DISCORDANT EDGES FOR THE VOTER MODEL
ON REGULAR RANDOM GRAPHS

LUCA AVENA, RANGEL BALDASSO, RAJAT SUBHRA HAZRA, FRANK DEN HOLLANDER,
MATTEO QUATTROPANI

ABSTRACT. We consider the two-opinion voter model on a regular random graph with n vertices and degree $d \geq 3$. It is known that consensus is reached on time scale n and that on this time scale the volume of the set of vertices with one opinion evolves as a Fisher-Wright diffusion. We are interested in the evolution of the number of discordant edges (i.e., edges linking vertices with different opinions), which can be thought as the perimeter of the set of vertices with one opinion. We show that if initially the two opinions are drawn independently from a Bernoulli distribution with parameter $u \in (0, 1)$, then on time scale 1 the fraction of discordant edges decreases and stabilises to a value that depends on d and u, and is related to the meeting time of two random walks on an infinity tree of degree d starting from two neighbouring vertices. Moreover, we show that on time scale n the fraction of discordant edges moves away from the constant plateau and converges to zero in an exponential fashion. Our proofs exploit the classical dual system of coalescing random walks and use ideas from Cooper et al. (2010) built on the so-called First Visit Time Lemma. We further introduce a novel technique to derive concentration properties from weak-dependence of coalescing random walks on moderate time scales.

Key words. Regular random graph, voter model, random walks, discordant edges, concentration.

MSC2020: 05C81, 60K35.

1. MODEL, LITERATURE AND RESULTS

Random processes on random graphs constitute a research area that poses many challenges. In the past decade, considerable progress has been made in understanding how the geometry of the graph affects the evolution of the process. In terms of the choice of graph, the focus has been on the Erdös-Rényi random graph, the configuration model, the preferential attachment model, and the exponential random graph, either directed or undirected, and in regimes ranging from sparse to dense. In terms of the choice of process, the focus has been on percolation, random walk, the stochastic Ising model, the contact process, and the voter model. What makes the area particularly interesting is that there is a delicate interplay between the *size* of the graph and the *time scale* on which the process is observed. For short times, the process behaves as if it lives on an infinite graph. For instance, many sparse graphs are locally tree-like and therefore the process behaves as if it evolves on an infinite Galton-Watson tree. For long times, however, the process sees that the graph is finite and exhibits a crossover in its behaviour. For instance, the voter model, which will be the process of interest in the present paper, eventually reaches consensus

on any finite connected graph, but the time at which it does depends on the size of the graph. Many such instances can be captured under the name of *finite-systems scheme*, i.e., the challenge to identify how a finite truncation of a stochastic system behaves as both the time and the truncation level tend to infinity, properly tuned together (Cox and Greven (1990), Cox and Greven (1994)).

Random processes on random graphs are part of the larger research area of random processes in *random environment*, where the environment selects the random transition probabilities or transition rates. Another name is that of *interacting particle systems* in random environment. For random process on lattices a more or less complete theory has been developed over the past four decades. The challenge is to extend this patrimony to random graphs. In the present paper we focus on the voter model on the regular random graph in the sparse regime. We track how the fraction of *discordant edges* evolves over time in the limit as the size of the graph tends to infinity, and identify its scaling behaviour on *three time scales*: short, moderate, and long.

Voter models were introduced and studied in Clifford and Sudbury (1973) and Holley and Liggett (1975). Voter models and their consensus times on finite graphs were analysed in Donnelly and Welsh (1983) and Cox (1989). The behaviour of voter models on random networks depends on the realisation of the network: Sood et al. (2008) made various predictions for the expected consensus time on heterogeneous random networks (including power-law random graphs). More recently, Fernley and Ortgiese (2019) study the asymptotics of the consensus time on inhomogeneous random graph models like the Chung-Lu model and the Norros-Reitu model. The expected consensus time for regular random graphs was studied in Cooper et al. (2010), in the discrete-time synchronous setting. As far as we know, there is no literature on the evolution of the number of discordant edges on random graphs.

