BIROn - Birkbeck Institutional Research Online

ASYMPTOTIC BEHAVIOUR OF THE NOISY VOTER MODEL DENSITY PROCESS

BY RICHARD PYMAR[1] AND NICOLÁS RIVERA[2]

[1]*Department of Economics, Mathematics and Statistics, Birkbeck, University of London, London, WC1E 7HX, UK,* *r.pymar@bbk.ac.uk*

[2]*Instituto de Ingeniería Matemática, Universidad de Valparaíso, Valparaíso, 2362905, Chile, n.a.rivera.aburto@gmail.com*

Given a transition matrix P indexed by a finite set V of vertices, the voter model is a discrete-time Markov chain in $\{0,1\}^V$ where at each time-step a randomly chosen vertex x imitates the opinion of vertex y with probability $P(x,y)$. The noisy voter model is a variation of the voter model in which vertices may change their opinions by the action of an external noise. The strength of this noise is measured by an extra parameter $p \in [0,1]$.

In this work we analyse the density process, defined as the stationary mass of vertices with opinion 1, i.e. $S_t = \sum_{x \in V} \pi(x)\xi_t(x)$, where π is the stationary distribution of P, and $\xi_t(x)$ is the opinion of vertex x at time t. We investigate the asymptotic behaviour of S_t when t tends to infinity for different values of the noise parameter p. In particular, by allowing P and p to be functions of the size $|V|$, we show that, under appropriate conditions and small enough p a normalised version of S_t converges to a Gaussian random variable, while for large enough p, S_t converges to a Bernoulli random variable. We provide further analysis of the noisy voter model on a variety of specific graphs including the complete graph, cycle, torus and hypercube, where we identify the critical rate p (depending on the size $|V|$) that separates these two asymptotic behaviours.

1. Introduction.

1.1. *Background.* The voter model is a classical interacting particle system, first described independently in [11] and [23]. It has found applications across the sciences including statistical physics [7], social sciences [18], chemistry [27], ecology [38], and sociophysics [19]. The process can be described simply: a set of vertices (or sites) V are initially each endowed with an opinion (0 or 1) and vertices update their opinions over time. At the update time of a vertex x, an imitation step occurs: x changes its opinion to that of a randomly chosen vertex (according to some transition matrix $P(x,\cdot)$, usually the transition matrix of the simple random walk on a graph G). The voter model can evolve in discrete or continuous time: in discrete time at each time step one vertex is chosen (uniformly) to update its opinion, whereas in the continuous time setting each vertex updates at the ring times of an independent exponential clock. After its conception, the voter model attracted the attention of probabilists (and later theoretical computer scientists) and several results have been proved in both the finite-vertex set and infinite-vertex set regimes. In this paper we only consider the finite-vertex setting, where the focus has predominantly been on estimating the consensus time defined as the time at which all vertices have the same opinion [4, 12, 13, 22, 25, 34].

Variations of the voter model are ubiquitous, and have been studied to investigate robustness to perturbations in the dynamics, or to allow for increased applicability. These variations include models with multiple opinions [43], the existence of one or more zealots/stubborn

vertices [24], constrained voter models [28], the antivoter model [39], and the process on dynamic graphs [3] or random graphs [42].

In this work we focus on analysing the noisy voter model introduced in [21] (and independently in [26]), which is a modification of the voter model in which the opinions are affected by an external noise. The dynamics are very similar to the voter model: upon selecting a vertex x for updating, with some probability p, x independently re-randomizes its opinion (choosing either opinion 0 or 1 with equal probability), otherwise (with probability $1 - p$) x performs an imitation step. Note that by choosing $p = 0$ we recover the standard voter model, whereas with $p = 1$ only re-randomization occurs. The noise allows for more interesting large-time behaviour in finite systems, as configurations of complete consensus are no longer absorbing. The external noise causes spontaneous opinion changes and provides an opposing force to the standard voter dynamics which seeks consensus.

To the best of our knowledge, for finite vertex set V, the probability literature has so far only considered mixing times. Suppose $|V| = n$ and let P be a transition matrix. If $p = 1$ (only re-randomization occurs) the process is equivalent to a lazy Markov chain on a hypercube on 2^n vertices, and so the mixing time is $\Theta(n \log n)$. When the noise parameter $p \in (0, 1)$ is constant and P is the transition matrix of a random walk on a graph, [37] shows that in the continuous time setting (where imitation steps occur at rate 1 and re-randomization at constant rate) the mixing time of the noisy voter model is $O(\log n)$. This translates to a $O(n \log n)$ mixing time for the discrete case. Recently Cox, Peres and Steif [15] proved that, under appropriate conditions, the mixing time of the noisy voter model (in discrete time[1]) is $n \log n / (2p)$, and that it exhibits total variation cut-off. The noisy voter model has been studied more extensively in the physics literature. Of particular relevance here is the observation – obtained through numerical analyses – of the possible existence of a critical rate for the noise parameter such that if $p = p(n)$ tends to 0 slower than the critical rate then the noise term dominates the dynamics, whereas if p tends to 0 faster than this rate the dynamics is closer to the voter model [6, 35]. We provide theoretical results demonstrating this phenomenon and conjecture its precise nature for general settings (Conjecture 7).

Our focus is on a particular function of the noisy voter model known as the *density process* in the probability literature [10, 14] and the *link magnetisation* [36] or the *degree-weighted magnetisation* [45] in the physics literature. Consider a transition matrix P on n vertices, let $p \in [0, 1]$ and $\xi_t(v)$ denote the opinion of vertex $v \in V$ at time t in the noisy voter model (formal definitions appear in the next section). The time t state of the density process is defined as $S_t := \sum_{x \in V} \pi(x) \xi_t(x)$ where π is the stationary distribution of P (assuming it exists). Thus the density process models the evolution of the weight (according to π) of opinion 1. Our motivation for choosing the pre-factor $\pi(x)$ is that, for the non-noisy voter model, the quantity S_t is then conserved under the dynamics, that is, $\mathbb{E}[S_{t+1} \mid S_t] = S_t$ almost surely. The quantity $\frac{1}{n} \sum_x \xi_t(x)$ would be natural in the context of edge-updates (i.e. instead of selecting a vertex for opinion-updating, we select an edge) as it is conserved under this dynamic. Further, if we assume that P is reversible and that each vertex has a spin value $+1$ or -1 (instead of an opinion), then the magnetisation of an edge e is the average spin of both endpoints, and the *link magnetisation* is the expected magnetisation of a random edge (x, y) chosen with probability $\pi(x) P(x, y)$. When P is the (lazy) transition matrix of a random walk on a graph, the link magnetisation is just the expected average spin of a randomly chosen edge.

The analysis of the evolution of S_t over time, and especially its large-time behaviour, provides insight into the dependency of the noisy voter model on the noise parameter p, allowing

---

[1] their main results are presented in continuous time, but for comparison purposes we present their result in discrete time

us to identify different regimes as p varies in $[0, 1]$. In particular, we show a transition from an ordered regime for p small to a disordered regime for p large, which gives support to many of the results empirically found in [36], especially when P is the transition matrix of a simple random walk on a large graph.

Our results are expressed in terms of $S = \lim_{t \to \infty} S_t$ (which exists in distribution). We prove that for p sufficiently small (depending on the transition matrix P), S converges (as $|V|$ tends to infinity) to a Bernoulli distribution (Proposition 6), indicating that in equilibrium almost all vertices have the same opinion as in the usual voter model. Conversely we show that for p sufficiently large (converging to a non-zero constant or to zero sufficiently slowly) S (normalised appropriately) converges to a Gaussian random variable (Theorem 1, Corollaries 3, 4, and 5) and so in particular there are positive proportions (according to the stationary distribution π of P) of each opinion present at equilibrium. As far as we are aware, our work is the first to study the asymptotic behaviour of the noisy voter model in which the noise parameter p can be unbounded from below when the size of the graph tends to infinity.

Throughout, P is chosen as the transition matrix of the simple random walk on a graph G, and our results are expressed in terms of the graph structure. For several graph families we obtain the precise critical rate $p_c = p_c(|V|)$, in the sense that if $p \ll p_c$ then S converges to Bernoulli, and $p \gg p_c$ implies that (normalised) S converges to Gaussian (Proposition 8). In all cases we find that the critical probability p_c is proportional to the inverse of the meeting time of two independent walks (started from stationarity). Obtaining the critical rate for the cycle and the 2D torus is particularly challenging and requires special treatment and careful analysis (Proposition 9). The cycle is particularly interesting as on the cycle the noisy voter model is identical to the stochastic Ising model, see for example [15] (the relationship between the inverse temperature β and the noise parameter p is $4\beta = \log(2/p - 1)$.)

There are two key ideas behind the proof of Theorem 1 (which is used for all results presented here pertaining to Gaussian convergence): duality and Stein's method. Duality for the voter model has been known since the original works on the model [11, 14, 23, 31] and allows for the application of the rich theory of random walks. Stein's method (introduced in [44]) is a powerful technique for proving approximation results (see for example the survey [41]). An application of a particular version in [40] is a key contribution to the proof of Theorem 1, providing general conditions for Gaussian convergence. This approximation has been used in [32] to obtain conditions for Gaussian convergence in a voter-like model. Stein's method has also been used to obtain Gaussian convergence in the antivoter model for the proportion of vertices with opinion 1 at stationarity [39]. More generally, versions of Stein's method have been used with other particle system/statistical mechanics problems: to various systems including the voter model [20]; to discrete Gibbs measures [17]; to the magnetisation in the Curie–Weiss model (Ising model on the complete graph) [5, 8, 9, 16].

In the remainder of this section we present the main results. Then, in Section 2 we construct a dual process to the noisy voter model. While only a small modification is required to the standard dual (to the voter model), for completeness we include the argument. We then use this duality to prove some key properties of the noisy voter model in Section 3 including identities for the mean and variance of S and the proof of all the results of Section 1.3, except Theorem 1 which features in Section 4, and Proposition 9 (results for the cycle and torus). The second half of the paper, Section 5, is devoted to the cycle and torus which requires ad-hoc arguments utilising structural properties of these graphs. The Appendix collects useful results used throughout.

1.2. *Formal description of the model and object of interest.* Here we present a formal definition of the noisy voter model and the statistic S. Let P be the transition matrix of an irreducible discrete-time random walk on a finite set of vertices V and let π denote the unique

stationary distribution of P. The noisy voter model on V associated with P is a discrete-time Markov chain $(\xi_t)_{t \in \mathbb{N}}$ in the space $\{0,1\}^V$ in which all vertices have an initial opinion in $\{0,1\}$. At each time-step, with probability p select a uniform vertex x and re-randomise its opinion, i.e.choose an opinion from $\{0,1\}$ uniformly. Otherwise, (with probability $q = 1 - p$) perform a standard (imitation) step of the voter model.

The transition matrix of $(\xi_t)_{t \in \mathbb{N}}$, denoted Q, can be given explicitly. Suppose the configurations ξ and ξ' differ only in the value of $x \in V$, i.e. $\xi(y) \neq \xi'(y) \iff y = x$, then

$$(1) \qquad Q(\xi, \xi') = \mathbb{P}(\xi_1 = \xi' | \xi_0 = \xi) = \begin{cases} \frac{p}{2n} + \frac{q}{n} \sum_{y \in V} P(x,y)\xi(y) & \text{if } \xi(x) = 0, \\ \frac{p}{2n} + \frac{q}{n} \sum_{y \in V} P(x,y)(1 - \xi(y)) & \text{if } \xi(x) = 1, \end{cases}$$

where $n = |V|$. If ξ and ξ' differ at more than one vertex then $Q(\xi, \xi') = 0$. Finally, $Q(\xi, \xi) = 1 - \sum_{\xi' \neq \xi} Q(\xi, \xi')$. We write $\mathrm{NVM}(V, P, p)$ for this process.

Observe that if $p > 0$ then all configurations in $\{0,1\}^V$ are reachable due to re-randomization, i.e.Q is irreducible. Moreover, with probability at least $p/2$, no vertex changes its opinion at a given time step, so Q is aperiodic. It follows that $\mathrm{NVM}(V, P, p)$ has a unique stationary distribution, which we denote by Γ.

Given a vector $\xi \in \{0,1\}^V$, we define the object of interest $S(\xi) := \sum_{x \in V} \pi(x)\xi(x)$ where π is the stationary distribution of P and ξ is a random $\{0,1\}$-valued vector with distribution Γ. Thus $S = S(\xi)$ is a random variable, and the value of S is given by the π-measure of the set of vertices with opinion 1.

When $p = 0$, the noisy voter model reverts to the voter model with transition matrix P. In such a case, the voter model reaches an ordered state of consensus, that is, a fixed state where all opinions are the same. Thus the process $\mathrm{NVM}(V, P, p)$ does not have a unique stationary distribution; indeed, it has two absorbing states: the state where all opinions are 1, and another where all opinions are 0.

On the other hand, if $p = 1$, then only re-randomization occurs, and thus the chain is a lazy random walk on the hypercube. It follows that $\Gamma(\xi) = 1/2^n$ for all $\xi \in \{0,1\}^V$, that is, the components of ξ are i.i.d. Bernoulli random variables of parameter $1/2$. In this case we have $\mathbb{E}(S) = 1/2$ and $\mathbb{V}\mathrm{ar}(S) = \sum_{x \in V} \pi^2(x)/4$, and so $\frac{S - 1/2}{\sqrt{\mathbb{V}\mathrm{ar}(S)}}$ converges in distribution (when n tends to infinity) to a standard Gaussian if $\max_{y \in V}\{\pi(y)^2 / \sum_{x \in V} \pi(x)^2\} \to 0$ by Lindeberg's central limit theorem.

Our results, presented below, provide much finer detail on the transition from order to disorder.

1.3. *Results.* We introduce some more notation before presenting the main results. Set $\pi^* := \max_{x \in V} \pi(x)$, $\nu^2 := \sum_{x \in V} \pi(x)^2$, and for $x, y \in V$, $\mu(x,y) := \pi(x)^2 P(x,y)/\nu^2$. We use the letter ξ to denote a sample from Γ, the stationary distribution of $\mathrm{NVM}(V, P, p)$. Finally, let $\sigma^2 := \mathbb{V}\mathrm{ar}(S)$ denote the variance of S, and $W := \frac{S - \mathbb{E}(S)}{\sigma}$ denote the standardisation of S. Sometimes, we will consider sequences $\{\mathrm{NVM}(V_n, P_n, p_n) : n \geq 1\}$ of the noisy voter model indexed by $|V_n| = n \geq 1$ and, in this case, we will include the subindex n for all above definitions (e.g. $S_n, \sigma_n, W_n, \pi_n$). We will always assume for each $n \geq 1$ that P_n is irreducible, and that p_n is non-increasing in n.

We denote convergence in distribution by $\overset{\mathcal{D}}{\Rightarrow}$, e.g. $X_n \overset{\mathcal{D}}{\Rightarrow} X$, and all limits given are taken when n tends to infinity (unless otherwise stated).

Our first result gives sufficient conditions for W_n converging to a Gaussian random variable as n tends to infinity. While the statement may appear a little cumbersome, the conditions required follow naturally from a version of Stein's method and the result in this form increases the ease of application. We present simplified versions of the result for particular situations in the ensuing corollaries.

THEOREM 1 (General conditions for Gaussian convergence). *Consider a noisy voter model* $\mathrm{NVM}(V, P, p)$ *with* $|V| = n$, *and let* $\xi \sim \Gamma$. *Let* Φ *denote the cumulative distribution function of a standard Gaussian random variable on* $\mathbb{R}$. *Then there exists a universal constant* $C > 0$ *such that*

$$(2) \qquad \sup_{t \in \mathbb{R}} \left| \mathbb{P}\left(\frac{S - 1/2}{\sigma} \leq t \right) - \Phi(t) \right| \leq C \left\{ \left(\frac{\pi^*}{\sigma} \right)^3 \frac{n}{p} + \left(\frac{\pi^*}{\sigma} \right)^2 \sqrt{\frac{n}{p}} + \frac{\nu^2}{p\sigma^2} \sqrt{\mathbb{V}\mathrm{ar}\Psi} \right\},$$

where the random variable Ψ *is given by* $\Psi := \sum_{x \in V} \sum_{y \in V} \mu(x, y) \mathbf{1}_{\{\xi(x) \neq \xi(y)\}}$. *Moreover, we have that*

$$(3) \qquad \mathbb{V}\mathrm{ar}(\Psi) \leq 16\sigma^2 (\pi^*)^2 \nu^{-4}.$$

In particular, if a sequence of noisy voter models $\mathrm{NVM}(V_n, P_n, p_n)$ *(with* $|V_n| = n$*) satisfies*

$$(4) \qquad \frac{n}{p_n} \left(\frac{\pi_n^*}{\sigma_n} \right)^3 + \frac{\pi_n^*}{\sigma_n p_n} \to 0$$

then $W_n \overset{\mathcal{D}}{\Rightarrow} N(0, 1)$.

The term $1/2$ in the standardisation of S in equation (2) comes from the fact that $\mathbb{E}(S) = 1/2$, see Proposition 13. We also remark that (4) follows from substituting the bound on the variance of Ψ from (3) into (2).

The following lemma gives two bounds that can be used to obtain immediate applications of Theorem 1.

LEMMA 2 (Variance lower bounds). *Consider the noisy voter model* $\mathrm{NVM}(V, P, p)$. *The following two lower bounds for* σ^2 *hold:*

1. $4\sigma^2 \geq \nu^2$,
2. *for* $p \leq 1/2$, $4\sigma^2 \geq (1 + 4pt_{\mathrm{hit}})^{-1}$, *where* $t_{\mathrm{hit}} = \max_{x, y \in V} \mathbb{E}_x(T_y)$, *and* T_y *is the hitting time of* $y \in V$ *by the random walk associated with* P.

We present the proof of Lemma 2 in Section 3. Combining Theorem 1 with Lemma 2 yields simple conditions to establish a Gaussian limit regime for a sequence of noisy voter models $\mathrm{NVM}(V_n, P_n, p_n)$, which we present next.

COROLLARY 3 (Simple conditions for Gaussian convergence). *For each* $n \geq 1$ *consider a noisy voter model* $\mathrm{NVM}(V_n, P_n, p_n)$ *with* $|V_n| = n$. *If*

$$(5) \qquad \left(\frac{\pi_n^*}{\nu_n} \right)^3 \frac{n}{p_n} \to 0 \quad as \ n \to \infty,$$

then $W_n \overset{\mathcal{D}}{\Rightarrow} \mathcal{N}(0, 1)$.

Now consider the case in which V is the vertex set of a connected graph and P the transition matrix of the random walk on this graph. It is standard that the stationary distribution of P is given by $\pi(x) = d(x)/(2m)$ where $d(x)$ is the degree of x, and m is the number of edges of the graph. In this setting, we can rewrite the conditions of Corollary 3 to obtain the following:

COROLLARY 4 (Gaussian convergence for random walks). *Let* G_n *be a sequence of connected graphs with vertex sets* V_n *of size* n, *and let* $d_n(x)$ *be the degree of vertex* x *in* G_n.

Consider the noisy voter model $\mathrm{NVM}(V_n, P_n, p_n)$ *where* P_n *is the transition matrix of the random walk on* G_n. *If*

$$(6) \qquad \left(\frac{\max_{x \in V} d_n(x)}{\sqrt{\sum_{x \in V} d_n(x)^2}} \right)^3 \frac{n}{p_n} \to 0 \quad as\ n \to \infty,$$

then $W_n \overset{\mathcal{D}}{\Rightarrow} \mathcal{N}(0, 1)$.

In the case that p_n is a fixed constant in $(0, 1)$, the condition $\pi_n^\star / \nu_n = o(n^{-1/3})$ implies (6), while (6) implies $\pi_n^\star / \nu_n \to 0$. The condition $\pi_n^\star / \nu_n = o(n^{-1/3})$ is satisfied by many families of graphs, e.g. for n-vertex regular graphs, we have $\pi_n^\star / \nu_n = n^{-1/2}$, in which case we need $p_n \gg n^{-1/2}$ to satisfy equation equation (6). On the other hand it is not satisfied by, for example, the star graph, since $\nu_n^2 / 4 \geq 1/16$.

COROLLARY 5 (Gaussian convergence for random walks on regular graphs). *Let* G_n *be a sequence of regular graphs with vertex sets* V_n *of size* n, *and let* P_n *be the transition matrix of the random walk on* G_n, *and consider a sequence of noisy voter models* $\mathrm{NVM}(V_n, P_n, p_n)$. *If* $p_n \gg n^{-1/2}$, *then* $W_n \overset{\mathcal{D}}{\Rightarrow} \mathcal{N}(0, 1)$.

We also show that for very small values of p_n, the distribution of S_n converges in distribution to a Bernoulli random variable:

PROPOSITION 6 (Conditions for Bernoulli convergence). *For each* $n \geq 1$, *consider a noisy voter model* $\mathrm{NVM}(V_n, P_n, p_n)$. *Let* M_n *be the meeting time of two independent continuous-time random walks (moving at rate 1 each according to* P_n*) started from stationarity. If* $p_n \mathbb{E}(M_n) \to 0$ *then* $\sigma_n^2 \to 1/4$ *and* $S_n \overset{\mathcal{D}}{\Rightarrow} \mathcal{B}er(1/2)$.

The following conjecture is natural; however it seems hard to prove in general, even for regular graphs.

CONJECTURE 7. *For each* $n \geq 1$, *consider a noisy voter model* $\mathrm{NVM}(V_n, P_n, p_n)$ *such that* $\pi_n^* / \nu_n \to 0$. *If* $p_n \mathbb{E}(M_n) \to \infty$ *then* $W_n \overset{\mathcal{D}}{\Rightarrow} \mathcal{N}(0, 1)$.

We expect our techniques are not sufficient to obtain a proof of Conjecture 7, as the condition in the conjecture involves (expectations of) meeting times of only two particles; on the other hand the conditions appearing in Theorem 1 (in particular for the quantity $\mathbb{V}\mathrm{ar}(\Psi)$) involve meeting times of four particles. Thus the conditions in Theorem 1 are likely stronger than needed as probabilities of events involving four particles typically can not be bounded by probabilities of events involving just two particles, except if the graph has nice structural properties such as fast mixing.

Note that Proposition 6 combined with the previous conjecture suggests that for transition matrices P_n with $\pi_n^* / \nu_n \to 0$ the inverse expected meeting time is the critical rate of p_n which separates the ordered and disordered phases.

The next result identifies conditions for which the critical rate p_c is equal to $1/n$. Recall that $t_{\mathrm{hit}} = \max_{x, y \in V} \mathbb{E}_x(T_y)$, with T_y is the hitting time of $y \in V$ by the random walk associated with P.

PROPOSITION 8 (General conditions for order-disorder transition). *For each* $n \geq 1$, *consider a noisy voter model* $\mathrm{NVM}(V_n, P_n, p_n)$ *such that* $t_{\mathrm{hit}} = O(n)$ *and* $\pi_n^* = O(1/n)$.

