Open access • Book Chapter • DOI:10.1007/978-3-662-48653-5_17

Fast Consensus for Voting on General Expander Graphs — Source link

Colin Cooper, Robert Elsässer, Tomasz Radzik, Nicolás Rivera ...+1 more authors

Institutions: King's College London, University of Salzburg, Kyushu University

Published on: 07 Oct 2015 - International Symposium on Distributed Computing

Topics: Vertex (geometry), Regular graph, Random graph, Expander graph and Voting

Fast consensus for voting on general expander graphs

Colin Cooper, Robert Elsässer, Tomasz Radzik, Nicolas Rivera, Takeharu Shiraga

Fast consensus for voting on general expander graphs*

Colin Cooper[1], Robert Elsässer[2], Tomasz Radzik[1], Nicolás Rivera[1], and Takeharu Shiraga[3]

[1] Department of Informatics, King's College London, United Kingdom
colin.cooper@kcl.ac.uk, tomasz.radzik@kcl.ac.uk, nicolas.rivera@kcl.ac.uk

[2] Department of Computer Sciences, University of Salzburg, Austria
elsa@cosy.sbg.ac.at

[3] Theoretical Computer Science Group, Department of Informatics, Kyushu University, Fukuoka, Japan
shiraga@tcslab.csce.kyushu-u.ac.jp

Abstract. Distributed voting is a fundamental topic in distributed computing. In the standard model of pull voting, at each step every vertex chooses a neighbour uniformly at random and adopts its opinion. The voting is completed when all vertices hold the same opinion. In the simplest case, each vertex initially holds one of two different opinions. This partitions the vertices into arbitrary sets A and B. For many graphs, including regular graphs and irrespective of their expansion properties, if both A and B are sufficiently large sets, then pull voting requires $\Omega(n)$ expected steps, where n is the number of vertices of the graph.

In this paper we consider a related class of voting processes based on sampling two opinions. In the simplest case, every vertex v chooses two random neighbours at each step. If both these neighbours have the same opinion, then v adopts this opinion. Otherwise, v keeps its own opinion. Let G be a connected graph with n vertices and m edges. Let P be the transition matrix of a simple random walk on G with second largest eigenvalue $\lambda < 1/\sqrt{2}$. We show that if the initial imbalance in degree between the two opinions satisfies $|d(A) - d(B)|/2m \geq 2\lambda^2$, then with high probability voting completes in $O(\log n)$ steps, and the opinion with the larger initial degree wins.

The condition that $\lambda < 1/\sqrt{2}$ includes many classes of expanders, for example random d-regular graphs where $d \geq 10$. If however $1/\sqrt{2} \leq \lambda(P) \leq 1 - \epsilon$ for a constant $\epsilon > 0$, or only a bound on the conductance of the graph is known, the sampling process can be modified so that voting still provably completes in $O(\log n)$ steps with high probability. The modification uses two sampling based on probing to a fixed depth $O(1/\epsilon)$ from any vertex.

In its most general form our voting process allows vertices to bias their sampling of opinions among their neighbours to achieve a desired outcome. This is done by allocating weights to edges.

* This work was supported in part by EPSRC grant EP/M005038/1, "Randomized algorithms for computer networks". N. Rivera was supported by funding from Becas CHILE.

1 Introduction

1.1 Background on distributed pull voting

Distributed voting has applications in various fields including consensus and leader election in large networks [3, 14], serialisation of read-write in replicated databases [13] and the analysis of social behaviour in game theory [11]. Voting algorithms are usually simple, fault-tolerant, and easy to implement [14, 16].

One simple form of distributed voting is *pull voting*. In the beginning each vertex of a connected undirected graph has an initial opinion. The voting process proceeds synchronously in discrete time steps called rounds. During each round, each vertex independently contacts a random neighbour and adopts the opinion of that neighbour.

In the *two-opinion voter model*, all vertices initially hold one of two opinions. Hassin and Peleg [14] and Nakata *et al.* [20] considered the two-opinion voter model and its application to consensus problems in distributed systems. Let $G = (V, E)$ be an undirected connected graph with n vertices and m edges. Let the opinions be labeled 0 and 1, and let A be the set of vertices with opinion 0 and B the set of vertices with opinion 1; where $A \cup B = V$. Let $d(v)$ be the degree of a vertex v and $d(S) = \sum_{v \in S} d(v)$ the degree of a set S. Thus $d(A)$ is the initial degree of opinion 0 and $d(A) + d(B) = 2m$. We say that A wins (equiv. opinion 0 wins), if all vertices eventually adopt the opinion held initially by the set A. Let P_A be the probability that opinion A wins the vote in the two-opinion model. The central result of [14] and [20] is that

$$P_A = \frac{d(A)}{2m}. \tag{1}$$

Thus in the case of connected regular graphs, the probability that A wins is proportional to the original size of A, irrespective of the graph structure.

Apart from the probability of winning the vote, another quantity of interest is the time taken for voting to complete. The completion time T of a voting process is the number of rounds needed for a single opinion to emerge. This is normally measured in terms of its expectation $\mathbf{E}T$. It is proven in [14] that $\mathbf{E}T = O(n^3 \log n)$ for general graphs.

It was shown in [8] that the completion time on any connected graph G is upper bounded with high probability (w.h.p.) by $O(n/(\nu(1-\lambda))$, where λ is the second largest eigenvalue of the transition matrix of random walk on G and $\nu = n \sum_{v \in V} d^2(v)/(2m)^2$ indicates the regularity of G ($1 \leq \nu \leq n^2/(2m)$, with $\nu = 1$ for regular graphs). Tighter bounds can be derived for some specific classes of graphs. For example, it is proven in [7] that in the case of random d-regular graphs, w.h.p. $\mathbf{E}T \sim 2n(d-1)/(d-2)$. This means that two-opinion voting (almost always) needs $\Theta(n)$ time to complete on random d-regular graphs.

Thus the performance of the two-opinion pull-voting seems unsatisfactory in two ways. Firstly, it is reasonable to require that a clear majority opinion should win with high probability. From (1), even if initially only a single vertex v holds opinion A, then this opinion wins with probability $P_A = d(v)/2m$. Secondly, the expected completion time is $\Omega(n)$ on many classes of graphs, including regular expanders and complete graphs. This seems a long time to wait to resolve a dispute between two opinions in the context of distributed systems. A more reasonable waiting time would depend on the graph diameter, which is $O(\log n)$ for many classes of expanders.