1.1. **Model and background.** Given a connected graph $G = (V, E)$, the voter model is the Markov process $(\eta_t)_{t \geq 0}$ on state space $\{0, 1\}^V$, with $\eta_t = \{\eta_t(x) \colon x \in V\}$, where $\eta_t(x)$ represents the opinion at time t of vertex x. Each vertex is equipped with an independent exponential clock that rings at rate one, after which the vertex selects one of the neighbouring vertices uniformly at random and copies its opinion. (A formal description of this interacting particle system, its generator, and how its dynamics can be built up via the so-called graphical representation is postponed to Section 2.)

As mentioned previously, while the voter dynamics has been widely studied on periodic lattices, periodic tori, and complete graphs, only recently its evolution on general connected graphs has been considered. In particular, as discussed below, if the underlying (random or non-random) graph has sufficiently nice properties, then it is possible to identify the time scale on which consensus takes place, and to determine how the process behaves on this time scale. In the present paper we specialise to the d-regular random graph ensemble with $d \geq 3$. In particular, we consider the sequence of random graphs $(G_{d,n}(\omega))_{n \in \mathbb{N}}$, with law denoted by $\mathbb{P}$, where each element $G_{d,n}(\omega) = (V, E(\omega))$ is a regular random graph of degree $d \geq 3$, consisting of $|V| = n$ vertices and $|E(\omega)| = m = dn/2$ edges, uniformly sampled from the set of simple d-regular graphs with n vertices (to guarantee that dn is even, for d odd we restrict to n even). These graphs are locally tree-like and have good expansion properties (see Section 2.3), lying within the realm of so-called *mean-field geometries* (see Section 1.2). Therefore the voter model on the d-regular random

graph ensemble can be investigated in some depth, and refined statements clarifying how consensus is reached can be derived. In the present paper we analyse the evolution of the density of *discordances*. In other words, denoting the set of discordant edges at time t by

$$D_t^n = \{e = (x, y) \in E \colon \eta_t(x) \neq \eta_t(y)\},$$

we will study the fraction of discordant edges at time $t \geq 0$ given by

$$\mathcal{D}_t^n = \frac{|D_t^n|}{m} \tag{1.1}$$

in the limit as $n \to \infty$.

1.2. **Voter model on mean-field geometries.**

• *Voter model.* A classical model of relevance in population genetics is the voter model on the *complete graph*, which is referred to as the *Moran model* or the *Fisher-Wright model* (depending on whether continuous-time asynchronous or discrete-time parallel updates are considered). The states $\{0, 1\}$ represent two different alleles, each coding for a specific genetic trait, and the n vertices of the complete graph K_n represent individuals in a population of size n. In the limit as $n \to \infty$, the consensus time (to be interpreted as the extinction time of one of the two traits) is known to scale linearly in n. Furthermore, on time scale n the fraction of the individuals of, say, type 1 converges as a process to the so-called *Fisher-Wright diffusion*. These results can be derived by analysing the backward genealogical progeny, which amounts to studying n coalescing random walks evolving on the same graph, which in turn is related to the so-called *Kingman coalescent* (we refer to Durrett (2008) for details). In Section 2 we give a precise description of this *duality* on arbitrary graphs.

These by now classical results have been recently extended to general *mean-field geometries* under two main conditions that can be roughly described as follows: (I) the stationary distribution of the random walk on the random graph must be *not too concentrated*; (II) *fast-mixing* in the sense of an asymptotically vanishing ratio between the mixing time of a single random walk and the expected meeting time of two independent random walks starting from stationarity.

