Asynchronous Majority Dynamics in Preferential Attachment Trees

Maryam Bahrani
Princeton University, NJ, USA
mbahrani@alumni.princeton.edu

Nicole Immorlica
Microsoft Research, Cambridge, MA, USA
nicimm@microsoft.com

Divyarthi Mohan
Princeton University, NJ, USA
dm23@cs.princeton.edu

S. Matthew Weinberg
Princeton University, NJ, USA
smweinberg@princeton.edu

Abstract

We study information aggregation in networks where agents make binary decisions (labeled incorrect or correct). Agents initially form independent private beliefs about the better decision, which is correct with probability $1/2 + \delta$. The dynamics we consider are asynchronous (each round, a single agent updates their announced decision) and non-Bayesian (agents simply copy the majority announcements among their neighbors, tie-breaking in favor of their private signal).

Our main result proves that when the network is a tree formed according to the preferential attachment model [5], with high probability, the process stabilizes in a correct majority within $O(n \log n / \log \log n)$ rounds. We extend our results to other tree structures, including balanced M-ary trees for any M.

2012 ACM Subject Classification Theory of computation → Algorithmic game theory and mechanism design; Networks → Network dynamics

Keywords and phrases Opinion Dynamics, Information Cascades, Preferential Attachment, Majority Dynamics, non-Bayesian Asynchronous Learning, Stochastic Processes

Digital Object Identifier 10.4230/LIPIcs.ICALP.2020.8

Category Track A: Algorithms, Complexity and Games

Related Version A full version of the paper is available at `https://arxiv.org/abs/1907.05823`.

Funding *S. Matthew Weinberg*: Supported by NSF CCF-1717899.

1 Introduction

Individuals form opinions about the world both through private investigation and through discussion with one-another. A citizen, trying to decide which candidate's economic policies will lead to more jobs, might form an initial belief based on her own employment history. However, her stated opinion might be swayed by the opinions of her friends. The dynamics of this process, together with the social network structure of the individuals, can result in a variety of societal outcomes. Even if individuals are well-informed, i.e., are more likely to have correct than incorrect initial beliefs, certain dynamics and/or network structures can cause large portions of the population to form mistaken opinions.

47th International Colloquium on Automata, Languages, and Programming (ICALP 2020).
Editors: Artur Czumaj, Anuj Dawar, and Emanuela Merelli; Article No. 8; pp. 8:1–8:14
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

A substantial body of work exists modeling these dynamics mathematically, which we overview in Section 1.2. This paper focuses on the model of *asynchronous majority dynamics*. Initially, individuals have private beliefs over a binary state of the world, but no publicly stated opinion. Initial beliefs are independent: Correct with probability $1/2 + \delta$, and Incorrect with probability $1/2 - \delta$. In each time step, a random individual is selected to announce a public opinion. Each time an individual announces a public opinion, they simply copy the majority of their neighbors' announcements, tie-breaking in favor of their private belief. This is clearly naive: a true Bayesian would reason about the redundancy of information among the opinions of her friends, for example. Majority (or other non-Bayesian) dynamics are generally considered a more faithful model of agents with bounded rationality (e.g. voters), whereas Bayesian dynamics are generally considered a more faithful model of fully rational actors (e.g. financial traders). We consider asynchronous announcements[1] which are a more faithful model of human decisions (e.g. citizens deciding which candidate is better).

It's initially tempting to conjecture that these dynamics in a connected network should result in a Correct consensus; after all, the majority is initially Correct (with high probability) by assumption. Nonetheless, it's well-understood that individuals can fail miserably to learn. Suppose for instance that the individuals form a complete graph. Then in asynchronous majority dynamics, whichever individual is selected to announce first will have their opinion copied by the entire network. As this opinion is Incorrect with constant probability, there's a good chance that the entire network makes the wrong decision (this is known as an *information cascade*, and is not unique to asynchonous majority dynamics [4, 7]). So the overarching goal in these works is to understand in *which* graphs the dynamics stabilize in correctness with high probability.

For most previously studied dynamics (discussed in Section 1.2), "correctness" means a Correct consensus. This is because the models terminate in a consensus with probability 1, and the only question is whether this consensus is correct or not. With majority dynamics, it is certainly possible that the process stabilizes without a consensus. To see this, suppose individuals form a line graph. In this case, two adjacent individuals with the same initial belief are likely to form a "road block" (if both announce before their other neighbors), sticking to their initial beliefs throughout the process. In this case, with high probability a constant fraction of individuals terminate with a Correct opinion, but also a constant fraction terminate with an Incorrect opinion. As consensus is no longer guaranteed, we're instead interested in understanding network structures for which the dynamics converge, with high probability, to a *majority* of nodes having the Correct opinion (i.e., if a majority vote were to be taken, would it be correct w.h.p.?).

Prior work shows that, to reach a Correct consensus, it's sufficient for the social network to be sparse (every individual has only a constant number of neighbors) and expansive (every group of individuals have many friends outside the group) [12], and the tools developed indeed make strong use of both assumptions. Many networks of interest, however, like the hierarchy of employees in a corporation, are neither sparse nor expansive. Therefore, the focus of this paper is to push beyond these assumptions and develop tools for more general graphs.

[1] Unlike synchronous models where all agents announce simultaneously.