1. *If $p_n n \to \infty$ then $W_n \overset{\mathcal{D}}{\Rightarrow} \mathcal{N}(0,1)$.*
2. *If $p_n n \to 0$ then $S_n \overset{\mathcal{D}}{\Rightarrow} \mathcal{B}er(1/2)$.*

The conditions of Proposition 8 are satisfied by P_n the random walk on several graph families including hypercubes, regular expanders, grids and tori in dimension larger than 2, and complete graphs.

On the other hand, these general results do not cover two important cases: the cycle, which we denote C_n to indicate it has n vertices, and the two-dimensional torus, denoted by $\mathbb{T}_n$ to indicate it has n vertices (note that when writing $\mathbb{T}_n$ we are implicitly assuming that n is a square number). For these graphs, we do not only need a sharp lower bound for σ_n^2 but also better control of the term at the right-hand side of (2), which requires some ad-hoc arguments. For the cycle the proof features several couplings with a random walk on the integers, and thus it cannot be extended to other graphs, whereas our argument on the torus may be extended to other transitive graphs.

PROPOSITION 9 (Order-disorder transition for cycle and 2D torus). *Consider the noisy voter model* $\mathrm{NVM}(V_n, P_n, p_n)$ *where P_n is a sequence of transition matrices associated with the random walk on the graph $G_n = (V_n, E_n)$, and p_n is non-increasing. Then the following holds true:*

1. *Let $G_n = C_n$ be the cycle on n vertices. If $p_n n^2 \to \infty$ then $W_n \overset{\mathcal{D}}{\Rightarrow} \mathcal{N}(0,1)$, and if $p_n n^2 \to 0$ then $S_n \overset{\mathcal{D}}{\Rightarrow} \mathcal{B}er(1/2)$.*
2. *Let $G_n = \mathbb{T}_n$ be a 2-dimensional torus on n vertices. If $p_n n \log n \to \infty$ then $W_n \overset{\mathcal{D}}{\Rightarrow} \mathcal{N}(0,1)$, and if $p_n n \log n \to 0$ then $S_n \overset{\mathcal{D}}{\Rightarrow} \mathcal{B}er(1/2)$.*

Note that on both the cycle and the torus, the cases of convergence of S_n to a $\mathcal{B}er(1/2)$ random variable follow directly from Proposition 6.

In order to apply Theorem 1 to obtain the Gaussian convergence results in Proposition 9, we require bounds on σ_n^2 and $\mathbb{V}\mathrm{ar}(\Psi_n)$. Consider the case of the cycle. Using duality, we can relate σ_n^2 to the probability that two independent random walks on the cycle – started from uniformly chosen locations – have meeting time less than an independent $\mathrm{Geo}(p)$ random variable (which plays the role of the noise). In fact, σ_n^2 can be computed explicitly in terms of p and n (for the cycle), see Lemma 27. Duality can also be used to relate $\mathbb{V}\mathrm{ar}(\Psi_n)$ to events involving four coalescing random walks. The relevant events are of the form: two particular walkers meet before any other walkers meet and this first meeting time is less than a $\mathrm{Geo}(p)$ random variable which itself is less than the meeting time of another particular pair of walkers. Controlling the probability of such events is challenging due to the complicated dependencies and we make use of several martingale optional stopping arguments together with the FKG inequality. This type of problem is similar to those studied in [1].

REMARK 10 (Discrete-time vs continuous-time). Although the results in this paper are given for the discrete-time version of the noisy voter model, they can also be established for a continuous-time version. In the continuous-time setting, we have a transition-rate matrix $R = (r(x,y) : x, y \in V)$, and a noise rate δ. Each vertex x adopts the opinion of vertex y at rate $r(x,y)$, and re-randomises its opinion at rate δ (equivalently, flips its opinion at rate $\delta/2$). We can then transform the continuous-time process into a discrete one with the same stationary distribution by using the standard uniformalisation method. Indeed, for given transition-rate R, and noise rate δ, the analogue discrete process has transition matrix $P = I + \frac{R}{r_{\max}}$, where $r_{\max} = \max\{|r(x,x)| : x \in V\}$ is the maximum transition rate from any $x \in V$, and re-randomisation probability $p = \delta/(r_{\max} + \delta)$. Note the stationary distribution of P and R are

the same, so the results of this paper are consistent in both settings. We can also transform the discrete model into a continuous one; for this, each vertex x re-randomises its opinion at rate p, and adopts the opinion of vertex y at rate $(1-p)P(x,y)$.

NOTATION. We make use of standard asymptotic notation. Given positive real-valued functions f and g, we say that $f(x) = O(g(x))$ or $f(x) \lesssim g(x)$ if $\limsup_{x\to\infty} f(x)/g(x) \leq C$ for some constant $C > 0$.

Similarly $f(x) = \Omega(g(x))$ or $f(x) \gtrsim g(x)$ if $\liminf_{x\to\infty} f(x)/g(x) \geq c > 0$ for some constant c; and $f(x) = \Theta(g(x))$ or $f(x) \asymp g(x)$ if and only if $f(x) = O(g(x))$ and $f(x) = \Omega(g(x))$. Finally, $f(x) = o(g(x))$ if $\lim_{x\to\infty} f(x)/g(x) = 0$, in particular $f(x) = o(1)$ implies that $\lim_{x\to\infty} f(x) = 0$.

We use the following set notation: if S is a finite set and $k \in \mathbb{N}$, we write S^k for the set of k-tuples from S and $(S)_k$ for the set of k-tuples from S without repeats. Additionally, we use standard language and notation of basic graph theory, in particular $x \sim y$ means that vertices x and y are neighbours.

2. The dual process. It is classical that the voter model has a dual process of coalescing random walks, see for example [31]. The noisy voter model also enjoys a coalescing random walk dual process that allows us to recast probabilities involving the noisy voter model in terms of probabilities involving the dual (see Proposition 11) which are often more tractable, and allows for more informative sampling of the stationary distribution Γ.

To be more specific, we use a version of Stein's method (Theorem 34) to obtain conditions (involving variance and covariance) for Gaussian convergence. The dual process will allow us to transform the variance, respectively covariance, estimation problem into one of estimating probabilities of events related to interactions between two, respectively four, particles.

2.1. *Voter model with stubborn vertices.* Consider a noisy voter model $\mathrm{NVM}(V, P, p)$. Before constructing its dual, we introduce a process called the *voter model with stubborn vertices*, so-called as it features vertices whose opinions remain unchanged over time (see [47] for further discussion on the voter model with stubborn vertices).

The role of the stubborn vertices is to encode each re-randomisation step of the noisy voter model as instead a step of the (non-noisy) voter model. To this end, we augment the vertex set V with two additional (stubborn) vertices denoted $\mathbf{0}$ and $\mathbf{1}$, set $\widetilde{V} = V \cup \{\mathbf{0}, \mathbf{1}\}$, and define a transition matrix $\widetilde{P}$ on $\widetilde{V}$ as

$$(7) \qquad \widetilde{P}(x,y) = \begin{cases} qP(x,y) & x,y \in V, \\ p/2 & x \in V, y \in \{\mathbf{0},\mathbf{1}\}, \\ 1 & y = x, x \in \{\mathbf{0},\mathbf{1}\}, \\ 0 & \text{otherwise.} \end{cases}$$

Observe that the state of all 0s and the state of all 1s are the unique absorbing states of $\widetilde{P}$.

The voter model with stubborn vertices is the discrete-time model in which each vertex in $\widetilde{V}$ has an opinion in $\{0,1\}$ and which at each step chooses uniformly a vertex x from V (not $\widetilde{V}$) and updates the opinion of the chosen vertex to the opinion of vertex y with probability $\widetilde{P}(x,y)$. Stubborn vertex $\mathbf{0}$ (respectively $\mathbf{1}$) has opinion 0 (respectively 1) for all time.

Writing $(\xi_t)_{t\geq 0}$ for this process of opinions (we take its state space to be $\{0,1\}^V$ but can extend ξ_t to a function on $\widetilde{V}$ by setting $(\xi_t(\mathbf{0}), \xi_t(\mathbf{1})) = (0,1))$ we see that $(\xi_t)_{t\geq 0}$ is precisely the noisy voter model. The purpose of this new process will become apparent in the next section.

2.2. *The dual process.*   Consider the following discrete-time process: initialise by placing a particle on each $x \in V$ and label it with its host vertex. At each time step choose a vertex x from V uniformly and move all particles at x to $y \in \widetilde{V}$ with probability $\widetilde{P}(x, y)$ with $\widetilde{P}$ defined in (7). Note that particles stay together upon meeting.

Observe that when a particle reaches $\mathbf{0}$ or $\mathbf{1}$ it never moves again, and we say the particle is absorbed. Note that all particles are eventually absorbed. Indeed, consider particle x: at each time-step, the vertex containing x is chosen with probability $1/n$, and then it is absorbed into $\{\mathbf{0}, \mathbf{1}\}$ with probability p. Thus, particle x is absorbed in $\mathrm{Geometric}(p/n)$ steps, and so all particles are absorbed in finite time almost surely. We define the absorption time τ_{abs} as the time that all particles are absorbed. Formally, denote by Y_t the vector whose entries are the positions of all particles at time t, i.e. $Y_t(x)$ denotes the position of particle x at time t (note that $Y_0(x) = x$). The absorption time is

$$\tau_{\mathrm{abs}} := \min\{t : Y_t(i) \in \{\mathbf{0}, \mathbf{1}\} \text{ for all } i \in V\}.$$

Finally, define the random variable $B(x)$ to be 1 if particle x is absorbed by $\mathbf{1}$, otherwise $B(x) = 0$. For purposes of the analysis, it is convenient to define $B(\mathbf{1})$ and $B(\mathbf{0})$ as constants taking the values 1 and 0 respectively (i.e. if a particle starts in $\mathbf{1}$, then it is already absorbed by $\mathbf{1}$).

The relation between the voter model with stubborn vertices $(\xi_t)_{t \geq 0}$, and the coalescing particle system $(Y_t)_{t \geq 0}$ is established in the following result.

PROPOSITION 11 (Duality).   *Let Γ be the stationary distribution of the noisy voter model (or voter model with stubborn vertices). For each $\xi \in \{0, 1\}^V$, it holds that*

$$(8) \qquad\qquad \mathbb{P}(B(i) = \xi(i), \forall i \in V) = \Gamma(\xi),$$

i.e. the distribution of the random vector $(B(i))_{i \in V}$ is the same as the stationary distribution of the noisy voter model (or voter model with stubborn vertices).

The proof of Proposition 11 is a straightforward application of the graphical representation of the voter model/coalescing random walks; however, for completeness we include a proof.

PROOF.   Let $(\xi_t)_{t \geq 0}$ be the voter model with stubborn vertices and $(Y_t)_{t \geq 0}$ the coalescing particle system. Recall that $\xi_t(\mathbf{0}) = 0$ and $\xi_t(\mathbf{1}) = 1$.

For a pair of vertices $(v, w) \in V \times \widetilde{V}$ define $M_{v,w} : \{0, 1\}^{\tilde{V}} \to \{0, 1\}^{\tilde{V}}$ such that $\xi' = M_{v,w}\xi$ satisfies $\xi'(u) = \xi(u)$ for every $u \in \tilde{V} \setminus \{v\}$, and

$$\xi'(v) = \begin{cases} \xi(w) & \text{if } w \in V \\ 1 & \text{if } w = \mathbf{1} \\ 0 & \text{if } w = \mathbf{0} \end{cases},$$

in other words, ξ' is the result of vertex v imitating the opinion of vertex w in state ξ.

Let Y be a state of the dual process, i.e. $Y \in \widetilde{V}^V$ where $Y(i)$ indicates the current position of particle i. For $(v, w) \in V \times \widetilde{V}$, define the operator $M'_{v,w} : \widetilde{V}^V \to \widetilde{V}^V$ that moves all particles located in vertex v to vertex w. that is, for any particle i such that $Y(i) = v$, then $Y' = M'_{v,w}Y$ is such that $Y'(i) = w$.

We start by proving that for each $T \geq 0$ and each $\xi, \xi' \in \{0, 1\}^V$, it holds that

$$(9) \qquad\qquad \mathbb{P}(\xi_T = \xi' | \xi_0 = \xi) = \mathbb{P}(\xi(Y_T) = \xi'),$$

where $\xi(Y_t)(i) = \xi(Y_t(i))$. To prove equation (9) we couple the voter model and its dual on $t = 0, \ldots, T$. Recall that a coupling $((\xi_t^c, Y_t^c) : t \in \{0, \ldots, T\})$ is such that marginally the

distributions of $(\xi_t)_{t=0}^T$ and $(\xi_t^c)_{t=0}^T$ are the same and the distributions of $(Y_t)_{t=0}^T$ and $(Y_t^c)_{t=0}^T$ are also the same.

Define a probability measure m on $V \times \widetilde{V}$ as

$$m(v, w) = \frac{\widetilde{P}(v, w)}{n}. \tag{10}$$

Let $\mathbf{u} = ((v_t, w_t) : t \in \{1, 2, \ldots, \})$ be an i.i.d collection of random elements in $V \times \widetilde{V}$ sampled with distribution m. We use $\mathbf{u}$ to couple the processes.

For $t \in \{1, \ldots, T\}$ define $M_t = M_{v_t, w_t}$. Define the $\{0, 1\}^V$-valued process $(\xi_t^c : t \in \{0, 1, \ldots, T\})$ as

$$\xi_t^c = \begin{cases} \xi, & \text{if } t = 0, \\ M_t \xi_{t-1}^c, & \text{if } t \in \{1, \ldots, T\}, \end{cases}$$

where $\xi \in \{0, 1\}^V$ is given. Observe that $(\xi_t^c : t \in \{0, 1, \ldots, T\})$ is distributed as the first T rounds $(\xi_t)_{t \geq 0}$ starting with initial opinion ξ.

On the other hand, for $t \in \{1, \ldots, T\}$, define $M_t' = M_{v_{T-t+1}, w_{T-t+1}}'$. Define the $\widetilde{V}^V$-valued process $(Y_t^c : t \in \{0, 1, \ldots, T\})$ as

$$Y_t^c = \begin{cases} \iota, & \text{if } t = 0, \\ M_t' Y_{t-1}^c, & \text{if } t \in \{1, \ldots, T\}, \end{cases} \tag{11}$$

where ι is the identity map. In this case $(Y_t^c : t \in \{0, \ldots, T\})$ is distributed as the first t rounds of the dual coalescing process $(Y_t)_{t \geq 0}$.

The key observation is that for every i and for all $t \in \{0, \ldots, T\}$

$$\xi_{T-t}^c(Y_t^c(i)) = \xi_T^c(i), \tag{12}$$

where we assume that $\xi_t(\mathbf{0}) = 0$ and $\xi_t(\mathbf{1}) = 1$. To prove (12) we fix in advance the sequence $\mathbf{u}_T = ((v_t, w_t) \in V \times \widetilde{V} : t \in \{1, \ldots, T\})$ and use induction.

For $t = 0$, since $Y_0^c(i) = i$ we have $\xi_T^c(Y_0^c(i)) = \xi_T^c(i)$. Assume the result holds for some $t < T$. We prove it holds for $t + 1$. Using that $M_{T-t} = M_{v_{T-t}, w_{T-t}}$ from the definition of $\xi_{T-t}^c = M_{T-t} \xi_{T-t-1}^c$, we have that

$$\xi_T^c(i) = \xi_{T-t}^c(Y_t^c(i)) = \begin{cases} \xi_{T-t-1}^c(w_{T-t}), & \text{if } Y_t^c(i) = v_{T-t}, \\ \xi_{T-t-1}^c(Y_t^c(i)), & \text{if } Y_t^c(i) \neq v_{T-t}. \end{cases} \tag{13}$$

On the other hand, by using that $M_{t+1}' = M_{v_{T-t}, w_{T-t}}'$,

$$Y_{t+1}^c(i) = \begin{cases} w_{T-t}, & \text{if } Y_t^c(i) = v_{T-t}, \\ Y_t^c(i), & \text{if } Y_t^c(i) \neq v_{T-t}. \end{cases} \tag{14}$$

Therefore, from equations (13) and (14) we get that $\xi_{T-t-1}^c(Y_{t+1}^c(i)) = \xi_T^c(i)$ for all $i \in V$, completing the proof of (12).

A particular case of equation (12) for our coupled variables is that $\xi_0^c(Y_T^c) = \xi_T^c$. Then, for all $\xi, \xi' \in \{0, 1\}^V$ we have

$$\mathbb{P}(\xi_T = \xi' | \xi_0 = \xi) = \mathbb{P}(\xi_T^c = \xi' | \xi_0^c = \xi) = \mathbb{P}(\xi(Y_T^c) = \xi') = \mathbb{P}(\xi(Y_T) = \xi'). \tag{15}$$

Now that equation (9) is proved, we can forget about the coupling and work with the original processes. By taking limits in both sides of equation (9) when T tends to infinity, we can recover equation (8). Indeed, on the left-hand side of equation (9) we have that

$$\lim_{T \to \infty} \mathbb{P}(\xi_T = \xi' | \xi_0 = \xi) = \Gamma(\xi') \tag{16}$$

since the stationary distribution of ξ_T is Γ. On the other hand, since $\lim_{T\to\infty} \mathbb{P}(\tau_{\mathrm{abs}} \leq T) = 1$, by taking limit on the right-hand side of equation (9) we have

$$
\begin{aligned}
\lim_{T\to\infty} \mathbb{P}(\xi(Y_T) = \xi') &= \lim_{T\to\infty} \mathbb{P}(\xi(Y_T) = \xi', \tau_{\mathrm{abs}} \leq T) \\
&= \lim_{T\to\infty} \mathbb{P}(\xi(Y_{\tau_{\mathrm{abs}}}) = \xi', \tau_{\mathrm{abs}} \leq T) \\
&= \mathbb{P}(\xi(Y_{\tau_{\mathrm{abs}}}) = \xi') \\
&= \mathbb{P}(B(i) = \xi'(i), \forall i). \qquad \square
\end{aligned}
$$

REMARK 12 (Dual time relation).   Since the speed of the dual process is not relevant (in the sequel we are only concerned with the orderings of meeting and absorption times), sometimes it will be convenient for us to consider the continuous-time version of the process defined as follows. Each vertex $x \in V$ is provided with an independent exponential clock of rate $1/(1-p)$. Each time the clock of $x \in V$ rings, it moves all its particles, if any, using the transition matrix $\widetilde{P}$. Note that with probability $\widetilde{P}(x,x)$ the particles do not move. Thus, a vertex rings and any present particles jump at rate 1 (deciding where to move by using P), and at rate $p/(1-p)$ any particles on a particular vertex are absorbed into one of the stubborn states.

The continuous-time version is slightly more convenient as particles are independent until they meet or are absorbed by one of the stubborn vertices.

We finally note that moving between continuous and discrete-time does not affect the validity of Proposition 11.

3. Basic properties of $\mathrm{NVM}(V, P, p)$.   In this section we prove some basic properties of $\mathrm{NVM}(V, P, p)$, in particular properties of the random variable $S = \sum_{x \in V} \pi(x)\xi(x)$ where ξ is sampled from the stationary distribution Γ of $\mathrm{NVM}(V, P, p)$, and π is the stationary distribution of P. These properties feature in the proofs of the results presented in Section 1.3. In all cases we assume $p \in (0, 1)$ to avoid trivial cases.

Associated with the dual process of coalescing random walks we define the following random variables. Under the standard convention that $\inf\{\varnothing\} = \infty$, we define the absorption time of particle i as $\tau^{Y(i)} = \inf\{t \geq 0 : Y_t(i) \in \{\mathbf{0}, \mathbf{1}\}\}$, i.e. $\tau^{Y(i)}$ is the hitting time of the set $\{\mathbf{0}, \mathbf{1}\}$. Also, for $i, j \in V$ define $\tau^{Y(i), Y(j)} = \inf\{\tau^{Y(i)}, \tau^{Y(j)}\}$. We write $M^{Y(i), Y(j)} = \inf\{t \geq 0 : Y_t(i) = Y_t(j) \in V\}$, i.e $M^{Y(i), Y(j)}$ is the meeting time of particles i and j. Notice that we are only considering meetings that happen in vertices of V and not in the stubborn vertices $\mathbf{0}$ nor $\mathbf{1}$. Hence, $M^{Y(i), Y(j)}$ can be infinity, even if both particles are absorbed by the same vertex. For $i, j \in V$ define the event $E_{ij} = \{M^{Y(i), Y(j)} < \tau^{Y(i), Y(j)}\} = \{M^{Y(i), Y(j)} < \infty\}$. For convenience we define $M^{Y(i), Y(i)} = 0$, and thus $\mathbb{P}(E_{ii} = 1)$.

Finally, we will frequently couple the dual process with an independent random walks process. Given an integer $m \geq 1$, we say that $(X_t(1), \ldots, X_t(m))_{t \geq 0}$ is a $\mathrm{RW}(m, P)$ process if all the random walks $X(i)$ are independent, and they move according to the transition matrix P. We will allow time to be continuous or discrete (which shall be made clear when we use the process); in the case of continuous time the walks jump at rate 1, unless stated otherwise. It follows from independence that the particles in $\mathrm{RW}(m, P)$ do not coalesce upon meeting and there is no absorption (in particular they always move in V). Given $\boldsymbol{x} \in V^m$ (or a measure $q \in V^m$), we use $\mathbb{P}_{\boldsymbol{x}}^{\mathrm{RW}}$ ($\mathbb{P}_q^{\mathrm{RW}}$) to denote the measure with respect to $\mathrm{RW}(m, P)$ when the walks start from $\boldsymbol{x}$ (from q). Finally, we write RW or $\mathrm{RW}(m)$ when m and P are clear from the context. We also write M^{X_i, X_j} to denote the meeting time of the walk X_i and X_j in $\mathrm{RW}(m, P)$.

