Local majority dynamics on preferential attachment graphs

Mohammed Amin Abdullah[1]⋆, Michel Bode[2], and Nikolaos Fountoulakis[3]⋆

[1] Mathematical and Algorithmic Sciences Lab, Huawei Technologies Ltd,
`mohammed.abdullah@huawei.com`
[2] School of Mathematics, University of Birmingham,
`michel.bode@gmx.de`
[3] School of Mathematics, University of Birmingham,
`n.fountoulakis@bham.ac.uk`

Abstract. Suppose in a graph G vertices can be either red or blue. Let k be odd. At each time step, each vertex v in G polls k random neighbours and takes the majority colour. If it doesn't have k neighbours, it simply polls all of them, or all less one if the degree of v is even. We study this protocol on the preferential attachment model of Albert and Barabási [3], which gives rise to a degree distribution that has roughly power-law $P(x) \sim \frac{1}{x^3}$, as well as generalisations which give exponents larger than 3. The setting is as follows: Initially each vertex of G is red independently with probability $\alpha < \frac{1}{2}$, and is otherwise blue. We show that if α is sufficiently biased away from $\frac{1}{2}$, then with high probability, consensus is reached on the initial global majority within $O(\log_d \log_d t)$ steps. Here t is the number of vertices and $d \geq 5$ is the minimum of k and m (or $m-1$ if m is even), m being the number of edges each new vertex adds in the preferential attachment generative process. Additionally, our analysis reduces the required bias of α for graphs of a given degree sequence studied in [1] (which includes, e.g., random regular graphs).

Keywords: Local majority dynamics, preferential attachment, power-law graphs, voting, consensus

1 Introduction

Let $G = (V, E)$ be a graph where each vertex maintains an opinion, which we speak of in terms of two colours - red and blue. We make no assumptions about the properties of the colours/opinions except that vertices can distinguish between them. We are interested in distributed protocols on G that can bring about consensus to a single opinion.

One of the simplest and most widely studied distributed consensus algorithms is the *voter model* (see, e.g., [4, ch. 14]). In the discrete time setting, at each time

---

step τ, 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, [13] determined the probability of consensus to a particular colour. In our context this would be the sum of the degrees of vertices which started with that colour, as a fraction of the sum of degrees over all vertices. Thus, on regular graphs, for example, if the initial proportion of reds is a constant α, the probability of a red consensus is α. 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.

The *local majority* protocol in a synchronous discrete time setting does the following: At each time step, each vertex v polls all its neighbours and assumes the majority colour in the next time step. This can be motivated by both a prescriptive and a descriptive view. In the former, as a consensus protocol, it can be seen as a distributed co-ordination mechanism for networked 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.

Let k be odd. Suppose at time step $\tau = 0$ each vertex of a graph $G = (V, E)$ is either red or blue. In this paper we study the following generalisation of the local majority protocol (also in a synchronous, discrete time setting):

Definition 1 (k-choice Local Majority Protocol $\mathcal{MP}^k$). *For each vertex $v \in V$, for each time step $\tau = 1, 2, \ldots$ do the following: choose a set of k neighbours of v uniformly at random. The colour of v at time step τ is the majority colour of this set at time step $\tau - 1$. If v does not have k neighbours, then choose a random set of largest possible odd cardinality.*

Clearly, we can retrieve the local majority protocol by setting k to be the maximum degree, for example.

In addition to which colour dominates, 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 random walks on the graph. The time it takes for a single opinion to emerge 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. However, the analyses of local-majority-type protocols 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 duality does not readily transfer, nor is there an obvious way of altering the walks appropriately. Thus, ad-hoc techniques and approaches have been developed.

We say a sequence of events $(\mathcal{E}_t)_t$ occurs *with high probability* (**whp**) if $\mathbf{Pr}(\mathcal{E}_t) \to 1$ as $t \to \infty$. In this paper, the underlying parameter t which goes to infinity will be the number of vertices in the sequence of graphs $\mathrm{PA}_t(m, \delta)$ we consider.

The main result in this paper will be to show that when each vertex of a preferential attachment graph $\mathrm{PA}_t(m, \delta)$ (introduced in the next section) is red independently with probability $\alpha < 1/2$, where α is sufficiently biased away from $1/2$, then the system will converge to the majority colour with high probability, and we give an upper bound for the number of steps this takes.

2   Preferential attachment graphs

The preferential attachment models have their origins in the work of Yule [17], where a growing model is proposed in the context of the evolution of species. A similar model was proposed by Simon [16] in the statistics of language. The principle of these models was used by Albert and Barabási [3] to describe a random graph model where vertices arrive one by one and each of them throws a number of half-edges to the existing graph. Each half-edge is connected to a vertex with probability that is proportional to the degree of the latter. This model was defined rigorously by Bollobás, Riordan, Spencer and Tusnády [6] (see also [5]). We will describe the most general form of the model which is essentially due to Dorogovtsev et al. [11] and Drinea et al. [12]. Our description and notation below follows that from the book of van der Hofstad [14].

The random graph $\mathrm{PA}_t(m, \delta) = (V, E)$ where $V = [t]$ is parameterised by two constants: $m \in \mathbb{N}$, and $\delta \in \mathbb{R}$, $\delta > -m$. It gives rise to a random graph sequence (i.e., a sequence in which each member is a random graph), denoted by $(\mathrm{PA}_t(m, \delta))_{t=1}^{\infty}$. The tth term of the sequence, $\mathrm{PA}_t(m, \delta)$ is a graph with t vertices and mt edges. Further, $\mathrm{PA}_t(m, \delta)$ is a subgraph of $\mathrm{PA}_{t+1}(m, \delta)$. We define $\mathrm{PA}_t(1, \delta)$ first, then use it to define the general model $\mathrm{PA}_t(m, \delta)$ (the Barabási-Albert model corresponds to the case $\delta = 0$).