1.1 Our Results and Techniques

We focus our attention on trees, the simplest graphs outside the reach of prior techniques. In addition to modeling certain types of social networks (including hierarchical ones, or communication networks in which redundancy is expensive), and forming the backbone of many more, trees already present a number of technical challenges whose absence enabled the prior results. We study preferential attachment trees, which are well-studied graphs with rich structure[2]. Our main result is the following:

▶ **Theorem 1.** *Let G be a tree. Then with probability $1 - o(1)^3$, asynchronous majority dynamics in G stabilizes in a* Correct *majority if:*
- *G is formed according to the preferential attachment model.*[4]
- *G is a balanced, M-ary tree of any degree.*[5]

Beyond Prior Tools

In prior work [12], the authors have two key ideas. Without yet getting into full details, one key idea crucially invokes sparsity to claim that most pairs of nodes u, v have distance $d(u, v) = \Omega(\ln n/ \ln\ln n)$, which allows them to conclude that after $O(n \ln n/ \ln\ln n)$ steps, most nodes are announcing a Correct opinion.

Still, just the fact that the dynamics hit a Correct majority along the way does not imply that the Correct majority will hold thru termination. To wrap up, they crucially invoke expansiveness (building off an argument of [24]) to claim that once there is a Correct majority, it spreads to a Correct consensus with high probability.

Both properties are necessary for prior work, and both properties fail in trees. For instance, the star graph is a tree, but $d(u, v) \leq 2$ for all u, v (precluding their "majority at $O(n \ln n/ \ln\ln n)$" argument). Additionally, trees are not expansive. In particular, the line graph discussed earlier is a tree which hits a Correct majority at some point (as this tree happens to be sparse), but does not converge to consensus, so there is no hope for an argument like this. However, we believe that the process stabilizes in a correct majority in all trees.

▶ **Conjecture.** *Let G be any tree. Then the asynchronous majority dynamics stabilizes in a* Correct *majority with probability $1 - o(1)$.*

New Tools

Our main technical innovation is an approach to reason about majority without going through consensus. Specifically, we show in Sections 5 and 6 for preferential attachment trees, or balanced M-ary trees, that with probability $1 - o(1)$, a $1 - o(1)$ fraction of nodes have *finalized* after $O(n \ln n/ \ln\ln n)$ steps. That is, after $O(n \ln n/ \ln\ln n)$ steps, most (but not all) of the network has stabilized. The main barrier to extending our results to general trees is Section 5, as we require additional structure on the graphs to prove that the process stabilizes quickly. We postpone further details to Section 5, but just wish to highlight this approach as a fairly significant deviation from prior work.

[2] The more general preferential attachment graphs are a popular model of real networks.
[3] As $n \to \infty$ the probability converges to 1, where n is the number of nodes in G.
[4] That is, G is created by adding nodes one at a time. When a node is added, it attaches a single edge to a random previous node, selected proportional to its degree.
[5] That is, G can be rooted at some node v. All non-leaf nodes have M children, and all leaves have the same distance to v.

From here, our task is now reduced to showing that a CORRECT majority exists w.h.p. after $O(n \ln n / \ln \ln n)$ steps. Our main insight here is that most nodes with $d(u, v) = O(\ln n / \ln \ln n)$ must have some high-degree nodes along the path from u to v. We prove that such nodes act like a "road block," causing announcements on either side to be independent with high probability (and all nodes with $d(u, v) = \Omega(\ln n / \ln \ln n)$ can be handled with similar arguments to prior work).

1.2 Related Work

Information aggregation in social networks is an enormous field, and we will not come close to overviewing it in its entirety. Below, we'll briefly summarize the most related literature, restricting attention to works that consider two states of the world and independent initial beliefs are independently CORRECT with probability $1/2 + \delta$.

Bayesian Dynamics

In Bayesian models, agents are fully rational and sequentially perform Bayesian updates to their public opinion based on the public opinions of their neighbors. Seminal works of Banerjee [4] and Bikchandani, Hirshleifer, and Welch [7] first identified the potential of information cascades in this model. Subsequent works consider numerous variations, aiming to understand what assumptions on the underlying network or information structure results in CORRECT consensus [28, 3, 2]. Many other works studied repeated interactions of Bayesian agents in Social Networks [13, 27, 20, 26, 25, 23]. While the high-level goals of these works align with ours, technically they are mostly unrelated as we consider non-Bayesian dynamics.

Voter and DeGroot Dynamics

Prior work also considered other non-Bayesian dynamics. In voter dynamics, individuals update by copying a random neighbor [10, 18]. Similar dynamics (such as 3-majority, or k-majority) are analyzed from a distributed computing perspective with an emphasis on the rate of convergence to consensus [6, 15, 14]. In the DeGroot model, individuals announce an opinion in $[0, 1]$ (as opposed to $\{0, 1\}$), and update by averaging their neighbors [11, 16]. The biggest difference between these works and ours is that consensus is reached with probability 1 in these models on any connected graph, which doesn't hold for majority dynamics.

Majority Dynamics

The works most related to ours consider majority dynamics. Even synchronous majority dynamics may not result in a consensus (consider again the line graph). These works, like ours, therefore seek to understand what graph structures result in a CORRECT majority. Mossel, Neeman, and Tamuz study synchronous majority dynamics and prove that a CORRECT majority arises as long as the underlying graph is sufficiently symmetric, or sufficiently expansive (in the latter case, they prove that the network further reaches consensus) [24]. Feldman et al. study asynchronous majority dynamics and prove that a CORRECT consensus arises when the underlying graph is sparse and expansive [12]. Work of [29] further studies "retention of information," which asks whether *any* recovery procedure (not necessarily a majority vote) at stabilization can recover the ground truth with high probability. In connection to these, our work simply pushes the boundary beyond what classes of graphs are understood in prior work.

The key difference between the synchronous and asynchronous models is captured by the complete graph. In asynchronous dynamics, a CORRECT majority occurs only with probability $1/2 + \delta$, whereas in synchronous dynamics a CORRECT consensus occurs with probability $1 - \exp(-\Omega(n))$. This is because in step one, every node simply announces their private belief, and in step two everyone updates to the majority, which is CORRECT with probability $1 - \exp(-\Omega(n))$. So while the models bear some similarity, and some tools are indeed transferable (e.g. the expansiveness lemma of [24] used in [12]), much of the anlayses will necessarily diverge.

