrices Q_S that depend on the transition matrix Q .

Let S be a non-empty subset of V , the set of vertices, of size smaller than n . We consider the matrix Q_S defined by, for $i, j \in V$,

$$Q_S(i, j) = \begin{cases} -\sum_{l \in V} q_{i,l}, & i = j \\ q_{i,j}, & i \notin S, j \neq i \\ 0, & i \in S, j \neq i. \end{cases} \quad (1)$$

Our first result provides a general bound on the eigenvalues of the matrices Q_S that depends on the graph G and the parameter α . As it will appear later in this section, this result is crucial for our analysis based on tracking the number of nodes holding the wrong state, i.e. either 1 or e_1 .

Lemma III.1. *For any finite graph G , there exists $\delta(G, \alpha) > 0$ such that, for any non-empty subset of vertices S with $|S| < n$, if λ is an eigenvalue of the the matrix Q_S defined in (1), then it satisfies*

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

Proof is in Appendix A.

We now state our main result which provides an upper bound for the duration of the two phases of the process and yields a general bound for the time to convergence in terms of the parameter $\delta(G, \alpha)$ defined in Lemma III.1.

Theorem III.1. *Let T_1 be the smallest time at which all the nodes in state 1 are depleted. Then,*

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

Moreover, 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

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

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

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

The process evolves in two phases with rather similar dynamics as it will appear below. Starting from a majority of nodes in state 0 and the rest in state 1, the first phase ends when all nodes in state 1 disappear after interacting with nodes in state 1. This is followed by a second phase that ends when all nodes in state e_1 disappear under the pressure of nodes in state 0.

We take the initial condition αn nodes in state 0 and $(1 - \alpha)n$ nodes in state 1 with $\alpha > 1/2$. The first phase ends when there are no more nodes in state 1, being all moved to the intermediate states e_0 and e_1 (*depletion of nodes in state 1*). It is not difficult to note that, at the end of this phase, there are exactly $(2\alpha - 1)n$ nodes in state 0 and $2(1 - \alpha)n$ nodes in either state e_0 or state e_1 .

The second phase ends when there are no more nodes in state e_1 , being all moved to the states e_0 (*depletion of nodes in state e_1*). It is not difficult to note that, at the end of this phase, there are exactly $(2\alpha - 1)n$ nodes in state 0 and $2(1 - \alpha)n$ nodes in state e_0 .

In the remainder of this section, we prove Theorem III.1; the reader who is more interested in the applications of the result may skip to Section IV with no loss of continuity.

Phase 1: Depletion of nodes in state 1. We describe the dynamics of the first phase through the following indicator functions: Let $i \in V$, we define $Z_i(t)$ be the indicator that node i is in state 0 at time t and $A_i(t)$ be the indicator that node i is in state 1 at time t . The indicator for being in either state e_0 or state e_1 at time t is encoded by $A_i(t) = Z_i(t) = 0$.

For $t \geq 0$, let $A(t) = (A_i(t))_{i \in V}$ and $Z(t) = (Z_i(t))_{i \in V}$. The above dynamics reduces to the transitions of the Markov process (Z, A) given by

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

where e_i is the n -dimensional vector with all entries equal to 0 except the i -th entry which is set to 1.

From this, as Q is symmetric, we have

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

or, equivalently,

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

Let us now consider the behaviour 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 evolve (the instants t_k are stopping times). Consider such interval during which $S_0(t) = S_k$ for $t \in [t_k, t_{k+1})$. We have, for $t \in [t_k, t_{k+1})$

$$\begin{aligned} \frac{d}{dt} \mathbb{E}_k(A_i(t)) &= - \left(\sum_{l \in V} q_{i,l} \right) \mathbb{E}_k(A_i(t)) \\ &\quad + \begin{cases} \sum_{j \in V} q_{i,j} \mathbb{E}_k(A_j(t)) , & i \notin S_k \\ 0 , & i \in S_k \end{cases} \end{aligned} \quad (2)$$

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

$$\frac{d}{dt} \mathbb{E}_k(A(t)) = Q_{S_k} \mathbb{E}_k(A(t)) ,$$

where Q_{S_k} is given by (1).

Solving the above differential equation for $t \in [t_k, t_{k+1})$, we have

$$\mathbb{E}_k(A(t)) = e^{Q_{S_k}(t-t_k)} \mathbb{E}_k(A(t_k)) .$$

Using the strong Markov property [17], it is not difficult to see that,

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

where

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

Phase 2: Depletion of nodes in state e_1 . To describe the

dynamics of the second phase, we introduce, for $i \in V$, $B_i(t)$ be the indicator that node i is in state e_1 at time t . The indicator for being in state e_0 at time t is encoded by $B_i(t) = Z_i(t) = 0$.

For $t \geq 0$, let $B(t) = (B_i(t))_{i \in V}$. Then the dynamics in this phase reduces to the transitions of the Markov process (Z, B) given by

$$(Z, B) \rightarrow \begin{cases} (Z - e_i + e_j, B - e_j) & : q_{i,j} Z_i B_j \\ (Z - e_i + e_j, B) & : q_{i,j} Z_i (1 - B_j - Z_j) \\ (Z, B - e_i + e_j) & : q_{i,j} B_i (1 - B_j - Z_j) \end{cases}$$

From this, we have

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

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 evolve (the instants t'_k are stopping times). Consider such interval during which $S_0(t) = S'_k$ for $t \in [t'_k, t'_{k+1})$. As in the first phase for $t \in [t'_k, t'_{k+1})$, we have

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

where

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

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

$$\begin{aligned} \mathbb{E}(Y(t)) &= \mathbb{E} \left[e^{\lambda(t)} Y(0) \right] \\ \lambda(t) &= Q_{S_k}(t - t_k) + \sum_{l=0}^{k-1} Q_{S_l}(t_{l+1} - t_l) , \end{aligned}$$

where $0 = t_0 \leq t_1 \leq \dots \leq t_k \leq t \leq t_{k+1}$. First, we have

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

since Y is an n -dimensional vector with entries in $\{0, 1\}$ and where $\|\cdot\|$ denotes the matrix norm associated to the Euclidean norm $\|\cdot\|_2$.

Moreover, by Cauchy-Schwartz,

$$\begin{aligned} \sum_{i \in V} \mathbb{E}(Y_i(t)) &\leq \|\mathbb{E}(Y(t))\|_2 \|1\|_2 \\ &\leq n e^{-\delta(G, \alpha)t}. \end{aligned}$$

Therefore, we have

$$\begin{aligned} \mathbb{P}(Y(t) \neq \mathbf{0}) &\leq \sum_{i \in V} \mathbb{E}(Y_i(t)) \\ &\leq n e^{-\delta(G, \alpha)t}. \end{aligned}$$

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

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

IV. APPLICATION TO PARTICULAR GRAPHS

In this section, we instantiate the bounds of Theorem III.1 for particular networks of interest, including complete graphs, star-shaped networks, and Erdős-Rényi random graphs. In the former two cases, we compare with alternative computations as well as simulation results, and provide simulation results for the Erdős-Rényi random graphs. We find that for all these examples our bound is of the exact order with respect to the number of nodes.

