Using Three States for Binary Consensus on Complete Graphs

Etienne Perron*, Dinkar Vasudevan*, and Milan Vojnović†
*EPFL, Lausanne, Switzerland †Microsoft Research, Cambridge, UK

September 2008

Technical Report
MSR-TR-2008-114

Microsoft Research
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052
http://www.research.microsoft.com

Abstract – We consider the binary consensus problem where each node in the network initially observes one of two states and the goal for each node is to eventually decide which one of the two states was initially held by the majority of the nodes. Each node contacts other nodes and updates its current state based on the state communicated by the last contacted node. We assume that both signaling (the information exchanged at node contacts) and memory (computation state at each node) are limited and restrict our attention to systems where each node can contact any other node (i.e., complete graphs). It is well known that for systems with binary signaling and memory, the probability of reaching incorrect consensus is equal to the fraction of nodes that initially held the minority state. We show that extending both the signaling and memory by just one state dramatically improves the reliability and speed of reaching the correct consensus. Specifically, we show that the probability of error decays exponentially with the number of nodes N and the convergence time is logarithmic in N for large N. We also examine the case when the state is ternary and signaling is binary. The convergence of this system to consensus is again shown to be logarithmic in N for large N, and is therefore faster than purely binary systems. The type of distributed consensus problems that we study arises in a variety of applications including those of sensor networks and opinion formation in social networks – our results suggest that robust and efficient protocols can be built with rather limited signaling and memory.

I. Introduction

A. Problem and Motivation

The binary consensus problem is the following – Given a network where each node initially observes one of two states, 0 or 1, how to construct a robust distributed protocol which ensures that the nodes reach the right consensus, i.e., the *majority* observation at the start of the protocol. We are interested in the binary consensus problem when there is a limitation on the memory and the communication between the nodes of the network. In particular, we analyze two protocols for the binary consensus problem. In both the protocols, we restrict the nodes to store one of three values 0, 1 and e. In the first protocol, we restrict the signaling to be ternary, i.e, a node can communicate only one of three states, and in the second, we consider the case when the signaling is binary. We call the two protocols respectively, the ternary signaling and the binary signaling protocol. The extra state e corresponds to an "undecided" state where the node is unsure of the majority value. This state can also be thought of as an extra quantization level which corresponds to the averaging of a 0 and a 1. In [18], the authors also introduce nodes with an "undecided" state, but only in the beginning of the protocol execution. The dynamics of the two protocols that we propose are very different from the dynamics in [18]. We are interested in characterizing our two protocols with respect to error probability of final consensus and convergence rate. Our results are for the case when the underlying graph is *complete* – analysis for general graphs is of interest but is out of the scope of this paper and is left for future work.

The distributed binary consensus problem arises in several applications. For instance, consider a ranking application in networks where each node has personally ranked two items. A node can observe from some other node how this node has ranked the items at current time, which again depends on the observations of this encountered node from other nodes. The objective is to get all the nodes to agree on the rank of the item based on the initial majority opinion. Other examples are sensor networks where the binary observations could be some state of nature, or social networks, where the observations reflect an opinion held on some recently released media item (video or audio or piece of news), when this opinion as well as the media item is displayed on publicly accessible web pages. The protocol with binary signaling is not only of technical interest to understand the performance under even further limited information exchange, but is of interest also from a practical stance. For example, in our ranking application scenario, note that the nature of the application may well be such that each node can only signal one of two states, e.g. each node must display one of two media items. The user may be in the "indifferent state" e with respect to both items but will still have to display one of the media items in her profile page and thus signal a preference for this media item to other nodes. In the ternary model, the user could have signalled indifference to other nodes through an appropriate display that indicates state e (e.g. show both videos or show equal preference scores).

B. Related Work

One approach for binary consensus is the voter model, where at a sampling instant, a node picks up the opinion of a randomly chosen neighbouring node. At any given time, nodes store binary values and communicate binary observations. The voter model has been extensively studied in the context of infinite lattices [13], finite graphs [6], [1], [7], heterogenous random graphs [16] and social networks [18]. While the voter model guarantees consensus, the probability of incorrect consensus is a constant bounded away from zero depending on the initial fraction (and location) of the minority observations – "the proportional agreement" [7]. However, the voter model is economical with respect to memory and signaling.

At the other extreme, if there are no memory or signaling constraints, any robust averaging algorithm would guarantee reliable consensus. Indeed, the average value of the initial observations indicates the majority observation. Various approaches to averaging have been analyzed, such as gossip based algorithms [3] and belief propagation [14]. These averaging protocols requires that real values be stored and exchanged between nodes. This appears to be an excessive overhead when the observations are binary and the objective is to obtain the majority observation.

More recently, in the context of averaging algorithms, the effects of quantization of the values exchanged between nodes has been studied [8],[5]. In [5], the quantized algorithm guarantees consensus, in that the nodes agree on a final value. However the algorithm does not preserve the average at every step. Bounds on the error between the final consensus value and the initial average are provided as a function of the number of quantization levels. In [8], randomized gossip-type quantized averaged algorithms are studied under the assumption that the initial observations are integer values. However, the algorithms preserve the average of the initial observations, and the consensus attained is approximate, i.e., the final readings can differ by 1.

In binary distributed hypothesis testing (distributed detection), nodes observe a binary hypothesis through independent noisy channels and communicate with each other in a rate constrained manner, the objective being that one or several nodes should agree on a reliable estimate of the hypothesis. Binary detection is related to the consensus problem in two ways. On one hand, one can derive a distributed detection scheme by first applying locally optimal detection rules to map the observation of each sensor node into an estimate of the hypothesis, and then running a consensus protocol to disseminate the majority estimate to all the nodes. On the other hand, the binary consensus problem can be viewed as a particular distributed detection problem, where the channel for the observations is a binary symmetric channel which flips the binary hypothesis with a given probability. All nodes are required to reliably construct the hypothesis. However, to our knowledge, such binary channels have never been considered in the distributed detection literature. See [17] and [2] for a survey of this research area.

Our work also relates to hypothesis testing with limited memory considered in [11] – Given a sequence $X_1, X_2, \ldots$ of i.i.d. Bernoulli random variables with unknown mean p, the goal is to identify, in which of given m disjoint intervals that cover the interval $[0, 1]$, the parameter p lies. Limited information can be stored about the sequence $\{X\}_1^i$ at any time i. It was found that to identify the correct hypothesis with diminishing probability of error with the number of samples, it is *necessary and sufficient* to maintain $m+1$ states. Hence, for binary hypothesis testing, 3 states are sufficient. Our problem is different as we consider *a dynamical system* where observations are taken from other nodes in the network and thus not necessarily i.i.d.

Our work also relates to diffusion of innovations and cascading, which is one of the central questions in social sciences and of interest in a number of on-line settings. The specific question is that of understanding how an initial idea or behavior attains wide adoption across the network – see [10] for a survey of results and models.

In summary, to the best of our knowledge, the problem studied in this paper has not been addressed previously and the results that we establish appear novel.

C. Summary of our Results

We show that adding one extra state increases both the reliability and the speed of reaching the correct consensus. We show that if α is the initial fraction of nodes observing the majority value, then, under the ternary signaling protocol, the probability of reaching the false consensus decays *exponentially* with rate $N \log 2(1 - H(\alpha))$ (Corollary 1). Here, N is the number of nodes and $H(\cdot)$ is the binary entropy function. This result is in contrast to previous quantized algorithms, where reliable binary consensus is only possible as the number of quantization levels gets large. When N is large, the convergence time is shown to be (Theorem 2) *logarithmic* in the number of nodes in the network. This result says that to reach consensus, a node need not sample all the other nodes in the network. It suffices for the node to sample, uniformly at random, a logarithmic number of nodes. In [9], it is shown that the convergence of a gossip-type averaging algorithm is logarithmic in the number of nodes on a complete graph. Thus, our results show that on a complete graph, three states are sufficient to obtain a convergence as fast as for real-valued states. For binary signaling, we show (Theorem 4) that the error probability is no worse than the classical voter model but is worse (Corollary 2) than ternary signaling by a factor that increases exponentially with N. For large N, we show (Corollary 3) that the convergence time under binary signaling is slower by at least a factor 2 than ternary signaling. However, we establish (Corollary 3) that this slow-down is no worse than a factor 3. While we are not able to obtain an exponential upper bound on the error probability under binary signaling, our simulation results indicate that even under binary signaling the error probability may be decaying exponentially.