PROPOSITION 13.   *In $\mathrm{NVM}(V, P, p)$ we have $\mathbb{E}(S) = 1/2$.*

PROOF. By Proposition 11, we know that $(B(i))_{i \in V}$ and ξ have the same distribution, and so the same marginal distributions. In particular, for all $x \in V$, it holds that $\mathbb{P}(\xi(x) = 1) = \mathbb{P}(B(x) = 1)$. Moreover, note that $\mathbb{P}(B(x) = 1) = 1/2$ since in the dual process, each particle is absorbed by either **0** or **1** with equal probability. (Equation (7) gives the probability that a particle at v transitions to w in a given step.) Therefore

$$\mathbb{E}(S) = \sum_{x \in V} \pi(x) \mathbb{E}(\xi(x)) = \sum_{x \in V} \pi(x) \mathbb{P}(\xi(x) = 1) = 1/2. \qquad \square$$

LEMMA 14. *Consider* $\mathrm{NVM}(V, P, p)$, *and let* $\sigma^2 = \mathbb{V}\mathrm{ar}(S)$. *Then*

$$\sigma^2 = \frac{1}{4} \sum_{x \in V} \sum_{y \in V} \pi(x) \pi(y) \mathbb{P}(E_{xy})$$

PROOF. Recall that

$$\sigma^2 = \mathbb{V}\mathrm{ar}\left(\sum_{x \in V} \pi(x) \xi(x) \right) = \sum_{x \in V} \sum_{y \in V} \pi(x) \pi(y) \mathbb{C}ov(\xi(x), \xi(y)).$$

By Proposition 11, the vector $(B(x))_{x \in V}$ has the same stationary distribution as ξ. Thus

$$\sigma^2 = \sum_{x \in V} \sum_{y \in V} \pi(x) \pi(y) \mathbb{C}ov(B(x), B(y)).$$

From the definition of covariance,

$$(17) \quad \mathbb{C}ov(B(x), B(y)) = \mathbb{E}(B(x)B(y)) - \mathbb{E}(B(x))\mathbb{E}(B(y)) = \mathbb{E}(B(x)B(y)) - (1/4).$$

Given the event $E_{xy} = \{ M^{Y(x),Y(y)} < \tau^{Y(x),Y(y)} \}$, observe that $\mathbb{E}(B(x)B(y)|E_{xy}) = 1/2$: if both particles meet before either one is absorbed, they move together and their final state is the same. They are absorbed by **1** with probability $1/2$, and by **0** otherwise. On the other hand, $\mathbb{E}(B(x)B(y)|E_{xy}^c) = 1/4$ since both particles are absorbed at different times and thus their final states are independent given E_{xy}^c. Hence

$$\mathbb{C}ov(B(x), B(y)) = (1/2)\mathbb{P}(E_{xy}) + (1/4)\mathbb{P}(E_{xy}^c) - 1/4 = (1/4)\mathbb{P}(E_{xy}). \qquad \square$$

PROOF OF LEMMA 2. We begin with the first item. Since $\mathbb{P}(E_{xx}) = 1$ for all $x \in V$, then Lemma 14 yields

$$\sigma^2 = \mathbb{V}\mathrm{ar}(S) \geq (1/4) \sum_{x \in V} \pi(x)^2 \mathbb{P}(E_{xx}) = (1/4) \sum_{x \in V} \pi(x)^2 = \nu^2/4.$$

For the second item, as was discussed previously $\mathbb{P}(E_{xy})$ has the same value if we consider discrete or continuous time since this only affects the speed of the process. For this proof it is more convenient to work in continuous time for which particles in the dual move at rate 1 and are absorbed by the stubborn vertices at rate $\delta := p/(1-p)$. Note that up to time $M^{Y(x),Y(y)} \wedge \tau^{Y(x),Y(y)}$, the particles $Y(x), Y(y)$ move like independent continuous-time random walks, and the time $\tau^{Y(x),Y(y)}$ is independent of $Y(x), Y(y)$ and is exponentially distributed with rate 2δ. Hence, we can naturally couple $(Y(x), Y(y))$ with a $\mathrm{RW}(2, P)$ process (X_1, X_2) up to time $M^{Y(x),Y(y)} \wedge \tau^{Y(x),Y(y)}$. Thus

$$(18) \qquad 1 - \mathbb{P}(E_{x,y}) = \mathbb{P}(M^{Y(x),Y(y)} > \tau^{Y(x),Y(y)}) = \mathbb{P}^{\mathrm{RW}(2)}_{(x,y)}(M^{X_1,X_2} > Z)$$

where M^{X_1,X_2} denotes the meeting time of the independent random walks X_1, X_2 and Z is an independent exponential random variable of rate 2δ.

By multiplying (18) by $\pi(x)\pi(y)$ and summing over all $x, y \in V$, Lemma 14 yields

$$(19) \qquad 1 - 4\sigma^2 = \mathbb{P}^{\mathrm{RW}}_{\pi^2}(M^{X_1,X_2} > Z) = \int_0^\infty \mathbb{P}^{\mathrm{RW}}_{\pi^2}(M^{X_1,X_2} > t)2\delta e^{-2\delta t}dt,$$

Recall that X_1 and X_2 move on state space V at rate 1 according to the transition matrix P. Then from [33, Lemma 1.7], we have

$$\mathbb{P}_{\pi^2}(M^{X_1,X_2} > t) \leq e^{-t/t_{\mathrm{hit}}},$$

where we recall that $t_{\mathrm{hit}} = \max_{x,y \in V} \mathbb{E}_x(T_y)$, with T_y is the hitting time of vertex $y \in V$. Therefore,

$$\int_0^\infty \mathbb{P}^{\mathrm{RW}}_{\pi^2}(M^{X_1,X_2} > t)2\delta e^{-2\delta t}dt \leq 2\delta \int_0^\infty \exp\left(-t\left(\frac{1}{t_{\mathrm{hit}}} + 2\delta\right)\right)dt$$

$$= \frac{2\delta}{\frac{1}{t_{\mathrm{hit}}} + 2\delta} = \frac{2\delta t_{\mathrm{hit}}}{1 + 2\delta t_{\mathrm{hit}}}.$$

By plugging this inequality into (19) we obtain

$$4\sigma^2 \geq 1 - \frac{2\delta t_{\mathrm{hit}}}{1 + 2\delta t_{\mathrm{hit}}} = \frac{1}{1 + 2\delta t_{\mathrm{hit}}}.$$

Note that t_{hit} is the same in either continuous or discrete time, so the identical bound holds in either setting. Using the previous inequality and the assumption $p \leq 1/2$, we obtain

$$4\sigma^2 \geq \frac{1}{1 + 2\delta t_{\mathrm{hit}}} \geq \frac{1}{1 + 4pt_{\mathrm{hit}}}. \qquad \square$$

Using a similar argument we next prove Proposition 6.

PROOF OF PROPOSITION 6. We shall show that $\sigma_n^2 \to 1/4$, and so, since the Bernoulli random variable of parameter $1/2$ is the only distribution taking values in $[0, 1]$ with variance $1/4$, we will have proven the result.

Following the same argument of the proof of Lemma 2 we just have to prove that

$$\mathbb{P}^{\mathrm{RW}(2)}_{\pi_n^2}(M^{X_1,X_2} > Z) \to 0,$$

where (X_1, X_2) is a continuous-time $\mathrm{RW}(2, P_n)$ process on V_n moving at rate 1 according to P_n, and Z is an exponential random variable of rate $2p_n/(1 - p_n)$.

In the following computations we omit the subscript n to ease notation. Let $t_{\mathrm{meet}} = \mathbb{E}^{\mathrm{RW}(2)}_{\pi^2}(M^{X_1,X_2})$, and $\delta = p/(1 - p)$, then

$$\mathbb{P}^{\mathrm{RW}(2)}_{\pi^2}(M^{X_1,X_2} > Z) = \int_0^\infty \mathbb{P}_{\pi^2}(M^{Y(x),Y(y)} > t)2\delta e^{-2\delta t}dt$$

$$(20) \qquad\qquad \leq (1 - e^{-2\delta t_{\mathrm{meet}}}) + t_{\mathrm{meet}}\int_{t_{\mathrm{meet}}}^\infty \frac{2\delta e^{-2\delta t}}{t}dt.$$

Now,

$$\int_{t_{\mathrm{meet}}}^\infty \frac{2\delta e^{-2\delta t}}{t}dt = \sum_{i=1}^\infty \int_{it_{\mathrm{meet}}}^{(i+1)t_{\mathrm{meet}}} \frac{2\delta e^{-2\delta t}}{t}dt \leq 2\delta \sum_{i=1}^\infty \frac{e^{-2\delta it_{\mathrm{meet}}}}{i}.$$

Observe that $\sum_{i=1}^{\infty} e^{-ai}/i = -\log(1 - e^{-a})$, and that $e^{-a} \leq 1/(1+a)$ for $|a| < 1$. Thus for all n sufficiently large (so that $\delta t_{\text{meet}} < 1$), we obtain

$$\int_{t_{\text{meet}}}^{\infty} \frac{2\delta e^{-2\delta t}}{t} dt \leq 2\delta \log(1 + 1/(2\delta t_{\text{meet}})).$$

We now use that $e^{-x} \leq 1 - x + x^2/2$ for $x > 0$, to give that for $0 < x < 1$, $1 \leq \frac{1-e^{-x}}{x}\frac{1}{1-x/2}$, i.e. for n sufficiently large,

$$1 \leq \frac{1 - e^{-2\delta t_{\text{meet}}}}{2\delta t_{\text{meet}}}\frac{1}{1 - \delta t_{\text{meet}}}.$$

We thus obtain

$$\int_{t_{\text{meet}}}^{\infty} \frac{2\delta e^{-2\delta t}}{t} dt \leq \frac{1 - e^{-2\delta t_{\text{meet}}}}{t_{\text{meet}}}\frac{1}{1 - \delta t_{\text{meet}}} \log(1 + 1/(2\delta t_{\text{meet}})),$$

and plugging this into (20) yields

$$(21) \qquad \mathbb{P}_{\pi_n^2}^{\text{RW}(2)}(M^{X_1,X_2} > Z) \leq (1 - e^{-2\delta t_{\text{meet}}})\left(1 + \frac{\log\left(1 + 1/(2\delta t_{\text{meet}})\right)}{1 - \delta t_{\text{meet}}}\right).$$

The result follows since $\lim_{x\to 0}(1 - e^{-x})(1 + \log(1 + 1/x)/(1 - x/2)) = 0$ and $\delta t_{\text{meet}} \to 0$ by assumption. $\qquad\square$

PROOF OF COROLLARY 3. We will prove that (5) implies (4). We consider two cases: first, suppose that $\left(\frac{\pi_n^*}{\sigma_n}\right)^2 n \geq 1$, then, by using the first item of Lemma 2 in equation (2) we have

$$\frac{n}{p_n}\left(\frac{\pi_n^*}{\sigma_n}\right)^3 + \frac{\pi_n^*}{\sigma_n p_n} = \frac{\pi_n^*}{\sigma_n p_n}\left(\left(\frac{\pi_n^*}{\sigma_n}\right)^2 n + 1\right) \leq 2\left(\frac{\pi_n^*}{\sigma_n}\right)^3 \frac{n}{p_n} \leq 2\left(\frac{2\pi_n^*}{\nu_n}\right)^3 \frac{n}{p_n} = o(1).$$

Secondly, consider $\left(\frac{\pi_n^*}{\sigma_n}\right)^2 n < 1$, thus

$$\frac{n}{p_n}\left(\frac{\pi_n^*}{\sigma_n}\right)^3 + \frac{\pi_n^*}{\sigma_n p_n} = \frac{\pi_n^*}{\sigma_n p_n}\left(\left(\frac{\pi_n^*}{\sigma_n}\right)^2 n + 1\right) \leq 2\frac{\pi_n^*}{\sigma_n p_n} < \frac{2}{\sqrt{n}p_n} = o(1),$$

since $\sqrt{n}p_n \to \infty$ which is implicit in statement (5) as $(\pi_n^*/\nu_n)^2 \geq n^{-1}$. $\qquad\square$

PROOF OF PROPOSITION 8. Recall that p_n is non-increasing, hence for the first item, we divide into the cases that eventually $p_n \leq 1/2$, and $p_n > 1/2$. If $p_n > 1/2$ we apply directly Corollary 3 since $\nu_n = \Theta(1/\sqrt{n})$ and $\pi_n^* = O(1/n)$. For the case $p_n \leq 1/2$, by Lemma 2 we have that $\sigma_n = \Omega\left(\frac{1}{np_n}\right)$, which replaced on the left-hand side of equation (4) yields

$$O\left(\sqrt{p_n/n} + 1/\sqrt{np_n}\right) = o(1),$$

and so Theorem 1 yields the result.

For the second item, by [2, Proposition 14.5] we have that $\mathbb{E}(M_n) \leq t_{\text{hit}} = O(n)$ where M_n is the meeting time of two independent continuous-time random walks where the initial state is sampled according to π. The result follows then from Proposition 6. $\qquad\square$

The next result will be useful to prove the Gaussian approximation of the random variable W. Suppose we consider $\text{NVM}(V, P, p)$ with $|V| = n$ starting from $\xi_0 = \xi \sim \Gamma$. Denote $\xi' = \xi_1$. Recall the definition $W = (S - 1/2)/\sigma$, where $S = \sum_{x \in V} \pi(x)\xi(x)$, and similarly define W' (replacing ξ with ξ'). Since $\xi \sim \Gamma$ we have $\text{Var}(S) = \text{Var}(S') = \sigma^2$ and $\mathbb{E}(S) = \mathbb{E}(S') = 1/2$. The following lemma relates W' with W.

LEMMA 15. $\mathbb{E}(W'|W) = \left(1 - \frac{p}{n}\right)W$.

PROOF. Set $A = \{x \in V : \xi(x) = 1\}$ and let $P(x, A) = \sum_{y \in A} P(x, y)$. Then, for $x \in V$, we have

$$\mathbb{E}(\xi'(x)|\xi) = \frac{(n-1)}{n}\xi(x) + \frac{1}{n}\left(\frac{p}{2} + qP(x, A)\right).$$

To see this consider the following argument. With probability $1/n$, vertex x is chosen to update its opinion. If this happens, then with probability p this update is a re-randomisation and further with probability $1/2$ its opinion becomes 1. Otherwise, with probability $q = 1 - p$, a standard voting step is performed, and thus x samples a vertex with opinion 1 with probability $P(x, A)$. By multiplying by $\pi(x)$ and summing over V, we have

$$(22) \qquad \mathbb{E}(S'|\xi) = \frac{n-1}{n}S + \frac{p}{2n} + \frac{q}{n}\sum_{x \in V}\pi(x)P(x, A).$$

Denote $Q(A, B) = \sum_{x \in A}\pi(x)P(x, B)$. For any irreducible transition matrix P, it holds that $Q(A, A^c) = Q(A^c, A)$ [30, Exercise 7.2]. Then $Q(V, A) = Q(A, A) + Q(A^c, A) = Q(A, A) + Q(A, A^c) = Q(A, V) = \pi(A)$. Thus

$$\sum_{x \in V}\pi(x)P(x, A) = \pi(A) = \sum_{x \in V}\pi(x)\xi(x) = S.$$

Plugging into equation (22) we obtain

$$\mathbb{E}(S'|\xi) = S\left(1 - \frac{p}{n}\right) + \frac{p}{2n}.$$

Observe that since S is a function of ξ we have $\mathbb{E}(\mathbb{E}(S'|\xi)|S) = \mathbb{E}(S'|S)$. We conclude

$$\mathbb{E}(S'|S) - \frac{1}{2} = \left(S - \frac{1}{2}\right)\left(1 - \frac{p}{n}\right).$$

Dividing by $\sigma = \sqrt{\mathbb{V}\mathrm{ar}(S)}$ on both sides completes the proof. $\qquad\square$

4. Gaussian approximation - proof of Theorem 1.

4.1. *Proof of equation* (2). In this section we prove equation (2) of Theorem 1. For that, we combine our previous results with Stein's method, leading not only to the asymptotic normality of sequences of NVM processes, but also to error rates. For such, we will employ a result from Röllin [40] (Theorem 34 in the Appendix) that gives an upper bound for the Kolmogorov distance between the random variable W and a standard real-valued Gaussian random variable. Röllin simplified the conditions of the so-called 'exchangeable pair' method developed in [39] which has been used to study the antivoter model.

Recall the definition of Ψ from Theorem 1 as

$$\Psi := \sum_{x \in V}\sum_{y \in V}\mu(x, y)\mathbf{1}_{\{\xi(x) \neq \xi(y)\}}.$$

PROPOSITION 16. *Consider a noisy voter model* NVM(V, P, p) *where* $|V| = n$. *Let* Φ *be the cumulative distribution function of a standard normal random variable. There exists a universal constant* $C > 0$ *such that*

$$(23) \qquad \sup_{t \in \mathbb{R}}|\mathbb{P}(W \leq t) - \Phi(t)| \leq C\left\{\left(\frac{\pi^*}{\sigma}\right)^3\frac{n}{p} + \left(\frac{\pi^*}{\sigma}\right)^2\sqrt{\frac{n}{p}} + \frac{\nu^2}{p\sigma^2}\sqrt{\mathbb{V}\mathrm{ar}(\Psi)}\right\}.$$

PROOF. Consider the random vector of opinions ξ sampled from Γ, and let ξ' be the vector ξ after one iteration of the noisy voter model from ξ. We consider $S = \sum_x \pi(x)\xi(x)$ and $S' = \sum_x \pi(x)\xi'(x)$, and denote by W and W' the corresponding standardised versions.

Note that $\mathbb{E}(W) = 0$ and $\mathbb{V}\mathrm{ar}(W) = 1$, and that ξ' has the same distribution as ξ since ξ is sampled from the stationary distribution (so W' has the same mean and variance as W). Moreover, note that $|W' - W| \leq \pi^*/\sigma$ since only one vertex may change its opinion. Finally, by Lemma 15 we have $\mathbb{E}(W'|W) = (1 - p/n)W$. Then, combining the three previous facts with Theorem 34 yields

$$(24) \qquad \sup_{t \in \mathbb{R}} |\mathbb{P}(W \leq t) - \Phi(t)| \leq \frac{12n}{p}\sqrt{\mathbb{V}\mathrm{ar}(\mathbb{E}((W' - W)^2|W))} + 32\frac{A^3}{\lambda} + 6\frac{A^2}{\sqrt{\lambda}},$$

where $A = \pi^*/\sigma$ and $\lambda = p/n$, and Φ is the cumulative distribution function of a standard Gaussian random variable; and further

$$(25) \qquad 32\frac{A^3}{\lambda} + 6\frac{A^2}{\sqrt{\lambda}} \leq 32\left(\left(\frac{\pi^*}{\sigma}\right)^3\frac{n}{p} + \left(\frac{\pi^*}{\sigma}\right)^2\sqrt{\frac{n}{p}}\right).$$

We now proceed to find an upper bound for $\mathbb{V}\mathrm{ar}(\mathbb{E}((W' - W)^2|W))$. We start by noticing that

$$\mathbb{V}\mathrm{ar}(\mathbb{E}((W' - W)^2|W)) \leq \mathbb{V}\mathrm{ar}(\mathbb{E}((W' - W)^2|\xi)),$$

since W is a function of ξ. Next, notice that since at most one vertex changes its opinion in one iteration of the noisy voter model we have

$$(W' - W)^2 = \sum_{x \in V}\left(\frac{\pi(x)}{\sigma}\right)^2 \mathbf{1}_{\{\xi'(x) \neq \xi(x)\}}$$

and thus

$$(26) \qquad \mathbb{E}((W' - W)^2|\xi) = \sum_{x \in V}\left(\frac{\pi(x)}{\sigma}\right)^2 \mathbb{P}(\xi'(x) \neq \xi(x)|\xi).$$

Denote by $D_x = D_x(\xi) := \{y \in V : \xi(x) \neq \xi(y)\} \subseteq V$ the (random) set of vertices whose opinion differs from the opinion of x, and then we claim that

$$(27) \quad \mathbb{P}(\xi'(x) \neq \xi(x)|\xi) = \frac{1}{n}\left(\frac{p}{2} + qP(x, D_x)\right) = \frac{1}{n}\left(\frac{p}{2} + q\sum_{y \in V}P(x,y)\mathbf{1}_{\{\xi(x) \neq \xi'(x)\}}\right).$$

To see this note first of all that for $\xi'(x)$ to differ from $\xi(x)$, we have to choose vertex x in the round, which has probability $1/n$. After that, with probability $p/2$ we re-randomise the opinion of x to the opposite opinion. Otherwise (with probability $q = 1 - p$), x chooses a random y vertex using $P(x, \cdot)$ and imitates its opinion, obtaining the opposite opinion if and only if $y \in D_x$. By plugging equation (27) into equation (26) we obtain

$$\mathbb{E}((W' - W)^2|\xi) = \sum_{x \in V}\left(\frac{\pi(x)}{\sigma}\right)^2 \mathbb{P}(\xi'(x) \neq \xi(x)|\xi)$$

$$= \frac{1}{n}\sum_{x \in V}\left(\frac{\pi(x)}{\sigma}\right)^2\left(\frac{p}{2} + q\sum_{y \in V}P(x,y)\mathbf{1}_{\{\xi(x) \neq \xi(y)\}}\right)$$

$$= \frac{p}{2n} \sum_{x \in V} \left(\frac{\pi(x)}{\sigma} \right)^2 + \frac{q}{n\sigma^2} \left(\sum_{x \in V} \sum_{y \in V} \pi(x)^2 P(x,y) \mathbf{1}_{\{\xi(x) \neq \xi(y)\}} \right)$$

$$(28) \qquad = \frac{p}{2n} \sum_{x \in V} \left(\frac{\pi(x)}{\sigma} \right)^2 + q \frac{\nu^2}{n\sigma^2} \left(\sum_{x \in V} \sum_{y \in V} \mu(x,y) \mathbf{1}_{\{\xi(x) \neq \xi(y)\}} \right)$$

$$(29) \qquad = \frac{p}{2n} \sum_{x \in V} \left(\frac{\pi(x)}{\sigma} \right)^2 + q \frac{\nu^2}{n\sigma^2} \Psi,$$

where in equation (28) have used the definition of μ, and in (29) the definition of Ψ.

Taking variance we have

$$\mathbb{V}\mathrm{ar}(\mathbb{E}((W' - W)^2 | \xi)) = (1-p)^2 \frac{\nu^4}{n^2 \sigma^4} \mathbb{V}\mathrm{ar}(\Psi) \leq \frac{\nu^4}{n^2 \sigma^4} \mathbb{V}\mathrm{ar}(\Psi),$$

concluding that

$$(30) \qquad \frac{12n}{p} \sqrt{\mathbb{V}\mathrm{ar}(\mathbb{E}((W' - W)^2 | W))} \leq \frac{12n}{p} \frac{\nu^2}{n\sigma^2} \sqrt{\mathbb{V}\mathrm{ar}(\Psi)}.$$

We complete the proof by combining equation (30) with equation (25), and by choosing $C = 32$. $\qquad \square$

4.2. *Bounding* $\mathbb{V}\mathrm{ar}(\Psi)$. In this section we consider the $\mathrm{NVM}(V, P, p)$ process with $|V| = n$, and provide a bound on $\mathbb{V}\mathrm{ar}(\Psi)$. Recall that in the (discrete-time) dual process $Y_t(x)$ denotes the location of particle initially at vertex $x \in V$ and we denote by $B(x)$, the final location of the particle that started at vertex x.

