

Fast Plurality Consensus in Regular Expanders*

Colin Cooper¹, Tomasz Radzik², Nicolás Rivera³, and
Takeharu Shiraga⁴

1 Department of Informatics, King's College London, UK
colin.cooper@kcl.ac.uk

2 Department of Informatics, King's College London, UK
tomasz.radzik@kcl.ac.uk

3 Department of Informatics, King's College London, UK
nicolas.rivera@kcl.ac.uk

4 Department of Information and System Engineering, Chuo University, Japan
shiraga@ise.chuo-u.ac.jp

Abstract

In a *voting process* on a graph vertices revise their opinions in a distributed way based on the opinions of nearby vertices. The voting completes when the vertices reach consensus, that is, they all have the same opinion. The classic example is synchronous pull voting where at each step, each vertex adopts the opinion of a random neighbour. This very simple process, however, can be slow and the final opinion is not necessarily the one with the initial largest support. It was shown earlier that if there are initially only two opposing opinions, then both these drawbacks can be overcome by a synchronous *two-sample* voting, in which at each step each vertex considers its own opinion and the opinions of two random neighbours.

If there are initially three or more opinions, a problem arises when there is no clear majority. One class of opinions may be largest (the plurality opinion), although its total size is less than that of two other opinions put together. We analyse the performance of the two-sample voting on d -regular graphs for this case. We show that, if the difference between the initial sizes A_1 and A_2 of the largest and second largest opinions is at least $Cn \max\{\sqrt{(\log n)/A_1}, \lambda\}$, then the largest opinion wins in $O((n \log n)/A_1)$ steps with high probability. Here C is a suitable constant and λ is the absolute second eigenvalue of transition matrix $P = \text{Adj}(G)/d$ of a simple random walk on the graph G . Our bound generalizes the results of Becchetti et al. [SPAA 2014] for the related *three-sample voting process* on complete graphs. Our bound implies that if $\lambda = o(1)$, then the two-sample voting can consistently converge to the largest opinion, even if $A_1 - A_2 = o(n)$. If λ is constant, we show that the case $A_1 - A_2 = o(n)$ can be dealt with by sampling using short random walks. Finally, we give a simple and efficient *push* voting algorithm for the case when there are a number of large opinions and any of them is acceptable as the final winning opinion.

1998 ACM Subject Classification F.2.2 Nonnumerical Algorithms and Problems

Keywords and phrases Plurality consensus; Regular expanders

Digital Object Identifier 10.4230/LIPIcs.DISC.2017.13

* Colin Cooper and Tomasz Radzik were supported in part by EPSRC grant EP/M005038/1, "Randomized algorithms for computer networks". Nicolás Rivera was supported by funding from Becas CHILE. Takeharu Shiraga was supported by JSPS KAKENHI Grant Number 15J03840. Work carried out while Takeharu Shiraga was visiting King's College London with the support of the ELC project (Grant-in-Aid for Scientific Research on Innovative Areas MEXT Japan).



1 Introduction

The problem of reaching consensus in a graph by means of local interactions is an abstraction of such behavior in human society as well as some processes in computer networks. In a *voting process* on a graph, vertices revise their opinions in a systematic and distributed way based on opinions of other vertices, typically on the opinions of a sample of their local neighbours. The aim is that eventually a single opinion will emerge, and that this opinion will reflect the relative importance of the original mix of opinions in some way. Voting processes have application in various fields of computing including consensus and leader election in large networks [9, 22], serialisation of read/write in replicated data-bases [21], and analysis of social behavior [16]. In general, a voting process should be conceptually simple, fast, fault-tolerant and straightforward to implement [22, 23].

In a synchronous voting process each vertex of a connected graph has one of several possible opinions. In each time-step, each vertex, using the same protocol, queries the opinion of one or more of its neighbours and decides whether to modify or to keep its current opinion. A simple voting protocol is ideally memoryless: in the current step, each vertex uses only its current opinion and the current opinions of the queried neighbours. When all vertices have a common (and thus final) opinion, we say a consensus has been reached. For a given voting process, the main quantities of interest are the probability that a particular opinion wins and the expected time to reach consensus.

In the classical *voter model* each vertex initially has a distinct opinion, but in general we assume that each vertex holds one of k different opinions. The simplest case, *two party voting*, is when there are initially two opinions ($k = 2$). If there are at least three opinions ($k \geq 3$), then the problem is often referred to as *plurality consensus*. We would like the dominant opinion to eventually become the final opinion of all vertices. The probability of this, however, strongly depends on the voting process.

Pull voting

The most well known voting process is synchronous pull voting. In this model, at each step each vertex changes its opinion to that of a random neighbour. We assume henceforth that the graphs which we consider are connected and non-bipartite, so that a consensus is possible. For such a graph, the probability that pull voting ends with a particular opinion taking over the whole graph is proportional to the initial degree of this opinion in the graph [22]. More precisely, if A is the set of vertices initially holding a given opinion, then

$$\Pr(A \text{ wins in the voting process}) = \sum_{v \in A} \frac{d(v)}{2m} = \frac{d(A)}{2m}, \quad (1)$$

where $d(v)$ is the degree of vertex v and m is the number of edges in the graph. Surprisingly, the probability here depends only on the voting process and the total degree $d(A)$, but does not depend on the details of the initial arrangement of opinions on the graph.

For an n -vertex graph, let $\mathbf{E}(T)$ be the expected value of the time to consensus T . Much of the early work was on analysing $\mathbf{E}(T)$ for classical pull voting in an asynchronous model in a continuous time setting. Here the vertices have independent exponentially distributed waiting times (Poisson clocks); see e.g. Cox [15] and Aldous [1]. In the synchronous model the expected time to consensus is $O(H_{\max} \log n)$, where $H_{\max} = O(n^3)$ is the maximum hitting time of any vertex by a random walk; see Aldous and Fill [2]. For regular expanders the expected consensus time is $\Theta(n)$, see [12].

Two- and three-sample voting

Since the classical pull voting tends to be slow ($\mathbf{E}(T) = \Theta(n)$ for regular expanders) and may be viewed as undemocratic (giving only weak preference for the largest initial opinion as shown in (1)), there has been considerable interest in modifying this simple voting process to avoid these two problems. Instead of taking the opinion of only one neighbour, the next simplest approach is to sample the opinions of a larger number of neighbours (say two or three), compare them in some way, and hope that the so-called ‘power of two choices’ improves the performance of voting. The consequences of this approach are as follows. Firstly, the number of neighbours queried affects the consensus time and the voting outcome. Secondly, the relative size of the opinions affects the ability of the process to ensure that the largest initial opinion wins.

In this setting we study the following protocols for two-sample and three-sample voting. In the two-sample voting model, at each step, each vertex v chooses two random neighbours with replacement, and if the selected vertices have the same opinion, then v adopts it; otherwise v keeps its current opinion. In the three-sample voting model, each vertex v chooses three random neighbours with replacement and adopts the majority opinion among them. If there is no majority, v picks the opinion of the first sampled neighbour. Other rules are equally possible here, e.g. v keeps its opinion. The rule we choose is the one used by Becchetti et al. [4].

For d -regular expanders, two-sample voting was studied in [13] for the case where there are initially two opinions ($k = 2$). Provided the initial difference between the sizes of the two opinions is sufficiently large, the initial majority wins with high probability (w.h.p.)¹ and voting is completed in $\mathcal{O}(\log n)$ steps. This is tight since the diameter of a d -regular graph is $\Omega(\log n)$ for constant d . This result is extended in [14] to non-regular expander graphs.