To address these issues, we consider a modified version of pull voting in which each vertex v randomly queries two neighbours at each step. On the basis of the sample taken, vertex v revises its opinion as follows. If both neighbours have the same opinion, the calling vertex v adopts this opinion. If the two opinions differ, the calling vertex v retains its current opinion in this round. To distinguish this process from the conventional pull voting, we refer to it as *two-sample voting.* The aim of two-sample voting is to ensure that voting finishes quickly and the initial majority wins. Two-sample voting is intrinsically attractive, as it seems to mirror the way people behave. If you hear it twice it must be true.

1.2 Main Results

Two-sample voting is used in [9] to speed up time to consensus for pull voting on d-regular expander graphs. It was shown that synchronous two-sample-voting completed in $O(\log n)$ time w.h.p. even under adversarial conditions, and also that the initial majority opinion wins, provided sufficient initial imbalance between the sizes of the two opinions. In related work, in a non-adversarial context, Abdullah and Draief [1] obtained a $O(\log_d \log_d n)$ bound for the majority multi-sample-voting on d-regular graphs where at least five neighbours are consulted (hence requiring $d \geq 5$). They also proved that this bound is asymptotically best possible for a wide class of voting protocols. For the case of the complete graph, Cruise and Ganesh [10] made a more general analysis of multi-sample-voting strategies.

In this paper we extend the analysis of two-sample voting from [9] to general (inhomogeneous) expander graphs, with no regularity restriction on the vertex degrees, and prove that the speed of this protocol remains $O(\log n)$ (Theorem 1). However, the property that the initial majority opinion wins is found to be restricted to regular graphs. For inhomogeneous graphs, the party with the largest initial degree wins, provided sufficient initial imbalance between the degrees of the two opinions. As a special case, we get a stronger result for two sample-voting on regular expander graphs than in [9] by significantly reducing the required initial imbalance between the sizes of the two opinions.

Our analysis uses a different approach from previous work on two sample voting. The main technical theorem (Theorem 3) is based on the connection between the voting and the related random walk process. Using this theorem, we can obtain results for a wide range of protocols in which vertices sample neighbours at random according to predetermined edge weights. We refer to this generalization of two-sample-voting as *best-of-two voting*, and reserve two-sample-voting for the special case where neighbours are chosen uniformly at random (equiv. the edges weights are uniform). We show that the speed of best-of-two voting is $O(\log n)$ for general weighted expanders (Theorem 2).

Additionally, we consider an extension of the best-of-two voting, which we refer to as *k-extended best-of-two voting*. In this process in each round every vertex v performs two independent k step random walks. If the vertices visited by the walks at the k-th step have the same opinion, vertex v adopts this opinion; otherwise v keeps its current opinion in this round. The case $k = 1$ is best-of-two voting, and $k \geq 2$ extends the model by allowing vertices to obtain opinions beyond their immediate neighbourhood. Once again the protocol takes $O(\log n)$ rounds for general expanders (Corollary 1). It will emerge that k-extended best-of-two voting can be seen as best-of-two voting in a different weighted graph. A major advantage is that by increasing the value of k, Corollary 1 can be applied to graphs with poor expansion, which are not covered by Theorems 1 and 2.

Voting in weighted graphs. For an undirected connected weighted graph $G = (V, E)$, let $w(u, v)$ denote the positive weight assigned to an edge $(u, v) \in E$. We use $N(v)$ for the set of neighbours of v and define $w(v) = \sum_{x \in N(v)} w(v, x)$ the weight of v, $w(S) = \sum_{u \in S} w(u)$ the weight of a set $S \subseteq V$, and $w(G) = \sum_{v \in V} w(v)$ the (total) weight of the graph. Best-of-two voting is a synchronous process in which during each step, every vertex $v \in V$ independently queries two neighbours u' and u'', not necessary distinct, which are chosen randomly using the selection probabilities proportional to the edge weights. If u' and u'' have at the beginning of the step the same opinion X, then at the end of this step v also has opinion X. If u' and u'' have different opinions, then at the end of the step v has the same opinion as it had at the beginning of this step. Using the selection probabilities proportional to the edge weights means that v selects an ordered pair of its neighbours $\langle u', u'' \rangle$ (not necessarily distinct) with probability $P(v, u')P(v, u'')$, where $P(v, u) = w(v, u)/w(v)$. The probability that a vertex v in A moves to B at a given step is equal to

$$\mathbf{Pr}(v \text{ chooses twice in } B) = \left(\sum_{u \in B \cap N(v)} \frac{w(v, u)}{w(v)} \right)^2 = \left(\sum_{u \in B \cap N(v)} P(v, u) \right)^2 .$$

Two-sample voting can be viewed as the special case of best-of-two voting when the edge weights are uniform: $w(e) = 1$ for each $e \in E$, $w(v) = d(v)$, $w(S) = d(S)$, $w(G) = 2m$ and $P(v, u) = 1/d(v)$.

Observe that P is the transition matrix of a reversible random walk on G. We assume that G is not bipartite so that this random walk is aperiodic and has a well defined stationary distribution π: $\pi(u) = w(u)/w(G)$. (For a bipartite graph $G = (V_1 \cup V_2, E)$, the voting would never converge, if one opinion resided on V_1 and the other on V_2.) Conversely, if P is the transition matrix of a reversible random walk on G with the stationary distribution π (that is, $P(u, v) > 0$ iff $(u, v) \in E$, $\sum_{u \in N(v)} P(v, u) = 1$ for each $v \in V$, and $\pi(u)P(u, v) = \pi(v)P(v, u)$ for each $(u, v) \in E$), then we can associate positive edge weights $\boldsymbol{w} = (w(e), e \in E)$ with P so that the transition probabilities of P are proportional to these weights (set $w(u, v) = \pi(u)P(u, v)$ to have $P(u, v) = w(v, u)/w(v)$). For a set $S \subseteq V$, we have $\pi(S) = \sum_{u \in S} \pi(u) = w(S)/w(G)$, so π and w are the same measures of the subsets of vertices up to the scaling factor $w(G)$. For the transition matrix P of simple (uniform) random walk, which corresponds to two-sample voting, $\pi(v) = d(v)/(2m)$ for each $v \in V$, and $\pi(S) = d(S)/(2m)$ for a subset $S \subseteq V$.

Thus best-of-two voting can be defined equivalently either by specifying edge weights or a transition matrix P of a reversible aperiodic random walk. The transition matrix of simple (uniform) random walk gives two-sample voting. We stress that we do not establish a relation between the best-of-two voting process based on matrix P of selection probabilities and the random walk process based on matrix P of transition probabilities other than that both processes use the same matrix P (but for somewhat different purposes). Some properties of such matrices, which have been developed largely in the context of analysing random walks, turn out to be useful for studying the best-of-two voting.