The organization of the paper is as follows. Section II gives the preliminaries, Section III analyzes the ternary signaling protocol while Section IV analyzes the binary signaling protocol. Section V compares the analysis with simulations and is followed by the conclusion in Section VI.

II. Preliminaries

Consider an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ is the set of vertices and $\mathcal{E}$ is the set of edges. Let vector $\mathbf{X}(0) = [X_1(0), \ldots, X_N(0)]^\top$ represent the vector of *binary* $0, 1$ observations $X_i(0)$ observed by node i. Let

$$\hat{X}(0) = \begin{cases} 0 & \frac{1}{N}\sum_{i=1}^{N} X_i(0) \leq 0.5 \\ 1 & \frac{1}{N}\sum_{i=1}^{N} X_i(0) > 0.5 \end{cases}$$

represent the majority of the initial observations at the nodes. Our goal is to construct a reliable distributed protocol with *minimal communication and memory overhead* that computes the majority $\hat{X}(0)$ at each node.

We use the asynchronous time model defined in [3]. Each node has a clock which ticks at the times of a rate 1 Poisson process. The inter-tick times at each node are rate 1 exponentials, independent across nodes and over time. Equivalently, this corresponds to a single clock ticking according to a rate N Poisson process at times Z_k, $k \geq 1$, where $\{Z_{k+1} - Z_k\}$ are i.i.d. exponentials of rate N. At time Z_k, a node i chosen uniformly at random from $\mathcal{V}$ contacts a neighbouring vertex j, again chosen at random and updates its value based on the signal received from j.

Throughout the paper for two sequences a_N and b_N we write $a_N \sim b_N$, large N, meaning that a_N/b_N tends to 1 as N tends to infinity. Also, U and V denote the number of nodes that are in state 1 and 0, respectively, while $u = \lim_{N\to\infty} \frac{U}{N}$ and $v = \lim_{N\to\infty} \frac{V}{N}$ denote the asymptotic fraction of the nodes that are in state 1 and 0, respectively.

III. Ternary Signaling

The ternary signaling model corresponds to the following two constraints, respectively, on the communication and the state at every node in $\mathcal{G}$.

- Communication is ternary – a node can communicate only one of three states to its neighbouring node.
- State is ternary – a node can store only one of three states.

Our protocol under this model is the following – At any time, a node can store one of three values $0, 1$ or e. The value e implies that the node is undecided about the majority value. Let $\mathcal{U}, \mathcal{V}$ and $\mathcal{S}$ represent, respectively, the set of nodes storing $0, 1$ and e. If a node in $\mathcal{U}$ (resp. $\mathcal{V}$) contacts a node in $\mathcal{U}, \mathcal{S}$ (resp. $\mathcal{V}, \mathcal{S}$), then it does not update its value. If a node in $\mathcal{U}$ (resp. $\mathcal{V}$) contacts a node in $\mathcal{V}$ (resp. $\mathcal{U}$), it updates its value to e(resp. e). If a node in $\mathcal{S}$ contacts a node in $\mathcal{U}$ (resp.$\mathcal{V}$), then it updates it's value to 1 (resp. 0).

A. System Dynamics

We describe the dynamics for a general graph $\mathcal{G}$. Let $U_i = 1$ if node i is in state 1 and $V_i = 1$ if node i is in state 0. We encode the state e by $U_i = V_i = 0$. Let $p(i,j) \geq 0$ be given for each node pair i and j. At a sampling instance, a node i samples node j with probability $p(i,j)$. The state of the system evolves according to the continuous-time Markov process $(\mathbf{U}, \mathbf{V})$ specified by the transition rates:

$$(\mathbf{U}, \mathbf{V}) \rightarrow \begin{cases} (\mathbf{U} + \mathbf{e}_i, \mathbf{V}) & : (1 - U_i - V_i) \sum_j p(i,j) U_j \\ (\mathbf{U} - \mathbf{e}_i, \mathbf{V}) & : U_i \sum_j p(i,j) V_j \\ (\mathbf{U}, \mathbf{V} + \mathbf{e}_i) & : (1 - U_i - V_i) \sum_j p(i,j) V_j \\ (\mathbf{U}, \mathbf{V} - \mathbf{e}_i) & : V_i \sum_j p(i,j) U_j, \end{cases}$$

where $\mathbf{e}_i$ is a vector of dimension N with all coordinates equal to 0 but the i-th coordinate equal to 1.

We now focus our analysis of the ternary signaling protocol on the complete graph with N nodes. In particular, we choose $p(i,j) = 1/N$, for all $i, j \in 1, \ldots, N$. Let $U = \sum_i U_i$ and $V = \sum_i V_i$. We then have that (U, V) is a continuous-time Markov process specified by the transition rates

$$(U, V) \rightarrow \begin{cases} (U+1, V) & : (N - U - V) U/N \\ (U-1, V) & : UV/N \\ (U, V+1) & : (N - U - V) V/N \\ (U, V-1) & : VU/N \end{cases} \tag{1}$$

Note that this is a Markov process on finite state space $S_N \triangleq \{(U, V) \in \mathbb{N}_+^2 : \quad U + V \leq N\}$ and it therefore terminates in one of the absorbing states $(N, 0)$ or $(0, N)$. Above and hereafter, (U, V) sometimes denotes the random process $(U(t), V(t))$, and sometimes a deterministic value taken by the random process. The meaning should be clear from the context.

We are interested in the probability of error of our ternary signaling protocol for the complete graph as well as the expected time to convergence. The two quantities are examined in the following two sections.

B. Probability of Error

For any (U, V), define

$$\begin{aligned} f_{U,V} \triangleq \mathbb{P}((U(t), V(t)) &= (N, 0) \text{ for some } t \geq 0 \\ &|(U(0), V(0)) = (U, V)). \end{aligned} \tag{2}$$

From (1), using the first-step analysis [4] we have that $f_{U,V}$ satisfies the following recursion:

$$\begin{aligned} (\epsilon U + \epsilon V + 2UV) f_{U,V} = \epsilon U f_{U+1,V} + UV f_{U-1,V} \\ + \epsilon V f_{U,V+1} + UV f_{U,V-1} \end{aligned} \tag{3}$$

where $\epsilon = N - U - V$. The boundary conditions of $f_{U,V}$ are given by $f_{0,V} = 0$ for $V \geq 1$ and $f_{U,0} = 1$ for $U \geq 0$.

An error occurs when the protocol converges to the false consensus i.e., $U(0) > V(0)$ and $(U(t), V(t))$ hits $(0, N)$ or vice versa. i.e., $U(0) < V(0)$ and $(U(t), V(t))$ hits $(N, 0)$. Without loss of generality, we focus on the case $U(0) < V(0)$, for which $f_{U,V}$ is the error probability. Note that by the symmetry of the protocol, $f_{U,U} = \frac{1}{2}$ for $U = 1, \ldots, \lfloor \frac{N}{2} \rfloor$.

The following theorem provides the solution to $f_{U,V}$ and thus establishes an exact expression for the error probability.