Two- and three-sample voting for plurality consensus

Not so much is known about improving the performance of voting by using two or more samples in the case where there are initially three or more opinions ($k \geq 3$). Generally, analysing plurality voting protocols tends to be more involved than analysing two party voting. The case when some opinion has an absolute majority can usually be reduce to two-party voting by grouping the other opinions into a single minority class. Difficulties arise when there is no clear majority, that is, when the largest opinion is smaller than two (or more) other opinions put together. We note that the well established techniques used in analysis of the classical pull voting (‘single-sample’ voting), for example the correspondence with multiple coalescing random walks [1, 12], do not have ready extensions or generalisations to multi-sample voting.

Plurality consensus using the three-sample voting protocol given above was studied by Becchetti et al. [4]. They proved that for the complete graph, if the difference between the initial sizes A_1 and A_2 of the largest and second largest opinions is at least $24n\sqrt{2(\log n)}/A_1$, then the largest opinion wins in $O((n \log n)/A_1)$ steps w.h.p. They also showed that this result is tight for some ranges of the parameters. Subsequently Becchetti et al. [3] analyse another simple plurality voting protocol, which can be viewed as a variation of two-sampling, showing that polylog convergence on complete graphs can be achieved even if $A_1 = o(n/\text{polylog}n)$, provided that $A_1/A_2 \geq \alpha > 1$ for a constant α and only $O(\text{polylog}n)$ opinions are initially ‘comparable’ in size with the largest opinion. Detailed parameterized bounds are presented

¹ “With high probability” (w.h.p.) means in this paper probability at least $1 - n^{-\alpha}$, for a constant $\alpha > 0$.

in [3], but they do not improve the worst-case general bound given in [4]. Two- and three-sample voting is memoryless, so requires only k states ($\log k$ bits) per vertex to store the current opinion held by this vertex. The protocol in [3] requires $k + 1$ states, as it allows each vertex to hold the “undecided” opinion.

Elsässer et al. [17] consider asynchronous two-sample voting and show that it converges on complete graphs within $O(k \log n)$ rounds, subject to suitable bounds on the number of opinions and the initial difference between the largest and the second largest opinions. Becchetti et al. [5] analyse robustness of three-sample voting and show that it converges on complete graphs in the number of rounds polynomial in k and $\log n$, even if an adversary corrupts $o(\sqrt{n})$ vertices in each round. Subsequently, Ghaffari and Lengler [19] improve the number of rounds to $O(k \log n)$, allowing at the same time for a stronger adversary.

Berenbrink et al. [6] analyse two-sample and three-sample voting for the case of a large number of initial opinions. They show that three-sample voting converges on complete graphs in $O(n^{3/4} \log n)$ rounds for any number of initial opinions $k \leq n$, but two-sample voting requires $\Omega(n/\log n)$, if initially each opinion is supported by only $O(\log n)$ vertices.

Push voting

The voting processes which we have discussed so far are examples of *pull protocols*: each vertex ‘pulls’ the information from its (selected) neighbours. In a *push protocol*, a vertex ‘pushes’ its own information onto its neighbours. While the push communication paradigm is natural and effective in rumor spreading (broadcasting) protocols, it has found so far only limited use in voting protocols. It is not clear how synchronous push voting could be defined, so push voting has been mostly confined to asynchronous processes. An example is the work of Copper et al. [11], who consider both pull and push voting in the context of asynchronous ‘discordant voting.’ Elsässer et al. [17] develop an asynchronous plurality consensus algorithm which combines two-sample voting with push-pull broadcasting. In the current paper, we propose a general simple framework for push voting.

Other related previous work

Berenbrink et al. [7] propose two synchronous pull-based plurality-consensus protocols for complete graphs. Their protocols achieve plurality consensus in $O(\log k \log \log n)$ rounds with $\log k + \Theta(\log \log k)$ bits of memory per vertex and in $O(\log n \log \log n)$ rounds with $\log k + 4$ bits of memory per vertex, provided a sufficient initial bias towards the largest opinion. Independently, Ghaffari and Parter [20] showed a plurality consensus algorithm of a similar type, which converges on complete graphs in $O(\log n \log k)$ rounds and requires $\log k + O(1)$ bits of memory per vertex.

While [3, 4, 5, 7, 17, 20] analyse plurality voting only on complete graphs, Berenbrink et al. [8] consider arbitrary connected graphs. They show two protocols, which are based on earlier work on distributed load balancing, and show a detailed analysis of their performance in various communication models. They achieve, for example, a w.h.p. $O(\log n)$ bound on the number of rounds for expanders in *diffusion model* (in each round each vertex exchanges messages with all its neighbours), but each vertex requires $\Theta((n/(A_1 - A_2))^2 \log n \log k)$ bits of memory. In contrast, two- and three-sample voting, as well as protocols of the type considered in [3], are very simple protocols requiring only $O(\log n)$ bits of memory per vertex.

2 Our contributions

Our contributions are two-fold. Firstly we extend the analysis of plurality consensus on the complete graph (using voting) to the case of regular graphs. Secondly we give a push based algorithm which reaches consensus on the complete graph in four rounds, and apply this to the problem of approximate plurality consensus.

Plurality consensus on regular graphs

The earlier analysis of plurality consensus on the complete graph can be extended to regular graphs. For regular expanders our results have the same asymptotic convergence time as given for K_n in [4]. We use the two-sample voting process of [14], but generalize the analysis from two-party voting to k -party voting.

Let G be a connected regular n -vertex graph and let λ be the second largest absolute eigenvalue of the transition matrix $P = P(G)$ of a random walk on G . Let A_1 be the set of vertices with the largest initial opinion and A_2 the set with the second largest opinion. If no confusion arises, we also let A stand for the size of set A .

► **Theorem 1.** *Let G be a regular n -vertex graph and let the initial sizes of the opinions be A_1, A_2, \dots, A_k in non-increasing order. Assume that $A_1 - A_2 \geq Cn \max\{\sqrt{(\log n)/A_1}, \lambda\}$, where λ is the absolute second eigenvalue of $P(G)$ and $C > 0$ is a suitably large constant. With probability at least $1 - 1/n$, after at most $O((n/A_1) \log(A_1/(A_1 - A_2)) + \log n)$ rounds, the two-sample voting completes and the final opinion is the largest initial opinion.*

We note the following w.h.p. property of the second eigenvalue λ for random d -regular graphs for $d = o(n^{1/2})$. For d constant it is a result of Friedman [18] that $\lambda \leq \gamma/\sqrt{d}$, where $\gamma = 2 + \epsilon$ for some small $\epsilon > 0$. For d growing with n , the following estimate of λ is given in [10]. Provided $d = o(n^{1/2})$ there exists constant $\gamma > 0$ such that w.h.p. $\lambda \leq \gamma/\sqrt{d}$. In either case the size separation condition in Theorem 1 is $A_1 - A_2 \geq C'n/\sqrt{d}$.

Theorem 1 can be applied to a number of specific scenarios. Consider, for example, the case where all k opinions are fairly evenly represented, but with one opinion slightly larger than the average n/k . More specifically, assume that $A_1 \geq (n/k)(1 + \epsilon)$, for some $0 < \epsilon \leq 1$, and that $A_2 \leq A_1/(1 + \epsilon)$. Theorem 1 implies the following corollary for this case.