Let the eigenvalues of matrix P be ordered in decreasing value $1 = \lambda_1(P) > \lambda_2(P) \geq \cdots \geq \lambda_n(P) > -1$, and let $\lambda = \lambda(P) = \max(|\lambda_2(P)|, |\lambda_n(P)|)$. An expander graph G (or simply, an expander) is commonly defined as a graph with $\lambda(P)$ bounded away from 1, where P is the transition matrix of the simple random walk G. Generalising this, a weighted expander is a weighted graph with $\lambda(P)$ bounded away from 1, where P is the transition matrix of the random walk G with transition probabilities proportional to the edge weights (see, e.g. [2]).

In the formal statements of our results, "with high probability" (w.h.p.) means with probability at least $1 - 1/n^{\alpha}$ for some constant α. Before discussing our results in their most general form, we give the findings for two-sample-voting, and also some specific examples.

Theorem 1. (Two-sample voting) *Let G be a connected non-bipartite graph with n vertices and m edges, let P be the transition matrix of a simple random walk on G, and let $\nu = n\left(\sum_{vV} d^2(v)\right)/(2m)^2$. Let A and B denote the sets of vertices of G with initial opinions of the two types, and let $\epsilon_0 = |d(A) - d(B)|/2m$ denote the initial degree imbalance between these sets.*

Provided $\lambda = \lambda(P) \leq 1/\sqrt{2} - \delta$ for arbitrarily small constant $\delta > 0$, $\epsilon_0 \geq 2\lambda^2$ and $n\epsilon_0^2/\nu \geq K \log n$ for sufficiently large constant K, then

(a) w.h.p. two-sample voting is completed in $O(\log n)$ rounds and the winner is the opinion with the larger initial degree;

(b) if $\lambda = o(1)$ and $n\lambda^{\xi}/\nu \geq K \log n$, for arbitrarily small constant $\xi > 0$ and sufficiently large constant K, then w.h.p. two-sample voting is completed in $O(\log 1/\epsilon_0) + O(\log\log(1/\lambda)) + O(\log_{1/\lambda} n)$ rounds and the winner is the opinion with the larger initial degree.

Examples of graphs with $\lambda < 1/\sqrt{2}$ include random d-regular graphs with $d \geq 10$. The analysis of two-sample-voting on such graphs given in [9] required the initial imbalance between the opinions $\epsilon_0 \geq K\lambda$, for a large constant K, while the above theorem requires a weaker bound $\epsilon_0 \geq \max\{2\lambda^2, (K \log n)/n\}$ (as $\nu = 1$ for regular graphs). Examples of graphs with $\lambda(P) = o(1)$ include random d-regular graphs $d \to \infty$, pseudo-regular graphs of high degree, random graphs $G(n, p)$ when $np = \Omega(\log n)$, and Chung-Lu random graphs [4] satisfying certain conditions on minimum, average and maximum degree. The Chung-Lu graphs include many classes of inhomogeneous random graphs with wide variation in vertex degree. A more complete description of these classes of graphs, and proofs or descriptions of the results are given in Section 4.

Theorem 2. (Best-of-two voting) *Let $G = (V, E)$ be a connected non-bipartite graph, let P be the transition matrix of a reversible random walk on G with stationary distribution π, let $\boldsymbol{w} = (w(e), e \in E)$ be positive edge weights associated with P, and let $\nu = \nu(\boldsymbol{w}) = n\left(\sum_{v\in V} w^2(v)\right)/w^2(G)$. Let A and B denote the sets of vertices of G with initial opinions of the two types, and let $\epsilon_0 = |w(A) - w(B)|/w(G)$ denote the initial weight imbalance between these sets.*

Provided $\lambda = \lambda(P) \le 1/\sqrt{2} - \delta$ for an arbitrarily small constant δ, $\epsilon_0 \ge 2\lambda^2$ and $n\epsilon_0^2/\nu \ge K \log n$ for a sufficiently large constant K, then

(a) w.h.p. best-of-two voting is completed in $O(\log n)$ rounds and the winner is the opinion with the larger initial weight;

(b) if $\lambda = o(1)$ and $n\lambda^{\xi}/\nu \ge K \log n$ for arbitrarily small constant $\xi > 0$ and sufficiently large constant K, then w.h.p. best-of-two voting is completed in $O(\log 1/\epsilon_0) + O(\log\log(1/\lambda)) + O(\log_{1/\lambda} n)$ rounds and the winner is the opinion with the larger initial weight.

Regarding the conditions of Theorems 1 and 2, we need the lower bound on the initial imbalance of the opinions $\epsilon_0 \ge 2\lambda^2$ to show that in each step the majority opinion is expected to increase. We need the additional bound $\epsilon_0 \ge \sqrt{(K\nu \log n)/n}$ (and the condition $n\lambda^{\xi}/\nu \ge K \log n$ for the part (b) of the theorems) to argue that this increase happens w.h.p.

The advantage of best-of-two voting is that by choosing neighbours in the voting process based on assigning suitable weights to the edges we can tailor the outcome to our needs. In the simplest case that all edges are weighted equally we have the ordinary two-sample voting. The set with the largest initial degree wins w.h.p. The weights $w(u, v) = d(u) + d(v)$ biass voting towards the opinions of high degree vertices. The weights $w(u, v) = \max\{1/d(u), 1/d(v)\}$ biass voting towards the opinions of low degree vertices. To completely remove the effect of vertex degree on the voting process, we can use the following Metropolis process. Let $M = \max_{v\in V} d(v)$ be the maximum degree of G. Let each edge of G have weight one, and each vertex v introduce a self-loop of weight $M - d(v)$. Then $\pi(v) = 1/n$, so $w(A)/w(B) = \pi(A)/\pi(B) = |A|/|B|$ and the majority wins.

Theorems 1 and 2 both require the upper bound $1/\sqrt{2}$ on λ and the lower bound $2\lambda^2$ on the initial imbalance of the two opinions. We introduce the *k-extended best-of-two voting*, which can deal with cases when one or both of these conditions are not satisfied. This voting is a synchronous process in which during each round, every vertex v performs k steps of two independent weighted random walks starting at v. If the two vertices visited at step k of these two random walks have the same opinion, then vertex v adopts such opinion. This voting can be viewed as the best-of-two voting which uses P^k as the matrix of the sampling probabilities, where P is the transition matrix of the weighted random walk. Since P^k is reversible and $\lambda(P^k) = (\lambda(P))^k$, Theorem 2 implies the following corollary. Note that one round of the *k-extended best-of-two voting* involves k random-walk steps. This is the price to pay, if λ is poor and/or the initial imbalance of the two opinions is small.

