Consensus on the Initial Global Majority by Local Majority Polling for a Class of Sparse Graphs

Mohammed Amin Abdullah
Imperial College London
m.abdullah@imperial.ac.uk

Moez Draief
Imperial College London
m.draief@imperial.ac.uk

May 2, 2014

Abstract

We study the local majority protocol on simple graphs of a given degree sequence, for a certain class of degree sequences. We show that for almost all such graphs, subject to a sufficiently large bias, within time $A \log_d \log_d n$ the local majority protocol achieves consensus on the initial global majority with probability $1 - n^{-\Omega((\log n)^\varepsilon)}$, where $\varepsilon > 0$ is a constant. $A$ is bounded by a universal constant and $d$ is a parameter of the graph; the smallest integer which is the degree of $\Theta(n)$ vertices in the graph. We further show that under the assumption that a vertex $v$ does not change its colour if it and all of its neighbours are the same colour, any local protocol $\mathcal{P}$ takes time at least $(1 - o(1)) \log_d \log_d n$, with probability $1 - e^{-\Omega((n^{1-o(1)})}$ on such graphs. We further show that for almost all $d$-regular simple graphs with $d$ constant, we can get a stronger probability to convergence of initial majority at the expense of time. Specifically, with probability $1 - O(c^{-n})$, the local majority protocol achieves consensus on the initial majority by time $O(\log n)$. Finally, we show how the technique for the above sparse graphs can be applied in a straightforward manner to get bounds for the Erdős–Rényi random graphs in the connected regime.

1 Introduction

We are interested in protocols $\mathcal{P}$ on a graph $G$ that can bring about a consensus of opinions, which we will speak of in terms of two colours, red and blue. At each time step, a vertex is precisely one of the colours. Consensus then, is for all the vertices to be the same colour. Additionally, we would like the final colour to be the one that was the initial majority. We make no assumptions about the properties of the colours/opinions except that vertices can distinguish between them. We impose the restriction that at each time step, a vertex $v$ can only directly exchange information with its neighbours in $G$.

We study a particular example of such a protocol, the local majority protocol in synchronous discrete time with the following initial setting: Let $\alpha \in (0, \frac{1}{2})$ be a constant. For the graph $G$ in question, initially at time $t = 0$, each vertex in the vertex set $V(G)$ is red with probability $\alpha$ independently of other vertices. Hence this represents a setting where opinions are uniformly distributed across the vertices of the graph. It can easily be shown that the probability that red is not the minority is at most $c^{-n}$ for some constant $c > 1$. Furthermore, by the law of large numbers,
as \( n = |V(G)| \to \infty \), the proportion of vertices starting in red will be close to \( \alpha \). At each time step, each vertex picks a subset of its neighbours and assumes the majority colour of this subset. We show that when \( G \) belongs to a certain class of sparse graphs - those of a prescribed degree sequence - this protocol reaches consensus on the initial majority with probability \( 1 - n^{-\Omega((\log n)^{\epsilon})} \) where \( \epsilon > 0 \) is a constant. Furthermore, we show that this convergence happens within time \( t = A \log_d \log_d n \) where \( A = A(\epsilon, d) \) is upper bounded by a universal constant. Here \( d \) is a property of the graph which we call the effective minimum degree. It is the smallest integer \( d \) such that \( \Theta(n) \) of the vertices have degree \( d \). It signifies that - if there are not too many of them - one can have smaller degree nodes in the graph without them being of relevance to the process. Additionally we show a convergence time lower bound of \( (1 - o(1)) \log_d \log_d n \) on this class of graphs. The lower bound applies to any protocol where a vertex can only communicate with its neighbours and additionally, does not change in the next step if its colour is the same as all those around it. It is worth noting that in these classes of graphs, which we elaborate upon below, \( d \) can go to infinity with \( n \).

We give further results for regular degree special cases of the above, and study Erdős–Renyi random graphs in the connected regime.

Consensus can be studied both with a prescriptive and a descriptive view. In the former, consensus is a fundamental co-ordination mechanism for distributed systems. In the latter, it can be seen as a natural process occurring, for example in social networks where it may represent the spread of influence. In this case, a local majority protocol is particularly pertinent. The voter model (see, e.g., \[2\]) is one of the simplest and most widely studied consensus algorithms. In the discrete time setting, at each time step \( t \), each vertex chooses a single neighbour uniformly at random (uar) and assumes its opinion. The number of different opinions in the system is clearly non-increasing, and consensus is reached almost surely in finite, non-bipartite, connected graphs. Using an elegant martingale argument, \[8\] determined the probability of consensus to a particular colour. Applied to our context, the probability of being absorbed in colour say, red, is \( \sum_{v \in R_0} \frac{d(v)}{2m} \) where \( R_0 \) is the set of vertices initially red, \( m = |E(G)| \) is the number of edges, and \( d(v) \) is the degree of vertex \( v \) (their result is more general than this, allowing for weights on edges). Thus, on regular graphs, for example, if the initial proportion of reds is a constant \( \alpha \), the probability of a red consensus is \( \alpha \). This probability is increased on non-regular graphs if the minority is “privileged” by sitting on high degree vertices (as in say, for example, the small proportion of high degree vertices in a graph with power-law distribution). This motivates an alternative where the majority is certain, or highly likely, to win.

In addition to probability of a particular colour/opinion dominating, one is also interested in how long it takes to reach consensus. In the voter model, there is a duality between the voting process and multiple randoms walk on the graph. The time it takes for a single opinion to remain is the same as the time it takes for \( n \) independent random walks - one starting at each vertex - to coalesce into a single walk, where two or more random walks coalesce if they are on the same vertex at the same time. Thus, consensus time can be determined by studying this multiple walk process. In fact, the vertex at which the final coalescence occurs is also the vertex whose initial opinion is the one that becomes the consensus opinion, so this probabilities of converging to some opinion can be approached through this duality. See, e.g., \[2\].