For $u, v, x, y \in V$, we set

$$h(u,v,x,y) := \mathbb{P}(B(x) \neq B(u), B(y) \neq B(v)) - \mathbb{P}(B(x) \neq B(u))\mathbb{P}(B(y) \neq B(v))$$

$$= \mathbb{C}ov(\mathbf{1}_{\{B(x) \neq B(u)\}}, \mathbf{1}_{\{B(y) \neq B(v)\}})$$

$$(31) \qquad = \mathbb{C}ov(\mathbf{1}_{\{\xi(x) \neq \xi(u)\}}, \mathbf{1}_{\{\xi(y) \neq \xi(v)\}})$$

where the last equality follow from Proposition 11 assuming $\xi \sim \Gamma$. With this notation we can write

$$\mathbb{V}\mathrm{ar}(\Psi) = \mathbb{V}\mathrm{ar} \left(\sum_{x \in V} \sum_{u \in V} \mu(x,u) \mathbf{1}_{\{\xi(x) \neq \xi(u)\}} \right)$$

$$= \sum_{x,u,y,v \in V} \mu(x,u)\mu(y,v) \mathbb{C}ov \left(\mathbf{1}_{\{\xi(x) \neq \xi(u)\}}, \mathbf{1}_{\{\xi(y) \neq \xi(v)\}} \right)$$

$$= \sum_{x,u,y,v \in V} \mu(x,u)\mu(y,v) h(u,v,x,y).$$

Denote $h_0(u,v,x,y) = \max\{0, h(u,v,x,y)\}$, then

$$(32) \qquad \mathbb{V}\mathrm{ar}(\Psi) \leq \frac{(\pi^*)^2}{\nu^4} \sum_{x,u,y,v \in V} \pi(x)P(x,u)\pi(y)P(y,v)h_0(u,v,x,y).$$

A careful analysis of the function h_0 leads to the following result.

LEMMA 17. *In* $\mathrm{NVM}(V, P, p)$ *we have* $\mathbb{V}\mathrm{ar}(\Psi) \leq 16(\pi^*)^2 \nu^{-4} \sigma^2$.

PROOF OF THEOREM 1. Note that Lemma 17 is exactly the bound of equation (3), which applied to (2) leads to the condition stated in equation (4), and thus Theorem 1 is proved. $\square$

We remark that to obtain better bounds on $\mathbb{V}\mathrm{ar}(\Psi)$ we need to rely on equation (32) in combination with estimates of h_0 which depend on the geometry of the graph/transition matrix. Indeed, this is how we proceed for the 2D torus and the cycle.

4.3. *Proof of Lemma 17.* The proof of Lemma 17 needs a subtle analysis of the function h_0 defined in (31). Our analysis makes heavy use of the dual process in discrete time– recall the notation for meeting and absorption times from Section 3. For ease of notation in this section we shall write $M^{x,u}$ for $M^{Y(x),Y(u)}$ and τ^x for $\tau^{Y(x)}$. For $u, v, x, y \in V$, recall the definition of $h_0(u, v, x, y)$ and set

$$(33) \qquad \mathcal{T}^{u,v,x,y} := \{\tau^u, \tau^v, \tau^x, \tau^y, M^{u,v}, M^{v,x}, M^{u,x}, M^{u,y}, M^{v,y}, M^{x,y}\}$$

The following proposition is crucial.

PROPOSITION 18. *Fix $u, v, x, y \in V$ and denote by $T_1 < T_2 < T_3$ the three smallest elements of $\mathcal{T}^{u,v,x,y}$. Then*

$$h_0(u,v,x,y) \leq \mathbb{P}(T_1 = M^{x,y}, T_2 = \tau^x, T_3 = M^{u,v}) + \mathbb{P}(T_1 = M^{x,y}, T_2 = M^{u,v})$$
$$+ \mathbb{P}(T_1 = M^{x,v}, T_2 = \tau^x, T_3 = M^{u,y}) + \mathbb{P}(T_1 = M^{x,v}, T_2 = M^{u,y})$$
$$+ \mathbb{P}(T_1 = M^{u,v}, T_2 = \tau^u, T_3 = M^{x,y}) + \mathbb{P}(T_1 = M^{u,v}, T_2 = M^{x,y})$$
$$+ \mathbb{P}(T_1 = M^{u,y}, T_2 = \tau^u, T_3 = M^{x,v}) + \mathbb{P}(T_1 = M^{u,y}, T_2 = M^{x,v}).$$

(34)

For the proof of Lemma 17, the following corollary will suffice.

COROLLARY 19. *For each $u, v, x, y \in V$, $h_0(u,v,x,y) \leq \mathbb{P}(M^{x,y} < \tau^x \wedge \tau^y) + \mathbb{P}(M^{x,v} < \tau^x \wedge \tau^v) + \mathbb{P}(M^{u,v} < \tau^u \wedge \tau^v) + \mathbb{P}(M^{u,y} < \tau^u \wedge \tau^y).$*

PROOF OF LEMMA 17. For $x, y \in V$ define $E_{x,y} := \{M^{x,y} < \tau^x \wedge \tau^y\}$, then by combining (32) with Corollary 19 we obtain

$\mathbb{V}\mathrm{ar}(\Psi)$

$$\leq \frac{(\pi^*)^2}{\nu^4} \sum_{x,u,y,v \in V} \pi(x) P(x,u) \pi(y) P(y,v) \left(\mathbb{P}(E_{x,y}) + \mathbb{P}(E_{x,v}) + \mathbb{P}(E_{u,v}) + \mathbb{P}(E_{u,y}) \right).$$

We claim that the right-hand side above is equal to $4 \frac{(\pi^*)^2}{\nu^4} \sum_{x,y \in V} \pi(x) \pi(y) \mathbb{P}(E_{x,y})$. To see this, observe that

$$\sum_{x,u,y,v \in V} \pi(x) P(x,u) \pi(y) P(y,v) \mathbb{P}(E_{x,v}) = \sum_{x,v \in V} \pi(x) \sum_{y \in V} \pi(y) P(y,v) \mathbb{P}(E_{x,v})$$
$$= \sum_{x,v \in V} \pi(x) \pi(v) \mathbb{P}(E_{x,v}),$$

and similarly

$$\sum_{x,u,y,v \in V} \pi(x) P(x,u) \pi(y) P(y,v) \mathbb{P}(E_{u,v}) = \sum_{u,v \in V} \pi(u) \pi(v) \mathbb{P}(E_{u,v}),$$
$$\sum_{x,u,y,v \in V} \pi(x) P(x,u) \pi(y) P(y,v) \mathbb{P}(E_{u,y}) = \sum_{u,y \in V} \pi(u) \pi(y) \mathbb{P}(E_{u,y}),$$

which proves the claim. We deduce that

$$\mathbb{V}\mathrm{ar}(\Psi) = 4\frac{(\pi^*)^2}{\nu^4}\sum_{x,u,y,v\in V}\pi(x)P(x,u)\pi(y)P(y,v)\mathbb{P}(E_{x,y})$$

$$= 4\frac{(\pi^*)^2}{\nu^4}\sum_{x,y\in V}\pi(x)\pi(y)\mathbb{P}(E_{x,y}) = 16\frac{(\pi^*)^2}{\nu^4}\sigma^2,$$

where the final equality holds due to Lemma 14. $\qquad\square$

We now focus on proving Proposition 18. For that, we require some preliminary results that use similar martingale arguments. We begin with a simple lemma.

LEMMA 20. *For distinct $w, z \in V$, define $f(w,z) := \mathbb{P}(B(w) = B(z)) - 1/2$ and for each $t \in \mathbb{N}_0$ set $F_t = f(Y_t(w), Y_t(z))$. Then (F_t) is a martingale up to time $R = \min\{\tau^w, \tau^z, M^{w,z}\}$. Moreover, $f(w,z) = \frac{1}{2}\mathbb{P}(R = M^{w,z})$.*

PROOF. It is immediate that F_t is a martingale. We then apply optional stopping at time R which is valid since the martingale is bounded and $R < \infty$ almost surely (the chain P is finite and irreducible). $\qquad\square$

LEMMA 21. *For distinct a, c, d, define $g(a,c,d) := \mathbb{P}(B(d) \neq B(a), B(d) \neq B(c)) - \mathbb{P}(B(d) \neq B(a))\mathbb{P}(B(d) \neq B(c))$ and for each $t \in \mathbb{N}_0$ set $G_t := g(Y_t(a), Y_t(c), Y_t(d))$. Then (G_t) is a martingale up to time $T := \min\{\tau^a, \tau^c, \tau^d, M^{a,c}, M^{a,d}, M^{c,d}\}$. Moreover*

$$g(a,c,d) = \mathbb{E}[G_T\mathbf{1}_{\{T=\tau^d\}} + G_T\mathbf{1}_{\{T=M^{a,c}\}}].$$

PROOF. The fact that G_t is a martingale is immediate. We apply optional stopping at time T to obtain $g(a,c,d) = \mathbb{E}[G_T]$. Note that almost surely,

$$G_T\mathbf{1}_{\{T=\tau^a\}} = \left(\frac{1}{2}\mathbb{P}(B(Y_T(d)) \neq B(Y_T(c))) - \frac{1}{2}\mathbb{P}(B(Y_T(d)) \neq B(Y_T(c)))\right)\mathbf{1}_{\{T=\tau^a\}} = 0.$$

Similarly $G_T\mathbf{1}_{\{T=\tau^c\}} = 0$ almost surely. Furthermore, $G_T\mathbf{1}_{\{T=M^{a,d}\}} = G_T\mathbf{1}_{\{T=M^{d,c}\}} = 0$. Hence $g(a,c,d) = \mathbb{E}[G_T\mathbf{1}_{\{T=\tau^d\}} + G_T\mathbf{1}_{\{T=M^{a,c}\}}]$. $\qquad\square$

Combining these two lemmas we can obtain the following:

COROLLARY 22. *For distinct $a, c, d \in V$ set $G_t := g(Y_t(a), Y_t(c), Y_t(d))$. Define times*

$$T := \min\{\tau^a, \tau^c, \tau^d, M^{a,c}, M^{a,d}, M^{c,d}\},$$

$$R := \inf\{t > T : t \in \{\tau^a, \tau^c, \tau^d, M^{a,c}, M^{a,d}, M^{c,d}\}\}.$$

Then

$$g(a,c,d) \leq (1/4)\mathbb{P}(T = \tau^d, R = M^{a,c}) + \mathbb{P}(T = M^{a,c}).$$

PROOF. This follows by applying Lemmas 20 and 21, the strong Markov property at time T, and the bound $G_t \leq 1$ which holds for all times t. Indeed, we have

$$\mathbf{1}_{\{T=\tau^d\}}G_T = \frac{1}{2}\mathbf{1}_{\{T=\tau^d\}}f(Y_T(a), Y_T(c)),$$

and by the strong Markov property at time T and Lemma 20

$$\mathbb{E}[\mathbf{1}_{\{T=\tau^d\}} f(Y_T(a), Y_T(c))] = \frac{1}{2}\mathbb{P}(T = \tau^d, R = M^{a,c}).$$

Thus by Lemma 21 and the bound $G_T \leq 1$,

$$g(a, c, d) \leq \frac{1}{4}\mathbb{P}(T = \tau^d, R = M^{a,c}) + \mathbb{P}(T = M^{a,c}). \qquad \square$$

LEMMA 23. *For distinct $u, v, x, y \in V$, define $H_t := h(Y_t(u), Y_t(v), Y_t(x), Y_t(y))$ and recall the definition of $\mathcal{T}^{u,v,x,y}$ in equation* (33). *Then (H_t) is a martingale up to time $\varrho :=$ $\min \mathcal{T}^{u,v,x,y}$. Moreover*

$$h(u, v, x, y) = \mathbb{E}\left[H_\varrho \mathbf{1}_{\{\varrho \in \{M^{x,y}, M^{x,v}, M^{u,v}, M^{u,y}\}\}}\right].$$

PROOF. The fact that H_t is a martingale is immediate. Applying optional stopping gives $h(u, v, x, y) = \mathbb{E}[H_\varrho]$. However note that almost surely

$$H_\varrho \mathbf{1}_{\{\varrho=\tau^u\}} = \left(\frac{1}{2}\mathbb{P}(B(Y_\varrho(y)) \neq B(Y_\varrho(v))) - \frac{1}{2}\mathbb{P}(B(Y_\varrho(y)) \neq B(Y_\varrho(v)))\right) \mathbf{1}_{\{\varrho=\tau^u\}} = 0,$$

and similarly $H_\varrho \mathbf{1}_{\{\varrho=\tau^v\}} = H_\varrho \mathbf{1}_{\{\varrho=\tau^x\}} = H_\varrho \mathbf{1}_{\{\varrho=\tau^y\}} = 0$ almost surely. Furthermore, $H_\varrho \mathbf{1}_{\{\varrho=M^{x,u}\}} = H_\varrho \mathbf{1}_{\{\varrho=M^{y,v}\}} = 0$. This gives the stated identity. $\qquad \square$

PROOF OF PROPOSITION 18. First observe that if $h < c$ for some $c > 0$ then it follows that $h_0 < c$. From Lemma 23 we have

$$h(u, v, x, y)$$
$$= \mathbb{E}\left[H_{T_1}\mathbf{1}_{\{T_1=M^{x,y}\}} + H_{T_1}\mathbf{1}_{\{T_1=M^{x,v}\}} + H_{T_1}\mathbf{1}_{\{T_1=M^{u,v}\}} + H_{T_1}\mathbf{1}_{\{T_1=M^{u,y}\}}\right]$$
$$= \mathbb{E}\Big[g(Y_{T_1}(u), Y_{T_1}(v), Y_{T_1}(x))\mathbf{1}_{\{T_1=M^{x,y}\}} + g(Y_{T_1}(u), Y_{T_1}(y), Y_{T_1}(x))\mathbf{1}_{\{T_1=M^{x,v}\}}$$
$$+ g(Y_{T_1}(x), Y_{T_1}(y), Y_{T_1}(u))\mathbf{1}_{\{T_1=M^{u,v}\}} + g(Y_{T_1}(x), Y_{T_1}(v), Y_{T_1}(u))\mathbf{1}_{\{T_1=M^{u,y}\}}\Big].$$

By the strong Markov property at time T_1 together with Corollary 22 we obtain the claimed bound. $\qquad \square$

5. Torus and Cycle - Proof of Proposition 9.

5.1. *The Torus.* In this section, we consider a graph $G_n = (V_n, E_n)$ the torus $\mathbb{T}_n$ on n vertices, and we let P_n be the simple random walk on $\mathbb{T}_n$.

PROPOSITION 24. *There exists a universal constant $C > 0$ (not depending on n) such that for any torus $\mathbb{T}_n$ with $n \geq 9$,*

$$\mathbb{Var}(\Psi) \leq C\left(\frac{1}{\sqrt{n}} + \frac{\sigma^2}{(\log n)^2}\right).$$

For the proof of this proposition we will consider the continuous-time version of the dual process (see Remark 12), where particles move at rate 1 in the graph, and are absorbed at rate $\delta = p/(1-p)$. Further, recall that walks move independently until they meet (or are absorbed), and so couplings with the process $\mathrm{RW}(m)$ of m independent random walks as defined at the beginning of Section 3 will feature in our analysis, indeed the following two results about $\mathrm{RW}(2)$ on the torus are needed.

LEMMA 25.   *Let (X_t, Y_t) be a realisation of $\mathrm{RW}(2)$ on the torus in continuous time. Let $M^{X,Y}$ denote the meeting time of the two walks. Then for all $t \geq 0$,*

$$\mathbb{P}_{\pi^2}^{\mathrm{RW}(2)}(M^{X,Y} \leq t) \leq (2t+1)\nu^2.$$

LEMMA 26.   *Let (X_t, Y_t) be a realisation of $\mathrm{RW}(2)$ on the torus in continuous time. Let x and y be two adjacent vertices on the torus. Then there exists a universal constant $C > 0$ (not depending on n) such that for each $t \geq e$,*

$$\mathbb{P}_{(x,y)}^{\mathrm{RW}(2)}(M^{X,Y} > t) \leq C/\log(t).$$

PROOF OF PROPOSITION 9 ITEM 2. To start with, note that the case $p_n n \log n \to 0$ follows immediately from Proposition 6, since the meeting time of two independent stationary walks is $\Theta(n \log n)$.

Now, suppose that $p_n n \log n \to \infty$. Recall that p_n is non-increasing hence $\lim_n p_n$ exists. If $\lim_n p_n > 0$ then the result follows from Corollary 3. So suppose $\lim_n p_n = 0$. By Theorem 1 it suffices to show

$$\frac{n}{p}\left(\frac{\pi^*}{\sigma}\right)^3 + \frac{\nu^2}{p\sigma^2}\sqrt{\mathbb{V}\mathrm{ar}(\Psi)} \to 0 \quad \text{as } n \to \infty,$$

(to ease notation, we are not writing the subindex n) i.e.

$$\frac{1}{n^2 p\sigma^3} + \frac{\sqrt{\mathbb{V}\mathrm{ar}(\Psi)}}{np\sigma^2} \to 0 \quad \text{as } n \to \infty.$$

Eventually $p \leq 1/2$ so by Lemma 2, $\sigma^2 \gtrsim \frac{1}{pt_{\mathrm{hit}}} \gtrsim \frac{1}{pn\log n}$ (see [30, Proposition 10.13]), hence

$$\frac{1}{n^2 p\sigma^3} \lesssim \sqrt{p(\log n)^3/n} \to 0 \quad \text{as } n \to \infty.$$

By Proposition 24 it remains to show that

$$\frac{\sqrt{\frac{1}{\sqrt{n}} + \frac{\sigma^2}{(\log n)^2}}}{np\sigma^2} \leq \frac{1}{n^{5/4}p\sigma^2} + \frac{1}{np\sigma\log n} \to 0 \quad \text{as } n \to \infty.$$

Using again $\sigma^2 \gtrsim \frac{1}{pn\log n}$ we have the bound

$$\frac{1}{n^{5/4}p\sigma^2} + \frac{1}{np\sigma\log n} \lesssim \frac{\log n}{n^{1/4}} + \frac{1}{\sqrt{np\log n}} \to 0 \quad \text{as } n \to \infty,$$

by the assumption $pn\log n \gg 1$, completing the proof. $\square$

We proceed to prove Proposition 24. Our proof uses the dual process. For ease of notation in we shall write $M^{x,u}$ for $M^{Y(x),Y(u)}$ and τ^x for $\tau^{Y(x)}$, etc.

PROOF OF PROPOSITION 24. To bound h_0 we use Proposition 18. Consider two terms at the top of the right-hand side of equation (34), given by

$$\mathbb{P}(T_1 = M^{x,y}, T_2 = \tau^x, T_3 = M^{u,v}) + \mathbb{P}(T_1 = M^{x,y}, T_2 = M^{u,v}).$$

Here we are considering the continuous-time version of the dual. Note the previous events are disjoint and establish that the meeting of x and y occurs before any other meeting or absorption, so, in particular, we have

$$\mathbb{P}(T_1 = M^{x,y}, T_2 = \tau^x, T_3 = M^{u,v}) + \mathbb{P}(T_1 = M^{x,y}, T_2 = M^{u,v}) \leq \mathbb{P}(T_1 = M^{x,y}).$$

Note that up to time T_1 no meeting has occurred, so the four particles involved move like independent random walks up to this time, thus it is natural to couple $(Y_t(x), Y_t(y), Y_t(u), Y_t(v))$ up to time T_1 with an independent random walks process RW(4) that is stopped when a meeting occurs or when an independent exponential clock of rate $4p/(1-p)$ rings (representing the first absorption time of the four walks). Let (X_t, Y_t, U_t, V_t) be a RW(4) process, let T be the time of the first meeting between any pair of the walks in RW(4), let $M^{X,Y}$ be the meeting time between (X_t) and (Y_t), and let Z be an exponential random variable of parameter 4δ with $\delta = p/(1-p)$, which is independent of the walks. Then we have

$$\mathbb{P}(T_1 = M^{x,y}) = \mathbb{P}^{\mathrm{RW}(4)}_{(x,y,u,v)}(M^{X,Y} = T, M^{X,Y} < Z)$$

$$\leq \mathbb{P}^{\mathrm{RW}(4)}_{(x,y,u,v)}(M^{X,Y} < M^{X,U} \wedge M^{Y,V} \wedge Z)$$

Repeating the same argument to the other terms of equation (34), using reversibility of P, and plugging into equation (32) yields

$$\mathbb{V}\mathrm{ar}(\Psi)$$

$$\leq 4\frac{(\pi^*)^2}{\nu^4} \sum_{x,y,u,v \in V} \pi(x)P(x,u)\pi(y)P(y,v)\mathbb{P}^{\mathrm{RW}(4)}_{(x,y,u,v)}(M^{X,Y} \leq M^{X,U} \wedge M^{Y,V} \wedge Z)$$

$$(35) \quad = 4 \sum_{x,y,u,v \in V} \frac{\mathbf{1}_{\{x \sim u\}}\mathbf{1}_{\{y \sim v\}}}{16n^2}\mathbb{P}^{\mathrm{RW}(4)}_{(x,y,u,v)}(M^{X,Y} \leq M^{X,U} \wedge M^{Y,V} \wedge Z).$$

We proceed to bound the probability in (35). We write members of V^4 in bold, i.e. $\boldsymbol{x} = (x,y,u,v) \in V^4$, and write $\boldsymbol{X}_t$ for (X_t, Y_t, U_t, V_t). As previously, $M^{X,Y}$ refers to the meeting time of random walks X and Y.

Let $\boldsymbol{x} \in V^4$. If $x \neq y$ then $\mathbb{P}^{\mathrm{RW}(4)}_{\boldsymbol{x}}(M^{X,Y} = 0, M^{X,Y} \leq M^{X,U} \wedge M^{Y,V}) = 0$, so in such case we denote by $f_{\boldsymbol{x}}(t)$ the derivative of

$$t \mapsto \mathbb{P}^{\mathrm{RW}(4)}_{\boldsymbol{x}}(M^{X,Y} \leq t, M^{X,Y} \leq M^{X,U} \wedge M^{y,v}),$$

which exists since the meeting between two particles is a continuous random variable when the particle start in different locations.

Then we can write $f_{\boldsymbol{x}}$ in terms of the four independent random walks. Indeed:

$$f_{\boldsymbol{x}}(t) = \frac{1}{2}\mathbb{P}^{\mathrm{RW}(4)}_{\boldsymbol{x}}(\{M^{X,Y} > t, M^{X,U} > t, M^{Y,V} > t\} \cap \{X_t \sim Y_t\}),$$

which holds because the instantaneous probability that X_t and Y_t meet for first time at time t is exactly the probability that they do not meet up to time t, then at time t one walker moves to the location of the other (so they have to be neighbours just prior to time t). The rate of jumping to an adjacent neighbour is $1/4$, giving the equation above.

