MAJORITY DYNAMICS ON RESAMPLED SPARSE ERDŐS–RÉNYI GRAPHS: GAUSSIAN WINNER SELECTION IN COLLABORATIVE AGENT NETWORKS

Anonymous authors
Paper under double-blind review

ABSTRACT

Collective decisions in agent networks emerge from repeated local interactions, as agents update their opinions (colors) based on their neighbors, potentially leading to consensus. In sparse communication graphs, repeated interactions can amplify both the initial majority preference and the random fluctuations caused by individual agents. Understanding when the initial majority determines the final preference and how quickly an agreement emerges is, therefore, a basic question in collective decision-making. We study binary majority dynamics as an idealized model of this process. At each round, the interaction graph is independently resampled from the Erdős–Rényi model $\mathbb{G}(N,p)$, with $p = b\log N/N$ and fixed $b > 1$. Each agent adopts the majority opinion (color) among its neighbors, retaining its current opinion if there is a tie. For a fixed initial configuration, let Δ_0 denote the initial difference between the number of blue agents and the number of red agents. We identify three regimes for winner selection and consensus time, determined by the size of Δ_0. First, in the critical window $\Delta_0\sqrt{p} = O(1)$, the probability of blue unanimity is

$$\Phi\left(\Delta_0\sqrt{2p/\pi}\right) + o(1),$$

where Φ is the cumulative distribution function of a standard Gaussian random variable. Moreover, consensus is reached after $(1+o(1))\log N/\log\log N$ rounds with high probability. Second, in the intermediate regime $1 \ll \Delta_0\sqrt{p} \lesssim \sqrt{N}$, we establish explicit high probability upper and lower bounds on the time to blue-unanimity. Lastly, when the initial blue advantage is above an explicit constant multiple of $N/\sqrt{\log N}$, blue unanimity is guaranteed within two rounds with high probability. Our findings provide a complete description of how the initial advantage influences winner selection and the speed of consensus in this type of majority dynamics with sparse interactions.

1 INTRODUCTION

In a large collaboration network, agents exchange opinions repeatedly through local interactions until they reach a collective decision. Due to communication constraints, the number of contacts with which an agent can interact each round may be limited when compared to the entire population, leading to a *sparse* interaction graph. Hence, the opinion favored in its neighborhood may differ from the opinion favored globally. While repeated interactions can reinforce the initial population preference, they may also amplify random fluctuations or propagate contrarian opinions. Recent work on the collective behavior of AI agents motivates studying these effects through models of interacting opinions (El et al., 2026). These considerations raise the following two questions regarding communication dynamics and the correctness of information propagation.

Question 1 (Consensus)**.** Starting from any fixed configuration, under what conditions is the initial preference preserved, and how quickly is consensus reached?

Question 2 (Correctness)**.** When initial opinions are random choices between two options, slightly in favor of the correct one, how likely is it that the final consensus is aligned with the correct choice?

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

We explore answers to both questions by studying *majority dynamics*, a fundamental model for opinion evolution introduced by Mossel & Tamuz (2017). Each agent holds one of two possible opinions, *blue* or *red*. Let $\mathcal{B}_t$ and $\mathcal{R}_t$ denote the sets of blue and red agents after round $t \geq 0$, respectively, and $\Delta_0 = |\mathcal{B}_0| - |\mathcal{R}_0|$ denote the initial blue advantage. In each round, all agents update their opinions synchronously by following the majority opinion among their neighbors, retaining their current opinion if there is a tie. We say that *blue unanimity* is achieved after round T if $\mathcal{B}_T = \mathcal{V}$. For the sake of theoretical analysis, we assume that the interaction graph $\mathcal{G}$ is sampled from the Erdős–Rényi model $\mathbb{G}(N, p)$, where each pair of distinct agents is independently connected with probability p. We say that $\mathcal{G}$ is a *dense* graph when $p \in (0, 1)$ is a fixed constant; in contrast, it *is sparse* when $p = o(1)$.

Previous literature largely focuses on the scenario that $\mathcal{G}$ is *dense* and *static*. Tran & Vu (2023) proved that, with high probability, an initial blue advantage of at least 6 agents ($\Delta_0 \geq 6$) leads to blue unanimity after 4 rounds. Berkowitz & Devlin (2022) established a central limit theorem for the number of agents of each opinion after one round and improved the result of Tran & Vu (2023) by decreasing the necessary initial blue advantage to $\Delta_0 \geq 3$. Sah & Sawhney (2025) established an exact winner-selection law and proved that any nonzero initial blue advantage ($\Delta_0 \geq 1$) will lead to blue unanimity after 3 rounds. For the sparse case where $pN \asymp \log(N)$ but $\mathcal{G}$ remains fixed, Tran & Vu (2025) proved that an initial blue advantage satisfying $p\Delta_0 \geq 10$ will lead to blue unanimity after $\log(N)/\log\log(N)$ rounds. They conjectured that, similar to the dense case, a winner-selection law can be established when $\Delta_0\sqrt{p} \asymp 1$. There has been progress toward resolving this conjecture for sparse, static interaction graphs (Kim & Tran, 2025; Jaffe, 2025; Goel & Sah, 2026), but it has not yet been fully solved. Readers may refer to Section G for a more detailed literature review.

Despite the remarkable progress noted above, we note that two major constraints remain. First, in large-scale collaborative networks, each agent is restricted to interacting with only a vanishing fraction of the overall population due to limited attention and communication capacity. This motivates an understanding of consensus formation under sparse communications. Second, the static model restricts the potential neighbors of each agent throughout the evolution. The assumption of the same contacts every day is somewhat unrealistic in many communication settings, as interaction partners typically change over time, especially in scenarios where opinions evolve dynamically.

Our contributions. Suppose that the interaction graph $\mathcal{G}$ is independently resampled from $\mathbb{G}(N, p)$ every round. We focus on the *critically sparse* regime[1] where $pN \asymp \log(N)$ and each agent interacts with $\log(N)$ many peers per round. Our main results can be summarized as follows.

1. Gaussian winner selection (Theorem 3.4). When $\Delta_0\sqrt{p} = O(1)$, the probability of blue unanimity is $\Phi(\Delta_0\sqrt{2p/\pi}) + o(1)$. The consensus time is $\log N/\log\log N$, regardless of the final winner.

2. Consensus time and majority amplification (Theorem 3.2). When $\Delta_0\sqrt{p} \gg 1$, blue wins with high probability. We also provide explicit upper and lower bounds on the consensus time, based on a sharp two-sided amplification estimate for blue advantage per round, which is typically of order $\sqrt{\log N}$.

3. Rapid consensus (Theorem 3.1). When the initial advantage Δ_0 is above an explicit constant multiple of $N/\sqrt{\log N}$, blue unanimity is reached within two rounds with high probability. If a further lower bound on Δ_0 is satisfied, blue unanimity is achieved after one round. The critical thresholds for two- and one round unanimity are explicitly characterized.

4. Correctness from random votes (Corollary 3.6). When the initial opinions are random votes on colors, each correct with probability $1/2 + \theta/\sqrt{N}$ for some fixed $\theta \geq 0$, the correct unanimity occurs with probability $\Phi(2\theta) + o(1)$. This matches the asymptotic accuracy of centralized majority voting on the initial opinions, with consensus time again asymptotic to $\log N/\log\log N$.

Organization. The rest of the paper is organized as follows. We introduce the mathematical definitions of our model in Section 2, Our main results are presented in Section 3. Numerical experiments that complement our theoretical findings are provided in Section 4 with a technical overview in Subsection 3.5. Some possible future directions are discussed in Section 5. Detailed proofs of the main results are deferred to the Appendix.

[1]The $\mathbb{G}(N, p)$ is connected if and only if $pN > \log(N)$,

2 MODEL AND PROBLEM FORMULATION

We introduce the precise definitions of the model and the problem here.

2.1 RESAMPLED MAJORITY DYNAMICS

Let $\mathcal{V} = [N] = \{1, \dots, N\}$ and $\mathbf{y}_t \in \{\pm 1\}^{\mathcal{V}}$ record the opinions after round $t \geq 0$, with blue represented by $+1$ and red by -1. Define the two camps by

$$\mathcal{B}_t = \{v : \mathbf{y}_t(v) = +1\}, \qquad \mathcal{R}_t = \{v : \mathbf{y}_t(v) = -1\}. \tag{2.1}$$

At round $t \geq 0$, an undirected graph $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t)$ is sampled from $\mathbb{G}(N, p)$, independent of the initial opinions and all previous graphs.

Assumption 2.1 (Critical sparsity)**.** Suppose $p = d/N$ with $d = b\log(N)$ for some constant $b > 1$.

For each $v \in \mathcal{V}$, we define its blue and red neighbors at round $t \geq 0$ by

$$N_t^{\mathcal{B}}(v) = \#\{w \in \mathcal{B}_t : \{v, w\} \in \mathcal{E}_{t+1}\}, \qquad N_t^{\mathcal{R}}(v) = \#\{w \in \mathcal{R}_t : \{v, w\} \in \mathcal{E}_{t+1}\}. \tag{2.2}$$

All opinions update simultaneously, with ties preserving the previous opinion:

$$\mathrm{y}_{t+1}(v) = \mathbb{1}_{\{N_t^{\mathcal{B}}(v) > N_t^{\mathcal{R}}(v)\}} - \mathbb{1}_{\{N_t^{\mathcal{B}}(v) < N_t^{\mathcal{R}}(v)\}} + \mathrm{y}_t(v) \cdot \mathbb{1}_{\{N_t^{\mathcal{B}}(v) = N_t^{\mathcal{R}}(v)\}}. \tag{2.3}$$

The population advantage satisfies

$$\Delta_t = |\mathcal{B}_t| - |\mathcal{R}_t|, \qquad |\mathcal{B}_t| = (N + \Delta_t)/2, \qquad |\mathcal{R}_t| = (N - \Delta_t)/2. \tag{2.4}$$

We define the blue and red unanimity times by

$$\tau_{\mathrm{B}} = \inf\{t \geq 0 : \mathcal{R}_t = \emptyset\}, \qquad \tau_{\mathrm{R}} = \inf\{t \geq 0 : \mathcal{B}_t = \emptyset\}, \tag{2.5}$$

with the convention $\inf \emptyset = \infty$. We denote the consensus time by $T = \min\{\tau_{\mathrm{B}}, \tau_{\mathrm{R}}\}$. The opinion process is a time-homogeneous Markov chain whose next camp sizes depend in distribution only on the current camp sizes. Its natural filtration is

$$\mathcal{F}_0 = \sigma(\mathbf{y}_0), \qquad \mathcal{F}_t = \sigma(\mathbf{y}_0, \mathcal{G}_1, \dots, \mathcal{G}_t), \quad t \geq 1. \tag{2.6}$$

Conditional on $\mathcal{F}_t$, the next graph is independent of the past and has law $\mathbb{G}(N, p)$.

2.2 INITIAL OPINION MODELS

Fixed initial configuration. For each N, fix any deterministic configuration $\mathbf{y}_0 \in \{\pm 1\}^{\mathcal{V}}$. We are interested in the probability of blue unanimity $\mathbb{P}(\tau_{\mathrm{B}} < \infty \mid \mathbf{y}_0)$, its red counterpart, and the distribution of T. These quantities describe the agreement and preservation of an initial majority; no opinion is designated as correct in this model.

Random Initial Vote Model. Fix a correct opinion $z \in \{-1, +1\}$ and a constant $\theta \geq 0$. Independently, over vertices and all interaction graphs, we draw the initial opinions according to

$$\mathbb{P}(\mathbf{y}_0(v) = z) = 1/2 + \theta/\sqrt{N}, \qquad v \in [N], \tag{2.7}$$

The correct opinion z is used to evaluate the outcome and is not supplied to the update rule. The votes are drawn once; subsequent rounds use only the shared majority dynamics. We ask for the probability of correct unanimity, $\mathbb{P}(\exists t \geq 0 : \mathbf{y}_t \equiv z)$, and the consensus time T. Here, probabilities include both the initial votes and the interaction graphs. The benchmark is centralized majority voting on the initial opinions, with a rule for breaking ties.

3 MAIN RESULTS

We first present our main results, followed by a technical overview of the proof.

3.1 RAPID UNANIMITY: QUANTITATIVE BOUNDS

Define the following constants

$$r_* = \big(1 - b^{-1}\sqrt{2b-1}\big)/2, \qquad K_2^{\mathrm{ER}}(b) = b^{-1/2} \cdot \Phi^{-1}(1 - r_*). \tag{3.1}$$

Here, $r_* \in (0, 1/2)$ is the solution to the equation $b(\sqrt{1-r} - \sqrt{r})^2 = 1$. Below, we provide sufficient conditions for rapid unanimity.

Theorem 3.1 (Two- and one-day unanimity)**.** *Under Assumption 2.1, the following statements hold for all sufficiently large N.*

(i) (Two days.) *Fix $K > K_2^{\mathrm{ER}}(b)$. There exists $\xi = \xi(b, K) > 0$ such that $\mathcal{R}_2 = \emptyset$ with probability at least $1 - 2\exp(-(\log N)^2) - 2N^{-\xi}$ whenever*

$$\Delta_0 \geq KN/\sqrt{\log N}, \tag{3.2}$$

(ii) (One day.) *Fix $0 < r < r_*$. If*

$$\Delta_0 \geq (1 - 2r)N, \tag{3.3}$$

then there exists $\xi = \xi(b, r) > 0$ such that $\mathcal{R}_1 = \emptyset$ with probability at least $1 - 2N^{-\xi}$.

The proof is in Appendix C. Intuitively, a large advantage first reduces the minority to a sufficiently small fraction; the next independent graph removes the remaining minority with high probability.

3.2 AMPLIFICATION AND THE UNANIMITY-TIME BRACKET

An initial advantage above the critical scale dominates the first-round fluctuations. The next theorem quantifies its subsequent amplification and brackets the blue-unanimity time τ_{B}.

Fix $K > K_2^{\mathrm{ER}}(b)$, and let $c_1 = c_1(b, K) > 0$, $C_2 = C_2(b, K) > 0$, and $c = c(b, K) > 0$ be the constants supplied by Lemma B.1 with upper-window coefficient K. Let $\xi = \xi(b, K) > 0$ be the constant in Theorem 3.1. For any $s \geq 0$, define

$$\eta_N(s) := 8\exp\left(-cs^2\log(N)/N\right) + 2\exp\left(-(\log N)^2\right) + 2N^{-\xi}. \tag{3.4}$$

For $\Delta_0 > 0$, define the deterministic horizons

$$\underline{T} := \max\left\{0, \left\lfloor \log\left(KN/(\Delta_0\sqrt{\log N})\right) / \log\left(C_2\sqrt{\log N}\right)\right\rfloor\right\}, \tag{3.5a}$$

$$\overline{T} := \max\left\{0, \left\lceil \log\left(KN/(\Delta_0\sqrt{\log N})\right) / \log\left(c_1\sqrt{\log N}\right)\right\rceil\right\}. \tag{3.5b}$$

Theorem 3.2 (Logarithmic time to blue unanimity)**.** *Under Assumption 2.1, suppose that $\mathcal{R}_0 \neq \emptyset$, $1 \leq h_N = o(\sqrt{\log N})$, and*

$$\Delta_0 \geq \sqrt{N}/h_N. \tag{3.6}$$

Then, for all sufficiently large N,

$$\mathbb{P}\left(\underline{T} + 1 \leq \tau_{\mathrm{B}} \leq \overline{T} + 2 \,\middle|\, \mathbf{y}_0\right) \geq 1 - \eta_N(\Delta_0). \tag{3.7}$$

Corollary 3.3. *Under the conditions of Theorem 3.2, we further suppose that*

$$\Delta_0 \asymp \sqrt{N}/h_N.$$

Then for every fixed $0 < \delta < 1$ and all sufficiently large N,

$$\mathbb{P}\left(\mathcal{R}_{\lfloor(1-\delta)\log N/\log\log N\rfloor} \neq \emptyset,\ \mathcal{R}_{\lceil(1+\delta)\log N/\log\log N\rceil} = \emptyset \,\middle|\, \mathbf{y}_0\right) \geq 1 - \eta_N(\Delta_0). \tag{3.8}$$

The theorem and corollary are proved in Appendix B. The theorem's initial-advantage condition is equivalent to $\Delta_0\sqrt{p} \to \infty$: for any such sequence, one may choose $h_N = \max\{1, \sqrt{N}/\Delta_0\} = o(\sqrt{\log N})$. In this regime, $\eta_N(\Delta_0) = o(1)$, so blue wins with high probability. The lower and upper time bounds describe the fastest and slowest amplification rates allowed by our estimates before the process reaches $KN/\sqrt{\log N}$, followed by at most two completion rounds. In the corollary's narrower regime, $\tau_{\mathrm{B}}/(\log N/\log\log N) \to 1$ in probability.

3.3 CRITICAL-WINDOW WINNER SELECTION: UNIFORM ERRORS AND HORIZON

At the critical scale $\Delta_0 = O(\sqrt{N/\log N})$, the first update's mean and fluctuations have comparable magnitudes. The normalized advantage $\Delta_1/\sqrt{N}$ has a nondegenerate Gaussian limit (Lemma A.2), whose sign determines the eventual winner. Outside a vanishing window around zero, its magnitude enters the amplification regime of Theorem 3.2.

To state the result, retain the fixed K and associated constants introduced above, and let h_N satisfy the following conditions:

$$h_N \to \infty \quad \text{and} \quad h_N = o(\sqrt{\log N}).$$

For $C_0 > 0$, the total error, combining the one-step Gaussian approximation, the probability of the vanishing central window, and the subsequent amplification failure probability, is

$$\mathfrak{r}_N(C_0) := \rho_N(C_0) + 1/h_N + \eta_N(\sqrt{N}/h_N), \tag{3.9}$$

where $\rho_N(C_0)$ is the uniform one-step Gaussian approximation error defined in (A.4). The corresponding worst-case amplification horizon from the threshold $\sqrt{N}/h_N$ is

$$T_N := \max\left\{0, \left\lceil \log\left(Kh_N\sqrt{N/\log N}\right)/\log\left(c_1\sqrt{\log N}\right)\right\rceil\right\}. \tag{3.10}$$

Theorem 3.4 (Critical-window winner selection)**.** *Under Assumption 2.1, retain the preceding choices of K and h_N, and fix $C_0 > 0$. Consider deterministic initial configurations satisfying*

$$x_0 := \Delta_0\sqrt{2b\log N/\pi N}, \qquad |x_0| \le C_0. \tag{3.11}$$

Uniformly over all such configurations, the blue- and red-unanimity probabilities at $T_N + 3$ satisfy

$$\mathbb{P}\left(\mathcal{R}_{T_N+3} = \emptyset \,|\, \mathbf{y}_0\right) = \Phi(x_0) + O\left(\mathfrak{r}_N(C_0)\right), \tag{3.12a}$$
$$\mathbb{P}\left(\mathcal{B}_{T_N+3} = \emptyset \,|\, \mathbf{y}_0\right) = \Phi(-x_0) + O\left(\mathfrak{r}_N(C_0)\right). \tag{3.12b}$$

In particular, $\mathfrak{r}_N(C_0) = o(1)$, and the time horizon satisfies $T_N + 3 = (1 + o(1))\log N/\log\log N$.

Since the two unanimous states are absorbing and the probability that neither has been reached by $T_N + 3$ is $O(\mathfrak{r}_N(C_0))$, the theorem also gives the eventual blue- and red-winning probabilities $\Phi(x_0) + O(\mathfrak{r}_N(C_0))$ and $\Phi(-x_0) + O(\mathfrak{r}_N(C_0))$, respectively. Thus, if $x_0 \to x > 0$, the initial blue majority loses with a limiting probability of $\Phi(-x) > 0$; at balance, each opinion wins with a probability of $1/2 + o(1)$. The following corollary shows that the same leading consensus-time scale holds throughout the critical window.

Corollary 3.5 (Critical-window consensus time)**.** *Under Assumption 2.1, fix $C_0 > 0$. For every fixed $0 < \varepsilon < 1$, uniformly over deterministic initial configurations satisfying (3.11),*

$$\mathbb{P}\left((1-\varepsilon)\log N/\log\log N \le \min\{\tau_B, \tau_R\} \le (1+\varepsilon)\log N/\log\log N \,|\, \mathbf{y}_0\right) = 1 - o(1). \tag{3.13}$$

The theorem and consensus-time corollary are proved in Appendix A. Figure 1 illustrates the agreement regimes for fixed initial configurations.

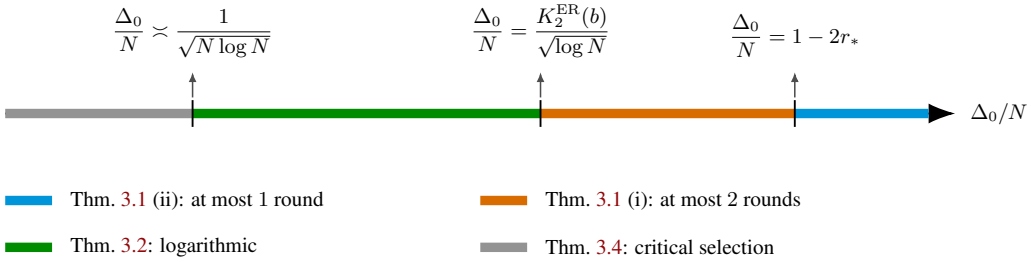


Figure 1: Schematic agreement regimes for the resampled ER model with fixed $b > 1$ and $\Delta_0 \ge 0$; exchanging blue and red gives the case $\Delta_0 < 0$. Horizontal lengths are illustrative. The markers show the critical Gaussian scale and the constants in the sufficient two- and one-round conditions. The latter guarantees require fixed margins $K > K_2^{\mathrm{ER}}(b)$ and $r < r_*$.