Preferential Attachment and Balanced M-ary Trees

There is also substantial prior work studying aggregation dynamics in trees. Here, the most related work is [19, 22], which studies synchronous majority dynamics in balanced M-ary trees. Less related are works which study "bottom-up" dynamics in balanced M-ary trees [21, 31, 30], k-majority dynamics in preferential attachment trees [1], or model cascades themselves as a preferential attachment tree [17]. While these works provide ample motivation for restricting attention to preferential attachment trees, or balanced M-ary trees, they bear no technical similarity to ours.

2 Model and Preliminaries

We consider an undirected tree $G = (V, E)$ with $|V(G)| = n$ and $|E(G)| = m$. We denote by $\deg(v)$ the degree of a node $v \in V(G)$, $N(v)$ to be its neighbors $\{u, (u, v) \in E\}$, and $d(u, v)$ to be the length of the unique path between u and v, and let $P(u, v)$ denote the ordered list of vertices on this path (i.e. starting with u and ending with v). We'll also denote by $D(G) = \max_{u,v}\{d(u, v)\}$ the diameter of G.

Individuals initially have one of two private beliefs, which we'll refer to as CORRECT (or 1) and INCORRECT (or 0). That is, each $v \in V(G)$ receives an independent private signal $X(v) \in \{0, 1\}$, and $\Pr[X(v) = 1] = 1/2 + \delta$, for some constant $0 < \delta < 1/2$.

Individuals also have a *publicly announced* opinion (which we will simply refer to as an announcement). We define $C^t(v) \in \{\perp, 0, 1\}$ to be the public announcement of $v \in V(G)$ at time t. Initially, no announcements have been made, *i.e.* $C^0(v) = \perp$ for all v. In each subsequent step, a *single node* v^t is chosen uniformly at random from $V(G)$ and updates her announcement (announcements of all other nodes stay the same)[6]. v^t updates her announcement using *majority dynamics*. That is, if $N_1^t(v)$ denotes the number of v's neighbors with a CORRECT announcement at time t, and $N_0^t(v)$ denotes the number of v's neighbors with an INCORRECT announcement, then:

$$C^t(v) = \begin{cases} 1 & \text{if } N_1^{t-1}(v) > N_0^{t-1}(v), \text{ and } v = v^t, \\ 0 & \text{if } N_1^{t-1}(v) < N_0^{t-1}(v), \text{ and } v = v^t, \\ X(v) & \text{if } N_1^{t-1}(v) = N_0^{t-1}(v), \text{ and } v = v^t, \\ C^{t-1}(v) & \text{if } v \neq v^t. \end{cases}$$

Note that we will treat δ as an absolute constant. Therefore, the only variable taken inside Big-Oh notation is n, the number of nodes (and, for instance, when we write $o(1)$ we mean any function of n that approaches 0 as n approaches ∞).

[6] This makes the process *asynchronous*.

As shown in [12], it is easy to see that in any network this process stabilizes with high probability in $O(n^2)$ steps. That is, the network reaches a state where no node will want to change its announcement and thus the process terminates.

2.1 Concentration Bounds and Tools from Prior Work

Our work indeed makes use of some tools from prior work to get started, which we state below. The concept of a *critical time*, defined below, is implicit in [12].

▶ **Definition 2.** *The* critical time[7] *from u to u, $T(u,u)$, is the first time that node u announces. The critical time from u to v, $T(u,v)$, is recursively defined as the first time that v announces* after *the critical time from u to x, where x is the neighbor of v in $P(u,v)$. We further denote the* critical chain *from u to v as the ordered list of critical times from u to x for all x on $P(u,v)$.*

The following lemma is a formal statement of ideas from prior work (a proof appears in Appendix A of the full version). To parse it, it will be helpful to think of the process as first drawing a countably infinite sequence S of nodes to announce, which then allows each $C^t(v)$ to be written as a deterministic function of the random variables $\{X(u), u \in V\}$. Lemma 3 below states that in fact, for early enough t, initial beliefs for only a proper subset of V suffice.

▶ **Lemma 3** ([12])**.** *For all t, and all v, $C^t(v)$ can be expressed as a function of the subset of signals $\{X(u), T(u,v) \leq t\}$.*

The final theorem we take from prior work is due to Mossel et al., and is used to claim that at minimum the *expected* number of CORRECT nodes at termination is at least $(1/2+\delta)n$.

▶ **Theorem 4** ([24])**.** *Let f be an odd, monotone Boolean function. Let $X_1, \dots, X_n$ be input bits, each sampled i.i.d. from a distribution that is 1 with probability $p \geqslant 1/2$ and 0 otherwise. Then $\mathbb{E}[f(X_1, \dots, X_n)] \geqslant p$.*

Note that, as long as v has announced at least once by t, $C^t(v)$ is an odd, monotone, Boolean function in variables $\{X(u), u \in V\}$,[8] and therefore $\Pr[C^t(v) = 1] \geq 1/2 + \delta$ for all v and $t \geq T(v,v)$.

Finally, we'll make use of the following concentration bound on $T(u,v)$ repeatedly. Its proof is a simple application of a Chernoff bound and appears in Appendix A of the full version.