For the complete graph, it yields the exact multiplicative constant and we show that our bound is tight by deriving concentration results.

We pinpoint the fact that the expected convergence time critically depends on the voting margin defined as the difference between the portions of the nodes that initially held the majority and the minority state, respectively. For the complete graph and star networks, we derive exact relations between the expected convergence times and the voting margin that, in particular, reveals how the expected convergence time tends to infinity as the voting margin approaches zero.

A. Complete Graphs

In this case, each node i contacts a node $j \neq i$ at instances of a Poisson process of rate $\frac{1}{2(n-1)}$, so that the contact rate of each node is 1 (node initiates contacts at rate $1/2$ and is contacted by other nodes at rate $1/2$). Hence, contacts between any pair of nodes occur at instances of a Poisson process of rate $\frac{1}{n-1}$, i.e. we have $q_{i,j} = \frac{1}{n-1}$ for all $i \neq j$.

Lemma IV.1. *For a complete graph on n nodes,*

$$\delta(G, \alpha) \geq 2\alpha - 1.$$

Proof: The matrix Q_S is such that

$$Q_S(i, j) = \begin{cases} -1, & i = j \\ \frac{1}{n-1}, & i \notin S, j \neq i \\ 0, & i \in S, j \neq i. \end{cases}$$

It is not difficult to see that

$$\underbrace{(1, \dots, 1)}_{S^c}, \underbrace{(0, \dots, 0)}_S)^T, \text{ where } S^c = V \setminus S,$$

is an eigenvector with eigenvalue $-\frac{|S|}{n-1}$. Since in each of the two phases $|S| \geq (2\alpha - 1)n$, we have $\frac{|S|}{n-1} \geq (2\alpha - 1)\frac{n}{n-1} \geq 2\alpha - 1$ and $\delta(G, \alpha) \geq 2\alpha - 1$. ■

Combining the last lemma with Theorem III.1, we have

Corollary IV.1. *For the complete graph on n nodes, the expected time in each of the phases $i = 1, 2$ satisfies*

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

Comparison with direct computation. For complete graphs, we can pursue a more direct approach to bound the expected convergence time, which we use for comparison with the result established using our spectral approach.

Let τ_i be the time of the i -th encounter of a node in state 0 and a node in state 1, where $i = 1, \dots, |S_1|$. Note that for any time $t \geq \tau_{|S_1|} = T_1$, the number of nodes in state 1 is zero. Note that $|S_0(t)| = |S_0| - i$ and $|S_1(t)| = |S_1| - i$, for $\tau_i \leq t < \tau_{i+1}$ and that

$$(|S_0| - i + 1, |S_1| - i + 1) \rightarrow (|S_0| - i, |S_1| - i)$$

at times τ_i , $i = 1, \dots, |S_1|$.

It is not difficult to observe that $L_i = \tau_{i+1} - \tau_i$ is the minimum of $(|S_0| - i)(|S_1| - i)$ random variables (corresponding to the $(|S_0| - i)(|S_1| - i)$ edges between nodes in state 0 and nodes in state 1), each with parameter $1/(n-1)$. Thus L_i is an exponential random variable with parameter

$$\gamma_i = \frac{(|S_0| - i)(|S_1| - i)}{n-1}, \quad i = 0, \dots, |S_1| - 1. \quad (3)$$

Proposition IV.1. *The expected value of T_1 is given by*

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

where $H_k = \sum_{i=1}^k \frac{1}{i}$. Moreover,

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

Proof: Simple summation $\mathbb{E}(T_1) = \sum_{i=0}^{|S_1|-1} \gamma_i^{-1}$. ■

Note that the asymptote in the proposition is exactly $\log(n)/\delta(G, \alpha)$ with $\delta(G, \alpha)$ given in Lemma IV.1. Recall that T_1 is the time to deplete nodes in state 1. Using similar arguments, we can upper bound the expected value of the time T_2 it takes to deplete nodes in state e_1 .

At time T_1 , there are $|S_0| - |S_1|$ nodes in state 0 and 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 time $\mathbb{E}(T_2)$ is largest if all $n - |S_0| + |S_1|$ nodes are in state e_1 . Note that T_2 is the sum of the exponential random variables L'_i with parameter γ'_i given by

$$\gamma'_i = \frac{(|S_0| - |S_1|)(n - |S_0| + |S_1| - i)}{n - 1},$$

for $i = 0, \dots, n - |S_0| + |S_1| - 1$. It follows

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

Summing the latter upper bound and that of Proposition IV.1, we recover the result of Corollary IV.1 which was established by the spectral method.

In fact it is not difficult to get concentration results for T_i , $i = 1, 2$ by analyzing the Laplace transforms of the random variables $W_i = T_i - \frac{1}{(2\alpha - 1)} \log(n)$, $i = 1, 2$ (Proof in Appendix B).

Proposition IV.2. *The random variable W_i*

$$W_i = T_i - \frac{1}{2\alpha - 1} \log(n)$$

is finite with high probability. And as a consequence,

$$T - \frac{2}{2\alpha - 1} \log(n)$$

is finite with high probability.

Mean-field approximation for complete graphs. For complete graphs, we can in fact derive closed-form expressions for the proportions of nodes in each of the states, for asymptotically large networks, i.e. large n . This in turn enables us to estimate the convergence time for large networks.

For the complete graph of n nodes, $\mathbf{X}(t) = (|S_0(t)|, |S_{e_0}(t)|, |S_{e_1}(t)|, |S_1(t)|)$ is a Markov process with the following transition rates

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

We consider the scaled process

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

Under the assumption that $\mathbf{x}^{(n)}(0) \rightarrow \mathbf{x}(0)$ where $\mathbf{x}(0) = (s_0(0), s_{e_0}(0), s_{e_1}(0), s_1(0))$ is fixed, by the Kurtz's convergence theorem [18], we have that $\mathbf{x}^{(n)}(t)$ converges to $\mathbf{x}(t)$ as

n tends to infinity¹, where $\mathbf{x}(t) = (s_0(t), s_{e_0}(t), s_{e_1}(t), s_1(t))$ is the solution to the system of ordinary differential equations

$$\frac{d}{dt} s_0(t) = -s_1(t)s_0(t) \quad (4)$$

$$\frac{d}{dt} s_1(t) = -s_0(t)s_1(t) \quad (5)$$

$$\frac{d}{dt} s_{e_1}(t) = s_1(t)(1 - s_1(t)) - (s_0(t) + s_1(t))s_{e_1}(t) \quad (6)$$

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

The system of equations (4)-(5) admits a closed-form solution (Proof in Appendix C1).

