Discrete Incremental Voting

Colin Cooper
Department of Informatics, King's College, University of London, UK

Tomasz Radzik
Department of Informatics, King's College, University of London, UK

Takeharu Shiraga
Department of Information and System Engineering, Faculty of Science and Engineering, Chuo University, Tokyo, Japan

Abstract

We consider a type of pull voting suitable for discrete numeric opinions which can be compared on a linear scale, for example, 1 ("disagree strongly"), 2 ("disagree"), ..., 5 ("agree strongly"). On observing the opinion of a random neighbour, a vertex changes its opinion incrementally towards the value of the neighbour's opinion, if different. For opinions drawn from a set $\{1, 2, \ldots, k\}$, the opinion of the vertex would change by $+1$ if the opinion of the neighbour is larger, or by -1, if it is smaller.

It is not clear how to predict the outcome of this process, but we observe that the total weight of the system, that is, the sum of the individual opinions of all vertices, is a martingale. This allows us analyse the outcome of the process on some classes of dense expanders such as complete graphs K_n and random graphs $G_{n,p}$ for suitably large p. If the average of the original opinions satisfies $i \le c \le i+1$ for some integer i, then the asymptotic probability that opinion i wins is $i+1-c$, and the probability that opinion $i+1$ wins is $c-i$. With high probability, the winning opinion cannot be other than i or $i+1$.

To contrast this, we show that for a path and opinions $0, 1, 2$ arranged initially in non-decreasing order along the path, the outcome is very different. Any of the opinions can win with constant probability, provided that each of the two extreme opinions 0 and 2 is initially supported by a constant fraction of vertices.

2012 ACM Subject Classification Theory of computation → Distributed algorithms

Keywords and phrases Random distributed processes, Pull voting

Digital Object Identifier 10.4230/LIPIcs.OPODIS.2023.10

Related Version A part of this paper appeared in the 2023 ACM Symposium on Principles of Distributed Computing (PODC 2023) as a brief announcement.
Full Version: https://arxiv.org/abs/2305.15632

Funding *Colin Cooper*: Research supported at the University of Hamburg, by a Mercator fellowship from DFG – Project 491453517.
Takeharu Shiraga: Research supported by JSPS KAKENHI Grant Number 23K16840.

1 Introduction

Background on distributed pull voting. Distributed voting has applications in various fields of computer science including consensus and leader election in large networks [7, 13]. Initially, each vertex has some value chosen from a set S, and the aim is that the vertices reach consensus on (converge to) the same value, which should, in some sense, reflect the initial distribution of the values. Voting algorithms are usually simple, fault-tolerant, and easy to implement [13, 14].

Pull voting is a simple form of distributed voting in connected graphs. At each step, a randomly chosen vertex (asynchronous process), or each vertex (synchronous process), replaces its opinion with that of randomly chosen neighbour. The probability a particular

27th International Conference on Principles of Distributed Systems (OPODIS 2023).
Editors: Alysson Bessani, Xavier Défago, Junya Nakamura, Koichi Wada, and Yukiko Yamauchi; Article No. 10; pp. 10:1–10:22
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

opinion, say opinion A, wins is $d(A)/2m$, where $d(A)$ is the sum of the degrees of the vertices initially holding opinion A, and m is the number of edges in the graph; see Hassin and Peleg [13] and Nakata *et al.* [16]. The pull voting process can be modified to consider two or more opinions at each step. The aim of this modification is twofold; to ensure the majority (or plurality) wins, and to speed up the run time of the process. Work on *best-of-k* models, where a vertex replaces its opinion with the opinion most represented in a sample of k opinions, includes [1, 4, 3, 5, 6, 8, 10, 12, 15, 18].

The general model of pull voting regards the opinions as incommensurate, and thus not comparable on a numeric scale. In contrast to this, Doerr at al. [11] consider opinions drawn from an ordered set and a process which aims to converge to the median. At each step a random vertex selects two neighbours and replaces its opinion by the median of all three values (including its own current value).

In this paper we consider another variant of pull voting, with opinions comparable on a linear scale. This variant – *discrete incremental voting* – can be seen as modelling the convergence to consensus of a group opinion, based on compromise during extended discussion. The final opinion may not be one held originally by anyone, but would reflect the compromise. If the initial (degree-weighted) average of the opinions is c, then the expectation of the final opinion (a random variable) is always equal to the initial average c. Furthermore, for some classes of expanders, w.h.p.[1] the process converges to an integer average, $\lfloor c \rfloor$ or $\lceil c \rceil$. Seen in this context, the pull voting processes above mirror the statistical measures of Mode, Median and Mean, for pull voting, median voting and discrete incremental voting, respectively.

Distributed processes for computing the exact average of the vertex values have been widely proposed and studied, but they would require calculating and storing fractional numbers and coordinated two-sided updates (both interacting vertices simultaneously update their states). The one-sided updates of pull-voting processes is an appealing simplicity.

Discrete incremental voting: An introduction. We assume the initial opinions of the vertices are chosen from among the integers $\{1, 2, ..., k\}$. As a simple example, suppose the entries reflect the views of the vertices about some issue, and range from 1 ('disagree strongly') to k ('agree strongly'). Then it seems unrealistic that a vertex would change its opinion to that of a neighbour, (as in pull voting), based only on observing what the neighbour thinks. However, people being what they are, it seems possible that they may modify their opinion slightly towards the opinion of their neighbour on observing it.

In the simplest case, suppose that a vertex v has opinion i and observes at its neighbour u opinion j. If $j > i$, then vertex v modifies its opinion to $i + 1$ (tends to agree more). Similarly, if the observed neighbour u has value $j < i$, vertex v changes its opinion to $i - 1$ (tends to disagree more). The neighbour u does not change its opinion at this interaction. That this process converges, and the value it converges to, is the topic of this paper.

We consider two related asynchronous and one synchronous variants of incremental voting. Given a connected graph $G = (V, E)$ with n vertices and m edges, let $X(t) = (X_v(t) : v \in V)$ be the vector of integer opinions at step t. The value of $X(t + 1)$ is obtained as follows.

Asynchronous vertex process: Given $X = X(t)$, pick a vertex v uniformly at random (u.a.r.) and an adjacent edge (v, w) u.a.r. The following update rule $X_v \to X'_v$ holds,

$$\left.\begin{array}{lll} X_v < X_w & \Longrightarrow & X'_v = X_v + 1 \\ X_v = X_w & \Longrightarrow & X'_v = X_v \\ X_v > X_w & \Longrightarrow & X'_v = X_v - 1 \end{array}\right\} \tag{1}$$

[1] With high probability, which in this paper means probability $1 - o(1)$.

Asynchronous edge process[2]: Pick a random endpoint v of a random edge $e = (v, w)$ selected u.a.r. The value X_v at vertex v is updated to X'_v using the rules in (1) above.

Synchronous vertex process: Given $X = X(t)$, each vertex v picks an adjacent edge (v, w) u.a.r. and updates its value $X_v \to X'_v$ according to (1).

Discrete incremental voting: Main results. If the initial set of opinions is $\{0, 1\}$ (or $\{i, i+1\}$ for an integer i) the incremental voting (with updates (1)) is equivalent to ordinary "two-value" pull voting as studied by [13] and others: when a vertex v updates its value using the value at a neighbour w, vertex v simply takes on the value from vertex w. The "two-value" pull voting comes in the same three variants: asynchronous or synchronous vertex process, or asynchronous edge process. In a vertex process (asynchronous or synchronous), the probability that opinion 0 wins, $d(0)/2m$, is proportional to the sum $d(0)$ of the degrees of the vertices initially holding this opinion. Discrete incremental voting generalises two-value pull voting, where the simplest case differing from pull voting is when opinion values are in $\{0, 1, 2\}$. In general we assume the initial values are in the range $\{0, 1, ..., k\}$ or $\{1, ..., k\}$, where appropriate bounds on k may be required in the analysis.

In order to reach a consensus opinion, all other opinions must be eliminated. The only way to irreversibly reduce the number of opinions, is to remove one of the extreme values in the order, leading to the next stage. The process continues through such stages until one opinion remains. Returning to our original example (given at the start of the paragraph "Discrete incremental voting: An introduction", with $k = 5$), the values have the following meanings: 1 ("disagree strongly"), 2 ("disagree"), 3 ("indifferent"), 4 ("agree"), 5 ("agree strongly"). Suppose we start with initial opinions $\{1, 2, 5\}$. Then a possible evolution of the system is

$$\{1, 2, 5\} \to \{1, 2, 3, 4\} \to \{2, 4\} \to \{2, 3\} \to \{3\},$$

where the sets of opinions at the beginning of each stage are indicated, and each "$\to$" represents a sequence of one or more steps constituting one stage. The intermediate values may disappear but then they appear again (in the above example, opinion 3 disappeared in stage 2 and they appeared again in stage 3). Eventually, as extreme values disappear, we reach the final stage of voting when only two adjacent values remain. In the example above the final stage has values $\{2, 3\}$. At this point the process reverts to ordinary two-value pull voting. Suppose only values $\{i, i+1\}$ remain. Let A_j, $j \in \{i, i+1\}$, be the set of vertices with value j at the start of this final stage, and $N_j = |A_j|$, so $N_i + N_{i+1} = n$. Let $d(A) = \sum_{v \in A} d(v)$ be the total degree of set A. The probability that i wins is

$$\mathbb{P}(i \text{ wins}) = \frac{N_i}{n} \quad \text{(Edge process)}, \qquad \mathbb{P}(i \text{ wins}) = \frac{d(A_i)}{2m} \quad \text{(Vertex process)}. \tag{2}$$

In Section 2, Lemma 4, we prove that the average weight of the process is a martingale in both the asynchronous and synchronous processes. (The vertex opinion, value, and weight refer to the same quantity.) This allows us to establish Theorem 5 which gives the distribution of winning opinions in the case where the initial average c is maintained throughout the process. In this case, when only two opinions $\{i, i+1\}$ remain, we have $i \le c \le i+1$ and their winning probability is determined as in (2) above.

As shown in Section 2, Lemma 6, the expected time for one of the two extreme opinions to disappear is $O(T_2)$, where T_2 is the (worst-case) expected time to consensus for two-value pull voting on the same graph. Thus the expected time to consensus is $O(kT_2)$. See [9] and

[2] The edge process is as a vertex process but with the leading vertex v sampled with prob. $\pi_v = d(v)/2m$.

[17] for graph specific bounds on the value of T_2, e.g., for the synchronous vertex model, $T_2 = O(n^3)$ for any connected graph, $T_2 = O(n^2)$ for regular graphs, and $T_2 = O(n)$ for regular expander graphs, with bounds for asynchronous models higher by an $O(n)$ factor. However, for the complete graph K_n and some classes of expanders[3], as the number of opinions k increases, the bound $O(kT_2)$ on consensus time becomes weak. For such graphs we show that with high probability the extreme values disappear faster than the time to complete two-value pull voting. Thus, with suitable bounds on k, the expected run time can be reduced from $O(kT_2)$ to $O(T_2)$, and is directly comparable with ordinary pull voting. See e.g., Lemma 10. Ideally (for easier analysis) we would like one of the two extreme opinions to disappear completely before moving on to considering the next extreme opinion. However, to obtain good bounds, in some cases we have to move on to, say, the next smallest opinion κ, while some small number of vertices may still hold opinions smaller than κ.

As the average opinion is a martingale (details of this are in the next Section 2), in cases where the process converges rapidly to two neighbouring states, martingale concentration allows us to use Theorem 5 to predict the outcome of the process. This is fundamental for our analysis on expander graphs. For the cases we studied, $G_{n,p}$ and K_n, essentially the process converges quickly to two neighbouring states $\{i, i+1\}$. Because the time to consensus in the final stage is determined by known results, i.e., two-value pull voting, we only need to estimate the time to reach a final pair of values $\{i, i+1\}$ where w.h.p. $i \leq c \leq i+1$. The overall expected time to consensus is determined by the (slower) final stage of two-value pull voting; namely $O(n)$ for the synchronous, and $O(n^2)$ for the asynchronous process.