▶ **Lemma 5.** *For all $0 < \beta < 1$:*

- $\Pr[T(u,v) > 8 \cdot \max\{\ln(1/\beta), d(u,v)+1\} \cdot n] \leq \beta^2$.
- $\Pr[T(u,v) < (d(u,v)+1) \cdot \beta \cdot n] \leq e^{-\beta d(u,v)(1-\beta)^2/3}$.
- $\Pr[T(u,v) < (d(u,v)+1) \cdot \beta \cdot n] \leq (e\beta)^{d(u,v)} = e^{(1+\ln\beta)\cdot d(u,v)}$.

[7] This definition can be naturally extended to any general graph.
[8] That is, flipping all $X(v)$ simultaneously to $1 - X(v)$ would cause $C^t(v)$ to flip (odd), and changing any subset of initial beliefs from 0 to 1 cannot change $C^t(v)$ from 1 to 0 (monotone).

3 Key Concepts

Before getting into our proofs, we elaborate some key concepts that will be used throughout. In Proposition 7 below, we analyze the connection between critical chains and switches in announcements. Intuitively, Proposition 7 is claiming that every fresh announcement can cause other nodes to switch a previous announcement along critical chains, but that these are the *only* switches that can occur.

▶ **Definition 6.** *Let v change her announcement at t, and her previous announcement be made at $t' > 0$. We say that node u is a* cause *of v changing her announcement at t if $C^t(u) = C^t(v)$, and $C^{t'}(u) \neq C^t(v)$. Observe that every such change in announcement has a cause.*

▶ **Proposition 7.** *If $C^t(v) \neq C^{t-1}(v)$, then there exists a node u such that:*

- *$t = T(u, v)$ (i.e. the influence of u just reaches v at time t).*
- *$C^{T(u,u)}(u) = C^t(v)$ (i.e. v is updating to match u's initial announcement).*
- *Denote $u = x_0, x_1, \dots, x_{d(u,v)} = v$ the path $P(u, v)$. Then* every *x_i, $i > 0$, has $C^{T(u,x_i)-1}(x_i) = C^{t-1}(v)$ and $C^{T(u,x_i)}(x_i) = C^t(v)$, and x_{i-1} caused this change (i.e.* every *node along the path from u to v* changed *to match u's initial announcement).*

Proof. The proof proceeds by induction on t. Consider $t = 1$ as a base case. If $C^1(v) \neq C^0(v) = \bot$, then it must be because v announced at time 1, meaning that $1 = T(v, v)$ as desired.

Now assume that for all v and all $t' < t$ the claim holds, and consider time t. If v does not announce at time t then the claim vacuously holds. If v announces at time t but does *not* change their announcement, then again the claim vacuously holds. If v announces at time t for the first time, then v itself is the desired u and the claim holds. The remaining case is if v changes their previous announcement that was made at time $t' < t$ (and v did not announce between t' and t).

Let's consider the state of affairs at time t', when v announced some opinion A. This means that, at time t', a majority (tie-breaking for $X(v)$) of v's neighbors were announcing A. Yet, at time t, a majority (tie-breaking for $X(v)$) of v's neighbors were announcing $B = 1 - A$. Therefore, *some* node adjacent to v must have switched its announcement to B at some $t'' \in (t', t)$, and stays B till time t (and caused the change). Call this node x. We now wish to invoke the inductive hypothesis for x at t''.

The inductive hypothesis claims there there is some u (maybe $u = x$) such that u made B as its first announcement, and then every node y along the critical chain from u to x switched from A to B at $T(u, y)$ (caused by its predecessor), and that $t'' = T(u, x)$. Let's first consider the case that v is not on the path from u to x (and therefore x is on the path from u to v, since they are adjacent). Then as $T(u, x) = t'' \in (t', t)$, and v does not announce in (t', t), we see that $T(u, v) = t$ (immediately by definition of critical times). Moreover, as $P(u, v)$ is simply $P(u, x)$ concatenated with v, the inductive hypothesis already guarantees that u announced B at $T(u, u)$, and that every node y on $P(u, v)$ switched from A to B at $T(u, y)$. So the last step is to show that in fact v *must* not be on the path from u to x, and then the inductive step will be complete.

Finally, we show that we cannot have v on the path from u to x, completing the inductive step. Assume for contradiction that v were on the path from u to x. Then as $t'' = T(u, x)$, we would necessarily have $t' \geq T(u, v)$ (immediately from definition of critical times). However, by hypothesis, $C^{t''}(v) = C^{t'}(v) = A$ (the first equality is simply because v does not announce in $(t', t'']$), contradicting the inductive hypothesis that v caused x to change (because of u), which would imply instead that $C^{T(u,v)}(v) = C^{t''}(v) = B$. So v cannot be on the path from u to x. ◀

Below, we make use of Proposition 7 to prove that, in any tree[9], the process terminates quickly (proof in Appendix B of the full version). Note that [12] already proves that the process on trees terminates with probability $1 - o(1)$ after $O(n^2)$ steps, so Corollary 8 is a strict improvement when the diameter $D(G) = o(n)$.

▶ **Corollary 8.** *Let T_{stable} denote the last time that a node changes its announcement. Then with probability $1 - o(1)$, $T_{stable} \leq 8 \cdot \max\{2\ln(n), D(G) + 1\} \cdot n$.*

Finally, we prove one last proposition which will be used in future sections regarding the probability that a single node announces Correct throughout the process (the proof appears in Appendix B of the full version). Beginning with v's first announcement, because the graph is a tree, prior to v's first announcement all of v's neighbors' announcements are independent. Therefore, it initially seems like we should expect v's initial announcement to be Correct except with probability exponentially small in $\deg(v)$ – indeed, this would hold if the dynamics were synchronous. However, since the dynamics are *asynchronous*, there's a good chance that v announces before any of its neighbors and simply announces $X(v)$. That is, the probability that v's initial announcement is Incorrect is at least $\frac{1/2-\delta}{\deg(v)}$, so we cannot hope for such strong guarantees. This observation highlights one (of several) crucial differences between synchronous and asynchronous dynamics. Still, the proposition below shows roughly that the only bad event is v announcing before many of its neighbors. Below for a set S, we'll use $C_S^t(v)$ to denote the following modified dynamics: First, set $C_S^t(v) = $ Incorrect for all $v \in S$, and all t. Then, run the asynchronous majority dynamics as normal. In other words, the modified dynamics hard-code an Incorrect announcement for all nodes in S and otherwise run asynchronous majority dynamics as usual (this extension will be necessary for a later argument).