Write $\mathcal{E}_t = \{M^{X,Y} > t, M^{X,U} > t, M^{Y,V} > t\}$; then $f_{\boldsymbol{x}}(t)$ can be written as

$$f_{\boldsymbol{x}}(t) = \frac{1}{2}\mathbb{P}^{\mathrm{RW}(4)}_{\boldsymbol{x}}(\mathcal{E}_t, X_t \sim Y_t).$$

By using the Markov property and reversibility we have

$$f_{\boldsymbol{x}}(t) = \frac{1}{2} \sum_{\boldsymbol{x}' \in V^4} \mathbb{P}^{\mathrm{RW}(4)}_{\boldsymbol{x}}(\mathcal{E}_{t/2}, \boldsymbol{X}_{t/2} = \boldsymbol{x}')\mathbb{P}^{\mathrm{RW}(4)}_{\boldsymbol{x}'}(\mathcal{E}_{t/2}, X_{t/2} \sim Y_{t/2})$$

$$= \frac{1}{2} \sum_{\boldsymbol{x}' \in V^4} \mathbb{P}^{\mathrm{RW}(4)}_{\boldsymbol{x}'}(\mathcal{E}_{t/2}, \boldsymbol{X}_{t/2} = \boldsymbol{x})\mathbb{P}^{\mathrm{RW}(4)}_{\boldsymbol{x}'}(\mathcal{E}_{t/2}, X_{t/2} \sim Y_{t/2})$$

$$\leq \frac{1}{2} \sum_{\boldsymbol{x'} \in V^4} \mathbb{P}^{\mathrm{RW}(4)}_{\boldsymbol{x'}} (M^{X,U} > t/2, M^{Y,V} > t/2, \boldsymbol{X}_{t/2} = \boldsymbol{x})$$

$$\times \mathbb{P}^{\mathrm{RW}(4)}_{\boldsymbol{x'}} (M^{X,Y} > t/2, X_{t/2} \sim Y_{t/2})$$

$$= \frac{1}{2} \sum_{\boldsymbol{x'} \in V^4} \mathbb{P}^{\mathrm{RW}(2)}_{(x',u')} (M^{X,U} > t/2, (X_{t/2}, U_{t/2}) = (x,y))$$

(36)
$$\times \mathbb{P}^{\mathrm{RW}(2)}_{(y',v')} (M^{Y,V} > t/2, (Y_{t/2}, V_{t/2}) = (y,v))$$

$$\times \mathbb{P}^{\mathrm{RW}(2)}_{(x',y')} (M^{X,Y} > t/2, X_{t/2} \sim Y_{t/2}),$$

where $\mathbb{P}^{\mathrm{RW}(2)}$ is the probability measure associated with two independent random walks.

Denote by $\mathcal{F}_t$ the event $\{M^{X,Y} > t, X_t \sim Y_t\}$, then from (36) we have

$$\sum_{\boldsymbol{x} \in V^4} \frac{\mathbf{1}_{\{x \sim u\}} \mathbf{1}_{\{y \sim v\}}}{16n^2} f_{\boldsymbol{x}}(t) \leq \frac{2}{64n^2} \sum_{\boldsymbol{x'} \in V^4} \mathbb{P}^{\mathrm{RW}(2)}_{(x',u')}(\mathcal{F}_{t/2}) \mathbb{P}^{\mathrm{RW}(2)}_{(y',v')}(\mathcal{F}_{t/2}) \mathbb{P}^{\mathrm{RW}(2)}_{(x',y')}(\mathcal{F}_{t/2})$$

$$= \frac{1}{32} \sum_{(x',y') \in V^2} \mathbb{P}^{\mathrm{RW}(2)}_{(x',\pi)}(\mathcal{F}_{t/2}) \mathbb{P}^{\mathrm{RW}(2)}_{(y',\pi)}(\mathcal{F}_{t/2}) \mathbb{P}^{\mathrm{RW}(2)}_{(x',y')}(\mathcal{F}_{t/2}).$$

Since the torus is transitive, we have that $\mathbb{P}^{\mathrm{RW}(2)}_{(x',\pi)}(\mathcal{F}_{t/2})$ does not depend on x', so $\mathbb{P}^{\mathrm{RW}(2)}_{(x',\pi)}(\mathcal{F}_{t/2}) = \mathbb{P}^{\mathrm{RW}(2)}_{(\pi,\pi)}(\mathcal{F}_{t/2})$, concluding from the previous equation that

$$\sum_{\boldsymbol{x} \in V^4} \frac{\mathbf{1}_{\{x \sim u\}} \mathbf{1}_{\{y \sim v\}}}{16n^2} f_{\boldsymbol{x}}(t) \leq \frac{1}{32} \sum_{(x',y') \in V^2} \mathbb{P}^{\mathrm{RW}(2)}_{(\pi,\pi)}(\mathcal{F}_{t/2}) \mathbb{P}^{\mathrm{RW}(2)}_{(\pi,\pi)}(\mathcal{F}_{t/2}) \mathbb{P}^{\mathrm{RW}(2)}_{(x',y')}(\mathcal{F}_{t/2})$$

(37)
$$= \frac{n^2}{32} \mathbb{P}^{\mathrm{RW}(2)}_{(\pi,\pi)}(\mathcal{F}_{t/2})^3.$$

Using reversibility again we obtain

$$\mathbb{P}^{\mathrm{RW}(2)}_{(\pi,\pi)}(\mathcal{F}_{t/2}) = \sum_{x',y' \in V} \frac{1}{n^2} \mathbb{P}^{\mathrm{RW}(2)}_{(x',y')}(\mathcal{F}_{t/2})$$

$$= \sum_{x',y' \in V} \sum_{x,y \in V} \mathbf{1}_{\{x \sim y\}} \frac{1}{n^2} \mathbb{P}^{\mathrm{RW}(2)}_{(x',y')}(M^{X,Y} > t/2, X_{t/2} = x, Y_{t/2} = y)$$

$$= \sum_{x',y' \in V} \sum_{x,y \in V} \mathbf{1}_{\{x \sim y\}} \frac{1}{n^2} \mathbb{P}^{\mathrm{RW}(2)}_{(x,y)}(M^{X,Y} > t/2, X_{t/2} = x', Y_{t/2} = y')$$

$$= \sum_{x,y \in V} \mathbf{1}_{\{x \sim y\}} \frac{1}{n^2} \mathbb{P}^{\mathrm{RW}(2)}_{(x,y)}(M^{X,Y} > t/2)$$

$$= \frac{4}{n} \sum_{x,y} \mu(x,y) \mathbb{P}^{\mathrm{RW}(2)}_{(x,y)}(M^{X,Y} > t/2).$$

By denoting $\mathbb{P}_\mu^{\mathrm{RW}(2)}(\cdot) = \sum_{x,y} \mu(x,y) \mathbb{P}_{(x,y)}^{\mathrm{RW}(2)}(\cdot)$, and substituting the previous equality into (37) we have

$$(38) \qquad \sum_{\boldsymbol{x} \in V^4} \frac{\mathbf{1}_{\{x \sim u\}} \mathbf{1}_{\{y \sim v\}}}{16n^2} f_{\boldsymbol{x}}(t) \leq \frac{1}{8} \mathbb{P}_\mu^{\mathrm{RW}(2)}(M^{X,Y} > t/2)^2 \mathbb{P}_{(\pi,\pi)}^{\mathrm{RW}}(\mathcal{F}_{t/2}).$$

Going back to equation (35), by (38) we have

$$\mathbb{V}\mathrm{ar}(\Psi) \leq 4 \sum_{x,y,u,v \in V} \frac{\mathbf{1}_{\{x \sim u\}} \mathbf{1}_{\{y \sim v\}}}{16n^2} \mathbb{P}_{(x,y,u,v)}^{\mathrm{RW}(4)}(M^{X,Y} \leq M^{X,U} \wedge M^{Y,V} \wedge Z)$$

$$\leq 4 \sum_{x,y,u,v \in V} \frac{\mathbf{1}_{\{x \sim u\}} \mathbf{1}_{\{y \sim v\}}}{16n^2} \mathbb{P}_{(x,y)}^{\mathrm{RW}(2)}(M^{X,X} \leq 2\sqrt{n})$$

$$+ 4 \sum_{x,y,u,v \in V} \frac{\mathbf{1}_{\{x \sim u\}} \mathbf{1}_{\{y \sim v\}}}{16n^2} \mathbb{P}_{(x,y,u,v)}^{\mathrm{RW}(4)}(2\sqrt{n} \leq M^{X,Y} \leq M^{X,U} \wedge M^{Y,V} \wedge Z)$$

$$= \frac{1}{4} \mathbb{P}_{\pi^2}^{\mathrm{RW}(2)}(M^{X,Y} \leq 2\sqrt{n}) + 4 \sum_{x,y,u,v \in V} \frac{\mathbf{1}_{\{x \sim u\}} \mathbf{1}_{\{y \sim v\}}}{16n^2} \int_{2\sqrt{n}}^\infty f_{\boldsymbol{x}}(t) e^{-4\delta t}\, dt$$

$$\leq \cdot \frac{5}{4\sqrt{n}} + 4 \int_{2\sqrt{n}}^\infty \frac{1}{8} \mathbb{P}_\mu^{\mathrm{RW}(2)}(M^{X,Y} > t/2)^2 \, \mathbb{P}_{(\pi,\pi)}^{\mathrm{RW}(2)}(\mathcal{F}_{t/2}) e^{-4\delta t}\, dt$$

$$\leq \frac{5}{4\sqrt{n}} + c \int_{2\sqrt{n}}^\infty (\log n)^{-2} \, \mathbb{P}_{(\pi,\pi)}^{\mathrm{RW}(2)}(\mathcal{F}_{t/2}) e^{-4\delta t}\, dt,$$

for a universal constant $c > 0$. The bound on $\mathbb{P}_{\pi^2}^{\mathrm{RW}(2)}(M^{X,Y} \leq 2\sqrt{n})$ follows from Lemma 25 and the bound on $\mathbb{P}_\mu^{\mathrm{RW}(2)}(M^{X,Y} > t/2)^2$ is from Lemma 26. Now recall that $\frac{1}{8} \mathbb{P}_{(\pi,\pi)}^{\mathrm{RW}}(\mathcal{F}_{t/2})$ is the density of $M^{X,Y}$ at time $t/2$ when both particles start independently with distribution π, which we shall denote by $g(t/2)$. Then,

$$\mathbb{V}\mathrm{ar}(\Psi)$$

$$\leq \frac{5}{4\sqrt{n}} + 8c \int_{2\sqrt{n}}^\infty (\log n)^{-2} g(t/2) e^{-4\delta t}\, dt \leq \frac{5}{4\sqrt{n}} + 8c \int_0^\infty (\log n)^{-2} g(s) e^{-2\delta s}\, ds$$

$$\leq \frac{5}{4\sqrt{n}} + 8c(\log n)^{-2} \mathbb{E}_{(\pi,\pi)}^{\mathrm{RW}(2)}\left[e^{-2\delta M^{X,Y}} \right] = \frac{5}{4\sqrt{n}} + 8c(\log n)^{-2} \mathbb{P}_{\pi^2}(M^{XY} < Z')$$

$$\leq \hat{c} \left(\frac{1}{\sqrt{n}} + \sigma^2 (\log n)^{-2} \right),$$

where Z' is an exponential random variable of parameter 2δ (independent of everything), $c > 0$ and $\hat{c} > 0$ are universal constants, and the last inequality holds due to Lemma 14. $\qquad \square$

We now complete this section by presenting the proofs of Lemmas 25 and 26.

PROOF OF LEMMA 25. To ease notation, and since there are no other processes in the proof, we omit the superscript $\mathrm{RW}(2)$ in the probability measures. Now, denote by $N(t)$ the total number of jumps made by the two walks by time t. Then we claim that for each $t \geq 0$ and $j \in \mathbb{N}$,

$$\mathbb{P}_{\pi^2}(M^{X,Y} \leq t \mid N(t) = j) \leq (j+1)\nu^2.$$

To see this, denote by $T_1, \ldots, T_j$ the jump times. Then at these times the distribution of the locations of the walks is exactly the stationary distribution by the strong Markov property (and so the probability they are equal is $\nu^2 = \sum_{x \in V} \pi(x)^2$). Thus

$$\mathbb{P}_{\pi^2}(M^{X,Y} \leq t \mid N(t) = j) \leq \mathbb{P}_{\pi^2}(X_0 = Y_0) + \sum_{i=1}^{j} \mathbb{P}_{\pi^2}(X_{T_i} = Y_{T_i}) = (j+1)\nu^2.$$

We conclude that $\mathbb{P}_{\pi^2}(M^{X,Y} \leq t) \leq \nu^2(\mathbb{E}_{\pi^2}(N(t)+1)) = \nu^2(2t+1)$ since $N(t)$ has Poisson distribution of mean $2t$. $\qquad\square$

PROOF OF LEMMA 26. As in the proof of Lemma 25 we omit the superscript $\mathrm{RW}(2)$ in the probability measure. The first step is to assume that the random walks move on $\mathbb{Z}^2$ instead of $\mathbb{T}_n$, which clearly provides an upper-bound on the quantity of interest. Recall that $x = X_0$ and $y = Y_0$ are adjacent vertices in $\mathbb{Z}^2$, and by transitivity we can assume that $y = (0,0) \in \mathbb{Z}^2$ and $x = (1,0) \in \mathbb{Z}^2$. Also, by further using the symmetry of $\mathbb{Z}^2$, we can assume that only one walk moves but at rate 2, while the other is fixed. Then, denote by Z_t a simple random walk on $\mathbb{Z}^2$ moving at rate 2, let o be the origin $(0,0) \in \mathbb{Z}^2$ and $x = (1,0) \in \mathbb{Z}^2$, and let H_o be the hitting time of o by the random walk Z_t. By our previous discussion, we deduce that

$$\mathbb{P}_{(x,y)}(M^{X,Y} > t) \leq \mathbb{P}_x(H_o > t),$$

We condition on the number of jumps $N(t)$ the walk Z_t performs up to time t which is distributed as a Poisson random variable with mean $2t$. Denote by $\widehat{H}_o$ the number of jumps until Z_t hit vertex o, and by $M(j)$ the number of times the random walk jumps into vertex o up to (and including) the j-th jump. Then

$$(39) \qquad \mathbb{P}_x(H_o > t \mid N(t) = j) = \mathbb{P}_x(\widehat{H}_o > j) = 1 - \mathbb{P}_x(M(j) > 0).$$

Note that conditioning on $\{N(t) = 2j\}$ or $\{N(t) = 2j+1\}$ is the same since $\mathbb{Z}$ is bipartite, i.e. $\mathbb{P}_x(M(2j) > 0) = \mathbb{P}_x(M(2j+1) > 0)$, so without loss of generality we focus just on $\mathbb{P}_x(M(2j) > 0)$.

We claim that for $j \geq 1$,

$$(40) \qquad 1 - \mathbb{P}_x(M(2j) > 0) \leq \frac{1}{2(1 + c' \log(j))},$$

for some constant $c' > 0$. By using this we have

$$\mathbb{P}_x(H_o > t) = \sum_{j=0}^{\infty} (1 - \mathbb{P}_x(M(2j) > 0))\mathbb{P}_x(N(t) \in \{2j, 2j+1\})$$

$$\leq \frac{1}{2} \sum_{j=0}^{\infty} \frac{\mathbb{P}_x(N(t) \in \{2j, 2j+1\})}{1 + c' \log(j+1)}$$

$$\leq \frac{1}{2}\mathbb{P}_x(N(t) \in \{0,1\}) + \frac{1}{2} \sum_{j=2}^{\infty} \frac{\mathbb{P}_x(N(t) = j)}{1 + c' \log((j+1)/2)}$$

$$\leq \frac{1}{2}\mathbb{P}_x(N(t) \in \{0,1\}) + \frac{1}{2} \sum_{j=2}^{\sqrt{t}} \mathbb{P}_x(N(t) = j) + \frac{1}{2} \sum_{j=\sqrt{t}+1}^{\infty} \frac{\mathbb{P}_x(N(t) = j)}{1 + c' \log(\sqrt{t}/2 + 1)}$$

$$\leq \frac{1}{2}\mathbb{P}_x(N(t) \leq \sqrt{t}) + \frac{1}{2(1 + c' \log(\sqrt{t}/2 + 1))} \leq \frac{C}{\log t}$$

for some $C > 0$, using in the last bound Chebyshev's inequality, for example.

To verify equation (40) we start by noting that

$$\mathbb{P}_x(M(2j) > 0) = \frac{\mathbb{E}_x(M(2j))}{\mathbb{E}_x(M(2j)|M(2j) > 0)}.$$

For the denominator, we can assume that Z_t hit vertex o for first time at time-step 1, obtaining the following upper bound $\mathbb{E}_x(M(2j) \mid M(2j) > 0) \leq \mathbb{E}_o(M(2j-1)) = \sum_{i=0}^{2j-1} P^i(o,o)$.

For the numerator, note that $\mathbb{E}_x(M(2j)) = \mathbb{E}_z(M(2j))$ for any $z \sim o$, by the symmetry of $\mathbb{Z}^2$ (recall that $x \sim o$), and so

$$\mathbb{E}_x(M(2j)) = \frac{1}{d} \sum_{z:z\sim o} \mathbb{E}_z(M(2j)) = \sum_{z:z\sim o} P(o,z)\mathbb{E}_z(M(2j)) = \sum_{i=1}^{2j+1} P^i(o,o).$$

Then

$$1 - \mathbb{P}_x(M(j) > 0) \leq 1 - \frac{\sum_{i=1}^{2j+1} P^i(o,o)}{\sum_{i=0}^{2j-1} P^i(o,o)} = \frac{1 - P^{2j}(o,o) - P^{2j+1}(o,o)}{\sum_{i=0}^{2j-1} P^i(o,o)}.$$

Note that $P^{2j}(o,o) + P^{2j+1}(o,o) \leq 1/2$, so

$$1 - \mathbb{P}_x(M(2j) > 0) \leq \frac{1}{2 \sum_{i=0}^{2j-2} P^i(o,o)} = \frac{1}{2 \sum_{i=0}^{j-1} P^{2i}(o,o)} \leq \frac{1}{2(1 + c' \log(j))},$$

where $c' > 0$ is a constant independent of i. The last equality follows since $P^{2i}(o,o) \geq \frac{c'}{i}$ for all $i \geq 1$ where $c' > 0$ is independent of the number of vertices and i. To see this, note that for a random walk on $\mathbb{Z}^2$, $P^{2i}(o,o) \geq \frac{1}{\pi i} - \frac{c}{i^2}$ by Theorem 35, where $c > 0$ is independent of i; so by choosing c' small enough we have $P^{2i}(o,o) \geq \frac{c'}{i}$, which gives

$$\sum_{i=0}^{j-1} P^{2i}(o,o) \geq 1 + \sum_{i=1}^{j-1} P^{2i}(o,o) \geq 1 + \sum_{i=1}^{j-1} c'/i \geq 1 + c' \log j. \qquad \square$$

5.2. *The Cycle C_n.* Our interest in this section is on the noisy voter model on the cycle[2] C_n with P_n corresponding to the simple random walk. Specifically, we shall prove Item 1 of Proposition 9. We first obtain the precise value of the variance, $\sigma^2 = \mathbb{Var}(S_n)$.

LEMMA 27 (Cycle variance). *For each $p \in (0,1)$ set $\theta = \frac{1 - \sqrt{p(2-p)}}{1-p} \in (0,1)$. Then*

$$\sigma^2 = \frac{1}{4n}\left(1 + \frac{2\theta(1 - \theta^{n-1})}{(1-\theta)(1+\theta^n)}\right).$$

PROOF. We appeal to martingale arguments. By Lemma 14, $\sigma^2 = \frac{1}{4n^2} \sum_{x,y \in V} \mathbb{P}(E_{xy})$ and recall that $E_{x,y} = \mathbb{P}(M^{Y(x),Y(y)} < \tau^{Y(x),Y(y)})$. Here we are considering the discrete-time dual process. For $t \in \mathbb{N}_0$, set X_t to be the clockwise distance between the two particles $Y(x)$ and $Y(y)$ initialised at x and y in the dual process just after the tth time that one of these two particles is selected to move/be absorbed (we are essentially looking at the dual when particles $Y(x)$ and $Y(y)$ move, ignoring the other particles in the process). Note that (X_t) behaves like a discrete-time simple random walk on the integers. Let $T_{0,n}$ be the hitting time of $\{0, n\}$ by (X_t) and G an independent Geometric random variable with parameter p. Then we deduce that $\sigma^2 = \frac{1}{4} \sum_{a,b \in \{0,\ldots,n-1\}} n^{-2} \mathbb{P}_{|a-b|}^{\text{RW}}(T_{0,n} < G)$, where $\mathbb{P}^{\text{RW}}$ is the probability

---

[2]C_n has vertex set $[n]$ and edge set $\{\{1,2\}, \{2,3\}, \ldots, \{n-1,n\}, \{n,1\}\}$

measure associated with (X_t) (a simple random walk on $\mathbb{Z}$). We now consider the bounded process $M_t := \theta^{X_t}(1-p)^t$. Since $\theta + \theta^{-1} = 2/(1-p)$, this is a martingale up to $T_{0,n}$. Thus by the optional stopping theorem, for each $k \in \{1, \ldots, n-1\}$,

$$\theta^k = \mathbb{E}_k^{\mathrm{RW}}\left[(1-p)^{T_{0,n}}\left(\mathbf{1}_{\{X_{T_{0,n}}=0\}} + \theta^n \mathbf{1}_{\{X_{T_{0,n}}=n\}}\right)\right].$$

By symmetry

$$\mathbb{E}_k^{\mathrm{RW}}\left[(1-p)^{T_{0,n}}\left(\mathbf{1}_{\{X_{T_{0,n}}=0\}} + \theta^n \mathbf{1}_{\{X_{T_{0,n}}=n\}}\right)\right]$$
$$= \mathbb{E}_{n-k}^{\mathrm{RW}}\left[(1-p)^{T_{0,n}}\left(\mathbf{1}_{\{X_{T_{0,n}}=n\}} + \theta^n \mathbf{1}_{\{X_{T_{0,n}}=0\}}\right)\right],$$

thus

$$\theta^k + \theta^{n-k}$$
$$= \mathbb{E}_k^{\mathrm{RW}}\left[(1-p)^{T_{0,n}}\left(\mathbf{1}_{\{X_{T_{0,n}}=0\}} + \theta^n \mathbf{1}_{\{X_{T_{0,n}}=n\}}\right)\right]$$
$$+ \mathbb{E}_{n-k}^{\mathrm{RW}}\left[(1-p)^{T_{0,n}}\left(\mathbf{1}_{\{X_{T_{0,n}}=0\}} + \theta^n \mathbf{1}_{\{X_{T_{0,n}}=n\}}\right)\right]$$
$$= \mathbb{E}_k^{\mathrm{RW}}\left[(1-p)^{T_{0,n}}(1 + \theta^n)\right],$$

i.e.