The analyses of protocols dealing with majority consensus have not been readily amenable to the established techniques for the voter model, namely, martingales and coalescing random walks.
Martingales have proved elusive and the random walks dual does not readily transfer, nor is there an obvious way of altering the walks appropriately. Thus, ad-hoc techniques and approaches have been developed. In the next section, we review and compare work closely related to our own, and elaborate upon the significance of it.

2 Related Work and Our Contribution

We first review some closely related work before discussing our contribution. In [3] a protocol for general graphs is given that is shown to reach majority consensus almost surely. They demonstrate the convergence time is finite, but don’t give a bound. Subsequently, [6] gave results on expected time to convergence and recently, [13] gave a general bound of $O(n^4 \log n)$.

In [12] an asynchronous continuous time polling protocol is analysed on the complete graph where vertices contact neighbours to query their colours. A vertex can be in one of three states, two of which are opinions and the third is an “uncertain” state ‘e’. Suppose $u$ has an opinion and contacts $v$. If $v$ is the same opinion as $u$, or is in state $e$, then $u$ does not change its opinion, but if $v$ has a different opinion, $u$ goes into state $e$. If $u$ is in state $e$, it just assumes the state of $v$ (i.e., remains in $e$ if $v$ has state $e$, or takes $v$’s opinion). They show convergence to consensus in $O(\log n)$ time with error probability (probability of converging to the initial minority state) being exponentially small in $n$.

In [5] an asynchronous continuous time protocol on the complete graph is analysed. The protocol, parameterised on a pair of integers $(m, d)$ is a generalisation of the local majority protocol; a vertex contacts $m$ others and changes its own colour if more than $d$ of them disagree. As with [12] convergence time is $O(\log n)$ and probability of converging to the initial minority decays exponentially with $n$.

In [10] a local majority process is analysed on infinite regular trees. The initial setting is that each vertex is independently red with some probability. They study the bias threshold toward blue that creates a blue consensus. In [11] a local majority dynamic is studied. They show that on regular expanders, the correct consensus is reached and they give the bias threshold for this. In the case of expanders with large girth, they utilise [10].

2.1 Our contribution and how it compares

The works of [3], [13], [12] and [5], amongst others, analyse various distributed majority consensus protocols (those that converge on the initial majority) but with different aspects of success. Whilst the protocol in [3] is general and correct almost surely, the convergence time bound given in [13] (the best general bound currently available) is polynomial in $n$. In contrast, [12] and [5] give (different) protocols that are shown to converge in logarithmic time with exponentially small error probability, but they analyse only the complete graph. In contrast, we analyse a large class of sparse graphs, as well as Erdős-Rényi random graphs, with complete graphs as a special case of the latter. The error probability we give is not as strong ($n^{-\Omega(\log n)^2}$ compared to their exponentially small bound), but still very strong. Furthermore, the convergence time we give is much smaller, $O(\log_d \log_d n)$. We re-iterate that $d$ can go to infinity with $n$, so this bound can be much smaller than $O(\log \log n)$. 

3
We also give a lower bound on the convergence time for a more general class of protocols than local majority. The lower bound is within a constant factor of the upper bound. Additionally, we show that for almost all simple \(d\)-regular graphs with \(d\) constant, we can get a stronger error probability at the expense of time. Specifically, with probability \(1 - O(c^{-n^2})\), the vertices have converged to the initial majority by time \(O(\log n)\). Lastly, we analyse the result to Erdős–Renyi Random graphs, where the results we get show that for the particular case of the complete graph, we get a result complementary to those given in [5], with a quicker convergence traded off for a higher error probability.

Our work has some overlap with [11], but not a great deal. The overlap is mainly in part 6 of that paper, which is on \(d\)-regular \(\lambda\)-expanders. They get several results: For sufficiently large bias there is convergence to initial majority. This is not a probabilistic result; For the special case of \(d\)-regular random graphs, a bias equivalent to our \(\alpha = 1/2 - 1/\sqrt{d}\) is necessary. Our condition is more demanding, i.e., whenever their condition is satisfied, ours is too; For expanders with girth going to infinity, they use the infinite tree result from [10] to get a better bias condition. However, they state that \(d\) needs to be sufficiently large. Our bias condition can be arbitrarily small if we can make \(d\) as large as we want. This is not very relevant to our result in any case, since random regular graphs do have small cycles (the number of \(k\)-cycles has a Poisson distribution, see [9]). They do not address timings.

As far as we know, our analysis is the first to demonstrate (sub)logarithmic, distributed, consensus on the initial majority with high probability, on sparse graphs.

### 3 Outline of key ideas

We give a brief outline of the key ideas that we use. Core to our analysis is the tree subgraph. Consider a tree \(T(v)\) with root \(v\). Suppose for simplicity that all non-leaf vertices have degree \(d\) and that the tree is of depth \(h\). At \(t = 0\), each vertex is red independently with probability \(\alpha < 1/2\). Now consider a vertex \(x\) at depth \(h - 1\), meaning \(x\) is parent to leaves. At time \(t = 1\), \(x\) polls its neighbours and takes the majority colour. \(x\)’s own colour at time \(t = 0\) is irrelevant. Furthermore, if \(d\) is fairly large, \(x\) can ignore the colour of its parent since it is unlikely to matter much to the outcome. Thus, let us consider a protocol where non-leaf vertices only poll their children and take the majority colour from that. Returning to \(x\), at time \(t = 1\), it will be red with probability \(p_1 = \Pr(\text{Bin}(d, \alpha) > d/2)\), which is considerably less than \(p_0 = \alpha\). All the other vertices at depth \(h - 1\) have the same probability of being red. Furthermore, since vertices only poll their children, those at depth \(h - 1\) are independent of each other. Thus, from the point of view of vertices at depth \(h - 2\), at time \(t = 2\), they poll children, each of which is independently red with probability \(p_1\). Hence, at \(t = 2\), a depth \(h - 2\) vertex is red with probability \(p_2 = \Pr(\text{Bin}(d, p_1) > d/2)\) which is considerably less than \(p_1\). Continuing in this manner, by time \(t = h\), the root vertex will be sampling children which have a very low probability of being red such that taking the union bound over all such “tree-like” vertices \(v\), means with high probability none of them are red at time \(h\).