► **Corollary 2.** *Let the number of opinions be $k \leq ((1/C)^2 n / \log n)^{1/3}$, $A_1 \geq (n/k)(1 + \epsilon)$, $A_2 \leq A_1/(1 + \epsilon)$, and $\lambda \leq \epsilon/(Ck)$, where $C > 0$ is the constant from Theorem 1 and $\epsilon = k^{3/2}/((1/C^2)n/\log n)^{1/2} \leq 1$. Then with probability at least $1 - 1/n$, after $O(k \log n)$ rounds the two-sample voting completes and the final opinion is the largest initial opinion.*

In Section 5 we show that the statements of Theorem 1 and Corollary 2 also hold for the three-sample voting protocol analyzed by Becchetti et. al. [4]. We note that the bound on the running time in Theorem 1 is $O(\log n)$, if A_1 is $\Omega(n/\log n)$, provided that $A_1 - A_2$ is also $\Omega(n/\log n)$ and the graph has the λ parameter appropriately small. This includes complete graphs. The bound obtained in [4] for complete graphs is $O(\log n)$ only if $A_1 = \Theta(n)$.

If the graph is not very expansive (λ is too large), we can improve the ability of two-sample voting to discriminate the plurality as follows. In ℓ -extended two-sample voting model (see [14]), each vertex makes two independent random walks of length ℓ and carries out two-sample voting using the opinions on the terminal vertices of these walks. Random walks of length ℓ replace the transition matrix P used in the proof of Theorem 1 by P^ℓ . If the graph is regular, then the only effect on the proofs is to replace all eigenvalues by their ℓ -th power. This reduces the absolute second eigenvalue from λ to λ^ℓ . The effect is to replace the condition $A_1 - A_2 \geq Cn \max\{\sqrt{(\log n)/A_1}, \lambda\}$ of Theorem 1 by the improved condition $A_1 - A_2 \geq Cn \max\{\sqrt{(\log n)/A_1}, \lambda^\ell\}$, and thus diminishing the influence of λ .

Approximate plurality consensus by push voting

In Section 6 we describe a push based algorithm (*Algorithm Pushy*) and show that it reaches consensus on K_n in four push rounds whatever the initial mix of opinions. In each round each vertex activates with some (small) probability p and sends its opinion to all its neighbours (so to all other vertices, if the graph is complete). Each vertex, after receiving the opinions of the activated vertices, tallies up different opinions and adopts the opinion with the highest count, breaking ties uniformly at random.

► **Theorem 3.** *On the complete graph, and irrespective of the initial number and distribution of opinions, after four rounds of Algorithm Pushy with the parameter $p = \Theta((\log n)/n)$, there is a consensus opinion with probability $1 - o(1)$.*

If the opinion sets A_i , $1 \leq i \leq k$, are all of the same size, then in algorithm Pushy each opinion has the same probability of becoming the consensus opinion, so in this case the algorithm gives a completely fair solution to plurality. If there are few large opinion classes of similar sizes, then while there is no guarantee that the largest opinion wins, there is only small probability that the winning opinion is outside of those large ones. As an example, suppose there are $\ell \geq 1$ classes of size at least $\Theta(n/\ell)$, where $A_1 \geq A_2 \geq \dots \geq A_\ell$ and $\ell = O(\log n / \log \log n)$, and let $\delta = \sqrt{\ell \log \log n / \log n}$. Then the probability that the winning opinion comes from opinions for which $A_j(1 + O(\delta)) \leq A_1$ is only $o(1)$. In particular, if $A_2(1 + O(\delta)) \leq A_1$, then the largest opinion A_1 wins with probability $1 - o(1)$.

We can either use Pushy on its own, or combine it with another consensus algorithm to try to solve plurality approximately but quickly in the following sense. Suppose there is a given value s such that if possible we would like to choose a consensus opinion which was initially supported by at least s vertices. According to (the proof of) Theorem 1, and assuming $s = \Omega(n^{2/3} \log n)$ and that the initial largest opinion size is at least s , then after $T = \Theta((n/s) \log n)$ rounds of two-sample voting, w.h.p. we will have discarded any opinion of initial size at most $s/(1 + \varepsilon)$. We then use Pushy to return quickly a consensus opinion chosen from the remaining opinions.

On the complete graph K_n , algorithm Pushy requires $O(\log^2 n)$ storage per vertex, and the constant number of push rounds mean that the total number of message transmissions is $O(n \log n)$, which is the same order as in $O(\log n)$ rounds of two-sample voting. In this paper we focus on proving Theorem 3, leaving for separate investigations possibilities of implementing algorithm Pushy on other interaction models.

3 Preliminary Markov chain results

In this section we establish some technical results used in our proof of Theorem 1. Consider a connected and non-bipartite graph $G = (V, E)$ with n vertices and m edges. Let P be the transition matrix of a simple random walk on G . A random walk on a connected and non-bipartite graph defines a reversible Markov chain with stationary distribution $\pi(x) = d(x)/(2m)$, where $d(x)$ denotes the degree of vertex x . The reversibility of P means that $\pi(x)P(x, y) = \pi(y)P(y, x)$, for all vertices x, y .

Let $1 = \lambda_1 > \lambda_2 \geq \dots \geq \lambda_n > -1$ be the eigenvalues of P and define $\lambda = \lambda(P)$ by $\lambda = \max\{|\lambda_2|, |\lambda_n|\}$. We also consider the matrix $P^2 = P \times P$ (standard matrix product), which is the transition matrix of the two-step random walk, is also reversible and has the same stationary distribution and eigenvectors as P . Moreover, the eigenvalues of P^2 are the squares of the eigenvalues of P . In particular, $\lambda(P^2) = (\lambda(P))^2$. Given $A, B \subseteq V$ and $x \in V$,

we define $P(x, A) = \sum_{y \in A} P(x, y)$ and the *flow function* $Q(A, B)$ from A to B as

$$Q(A, B) = \sum_{x \in A} \pi(x)P(x, B). \quad (2)$$

$Q(A, B)$ is the probability a random walk starting from the stationary distribution makes a transition from a vertex in A to a vertex in B . Due to reversibility of P , $Q(A, B) = Q(B, A)$.

We use the following standard inequalities, often referred to as the *Expander Mixing Lemma for Inhomogeneous Graphs*. Let $A, B \subseteq V$, and $A^c = V \setminus A$, then

$$|Q(A, A^c) - \pi(A)\pi(A^c)| \leq \lambda\pi(A)\pi(A^c), \quad (3)$$

$$|Q(A, B) - \pi(A)\pi(B)| \leq \lambda\sqrt{\pi(A)\pi(B)\pi(A^c)\pi(B^c)}. \quad (4)$$

We also need the following lower bounds for Q^2 , which can be proven using (4).

► **Lemma 4.** *For any $A, B \subseteq V$, we have*

$$Q(A, B)^2 \geq (\pi(A)\pi(B))^2 - 2\lambda(\pi(A)\pi(B))^{3/2}(\pi(A^c)\pi(B^c))^{1/2}. \quad (5)$$

In two-sample voting, the probability vertex x adopts the opinion B in one step is $P(x, B)^2 = (d_B(x)/d(x))^2$. Given $A, B \subseteq V$, define the quantity

$$R(A, B) = \sum_{x \in A} \pi(x)(P(x, B))^2.$$

$R(A, B)$ is the stationary measure π resulting from vertices in set A choosing opinion B in one round of two-sample voting. We do not require here A and B to be disjoint, so $A = B$ is possible. The next lemma shows that there is a connection between two-sample voting and two-step random walks. While two-sample voting does not refer to random walks in any explicit way, the transition matrix P of a random walk appears in the analysis of this voting process because of Lemma 5.