• *Dual process.* For what concerns the dual process of coalescing random walks, Oliveira (2013) computes the distributional limit of the coalescence time under the above mentioned assumptions. In particular, it follows from (Oliveira, 2013, Theorem 1.2) that

$$\lim_{n \to \infty} \frac{\mathbf{E}[\tau_{\mathrm{coal}}]}{\mathbf{E}[\tau_{\mathrm{meet}}^{\pi \otimes \pi}]} = 2, \tag{1.2}$$

where τ_{coal} is the coalescence time of n random walks, each starting from a different vertex, and $\tau_{\mathrm{meet}}^{\pi \otimes \pi}$ is the meeting time of two random walks independently starting from stationarity. This offers insight into how the dual voter model behaves in mean-field like geometries. In particular, (1.2) allows us to translate the question of how long it takes for the voter model to achieve consensus to the question of controlling the meeting time of two random walks starting from stationarity. Indeed, as a consequence of this duality, it is immediate that the n-coalescence time τ_{coal} stochastically dominates the above mentioned *consensus time*, defined as

$$\tau_{\mathrm{cons}} = \inf\{t \geq 0 \colon \eta_t(x) = \eta_t(y), \text{ for all } x, y \in V\}. \tag{1.3}$$

3

For what concerns the convergence to a Fisher-Wright diffusion after proper scaling, the recent work by Chen et al. (2016) considers the *fraction of vertices in state* 1 *at time* t in the voter model,

$$\mathcal{B}_t^n = \frac{1}{n} \sum_{x \in V} \eta_t(x), \qquad t \geq 0, \tag{1.4}$$

under the above mentioned *mean-field conditions* (see (Chen et al., 2016, Theorem 2.2) for a precise formulation). The main result states that, when time is speeded up by a factor $\gamma_n := \mathbf{E}[\tau_{\mathrm{meet}}^{\pi \otimes \pi}]$, the rescaled density of vertices in state 1, i.e., $(\mathcal{B}_{\gamma_n t}^n)_{t \geq 0}$, converges as $n \to \infty$ to a Fisher-Wright diffusion in the Skorokhod topology.

• *Regular random graphs as mean-field geometries.* Let us next discuss the implications of the results mentioned above within the specific framework of d-regular random graphs. The latter is a geometric setting that satisfies the aforementioned *mean-field conditions* with high probability with respect to the law $\mathbb{P}$ of the environment. Indeed, it is well known that a typical realisation of the graph is connected with high probability as soon as $d \geq 3$. Due to the undirectedness of the edges, the corresponding stationary distribution is uniform over the vertex set. Furthermore, with high probability under $\mathbb{P}$, the mixing time is of order $\log n$ (see Lubetzky and Sly (2010)), while the expected meeting time (see Chen (2021)) satisfies

$$\frac{\mathbf{E}[\tau_{\mathrm{meet}}^{\pi \otimes \pi}]}{n} \xrightarrow{\mathbb{P}} \frac{1}{2\theta_d}$$

with

$$\theta_d = \frac{d-2}{d-1}. \tag{1.5}$$

Hence, by applying (1.2) to this setting, we find that

$$\frac{\mathbf{E}[\tau_{\mathrm{coal}}]}{n} \xrightarrow{\mathbb{P}} \frac{1}{\theta_d}.$$

Moreover, for the voter model starting from independent Bernoulli opinions with parameter $u \in (0,1)$, by Chen et al. (2016) we also have that the fraction of opinions of type 1, i.e., $\mathcal{B}_t^n$ converges, after time is speeded up by a factor n, to the Fisher-Wright diffusion $(\bar{\mathcal{B}}_s)_{s \geq 0}$ given by the SDE

$$\mathrm{d}\bar{\mathcal{B}}_s = \sqrt{2\theta_d \bar{\mathcal{B}}_s (1 - \bar{\mathcal{B}}_s)} \, \mathrm{d}W_s, \qquad \bar{\mathcal{B}}_0 = u, \tag{1.6}$$

where $(W_s)_{s \geq 0}$ denotes the standard Brownian motion (with generator $\frac{1}{2}\Delta$).