In our analysis, we use this tree local subgraph and analyse a modification of the majority protocol where vertices in the tree make the conservative assumption that their parents are red. This can only do worse in terms of error probability so provides a bound. Of course not all vertices in the graphs under consideration are regular and locally tree-like, and we make use of results on
structural properties of these graphs to handle those vertices close to cycles or which have small degree vertices in the locality.

4 Notation, definitions and models

For a graph \( G = (V, E) \), denote by \( V(G) = V \) and \( E(G) = E \) the vertex and edge set respectively. For a vertex \( v \in V(G) \) denote by \( N(v) \) the neighbours of \( v \) in \( G \) and let \( d(v) = d_v = |N(v)| \). In this paper the log function is to base \( e \) when no base is stated.

We will denote by \( G \in \mathcal{G}(n, p) \) the Erdős–Rényi random graph with \( n \) vertices and edge probability \( p \).

4.1 Graph model: Graphs of a given degree sequence

Let \( V = [n] \) be a vertex set and define \( \mathcal{G}_n(d) \) to be the set of connected simple graphs with degree sequence \( d = (d_1, d_2, \ldots, d_n) \), where \( d_i \) is the degree of vertex \( i \in V \). Clearly, there are restrictions on degree sequences for the model to make sense. An obvious one is that the sum of the degrees in the sequence cannot be odd. Even then, not all degree sequences are graphical and not all graphical sequences can produce simple graphs. Take for example the two vertices \( v \) and \( w \) where \( d_v = 3 \) and \( d_w = 1 \). In order to study this model, we restrict the degree sequences to those which are nice and graphs which have nice degree sequences are termed the same. The precise definition will be given in section 6.1, but basically, nice graphs are sparse, with not too many high degree vertices. They also have the property that there is a constant \( 0 < \kappa \leq 1 \) and an integer \( d \) (that may grow with \( n \)) that occurs \( \kappa n + o(n) \) times in \( d \), and any integer smaller than \( d \) occurs \( o(n) \) times. We call \( d \) the effective minimum degree. It may be assumed where it appears below that \( d \) is nice.

Our analysis also requires that graphs \( G \) taken from \( \mathcal{G}_n(d) \) have certain structural properties. The subset of graphs \( \mathcal{G}'_n(d) \) having these properties form a large proportion of \( \mathcal{G}_n(d) \), in fact, \( |\mathcal{G}'_n(d)|/|\mathcal{G}_n(d)| = 1 - n^{-\Omega(1)} \) when \( d \) is nice (\cite{ref1}). We term graphs in \( \mathcal{G}'_n(d) \) for \( d \) nice as typical. Further details will be given in 6.2.

For the special case of \( d \)-regular graphs \( G \) with \( d \geq 5 \) constant, we can give stronger bounds on the error probability at the expense of time. To do so we require that \( G \) be typical regular, which almost all connected simple \( d \)-regular graphs are. More precisely, if \( \mathcal{G}_n(d) \) is the set of all connected simple \( d \)-regular graphs on \( n \) vertices, and \( \mathcal{G}'_n(d) \subseteq \mathcal{G}_n(d) \) is the subset of typical regular graphs, then \( |\mathcal{G}'_n(d)|/|\mathcal{G}_n(d)| \to 1 \) as \( n \to \infty \), as shown in \cite{ref4}.

Although this model is typically framed as a random graph, randomness here is superfluous. We assume that the graph \( G \) that the protocol acts upon is from the typical subset \( \mathcal{G}'_n(d) \) of the set \( \mathcal{G}_n(d) \) of simple graphs with nice degree sequence \( d \). As long as \( G \) has the typical properties, the time and error probability bounds will hold. The fact that the typical subset is almost the same size as the general set is demonstrated via the configuration model. See \cite{ref1} for a detailed explanation.
4.2 Protocol

Time is indexed by the non-negative integers \( t = 0, 1, 2, \ldots \). For a given graph \( G \), let \( X^P_t(v) \) be the indicator function for vertex \( v \) being blue at time \( t \) under protocol \( P \), i.e., \( X^P_t(v) = 1 \) iff \( v \) is blue at time step \( t \) when running protocol \( P \). For a positive integer \( k \), and a vertex \( v \), define \( k(v) = \min \left\{ k, 2 \left\lfloor \frac{d(v) - 1}{2} \right\rfloor + 1 \right\} \). Thus, \( k(v) \) is the minimum of \( k \) and the largest odd integer not greater than \( d(v) \).

**Definition 1** (\( k \)-choice Majority Protocol \( MP^k \)). Assume \( k \) is odd. At time step \( t + 1 \), each vertex \( v \in V(G) \) randomly picks a \( k(v) \)-subset \( N_v(t + 1) \) uniformly from the set of \( \binom{d(v)}{k(v)} \) possible subsets. \( v \) then assumes at time \( t + 1 \) the majority colour at time \( t \) of the vertices in \( N_v(t + 1) \). More more formally,

\[
X^P_{t+1}(v) = 1 \left\{ \left( \sum_{w \in N_v(t+1)} X^P_t(w) \right) > \frac{k(v)}{2} \right\}
\]

**Definition 2.** We call a protocol \( P \) local-stable (LS) if it has the following properties: (i) It is local meaning that a vertex \( v \) can only directly exchange information with its neighbours. (ii) It is stable, meaning that if a vertex \( v \) and all its neighbours \( N(v) \) are the same colour, then under \( P \), \( v \) will not change colour in the next step.