We illustrate incremental voting using three examples: the asynchronous process on $G_{n,p}$ (Theorem 2) and the synchronous process on K_n (Theorem 1), both of which work as one might expect, and an asynchronous process on the path which does not (Theorem 3).

Notation. For functions $a = a(n)$ and $b = b(n)$, $a \sim b$ denotes $a = b(1 + o(1))$, where $o(1)$ is a function of n which tends to zero as $n \to \infty$. We use ω to denote a generic quantity tending to infinity as $n \to \infty$, but suitably slowly as required in the given proof context. An event A on an n-vertex graph holds with high probability (w.h.p.), if $\mathbb{P}(A) = 1 - o(1)$.

► **Theorem 1.** SYNCHRONOUS INCREMENTAL VOTING ON K_n.
Let the initial values of the vertices of K_n be chosen from $\{1, 2, \ldots, k\}$, where $k = o(n/(\log n)^2)$, and let $S(0) = \sum_{v\in V} X_v(0) = cn$.

- **(i)** *If $i < c < i+1$, then $\mathbb{P}(i \text{ wins}) \sim i+1-c$ and $\mathbb{P}(i+1 \text{ wins}) \sim c-i$. If $c = i(1+o(1))$, then $\mathbb{P}(i \text{ wins}) \sim 1$.*
- **(ii)** *The number of opinions is reduced to at most three consecutive values in $O(k \log n)$ steps w.h.p., and the expected time for the whole process to finish is $O(n)$.*

A similar analysis for the asynchronous process, giving w.h.p. convergence to three adjacent values in $O(nk \log n)$ steps is given in the full version. The expected time for the asynchronous process to finish is $O(n^2)$.

► **Theorem 2.** ASYNCHRONOUS INCREMENTAL VOTING ON RANDOM GRAPHS.
Let $G \in G_{n,p}$, where $np \geq \log^{1+\varepsilon} n$ for some constant $\varepsilon > 0$. Let the initial values be in $\{1, 2, ..., k\}$ where k is a fixed positive integer, and $S(0) = \sum_{v\in V} X_v(0) = cn$ be the initial total weight.

- **(i)** *If $i < c < i+1$, then $\mathbb{P}(i \text{ wins}) \sim i+1-c$, and $\mathbb{P}(i+1 \text{ wins}) \sim c-i$. If $c = i(1+o(1))$, then $\mathbb{P}(i \text{ wins}) \sim 1$.*
- **(ii)** *The expected time for the asynchronous process to finish is $O(n^2)$.*

[3] We view expansion in terms of the relative number of edges between sets S and $V \setminus S$.

The intuitive basis of Theorems 1 and 2 is to prove that the "extremal" values from $\{1, 2, ..., k\}$ disappear rapidly leaving just two values $i, i + 1$ whose weighted average is c, and to which we can apply the results of two-value pull voting. For K_n this is essentially what happens, and for $G_{n,p}$ it is a reasonable approximation. We remark that the expected time to complete two-value pull voting on K_n and $G_{n,p}$ is $\Theta(n)$ in synchronous, and $\Theta(n^2)$ in asynchronous model (see [2] Chapter 14.3.3 and [9]) and the completion time of incremental voting is asymptotically of the same order.

To complement the above results, for graphs which are not expanders, we give an example on the path graph for which the final answer is quite different than in Theorems 1 and 2. Let P_n be the path with vertex set $\{1, 2, ..., n\}$, with initial values $\{0, 1, 2\}$ ordered on the path vertices in non-decreasing value: first $N_0 \geq 0$ zeroes, then $N_1 \geq 0$ ones and finally $N_2 \geq 0$ twos, where $N_0 + N_1 + N_2 = n$. We refer to such an arrangement as the ordered path.

▶ **Theorem 3.** Asynchronous incremental voting on the ordered path P_n.
If initially $N_0 = an, N_1 = (1 - (a + b))n$, and $N_2 = bn$, then

$$\begin{aligned}
\mathbb{P}(\textit{Opinion 0 wins}) &\sim a(1-b),\\
\mathbb{P}(\textit{Opinion 1 wins}) &\sim ab + (1-a)(1-b),\\
\mathbb{P}(\textit{Opinion 2 wins}) &\sim (1-a)b.
\end{aligned}$$

An example for comparison. We consider values in $\{0, 1, 2\}$. Initially 1/5 of the values are 0, none are 1, and 4/5 are 2. Thus $c = 8/5$ and $1 < c < 2$. In K_n and $G_{n,p}$,

$$\mathbb{P}(0 \text{ wins}) \sim 0, \qquad \mathbb{P}(1 \text{ wins}) \sim 2/5, \qquad \mathbb{P}(2 \text{ wins}) \sim 3/5,$$

whereas on the ordered path

$$\mathbb{P}(0 \text{ wins}) \sim 1/25, \qquad \mathbb{P}(1 \text{ wins}) \sim 8/25, \qquad \mathbb{P}(2 \text{ wins}) \sim 16/25.$$

2 Basic properties of incremental voting

Let $X(t) = (X_v(t) : v \in V)$ be the vector of integer opinions held by the vertices at step t; $X(0)$ is the vector of initial opinions. $A_i(t) = \{v \in V : X_v(t) = i\}$ is the set of vertices holding opinion $i \in \{1, ..., k\}$ at time t, and $N_i(t) = |A_i(t)|$. We may abbreviate by dropping the step index t, e.g., N_i and N_i' would refer to the number of vertices holding opinion i at the beginning and end, respectively, of the current step. Let $S(t)$ be the total weight at step $t \geq 0$: $S(t) = \sum_{v \in V} X_v(t) = \sum_j jN_j(t)$. Let $\pi_v = d(v)/2m$ where m is the number of edges of the graph, and let $Z(t) = n \sum_{v \in V} \pi_v X_v(t)$ be the degree biased weight. For regular graphs, $\pi_v = 1/n$, so $S(t) = Z(t)$. We also use notation $\|\pi\|_2 = \sqrt{\sum_i \pi^2}$ and $\|\pi\|_\infty = \max_{v \in V} \pi_v$.

A random variable $W(t), t = 0, 1, ...$ of the incremental voting process is a martingale if its expected value at the next step depends only on the current opinions $X(t)$, and it satisfies $\mathbf{E}(W(t + 1) \mid X(t)) = W(t)$.

▶ **Lemma 4.** The average weight is a martingale. *The following hold for each $t \geq 0$.*
 (i) Asynchronous edge process. *For arbitrary graphs, $S(t)$ is a martingale.*
 (ii) Asynchronous vertex process. *For arbitrary graphs, $Z(t)$ is a martingale.*
 (iii) Synchronous vertex process. *For arbitrary graphs, $Z(t)$ is a martingale.*

Proof.

Proof of (i). Consider step $t + 1$, take any edge (v, w) and let $\Delta_v(w)$ be the change in X_v if this edge and its endpoint v are chosen in this step. Thus $\Delta_v(w) \in \{-1, 0, +1\}$ and $\Delta_v(w) = -\Delta_w(v)$; see (1). Only one of these changes can occur at a given step in the asynchronous process. Let $e = (v, w)$ be the chosen edge, an event of probability $1/m$ in the edge process. Then,

$$\mathbf{E}(S(t+1) \mid X(t), e=(v,w) \text{ chosen}) = S(t) + \frac{1}{2}\Delta_v(w) + \frac{1}{2}\Delta_w(v) = S(t).$$

Proof of (ii). Let $A_i(t)$ be the vertices with value i at step t. For a vertex $u \in A_i$, let $s_{ij}(u)$ be the number of edges from u to A_j. Adding the edges between A_i and A_j in two ways,

$$\sum_{u \in A_i} s_{ij}(u) = \sum_{v \in A_j} s_{ji}(v). \tag{3}$$

For $1 \leq i < j \leq k$, let Δ_{ij} be the change in $Z(t)$ at step $t+1$ arising from an interaction between A_i and A_j (a vertex with opinion i picks up a neighbour with opinion j, or vice versa). We have $Z(t+1) = Z(t) + \sum_{i<j} \Delta_{ij}$. In the vertex process, we first pick a vertex u u.a.r and then an edge (u,v) from u u.a.r.. If $u \in A_i$ is the sampled vertex, the probability an edge from u to A_j is chosen is $s_{ij}(u)/d(u)$. As $\pi_u = d(u)/2m$, and then using (3),

$$\mathbf{E}\Delta_{ij} = \sum_{u \in A_i} \frac{1}{n}\frac{s_{ij}(u)}{d(u)}\pi_u - \sum_{v \in A_j} \frac{1}{n}\frac{s_{ji}(v)}{d(v)}\pi_v = \frac{1}{2nm}\left(\sum_{u \in A_i} s_{ij}(u) - \sum_{v \in A_j} s_{ji}(v)\right) = 0. \tag{4}$$

Proof of (iii). In the synchronous vertex process, with the notation as above in *(ii)*, the expected value of Δ_{ij} is as in (4) but without the $1/n$ factors (since each vertex selects a neighbour and updates its value). ◀

As the process is randomized, the final value is a random variable with distribution $D(i)$ on the initial values $\{1, ..., k\}$. The following theorem helps us to characterize this distribution in certain cases. If only two neighbouring opinions $i, i+1$ remain at some step t, the process is equivalent to two-value pull voting, and we say the voting is at the final stage.

▶ **Theorem 5.** Distribution of winning value. *Let $W(t)$ stand for $S(t)$, if we refer to the edge model, or for $Z(t)$, if we refer to the vertex model. Let $W(0) = cn$ be the total initial weight, where c is the initial average opinion.*

(i) *For any graph, the expected average opinion at any step is always the initial average: $\mathbf{E}[W(t)/n] = W(0)/n = c$. $W(t)$ converges to a time invariant random variable.*

(ii) *If at the start of the final stage only two opinions i and $i+1$ remain and the total weight W is $c'n$, then for any connected graph, the winning opinion is i with probability $p = i+1-c'$, or $i+1$ with probability $q = c'-i$.*

(iii) *Suppose the final stage is reached in T steps, where $T = o(1/\|\pi\|_2^2)$ for the synchronous vertex process (which reduces to $T = o(n)$ for regular graphs), $T = o(n^2)$ for the asynchronous edge process, and $T = o(1/\|\pi\|_\infty^2)$ for the asynchronous vertex process. Then w.h.p. $|W(T) - W(0)| = o(n)$ and the results of part (ii) hold with $c' \sim c$. That is, for i such that $i \leq c < i+1$, the winning opinion is i with probability $p \sim i+1-c$, or $i+1$ with probability $q \sim c-i$.*

Proof. **(i)** The first part follows from $\mathbf{E}W(t) = W(0)$ (Lemma 4). $\mathbf{E}W^2(t) \leq k^2$ and the limit random variable, for $t \to \infty$, exists by the martingale convergence theorem.
(ii) Using (2), we have $ipn + (i+1)qn = W$, implying that $p = i+1-c'$ and $q = c'-i$.
(iii) In the synchronous vertex process, using Lemmas 12 and 13, we have $|W(T) - W(0)| = o(n)$ w.h.p., provided $T = o(1/\|\pi\|_2^2)$, which reduces to $T = o(n)$ for regular graphs. In the asynchronous edge process, $|S(t+1) - S(t)| \leq 1$, and in the asynchronous vertex process,

$|Z(t+1) - Z(t)| \leq n \max_{v \in V} \pi_v = n\|\pi\|_\infty$, so using the Azuma-Hoeffding inequality for martingale concentration (Lemma 18), we obtain $|W(T) - W(0)| = o(n)$ w.h.p., provided $T = o(n^2)$, respectively $T = o(1/\|\pi\|_\infty^2)$. ◀

In the light of Theorem 5, there are two main ways of analysing the problem. On the (regular) expander graphs we consider the final stage of pull voting with two values i and $i+1$, takes $T = O(n)$ expected time (synchronous process) or $T = O(n^2)$ expected time (asynchronous process); see e.g. [2, 9]. If the time t for the other opinions to disappear, leaving only two neighbouring values, can be shown to be $t = o(T)$ then the total weight $S(t)$ is concentrated around $S(0) = cn$. In this case we can use Theorem 5 to give an asymptotic result on the distribution D. There are however challenges in the analysis, e.g., dealing with the fact that a small number of opinions other than i or $i+1$ may persist in the system for a longer time. This is the topic of Sec. 3 and 5. On graphs which are not expanders we cannot expect $S(t)$ (or $Z(t)$) to be concentrated, but we can try to explicitly obtain the distribution D of the limiting random variable, at least for some special cases. This is the topic of Sec. 4.