► **Lemma 5.** *For any $A \subseteq V$, we have $R(V, A) = Q_2(A, A)$, where Q_2 is the flow function for the two-step transition matrix P^2 .*

Proof. From definition of $R(V, A)$, reversibility of P and $P^2(x, y) = \sum_{z \in V} P(x, z)P(z, y)$:

$$\begin{aligned} R(V, A) &= \sum_{x \in V} \pi(x)P(x, A)^2 = \sum_{x \in V} \pi(x)P(x, A) \sum_{y \in A} P(x, y) \\ &= \sum_{y \in A} \pi(y) \sum_{x \in V} P(y, x)P(x, A) = \sum_{y \in A} \pi(y)P^2(y, A) = Q_2(A, A). \quad \blacktriangleleft \end{aligned}$$

If G is a complete graph (with a loop at each vertex), then $R(V, A) = \pi(A)^2 = (|A|/n)^2$ and $R(A, B) = \pi(A)\pi(B)^2 = |A| \cdot |B|^2/n^3$. The next two lemmas give bounds on deviations from these values in regular graphs.

► **Lemma 6.** *For $A \subseteq V$, we have*

$$|R(V, A) - \pi(A)^2| = |Q_2(A, A^c) - \pi(A)\pi(A^c)| \leq \lambda^2\pi(A)\pi(A^c). \quad (6)$$

Proof. By Lemma 5, $R(V, A) = Q_2(A, A)$. Also $Q_2(A, A) = Q_2(A, V) - Q_2(A, A^c) = \pi(A) - Q_2(A, A^c)$, so

$$R(V, A) - \pi(A)^2 = \pi(A) - Q_2(A, A^c) - \pi(A)^2 = \pi(A)\pi(A^c) - Q_2(A, A^c).$$

Taking the absolute value of both sides gives the first equality in (6). To obtain the inequality, apply (3) to P^2 , Q_2 and λ^2 as the second largest absolute eigenvalue of P^2 . ◀

► **Lemma 7.** *Let $A, B \subseteq V$, then*

$$R(A, B) \geq \frac{Q(A, B)^2}{\pi(A)} \geq \pi(A)\pi(B)^2 - 2\lambda\pi(A)^{1/2}\pi(B)^{3/2}\pi(A^c)^{1/2}\pi(B^c)^{1/2}.$$

Proof. The second inequality is from Lemma 4. From convexity of the function $z \mapsto z^2$,

$$R(A, B) = \pi(A) \sum_{x \in A} \frac{\pi(x)}{\pi(A)} (P(x, B))^2 \geq \pi(A) \left(\sum_{x \in A} \frac{\pi(x)}{\pi(A)} P(x, B) \right)^2 = \frac{1}{\pi(A)} Q(A, B)^2. \quad (7)$$

◀

Suppose the family of sets $\mathcal{C} = (A_1, \dots, A_k)$ is a partitioning of V . Define the quantity $S_{\mathcal{C}}(A) = \sum_{i=1}^k R(A, A_i)$. For a complete graph, $S_{\mathcal{C}}(V) = \sum_{i=1}^k \pi(A_i)^2$ and the following lemma bounds the deviation from this value in regular graphs.

► **Lemma 8.** *Consider a partitioning $\mathcal{C} = (A_1, \dots, A_k)$ of V . Then*

$$\left| S_{\mathcal{C}}(V) - \sum_{i=1}^k \pi(A_i)^2 \right| \leq \lambda^2 \left(1 - \sum_{i=1}^k \pi(A_i)^2 \right).$$

Proof. Using Lemma 6, we get

$$\left| S_{\mathcal{C}}(V) - \sum_{i=1}^k \pi(A_i)^2 \right| \leq \sum_{i=1}^k |R(V, A_i) - \pi(A_i)^2| \leq \lambda^2 \left(1 - \sum_{i=1}^k \pi(A_i)^2 \right). \quad \blacktriangleleft$$

► **Lemma 9.** *Let $\mathcal{C} = (A_1, \dots, A_k)$ be a partitioning of V . For any $A \subseteq V$,*

$$S_{\mathcal{C}}(A) \geq \pi(A) \sum_{i=1}^k \pi(A_i)^2 - 2\lambda\pi(A)^{1/2} \sum_{i=1}^k \pi(A_i)^{3/2}, \quad (8)$$

$$S_{\mathcal{C}}(A) \leq \pi(A) \sum_{i=1}^k \pi(A_i)^2 + 2\lambda\pi(A)^{1/2} \sum_{i=1}^k \pi(A_i)^{3/2} + \lambda^2. \quad (9)$$

Proof. Lemma 7 gives:

$$S_{\mathcal{C}}(A) = \sum_{i=1}^k R(A, A_i) \geq \pi(A) \sum_{i=1}^k \pi(A_i)^2 - 2\lambda\pi(A)^{1/2} \sum_{i=1}^k \pi(A_i)^{3/2}, \quad (10)$$

and Inequality (8) follows. To show Inequality (9), observe that $S_{\mathcal{C}}(A) + S_{\mathcal{C}}(A^c) = S_{\mathcal{C}}(V)$ and use Lemma 8 and (10) applied to A^c

$$\begin{aligned} S_{\mathcal{C}}(A) &= S_{\mathcal{C}}(V) - S_{\mathcal{C}}(A^c) \\ &\leq \pi(A) \sum_{i=1}^k \pi(A_i)^2 + 2\lambda\pi(A)^{1/2}\pi(A^c)^{1/2} \sum_{i=1}^k \pi(A_i)^{3/2} + \lambda^2 \left(1 - \sum_{i=1}^k \pi(A_i)^2 \right). \quad \blacktriangleleft \end{aligned}$$

4 Proof of Theorem 1

From now on we assume the graph is d -regular, so $\pi(x) = 1/n$, and for $A \subseteq V$, $\pi(A) = |A|/n$. Furthermore, $nR(A, B) = \sum_{x \in A} (d_B(x)/d)^2$ is the expected number of vertices in A which pick two opinions in B . When clear from the context, we use A instead of $|A|$.

Let A_j be the set of vertices with opinion j . At any step, the opinions are ordered according to their sizes: $A_1 \geq A_2 \geq \dots \geq A_k$. Thus $\mathcal{C} = (A_1, \dots, A_k)$ is a partition of V .

Let A'_j be the set of vertices with opinion j after one round. We have the following equality, where the second term in (11) is the expected weight of vertices changing their opinion to A_j and the third term is the expected weight of vertices changing their opinion from A_j (using the measure π as the weight of a set of vertices).

$$\mathbf{E}(\pi(A'_j)|\mathcal{C}) = \pi(A_j) + R(V \setminus A_j, A_j) - \sum_{i \neq j} R(A_j, A_i) \quad (11)$$

$$\begin{aligned} &= \pi(A_j) + R(V, A_j) - R(A_j, A_j) - \sum_{i \neq j} R(A_j, A_i) \\ &= \pi(A_j) + R(V, A_j) - S_{\mathcal{C}}(A_j). \end{aligned} \quad (12)$$

The next lemma shows that, given a sufficient advantage of opinion 1, after one round of voting opinion 1 remains the largest opinion. More precisely, the lemma gives lower bounds on the increase of the size of opinion 1 and on the increase of the advantage of this opinion over the other opinions.