3.4 Correctness under random initial votes

We now evaluate agreement against the correct opinion in the random initial vote model. With high probability, the initial majority has an advantage in the amplification regime. Applying the agreement bounds conditionally on those votes and then averaging gives the following correctness guaranty.

Corollary 3.6 (Correctness from random initial votes)**.** *Under Assumption 2.1, consider the random initial vote model of Section 2.2, with a fixed correct opinion z and parameter $\theta \geq 0$. Then*

$$\mathbb{P}\big(\exists t \geq 0 : \mathbf{y}_t \equiv z\big) = \Phi(2\theta) + o(1). \tag{3.14}$$

Moreover, (3.13) holds as well under the current random initial vote model.

3.5 Proof Overview and Technical Contributions

We summarize the proof outline here. Readers may refer to Appendices A–F for the complete proof.

1. The first update selects a sign. In the critical window, the mean estimate and the one-step central limit theorem in Berkowitz & Devlin (2022) lead to

$$\sup_{z\in\mathbb{R}} \left| \mathbb{P}\left(\Delta_1/\sqrt{N} \leq z \mid \mathbf{y}_0\right) - \Phi(z - x_0) \right| = o(1), \tag{3.15}$$

uniformly for bounded x_0. Thus a critical initial advantage of order $\sqrt{N/\log N}$ becomes a random advantage of order $\sqrt{N}$. The Gaussian law supplies both the sign probabilities and anti-concentration: for every $h_N \to \infty$, $\mathbb{P}(|\Delta_1| < \sqrt{N}/h_N \mid \mathbf{y}_0) = o(1)$. Choosing such a sequence with $h_N = o(\sqrt{\log N})$ ensures that $|\Delta_1|\sqrt{p} \geq \sqrt{b \log N}/h_N \to \infty$ outside this event. We therefore apply Theorem 3.2 starting from the configuration after round one, interchanging blue and red if $\Delta_1 < 0$. We use the CLT of Berkowitz & Devlin (2022) to obtain the first-round Gaussian approximation with an $o(1)$ error. For subsequent rounds, we use concentration inequalities to obtain explicit bounds on the probability that amplification fails. For $h_N = (\log N)^{1/4}$, Gaussian tightness also gives $|\Delta_1| \leq h_N\sqrt{N}$ with high probability. Applying both sides of the amplification bracket uniformly between these endpoints yields the consensus-time asymptotic in Corollary 3.5.

2. Sparse drift requires a local comparison. Let $p_t^{\mathcal{R}}$ and $p_t^{\mathcal{B}}$ denote the conditional probabilities of red-to-blue and blue-to-red flips. The exact drift identity is

$$\mathbb{E}[\Delta_{t+1} \mid \mathbf{y}_t] = N(p_t^{\mathcal{R}} - p_t^{\mathcal{B}}) + \Delta_t(1 - p_t^{\mathcal{R}} - p_t^{\mathcal{B}}). \tag{3.16}$$

Approximating each flip probability separately by a Gaussian and subtracting is insufficient near balance: the additive errors can exceed the difference being estimated. We instead express this difference as a sum of local probabilities.

Condition on $\mathbf{y}_t$, and suppose that $\Delta_t \geq 0$ are both nonempty camps. For each integer $0 \leq j \leq \Delta_t$, introduce auxiliary variables with distributions

$$\mathrm{D}_{t,j}^{\mathcal{R}} \stackrel{\mathrm{d}}{=} \mathbf{Bin}(|\mathcal{R}_t| + j, p) - \mathbf{Bin}(|\mathcal{R}_t| - 1, p),$$
$$\mathrm{D}_{t,j}^{\mathcal{B}} \stackrel{\mathrm{d}}{=} \mathbf{Bin}(|\mathcal{R}_t|, p) - \mathbf{Bin}(|\mathcal{R}_t| - 1 + j, p),$$

where the two binomial variables in each difference are independent. At $j = 0$, the two differences have the same distribution. At $j = \Delta_t$, their positive-tail probabilities are $p_t^{\mathcal{R}}$ and $p_t^{\mathcal{B}}$, respectively. Increasing j by one adds an independent $\mathbf{Ber}(p)$ trial to the first difference and subtracts one from the second. Their positive-tail probabilities, therefore, change by $p\mathbb{P}(\mathrm{D}_{t,j}^{\mathcal{R}} = 0)$ and $-p\mathbb{P}(\mathrm{D}_{t,j}^{\mathcal{B}} = 1)$. Telescoping gives the exact identity

$$p_t^{\mathcal{R}} - p_t^{\mathcal{B}} = p \sum_{j=0}^{\Delta_t - 1} \left[\mathbb{P}(\mathrm{D}_{t,j}^{\mathcal{R}} = 0) + \mathbb{P}(\mathrm{D}_{t,j}^{\mathcal{B}} = 1)\right], \tag{3.17}$$

where the probabilities refer to distributions conditional on $\mathbf{y}_t$, and the sum is zero when $\Delta_t = 0$.

A Poisson coupling and a local Gaussian estimate control these probabilities. Uniformly for $0 \leq \Delta_t \leq KN/\sqrt{\log N}$ with fixed K, each local approximation has error $O((\log N)^{-1})$. Its contribution

after summation in (3.17) is therefore $O(p\Delta_t/\log N)$; see Lemma D.4. In particular, for every fixed $C_0 > 0$, uniformly over initial configurations with $|x_0| \le C_0$,

$$p_0^{\mathcal{R}} - p_0^{\mathcal{B}} = \sqrt{\frac{2}{\pi}}\,\frac{p\Delta_0}{\sqrt{(N-1)p(1-p)}} + O\left(\frac{p|\Delta_0|}{\log N}\right). \tag{3.18}$$

For negative advantages the formula follows by color symmetry. Substitution into (3.16) identifies the critical mean $\mathbb{E}[\Delta_1 \mid \mathbf{y}_0]/\sqrt{N} = x_0 + o(1)$. More generally, throughout a fixed window $\sqrt{N/\log N} \le \Delta_t \le KN/\sqrt{\log N}$, the conditional mean is bounded above and below by positive constant multiples of $\sqrt{\log N}\,\Delta_t$.

3. Dependent updates concentrate and amplify. Conditional on $\mathbf{y}_t$, the indicator that a vertex is red after the next update is a function of its incident independent edges. Each edge enters only the two endpoint indicators, so these indicators form a read-2 family. The read-2 concentration bound (Gavinsky et al., 2015) gives, for every $s > 0$,

$$\mathbb{P}(||\mathcal{R}_{t+1}| - \mathbb{E}[|\mathcal{R}_{t+1}| \mid \mathbf{y}_t]| \ge s \mid \mathbf{y}_t) \le 2\exp\bigl(-s^2/N\bigr).$$

Since $\Delta_{t+1} = N - 2|\mathcal{R}_{t+1}|$, setting $\mu_t = \mathbb{E}[\Delta_{t+1} \mid \mathbf{y}_t]$ yields

$$\mathbb{P}(|\Delta_{t+1} - \mu_t| \ge u \mid \mathbf{y}_t) \le 2\exp\bigl(-u^2/4N\bigr), \qquad u > 0. \tag{3.19}$$

Combining (3.19) with the conditional-mean bounds above, and taking $u \asymp \sqrt{\log N}\,\Delta_t$, gives that for suitable constants $c_1, C_2, c > 0$ depending only on b, K, we have

$$\mathbb{P}\bigl(c_1\sqrt{\log N}\,\Delta_t \le \Delta_{t+1} \le C_2\sqrt{\log N}\,\Delta_t \mid \mathbf{y}_t\bigr) \ge 1 - 2\exp\bigl(-c\Delta_t^2\log(N)/N\bigr) \tag{3.20}$$

throughout $\sqrt{N/\log N} \le \Delta_t \le KN/\sqrt{\log N}$, The failure probability vanishes when the advantage is asymptotically above the lower endpoint of this window. Along successful steps, the advantage grows geometrically, making the later failure probabilities summable. Iterating the two bounds yields the deterministic time bracket in Theorem 3.2.

4. Fresh communication completes the decision. We first identify a minority fraction that disappears after one round. Suppose that $|\mathcal{R}_t| \le rN$ for a fixed $0 < r < 1/2$, and define

$$J_b(r) = b\bigl(\sqrt{1-r} - \sqrt{r}\bigr)^2.$$

For each agent, the probability of being red after the next update is nondecreasing in the current red-camp size. It therefore suffices to consider $|\mathcal{R}_t| = \lfloor rN \rfloor$. The one-vertex large-deviation estimate bounds this probability by $N^{-J_b(r)+o(1)}$. A union bound over vertices gives

$$\mathbb{P}(\mathcal{R}_{t+1} \ne \emptyset \mid \mathbf{y}_t) \le N^{1-J_b(r)+o(1)}, \tag{3.21}$$

uniformly over configurations with $|\mathcal{R}_t| \le rN$. The right-hand side vanishes when $J_b(r) > 1$, equivalently $r < r_*$. Thus $J_b(r_*) = 1$ identifies where the exponent in (3.21) changes sign, yielding the sufficient condition $r < r_*$ in the one-round statement of Theorem 3.1; the quantitative failure probability bound is given in Lemma C.1.

To reach this small-minority regime, suppose that $\Delta_t \ge KN/\sqrt{\log N}$ for fixed $K > 0$. At $\Delta_t = KN/\sqrt{\log N} + O(1)$, a Gaussian estimate gives an expected red fraction $\Phi(-\sqrt{b}\,K) + o(1)$. Monotonicity extends this to all larger advantages:

$$\mathbb{E}[|\mathcal{R}_{t+1}| \mid \mathbf{y}_t]/N \le q_K + o(1), \qquad q_K = \Phi(-\sqrt{b}\,K). \tag{3.22}$$

The condition $K > K_2^{\mathrm{ER}}(b)$ is precisely $q_K < r_*$. Choose a fixed r with $q_K < r < r_*$. The concentration bound above then gives $|\mathcal{R}_{t+1}| \le rN$ with high probability, as quantified in Lemma C.2. Applying (3.21) to the next independent graph completes blue unanimity in one further round.

Independence of future graphs allows the amplification and completion estimates to be applied conditionally to the random configuration selected after round one. Combining this with (3.15) transfers the Gaussian sign probability to the eventual winner and proves Theorem 3.4.

4 Numerical Experiments

We simulate the majority dynamics with independently resampled ER graphs. The experiments examine Gaussian winner selection, amplification of the population advantage, and rapid unanimity. All runs use $N = 10^4$, $p = b \log N/N$, and deterministic, parity-admissible initial camp sizes.

Gaussian winner selection. For each $b \in \{1.5, 2, 4\}$, we use 13 equally spaced target values of x_0 in $[-3, 3]$. We round $\Delta_0 = \sqrt{\pi N/(2b \log N)}\, x_0$ to the nearest integer satisfying $\Delta_0 \equiv N \pmod 2$ and plot the resulting actual value of x_0. Each setting uses 200 independent trajectories, simulated until either opinion reaches unanimity or through round 100. Figure 2 compares these empirical probabilities with $\Phi(x_0)$. Among the three densities, the curves track the Gaussian prediction, with a crossover near balance, supporting the normalization in Theorem 3.4.

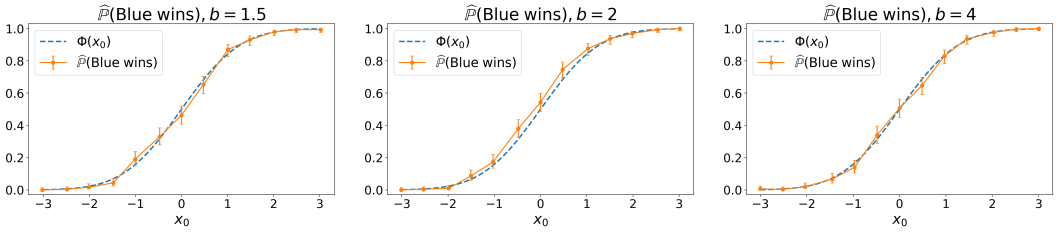


Figure 2: Gaussian winner selection for $b = 1.5$ (left), $b = 2$ (middle), and $b = 4$ (right), with 200 independent paths per setting. Solid curves show empirical probabilities of blue unanimity by round 100, and dashed curves show $\Phi(x_0)$. The pointwise 95% Wilson intervals quantify Monte Carlo uncertainty at the simulated population size.

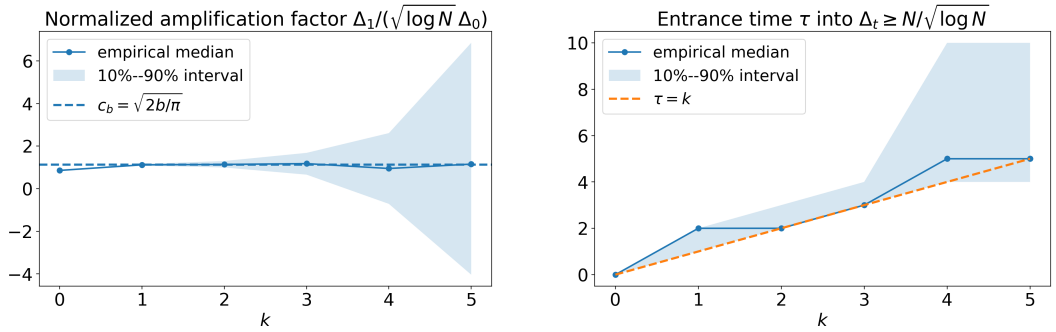


Figure 3: Amplification and entrance into $N/\sqrt{\log N}$, with $b = 2$ and 300 paths per k. Solid curves are medians, and shaded regions are 10%–90% pathwise intervals. Dashed references are $c_b = \sqrt{2b/\pi}$ (left) and $\tau = k$ (right). Paths run through round 9; nonentrance is recorded at $\tau = 10$.

Amplification and entrance time. For $b = 2$ and $k = 0, \ldots, 5$, choose

$$\Delta_0 \approx N/\left[\sqrt{\log N}(c_b\sqrt{\log N})^k\right], \qquad c_b = \sqrt{2b/\pi},$$

rounding the initial blue camp upward to an integer. Define the entrance time

$$\tau = \inf\{t \geq 0 : \Delta_t \geq N/\sqrt{\log N}\}.$$

Near balance, the linear-response prediction $\Delta_{t+1} \approx c_b\sqrt{\log N}\,\Delta_t$ suggests entrance near round k. Figure 3 shows the normalized first-step growth and the entrance time, with references c_b and $\tau = k$. Under each setting, 300 independent paths are simulated through round 9. A path that has not reached the target scale is recorded at $\tau = 10$. Solid curves are empirical medians, and shaded bands are 10%–90% pathwise intervals. The variation increases for larger k, when the initial advantage approaches or falls below the critical scale. These settings illustrate the boundary of reliable amplification. Entrance into $N/\sqrt{\log N}$ alone does not imply the two-round guaranty in Theorem 3.1, which requires a coefficient $K > K_2^{\mathrm{ER}}(b)$.

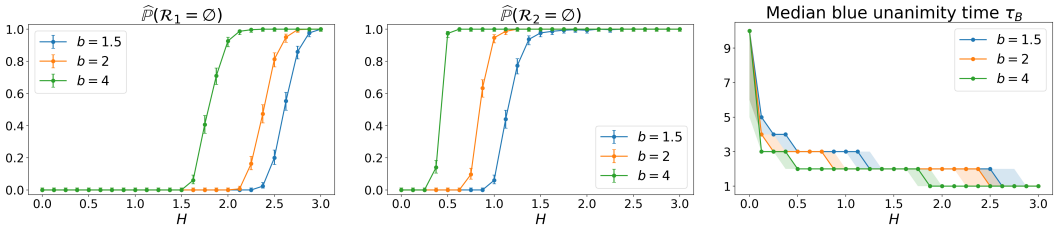


Figure 4: Rapid unanimity for $N = 10^4$ and $b \in \{1.5, 2, 4\}$, with 300 independent paths per parameter value. The left and middle panels show empirical probabilities of blue unanimity by rounds 1 and 2, respectively, with pointwise 95% Wilson intervals. The right panel shows the median of $\min\{\tau_B, 10\}$, with the 10%–90% pathwise interval.

Rapid unanimity. We vary the normalized initial advantage for $b \in \{1.5, 2, 4\}$ by changing it.

$$H = \Delta_0 \sqrt{\log N}/N, \tag{4.1}$$

using 300 independent paths per parameter value. Figure 4 shows the empirical probabilities of blue unanimity by rounds 1 and 2, together with the capped blue unanimity time $\min\{\tau_B, 10\}$. The observed transitions occur at smaller values of H as b increases. The transition for blue unanimity by round 2 also occurs at a smaller initial advantage than the transition by round 1. These observations illustrate the rapid-completion behavior described by Theorem 3.1. The final panel shows how the capped blue-unanimity time decreases as the initial advantage grows.

5 Future Directions

Effect of spatial locality. The Erdős–Rényi (ER) model treats vertices as statistically interchangeable, whereas a *Random Geometric Graph* (RGG), denoted by $\mathbb{G}(N, p, d)$, embeds vertices as points in a d-dimensional feature space and introduces geometric dependence and local bottlenecks. For fixed $p \in (0, 1)$, Bubeck, Ding, Eldan, and Rácz Bubeck et al. (2016) proved that the total variation distance $\mathrm{TV}(\mathbb{G}(N, p), \mathbb{G}(N, p, d))$ converges to 1 when $d \ll N^3$, whereas the same distance tends uniformly to zero when $d \gg N^3$. For the sparse regime $p = c/N$, the same work proved distinguishability in terms of total variation distance when $d \ll (\log N)^3$, while Brennan, Bresler, and Nagaraj Brennan et al. (2020), and subsequently Liu, Mohanty, Schramm, and Yang Liu et al. (2022) established equivalence also in terms of total variation distance for $d \gg N^{3/2}$ and $d \gg (\log N)^{36}$, respectively, narrowing the gap toward the conjectured threshold $d \asymp (\log N)^3$. At the spectral level, Cao and Zhu Cao & Zhu (2025) proved that when $d \gtrsim Np \log(1/p)$ and $Np \to \infty$, the *Empirical Spectral Distribution* (ESD) of the normalized adjacency matrix $\mathbf{A}/\sqrt{Np}$ under $\mathbb{G}(N, p, d)$ converges to the semicircle law; and when $d \gg \log N$, $p = c/N$ for some fixed $c > 0$, the ESD of $\mathbf{A}/\sqrt{c}$ converges to the limiting spectral distribution under $\mathbb{G}(N, c/N)$. Preceding results indicate that RGGs in sufficiently high dimensions behave similarly to ER graphs, while those in lower dimensions preserve geometric features and can display more intricate dynamics. This leads us to investigate how spatial locality influences both the consensus threshold and the rate at which consensus is reached in the RGG model relative to the ER benchmark.

Other interaction mechanisms. Beyond resampled random graphs, it is of interest to explore how different interaction mechanisms affect winner selection and the speed of reaching unanimity. For instance, one may consider exact-majority population protocols featuring random pairwise interactions Alistarh et al. (2015), systems in which social and opinion variables co-evolve continuously Zimper et al. (2026), or preferential-attachment networks that grow and incorporate reinforced opinion adoption Podder & Roy (2026). An open question is whether phenomena such as the critical window for winner selection, stepwise amplification, and sharp regimes for convergence time continue to manifest in these alternative settings, and to what extent these effects rely on the structure of synchronous, graph-based majority updates.

STATEMENT OF AI USE

In this work, we used generative AI tools to <generate synthetic data sets and assist in the writing of proofs>. We have not used generative AI tools to <help develop theoretical models or conceptual frameworks, formulate mathematical claims, provide critical ingredients for proving mathematical claims, propose or refine hypotheses, and interpret results>, and <design or provide feedback on research methodology or experiments, implement methods, assist with translation, clean and reformat datasets, support qualitative and thematic data analysis> are not applicable to this work. Additionally, we used generative AI tools to <identify relevant literature>. We have reviewed all AI-assisted work. We take full responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

REFERENCES

Dan Alistarh, Rati Gelashvili, and Milan Vojnović. Fast and exact majority in population protocols. In *Proceedings of the 2015 ACM Symposium on Principles of Distributed Computing*, PODC'15, pp. 47–56. Association for Computing Machinery, 2015. ISBN 9781450336178. doi: 10.1145/2767386.2767429. URL https://doi.org/10.1145/2767386.2767429.

Itai Benjamini, Siu-On Chan, Ryan O'Donnell, Omer Tamuz, and Li-Yang Tan. Convergence, unanimity and disagreement in majority dynamics on unimodular graphs and random graphs. *Stochastic Processes and their Applications*, 126(9):2719–2733, 2016.

Ross Berkowitz and Pat Devlin. Central limit theorem for majority dynamics: Bribing three voters suffices. *Stochastic Processes and their Applications*, 146:187–206, 2022. doi: 10.1016/j.spa.2022.01.010.

Matthew Brennan, Guy Bresler, and Dheeraj Nagaraj. Phase transitions for detecting latent geometry in random graphs. *Probability Theory and Related Fields*, 178(3):1215–1289, 2020.