The random graph $\mathrm{PA}_1(1, \delta)$ consists of a single vertex with one self-loop. We denote the vertices of $\mathrm{PA}_t(1, \delta)$ by $\{v_1^{(1)}, v_2^{(1)}, \ldots, v_t^{(1)}\}$. We denote the degree of vertex $v_i^{(1)}$ in $\mathrm{PA}_t(1, \delta)$ by $D_i(t)$. Then, conditionally on $\mathrm{PA}_t(1, \delta)$, the growth rule to obtain $\mathrm{PA}_{t+1}(1, \delta)$ is as follows: We add a single vertex $v_{t+1}^{(1)}$ having a single edge. The other end of the edge connects to $v_{t+1}^{(1)}$ itself with probability $\frac{1+\delta}{t(2+\delta)+(1+\delta)}$, and connects to a vertex $v_i^{(1)} \in \mathrm{PA}_t(1, \delta)$ with probability $\frac{D_i(t)+\delta}{t(2+\delta)+(1+\delta)}$ – we write $v_{t+1}^{(1)} \to v_i^{(1)}$. For any $t \in \mathbb{N}$, let $[t] = \{1, \ldots, t\}$. Thus,

$$\mathbf{Pr}\left(v_{t+1}^{(1)} \to v_i^{(1)} \mid \mathrm{PA}_t(1, \delta)\right) = \begin{cases} \frac{1+\delta}{t(2+\delta)+(1+\delta)} & \text{for } i = t+1, \\ \frac{D_i(t)+\delta}{t(2+\delta)+(1+\delta)} & \text{for } i \in [t] \end{cases}$$

The model $PA_t(m, \delta)$, $m > 1$, with vertices $\{1, \ldots, t\}$ is derived from $PA_{mt}(1, \delta/m)$ with vertices $\{v_1^{(1)}, v_2^{(1)}, \ldots, v_{mt}^{(1)}\}$ as follows: For each $i = 1, 2, \ldots, t$, we contract the vertices $\{v_{(i-1)+1}^{(1)}, v_{(i-1)+2}^{(1)}, \ldots, v_{(i-1)+t}^{(1)}\}$ into one super-vertex, and identify this super-vertex as i in $PA_t(m, \delta)$. When a contraction takes place, all loops and multiple edges are retained. Edges shared between a set of contracted vertices become loops in the contracted super-vertex. Thus, $PA_t(m, \delta)$ is a graph on $[t]$.

The above process gives a graph whose degree distribution follows a power law with exponent $3 + \delta/m$. This was suggested by the analyses in [11] and [12]. It was proved rigorously for integral δ by Buckley and Osthus [7]. For a full proof for real δ see [14]. In particular, when $-m < \delta < 0$, the exponent is between 2 and 3. Experimental evidence has shown that this is the case for several networks that emerge in applications (cf. [3]). Furthermore, when $m \geq 2$, then $PA_t(m, \delta)$ is **whp** connected, but when $m = 1$ this is not the case, giving rise to a logarithmic number of components (see [14]).

3   Results and related work

Our main result is the following:

Theorem 1. *Let $k \geq 5$ be odd and let $d = m \wedge k$ if m is odd and $d = (m - 1) \wedge k$ if m is even. Let α^* be the smallest positive solution for x in the equation $\mathbf{Pr}\left(Bin(d - 1, x) \geq \frac{d-1}{2}\right) = x$. If $\delta \geq 0$ and each vertex in $PA_t(m, \delta)$ is red independently with probability $\alpha < \alpha^*$, then for any constant $\varepsilon > 0$, **whp** under $\mathcal{MP}^k$ every vertex in $PA_t(m, \delta)$ is blue at all time steps $\tau \geq \frac{1+\varepsilon}{\log_d\left(\frac{d-1}{2}\right)} \log_d \log_d t$.*

Note that $\delta = 0$ gives the model proposed in the seminal work of Albert and Barabási [3], giving power law exponent 3, and that $\delta > 0$ gives exponents larger than 3. We refer the reader to [14] for further details.

Note, for $d = 5, 7, 9, 11$ we have $\alpha^* = 0.232, 0.347, 0.396, 0.421$ to 3 significant figures (s.f.), respectively. Of course, $\alpha^* \to \frac{1}{2}$ as $d \to \infty$.

The most closely related work is [1]. Here, the same protocol was studied on random graphs of a given degree sequence (which includes random regular graphs) and Erdős–Rényi random graphs slightly above the connectivity threshold. Similar results similar to Theorem 1 were obtained, and in the case of the former model, an almost matching lower bound was shown. It should be noted that the thresholds for α obtained in this work apply equally to the models in [1], and improve the thresholds for α. To contrast, in that paper, the thresholds for $d = 5, 7, 9, 11$ were $0.092, 0.182, 0.234, 0.268$ to 3 s.f., respectively.

In [15], (full) local majority dynamics on d-regular λ-expanders on n vertices is studied. In our notation, they show that when $\alpha \leq 1/2 - \frac{2\lambda n}{d}$, there is convergence

to the initial majority, so long as $\frac{\lambda}{d} \leq \frac{3}{16}$. Since $\lambda \geq (1 - o(1))\sqrt{d}$ for a d-regular graph, this condition implies $d \geq 29$. In contrast, our results apply for $d \geq 5$.

In [10] a variant of local majority is studied where a vertex contacts m others and if d of them have the same colour, the vertex subsequently assumes this colour. They demonstrate convergence time of $O(\log n)$ and error probability – the probability of converging on the initial minority – decaying exponentially with n. However, the analysis is done only for the complete graph; our analysis of sparse graphs is a crucial difference, because the techniques employed for complete graphs do not carry through to sparse graphs, nor are they easily adapted. The error probability we give is not as strong but still strong, nevertheless. Furthermore, the convergence time we give is much smaller.

In [8] the authors study the following protocol on random regular graphs and regular expanders: Each vertex picks two neighbours at random, and takes the majority of these with itself. They show convergence to the initial majority in $O(\log n)$ steps with high probability, subject to sufficiently large initial bias and high enough vertex degree. However, in their setting, the placement of colours can be made adversarially.

In summary, our contribution is demonstrating convergence and time of convergence to initial majority for a generalisation of local majority dynamics for preferential attachment graphs with power-law exponent 3 and above. As far as we know this is the only such result for power-law graphs (by preferential attachment or otherwise). Furthermore, we have improved the bias thesholds for graphs of a given degree sequence studied by the first author in [1], which, to the best of our knowledge, were already the best or only known results for small degree graphs in this class (which includes, e.g., random regular graphs).

4   Structural results

Throughout this paper we let $\gamma = \gamma(m, \delta) = \frac{1}{2 + \delta/m}$. Observe the condition $\delta > -m$ (which must be imposed), implies $0 < \gamma < 1$.

Furthermore, for two non-negative functions $f(t), g(t)$ on $\mathbb{N}$ we write $f(t) \lesssim g(t)$ to denote that $f(t) = O(g(t))$. The underlying asymptotic variable will always be t, the number of vertices in $\mathrm{PA}_t(m, \delta)$.