Theorem 1. *The solution to (3) when $V > U$ is given by*

$$f_{U,V} = \frac{1}{2} \sum_{j=1}^{U} \frac{a_{U,V}(j)}{2^{(U-j)+(V-j)}} \tag{4}$$

where

$$a_{U,V}(j) = \frac{V - U}{(U-j) + (V-j)} \binom{(U-j) + (V-j)}{(U-j)}. \tag{5}$$

The proof is in the appendix. It turns out that $f_{U,V}$ satisfies a recursion for the error probability of the auxiliary Markov process (X, Y) specified by the transition probabilities

$$(X, Y) \rightarrow \begin{cases} (X, Y-1) & : \frac{1}{2} 1_{\{Y>0\}} \\ (X-1, Y) & : \frac{1}{2} 1_{\{X>0\}} \end{cases}$$

with the same boundary conditions. This recursion is solved by a path counting argument using the Ballot theorem [15].

Theorem 1 implies that for a large number of nodes N with initial state (U, V) such that (U, V) scales linearly with N, we have that the error probability $f_{U,V}$ decays exponentially with N. The rate of this decay is given in the following corollary.

Corollary 1. *Let the initial state (U, V) be such that there exists $\alpha \in (1/2, 1]$ for which $(U, V)/N \rightarrow (1 - \alpha, \alpha)$ as N tends to infinity. We have*

$$\frac{1}{N} \log_2 f_{U,V} \sim -\left[1 - H(\alpha)\right], \quad \textit{large } N,$$

where $H(x)$ is the entropy of a binary random variable with mean x.[1]

Therefore, the probability of error decays *exponentially* with N at a rate which depends on the portion of nodes α that

[1] I.e. $H(x) = -x \log_2(x) - (1 - x) \log_2(1 - x)$ for $x \in [0, 1]$.

hold the initial majority opinion. This is in sharp contrast to the classical voter model for the complete graph where the probability of error is a *constant* $(1-\alpha)$ and is *independent* of N. Thus, the addition of a state e into the communication has the effect of making consensus far more robust.

The next section examines the speed of convergence of the ternary protocol.

C. Convergence Time

In contrast to the previous section, where the analysis is done for finite N, in this section we examine the rate of convergence of the protocol in the asymptotic setting, i.e., for large N. At time t, there are $U(t)$ nodes in state 1, $V(t)$ nodes in state 0, and $S(t)$ nodes in state e. We have $U(t)+V(t)+S(t)=N$ for all $t\geq 0$.

Define the scaled state $u^N(t)=U(t)/N$, $v^N(t)=V(t)/N$, and $s^N(t)=S(t)/N$. The Markov process (U,V,S) is a density dependent Markov jump process, so by the known convergence result by Kurtz [12], we know that under the assumption that $(u^N(0),v^N(0),s^N(0))$ goes to a fixed $(u(0),v(0),s(0))$, then $(u^N(t),v^N(t),s^N(t))$ uniformly converges on any compact time interval to $(u(t),v(t),s(t))$, given by the system of ordinary differential equations:

$$\frac{du(t)}{dt}=u(t)s(t)-u(t)v(t)$$
$$\frac{dv(t)}{dt}=v(t)s(t)-v(t)u(t)$$
$$\frac{ds(t)}{dt}=2u(t)v(t)-s(t)(u(t)+v(t)).$$

As $s(t)=1-u(t)-v(t)$, it suffices to consider:

$$\frac{du(t)}{dt}=u(t)(1-u(t)-2v(t)) \tag{6}$$
$$\frac{dv(t)}{dt}=v(t)(1-v(t)-2u(t)). \tag{7}$$

Theorem 2. *The system (6)–(7) has the following properties:*

1) *If* $u(0)<v(0)$ *[resp.* $u(0)>v(0)$*] then* $(u(t),v(t),s(t))$ *goes to* $(0,1,0)$ *[resp.* $(1,0,0)$*].*
2) *If* $u(0)=v(0)$*, then* $(u(t),v(t),s(t))$ *goes to* $(\frac{1}{3},\frac{1}{3},\frac{1}{3})$.
3) *The time* t *to reach* $(u(t),v(t))$ *is given by*

$$t=\log\left(\frac{(v(t)-u(t))^3}{u(t)v(t)}\right)-\log\left(\frac{(v(0)-u(0))^3}{u(0)v(0)}\right)$$

4) *For* $v(0)>u(0)$*, the time* $t(N)$ *to reach* $(u(t),v(t))$ *so that* $u(t)$ *and* $1-v(t)$ *are order* $1/N$ *is such that*

$$t(N)\sim\log N,\ \textit{large}\ N. \tag{8}$$

Item 4 tells us that the convergence time is *logarithmic* in the number of nodes N with multiplicative constant 1.

While we established good properties for the ternary protocol with respect to the error probability and convergence time, we point out that these established results are for complete graphs. In the following, we give an example of a graph for which the ternary protocol provides no additional benefits over the voter model with respect to either the error probability or the convergence time.

Consider a line graph. Suppose that the initial observations at the nodes are given as: The first U nodes from the left observe 1 and the remaining $V=N-U$ nodes observe 0. The graph and the initial observations at the nodes are illustrated in Figure 1. It can be shown that the dynamics of ternary signaling parallels that of the classical voter model. The probability of false consensus is a constant (given by $\frac{U}{N}$ if the nodes observing 0 are in a majority) and the convergence time is quadratic in N. The detailed proof is in Appendix VII-H.

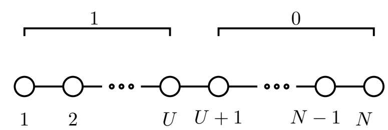

Fig. 1. Initial state on the line graph.

IV. Binary Signaling

So far, we showed that for complete graphs, increasing the computation and signaling state from binary to ternary yields significant improvements for both the error probability and the convergence time. However, it is a priori unclear whether this benefit is because of augmenting both the computation state and the signaling or just the state. In this section, we study the improvements that are achieved if only the computation state is ternary and the signaling remains binary. Our protocol under this model is the following – as in Section III, a node can store one of three values 0, 1 or e. Signaling, however, is binary, *i.e.*, nodes are only allowed to display one of two values 0 or 1. If a node in state 0 (respectively 1) is contacted by another node, it displays its state. If a node in state e is contacted by another node, it draws one of the values $\{0,1\}$ at random (with uniform probabilities) and displays that value. The updating rules are as in Section III, *i.e.*, if a node in state 0 (respectively 1) contacts a node that displays a 1 (respectively a 0), then the contacting node changes its state to e. If the contacting node is in state e, it changes its state to the displayed value. Otherwise, the state of the contacting node remains unchanged.

A. System Dynamics

It can be readily checked that the system dynamics are fully described by the Markov process (U,V) specified by the transition rates:

$$(U,V)\to\begin{cases}(U+1,V) & :\frac{1}{2}(N-U-V)(1+\frac{U-V}{N})\\ (U-1,V) & :\frac{1}{2}U(1+\frac{V-U}{N})\\ (U,V+1) & :\frac{1}{2}(N-U-V)(1+\frac{V-U}{N})\\ (U,V-1) & :\frac{1}{2}V(1+\frac{U-V}{N}).\end{cases} \tag{9}$$

Note that this is a Markov process on a finite state space $S_N\triangleq\{(U,V)\in\mathbb{N}_+^2:\ U+V\leq N\}$ and it therefore terminates in one of the absorbing states $(N,0)$ or $(0,N)$. We are interested in the probability of error of the binary signaling protocol for the complete graph as well as the expected time of convergence. The two quantities are examined in the following two sections.