► **Lemma 10.** *Assume $A_1 \leq 2n/3$, $A_1 - A_2 \geq Cn\sqrt{(\log n)/A_1}$ (requiring $A_1 \geq C^{2/3}n^{2/3}\log^{1/3}n$), where $C = 240\sqrt{2}$, and $\lambda \leq (A_1 - A_2)/(32n)$. Then with probability at least $1 - 1/n^2$,*

$$A'_1 \geq A_1 \left(1 + \frac{A_1 - A_2}{5n}\right). \quad (13)$$

$$\min_{2 \leq j \leq k} \{A'_1 - A'_j\} \geq (A_1 - A_2) \left(1 + \frac{A_1}{10n}\right), \quad (14)$$

Proof. Several times in this proof we use that $\pi(A_1) \leq 2/3$, which implies that $\pi(A'_1) \geq 1/3$. Our proof uses the following Chernoff bounds. If X is the sum of independent Bernoulli random variables, then for $\varepsilon \in (0, 1)$ and $\delta \geq 1$,

$$\Pr(X \geq (1 + \varepsilon)\mathbf{E}(X)), \Pr(X \leq (1 - \varepsilon)\mathbf{E}(X)) \leq \exp(-\varepsilon^2\mathbf{E}(X)/3), \quad (15)$$

$$\Pr(X \geq (1 + \delta)\mathbf{E}(X)) \leq \exp(-\delta\mathbf{E}(X)/3). \quad (16)$$

From Equation (12) and Lemmas 6 and 9, we have the following lower and upper bounds on $\mathbf{E}(\pi(A'_j)|\mathcal{C})$ for any $j \in [k]$.

$$\begin{aligned} \mathbf{E}(\pi(A'_j)|\mathcal{C}) &= \pi(A_j) + R(V, A_j) - S_{\mathcal{C}}(A_j) \\ &\geq \pi(A_j) + \pi(A_j)^2 - \lambda^2\pi(A_j)\pi(A'_j) \\ &\quad - \pi(A_j) \sum_{i=1}^k \pi(A_i)^2 - 2\lambda\pi(A_j)^{1/2} \sum_{i=1}^k \pi(A_i)^{3/2} - \lambda^2 \\ &\geq \pi(A_j) \left(1 + \pi(A_j) - \sum_{i=1}^k \pi(A_i)^2\right) - 2\lambda\pi(A_j)^{1/2}\pi(A_1)^{1/2} - (5/4)\lambda^2. \end{aligned} \quad (17)$$

$$\begin{aligned} \mathbf{E}(\pi(A'_j)|\mathcal{C}) &= \pi(A_j) + R(V, A_j) - S_{\mathcal{C}}(A_j) \\ &\leq \pi(A_j) + \pi(A_j)^2 + \lambda^2\pi(A_j)\pi(A'_j) - \pi(A_j) \sum_{i=1}^k \pi(A_i)^2 + 2\lambda\pi(A_j)^{1/2} \sum_{i=1}^k \pi(A_i)^{3/2} \\ &\leq \pi(A_j) \left(1 + \pi(A_j) - \sum_{i=1}^k \pi(A_i)^2\right) + (1/4)\lambda^2 + 2\lambda\pi(A_j)^{1/2}\pi(A_1)^{1/2}. \end{aligned} \quad (18)$$

13:10 Fast Plurality Consensus in Regular Expanders

We have $\lambda \leq \pi(A_1)/32$ and $\pi(A_1) \leq 2/3$, by assumption, so (17) and (18) imply

$$\pi(A_1)/2 \leq \mathbf{E}(\pi(A'_1)|\mathcal{C}) \leq 2\pi(A_1). \quad (19)$$

Define $\varepsilon_1 = \sqrt{\frac{9 \log n}{\mathbf{E}(A'_1|\mathcal{C})}} \leq \sqrt{\frac{18 \log n}{A_1}} < 1$. Using the Chernoff bounds (15), we get

$$\Pr(A'_1 \leq (1 - \varepsilon_1)\mathbf{E}(A'_1|\mathcal{C})|\mathcal{C}) \leq e^{-3 \log(n)} = n^{-3}. \quad (20)$$

For a fixed j , $2 \leq j \leq k$, define $\varepsilon_j = \sqrt{9(\log n)\mathbf{E}(A'_1|\mathcal{C})/\mathbf{E}(A'_j|\mathcal{C})}$ and show the following bound, using (15), for $\varepsilon_j \leq 1$, and (16), for $\varepsilon_j \leq 1$.

$$\Pr(A'_j \geq (1 + \varepsilon_j)\mathbf{E}(A'_j|\mathcal{C})|\mathcal{C}) \leq n^{-3}. \quad (21)$$

The bounds (20) and (21) imply that with probability at least $1 - n^{-2}$, for all $2 \leq j \leq k$,

$$\begin{aligned} A'_1 - A'_j &\geq (1 - \varepsilon_1)\mathbf{E}(A'_1|\mathcal{C}) - (1 + \varepsilon_j)\mathbf{E}(A'_j|\mathcal{C}) \\ &= \mathbf{E}(A'_1 - A'_j|\mathcal{C}) - 2\sqrt{9(\log n)\mathbf{E}(A'_1|\mathcal{C})}, \end{aligned} \quad (22)$$

and thus

$$\pi(A'_1) - \pi(A'_j) \geq \mathbf{E}(\pi(A'_1) - \pi(A'_j)|\mathcal{C}) - 2\sqrt{\frac{9(\log n)\mathbf{E}(\pi(A'_1)|\mathcal{C})}{n}}. \quad (23)$$

The right-hand side of (18) is non-increasing with increasing j , so for each $2 \leq j \leq k$,

$$\mathbf{E}(\pi(A'_j)|\mathcal{C}) \leq \pi(A_2) \left(1 + \pi(A_2) - \sum_{i=1}^k \pi(A_i)^2 \right) + (1/4)\lambda^2 + 2\lambda\pi(A_1). \quad (24)$$

Let $\Delta = \pi(A_1) - \pi(A_2)$. Inequalities (17) and (24) give for each $2 \leq j \leq k$,

$$\begin{aligned} \mathbf{E}(\pi(A'_1) - \pi(A'_j)|\mathcal{C}) &\geq \Delta \left(1 + \pi(A_1) + \pi(A_2) - \sum_{i=1}^k \pi(A_i)^2 \right) - 4\lambda\pi(A_1) - (3/2)\lambda^2 \\ &\geq \Delta(1 + \pi(A_1)\pi(A_1^c)) - 4\lambda\pi(A_1) - 2\lambda^2 \end{aligned} \quad (25)$$

$$\geq \Delta + \Delta\pi(A_1)/7. \quad (26)$$

Inequality (25) holds because $\sum_{i=2}^k \pi(A_i)^2 \leq \pi(A_2)$. In the last step we used that $\pi(A_1^c) \geq 1/3$ and $\lambda \leq \Delta/32$. From (23), (26) and (19), with probability at least $1 - n^{-2}$,

$$\begin{aligned} \min_{2 \leq j \leq k} \{\pi(A'_1) - \pi(A'_j)\} &\geq \mathbf{E}(\pi(A'_1) - \pi(A'_j)|\mathcal{C}) - \frac{\varepsilon_1}{n}\mathbf{E}(A'_1|\mathcal{C}) - \frac{\varepsilon_j}{n}\mathbf{E}(A'_j|\mathcal{C}) \\ &\geq \Delta \left(1 + \pi(A_1)/7 - \frac{6}{\Delta} \sqrt{\frac{2 \log n}{n} \pi(A_1)} \right). \end{aligned}$$

By assumption, $\Delta \geq 240\sqrt{2 \log(n)/A_1}$, so with probability at least $1 - n^{-2}$,

$$\min_{2 \leq j \leq k} \{\pi(A'_1) - \pi(A'_j)\} \geq \Delta(1 + \pi(A_1)/10), \quad (27)$$

and we get we get (14). This also proves that w.h.p. opinion 1 remains the majority opinion. The order between the other opinions might change.