Let A be a large constant and let

$$\omega = A \log \log t.$$

Let

$$\kappa = (\log t)^{7\omega}$$

and define as the *inner core* the vertices $[\kappa]$, and refer to them as *heavy* vertices. We also refer to vertices outside the inner core as *light* vertices.

Let

$$\kappa_o = (\log t)^{999\omega}$$

and define as the *outer core* the vertices $[\kappa_o]$.

Call a path *short* if it has length at most ω. Call a cycle *short* if it has at most $2\omega + 1$ vertices. Here "cycle" includes a pair of vertices connected by parallel edges and a vertex with a self-loop.

Below, we repeatedly apply the following, which is proved in [2] (and for the case $k = 1$ was given in [14]):

Proposition 1. *Suppose $i_1, j_1, i_2, j_2, \ldots, i_k, j_k$ are vertices in $PA_t(m, \delta)$ where $i_s < j_s$ for $s = 1, 2, \ldots, k$. Then*

$$\mathbf{Pr}(j_1 \to i_2 \cap j_2 \to i_2, \ldots, j_k \to i_k) \leq M^k \frac{1}{i_1^\gamma j_1^{1-\gamma}} \frac{1}{i_2^\gamma j_2^{1-\gamma}} \cdots \frac{1}{i_k^\gamma j_k^{1-\gamma}}$$

where $M = M(m, \delta)$ is a constant that depends only on m and δ.

Below, we also use the fact that $\frac{1}{i^\gamma j^{1-\gamma}} \leq \frac{1}{(ij)^{\frac{1}{2}}}$ when $\delta \geq 0$ and $i \leq j$. A similar counting approach was used in [9].

For a vertex v define $\mathcal{B}(v, r)$ to be the r ball of v in $PA_t(m, \delta)$, the subgraph within distance r.

Lemma 1. *With high probability, every vertex $v > \kappa$ has the following property: $\mathcal{B}(v, \omega)$ contains at most one cycle consisting entirely of light vertices.*

Proof. Define a *cycle-path-cycle* (CPC) structure as a pair of cycles connected by a path. We consider CPC structures where the cycles and paths are short, that is, cycles have sizes $1 \leq r, s \leq 2\omega + 1$, and the path has length $0 \leq \ell \leq \omega$. Note $r = 1$ denotes a self-loop and $r = 2$ denotes a pair of parallel edges between two vertices.

We denote by $a_1, \ldots, a_r$ and $b_1, \ldots, b_s$ the vertices of the cycles, and $c_0, \ldots, c_\ell$ the vertices of the path. Without loss of generality, we may assume $a_1 = c_0$ and $b_1 = c_\ell$. Thus, the structure has $r + s + \ell$ edges and $r + s + \ell - 1$ vertices.

Applying Proposition 1, the expected number of such structures lying entirely in $[t] \setminus [\kappa]$ is bounded by

$$\sum_{r=1}^{2\omega+1} \sum_{\ell=0}^{\omega-1} \sum_{s=1}^{2\omega+1} \sum_{\kappa < a_1, \ldots, a_r} \sum_{\kappa < b_1, \ldots, b_s} \sum_{\kappa < c_1, \ldots, c_{\ell-1}} \frac{M^{r+s+\ell}}{(a_1 b_1)^{3/2}} \prod_{i=2}^{r} \frac{1}{a_i} \prod_{j=2}^{s} \frac{1}{b_j} \prod_{k=1}^{\ell-1} \frac{1}{c_k}$$

$$\lesssim \left(\int_\kappa^t x^{-3/2} \, \mathrm{d}x \right)^2 \sum_{r=1}^{2\omega+1} \sum_{\ell=0}^{\omega-1} \sum_{s=1}^{2\omega+1} (M \log t)^{r+s+\ell}$$

$$\lesssim \frac{(\log t)^{6\omega}}{\kappa} = o(1).$$

The rest of the proof is of a similar nature and is continued in the appendix.   □

Lemma 2. *With high probability, the following hold for all $v > \kappa_0$:*

(i) *v has at most 2 edges on short paths into $[\kappa]$.*

(ii) *If v is on a short light cycle, then v has no edge that is on a short (light) path into $[\kappa]$ but that is not part of the cycle.*

(iii) *If v is connected to a short light cycle C by a short light path P, then v has at most one edge e such that e is on a short path into $[\kappa]$ but $e \notin P$.*

Proof. **(i)** Suppose v has three edges e_1, e_2, e_3 (possibly parallel) on short paths to $[\kappa]$ to vertices $i_1, i_2, i_3 \in [\kappa]$ (not necessarily distinct). Then there is a minimal structure S which contains $v, e_1, e_2, e_3, i_1, i_2, i_3$, and a short path from v to $[\kappa]$ via each edge e_1, e_2, e_3. Since S is minimal, there are $0 \leq r \leq 3(\omega - 1)$ light vertices $a_1, \dots, a_r$ in S which form the short paths from v to $[\kappa]$ via e_1, e_2, e_3. Also, since S is minimal, it contains at most 3ω edges. To consider two extremes, for example, S might be three non-intersecting paths, or a single path with e_1, e_2, e_3 being parallel and all other vertices connected connecting by non-parallel edges.

Observe that each vertex a_i has at least two edges, meaning in the application of Proposition 1 it incurs a fraction $\frac{1}{a_i}$ or less. Applying Proposition 1, the expected number of structures S is asymptotically bounded from above by

$$\kappa^3 M^{3\omega} \sum_{r=0}^{3\omega} \sum_{\kappa < a_1, \dots, a_r} \frac{1}{v^{3/2}} \prod_{i=1}^{r} \frac{1}{a_i} \lesssim \frac{3\omega \kappa^3 (M \log t)^{3\omega}}{v^{3/2}}.$$

Hence, taken over all $v > \kappa_o$, this is $O\left(\frac{3\omega \kappa^3 (M \log t)^{3\omega}}{\kappa_o^{1/2}}\right) = o(1)$.

The rest of the proof is of a similar nature and is continued in the appendix.   □

We define the *truncated r-ball around v*, denoted by $\widetilde{\mathcal{B}}(v, r)$, as follows:

1. Delete from $\mathcal{B}(v, r)$ all edges incident to vertices in $[\kappa]$, denote by $\mathcal{B}_{-}(v, r)$ the resulting graph.

2. Let $\mathcal{C}_v(\mathcal{B}_{-}(v, r))$ be the connected component in $\mathcal{B}_{-}(v, r)$ that contains v. Add to $\mathcal{C}_v(\mathcal{B}_{-}(v, r))$ all edges (u, v) deleted in the previous step such that $u \in [\kappa]$ and $v \in \mathcal{C}_v(\mathcal{B}_{-}(v, r))$. The resulting graph is $\widetilde{\mathcal{B}}(v, r)$.