B. Probability of Error

Let $f_{U,V}$ be defined as in (2). From (9), using first-step analysis, one can show that $f_{U,V}$ satisfies

$$\begin{aligned} & \left[(U+\epsilon)\left(V+\frac{\epsilon}{2}\right)+(V+\epsilon)\left(U+\frac{\epsilon}{2}\right)\right] f_{U,V} \\ = \quad & U\left(V+\frac{\epsilon}{2}\right) f_{U-1,V}+V\left(U+\frac{\epsilon}{2}\right) f_{U,V-1} \\ & +\epsilon\left(U+\frac{\epsilon}{2}\right) f_{U+1,V}+\epsilon\left(V+\frac{\epsilon}{2}\right) f_{U,V+1} \end{aligned} \tag{10}$$

where $\epsilon \triangleq N-(U+V)$ with the boundary conditions given by $f_{U,U}=\frac{1}{2}$ for $U \in \{0,\ldots,\frac{N}{2}\}$, $f_{0,N}=0$, and $f_{N,0}=1$. As before, $f_{U,V}$ gives the error probability when $U<V$. The following theorem provides a lower bound on $f_{U,V}$.

Theorem 3. *The solution to (10) for $V>U$ satisfies $f_{U,V} \geq p_{V-U}$ where*

$$p_{V-U}=\frac{1}{2} \frac{\sum_{i=N+V-U}^{2N-1} \frac{N^i}{i!}}{\sum_{i=N}^{2N-1} \frac{N^i}{i!}}. \tag{11}$$

The lower bound decays exponentially with N with rate specified in the following.

Corollary 2. *Let the initial state (U,V) be such that there exists $\alpha \in (1/2,1]$ for which $(U,V)/N \to (1-\alpha,\alpha)$ as N tends to infinity. We have*

$$\frac{1}{N}\log(p_{V-U}) \sim -\left[1-2\alpha(1-\log(2\alpha))\right], \text{ large } N. \tag{12}$$

The function $1-2\alpha(1-\log(2\alpha))$ is increasing convex with α on $[1/2,1]$ (as the derivative $2\log(2\alpha)$ is increasing with α) and we have

$$0 \leq 1-2\alpha(1-\log(2\alpha)) \leq 2\log 2-1 \approx 0.3863.$$

For the ternary signaling protocol, we established that the decay rate of the error probability was $a(\alpha)=[1-H(\alpha)]\log 2$. If we let $b(\alpha)=1-2\alpha(1-\log(2\alpha))$, then it can be checked that the difference $a(\alpha)-b(\alpha)$ is increasing in α and

$$0 \leq a(\alpha)-b(\alpha) \leq 1-\log 2.$$

In summary, for any fixed fraction $V/N > 1/2$, the gap between the error probabilities under binary signaling and ternary signaling is exponentially large.

We are not able to prove an upper bound to the error probability that is exponentially decaying in N. However, we can show the following:

Theorem 4. *For $V \geq U$, $f_{U,V} \leq \frac{U}{N}$.*

The above theorem says that binary signaling is at least as reliable as the classical voter model. The proof follows by considering the recursion (10) only on the line $U+V=N$ and showing that the probability of reaching $(N,0)$ (when restricted to this line) upper bounds $f_{U,V}$. The detailed proof can be found in Appendix VII-I.

C. Convergence Time

As in the analysis of the convergence time for the ternary signaling protocol in Section III-C we consider the asymptotic behaviour of the system for a large number of nodes N. The limit dynamics are given by the following system of ordinary differential equations:

$$\frac{du(t)}{dt}=\frac{1}{2}\big((1-v(t))^2-(1+v(t))u(t)\big) \tag{13}$$

$$\frac{dv(t)}{dt}=\frac{1}{2}\big((1-u(t))^2-(1+u(t))v(t)\big). \tag{14}$$

Let $z=u+v$ and $w=v-u$. By simple calculus,

$$\frac{d}{dt}z(t) \quad = \quad 1-\frac{3}{2}z(t)+\frac{1}{2}w(t)^2 \tag{15}$$

$$\frac{d}{dt}w(t) \quad = \quad \frac{1}{2}(1-z(t))w(t). \tag{16}$$

An interesting observation is the following:

Proposition 1. *For the system (13)–(14), the initial majority remains the majority forever.*

Proof: The result follows by noting the fact $z(t) \leq 1$, for all $t \geq 0$, thus from (16), the sign of $w(t)=v(t)-u(t)$ remains unchanged for all $t \geq 0$. ■

We are not able to provide an exact derivation of the convergence time for this system. However, the following theorem and its corollary give lower and upper bounds, both of the same order as in ternary signaling. The proofs are given in the appendix.

Theorem 5. *The solution $(u(t),v(t))$ of the system (13)–(14) satisfies:*

1) *For any initial point $(u(0),v(0))$ such that $v(0)>u(0)$ we have that for a finite $t_0 \geq 0$, and all $t \geq t_0$,*

$$u(t) \quad \geq \quad \frac{(1-v(t))^2}{1+v(t)} \tag{17}$$

$$u(t) \quad \leq \quad \frac{3}{2}+v(t)-\frac{1}{2}\sqrt{1+24v(t)}. \tag{18}$$

2) *Time lower bound: for any $t \geq t_0$,*

$$t-t_0 \geq \log\left(\frac{v(t_0)}{v(t)}\left(\frac{3v(t)-1}{3v(t_0)-1}\right)^{\frac{8}{3}}\left(\frac{1-v(t_0)}{1-v(t)}\right)^2\right). \tag{19}$$

3) *Time upper bound: for any $t \geq t_0$,*

$$t-t_0 \leq 3\log\left(\frac{(v(t)-u(t))^2}{(v(t_0)-u(t_0))^2}\frac{1-(v(t_0)-u(t_0))^2}{1-(v(t)-u(t))^2}\right). \tag{20}$$

Corollary 3. *From any initial point $(u(0),v(0))$ such that $v(0)>u(0)$ we have that the time $t(N)$ for $(u(t),v(t))$ to reach the state such that $u(t)$ and $1-v(t)$ are of order $1/N$ satisfies*

$$2\log(N)+A \leq t(N) \leq 3\log(N)+B, \text{ large } N \tag{21}$$

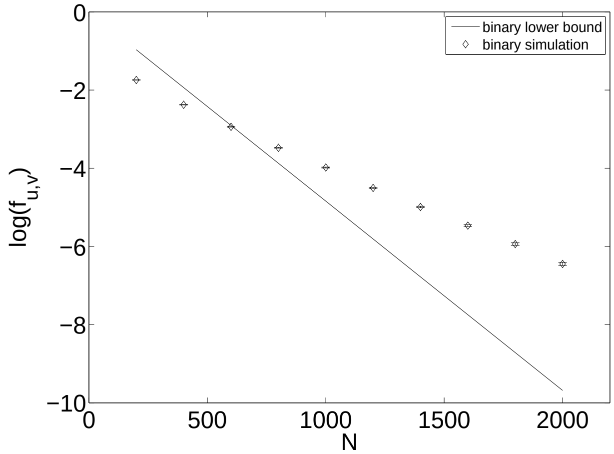

Fig. 2. Plot of $\log f_{u,v}$ versus the number of nodes N for binary signaling with $(u(0), v(0)) = (0.45, 0.55)$. The exponent of the error probability corresponds to the slope. Confidence intervals are for 95% of confidence.

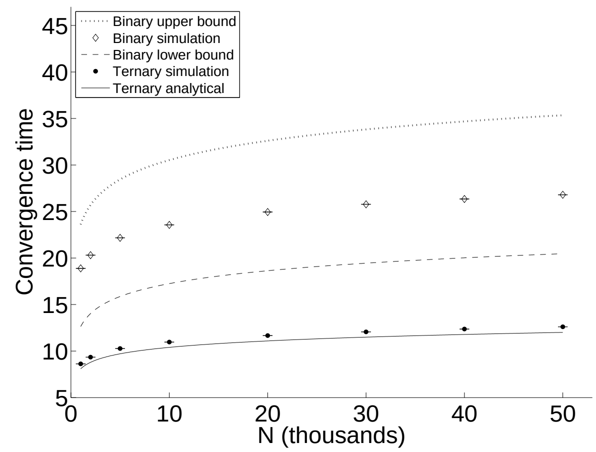

Fig. 3. Convergence time versus the number of nodes N for binary and ternary signaling with $(u(0), v(0)) = (0.3, 0.7)$. Confidence intervals are for 95% of confidence.

where A and B are constants that depend on the initial point $(u(0), v(0))$ but not on N. If the initial point $(u(0), v(0))$ satisfies the inequalities (17)–(18), then we can set,

$$A = \log\left(\frac{v(0)(1-v(0))^2}{\left(\frac{3v(0)-1}{2}\right)^{\frac{8}{3}}}\right) \tag{22}$$

and

$$B = 3\log\left(\frac{1-(v(0)-u(0))^2}{2(v(0)-u(0))^2}\right). \tag{23}$$

We established that the binary protocol is slower than the ternary protocol by at least a factor 2 for large complete graphs. We also established that this slow-down is by at most a factor 3. In the next section, we provide numerical results that validate our convergence time analysis for the ternary and binary signaling protocol.