In all cases, the winning value of an incremental voting process will be given by a probability distribution D on $\{1, ..., k\}$. As the total weight $W(t)$ is a martingale, the distribution D must have expected value $c = W(0)/n$. Indeed, for the vertex process and $Z(t)$ (and similarly for the edge process and $S(t)$),

$$c = \frac{Z(0)}{n} = \lim_{t\to\infty} \frac{\mathbf{E}Z(t)}{n} = \sum_v \pi_v \lim_{t\to\infty} \mathbf{E}X_v(t) = \sum_v \pi_v \mathbf{E}D = \mathbf{E}D. \tag{5}$$

▶ **Lemma 6.** Completion time, a general bound. *For any connected graph and any of the three types of incremental voting (asynchronous or synchronous vertex process, or asynchronous edge process), the worst-case expected time to eliminate one of the two extreme opinions (over all initial configurations) is at most the worst-case expected completion time of the corresponding (asynchronous, synchronous vertex, or asynchronous edge, respectively) standard 2-opinion voting process.*

Proof. Let $A_i = \{v \in V : X_v = i\}$. We consider our process $(A_1(t), A_1(t), \ldots, A_k(t))_{t\geq 0}$ and the standard 2-opinion voting $B(t)_{t\geq 0}$, where $B(t)$ and $V \setminus B(t)$ are the supports of the two opinions at time t.

We set $B(0) = A_\ell(0)$, where ℓ is the minimum opinion: $\ell = \min\{\kappa : A_\kappa(0) \neq \emptyset\}$, and couple processes A and B, running them on the same random selection of vertices in each step. The two opinions in process B are opinions ℓ and non-ℓ, that is, for process B, each opinion other than opinion ℓ is viewed as the same opinion non-ℓ. While initially the vertices with opinion ℓ in process B are exactly the vertices with this opinion in process A, this does not need to be the case later during the computation. If in the first step a vertex v with A-opinion (its opinion in process A) equal to $q \geq \ell+2$ picks up a neighbour with A-opinion ℓ, then it updates its A-opinion to $q-1 \geq \ell+1$, but its B-opinion becomes ℓ.

Throughout the computation, however, the following relations for the two extreme opinions ℓ and $r = \max\{\kappa : A_\kappa(0) \neq \emptyset\}$ hold by induction,

$$A_\ell(t) \subseteq B(t), \quad A_r(t) \subseteq V \setminus B(t) \tag{6}$$

If at step t the vertex v changes its A-opinion to ℓ, (and consequently $v \in A_\ell(t+1)$), this happens because $v \in A_{\ell+1}(t)$ and picks a neighbour $w \in A_\ell(t)$. Then we must also have $v \in B(t+1)$, because by induction $w \in B(t)$, so its B-opinion is ℓ. Similarly, if $v \in A_{r-1}(t)$ changes its A-opinion to r, this happens because v picks a neighbour $w \in A_r(t)$; By induction $w \in V \setminus B(t)$, and so $v \in V \setminus B(t+1)$.

Let T be the step when the two-voting process B completes, that is, the first step when $B(T)$ is either empty or the whole set V. In the former case, $A_\ell(T) = \emptyset$ and in the latter $A_r(T) = \emptyset$, from (6), so by step T, either opinion ℓ or r must have been eliminated. ◀

▶ **Corollary 7.** *The expected completion time of the discrete incremental voting is $O(k \cdot \mathcal{T}_{2-vote})$, where $\mathcal{T}_{2-vote}$ is the worst-case expected completion time of the 2-opinion voting.*

3 Analysis of asynchronous process for $G_{n,p}$: proof of Theorem 2

In this section we analyse the asynchronous incremental voting on random graphs $G_{n,p}$ above the connectivity threshold, which we view as examples of expanders. Much of the analysis is general, so the results should hold equally for other classes of graphs with expansion properties similar to those in Lemma 8. To indicate why this should be the case, we use the example of the asynchronous edge process on a d-regular graph with opinions in $\{0, 1, 2\}$. The expected change in the two 'extremal' values $0, 2$ at any step is given by

$$\mathbf{E}(N_0' + N_2') = N_0 + N_2 - \frac{2}{dn} M_{0,2}, \tag{7}$$

where $M_{0,2}$ is the number of edges between vertices holding opinion 0 and those holding opinion 2. This is because when an edge in $M_{0,2}$ is selected, then $N_0 + N_2$ is reduced by 1, and when an edge in $M_{0,1} \cup M_{1,2}$ is selected, then the expected change of $N_0 + N_2$ is zero.

For a d-regular connected graph G and any two vertex sets S and T in G,

$$\left| e(S,T) - \frac{d|S||T|}{n} \right| \le \lambda \sqrt{|S|(1 - |S|/n)|T|(1 - |T|/n)}, \tag{8}$$

where $e(S,T)$ is the number of edges between S and T and λ is the absolute value of the second eigenvalue of the adjacency matrix of G. Assume that $\lambda = \varepsilon d$ for some constant $\varepsilon < 1$ (the expander assumption) and take $|S| = N_0$ and $|T| = N_2$. If N_0 and N_2 are large, then so is $M_{0,2} = e(A_0, A_2)$, since from (8), $M_{0,2}$ is close to dN_0N_2/n. Thus from (7), $N_0 + N_2$ quickly decreases. Although this type of argument may not allow us to completely eliminate one of the extremal values 0 and 2 (as $M_{0,2}$ becomes too small, eventually 0, while both A_0 and A_2 are still non-empty), we can use it to make one of N_0 and N_2 small relative to N_1.

Returning to graphs $G_{n,p}$, we assume all opinions are in $\{1, 2, \dots, k\}$, for an arbitrary but fixed integer k (constant, while n grows to infinity). The entire point of the proof in this section is to ensure that within $T = o(n^2)$ steps, all but $o(n)$ vertices have two adjacent opinions in $\{i, i+1\}$. This will allow us to apply Theorem 5 (ii)-(iii), with $c' \sim c$.

In the edge process (and in the vertex process in regular graphs) the expected change in the number of vertices with any given value can be characterised as follows. Let $M_{i,j} = M_{i,j}(t)$ be the number of edges between sets A_i and A_j at step t. Letting $N_i = N_i(t)$ and $N_i' = N_i(t+1)$,

$$\mathbf{E}N_i' = N_i + \frac{1}{2m} \left(\sum_{j \ge i+1} M_{i-1,j} + \sum_{j \le i-1} M_{i+1,j} - \sum_{j \ne i-1,i,i+1} M_{i,j} \right). \tag{9}$$

If $k > 2$, then the number of vertices with an extreme value 1 or k exhibits downward drift (the first two sums in (9) are equal to 0 for i equal to 1 or k), provided there are edges between A_1 and $\bigcup_{j \ge 3} A_j$, or between A_k and $\bigcup_{j \le k-2} A_j$. Thus if there is enough connectivity (expansion) in the graph, then the support of one of the extremal values reduces relatively quickly to $o(n)$. If this was, say, opinion 1, then, still relatively quickly, the support of the next extremal opinion, either 2 or k, reduces to $o(n)$; and so on, until the support of all opinions other than some two consecutive opinions i and $i+1$ is reduced to $o(n)$. The analysis of the completion of the process from such a state will requires another approach.

In what follows, ω and ω' denote functions tending to infinity with n, with $\omega' = o(\omega)$. In general the exact values are not important but the growth to infinity has to be sufficiently slow to satisfy the bounds arising in the analysis. In the final part of the analysis, we will choose $\omega = \log n$, and $\omega' = \log\log n$.

Required properties of $G_{n,p}$. The following are the expansion properties of $G_{n,p}$ needed for our proofs. The lower bound on np ensures that w.h.p. the graph is connected. To maintain continuity of discussion the proof of the following lemma is given in Appendix A.

▶ **Lemma 8.** *Let $G \in G_{n,p}$, where $np \geq \log^{1+\varepsilon} n$ for some constant $\varepsilon > 0$. The following properties hold w.h.p..*

P1. (Almost regular graphs) *G is connected and all vertices v have degree $d(v) = np + O(\sqrt{np \log n})$ and stationary distribution $\pi_v = \frac{1}{n} + O\left(\frac{1}{n \log^{\varepsilon/2} n}\right)$.*

P2. (Large number of edges between large subsets of vertices) *Let $\delta \geq 5/\sqrt{np}$. For any pair of disjoint vertex sets A, B, with $|A| \geq \delta n$, $|B| \geq \delta n$, the number of edges X_{AB} between A and B satisfies $\mu/2 \leq X_{AB} \leq 3\mu/2$, where $\mu = |A||B|p$, the expected number of edges between the sets A and B in $G_{n,p}$.*

P3. (Not too many edges within small subsets of vertices)
- **(i)** *For $\omega \geq e$, $\omega \log \omega \leq np$, no vertex set S of size $s = n/\omega$ induces more than $X_S = e^2 s^2 p$ edges.*
- **(ii)** *Let $d = np$. No set S, $|S| \leq n/\omega$ induces more than $X_S = s\sqrt{4d \log(ne/s)}$ edges.*
- **(iii)** *Provided $\omega = O(\log n)$ and $np = d \geq \log^{1+\theta} n$, the ratio $X_S/X_{S,V-S}$ is $O(1/\omega)$, the value achieved in P3.(i) above.*

Outline to the analysis of the process. Our analysis of $G_{n,p}$ is for the edge process, but as w.h.p. vertex degrees are concentrated for the range of p we consider, the vertex process and edge process are asymptotically equivalent. Indeed, by property P1 of Lemma 8, the degree weighted total $Z(t)$ and the unweighted total $S(t)$ satisfy $|Z(t) - S(t)| \leq c/\log^{\varepsilon/2} n$, and it suffices to analyse the convergence of $S(t)$.

Ideally we would like to keep completely removing the values $\{1, ..., k\}$ one by one in some order, as in the proof of Theorem 1. As can be seen from (9), the drift on the extremal values $1, k$ is negative or zero.

$$\mathbf{E}N_1' + \mathbf{E}N_k' = N_1 + N_k - \frac{1}{2m}\left(\sum_{j\geq 3} M_{1,j} + \sum_{j\leq k-2} M_{k,j}\right). \tag{10}$$

Thus at least one extremal value 1 or k should disappear, allowing us then to repeat the analysis with e.g., values $\{2, ..., k\}$. However, the time taken for such an approach is $\Omega(n^2)$, which is too long for the total weight $S(t)$ to remain concentrated around $S(0)$. Therefore, in our analysis, we settle for making one of the extremal values sufficiently small, which can be done in $o(n^2)$ steps and then repeat the analysis for the remaining large values. Finally one value dominates, and w.h.p. all other values disappear at some subsequent step. It remains to be proved below that such an approach can be made to work.