The following is a corollary of the above.

Corollary 1. *With high probability, for every vertex $v > \kappa_o$, $\widetilde{\mathcal{B}}(v, \omega)$ belongs to one of the following categories:*

(i) *$\widetilde{\mathcal{B}}(v, \omega)$ is a tree and all vertices are light.*

(ii) *$\widetilde{\mathcal{B}}(v, \omega)$ has no cycles and one or two heavy vertices.*

(iii) *In $\widetilde{\mathcal{B}}(v, \omega)$, v is part of a short cycle of light vertices, and any heavy vertex in $\widetilde{\mathcal{B}}(v, \omega)$ only connects to v via edges that are part of that cycle.*

(iv) *In $\widetilde{\mathcal{B}}(v, \omega)$, there is a short cycle of light vertices which v is not part of, which connects to v through a short path P, and there is at most one edge e on a path from from v to a heavy vertex in $\widetilde{\mathcal{B}}(v, \omega)$ such that e is not part of P.*

Degree of outer-core vertices For $i \in [t]$ consider the vertex i and the core $[i]$. Immediately after the vertex i is added, the graph under construction at that point, $\text{PA}_i(m, \delta)$, has total degree $2mi$, and $D_i(i)$ is a random variable taking integral value between m and $2m$. We may ask, given $D_i(i) = a$, what is the probability that $D_i(t) = a + d$? The question can be framed as one about a Polya urn process in which the urn initially contains a red balls and $2mi - a$ black balls, and the selection process has weighting functions $W_R(k) = k + \delta$ and $W_B(k) = k - (i-1)\delta$ for red and black balls respectively (see, e.g., [14]).

Notation: $S_i(t)$ is the sum of degrees of vertices in $[i]$ in $\text{PA}_t(m, \delta)$. The following was shown in [2]:

Lemma 3. *There is a constant $C(m, \delta)$ that depends only on m and δ, such that for $1 \le d \le n \le m(t - i)$,*

$$\mathbf{Pr}\left(D_i(t) = a + d \mid S_i(t) - 2mi = n, D_i(i) = a\right) \le C(m, \delta) \frac{1}{d} \left(\frac{Id}{I + n - d}\right)^{a+\delta} e^{-\frac{dI}{I+n}}$$

and

$$\mathbf{Pr}(X_R(n, a) = 0) \le \left(\frac{I}{I + n}\right)^{a+\delta},$$

where $I = I(i, m, \delta) = i(2m + \delta) - 1$.

Furthermore, the following was also given in [2]:

Lemma 4. *Suppose $\delta \ge 0$ and for a vertex $i \in [t]$, $i = i(t) \to \infty$. There exists a constant $K_0 > 0$ that depends only on m and δ, such that the following holds for any constant $K > K_0$ and h which is smaller than a constant that depends only on m, δ,*

$$\mathbf{Pr}\left(S_i(t) < \frac{1}{K}\mathbf{E}[S_i(t)]\right) \le e^{-hi}.$$

We use these lemmas to prove the following:

Lemma 5. *With high probability, for every $i \in [\kappa_o]$, $D_i(t) \ge \left(\frac{t}{\kappa_o}\right)^\gamma \frac{1}{\kappa_o^2}$.*

Proof. Letting $h = \frac{\log \kappa_o}{\kappa_o}$ in Lemma 4, we see that for some constant K, $S_{\kappa_o}(t) \geq Kt^\gamma \kappa_o^{1-\gamma}$. Let $z = z(t) \to \infty$ as $t \to \infty$ to be determined later. Letting $n = Kt^\gamma \kappa_o^{1-\gamma} - 2m\kappa_o$ and applying 3,

$$\mathbf{Pr}\left(D_{\kappa_o}(t) \leq \frac{n}{\kappa_o z} \mid S_{\kappa_o}(t) - 2m\kappa_o \geq n, D_{\kappa_o}(\kappa_o) = a\right)$$

$$\leq \sum_{d=0}^{n/(\kappa_o z)} C(m, \delta) \frac{1}{d}\left(\frac{Id}{I+n-d}\right)^{a+\delta} e^{-\frac{dI}{I+n}}$$

$$\lesssim \left(\frac{I}{I+n}\right)^{a+\delta} + \sum_{d=1}^{n/(\kappa_o z)} \left(\frac{I}{I+n-d}\right)^{a+\delta} d^{a+\delta-1} e^{-\frac{dI}{I+n}}$$

$$\leq \left(\frac{I}{I+n}\right)^{a+\delta} + \frac{I^{a+\delta}}{(I+n-n/(\kappa_o z))^{a+\delta}} \sum_{d=0}^{n/(\kappa_o z)} d^{a+\delta-1}.$$

Since $\kappa_o \to \infty$ and $z \to \infty$ as $t \to \infty$, we have $n/(\kappa_o z) = o(n)$, so $\frac{1}{(I+n-n/(\kappa_o z))^{a+\delta}} \lesssim \frac{1}{(I+n)^{a+\delta}}$.

Furthermore,

$$\sum_{d=0}^{n/(\kappa_o z)} d^{a+\delta-1} \lesssim \int_0^{n/(\kappa_o z)} x^{a+\delta-1}\,\mathrm{d}x \leq \frac{1}{a+\delta}\left(\frac{n}{\kappa_o z}\right)^{a+\delta}.$$

Hence,

$$\mathbf{Pr}\left(D_{\kappa_o}(t) \leq \frac{n}{\kappa_o z} \mid S_{\kappa_o}(t) - 2m\kappa_o \geq n, D_{\kappa_o}(\kappa_o) = a\right)$$

$$\lesssim \left(\frac{I}{I+n}\right)^{a+\delta} + \left(\frac{I}{I+n}\right)^{a+\delta}\left(\frac{n}{\kappa_o z}\right)^{a+\delta}$$

$$\lesssim \left(\frac{I}{I+n}\right)^{a+\delta} + \frac{1}{z^{a+\delta}}.$$

Now, we choose $z(t) = \kappa_o^2$. Then,

$$\left(\frac{I}{I+n}\right)^{a+\delta} = \left(\frac{\kappa_o(2m+\delta-1)}{\kappa_o(2m+\delta-1) + Kt^\gamma \kappa_o^{1-\gamma} - 2m\kappa_o}\right)^{a+\delta} \lesssim \left(\frac{\kappa_o}{t}\right)^{\gamma(a+\delta)} = o\left(\frac{1}{z^{a+\delta}}\right)$$

since $a \geq m \geq 5$ and $\delta \geq 0$.