V. Numerical Results

A. Probability of Error for Binary Signaling

Figure 2 shows the exponent of the error probability of binary signaling obtained from simulations and also the lower bound (12). The simulation plot indicates that the probability of error for binary signaling decays *exponentially* with N, but with slower rate than that of the lower bound (12).

The simulations confirm that (12) is an asymptotic lower bound. Note that although the lower bound on the error probability (11) is valid for all N, the approximations used to get the exponent (12) are valid only for large values of N. This is visible in Figure 2, where the so-called lower bound is actually larger than the estimated error probability for N below 600.

B. Convergence Time

Figure 3 shows the convergence time for binary and ternary signaling and validates (8) and (21). In particular, the simulations confirm that (a) convergence time grows logarithmically for both binary and ternary signaling, (b) the multiplicative constant for ternary signaling is 1, and (c) the multiplicative constant for binary signaling lies between 2 and 3.

VI. Conclusion

The binary consensus problem has been studied for complete graphs. It is shown how adding an extra state at the nodes increases the reliability and speed of consensus. This is primarily because the resulting mean field equations (6) and (7) for ternary signaling and (13) and (14) are such that the state dynamics are steered towards reliable consensus. We are in the process of studying the performance of the protocols for more general classes of graphs. Another direction of future work would be to establish analogous results for the n-ary consensus problems.

References

[1] David Aldous and James Allen Fill. *Reversible Markov Chains and Random Walks on Graphs*, volume 26. monograph in preparation (http://www.stat.berkeley.edu/ aldous/RWG/book.html), 1999.

[2] R. S. Blum, S. A. Kassam, and H. V. Poor. Distributed detection with multiple sensors – part ii: Advanced topics. *Proc. IEEE*, 85:64–79, 1997.

[3] Stephen Boyd, Arpita Ghosh, Balaji Prabhakar, and Devavrat Shah. Randomized gossip algorithms. *IEEE Trans. on Information Theory*, 52(6):2508–2530, June 2006.

[4] Pierre Brémaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation and Queues*. Springer, 1999.

[5] Angelia Nedić, Alex Olshevsky, Asuman Ozdaglar, and John N. Tsitsiklis. On distributed averaging algorithms and quantization effects. *under submission*, 2007.

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

[7] Yehuda Hassin and David Peleg. Distributed probabilistic polling and applications to proportionate agreement. *Information and Computation*, 171(2):248–268, December 2002.

[8] Akshay Kashayap, T. Basar, and R. Srikant. Consensus with quantized information updates. In *Proc. of the 45th IEEE Conference on Decision & Control*, San Diego, CA, December 2006.

[9] D. Kempe, A. Dobra, and J. Gehrke. Gossip-based computation of aggregate information. *Foundations of Computer Science*, 2003.

[10] Jon Kleinberg. *Cascading Behavior in Networks: Algorithmic and Economic Issues*. In Algorithmic Game Theory (N. Nisan, T. Roughgarden, E. Tardos, V. Vaizrani, eds.), Cambridge University Press, 2007.

[11] Jack Koplowitz. Necessary and sufficient memory size for m-hypothesis testing. *IEEE Trans. on Information Theory*, IT-21(1):44–46, Jan 1975.

[12] Thomas G. Kurtz. *Approximation of Population Processes*. Society for Industrial and Applied Mathematics - Social Science, 1981.

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

[14] Ciamac C. Moallemi and Benjamin Van Roy. Consensus propagation. *IEEE Trans. on Information Theory*, 52(11):4753–4766, Nov 2006.

[15] A.N. Shiryaev. *Probability*. Springer, Berlin, 1996.

[16] V Sood and S Redner. Voter model on heterogenous graphs. In *Physical Review Letters, 2005*, May 2005.

[17] R. Viswanathan and P. K. Varshney. Distributed detection with multiple sensors – part i: Fundamentals. *Proc. IEEE*, 85:54–63, 1997.

[18] Fang Wu and Bernardo A. Huberman. Social structure and opinion formation. *Computational Economics, http://www.hpl.hp.com/research/idl/papers/opinions/opinions.pdf*, 2004.

VII. Appendix

A. Proof of Theorem 1

The proof follows from the following two lemmas.

Lemma 1. *The solution to (3), for any $N \geq 1$, is $f_{U,V}$ given by*

$$f_{U,V} = \frac{1}{2} f_{U,V-1} + \frac{1}{2} f_{U-1,V} \tag{24}$$

with boundary conditions given by $f_{U,0} = 1$ for $U \geq 0$ and $f_{0,V} = 0$ for $V \geq 1$.

Proof: Assume that $f_{U,V}$ satisfies (24) for all (U,V). We show that this $f_{U,V}$ also satisfies the recursion (3) with the same boundary conditions. We do this in several steps. First, we show by induction that $f_{U,V}$ satisfies the recursion

$$(U+V) f_{U,V} = U f_{U+1,V-1} + V f_{U-1,V+1}, \tag{25}$$

again with the same boundary conditions.

Base case: Let $n \triangleq U+V = 2$. It is easy to check that for $(U,V) = (1,1)$, both (24) and (25) yield $f_{1,1} = \frac{1}{2}$.

Induction step: Assume that $f_{U,V}$ satisfies (25) for all (U,V) such that $U+V \leq n-1$. Now, for any (U,V) satisfying $U+V = n$, the induction assumption implies that

$$(U-1+V) f_{U-1,V} = (U-1) f_{U,V-1} + V f_{U-2,V+1}$$

and

$$(U+V-1) f_{U,V-1} = U f_{U+1,V-2} + (V-1) f_{U-1,V}.$$

Summing these two equations, we obtain

$$U f_{U-1,V} + V f_{U,V-1} = U f_{U+1,V-2} + V f_{U-2,V+1}.$$

Multiplying both sides by $\frac{1}{2}$ and adding the same term $\frac{U}{2} f_{U,V-1} + \frac{V}{2} f_{U-1,V}$ on both sides, we obtain

$$(U+V)\left(\frac{1}{2} f_{U,V-1} + \frac{1}{2} f_{U-1,V}\right) =$$
$$U\left(\frac{1}{2} f_{U,V-1} + \frac{1}{2} f_{U+1,V-2}\right) + V\left(\frac{1}{2} f_{U-2,V+1} + \frac{1}{2} f_{U-1,V}\right).$$

Using (24), we see that the above is equivalent to (25). This concludes the induction step.