In conclusion, on d-regular ensembles consensus is reached on average on time scale n and, as far as the law is concerned, on the time scale of the consensus the process can be well approximated by the Fisher-Wright diffusion in (1.6). Interestingly, as pointed out by Chen et al. (2016), the diffusion coefficient in (1.6) (which is referred to as the *genetic variability* in a population genetics context) can be related to the fraction of discordant edges. This will be the starting point of our investigation, which is devoted to a deeper understanding of the evolution of the voter model beyond the consensus time scale n, via a detailed analysis of the discordant edges .

1.3. **Main theorems: evolution of discordances and stabilisation before consensus.** Before we present our main results, we introduce some notation. The symbols $\mathbb{P}$ and $\mathbb{E}$ will be reserved for the probability space of the d-regular random graph. The abbreviation *whp* refers to events that occur with probability $\mathbb{P}$ tending to 1 as $n \to \infty$.

Thanks to the duality between the n-vertex voter model and a system of n-coalescing random walks (to be described in more detail in Section 2), it is sufficient to look at a collection of $2m$ Poisson processes, each associated with an (oriented) edge of the graph, together with the initial assignment of the opinions. In this way, we can use the very same source of randomness to sample the evolution of the voter model, the dual system of coalescing random walks, and a system of n independent random walks, each starting at a different vertex of the graph. For this reason we adopt the symbols $\mathbf{P}$ and $\mathbf{E}$ to refer to any of these three stochastic processes evolving on a quenched realisation of the graph. To distinguish between the random walks and the voter model, in the latter case we use a sub-index to refer to the initial distribution of the process. For instance, $\mathbf{P}_\xi$ denotes the law of the voter model starting from the configuration $\eta_0 = \xi \in \{0,1\}^V$. Similarly, when we write $\mathbf{P}_u$, we refer to the voter model initialised by a product of Bernoulli random variables of mean $u \in (0,1)$. On the other hand, when considering a system of two or more independent or coalescing random walks, we simply use the symbols $\mathbf{P}$ and $\mathbf{E}$, and the starting positions of the random walks will be clear from the context (with the exception of Section 5, where an *ad-hoc* notation will be introduced). Finally, we write $\mathrm{Ber}(u)$ for a Bernoulli distribution of parameter u.

We now present our main results. The first theorem identifies *whp* the first order asymptotics of the expected number of discordant edges on all time scales.

Theorem 1.1 (Expected density of discordances at all time scales). *Fix $d \geq 3$, and let θ_d be as in (1.5). Fix $u \in (0,1)$, and consider the voter model on a d-regular random graph $G_{d,n}(\omega)$ with initial distribution $[\mathrm{Ber}(u)]^{\otimes V}$. Then, for any non-negative time t_n, the density of discordant edges in (1.1) satisfies*

$$\left| \mathbf{E}_u \left[\mathcal{D}^n_{t_n} \right] - 2u(1-u)(1 - f_d(t_n)) \mathrm{e}^{-2\theta_d \frac{t_n}{n}} \right| \xrightarrow{\mathbb{P}} 0, \tag{1.7}$$

where

$$f_d(t) = \mathbf{P}^{\mathcal{T}_d}(\tau^{x,y}_{\mathrm{meet}} \leq t) = \sum_{\kappa=0}^{\infty} \mathrm{e}^{-2t} \frac{(2t)^\kappa}{\kappa!} \sum_{s=0}^{\lfloor \frac{\kappa-1}{2} \rfloor} \binom{2s}{s} \frac{1}{s+1} \left(\frac{1}{d}\right)^{s+1} \left(\frac{d-1}{d}\right)^s, \tag{1.8}$$

with $\mathbf{P}^{\mathcal{T}_d}$ denoting the law of two independent random walks on the infinite d-regular tree $\mathcal{T}_d$ starting from the endpoints of an edge $e = (x,y)$ in $\mathcal{T}_d$.