Thus,

$$\mathbf{Pr}\left(D_{\kappa_o}(t) \leq \frac{n}{\kappa_o z} \mid S_{\kappa_o}(t) - 2m\kappa_o \geq n, D_{\kappa_o}(\kappa_o) = a\right) \lesssim \frac{1}{\kappa_o^{2(m+\delta)}}.$$

A simple coupling argument shows that $D_{\kappa_o}(t)$ is stochastically dominated by $D_i(t)$ for any $i \in [\kappa_o]$. Therefore, taking the union bound over $[\kappa_o]$ we get $\frac{\kappa_o}{\kappa_o^{2(m+\delta)}} = o(1)$ since $a \geq m \geq 5$ and $\delta \geq 0$.   □

5 Convergence of the majority dynamics

In this section we show that the system converges to the initial majority opinion and bound the time it takes. Informally, Lemma 6 shows convergence for a tree when the bias away from $1/2$ is large enough, Lemma 7 demonstrates for vertices outside the outer core, it only takes a constant number of steps for the probability of being red to get below the bias threshold that Lemma 6 requires. It also uses the fact that vertices in this range are almost tree-like. The conclusion is that there is a certain contiguous set of steps when all the vertices outside the outer core are blue. Finally, Lemma 8 shows that when this happens, the vertices in the outer core are all blue. Since there is a time step in which all vertices are blue, the graph remains blue thereafter.

For real p and integer $n > 3$ define

$$f(n,p) = \left[\left(1 + \frac{1}{\sqrt{n-1}}\right)2\right]^{\frac{2}{n-3}} 4p(1-p). \tag{1}$$

The following lemma was essentially first proved by the first author in [1]. Due to space restrictions, we give an informal overview here and leave the full proof in the appendix.

Lemma 6. *Let $T_u(h, d^+)$ be a depth-h tree rooted at u where all non-leaf vertices have degree at least $d^+ \geq 5$ odd. Let $p \in (0, \frac{1}{2})$, $k \geq 5$ and $d = k \wedge d^+$. Suppose at time $\tau = 0$ each vertex of $T_u(h, d^+)$ is assigned red with probability p. Under $\mathcal{MP}^k$ the probability that the root u is red at time step h is at most $\frac{1}{4}\left(f(d,p)\right)^{\left(\frac{d-1}{2}\right)^h}$.*

Proof (Overview). Suppose instead of $\mathcal{MP}^k$ we had a modified version $\mathcal{MMP}^k$ on the tree in which each vertex other than the root u assumes its parent is red. Under the same sequence of random choices of which neighbours to poll, $\mathcal{MMP}^k$ can only make it more likely that u ends up being red at time step $\tau = h$. It also has the advantage of breaking dependencies between vertices at the same depth in the tree. Denoting $p_\tau(v)$ the probability of vertex v being red at time step τ, we show that under $\mathcal{MMP}^k$, we get $p = p_0(v_h) > p_1(v_{h-1}) > \ldots > p_{h-1}(v_1) > p_h(v_0)$ where v_i is a child of v_{i-1} in the tree and $v_0 = u$. In fact, the sequence of probabilities decays very rapidly, and we find that $p_h(v_0) < \frac{1}{4}\left(f(d,p)\right)^{\left(\frac{d-1}{2}\right)^h}$. $\qquad\square$

Lemma 7. *Let $k \geq 5$ be odd and let $d = m \wedge k$ if m is odd and $d = (m-1) \wedge k$ if m is even. Let ε be any positive constant, let $\tau_* = B \log_d \log_d t$ where $B = B(d, \varepsilon) = \frac{1+\varepsilon}{\log_d\left(\frac{d-1}{2}\right)}$ and let α^* be the smallest positive solution for x in the equation $\mathbf{Pr}\left(Bin(d-1, x) \geq \frac{d-1}{2}\right) = x$. If each vertex in $PA_t(m, \delta)$ is red independently with probability $\alpha < \alpha^*$, then **whp** under $\mathcal{MP}^k$ every vertex $v \in [t] \setminus [k_o]$ is blue at time steps $\tau = \tau_* + 1, \tau_* + 2$.*

Proof. Let integer $n \geq 2$, $f(x) = \mathbf{Pr}\left(\text{Bin}(2n, x) \geq n\right)$ and $g(x) = f(x) - x$. Observe $g(0) = 0$ and $g(1/2) > 0$. Furthermore, $g'(x) = \binom{2n}{n} n x^{n-1}(1-x)^n - 1$, whence $g'(0) = -1$. Therefore $g(x)$ has a root x^* in $(0, 1/2)$. Now $g''(x) = \binom{2n}{n} n x^{n-2}(1-x)^{n-1}\left[(n-1) - x(2n-1)\right]$ which is strictly positive on $0 < x < \frac{1}{2} - \frac{1}{2(2n-1)}$ and non-positive on $\frac{1}{2} - \frac{1}{2(2n-1)} \leq x < 1$. We can therefore deduce that x^* is the unique root of $g(x)$ in $(0, 1/2)$, and that for the interval $[c_1, c_2]$ where $0 < c_1 < c_2 < x^*$, g attains a maximum at c_1 or c_2. Hence, for a given $x \in [c_1, c_2]$, we have $0 < f(x) < x$ and $x - f(x) = -g(x) > -\max\{g(c_1), g(c_2)\} > 0$. Therefore, we need only a constant number of iterations of f until $f(f(\ldots f(x))\ldots) < c_1$. When $2n = d - 1$, we write $\alpha^* = x^*$.

Now consider a rooted tree of depth h where non-leave vertices have $2n = d - 1$ children, and suppose that each vertex is coloured red independently with probability $\alpha < \alpha^*$ at time $\tau = 0$. By the same argument as in the proof of Lemma 6, at time $\tau = 1$ the depth $h - 1$ vertices are red independently with probability $f(\alpha) < \alpha - c_2$. Continuing in this way, the probability the root is red is at most c_1 if $h > c_3$ where c_3 is a large enough finite constant.

Let $\tau' = \tau_* - c_3$ and suppose that $\mathcal{B}(v, \omega)$ is a tree. Since $\omega = A \log \log t$ with A arbitrarily large, then we may assume $\omega = a \log_d \log_d t$ where a is a constant such that $\omega \geq \tau_* + 3$. This means $\mathcal{B}(w, \tau_*)$ is a tree if w is a neighbour of v or v itself. By the above, we may therefore assume that at time $t = c_3$, the depth $\tau_* - c_3 = \tau'$ vertices are red independently with probability at most c_1.