To get information about the increase in the number of vertices with opinion 1, we use Equation (17) with $j = 1$ and the assumption that $\lambda \leq \Delta/32$. We obtain

$$\begin{aligned} \mathbf{E}(\pi(A'_1)|\mathcal{C}) &\geq \pi(A_1)(1 + \pi(A_1) - \sum_{i=1}^k \pi(A_i)^2) - \Delta\pi(A_1)/16 - \Delta^2/(32)^2 \\ &\geq \pi(A_1)(1 + \pi(A_1) - \pi(A_1)^2 - \pi(A_2)\pi(A_1^c) - \Delta/16 - \Delta/(32)^2) \\ &> \pi(A_1)(1 + \Delta/4). \end{aligned} \quad (28)$$

By using Chernoff bounds (15) with $\varepsilon = \sqrt{\frac{9 \log n}{\mathbf{E}(A'_1|\mathcal{C})}}$ and Inequalities (28) and (19), with probability at least $1 - n^{-2}$,

$$\begin{aligned} A'_1 &\geq A_1(1 + \Delta/4) - \sqrt{\mathbf{E}(A'_1|\mathcal{C})9 \log n} \geq A_1(1 + \Delta/4) - \sqrt{18A_1 \log n} \\ &= A_1(1 + \Delta/4 - 3\sqrt{2}\sqrt{\log n/A_1}). \end{aligned} \quad (29)$$

From the assumptions of the lemma, we have $\Delta/20 = (A_1 - A_2)/(20n) \geq 3\sqrt{2}\sqrt{\log n/A_1}$. Therefore (29) implies $A'_1 \geq A_1(1 + \Delta/5)$, which is the same as (13). \blacktriangleleft

► Lemma 11. *Assume $A_1 \leq 2n/3$, $A_1 - A_2 \geq Cn\sqrt{(\log n)/A_1}$ and $\lambda \leq (A_1 - A_2)/(32n)$. With probability at least $1 - 1/n$, after at most $O((n/A_1) \log(A_1/(A_1 - A_2)))$ rounds, the number of vertices with opinion 1 is at least $2n/3$.*

Proof. We apply Lemma 10 to consecutive rounds until the size of opinion 1 reaches $2n/3$. Since w.h.p. the difference between the size of opinion 1 and the size of the second largest opinion increases, our assumption about λ in Lemma 10 is maintained from round to round. If the ordering of the opinions according to size changes at any step t , we relabel the opinions so that $A_1(t) \geq A_2(t) \cdots \geq A_k(t)$. Lemma 10 implies that w.h.p. opinion 1 remains the largest opinion, and thus never relabeled.

Denote by $x(i)$ the fraction of vertices with opinion 1 at the end of round i , and by $y(i)$ the difference between the fraction of vertices with opinion 1 and the fraction of vertices with the second largest opinion. Thus $x(0) = \pi(A_1)$, and $y(0) = \Delta = \pi(A_1) - \pi(A_2) < x(0)$. By (13) and (14) and induction on the number of rounds, with probability at least $1 - 1/n$, for each round $1 \leq i \leq n$, if $x(i) < 2/3$, then

$$x(i) \geq x(i-1)(1 + y(i-1)/5), \quad (30)$$

$$y(i) \geq y(i-1)(1 + x(i-1)/10). \quad (31)$$

Iterating (30) and (31) for $j = \lceil 10/x(0) \rceil < n$ rounds, we get $y(j) \geq 2y(0)$ and $x(j) \geq x(0) + y(0)$, or $x(i) \geq 2/3$ for some $i \leq j$. Repeating this $r = \lceil \log_2(x(0)/y(0)) \rceil$ times, we get for round $i_1 = rj < n$, $y(i_1) \geq x(0)$ and $x(i_1) \geq x(0) + y(0) + 2y(0) + 4y(0) \cdots + 2^{r-1}y(0) \geq 2x(0)$, or $x(i) \geq 2/3$ for some $i \leq i_1$.

If for some $q \geq 1$, $y(i_q) \geq 2^{q-1}x(0)$ and $x(i_q) \geq 2^q x(0)$, or $x(i) \geq 2/3$ for some $i \leq i_q$, then at the end of round $i_{q+1} = i_q + \lceil 10/(2^q x(0)) \rceil$, either $y(i_{q+1}) \geq 2^q x(0)$ and $x(i_{q+1}) \geq 2^{q+1}x(0)$, or $x(i) \geq 2/3$ for some $i \leq i_{q+1}$, or $i_{q+1} > n$. Taking $q = \lceil \log_2(1/x(0)) \rceil$, we have $i_q = O((1/x(0)) \log(x(0)/y(0))) = O((n/A_1) \log(A_1/(A_1 - A_2)))$ (observe that $i_q < n$) and $2^q x(0) \geq 1$, so we must have $x(i) \geq 2/3$ for some $i \leq i_q$. \blacktriangleleft

When the largest opinion reaches the size $2n/3$, it will take over the whole graph within additional $O(\log n)$ rounds. The progress of voting in this final stage would be slowest, if all minority opinions were joined together into a single “second” opinion. The next lemma can

be proven in the similar way as it is proven in [14] that two-sample voting finishes in $O(\log n)$ rounds, if there are two opinions, the majority opinion has size at least cn , for a constant $c > 1/2$, and λ is sufficiently small. Theorem 1 follows immediately from Lemmas 11 and 12.

► **Lemma 12.** *Let G be a connected regular graph with $\lambda \leq 1/4$. If the majority opinion has size at least $2n/3$, then with probability at least $1 - n^{-2}$, the voting finishes within $O(\log n)$ rounds.*

5 Reducing Three-sample voting to Two-sample voting

In the three-sample voting process considered in [4], each vertex v samples in each round three random neighbours with replacement, collecting their opinions, say, $Y_{v,1}, Y_{v,2}, Y_{v,3}$. Vertex v changes its opinion to the majority of $\{Y_{v,1}, Y_{v,2}, Y_{v,3}\}$, or, if there is no majority, to $Y_{v,1}$. We show that our analysis of two sampling applies, with small modifications, also to three sampling, giving the same bounds as in Theorem 1. The crucial idea is to view the three sampling process in the following equivalent way.

Suppose in a given round we have k opinions, let $\mathcal{C} = (A_1, A_2, \dots, A_n)$ be the partition of the vertices into the opinion classes and let A'_j be the vertices with opinion j at the next round. Each vertex v decides on its opinion for the next round in the following way. First, v takes on the opinion $Y_{v,1}$. Let A''_j , $1 \leq j \leq k$, be the opinion classes after this initial update. Now v obtains its final opinion by the two-sampling decision using opinions $Y_{v,2}, Y_{v,3}$ (taken in the original partition \mathcal{C}). Observe that classes A''_j result from choosing only one (random) vertex, that is, they are obtained by one round of the standard (single-sample and slow) pull voting. While the sizes of opinions may change in one round of the standard pull voting, the expected sizes are equal to the initial sizes; see, for example [22]. That is, we have $\mathbf{E}(\pi(A''_j|\mathcal{C})) = \pi(A_j)$.

The following lemma implies that the analysis of two-sample voting can be updated to three-sample voting by putting $\mathbf{E}(S_{\mathcal{C}}(A''_j)|\mathcal{C})$ in place of $S_{\mathcal{C}}(A_j)$ and using the bounds (33) and (34) on $\mathbf{E}(S_{\mathcal{C}}(A''_j)|\mathcal{C})$, which are exactly the same as the bounds on $S_{\mathcal{C}}(A_j)$ in Lemma 9. The proof of the lemma and further details will be included in the full version of the paper.