▶ **Definition 9.** *We say that a node v is* safe thru T *if $C^t(v) \in \{\perp, 1\}$ for all $t \leq T$. We further say that a node v is* safe thru T, even against S *if $C_S^t(v) \in \{\perp, 1\}$ for all $t \leq T$.*

▶ **Proposition 10.** *For all a, there exist constants[10] b, c such that for any S with $|S| = a$, $v \notin S$ and $T \leq n \cdot e^{b \deg(v)}$, v is safe thru T, even against S, with probability at least $1 - c/\deg(v)$.*

4 Forming an Initial Majority

In this section, we prove that in any tree, a Correct majority forms after a near-linear number of steps (but may later fade). The main idea is to show that the announcements of most pairs of nodes are independent with probability $1 - o(1)$, and use Chebyshev's inequality to show that the number of Correct announcements therefore concentrates around its expectation. The independence argument is the crux of the proof. To show it, we consider three cases depending on the length and degree sequence of the path between a pair of nodes. If the path is long, then, similar to prior work, there is simply not enough time for the pair to influence each other. If the path is short, but (some of) the intermediate nodes have high degrees, then these effectively block influence because the announcements of these high-degree nodes is effectively independent of what's happening on the path. Finally, if the path is short and the intermediate nodes have low degrees, then the pair certainly may influence each other. However, a counting argument shows there can only be a vanishingly small fraction of such pairs. The main result of this section is the following:

[9] Proposition 7 and Corollary 8 hold for any general graph. Since we are only interested in trees, we restrict our proofs to just trees for brevity.
[10] does not depend on n, but may depend on δ.

▶ **Theorem 11** (Majority in trees). *For sufficiently large n, any tree on n nodes and any $T \leq \frac{n \ln n}{32 \ln \ln n}$, after T steps, with probability at least $1-O(e^{-\ln n/(24 \ln \ln n)})$, the announcements of at least $(\frac{1}{2} + \frac{\delta}{2} - e^{-T/n}) \cdot n$ nodes are* Correct *.*

Further, for all constants $\gamma > 0$, there exists a constant $\alpha > 0$ such that for sufficiently large n, when $T \leq \frac{n \ln^{1-\gamma} n}{\ln \ln n}$, after T steps, with probability at least $1-n^{-\alpha}$, the announcements of at least $(\frac{1}{2} + \frac{\delta}{2} - e^{-T/n}) \cdot n$ nodes are Correct *.*

First, we analyze the expected number of Correct nodes using Theorem 4 (proof in Appendix C of the full version). Note that, the probability that a node v has announced by T is at least $(1 - e^{-T/n})$[11].

▶ **Lemma 12.** *At any time T, the expected number of nodes v with $C^T(v) =$ Correct is at least $(1/2 + \delta - e^{-T/n})n$.*

From here, we now need to show that the number of Correct announcements concentrates around its expectation. To this end, we'll show that most pairs of nodes can be written as functions of disjoint initial beliefs, and are therefore independent. Ideas from [12] formally show that this suffices:

▶ **Definition 13.** *We say that two nodes u, v are ε-disjoint at t if there exist random variables X_u, X_v, written as functions of disjoint sets of initial beliefs (and therefore independent), such that $\Pr[C^t(u) = X_u] \geq 1 - \varepsilon$ and $\Pr[C^t(v) = X_v] \geq 1 - \varepsilon$.*

▶ **Lemma 14** (Inspired by [12]). *Let ε^t_{uv} be such that u, v are ε^t_{uv}-disjoint at t. Then if $\sum_{u,v} \varepsilon^t_{uv} = D$, the number of* Correct *nodes at time t is within $\delta n/2$ of its expectation with probability $1 - \frac{4n+16D}{\delta^2 n^2}$.*

So our remaining task is to upper bound $\sum_{u,v} \varepsilon^t_{uv}$, and this is the point where we diverge from prior work. For a given pair u, v, there are three possible cases. Below, case one is most similar to prior work, and cases two/three are fairly distinct.

Case One: Long Paths

One possibility is that $d(u, v) \geq f(n)$, for some $f(n)$ to be decided later. The following proposition implies that at any time T, the announcements of pairs of nodes at a large enough distance are almost independent. The proof is provided in Appendix C of the full version.

▶ **Proposition 15.** *Let $d(u, v) \geq \max\{kT, f(n)\}$ for $k \geq 4$. Then $\varepsilon^T_{uv} \leq 2e^{-f(n)/24}$, and $\varepsilon^T_{uv} \leq 2e^{(1-\ln k)\cdot f(n)}$.*

Case Two: Short Paths A

Another possibility is that $d(u, v) < f(n)$. Here, there will be two subcases. First, maybe it's the case that $d(u, v)$ is small *and* the product of degrees on the path from u to v is small. In this case, it very well could be that ε^t_{uv} is large, which is bad. However, we prove that there cannot be many such pairs (and so in total they contribute $o(n^2)$ to the sum). The following lemma shows in fact that even if we remove the restriction that $d(u, v) = O(T)$, there simply cannot be many pairs of nodes such that the product of degrees on $P(u, v)$ is small (proof in Appendix C of the full version).