Sébastien Bubeck, Jian Ding, Ronen Eldan, and Miklós Z Rácz. Testing for high-dimensional geometry in random graphs. *Random Structures & Algorithms*, 49(3):503–532, 2016.

Yifan Cao and Yizhe Zhu. Spectra of high-dimensional sparse random geometric graphs. *arXiv preprint arXiv:2507.06556*, 2025.

Debsoumya Chakraborti, Jeong Han Kim, Joonkyung Lee, and Tuan Tran. Majority dynamics on sparse random graphs. *Random Structures & Algorithms*, 63(1):171–191, 2023.

Ioana Dumitriu, Muchen Ju, and Hai-Xiao Wang. Majority dynamics on assortative sparse stochastic block models. *arXiv preprint arXiv:2607.24652*, 2026.

Batu El, Jinhee Paeng, Fatih Dinc, Shiye Su, Mete Erdogan, Aneesh Pappu, Haotian Ye, Wanjia Zhao, Surya Ganguli, and James Zou. Physics of agents: Statistical mechanics predicts collective behavior of AI agents. *arXiv preprint arXiv:2608.16578*, 2026.

Carl-Gustav Esseen. A moment inequality with an application to the central limit theorem. *Scandinavian Actuarial Journal*, 1956(2):160–170, 1956. doi: 10.1080/03461238.1956.10414946. URL https://doi.org/10.1080/03461238.1956.10414946.

Nikolaos Fountoulakis, Mihyun Kang, and Tamás Makai. Resolution of a conjecture on majority dynamics: Rapid stabilization in dense random graphs. *Random Structures & Algorithms*, 57(4):1134–1156, 2020.

Dmitry Gavinsky, Shachar Lovett, Michael E. Saks, and Srikanth Srinivasan. A Tail Bound for Read-k Families of Functions. *Random Structures & Algorithms*, 47(1):99 – 108, 2015. doi: 10.1002/rsa.20532. URL https://doi.org/10.1002/rsa.20532.

Gopal Goel and Ashwin Sah. Majority dynamics on sparse random graphs. *arXiv preprint arXiv:2609.14957*, 2026. URL https://arxiv.org/abs/2609.14957.

Sean Jaffe. A new bound in majority dynamics on random graphs. *arXiv preprint arXiv:2503.14401*, 2025.

Jeong Han Kim and BaoLinh Tran. A new density limit for unanimity in majority dynamics on random graphs. *arXiv preprint arXiv:2503.07447*, 2025.

Lucien Le Cam. An approximation theorem for the Poisson binomial distribution. *Pacific Journal of Mathematics*, 10(4):1181–1197, 1960. doi: 10.2140/pjm.1960.10.1181.

Siqi Liu, Sidhanth Mohanty, Tselil Schramm, and Elizabeth Yang. Testing thresholds for high-dimensional sparse random geometric graphs. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pp. 672–677, 2022.

Elchanan Mossel and Omer Tamuz. Opinion exchange dynamics. *Probability Surveys*, 14:155–204, 2017.

Sooraj M Moumanti Podder and Archi Roy. A model of opinion dynamics evolving via a preferential attachment mechanism involving multiple extractions. *arXiv preprint arXiv:2608.01419*, 2026. URL https://arxiv.org/abs/2608.01419.

Ashwin Sah and Mehtaab Sawhney. Majority dynamics: The power of one. *Israel Journal of Mathematics*, 267:85–133, 2025. doi: 10.1007/s11856-024-2690-1.

J. Michael Steele. Le Cam's inequality and Poisson approximations. *The American Mathematical Monthly*, 101(1):48–54, 1994.

BaoLinh Tran and Van Vu. The "power of few" phenomenon: The sparse case. *Random Structures & Algorithms*, 66(1):e21260, 2025.

Linh Tran and Van Vu. Reaching a consensus on random networks: The power of few. *Theory of Computing*, 19(6):1–21, 2023. doi: 10.4086/toc.2023.v019a006. URL https://theoryofcomputing.org/articles/v019a006.

Haoyu Wang, Jiaheng Wei, and Zhenyuan Zhang. Consensus on dynamic stochastic block models: Fast convergence and phase transitions. *arXiv preprint arXiv:2209.03999*, 2022. URL https://arxiv.org/abs/2209.03999.

Ahad N Zehmakan. Opinion forming in Erdős–Rényi random graph and expanders. *Discrete Applied Mathematics*, 277:280–290, 2020.

Sebastian Zimper, Nataša Djurdjevac Conrad, Federico Cornalba, and Ana Djurdjevac. Clustering in co-evolving opinion dynamics: reduced SPDE models. *arXiv preprint arXiv:2604.27961*, 2026.

CONTENTS

1 Introduction **1**

2 Model and Problem Formulation **3**

3 Main results **3**

4 Numerical Experiments **8**

5 Future Directions **9**

A Proof of Gaussian Winner Selection **13**

B Proofs of Amplification and Consensus-Time Bounds **19**

C Proofs of Rapid Unanimity **23**

D One-Step Probability Estimates **26**

E Continuity-corrected Gaussian estimates via Poisson approximation **30**

F Auxiliary Probabilistic Tools **36**

G Related Literature **38**

A Proof of Gaussian Winner Selection

We present the proof of Theorem 3.4 in this section. The proofs of Lemmas A.1, A.2, and A.3 are given in Section A.2 below. For this winner-selection analysis, we condition on a deterministic initial configuration $\mathbf{y}_0$ satisfying (3.11) and work under Assumption 2.1. The final subsection instead considers the random initial votes of Corollary 3.6.

The main idea is that the first update selects the sign of the eventual winner according to a Gaussian crossover and that the selected color subsequently reaches unanimity through the amplification mechanism of Theorem 3.2. We first provide the three one-step estimates needed for the proof below.

A.1 One-step estimates in the critical window

We begin with the conditional mean of the first-update advantage in the critical window (3.11).

Lemma A.1 (Critical-window mean advantage)**.** *For every fixed $C_0 < \infty$, uniformly over deterministic initial configurations satisfying (3.11), we have*

$$\frac{\mathbb{E}[\Delta_1 \mid \mathbf{y}_0]}{\sqrt{N}} = x_0 + O\left(\frac{1}{\sqrt{\log N}}\right). \tag{A.1}$$

The implicit constant depends only on b and C_0.

To upgrade the one-step estimate to a uniform approximation, we follow the approach in Berkowitz & Devlin (2022) to introduce the configuration-dependent factor and the two uniform approximation errors that appear in Theorem 3.4. For a deterministic initial configuration $\mathbf{y}_0$, let $\mathrm{W}_0^{\mathcal{R}}$ and $\mathrm{W}_0^{\mathcal{B}}$ be independent auxiliary random variables with the following distributions:

$$\mathrm{W}_0^{\mathcal{R}} \sim \mathbf{Bin}(|\mathcal{R}_0|, p), \qquad \text{and} \qquad \mathrm{W}_0^{\mathcal{B}} \sim \mathbf{Bin}(|\mathcal{B}_0|, p).$$

The corresponding normalization factor, introduced in Berkowitz & Devlin (2022), is defined as

$$\chi_N(\mathbf{y}_0) \coloneqq \mathbb{P}(\mathrm{W}_0^{\mathcal{R}} \geq \mathrm{W}_0^{\mathcal{B}}) \cdot \mathbb{P}(\mathrm{W}_0^{\mathcal{R}} \leq \mathrm{W}_0^{\mathcal{B}}). \tag{A.2}$$

For $C_0 < \infty$, the centered approximation error, uniform over the critical window $|x_0| \leq C_0$, is

$$\eta_N^{\mathrm{BD}}(C_0) \coloneqq \sup_{\substack{\mathbf{y}_0 \text{ deterministic:} \\ |x_0| \leq C_0}} \sup_{z \in \mathbb{R}} \left| \mathbb{P}\left(\frac{\Delta_1 - \mathbb{E}[\Delta_1 \mid \mathbf{y}_0]}{2\sqrt{N \chi_N(\mathbf{y}_0)}} \leq z \,\middle|\, \mathbf{y}_0 \right) - \Phi(z) \right|. \tag{A.3}$$

The corresponding uniform error after restoring the original centering and normalization is

$$\rho_N(C_0) \coloneqq \sup_{\substack{\mathbf{y}_0 \text{ deterministic:} \\ |x_0| \leq C_0}} \sup_{z \in \mathbb{R}} \left| \mathbb{P}\left(\frac{\Delta_1}{\sqrt{N}} \leq z \,\middle|\, \mathbf{y}_0 \right) - \Phi\left(z - x_0\right) \right|. \tag{A.4}$$

Lemma A.2 (Uniform critical-window one-step CLT)**.** *For every fixed $C_0 < \infty$, $\eta_N^{\mathrm{BD}}(C_0) = o(1)$. Moreover, there exists $C = C(b, C_0) > 0$ such that, for all sufficiently large N,*

$$\rho_N(C_0) \leq \eta_N^{\mathrm{BD}}(C_0) + \frac{C}{\sqrt{\log N}} = o(1). \tag{A.5}$$

We then provide the uniform anti-concentration estimate for Δ_1.

Lemma A.3 (Uniform critical-window anti-concentration)**.** *Fix $C_0 < \infty$. For every deterministic positive sequence $\ell_N \to 0$, uniformly over deterministic initial configurations satisfying (3.11),*

$$\sup_{u \in \mathbb{R}} \mathbb{P}\left(u\sqrt{N} \leq \Delta_1 \leq (u + \ell_N)\sqrt{N} \,\middle|\, \mathbf{y}_0 \right) \leq \frac{\ell_N}{\sqrt{2\pi}} + 2\rho_N(C_0) = o(1). \tag{A.6}$$

A.2 Proofs of the one-step estimates

A.2.1 Proof of Lemma A.1

Proof of Lemma A.1. By symmetry under interchanging blue and red, it suffices to consider $\Delta_0 \geq 0$. We denote

$$s_N \coloneqq \sqrt{(N-1)p(1-p)}, \qquad z_N \coloneqq \frac{p\Delta_0}{s_N}, \qquad \delta_N \coloneqq \frac{p - 1/2}{s_N}.$$

Condition (3.11) gives

$$0 \le \Delta_0 \le C_0\sqrt{\frac{\pi N}{2b\log N}}, \qquad z_N = O(N^{-1/2}), \qquad \delta_N = O((\log N)^{-1/2}). \tag{A.7}$$

Thus both color classes have size at least $N/3$ for all sufficiently large N, and the standardized means remain uniformly bounded. Hence Lemma D.4 applies uniformly.

Its critical-window estimate (D.15) gives

$$p_0^{\mathcal{R}} - p_0^{\mathcal{B}} = \sqrt{\frac{2}{\pi}}\frac{p\Delta_0}{s_N} + O\left(\frac{p\Delta_0}{\log N}\right). \tag{A.8}$$

Moreover, Lemma D.1 gives the exact identities

$$x_0^{\mathcal{R}} - \frac{1}{2\sqrt{v_0^{\mathcal{R}}}} = z_N + \delta_N, \qquad x_0^{\mathcal{B}} - \frac{1}{2\sqrt{v_0^{\mathcal{B}}}} = -z_N + \delta_N.$$

Therefore, (D.13a)–(D.13b), the identity $\Phi(z) + \Phi(-z) = 1$, and the Lipschitz continuity of Φ yield

$$p_0^{\mathcal{R}} + p_0^{\mathcal{B}} = \Phi(z_N + \delta_N) + \Phi(-z_N + \delta_N) + O((\log N)^{-1}) = 1 + O((\log N)^{-1/2}). \tag{A.9}$$

Substituting (A.8) and (A.9) into (D.10b) and dividing by $\sqrt{N}$ gives

$$\frac{\mathbb{E}[\Delta_1 \mid \mathbf{y}_0]}{\sqrt{N}} = \sqrt{\frac{2}{\pi}}\frac{\sqrt{N}\,p\Delta_0}{s_N} + O\left(\frac{\sqrt{N}\,p\Delta_0}{\log N} + \frac{\Delta_0}{\sqrt{N\log N}}\right). \tag{A.10}$$

Since $p = b\log N/N$, the leading term satisfies

$$\sqrt{\frac{2}{\pi}}\frac{\sqrt{N}\,p\Delta_0}{s_N} = x_0\sqrt{\frac{N}{N-1}}\frac{1}{\sqrt{1-p}} = x_0 + O\left(\frac{\log N}{N}\right),$$

whereas the two remainders in (A.10) are respectively $O((\log N)^{-1/2})$ and $O((\log N)^{-1})$. This proves (A.1) for $\Delta_0 \ge 0$.

If $\Delta_0 < 0$, interchange the two colors. This changes both x_0 and Δ_1 to their negatives, so the same conclusion follows. □

A.2.2 Proof of Lemma A.2

Proof of Lemma A.2. Fix an admissible initial configuration, abbreviate $\chi_N = \chi_N(\mathbf{y}_0)$, and set

$$\sigma_N := \sqrt{Np(1-p)}, \qquad D_N := \mathrm{W}_0^{\mathcal{R}} - \mathrm{W}_0^{\mathcal{B}}, \qquad F_N(k) := \mathbb{P}(D_N \le k).$$

Condition (3.11) gives

$$\frac{p|\Delta_0|}{\sigma_N} = O(N^{-1/2}), \qquad \sigma_N \sim \sqrt{b\log N} \longrightarrow \infty. \tag{A.11}$$

The variable D_N has mean $-p\Delta_0$, variance σ_N^2, and total sparse-binomial density parameter b. Hence Lemma D.5, evaluated at $k \in \{-1, 0\}$, gives

$$F_N(k) = \Phi\left(\frac{k + p\Delta_0}{\sigma_N}\right) + O((\log N)^{-1/2}) = \frac{1}{2} + O((\log N)^{-1/2}),$$

uniformly over the initial configurations under consideration. Therefore,

$$\mathbb{P}(\mathrm{W}_0^{\mathcal{R}} \le \mathrm{W}_0^{\mathcal{B}}) = F_N(0) = \frac{1}{2} + O((\log N)^{-1/2}),$$

and

$$\mathbb{P}(\mathrm{W}_0^{\mathcal{R}} \ge \mathrm{W}_0^{\mathcal{B}}) = 1 - F_N(-1) = \frac{1}{2} + O((\log N)^{-1/2}).$$

Multiplying these estimates yields

$$\chi_N = \frac{1}{4} + O((\log N)^{-1/2}). \tag{A.12}$$

In particular, $\log(1/\chi_N) = O(1) = o(\log \sigma_N)$, so the hypotheses of Theorem F.1 hold.

Consider any sequence of admissible deterministic initial configurations. Conditional on each $\mathbf{y}_0$, the next update uses a graph sampled independently from $\mathbb{G}(N,p)$, so Theorem F.1 applies with $W_R = \mathrm{W}_0^{\mathcal{R}}$, $W_B = \mathrm{W}_0^{\mathcal{B}}$, and normalization factor $\chi_N = \chi_N(\mathbf{y}_0)$. The hypotheses verified above therefore give

$$\frac{|\mathcal{R}_1| - \mathbb{E}[|\mathcal{R}_1| \mid \mathbf{y}_0]}{\sqrt{N\chi_N}} \xrightarrow{\mathrm{d}} \mathcal{N}(0,1).$$

On the other hand, $\Delta_1 = N - 2|\mathcal{R}_1|$, and hence

$$\Delta_1 - \mathbb{E}[\Delta_1 \mid \mathbf{y}_0] = -2\left(|\mathcal{R}_1| - \mathbb{E}[|\mathcal{R}_1| \mid \mathbf{y}_0]\right).$$

Thus Y_N below is the negative of the standardized red-camp size in the preceding display. Since the standard normal distribution is symmetric,

$$Y_N := \frac{\Delta_1 - \mathbb{E}[\Delta_1 \mid \mathbf{y}_0]}{2\sqrt{N\chi_N}} \xrightarrow{\mathrm{d}} \mathcal{N}(0,1).$$

Because Φ is continuous, this convergence is also convergence in Kolmogorov distance. It is uniform over the admissible configurations: otherwise, for some $\varepsilon > 0$, one could select a sequence of admissible configurations whose Kolmogorov distances are at least ε, contradicting the preceding sequential conclusion. Thus

$$\eta_N^{\mathrm{BD}}(C_0) = o(1).$$

It remains to restore the original centering and normalization. Define

$$m_N := \frac{\mathbb{E}[\Delta_1 \mid \mathbf{y}_0]}{\sqrt{N}}, \qquad \alpha_N := 2\sqrt{\chi_N}.$$

Equations (A.12) and (A.1) give, uniformly,

$$\alpha_N = 1 + O((\log N)^{-1/2}), \qquad m_N = x_0 + O((\log N)^{-1/2}).$$

For every $z \in \mathbb{R}$,

$$\mathbb{P}\left(\frac{\Delta_1}{\sqrt{N}} \le z \,\middle|\, \mathbf{y}_0\right) = \mathbb{P}\left(Y_N \le \frac{z - m_N}{\alpha_N} \,\middle|\, \mathbf{y}_0\right).$$

For α in a fixed neighborhood of 1,

$$\sup_{w \in \mathbb{R}} |\Phi(w/\alpha) - \Phi(w)| \le C|\alpha - 1|,$$

while the Lipschitz continuity of Φ controls translations. Hence

$$\begin{aligned}\sup_{z \in \mathbb{R}} \left|\mathbb{P}\left(\frac{\Delta_1}{\sqrt{N}} \le z \,\middle|\, \mathbf{y}_0\right) - \Phi(z - x_0)\right| &\le \eta_N^{\mathrm{BD}}(C_0) + C|\alpha_N - 1| + C|m_N - x_0| \\ &\le \eta_N^{\mathrm{BD}}(C_0) + \frac{C}{\sqrt{\log N}}.\end{aligned}$$

Taking the supremum over the admissible initial configurations and using the definition (A.4) proves (A.5). □

A.2.3 Proof of Lemma A.3

Proof of Lemma A.3. For a fixed admissible initial configuration, define

$$F_{\mathbf{y}_0}(z) := \mathbb{P}\left(\frac{\Delta_1}{\sqrt{N}} \le z \,\middle|\, \mathbf{y}_0\right), \qquad G_{\mathbf{y}_0}(z) := \Phi(z - x_0).$$

By the definition of $\rho_N(C_0)$ in (A.4),

$$\sup_{z\in\mathbb{R}} |F_{\mathbf{y}_0}(z) - G_{\mathbf{y}_0}(z)| \le \rho_N(C_0).$$

Since $G_{\mathbf{y}_0}$ is continuous, taking the left limit at u gives

$$F_{\mathbf{y}_0}(u^-) \ge G_{\mathbf{y}_0}(u) - \rho_N(C_0).$$

Therefore, for every $u \in \mathbb{R}$,

$$\begin{aligned}\mathbb{P}\left(u\sqrt{N} \le \Delta_1 \le (u+\ell_N)\sqrt{N} \,\middle|\, \mathbf{y}_0\right) &= F_{\mathbf{y}_0}(u+\ell_N) - F_{\mathbf{y}_0}(u^-) \\ &\le G_{\mathbf{y}_0}(u+\ell_N) - G_{\mathbf{y}_0}(u) + 2\rho_N(C_0) \\ &\le \frac{\ell_N}{\sqrt{2\pi}} + 2\rho_N(C_0),\end{aligned}$$

where the last inequality uses $\sup_x \varphi(x) = 1/\sqrt{2\pi}$. The bound is uniform in u and in the admissible initial configuration. Taking the supremum over u, and then using $\ell_N \to 0$ and $\rho_N(C_0) = o(1)$, proves (A.6). □

A.3 Proof of Theorem 3.4

The proof separates winner selection on the first update from the subsequent amplification to unanimity. Lemma A.1 and Lemma A.2 show uniformly that $\Delta_1/\sqrt{N}$ is approximated by a Gaussian random variable with mean x_0 and variance one. In particular, along any sequence with $x_0 \to x \in \mathbb{R}$,

$$\frac{\Delta_1}{\sqrt{N}} \xrightarrow{\mathrm{d}} x + Z, \qquad Z \sim \mathcal{N}(0,1),$$

so the first update produces a positive or negative advantage with limiting probabilities $\Phi(x)$ and $\Phi(-x)$, respectively. Lemma A.3 shows that the probability of the event $|\Delta_1| < \sqrt{N}/h_N$ is vanishing. Consequently, with probability $1 - o(1)$, the first update either has already reached unanimity or produces a signed advantage of magnitude at least $\sqrt{N}/h_N$. Conditional on the first update, the future graphs are independent of the resulting configuration; thus the time-shifted form of Theorem 3.2 drives a positive advantage to blue unanimity and, after exchanging the colors, a negative advantage to red unanimity. The proof below makes this argument uniform and tracks its total error precisely.