► **Lemma 13.** *Let G be a connected graph and let $\mathcal{C} = (A_1, \dots, A_k)$ partition V . Then*

$$\mathbf{E}(\pi(A'_j|\mathcal{C})) = \pi(A_j) + R(V, A_j) - \mathbf{E}(S_{\mathcal{C}}(A''_j)|\mathcal{C}), \quad (32)$$

$$\mathbf{E}(S_{\mathcal{C}}(A''_j)|\mathcal{C}) \geq \pi(A_j) \sum_{i=1}^k \pi(A_i)^2 - 2\lambda\pi(A_j)^{1/2} \sum_{i=1}^k \pi(A_i)^{3/2}, \quad (33)$$

$$\mathbf{E}(S_{\mathcal{C}}(A''_j)|\mathcal{C}) \leq \pi(A_j) \sum_{i=1}^k \pi(A_i)^2 + 2\lambda\pi(A_j)^{1/2} \sum_{i=1}^k \pi(A_i)^{3/2} + \lambda^2. \quad (34)$$

If $\mathcal{C} = (A, B)$, then $\mathbf{E}(S_{\mathcal{C}}(B'')|\mathcal{C}) \geq \pi(B)/2$.

6 Algorithm Pushy and Proof of Theorem 3

We describe a push-voting algorithm for reaching consensus on a graph G .

Algorithm Pushy

Repeat L times

begin

Each vertex activates with probability p and sends its opinion to all neighbours

Each vertex keeps a tally of how many opinions of each type it has received
 Let $M(v)$ be the set of opinions received by v which have the maximal count
 If $|M(v)| \geq 1$, vertex v picks an opinion u.a.r. from $M(v)$
 If $|M(v)| = 0$ (i.e. v has not received any opinion), v keeps its current opinion
end

The following lemma rephrases Theorem 3.

► **Lemma 14.** *Let $G = K_n$ and $p = \alpha \log n/n$, for a sufficiently large constant $\alpha > 0$. Then with probability $1 - o(1)$, after $L = 4$ rounds there is a unique remaining opinion, irrespective of the initial number and distribution of opinions.*

Proof. On the complete graph K_n , for each round, the sets $M(v) \equiv M$ are the same for all vertices. Consider one round and let $m = |M|$. The expected number of activated vertices is $\alpha \log n$ and the actual number is concentrated around this value (see (15)), so w.h.p. at least one but fewer than $\beta \log n$ vertices activate in this round, where $\beta = \beta(\alpha)$ is a constant sufficiently larger than α . Thus w.h.p. $1 \leq m < \beta \log n$. If n vertices pick u.a.r. from the set M , then w.h.p. each opinion i is chosen $N_i = (1 + o(1))n/m$ times (use again (15)). Thus w.h.p. the number of opinions in the next round is $k = m < \beta \log n$.

Consider now a round $j \geq 2$ with the number of opinions $k \geq 2$ (no consensus yet). We have w.h.p. $2 \leq k < \beta \log n$ and each opinion i is represented by $N_i = (1 + o(1))n/k$ vertices. We will upper bound m , which is the number of opinions with the maximal number of activations, giving the number of opinions for the next round. The number of activations of opinion i is a binomial random variable $X_i \sim \text{Bin}(N_i, p)$ independent of any other opinion. Henceforth we use X and $N = (1 + o(1))n/k$ to denote X_i and N_i . We use the inequalities for binomial probabilities given below, which hold as follows: (35) for any h , (36) for $h > Np$, (37) for $h > Np$, $h = o(\sqrt{N})$, $hp = o(1)$ and $\theta = (1 + O(1/h) + O(h^2/N))/\sqrt{2\pi}$.

$$\Pr(X = h) \leq 1/\sqrt{Np(1-p)}, \quad (35)$$

$$\Pr(X = h) \leq \Pr(X \geq h) \leq \Pr(X = h) \cdot \left(1 - \frac{N-h}{h+1} \frac{p}{1-p}\right)^{-1}, \quad (36)$$

$$\Pr(X = h) = \binom{N}{h} p^h (1-p)^{N-h} = \frac{\theta}{\sqrt{h}} \left(\frac{eNp}{h}\right)^h e^{-Np}. \quad (37)$$

We consider the following four cases, where $\omega = (\log \log n)/\log \log \log n$.

- C1:** $(\beta \log n)/\omega \leq k < \beta \log n$,
- C2:** $(\log n)/\log \log n \leq k < (\beta \log n)/\omega$,
- C3:** $\beta(\log n)^{1/6} \leq k < (\log n)/\log \log n$,
- C4:** $2 \leq k < \beta(\log n)^{1/6}$.

Case C1: $(\beta \log n)/\omega \leq k < \beta \log n$. Let $h = (11/12)\omega = (11/12)(\log \log n)/\log \log \log n$. Referring to (36) and recalling that β is sufficiently larger than α ,

$$\frac{N-h}{h+1} \frac{p}{1-p} \leq (1 + o(1)) \frac{Np}{h} \leq (1 + o(1)) \frac{\alpha}{(11/12)\beta} < 1/2.$$

Thus from (36), $\Pr(X \geq h) \leq 2 \cdot \Pr(X = h)$.

13:14 Fast Plurality Consensus in Regular Expanders

Let $N = (1 + o(1))n/k$, $p = f/n$ where $f = \alpha \log n$. Let $k = (\log n)/\sigma$ where $1/\beta \leq \sigma \leq \omega/\beta$. Using (37),

$$\begin{aligned} \Pr(X = h) &= \frac{\theta}{\sqrt{h}} \left(\frac{ne}{kh}p\right)^h e^{-np/k} = \frac{\theta}{\sqrt{h}} \left(\frac{ef}{hk}\right)^h e^{-f/k} = \frac{\theta}{\sqrt{h}} \left(\frac{e\alpha\sigma}{h}\right)^h e^{-\alpha\sigma} \\ &= \theta \exp\left(-\alpha\sigma - h \log\left(\frac{h}{e\alpha\sigma}\right) - (1/2) \log h\right) \\ &= \theta \exp\left(-h \log h \left(1 + \frac{\alpha\sigma}{h \log h} + \frac{1}{2h} - \frac{\log(e\alpha\sigma)}{h \log h}\right)\right) \\ &= \exp(-h \log h (1 + o(1))) \\ &= (\log n)^{-(11/12)(1+o(1))} = \rho. \end{aligned}$$

The probability that no opinion is represented at least h times is $(1 - \rho)^k = o(1)$. The expected number of opinions represented at least h times is $k\Pr(X \geq h)$, and

$$k\Pr(X \geq h) \leq 2k\Pr(X = h) = O(\log n)^{(1/12)(1+o(1))} = o((\log n)^{1/6}).$$

Thus with probability $1 - o(1)$, there is an opinion activated at least h times and the number m of opinions with the maximal number of activations (that is, the number of opinions for the next round) is less than $\beta(\log n)^{1/6}$. That is, in the next round we will have Case C4 or $k = 1$.