(41) $$\mathbb{E}_k^{\mathrm{RW}}\left[(1-p)^{T_{0,n}}\right] = (\theta^k + \theta^{n-k})(1 + \theta^n)^{-1}.$$

To relate this to σ^2, note that we can write

$$\sigma^2 = \frac{1}{4}\sum_{a,b\in\{0,\ldots,n-1\}} n^{-2}\mathbb{E}_{|a-b|}^{\mathrm{RW}}\left[(1-p)^{T_{0,n}}\right] = \frac{1}{4}n^{-2}\left(n\cdot 1 + n\sum_{k=1}^{n-1}\mathbb{E}_k^{\mathrm{RW}}\left[(1-p)^{T_{0,n}}\right]\right)$$

and so plugging in (41) and simplifying we obtain

$$\sigma^2 = \frac{1}{4n}\left(1 + \frac{2\theta(1-\theta^{n-1})}{(1-\theta)(1+\theta^n)}\right). \qquad \square$$

We introduce an auxiliary process which, up to a time change, is almost identical to the dual process on the cycle C_n. Given $n \in \mathbb{N}$ and $p \in (0,1)$, we define CAB (**c**oalescing walk with **ab**sorption) as a process $(\mathbf{X}_t)_{t\in\mathbb{N}_0} = (X_t, U_t, V_t, Y_t)_{t\in\mathbb{N}_0}$ taking values in $(C_n \cup \Delta)^4$ where Δ a cemetery state. The process starts at time 0 from a state in $(C_n)_4$ and we say that all 4 particles are alive. At each time $t \in \mathbb{N}$, we choose uniformly an alive particle and this particle with probability p gets absorbed into the cemetery state and dies, otherwise the particle jumps clockwise or anti-clockwise on the cycle with equal probability. If this jump results in particles meeting, these particles coalesce into a single alive particle (i.e. the number of alive particles decreases by 1). Any coalesced particles evolve in the same way at the same times, including having the same absorption times into the cemetery state.

The CAB process is similar to the dual process except time is sped-up, we only observe the movement of four particles, and the two stubborn vertices considered in the dual process are glued together into a single cemetery state Δ. To see this, note that if we consider the dual process, then when a particle $Y(z)$ coalesces with one of the other three particles, say $Y(x)$, then $Y(z)$ follows the trajectory of $Y(x)$ (this does not change the distribution of the dual process but allows us to ignore all particles different from $Y(x), Y(y), Y(u)$ and $Y(v)$). As a result probabilities of statements involving orders of meetings/absorption times are identical in both processes.

For a CAB process $(\mathbf{X}_t)_{t \in \mathbb{N}_0} = (X_t, U_t, V_t, Y_t)_{t \in \mathbb{N}_0}$, we use τ^A to denote the absorption time into Δ of particle $A \in \{X, U, V, Y\}$, $M^{A,B}$ to denote meeting time of particles $A, B \in \{X, U, V, Y\}$ and $\tau := \min\{\tau^X, \tau^U, \tau^V, \tau^Y, M^{X,U}, M^{V,Y}\}$. We also use $\mathbb{P}_{\mathbf{x}}^{CAB}$ to denote probability with respect to the CAB process when the four particle start from $\mathbf{x} \in C_n^4$. Additionally, for $a, b \in C_n$ we define $d(a, b)$ to be the clockwise distance from a to b and say that $(x_1, x_2, x_3, x_4) \in (C_n)_4$ is *clockwise-oriented* if and only if for $i \in \{1, 2\}$ and $j \notin \{i, i+1\}$, $d(x_i, x_{i+1}) < d(x_i, x_j)$.

We present a refinement of Proposition 18 to be used for the cycle, presented in terms of a CAB process. This provides us with the probabilities we need to control to bound the covariance term. Recall the function h defined in (31).

COROLLARY 28 (Cycle covariance). *Let* $\mathbf{x} = (x, u, v, y) \in (C_n)_4$ *be clockwise-oriented. Let* $(\mathbf{X}_t)_{t \geq 0} = (X_t, U_t, V_t, Y_t)_{t \geq 0}$ *be a CAB process initialised from* (x, u, v, y). *Then*

$$h_0(u, v, x, y)$$
$$\leq \mathbb{P}_{\mathbf{x}}^{CAB}(M^{X,Y} < \tau) + \mathbb{P}_{\mathbf{x}}^{CAB}(M^{U,V} < \tau^U = \min\{\tau, M^{X,Y}\}).$$

PROOF. Proposition 18 gives a bound on $h_0(u, v, x, y)$ in terms of the dual process, the probabilities in this bound being identical to probabilities for the CAB process, as already remarked. Further, due to the structure of the cycle, and the clockwise orientation of $\mathbf{x}$, certain probabilities are zero (those involving $T_1 = M^{x,v}$ or $T_1 = M^{u,y}$). The remaining probabilities can be easily bounded to give the claimed result. □

We control the probabilities in Corollary 28 using the following two propositions.

PROPOSITION 29. *Let* $\mathbf{x} = (x, u, v, y) \in (C_n)_4$ *be clockwise-oriented and distinct with* $x \sim u$, $v \sim y$ *and* $d := d(y, x) \geq d(u, v)$. *Let* $(\mathbf{X}_t)_{t \geq 0} = (X_t, U_t, V_t, Y_t)_{t \geq 0}$ *be a CAB process initialised from* (x, u, v, y). *Then there exists a universal constant* $C > 0$ *such that for any* $p \in (0, 1)$ *and* $n \geq 8$,

$$\mathbb{P}_{\mathbf{x}}^{CAB}(M^{X,Y} < \tau) \leq Cd^{-2}\left(\exp(-Cd\sqrt{p}) + d^{-1}\right).$$

We give the proof towards the end of this section.

PROPOSITION 30. *Let* $\mathbf{x} = (x, u, v, y) \in (C_n)_4$ *be clockwise-oriented and distinct with* $x \sim u$, $v \sim y$ *and* $d(y, x) \geq d(u, v)$. *Let* $(\mathbf{X}_t)_{t \geq 0} = (X_t, U_t, V_t, Y_t)_{t \geq 0}$ *be a CAB process initialised from* (x, u, v, y). *Then for any* $p \in (0, 1)$,

$$\mathbb{P}_{\mathbf{x}}^{CAB}(M^{U,V} < \tau^U = \min\{\tau, M^{X,Y}\}) \leq \mathbb{P}_{\mathbf{x}}^{CAB}(M^{U,V} < \tau) \wedge p.$$

Before proving this, we show how these propositions combined with Theorem 1 yield Item 1 of Proposition 9.

PROOF OF PROPOSITION 9 ITEM 1. The case $pn^2 \to 0$ follows immediately from Proposition 6 (recall that p depends on n), since the expected meeting time of two independent stationary walks is $\Theta(n^2)$.

Suppose now that $pn^2 \to \infty$. Recall that p is non-increasing in n, hence $\lim_n p$ exists. If $\lim_n p > 0$ then the result follows from Corollary 3. Next, by Theorem 1 and equation 32, it suffices to show that if $n^{-2} \ll p \ll 1$ then

$$\frac{1}{n^2 p \sigma^3} + \frac{1}{np\sigma^2}\sqrt{\sum_{x,y,u,v} \pi(x)\pi(y)P(x,u)P(y,v)h_0(u,v,x,y)} = o(1)$$

as $n \to \infty$. For the first term of the sum above, by Lemma 27, we have $\sigma^2 = \frac{1}{2\sqrt{2pn}}(1+o(1))$ and so $(n^2 p \sigma^3)^{-1} = o(1)$ as $n \to \infty$. The second term in the sum is more complicated. We begin by noticing that

$$(42) \qquad \sum_{x,y,u,v} \pi(x)\pi(y)P(x,u)P(y,v)h_0(u,v,x,y) = \mathbb{E}\left(h_0(U,V,X,Y)\right)$$

where $X, Y \sim \pi$, U is chosen according to $P(X,\cdot)$ and V according to $P(Y,\cdot)$ all independently. Now observe that from the definition of h,

$$h(u,v,x,y) = h(x,v,u,y) = h(u,y,x,v) = h(x,y,u,v),$$

and thus the same is true also for h_0. Therefore we have

$$\mathbb{E}[h_0(U,V,X,Y)] = \frac{1}{4n^2} \sum_{x,y \in C_n} \big[h_0(x_+, y_+, x, y) + h_0(x_+, y_-, x, y)$$

$$+ h_0(x_-, y_+, x, y) + h_0(x_-, y_-, x, y)\big]$$

$$= \frac{1}{n^2} \sum_{x,y \in C_n} h_0(x_+, y_-, x, y),$$

where x_+, x_- for $x \in C_n$ refers to the element of C_n neighbouring x clockwise, counterclockwise respectively. Notice also that $h_0(u,v,x,y) = h_0(y,x,v,u)$ and hence the above can be written as

$$\mathbb{E}[h_0(U,V,X,Y)] = \frac{1}{n^2} \sum_{x \in C_n} \sum_{\substack{y \in C_n: \\ 1 \le d(y,x) < \lfloor n/2 \rfloor}} h_0(y, x, y_-, x_+)$$

$$(43) \qquad\qquad + \frac{1}{n^2} \sum_{x \in C_n} \sum_{\substack{y \in C_n: \\ d\lfloor n/2 \rfloor \le d(y,x) \le n-3}} h_0(x_+, y_-, x, y)$$

$$+ \frac{1}{n^2} \sum_{x \in C_n} \sum_{\substack{y \in C_n: \\ n-2 \le d(y,x) \le n}} h_0(x_+, y_-, x, y).$$

If $d(y,x) = n-2$ then $x_+ = y_-$ and $h_0(x_+, y_-, x, y) = g_0(x, y, x_+)$ where g is as in Lemma 21 and $g_0 = g \vee 0$. To bound these terms we construct a discrete-time process $(M_t)_{t \in \mathbb{N}_0}$ defined in terms of the locations of the particles in the dual process as

$$(44) \qquad M_t = \left(1 - \zeta^{d(X_t,(X_+)_t)}\right)\left(1 - \zeta^{d((X_+)_t,Y_t)}\right)\mathbf{1}_{\{t < \tau_{\mathrm{abs}}\}} + \mathbf{1}_{\{t \ge \tau_{\mathrm{abs}}\}},$$

where $\tau_{\mathrm{abs}} = \min\{\tau^x, \tau^y, \tau^{x_+}\}$. This process is a bounded martingale until time

$$T := \min\{\tau_{\mathrm{abs}}, M^{x,y}, M^{x,x_+}, M^{y,x_+}\}$$

with $\zeta = \frac{2+p-\sqrt{3p(4-p)}}{2(1-p)} = 1 - \sqrt{3p} + O(p)$ as $p \to 0$ (details are given later for the similar martingale appearing in (47)). Then by the optional stopping theorem we have

$$(1-\zeta)^2 = \mathbb{E}\left[\mathbf{1}_{\{T=\tau_{\mathrm{abs}}\}} + \left(1 - \zeta^{d(X_T,(X_+)_T)}\right)\left(1 - \zeta^{d((X_+)_T,X_T)}\right)\mathbf{1}_{\{T=M^{x,y}\}}\right]$$

$$\ge \mathbb{E}\left[\mathbf{1}_{\{T=\tau_{\mathrm{abs}}\}} + \left(1 - \zeta^{d(X_T,(X_+)_T)} - \zeta^{d((X_+)_T,X_T)}\right)\mathbf{1}_{\{T=M^{x,y}\}}\right]$$

$$\geq \mathbb{E}\left[\mathbf{1}_{\{T=\tau_{\mathrm{abs}}\}} + (1-\zeta-\zeta^{n-1})\mathbf{1}_{\{T=M^{x,y}\}}\right]$$

$$= \mathbb{E}\left[\mathbf{1}_{\{T=\tau_{\mathrm{abs}}\}} + \sqrt{3p}(1+o(1))\mathbf{1}_{\{T=M^{x,y}\}}\right]$$

$$\geq \sqrt{3p}(1+o(1))\left(\mathbb{P}(T=\tau_{\mathrm{abs}}) + \mathbb{P}(T=M^{x,y})\right)$$

$$\geq \sqrt{p/3}(1+o(1))\left(\mathbb{P}(T=\tau^{x_+}) + \mathbb{P}(T=M^{x,y})\right)$$

$$\geq \sqrt{p/3}(1+o(1))g(x,y,x_+)$$

as $n \to \infty$. We deduce that there exists a universal $c>0$ such that for all n sufficiently large $g(x,y,x_+) \leq c\sqrt{p}$. It follows that also $g_0(x,y,x_+) \leq c\sqrt{p}$.

If $d(y,x)=n-1$ then $x=y_-$ and $y=x_+$ so $h(x_+,y_-,x,y) = \mathbb{P}(B(x) \neq B(x_+)) - \mathbb{P}(B(x) \neq B(x_+))^2 \leq \mathbb{P}(B(x) \neq B(x_+)) = \frac{1}{2}\left(1 - \mathbb{P}\left(M^{x,x_+} < \tau^x \wedge \tau^{x_+}\right)\right)$. We can now appeal to (41) to deduce that $\mathbb{P}\left(M^{x,x_+} < \tau^x \wedge \tau^{x_+}\right) = (\theta + \theta^{n-1})(1+\theta^n)^{-1}$ where $\theta = (1 - \sqrt{p(2-p)})(1-p)^{-1}$, i.e. $\mathbb{P}\left(M^{x,x_+} < \tau^x \wedge \tau^{x_+}\right) = 1 - \sqrt{2p} + O(p)$ as $p \to 0$, and so there exists a universal $c>0$ such that for all n sufficiently large $h(x_+,y_-,x,y) \leq c\sqrt{p}$. It follows that also $h_0(x_+,y_-,x,y) \leq c\sqrt{p}$.

If $d(y,x)=n$ then $x=y$ and $h_0(x_+,y_-,x,y)=g_0(x_+,y_-,x)$. By the same argument as the $d(y,x)=n-2$ case, we have $g_0(x_+,y_-,x) \leq c\sqrt{p}$ eventually.

If $1 \leq d(x,y) \leq n-3$ (which coincides with $\{x,y,x_+,y_-\}$ being distinct), we can apply Corollary 28 followed by Propositions 29 and 30 to each of these above terms. If further $d(y,x) < \lfloor n/2 \rfloor$ then $d(x_+,y_-) \geq d(y,x)$ and so we obtain

$$h_0(y,x,y_-,x_+)$$
$$\leq \mathbb{P}_{\mathbf{y}}^{\mathrm{CAB}}\left(M^{X_+,Y_-} < \tau\right) + \mathbb{P}_{\mathbf{y}}^{\mathrm{CAB}}\left(M^{X,Y} < \tau^X = \min\{\tau, M^{X_+,Y_-}\}\right)$$
$$\leq Cd(x_+,y_-)^{-2}\left(e^{-C\sqrt{p}d(x_+,y_-)} + d(x_+,y_-)^{-1}\right)$$
$$+ \left\{Cd(y,x)^{-2}\left(e^{-C\sqrt{p}d(y,x)} + d(y,x)^{-1}\right)\right\} \wedge p,$$

where $\mathbf{y} = (y_-,y,x,x_+)$ and $\tau = \min\{\tau^X, \tau^{X_+}, \tau^Y, \tau^{Y_-}, M^{X,X_+}, M^{Y,Y_-}\}$. Since $d(y,x) < \lfloor n/2 \rfloor$ we deduce that for $n \geq 4$, $d(x_+,y_-) \geq n/4$, hence there exists a universal constant $\hat{C} > 0$ such that for all n sufficiently large, we have the bound

(45)
$$h_0(y,x,y_-,x_+) \leq \frac{\hat{C}}{n^2}\left(e^{-\hat{C}\sqrt{p}n} + \frac{1}{n}\right) + \left\{Cd(y,x)^{-2}\left(e^{-C\sqrt{p}d(y,x)} + d(y,x)^{-1}\right)\right\} \wedge p.$$

If instead we have $d(y,x) \geq \lfloor n/2 \rfloor$ (with $\{x,y,x_+,y_-\}$ still distinct) then $d(y,x) \geq d(x_+,y_-)$ and we similarly obtain for all n sufficiently large

$$h_0(x_+,y_-,x,y)$$
$$\leq \mathbb{P}_{\mathbf{x}}^{\mathrm{CAB}}\left(M^{X,Y} < \tau\right) + \mathbb{P}_{\mathbf{x}}^{\mathrm{CAB}}\left(M^{X_+,Y_-} < \tau\right) \wedge \mathbb{P}_{\mathbf{x}}^{\mathrm{CAB}}\left(M^{X_+,Y_-} < \tau^{X_+} = \min\{\tau, M^{X,Y}\}\right)$$
$$\leq Cd(y,x)^{-2}\left(e^{-C\sqrt{p}d(y,x)} + d(y,x)^{-1}\right)$$
$$+ \left\{Cd(x_+,y_-)^{-2}\left(e^{-C\sqrt{p}d(x_+,y_-)} + d(x_+,y_-)^{-1}\right)\right\} \wedge p,$$

(46)
$$\leq \frac{\hat{C}}{n^2}\left(e^{-\hat{C}\sqrt{p}n} + \frac{1}{n}\right) + \left\{Cd(x_+,y_-)^{-2}\left(e^{-C\sqrt{p}d(x_+,y_-)} + d(x_+,y_-)^{-1}\right)\right\} \wedge p,$$

where $\mathbf{x} = (x, x_+, y, y_-)$ and $\tau = \min\{\tau^X, \tau^{X_+}, \tau^Y, \tau^{Y_-}, M^{X,X_+}, M^{Y,Y_-}\}$. Putting the bounds from (45) and (46) together with the bound of $c\sqrt{p}$ for the other cases into (43) we obtain that there exist constants $c_1, c_2, c_3, c_4 > 0$ such that for all n sufficiently large

$$\mathbb{E}[h_0(U, V, X, Y)]$$

$$\leq \frac{c_1}{n} \sum_{d=1}^{\lfloor n/2 \rfloor} \left[n^{-2} \left(e^{-c_1\sqrt{p}n} + n^{-1} \right) + \left(d^{-2} \left(e^{-c_1\sqrt{p}d} + d^{-1} \right) \right) \wedge p \right] + \frac{\sqrt{p}}{n}$$

$$\leq \frac{c_1}{n^2} + \frac{c_2}{n} \sum_{d=1}^{\lfloor (\log 1/p)/\sqrt{p} \rfloor} \left(d^{-2} e^{-c_1\sqrt{p}d} \right) \wedge p + \frac{c_2}{n} \sum_{d=\lfloor (\log 1/p)/\sqrt{p} \rfloor}^{\lfloor n/2 \rfloor} d^{-3} \wedge p + \frac{\sqrt{p}}{n}$$

$$\leq \frac{c_1}{n^2} + \frac{c_3}{n} \sum_{d=1}^{\lfloor 1/\sqrt{p} \rfloor} p + \frac{c_3}{n} \sum_{d=\lfloor 1/\sqrt{p} \rfloor}^{\lfloor (\log 1/p)/\sqrt{p} \rfloor} d^{-2} + \frac{c_2}{n} \sum_{d=\lfloor (\log 1/p)/\sqrt{p} \rfloor}^{\lfloor n/2 \rfloor} d^{-3} + \frac{\sqrt{p}}{n}$$

$$\leq \frac{c_1}{n^2} + \frac{c_3\sqrt{p}}{n} + \frac{c_4\sqrt{p}}{n} + \frac{c_4 p}{n} + \frac{\sqrt{p}}{n}.$$

Recalling that $\sigma^2 = \frac{1}{2\sqrt{2pn}}(1 + o(1))$ as $n \to \infty$, it follows that there exists a constant $c_5 > 0$ such that for all n sufficiently large,

$$\frac{1}{np\sigma^2} \sqrt{\mathbb{E}[h_0(U, V, X, Y)]} \leq \frac{c_5}{\sqrt{p}} \sqrt{n^{-2} + \sqrt{p}n^{-1}}$$

and the right-hand side tends to 0 as $n \to \infty$ by the assumption that $p \gg n^{-2}$. $\qquad \square$

PROOF OF PROPOSITION 30. The bound

$$\mathbb{P}_{\mathbf{x}}^{\mathrm{CAB}}\left(M^{U,V} < \tau^U = \min\{\tau, M^{X,Y}\}\right) \leq \mathbb{P}_{\mathbf{x}}^{\mathrm{CAB}}\left(M^{U,V} < \tau\right)$$

is immediate. To bound the probability on the left-hand side above by p, we shall use martingale optional stopping arguments. Before defining the appropriate martingales, we introduce a process on the line $\mathbb{Z}$. We define a process WAB (**w**alks with **ab**sorption) as a process $(\mathbf{X}_t^*)_{t \in \mathbb{N}_0} = (X_t^*, U_t^*, Y_t^*)_{t \in \mathbb{N}_0}$ taking values in $(\mathbb{Z} \cup \Delta)^3$ where Δ denotes the cemetery state. At each step, one of the three particles is chosen uniformly. The chosen particle jumps to the cemetery state with probability p, jumps down 1 with probability $(1-p)/2$ and otherwise jumps up 1. The process terminates at the first time two particles meet or any particle jumps to the cemetery state.