Proof of Theorem 3.4. We first verify that the total error vanishes. Since $h_N = o(\sqrt{\log N})$, we have $\log N/h_N^2 \to \infty$. Hence (3.4) gives

$$\eta_N\left(\frac{\sqrt{N}}{h_N}\right) = 8\exp\left(-c\frac{\log N}{h_N^2}\right) + 2\exp\left(-(\log N)^2\right) + 2N^{-\xi} = o(1). \tag{A.13}$$

Together with $h_N \to \infty$, (A.5), and (3.9), this gives

$$\mathfrak{r}_N(C_0) = \rho_N(C_0) + \frac{1}{h_N} + \eta_N\left(\frac{\sqrt{N}}{h_N}\right) = o(1). \tag{A.14}$$

Recall that $(\mathcal{F}_t)_{t\ge 0}$ in (2.6) denotes the natural filtration generated by the initial configuration and the resampled graphs through time t. Define the following three $\mathcal{F}_1$-measurable events, which partition the sample space:

$$\mathcal{E}_N^+ := \left\{\Delta_1 \ge \frac{\sqrt{N}}{h_N}\right\}, \quad \mathcal{E}_N^- := \left\{\Delta_1 \le -\frac{\sqrt{N}}{h_N}\right\}, \quad \mathcal{E}_N^0 := \left\{|\Delta_1| < \frac{\sqrt{N}}{h_N}\right\}.$$

Lemma A.3 with $u = -h_N^{-1}$ and $\ell_N = 2/h_N$ gives, uniformly under (3.11),

$$\mathbb{P}(\mathcal{E}_N^0 \mid \mathbf{y}_0) \le \frac{2}{\sqrt{2\pi}\,h_N} + 2\rho_N(C_0). \tag{A.15}$$

The definition of $\rho_N(C_0)$ controls the distribution function of $\Delta_1/\sqrt{N}$ and, because the Gaussian comparison distribution is continuous, also its left limits. Evaluating at the thresholds $\pm h_N^{-1}$ therefore gives

$$\mathbb{P}(\mathcal{E}_N^+ \mid \mathbf{y}_0) = 1 - \Phi(h_N^{-1} - x_0) + O(\rho_N(C_0)) = \Phi(x_0) + O\left(h_N^{-1} + \rho_N(C_0)\right), \tag{A.16a}$$

$$\mathbb{P}(\mathcal{E}_N^- \mid \mathbf{y}_0) = \Phi(-h_N^{-1} - x_0) + O(\rho_N(C_0)) = \Phi(-x_0) + O\left(h_N^{-1} + \rho_N(C_0)\right). \tag{A.16b}$$

Here we used $1 - \Phi(z) = \Phi(-z)$ and the Lipschitz continuity of Φ. All three estimates are uniform over the configurations under consideration.

Given $\mathcal{F}_1$, the configuration $\mathbf{y}_1$ is fixed and the future interaction graphs are independent of it. Moreover, for all sufficiently large N, we have $h_N \geq 1$ and $c_1\sqrt{\log N} > 1$. Hence the upper horizon in (3.5b) is nonincreasing in the initial advantage.

Blue unanimity either already holds at time 1, or Theorem 3.2 applies from time 1. On $\mathcal{E}_N^+$, we have $\Delta_1 \geq \sqrt{N}/h_N$, so its upper horizon is at most T_N and the failure probability is at most $\eta_N(\sqrt{N}/h_N)$. Since $\mathcal{E}_N^+ \in \mathcal{F}_1$, this gives

$$\mathbb{P}\left(\mathcal{R}_{T_N+3} \neq \emptyset, \mathcal{E}_N^+ \,\middle|\, \mathcal{F}_1\right) = \mathbb{1}_{\{\mathcal{E}_N^+\}}\mathbb{P}\left(\mathcal{R}_{T_N+3} \neq \emptyset \,|\, \mathcal{F}_1\right) \leq \eta_N\left(\frac{\sqrt{N}}{h_N}\right)\mathbb{1}_{\{\mathcal{E}_N^+\}}.$$

The same argument with the colors exchanged gives

$$\mathbb{P}\left(\mathcal{B}_{T_N+3} \neq \emptyset, \mathcal{E}_N^- \,\middle|\, \mathcal{F}_1\right) = \mathbb{1}_{\{\mathcal{E}_N^-\}}\mathbb{P}\left(\mathcal{B}_{T_N+3} \neq \emptyset \,|\, \mathcal{F}_1\right) \leq \eta_N\left(\frac{\sqrt{N}}{h_N}\right)\mathbb{1}_{\{\mathcal{E}_N^-\}}.$$

Here we used that $\eta_N(s)$ is nonincreasing for $s \geq 0$. Taking conditional expectations yields

$$\mathbb{P}\left(\mathcal{R}_{T_N+3} \neq \emptyset, \mathcal{E}_N^+ \,\middle|\, \mathbf{y}_0\right) \leq \eta_N\left(\frac{\sqrt{N}}{h_N}\right)\mathbb{P}(\mathcal{E}_N^+ \mid \mathbf{y}_0) \leq \eta_N\left(\frac{\sqrt{N}}{h_N}\right),$$

$$\mathbb{P}\left(\mathcal{B}_{T_N+3} \neq \emptyset, \mathcal{E}_N^- \,\middle|\, \mathbf{y}_0\right) \leq \eta_N\left(\frac{\sqrt{N}}{h_N}\right)\mathbb{P}(\mathcal{E}_N^- \mid \mathbf{y}_0) \leq \eta_N\left(\frac{\sqrt{N}}{h_N}\right).$$

The all-blue and all-red events are disjoint. Therefore, the second bound also implies

$$\mathbb{P}\left(\mathcal{R}_{T_N+3} = \emptyset, \mathcal{E}_N^- \,\middle|\, \mathbf{y}_0\right) \leq \eta_N\left(\frac{\sqrt{N}}{h_N}\right).$$

Using the partition $\mathcal{E}_N^+, \mathcal{E}_N^0, \mathcal{E}_N^-$, the lower bound follows from

$$\begin{aligned}\mathbb{P}\left(\mathcal{R}_{T_N+3} = \emptyset \,|\, \mathbf{y}_0\right) &\geq \mathbb{P}\left(\mathcal{R}_{T_N+3} = \emptyset, \mathcal{E}_N^+ \,\middle|\, \mathbf{y}_0\right) \\ &= \mathbb{P}(\mathcal{E}_N^+ \mid \mathbf{y}_0) - \mathbb{P}\left(\mathcal{R}_{T_N+3} \neq \emptyset, \mathcal{E}_N^+ \,\middle|\, \mathbf{y}_0\right) \\ &\geq \mathbb{P}(\mathcal{E}_N^+ \mid \mathbf{y}_0) - \eta_N\left(\frac{\sqrt{N}}{h_N}\right).\end{aligned}$$

For the upper bound, the same partition gives

$$\begin{aligned}&\mathbb{P}\left(\mathcal{R}_{T_N+3} = \emptyset \,|\, \mathbf{y}_0\right) \\ &= \mathbb{P}\left(\mathcal{R}_{T_N+3} = \emptyset, \mathcal{E}_N^+ \,\middle|\, \mathbf{y}_0\right) + \mathbb{P}\left(\mathcal{R}_{T_N+3} = \emptyset, \mathcal{E}_N^0 \,\middle|\, \mathbf{y}_0\right) + \mathbb{P}\left(\mathcal{R}_{T_N+3} = \emptyset, \mathcal{E}_N^- \,\middle|\, \mathbf{y}_0\right) \\ &\leq \mathbb{P}(\mathcal{E}_N^+ \mid \mathbf{y}_0) + \mathbb{P}(\mathcal{E}_N^0 \mid \mathbf{y}_0) + \eta_N\left(\frac{\sqrt{N}}{h_N}\right).\end{aligned}$$

Combining these two bounds with (A.15) and (A.16a) gives

$$\mathbb{P}\left(\mathcal{R}_{T_N+3} = \emptyset \,|\, \mathbf{y}_0\right) = \Phi(x_0) + O\left(\mathfrak{r}_N(C_0)\right).$$

By exchanging the colors and using (A.16b), we similarly obtain

$$\mathbb{P}\left(\mathcal{B}_{T_N+3} = \emptyset \,|\, \mathbf{y}_0\right) = \Phi(-x_0) + O\left(\mathfrak{r}_N(C_0)\right).$$

These are (3.12a) and (3.12b).

Finally, the assumptions on h_N imply $1 \le h_N \le \sqrt{\log N}$ for all sufficiently large N, and hence $\log h_N = O(\log\log N)$. Therefore,

$$\begin{aligned}\log\left(Kh_N\sqrt{N/\log N}\right) &= \frac{1}{2}\log N - \frac{1}{2}\log\log N + \log h_N + O(1)\\ &= \frac{1}{2}\log N + O(\log\log N),\\ \log\left(c_1\sqrt{\log N}\right) &= \frac{1}{2}\log\log N + O(1).\end{aligned}$$

The quotient in (3.10) is positive for all sufficiently large N, while the maximum and ceiling affect it by at most $O(1)$. Therefore, $T_N + 3 = (1 + o(1))\log N/\log\log N$, completing the proof. □

A.4 Proof of Corollary 3.5

Proof. Fix $C_0 > 0$ and $K > K_2^{\mathrm{ER}}(b)$, set $h_N = (\log N)^{1/4}$, and define

$$\ell_N = \frac{\sqrt{N}}{h_N}, \qquad u_N = h_N\sqrt{N}, \qquad \mathcal{H}_N = \{\ell_N \le |\Delta_1| \le u_N\}.$$

The uniform CLT and anti-concentration estimate in Lemmas A.2 and A.3 give

$$\mathbb{P}(\mathcal{H}_N^c \mid \mathbf{y}_0) \le \frac{2}{\sqrt{2\pi}\,h_N} + 4\rho_N(C_0) + 2\Phi(C_0 - h_N) = o(1),$$

uniformly over $|x_0| \le C_0$. Here the last term bounds the two Gaussian tails outside $[-h_N, h_N]$.

Write $\underline{T}(s), \overline{T}(s)$ for the deterministic horizons in (3.5a)–(3.5b) with initial advantage s. On $\mathcal{H}_N$, select blue if $\Delta_1 > 0$ and red if $\Delta_1 < 0$, and let $\widehat{\tau}$ be the number of further rounds after time 1 until this selected color is unanimous, with value ∞ if it never is. Both the initial and round-one configurations are nonunanimous for all sufficiently large N on $\mathcal{H}_N$, since $|\Delta_0| = O(\sqrt{N/\log N})$ and $u_N = o(N)$. Conditional on $\mathcal{F}_1$, future graphs are independent of the round-one configuration. Applying Theorem 3.2, with colors exchanged when necessary, gives

$$\mathbb{P}\left(\underline{T}(|\Delta_1|) + 1 \le \widehat{\tau} \le \overline{T}(|\Delta_1|) + 2 \,\middle|\, \mathcal{F}_1\right) \ge 1 - \eta_N(\ell_N) \qquad \text{on } \mathcal{H}_N.$$

The error tends to zero because $\ell_N^2 \log N/N = \sqrt{\log N} \to \infty$. On the event in this display, $\widehat{\tau} < \infty$, so the opposite color cannot have reached unanimity earlier: either unanimous state is absorbing. Therefore $\min\{\tau_{\mathrm{B}}, \tau_{\mathrm{R}}\} = 1 + \widehat{\tau}$. Both horizons are nonincreasing in s for all sufficiently large N. Taking conditional expectations consequently yields, with probability $1 - o(1)$ uniformly in the initial configuration,

$$\underline{T}(u_N) + 2 \le \min\{\tau_{\mathrm{B}}, \tau_{\mathrm{R}}\} \le \overline{T}(\ell_N) + 3.$$

Finally,

$$\begin{aligned}\log\frac{KN}{u_N\sqrt{\log N}} &= \frac{1}{2}\log N - \frac{3}{4}\log\log N + \log K,\\ \log\frac{KN}{\ell_N\sqrt{\log N}} &= \frac{1}{2}\log N - \frac{1}{4}\log\log N + \log K,\end{aligned}$$

whereas each denominator defining the horizons is $\frac{1}{2}\log\log N + O(1)$. Thus both endpoints of the preceding time bracket are $(1 + o(1))\log N/\log\log N$, proving (3.13). □

A.5 Proof of Corollary 3.6

Proof. Let $X_v = z\mathbf{y}_0(v)$, so $X_v = +1$ represents a correct initial vote, and set

$$M_N = \sum_{v=1}^{N} X_v = z\Delta_0.$$

The variables X_v are independent, with $\mathbb{E}M_N = 2\theta\sqrt{N}$, $\mathbb{V}\mathrm{ar}(M_N) = N - 4\theta^2$, and uniformly bounded centered absolute third moments. Berry–Esseen (Lemma D.3), followed by the change of scale from $\sqrt{N-4\theta^2}$ to $\sqrt{N}$, gives

$$\sup_{w\in\mathbb{R}} \left|\mathbb{P}(M_N/\sqrt{N} \le w) - \Phi(w - 2\theta)\right| = O_\theta(N^{-1/2}).$$

In particular, $\mathbb{P}(M_N > 0) = \Phi(2\theta) + o(1)$ and $\mathbb{P}(M_N = 0) = o(1)$, establishing the claimed limiting accuracy of the centralized initial majority.

Choose $K > K_2^{\mathrm{ER}}(b)$, set $h_N = (\log N)^{1/4}$, and write

$$\ell_N = \sqrt{N}/h_N, \qquad u_N = h_N\sqrt{N}, \qquad \mathcal{H}_N = \{\ell_N \le |M_N| \le u_N\}.$$

The preceding Gaussian approximation gives

$$\mathbb{P}(\mathcal{H}_N^c) \le \frac{2}{\sqrt{2\pi}\, h_N} + 2\Phi(2\theta - h_N) + O_\theta(N^{-1/2}) = o(1).$$

On $\mathcal{H}_N$, both camps are nonempty and the initial majority has advantage $|\Delta_0| = |M_N| \ge \ell_N$. Conditional on the entire initial configuration, the interaction graphs retain their independent $\mathbb{G}(N,p)$ laws. By Theorem 3.2, with colors exchanged if necessary, the initial majority reaches unanimity within its two-sided time bracket except with conditional probability at most $\eta_N(\ell_N) = o(1)$. Hence, with unconditional probability $1 - o(1)$, the eventual unanimous opinion agrees with the initial majority. Absorption excludes earlier unanimity of the opposite opinion on this event. Correct unanimity therefore occurs exactly when $M_N > 0$, apart from an event of probability $o(1)$. This proves (3.14).

For the time claim, use the horizon notation $\underline{T}(s), \overline{T}(s)$ from the preceding proof. Their monotonicity gives, on the same event of probability $1 - o(1)$,

$$\underline{T}(u_N) + 1 \le \min\{\tau_{\mathrm{B}}, \tau_{\mathrm{R}}\} \le \overline{T}(\ell_N) + 2.$$

The horizon expansions in the proof of Corollary 3.5 show that both endpoints are $(1 + o(1))\log N/\log\log N$, completing the proof. □

B Proofs of Amplification and Consensus-Time Bounds

We present the proof of Theorem 3.2 in this section. The proof proceeds by first upgrading the one-step estimates from Section D.1 into a high-probability amplification bound. We then iterate this bound until the advantage reaches the constant-time regime of Theorem 3.1, thereby proving Theorem 3.2 and Corollary 3.3. The one-step failure probabilities are summable because every successful update multiplies both the advantage and the exponent in the next concentration bound.

B.1 Stepwise advantage amplification

We first present the stepwise advantage amplification lemma, with its proof deferred to Section B.2.

Lemma B.1 (Stepwise advantage amplification)**.** *For every fixed $K > 0$, there are constants $c_1 = c_1(b,K) > 0$, $C_2 = C_2(b,K) > 0$, and $c = c(b,K) > 0$ such that, whenever*

$$\sqrt{N/\log N} \le \Delta_t \le KN/\sqrt{\log N}, \tag{B.1}$$

the following two-sided amplification bound holds:

$$\mathbb{P}\left(c_1\sqrt{\log N}\,\Delta_t \le \Delta_{t+1} \le C_2\sqrt{\log N}\,\Delta_t \,\middle|\, \mathbf{y}_t\right) \ge 1 - 2\exp\left(-c\frac{\Delta_t^2\log N}{N}\right). \tag{B.2}$$

The main idea of the proof of Lemma B.1 can be summarized as follows: First, Lemma D.1 places the standardized degree differences in a bounded Gaussian window. Second, Lemma D.4 and Lemma D.2 show that the conditional mean of the next advantage is of order $\sqrt{\log N}\,\Delta_t$. Finally, we apply the read-2 concentration inequality in Lemma F.6 to upgrade this mean estimate to a high-probability two-sided amplification bound. The detailed proof is deferred to Section B.2.

B.2 Proof of the amplification lemma

Proof of Lemma B.1. Condition throughout on $\mathbf{y}_t$, and write

$$s_N := \sqrt{(N-1)p(1-p)}.$$

The upper bound in (B.1) gives $\Delta_t = o(N)$, so both camps have size at least $N/3$ for all sufficiently large N. Moreover, Lemma D.1 and (B.1) give

$$\max\{|x_t^{\mathcal{R}}|, |x_t^{\mathcal{B}}|\} \le \frac{p(\Delta_t + 1)}{s_N} \le M(b, K).$$

The lower bound in (B.1) also gives $\Delta_t \ge 1$. Hence Lemma D.4 applies, and (D.14), together with $p = b \log N/N$, yields

$$N\big(p_t^{\mathcal{R}} - p_t^{\mathcal{B}}\big) \asymp \frac{Np}{s_N}\Delta_t \asymp \sqrt{\log N}\,\Delta_t,$$

where the implicit constants depend only on b and K.

We denote $\mu_t := \mathbb{E}[\Delta_{t+1} \mid \mathbf{y}_t]$. By (D.10b) and $0 \le p_t^{\mathcal{R}}, p_t^{\mathcal{B}} \le 1$,

$$\big|\mu_t - N\big(p_t^{\mathcal{R}} - p_t^{\mathcal{B}}\big)\big| = \Delta_t\,\big|1 - p_t^{\mathcal{R}} - p_t^{\mathcal{B}}\big| \le \Delta_t.$$

Since $\sqrt{\log N} \to \infty$, we may therefore choose constants $c_1, C_2 > 0$, depending only on b, K, such that

$$2c_1\sqrt{\log N}\,\Delta_t \le \mu_t \le (C_2 - c_1)\sqrt{\log N}\,\Delta_t. \tag{B.3}$$

Finally, Lemma F.6 gives, for every $u > 0$,

$$\mathbb{P}\left(|\Delta_{t+1} - \mu_t| \ge u \,|\, \mathbf{y}_t\right) \le 2\exp\left(-\frac{u^2}{4N}\right).$$

Taking $u = c_1\sqrt{\log N}\,\Delta_t$ and using (B.3), we obtain

$$c_1\sqrt{\log N}\,\Delta_t \le \Delta_{t+1} \le C_2\sqrt{\log N}\,\Delta_t$$

except with conditional probability at most

$$2\exp\left(-\frac{c_1^2}{4}\frac{\Delta_t^2 \log N}{N}\right).$$

Taking $c = c_1^2/4$ proves (B.2). □

B.3 Proof of Theorem 3.2

Proof of Theorem 3.2. Let $c_1, C_2, c > 0$ be the constants from the preceding Lemma B.1, and define

$$a_N := c_1\sqrt{\log N}, \qquad A_N := C_2\sqrt{\log N}, \qquad \lambda_N := \frac{\Delta_0^2 \log N}{N}. \tag{B.4}$$

Condition (3.6) gives

$$\lambda_N \ge \frac{\log N}{h_N^2} \longrightarrow \infty, \tag{B.5}$$

because $h_N = o(\sqrt{\log N})$. In particular, for all sufficiently large N,

$$a_N > 1, \qquad A_N > 1, \qquad (a_N)^2 \ge 2, \qquad \Delta_0 \ge \sqrt{N/\log N}.$$

For both the lower and upper bounds, we will use the following geometric estimate: for every integer $m \ge 0$, with the sum understood as zero when $m = 0$,

$$2\sum_{t=0}^{m-1}\exp\left(-c\lambda_N (a_N)^{2t}\right) \le 2\sum_{t=0}^{\infty}\exp\left(-c\lambda_N(t+1)\right) = \frac{2\exp(-c\lambda_N)}{1-\exp(-c\lambda_N)} \le 4\exp(-c\lambda_N), \tag{B.6}$$

where we used $(a_N)^{2t} \geq 2^t \geq t+1$ and (B.5).