Corollary 1. (Extended best-of-two voting) *Assume the same conditions as in Theorem 2 but $\lambda = \lambda(P) \geq 1/\sqrt{2}$ or $\epsilon_0 < 2\lambda^2$. Let an integer $k \geq 1$ be such that $\lambda^k < 1/\sqrt{2} - \delta$ for an arbitrarily small constant δ and $\epsilon_0 \geq 2\lambda^2$. Then*

(a) w.h.p. k-extended best-of-two voting is completed in $O(\log n)$ rounds and the winner is the opinion with the larger edge weight;
(b) if $\lambda^k = o(1)$ and $n\lambda^{\xi k}/\nu \geq K \log n$, for arbitrarily small constant $\xi > 0$ and sufficiently large constant K, then with high probability k-extended best-of-two voting is completed in $O(\log 1/\epsilon_0) + O(\log\log(1/\lambda)) + O(\log_{1/\lambda} n)$ rounds and the winner is the opinion with the larger weight.

An example where Corollary 1 can be applied is preferential attachment graphs generated by a scale-free process model in which each new vertex attaches d edges to the existing graph. The endpoints of the edges are chosen proportional to their current degree. For large d, $k = 7$ steps of random walks are enough for the corollary to hold. The details are given in Section 4.

2 Expected change in weight after one step of voting

In this section we derive a lower bound on the expected increase in the weight of the larger of the two sets A and B after one step of the voting process. The bound is very general and requires only the following two assumptions. (i) Each vertex v makes two choices at each step, and the choices are made independently among the vertices u of the graph with a fixed probability $P(v, u)$. (ii) The matrix P of probabilities $P(v, u)$ is the transition matrix of an irreducible aperiodic reversible random walk, and thus has a unique stationary distribution $\pi = (\pi(v), v \in V)$. We assume that there are always weights associated with the edges of the underlying graph, as explained in the previous section.

As an example of our approach, consider the transition matrix of a simple random walk. To make a transition from vertex v, the walk chooses a random neigbour $u \in N(v)$ with probability $P(v, u) = 1/d(v)$. Using this transition matrix P in the voting process corresponds to v choosing two neighbours uniformly at random with replacement. If $v \in A$ and the chosen neighbours are in B, then v changes its opinion to B. The degree $d(B)$ of B and the stationary probability (in the context of random walks) of B thus increase by $d(v)$ and $\pi(v) = d(v)/2m$, respectively.

Scaling the edge weights does not change the matrix P (hence does not change the random walk or the voting processes), so in our analysis we can use either the weights of sets $w(S)$ or the "normalised weights" $\pi(S)$, whichever is more convenient. Bearing this in mind, let for $x \in A$,

$$X_x^B = \begin{cases} \pi(x), & \text{if } x \text{ chooses twice in } B, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Thus X_x^B is the contribution of the vertex $x \in A$ to the increase of the (normalised) weight of B at the end of the step. Similarly, for $x \in B$ define $X_x^A = \pi(x)$, if x chooses twice in A, and zero otherwise. Adopting the notation $P(x, B) = \sum_{y \in B} P(x, y)$, we have for $x \in A$,

$$\mathbf{E}(X_x^B) \;=\; \pi(x)\mathbf{Pr}(X_x^B = \pi(x)) \;=\; \pi(x)\left(\sum_{y \in B} P(x, y)\right)^2 \;=\; \pi(x)(P(x, B))^2.$$

Let $X_A^B = \sum_{x \in A} X_x^B$, and let $R(A, B) = \mathbf{E}(X_A^B)$ be the expected increase of the weight of B in the current step (which is equal to the expected decrease of the weight of A) due to vertices moving from A to B. Then

$$R(A, B) = \mathbf{E}(X_A^B) = \sum_{x \in A} \mathbf{E}(X_x^B) = \sum_{x \in A} \pi(x)(P(x, B))^2. \tag{3}$$

Similarly, $R(B, A) = \sum_{x \in B} \pi(x)(P(x, A))^2$ is the expected increase of the weight of A due to vertices moving from B to A. If P is the transition matrix of simple random walk on G, then (3) can be written as

$$R(A, B) = \sum_{x \in A} \frac{d(x)}{2m}\left(\sum_{y \in B \cap N(x)} \frac{1}{d(x)}\right)^2 = \frac{1}{2m}\sum_{x \in A} \frac{(d^B(x))^2}{d(x)},$$

where $d^S(x) = |N(x) \cap S|$.

The next theorem and its corollary are the fundamental observations of this paper. They give lower bounds on $R(B, A) - R(A, B)$, which is the expected increase of the weight of set A in the current step. We use the notation $Q(A, B)$, which can be viewed as the normalised weight of the cut between A and B:

$$Q(A, B) = \sum_{x \in A}\sum_{y \in B} \pi(x)P(x, y) \;=\; \sum_{x \in A} \pi(x)P(x, B). \tag{4}$$

Note that for a reversible matrix P, $\pi(x)P(x, y) = \pi(y)P(y, x)$ implies $Q(A, B) = Q(B, A)$, and from the point of view of edge weights,

$$Q(A, B) = \sum_{x \in A}\sum_{y \in B \cap N(x)} w(x, y)/w(G) = \frac{w(A, B)}{w(G)}.$$

The proof of Theorem 3 refers to the inner product $\langle f, g\rangle_\pi$ of two vectors f, g of length n, defined by

$$\langle f, g\rangle_\pi = \sum_{x \in V} \pi(x)f(x)g(x).$$

Let $f_1, f_2, \ldots, f_n$ be (right) eigenvectors of P associated with the eigenvalues $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n > -1$. As we suppose P is reversible, we can assume that the eigenvectors $\{f_j\}_{j=1}^n$ are orthonormal with respect to the inner product $\langle\cdot,\cdot\rangle_\pi$ (see [17], Lemma 12.2); in particular, $f_1 = \mathbf{1}$. Thus $\langle f_i, f_j\rangle_\pi = 0$, if $i \neq j$, $\langle f_i, f_i\rangle_\pi = 1$ and for any $h \in \mathbb{R}^n$,

$$h = \sum_{j=1}^{n} \langle h, f_j\rangle_\pi f_j, \quad \text{and} \quad P^t h = \sum_{j=1}^{n} \lambda_j^t \langle h, f_j\rangle_\pi f_j. \tag{5}$$

Theorem 3. *Let P be a reversible transition matrix on G with stationary distribution π, $A \subseteq V$, $B = V \setminus A$ and let $\phi = Q(A,B)/\pi(B)$. Then*

$$R(B,A) - R(A,B) \geq \pi(B)\left((1-\lambda^2)\pi(A) - 2\phi(1-\phi)\right). \tag{6}$$

Since $0 < \phi < 1$, Theorem 3 gives immediately the following corollary.