Note that the limit is different depending on the time scale we are looking at:

- (**Short time scale**) For $t_n = \Theta(1)$, the first order of the above expectation is given by $2u(1-u)(1-f_d(t_n))$, and the behaviour is governed by the non-decreasing function captured in (1.8), representing the meeting time of two random walks starting from adjacent vertices on an infinite d-regular tree. See Figure 1.
- (**Moderate time scale**) For $t_n = \omega(1) \cap o(n)$, the discordances stabilise at a limiting plateau $2u(1-u)(1-f_d(\infty))$ characterised by the probability that two adjacent random walks on an infinite tree never meet, i.e., $1 - f_d(\infty) = \theta_d$.

5

- (**Long time scale**) For $t_n = sn$, $s \in (0, \infty)$, the dual system coalesces and the voter model reaches consensus. The evolution of the density of opinions is approximated by the Fisher-Wright diffusion in (1.6), and the expected density of discordances is characterised by the mean of the genetic variability $\mathbf{E}_u \left[\bar{\mathcal{B}}_s(1 - \bar{\mathcal{B}}_s) \right] = 2u(1-u)\theta_d \mathrm{e}^{-2\theta_d s}$. See Figure 2.
- (**Consensus**) For $t_n = \omega(n)$, the system has reached consensus and the limiting formula in (1.7) degenerates to zero.
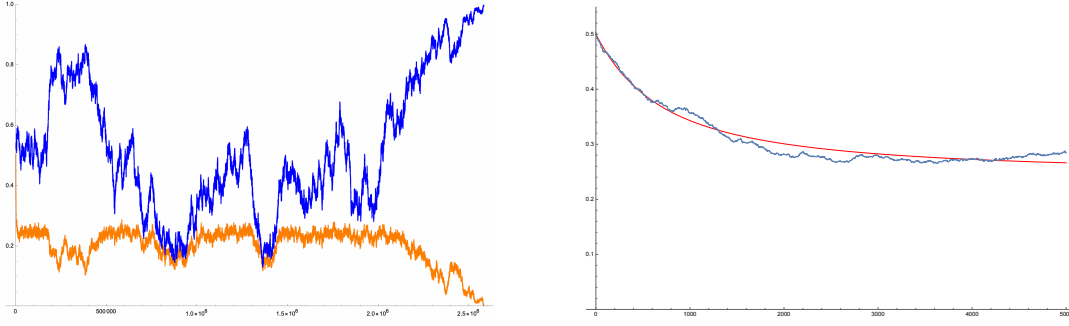
Figure 1. The two plots show a single simulation of the voter model on a regular random graph of size $n = 1000$, degree $d = 3$ and initial density $u = 0.5$. *Left*: In blue the density of blue opinions up to consensus ($\tau_{\mathrm{cons}} \approx 2.6 \times 10^3$), in orange the density of discordant edges up to consensus. *Right*: In blue the density of discordant edges up to time $t = 5$ (corresponding to a zoom-in of the plot to the left), in red the function $t \mapsto 2u(1-u)f_d(t)$.
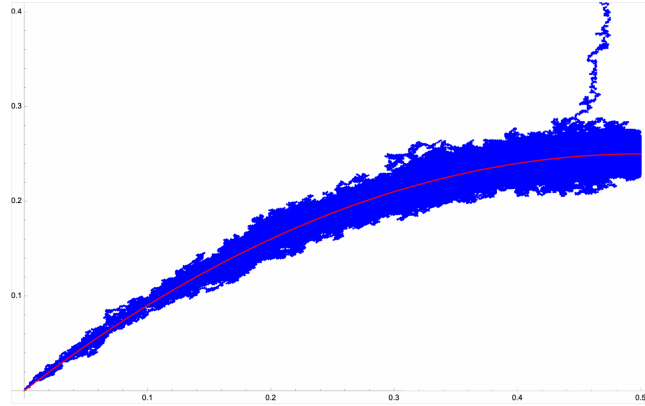
Figure 2. Scatter plot in blue for the same simulation as in Figure 1: the density of discordant edges versus the density of the minority opinion. The piece sticking out corresponds to short times. The curve in red is $x \mapsto x(1-x)$.