[11] The probability that v wasn't chosen in all T rounds is $\left(1 - \left(1 - \frac{1}{n}\right)^T\right)$.

▶ **Lemma 16.** *Let K be the set of pairs of nodes (u, v) such that $\prod_{w \in P(u,v)} \deg(w) \leq X$. Then $|K| \leq Xn/2$.*

Case Three: Short Paths B

The final possibility is that $d(u, v) < f(n)$, and also that $\prod_{w \in P(u,v)} \deg(w)$ is large. In this case, we will prove that with probability $1 - o(1)$ there is some block in $P(u, v)$ causing u's and v's announcemnts to be independent (proof in Appendix C of the full version).

▶ **Definition 17.** *We say that a node $x \in P(u, v)$* cuts *u from v thru T if some node y in $P(u, x)$ is safe thru T even against S_y, where S_y are y's (at most) two neighbors in $P(u, v)$.*

▶ **Lemma 18.** *Let T be any time and p_x be the probability that x cuts u from v thru T* and also *cuts v from u thru T. Then u and v are p_x-disjoint at T.*

Next, we wish to show that with good probability there is indeed a node on $P(u, v)$ that cuts u from v and also v from u (proofs in Appendix C of the full version).

▶ **Lemma 19.** *There exist absolute constants b, d such that for any pairs of nodes $u, v \in V$ with $\prod_{w \in P(u,v)} \deg(w) = X$, $d(u, v) \leq \frac{\ln(X)}{d}$, and $T \leq ne^{bX^{1/(4d(u,v))}}$, there exists an x such that that with probability $1 - X^{-1/8}$, x cuts u from v thru T and also v from u thru T.*

▶ **Corollary 20.** *There exist absolute constants b, d such that for pairs of nodes any $u, v \in V$ with $\prod_{w \in P(u,v)} \deg(w) = X$, $d(u, v) \leq \frac{\ln(X)}{d}$, and $T \leq ne^{bX^{1/(4d(u,v))}}$), u and v are $X^{-1/8}$-disjoint at T.*

Now, we'll put together case one, Lemma 18 and Corollary 20 together to prove Theorem 11, which is mostly a matter of setting parameters straight (and appears in Appendix C of the full version).

To conclude, at this point we have proven that a majority takes hold after $n\frac{\ln n}{32 \ln \ln n}$ steps for any tree. The remaining work is to prove that it does not disappear.

5 Stabilizing Quickly

In this section, we identify properties of a tree which cause it to stabilize quickly. Our main theorem will then follow by proving that both balanced M-ary trees and preferential attachment trees have this property. The main idea is to consider nodes that are "close" to leaves in the following formal sense:

▶ **Definition 21.** *We say that a node v is an (X, Y)-leaf in G if there exists a rooting of G such that v has $\leq X$ descendants, and the longest path from v to one of its descendants is at most Y. Note that leaves are $(0, 0)$-leaves. When we refer to a node's parent, children, or descendants, it will be with respect to this rooting.*

▶ **Definition 22.** *We say that a node v is:*

- finalized *at T, if $C^t(v) = C^T(v)$ for all $t \geq T$.*
- nearly-finalized *at T with respect to u if there exists a $t' \geq T$ such that v is finalized at t' and for all $t \in (T, t')$ when v announces, it either updates $C^t(v) = C^t(u)$, if $C^t(u) \neq \perp$, or $C^t(v) = C^{t-1}(v)$, if $C^t(u) = \perp$.*

Intuitively, a node is finalized if it is done changing its announcement. A node v is nearly-finalized with respect to u if v is not quite finalized, but changes in u are the only reason why v would change its announcement (and moreover, v will copy u every announcement until v finalizes).

The main result of this section is as follows:

▶ **Theorem 23.** *Let v be an (X, Y)-leaf. Then with probability $1 - Xe^{-T/nY}$, v is nearly-finalized at T with respect to its parent.*

The main insight for the proof of Theorem 23 will be the following lemmas. Below, Lemma 24 asserts that once all of v's children are nearly-finalized with respect to v, any changes in v's opinion are to copy its parent, and Lemma 25 builds off this to claim that we can relate the time until v nearly-finalizes to its critical times. Importantly, Lemma 25 does *not* require *all* critical paths to hit v, but only those from its descendents.

▶ **Lemma 24.** *Let all of v's children be nearly-finalized with respect to v at T, and let u be v's parent. Let also $t > t' \geq T$ be two timesteps during which v announced. Then if $C^t(v) \neq C^{t-1}(v)$, we must have $C^t(v) = C^t(u)$.*

▶ **Lemma 25.** *Let $T_v := \max\{T(x, v), x \text{ is a descendant of } v\}$. Then v is nearly-finalized at T_v with respect to its parent.*

These above lemmas suffice to prove Theorem 23. The proofs of these lemmas and Theorem 23 appear in Appendix D of the full version.

We will also need the following implications of Theorem 23. Below, Lemma 26 will be helpful in proving Corollary 27. Corollary 27 lets us claim that while nearly-finalized nodes are not themselves finalized, their existence implies the existence of other finalized nodes. This will be helpful in wrapping up in the following section, since the process only terminates once nodes are finalized. The proofs of Lemma 26 and Corollary 27 appear in Appendix D of the full version.

▶ **Lemma 26.** *For any $t > T_v$, if a child of v changes their announcement at t, v becomes finalized at t.*

▶ **Corollary 27.** *For any T, with probability $1 - e^{-T/n}$, v has $\lfloor(\deg(v) - 1)/2\rfloor$ children who are finalized at $T_v + T$.*

Moreover, if v is an (X, Y)-leaf and finalized at $t \geq T_v$, then with probability $1 - Xe^{-T/nY}$, all of v's descendants are finalized at $t + T$.