Lemma IV.2. *Let $\Delta = s_0(0) - s_1(0) > 0$. The solution to (4)-(6) is given by*

$$s_0(t) = s_0(0) \frac{\Delta}{s_0(0) - s_1(0)e^{-\Delta t}} \quad (7)$$

$$s_1(t) = s_1(0) \frac{\Delta}{s_0(0) - s_1(0)e^{-\Delta t}} e^{-\Delta t} \quad (8)$$

$$\begin{aligned} s_{e_1}(t) &= s_{e_1}(0) \left(\frac{\Delta}{s_0(0) - s_1(0)e^{-\Delta t}} \right)^2 e^{-\Delta t} \\ &+ s_1(0) \frac{s_0(0)\Delta t - s_1(0)(1 + \Delta)(1 - e^{-\Delta t})}{(s_0(0) - s_1(0)e^{-\Delta t})^2} e^{-\Delta t}. \end{aligned} \quad (9)$$

In Fig. 2, we illustrate the depletion of state e_1 and state 1 nodes over time by plotting the solution in Lemma IV.2 for $\alpha = 0.55$.

We have the following characterization of the convergence time.

Theorem IV.1. *Assume $(s_0(0), s_1(0)) = (\alpha, 1 - \alpha)$ with $\alpha \in (1/2, 1]$. We have*

$$s_{e_1}(t) + s_1(t) \sim (2\alpha - 1) \frac{\alpha}{1 - \alpha} t e^{-(2\alpha - 1)t}, \quad \text{large } t.$$

Moreover, the time $t_{\alpha, n}$ for $s_{e_1}(t) + s_1(t)$ to reach $1/n$ satisfies

$$t_{\alpha, n} \sim \frac{1}{2\alpha - 1} \log(n), \quad \text{large } n. \quad (10)$$

Proof is available in Appendix C. From (8) and (9) it is not difficult to note that for large t ,

$$s_{e_1}(t) \sim (2\alpha - 1) \frac{1 - \alpha}{\alpha} t e^{-(2\alpha - 1)t}$$

$$s_1(t) \sim (2\alpha - 1) \frac{1 - \alpha}{\alpha} e^{-(2\alpha - 1)t}.$$

Hence, the time for both $s_{e_1}(t)$ and $s_1(t)$ to reach $1/n$ is asymptotically (10) which is equal to the bound obtained by the spectral method (Theorem III.1).

B. Star-Shaped Networks

Consider the star network consisting of a hub and $n - 1$ leaves, each of which is attached only to the hub. More

¹More precisely, for any finite $T > 0$, $\sup_{t \in [0, T]} |\mathbf{x}^{(n)}(t) - \mathbf{x}(t)| \rightarrow 0$ with probability 1 as n tends to infinity.

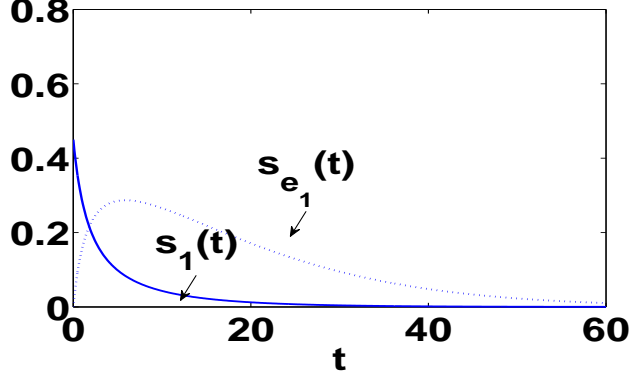


Fig. 2. Depletion of state e_1 and state 1 nodes for large complete graphs.

precisely, we assume that the hub corresponds to node 1 and $q_{1,i} = q_{i,1} = \frac{1}{n-1}$, $i \neq 1$ and $q_{i,j} = 0$, $i, j \neq 1$.

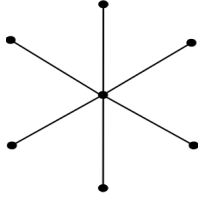


Fig. 3. A star-shaped network.

Lemma IV.3. For a star network of n nodes, we have

$$\delta(G, \alpha) = \frac{n - \sqrt{n^2 - 4(2\alpha - 1)n}}{2(n - 1)}.$$

The above lemma yields the following corollary.

Corollary IV.2. For the star network with $n > 2$ nodes, the expected duration of phase $i = 1, 2$ satisfies

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

Comparison with a direct computation. For the star-shaped network, we can compute exactly the expected time to deplete the nodes in state 1. Recall that T_1 is the smallest time at which the number of nodes in state 1 is equal to zero.

Lemma IV.4. For the star-shaped network with n nodes, the expected time to deplete nodes in state 1 satisfies

$$\mathbb{E}(T_1) = \frac{1}{(2\alpha - 1)(3 - 2\alpha)} n \log(n) + O(n). \quad (11)$$

Notice that the dominant term in Lemma IV.4 is smaller than the upper bound in Corollary IV.2 by a factor $1/(3 - 2\alpha)$.

Proof of the lemma is given in Appendix D3. It is based on computing the expected times between depletions of nodes

in state 1 and then summing up these expected times. Upon depletion of a node in state 1, the hub is in either state e_0 or state e_1 . Since the hub switches to state 0 or 1 with the same rates starting from either e_0 or e_1 , it is irrelevant whether the hub is in state e_0 or e_1 and thus the two states can be lumped into one. A similar analysis could be pursued for phase 2 to estimate the expected time until depletion of the nodes in state e_1 . However, upon depletion of a node in state e_1 , the hub is in either state 0 or state e_0 and now the expected time until depletion of a node in state e_1 does depend on whether the hub is in state 0 or e_0 , which makes analysis more complicated. For this reasons, we do not provide analytical estimates for the expected duration of phase 2 but resort to simulations.

Simulations. We compare our asymptotic results with simulations in Fig. 4. The results in Fig. 4-left demonstrate tightness of the asymptote for the expected duration of phase 1 (Lemma IV.4). Fig. 4-right confirms that $[2/(2\alpha - 1)]n \log(n)$ is indeed an upper bound on the expected convergence time.

C. Erdős-Rényi Random Graphs

This case corresponds to

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

where $0 < p_n < 1$ and $X_{i,j}$ is an i.i.d. sequence of Bernoulli random variables with mean p_n , the normalization with $1/(np_n)$ ensures that for each node, the mean contact rate is equal to 1, where

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

It is well known that, for $c > 1$, the underlying graph defined by the sequence of random variables $X_{i,j}$ is connected with high probability. In what follows, our results are satisfied under the condition $c > \frac{2}{2\alpha - 1} > 1$.

Lemma IV.5. For an Erdős-Rényi random graph with n nodes and for $c(2\alpha - 1) > 2$, we have

$$\frac{1}{\delta(G, \alpha)} \leq \frac{1}{(2\alpha - 1)h^{-1}\left(\frac{2}{c(2\alpha - 1)}\right)} + O\left(\frac{1}{\log(n)}\right) \quad (12)$$

with high probability, where $h^{-1}(\cdot)$ is the inverse function of $h(x) = x \log(x) + 1 - x$.

Proof is in Appendix E.

The above lemma and Theorem III.1 imply the following corollary.

Corollary IV.3. Under $c(2\alpha - 1) > 2$, for phase $i = 1, 2$, we have,

$$\mathbb{E}(T_i) \leq \frac{1}{(2\alpha - 1)h^{-1}\left(\frac{2}{c(2\alpha - 1)}\right)} \log(n) + O(1), \quad (13)$$

for large n .