Corollary 2. *Let P be a reversible transition matrix on G with stationary distribution π, $A \subseteq V$ and $B = V \setminus A$. Then*

$$R(B,A) - R(A,B) \ \geq \ \pi(B)\left((1-\lambda^2)\pi(A) - 1/2\right). \tag{7}$$

Proof of Theorem 3. Let

$$g(x) = \begin{cases} \pi(A), & \text{if } x \in B, \\ -\pi(B), & \text{if } x \in A; \end{cases} \tag{8}$$

The x-coordinate of the vector Pg is equal to

$$\begin{aligned} (Pg)(x) = P(x,.)\cdot g \ &= \sum_{y\in V} P(x,y)g(y) = \sum_{y\in A} P(x,y)g(y) + \sum_{y\in B} P(x,y)g(y) \\ &= \sum_{y\in A} P(x,y)(-\pi(B)) + \sum_{y\in B} P(x,y)\pi(A) \\ &= -\pi(B)P(x,A) + \pi(A)P(x,B) \\ &= \pi(A) - P(x,A) = P(x,B) - \pi(B). \end{aligned} \tag{9}$$

Using (9) in (10) and (3) and (4) in (11), we have

$$\begin{aligned} \langle Pg, Pg\rangle_\pi &= \sum_{x\in V} \pi(x)((Pg)(x))^2 = \sum_{x\in A} \pi(x)((Pg)(x))^2 + \sum_{x\in B} \pi(x)((Pg)(x))^2 \\ &= \sum_{x\in A} \pi(x)\Big(P(x,B) - \pi(B)\Big)^2 + \sum_{x\in B} \pi(x)\Big(\pi(A) - P(x,A)\Big)^2 \qquad (10) \\ &= \sum_{x\in A} \pi(x)P(x,B)^2 + \sum_{x\in A} \pi(x)\pi(B)^2 + \sum_{x\in A} \pi(x)\Big(-2P(x,B)\pi(B)\Big) \\ &\quad + \sum_{x\in B} \pi(x)P(x,A)^2 + \sum_{x\in B} \pi(x)\pi(A)^2 + \sum_{x\in B} \pi(x)\Big(-2P(x,A)\pi(A)\Big) \\ &= R(A,B) + R(B,A) + \pi(A)\pi(B)\Big(\pi(B) + \pi(A)\Big) \\ &\quad - 2Q(B,A)\Big(\pi(B) + \pi(A)\Big) \qquad (11) \\ &= R(A,B) + R(B,A) + \pi(A)\pi(B) - 2Q(B,A). \qquad (12) \end{aligned}$$

Equation (12) is equivalent to:

$$R(B,A) - R(A,B) = \pi(A)\pi(B) - \langle Pg, Pg\rangle_\pi - 2\Big(Q(B,A) - R(B,A)\Big). \quad (13)$$

We find that

$$\begin{aligned} Q(B,A) - R(B,A) = \\ &= \sum_{x\in B} \pi(x)P(x,A) - \sum_{x\in B} \pi(x)\left(P(x,A)\right)^2 \\ &= \pi(B)\sum_{x\in B}\frac{\pi(x)}{\pi(B)}P(x,A)(1-P(x,A)) \\ &\le \pi(B)\left(\sum_{x\in B}\frac{\pi(x)}{\pi(B)}P(x,A)\right)\left(1-\sum_{x\in B}\frac{\pi(x)}{\pi(B)}P(x,A)\right) \quad (14) \\ &= \pi(B)\frac{Q(B,A)}{\pi(B)}\left(1-\frac{Q(B,A)}{\pi(B}\right) \;=\; \pi(B)\phi(1-\phi), \quad (15) \end{aligned}$$

where (14) follows from the fact that the function $z(z-1)$ is concave. The claimed bound (6) follows from (13), (15) and the following result:

$$\langle Pg, Pg\rangle_\pi \le \lambda^2\pi(A)\pi(B). \quad (16)$$

To verify (16), check first that $\langle Pg, Pg\rangle_\pi = \langle P^2g, g\rangle_\pi$, using reversibility of P, and $\langle g, g\rangle_\pi = \pi(A)\pi(B)$, using the definition of g. Then using (5) and $\langle g, f_1\rangle_\pi = 0$ (since $f_1 = \mathbf{1}$), derive $\langle g, g\rangle_\pi = \sum_{j=2}^n \langle g, f_j\rangle_\pi^2$ and finally

$$\langle P^2g, g\rangle_\pi \;=\; \sum_{j=2}^n \lambda_j^2\langle g, f_j\rangle_\pi^2 \;\le\; \lambda^2\sum_{j=2}^n\langle g, f_j\rangle_\pi^2 \;=\; \lambda^2\langle g, g\rangle_\pi \;=\; \lambda^2\pi(A)\pi(B)$$

. □

The following known result generalizes the Expander Mixing Lemma for undirected graphs to weighted graphs. While bound (7) given in Corollary 2 will be sufficient in the proofs of part (a) of Theorems 1 and 2, the tighter bound (6) given in Theorem 3 together with Lemma 1 will be needed to prove part (b).

Lemma 1. *Let P be the transition matrix of the weighted random walk on a connected undirected graph $G = (V, E)$ with edge weights $\boldsymbol{w} = (w(e), e \in E)$, and let $\lambda = \max(|\lambda_2(P)|, |\lambda_n(P)|)$. Then*

$$|w(A,B) - w(A)w(B)/w(G)| \le \lambda\, w(A)w(B)/w(G). \quad (17)$$

3 Proof of Theorem 1

In this section we give the proof of Theorem 1. The proof of Theorem 2 is very similar. Assume $\lambda^2 \le 1/2 - \delta$ for small constant $\delta > 0$, $2\lambda^2 \le \epsilon_0 < 1$

and $\epsilon_0^2 \geq (K\nu \log n)/n$, for some large constant K. We first prove the part (a) of Theorem 1. We assume that B is the minority set with the initial degree $d(B) = m(1-\epsilon_0)$. The proof is in two phases. Phase I reduces $d(B)$ to cm in $T_{\rm I} = O(\log 1/\epsilon_0)$ steps, w.h.p., where $c > 0$ is an arbitrarily small constant. Then Phase II reduces $d(B)$ to zero in $T_{\rm II} = O(\log n)$ steps, w.h.p.