5 Results

The proof of Theorem 2 is given in the main text. Due to space restrictions, the proofs not given in the main text are put in the appendix.

**Theorem 1.** Suppose \( G \in \mathcal{G}_n(d) \) is typical with effective minimum degree \( d \). For any local-stable protocol \( P \), the following holds: With probability \( 1 - e^{-\Omega(n^{1-o(1)})} \), at time step \( (1 - o(1)) \log_d \log_d n \), \( P \) will not have reached consensus on the initial majority.

**Theorem 2.** Suppose \( G \in \mathcal{G}_n(d) \) is typical with effective minimum degree \( d \). Let \( \nu = \left\lfloor \frac{d-1}{2} \right\rfloor \) and suppose that for some constant \( \beta < 1 \) the following condition holds:

\[
\left( 1 + \frac{1}{\sqrt{2\nu}} \right) 2^{\frac{1}{\nu-1}} 4\alpha(1 - \alpha) < \beta. \tag{1}
\]

Then for an arbitrarily small constant \( \varepsilon > 0 \), the following holds: With probability \( 1 - n^{-\Omega((\log n)^{\varepsilon/2})} \), at time step \( \frac{1+\varepsilon}{\log_d(d-1)-\log_d n} \log_d \log_d n \), \( MP^{2\nu+1} \) will have reached consensus on the initial majority.

**Remark** Consequent of Theorem 1 and Theorem 2, \( MP \) is asymptotically optimal. Note the upper bound in this interval is at most \( 4(1 + \varepsilon) \) by the assumption \( d \geq 5 \), and it is \( (1 + o(1))(1 + \varepsilon) \) if \( d \to \infty \) with \( n \), which nice degree sequences allow for. Thus, in this case, the upper bound is within factor \( 1 + \varepsilon \) of the lower bound, for arbitrarily small constant \( \varepsilon > 0 \).

**Theorem 3.** Suppose \( G \in \mathcal{G}_n(d) \) is typical regular. For some some constants \( c > 1 \) and \( \varepsilon > 0 \), with probability \( 1 - O\left(c^{-n^\varepsilon}\right) \), by time \( O(\log n) \), \( MP^d \) will have reached consensus on the initial majority.

**Theorem 4.** Suppose \( p = \frac{c \log n}{n} \) where \( c > 2 + \varepsilon \) for some constant \( \varepsilon > 0 \), and suppose \( d \geq 5 \) is an odd constant. Pick a graph \( G \in G(n, p) \) and run \( MP^d \). Let \( A = \frac{1+\varepsilon}{\log_d(d-1)-\log_d n} \) where \( \varepsilon > 0 \) is
a small constant. Subject to condition (i), with probability $1 - n^{-\Omega(1)}$, by time $A \log_d \log_d n$, $\mathcal{MP}^d$ will have reached consensus on the initial majority.

Remark Observe Theorem 4 gives a probability for a trial whereby a graph $G$ from $\mathcal{G}(n, p)$ is picked then the protocol run on it, i.e., there is randomness in the actual graph the protocol is run on. This is in contrast to Theorems 1, 2 and 3 where there was no randomness in picking the graph; it merely had to be a graph from the typical set. When $p = 1$, however, $\mathcal{G}(n, p)$ will be the complete graph with probability 1. In this case, it is interesting to compare this result with those given in [5]; the error probability in Theorem 4 is not as strong as the exponentially small one given in [5] but the convergence time is $O(\log \log n)$ compared to their $O(\log n)$.

6 Graphs of a given degree sequence

6.1 Assumptions about the degree sequence

We denote by $\mathcal{G}_n(d)$ the set of simple graphs with vertex set $V = [n]$ and degree sequence $d = (d_1, d_2, \ldots, d_n)$, where $d_i$ is the degree of vertex $i \in V$. We make the following definitions: Let $V_j = \{i \in V : d_i = j\}$ and let $n_j = |V_j|$. Let $\sum_{i=1}^n d_i = 2m$ and let $\theta = 2m/n$ be the average degree. We use the notations $d_i$ and $d(i)$ for the degree of vertex $i$.

Let $0 < \kappa \leq 1$ be constant, $0 < c < 1/8$ be constant and let $\gamma = (\sqrt{\log n/\theta})^{1/3}$. We suppose the degree sequence $d$ satisfies the following conditions:

(i) Average degree $\theta = o(\sqrt{\log n})$.

(ii) Minimum degree $\delta \geq 3$.

(iii) Let $d$ be such that $n_d = \kappa n + o(n)$. We call $d$ the effective minimum degree.

(iv) Number of little vertices $\sum_{d-1}^{d} n_j = O(n^{c(d-1)/d})$; a vertex $v$ is little if $d(v) \leq d - 1$.

(v) Maximum degree $\Delta = O(n^{c(d-1)/d})$.

(vi) Upper tail size $\sum_{j=\gamma \theta}^{\Delta} n_j = O(\Delta)$.

Any degree sequence with constant maximum degree, and for which $d = \delta$ is nice. The conditions hold in particular, for $d$-regular graphs, $d \geq 3$, $d = \delta = o(\sqrt{\log n})$, as condition (iii) holds with $\kappa = 1$. The spaces of graphs we consider are somewhat more general. The condition nice, allows for example, bi-regular graphs where half the vertices are degree $\delta = \delta = o(\sqrt{\log n})$.