The analysis proceeds in three phases, which in outline are as follows.

I. One by one, the extremal values are made small. By the beginning of iteration r, $1 \leq r \leq k-2$, the support for $r - 1 = i - 1 + (k - j)$ extremal values $\{1, 2, \ldots, i - 1\} \cup \{j + 1, j + 3, \ldots, k\}$ has been made small, but $N_i > \delta_r n$ and $N_j > \delta_r n$. During iteration r, the next extremal value, either i or j, is made small. As we progress through the iterations, our analysis loses accuracy, so δ_r increases with r (but remains $o(1)$).

II. For two adjacent values i and $i+1$, $N_i, N_{i+1} > n/\omega$ and $N_i + N_{i+1} = n(1-o(1))$.
III. There is a unique value i with $N_i = n(1-o(1))$.

Arriving at Phase II, the process corresponds (in general principle) to ordinary pull voting with two values. If at the completion of Phase I $\min\{N_i, N_{i+1}\} < n/k\omega$, we skip Phase II. Phase III is a clean up phase, removing any remaining small sets. At the end of Phase III, $N_i = n$, and the analysis is completed.

Phase I. Making small the first extremal value (either N_1 or N_k).

▶ **Lemma 9.** *Let $\delta = \max\left(1/n^{1/4}, 5/\sqrt{np}\right)$. Let T_1 be the number of process steps to reduce one of A_1 or A_k to size at most δn and let α be the extreme opinion (1 or k) with the size of support at most δn at step T_1. Then the following hold w.h.p.*
(i) $\mathbf{E}T_1 = O(n^{3/2})$.
(ii) $|S(T_1) - S(0)| = o(n)$.
(iii) *$N_\alpha(t) \le \omega\delta n$ at all steps $t > T_1$ (for some $\omega \to \infty$).*

Proof. Let $A = A_1$, $B = A_k$, $|A| = N_1 > \delta n$ and $|B| = N_k > \delta n$. We proceed in stages indexed by *decreasing* ℓ. At the beginning of the current stage, we assume w.o.l.g. that $N_1 \le N_k$ and integer $\ell \ge 1$ is such that $\ell\delta n \le |A| < (\ell+1)\delta n$ (that is, $\ell = \lfloor |A|/(\delta n)\rfloor$). This stage continues until N_1 or N_k drops below $\ell\delta n$, when the next stage $\ell - 1$ starts. At the end of the final stage, for $\ell = 1$, one of N_1 or N_k is less than δn.

We estimate the expected time of one stage ℓ by first estimating the number of *active* steps in this stage, defined as the steps when N_1 changes, either by -1 or $+1$. We estimate the number of active steps by comparing the random variable N_1 with the biased random walk on the integer line. We then factor in the expected number of process steps between two consecutive active steps.

We view the random selection in the current step as selection of a uniformly random *oriented* edge (v,u), where v is the vertex which updates its value. The value of N_1 decreases, resp. increases, by 1, if and only if, the selected oriented edge belongs to $(A, \overline{A})$, resp. (A_2, A). Thus in the current *active* step N_1, the ratio of the probability q that N_1 decreases by 1 to the probability $r = 1-q$ that N_1 increases by 1 is equal to

$$\frac{q}{r} = \frac{|(A,\overline{A})|}{|(A_2,A)|} \ge \frac{|(A,\overline{A})|}{|(\overline{A},A)| - |(B,A)|} \ge 1 + \frac{|(B,A)|}{|(A,\overline{A})|} \ge 1 + \frac{N_1\, N_k\, p/2}{N_1\, n\, 3p/2} \ge 1 + \frac{\ell\delta}{3} = \frac{1+\varepsilon}{1-\varepsilon}, \quad (11)$$

where $\varepsilon \ge \ell\delta/4$. We have used Lemma 8.P2 to bound the number of edges between sets A and B and sets A and $\overline{A}$.

For a biased random walk on the integer line $\{0, 1, ..., L\}$, where in each step the probabilities of moving left (towards 0) or right are equal to $q > 1/2$ and $r = 1-q$, respectively, the probability q_z of ruin (absorption at zero) and the expected number d_z of steps to ruin when starting from position z are equal to

$$q_z = \frac{\rho^L - \rho^z}{\rho^L - 1} = 1 - \frac{\rho^z - 1}{\rho^L - 1}, \qquad d_z = \frac{z}{q-r}\left(1 - \frac{L}{z}\cdot\frac{\rho^z - 1}{\rho^L - 1}\right). \quad (12)$$

where $\rho = q/r > 1$. We view N_1 as a random walk on the integer line $\{\ell\delta n, \ldots, (\ell+2)\delta n\}$, starting at $\ell\delta n + z$, for $z = z(\ell) < \delta n$. At an active step, the probabilities of moving left or right are at least $q = \frac{1}{2}(1+\varepsilon)$ and at most $r = \frac{1}{2}(1-\varepsilon)$, respectively. Then the probability $q'_{z(\ell)}$ that N_1 reaches $\ell\delta n$ before it reaches $(\ell+2)\delta n$ or N_k reaches $\ell\delta n$ is such that, from (12),

$$1 - q'_{z(\ell)} \le \frac{\rho^{\delta n} - 1}{\rho^{2\delta n} - 1} = \frac{1}{\rho^{\delta n} + 1} \le \frac{1}{(1+\ell\delta/3)^{\delta n}} \le e^{-\ell\delta^2 n/4} \le e^{-\sqrt{n}/4}, \quad (13)$$

where the last inequality follows from $\delta \geq 1/n^{1/4}$. The expected duration $d'_{z(\ell)}$ of this stage, in terms of the number of active steps, is at most, from (12),

$$d'_{z(\ell)} \leq \frac{z}{q-r} \leq \frac{\delta n}{\varepsilon} = \frac{4n}{\ell}. \tag{14}$$

By Lemma 8.P2, at step t of the current stage ℓ,

$$\mathbb{P}(t \text{ is an active step}) \geq \frac{|(A, \overline{A})|}{2m} \geq \frac{(\ell\delta n)(n/2)p/2}{(1+o(1))n^2 p} \geq \frac{\ell\delta}{5} \equiv P_\ell. \tag{15}$$

Thus for the first process step T_1 at which $N_1 < \delta n$ or $N_k < \delta n$, we have, from (14) and (15),

$$\mathbf{E}(T_1) = \sum_{\ell} \frac{1}{P_\ell} d'_{z(\ell)} \leq \sum_{\ell \geq 1} \frac{5}{\ell\delta} \cdot \frac{4n}{\ell} = O\left(\frac{n}{\delta}\right) = O(n^{5/4}).$$

As $S(t)$ is a martingale, $\mathbf{E}S(t) = S(0)$. Apply the Azuma martingale inequality (Lemma 18) to the sequence of oriented edges $(e_1, ..., e_t)$ inspected at steps $1, \ldots, t$. At each step, S changes by at most 1. At step $T_1^* = \omega \mathbf{E} T_1$, with $h = \sqrt{3T_1^*(\omega + \log T_1^*)} = o(n)$,

$$\begin{aligned} \mathbb{P}((|S(T_1) - S(0)| \geq h) &\leq \mathbb{P}(\exists T < T_1^* : |S(T) - S(0)| \geq h) + \mathbb{P}(T_1 > T_1^*) \\ &\leq T_1^* e^{-h^2/(3T_1^*)} + o(1) = e^{-\omega} + o(1) = o(1). \end{aligned}$$

Thus w.h.p. $|S(T_1) - S(0)| = o(n)$, as required.

For part (iii) of the lemma, beyond step T_1, we bound N_α (where α is opinion 1 or k and $N_\alpha \leq \delta n$ at step T_1) with the progress of *unbiased* random walk on $\{0, 1, \ldots, L = \omega\delta n\}$ starting at $z \leq \omega\delta n$, which with probability $1 - 1/\omega$ is absorbed at zero before reaching L. ◀

Phase I. Making the next extremal value small. Having completed the first iteration, we continue inductively for general iteration g, $1 < g \leq k-2$. By the beginning of this iteration, the support for $g-1$ extremal values $\{1, 2, \ldots, i-1\} \cup \{j+1, j+2, \ldots, k\}$ has been made small, but $N_i > \delta_g n$ and $N_j > \delta_g n$. Note that $j \geq i+2$. During iteration g, the next extremal value is made small, that is, N_i or N_j is reduced to at most $\delta_g n$.

Let $\delta_1 = \delta$, and in general, for $1 < g \leq k-2$, $\delta_g = \omega^{2(g-1)}\delta$, where $\omega \to \infty$ but sufficiently slowly so that $\omega^3 \delta_{k-2} = o(1)$. The argument is similar as in Lemma 9 for the first iteration, replacing N_1, N_k and δ with N_i, N_j and δ_g, and assuming by induction that the support for the reduced $g-1$ values is, and w.h.p. will remain, at most $\omega\delta_1 n + \omega\delta_2 n + \cdots \omega\delta_{g-1} n < 2\delta_g/\omega$. As in (11), in an active step of stage ℓ of this iteration, the ratio of the probability q that N_i decreases by 1 to the probability $r = 1-q$ that it increases by 1 is equal to

$$\frac{q}{r} = \frac{|(A_i, \overline{A_i})|}{|(A_{i+1}, A_i)| + |(A_{i-1}, A_i)|} \geq \frac{|(A_i, \overline{A_i})|}{|(\overline{A_i}, A_i)| - |(A_j, A_i)|} \geq 1 + \frac{\ell\delta_g}{3} = \frac{1+\varepsilon}{1-\varepsilon},$$

where $\varepsilon \geq \ell\delta_g/4$. Proceeding as in the proof of Lemma 9, we conclude that w.h.p.: (i) the expected time $\mathbf{E}(T_g - T_{g-1})$ of iteration g is $O(n/\delta_g) = O(n^{5/4})$; (ii) at step T_g when this iteration ends, $|S(T_g) - S(0)| = o(n)$; (iii) $N_\alpha(T_g) \leq \delta_g n$ and $N_\alpha(t) \leq \omega\delta_g n$ at all steps $t > T_g$, where α is the reduced opinion i or j. (Recall that we assume that k is constant.)

For the final point (iii), the argument is more subtle for the general iteration than it was for the first iteration. Assuming that $\alpha = i$, N_i may have a tendency to increase (a positive drift) after step T_g, if $M_{i-1,i+1}$ happens to become larger than the number of edges adjacent to A_i (see (9)). To deal with this, while there are opinions smaller than i we only care that if N_i increases to $\omega\delta_g n/2$, then the probability that it further increases to $\omega\delta_g n$ before going back below $\delta_g n$ is exponentially small (as in (13)). When i becomes the smallest surviving opinion, than we start comparing the changing N_i with the unbiased random walk on $\{0, 1, \ldots, L = \omega^2\delta_g n\}$ starting from $z \leq \omega\delta_g n$ (similalry as in the first iteration).

Phase II analysis. When Phase I ends at a step t_I, then for some i, the total support for $k-2$ values $\{1, 2, \dots, i-1\} \cup \{i+2, i+3, \dots, k\}$ is at most $2\omega\delta_{k-2}n \le n/\omega^2$. Thus Phase II starts with $N_i(t_I) + N_{i+1}(t_I) = n(1-o(1))$. We also have $|S(t_I) - S(0)| = o(n)$ (since $S(t)$ is a martingale and $t_I = O(n^{3/2})$), so we can take $i = \lfloor c \rfloor$, where $c = S(0)/n$, and have $N_i = (i+1-c)n + o(n)$, $N_{i+1} = (c-i)n + o(n)$. During Phase II, we compare N_i with unbiased random walk on $\{0, \dots n\}$ which is driven by selecting edges in $M_{i,i+1}$. Each such edge increases or decreases N_i by 1 with equal probability. We continue Phase II only while $\min(N_i, N_{i+1}) \ge \omega^2\delta_{k-2}n$ to ensure that the opinions other than i and $i+1$ remain small and distort the unbiased random walk of N_i only in a negligible way.