Now we construct a discrete-time process $(M_t)_{t \in \mathbb{N}_0}$ defined in terms of the locations of the particles in the WAB process. This process is almost identical to the one defined in (44) except the underlying process here is a WAB process on the line and the prior martingale is the dual process on the cycle. Specifically, for each $t \in \mathbb{N}_0$ we set

$$(47) \qquad M_t = \left(1 - \zeta^{U_t^* - X_t^*}\right)\left(1 - \zeta^{Y_t^* - U_t^*}\right) \mathbf{1}_{\{t < \tau_{\mathrm{abs}}\}} + \mathbf{1}_{\{t \geq \tau_{\mathrm{abs}}\}},$$

for $\zeta = \zeta(p) \in (0, 1)$ to be determined and where $\tau_{\mathrm{abs}} = \min\{\tau^{X^*}, \tau^{U^*}, \tau^{Y^*}\}$ (with this notation again referring to absorption times). The choice of ζ will ensure (M_t) is a martingale. Let $\mathcal{F}_t$ be the filtration generated by the trajectories of all particles in WAB up to time t. Then for each $t \in \mathbb{N}_0$,

$$\mathbb{E}[M_{t+1} \mid \mathcal{F}_t]$$

$$= \mathbb{E}\left[\left(1 - \zeta^{U_{t+1}^* - X_{t+1}^*}\right)\left(1 - \zeta^{Y_{t+1}^* - U_{t+1}^*}\right) \mathbf{1}_{\{t+1 < \tau_{\mathrm{abs}}\}} + \mathbf{1}_{\{t+1 \geq \tau_{\mathrm{abs}}\}} \mid \mathcal{F}_t\right]$$

$$= \mathbf{1}_{\{t \geq \tau_{\text{abs}}\}} + p\mathbf{1}_{\{t < \tau_{\text{abs}}\}}$$

$$+ \frac{1-p}{6}\mathbf{1}_{\{t < \tau_{\text{abs}}\}}\left\{ \left(1 - \zeta^{U_t^* - X_t^* - 1}\right)\left(1 - \zeta^{Y_t^* - U_t^*}\right) + \left(1 - \zeta^{U_t^* - X_t^* + 1}\right)\left(1 - \zeta^{Y_t^* - U_t^*}\right) \right\}$$

$$+ \frac{1-p}{6}\mathbf{1}_{\{t < \tau_{\text{abs}}\}}\left\{ \left(1 - \zeta^{U_t^* - X_t^* + 1}\right)\left(1 - \zeta^{Y_t^* - U_t^* - 1}\right) \right.$$

$$\left. + \left(1 - \zeta^{U_t^* - X_t^* - 1}\right)\left(1 - \zeta^{Y_t^* - U_t^* + 1}\right) \right\}$$

$$+ \frac{1-p}{6}\mathbf{1}_{\{t < \tau_{\text{abs}}\}}\left\{ \left(1 - \zeta^{U_t^* - X_t^*}\right)\left(1 - \zeta^{Y_t^* - U_t^* + 1}\right) + \left(1 - \zeta^{U_t^* - X_t^*}\right)\left(1 - \zeta^{Y_t^* - U_t^* - 1}\right) \right\}$$

$$= \mathbf{1}_{\{t \geq \tau_{\text{abs}}\}} + \mathbf{1}_{\{t < \tau_{\text{abs}}\}}\left[1 + \frac{1-p}{3}\left(1 + \zeta + 1/\zeta\right)\left(\zeta^{Y_t^* - X_t^*} - \zeta^{U_t^* - X_t^*} - \zeta^{Y_t^* - U_t^*}\right) \right].$$

Thus for (M_t) to be a martingale it suffices that $1 + \zeta + 1/\zeta = 3/(1-p)$, i.e. we take

(48)
$$\zeta = \frac{2 + p - \sqrt{3p(4-p)}}{2(1-p)} \geq 1 - \sqrt{3p}.$$

Let $T^* := \tau_{\text{abs}} \wedge M^{X^*, U^*} \wedge M^{Y^*, U^*}$ and observe that $|M_t| \leq 1$ for all $t \leq T^*$. Thus by the optional stopping theorem,

(49)
$$\mathbb{E}[(1 - \zeta^{U_0^* - X_0^*})(1 - \zeta^{Y_0^* - U_0^*})] = \mathbb{E}[M_0] = \mathbb{E}[M_{T^*}] = \mathbb{P}^{\text{WAB}}(\tau_{\text{abs}} < M^{X^*, U^*} \wedge M^{Y^*, U^*}).$$

We turn to bounding the probability of interest. It is convenient to couple the CAB process to a process denoted CAB$'$ which is initialised from the same state and which evolves in the same way except that there can be no absorptions until the first meeting time (one way to construct is to take the CAB process and set $p = 0$ until the first meeting occurs). The coupling is successful (i.e. the two processes agree for all times) for CAB processes which satisfy that the first absorption is after the first meeting. Thus

$$\mathbb{P}_{\mathbf{x}}^{\text{CAB}}\left(M^{U,V} < \tau^U = \min\{\tau, M^{X,Y}\}\right)$$

$$\leq \mathbb{P}_{\mathbf{x}}^{\text{CAB}'}\left(M^{U,V} < \tau^U = \min\{\tau, M^{X,Y}\}\right)$$

$$= \mathbb{P}_{\mathbf{x}}^{\text{CAB}'}\left(M^{U,V} < \min\{M^{X,U}, M^{X,Y}, M^{U,Y}\}, \tau^U = \min\{\tau, M^{X,Y}\}\right).$$

Now let $\mathcal{F}_t^{\text{CAB}'}$ be the sigma algebra generated by the CAB$'$ process up to time t and $R := \min\{\tau^X, \tau^U, \tau^Y\}$. Then

$$\mathbb{P}_{\mathbf{x}}^{\text{CAB}}\left(M^{U,V} < \tau^U = \min\{\tau, M^{X,Y}\}\right)$$

$$\leq \mathbb{E}_{\mathbf{x}}^{\text{CAB}'}\left[\mathbb{E}\left[\mathbf{1}_{\{M^{U,V} < \min\{M^{X,U}, M^{X,Y}, M^{V,Y}\}\}}\mathbf{1}_{\{\tau^U = \min\{\tau, M^{X,Y}\}\}} \mid \mathcal{F}_{M^{U,V}}^{\text{CAB}'}, M^{U,V} \right] \right]$$

$$\leq \frac{1}{3}\mathbb{E}_{\mathbf{x}}^{\text{CAB}'}\left[\mathbf{1}_{\{M^{U,V} < \min\{M^{X,U}, M^{X,Y}, M^{V,Y}\}\}} \right.$$

(50)
$$\left. \cdot \mathbb{E}\left[\mathbf{1}_{\{R < \min\{M^{X,U}, M^{X,Y}, M^{V,Y}\}\}} \mid \mathcal{F}_{M^{U,V}}^{\text{CAB}'}, M^{U,V} \right] \right].$$

The random variable

$$\mathbf{1}_{\{M^{U,V} < \min\{M^{X,U}, M^{X,Y}, M^{V,Y}\}\}}\mathbb{E}\left[\mathbf{1}_{\{R < \min\{M^{X,U}, M^{X,Y}, M^{V,Y}\}\}} \mid \mathcal{F}_{M^{U,V}}^{\text{CAB}'}, M^{U,V} \right]$$

is clearly only non-zero for trajectories satisfying $M^{U,V} < \min\{M^{X,U}, M^{X,Y}, M^{V,Y}\}$. For these, we have by the memoryless property (R is geometrically distributed) and the strong Markov property at time $S := M^{U,V}$, almost surely

$$\mathbb{E}\left[\mathbf{1}_{\{R < \min\{M^{X,U}, M^{X,Y}, M^{V,Y}\}\}} \mid \mathcal{F}_S^{\text{CAB}'}, S\right]$$
$$\leq \mathbb{P}_{(X_S, U_S, V_S, Y_S)}^{\text{CAB}'}(R < \min\{M^{X,U}, M^{X,Y}, M^{V,Y}\}).$$

The CAB$'$ process in the last probability is initialised from the random state (X_S, U_S, V_S, Y_S), and as such U and V will remain together for all later times; thus we can write the preceding probability as (we change $M^{V,Y}$ to $M^{U,Y}$)

$$\mathbb{P}_{(X_S, U_S, V_S, Y_S)}^{\text{CAB}}(R < \min\{M^{X,U}, M^{X,Y}, M^{U,Y}\}),$$

where (X_S, U_S, V_S, Y_S) is the time S state of a CAB$'$ process (from time S onwards the process behaves as a CAB process hence the above probability measure is $\mathbb{P}^{\text{CAB}}$).

This CAB can be coupled with a WAB process started from $(-d(X_S, U_S), 0, d(U_S, Y_S))$ such that a jump clockwise/anti-clockwise in CAB corresponds to a jump up/down in WAB, and jumps to cemetery states coincide in the two processes.

This coupling gives that, almost surely

$$\mathbb{P}_{(X_S, U_S, V_S, Y_S)}^{\text{CAB}}\left(R < \min\{M^{X,U}, M^{X,Y}, M^{U,Y}\}\right)$$
$$\leq \mathbb{P}_{(-d(X_S, U_S), 0, d(U_S, Y_S))}^{\text{WAB}}\left(\tau_{\text{abs}} < \min\{M^{X^*, U^*}, M^{Y^*, U^*}\}\right),$$

and combining this with (49) we obtain

$$\mathbb{P}_{(X_S, U_S, V_S, Y_S)}^{\text{CAB}}\left(R < \min\{M^{X,U}, M^{X,Y}, M^{U,Y}\}\right)$$
$$\leq \mathbb{E}^{\text{WAB}}\left[(1 - \zeta^{d(X_S, U_S)})(1 - \zeta^{d(U_S, Y_S)}) \mid X_S, U_S, Y_S\right]$$
$$\leq 3p\, d(X_S, U_S)\, d(U_S, Y_S)$$

almost surely, where we have used the bound (48) in the final inequality.

Plugging this into (50) we obtain

$$\mathbb{P}_{\mathbf{x}}^{\text{CAB}}\left(M^{U,V} < \tau^U = \min\{\tau, M^{X,Y}\}\right)$$

(51) $$\leq p\, \mathbb{E}_{\mathbf{x}}^{\text{CAB}'}\left[\mathbf{1}_{\{M^{U,V} < \min\{M^{X,U}, M^{X,Y}, M^{V,Y}\}\}} d(X_{M^{U,V}}, U_{M^{U,V}})\, d(U_{M^{U,V}}, Y_{M^{U,V}})\right].$$

Now the reason for introducing the process CAB$'$ becomes apparent: we know that, on event $\{M^{U,V} < \min\{M^{X,U}, M^{X,Y}, M^{V,Y}\}\}$, until time $M^{U,V}$ there are no absorptions (by definition of this process). Also on this event we have (by the nature of the cycle) that for all $t \leq M^{U,V}$, $d(X_{M^{U,V}}, U_{M^{U,V}})\, d(U_{M^{U,V}}, Y_{M^{U,V}}) < n^2$. Thus we can bound the above expectation by coupling to a process of simple random walks on the line, RW(4), defined as follows: RW(4) takes values in $\mathbb{Z}^4$ and at each step a uniformly chosen particle jumps up or down equally likely. Let (A_t, B_t, C_t, D_t) be the time-t state of RW(4). The coupling is again simple: clockwise/anti-clockwise movements of particles correspond to jumps up/down. We obtain the inequality

$$\mathbb{E}_{\mathbf{x}}^{\text{CAB}'}\left[\mathbf{1}_{\{M^{U,V} < \min\{M^{X,U}, M^{X,Y}, M^{V,Y}\}\}} d(X_{M^{U,V}}, U_{M^{U,V}})\, d(U_{M^{U,V}}, Y_{M^{U,V}})\right]$$
$$\leq \mathbb{E}_{(-d(x,u), 0, d(u,v), d(u,y))}^{\text{RW}(4)}\left[\mathbf{1}_{\{M^{B,C} < M^{A,B} \wedge M^{C,D}\}}\right.$$
$$\left. \cdot \left\{(B_{M^{B,C}} - A_{M^{B,C}})(D_{M^{B,C}} - B_{M^{B,C}}) \wedge n^2\right\}\right]$$

$$= \mathbb{E}^{\mathrm{RW}(4)}_{(-d(x,u),0,d(u,v),d(u,y))}\Big[\mathbf{1}_{\{M^{B,C}<M^{A,B}\wedge M^{C,D}\}}$$
$$\cdot\big\{(B_{M^{B,C}}-A_{M^{B,C}})(D_{M^{B,C}}-C_{M^{B,C}})\wedge n^2\big\}\Big]$$

$$= \mathbb{E}^{\mathrm{RW}(4)}_{(-d(x,u),0,d(u,v),d(u,y))}[N_M],$$

where the process, $(N_t)_{t\in\mathbb{N}_0}$, defined as

$$N_t = (B_t-A_t)(D_t-C_t)\wedge n^2$$

is a bounded submartingale until the first meeting time M of any two particles in $\mathrm{RW}(4)$. Thus by the optional stopping theorem,

$$1 = d(x,u)\,(d(u,y)-d(u,v))$$
$$= \mathbb{E}^{\mathrm{RW}(4)}_{(-d(x,u),0,d(u,v),d(u,y))}[N_0] \geq \mathbb{E}^{\mathrm{RW}(4)}_{(-d(x,u),0,d(u,v),d(u,y))}[N_M].$$

Combining with (51) we obtain

$$\mathbb{P}^{\mathrm{CAB}}_{\mathbf{x}}\big(M^{U,V}<\tau^U=\min\{\tau,M^{X,Y}\}\big)\leq p,$$

which completes the proof. $\hfill\square$

We now turn to proving Proposition 29. The first step is a preliminary lemma on the decay rate of an expectation involving the hitting time of zero by a discrete-time random walk on $\mathbb{Z}$. To this end, we write $\mathbb{E}^{\mathrm{RW}}$ and $\mathbb{P}^{\mathrm{RW}}$ for the expectation and probability measure associated with such a random walk and τ_0^X for the hitting time of zero by process X.

LEMMA 31. *There exists a constant $c>0$ such that for each $p\in[0,1)$,*

$$\mathbb{E}^{\mathrm{RW}}_{\lfloor d/8\rfloor}[(\tau_0^X)^{-1/2}(1-p)^{\tau_0^X/3}] \leq \begin{cases} cd^{-1} & \text{if } p=0, \\ c\big(d^{-1}e^{-cd\sqrt{p}}+d^{-2}\big) & \text{if } p\in(0,1). \end{cases}$$

PROOF. We first decompose over the hitting time of 0:

$$\mathbb{E}^{\mathrm{RW}}_{\lfloor d/8\rfloor}[(\tau_0^X)^{-1/2}(1-p)^{\tau_0^X/3}] = \sum_{i=\lfloor d/8\rfloor}^{\infty} i^{-1/2}(1-p)^{i/3}\mathbb{P}^{\mathrm{RW}}_{\lfloor d/8\rfloor}[\tau_0^X=i].$$

Next, by the hitting time theorem (Theorem 36) for each $i\geq\lfloor d/8\rfloor$,

$$\mathbb{P}^{\mathrm{RW}}_{\lfloor d/8\rfloor}(\tau_0^X=i) = \frac{1}{i}\left\lfloor\frac{d}{8}\right\rfloor\mathbb{P}^{\mathrm{RW}}_{\lfloor d/8\rfloor}(X_i=0).$$

We can now apply the local central limit theorem which gives the existence of constants $c_1,c_2>0$ such that, uniformly in $i\geq\lfloor d/8\rfloor$,

$$\mathbb{P}^{\mathrm{RW}}_{\lfloor d/8\rfloor}(X_i=0) \leq c_1 i^{-1/2}\left(\exp(-c_2 d^2/i)+d^{-2}\right).$$

Using this gives the existence of constants $c_2,c_3>0$ such that

$$\sum_{i=\lfloor d/8\rfloor}^{\infty} i^{-1/2}(1-p)^{i/3}\mathbb{P}^{\mathrm{RW}}_{\lfloor d/8\rfloor}[\tau_0^X=i] \leq c_3 d\sum_{i=\lfloor d/8\rfloor}^{\infty} i^{-2}(1-p)^{i/3}\left(\exp(-c_2 d^2/i)+d^{-2}\right)$$

We treat separately the case $p = 0$. In this case we have the bound

$$\mathbb{E}^{\mathrm{RW}}_{\lfloor d/8 \rfloor}[(\tau_0^X)^{-1/2}] \leq c_3 d \sum_{i=\lfloor d/8 \rfloor}^{\infty} i^{-2} \left(\exp(-c_2 d^2/i) + d^{-2} \right) \leq c_4 d^{-1},$$

for some $c_4 > 0$. For $p \in (0,1)$, we have the bound

$$\mathbb{E}^{\mathrm{RW}}_{\lfloor d/8 \rfloor}[(\tau_0^X)^{-1/2}(1-p)^{\tau_0^X/3}]$$

$$\leq c_3 d \sum_{i=\lfloor d/8 \rfloor}^{\infty} i^{-2} e^{-ip/3} \exp(-c_2 d^2/i) + c_5 d^{-2}$$

$$\leq c_3 d \sum_{i=\lfloor d/8 \rfloor}^{d/\sqrt{p}} i^{-2} \exp(-c_2 d^2/i) + c_3 d \sum_{i=d/\sqrt{p}}^{\infty} i^{-2} e^{-ip/3} + c_5 d^{-2}$$

$$\leq c_6 \left(d^{-1} e^{-c_6 d \sqrt{p}} + d^{-2} \right),$$

for some $c_5, c_6 > 0$. $\qquad\square$

Given $n \in \mathbb{N}$ and $p \in (0,1)$ (possibly depending on n), we define two new processes, denoted WAG and WAIG. Process WAG (**w**alks with **a**bsorption and **g**host) is a process $(\mathbf{X}_t)_{t \in \mathbb{N}_0} = (X_t, U_t, V_t, Y_t)_{t \in \mathbb{N}_0}$ taking values in $(\mathbb{Z} \cup \Delta)^4$ where Δ denotes the cemetery state. This process starts at time 0 from a state $\mathbf{x} = (x, u, v, y)$ with $u \leq x \leq 0 \leq y \leq v$ and we say that all 4 particles are *alive* unless $v - u = n$ in which case all particles except V are alive, with V being a *ghost*. At each time $t \in \mathbb{N}$ we choose uniformly from the alive particles and this chosen particle with probability p gets absorbed into the cemetery state Δ and dies (is no longer alive), otherwise the particle chooses to jump up by one or down by one with equal probability. If this jump results in $V - U = n$, we say that V becomes a ghost (and is no longer alive). If V was already a ghost and the particle chosen is U, then V evolves in the same way as U, i.e. if U is absorbed into Δ then so is V, if U jumps up/down then V jumps up/down. The process terminates (no further evolution occurs) if U and X meet, or Y and V meet, either of X and Y hits 0, or any particle is absorbed into the cemetery state.

Process WAIG (**w**alks with **a**bsorption and **i**ndependent **g**host) is almost identical to WAG. The only difference is that if V is a ghost and U is chosen and not absorbed into Δ, then V jumps up/down independently of U, rather than copying its move.

In each of these systems we use notation τ^A to denote the absorption time into Δ of particle $A \in \{X, U, V, Y\}$, τ_0^A the time for particle A to hit 0, $M^{A,B}$ to denote meeting time of particles $A, B \in \{X, U, V, Y\}$, and G^V to denote the time that V becomes a ghost (any of these times is infinite if the process terminates prior to their occurrence). The following lemma relates the probability that X or Y hits 0 before time $\tau := M^{X,U} \wedge M^{V,Y} \wedge \tau^Y$ in these two processes.

LEMMA 32. *For each* $\mathbf{x} = (x, u, v, y) \in \mathbb{Z}^4$ *with* $u \leq x \leq y \leq v$,

$$\mathbb{P}^{\mathrm{WAG}}_{\mathbf{x}}(\tau_0^X \wedge \tau_0^Y < M^{X,U} \wedge M^{V,Y} \wedge \tau^Y) \leq \mathbb{P}^{\mathrm{WAIG}}_{\mathbf{x}}(\tau_0^X \wedge \tau_0^Y < M^{X,U} \wedge M^{V,Y} \wedge \tau^Y).$$

PROOF. We first note that $(\mathbb{P}^{\mathrm{WAG}}_{\mathbf{X}_t}(\tau_0^X \wedge \tau_0^Y < \tau))_{t \in \mathbb{N}_0}$ is a bounded martingale until time $T := \tau_0^X \wedge \tau_0^Y \wedge \tau$ and so by the optional stopping theorem at time $S := T \wedge G^V$,

$$\mathbb{P}^{\mathrm{WAG}}_{\mathbf{x}}(\tau_0^X \wedge \tau_0^Y < \tau) = \mathbb{E}^{\mathrm{WAG}}[\mathbb{P}^{\mathrm{WAG}}_{\mathbf{X}_S}(\tau_0^X \wedge \tau_0^Y < \tau)]$$

$$(52) \qquad\qquad = \mathbb{E}^{\mathrm{WAG}}\left[\mathbf{1}_{\{\tau_0^X \wedge \tau_0^Y < \tau \wedge G^V\}} + \mathbf{1}_{\{G^V < T\}} \mathbb{P}^{\mathrm{WAG}}_{\mathbf{X}_{G^V}}(\tau_0^X \wedge \tau_0^Y < \tau) \right].$$

By the same reasoning we also have

(53)
$$\mathbb{P}_{\mathbf{x}}^{\text{WAIG}}(\tau_0^X \wedge \tau_0^Y < \tau) = \mathbb{E}^{\text{WAIG}}\left[\mathbf{1}_{\{\tau_0^X \wedge \tau_0^Y < \tau \wedge G^V\}} + \mathbf{1}_{\{G^V < T\}}\mathbb{P}_{\mathbf{X}_{G^V}}^{\text{WAIG}}(\tau_0^X \wedge \tau_0^Y < \tau)\right].$$

Notice that at time G^V, we know that $V - U = n$. So consider a state $\mathbf{a} = (a, b, n + b, c)$. We will show that

(54)
$$\mathbb{P}_{\mathbf{a}}^{\text{WAG}}(\tau_0^X \wedge \tau_0^Y < \tau) \leq \mathbb{P}_{\mathbf{a}}^{\text{WAIG}}(\tau_0^X \wedge \tau_0^Y < \tau).$$

Once we have this inequality we can complete the proof by taking a coupling of WAG and WAIG: we can couple them so that the absorption times, G^V and the trajectories of all particles are the same in both processes up to time G^V. It follows that under this coupling if each of the events $\{\tau_0^X \wedge \tau_0^Y < \tau \wedge G^V\}$ and $\{G^V < T\}$ hold in WAG, they hold in WAIG and vice-versa. Hence the indicators in equations (52) and (53) are equal.

It remains to show (54). Note that starting WAG from $\mathbf{a}$ means that V is immediately a ghost. To show (54) we will make use of the FKG inequality for partially ordered sets. In process WAG, we denote by $(Q_i)_{i \in \mathbb{N}}$ a sequence of Rademacher random variables (± 1 with probability 1/2 each), which will be used (in the obvious way) to generate the movement of U (and hence, since V starts as a ghost, of V). To highlight the dependence of U and V on $\mathbf{Q}$ we will sometimes write these processes as $U(\mathbf{Q})$ and $V(\mathbf{Q})$ (note that if U has jumped k times by time t then $U_t(\mathbf{Q}) = u + \sum_{i=1}^{k} Q_i$).

Let J^U be the sequence of times at which particle U is chosen to evolve (jump or be absorbed) and $W = ((X_t, Y_t)_{t \in \mathbb{N}_0}, \tau^U, \tau^V, J^U)$ and note that τ^X and τ^Y are measurable with respect to the sigma algebra generated by the trajectories of X and Y. For $\mathbf{z} \in \{-1, +1\}^{\mathbb{N}}$, consider the functions

$$f(\mathbf{z}) := \mathbf{1}_{\{\tau_0^X \wedge \tau_0^Y < M^{X,U(\mathbf{z})} \wedge \tau^X \wedge \tau^{U(\mathbf{z})}\}}$$

$$g(\mathbf{z}) := \mathbf{1}_{\{\tau_0^X \wedge \tau_0^Y < M^{V(\mathbf{z}),Y} \wedge \tau^{V(\mathbf{z})} \wedge \tau^Y\}}.$$

For each $m \in \mathbb{N}$, we define a partial order on the sample space $\Omega_m = \{-1, +1\}^m$ of $(Q_i)_{1 \leq i \leq m}$. For $\omega_1, \omega_2 \in \Omega_m$, we say that $\omega_1 \leq \omega_2$ iff ω_2 can be obtained from ω_1 by flipping some number of -1s in ω_1 to $+1$s. Writing $\mathbf{1}$ for an infinite vector of 1s and $\mathbf{Q} \in \Omega_m$, we observe that $f((\mathbf{Q}, \mathbf{1}))$ is decreasing with respect to this partial order. Conversely, g_m is increasing. Let $\hat{\mu}_m$ be the measure on $\{-1, +1\}^{\mathbb{N}}$ which assigns non-zero and equal mass to vectors $\mathbf{z}$ with $\mathbf{z}(i) = 1$ for all $i > m$, and let μ_m be $\hat{\mu}_m$ conditioned on W. By the FKG inequality we obtain that for each $m \in \mathbb{N}$,

$$\mu_m(fg) \leq \mu_m(f)\mu_m(g).$$

Sending $m \to \infty$[3] we deduce that

$\mathbb{P}_{\mathbf{a}}^{\text{WAG}}(\tau_0^X \wedge \tau_0^Y < \tau \mid W)$

$\leq \mathbb{P}_{\mathbf{a}}^{\text{WAG}}(\tau_0^X \wedge \tau_0^Y < M^{X,U} \wedge \tau^X \wedge \tau^U \mid W)\,\mathbb{P}_{\mathbf{a}}^{\text{WAG}}(\tau_0^X \wedge \tau_0^Y < M^{V,Y} \wedge \tau^V \wedge \tau^Y \mid W)$

$= \mathbb{P}_{\mathbf{a}}^{\text{WAIG}}(\tau_0^X \wedge \tau_0^Y < M^{X,U} \wedge \tau^X \wedge \tau^U \mid W)\,\mathbb{P}_{\mathbf{a}}^{\text{WAIG}}(\tau_0^X \wedge \tau_0^Y < M^{V,Y} \wedge \tau^V \wedge \tau^Y \mid W)$

$= \mathbb{P}_{\mathbf{a}}^{\text{WAIG}}(\tau_0^X \wedge \tau_0^Y < \tau \mid W),$

where the second equality comes from the fact that given W, U and V move independently under WAIG. Taking an expectation gives (54) and completes the proof. $\qquad\square$

---

[3] the limits exist since $\mu_m(fg)$, $\mu_m(f)$ and $\mu_m(g)$ remain unchanged for all $m \geq \tau_0^X \wedge \tau_0^Y$

Now we bound the probability that, under WAIG, at least one of particles X and Y hits 0 before time τ.