We first prove the lower bound on the time to unanimity. If $\underline{T} = 0$, then $\mathcal{R}_0 \neq \emptyset$ gives $\tau_{\mathrm{B}} \geq 1 = \underline{T}+1$, so the desired bound is immediate. Assume therefore that $\underline{T} \geq 1$. For $m \geq 0$, define

$$\mathcal{E}_m^- := \bigcap_{s=0}^{m-1} \{a_N \Delta_s \leq \Delta_{s+1} \leq A_N \Delta_s\}. \tag{B.7}$$

Here and below, an intersection over an empty index set is the sure event; in particular, $\mathcal{E}_0^-$ is the sure event. By (B.7), on $\mathcal{E}_t^-$ every update from time 0 through time $t-1$ satisfies the two-sided amplification inequality. Consequently,

$$(a_N)^t \Delta_0 \leq \Delta_t \leq (A_N)^t \Delta_0. \tag{B.8}$$

For every $0 \leq t < \underline{T}$, the definition of $\underline{T}$ gives

$$(A_N)^t \Delta_0 \leq (A_N)^{\underline{T}} \Delta_0 \leq K \frac{N}{\sqrt{\log N}}.$$

Together with (B.8) and the initial condition, this shows that, on $\mathcal{E}_t^-$,

$$\sqrt{\frac{N}{\log N}} \leq \Delta_t \leq K \frac{N}{\sqrt{\log N}}.$$

Thus the amplification lemma applies conditionally on $\mathcal{F}_t$. Moreover, the definition in (B.7) gives

$$(\mathcal{E}_{t+1}^-)^{\complement} \cap \mathcal{E}_t^- = \mathcal{E}_t^- \cap (\{\Delta_{t+1} < a_N \Delta_t\} \cup \{\Delta_{t+1} > A_N \Delta_t\}).$$

Using both sides of (B.2) and the tower property, we therefore obtain, for every $0 \leq t < \underline{T}$,

$$\begin{aligned} \mathbb{P}\left((\mathcal{E}_{t+1}^-)^{\complement} \cap \mathcal{E}_t^- \,\middle|\, \mathbf{y}_0\right) &= \mathbb{E}\left[\mathbb{1}_{\{\mathcal{E}_t^-\}} \mathbb{P}\left(\Delta_{t+1} \notin [a_N \Delta_t, A_N \Delta_t] \,|\, \mathcal{F}_t\right) \,\middle|\, \mathbf{y}_0\right] \\ &\leq 2 \exp\left(-c \frac{((a_N)^t \Delta_0)^2 \log N}{N}\right) = 2 \exp\left(-c \lambda_N (a_N)^{2t}\right). \end{aligned} \tag{B.9}$$

By (B.7), the events $\{\mathcal{E}_m^-\}_{m \geq 0}$ are nested. Hence a union bound over the first failed update, followed by (B.6), gives

$$\mathbb{P}\left((\mathcal{E}_{\underline{T}}^-)^{\complement} \mid \mathbf{y}_0\right) \leq 2 \sum_{t=0}^{\underline{T}-1} \exp\left(-c \lambda_N (a_N)^{2t}\right) \leq 4 \exp(-c \lambda_N). \tag{B.10}$$

On the event $\mathcal{E}_{\underline{T}}^-$, we have

$$\Delta_{\underline{T}} \leq (A_N)^{\underline{T}} \Delta_0 \leq K \frac{N}{\sqrt{\log N}} < N$$

for all sufficiently large N. Hence $\mathcal{R}_{\underline{T}} \neq \emptyset$. Since time is integer-valued and the all-blue configuration is absorbing, this implies $\tau_{\mathrm{B}} \geq \underline{T} + 1$. Together with the zero-horizon case, this proves

$$\mathbb{P}\left(\tau_{\mathrm{B}} < \underline{T} + 1 \mid \mathbf{y}_0\right) \leq 4 \exp(-c \lambda_N). \tag{B.11}$$

We then prove the upper bound. Define the random entrance time into the two-day completion window by

$$\tau_K := \inf\left\{t \geq 0 : \Delta_t \geq K \frac{N}{\sqrt{\log N}}\right\}, \tag{B.12}$$

with the convention $\inf \emptyset = \infty$. For $m \geq 0$, define

$$\mathcal{E}_m := \bigcap_{s=0}^{m-1} (\{\tau_K \leq s\} \cup \{\Delta_{s+1} \geq a_N \Delta_s\}). \tag{B.13}$$

By the empty-intersection convention, $\mathcal{E}_0$ is the sure event. If $\mathcal{E}_t \cap \{\tau_K > t\}$ occurs, then $\tau_K > s$ for every $0 \le s < t$; hence (B.13) forces every preceding update to satisfy the lower-amplification inequality. Therefore,

$$\Delta_t \ge (a_N)^t \Delta_0. \tag{B.14}$$

Moreover, $\tau_K > t$ gives $\Delta_t < KN/\sqrt{\log N}$, while (B.14) and the initial condition give $\Delta_t \ge \sqrt{N/\log N}$. Thus the amplification lemma applies conditionally on $\mathcal{F}_t$. Using only the lower side of (B.2) and then the tower property, we obtain, for every $0 \le t < \overline{T}$,

$$\begin{aligned}
\mathbb{P}\left(\mathcal{E}_{t+1}^{\complement} \cap \mathcal{E}_t \,\middle|\, \mathbf{y}_0\right) &= \mathbb{P}\left(\mathcal{E}_t \cap \{\tau_K > t\} \cap \{\Delta_{t+1} < a_N \Delta_t\} \,|\, \mathbf{y}_0\right) \\
&= \mathbb{E}\left[\mathbb{1}_{\{\mathcal{E}_t \cap \{\tau_K > t\}\}} \mathbb{P}\left(\Delta_{t+1} < a_N \Delta_t \,|\, \mathcal{F}_t\right) \,\middle|\, \mathbf{y}_0\right] \\
&\le 2\exp\left(-c\frac{((a_N)^t \Delta_0)^2 \log N}{N}\right) = 2\exp\left(-c\lambda_N (a_N)^{2t}\right).
\end{aligned} \tag{B.15}$$

The definition of $\overline{T}$ in (3.5b) ensures that

$$(a_N)^{\overline{T}} \Delta_0 \ge K \frac{N}{\sqrt{\log N}}. \tag{B.16}$$

This also covers $\overline{T} = 0$, in which case the initial configuration already lies in the completion window. If $\mathcal{E}_{\overline{T}}$ occurs and $\tau_K > \overline{T}$, then (B.14) and (B.16) contradict the definition of τ_K. Consequently, $\{\tau_K > \overline{T}\} \subseteq \mathcal{E}_{\overline{T}}^{\complement}$. By (B.13), the events $(\mathcal{E}_m)_{m \ge 0}$ are nested. Hence a union bound over the first failed update, followed by (B.6), gives

$$\mathbb{P}(\tau_K > \overline{T} \mid \mathbf{y}_0) \le 2 \sum_{t=0}^{\overline{T}-1} \exp\left(-c\lambda_N (a_N)^{2t}\right) \le 4\exp(-c\lambda_N), \tag{B.17}$$

where the sum is zero when $\overline{T} = 0$.

It remains to pass from entrance to unanimity. Let $\xi = \xi(b, K) > 0$ be supplied by Theorem 3.1. For each deterministic $s \ge 0$, the graphs sampled after time s are independent of $\mathcal{F}_s$ and have the same law as the graphs sampled from time zero. On $\{\tau_K = s\}$, we have $\Delta_s \ge KN/\sqrt{\log N}$, so the two-day theorem, shifted to time s, gives

$$\mathbb{P}\left(\mathcal{R}_{s+2} \ne \emptyset \mid \mathcal{F}_s\right) \le 2\exp\left(-(\log N)^2\right) + 2N^{-\xi} \qquad \text{on } \{\tau_K = s\}.$$

Therefore, another application of the tower property yields

$$\begin{aligned}
&\mathbb{P}\left(\tau_K \le \overline{T},\ \mathcal{R}_{\tau_K+2} \ne \emptyset \,\middle|\, \mathbf{y}_0\right) \\
&= \sum_{s=0}^{\overline{T}} \mathbb{E}\left[\mathbb{1}_{\{\{\tau_K = s\}\}} \mathbb{P}\left(\mathcal{R}_{s+2} \ne \emptyset \mid \mathcal{F}_s\right) \,\middle|\, \mathbf{y}_0\right] \le 2\exp\left(-(\log N)^2\right) + 2N^{-\xi}.
\end{aligned} \tag{B.18}$$

The all-blue configuration is absorbing. Hence

$$\{\mathcal{R}_{\overline{T}+2} \ne \emptyset\} \subseteq \{\tau_K > \overline{T}\} \cup \{\tau_K \le \overline{T},\ \mathcal{R}_{\tau_K+2} \ne \emptyset\}.$$

Combining (B.17) and (B.18) gives

$$\mathbb{P}\left(\tau_{\mathrm{B}} > \overline{T} + 2 \mid \mathbf{y}_0\right) \le 4\exp(-c\lambda_N) + 2\exp\left(-(\log N)^2\right) + 2N^{-\xi}. \tag{B.19}$$

Finally, a union bound with (B.11) gives

$$\begin{aligned}
&\mathbb{P}\left(\{\underline{T} + 1 \le \tau_{\mathrm{B}} \le \overline{T} + 2\}^{\complement} \,\middle|\, \mathbf{y}_0\right) \\
&\le 8\exp\left(-c\frac{\Delta_0^2 \log N}{N}\right) + 2\exp\left(-(\log N)^2\right) + 2N^{-\xi} = \eta_N(\Delta_0).
\end{aligned}$$

Taking complements proves (3.7). □

B.4 Proof of Corollary 3.3

Proof of Corollary 3.3. Suppose that

$$\Delta_0 \asymp \frac{\sqrt{N}}{h_N}.$$

Since $h_N \geq 1$ and $h_N = o(\sqrt{\log N})$,

$$\log\left(\frac{KN}{\Delta_0\sqrt{\log N}}\right) = \frac{1}{2}\log N - \frac{1}{2}\log\log N + \log h_N + O(1) = \frac{1}{2}\log N + O(\log\log N),$$

$$\log\left(C_2\sqrt{\log N}\right) = \frac{1}{2}\log\log N + O(1),$$

$$\log\left(c_1\sqrt{\log N}\right) = \frac{1}{2}\log\log N + O(1).$$

Substitution into (3.5a) gives

$$\underline{T} = (1+o(1))\,\frac{\log N}{\log\log N}.$$

Separately, substitution into (3.5b) gives

$$\overline{T} = (1+o(1))\,\frac{\log N}{\log\log N}.$$

For every fixed $0 < \delta < 1$, these two asymptotics imply, for all sufficiently large N,

$$\left\lfloor (1-\delta)\frac{\log N}{\log\log N}\right\rfloor \leq \underline{T}, \qquad \overline{T} + 2 \leq \left\lceil (1+\delta)\frac{\log N}{\log\log N}\right\rceil.$$

The all-blue configuration is absorbing, so (3.8) follows from (3.7). This completes the proof. □

C Proofs of Rapid Unanimity

We present the proof of Theorem 3.1 in this section. The proof proceeds backward from unanimity. We first identify a red-camp size below which all remaining red vertices disappear in one update. We then show that an advantage of order $N/\sqrt{\log N}$, with coefficient strictly above $K_2^{\mathrm{ER}}(b)$, reaches this extinction region in one update.

We first present two Lemmas needed, with their proofs deferred to Section C.2.

C.1 Reduction to the extinction regime

Lemma C.1 (A small red camp disappears in one day)**.** *Fix a deterministic time $t \geq 0$ and a constant $0 < r < r_*$, where r_* is defined in (3.1). If*

$$|\mathcal{R}_t| \leq rN, \tag{C.1}$$

then there exists $\xi = \xi(b,r) > 0$ such that, for all sufficiently large N,

$$\mathbb{P}\left(\mathcal{R}_{t+1} \neq \emptyset \,|\, \mathbf{y}_t\right) \leq 2N^{-\xi}. \tag{C.2}$$

For a fixed $K > K_2^{\mathrm{ER}}(b)$, define

$$q_K := \Phi(-\sqrt{b}\,K). \tag{C.3}$$

By (3.1), $K > K_2^{\mathrm{ER}}(b)$ is equivalent to $q_K < r_*$, so one can choose an extinction density strictly between these two quantities.

Lemma C.2 (One-step reduction below the extinction threshold)**.** *Fix a deterministic time $t \geq 0$, condition on $\mathbf{y}_t$, and suppose that for some fixed $K > K_2^{\mathrm{ER}}(b)$,*

$$\Delta_t \geq K\frac{N}{\sqrt{\log N}}. \tag{C.4}$$

Then, for any fixed r satisfying

$$q_K < r < r_*, \tag{C.5}$$

we have, for all sufficiently large N,

$$\mathbb{P}\left(|\mathcal{R}_{t+1}| \leq rN \,|\, \mathbf{y}_t\right) \geq 1 - 2\exp\left(-(\log N)^2\right). \tag{C.6}$$

C.2 Proofs of the reduction lemmas

C.2.1 Proof of Lemma C.1

Proof of Lemma C.1. Condition on $\mathbf{y}_t$ and assume (C.1). If $\mathcal{R}_t = \emptyset$, then no blue vertex has a red neighbor, and hence $\mathcal{R}_{t+1} = \emptyset$ deterministically. We may therefore assume $1 \le |\mathcal{R}_t| \le rN$, so $|\mathcal{B}_t| \ge (1-r)N$.

To use the uniform large-deviation estimate even when $|\mathcal{R}_t|$ is sublinear, we choose

$$m_N := \lfloor rN \rfloor.$$

By (D.2b) and (D.2a), both the probability that a red vertex stays red and the probability that a blue vertex flips to red are nondecreasing in the current red-camp size. Since $|\mathcal{R}_t|$ is an integer, $|\mathcal{R}_t| \le m_N$. It therefore suffices to bound the two probabilities when $|\mathcal{R}_t| = m_N$, which we assume for the remainder of this estimate.

In this worst-case configuration, the scaled camp sizes defined in (D.6) satisfy

$$b_t^{\mathcal{R}} = \frac{bm_N}{N}, \qquad b_t^{\mathcal{B}} = \frac{b(N-m_N)}{N}.$$

Since $m_N/N \to r \in (0, 1/2)$, this worst-case configuration satisfies (D.11), for example with $c_0 = r/2$, for all sufficiently large N. Corollary D.8 therefore gives

$$q_t^{\mathcal{R}} = N^{-I(b_t^{\mathcal{B}}, b_t^{\mathcal{R}})+o(1)}, \qquad p_t^{\mathcal{B}} = N^{-I(b_t^{\mathcal{B}}, b_t^{\mathcal{R}})+o(1)}.$$

Moreover,

$$I(b_t^{\mathcal{B}}, b_t^{\mathcal{R}}) = \left(\sqrt{b_t^{\mathcal{B}}} - \sqrt{b_t^{\mathcal{R}}}\right)^2 = b\left(\sqrt{1-r} - \sqrt{r}\right)^2 + o(1).$$

Because $r < r_*$, the definition of r_* implies that

$$\kappa := b\left(\sqrt{1-r} - \sqrt{r}\right)^2 - 1 > 0.$$

Thus, with $\xi := \kappa/2$, the uniform $o(1)$ terms in Corollary D.8 yield, for all sufficiently large N,

$$q_t^{\mathcal{R}} \le N^{-1-\xi}, \qquad p_t^{\mathcal{B}} \le N^{-1-\xi}.$$

By the preceding monotonicity reduction, these bounds hold for every original configuration satisfying $|\mathcal{R}_t| \le rN$. A vertex is red at time $t+1$ only if it was red and stayed red, or it was blue and flipped to red. The union bound therefore gives

$$\begin{aligned}\mathbb{P}\left(\mathcal{R}_{t+1} \neq \emptyset \,|\, \mathbf{y}_t\right) &\le |\mathcal{R}_t| q_t^{\mathcal{R}} + |\mathcal{B}_t| p_t^{\mathcal{B}} \\ &\le N \cdot N^{-1-\xi} + N \cdot N^{-1-\xi} = 2N^{-\xi}.\end{aligned}$$

The constant $\xi > 0$ depends only on b and r, which proves the lemma. □

C.2.2 Proof of Lemma C.2

Proof of Lemma C.2. Throughout the proof, condition on $\mathbf{y}_t$ and assume (C.4). Fix r satisfying (C.5), and define

$$\rho := \frac{q_K + r}{2}.$$

Then $q_K < \rho < r < r_*$. We claim that, uniformly over every configuration satisfying (C.4),

$$\mathbb{E}\left[|\mathcal{R}_{t+1}| \,|\, \mathbf{y}_t\right] \le \rho N + o(N). \tag{C.7}$$

Assuming (C.7), the desired high-probability bound follows directly from Lemma F.6, applied at time t with $s = \sqrt{N}\log N$:

$$\mathbb{P}\left(\left||\mathcal{R}_{t+1}| - \mathbb{E}\left[|\mathcal{R}_{t+1}| \,|\, \mathbf{y}_t\right]\right| \ge \sqrt{N}\log(N) \,\middle|\, \mathbf{y}_t\right) \le 2\exp\left(-(\log(N))^2\right).$$

Since $\rho < r$ and $\sqrt{N}\log N = o(N)$, (C.7) implies that $|\mathcal{R}_{t+1}| \le rN$ for all sufficiently large N, except with conditional probability at most $2\exp\left(-(\log N)^2\right)$.

It remains to prove (C.7). We divide the proof into two cases according to the size of $\mathcal{R}_t$.

Case 1: $|\mathcal{R}_t| \le \rho N$. In this case, the red camp is already below the intermediate density $\rho < r_*$. Applying Lemma C.1 with extinction density ρ, we obtain some $\xi_\rho > 0$ such that

$$\mathbb{E}\left[|\mathcal{R}_{t+1}| \,|\, \mathbf{y}_t\right] \le N\mathbb{P}\left(\mathcal{R}_{t+1} \neq \emptyset \,|\, \mathbf{y}_t\right) \le 2N^{1-\xi_\rho} = o(N).$$

This proves (C.7) in Case 1.

Case 2: $|\mathcal{R}_t| > \rho N$. Both camps have linear size. Since (C.4) implies $\Delta_t > 0$, (2.4) gives

$$|\mathcal{R}_t| = \frac{N - \Delta_t}{2} \le \frac{N}{2}, \qquad |\mathcal{B}_t| = \frac{N + \Delta_t}{2} \ge \frac{N}{2}.$$

Together with $|\mathcal{R}_t| > \rho N$, this allows us to apply Corollary D.6 uniformly in the present case.

Set $\gamma_N := \Delta_t/N$. Then $0 < \gamma_N \le 1$, and (C.4) implies

$$\gamma_N\sqrt{\log(N)} \ge K.$$

We first bound the probability $q_t^{\mathcal{R}}$ that a red vertex remains red. By (D.7),

$$m_t^{\mathcal{R}} = p(\Delta_t + 1) = b\gamma_N \log(N) + \frac{b\log(N)}{N}.$$

Moreover, the variance formula in the proof of Lemma D.1 gives

$$v_t^{\mathcal{R}} \le b\log N.$$

Combining all the estimates above, we obtain

$$x_t^{\mathcal{R}} \ge \sqrt{b}\,\gamma_N\sqrt{\log(N)} + o(1) \ge \sqrt{b}\,K + o(1).$$

Therefore, by (D.19b) in Corollary D.6,

$$q_t^{\mathcal{R}} \le \Phi(-x_t^{\mathcal{R}}) + O(1/\sqrt{\log(N)}) \le \Phi(-\sqrt{b}\,K + o(1)) + O(1/\sqrt{\log(N)}) = q_K + o(1).$$

We next bound $p_t^{\mathcal{B}}$, the probability that a blue vertex flips to red. By (D.7),

$$-m_t^{\mathcal{B}} = b\gamma_N \log(N) - \frac{b\log(N)}{N}.$$

The same variance computation gives

$$v_t^{\mathcal{B}} \le b\log N.$$

Combining the preceding displays gives

$$-x_t^{\mathcal{B}} \ge \sqrt{b}\,\gamma_N\sqrt{\log(N)} - o(1) \ge \sqrt{b}\,K - o(1).$$

Equivalently, $x_t^{\mathcal{B}} \le -\sqrt{b}\,K + o(1)$. Therefore, by (D.19c) in Corollary D.6,

$$p_t^{\mathcal{B}} \le \Phi(x_t^{\mathcal{B}}) + O(1/\sqrt{\log(N)}) \le \Phi(-\sqrt{b}\,K + o(1)) + O(1/\sqrt{\log(N)}) = q_K + o(1).$$

Combining $q_t^{\mathcal{R}} \le q_K + o(1)$ and $p_t^{\mathcal{B}} \le q_K + o(1)$ with (D.9b), we obtain

$$\begin{aligned}\mathbb{E}\left[|\mathcal{R}_{t+1}| \,|\, \mathbf{y}_t\right] &= |\mathcal{R}_t|q_t^{\mathcal{R}} + |\mathcal{B}_t|p_t^{\mathcal{B}} \\ &\le (q_K + o(1))N \le \rho N + o(N),\end{aligned}$$

where the last inequality uses $q_K < \rho$. This proves (C.7) in Case 2.

Therefore, (C.7) holds in both cases, completing the proof. □

C.3 Proof of Theorem 3.1

The preceding lemmas form a nested constant-time mechanism. Lemma C.1 eliminates a sufficiently small red camp in one update, while Lemma C.2 reaches that extinction region in one update from a sufficiently large advantage. We now combine these two steps to prove Theorem 3.1.

Proof of Theorem 3.1. We first prove one-day unanimity. Fix $0 < r < r_*$ and assume (3.3). By (2.4),

$$|\mathcal{R}_0| = \frac{N - \Delta_0}{2} \le rN.$$

Thus (C.1) holds at time $t = 0$. Lemma C.1 gives a constant $\xi = \xi(b, r) > 0$ such that

$$\mathbb{P}(\mathcal{R}_1 \neq \emptyset \mid \mathbf{y}_0) \le 2N^{-\xi}.$$

Taking complements proves part (ii).