Proof of Theorem 1(a), Phase I. Let $\varDelta_{AB}$ be the increase in degree of the vertices of A at a given step of the voting. Then by Corollary 2,

$$\mathbf{E}\varDelta_{AB} = 2m(R(B,A) - R(A,B)) \geq d(B)\left((1-\lambda^2)d(A)/(2m) - 1/2\right). \quad (18)$$

Let $\epsilon = (d(A)-d(B))/2m$. Thus $d(A) = m(1+\epsilon)$ and $d(B) = m(1-\epsilon) > cm$. We assume $\epsilon \geq \epsilon_0$ (by induction, the imbalance ϵ increases in each step in Phase I w.h.p.). Thus $\epsilon \geq 2\lambda^2$, which together with $\delta \leq 1/2 - \lambda^2$ gives

$$\mathbf{E}\varDelta_{AB} \;\geq\; \frac{d(B)}{2}\left((1-\lambda^2)(1+\epsilon) - 1\right) \;=\; \frac{d(B)}{2}(\epsilon - \lambda^2(1+\epsilon)) \;\geq\; d(B)\frac{\epsilon\delta}{2}. \quad (19)$$

The following version of the Hoeffding Lemma can be found in e.g. [18]. Let $X_k,\ k = 1, ..., N$ be independent random variables, where for each k, $a_k \leq X_k \leq b_k$. Let $X = \sum_{k=1}^N X_k$ and let $\mu = \mathbf{E}X$. Then for any $t > 0$

$$\mathbf{Pr}(|X-\mu| \geq Nt) \leq 2\exp\left(-2N^2t^2 / \sum_{k=1}^{N}(b_k - a_k)^2\right). \quad (20)$$

Let C be the vertices which have a neighbour in the other vote set, that is, the wertices which have positive probability of changing their vote. Let $A_C = C \cap A$ and $B_C = C \cap B$. We use (20) with $N = |C|$ and take X_v, for $v \in C$, as the signed degree of v based on (2). For $v \in A_C$, $X_v = -X_v^B \cdot 2m$, which is either $-d^B(v)$ or 0, and for $v \in B_C$, $X_v = X_v^A \cdot 2m$, which is either $d^A(v)$ or 0. Thus $\sum_{v\in C} X_v = \varDelta_{AB}$ and the sum $\sum_{v\in C}(b_v - a_v)^2$ in (20) is

$$\sum_{v\in C}(b_v - a_v)^2 = \sum_{v\in A_C}(d^B(v))^2 + \sum_{v\in B_C}(d^A(v))^2 \leq \sum_{v\in V} d^2(v) = (2m)^2\nu/n.$$

From (19), (20), $d(B) \geq cm$ and $n\epsilon^2/\nu \geq K\log n$, we find

$$\begin{aligned}\mathbf{Pr}(\varDelta_{AB} \leq \mathbf{E}\varDelta_{AB}/2) \leq\; & \mathbf{Pr}(|\varDelta_{AB} - \mathbf{E}\varDelta_{AB}| \geq \mathbf{E}\varDelta_{AB}/2)\\ \leq\; & 2\exp\left(\frac{-2(\mathbf{E}\varDelta_{AB}/2)^2}{(2m)^2\nu/n}\right) \;\leq\; 2\exp\left(-\frac{n(d(B)\epsilon\delta)^2}{32m^2\nu}\right)\\ \leq\; & 2\exp\left(-\frac{n\epsilon^2}{\nu}\,\frac{(c\delta)^2}{32}\right) \;\leq\; \frac{1}{n^\alpha},\end{aligned} \quad (21)$$

for constant $\alpha = K(c\delta)^2/32$.

Let B and B' be the set of vertices with the B vote at the beginning of the current and next step, respectively. If $\varDelta_{AB} \geq \mathbf{E}\varDelta_{AB}/2$, then it follows from (19) that the size of $d(B')$ is

$$d(B') = d(B) - \varDelta_{AB} \leq d(B) - \mathbf{E}\varDelta_{AB}/2 \leq d(B)(1-\epsilon\delta/4). \quad (22)$$

Suppose firstly that $\epsilon \le 1/2$, then in one step $d(B)$ decreases w.h.p. from $m(1-\epsilon)$ to at most $m(1-\epsilon(1+\delta/8))$. Starting from $d(B_0) = m(1-\epsilon_0)$, after j steps we have that $d(B_j) \le m(1-(1+\delta/4)^j \epsilon_0)$. On the other hand, if $\epsilon > 1/2$, that is, $d(B) \le m/2$, then (22) implies that $d(B)$ reduces to size cm in a constant number of steps. Thus after $T_{\mathrm{I}} = O(\log 1/\epsilon_0)$ steps, w.h.p. $d(B_{T_{\mathrm{I}}}) \le cm$.

Proof of Theorem 1(a), Phase II. Let B and B' denote the set of vertices with the B vote at the beginning of the current and the next step, respectively. At the end of Phase I, $d(B) \le cm$, so $\pi(B) \le c/2$ for some small constant $c > 0$. Firstly, using (21), we observe that $d(B)$ remains below cm w.h.p. for polylogarithmic number of steps:

$$\mathbf{Pr}(d(B') \ge cm \,|\, d(B) \le cm) \le \mathbf{Pr}(d(B') \ge cm \,|\, d(B) = cm) \le \frac{1}{n^\alpha}. \tag{23}$$

Using (19) (which, as (18), applies to A and $B = V \setminus A$ of any sizes) and noting that $d(B) \le cm$ implies $\epsilon \ge 1 - c \ge 2/3$, we have for any $0 \le q \le cm$,

$$\mathbf{E}(d(B')|d(B) = q) \;\le\; (1-\delta/3)q. \tag{24}$$

Let B_0 be the B-set at the beginning of Phase II and let B_i be the B-set after i steps. We assume that $d(B_0) \le cm$, and generally $0 \le d(B_i) \le 2m$, for each $i \ge 1$. We now bound $\mathbf{E}(d(B_i))$, for $i \ge 1$. Denoting $\mathcal{B}_i \equiv \{d(B_i) \le cm\}$, for $i \ge 0$, we have

$$\mathbf{E}(d(B_i)) \le \; \mathbf{E}(d(B_i)|\mathcal{B}_{i-1}) \cdot \mathbf{Pr}(\mathcal{B}_{i-1}) \;+\; (2m) \cdot \mathbf{Pr}(\neg\mathcal{B}_{i-1}). \tag{25}$$