Phase II ends within the expected $O(n^2)$ steps, and w.h.p. $N_{i'} \ge n - n/\omega$, and $\sum_{j \ne i'} N_j \le n/\omega$, where $i' = i$ with probability $i+1-c+o(1)$, or $i+1$ with probability $c-i+o(1)$.

Phase III analysis. At the start of Phase III, there is one opinion i for which $N_i = n(1-o(1))$. Let $A = A_i$, and $B = \cup_{j \ne i} A_j$. We show that w.h.p. opinion i wins.

The process during the final Phase III resembles a "balls in boxes" system in which a vertex with value j is a box with j balls. When a vertex is selected, then it may get one ball added or removed, depending on the number of balls in the selected neighbour. When all vertices have the same value (the same number of balls), the process ends. If an edge with values $(i,j), j \ne i$ is chosen, then we call this a Type 1 event. Choosing an edge with values $(j,j'), j \ne i \ne j'$ is a Type 2 event. We ignore (i,i) and (j,j) events.

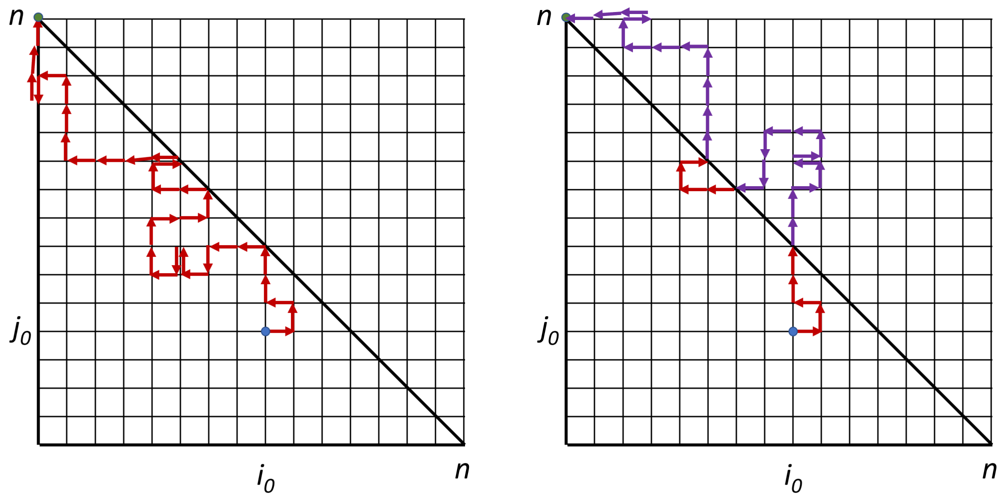
We compare this process with an unbiased walk on integers $\{0, 1, ..., L\}$ starting from $z = \hat{S}(0) = \sum_{j \ne i} |j-i| N_j$, the weight of set B at the start of Phase III w.r.t. value i, representing the distance between the starting configuration and the target configuration when all vertices have value i. Thus initially $z \le k|B| \le n/\omega$. Each Type 1 event changes z by $+1$ or -1 with equal probability (once an edge is selected, one of its end vertices is chosen for an update with equal probability). If only Type 1 events occurred, then z would be an unbiased random walk on $\{0, 1, ..., L\}$, and we put $L = \omega z$. (The process does not stop before $\hat{S}(t)$ reaches 0 or L.) Value i wins when the walk is absorbed at zero. The probability of this is $1 - z/L = 1 - 1/\omega$ and the expected duration is $z(L-z)$.

For steps $t = 0, 1, ...,$ as Phase III proceeds, the value $\hat{S}(t)$ will change due to both Type 1 and Type 2 events. The change due to Type 1 events is directly included in the random walk given above, and with probability $1 - \omega' z/n$ the walk will not increase above $\omega'\hat{S}(0) \le \omega' n/\omega$, where $\omega' \to \infty$ but $\omega' = o(\omega)$.

Type 2 events are not represented directly by the random walk. Each Type 2 occurrence can change $\hat{S}$ by $+1$ or -1. A Type 2 event on an edge (j,j'), where $j < j'$, increases or decreases the number of balls in the system by one with equal probability. After T events of Type 2, the additional change in $\hat{S}$ due to this is $(+1)X + (-1)(T-X)$ where $X \sim Bin(T, 1/2)$. Thus w.h.p. X will not exceed $O(\sqrt{T \log T})$.

Only Type 1 moves can increase the size of B, whereas Type 2 moves may decrease it, if $j = i-1$ or $j' = i+1$. W.h.p. the maximum size of B due to the Type 1 walk is at most $s = \omega' n/\omega$. By Lemma 8, property P2, w.h.p. no set of size s induces more than $O(s^2 p)$ edges, whereas by property P1, there are at least $nsp/3$ edges between A_i and B. Thus the probability of a Type 2 event is at most $O(s/n)$, and the number of Type 2 events in the duration of the Type 1 random walk is w.h.p. of order at most $T = zL\frac{s}{n}\omega' = O((n\omega'/\omega)^2)$. W.h.p. the maximum increase of $\hat{S}$ due to Type 2 events is $O(\sqrt{T \log T}) = O(n\omega'(\log n)^{1/2}/\omega)$. Taking $\omega = \log n$ and $\omega' = \log\log n$ (so $\omega \ge (\omega')^2(\log n)^{1/2}$), we can increase $\hat{S}(0)$ to $z' = n/\omega'$ and conclude that with probability $1 - 1/\omega'$, at the end of Phase III $N_i = n$, as required.

It can be shown, by counting in the ignored events (i,i) and (j,j) and considering stages of halving the value z, that Phase III ends within the expected $O(n^2)$ steps.



(a) Random walk on the triangle grid. **(b)** Random walk on the square grid.

Figure 1 The random walk on the square grid (right diagram) and the corresponding walk on the triangle grid (left diagram). The random walk of the triangle grid defines the evolution of the incremental voting on the ordered path. In this example, the averaging process starts with i_0 vertices with opinion 0 and j_0 vertices with opinion 2, and stabilises with all vertices having opinion 2.

4 Asynchronous incremental voting on the line: proof of Theorem 3

To indicate that Theorems 5 and 2 do not hold for general graphs, we consider the following specific example of an *ordered path*. The graph is a path with n vertices $\{1, 2, \ldots, n\}$. There are three opinions $0, 1, 2$ and initially they are ordered along the path: vertices $\{1, \ldots, i_0\}$ have opinion 0, vertices $\{i_0 + 1, \ldots, n - j_0\}$ have opinion 1, and vertices $\{n - j_0 + 1, \ldots, n\}$ have opinion 2. Thus, the $0 \le i_0 \le n$ vertices in the initial segment of the path have opinion 0, the $0 \le j_0 \le n - i_0$ vertices in the final segment of the path have opinion 2, and the remaining $n - (i_0 + j_0)$ vertices in the middle of the path have opinion 1.

We show that the probability that opinion 0 wins is equal to $a(1 - b)$, where $a = i_0/n$ and $b = j_0/n$. By symmetry, the probability that opinion 2 wins is equal to $(1 - a)b$, leaving the probability of $ab + (1 - a)(1 - b)$ for opinion 1 to win.

The non-decreasing order of the opinions along the path is an invariant of the process, so each intermediate configuration is characterised by the number $i = N_0$ of vertices at the beginning of the path with opinion 0, and the number $j = N_2$ of vertices at the end of the path with opinion 2, where $0 \le i \le n$, $0 \le j \le n - i$. The process ends when opinion 0 wins (i becomes n), or opinion 2 wins (j becomes n), or opinion 1 wins (both i and j become 0).

The process of changing from one configuration to the next one is a random walk on the integral points of the triangle $i \ge 0, j \ge 0, i + j \le n$; see Figure 1a. Considering only the steps when the configuration changes, a configuration (i, j) which is strictly inside this triangle (that is, $i > 0, j > 0, i + j < n$) changes to any of the four configurations $(i + 1, j)$, $(i - 1, j)$, $(i, j + 1)$ and $(i, j - 1)$ with equal probability of $1/4$. Indeed, configuration (i, j) changes when either edge $(i, i + 1)$ (with opinions 0 and 1) or edge $(n - j, n - j + 1)$ (with opinions 1 and 2) is selected (equal probability). If edge $(i, i + 1)$ is selected, then the configuration changes to $(i + 1, j)$ or $(i - 1, j)$, depending which vertex i or $i + 1$ updates its opinion. With equal probability, either vertex $i + 1$ decreases its opinion from 1 to 0, or vertex i increases its opinion from 0 to 1. Analogously when edge $(n - j, n - j + 1)$ is selected.

From a configuration $(0, j)$, where $0 < j < n$, we have equally probable transitions to configuration $(0, j+1)$ or $(0, j-1)$. Analogously, a configuration $(i, 0)$, where $0 < i < n$, transitions to $(i+1, 0)$ or $(i-1, 0)$ with equal probability.

Finally, consider the *diagonal* configurations lying on the side of the triangle formed by the line segment from $(0, n)$ to $(n, 0)$. For a non-final configuration (i, j): $i + j = n$, $i > 0$, $j > 0$, the vertex i has opinion 0 and vertex $i+1$ has opinion 2. The configuration changes when the unique edge $(i, i+1)$ is selected. In this case the configuration transitions to $(i-1, j)$, when vertex i increases its opinion from 0 to 1, or to $(i, j-1)$, when vertex $i+1$ decreases its opinion from 2 to 1 (equal probability for either of these two transitions).

For convenience, we view this random walk W on the triangle as a random walk W' on the full square $0 \leq i \leq n, 0 \leq j \leq n$, unifying the pairs of states (i, j) and $(n-j, n-i)$, these being identical on the diagonal of the triangle. See Figure 1, where the right diagram gives an example of the random walk W' on the square grid, and the left diagram shows the corresponding walk on the triangle. The transition probabilities for walk W' are the same as for W for all non-diagonal states (i, j). For such a state, if it is not on the boundary of the square, then one of the coordinates increases or decreases by 1, with all four possibilities equally probable. For a state on the boundary of the square, the configuration changes, with equal probability, to one of the two neighbouring boundary states.

For a diagonal non-final state (i, j), the random walk W' moves also to any of the four neighbouring states $(i+1, j)$, $(i-1, j)$, $(i, j+1)$ or $(i, j-1)$, with equal probability. In this case, the transition of W' with probability $1/2$ to either $(i-1, j)$ or $(i, j+1)$ corresponds to random walk W moving with probability $1/2$ from configuration (i, j) to configuration $(i-1, j)$. Thus the pair of states $(i-1, j)$ and $(i, j+1)$ in W' correspond to the configuration $(i-1, j)$ in W. The below diagonal state (i, j) and above diagonal state $(n-j, n-i)$ in the square, both correspond to state (i, j) in the triangle.

Thus our incremental voting on the path corresponds to the random walk W' on the square grid. Inside the square the walk transitions with equal probability from one state to any of the four neighbouring states. A transition on one coordinate is completely independent of the value of the other coordinate. When the walk hits a side of the square, this corresponds to one of the two extreme values 0 or 2 being eliminated. The walk then remains within this side of the square, moving independently to one of the two neighbouring boundary states. The final absorbing states are the four corners of the square. State $(n, 0)$ corresponds to opinion 0 winning, state $(0, n)$ corresponds to opinion 2 winning, and states $(0, 0)$ and (n, n) (corresponding to state $(0, 0)$ in the triangle) correspond to opinion 1 winning.