Condition (ii) ensures connectivity. Conditions (i), (iv), (v) and (vi) allow the structural properties in Lemma 5 to be inferred via the configuration model, as was done in [1]. The effective minimum degree condition (iii), ensures that some entry in the degree sequence occurs order $n$ times.

Finally, as we shall see, our analysis of the consensus protocol requires

(vii) $d \geq 5$. 7
Definition 3. A nice degree sequence \( \mathbf{d} \) is one that satisfies conditions (i)-(vii) above, and we apply the same adjective to any graph \( G \in \mathcal{G}_n(\mathbf{d}) \) with a nice \( \mathbf{d} \).

6.2 Structural Properties of \( G \)

Let \( C \) be a large constant, and let
\[
\omega = C \log \log n. \tag{2}
\]
A cycle or path is small, if it has at most \( 2\omega + 1 \) vertices, otherwise it is large. Let
\[
\ell = B \log^2 n \tag{3}
\]
for some large constant \( B \). A vertex \( v \) is light if it has degree at most \( \ell \), otherwise it is heavy. A cycle or path is light if all vertices are light.

Lemma 5 and 6 are from [1].

Lemma 5 ([1]). Let \( \mathbf{d} \) be a nice degree sequence and let \( G \) be chosen uar from \( \mathcal{G}(\mathbf{d}) \). With probability \( 1 - O(n^{-1/4}) \),

(a) No vertex disjoint pair of small light cycles are joined by a small light path.
(b) No light vertex is in two small light cycles.
(c) No small cycle contains a heavy vertex or little vertex, or is connected to a heavy or little vertex by a small path.
(d) No pair of little or heavy vertices is connected by a small path.

For a vertex \( v \), let \( G[v, s] \) be the subgraph induced by the set of vertices within a distance \( s \) of \( v \). A vertex \( v \) is \( d \)-tree-like to depth \( h \) if \( G[v, h] \) is a \( d \)-regular tree, (i.e. all vertices on levels 0, 1, ..., \( h-1 \) have degree \( d \)). We choose the following value for \( h \), which depends on \( \theta \).
\[
h = \frac{1}{\log d} \log \left( \frac{\log n}{(\log \log n) \log \theta} \right) \tag{4}
\]
A vertex \( v \) is \( d \)-compliant, if \( G[v, \omega] \) is a tree, and all vertices of \( G[v, \omega] \) have degree at least \( d \). A vertex \( v \) is \( d \)-tree-regular, if it is \( d \)-tree-like to depth \( h \), \( d \)-compliant to depth \( \omega \) and all vertices of \( G[v, \omega] \) are light. For such a vertex \( v \), the first \( h \) levels of the BFS tree, really are a \( d \)-regular tree, and the remaining \( \omega - h \) levels can be pruned to a \( d \)-regular tree.

Lemma 6 ([1]). Let \( \mathbf{d} \) be a nice degree sequence and let \( G \) be chosen uar from \( \mathcal{G}(\mathbf{d}) \). There exists \( \epsilon > 0 \) constant such that with probability \( 1 - O(n^{-\epsilon}) \),

(e) there are \( n^{1-O(1/\log \log n)} \) \( d \)-tree-regular vertices.

Definition 4. A typical graph \( G \) is one that is nice and also satisfies conditions (a)-(e) of Lemmas 5 and 6.

Definition 5. Let \( L_1 = \epsilon_1 \log d n \), where \( \epsilon_1 > 0 \) is a sufficiently small constant. A typical regular graph \( G \) of degree \( d \) for some constant \( d \) is one that is typical and and additionally has the following property:
(f) No pair of cycles $C_1, C_2$ with $|C_1|, |C_1| \leq 100L_1$ are within distance $100L_1$ of each other.

From [4] we have the following,

**Lemma 7** ([4]). Let $G$ be chosen uniform from $G_n(d)$, the set of all simple connected $d$-regular graphs. With probability tending to 1 as $n \to \infty$, no pair of cycles $C_1, C_2$ with $|C_1|, |C_1| \leq 100L_1$ are within distance $100L_1$ of each other.

Thus, we see that a fraction $1 - n^{-\Omega(1)}$ of graphs $G \in G_n(d)$ are typical and a fraction $1 - o(1)$ of $G \in G_n(d)$, the set of $n$-vertex simple connected $d$-regular graphs, are typical regular.

### 6.3 Proof of Theorem [2]

For any $v \in V(G)$, there is some integer $s \geq 0$ such that $T := G[v, s]$ is a tree rooted at $v$. We define the modified majority protocol ($\text{MMP}$) with respect to a vertex $v \in V(G)$. The initial setting is as before; each vertex is independently red with probability $\alpha < 1/2$. Recall $X_t^{\text{MMP}}(x) = 1$ if, under $\text{MMP}$, $x$ is blue at time $t$ and 0 if it is red. Let $X_t^{\text{MMP}(v,s)}(x)$ be the same for $\text{MMP}(v,s)$.

**Definition 6** ($k$-choice Modified Majority Protocol $\text{MMP}^k(v,s)$). Assume $k$ is odd and let $T := G[v,s]$. At time step $t+1$, each vertex $x \in V(G)$ randomly picks a $k(x)$-subset $N_x(t+1)$ uniformly from the set of $\binom{k(x)}{k^2}$ possible subsets. If $x \notin T$ then $x$ becomes at time $t+1$ the majority colour at time of the vertices in $N_x(t+1)$. More more formally,

$$X_{t+1}^{\text{MMP}^k(v,s)}(x) = 1\left\{ \sum_{y \in N_x(t+1)} X_t^{\text{MMP}^k(y)} > k(x)/2 \right\}$$

If $x \in T$ then denote by $\text{Par}(x)$ the parent of $x$ in $T$. At time $t+1$, $x$ becomes the majority colour at time $t$ of the vertices in $N_x(t+1)$, with the added assumption that $\text{Par}(x)$ was red at time $t$. More more formally,

$$X_{t+1}^{\text{MMP}^k(v,s)}(x) = 1\left\{ \sum_{y \in N_x(t+1) \setminus \{\text{Par}(x)\}} X_t^{\text{MMP}^k(y)} > k(x)/2 \right\}$$

Thus, $\text{MMP}$ is the same as $\text{MP}$ except that vertices in $T$ effectively make the conservative assumption that if a parent is picked, it is red. This will help in getting an upper bound on the probability of red.