Next, we show that (24) and (25) together imply

$$(U+V) f_{U,V} = U f_{U+1,V} + V f_{U,V+1}. \tag{26}$$

This follows by applying the recurrence (24) on the points $(U, V+1)$ and $(U+1, V)$:

$$f_{U+1,V} = \frac{1}{2} f_{U,V} + \frac{1}{2} f_{U+1,V-1} \tag{27}$$

$$f_{U,V+1} = \frac{1}{2} f_{U-1,V+1} + \frac{1}{2} f_{U,V}. \tag{28}$$

From (25), it holds that

$$\begin{aligned}(U+V) f_{U,V} &= U f_{U+1,V-1} + V f_{U-1,V+1} \\ &\overset{(a)}{=} U(2 f_{U+1,V} - f_{U,V}) + V(2 f_{U,V+1} - f_{U,V})\end{aligned}$$

where (a) follows from (27) and (28). The above rearranges to (26). The proof of the lemma follows by noting that (3) is a linear combination of (24) and (26). Since we know $f_{U,V}$ is a solution to both these recursions, it follows that $f_{U,V}$ is also a solution of (3). ■

Lemma 2. *For $U \leq V$, the solution to (24) is given by (4).*

Proof: Recursing successively Eq. (24),

$$\begin{aligned}f_{U,V} &= \frac{1}{2} f_{U,V-1} + \frac{1}{2} f_{U-1,V} \\ &= \frac{1}{2}\left(\frac{1}{2} f_{U,V-2} + \frac{1}{2} f_{U-1,V-1}\right) + \frac{1}{2}\left(\frac{1}{2} f_{U-1,V-1} + \frac{1}{2} f_{U-2,V}\right) \\ &\ldots \\ &= \sum_{j=1}^{U} c_j f_{j,j} + \sum_{j=1}^{V} b_j f_{0,j} = \frac{1}{2} \sum_{j=1}^{U} c_j,\end{aligned}$$

we note that $f_{U,V}$ can be expressed as a linear combination of the $U+V$ boundary terms. Since $f_{0,j} = 0$, we focus on computing c_j. Consider paths on the lattice S_N that for any two sites of S_N are defined as a concatenation of downward and leftward edges between neighbouring sites of S_N. It is easy to check that the coefficient c_j is the product of the number of paths from the site (U,V) to the site (j,j) that do not intersect with the $U = V$ line, and $1/2^{(U-j)+(V-j)}$. The latter term is due to the accumulation of the $\frac{1}{2}$ factor while applying the recursion successively. The number of such paths is given by the Ballot theorem [15]. Indeed, let the number of ballots given to candidate 1 and candidate 2 be $V-j$ and $U-j$, respectively. The number of paths that do not intersect the $U = V$ line until the point (j,j) is equivalent to the number of permutations for which candidate 1 is ahead of candidate 2 throughout the counting of ballots. It follows from the Ballot theorem that the number of such permutations is

$$\begin{aligned} & \frac{(V-j)-(U-j)}{(V-j)+(U-j)}\binom{(V-j)+(U-j)}{(V-j)} \\ = & \frac{V-U}{(U-j)+(V-j)}\binom{(U-j)+(V-j)}{(V-j)}. \end{aligned}$$

The result follows. ∎

B. Proof of Corollary 1

Let $x=j/N$, $V=\alpha N$, and $U=(1-\alpha)N$. From Eq. (5) and Stirling's approximation, we have

$$\begin{aligned} & \frac{1}{N}\log\left(a_{U,V}(j)2^{-((U-j)+(V-j))}\right) \\ \sim & \frac{1}{N}\log\binom{(1-2x)N}{(\alpha-x)N}-(1-2x),\ \text{large } N \\ \sim & -\nu(x),\ \text{large } N \end{aligned}$$

where

$$\nu(x)\triangleq(1-2x)\left[1-H\left(\frac{\alpha-x}{1-2x}\right)\right].$$

By the principle of the largest term, we have

$$\frac{1}{N}\log_2 f_{U,V}\sim-\min_{x\in[0,1-\alpha]}\nu(x),\ \text{large } N.$$

It is readily checked that $\nu(x)$ is increasing on $[0,1-\alpha]$ hence it achieves minimal value at $x=0$. The result follows.

C. Proof of Theorem 2

Subtracting (6) from (7), we have

$$\frac{d((v(t)-u(t))}{dt}=(v(t)-u(t))(1-u(t)-v(t)) \tag{29}$$

The relation (29) says that the difference $v(t)-u(t)$ is increasing with t. Therefore the initial majority is the final majority. (29) can equivalently be written as

$$d\log(v(t)-u(t))=(1-u(t)-v(t))dt \tag{30}$$

Furthermore, (6) and (7) can be written, respectively as

$$d\log u(t)=(1-u(t)-2v(t))dt \tag{31}$$

$$d\log v(t)=(1-v(t)-2u(t))dt \tag{32}$$

Adding (31) and (32), we get

$$d\log u(t)v(t)=\big(2-3(u(t)+v(t))\big)dt \tag{33}$$

Integrating (30) and (33) from 0 to t and rearranging, we obtain

$$\int_0^t\big(u(x)+v(x)\big)dx=x-\log\big[v(x)-u(x)\big]_0^t$$

$$\int_0^t\big(u(x)+v(x)\big)dx=\frac{2x}{3}-\frac{1}{3}\log\big[v(x)u(x)\big]_0^t.$$

From the above two identities, we get

$$t=3\log\frac{v(t)-u(t)}{v(0)-u(0)}-\log\frac{v(t)u(t)}{v(0)u(0)}.$$

If $t(N)$ is chosen so that $u(t(N))\approx\frac{1}{N}$ and $v(t(N))\approx 1-\frac{1}{N}$, then for a constant C independent of N,

$$t(N)\approx\log N+C.$$

Finally, when $u(0)=v(0)$, then from the differential equations, we have $u(t)=v(t)$ for all $t\geq 0$ and

$$\frac{du(t)}{dt}=2u(t)(1-3u(t))$$

This is a logistic differential equation with the limit point $(1/3,1/3,1/3)$.

D. Proof of Theorem 3

Define $K(t)=V(t)-U(t)$. At every step, the value of $K(t)=K$ updates to one of $K-1,K,K+1$. The transition probabilities conditional on K updating to $K-1$ or $K+1$ are given by

$$\begin{aligned} & \mathbb{P}(K(t+1)=K+1|(U(t),V(t))=(U,V)) \\ & \quad=\frac{(U+\epsilon)(V+\frac{\epsilon}{2})}{(U+\epsilon)(V+\frac{\epsilon}{2})+(V+\epsilon)(U+\frac{\epsilon}{2})} \\ & \mathbb{P}(K(t+1)=K-1|(U(t),V(t))=(U,V)) \\ & \quad=\frac{(V+\epsilon)(U+\frac{\epsilon}{2})}{(U+\epsilon)(V+\frac{\epsilon}{2})+(V+\epsilon)(U+\frac{\epsilon}{2})}. \end{aligned}$$

We rewrite the first probability as

$$\mathbb{P}(K(t+1)=K+1|(U(t),V(t))=(U,V)) \tag{34}$$

$$=\frac{1}{1+\frac{(N-K)(N-U)}{(N+K)(N-V)}}. \tag{35}$$

For a fixed K, one can check that the above probability is maximum when $(U=0,V=K)$.

The probability of error is one-half times the the probability of reaching the state $K=0$ before reaching $K=N$. We obtain a lower bound by assuming that the bias towards the larger K (i.e., $K=K+1$) is always maximum. Hence, we consider a new Markov chain K' with transition probabilities

$$\mathbb{P}(K'(t+1)=K-1|K'(t)=K)=\frac{N}{2N+K}$$

$$\mathbb{P}(K'(t+1)=K+1|K'(t)=K)=\frac{N+K}{2N+K}.$$

The probability of error p_k for this Markov chain satisfies

$$(2N+K)p_K=(N+K)p_{K+1}+Np_{K-1}, \tag{36}$$

for $K=1,\ldots,N-1$ with boundary conditions $p_0=\frac{1}{2}$, $p_N=0$. The lower bound to $f_{U,V}$ is given by

$$f_{U,V}\geq p_{V-U}\quad\text{for } U\leq V.$$

It can be verified that the solution to (36) is

$$p_K=\frac{1}{2}\frac{\sum_{i=N+K}^{2N-1}\frac{N^i}{i!}}{\sum_{i=N}^{2N-1}\frac{N^i}{i!}}. \tag{37}$$

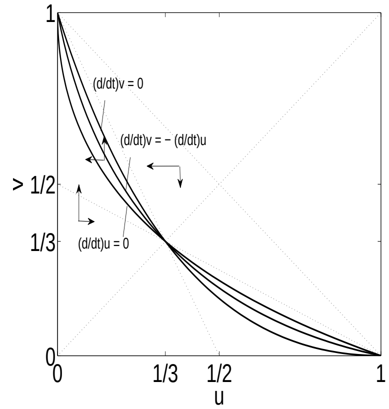

Fig. 4. The vector field of (13)–(14).