The second theorem characterises the process $\mathcal{D}^n$ in (1.1) beyond its expectation. In particular, we show that on time scale $o(n)$ the density of discordances concentrates around the expectation, while on time scale $\Theta(n)$ it behaves as a deterministic function of the Fisher-Wright diffusion.

6

Theorem 1.2 (Beyond expectation). *Consider the voter model on a d-regular random graph $G_{d,n}(\omega)$ with $d \geq 3$.*

(i) **Concentration before coalescence.** *Let t_n be such that $t_n/n \to 0$. Then, for every $\varepsilon > 0$,*

$$\sup_{\xi \in \{0,1\}^V} \mathbf{P}_\xi \left(\left| \mathcal{D}_{t_n}^n - \mathbf{E}_\xi[\mathcal{D}_{t_n}^n] \right| > \varepsilon \right) \xrightarrow{\mathbb{P}} 0.$$

(ii) **Discordance on consensus time scale.** *Let t_n be such that $t_n/n \to s \in (0,\infty)$. Then, for every $u \in (0,1)$,*

$$\sup_{x \in [0,1]} \left| \mathbf{P}_u \left(\mathcal{D}_{t_n}^n \leq x \right) - \mathbf{P}_u \left(\bar{\mathcal{B}}_{2s\theta_d}(1 - \bar{\mathcal{B}}_{2s\theta_d}) \leq x \right) \right| \xrightarrow{\mathbb{P}} 0,$$

where $(\bar{\mathcal{B}}_s)_{s \geq 0}$ is the solution of (1.6).

The third and last theorem is a strengthening of Theorem 1.2*(i)*. More precisely, we show that as soon as we focus on the process for times that are polynomially smaller than the consensus time, the concentration around the mean is actually uniform in time.

Theorem 1.3 (Uniform concentration on moderate time scale). *Consider the voter model on a d-regular random graph $G_{d,n}(\omega)$. Then, for every $u \in (0,1)$ and $\delta, \varepsilon > 0$,*

$$\mathbf{P}_u \left(\sup_{0 \leq t \leq n^{1-\delta}} \left| \mathcal{D}_t^n - \mathbf{E}_u[\mathcal{D}_t^n] \right| > \varepsilon \right) \xrightarrow{\mathbb{P}} 0.$$

1.4. **Open problems.** We point out two open problems.

- Theorem 1.3 says that the concentration of the fraction of discordant edges is uniform up to times $n^{1-\delta}$, for any $\delta > 0$. We expect that Theorem 1.3 can be strengthened to the statement that for every t_n such that $t_n/n \to 0$ and every $C_n \to \infty$,

$$\mathbf{P}\left(\left| \mathcal{D}_{t_n}^n - \mathbf{E}[\mathcal{D}_{t_n}^n] \right| > C_n \sqrt{t_n/n} \right) \xrightarrow{\mathbb{P}} 0. \tag{1.9}$$

Note that, thanks to Azuma inequality, the concentration in (1.9) holds for the quantity $\mathcal{B}_t^n$ in (1.4), because this is a martingale. Note further that the concentration in Theorem 1.3 follows from (1.9) and a union bound. For now, (1.9) is beyond our reach.

- We expect that Theorems 1.1–1.3 can be extended to *non-regular* sparse random graphs. We do not have a conjecture on how the function f_d and the diffusion constant θ_d modify in this more general setting.

1.5. **Description of techniques.** All our proofs are based on the classical notion of duality between the voter model and a collection of coalescent random walks. In particular, on time scales $o(\log n)$, i.e., below the typical distance between two vertices, the analysis can be carried out by coupling a system of two random walks, starting at adjacent vertices and evolving on the d-regular random graph, with two random walks on the d-regular tree. Because the tree is regular, the distance of the two random walks can be viewed as the distance to the origin of a single biased random walk on $\mathbb{N}_0$. Note that the same does not hold when the tree is not regular.