Then by Lemma 6 the probability v is red at time step τ_* is at most $\frac{1}{4} f(d, c_1)^{\left(\frac{d-1}{2}\right)^{\tau'}}$. For large enough t, $\tau' \geq \frac{1+\varepsilon/2}{\log_d\left(\frac{d-1}{2}\right)} \log_d \log_d t$, therefore

$$\left(\frac{d-1}{2}\right)^{\tau'} \geq \left(\frac{d-1}{2}\right)^{\frac{1+\varepsilon/2}{\log_d\left(\frac{d-1}{2}\right)} \log_d \log_d t} = d^{(1+\varepsilon/2)\log_d \log_d t} = (\log_d t)^{1+\varepsilon/2}.$$

Thus,

$$f(d, c_1)^{\left(\frac{d-1}{2}\right)^{\tau'}} \leq d^{-\log_d\left(\frac{1}{f(d,c_1)}\right)(\log_d t)^{1+\varepsilon/2}} = t^{-\log_d\left(\frac{1}{f(d,c_1)}\right)(\log_d t)^{\varepsilon/2}}.$$

If $f(d, c_1) < \beta < 1$ where β is a constant then the above is at most $t^{-(\log_d t)^{\varepsilon/4}}$ when t is large enough. By the same logic, and since each of the children of v are also trees out to distance τ_*, the the same probability bound applies to them. Thus, taking the union bound, we see that all vertices v such that $\mathcal{B}(v, \omega)$ is a tree are blue at times $\tau_*, \tau_* + 1, \tau_* + 2$.

We extend the above to other vertices outside $[\kappa_o]$. From Corollary 1, we see that v always has at most two "bad" edges that it can assume are always red. Since $m \geq 5$, this leaves $m - 2 \geq 3$ "good" edges which, if they are blue, will outvote the bad edges, regardless of what their actual colours are. Thus, suppose $e_1 = (v, w_1), \ldots, e_{m-2} = (v, w_{m-2})$ are good edges. As per the proof of Lemma 6, the random variables $Y_\tau(w_i)$ for $i \in \{1, \ldots, m-2\}$ depend only on vertices in

the subtree of $\widetilde{\mathcal{B}}(v,\omega)$ for which w_i is a root. This is a depth–$(\omega-1)$ tree where each vertex not a leaf nor root has at least $m-1$ children. Since we may assume that $\omega \geq \tau_* + 2$, it follows by the above that **whp**, all such w_i are blue at time steps $\tau_*, \tau_* + 1, \tau_* + 2$. This forces v to be blue at time steps $\tau_* + 1, \tau_* + 2, \tau_* + 3$. Thus, we have proved that **whp**, all vertices $v \in [t] \setminus [\kappa_o]$ are blue at time steps $\tau_* + 1, \tau_* + 2$. □

It remains to consider the vertices in $[\kappa_o]$:

Lemma 8. *If every vertex in $v \in [t] \setminus [\kappa_o]$ is blue at time step $\tau_* + 1$, then* **whp** *every $v \in [\kappa_o]$ is blue at time step $\tau_* + 2$.*

Proof. Consider a vertex $v \in [\kappa_o]$. We partition v's set of incident edges E_v in $\mathrm{PA}_t(m,\delta)$ into two sets $E_{v1} = \{(v,w) : w \in [\kappa_o]\}$ and $E_{v2} = E_v \setminus E_{v1}$. Clearly, $|E_{v1}| \leq m\kappa_o$, so by Lemma 5, we may assume that $|E_{v2}| \geq \left(\frac{t}{\kappa_o}\right)^{\gamma} \frac{1}{\kappa_o^2} - m\kappa_o$ for every $v \in [\kappa_0]$. Consequently, the probability that at time step $\tau_* + 1$ the majority of edges picked by v are in E_{v1} is zero if $d \geq 2|E_{v1}| + 1$ and $O\left(\mathbf{Pr}\left(\mathrm{Bin}(d, \frac{\kappa_o^4}{t}) > \frac{d}{2}\right)\right) = O\left(\kappa_o^4/t\right)$ if $d \leq 2|E_{v1}|$. Taken over all vertices in $[\kappa_o]$ this is $o(1)$. □

Corollary 2. *With high probability, $PA_t(m,\delta)$ is entirely blue at all time steps $\tau \geq \tau_* + 2$.*

6   Conclusion and open problems

We have seen that with high probability, local majority dynamics on preferential attachment graphs with power law exponent at least 3 very rapidly converges to the initial majority when the initial distribution of red vs. blue opinions is sufficiently biased away from equality. The speed of convergence is affected both by the number of neighbours polled at each step as well structural parameters of the graph, specifically, how many edges are added when a new vertex joins in the construction process of the graph.

A natural next step would be to analyse the process for $-m < \delta < 0$, which generates graphs with power-law exponents between 2 and 3. These appear to better reflect "real world" networks, but our experience suggests that structural differences make the techniques of this paper ineffective in this regime.

Another direction would be to explore how adversarial placements of opinions affects outcome, as studied in [8] for random regular graphs.

References

1. M.A. Abdullah and M. Draief, Global majority consensus by local majority polling on graphs of a given degree sequence, *Discrete Applied Mathematics* 180:1–10, 2015