Further,

$$\begin{aligned}
&\mathbf{E}(d(B_i)|\mathcal{B}_{i-1}) \cdot \mathbf{Pr}(\mathcal{B}_{i-1}) \\
&\quad = \sum_{0\le q\le cm} \mathbf{E}(d(B_i)|d(B_{i-1}) = q) \cdot \mathbf{Pr}(d(B_{i-1}) = q|\mathcal{B}_{i-1}) \cdot \mathbf{Pr}(\mathcal{B}_{i-1}) \\
&\quad = \sum_{0\le q\le cm} \mathbf{E}(d(B_i)|d(B_{i-1}) = q) \cdot \mathbf{Pr}(d(B_{i-1}) = q) \\
&\quad \le \sum_{0\le q\le cm} (1-\delta/3) \cdot q \cdot \mathbf{Pr}(d(B_{i-1}) = q) \;\;\le\;\; (1-\delta/3)\, \mathbf{E}(d(B_{i-1}),
\end{aligned} \tag{26}$$

and, using (23),

$$\mathbf{Pr}(\neg\mathcal{B}_{i-1}) \;\le\; \sum_{j=1}^{i-1} \mathbf{Pr}(\mathcal{B}_{j-1} \text{ and } \neg\mathcal{B}_j) \;\le\; \sum_{j=1}^{i-1} \mathbf{Pr}(\neg\mathcal{B}_j \,|\, \mathcal{B}_{j-1}) \;\le\; \frac{i}{n^\alpha}. \tag{27}$$

Putting (26) and (27) in (25), we get

$$\begin{aligned}
\mathbf{E}(d(B_i)) &\le\; (1-\delta/3)\, \mathbf{E}(d(B_{i-1}) + 2mi/n^\alpha, \quad \text{and} \\
\mathbf{E}(d(B_i)) &\le\; (1-\delta/3)^i\, d(B_0) + (3/\delta)2mi/n^\alpha.
\end{aligned}$$

Thus for $T = T_{II} = (3/\delta)(2+\alpha)\ln n$,

$$\mathbf{E}(d(B_T)) \leq \; (1-\delta/3)^T\, cm + (3/\delta) 2mT/n^{\alpha} \;\leq\; 1/n^{\alpha/2},$$

so

$$\mathbf{Pr}(d(B_T)=0) \;=\; 1-\mathbf{Pr}(d(B_T)\geq 1) \;\geq\; 1-\mathbf{E}(d(B_T)) \;\geq\; 1-n^{-\alpha/2}.$$

This means that w.h.p. Phase II completes in $T = T_{II} = O(\log n)$ steps and the winner is vote A.

Proof of Theorem 1(b). For a simple random walk, all edges have weight one, so in Lemma 1, $w(A)$ and $w(B)$ are $d(A)$ and $d(B)$, $w(G) = 2m$ and $w(A,B) = d(A,B)$, the number of edges between sets A and B. Thus (17) gives the following inequality for any sets A and $B = V \setminus A$.

$$\left| \frac{d(A,B)}{d(B)} - \frac{d(A)}{2m} \right| \leq \lambda \frac{d(A)}{2m}. \tag{28}$$

In Theorem 3, $\phi = Q(A,B)/\pi(B) = d(A,B)/d(B)$ and $\pi(A) = d(A)/(2m)$, so (28) implies that $\pi(A)(1-\lambda) \leq \phi \leq \pi(A)(1+\lambda)$. For this range of ϕ, if $\pi(A) \geq 1/2$, then $\phi(1-\phi)$ in (6) is maximised at $\phi = \pi(A)(1-\lambda)$, so (6) implies

$$R(B,A) - R(A,B) \geq \pi(B)\pi(A)(1-\lambda)^2(1-2\pi(B)).$$

Hence after one step, the set B is replaced by a set B' of expected degree

$$\begin{aligned}\mathbf{E}(d(B') \mid d(B)) \;&=\; d(B) - 2m(R(B,A)-R(A,B)) \\ &\leq d(B)\left(1-(1-\pi(B))(1-\lambda)^2(1-2\pi(B))\right) \\ &\leq\; d(B)(2\lambda + 3\pi(B)). \end{aligned} \tag{29}$$

In the analysis of Phase II, we use now the bound (29) on $\mathbf{E}(d(B')|d(B))$ instead of the bound (24). We split Phase II into two parts. First $d(B)$ keeps decreasing from cm to $\lambda^{\xi/4}m$. For this range of $d(B)$, $\pi(B) \geq \lambda$, so (29) implies that $\mathbf{E}(\pi(B')) \leq 5(\pi(B))^2$. If $\pi(B') \geq \mathbf{E}(\pi(B')) + (\pi(B))^2$, then $\varDelta_{AB} \leq \mathbf{E}(\varDelta_{AB}) - 2m(\pi(B))^2$, so we have, in a similarly way as in (21) and using $\pi(B) \geq \lambda^{\xi/4}/2$ and the assumption that $n\lambda^{\xi}/\nu \geq K\log n$,

$$\begin{aligned}\mathbf{Pr}\left(\pi(B') \geq 6(\pi(B))^2\right) &\leq \mathbf{Pr}\left(|\varDelta_{AB} - \mathbf{E}\varDelta_{AB}| \geq 2m(\pi(B))^2\right) \\ &\leq 2\exp\left(-2(2m(\pi(B))^2)^2/((2m)^2\nu/n)\right) \\ &\leq 2\exp\left(-\frac{n\lambda^{\xi}}{8\nu}\right) \leq \frac{1}{n^{K/8}}.\end{aligned}$$

Thus w.h.p. in each step of the first part of Phase II, $\pi(B') \geq 6(\pi(B))^2$, giving $O(\log\log(1/\lambda))$ steps. Then $d(B)$ decreases from $\lambda^{\xi/4}m$ to zero and for this part of Phase II, (29) implies that $d(B') \leq 5\lambda^{\xi/4}d(B)$, leading to the $O(\log_{1/\lambda} n)$ bound on the number of rounds.

4 Specific examples and notes on eigenvalue gaps

We give various examples of graphs which satisfy our theorems. In some cases, additional work, not discussed here, is required to relate known results to the second eigenvalue $\lambda(P)$ of the transition matrix P.