Clearly, in order to analyse the process on time scales $\Theta(\log n)$, i.e., on the typical distance between two vertices, the coupling argument must be combined with a finer control of the two random walks on the regular random graph. To this aim, we exploit the strategy developed in Cooper et al. (2010), which uses the so-called *First Visit Time Lemma*

(see Theorem 2.2). Such control, together with a first-moment argument, is enough to compute the evolution of the expected number of discordant on every time scale.

On the other hand, in order to obtain the concentration result in Theorem 1.3, a much deeper analysis is required. Roughly, in order to have proper control on the correlations between the edges constituting $\mathcal{D}_t^n$, we must analyse a dual system of random walks whose number grows with n. Exploiting a classical result by Aldous and Brown (1992), we derive upper bounds for the number of meetings of a poly-logarithmic number of independent random walks evolving on the random graph for a time $n^{1-o(1)}$. Such a bound will be exploited in the forthcoming Proposition 5.1 to deduce an upper bound for the deviation from the mean that is exponentially small in n and uniform in time. This upper bound can in turn be translated to the result in Theorem 1.3 by taking a union bound.

1.6. **Outline.** Section 2 lists definitions and notations, recalls the graphical construction and duality, states a key lemma, and collects a few basic facts about regular random graphs and random walks on such geometries. Theorems 1.1–1.3 are proved in Sections 3–5, respectively. Appendix A contains two auxiliary facts for càdlàg processes.

2. NOTATION AND PREPARATION

In this section we properly introduce the voter model on general graphs, and collect some results that will be needed along the way. In Section 2.1 we introduce the graphical construction for the voter model and the associated dual process, known as *coalescing random walks*. Section 2.2 contains the *First Visit Time Lemma*, which has been introduced by Cooper and Frieze (2005). Section 2.3 collects some useful facts about the geometry of d-regular random graphs and the behaviour of random walks on them. In what follows we drop the upper index n to lighten the notation.

2.1. **The voter model and coalescing random walks.** Let $G(V,E)$ be a connected (possibly infinite and locally-finite) undirected graph. For each $x \in V$, let d_x denote the degree of vertex x. The voter model on G is defined as the interacting particle system with state space $\{0,1\}^V$ and generator $\mathcal{L}$ acting on local functions $f \colon V \to \mathbb{R}$ as

$$(\mathcal{L}f)(\eta) = \sum_{x \in V} \sum_{y \sim x} \frac{1}{d_x} \big(f(\eta^{x \leftarrow y}) - f(\eta)\big), \qquad \eta \in \{0,1\}^V,$$

where $\eta^{x \leftarrow y}$ is obtained from η as

$$\eta^{x \leftarrow y}(z) = \begin{cases} \eta(y), & \text{if } z = x, \\ \eta(z), & \text{otherwise.} \end{cases}$$

In this system, agents (represented by the vertices of G) start with binary opinions and update these at rate one by copying the opinion of a uniformly chosen neighbour. For $t \geq 0$ and $x \in V$, denote by $\eta_t(x) \in \{0,1\}$ the opinion of vertex x at time t.

When G is finite and regular, i.e., $d_x = d$ for every $x \in V$, the number of vertices that have opinion 1,

$$B_t = \sum_{x \in V} \eta_t(x),$$

8

is a martingale. Since B_t is bounded by $|V|$, it converges, which implies that the voter model eventually fixates on a constant configuration, composed entirely of 0's or 1's, denoted by $\bar{0}$ and $\bar{1}$, respectively. Define the consensus time of G as the stopping time

$$\tau_{\text{cons}} = \inf \left\{ t \geq 0 \colon \eta_t \in \{\bar{0}, \bar{1}\} \right\}.$$