In the next lemma, we compare $\text{MMP}^k$ and $\text{MP}^k$. In order to do so, we make use of the fact that the randomness of the system is not affected by the actions of the protocols. To reiterate, given a graph $G$, there are two sources of randomness. There is the initial random assignment of colours, and there is the sequence of choices of neighbours $(N_v(t))_{t=1}^\infty$ each vertex $v$ makes. Thus, the Cartesian product $\Omega$ of the possible initial colourings with each of the infinite sequences of neighbour choices each vertex creates a probability space, where an element $\sigma \in \Omega$ is a particular initial colouring of the vertices and, for each vertex $v$, a particular infinite sequence of neighbour choices made by $v$. The next lemma compares $\text{MMP}^k$ and $\text{MP}^k$ under the same $\sigma \in \Omega$.

**Lemma 8.** Suppose $T := G[v,s]$ is a tree. Fix a $\sigma \in \Omega$ and consider $\text{MP}^k$ and $\text{MMP}^k(v,s)$ under this $\sigma$. For all $t \geq 0$, we have $X_t^{\text{MMP}^k(v,s)}(x) \leq X_t^{\text{MP}^k}(x)$ for every vertex $x$. 

9
Proof. By induction on $t$. See appendix. \hfill \square

**Corollary 9.** Suppose $G[v, s]$ is a tree. $\Pr(X^\text{MMP}_k(v) = 0) \leq \Pr(X^\text{MMP}_k(v, s)(v) = 0)$.

Let
\[
\omega' = \log_d \log_d n.
\]

**Lemma 10.** Let $\nu = \lfloor \frac{d-1}{2} \rfloor$ and $k = 2\nu + 1$. Let $A = \frac{1 + \varepsilon}{\log \omega'}$ where $\varepsilon > 0$ is a small constant. Suppose a vertex $v$ is such that $G[v, A\omega']$ is a tree and each non-leaf vertex in this tree has degree at least $d$. Subject to condition (I), we have $\Pr(X^\text{MMP}_{A\omega'}(v) = 0) = n^{-\Omega((\log n)^{\varepsilon/2})}$.

Proof. $k$ is the largest odd integer not greater than $d$. We bound $\Pr(X^\text{MMP}_{A\omega'}(v, s)(v) = 0)$ and use Corollary 9. For convenience we forgo notation that is obvious.

$\Pr(X^\text{MMP}_0(x) = 0) = \alpha$, and in particular, this holds for any vertex $x$ such that $d(v, x) = A\omega'$. Now for $x$ with $d(v, x) = A\omega' - 1$,

\[
p_1 = \Pr(X^\text{MMP}_1(x) = 0) = \Pr \left( \sum_{y \in N_v(1) \setminus \{\text{Par}(x)\}} X^\text{MMP}_0(y) \leq \nu \right) \leq \sum_{i=\nu}^{2\nu} \binom{2\nu}{i} \alpha^i (1 - \alpha)^{2\nu - i}. \tag{5}
\]

Since $\alpha < \frac{1}{2}$, $\sum_{i=\nu}^{2\nu} \binom{2\nu}{i} \alpha^i (1 - \alpha)^{2\nu - i} \leq \alpha^\nu (1 - \alpha)^\nu \sum_{i=\nu}^{2\nu} \binom{2\nu}{i} = \alpha^\nu (1 - \alpha)^\nu \left( \frac{1}{2} \right)^{2\nu} + \frac{1}{2} \binom{2\nu}{\nu}$) and using the inequality $\binom{2\nu}{\nu} \leq \frac{2^{2\nu}}{\sqrt{2\nu}}$, we have $p_1 \leq \frac{1}{2} (1 + \frac{1}{\sqrt{2\nu}}) (4\alpha(1 - \alpha))^\nu$.

Assume for $t < A\omega'$ and all $x$ such that $d(v, x) = A\omega' - t$,

\[
p_t \leq \frac{1}{4} \left[ \left( 1 + \frac{1}{\sqrt{2\nu}} \right) 2 \right] \sum_{i=0}^{\nu t} (4\alpha(1 - \alpha))^\nu .
\]