We then prove the two-day unanimity case. Fix $K > K_2^{\mathrm{ER}}(b)$ and assume (3.2). Since $q_K < r_*$, we choose r as

$$r := \frac{q_K + r_*}{2},$$

which satisfies (C.5). Define the event

$$\mathcal{E}_1 := \{|\mathcal{R}_1| \le rN\}.$$

The hypothesis (3.2) is precisely (C.4) at time $t = 0$. Hence Lemma C.2 gives

$$\mathbb{P}(\mathcal{E}_1^c \mid \mathbf{y}_0) \le 2\exp\left(-(\log N)^2\right).$$

On the event $\mathcal{E}_1$, we have

$$\Delta_1 = N - 2|\mathcal{R}_1| \ge (1 - 2r)N.$$

Because the graph used for the update from time 1 to time 2 is freshly resampled independently of $\mathbf{y}_1$, the one-day result proved above applies conditionally at time $t = 1$. Hence there is a constant $\xi = \xi(b, r) > 0$ such that

$$\mathbb{P}(\mathcal{R}_2 \neq \emptyset \mid \mathbf{y}_1) \le 2N^{-\xi} \qquad \text{on } \mathcal{E}_1.$$

Taking conditional expectations with respect to $\mathbf{y}_1$ and separating according to $\mathcal{E}_1$, we obtain

$$\begin{aligned}\mathbb{P}(\mathcal{R}_2 \neq \emptyset \mid \mathbf{y}_0) &\le \mathbb{P}(\mathcal{E}_1^c \mid \mathbf{y}_0) + \mathbb{E}\left[\mathbb{1}_{\{\mathcal{E}_1\}}\mathbb{P}(\mathcal{R}_2 \neq \emptyset \mid \mathbf{y}_1) \,\middle|\, \mathbf{y}_0\right] \\ &\le 2\exp\left(-(\log N)^2\right) + 2N^{-\xi}.\end{aligned}$$

Because the chosen r depends only on b and K, we may write $\xi = \xi(b, K)$. Taking complements proves part (i) and completes the proof. □

D One-Step Probability Estimates

This appendix develops the local estimates used throughout the proofs: conditional degree differences and their moments, continuity-corrected Gaussian formulas, and large-deviation bounds for the final extinction step.

D.1 Degree differences and conditional expectations

In this section, we present the necessary preliminaries to study the evolution of the advantage. The proof of Lemma D.4 is deferred to Section E.

Recall the opinion update rule in (2.3). To facilitate our analysis, whenever the corresponding camp is nonempty, we introduce the degree differences for a blue vertex $u \in \mathcal{B}_t$ and a red vertex $v \in \mathcal{R}_t$:

$$\mathrm{D}_t^{\mathcal{B}}(u) := N_t^{\mathcal{R}}(u) - N_t^{\mathcal{B}}(u), \qquad \mathrm{D}_t^{\mathcal{R}}(v) := N_t^{\mathcal{B}}(v) - N_t^{\mathcal{R}}(v). \tag{D.1}$$

Conditioned on $\mathbf{y}_t$, u flips to red on day $t+1$ if and only if $\mathrm{D}_t^{\mathcal{B}}(u) > 0$, while v flips to blue on day $t+1$ if and only if $\mathrm{D}_t^{\mathcal{R}}(v) > 0$. Since all edges in $\mathcal{G}_{t+1}$ have the same probability p, the laws of these

differences depend only on the current camp sizes. For convenience, we omit the vertex argument and write

$$\mathrm{D}_t^{\mathcal{B}} \overset{d}{=} \mathbf{Bin}(|\mathcal{R}_t|, p) - \mathbf{Bin}(|\mathcal{B}_t| - 1, p), \tag{D.2a}$$

$$\mathrm{D}_t^{\mathcal{R}} \overset{d}{=} \mathbf{Bin}(|\mathcal{B}_t|, p) - \mathbf{Bin}(|\mathcal{R}_t| - 1, p). \tag{D.2b}$$

We then introduce the flip probabilities as follows whenever the corresponding camp is nonempty:

$$p_t^{\mathcal{B}} \coloneqq \mathbb{P}(\mathrm{D}_t^{\mathcal{B}} > 0 \mid \mathbf{y}_t), \qquad q_t^{\mathcal{B}} \coloneqq \mathbb{P}(\mathrm{D}_t^{\mathcal{B}} \le 0 \mid \mathbf{y}_t), \tag{D.3a}$$

$$p_t^{\mathcal{R}} \coloneqq \mathbb{P}(\mathrm{D}_t^{\mathcal{R}} > 0 \mid \mathbf{y}_t), \qquad q_t^{\mathcal{R}} \coloneqq \mathbb{P}(\mathrm{D}_t^{\mathcal{R}} \le 0 \mid \mathbf{y}_t), \tag{D.3b}$$

where $p_t^{\mathcal{R}}$ (resp. $p_t^{\mathcal{B}}$) denotes the probability that a red (resp. blue) vertex flips to blue (resp. red) on day $t+1$, while $q_t^{\mathcal{R}}$ (resp. $q_t^{\mathcal{B}}$) denotes the probability that a red (resp. blue) vertex stays red (resp. blue) on day $t+1$. If a camp is empty, we set its flip probability to zero and its stay probability to one; with these conventions, the camp-size expectation identities below remain valid at unanimity. When both camps are nonempty, we define the means and variances of $\mathrm{D}_t^{\mathcal{B}}$ and $\mathrm{D}_t^{\mathcal{R}}$ via

$$m_t^{\mathcal{B}} \coloneqq \mathbb{E}\left[\mathrm{D}_t^{\mathcal{B}} \mid \mathbf{y}_t\right], \qquad v_t^{\mathcal{B}} \coloneqq \mathbb{V}\mathrm{ar}\left(\mathrm{D}_t^{\mathcal{B}} \mid \mathbf{y}_t\right), \tag{D.4a}$$

$$m_t^{\mathcal{R}} \coloneqq \mathbb{E}\left[\mathrm{D}_t^{\mathcal{R}} \mid \mathbf{y}_t\right], \qquad v_t^{\mathcal{R}} \coloneqq \mathbb{V}\mathrm{ar}\left(\mathrm{D}_t^{\mathcal{R}} \mid \mathbf{y}_t\right). \tag{D.4b}$$

We also define the standardized means

$$x_t^{\mathcal{R}} \coloneqq \frac{m_t^{\mathcal{R}}}{\sqrt{v_t^{\mathcal{R}}}}, \qquad x_t^{\mathcal{B}} \coloneqq \frac{m_t^{\mathcal{B}}}{\sqrt{v_t^{\mathcal{B}}}}. \tag{D.5}$$

We also define the scaled camp sizes

$$b_t^{\mathcal{R}} \coloneqq \frac{b|\mathcal{R}_t|}{N}, \qquad b_t^{\mathcal{B}} \coloneqq \frac{b|\mathcal{B}_t|}{N}. \tag{D.6}$$

Lemma D.1 (Expectations and variances of the degree differences)**.** *Under Assumption 2.1, whenever both camps are nonempty, the expected degree differences on day t are*

$$m_t^{\mathcal{R}} = p(\Delta_t + 1), \qquad m_t^{\mathcal{B}} = p(1 - \Delta_t). \tag{D.7}$$

Furthermore, the variances of the degree differences on day t satisfy

$$v_t^{\mathcal{R}} = v_t^{\mathcal{B}} = (N-1)p(1-p) \asymp \log(N). \tag{D.8}$$

Proof of Lemma D.1. The proof follows directly from linearity of expectation and variance. By (D.2b),

$$m_t^{\mathcal{R}} = p|\mathcal{B}_t| - p(|\mathcal{R}_t| - 1) = p(\Delta_t + 1),$$
$$v_t^{\mathcal{R}} = \big(|\mathcal{B}_t| + |\mathcal{R}_t| - 1\big)p(1-p) = (N-1)p(1-p).$$

Similarly, by (D.2a),

$$m_t^{\mathcal{B}} = p|\mathcal{R}_t| - p(|\mathcal{B}_t| - 1) = p(1 - \Delta_t),$$
$$v_t^{\mathcal{B}} = \big(|\mathcal{R}_t| + |\mathcal{B}_t| - 1\big)p(1-p) = (N-1)p(1-p).$$

By Assumption 2.1, the common variance is asymptotic to $b \log N$. □

The camp-size identities in (2.4), together with (D.1), give the following conditional expectations.

Lemma D.2 (Conditional expectations of the next day's camp sizes and advantage)**.** *For all $t \ge 0$, the conditional expectations of the camp sizes on the next day can be expressed as*

$$\mathbb{E}\left[|\mathcal{B}_{t+1}| \,|\, \mathbf{y}_t\right] = |\mathcal{R}_t|p_t^{\mathcal{R}} + |\mathcal{B}_t|q_t^{\mathcal{B}}, \tag{D.9a}$$

$$\mathbb{E}\left[|\mathcal{R}_{t+1}| \,|\, \mathbf{y}_t\right] = |\mathcal{R}_t|q_t^{\mathcal{R}} + |\mathcal{B}_t|p_t^{\mathcal{B}}. \tag{D.9b}$$

Furthermore, the conditional expectation of the advantage on the next day can be expressed as

$$\mathbb{E}\left[\Delta_{t+1} \,|\, \mathbf{y}_t\right] = \Delta_t + 2\left(|\mathcal{R}_t|p_t^{\mathcal{R}} - |\mathcal{B}_t|p_t^{\mathcal{B}}\right). \tag{D.10a}$$

$$= N\big(p_t^{\mathcal{R}} - p_t^{\mathcal{B}}\big) + \Delta_t\big(1 - p_t^{\mathcal{R}} - p_t^{\mathcal{B}}\big). \tag{D.10b}$$

Proof of Lemma D.2. Condition on $\mathbf{y}_t$. By the update rule in (2.3) and the degree differences in (D.1), a red vertex $v \in \mathcal{R}_t$ contributes to $\mathcal{B}_{t+1}$ if and only if $\mathrm{D}_t^{\mathcal{R}}(v) > 0$, while it contributes to $\mathcal{R}_{t+1}$ if and only if $\mathrm{D}_t^{\mathcal{R}}(v) \le 0$. Similarly, a blue vertex $u \in \mathcal{B}_t$ contributes to $\mathcal{R}_{t+1}$ if and only if $\mathrm{D}_t^{\mathcal{B}}(u) > 0$, while it contributes to $\mathcal{B}_{t+1}$ if and only if $\mathrm{D}_t^{\mathcal{B}}(u) \le 0$. Therefore,

$$|\mathcal{B}_{t+1}| = \sum_{v\in\mathcal{R}_t} \mathbb{1}_{\{\mathrm{D}_t^{\mathcal{R}}(v)>0\}} + \sum_{u\in\mathcal{B}_t} \mathbb{1}_{\{\mathrm{D}_t^{\mathcal{B}}(u)\le 0\}},$$
$$|\mathcal{R}_{t+1}| = \sum_{v\in\mathcal{R}_t} \mathbb{1}_{\{\mathrm{D}_t^{\mathcal{R}}(v)\le 0\}} + \sum_{u\in\mathcal{B}_t} \mathbb{1}_{\{\mathrm{D}_t^{\mathcal{B}}(u)>0\}}.$$

Taking conditional expectations and using (D.3b) and (D.3a), we obtain (D.9a) and (D.9b).

We now compute the conditional expectation of the advantage. By (2.4), (D.9a), and (D.9b),

$$\begin{aligned}\mathbb{E}\left[\Delta_{t+1} \,|\, \mathbf{y}_t\right] &= \mathbb{E}\left[|\mathcal{B}_{t+1}| - |\mathcal{R}_{t+1}| \,|\, \mathbf{y}_t\right] \\ &= |\mathcal{R}_t| p_t^{\mathcal{R}} + |\mathcal{B}_t| q_t^{\mathcal{B}} - \left(|\mathcal{R}_t| q_t^{\mathcal{R}} + |\mathcal{B}_t| p_t^{\mathcal{B}}\right) \\ &= \Delta_t + 2|\mathcal{R}_t| p_t^{\mathcal{R}} - 2|\mathcal{B}_t| p_t^{\mathcal{B}} \\ &= \Delta_t + 2\left(|\mathcal{R}_t| p_t^{\mathcal{R}} - |\mathcal{B}_t| p_t^{\mathcal{B}}\right),\end{aligned}$$

where the third equality uses $q_t^{\mathcal{R}} = 1 - p_t^{\mathcal{R}}$ and $q_t^{\mathcal{B}} = 1 - p_t^{\mathcal{B}}$. This proves (D.10a). It remains only to rewrite the same expression in terms of Δ_t and N. From (2.4), $2|\mathcal{R}_t| = N - \Delta_t$ and $2|\mathcal{B}_t| = N + \Delta_t$. Substituting these identities into (D.10a) yields

$$\begin{aligned}\mathbb{E}\left[\Delta_{t+1} \,|\, \mathbf{y}_t\right] &= \Delta_t + (N - \Delta_t)\, p_t^{\mathcal{R}} - (N + \Delta_t)\, p_t^{\mathcal{B}} \\ &= N\left(p_t^{\mathcal{R}} - p_t^{\mathcal{B}}\right) + \Delta_t\left(1 - p_t^{\mathcal{R}} - p_t^{\mathcal{B}}\right),\end{aligned}$$

which is (D.10b). This completes the proof. □

Therefore, to understand the one-step evolution of the advantage, we need to understand the flip probabilities $p_t^{\mathcal{R}}$ and $p_t^{\mathcal{B}}$ in a very detailed manner. A natural idea is to approximate them using a Gaussian distribution via the Berry–Esseen theorem below.

Lemma D.3 (Berry–Esseen Esseen, 1956)**.** *Let $\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_n$ be independent random variables with zero means, variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$, respectively, and finite absolute third moments $\mathbb{E}[|\mathrm{X}_j|^3] = \rho_j < \infty$. Define the normalized random variable $\mathrm{S}_n := \sum_{j=1}^n \mathrm{X}_j / \sqrt{\sum_{j=1}^n \sigma_j^2}$. Let $F_n(x)$ and $\Phi(x)$ denote the cumulative distribution functions of S_n and the standard normal distribution, respectively. There exists a universal constant $\mathcal{C}_{\mathrm{BE}}$ such that, for any n,*

$$\sup_{x\in\mathbb{R}} |F_n(x) - \Phi(x)| \le \mathcal{C}_{\mathrm{BE}} \sum_{j=1}^n \rho_j \Big(\sum_{j=1}^n \sigma_j^2\Big)^{-3/2},$$

However, a straightforward application of Lemma D.3 results in an error term of order $O((\log N)^{-1/2})$, which neither accounts for the half-integer lattice correction nor achieves the sharper $O((\log N)^{-1})$ error required in (D.13a) and (D.13b). To address this, we instead employ the Poisson approximation, as developed in Section E below. Under the current sparsity Assumption 2.1, the Poisson approximation is more accurate than the final approximation requires. Indeed, Le Cam's inequality and maximal coupling couple $\mathbf{Bin}(m, p)$ and $\mathbf{Pois}(mp)$ with mismatch probability at most mp^2. The two binomial variables in either one-vertex degree difference contain $N - 1$ trials in total, so two independent couplings and a union bound give $(N-1)p^2$ as an upper bound, which is

$$O\left(\frac{(\log N)^2}{N}\right) = o\left(\frac{1}{\log N}\right),$$

under Assumption 2.1. Thus the coupling error is smaller than the accuracy required in Lemma D.4. The remaining task is to obtain an $O((\log N)^{-1})$ continuity-corrected approximation for the corresponding difference of independent Poisson variables.

Lemma D.4 (Continuity-corrected Gaussian estimates in the one-vertex Gaussian window)**.** *Recall $x_t^{\mathcal{R}}, x_t^{\mathcal{B}}$ from (D.5). Fix $c_0 > 0$ and $M < \infty$, condition on $\mathbf{y}_t$, and suppose that*

$$|\mathcal{R}_t| \ge c_0 N, \qquad |\mathcal{B}_t| \ge c_0 N, \tag{D.11}$$

while $x_t^{\mathcal{R}}$ and $x_t^{\mathcal{B}}$ in (D.5) satisfy

$$|x_t^{\mathcal{R}}| \le M, \qquad |x_t^{\mathcal{B}}| \le M. \tag{D.12}$$

Then, uniformly over such configurations,

$$p_t^{\mathcal{R}} = \Phi\left(x_t^{\mathcal{R}} - \frac{1}{2\sqrt{v_t^{\mathcal{R}}}}\right) + O\left(\frac{1}{\log N}\right), \tag{D.13a}$$

$$p_t^{\mathcal{B}} = \Phi\left(x_t^{\mathcal{B}} - \frac{1}{2\sqrt{v_t^{\mathcal{B}}}}\right) + O\left(\frac{1}{\log N}\right). \tag{D.13b}$$

For $\Delta_t \ge 1$, there are constants $c, C > 0$ such that

$$c\frac{p\Delta_t}{\sqrt{(N-1)p(1-p)}} \le p_t^{\mathcal{R}} - p_t^{\mathcal{B}} \le C\frac{p\Delta_t}{\sqrt{(N-1)p(1-p)}}. \tag{D.14}$$

Furthermore, for every fixed $C_0 < \infty$, when $0 \le \Delta_t \le C_0\sqrt{N/\log N}$, we have

$$p_t^{\mathcal{R}} - p_t^{\mathcal{B}} = \sqrt{\frac{2}{\pi}}\frac{p\Delta_t}{\sqrt{(N-1)p(1-p)}} + O\left(\frac{p\Delta_t}{\log N}\right). \tag{D.15}$$

All implicit constants and the constants c, C depend only on b, c_0, M, and additionally on C_0 in (D.15).

D.2 Gaussian and large-deviation estimates for sparse binomial differences

Under Assumption 2.1, let

$$\mathrm{D} := \mathrm{X} - \mathrm{Y}, \qquad \mathrm{X} \sim \mathbf{Bin}(s, p), \qquad \mathrm{Y} \sim \mathbf{Bin}(r, p), \tag{D.16}$$

where X and Y are independent and $r = r(N), s = s(N)$ are nonnegative integers at most N. Set

$$A_N := \frac{br}{N}, \qquad B_N := \frac{bs}{N}, \qquad \mu_N := p(s-r), \qquad \sigma_N^2 := p(1-p)(s+r). \tag{D.17}$$

The results below are obtained by specializing (Dumitriu et al., 2026, Lemma A.1, Corollary A.2, Lemma A.3, and Corollary A.4): the two edge probabilities there are both set equal to p, and the two logarithmic density constants are both set equal to b. We state the resulting ER estimates without repeating their proofs.

D.2.1 Gaussian window: Berry–Esseen approximation

The next two statements are the ER specializations of (Dumitriu et al., 2026, Lemma A.1 and Corollary A.2).

Lemma D.5 (Gaussian approximation for sparse binomial differences)**.** *Suppose that $A_N + B_N \ge c_0$ for some fixed constant $c_0 > 0$. Then, uniformly over all r, s satisfying this condition,*

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left(\frac{\mathrm{D}-\mu_N}{\sigma_N} \le x\right) - \Phi(x)\right| \le \frac{C}{\sqrt{\log N}}, \tag{D.18}$$

where $C = C(b, c_0) > 0$.

Corollary D.6 (Gaussian estimates for one-vertex flip probabilities)**.** *Recall $p_t^{\mathcal{R}}, q_t^{\mathcal{R}}$ from (D.3b), $p_t^{\mathcal{B}}, q_t^{\mathcal{B}}$ from (D.3a), $x_t^{\mathcal{R}}, x_t^{\mathcal{B}}$ from (D.5). Fix $t \ge 0$, and condition on a configuration $\mathbf{y}_t$ satisfying (D.11) for some fixed constant $c_0 > 0$. Uniformly over all such configurations,*

$$p_t^{\mathcal{R}} = \Phi(x_t^{\mathcal{R}}) + O(1/\sqrt{\log N}), \tag{D.19a}$$

$$q_t^{\mathcal{R}} = \Phi(-x_t^{\mathcal{R}}) + O(1/\sqrt{\log N}), \tag{D.19b}$$

$$p_t^{\mathcal{B}} = \Phi(x_t^{\mathcal{B}}) + O(1/\sqrt{\log N}), \tag{D.19c}$$

$$q_t^{\mathcal{B}} = \Phi(-x_t^{\mathcal{B}}) + O(1/\sqrt{\log N}). \tag{D.19d}$$

The constants in the $O(\cdot)$ terms depend only on b and c_0.

D.2.2 LARGE-DEVIATION TAIL: RATE-FUNCTION EXPONENTS

Recall $\mathbf{ReLU}(z) = \max\{0, z\}$, and define

$$I(x, y) := \left(\mathbf{ReLU}\left(\sqrt{x} - \sqrt{y}\right)\right)^2. \tag{D.20}$$

The next two statements are the ER specializations of (Dumitriu et al., 2026, Lemma A.3 and Corollary A.4).

Lemma D.7 (Sparse-binomial large deviations)**.** *Suppose that $A_N, B_N \geq c_0$ for some fixed constant $c_0 > 0$. Then, uniformly over all r, s satisfying this condition,*

$$\mathbb{P}(\mathrm{D} > 0) = N^{-I(A_N, B_N)+o(1)}, \tag{D.21a}$$

$$\mathbb{P}(\mathrm{D} \leq 0) = N^{-I(B_N, A_N)+o(1)}. \tag{D.21b}$$

Corollary D.8 (Large-deviation estimates for one-vertex flip probabilities)**.** *Recall $b_t^{\mathcal{R}}, b_t^{\mathcal{B}}$ from (D.6). Fix $t \geq 0$, condition on $\mathbf{y}_t$, and suppose that (D.11) holds for some fixed constant $c_0 > 0$. Then, uniformly over all such configurations,*

$$p_t^{\mathcal{B}} = N^{-I(b_t^{\mathcal{B}}, b_t^{\mathcal{R}})+o(1)}, \tag{D.22a}$$

$$q_t^{\mathcal{R}} = N^{-I(b_t^{\mathcal{B}}, b_t^{\mathcal{R}})+o(1)}, \tag{D.22b}$$

$$p_t^{\mathcal{R}} = N^{-I(b_t^{\mathcal{R}}, b_t^{\mathcal{B}})+o(1)}, \tag{D.22c}$$

$$q_t^{\mathcal{B}} = N^{-I(b_t^{\mathcal{R}}, b_t^{\mathcal{B}})+o(1)}. \tag{D.22d}$$

The missing self-edge changes the relevant camp size by one and hence its scaled parameter by b/N, which is absorbed by the uniform $o(1)$ terms.