E. Proof of Corollary 2

We consider the asymptotics of p_K defined in (37) for large N where $(U,V)/N$ tends to $(1-\alpha,\alpha)$ for fixed $\alpha \in (1/2,1]$.

From (37),

$$\begin{aligned}\log(p_K) &= -\log 2 + \log\left(\sum_{i=N+K}^{2N-1}\frac{N^i}{i!}\right) - \\ &\quad -\log\left(\sum_{i=N}^{2N-1}\frac{N^i}{i!}\right).\end{aligned} \tag{38}$$

Now, $N^i/i!$ is decreasing with i for $N \le i \le 2N$. Hence, by the principle of the largest term,

$$\frac{1}{N}\log\left(\sum_{i=N+K}^{2N}\frac{N^i}{i!}\right) \sim \frac{1}{N}\log\left(\frac{N^{N+K}}{(N+K)!}\right)$$

for large N. By Stirling's approximation, we have that

$$\frac{1}{N}\log\left(\frac{N^m}{m!}\right) \sim \frac{m}{N}\left(1-\log\left(\frac{m}{N}\right)\right), \text{ large } N.$$

Using the last asymptote in (38), we have

$$\frac{1}{N}\log(p_K) \sim \frac{K}{N} - \left(1+\frac{K}{N}\right)\log\left(1+\frac{K}{N}\right), \text{ large } N.$$

The result follows by noting that $K/N \sim 2\alpha - 1$.

F. Proof of Theorem 5

1) Item 1: Figure 4 illustrates the vector fields $(du/dt, dv/dt)$ in the region $\{v > u, v+u \le 1\}$. Item 1 says that after a finite time t_0, $(u(t),v(t))$ lies between the curves representing the points where, respectively, $du/dt = 0$ and $d(u+v)/dt = 0$.

The claim follows by direct inspection of the vector field of the system (13)–(14) – see Fig. 4. It suffices to consider only $(u,v) \in \Omega$ defined by $\Omega = \{(u,v) \in [0,1]^2 : \ v > u,\ u+v \le 1\}$. Indeed, for the system (13)–(14), if $(u(0),v(0)) \in \Omega$, then $(u(t),v(t)) \in \Omega$, for any $t \ge 0$. This follows from Proposition 1. The claim in the theorem says that for any $(u(0),v(0)) \in \Omega$ there exists a finite $t_0 \ge 0$ such that $(u(t),v(t)) \in A$, for all $t \ge t_0$, where the set A is defined by

$$A = \{(u,v) \in \Omega :\ (1-v)^2/(1+v) \le u \le 3/2 + v - \sqrt{1+24v}/2\}.$$

We first note that if $(u(0),v(0)) \in A$, then $(u(t),v(t)) \in A$, for all $t \ge 0$. To see this, note that at the boundary $(1-v)^2/(1+v) = u$ the vector field is such that $(d/dt)u = 0$ and $(d/dt)v > 0$, thus points inwards into the set A. Similarly, note that the boundary $u = 3/2 + v - \sqrt{1+24v}/2$ is the same as $v = f(u)$, where $f(u) \triangleq 3/2 + u - \sqrt{1+24u}/2$ at which the vector field is such that $(d/dt)v = -(d/dt)u$. It suffices to show that $(d/du)f(u) \le -1$, for all $u \in A$. The last inequality is equivalent to $u \le 1/3$, which is indeed true for $u \in A$. Thus, the vector field also points inwards into the set A at the boundary $u = f(v)$.

It remains only to show that for $(u(0),v(0)) \in \Omega \setminus A$, we have $(u(t_0),v(t_0)) \in A$, for some finite $t_0 \ge 0$. Recall that we defined $w(t) = v(t) - u(t)$, and $z(t) = u(t) + v(t)$. We consider the following cases.

- **Case 1:** $(u(0),v(0)) \in B_1$ where $B_1 = \{(u,v) \in \Omega : u \le (1-v^2)/(1+v)\}$. In this region, we have that $(du/dt) \ge 0, (dv/dt) \ge 0$. Extend a vertical line upward (90^0) and a unit slope 45^0 line from $(u(0),v(0))$ until the lines intersect the curve $u = (1-v^2)/(1+v)$. Let C be the region enclosed by the two lines and the curve. Since both u and v are increasing in this region and since v is increasing more than u, the process is constrained to lie in C before hitting the set A. Clearly, $dz/dt > 0$ in C as the region C is bounded away from the curve $v = 3/2 + u - \sqrt{1+24u}/2$ (at which points, $dz/dt = 0$). Hence the time t_0 taken to hit the curve $u = (1-v^2)/(1+v)$ is upper bounded by
$$t_0 \le K_1(1 - u(0) - v(0))$$
where $K_1 < \infty$ is a constant.
- **Case 2:** $(u(0),v(0)) \in B_2$ where $B_2 = \{(u,v) \in \Omega :\ v \ge (1-u)^2/(1+u)\}$. In this region, we have that $(du/dt) \le 0, (dv/dt) \le 0$. Extend a horizontal line leftward (180^0) and a unit slope (-135^0) line from $(u(0),v(0))$ until the lines intersect the curve $v = (1-u^2)/(1+u)$. Let D be the region enclosed by the two lines and the curve. Since both u and v are decreasing in this region and since u is decreasing more than v, the process is constrained to lie in D before hitting the set A. Clearly, $dz/dt < 0$ in D as the region D is bounded away from the curve $v = 3/2 + u - \sqrt{1+24u}/2$. Hence the time t_0 taken to hit the curve $v = (1-u^2)/(1+u)$ is upper bounded by
$$t_0 \le K_2(u(0) + v(0))$$
where $K_2 < \infty$ is a constant.
- **Case 3:** $(u(0),v(0)) \in B_3$ where $B_3 = \{(u,v) \in \Omega : 3/2 + u - \sqrt{1+24u}/2 \le v \le (1-u)^2/(1+u)\}$. In this region, we have that $(d/dt)u \le 0 \le (d/dt)v$, and we also have $(d/dt)(v+u) \le 0$. Let t_0 be the time when $u(t), v(t)$ intersects with the curve $v = 3/2 + u - \sqrt{1+24u}/2$. For $t \le t_0$,
$$\begin{aligned}\frac{dw}{dt} &= \frac{1}{2}(1-z(t))w(t) \\ &\ge \frac{1}{2}(1-z(0))w(0) > 0\end{aligned}$$

The inequality follows since $z(t)$ is decreasing and $w(t)$ is increasing with time. Therefore

$$t_0 \leq K_3(1 - w(0))$$

where $K_3 < \infty$ is a constant.