Remark IV.1. In the case when the mean degree of a node is large, we recover the bound on the convergence time for a

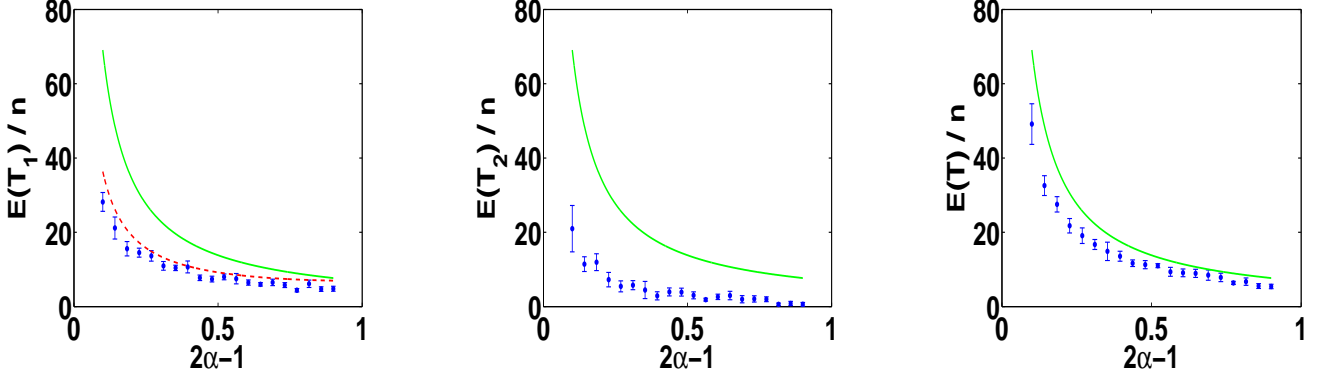


Fig. 4. Expected convergence time for the star-shaped network versus the voting margin $2\alpha - 1$ with $n = 1000$. (Left) expected time in phase 1; (Middle) expected time in phase 2; (Right) expected convergence time. The solid curves indicate $\log(n)/(2\alpha - 1)$; the dashed line indicates $\log(n)/[(2\alpha - 1)(3 - 2\alpha)]$; the bars indicate 95%-confidence estimates obtained by simulations.

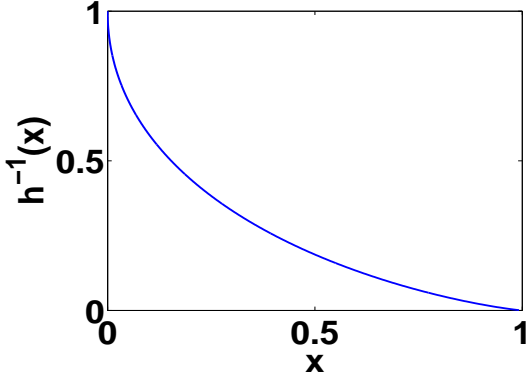


Fig. 5. Function $h^{-1}(x)$.

complete graph. To see this, note that the function $h^{-1}(x)$ is decreasing on $[0, 1]$ with $h^{-1}(0) = 1$ and $h^{-1}(1) = 0$ (see Figure 5). In the case when the mean degree of a node is large, $c(2\alpha - 1) \gg 2$, and we have $h^{-1}(2/(c(2\alpha - 1))) \approx 1$. In this regime, the right-hand side in (13) is approximately equal to $\frac{2}{2\alpha - 1} \log(n)$ which coincides with the bound for a complete graph of n nodes.

Simulations. We compare the bound of Corollary IV.3 with estimates obtained by simulations in Fig. 6. The results confirm that Corollary IV.3 indeed provides a bound. The results also indicate that the bound is not tight which is due to the bound on the eigenvalues that we use in our proof (Appendix E); tighter results may be obtained by using a tighter bound on the eigenvalues.

V. DISCUSSION

In this section, we discuss the convergence speed of the above binary interval consensus, based on the results established in Section IV-A, and compare with the previously-proposed ternary protocol [2] for the case of the complete graph.

In Section IV, we established that for the complete graph on n nodes, the convergence time of the binary interval consensus is $\log(n)/(2\alpha - 1)$, for large n . This is in contrast to the convergence time of the ternary protocol which is given by $\log(n)$, and is independent of the voting margin $2\alpha - 1$, provided that the convergence is to the correct consensus. The convergence to the incorrect consensus occurs with probability p_e that is asymptotically $2^{-D(\alpha||\frac{1}{2})n}$, for large n , (Corollary 1 [2]), where $D(\alpha||\frac{1}{2})$ is the Kullback-Lieber divergence between two Bernoulli distributions with means α and $\frac{1}{2}$, respectively. Therefore, the ternary protocol converges faster than the binary interval consensus by the factor $1/(2\alpha - 1)$, with probability $1 - p_e$.

For any fixed voting margin $\alpha \in (1/2, 1]$ and accuracy parameter $\delta > 0$, we have that $p_e \leq \delta$ and the convergence time is $\log(n)$, for a sufficiently large network of size n . Furthermore, even for a given network size n , the accuracy can be improved by repeatedly running the ternary protocol as follows: run k rounds of the ternary protocol; for each such round, the nodes record their decision; after k rounds have been completed, each node takes as the final decision the majority over the k decisions made by this node. The repeated ternary protocol fails to converge to the correct consensus if it fails to converge for at least $\frac{k}{2}$ rounds. Whenever an incorrect final decision is made by a node, the following event holds $\{X \geq \frac{k}{2}\}$ where X is a binomial random variable with parameters k and $2^{-D(\alpha||\frac{1}{2})n}$. It is not difficult to show that

$$\mathbb{P}\left(X \geq \frac{k}{2}\right) \leq 2^{-\frac{k}{2}(D(\alpha||\frac{1}{2})n-2)}.$$

Suppose $D(\alpha||\frac{1}{2}) > \frac{2}{n}$, then it suffices to take k_δ rounds so that the probability of error is at most δ where

$$k_\delta = \left\lceil \frac{\log_2 \frac{1}{\delta^2}}{D(\alpha||\frac{1}{2}) - 2} \right\rceil.$$