What is the probability that the random walk W' terminates in the state $(n, 0)$, meaning the win for opinion 0? We generate the two-dimensional random walk W' from two independent one-dimensional walks, one walk for each of the two coordinates, both walks with the range $\{0, 1, \ldots, n\}$. To move walk W', we take, with equal probability, the next step from one of the two one-dimensional walks. Walk W' ends in the state $(n, 0)$ if, and only if, the one-dimensional walk for the coordinate i ends in state n and the one-dimensional walk for the coordinate j ends in state 0. Indeed, for the "if" part, if the one dimensional random walks for coordinates i and j end in states n and 0, respectively, then walk W' must end in the state $(n, 0)$, For the "only if" part, if walk W' ends in $(n, 0)$, then the walk for coordinate i cannot end in 0. Otherwise, if the walk for coordinate i ended in 0, then walk W' would reach a state $(0, y)$, for $0 < y < n$, and then end in either $(0, 0)$ or $(0, n)$, or would reach a state $(x, 0)$, for $0 < x < n$ and then end in $(0, 0)$, or would reach a state (x, n), for $0 < x < n$, and then end in $(0, n)$. Analogously, if walk W' ends in $(n, 0)$, then the walk for coordinate j cannot end in n.

For an unbiased random walk on $\{0, 1, \ldots, n\}$ staring at position X, the probability that the walk ends in the state 0 is equal to $(n - X)/n$. The one-dimensional random walks for the coordinates i and j start at positions i_0 and j_0, respectively. Thus the probability that the first walk ends in n is equal to $i_0/n = a$ and the probability that the second walk ends in 0 is equal to $(n - j_0)/n = 1 - b$.

5 Synchronous incremental voting on K_n: Theorem 1

In this section, we show Theorem 1, which refers to the synchronous process on the complete graph K_n. At each discrete time step, each vertex v chooses a vertex w independently and uniformly at random, and updates its opinion X_v to X'_v as in (1). We are interested in the evolution of $(X(t))_{t \geq 0}$.

Firstly, we show that the smallest opinion $s = \min_{v \in V} X_v(0)$ or the largest opinion $\ell = \max_{v \in V} X_v(0)$ vanishes w.h.p. within $O(\log n)$ steps, while $s + 3 \leq \ell$ (Lemma 10). Hence, after the smallest or largest opinion disappears $k - 3$ times, which occurs w.h.p. in $T = O(k \log n) = o(n/\log n)$ steps, at most three consecutive opinions $\{i - 1, i, i + 1\}$ are left. Using a Martingale concentration argument (Lemma 12) we further show that w.h.p. $|S(T) - S(0)| = O(\sqrt{nT \log n}) = o(n)$.

At this point only three adjacent values $\{i - 1, i, i + 1\}$ remain. In Lemma 15, we next either reduce the number of remaining opinions to two consecutive opinions, or if not, and we still have three opinions, then the sizes of opinions $i - 1$ and $i + 1$ are $o(n)$. This reduction takes $o(n)$ steps w.h.p., so we still have $|S(T) - S(0)| = o(n)$). In either case, the next, final phase completes in $O(n)$ expected steps, by comparing with pull voting. The comparison is straightforward, if only two consecutive opinions i and $i + 1$ remain.

When there are three opinions $i - 1, i, i + 1$, where $|A_{i-1} \cup A_{i+1}| = o(n)$, then $S(t)/n \sim i$, and we prove that w.h.p. i wins by coupling the process with pull voting. Let $A = A_i =$ and $B = A_{i-1} \cup A_{i+1}$. In pull voting value i wins with probability $|A|/n = 1 - o(1)$. After the first step of synchronous pull voting, $|A'_P| = Bin(n, |A|/n)$.

There is a coupling between incremental voting on three values, and pull voting such that $|A'_I|$ stochastically dominates $|A'_P|$. Firstly the number of vertices which choose in A directly is $Bin(n, |A|/n)$. Denote this set by $A' = A'_P$ and let $|A'_P| = X_P$. Given the set A'_P, a further non-negative number Y_I of vertices take the value i indirectly. This number Y_I is a sum of two binomials: $Y_I = Bin(|A_{i+1} \setminus A'_P|, N_{i-1}/n) + Bin(|A_{i-1} \setminus A'_P|, N_{i+1}/n)$. We have $A'_P \subseteq A'_I$, and the coupling can be extended to subsequent steps. Thus the probability that i wins in incremental voting is at least the probability that i wins in pull voting, so $1 - o(1)$.

5.1 Many opinions case

First, we show that one of the extreme opinions disappears within $O(\log n)$ steps.

▶ **Lemma 10.** *Let $s = \min_{v \in V} X_v(0)$ and $\ell = \max_{v \in V} X_v(0)$ be the smallest and the largest opinions in the initial round, respectively. Suppose $\ell \geq s + 3$. Then, $N_s(T)N_\ell(T) = 0$ w.h.p. within $T = O(\log n)$ steps.*

Applying Lemma 10 repeatedly, we immediately have the following.

▶ **Theorem 11.** *From any initial configuration of opinions from $[k] = \{1, 2, \ldots, k\}$, $X_v(T) \in \{i - 1, i, i + 1\}$ holds for some $1 < i < k$ and for any $v \in V$ within $T = O(k \log n)$ steps w.h.p.*

Proof of Lemma 10. By definition, we have that $N'_s \sim Bin\left(N_s + N_{s+1}, N_s/n\right)$ and $N'_\ell \sim Bin\left(N_{\ell-1} + N_\ell, N_\ell/n\right)$. Furthermore, N'_s and N'_ℓ are independent since $s+1 < \ell - 1$. Write $Z = N_s N_\ell$ and $Z' = N'_s N'_\ell$. Then, we have

$$\begin{aligned}\mathbf{E}\left[Z'\right] &= \mathbf{E}\left[N'_s N'_\ell\right] = \mathbf{E}\left[N'_s\right]\mathbf{E}\left[N'_\ell\right] = (N_s + N_{s+1})\frac{N_s}{n}(N_{\ell-1} + N_\ell)\frac{N_\ell}{n} \\ &= Z\frac{N_s + N_{s+1}}{n}\frac{N_{\ell-1} + N_\ell}{n} \le Z\frac{N_s + N_{s+1}}{n}\left(1 - \frac{N_s + N_{s+1}}{n}\right) \le \frac{1}{4}Z. \end{aligned} \tag{16}$$

The first inequality follows from $N_s + N_{s+1} + N_{\ell-1} + N_\ell \le n$. For $Z(t) = N_s(t)N_\ell(t)$, (16) implies that $\mathbf{E}\left[Z(t+1)\right] \le \mathbf{E}[Z(t)]/4$ holds for any $t \ge 0$. Taking $T = \lceil 3\log n\rceil$ and using the Markov inequality, we obtain

$$\mathbb{P}\left[Z(T) > 0\right] \le \mathbf{E}\left[Z(T)\right] \le \frac{1}{4}\mathbf{E}\left[Z(T-1)\right] \le \cdots \le \frac{1}{4^T}\mathbf{E}[Z(0)] \le \frac{n^2}{\mathrm{e}^{\lceil 3\log n\rceil}} \le \frac{1}{n}. \quad ◀$$

5.2 Difference from the initial average

Next, we show that the average of the opinions is concentrated around the initial average.

▶ **Lemma 12.** *Let $S(t) = \sum_{v\in V} X_v(t)$. For any $T \ge 0$ and $\epsilon > 0$,*

$$\mathbb{P}\left[|S(T) - S(0)| \ge \epsilon\right] \le 2\exp\left(-\frac{\epsilon^2}{2nT}\right).$$

Proof. First, by Lemma 4 we observe that $(S(t))_{t=0,1,2,\ldots}$ is a martingale. From definition, we have $X_v(t+1) - X_v(t) \in \{-1, 0, 1\}$. Furthermore, for any $v \ne v'$, $X_v(t+1) - X_v(t)$ and $X_{v'}(t+1) - X_{v'}(t)$ are independent. Write $\Delta_v(t) = X_v(t) - X_v(t-1)$. Applying Lemma 19, we have

$$\begin{aligned}\mathbf{E}\left[\mathrm{e}^{\lambda(S(t+1)-S(t))}\middle| X(t)\right] &= \mathbf{E}\left[\mathrm{e}^{\lambda((S(t+1)-S(t))-\mathbf{E}[S(t+1)-S(t)|X(t)])}\middle| X(t)\right] \\ &= \mathbf{E}\left[\mathrm{e}^{\lambda\sum_{v\in V}(\Delta_v(t+1)-\mathbf{E}[\Delta_v(t+1)|X(t)])}\middle| X(t)\right] \\ &= \prod_{v\in V}\mathbf{E}\left[\mathrm{e}^{\lambda(\Delta_v(t+1)-\mathbf{E}[\Delta_v(t+1)|X(t)])}\middle| X(t)\right] \\ &\le \prod_{v\in V}\mathrm{e}^{\frac{\lambda^2}{2}} = \mathrm{e}^{\frac{\lambda^2 n}{2}}. \end{aligned} \tag{17}$$

Combining (17) and Lemma 20, we obtain the claim. ◀

▶ Remark. Choose $T = \lceil 3\log n\rceil$ from the proof of Theorem 11 and $\epsilon = \sqrt{7n}\log n$ in Lemma 12 to obtain

$$\mathbb{P}(|S(T) - S(0)| \ge \sqrt{7n}\log n) < \frac{1}{n}.$$

Note that the $2\exp\left(-\frac{\epsilon^2}{2nT}\right)$ bound in Lemma 12 is better than the $2\exp\left(-\frac{\epsilon^2}{2n^2T}\right)$ bound obtained directly from the Azuma-Hoeffding inequality (Lemma 18).

▶ **Lemma 13.** *Consider a synchronous vertex process on an arbitrary graph. Let $Z(t) = n\sum_{v\in V}\pi_v X_v(t)$. Then, for any $T \ge 0$ and $\epsilon > 0$,*

$$\mathbb{P}\left[|Z(T) - Z(0)| \ge \epsilon\right] \le 2\exp\left(-\frac{\epsilon^2}{2n^2\|\pi\|_2^2 T}\right),$$

where $\|\pi\|_2 = \sqrt{\sum_{v\in V}\pi_v^2}$.

Proof. First, by Lemma 4 we observe that $(Z(t))_{t=0,1,2,\dots}$ is a martingale. From definition, we have $X_v(t+1) - X_v(t) \in \{-1, 0, 1\}$. Furthermore, for any $v \neq v'$, $X_v(t+1) - X_v(t)$ and $X_{v'}(t+1) - X_{v'}(t)$ are independent. Write $\Delta_v(t) = X_v(t) - X_v(t-1)$. Applying Lemma 19, we have

$$\begin{aligned}
\mathbf{E}\left[\mathrm{e}^{\lambda(Z(t+1)-Z(t))}\middle|X(t)\right] &= \mathbf{E}\left[\mathrm{e}^{\lambda((Z(t+1)-Z(t))-\mathbf{E}[Z(t+1)-Z(t)|X(t)])}\middle|X(t)\right] \\
&= \mathbf{E}\left[\mathrm{e}^{\lambda n \sum_{v\in V} \pi_v(\Delta_v(t+1)-\mathbf{E}[\Delta_v(t+1)|X(t)])}\middle|X(t)\right] \\
&= \prod_{v\in V} \mathbf{E}\left[\mathrm{e}^{\lambda n \pi_v(\Delta_v(t+1)-\mathbf{E}[\Delta_v(t+1)|X(t)])}\middle|X(t)\right] \\
&\leq \prod_{v\in V} \mathrm{e}^{\frac{4\lambda^2 n^2 \pi_v^2}{8}} = \mathrm{e}^{\frac{\lambda^2 n^2 \|\pi\|_2^2}{2}}. \qquad (18)
\end{aligned}$$

Combining (18) and Lemma 20, we obtain the claim. ◀

For regular graphs, both π_v and $\|\pi\|_2^2$ are $1/n$. So Lemma 13 generalizes Lemma 12.