Then for $t + 1$ and all $x$ such that $d(v, x) = A\omega' - t - 1$,

\[
p_{t+1} := \Pr(X^\text{MMP}_{t+1}(x) = 0) = \sum_{i=\nu}^{2\nu} \binom{2\nu}{i} p_i^t (1 - p_i)^{2\nu - i}
\]
\[
\leq \frac{1}{2} \left( 1 + \frac{1}{\sqrt{2\nu}} \right) (4p_t(1 - p_t))^\nu
\]
\[
\leq \frac{1}{2} \left( 1 + \frac{1}{\sqrt{2\nu}} \right) (4p_t)^\nu
\]
\[
\leq \frac{1}{2} \left( 1 + \frac{1}{\sqrt{2\nu}} \right) \left( \frac{1 + \frac{1}{\sqrt{2\nu}}}{2} \right) \sum_{i=0}^{\nu t} (4\alpha(1 - \alpha))^{\nu t + 1}.
\]
Hence, for any \( t \leq A\omega' \),
\[
p_t \leq \frac{1}{4}\left(\left(\frac{1}{1 + \frac{1}{\sqrt{2\epsilon}}}\right)^2\right)^{\frac{1}{\nu-1}} 4\alpha(1-\alpha)^{\nu^t}.
\]
In particular, when \( t = A\omega' \), \( \nu^t = \nu^{A\omega'} = (\log_d n)^{A\log_d \nu} \), and by condition (11),
\[
\Pr(X_{A\omega'}^{\text{MAMP}}(v) = 0) \leq \frac{1}{4\nu^{(\log_d n)^{A\log_d \nu}}} = \frac{1}{4n} - \frac{1}{\nu^{(\log_d n)^{A\log_d \nu-1}}}
\]
Now \( \log_d \nu = \log_d (\lceil \frac{d-1}{2}\rceil) \geq \log_d (d - 2) - \log_d 2 > 0.25 \) since we assume \( d \geq 5 \). Hence, if \( A = \frac{1+\epsilon}{\log_d \nu} \) for some arbitrarily small constant \( \epsilon > 0 \), the positive part of the exponent in (6) is
\[
\log_d \left(\frac{1}{\beta}\right)(\log_d n)^\epsilon = \log \left(\frac{1}{\beta}\right) \frac{(\log n)^\epsilon}{(\log d)^{1+\epsilon}} = \Omega \left(\frac{(\log n)^\epsilon}{(\log \log n)^{1+\epsilon}}\right) = \Omega((\log n)^{\epsilon/2})
\]
where the second equality holds because \( d = o(\sqrt{\log n}) \) by the degree sequence assumptions. Thus, (6) is \( n^{\Omega((\log n)^{\epsilon/2})} \) and applying Corollary 9 completes the proof.

The above lemma deals with vertices \( v \) for which \( G[v, A\omega'] \) is a tree and each non-leaf vertex in this tree has degree at least \( d \). We are left to deal with vertices \( v \) for which \( G[v, A\omega'] \) contains a cycle or a non-leaf vertex with degree less than \( d \).

**Lemma 11.** Suppose \( G \) is typical with effective minimum degree \( d \). Subject to condition (11), the conclusion of Lemma 10 holds for \( G[v, A\omega'] \) when \( G[v, A\omega'] \) contains a cycle or a non-leaf vertex with degree less than \( d \).

**Proof.** In Lemma 5 (c) says all vertices within distance \( 2\omega + 1 \) of a small cycle \( C_1 \) are light, so any other small cycle \( C_2 \) within \( 2\omega + 1 \) either connects to \( C_1 \) via a small light path or \( C_1 \) and \( C_2 \) intersect. The former case is precluded by (a) and the latter by (b) of the same lemma. Therefore, no pair of small cycles is within distance \( 2\omega + 1 \) of each other. Hence, if for some \( v \), \( G[v, A\omega'] \) is not a tree, then for some (unique) cycle \( C \), \( v \) is either on \( C \) or there is a unique small path from \( v \) to \( C \). Consider the latter case. Suppose \( x \in N(v) \) is on the small path. We may assume \( x \) always to be red and since \( v \) has degree at least \( d \) (by Lemma 5 (c)), the bound in (6) holds.

Now suppose \( v \) is on the cycle \( C \). Suppose \( \{x, y\} \) are the neighbours of \( v \) on \( C \). Each vertex \( u \in N(v) \setminus \{x, y\} \) has degree at least \( d \) (by Lemma 5 (c)) and can assume \( v \) is always red to get the bound in (6). Thus by time \( A\omega' \) all of \( N(v) \setminus \{x, y\} \) will be blue and in the next step they will out-vote \( \{x, y\} \) if \( d - 2 > 2 \).

We now deal with little vertices. By Lemma 5 (d) there can be at most one vertex \( u \in G[v, A\omega'] \) such that \( d(u) < d \). If \( u \equiv v \) then by the above argument all of \( N(v) \) will be blue by time \( A\omega' \) and so \( v \) will be in the next step. Otherwise, there is a unique path from \( v \) to \( u \) which can be cut off, and by the above argument the bound holds.

**Proof of Theorem 2.** Using Lemmas 10 and 11 apply a union bound to all \( n \) vertices in \( G \).
7 Appendix

Proof of Theorem 7. Suppose \( v \) is \( d \)-tree-regular. The probability all vertices in \( G[v,h] \) are initially red is \( \alpha^K \) where \( K = 1 + \frac{d}{d-2}((d-1)^h - 1) \). Since \( G \) is typical, there exist \( n^{1-O(1/\log \log n)} \) \( d \)-tree-regular vertices (Definition 4) so there will be \( \Omega(n^{1-O(1/\log \log n)}/\ell^{2\omega}) \) non-intersecting \( d \)-regular trees of depth \( h \). Therefore, the probability that at least one of these is initially all red is
at least

\[
\begin{align*}
1 - (1 - \alpha^k) &\Omega((n^{1-O(1/\log \log n)})/\ell^{2\omega}) \\
&\geq 1 - \exp\left\{-\Omega\left(\frac{\alpha^k n^{1-O(1/\log \log n)}}{\ell^{2\omega}}\right)\right\}
\end{align*}
\]

where \( c = \frac{1}{\alpha} > 2 \). The logarithm of the bracketed expression in (7) is

\[
\left(1 - O\left(\frac{1}{\log \log n}\right)\right) \log n \geq 1 - \exp\left\{-\Omega\left((n^{1-O(1/\log \log n)})\right)\right\}
\]

Thus (7) is

\[
1 - e^{-\Omega(n^{1-o(1)})}
\]

By the locality and stability conditions, \( G[v, h] \) being initially all red means it requires at least \( h \) time steps until \( v \) can become blue.

**Proof of Lemma 8** Below we will forgo ‘\( k \)’ and ‘(v, s)’ in the superscripts. \( N_x(t) \) the set of neighbours chosen by vertex \( x \) at time \( t \) under \( \sigma \).