2. M.A. Abdullah and N. Fountoulakis, A phase transition in the evolution of bootstrap percolation processes on preferential attachment graphs, arXiv:1404.4070
3. R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, *Reviews of Modern Physics* 74:47–97, 2002.
4. D. Aldous and J. Fill, *Reversible Markov Chains and Random Walks on Graphs*, (in preparation) `http://stat-www.berkeley.edu/pub/users/aldous/RWG/book.html`
5. B. Bollobás and O. Riordan, The diameter of a scale-free random graph, *Combinatorica* 24:5–34, 2004.
6. B. Bollobás, O. Riordan, J. Spencer and G. Tusnády, The degree sequence of a scale-free random graph process, *Random Structures and Algorithms* 18:279–290, 2001
7. P.G. Buckley and D. Osthus, Popularity based random graph models leading to a scale-free degree sequence, *Discrete Mathematics* 282:53–68, 2004
8. C. Cooper, R. Elsässer and T. Radzik, The power of two choices in distributed voting, *Proc. of The 41st International Colloquium on Automata, Languages, and Programming (ICALP)*, 2014
9. C. Cooper and A. Frieze, The cover time of the preferential attachment graph, *Journal of Combinatorial Theory Series B* 97:269–290, 2004
10. J. Cruise, A. Ganesh, Probabilistic consensus via polling and majority rules, *Proc. of Allerton Conference*, 2010
11. S.N. Dorogovtsev, J.F.F. Mendes and A.N. Samukhin, Structure of growing networks with preferential linking. *Physical Review Letters* 85: 4633–4636, 2000
12. E. Drinea, M. Enachescu and M. Mitzenmacher, Variations on random graph models for the web. Technical report TR-06-01, Harvard University, Department of Computer Science, 2001
13. Y. Hassin and D. Peleg, Distributed probabilistic polling and applications to proportionate agreement, *Information and Computation* 171:248–268, 2001
14. R. van der Hofstad, *Random Graphs and Complex Networks*, 2013 (book available at `http://www.win.tue.nl/ rhofstad/NotesRGCN.pdf`)
15. E. Mossel , J. Neeman and O. Tamuz, Majority dynamics and aggregation of nformation in social networks 2012, arXiv:1207.0893
16. H.A. Simon, On a class of skew distribution functions, *Biometrika*, 42:425–440, 1955.
17. G.U. Yule A mathematical theory of evolution, based on the conclusions of Dr. J.G. Willis F.R.S., *Phil. Trans. Roy. Soc. London, B* 213:21–87, 1925

7   Appendix

Proposition 2. *Let N be a natural number and $p \in (0, \frac{1}{2})$. Then*

$$\mathbf{Pr}\left(\mathrm{Bin}(2N, p) \geq N\right) \geq \mathbf{Pr}\left(\mathrm{Bin}(2N + 2, p) \geq N + 1\right).$$

Proof. Let X and Y be independent random variables with distributions $\mathrm{Bin}(2N, p)$ and $\mathrm{Bin}(2, p)$ respectively, and let $Z = X + Y$. Then $\mathbf{1}_{\{X \geq N\}} = \mathbf{1}_{\{Z \geq N+1\}}$ except when $X = N$ and $Y = 0$, or $X = N - 1$ and $Y = 2$. The former case occurs with probability $p_a = \binom{2N}{N}p^N(1 - p)^{N+2}$ and the latter with $p_b = \binom{2N}{N-1}p^{N+1}(1 - p)^{N+1}$. Observe $p_a \geq p_b$ if and only if $p \leq \frac{N+1}{2N+1}$, which is always the case when $p < \frac{1}{2}$.

Proof of Lemma 1 continued Now consider a pair of cycles $(a_1, \ldots, a_r)$ and $(b_1, \ldots, b_s)$ that share a single vertex $a_1 = b = 1$, and where $1 \leq r, s \leq 2\omega + 1$. Applying Proposition 1, the expected number of such structures lying entirely in $[t] \setminus [\kappa]$ is bounded by

$$\sum_{r=1}^{2\omega+1} \sum_{s=1}^{2\omega+1} \sum_{\kappa < a_1, \ldots, a_r} \sum_{\kappa < b_2, \ldots, b_s} \frac{M^{r+s}}{(a_1)^2} \prod_{i=2}^{r} \frac{1}{a_i} \prod_{j=2}^{s} \frac{1}{b_j}$$

$$\lesssim \left(\int_{\kappa}^{t} x^{-2} \, dx \right) \sum_{r=1}^{2\omega+1} \sum_{s=1}^{2\omega+1} (M \log t)^{r+s}$$

$$\lesssim \frac{(\log t)^{5\omega}}{\kappa} = o(1).$$

Finally, consider a cycle $a_1, \ldots, a_r$ and a connecting path $b_0, \ldots, b_\ell$ where $1 \leq r, \ell \leq 2\omega + 1$. Setting $b_0 = a_1$ and $b_\ell = a_x$ where a_x varies over the other $r - 1$ vertices of the cycle, the expected number of such structures lying entirely in $[t] \setminus [\kappa]$ is at most

$$\sum_{r=1}^{2\omega+1} \sum_{\ell=1}^{2\omega+1} \sum_{\kappa < a_1, \ldots, a_r} \sum_{\kappa < b_1, \ldots, b_\ell} \sum_{x=1}^{r} \frac{M^{r+\ell}}{(a_1 a_x)^{3/2}} \prod_{\substack{i=2 \\ i \neq x}}^{r} \frac{1}{a_i} \prod_{i=1}^{\ell-1} \frac{1}{b_j} \lesssim \frac{(\log t)^{5\omega}}{\kappa} = o(1).$$

$\square$

Proof of Lemma 2 continued (ii) Suppose the cycle in question (which may be a self-loop or a pair of parallel edges with another vertex) has length r and the path has length ℓ. Without loss of generality, assuming $a_1 = b_0 = v$, the expected number of such structures is bounded from above by

$$\kappa \sum_{r=1}^{2\omega+1} \sum_{\ell=1}^{\omega} \sum_{\kappa < a_1, \ldots, a_r} \sum_{\kappa < b_1, \ldots, b_{\ell-1}} \frac{M^{r+\ell}}{v^{3/2}} \prod_{i=2}^{r} \frac{1}{a_i} \prod_{j=1}^{\ell-1} \frac{1}{b_j} \lesssim \frac{\kappa (M \log t)^{4\omega}}{v^{3/2}}.$$

Summing over all $v > \kappa_0$ this is $O\left(\frac{\kappa (M \log t)^{4\omega}}{\kappa_o^{1/2}} \right) = o(1)$.

(iii) Suppose v has two edges e_1, e_2 on short paths into $[\kappa]$ and that neither is on P. By similar reasoning to part **(i)**, there is a minimal structure S with light vertices $a_1, \ldots, a_r$ where $r \leq 5\omega$, with at most 5ω edges and where each light vertex has at least edges. Applying Proposition 1, the expected number of structures S for a given v is asymptotically bounded from above by

$$\kappa^2 M^{5\omega} \sum_{r=1}^{5\omega} \sum_{\kappa < a_1, \ldots, a_r} \frac{1}{v^{3/2}} \prod_{i=1}^{r} \frac{1}{a_i} \lesssim \frac{5\omega \kappa^2 (M \log t)^{5\omega}}{v^{3/2}}.$$

Summing over all $v > \kappa_0$ this is $O\left(\frac{\kappa (M \log t)^{4\omega}}{\kappa_o^{1/2}} \right) = o(1)$. $\square$

Proof of Lemma 6

Let $N_v(\tau)$ be the (random) set of neighbours of v selected at time step τ. Let $d_v(k)$ be the minimum between k and the largest odd number not larger the degree of v. Observe $|N_v(\tau)| = d_v(k) \geq d$. For $v \neq u$, define $\text{Par}(v)$ to be the parent of v in $T_u(h, d)$.