6 Wrapping Up: Preferential Attachment and Balanced *M*-ary Trees

In this section, we show how to make use of Theorem 23 to conclude that a $1 - o(1)$ fraction of nodes are finalized by $\frac{n \ln n}{32 \ln \ln n}$. Proofs for the two cases follow different paths, but both get most of their mileage from the developments in Section 5.

6.1 Preferential Attachment Trees Stabilize Quickly

Let's first be clear what we mean by a preferential attachment tree.[12]

[12] Note that this is the standard definition of preferential attachment used for heuristic arguments, e.g. [5]. Most prior rigorous work uses a slightly modified definition that produces a forest instead of a tree in order to rigorously analyze (say) the degree distribution [9, 8]. As we are only interested in (fairly loose) bounds on the degrees, our results are rigorous in the standard model.

▶ **Definition 28** (Preferential Attachment Tree). *n nodes arrive sequentially, attaching a single edge to a pre-existing node at random proportional to its degree. Specifically:*
- *Let v_i denote the i^{th} node to arrive.*
- *Let $\deg_t(v_i)$ denote the degree of node v_i after a total of t nodes have arrived.*
- *There is a special node v_0, which only v_1 connects to upon arrival, and no future nodes.*
- *When v_{i+1} arrives, v_{i+1} attaches a single edge to a previous node, choosing node v_j, $j \in [1, i]$, with probability $\frac{\deg_i(v_j)}{2i-1}$.*

Our main argument for preferential attachment trees is that most nodes are in a "good" subtree, defined below. All subsequent proofs are in Appendix E of the full version. At a high level the plan is as follows: first, we prove that because most nodes are (X, Y)-leaves for small X, Y, these nodes quickly become nearly-finalized. Next, we prove that most such nodes are part of a small subtree whose parent is likely to be safe thru the entire process. Therefore, the parent of this subtree is finalized early, and once the subtree becomes nearly-finalized, it finalizes quickly as well.

▶ **Definition 29.** *Say that a subtree rooted at v is* good *if:*
- *v is a (X, Y)-leaf, for $X = \ln^{O(1)} n$ and $Y = O(\ln \ln n)$.*
- *v's parent has degree at least $\ln^{\Omega(1)} n$.*

▶ **Proposition 30.** *Let the subtree rooted at v be good, and let the diameter of the entire graph be $O(\ln n)$. Then with probability $1 - o(1)$, the entire subtree rooted at v is finalized by $n\frac{\ln n}{32 \ln \ln n}$.*

▶ **Proposition 31.** *For a tree built according to the preferential attachment model, the following simultaneously hold with probability $1 - o(1)$.*
- *$n - o(n)$ nodes are in good subtrees.*
- *The diameter of the entire graph is $O(\ln n)$.*

▶ **Theorem 32.** *A tree built according to the preferential attachment model stabilizes in a* Correct *majority with probability $1 - o(1)$.*

The proofs for Proposition 30, Proposition 31, and Theorem 32 appear in Appendix E of the full version.

6.2 Balanced *M*-ary Trees Stablize Quickly

Let's first be clear what we mean by a balanced M-ary tree.

▶ **Definition 33.** *We say a tree is a balanced M-ary tree if there is a root v such that all non-leaf nodes have exactly M children, and all root-leaf paths have the same length.*

Our plan of attack is as follows (all proofs are in Appendix E of the full version). First, the case for large M (say, $M > \ln n$) is actually fairly straight-forward as a result of Proposition 10. This is because every pair of nodes has a high-degree block on their path, meaning that the "Case Three" argument used in Section 4 actually applies all the way until the process terminates. The $M \leq \ln n$ case is more interesting, and requires the tools developed in Section 5.

Here, the plan is as follows. Corollary 27 roughly lets us claim that all nearly-finalized nodes must have a decent number of finalized children, and moreover that all these finalized children have finalized descendents. Iterating this counting inductively through children, we see that actually most descendents of nearly-finalized nodes of sufficient height must themselves be finalized.

Formally, the approach is to first get a bound on the height for which we can claim that nodes are indeed nearly-finalized with high probability (Corollary 34, immediately from Theorem 23). $\ln\ln n$ turns out to be a good choice.

▶ **Corollary 34.** *Let v be distance h from a leaf. Then v is an $(2M^h, h)$-leaf, and therefore v is nearly-finalized with respect to its parent at $\frac{n\ln n}{64\ln\ln n}$ with probability $1 - 2M^h \cdot e^{-\frac{\ln n}{64h\ln\ln n}} = 1 - e^{-\frac{\ln n}{64h\ln\ln n} + h\ln(2M)}$.*

In particular, if $h = o(\sqrt{\frac{\ln n}{\ln\ln n \cdot \ln M}})$, then v is nearly-finalized with respect to its parent at $\frac{n\ln n}{64\ln\ln n}$ with probability $1 - o(1)$.

▶ **Proposition 35.** *Let v have height $h = \ln\ln n$ in a balanced M-ary tree for $M \leq \ln n$. Then with probability $1 - o(1)$, at most $2M^h \cdot (2/3)^h = o(M^h)$ descendents of v are not finalized by $\frac{n\ln n}{32\ln\ln n}$.*

▶ **Theorem 36.** *Any M-ary tree stablizes in a* Correct *majority with probability $1 - o(1)$.*

The proofs of Proposition 35 and Theorem 36 appear in Appendix E of the full version.