5.3 At most three consecutive opinions remain

In this section, we suppose that $X_v(0) \in \{i-1, i, i+1\}$ holds for some i and for all $v \in V$, i.e., all initial opinions are from three consecutive integers. Without loss of generality, we assume that $i = 2$ throughout this section.

▶ **Lemma 14.** *Suppose that $X_v(0) \in \{1, 2, 3\}$ holds for all $v \in V$. Then, for any $t \geq 0$,*

$$\mathbf{E}[N_1(t+1)N_3(t+1) \mid X(t)] \leq \left(1 - \frac{N_1(t) + N_3(t)}{2n}\right) N_1(t)N_3(t).$$

Proof. Let $Y_{i\to j}$ denote the number of vertices that change their opinion from i to j. We have $N_1' = Y_{1\to 1} + Y_{2\to 1}$ and $N_3' = Y_{2\to 3} + Y_{3\to 3}$. Note that $Y_{3\to 1} = Y_{1\to 3} = 0$. It is easy to see that $Y_{1\to 1} \sim Bin(N_1, N_1/n)$ and $Y_{3\to 3} \sim Bin(N_3, N_3/n)$. An important observation is that $(Y_{2\to 1}, Y_{2\to 2}, Y_{2\to 3})$ follows a multinomial distribution with parameters N_2 and $(N_1/n, N_2/n, N_3/n)$. Hence, $Cov(Y_{2\to 1}, Y_{2\to 3}) \leq 0$ and we have $\mathbf{E}[Y_{2\to 1}, Y_{2\to 3}] \leq \mathbf{E}[Y_{2\to 1}]\mathbf{E}[Y_{2\to 3}]$. Thus,

$$\begin{aligned}
\mathbf{E}[N_1'N_3'] &= \mathbf{E}[Y_{1\to 1}(Y_{2\to 3} + Y_{3\to 3})] + \mathbf{E}[Y_{2\to 1}Y_{2\to 3}] + \mathbf{E}[Y_{2\to 1}Y_{3\to 3}] \\
&\leq \mathbf{E}[Y_{1\to 1}]\mathbf{E}[Y_{2\to 3} + Y_{3\to 3}] + \mathbf{E}[Y_{2\to 1}]\mathbf{E}[Y_{2\to 3}] + \mathbf{E}[Y_{2\to 1}]\mathbf{E}[Y_{3\to 3}] \\
&= (\mathbf{E}[Y_{1\to 1}] + \mathbf{E}[Y_{2\to 1}])(\mathbf{E}[Y_{2\to 3}] + \mathbf{E}[Y_{3\to 3}]) \\
&= \left(N_1\frac{N_1}{n} + N_2\frac{N_1}{n}\right)\left(N_2\frac{N_3}{n} + N_3\frac{N_3}{n}\right) \\
&= N_1\left(1 - \frac{N_3}{n}\right)N_3\left(1 - \frac{N_1}{n}\right). \qquad (19)
\end{aligned}$$

Note that $Y_{i\to j}$ and $Y_{k\to \ell}$ are independent for $i \neq k$. Combining (19) and the fact that $(1-x)(1-y) = 1 - x - y + xy \leq 1 - (x+y) + \frac{(x+y)^2}{2} \leq 1 - \frac{x+y}{2}$ holds for any $0 \leq x + y \leq 1$, we obtain the claim. ◀

Intuitively speaking, Lemma 14 implies that $N_1(t)N_3(t)$ continues to decrease by a factor of $1 - 1/\sqrt{n}$ while $N_1(t) + N_3(t) \geq 2\sqrt{n}$. Hence, within $T = O(\sqrt{n}\log n)$ steps, $N_1(t)N_3(t)$ reaches 0 or $N_1(t) + N_3(t) < 2\sqrt{n}$. In other words, either of the following events occurs: (1) either $N_1(t)$ or $N_3(t)$ is zero, (2) both $N_1(t)$ and $N_3(t)$ are less than $2\sqrt{n}$. The following lemma shows it formally.

▶ **Lemma 15.** *Suppose that $X_v(0) \in \{1,2,3\}$ holds for all $v \in V$. Let $T = \lceil 3\sqrt{n}\log n \rceil$. Then, for some $0 \le t \le T$, w.h.p. one of the following two events occurs.*
1. $N_1(t) = 0$ *or* $N_3(t) = 0$.
2. $N_1(t) \le 2\sqrt{n}$ *and* $N_3(t) \le 2\sqrt{n}$.

Proof. Let

$$\tau = \min\{t \ge 0 \mid N_1(t) + N_3(t) < 2\sqrt{n} \text{ or } N_1(t)N_3(t) = 0\},$$
$$Y_t = N_1(t)N_3(t)\left(1 - \frac{1}{\sqrt{n}}\right)^{-t}, \quad Z_t = Y_{t\wedge\tau} = Y_{\min\{t,\tau\}}.$$

Note that we have $Z_{t+1} - Z_t = \mathbb{1}_{\tau>t}(Y_{t+1} - Y_t)$. From Lemma 14,

$$\begin{aligned}
\mathbf{E}[Z_{t+1} - Z_t \mid X(t)] &= \mathbb{1}_{\tau>t}\left(\mathbf{E}[Y_{t+1} \mid X(t)] - Y_t\right) \\
&\le \mathbb{1}_{\tau>t}\left(\frac{\left(1 - \frac{N_1(t)+N_3(t)}{2n}\right)N_1(t)N_3(t)}{\left(1-\frac{1}{\sqrt{n}}\right)^{t+1}} - \frac{N_1(t)N_3(t)}{\left(1-\frac{1}{\sqrt{n}}\right)^{t}}\right) \\
&\le 0
\end{aligned}$$

holds, i.e., $\mathbf{E}[Z_{t+1}] \le \mathbf{E}[Z_t]$ and $\mathbf{E}[Z_T] \le \mathbf{E}[Z_0] = Z_0$. Let $T = \lceil 3\sqrt{n}\log n \rceil$. Then,

$$\mathbf{E}[Z_T \mid \tau > T]\mathbb{P}[\tau > T] \le \mathbf{E}[Z_T] \le Z_0 = N_1(0)N_3(0) \le n^2. \tag{20}$$

Furthermore,

$$\mathbf{E}[Z_T \mid \tau > T] = \mathbf{E}\left[N_1(T)N_3(T)\left(1 - \frac{1}{\sqrt{n}}\right)^{-T} \mid \tau > T\right] \ge \left(1 - \frac{1}{\sqrt{n}}\right)^{-3\sqrt{n}\log n} \ge n^3. \tag{21}$$

Note that the event $\tau > T$ implies that $N_1(T) + N_3(T) \ge 2\sqrt{n}$ and $N_1(T)N_3(T) \ge 1$. Combining Equations (20) and (21), we obtain $\mathbb{P}[\tau > T] \le 1/n$. ◀

Completing the proof of Theorem 1. If we have reached here at some step t, then at most three values $i-1, i, i+1$ remain, and one of the cases Lemma 15 (1) or Lemma 15 (2) holds. We next prove that w.h.p. the process will finish in $O(n)$ steps with the claimed results.

In either case, by Lemma 12 $S(t) = S(0)(1 + o(1))$. So if Lemma 15 (1) holds, there are two remaining values, say $i, i+1$, and we can use two-value pull voting with Theorem 5 directly.

However if Lemma 15 (2) holds, then there are three values $i-1, i, i+1$, where $|A_{i-1} \cup A_{i+1}| = O(n^{1/2})$. Thus $S(t)/n \sim i$, and we next prove that i wins w.h.p. by coupling the process with pull voting. For convenience let $\{i-1, i, i+1\} = \{1,2,3\}$, let $A_2 = A$ and $B = A_1 \cup A_3$. In pull voting value $i = 2$ wins with probability $|A|/n = 1 - o(1)$. In one step of synchronous pull voting, $|A'_P| = Bin(n, |A|/n)$.

There is a coupling between incremental voting on three values, and pull voting such that $|A'_I|$ stochastically dominates $|A'_P|$. Firstly the number of vertices which choose in $A = A_2$ directly is $Bin(n, |A|/n)$. Denote this set by $A' = A'_P$ and let $|A'_P| = X_P$. Given the set A'_P, a further non-negative number Y_I of vertices take the value $i = 2$ indirectly. The value of Y_I is a sum of binomials, namely

$$Y_I = Bin(|A_3 \setminus A'_P|, N_1/n) + Bin(|A_1 \setminus A'_P|, N_3/n).$$

It follows that under the coupling $|A'_I| = X_P + Y_I \ge X_P = |A'_P|$ and thus

$$\mathbb{P}(\text{Value } i \text{ wins in incremental voting}) \ge \mathbb{P}(\text{Value } i \text{ wins in pull voting}) = 1 - o(1).$$

6 Concluding comments

The incremental voting model offers an alternative type of pull voting suitable for discrete numeric opinions which can be compared on a linear scale. This may be appropriate for systems which need a very simple protocol which converges towards an average opinion. As the extremal values are discarded rapidly in some instances, it could also offer a faster alternative to remove outliers in some plurality systems.

The incremental voting process can be viewed as a form of discrete averaging of integer weights. The final answer is an integer (no fractions), obtained in finite expected time. For suitable expanders, w.h.p. the process returns the average rounded up or down to an integer. To increase the accuracy of the averaging, multiply all initial values by 10^h before averaging. The final answer, after re-scaling, will now be w.h.p. correct to the h-th decimal place. The cost is the increased convergence time.

In incremental voting, the weighted average remains a martingale under a wide range of conditions. Let P be any reversible transition matrix and π be its stationary distribution. Then, if the selected vertex v chooses u with probability $P(v,u)$, the random variable $W = \sum_{v \in V} \pi_v X_v$ is a martingale. As an example, if $P(u,u) = 1 - L_u$ and $P(u,v) = L_u/d(u)$, where $0 < L_u \leq 1$, then L_u can be viewed as the propensity for vertex u to change its opinion when selected in a given step. Here, $\pi(v) = d(v)/CL_v$, where $C = \sum L_v/d(v)$.

References

1 Mohammed Amin Abdullah and Moez Draief. Global majority consensus by local majority polling on graphs of a given degree sequence. *Discrete Applied Mathematics*, 180:1–10, 2015. doi:10.1016/J.DAM.2014.07.026.

2 David Aldous and James Allen Fill. Reversible markov chains and random walks on graphs. Unfinished monograph (recompiled version, 2014), 2002.

3 Luca Becchetti, Andrea Clementi, Emanuele Natale, Francesco Pasquale, and Riccardo Silvestri. Plurality consensus in the gossip model. In *Proceedings, 26th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 371–390, Philadelphia, PA, 2015. SIAM. doi:10.1137/1.9781611973730.27.

4 Luca Becchetti, Andrea Clementi, Emanuele Natale, Francesco Pasquale, Riccardo Silvestri, and Luca Trevisan. Simple dynamics for plurality consensus. In *Proceedings of the 26th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pages 247–256, New York, NY, USA, 2014. ACM. doi:10.1145/2612669.2612677.

5 Luca Becchetti, Andrea Clementi, Emanuele Natale, Francesco Pasquale, and Luca Trevisan. Stabilizing consensus with many opinions. In *Proceedings of the 27th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 620–635, Philadelphia, PA, 2016. SIAM. doi:10.1137/1.9781611974331.CH46.