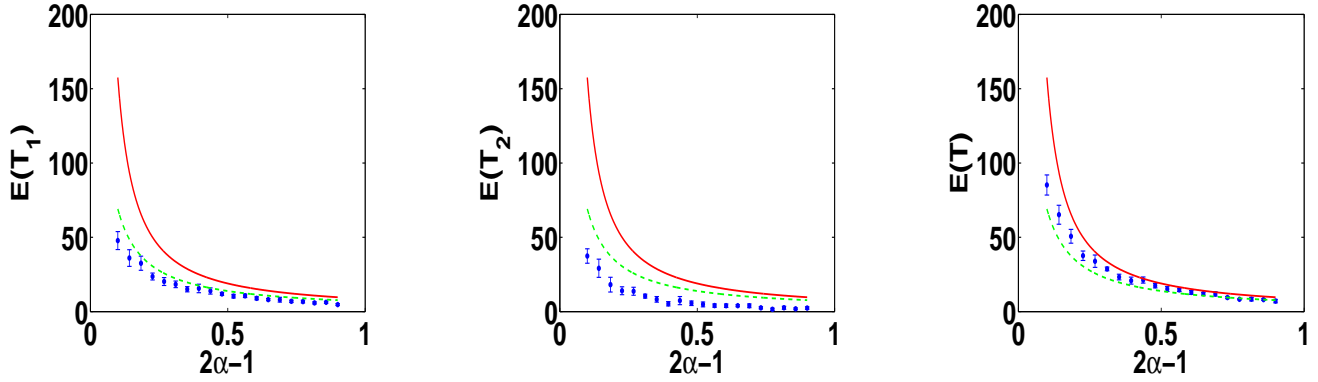


Fig. 6. Expected convergence time for Erdős-Rényi random network versus the voting margin $2\alpha - 1$ with $n = 1000$ and $c = 100$. (Left) expected time in phase 1; (Middle) expected time in phase 2; (Right) expected convergence time. The solid curves indicate the bound of Corollary IV.3; the dashed line indicates $\log(n)/(2\alpha - 1)$; the bars indicate 95%-confidence estimates obtained by simulations.

If the following holds

$$\delta > 2^{-\frac{1}{2}} \frac{D(\alpha|\frac{1}{2})^{n-2}}{2\alpha-1}$$

then the expected convergence time of the repeated ternary protocol is smaller than that of the binary interval consensus.

The above discussion suggests for future work to consider alternative algorithms for the consensus problem that are faster than the binary interval consensus under some accuracy guarantees. We should note that the above discussion is restricted to complete graphs and that, indeed, an appealing property of the binary interval consensus is that it guarantees convergence in finite time with probability 1.

VI. CONCLUSION

We established an upper bound on the expected convergence time of the binary interval consensus that can be applied to arbitrary, connected network topologies. The bound captures the effect of the network topology and the voting margin of the initial state held by the nodes. We showed that for a range of particular network topologies, the bound yields the exact order with respect to the number of the nodes and in some cases yields the exact asymptote. The results indicate that the convergence time of the binary interval consensus can assume large values as the voting margin approaches zero.

Future work could consider the expected convergence time using our spectral method and other approaches for other particular network topologies, e.g. cycle and path. An interesting direction for future work is to investigate the design of faster algorithms, possibly trading accuracy for speed.

REFERENCES

- [1] Akshay Kashyap, Tamer Basar, and Ravi Srikant. Quantized consensus. *Automatica*, 43:1192–1203, 2007.
- [2] Etienne Perron, Dinkar Vasudevan, and Milan Vojnović. Using Three States for Binary Consensus on Complete Graphs. In *IEEE Infocom 2009*, April 2009.
- [3] Florence Bénézit, Patrick Thiran, and Martin Vetterli. Interval Consensus: From Quantized Gossip to Voting. In *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] Yongxiang Ruan and Yasamin Mostofi. Binary consensus with soft information processing in cooperative networks. In *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.
- [7] Achour Mostéfaoui, Michel Raynal, and Frédéric Tronel. From Binary Consensus to Multivalued Consensus in asynchronous message-passing systems. *Information Processing Letters*, 73(5–6):207–212, 2000.
- [8] Jialin Zhanga and Wei Chenb. Bounded cost algorithms for multivalued consensus using binary consensus instances. *Information Processing Letters*, 109(17):1005–1009, 2009.
- [9] 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.
- [10] Peter Donnelly and Dominic Welsh. Finite particle systems and infection models. *Math. Proc. Camb. Phil. Soc.*, 94:167–182, 1993.
- [11] T.M. Liggett. *Interacting particle systems*. Springer, 2 edition, 2006.
- [12] V. Sood and S. Redner. Voter model on heterogenous graphs. *Physical Review Letters*, 94(17):178701, 2005.
- [13] Yehuda Hassin and David Peleg. Distributed probabilistic polling and applications to proportionate agreement. *Information and Computation*, 171(2):248–268, 2002.
- [14] Jon 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] Hamed Amini, Moez Draief, and M. Lelarge. Marketing in Random Networks. In *NetCoop08 (Int'l Workshop on Network Control and Optimization)*, LNCS 5425, pages 17–25, 2009.
- [16] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52(6):2508–2530, 2006.
- [17] P. Brémaud. *Markov Chains: Gibbs Fields, Monte Carlo simulations and Queues*. Springer, 1999.
- [18] Thomas G. Kurtz. *Approximation of Population Processes*. Society for Industrial and Applied Mathematics - Social Science, 1981.

APPENDIX

A. Proof of Lemma III.1

Let S be a non-empty subset of V of size xn , $x \in (0, 1)$. First it is clear from the form of the matrix Q_S that $-\sum_{l \in V} q_{i,l}$, $i \in S$ are eigenvalues of Q_S . Moreover the

remaining eigenvalues correspond to eigenvectors of the form $\underbrace{(x, \dots, x)}_{S^c}, \underbrace{(0, \dots, 0)}_S)^T$.

Let λ such eigenvalue of Q_S and x the corresponding eigenvector such that $\|x\|_2^2 = \sum_{i \in V} x_i^2 = \sum_{i \in S^c} x_i^2 = 1$. Then, as Q is symmetric, it is readily seen that

$$\begin{aligned} \lambda &= - \sum_{i \in S^c} \sum_{j \in V} q_{i,j} x_i^2 + \sum_{i,j \in S^c} q_{i,j} x_i x_j \\ &= - \sum_{i \in S^c} \sum_{j \in S} q_{i,j} x_i^2 - \sum_{i,j \in S^c} q_{i,j} x_i (x_i - x_j) \\ &= - \sum_{i \in S^c} \sum_{j \in S} q_{i,j} x_i^2 - \frac{1}{2} \sum_{i,j \in S^c} q_{i,j} (x_i - x_j)^2. \end{aligned}$$

It is therefore clear that $\lambda \leq 0$ with $\lambda = 0$ if

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

Let $W \subset S^c$ such that for $i \in W$, $x_i \neq 0$ and $S^c \setminus W$ its complement in S^c , i.e. the subset of S^c such that for $i \in S^c \setminus W$, $x_i = 0$. As x_i is an eigenvector then W is non empty. If $\lambda = 0$ then

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

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

It is clear that for any S , Q_S has negative eigenvalues. For n finite, there is a finite number of subsets S . Therefore there exists a $\delta(G, \alpha) > 0$ such that for any non-empty subset of S of size xn , $x \in (0, 1)$ the corresponding matrix Q_S is such that

$$0 > -\delta(G, \alpha) \geq \lambda,$$

where λ is an eigenvalue of Q_S .

B. Concentration results

Note that

$$(|S_0| - i + 1, |S_1| - i + 1) \rightarrow (|S_0| - i, |S_1| - i)$$

at times τ_i , $i = 1, \dots, |S_1|$.