References

1 Mohammed Amin Abdullah, Michel Bode, and Nikolaos Fountoulakis. Local majority dynamics on preferential attachment graphs. In *Proceedings of the 12th International Workshop on Algorithms and Models for the Web Graph - Volume 9479*, WAW 2015, pages 95–106, Berlin, Heidelberg, 2015. Springer-Verlag. doi:10.1007/978-3-319-26784-5_8.

2 Daron Acemoglu, Munther A. Dahleh, Ilan Lobel, and Asuman Ozdaglar. Bayesian learning in social networks. *Review of Economic Studies*, 78(4):1201–1236, 2011.

3 Abhijit Banerjee and Drew Fudenberg. Word-of-mouth learning. *Games and Economic Behavior*, 46(1):1–22, January 2004. URL: http://ideas.repec.org/a/eee/gamebe/v46y2004i1p1-22.html.

4 Abhijit V. Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817, 1992.

5 Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999. doi:10.1126/science.286.5439.509.

6 Luca Bechetti, Andrea E.F. Clementi, Emanuele Natale, Francesco Pasquale, and Luca Trevisan. Stabilizing consensus with many opinions. In *ACM Symposium on Discrete Algorithms (SODA)*, 2016.

7 Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change in informational cascades. *Journal of Political Economy*, 100(5):992–1026, October 1992.

8 Béla Bollobás and Oliver Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24(1):5–34, 2004. doi:10.1007/s00493-004-0002-2.

9 Béla Bollobás, Oliver Riordan, Joel Spencer, and Gábor E. Tusnády. The degree sequence of a scale-free random graph process. *Random Struct. Algorithms*, 18(3):279–290, 2001. doi:10.1002/rsa.1009.

10 Peter Clifford and Aidan Sudbury. A model for spatial conflict. *Biometrika*, 60(3):581–588, 1973. URL: http://www.jstor.org/stable/2335008.

11 Morris H. DeGroot. Reaching a consensus. *Review of Economic Studies*, 69(345):118–121, 1974.

12 Michal Feldman, Nicole Immorlica, Brendan Lucier, and S. Matthew Weinberg. Reaching consensus via non-bayesian asynchronous learning in social networks. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2014, September 4-6, 2014, Barcelona, Spain*, 2014. doi:10.4230/LIPIcs.APPROX-RANDOM.2014.192.

13 Douglas Gale and Shachar Kariv. Bayesian learning in social networks. *Games and Economic Behavior*, 45(2):329–346, 2003. Special Issue in Honor of Robert W. Rosenthal. doi:10.1016/S0899-8256(03)00144-1.

14 Mohsen Ghaffari and Johannes Lengler. Nearly-tight analysis for 2-choice and 3-majority consensus dynamics. In *ACM Symposium on Principles of Distributed Computing (PODC)*, 2018.

15 Mohsen Ghaffari and Merav Parter. A polylogarithmic gossip algorithm for plurality consensus. In *ACM Symposium on Principles of Distributed Computing (PODC)*, 2016.

16 Benjamin Golub and Matthew O. Jackson. Naïve learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1):112–149, 2010.

17 Vicenç Gómez, Hilbert J Kappen, and Andreas Kaltenbrunner. Modeling the structure and evolution of discussion cascades. In *Proceedings of the 22nd ACM conference on Hypertext and hypermedia*, pages 181–190. ACM, 2011.

18 Richard A. Holley and Thomas M. Liggett. Ergodic theorems for weakly interacting infinite systems and the voter model. *The Annals of Probability*, 3(4):643–663, 1975. URL: http://www.jstor.org/stable/2959329.

19 C Douglas Howard. Zero-temperature ising spin dynamics on the homogeneous tree of degree three. *Journal of applied probability*, 37(3):736–747, 2000.

20 Y. Kanoria and O. Tamuz. Tractable bayesian social learning on trees. *IEEE Journal on Selected Areas in Communications*, 31(4):756–765, 2013.

21 Yashodhan Kanoria and Andrea Montanari. Subexponential convergence for information aggregation on regular trees. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 5317–5322. IEEE, 2011.

22 Yashodhan Kanoria, Andrea Montanari, et al. Majority dynamics on trees and the dynamic cavity method. *The Annals of Applied Probability*, 21(5):1694–1748, 2011.

23 E. Mossel, N. Olsman, and O. Tamuz. Efficient bayesian learning in social networks with gaussian estimators. In *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2016.

24 Elchanan Mossel, Joe Neeman, and Omer Tamuz. Majority dynamics and aggregation of information in social networks. In *Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2013.

25 Elchanan Mossel, Allan Sly, and Omer Tamuz. Asymptotic learning on bayesian social networks. *Probability Theory and Related Fields*, 2014.

26 Manuel Mueller-Frank. A general framework for rational learning in social networks. *Theoretical Economics*, 8(1):1–40, 2013.

27 Dinah Rosenberg, Eilon Solan, and Nicolas Vieille. Informational externalities and emergence of consensus. *Games and Economic Behavior*, 66(2):979–994, 2009.

28 Lones Smith and Peter Sorensen. Pathological outcomes of observational learning. *Econometrica*, 68(2):371–398, March 2000. URL: http://ideas.repec.org/a/ecm/emetrp/v68y2000i2p371-398.html.

29 Omer Tamuz and Ran Tessler. Majority dynamics and the retention of information. In *Working paper*, 2013.

30 Y. P. Tian, X. J. Sun, and O. Tian. Detection performance of the majority dominance rule in m-ary relay trees with node and link failures. *IEEE Transactions on Signal Processing*, 66(6):1469–1482, March 2018. doi:10.1109/TSP.2017.2788418.

31 Zhenliang Zhang, Edwin KP Chong, Ali Pezeshki, William Moran, and Stephen D Howard. Learning in hierarchical social networks. *IEEE Journal of Selected Topics in Signal Processing*, 7(2):305–317, 2013.