E CONTINUITY-CORRECTED GAUSSIAN ESTIMATES VIA POISSON APPROXIMATION

In this section, we present the proof of Lemma D.4 in Section D.1, together with the necessary technical preparations.

The proof of Lemma D.4 proceeds in two steps. First, each sparse binomial variable in a one-vertex degree difference is coupled to a Poisson variable with the same mean, reducing the desired flip probability estimates to estimates for a difference of two independent Poisson variables. Second, a continuity-corrected Gaussian approximation for that Poisson difference is transferred back through the coupling. This yields the required Gaussian formulas while retaining the half-integer correction caused by the lattice support.

The current section develops those two ingredients. We first recall total variation and maximal coupling, and then combine them with Le Cam's inequality to couple a difference of independent binomial variables to the corresponding Poisson difference. Finally, we establish the continuity-corrected distributional and local estimates for that Poisson difference. Throughout, we use the standing edge probability p from Assumption 2.1 and the one-step notation from Section D.1.

E.1 TOTAL VARIATION AND MAXIMAL COUPLING

For probability measures μ and ν on a countable space $\mathcal{X}$, denote the total variation distance by

$$d_{\mathrm{TV}}(\mu, \nu) := \sup_{A \subseteq \mathcal{X}} |\mu(A) - \nu(A)|.$$

Theorem E.1 (Le Cam's inequality Le Cam, 1960; Steele, 1994)**.** *Let $\xi_1, \ldots, \xi_n$ be independent Bernoulli random variables with $\mathbb{P}(\xi_i = 1) = p_i$. We denote*

$$W := \sum_{i=1}^{n} \xi_i, \qquad \lambda := \sum_{i=1}^{n} p_i,$$

and let $P \sim \mathbf{Pois}(\lambda)$. *Then*

$$\sum_{k=0}^{\infty} \left| \mathbb{P}(W = k) - e^{-\lambda} \frac{\lambda^k}{k!} \right| \leq 2 \sum_{i=1}^{n} p_i^2. \tag{E.1}$$

Equivalently, under the previous total-variation convention, we have

$$d_{\mathrm{TV}}\big(\mathcal{L}(W), \mathcal{L}(P)\big) \leq \sum_{i=1}^{n} p_i^2, \tag{E.2}$$

where $\mathcal{L}(W)$ *and* $\mathcal{L}(P)$ *denote the laws of* W *and* P, *respectively.*

Lemma E.2 (Maximal coupling)**.** *Let* μ *and* ν *be probability measures on a countable space* $\mathcal{X}$. *There is a coupling* (U, V) *with* $\mathcal{L}(U) = \mu$ *and* $\mathcal{L}(V) = \nu$ *such that*

$$\mathbb{P}(U \neq V) = d_{\mathrm{TV}}(\mu, \nu) = \frac{1}{2} \sum_{x \in \mathcal{X}} |\mu(x) - \nu(x)|. \tag{E.3}$$

Proof of Lemma E.2. Denote $r(x) := \min\{\mu(x), \nu(x)\}$ and $q := \sum_x r(x) = 1 - d_{\mathrm{TV}}(\mu, \nu)$. With probability q, draw $U = V$ from the probability mass function r/q. On the complementary event, draw U and V from the normalized residual masses $(\mu - r)/(1 - q)$ and $(\nu - r)/(1 - q)$, respectively. The two residual supports are disjoint, so $U \neq V$ on this event. This coupling has the prescribed marginals and satisfies (E.3); the cases $q \in \{0, 1\}$ follow by the same construction with the empty component omitted. □

Lemma E.3 (Le Cam coupling)**.** *Let* $X \sim \mathbf{Bin}(N_X, p)$ *and* $Y \sim \mathbf{Bin}(N_Y, p)$ *be independent random variables. There exists a joint realization* $(\widehat{X}, \widehat{Y}, P_X, P_Y)$, *where the marginals* $\widehat{X} \overset{d}{=} X$, $\widehat{Y} \overset{d}{=} Y$ *are independent, and* $P_X \sim \mathbf{Pois}(pN_X)$, $P_Y \sim \mathbf{Pois}(N_Y p)$ *are independent, such that*

$$\mathbb{P}\left(\widehat{X} - \widehat{Y} \neq P_X - P_Y\right) \leq (N_X + N_Y)p^2. \tag{E.4}$$

Moreover,

$$\sup_{A \subseteq \mathbb{Z}} |\mathbb{P}(X - Y \in A) - \mathbb{P}(P_X - P_Y \in A)| \leq (N_X + N_Y)p^2. \tag{E.5}$$

Proof of Lemma E.3. We write X as the sum of N_X independent $\mathbf{Ber}(p)$ variables. By applying Theorem E.1 with $p_i = p$, for some $P_X \sim \mathbf{Pois}(pN_X)$, we have

$$d_{\mathrm{TV}}\left(\mathcal{L}(X), \mathbf{Pois}(pN_X)\right) \leq p^2 N_X.$$

Lemma E.2 thus yields a joint realization $(\widehat{X}, P_X)$ where $\widehat{X} \overset{d}{=} X$ such that

$$\mathbb{P}(\widehat{X} \neq P_X) \leq p^2 N_X.$$

Similarly, a joint realization $(\widehat{Y}, P_Y)$ can be constructed such that

$$\mathbb{P}(\widehat{Y} \neq P_Y) \leq p^2 N_Y,$$

where $\widehat{Y} \overset{d}{=} Y$ and $P_Y \sim \mathbf{Pois}(N_Y p)$. These two coupled pairs are drawn independently, which subsequently preserves the required independence within $(\widehat{X}, \widehat{Y})$ and within (P_X, P_Y). Moreover,

$$\left\{\widehat{X} - \widehat{Y} \neq P_X - P_Y\right\} \subseteq \{\widehat{X} \neq P_X\} \cup \{\widehat{Y} \neq P_Y\}.$$

The union bound proves (E.4). For every $A \subseteq \mathbb{Z}$, the two indicators of membership in A agree whenever the coupled differences agree. Hence their expectation difference is at most the mismatch probability, which proves (E.5). □

E.2 Continuity correction for a Poisson difference

Lemma E.4 (Continuity correction for a Poisson difference)**.** *Let* $P_X \sim \mathbf{Pois}(\lambda_X)$ *and* $P_Y \sim \mathbf{Pois}(\lambda_Y)$ *be independent, and set*

$$S \coloneqq P_X - P_Y, \qquad \lambda \coloneqq \lambda_X + \lambda_Y, \qquad m \coloneqq \lambda_X - \lambda_Y.$$

Suppose that for some fixed $M > 0$, $|m| \le M\sqrt{\lambda}$ *as* $\lambda \to \infty$*. Then uniformly over integers* k, ℓ*,*

$$\mathbb{P}(S \le \ell) = \Phi\left(\frac{\ell + 1/2 - m}{\sqrt{\lambda}}\right) + O_M(\lambda^{-1}), \tag{E.6a}$$

$$\mathbb{P}(S = k) = \frac{1}{\sqrt{\lambda}}\varphi\left(\frac{k - m}{\sqrt{\lambda}}\right) + O_M(\lambda^{-3/2}). \tag{E.6b}$$

Proof of Lemma E.4. Let U be uniform on $[-1/2, 1/2]$, independently of S, and define

$$Y \coloneqq \frac{S + U - m}{\sqrt{\lambda}}.$$

For every integer ℓ, the exact identity below follows from the continuity of U,

$$\mathbb{P}(S \le \ell) = \mathbb{P}\left(Y \le \frac{\ell + 1/2 - m}{\sqrt{\lambda}}\right), \tag{E.7}$$

which then leads to the half-integer continuity correction. We then denote

$$\operatorname{sinc}(x) \coloneqq \begin{cases} \sin(x)/x, & x \neq 0, \\ 1, & x = 0. \end{cases}$$

By independence of P_X, P_Y, and U, the characteristic function of Y factors as

$$\begin{aligned} \psi_\lambda(u) &\coloneqq \mathbb{E}\exp\left(\frac{\mathrm{i}u(S + U - m)}{\sqrt{\lambda}}\right) \\ &= \exp\left(-\frac{\mathrm{i}um}{\sqrt{\lambda}}\right)\mathbb{E}\exp\left(\frac{\mathrm{i}uP_X}{\sqrt{\lambda}}\right)\mathbb{E}\exp\left(-\frac{\mathrm{i}uP_Y}{\sqrt{\lambda}}\right)\mathbb{E}\exp\left(\frac{\mathrm{i}uU}{\sqrt{\lambda}}\right). \end{aligned}$$

Here we denote the imaginary unit by $\mathrm{i} = \sqrt{-1}$. The Poisson characteristic-function identity $\mathbb{E}e^{\mathrm{i}zP} = \exp\{\theta(e^{\mathrm{i}z} - 1)\}$ for $P \sim \mathbf{Pois}(\theta)$ gives

$$\mathbb{E}\exp\left(\frac{\mathrm{i}uP_X}{\sqrt{\lambda}}\right) = \exp\left\{\lambda_X\left(e^{\mathrm{i}u/\sqrt{\lambda}} - 1\right)\right\},$$

$$\mathbb{E}\exp\left(-\frac{\mathrm{i}uP_Y}{\sqrt{\lambda}}\right) = \exp\left\{\lambda_Y\left(e^{-\mathrm{i}u/\sqrt{\lambda}} - 1\right)\right\}.$$

Moreover, since U is uniform on $[-1/2, 1/2]$,

$$\mathbb{E}\exp\left(\frac{\mathrm{i}uU}{\sqrt{\lambda}}\right) = \int_{-1/2}^{1/2} e^{\mathrm{i}ux/\sqrt{\lambda}}\,dx = \frac{\sin\left(u/(2\sqrt{\lambda})\right)}{u/(2\sqrt{\lambda})} = \operatorname{sinc}\left(\frac{u}{2\sqrt{\lambda}}\right),$$

where the value at $u = 0$ follows by continuity. Combining these three factors and using $m = \lambda_X - \lambda_Y$, we obtain

$$\begin{aligned} \psi_\lambda(u) &= \exp\left\{-\frac{\mathrm{i}u(\lambda_X - \lambda_Y)}{\sqrt{\lambda}} + \lambda_X\left(e^{\mathrm{i}u/\sqrt{\lambda}} - 1\right) + \lambda_Y\left(e^{-\mathrm{i}u/\sqrt{\lambda}} - 1\right)\right\}\operatorname{sinc}\left(\frac{u}{2\sqrt{\lambda}}\right) \\ &= \exp\left\{\lambda_X\left(e^{\mathrm{i}u/\sqrt{\lambda}} - 1 - \frac{\mathrm{i}u}{\sqrt{\lambda}}\right) + \lambda_Y\left(e^{-\mathrm{i}u/\sqrt{\lambda}} - 1 + \frac{\mathrm{i}u}{\sqrt{\lambda}}\right)\right\}\operatorname{sinc}\left(\frac{u}{2\sqrt{\lambda}}\right). \end{aligned} \tag{E.8}$$

Fix a sufficiently large constant A, depending only on M, and let

$$T_\lambda \coloneqq \sqrt{A\log\lambda}$$

denote the truncation threshold.

Claim E.5. *Uniformly for* $|u| \le T_\lambda$,

$$\psi_\lambda(u) = e^{-u^2/2}\left[1 + O_M\left(\frac{u^2 + |u|^3 + u^4}{\lambda}\right)\right]. \tag{E.9}$$

We defer the proof of the claim until the end of the proof.

Since Y has a continuous distribution, Fourier inversion for distribution functions gives

$$\mathbb{P}(Y \le x) - \Phi(x) = -\frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ux}\frac{\psi_\lambda(u) - e^{-u^2/2}}{\mathrm{i}u}\,du. \tag{E.10}$$

The integral is absolutely convergent; the estimates below also verify this fact. On $|u| \le T_\lambda$, Claim E.5 implies

$$\int_{|u|\le T_\lambda}\frac{|\psi_\lambda(u) - e^{-u^2/2}|}{|u|}\,du \le \frac{C_M}{\lambda}. \tag{E.11}$$

It remains to control the noncentral Fourier aliases. For

$$I_j := \left[(2j-1)\pi\sqrt{\lambda}, (2j+1)\pi\sqrt{\lambda}\right], \qquad j \in \mathbb{Z},$$

write $u = 2\pi j\sqrt{\lambda} + v$, where $|v| \le \pi\sqrt{\lambda}$. The modulus of the exponential factor in (E.8) is

$$\exp\left\{\lambda\left(\cos(v/\sqrt{\lambda}) - 1\right)\right\} \le e^{-cv^2}$$

for an absolute $c > 0$. On the central interval I_0, this gives

$$\int_{I_0\cap\{|u|>T_\lambda\}}\frac{|\psi_\lambda(u)|}{|u|}\,du \le Ce^{-cT_\lambda^2}. \tag{E.12}$$

For $j \ne 0$, the zero of the sinc factor at $2\pi j\sqrt{\lambda}$ gives

$$\left|\operatorname{sinc}\left(\frac{u}{2\sqrt{\lambda}}\right)\right| \le \frac{C|v|}{|j|\sqrt{\lambda}}, \qquad |u| \ge c|j|\sqrt{\lambda}.$$

Consequently,

$$\int_{I_j}\frac{|\psi_\lambda(u)|}{|u|}\,du \le \frac{C}{j^2\lambda}\int_{\mathbb{R}}|v|e^{-cv^2}\,dv \le \frac{C}{j^2\lambda}. \tag{E.13}$$

After summing over $j \ne 0$, choosing A sufficiently large in (E.12), and using

$$\int_{|u|>T_\lambda}\frac{e^{-u^2/2}}{|u|}\,du \le Ce^{-T_\lambda^2/2},$$

we obtain

$$\int_{|u|>T_\lambda}\frac{|\psi_\lambda(u) - e^{-u^2/2}|}{|u|}\,du \le \frac{C_M}{\lambda}.$$

Together with (E.10) and (E.11), this proves

$$\sup_{x\in\mathbb{R}}|\mathbb{P}(Y \le x) - \Phi(x)| \le \frac{C_M}{\lambda}.$$

Equation (E.6a) now follows from (E.7).

For the local estimate, omit the uniform smoothing factor and let

$$\psi_\lambda^{(0)}(u) := \exp\left\{\lambda_X\left(e^{\mathrm{i}\frac{u}{\sqrt{\lambda}}} - 1 - \frac{\mathrm{i}u}{\sqrt{\lambda}}\right) + \lambda_Y\left(e^{-\mathrm{i}\frac{u}{\sqrt{\lambda}}} - 1 + \frac{\mathrm{i}u}{\sqrt{\lambda}}\right)\right\}.$$

Fourier inversion on one lattice period gives

$$\mathbb{P}(S = k) = \frac{1}{2\pi\sqrt{\lambda}}\int_{-\pi\sqrt{\lambda}}^{\pi\sqrt{\lambda}} e^{-\mathrm{i}u(k-m)/\sqrt{\lambda}}\psi_\lambda^{(0)}(u)\,du.$$

On $|u| \le T_\lambda$, the same expansion as above, without the $O(u^2/\lambda)$ sinc term, gives

$$\psi_\lambda^{(0)}(u) = e^{-u^2/2}\left[1 + O_M\left(\frac{|u|^3 + u^4}{\lambda}\right)\right].$$

The contribution of this error is $O_M(\lambda^{-3/2})$. On the remainder of the lattice period,

$$|\psi_\lambda^{(0)}(u)| \le e^{-cu^2},$$

and the Gaussian Fourier tail satisfies the same bound. Increasing A, if necessary, makes both tail contributions $O_M(\lambda^{-3/2})$. Extending the Gaussian integral to $\mathbb{R}$ therefore yields

$$\mathbb{P}(S = k) = \frac{1}{2\pi\sqrt{\lambda}} \int_{\mathbb{R}} e^{-\mathrm{i}u(k-m)/\sqrt{\lambda}} e^{-u^2/2}\, du + O_M(\lambda^{-3/2}),$$

which is (E.6b).

It remains to prove Claim E.5. Taylor's formula in the exponent of (E.8) gives, uniformly for $|u| \le T_\lambda$,

$$\begin{aligned}
&\lambda_X\left(e^{\mathrm{i}u/\sqrt{\lambda}} - 1 - \frac{\mathrm{i}u}{\sqrt{\lambda}}\right) + \lambda_Y\left(e^{-\mathrm{i}u/\sqrt{\lambda}} - 1 + \frac{\mathrm{i}u}{\sqrt{\lambda}}\right) \\
&\quad = -\frac{u^2}{2} - \frac{\mathrm{i}mu^3}{6\lambda^{3/2}} + \frac{u^4}{24\lambda} + O\left(\frac{|u|^5}{\lambda^{3/2}}\right).
\end{aligned}$$

Here the quadratic term uses $\lambda_X + \lambda_Y = \lambda$, while the cubic term uses $\lambda_X - \lambda_Y = m$. Since $|m| \le M\sqrt{\lambda}$ and $T_\lambda = O(\sqrt{\log \lambda})$, the preceding display becomes

$$\begin{aligned}
&\lambda_X\left(e^{\mathrm{i}u/\sqrt{\lambda}} - 1 - \frac{\mathrm{i}u}{\sqrt{\lambda}}\right) + \lambda_Y\left(e^{-\mathrm{i}u/\sqrt{\lambda}} - 1 + \frac{\mathrm{i}u}{\sqrt{\lambda}}\right) \\
&\quad = -\frac{u^2}{2} + O_M\left(\frac{|u|^3 + u^4}{\lambda}\right).
\end{aligned} \tag{E.14}$$

The error in (E.14) is $o(1)$ uniformly on the central range. Therefore,

$$\begin{aligned}
&\exp\left\{\lambda_X\left(e^{\mathrm{i}u/\sqrt{\lambda}} - 1 - \frac{\mathrm{i}u}{\sqrt{\lambda}}\right) + \lambda_Y\left(e^{-\mathrm{i}u/\sqrt{\lambda}} - 1 + \frac{\mathrm{i}u}{\sqrt{\lambda}}\right)\right\} \\
&\quad = e^{-u^2/2}\left[1 + O_M\left(\frac{|u|^3 + u^4}{\lambda}\right)\right].
\end{aligned} \tag{E.15}$$

Finally, Taylor expansion at the origin gives

$$\operatorname{sinc}\left(\frac{u}{2\sqrt{\lambda}}\right) = 1 - \frac{u^2}{24\lambda} + O\left(\frac{u^4}{\lambda^2}\right) = 1 + O\left(\frac{u^2}{\lambda}\right) \tag{E.16}$$

uniformly for $|u| \le T_\lambda$. Multiplying (E.15) and (E.16) proves (E.9). The same calculation without the sinc factor gives the unsmoothed expansion used in the local estimate, which completes the proof. □

E.2.1 Proof of Lemma D.4

Proof of Lemma D.4. Condition throughout on $\mathbf{y}_t$. We first prove (D.13a). By (D.2b), write

$$\mathrm{D}_t^{\mathcal{R}} = X_t^{\mathcal{R}} - Y_t^{\mathcal{R}}, \qquad X_t^{\mathcal{R}} \sim \mathbf{Bin}(|\mathcal{B}_t|, p), \qquad Y_t^{\mathcal{R}} \sim \mathbf{Bin}(|\mathcal{R}_t| - 1, p),$$

with the two binomial variables independent. Let

$$P_{X,t}^{\mathcal{R}} \sim \mathbf{Pois}(|\mathcal{B}_t|p), \qquad P_{Y,t}^{\mathcal{R}} \sim \mathbf{Pois}((|\mathcal{R}_t| - 1)p)$$

be independent Poisson variables, and denote

$$S_t^{\mathcal{R}} := P_{X,t}^{\mathcal{R}} - P_{Y,t}^{\mathcal{R}}, \qquad \lambda_t := (N-1)p.$$

The mean of $S_t^{\mathcal{R}}$ is $m_t^{\mathcal{R}}$, while its variance is λ_t. Applying Lemma E.3 with $N_X = |\mathcal{B}_t|$ and $N_Y = |\mathcal{R}_t| - 1$ gives

$$\sup_{A \subseteq \mathbb{Z}} \left| \mathbb{P}(\mathrm{D}_t^{\mathcal{R}} \in A \mid \mathbf{y}_t) - \mathbb{P}(S_t^{\mathcal{R}} \in A) \right| \le C(N-1)p^2 = O\left(\frac{(\log N)^2}{N}\right). \tag{E.17}$$

Recall that

$$v_t^{\mathcal{R}} = (N-1)p(1-p) = \lambda_t(1-p), \qquad \lambda_t \asymp \log N.$$

The assumed bound on $m_t^{\mathcal{R}}/\sqrt{v_t^{\mathcal{R}}}$ therefore implies a fixed bound on $m_t^{\mathcal{R}}/\sqrt{\lambda_t}$. Applying (E.6a) with $\ell = 0$ gives

$$\mathbb{P}(S_t^{\mathcal{R}} \le 0) = \Phi\left(\frac{1/2 - m_t^{\mathcal{R}}}{\sqrt{\lambda_t}}\right) + O\left(\frac{1}{\log N}\right).$$

Moreover,

$$\left| (1/2 - m_t^{\mathcal{R}}) \left(\frac{1}{\sqrt{\lambda_t}} - \frac{1}{\sqrt{v_t^{\mathcal{R}}}} \right) \right| \le C_M p = o\left(\frac{1}{\log N}\right).$$

Since Φ is Lipschitz and the coupling error in (E.17) is $o((\log N)^{-1})$, we conclude that

$$\mathbb{P}(\mathrm{D}_t^{\mathcal{R}} \le 0 \mid \mathbf{y}_t) = \Phi\left(\frac{1/2 - m_t^{\mathcal{R}}}{\sqrt{v_t^{\mathcal{R}}}}\right) + O\left(\frac{1}{\log N}\right).$$

Taking complements and using $1 - \Phi(x) = \Phi(-x)$ proves (D.13a). The proof of (D.13b) is identical, starting from (D.2a).