We argue by induction on \( t \). Clearly \( X_0^{MM^P}(x) = X_0^{MP}(x) \) for every \( x \). Suppose \( X_t^{MM^P}(x) \leq X_t^{MP}(x) \) for every \( x \). If \( x \notin \mathcal{T} \) then

\[
X_{t+1}^{MM^P}(x) = 1\left\{\sum_{y \in N_x(t+1)} X_t^{MM^P}(y) > \frac{k(x)}{4}\right\} \leq 1\left\{\sum_{y \in N_x(t+1)} X_t^{MP}(y) > \frac{k(x)}{4}\right\} = X_{t+1}^{MP}(x)
\]

If \( x \in \mathcal{T} \) then

\[
\sum_{y \in N_x(t+1) \setminus \{\text{Par}(x)\}} X_t^{MM^P}(y) \leq \sum_{y \in N_x(t+1) \setminus \{\text{Par}(x)\}} X_t^{MP}(y) \leq \sum_{y \in N_x(t+1)} X_t^{MP}(y),
\]

so

\[
X_{t+1}^{MM^P}(x) = 1\left\{\sum_{y \in N_x(t+1) \setminus \{\text{Par}(x)\}} X_t^{MM^P}(y) > \frac{k(x)}{4}\right\} \leq 1\left\{\sum_{y \in N_x(t+1)} X_t^{MP}(y) > \frac{k(x)}{4}\right\} = X_{t+1}^{MP}(x)
\]

**Proof of Theorem 3** This follows by the same reasoning as the proof of Theorem 2 except that now we use trees of depth \( L_1 \), giving a time \( L_1 \) and error bound (6) of

\[
\Pr(X_{L_1}^{MM^P}(v) = 0) \leq \frac{1}{4^L_1} = O(c^{-n^\varepsilon})
\]

for some constant \( c > 1 \) and constant \( \varepsilon > 0 \).
Erdős-Rényi Random Graphs

We study $G_n(p)$ for $p = \frac{c \log n}{n}$, where $c > 2 + \epsilon$ and $\epsilon > 0$ is a constant. We shall take $d$ to be an odd constant, though it should not be too difficult to extend the results for values which go to infinity with $n$.

**Proof of Theorem 4.** We use the following Chernoff bound from \[7\] eq (1.7): For $\epsilon > 0$

\[
\Pr(X < (1 - \epsilon)E[X]) \leq \exp \left( -\frac{\epsilon^2}{2}E[X] \right)
\]

Thus, if $X$ is the degree of a particular vertex $v$, $X \sim \text{Bin}(n-1, p)$. Let $\epsilon$ be a constant such that $1 > \epsilon > \frac{\sqrt{2}}{c}$. Then $\Pr(X < (1 - \epsilon)c\log n) \leq \exp \left( -\frac{\epsilon^2}{2}c\log n \right) = n^{-(1+\Omega(1))}$ and taking the union bound over all vertices means that with probability $1 - n^{-\Omega(1)}$ each vertex has degree at least $c^2\log n$ for some constant $c^2 > 0$.

For a vertex $v$, denote the set of $d$ vertices chosen by $v$ at time $t$ by $N_v(t)$. Consider a vertex $v$. We build a (multi)graph $T = T(v)$, which will be a tree for most vertices, with the following algorithm. As before, let $\omega' = \log_d \log_d n$. The data structure ‘map’ (also known as an associative array, or dictionary) associates a vertex with a value. $E(T)$ is the edge multiset of $T$. The vertex set $V(T)$ is not a multiset.

**Algorithm 1: $T$-BUILD**

1. $T \leftarrow \{v\}$
2. $\text{map}(v) \leftarrow 0$
3. for $i \leftarrow 0$ to $A\omega' - 1$ do
4.     foreach $x \in V(T)$ do
5.         if $\text{map}(x) = i$ then
6.             $V(T) \leftarrow V(T) \cup N_x(A\omega' - i)$
7.             foreach $y \in N_x(A\omega' - i)$ do
8.                 $\text{map}(y) \leftarrow i + 1$
9.                 $E(T) \leftarrow E(T) \cup (x, y)$

We show there is at most one cycle in $T$. Before two cycles have formed, no vertex has exposed more than $2d + 3$ edges. Given that $d$ is a constant and each vertex has degree $\Omega(\log n)$, this exposure is negligible. Since $|V(T)| \leq d^{A\omega'+1}$, the probability, at any given step before two cycles are formed, of connecting to $T$ is $O \left( d^{A\omega'+1}/n \right)$. Therefore, the probability that at any point in its construction $T$ is picked twice is at most

\[
O \left( \frac{d^{2(A\omega'+1)}}{n^2} \right) \times \left( \frac{d^{A\omega'+1}}{2} \right) = O \left( \frac{d^{4(A\omega'+1)}}{n^2} \right) = O \left( \frac{(\log_d n)^4}{n^2} \right)
\]

where we have used the assumption that $d$ is constant in the last equality. Taking the union bound over all $n$ originating vertices, we see that with probability $1 - n^{-\Omega(1)}$, $T(v)$ has at most one cycle.
for every $v$ in $G$.

We can now apply the same reasoning as in the proof of Theorem 2. If for a $v \mathcal{T}(v)$ has a cycle $\mathcal{C}$, then if $v$ is not on $\mathcal{C}$ we can cut off the (unique) branch containing $\mathcal{C}$ and the bounds in (6) holds, as per the proof of Lemma 11. If $v$ is on $\mathcal{C}$, then since $d - 2 > 2$, the two neighbours of $v$ on $\mathcal{C}$ are out-voted by those not on $\mathcal{C}$, again, as per the proof of Lemma 11.

\qed