2) Item 2: From (14) and (17), we have

$$\frac{d}{dt}v(t) \leq \frac{1}{2}\left[\left(1 - \frac{(1-v(t))^2}{1+v(t)}\right)^2 - \left(1 + \frac{(1-v(t))^2}{1+v(t)}\right)v(t)\right].$$

We can rewrite the last inequality as

$$\frac{d}{dt}v(t) \leq \frac{v(t)(1-v(t))(3v(t)-1)}{(1+v(t))^2}.$$

Hence,

$$\frac{(1+v(t))^2 dv(t)}{v(t)(1-v(t))(3v(t)-1)} \leq dt. \tag{39}$$

Note that

$$\frac{(1+v(t))^2}{v(t)(1-v(t))(3v(t)-1)} = \frac{-1}{v(t)} + \frac{2}{1-v(t)} + \frac{8}{3v(t)-1}.$$

It follows that

$$\begin{aligned} & \int_{v(t_0)}^{v(t)} \frac{(1+v(t))^2 dv(t)}{v(t)(1-v(t))(3v(t)-1)} \\ = \quad & -\int_{v(t_0)}^{v(t)} d(\log v(t)) - 2\int_{v(t_0)}^{v(t)} d(\log(1-v(t))) \\ & + \frac{8}{3}\int_{v(t_0)}^{v(t)} d(\log(3v(t)-1)). \end{aligned}$$

The result (19) follows from the above relation and (39).

3) Item 3: Since (18) holds when $t \geq t_0$, $dz(t)/dt = (d/dt)(u+v)(t) \geq 0$. Thus, from (15), we have

$$z(t) \leq \frac{2}{3} + \frac{1}{3}w(t)^2$$

where, recall, $w(t) = v(t) - u(t)$. Now, from the last inequality and (16), it follows that

$$\frac{dw(t)}{(1-w(t)^2)w(t)} \geq \frac{1}{6}dt.$$

Note that

$$\frac{1}{(1-w(t)^2)w(t)} = \frac{1}{2}\frac{1}{1-w(t)} - \frac{1}{2}\frac{1}{1+w(t)} + \frac{1}{w(t)}.$$

Integrating, it follows

$$t - t_0 \leq 3\log\left(\frac{w(t)^2}{w(t_0)^2}\frac{1-w(t_0)^2}{1-w(t)^2}\right),$$

and hence the result asserted in (20).

G. Proof of Corollary 3

By Item 1 of Theorem 5, it suffices to show (21), (22) and (23) for $(u(0), v(0))$ satisfying (17) and (18). The lower bound in (21), together with (22), follows from (19) and the condition $v(t(N)) \geq 1 - \frac{1}{N}$. The upper bound in (21), together with (23), follows from (20) and from the condition $v(t(N)) - u(t(N)) \approx 1 - \frac{1}{N}$.

H. Ternary Signaling on the Line Graph

We analyze the behaviour of the ternary signaling protocol for a line graph with a specific initial configuration. At a sampling instance, a node can contact only one of its neighbours, which in this case, are the nodes to the right and left of it.

Suppose that the initial observations at the nodes are given as – The first U nodes from the left observe 1 and the remaining $V = N - U$ nodes observe 0. The graph and the initial observations at the nodes are illustrated in Figure 1. We will refer to this configuration as $(U, 0, V)$ where the 0 indicates that there are no nodes in state e. Let $\mathcal{U}$ denote the set of nodes observing 1 and let $\mathcal{V}$ denote the set of nodes observing 0.

Under ternary signaling, there are three possible configurations that could result at the next step. The first configuration is the original configuration which occurs when either a node in $\mathcal{U}$ contacts a node in $\mathcal{U}$ or a node in $\mathcal{V}$ contacts a node in $\mathcal{V}$. For the second configuration (call it $(U-1, 1, V)$), node U contacts node $U+1$ and updates its value to e and in the third configuration (call it $U, 1, V-1$), node $U+1$ contacts node U and updates its value to e. The probability to transition from $(U, 0, V)$ to either of the last two configurations, i.e., $(U-1, 1, V)$ and $(U, 1, V-1)$ is equal. The last two configurations are illustrated in the Figures 5 and 6 respectively.

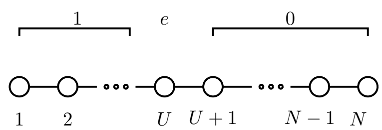

Fig. 5. Configuration $(U-1, 1, V)$. The node U updated to state e.

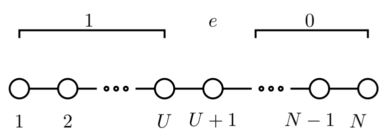

Fig. 6. Configuration $(U, 1, V-1)$. The node $U+1$ updated to state e.

Suppose that the configuration $(U-1, 1, V)$ occurs at the next step. There are three possible configurations which can occur at the subsequent step. The first is the configuration $(U-1, 1, V)$ itself when either a node in $\mathcal{U}$ contacts a node in $\mathcal{U}$ (or contacts node U) or a node in $\mathcal{V}$ contacts a node in $\mathcal{V}$ (or contacts node U). The second configuration is the starting configuration $(U, 0, V)$ and occurs when the node U in state e contacts node $U-1$ and updates its state to 1. The third configuration is a new configuration $(U-1, 0, V+1)$ and occurs when the node U contacts node $U+1$ and updates it's value to 0. Furthermore, the probability to transition from $(U-1, 1, V)$ to either of the last two configurations, i.e., $(U, 0, V)$ and $(U-1, 0, V+1)$ is equal.

Likewise, from the configuration $(U, 1, V-1)$, there are three possible configurations which can occur at the subsequent step – the configuration $(U, 1, V-1)$, the starting configuration $(U, 0, V)$ and the new configuration $(U+1, 0, V-1)$. The probability to transition from $(U, 1, V-1)$ to either of the last two configurations, i.e., $(U, 0, V)$ and $(U+1, 0, V-1)$ is equal.

Let $f_{U,V}$ denote the probability of reaching the all 0 state starting from the $(U, 0, V)$ configuration. From the arguments above, one can check that $f_{U,V}$ satisfies the recursion

$$f_{U,V} = \frac{1}{2} f_{U-1,V+1} + \frac{1}{2} f_{U+1,V-1},$$

with the boundary conditions given by $f_{0,N} = 1$ and $f_{N,0} = 0$. This recursion is just the recursion for the classical voter model and the solution for $f_{U,V}$ is given by

$$f_{U,V} = \frac{V}{N}.$$

Let $\tau_{U,V}$ denote the time to absorption. From standard arguments, it can be shown that

$$\tau_{U,V} = 6UV$$

Note that as opposed to the complete graph case where ternary signaling results in an exponentially low (in N) probability of error, in the line graph for a particular starting state, ternary signaling results in the *same* probability of error as the voter model. This is true because the voter model on the line graph for this particular starting configuration is equivalent to the gambler's ruin problem. The convergence time of the ternary signaling scheme is quadratic in the number of nodes N and is furthermore slower by a factor 6 as compared to the voter model (due to the occurrence of intermediate states with a node in state e).

I. *Proof of Theorem 4*

As in the proof of Theorem 3, we consider the embedded state $K = V - U$. Equation (34) gives the conditional transition probability of $K = k$ updating to $K = k+1$. It can be checked that this probability is minimum when $U + V = N$, i.e, when $(V = \frac{N+k}{2}, U = \frac{N-k}{2})$ and is equal to 0.5. The probability of error is one-half times the the probability of reaching the state $K = 0$ before reaching $K = N$. We obtain an upper bound by assuming that the bias towards the larger K (i.e., $K = k+1$) is always minimum. Hence, we consider a new Markov chain K' with transition probabilities

$$\mathbb{P}(K'(t+1) = k-1 | K'(t) = k) = \frac{1}{2} \tag{40}$$

$$\mathbb{P}(K'(t+1) = k+1 | K'(t) = k) = \frac{1}{2}. \tag{41}$$

The error probability p_k for this Markov chain is the solution to

$$p_k = \frac{1}{2} p_{k+1} + \frac{1}{2} p_{k-1}, \tag{42}$$

for $k = 1, \ldots, N-1$ with boundary conditions $p_0 = \frac{1}{2}$, $p_N = 0$. The upper bound to $f_{U,V}$ is given by

$$f_{U,V} \le p_{V-U} \quad \text{for } U \le V.$$

The recurrence (42) corresponds to the classical gambler's ruin problem and it is well known that $p_k = \frac{N-k}{2N} = \frac{U}{N}$.