It is not difficult to observe that $L_i = \tau_{i+1} - \tau_i$ is the minimum of $(|S_0| - i)(|S_1| - i)$ random variables (corresponding to the $(|S_0| - i)(|S_1| - i)$ edges between nodes in state 0 and nodes in state 1), each with parameter $2/(n-1)$. Thus L_i is an exponential random variable with parameter

$$\gamma_i = \frac{(|S_0| - i)(|S_1| - i)}{n-1}, \quad i = 0, \dots, |S_1| - 1. \quad (14)$$

The Laplace-Stieltjes transform of L_i equals

$$\begin{aligned} \mathbb{E}(e^{-\theta L_i}) &= \frac{\gamma_i}{\theta + \gamma_i} \\ &= \left(1 + \frac{\theta}{\gamma_i}\right)^{-1} \\ &= \left(1 + \frac{\theta(n-1)}{(|S_0| - i)(|S_1| - i)}\right)^{-1}. \end{aligned}$$

Hence

$$\mathbb{E}(e^{-\theta T_1}) = \prod_{i=0}^{|S_1|-1} \left(1 + \frac{\theta(n-1)}{(|S_0| - i)(|S_1| - i)}\right)^{-1}.$$

Letting $W_1 = T_1 - \frac{1}{(2\alpha-1)} \log(n)$, due to the independence of the L_i (Markov strong property), it follows that

$$\mathbb{E}(e^{-\theta W_1}) = n^{\frac{\theta}{(2\alpha-1)}} \prod_{i=0}^{|S_1|-1} \left(1 + \frac{\theta(n-1)}{(|S_0| - i)(|S_1| - i)}\right)^{-1}.$$

Recall that $|S_0| - |S_1| = (2\alpha - 1)n$. First note that, for $i = 0, \dots, |S_1| - 1$,

$$\frac{1}{(|S_0| - i)(|S_1| - i)} = \frac{1}{(2\alpha - 1)n((1 - \alpha)n - i)} - \frac{1}{(2\alpha - 1)n(\alpha n - i)}.$$

For n large, we have

$$\begin{aligned} &\prod_{i=0}^{|S_1|-1} \left(1 + \frac{\theta(n-1)}{(|S_0| - i)(|S_1| - i)}\right) \\ &\sim \prod_{i=0}^{|S_1|-1} \left(1 + \frac{\theta(n-1)}{(2\alpha - 1)n((1 - \alpha)n - i)}\right) \times \\ &\prod_{i=0}^{|S_1|-1} \left(1 - \frac{\theta(n-1)}{(2\alpha - 1)n(\alpha n - i)}\right) \\ &= \prod_{i=1}^{(1-\alpha)n} \left(1 + \frac{\theta(n-1)}{(2\alpha - 1)ni}\right) \prod_{i=(2\alpha-1)n+1}^{\alpha n} \left(1 - \frac{\theta(n-1)}{(2\alpha - 1)ni}\right) \\ &= \prod_{i=1}^n \left(1 + \frac{\theta(n-1)}{(2\alpha - 1)ni}\right) \prod_{i=(1-\alpha)n+1}^n \left(1 + \frac{\theta(n-1)}{(2\alpha - 1)ni}\right)^{-1} \\ &\quad \prod_{i=(2\alpha-1)n+1}^{\alpha n} \left(1 - \frac{\theta(n-1)}{(2\alpha - 1)ni}\right). \end{aligned}$$

From now on we assume that θ is positive constant. Using the above approximation together with elementary calculus arguments, we can show that, for n large

$$\mathbb{E}(e^{-\theta W_1}) \leq K(\alpha, \theta) \prod_{i=1}^n \left(1 + \frac{\theta(n-1)}{(2\alpha - 1)ni}\right)^{-1},$$

where $K(\alpha, \theta)$ is a positive constant that may depend on α and θ .

Let ϵ_n be a sequence that tends to $-\infty$ when n goes to $+\infty$. Then by the Chernoff bound we have

$$\begin{aligned}\mathbb{P}(W_1 < \epsilon_n) &\leq e^{\theta\epsilon_n} \mathbb{E}(e^{-\theta W}) \\ &\leq K(\alpha, \theta) \left(\Gamma\left(1 + \frac{\theta}{(2\alpha - 1)}\right) + O(1) \right) e^{\theta\epsilon_n},\end{aligned}$$

where we used Euler's formula $\Gamma(1 + z) = z\Gamma(z) = \lim_{n \rightarrow \infty} n^z \prod_{i=1}^n (1 + \frac{z}{i})^{-1}$. We conclude by noting that the last expression goes to 0 when n goes to ∞ .

A similar argument, based on the derivation of an upper bound for $\mathbb{E}(e^{\theta W_1})$, enables us to show that $\mathbb{P}(W_1 > \epsilon_n)$ goes to 0 when n goes to ∞ , for a sequence ϵ_n that tends to $+\infty$ when n goes to $+\infty$.

A similar argument, based on the fact that T_2 is a sum of exponential random variables L'_i with parameter γ'_i given by

$$\gamma'_i = \frac{(|S_0| - |S_1|)(n - |S_0| + |S_1| - i)}{n - 1},$$

for $i = 0, \dots, n - |S_0| + |S_1| - 1$, enables to derive concentration result for W_2 .

The concentration result for the convergence time T is a direct consequence of the above concentration results.

C. Mean-field Approximation for Complete Graphs

1) *Proof of Lemma IV.2:* From (4)-(5), note that $s_0(t)$ and $s_1(t)$ evolve independently of $s_{e_0}(t)$ and $s_{e_1}(t)$. Furthermore,

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

Hence,

$$s_0(t) - s_1(t) = \Delta, \quad \text{for all } t \geq 0 \quad (15)$$

where Δ is defined in the statement of the lemma.

Plugging in (4), we obtain

$$\frac{d}{dt}s_0(t) = (\Delta - s_0(t))s_0(t).$$

The solution to this differential equation is (7). Combining with (15) we obtain (8). Note that (6) is a linear differential equation with time inhomogeneous parameters, i.e.

$$\frac{d}{dt}s_{e_1}(t) + a(t)s_{e_1}(t) = b(t) \quad (16)$$

where

$$\begin{aligned}a(t) &= s_0(t) + s_1(t) \\ b(t) &= s_1(t)(1 - s_1(t)).\end{aligned}$$

From (7) and (8), it follows

$$\begin{aligned}a(t) &= \Delta \frac{s_0(0) + s_1(0)e^{-\Delta t}}{s_0(0) - s_1(0)e^{-\Delta t}} \\ b(t) &= s_0(0)\Delta \frac{s_0(0) - s_1(0)(1 + \Delta)e^{-\Delta t}}{(s_0(0) - s_1(0)e^{-\Delta t})^2} e^{-\Delta t}.\end{aligned}$$

It is well known that the solution to the linear differential equation (16) is

$$s_{e_1}(t) = s_{e_1}(0)e^{-\int_0^t a(s)ds} + \int_0^t b(s)e^{-\int_s^t a(x)dx} ds. \quad (17)$$

By integration we have

$$\exp\left(-\int_0^t a(s)ds\right) = \left(\frac{\Delta}{s_0(0) - s_1(0)e^{-\Delta t}}\right)^2 e^{-\Delta t}. \quad (18)$$

Using the latter identity we have

$$\int_0^t b(s)e^{-\int_s^t a(x)dx} ds = \frac{e^{-\Delta t}}{(s_0(0) - s_1(0)e^{-\Delta t})^2} I(t) \quad (19)$$

where

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

Combined with (19), (18), and (17), we obtain (9).