We define the indicator random variables $X_0(v) = Y_0(v) = 1$ if and only if vertex v is coloured blue at time $\tau = 0$, and for $\tau > 0$,

$$X_\tau(v) = \mathbf{1}_{\left\{\left(\sum_{w \in N_v(\tau)} X_{\tau-1}(w)\right) > \frac{|N_v(\tau)|}{2}\right\}}$$

for all v, and

$$Y_\tau(v) = \mathbf{1}_{\left\{\left(\sum_{w \in N_v(\tau)\backslash\{\text{Par}(v)\}} Y_{\tau-1}(w)\right) > \frac{|N_v(\tau)|}{2}\right\}}$$

for all $v \neq u$. Thus, $X_\tau(v)$ represent the outcome under $\mathcal{MP}^k$ and $Y_\tau(v)$ is blue if and only if the number of blue children forms the majority.

Observe $Y_\tau(v) \leq X_\tau(v)$ for all v and $\tau \geq 0$, for suppose it is the case for $\tau - 1$, then

$$Y_\tau(v) = \mathbf{1}_{\left\{\left(\sum_{w \in N_v(\tau)\backslash\{\text{Par}(v)\}} Y_{\tau-1}(w)\right) > \frac{|N_v(\tau)|}{2}\right\}} \leq \mathbf{1}_{\left\{\left(\sum_{w \in N_v(\tau)\backslash\{\text{Par}(v)\}} X_{\tau-1}(w)\right) > \frac{|N_v(\tau)|}{2}\right\}}$$

$$\leq \mathbf{1}_{\left\{\left(\sum_{w \in N_v(\tau)} X_{\tau-1}(w)\right) > \frac{|N_v(\tau)|}{2}\right\}} = X_\tau(v).$$

Let $p_\tau(v) = \mathbf{Pr}\left(Y_\tau(v) = 0\right)$.

We will show that $p = p_0(v_h) > p_1(v_{h-1}) > \ldots > p_{h-1}(v_1) > p_h(v_0)$ where v_i is a child of v_{i-1} in the tree and $v_0 = u$. In fact, the sequence of probabilities decays doubly-exponentially, with the implication that taking a union bound over many trees will still result in a value that is $o(1)$.

Consider a vertex v at depth $h - 1$. Its children are leaves which are coloured red at time $\tau = 0$ independently with probability p. Therefore,

$$p_1(v) \leq \mathbf{Pr}\left(\text{Bin}(d_v(k) - 1, p) \geq \frac{d_v(k) - 1}{2}\right) \leq \mathbf{Pr}\left(\text{Bin}(d - 1, p) \geq \frac{d - 1}{2}\right)$$

where the second inequality follows by Proposition 2 (see appendix).

Since $p < \frac{1}{2}$,

$$p_1(v) \leq \sum_{i=\frac{d-1}{2}}^{d-1} \binom{d-1}{i} p^i (1-p)^{d-1-i} \leq p^{\frac{d-1}{2}} (1-p)^{\frac{d-1}{2}} \sum_{i=\frac{d-1}{2}}^{d-1} \binom{d-1}{i}$$

$$= p^{\frac{d-1}{2}} (1-p)^{\frac{d-1}{2}} \left(\frac{1}{2} 2^{d-1} + \frac{1}{2} \binom{d-1}{\frac{d-1}{2}}\right).$$

Using the inequality $\binom{2n}{n} \leq \frac{2^{2n}}{\sqrt{2n}}$, we have $p_1(v) \leq \frac{1}{2}(1 + \frac{1}{\sqrt{d-1}})(4p(1-p))^{\frac{d-1}{2}}$.

Let $D(x)$ denote the depth of a vertex x, i.e., its distance from u in the tree. Observe that $\{Y_1(v) : D(v) = h - 1\}$, $\{Y_2(v) : D(v) = h - 2\}, \ldots, \{Y_{h-1}(v) : D(v) = 1\}$ are sets of independent random variables.

Suppose that at $\tau < h$ the following inequality held for all v such that $D(v) = h - \tau$:

$$p_\tau(v) \le \frac{1}{4}\left[\left(1 + \frac{1}{\sqrt{d-1}}\right)2\right]^{\sum_{i=0}^{\tau-1}\left(\frac{d-1}{2}\right)^i}(4p(1-p))^{\left(\frac{d-1}{2}\right)^\tau} < \frac{1}{2},$$

and define p_τ to be the RHS of the above inequality. Then for $\tau + 1$ and all v such that $D(v) = h - t - 1$,

$$
\begin{aligned}
p_{\tau+1}(v) &\le \sum_{i=\frac{d-1}{2}}^{d-1}\binom{d-1}{i}p_\tau^i(1-p_\tau)^{d-1-i} \\
&\le \frac{1}{2}\left(1 + \frac{1}{\sqrt{d-1}}\right)(4p_\tau(1-p_\tau))^{\frac{d-1}{2}} \\
&\le \frac{1}{2}\left(1 + \frac{1}{\sqrt{d-1}}\right)(4p_\tau)^{\frac{d-1}{2}} \\
&\le \frac{1}{2}\left(1 + \frac{1}{\sqrt{d-1}}\right)\left(\left[\left(1 + \frac{1}{\sqrt{d-1}}\right)2\right]^{\sum_{i=0}^{\tau-1}\left(\frac{d-1}{2}\right)^i}(4p(1-p))^{\left(\frac{d-1}{2}\right)^\tau}\right)^{\frac{d-1}{2}} \\
&= \frac{1}{4}\left[\left(1 + \frac{1}{\sqrt{d-1}}\right)2\right]^{\sum_{i=0}^{\tau}\left(\frac{d-1}{2}\right)^i}(4p(1-p))^{\left(\frac{d-1}{2}\right)^{\tau+1}} \\
&= p_{\tau+1}.
\end{aligned}
$$

Hence, for $\tau \le h$, and all v such that $D(v) = h - \tau$, we have

$$p_\tau(v) \le \frac{1}{4}\left(\left[\left(1 + \frac{1}{\sqrt{d-1}}\right)2\right]^{\frac{2}{d-3}}4p(1-p)\right)^{\left(\frac{d-1}{2}\right)^\tau} = \frac{1}{4}(f(d,p))^{\left(\frac{d-1}{2}\right)^\tau}.$$

$\square$