LEMMA 33.   *Suppose $d \geq 4$, u, x, y, v are integers satisfying $u \leq x \leq -\frac{d}{2} + 1 \leq \frac{d}{2} - \frac{1}{2} \leq y \leq v$ and set $\mathbf{x} = (x, u, v, y)$. Then there exists a constant $c > 0$ (not depending on d) such that*

$$\mathbb{P}_{\mathbf{x}}^{\text{WAIG}}(\tau_0^X \wedge \tau_0^Y < \tau) \leq cd^{-2}\left(e^{-cd\sqrt{p}} + d^{-1}\right).$$

PROOF.   We introduce a process denoted WMIG (**w**alks with **m**arking and **i**ndependent **g**host). This process is similar to WAIG except no particles are absorbed (when a particle is chosen it always jumps) and the process does not terminate when particles meet (and the dynamics of particles does not change when any two meet) or when X or Y hit 0. In addition, each time particle Y is chosen, with probability p it is marked (if not already) – this does not affect its movement at this time or afterwards. We denote by τ_m^Y the time at which Y is marked.

First we note that we can couple a WAIG process with a WMIG process if they start from the same configuration up until the time that the WAIG process terminates so that until this time particles move in the same way in the two processes. Since prior to time $\tau_0^X \wedge \tau_0^Y \wedge \tau$ the WAIG process does not terminate, we deduce that

$$\mathbb{P}_{\mathbf{x}}^{\text{WAIG}}(\tau_0^X \wedge \tau_0^Y < \tau)$$
$$\leq \mathbb{P}_{\mathbf{x}}^{\text{WAIG}}(\tau_0^X \wedge \tau_0^Y < M^{X,U} \wedge M^{V,Y} \wedge \tau^Y)$$
$$= \mathbb{P}_{\mathbf{x}}^{\text{WMIG}}(\tau_0^X \wedge \tau_0^Y < M^{X,U} \wedge M^{V,Y} \wedge \tau_m^Y)$$

$$(55) \qquad \leq \mathbb{P}_{\mathbf{x}}^{\text{WMIG}}(\tau_0^X < M^{X,U} \wedge M^{V,Y} \wedge \tau_m^Y) + \mathbb{P}_{\mathbf{x}}^{\text{WMIG}}(\tau_0^Y < M^{X,U} \wedge M^{V,Y} \wedge \tau_m^Y).$$

We now consider bounding $\mathbb{P}_{\mathbf{x}}^{\text{WMIG}}(\tau_0^X < M^{X,U} \wedge M^{V,Y} \wedge \tau^Y)$ (the other term can be bounded in the same way).

Given that $X_0 \leq -\frac{d}{2} - 1$, in order for X to hit 0 it must hit $-\lfloor d/4 \rfloor$, a time we denote $\tau_{-\lfloor d/4 \rfloor}^X$. We set $T^X := \tau_0^X - \tau_{-\lfloor d/4 \rfloor}^X$ be the time it takes for X to hit 0 from $-\lfloor d/4 \rfloor$. By the strong Markov property $\tau_{-\lfloor d/4 \rfloor}^X$ and T^X are independent. We have

$$(56)$$
$$\mathbb{P}_{\mathbf{x}}^{\text{WMIG}}(\tau_0^X < M^{X,U} \wedge M^{V,Y} \wedge \tau^Y) \leq \mathbb{P}_{\mathbf{x}}^{\text{WMIG}}(\tau_{-\lfloor d/4 \rfloor}^X < M^{X,U}, T^X < M^{V,Y} \wedge \tau^Y)$$

The two events appearing in the probability on the right-hand side above are in fact independent (which is the reason for introducing the WMIG process). We make this clear with the following argument, which also gives us a way to bound this probability.

Let $C^{X,U}$ denote the number of times during time interval $[0, \tau_{-\lfloor d/4 \rfloor}^X)$ that process WMIG chooses either X or U to evolve. Let $C^{V,Y}$ denote the number of times during time interval $[\tau_{-\lfloor d/4 \rfloor}^X, \tau_0^X)$ that V or Y jump under WMIG (if V becomes a ghost it can no longer be chosen but still jumps if U is chosen). Note that these two random variables are independent (they correspond to disjoint time intervals). Given $\tau_{-\lfloor d/4 \rfloor}^X$, we have $C^{X,U}$ stochastically dominates a $\text{Bin}(\tau_{-\lfloor d/4 \rfloor}^X, 1/2)$ random variable (the stochastic domination comes from the fact that once V becomes a ghost, we choose either X or U at each step with probability $2/3$). Similarly, given T^X, we have $C^{V,Y}$ stochastically dominates a $\text{Bin}(T^X, 1/2)$ random variable (once V becomes a ghost, we choose either Y or U (choosing U means V also moves) at each step with probability $2/3$).

For each $t \in \mathbb{N}_0$, for process WMIG set $d_t^1 = X_t - U_t$ and $d_t^2 = V_t - Y_t$ and for each $i \in \{1, 2\}$ let $(D_t^i)_{t \in \mathbb{N}_0}$ be the process $(d_t^i)_{t \in \mathbb{N}_0}$ observed only when d_t^i updates (i.e. the value changes). It is immediate that processes $(D_t^1)_{t \in \mathbb{N}_0}$ and $(D_t^2)_{t \in \mathbb{N}_0}$ are independent simple random walks on $\mathbb{Z}$ started from state 1. For each $i \in \{1, 2\}$, let $\tau_0^{D^i} = \min\{t \geq 0 : D_t^i = 0\}$ and let τ^D be the number of updates of d_t^2 until Y is marked under WMIG so that $\tau^D \sim \text{Geom}(1/2)$.

Then we have that under WMIG,

(57)
$$\{\tau_{-\lfloor d/4 \rfloor}^X < M^{X,U}\} = \{C^{X,U} < \tau_0^{D^1}\}, \qquad \{T^X < M^{V,Y} \wedge \tau^Y\} = \{C^{V,Y} < \tau_0^{D^2} \wedge \tau^D\},$$

and it is now clear that these events are independent.

It is standard that there exists $c_1 > 0$ such that $\mathbb{P}_1^{\text{RW}}(\tau_0^{D^1} > t) \sim c_1 t^{-1/2}$ as $t \to \infty$ (the superscript RW emphasises that $(D_t^1)_t$ is a discrete-time random walk on $\mathbb{Z}$). Hence there exists $c_2 > 0$ such that $\mathbb{P}_{\mathbf{x}}^{\text{WMIG}}(C^{X,U} < \tau_0^{D^1}) \leq c_2 \mathbb{E}_{\mathbf{x}}^{\text{WMIG}}[(C^{X,U})^{-1/2}]$. By the maximal Azuma's inequality, there exists $c_3 > 0$ such that $\tau_{-\lfloor d/4 \rfloor}^X$ is at least $d^2/(c_3 \log d)$ with probability at least $1 - d^{-3}$. Further, by a Chernoff bound, there exists $c_4 > 0$ such that given $\tau_{-\lfloor d/4 \rfloor}^X$, $C^{X,U}$ is at least $\frac{1}{3}\tau_{-\lfloor d/4 \rfloor}^X$ with probability at least $e^{-c_4 \tau_{-\lfloor d/4 \rfloor}^X}$. Putting these bounds together gives the existence of $c, \bar{c} > 0$ such that

$$\mathbb{P}_{\mathbf{x}}^{\text{WMIG}}(C^{X,U} < \tau_0^{D^1}) \leq c \left(\mathbb{E}_{\mathbf{x}}^{\text{WMIG}}[(\tau_{-\lfloor d/4 \rfloor}^X)^{-1/2}] + d^{-3} \right) \leq c \left(\mathbb{E}_{\lfloor d/4 \rfloor}^{\text{RW}}[(\tau_0^X)^{-1/2}] + d^{-3} \right)$$

(58)
$$\leq \bar{c} d^{-1},$$

where we have used Lemma 31 in the last inequality. Similarly, given T^X, $C^{V,Y}$ is at least $\frac{1}{3}T^X$ with probability at least $e^{-c_4 T^X}$. Note also that under WMIG $\tau_0^{D^2}$ and τ_0 are independent (since marking does not affect movement). Then we have the bound

$$\mathbb{P}_{\mathbf{x}}^{\text{WMIG}}(C^{V,Y} < \tau_0^{D^2} \wedge \tau^D) \leq c \left(\mathbb{E}_{\lfloor d/4 \rfloor}^{\text{RW}}[(\tau_0^X)^{-1/2}(1-p)^{\tau_0^X/3}] + d^{-3} \right)$$

(59)
$$\leq \hat{c} \left(d^{-1} e^{-\hat{c} d \sqrt{p}} + d^{-2} \right),$$

for some $\hat{c} > 0$, where we have applied Lemma 31 in the last inequality. Combining equations (55)–(59) completes the proof. $\qquad \square$

We now present the proof of Proposition 29.

PROOF OF PROPOSITION 29. Recall that $d := d(y, x) \geq d(u, v)$. The first step is to couple this CAB process with a WAG process denoted $(\mathbf{X}_t')_{t \geq 0} = (X_t', U_t', V_t', Y_t')_{t \geq 0}$. Let m be the midpoint between x and y, defined to satisfy

$$d(m, x) = \left\lfloor \frac{d}{2} \right\rfloor, \quad d(y, m) = \left\lceil \frac{d}{2} \right\rceil.$$

The coupled WAG process is initialised from

$$(X_0', U_0', V_0', Y_0') = \mathbf{x}' := (-d(m, x), -d(m, u), d(v, m), d(y, m)).$$

Up until WAG terminates, the absorption times of particles in $\mathbf{X}'$ to the cemetery state Δ are the same as their counterparts in $\mathbf{X}$. If particle X jumps clockwise (resp. anticlockwise), then particle X' jumps down (resp. up) by 1 (similarly for other particles). Then the meeting time of particles U and V coincides with the time that particle V' becomes a ghost.

Notice that for X and Y to meet before either X meets U or Y meets V, at least one of X and Y must hit the midpoint m, times we denote by τ_m^X and τ_m^Y. Therefore we have

$$\left\{ M^{X,Y} \vee M^{U,V} < \tau \right\} \subseteq \left\{ (\tau_m^X \wedge \tau_m^Y) \vee M^{U,V} < \tau \right\}.$$

However, the coupling forces

$$\left\{ (\tau_m^X \wedge \tau_m^Y) \vee M^{U,V} < \tau \right\} = \left\{ (\tau_0^{X'} \wedge \tau_0^{Y'}) \vee G^{V'} < \tau' \right\} \subseteq \left\{ \tau_0^{X'} \wedge \tau_0^{Y'} < \tau' \right\}$$

where $\tau' := \min\{\tau^{X'}, \tau^{U'}, \tau^{V'}, \tau^{Y'}, M^{X',U'}, M^{V',Y'}\}$. Hence we deduce that

$$\mathbb{P}_{\mathbf{x}}^{\mathrm{CAB}}(M^{X,Y} \vee M^{U,V} < \tau) \le \mathbb{P}_{\mathbf{x}'}^{\mathrm{WAG}}(\tau_0^{X'} \wedge \tau_0^{Y'} < \tau').$$

The proof is completed by combining this with Lemmas 32 and 33 (which can be applied since $n \ge 8$ and so $d \ge 4$, $d(m,x) \ge \frac{d}{2} - 1$ and $d(y,m) \ge \frac{d}{2} - \frac{1}{2}$). $\qquad\square$

APPENDIX

We present here some results which are used throughout. The first is an application of Stein's method to obtain conditions (involving variance and covariance) for Gaussian convergence.

THEOREM 34 (Röllin [40]). *Suppose W, W' are two random variables with the same distribution, $\mathbb{E}(W) = 0$, $\mathrm{Var}(W) = 1$ and $\mathbb{E}(W'|W) = (1 - \lambda)W$. Then for $\delta := \sup_{s \in \mathbb{R}} |\mathbb{P}(W \le s) - \Phi(s)|$,*

$$(60) \qquad \delta \le \frac{12}{\lambda} \sqrt{\mathrm{Var}(\mathbb{E}((W' - W)^2|W))} + 32\frac{A^3}{\lambda} + 6\frac{A^2}{\sqrt{\lambda}},$$

where A is such that $|W' - W| \le A$, and Φ is the cumulative distribution of a standard normal distribution on $\mathbb{R}$.

THEOREM 35 (Local central limit theorem [29, Theorem 1.2.1]). *Let $p_n(x)$ denote the probability that a lazy simple random walk on $\mathbb{Z}^d$ is located in $x \in \mathbb{Z}^d$ after n transitions starting from 0. Then*

$$(61) \qquad \left| p_n(x) - 2\left(\frac{d}{2\pi n}\right)^{d/2} e^{-\frac{d|x|^2}{2n}} \right| \le \frac{c}{n^{d/2}}(|x|^{-2} \wedge n^{-1}),$$

where $|x|^2 = \sum_{i=1}^d x_i^2$, and c is a universal constant.

THEOREM 36 (The hitting time theorem [46]). *Let $(X_t)_{t \ge 0}$ be a discrete-time simple random walk on $\mathbb{Z}$. Denote by T_0 the hitting time of $0 \in \mathbb{Z}$. Then*

$$(62) \qquad \mathbb{P}(T_0 = n|X_0 = k) = \frac{k}{n}\mathbb{P}(X_n = 0).$$

APPENDIX: ACKNOWLEDGMENTS

REFERENCES

[1] ADDARIO-BERRY, L., OLIVEIRA, R. I., PERES, Y. and SOUSI, P. (2016). Random walks colliding before getting trapped. *Electronic Journal of Probability* **21** 1–19.

[2] ALDOUS, D. and FILL, J. A. (2002). Reversible Markov Chains and Random Walks on Graphs. Unfinished monograph, recompiled 2014, available at http://www.stat.berkeley.edu/~aldous/RWG/book.html.

[3] BASU, R. and SLY, A. (2017). Evolving voter model on dense random graphs. *The Annals of Applied Probability* **27** 1235 – 1288.

[4] BERENBRINK, P., GIAKKOUPIS, G., KERMARREC, A.-M. and MALLMANN-TRENN, F. (2016). Bounds on the voter model in dynamic networks. *arXiv preprint arXiv:1603.01895*.

[5] BRESLER, G. and NAGARAJ, D. (2019). Stein's method for stationary distributions of Markov chains and application to Ising models. *The Annals of Applied Probability* **29** 3230–3265.

[6] CARRO, A., TORAL, R. and SAN MIGUEL, M. (2016). The noisy voter model on complex networks. *Scientific reports* **6** 1–14.

[7] CASTELLANO, C., FORTUNATO, S. and LORETO, V. (2007). Statistical physics of social dynamics. *Reviews of Modern Physics* **81**.

[8] CHATTERJEE, S. and DEY, P. S. (2010). Applications of Stein's method for concentration inequalities. *The Annals of Probability* **38** 2443–2485.

[9] CHATTERJEE, S. and SHAO, Q.-M. (2011). Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie–Weiss model. *The Annals of Applied Probability* **21** 464–483.

[10] CHEN, Y.-T., CHOI, J. and COX, J. T. (2016). On the convergence of densities of finite voter models to the Wright–Fisher diffusion. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* **52** 286–322. Institut Henri Poincaré.

[11] CLIFFORD, P. and SUDBURY, A. (1973). A model for spatial conflict. *Biometrika* **60** 581–588. MR343950

[12] COOPER, C., ELSASSER, R., ONO, H. and RADZIK, T. (2013). Coalescing random walks and voting on connected graphs. *SIAM Journal on Discrete Mathematics* **27** 1748–1758.

[13] COOPER, C., FRIEZE, A. and RADZIK, T. (2010). Multiple random walks in random regular graphs. *SIAM Journal on Discrete Mathematics* **23** 1738–1761.

[14] COX, J. T. (1989). Coalescing Random Walks and Voter Model Consensus Times on the Torus in $\mathbb{Z}^d$. *The Annals of Probability* **17** 1333 – 1366.

[15] COX, J. T., PERES, Y. and STEIF, J. E. (2016). Cutoff for the noisy voter model. *The Annals of Applied Probability* **26** 917–932.

[16] EICHELSBACHER, P. and LÖWE, M. (2010). Stein's method for dependent random variables occuring in statistical mechanics. *Electronic Journal of Probability* **15** 962–988.

[17] EICHELSBACHER, P. and REINERT, G. (2008). Stein's method for discrete Gibbs measures. *The Annals of Applied Probability* **18** 1588–1618.

[18] FERNÁNDEZ-GRACIA, J., SUCHECKI, K., RAMASCO, J. J., SAN MIGUEL, M. and EGUÍLUZ, V. M. (2014). Is the Voter Model a Model for Voters? *Phys. Rev. Lett.* **112** 158701.

[19] GALAM, S. (2005). Local dynamics vs.social mechanisms: A unifying frame. *http://dx.doi.org/10.1209/epl/i2004-10526-5* **70**.

[20] GOLDSTEIN, L. and WIROONSRI, N. (2018). Stein's method for positively associated random variables with applications to the Ising and voter models, bond percolation, and contact process. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* **54** 385–421. Institut Henri Poincaré.

[21] GRANOVSKY, B. L. and MADRAS, N. (1995). The noisy voter model. *Stochastic Processes and their applications* **55** 23–43.

[22] HASSIN, Y. and PELEG, D. (2001). Distributed probabilistic polling and applications to proportionate agreement. *Information and Computation* **171** 248–268.

[23] HOLLEY, R. A. and LIGGETT, T. M. (1975). Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probability* **3** 643–663. MR402985

[24] HUO, R. and DURRETT, R. (2019). The zealot voter model. *The Annals of Applied Probability* **29** 3128–3154.

[25] KANADE, V., MALLMANN-TRENN, F. and SAUERWALD, T. (2019). On coalescence time in graphs: When is coalescing as fast as meeting? In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms* 956–965. SIAM.

[26] KIRMAN, A. (1993). Ants, rationality, and recruitment. *The Quarterly Journal of Economics* **108** 137–156.

[27] KRAPIVSKY, P. L. (1992). Kinetics of monomer-monomer surface catalytic reactions. *Phys. Rev. A* **45** 1067–1072.

[28] LANCHIER, N. (2012). The Axelrod model for the dissemination of culture revisited. *The Annals of Applied Probability* **22** 860 – 880.

[29] LAWLER, G. F. (2013). *Intersections of random walks*. *Modern Birkhäuser Classics*. Birkhäuser/Springer, New York Reprint of the 1996 edition. MR2985195

[30] LEVIN, D. A., PERES, Y. and WILMER, E. L. (1999). *Markov chains and mixing times*. American Mathematical Soc.

[31] LIGGETT, T. M. (2005). *Interacting particle systems*. *Classics in Mathematics*. Springer-Verlag, Berlin Reprint of the 1985 original. MR2108619

[32] MARINOV, R. (2013). Counting Vertices in a Voter-type Model. *arXiv preprint arXiv:1311.4807*.

[33] OLIVEIRA, R. I. (2012). On the coalescence time of reversible random walks. *Trans. Amer. Math. Soc.* **364** 2109–2128. MR2869200

[34] OLIVEIRA, R. I. and PERES, Y. (2019). Random walks on graphs: new bounds on hitting, meeting, coalescing and returning. In *2019 Proceedings of the Sixteenth Workshop on Analytic Algorithmics and Combinatorics (ANALCO)* 119–126. SIAM.

[35] PERALTA, A. F., CARRO, A., SAN MIGUEL, M. and TORAL, R. (2018). Analytical and numerical study of the non-linear noisy voter model on complex networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **28** 075516.

[36] PERALTA, A. F., CARRO, A., SAN MIGUEL, M. and TORAL, R. (2018). Stochastic pair approximation treatment of the noisy voter model. *New Journal of Physics* **20** 103045.

[37] RAMADAS, H. (2014). Mixing of the noisy voter model. *Electronic Communications in Probability* **19**.

[38] RAVASZ, M., SZABÓ, G. and SZOLNOKI, A. (2004). Spreading of families in cyclic predator-prey models. *Phys. Rev. E* **70** 012901.

[39] RINOTT, Y. and ROTAR, V. (1997). On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted U-statistics. *Ann. Appl. Probab.* **7** 1080–1105. MR1484798

[40] RÖLLIN, A. (2008). A note on the exchangeability condition in Stein's method. *Statistics & Probability Letters* **78** 1800–1806.

[41] ROSS, N. (2011). Fundamentals of Stein's method. *Probability Surveys* **8** 210–293.

[42] SCHNEIDER-MIZELL, C. M. and SANDER, L. M. (2009). A generalized voter model on complex networks. *Journal of Statistical Physics* **136** 59–71.

[43] STARNINI, M., BARONCHELLI, A. and PASTOR-SATORRAS, R. (2012). Ordering dynamics of the multi-state voter model. *Journal of Statistical Mechanics: Theory and Experiment* **2012** P10027.

[44] STEIN, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the sixth Berkeley symposium on mathematical statistics and probability, volume 2: Probability theory* **6** 583–603. University of California Press.

[45] SUCHECKI, K., EGUÍLUZ, V. M. and MIGUEL, M. S. (2004). Conservation laws for the voter model in complex networks. *Europhysics Letters* **69** 228.

[46] VAN DER HOFSTAD, R. and KEANE, M. (2008). An elementary proof of the hitting time theorem. *Amer. Math. Monthly* **115** 753–756. MR2456097

[47] YILDIZ, E., OZDAGLAR, A., ACEMOGLU, D., SABERI, A. and SCAGLIONE, A. (2013). Binary Opinion Dynamics with Stubborn Agents. *ACM Trans. Econ. Comput.* **1** 19:1–19:30.