6 Petra Berenbrink, George Giakkoupis, and Peter Kling. Tight bounds for coalescing-branching random walks on regular graphs. In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1715–1733, Philadelphia, PA, 2018. SIAM. doi:10.1137/1.9781611975031.112.

7 Siddhartha Brahma, Sandeep Macharla, Sudebkumar Prasant Pal, and Sudhir Kumar Singh. Fair leader election by randomized voting. In *Proceedings of the 1st international conference on Distributed Computing and Internet Technology (ICDCIT)*, pages 22–31. Springer, 2004. doi:10.1007/978-3-540-30555-2_4.

8 Colin Cooper, R. Elsässer, and Tomasz Radzik. The power of two choices in distributed voting. In *Proceedings of the 41st International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 435–446. Springer, 2014. doi:10.1007/978-3-662-43951-7_37.

9 Colin Cooper, Robert Elsässer, Hirotaka Ono, and Tomasz Radzik. Coalescing random walks and voting on graphs. In *Proceedings of the 31st Annual ACM Symposium on Principles of Distributed Computing (PODC)*, pages 47–56, New York, NY, USA, 2012. ACM. doi:10.1145/2332432.2332440.

10 Colin Cooper, Tomasz Radzik, Nicolás Rivera, and Takeharu Shiraga. Fast plurality consensus in regular expanders. In *Proceedings of the 31st International Symposium on Distributed Computing (DISC)*, volume 91, pages 13:1–13:16. Springer, 2017. doi:10.4230/LIPICS.DISC.2017.13.

11 Benjamin Doerr, Leslie Ann Goldberg, Lorenz Minder, Thomas Sauerwald, and Christian Scheideler. Stabilizing consensus with the power of two choices. In *Proceedings of the 23rd ACM symposium on Parallelism in algorithms and architectures (SPAA)*, pages 149–158, New York, NY, USA, 2011. ACM. doi:10.1145/1989493.1989516.

12 Mohsen Ghaffari and Johannes Lengler. Nearly-tight analysis for 2-choice and 3-majority consensus dynamics. In *Proceedings of the 2018 ACM Symposium on Principles of Distributed Computing (PODC)*, pages 305–313, New York, NY, USA, 2018. ACM. URL: https://dl.acm.org/citation.cfm?id=3212738.

13 Yehuda Hassin and David Peleg. Distributed probabilistic polling and applications to proportionate agreement. *Information and Computation*, 171(2):248–268, 2001. doi:10.1006/INCO.2001.3088.

14 Barry Johnson. *Design and analysis of fault tolerant digital systems.* Addison-Wesley, Boston, MA, USA, 1989.

15 Nan Kang and Nicolás Rivera. Best-of-three voting on dense graphs. In *Proceedings of the 31st ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pages 115–121, New York, NY, USA, 2019. ACM. doi:10.1145/3323165.3323207.

16 Toshio Nakata, Hiroshi Imahayashi, and Masafumi Yamashita. Probabilistic local majority voting for the agreement problem on finite graph. In *Proceedings of the 5th Annual International Computing and Combinatorics Conference (COCOON)*, pages 330–338. Springer, 1999. doi:10.1007/3-540-48686-0_33.

17 Roberto Imbuzeiro Oliveira and Yuval Peres. Random walks on graphs: new bounds on hitting, meeting, coalescing and returning. In *Proceedings of the 16th Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, pages 119–126, Philadelphia, PA, 2019. SIAM. doi:10.1137/1.9781611975505.13.

18 Nobutaka Shimizu and Takeharu Shiraga. Phase transitions of best-of-two and best-of-three on stochastic block models. *Random Structures and algorithms*, 59(1):96–140, 2021. doi:10.1002/RSA.20992.

A Proof of Lemma 8

We repeat the lemma for convenience.

▶ **Lemma 16** (Lemma 8). *Let $G \in G_{n,p}$, where $np \geq \log^{1+\varepsilon} n$ for some constant $\varepsilon > 0$. The following properties hold w.h.p..*

P1. (Almost regular graphs) *G is connected and all vertices v have degree $d(v) = np + O(\sqrt{np \log n})$ and stationary distribution $\pi_v = \frac{1}{n} + O\left(\frac{1}{n \log^{\varepsilon/2} n}\right)$.*

P2. (Large number of edges between large subsets of vertices)
Let $\delta \geq 5/\sqrt{np}$. For any pair of disjoint vertex sets A, B, with $|A| \geq \delta n$, $|B| \geq \delta n$, the number of edges X_{AB} between A and B satisfies $\mu/2 \leq X_{AB} \leq 3\mu/2$, where $\mu = |A||B|p$, the expected number of edges between the sets A and B in $G_{n,p}$.

P3. (Not too many edges within small subsets of vertices)

(i) *For $\omega \geq e$, $\omega \log \omega \leq np$, no vertex set S of size $s = n/\omega$ induces more than $X_S = e^2 s^2 p$ edges.*

(ii) *Let $d = np$. No set S, $|S| \le n/\omega$ induces more than $X_S = s\sqrt{4d\log(ne/s)}$ edges.*

(iii) *Provided $\omega = O(\log n)$ and $np = d \ge \log^{1+\theta} n$, the ratio $X_S/X_{S,V-S}$ is $O(1/\omega)$, the value achieved in P3.(i) above.*

Proof.

P1. An application of the Chernoff-Hoeffding inequality (Lemma 17) shows that for all vertices v, $d(v) = np + O(\sqrt{np\log n})$.

P2. for given disjoint A, B Let $|A| = an$, $|B| = bn$ then the Chernoff-Hoeffding inequality (Lemma 17.3) with $\varepsilon = 1/2$ implies

$$P_{AB} = \mathbb{P}(X_{AB} \notin [\mu/2, 3\mu/2]) \le 2e^{-\mu/12}.$$

We say that A, B is a *bad pair*, if $|A| \ge \delta n$ and $|B| \ge \delta n$ but $X_{AB} \notin [\mu/2, 3\mu/2]$. Then

$$\mathbf{E}(\text{number of bad pairs}) = \sum_{A,B} P_{AB} \le$$

$$\le 4^n\, 2e^{-\delta^2 n^2 p/12} \le 2\left(4e^{-\delta^2 np/12}\right)^n \le 2\left(4e^{-2}\right)^n = o(1).$$

P3. (i) Let X_S denote the number of edges induced by a set S, and $\mu = \mathbf{E}(X_S) = \binom{s}{2}p$ the expected number. By the Chernoff-Hoeffding inequality (Lemma 17.4), for $\alpha \ge e$ and $s \ge 3$,

$$P_S = \mathbb{P}(X_S \ge \alpha\mu) \le (e/\alpha)^{\alpha\mu} \le (e/\alpha)^{\alpha s^2 p/3}. \tag{22}$$

Say a set S of size s is a *bad set*, if it induces more than $e^2 s^2 p \ge e^2\mu$ edges. Then, using (22) with $\alpha = e^2$,

$$\begin{aligned}\mathbf{E}(\text{number of bad sets of size } s) &\le \binom{n}{s} e^{-e^2 s^2 p/3}\\ &\le \left(\frac{ne}{s} e^{-e^2 sp/3}\right)^s = \left(\exp\{-e^2 sp/3 + \log ne/s\}\right)^s\\ &\le \left(e^{-(e^2/3)\log\omega + \log\omega e}\right)^s \le \left(e^{-(e^2/3-2)\log\omega}\right)^s = o(1).\end{aligned}$$

The last inequality follows from the assumption that $\omega \ge e$. The size $s = n/\omega$ is minimized when $\omega\log\omega = np$, implying that $s \ge (\log\log n)/2$. Sum the above over all s greater than this minimum value to conclude that the expected number of bad sets of sizes in the required range is $o(1)$.
(ii) Let $\mu = \mathbf{E}X_{S,V-S} = s(n-s)p$. Then, as $n - s = n(1 - o(1))$, $\mu = sd(1 - o(1))$ and

$$P_{S,V-S} = \mathbb{P}(X_{S,V-S} \le (1-\varepsilon)\mu) \le e^{-\varepsilon^2 sd/3}.$$

Thus

$$\mathbf{E}(\text{number of bad sets } S) \le \binom{n}{s} P_{S,V-S} \le \left(\frac{ne}{s} e^{-\varepsilon^2 d/3}\right)^s = o(1),$$

provided $\varepsilon \ge \sqrt{\frac{4\log(ne/s)}{d}}$.

As the total degree of S is $sd(1 + 0(1))$, no such S can induce as many as

$$X_S = \varepsilon s(n-s)p \le \varepsilon sd \le s\sqrt{4d\log(ne/s)}$$

edges.

(iii) Thus for $s \le n/\omega$,

$$\max \frac{X_S}{X_{S,V-S}} = O\left(\frac{s\sqrt{d\log n/s}}{sd}\right) = O\left(\sqrt{\frac{\log n}{d}}\right) = O\left(\frac{1}{\log^\theta n}\right) = O\left(\frac{1}{\omega}\right),$$

provided $\omega = O(\log n)$, and $np = d \ge \log^{1+\theta} n$. ◀

B Tools used in the analysis

▶ **Lemma 17** (The Chernoff-Hoeffding inequalities)**.** *Let $X_1, \dots, X_n$ be n independent random variables taking values in $[0,1]$. Let $X = \sum_{i=1}^n X_i$. Let $\mu^- \le \mathbf{E}[X] \le \mu^+$. Then, we have the following:*

1. $\mathbb{P}[X \ge (1+\varepsilon)\mu^+] \le \exp\left(-\frac{\min\{\varepsilon^2,\varepsilon\}\mu^+}{3}\right)$, *for* $\varepsilon \ge 0$.
2. $\mathbb{P}[X \le (1-\varepsilon)\mu^-] \le \exp\left(-\frac{\varepsilon^2\mu^-}{2}\right)$, *for* $0 \le \varepsilon \le 1$.
3. $\mathbb{P}[X \notin (\,(1-\varepsilon)\mu^-, (1+\varepsilon)\mu^+\,)] \le 2\exp\left(-\frac{\varepsilon^2\mu^-}{3}\right)$, *for* $0 \le \varepsilon \le 1$.
4. $\mathbb{P}[X \ge \alpha\mu^+] \le \left(\frac{e^{\alpha-1}}{\alpha^\alpha}\right)^\mu$, *for* $\alpha \ge 1$.

▶ **Lemma 18** (The Azuma-Hoeffding inequality)**.** *Let $(X_t)_{t=0,1,2,\dots}$ be a martingale. Suppose $|X_i - X_{i-1}| \le c_i$ holds for any $i \ge 0$. Then, for any $T \ge 0$ and $\epsilon > 0$,*

$$\mathbb{P}[|X_T - X_0| \ge \epsilon] \le 2\exp\left(-\frac{\epsilon^2}{2\sum_{i=1}^T c_i^2}\right).$$

The followings are the basic technical lemmas for the Hoeffding inequality.

▶ **Lemma 19.** *Let X be a random variable such that $\mathbf{E}[X] = 0$ and $a \le X \le b$. Then, for any $\lambda > 0$, $\mathbf{E}\left[\mathrm{e}^{\lambda X}\right] \le \mathrm{e}^{\lambda^2(b-a)^2/8}$.*

▶ **Lemma 20.** *For any $\alpha > 0$ and t, suppose that $\mathbf{E}\left[\mathrm{e}^{\alpha(Y_t - Y_{t-1})} \mid \mathcal{F}_{t-1}\right] \le \mathrm{e}^{\alpha^2 c_t^2}$ holds for some c_t. Then, for any $\epsilon > 0$, $\mathbb{P}[|Y_T - Y_0| \ge \epsilon] \le 2\exp\left(-\frac{\epsilon^2}{4\sum_{t=1}^T c_t^2}\right)$.*