2) *Proof of Theorem IV.1:* Under the assumptions of the theorem, $(s_0(0), s_{e_0}(0), s_{e_1}(0), s_1(0)) = (\alpha, 0, 0, 1 - \alpha)$. From (9) we obtain

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

From (8) and (9) note that for large t ,

$$\begin{aligned}s_{e_1}(t) &\sim (2\alpha - 1) \frac{1 - \alpha}{\alpha} t e^{-(2\alpha - 1)t} \\ s_1(t) &\sim (2\alpha - 1) \frac{1 - \alpha}{\alpha} e^{-(2\alpha - 1)t}.\end{aligned}$$

The first statement of the theorem follows from the latter two asymptotics. The second statement of the theorem follows from the first by noting that for large n ,

$$\frac{\log(n)}{t_{\alpha, n}} = -\frac{1}{t_{\alpha, n}} \log(s_{e_1}(t_{\alpha, n}) + s_1(t_{\alpha, n})) \sim 2\alpha - 1.$$

Finally, we conclude this section with a remark.

Remark A.1. For the complete graph, the dynamics $\mathbf{X}(t)$ is fully described by the following two-dimensional Markov process $\mathbf{Y} = (|S_{e_1}|, |S_1|)$. Indeed, from the above transition rates we have $|S_0(t)| - |S_1(t)| = |S_0(0)| - |S_1(0)|$, for all $t \geq 0$. We also have $|S_{e_0}(t)| = n - |S_0(t)| - |S_{e_1}(t)| - |S_1(t)|$, $t \geq 0$. From this it follows that

$$\begin{aligned}|S_0(t)| &= |S_0(0)| - |S_1(0)| + |S_1(t)|, \\ |S_{e_0}(t)| &= n - (|S_0(0)| - |S_1(0)|) - |S_{e_1}(t)| - |S_1(t)|.\end{aligned}$$

The Markov process \mathbf{Y} has the following transition rates $(|S_{e_1}|, |S_1|) \rightarrow$

$$\begin{cases} (|S_{e_1}| + 1, |S_1| - 1) & : \frac{|S_1|(|S_0(0)| - |S_1(0)| + |S_1|)}{|S_1|(n - |S_0(0)| + |S_1(0)| - |S_{e_1}| - |S_1|)} \\ (|S_{e_1}| + 1, |S_1|) & : \frac{|S_1|(n - |S_0(0)| + |S_1(0)| - |S_{e_1}| - |S_1|)}{n - 1} \\ (|S_{e_1}| - 1, |S_1|) & : \frac{|S_{e_1}|(|S_0(0)| - |S_1(0)| + |S_1|)}{n - 1}.\end{cases}$$

D. Star-shaped Networks

1) *Proof of Lemma IV.3:* We need to distinguish two cases depending on the hub being in S or not

Case 1: if the hub is in S then the unique eigenvalue of Q_S is $-\frac{1}{n-1}$.

Case 2: if the hub (without loss of generality we can assume that it is node 1) is not in S then λ is an eigenvalue of Q_S with eigenvector $x = (\underbrace{x}_{S^c}, \underbrace{0, \dots, 0}_S)^T$, then it must satisfy

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

which implies

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

Hence

$$(n-1)(\lambda+1)((n-1)\lambda+1) - |S^c| + 1 = 0$$

Let $k = |S^c|$ in our setting $k \leq 2(1-\alpha)n$. We have

$$(n-1)^2\lambda^2 + n(n-1)\lambda + n - k = 0$$

Then the only solutions for λ are $-\frac{n \pm \sqrt{n^2 - 4n + 4k}}{2(n-1)}$. Hence

$$\begin{aligned}\lambda &\leq -\frac{n - \sqrt{n^2 - 4n + 4k}}{2(n-1)} \\ &\leq -\frac{n - \sqrt{n^2 - 4n + 8(1-\alpha)n}}{2(n-1)} \\ &= -\frac{n - \sqrt{n^2 + 4(1-2\alpha)n}}{2(n-1)}.\end{aligned}$$

Therefore, in any case, $\delta(G, \alpha) = \frac{n - \sqrt{n^2 - 4(2\alpha - 1)n}}{2(n-1)}$.

2) *Proof of Corollary IV.2:* We have for $n > 1$,

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

where the inequality is tight for large n . The assertion of the corollary follows by combining the last inequality with Theorem III.1 and making use of the fact $(n-1)(\log n + 1) \leq n \log n$, for $n > 2$.

3) *Proof of Lemma IV.4:* Let $H(t)$ denote the state of the hub at time t . The system state evolves according to the Markov process $(H(t), |S_0(t)|, |S_1(t)|, |S_{e_0}(t)|, |S_{e_1}(t)|)$. We are interested in the hitting time of a state for which $|S_1| = 0$. To this end, it suffices to consider $(H(t), |S_0(t)|, |S_1(t)|, |S_e(t)|)$ where $S_e(t) = S_{e_0}(t) \cup S_{e_1}(t)$. We say that the system is in phase i at time t if the number

of state 1 nodes that has already been depleted is equal to i , i.e. $|S_0(t)| = |S_0(0)| - i$, $|S_1(t)| = |S_1(0)| - i$, and $|S_e(t)| = |S_e(0)| + 2i$.

For the sake of clarity we define an additional state e^* for the hub that corresponds to it reaching the state e following an interaction consisting of the hub at state 0 (resp. 1) and a leaf at state 1 (resp. 0). This corresponds to the start of a new phase following the depletion of both a node in state 0 and a node in state 1.

It is not difficult to observe that the mean time spent in phase i is equal to the mean hitting time of the state e^* , starting from the state e for a discrete-time Markov chain whose state indicates the state of the hub. The transition probabilities of this Markov chain are given by

$$p(x, y) = \begin{cases} \frac{x_e - 1}{n-1}, & x = e, y = e \\ \frac{x_0}{n-1}, & x = e, y = 0 \\ \frac{x_1}{n-1}, & x = e, y = 1 \\ \frac{x_0 - 1}{n-1}, & x = 0, y = 0 \\ \frac{x_e}{n-1}, & x = 0, y = e \\ \frac{x_1}{n-1}, & x = 0, y = e^* \\ \frac{x_1 - 1}{n-1}, & x = 1, y = 1 \\ \frac{x_e}{n-1}, & x = 1, y = e \\ \frac{x_0}{n-1}, & x = 1, y = e^* \end{cases}$$

where $x_0 = |S_0(0)| - i$, $x_e = |S_{e_0}(0)| + |S_{e_1}(0)| + 2i$, $x_1 = |S_1(0)| - i$. We assume that, initially, all the nodes are in either state 0 or state 1, i.e. $|S_{e_0}(0)| = |S_{e_1}(0)| = 0$. Note that $x_0 + x_1 + x_e = n$.