We next prove the estimates for $p_t^{\mathcal{R}} - p_t^{\mathcal{B}}$. Assume $\Delta_t \ge 1$, since both conclusions are immediate when $\Delta_t = 0$. For each integer $0 \le j \le \Delta_t$, let $\mathrm{D}_{t,j}^{\mathcal{R}}$ and $\mathrm{D}_{t,j}^{\mathcal{B}}$ have the respective distributions

$$\mathrm{D}_{t,j}^{\mathcal{R}} \overset{d}{=} \mathbf{Bin}(|\mathcal{R}_t| + j, p) - \mathbf{Bin}(|\mathcal{R}_t| - 1, p),$$
$$\mathrm{D}_{t,j}^{\mathcal{B}} \overset{d}{=} \mathbf{Bin}(|\mathcal{R}_t|, p) - \mathbf{Bin}(|\mathcal{R}_t| - 1 + j, p),$$

with independent binomial variables in each difference. Their means are $p(j+1)$ and $p(1-j)$, and both have variance $(2|\mathcal{R}_t| + j - 1)p(1-p)$.

The standardized-mean assumption in (D.12) gives

$$p\Delta_t \le C_M \sqrt{(N-1)p(1-p)}.$$

Consequently, uniformly over $0 \le j \le \Delta_t$,

$$(2|\mathcal{R}_t| + j - 1)p \asymp \log N, \qquad \frac{|p(j+1)| + |p(1-j)|}{\sqrt{(2|\mathcal{R}_t| + j - 1)p}} \le C_M.$$

For each j, Lemma E.3 replaces the two binomial variables by independent Poisson variables with total mean $(2|\mathcal{R}_t| + j - 1)p$ and coupling error at most Np^2. Applying (E.6b), and then replacing the Poisson variance by the binomial variance, gives, uniformly in $0 \le j \le \Delta_t$ and $k \in \{0, 1\}$,

$$\mathbb{P}(\mathrm{D}_{t,j}^{\mathcal{R}} = k) = \frac{\varphi\left(\dfrac{k - p(j+1)}{\sqrt{(2|\mathcal{R}_t| + j - 1)p(1-p)}}\right)}{\sqrt{(2|\mathcal{R}_t| + j - 1)p(1-p)}} + O_M\left(\frac{1}{\log N}\right),$$

$$\mathbb{P}(\mathrm{D}_{t,j}^{\mathcal{B}} = k) = \frac{\varphi\left(\dfrac{k - p(1-j)}{\sqrt{(2|\mathcal{R}_t| + j - 1)p(1-p)}}\right)}{\sqrt{(2|\mathcal{R}_t| + j - 1)p(1-p)}} + O_M\left(\frac{1}{\log N}\right).$$

Indeed, the local Poisson error is $O_M((\log N)^{-3/2})$, the coupling error is

$$Np^2 = O\left(\frac{(\log N)^2}{N}\right) = o\left(\frac{1}{\log N}\right),$$

and replacing p by $p(1-p)$ in the variance changes the Gaussian expression by $O_M(p/\sqrt{\log N})$.

At $j = 0$, the two interpolating differences have the same distribution. Adding one independent $\mathbf{Ber}(p)$ trial at each successive value of j gives

$$\mathbb{P}(\mathrm{D}^{\mathcal{R}}_{t,j+1} > 0) - \mathbb{P}(\mathrm{D}^{\mathcal{R}}_{t,j} > 0) = p\mathbb{P}(\mathrm{D}^{\mathcal{R}}_{t,j} = 0),$$
$$\mathbb{P}(\mathrm{D}^{\mathcal{B}}_{t,j} > 0) - \mathbb{P}(\mathrm{D}^{\mathcal{B}}_{t,j+1} > 0) = p\mathbb{P}(\mathrm{D}^{\mathcal{B}}_{t,j} = 1).$$

Summing these identities from $j = 0$ to $j = \Delta_t - 1$, and using the equality in distribution at $j = 0$, yields

$$p^{\mathcal{R}}_t - p^{\mathcal{B}}_t = p \sum_{j=0}^{\Delta_t - 1} \left[\mathbb{P}(\mathrm{D}^{\mathcal{R}}_{t,j} = 0) + \mathbb{P}(\mathrm{D}^{\mathcal{B}}_{t,j} = 1)\right].$$

If $0 \le \Delta_t \le C_0\sqrt{N/\log N}$, the standardized means in this identity converge to zero uniformly, so each paired local mass equals

$$\sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{(N-1)p(1-p)}} + O\left(\frac{1}{\log N}\right).$$

Substitution proves (D.15). In the full Gaussian window, the standardized arguments remain in a fixed compact interval. Each local mass is therefore bounded above and below by positive constant multiples of $((N-1)p(1-p))^{-1/2}$, which proves (D.14). □

F Auxiliary Probabilistic Tools

We record the external central limit theorem and concentration inequality used in the main argument, with their roles and assumptions stated explicitly.

F.1 The Berkowitz–Devlin One-Step Central Limit Theorem

For completeness, we record the external central limit theorem used in Section A. The formulation below is a restatement of Berkowitz & Devlin (2022, Theorem 1).

Theorem F.1 (Berkowitz–Devlin one-step CLT)**.** *Let $\mathcal{R}_0, \mathcal{B}_0$ be a deterministic partition of $[N]$. Independently sample $\mathcal{G} \sim \mathbb{G}(N,p)$ and perform one synchronous majority update with ties retaining their current colors. Define*

$$\sigma_N := \sqrt{Np(1-p)}.$$

For independent random variables

$$W_R \sim \mathbf{Bin}(|\mathcal{R}_0|, p), \qquad W_B \sim \mathbf{Bin}(|\mathcal{B}_0|, p),$$

set

$$\chi_N := \mathbb{P}(W_R \ge W_B)\mathbb{P}(W_R \le W_B).$$

Then the following conclusions hold.

(i) *For arbitrary parameter choices,*

$$\mathbb{V}\mathrm{ar}(|\mathcal{R}_1|) = N\chi_N + O\left(|\Delta_0| + \frac{N}{\sigma_N}\right). \tag{F.1}$$

(ii) *Suppose that*

$$\sigma_N \longrightarrow \infty, \qquad \log(1/\chi_N) = o(\log \sigma_N).$$

Then

$$\mathbb{V}\mathrm{ar}(|\mathcal{R}_1|) \sim N\chi_N,$$

and, for every fixed $a < c$,

$$\mathbb{P}\left(a \le \frac{|\mathcal{R}_1| - \mathbb{E}|\mathcal{R}_1|}{\sqrt{N\chi_N}} \le c\right) \longrightarrow \Phi(c) - \Phi(a). \tag{F.2}$$

In particular, the assumptions and conclusions in (ii) *hold if* $\sigma_N \to \infty$ *and*

$$|\Delta_0|p = O(\sigma_N).$$

Remark F.2 (Qualitative nature of the one-step CLT)**.** Theorem F.1 is a qualitative central limit theorem: its method-of-moments proof gives convergence to the standard normal distribution but does not provide a rate in Kolmogorov distance. This is sufficient for our critical-window winner-selection result, which only requires the uniform one-step approximation error to be $o(1)$. In Lemma A.2, we therefore keep this qualitative error separate from the explicit $O((\log N)^{-1/2})$ errors coming from the sparse-binomial approximation, the variance normalization, and the conditional mean.

F.2 Read-k Chernoff Bound

We first introduce the concept of a read-k family.

Definition F.3 (Read-k family)**.** Let $\mathrm{X}_1, \ldots, \mathrm{X}_m$ be independent random variables. For each $j \in [r]$, let $P_j \subseteq [m]$, and let f_j be a Boolean function of $\{\mathrm{X}_i\}_{i \in P_j}$. Suppose that each X_i is read by at most k of the functions f_j, i.e., $\big|\{j \in [r] : i \in P_j\}\big| \le k$ for every $i \in [m]$. Then, the random variables $\{\mathrm{Y}_j \coloneqq f_j(\{\mathrm{X}_i\}_{i \in P_j})\}_{j \in [r]}$ form a read-k family.

Then we state the read-k Chernoff bound.

Lemma F.4 (Read-k Chernoff bound Gavinsky et al., 2015, Theorem 1.1)**.** *Let* $\{\mathrm{Y}_j\}_{j \in [r]}$ *be a read-*k *family of indicator random variables with* $p_j \coloneqq \mathbb{P}(\mathrm{Y}_j = 1)$*, and denote their average by* $\bar{p} \coloneqq r^{-1} \sum_{j=1}^{r} p_j$*. If* $0 < \bar{p} < 1$*, then*

$$\mathbb{P}\left(\sum_{j=1}^{r} \mathrm{Y}_j \ge (\bar{p} + \varepsilon) r \right) \le \exp\left(-\frac{r}{k} \mathbb{D}_{\mathrm{KL}}(\bar{p} + \varepsilon \| \bar{p}) \right), \qquad 0 < \varepsilon \le 1 - \bar{p}, \tag{F.3a}$$

$$\mathbb{P}\left(\sum_{j=1}^{r} \mathrm{Y}_j \le (\bar{p} - \varepsilon) r \right) \le \exp\left(-\frac{r}{k} \mathbb{D}_{\mathrm{KL}}(\bar{p} - \varepsilon \| \bar{p}) \right), \qquad 0 < \varepsilon \le \bar{p}. \tag{F.3b}$$

where the Kullback-Leibler divergence is defined as

$$\mathbb{D}_{\mathrm{KL}}(u \| v) \coloneqq u \log(u/v) + (1-u) \log((1-u)/(1-v)).$$

We use the standard convention $0 \log 0 = 0$ *at* $u \in \{0, 1\}$*. If* $\bar{p} \in \{0, 1\}$*, all the indicators are almost surely constant and the corresponding tail statements are trivial.*

Corollary F.5 (Centered concentration for a read-k sum Dumitriu et al., 2026)**.** *Let* $\{\mathrm{Y}_j\}_{j \in [r]}$ *be a read-*k *family of indicator random variables, and set* $\mathrm{S} \coloneqq \sum_{j=1}^{r} \mathrm{Y}_j$*. Then, for every* $s > 0$*,*

$$\mathbb{P}\left(|\mathrm{S} - \mathbb{E}\mathrm{S}| \ge s \right) \le 2 \exp\left(-\frac{2s^2}{kr} \right). \tag{F.4}$$

Lemma F.6 (Read-2 concentration for one-step red camp and advantage Dumitriu et al., 2026)**.** *Fix* $t \ge 0$ *and condition on* $\mathbf{y}_t$*. Then, for every* $u > 0$*,*

$$\mathbb{P}\left(\big| |\mathcal{R}_{t+1}| - \mathbb{E}[|\mathcal{R}_{t+1}| \mid \mathbf{y}_t] \big| \ge u \,\middle|\, \mathbf{y}_t \right) \le 2 \exp(-u^2/N), \tag{F.5a}$$

$$\mathbb{P}\left(|\Delta_{t+1} - \mathbb{E}[\Delta_{t+1} \mid \mathbf{y}_t]| \ge u \,\middle|\, \mathbf{y}_t \right) \le 2 \exp\left(-\frac{u^2}{4N} \right). \tag{F.5b}$$

Proof. Condition on $\mathbf{y}_t$. For each vertex, the indicator that it is red at time $t+1$ is a Boolean function of the independent edge indicators incident to that vertex. Each edge indicator is therefore read by at most its two endpoints, so these N next-color indicators form a read-2 family. Applying Corollary F.5 with $r = N$ and $k = 2$ to their sum $|\mathcal{R}_{t+1}|$ gives (F.5a). Finally, $\Delta_{t+1} = N - 2|\mathcal{R}_{t+1}|$, so

$$|\Delta_{t+1} - \mathbb{E}[\Delta_{t+1} \mid \mathbf{y}_t]| = 2 \, \big| |\mathcal{R}_{t+1}| - \mathbb{E}[|\mathcal{R}_{t+1}| \mid \mathbf{y}_t] \big|.$$

Applying the first bound with threshold $u/2$ proves (F.5b). □

G Related Literature

G.1 Static Erdős–Rényi update graphs

Majority dynamics on an update graph sampled from $\mathbb{G}(N,p)$ and then held fixed is governed jointly by the density p, the initial advantage Δ_0, and the initial coloring scheme. For an i.i.d. initial vertex coloring with bias δ_N, Zehmakan (2020) proved high-probability unanimity for $p \geq (1+\varepsilon)\log N/N$, with fixed $\varepsilon > 0$, whenever $\delta_N \gg (Np)^{-1/2}$, equivalently $|\Delta_0| \gg \sqrt{N/p}$. Under the unbiased random initial coloring, in which every vertex is independently assigned either color with probability $1/2$, the initial advantage is typically of order $\sqrt{N}$. Benjamini et al. (2016) obtained winning probability at least 0.4 within four updates for $p \geq CN^{-1/2}$, and Fountoulakis et al. (2020) subsequently strengthened this to a high-probability result. The sufficient density was then lowered progressively by Chakraborti et al. (2023) to $p \geq CN^{-3/5}\log N$, by Kim & Tran (2025) to $p \gg N^{-2/3}(\log N)^{2/3}$, and by Jaffe (2025) to $p \geq CN^{-2/3}$, where C denotes a sufficiently large constant that may differ between results. More recently, Goel & Sah (2026) proved that, for every fixed $\varepsilon > 0$, unbiased random initial opinions reach unanimity with high probability within $O(1/\varepsilon)$ rounds whenever $pN \geq N^{\varepsilon}$. This extends the static-graph results to every fixed positive power of N for the expected degree. It does not cover the logarithmic expected degree $pN = b\log N$ considered here, or the Gaussian winner-selection law for prescribed initial camp sizes in our critical window.

For prescribed initial camp sizes and fixed p, Tran & Vu (2023) showed that a constant advantage can attain any fixed winning probability below one, while Berkowitz & Devlin (2022) showed that $\Delta_0 = 3$ on $\mathbb{G}(N,1/2)$ already gives a winning probability of at least 51%. Sah & Sawhney (2025) obtained the exact asymptotic winning probability

$$\Phi\left(\Delta_0\sqrt{\frac{2p}{\pi(1-p)}}\right) + o(1),$$

together with three-step unanimity for $(\log N)^{-1/16} \leq p \leq 1-(\log N)^{-1/16}$ and $0 < \Delta_0 \leq (\log N)^{1/4}$, thereby establishing the *power-of-one* phenomenon for fixed p.

For prescribed camp sizes and vanishing p, Tran & Vu (2025) proved that an advantage satisfying $|\Delta_0| \geq 20/p$ in our full-gap convention leads to unanimity within $O(\log_{Np} N)$ updates throughout the regime $(1+c)\log N \leq pN \leq N-10$ for some constant $c > 0$. Kim & Tran (2025) improved the sufficient scale to $|\Delta_0| \gg N^{-1/2}p^{-3/2}\log N$ for $N^{-1}(\log N)^2 \ll p \ll N^{-1/2}(\log N)^{1/4}$. More recently, Jaffe (2025) proved that throughout the regime $(1+c)\log N \leq pN \leq N-10$, the sufficient scale, where the majority dynamics reaches unanimity within $O(\log_{Np} N)$ updates, is

$$\Delta_0 \geq \max\left\{p^{-1/2}\exp\left(A\sqrt{\log(1/p)}\right),\ Bp^{-3/2}N^{-1/2}\right\}$$

for some suitable constants $A, B > 0$. These results establish a sparse *power-of-few* phenomenon, but the conjectured scale $|\Delta_0| \asymp p^{-1/2}$ (Tran & Vu, 2025) remains open. At $p = b\log N/N$, the sufficient static bound in that work has order $N/(\log N)^{3/2}$ (Jaffe, 2025, Theorem 4.1), rather than the conjectured $p^{-1/2} \asymp \sqrt{N/\log N}$.

Conjecture G.1 (Optimal power of few, Tran & Vu, 2025, Conjecture 1.11)**.** *There is a function f satisfying $f(x) > 1/2$ for $x > 0$ and $\lim_{x\to\infty} f(x) = 1$ with the following property: majority dynamics on $\mathbb{G}(N,p)$ with initial advantage Δ_0 for blue and $p \geq (1+c)\log(N)/N$ for some constant $c > 0$ ends unanimously blue with probability $f(\Delta_0\sqrt{p}) - o(1)$.*

For the dense regime where $p \geq (\log N)^{-1/16}$, a corollary of Sah & Sawhney (2025) verifies Conjecture G.1. However, for smaller values of p, the conjecture remains unresolved.

Our Gaussian selection law concerns independently resampled graphs. It establishes the corresponding critical scale for this model, while the sparse static-graph conjecture remains a distinct problem because the opinions become correlated with the graph reused in subsequent updates.

G.2 Resampled stochastic block models

An alternative choice for modeling the interactions between agents is the *Stochastic Block Model* (SBM), where vertices within the same community are connected with probability α, and vertices in

different communities are connected with probability β. For fixed $\alpha, \beta \in (0, 1]$, Wang et al. (2022) studied a Markovian model that resamples every edge and a non-Markovian model that resamples an edge only when one of its endpoints changes opinion. They proved a *power-of-one* winner-selection result for the Markovian model; for the non-Markovian model, they located, up to the second-order term, the initial-lead threshold between halting and unanimity within three updates and provided sufficient thresholds for unanimity in one or two updates.

In the sparse regime where $\alpha = a \log N/N$ and $\beta = b \log N/N$ for some constants $a, b > 0$, Dumitriu et al. (2026) analyzed majority dynamics under the assortative assumption ($a > b$). In that context, the progression toward unanimity is determined by the weighted advantage $\widetilde{\Delta}_t := b|\mathcal{B}_t| - a|\mathcal{R}_t|$, and new phenomena in the pace toward blue unanimity are uncovered, exhibiting constant-time, subpolynomial-time, and polynomial-time scales.

The homogeneous ER model considered here assigns the same edge probability to every pair of vertices. The sparse assortative analysis assumes $a > b$ and does not directly cover the case $a = b$. Our local comparison near balance resolves the smaller $p^{-1/2}$ initial-advantage scale and supports uniform Gaussian winner selection and two-sided consensus-time bounds. We reuse general one-step probability and concentration tools where applicable, with their sources identified in the preceding appendices.

G.3 Other models for opinion dynamics

In the population-protocol model, Alistarh et al. (2015) considered finite-state agents that interact in uniformly random pairs and seek exact consensus on the initial majority. Their *Average-and-Conquer* protocol provides a tunable state–time tradeoff and achieves polylogarithmic expected parallel time when the per-agent state space is sufficiently large relative to the initial bias.

Podder & Roy (2026) studied a growing directed preferential-attachment network in which each arriving agent samples past agents, adopts one of two immutable opinions through a reinforced stochastic rule, and links to sampled agents sharing that opinion. For fixed and suitably growing sample sizes, they characterized the almost-sure asymptotics of opinion shares, influence capital, and network activity; for fixed sample size, they also obtained second-order fluctuations.

Zimper et al. (2026) considered agent-based systems with continuous social positions and scalar opinions, allowing either one-way influence from social positions to opinions or two-way feedback. They derived reduced SPDEs for the resulting cluster dynamics and showed numerically that these reproduce the underlying systems at substantially lower computational cost, including in an application to General Social Survey data.

G.4 Spatial interaction graphs

The Erdős–Rényi (ER) model treats vertices as statistically interchangeable, whereas a *Random Geometric Graph* (RGG), denoted by $\mathbb{G}(N, p, d)$, embeds vertices as points in a d-dimensional feature space and introduces geometric dependence and local bottlenecks. For fixed $p \in (0, 1)$, Bubeck et al. (2016) proved that the total variation distance $\mathrm{TV}(\mathbb{G}(N, p), \mathbb{G}(N, p, d))$ converges to 1 when $d \ll N^3$, whereas the same distance tends uniformly to zero when $d \gg N^3$. For the sparse regime $p = c/N$, the same work proved distinguishability in terms of total variation distance when $d \ll (\log N)^3$, while Brennan et al. (2020), and subsequently Liu et al. (2022) established equivalence also in terms of total variation distance for $d \gg N^{3/2}$ and $d \gg (\log N)^{36}$, respectively, narrowing the gap toward the conjectured threshold $d \asymp (\log N)^3$. At the spectral level, Cao & Zhu (2025) proved that when $d \gtrsim Np \log(1/p)$ and $Np \to \infty$, the *Empirical Spectral Distribution* (ESD) of the normalized adjacency matrix $\mathbf{A}/\sqrt{Np}$ under $\mathbb{G}(N, p, d)$ converges to the semicircle law; and when $d \gg \log N$, $p = c/N$ for some fixed $c > 0$, the ESD of $\mathbf{A}/\sqrt{c}$ converges to the limiting spectral distribution under $\mathbb{G}(N, c/N)$.

The preceding results indicate that RGGs in sufficiently high dimensions behave similarly to ER graphs, while those in lower dimensions preserve geometric features and can display more intricate dynamics. This leads us to investigate how spatial locality influences both the consensus threshold and the rate at which consensus is reached in the RGG model relative to the ER benchmark.