Case C2: $(\log n)/\log \log n \leq k < (\beta \log n)/\omega$. For any h , using (35),

$$\Pr(X = h) \leq \frac{1}{\sqrt{np/k}} = \frac{1}{\sqrt{\omega\alpha/\beta}}.$$

Let $j = k\gamma/\sqrt{\omega}$, where $\gamma \geq 3e\sqrt{B/A}$. The probability that more than j opinions are transmitted the same (maximum) number of times is at most

$$\binom{k}{j} \left(\frac{1}{\sqrt{\omega\alpha/\beta}}\right)^{j-1} \leq O(\sqrt{\omega}) \left(\frac{ke}{j}\right)^j \left(\frac{1}{\sqrt{\omega\alpha/\beta}}\right)^j = O(\sqrt{\omega}) e^{-j} = o(1).$$

Thus the number of opinions in the next round will be $j = o(\log n / \log \log n)$, with probability $1 - o(1)$. That is, in the next round we will have Case C3 or Case C4 or $k = 1$.

Case C3: $\beta(\log n)^{1/6} \leq k < (\log n)/(\log \log n)$. For $Np \geq \alpha \log \log n$, using (36) and (37),

$$\Pr(X \geq eNp) \leq \frac{O(1)}{\sqrt{Np}} e^{-Np} = \frac{o(1)}{(\log n)^\alpha},$$

so with probability $1 - o(1)$ all opinions are transmitted at most eNp times. For $h = DNp$, $D > 1$, by direct calculation

$$\frac{\Pr(X = h+1)}{\Pr(X = h)} = \frac{p}{1-p} \frac{N-h}{h+1} = (1 + o(1)) \frac{1}{D}.$$

As $\Pr(X \geq \mathbf{E}(X)) = \gamma$ for some $\gamma > 0$ constant, it follows that with probability $1 - o(1)$ there exists $1 < D < e$ such that and for $h = DNp$, the expected number of opinions activated at least h times is

$$k\Pr(X \geq h) = \Theta((\log n)^{1/7}).$$

Thus with probability $1 - o(1)$, the number of opinions with the maximal number of activations is at most $\beta(\log n)^{1/7}$. That is, in the next round we will have Case C4 or $k = 1$.

Case C4: $2 \leq k < \beta(\log n)^{1/6}$. Using (35) with $Np = \Omega((\log n)^{5/6})$, we get

$$\Pr(X = h) = O\left(1/(\log n)^{5/12}\right).$$

Let X_i be the number of activated vertices of opinion $i = 1, \dots, k$, then

$$\Pr(\exists(i, j) \text{ s.t. } X_i = X_j) = O\left(\frac{k^2}{(\log n)^{5/12}}\right) = O\left(\frac{1}{(\log n)^{1/12}}\right) = o(1),$$

so, with probability $1 - o(1)$, a unique maximum opinion remains, that is, $k = 1$ in the next round.

The probability that no vertex activates in one round is $(1 - p)^n = O(n^{-\alpha})$. Thus after 4 rounds (if Case C2, then no C3 or C4) a unique opinion remains with probability $1 - o(1)$. ◀

References

- 1 D. Aldous. Meeting times for independent Markov chains. *Stochastic Processes and their Applications* 38(2):185–193, (1991).
- 2 D. Aldous and J. Fill. *Reversible Markov Chains and Random Walks on Graphs*, <http://stat-www.berkeley.edu/pub/users/aldous/RWG/book.html>.
- 3 L. Becchetti, A. Clementi, E. Natale, F. Pasquale and R. Silvestri. Plurality consensus in the gossip model. *SODA 2015*, pages 371–390.
- 4 L. Becchetti, A. Clementi, E. Natale, F. Pasquale, R. Silvestri and L. Trevisan. Simple dynamics for plurality consensus. In *Proceedings of the 26th ACM symposium on Parallelism in Algorithms and Architectures (SPAA 2014)*, pages 247–256, (2014).
- 5 L. Becchetti, A.E.F. Clementi, E. Natale, F. Pasquale and L. Trevisan. Stabilizing Consensus with Many Opinions. *SODA 2016*, pages 620–635.
- 6 P. Berenbrink, A.E.F. Clementi, R. Elsässer, P. Kling, F. Mallmann-Trenn, E. Natale. Ignore or Comply?: On Breaking Symmetry in Consensus. *PODC 2017*, pages 335–344.
- 7 P. Berenbrink, Tom Friedetzky, G. Giakkoupis, P. Kling. Efficient Plurality Consensus, Or: the Benefits of Cleaning up from Time to Time. *ICALP 2016*, pages 136:1–136:14.
- 8 P. Berenbrink, T. Friedetzky, P. Kling, F. Mallmann-Trenn and C. Wastell. Plurality Consensus in Arbitrary Graphs: Lessons Learned from Load Balancing. In *Proceedings of the 24th Annual European Symposium on Algorithms, (ESA 2016)*, pages 10:1–10:18, (2016).
- 9 S. Brahma, S. Macharla, S. P. Pal, S. R. Singh. Fair Leader Election by Randomized Voting. In *ICDCIT 2004*, pages 22–31, (2004).
- 10 A. Broder, A. Frieze, S. Suen and E. Upfal. Optimal construction of edge disjoint paths in random graphs. *SIAM Journal on Computing*, 28(2), pages 541–573, (1999).
- 11 C. Cooper, M.E. Dyer, A.M. Frieze, N. Rivera. Discordant Voting Processes on Finite Graphs. *ICALP 2016*, pages 145:1–145:13.
- 12 C. Cooper, R. Elsässer, H. Ono, T. Radzik. Coalescing Random Walks and Voting on Connected Graphs. *SIAM Journal of Discrete Math*, 27, pages 1748–1758, (2013).
- 13 C. Cooper, R. Elsässer and T. Radzik. The power of two choices in distributed voting, *ICALP 2014*, pages 435–446, (2014).
- 14 C. Cooper, R. Elsässer, T. Radzik, N. Rivera and T. Shiraga. Fast consensus for voting on general expander graphs. In *DISC 2015 – 29th International Symposium on Distributed Computing*, Springer-Verlag LNCS 9363, pages 248–262. (2015).
- 15 J. T. Cox. Coalescing random walks and voter model consensus times on the torus in \mathbb{Z}^d . *The Annals of Probability* 17(4):1333–1366, (1989).
- 16 X. Deng and C. Papadimitriou. On the Complexity of Cooperative Solution Concepts. *Mathematics of Operations Research* 19, pages 257–266, (1994).

13:16 Fast Plurality Consensus in Regular Expanders

- 17 R. Elsässer, T. Friedetzky, D. Kaaser, F. Mallmann-Trenn and H. Trinker. *Rapid Asynchronous Plurality Consensus*. [arXiv:1602.04667](https://arxiv.org/abs/1602.04667), February 2017.
- 18 J. Friedman. A proof of Alon's second eigenvalue conjecture. In *STOC 2003: Proc. 35th Annual ACM Symposium on Theory of Computing*, pages 720–724, (2003).
- 19 M. Ghaffari and J. Lengler. *Tight Analysis for the 3-Majority Consensus Dynamics*. [arXiv:1705.05583](https://arxiv.org/abs/1705.05583), May 2017.
- 20 M. Ghaffari and M. Parter. A Polylogarithmic Gossip Algorithm for Plurality Consensus. In *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing (PODC 2016)*, pages 117–126, (2016).
- 21 D. Gifford. Weighted Voting for Replicated Data. In *SOSP 1979: Proceedings of the 7th ACM Symposium on Operating Systems Principles*, pages 150–162, (1979).
- 22 Y. Hassin and D. Peleg. Distributed probabilistic polling and applications to proportionate agreement. *Information & Computation*, 171, pages 248–268, (2001).
- 23 B. Johnson, *Design and Analysis of Fault Tolerant Digital Systems*, Addison-Wesley, (1989).