Let $m_s(i)$ be the mean hitting time of state e^* , starting from state s , for $s = 0, 1, e$. Note that $m_e(i)$ is the mean time spent in phase i , for $i = 1, \dots, |S_1(0)| - 1$. By the first-step analysis, we have

$$\begin{aligned}m_0(i) &= \frac{x_0 - 1}{n-1} m_0(i) + \frac{x_e}{n-1} m_e(i) + 1 \\ m_e(i) &= \frac{x_0}{n-1} m_0 + \frac{x_e - 1}{n-1} m_e(i) + \frac{x_1}{n-1} m_1(i) + 1 \\ m_1(i) &= \frac{x_e}{n-1} m_e(i) + \frac{x_1 - 1}{n-1} m_1(i) + 1.\end{aligned}$$

The last system of equations can be rewritten as

$$(n - x_0)m_0(i) = x_e m_e(i) + n - 1 \quad (20)$$

$$(n - x_e)m_e(i) = x_0 m_0(i) + x_1 m_1(i) + n - 1 \quad (21)$$

$$(n - x_1)m_1(i) = x_e m_e(i) + n - 1. \quad (22)$$

From (20) and (22) we have

$$(n - x_0)m_1(i) = (n - x_1)m_0(i).$$

Plugging in (21), we obtain

$$\begin{aligned}&(n - x_e)(n - x_1)m_e(i) \\ &= [x_0(n - x_1) + x_1(n - x_0)]m_0(i) + (n - 1)(n - x_1).\end{aligned}$$

Plugging (20) in the last equation, we have

$$\begin{aligned} & [(n-x_0)(n-x_e)(n-x_1) - x_0x_e(n-x_1) \\ & - x_ex_1(n-x_0)]m_e(i) \\ = & (n-1)[(n-x_0)(n-x_1) + x_0(n-x_1) + x_1(n-x_0)]. \end{aligned}$$

From this, it follows

$$m_e(i) = (n-1) \frac{n^2 - x_0x_1}{x_0x_1(n+x_e)}.$$

Since $x_0 = \alpha n - i$, $x_e = 2i$, and $x_1 = (1-\alpha)n - i$, we have

$$m_e(i) = \frac{n^2(n-1)}{(\alpha n - i)((1-\alpha)n - i)(n+2i)} - \frac{n-1}{n+2i}.$$

Notice

$$\begin{aligned} & \frac{1}{(\alpha n - i)((1-\alpha)n - i)} \\ = & \frac{1}{(2\alpha - 1)n} \left(\frac{1}{(1-\alpha)n - i} - \frac{1}{\alpha n - i} \right) \end{aligned}$$

and

$$\begin{aligned} & \frac{1}{(\alpha n - i)(n+2i)} \\ = & \frac{1}{(1+2\alpha)n} \left(\frac{1}{\alpha n - i} + \frac{2}{n+2i} \right) \end{aligned}$$

and

$$\begin{aligned} & \frac{1}{((1-\alpha)n - i)(n+2i)} \\ = & \frac{1}{(3-2\alpha)n} \left(\frac{1}{(1-\alpha)n - i} + \frac{2}{n+2i} \right). \end{aligned}$$

Putting the pieces together,

$$\begin{aligned} m_e(i) = & \frac{n-1}{(2\alpha-1)(3-2\alpha)} \frac{1}{(1-\alpha)n-i} \\ & - \frac{n-1}{(2\alpha-1)(1+2\alpha)} \frac{1}{\alpha n-i} \\ & + \frac{1-4\alpha(1-\alpha)}{3+4\alpha(1-\alpha)} \frac{n-1}{n+2i}. \end{aligned}$$

It follows

$$\begin{aligned} & \sum_{i=1}^{(1-\alpha)n-1} m_e(i) \\ = & \frac{n-1}{(2\alpha-1)(3-2\alpha)} H_{(1-\alpha)n-1} \\ & - \frac{n-1}{(2\alpha-1)(1+2\alpha)} [H_{\alpha n-1} - H_{(2\alpha-1)n}] \\ & + \frac{1-4\alpha(1-\alpha)}{3+4\alpha(1-\alpha)} \sum_{i=1}^{(1-\alpha)n-1} \frac{n-1}{n+2i} \end{aligned}$$

where, recall, $H_k = \sum_{i=1}^k \frac{1}{i}$.

From the last above math display, it is readily observed that

$$\sum_{i=1}^{(1-\alpha)n-1} m_e(i) = \frac{1}{(2\alpha-1)(3-2\alpha)} n \log(n) + O(n). \quad (23)$$

Since the mean time spent in phase 0 is of the order 1^2 ,² by (23) we establish that (11) holds.

E. Proof of Lemma IV.5

Note that

$$\begin{aligned} \lambda &= - \sum_{i \in S^c} \sum_{j \in S} q_{i,j} x_i^2 - \frac{1}{2} \sum_{i,j \in S^c} q_{i,j} (x_i - x_j)^2 \\ &\leq - \min_{i \in S^c} \left(\sum_{j \in S} q_{i,j} \right). \end{aligned}$$

We establish the following inequality,

$$\begin{aligned} p_e &:= \mathbb{P} \left(\min_{i \in S^c} \left\{ \sum_{j \in S} q_{i,j} < x_n \right\} \right) \\ &\leq 2(1-\alpha)n \exp \left(-(2\alpha-1)np_n h \left(\frac{x_n}{2\alpha-1} \right) \right) \end{aligned}$$

where $h(x) = x \log(x) + 1 - x$, $x \geq 0$.

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

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

Now, since $|S^c| \leq 2(1-\alpha)n$ and $|S| \geq (2\alpha-1)n$, it follows

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

Furthermore, since $\log(1+x) \leq x$, for $x \geq 0$, we have

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

The right-hand side is minimized for

$$\theta = -np_n \log \left(\frac{x_n}{2\alpha-1} \right).$$

Hence, $p_e \leq p'_e$, where

$$p'_e = 2(1-\alpha)n \exp \left(-(2\alpha-1)np_n h \left(\frac{x_n}{2\alpha-1} \right) \right).$$

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

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

We have thus established that for

$$\lambda \leq -x_n$$

²The mean time spent in phase 0 is $(n-1)/|S_1(0)|$ and $(n-1)/|S_0(0)|$, conditional on that the initial state of the hub is 0 or 1, respectively.

to hold with high probability, it suffices that

$$h\left(\frac{x_n}{2\alpha - 1}\right) \geq \frac{2}{c(2\alpha - 1)} \left(1 + \frac{\log(2(1 - \alpha))}{2 \log(n)}\right). \quad (24)$$

Since $h(x) \leq 1$, for $x \in [0, 1]$, for (24) to hold it is easy to observe that it necessary that $c(2\alpha - 1) > 2$.

Finally, we show that Eq. (12) holds. This follows by noting that (24) amounts to

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