Random graphs $G(n,p)$**.** From Coja-Oghlan [6], Theorem 1.2, if $2(1+o(1))\log n \le np \le 0.99n$, then w.h.p. $\max_{j\ge 2}|\lambda_j(P)| \le (1+o(1))\frac{2}{\sqrt{np}}$.

Chung-Lu model. This model generalizes random graphs $G(n,p)$ to the space of random graphs $G(\boldsymbol{w})$ where $\boldsymbol{w}$ is a sequence of positive weights $\boldsymbol{w} = (w_1, w_2, ..., w_n)$. Edges are included independently, and edge $\{i,j\}$ has probability $p_{ij} = w_i w_j/\rho$ where $\rho = \sum_i w_i$. There is a further constraint that $\max_i w_i^2 < \rho$ to ensure $p_{ij} \le 1$. The average degree is $\overline{w} = \sum_{i=1}^n w_i/n = \rho/n$. The expected degree of vertex i is w_i, and the minimum expected degree $w_{\min} = \min_i w_i$.

The following result is from [5], where ω is any slowly growing function.

$$\max_{j\ge 2}|\lambda_j(P)| \le (1+o(1))\frac{4}{\sqrt{\overline{w}}} + \frac{\omega \log^2 n}{w_{\min}}.$$

Thus provided $w_{\min} \gg \omega \log^2 \sqrt{\overline{w}}$, the generated graphs have small $\lambda(P)$.

Pseudo-regular graphs. Take a random d-regular graph G and add extra edges, at most c at any vertex, where $c \le \epsilon d$ for some small constant ϵ. This gives $\lambda(P) \le (3\sqrt{d} + 2c)/(d+c)$.

Metropolis walks. Let G have degree bounded between d and $M = (1+a)d$. The transition matrix $\widetilde{P}$ of the Metropolis process has transition probabilities $\widetilde{P}_{ij} = 1/M$ if $\{i,j\}$ is an edge of G and loop probability $\widetilde{P}_{ii} = 1 - d(i)/M$. If P is the transition matrix of a simple random walk on G, then $|\lambda_k(\widetilde{P}) - \lambda_k(P)| \le 2a/(1+a)$.

Preferential Attachment model. The model $G_{m,t}$ generates a preferential attachment graph as follows. At any step $t \ge 1$ a new vertex v_t with m edges is attached to the existing graph $G_{m,t-1}$. The edges from v_t are attached to existing vertices chosen with probability proportional to their degree. The following result is given in [19]. For any $m \ge 2$, if positive constants a and c satisfy $c < 2(m-1) - 4a - 1$, then the conductance Φ of $G_{m,n}$ satisfies

$$\mathbf{Pr}(\Phi \le a/(m+a)) = o(n^{-c}).$$

Taking constants a and c such that $2c = 2(m-1) - 4a - 1$, we have w.h.p.

$$\Phi \ge \frac{2m-3-2c}{6m-3-2c}.$$

Choosing c small and using the relationship that $\lambda_2 \le 1 - \Phi^2/2$, we have that $\lambda_2(m)$ satisfies $\lambda_2(2) \le 199/200$ and for m large $\lambda_2(m) < 19/20$. In both cases, two-sample voting cannot provably guarantee the outcome. If we use k-extended best-of-two voting algorithm with $k = 70$ for the first case and $k = 7$ for the second, then we obtain $\lambda^k < 1/\sqrt{2}$ and thus Corollary 1 applies.

References

1. M. Abdullah and M. Draief. Consensus on the Initial Global Majority by Local Majority Polling for a Class of Sparse Graphs. (2013) `www.arXiv.org`
2. M. Bolla. Beyond the Expanders. International Journal of Combinatorics, vol. 2011, Article ID 787596, 11 pages (2011).
3. S. Brahma, S. Macharla, S. P. Pal, S. R. Singh. Fair Leader Election by Randomized Voting. In *ICDCIT 2004*, pages 22–31, (2004).
4. F. R. K. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. Annals of Combinatorics 6, pages 125–145 (2002).
5. F. R. K. Chung, L. Lu, and V. Vu. The spectra of random graphs with given expected degrees. Internet Mathematics 1, pages 257–275, (2003).
6. A. Coja-Oghlan. On the Laplacian Eigenvalues of $G_{n,p}$. Combinatorics, Probability and Computing, 16, pages 923–946 (2007).
7. C. Cooper, A. Frieze, B. Radzik. Multiple Random Walks in Random Regular Graphs. SIAM J. on Discrete Math. 23, pages 1738–1761, (2009).
8. C. Cooper, R. Elsässer, H. Ono, T. Radzik. Coalescing Random Walks and Voting on Connected Graphs. SIAM J. on Discrete Math. 27 pages 1748–1758, (2013).
9. C. Cooper, R. Elsässer and T. Radzik. The power of two choices in distributed voting. ICALP 2014. pages 435–446, (2014).
10. J. Cruise and A. Ganesh, Probabilistic consensus via polling and majority rules. Queueing Systems: Theory and Applications. 78, pages 99–120, (2014).
11. X. Deng and C. Papadimitriou. On the Complexity of Cooperative Solution Concepts. Mathematics of Operations Research 19, pages 257–266, (1994).
12. 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).
13. D. Gifford. Weighted Voting for Replicated Data. In *SOSP 1979: Proceedings of the 7th ACM Symposium on Operating Systems Principles*, pages 150–162, (1979).
14. Y. Hassin and D. Peleg. Distributed probabilistic polling and applications to proportionate agreement. Information & Computation, 171, pages 248–268, (2001).
15. R. A. Horn and C. R. Johnson. *Matrix Analysis.* CUP 2006.
16. B. Johnson, *Design and Analysis of Fault Tolerant Digital Systems*, Addison-Wesley, (1989).
17. D. A. Levine, Y. Peres, and E. L. Wilmer. *Markov Chain and Mixing Times.* American Mathematical Society, 2008.
18. C. McDiarmid. On the method of bounded differences. Pages 148–188 of *Surveys in Combinatorics, 1989*, Editor J. Siemons. CUP 1989.
19. M. Mihail, C. Papadimitriou, A. Sabieri. On Certain Connectivity Properties of the Internet Topology. Proceedings of Foundations of Computer Science, FOCS 2003, pages 28–35 (2003).
20. T. Nakata, H. Imahayashi, M. Yamashita. Probabilistic local majority voting for the agreement problem on finite graphs. In *COCOON 1999*, pages 330–338, (1999).