Let us next introduce the *graphical construction* of the voter model. Associate to every oriented edge $\vec{e} = (x, y)$ a Poisson point process $\mathcal{P}_{\vec{e}}$ on $\mathbb{R}$ with intensity $1/d_x$. When a clock from $\vec{e} = (x, y)$ rings, vertex x receives the opinion of vertex y. This constriction allows us to track the joint evolution of all the opinions via time duality. In order to determine $\eta_t(x)$, the state of a vertex x at time t, we start from the space time point (x, t), and go backwards in time, crossing every clock ring from an edge that reaches the current position of the path (see Figure 3).
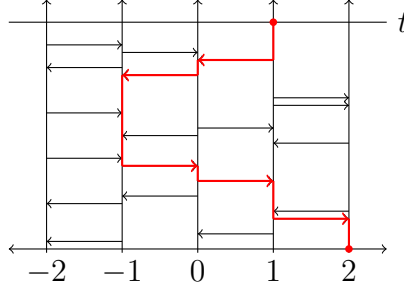
Figure 3. The graphical construction of the voter model on the integer lattice $\mathbb{Z}$. The red path allows us to determine that $\eta_t(1)$ is equal to $\eta_0(2)$.

The process that tracks the origin of the opinions can be viewed as a collection of *coalescing random walks*. To define this process we assume that G is finite and consider a family $\{(X^v_t)_{t \geq 0}\}_{v \in V}$ of independent rate-one continuous-time random walks such that $X^v_0 = v$ almost surely. In what follows we identify V with $[n] = \{1, \ldots, n\}$. When two walks meet at the same vertex they coalesce into a single walk (this can be seen immediately by the graphical representation in 3.)

Because G is finite, any two independent random walks meet in finite time. Consequently, τ_{coal}, the first time at which all the random walks sit at the same vertex called the *coalescence time*, satisfies $\tau_{\text{coal}} < \infty$ almost surely. In particular, the above mentioned duality implies that $\tau_{\text{cons}} \leq \tau_{\text{coal}}$, and therefore any upper bound on the coalescence time immediately yields an upper bound on the consensus time of the voter model.

Given $x, y \in [n]$, let

$$\tau^{x,y}_{\text{meet}} = \inf\{t \geq 0 \colon X^x_t = X^y_t\} \tag{2.1}$$

denote the *meeting time* of two independent continuous-time random walks starting from x and y, respectively. More generally, if the two random walks are initialised randomly and independently according to two distributions μ and ν on $[n]$, then we denote their meeting time by $\tau^{\mu \otimes \nu}_{\text{meet}}$. We note that the extension to infinite graphs G follows from standard estimates on the probability that the process can be approximated by a finite dynamics in finite portions of space during a bounded amount of time.

2.2. The First Visit Time Lemma.

In this section we state the *First Visit Time Lemma* (FVTL), which was introduced by Cooper and Frieze (2005) and recently simplified by Manzo et al. (2021). This lemma contains estimates on the time a random walk takes to reach a distinguished site when starting from the stationary distribution.

Remark 2.1. Since the FVTL is defined in discrete time, in what follows it will be easier to derive the results in discrete time and then translate them to continuous time. For this reason, when considering two random walks in discrete time, where at each step the random walk that moves is selected uniformly at random, we use the notation $\mathbf{P}^{\mathrm{dt}}$ to refer to their law.

Theorem 2.2 (First Visit Time Lemma; Manzo et al. (2021)). *Consider a sequence of ergodic Markov chains labelled by $N \in \mathbb{N}$, living on a finite state space Ω_N of size N, with transition matrix Q_N and stationary distribution μ_N and a target state $x \in \Omega_N$. Define*

$$T_{\mathrm{mix}}^{(N)} = \inf \left\{ t \geq 0 \colon \max_{x,y \in \Omega_N} |Q_N^t(x,y) - \mu_N(y)| \leq \frac{1}{N^3} \right\}, \tag{2.2}$$

and assume that

(1) $\lim_{N \to \infty} N^2 \min_{x \in \Omega_N} \mu_N(x) = \infty.$
(2) $\lim_{N \to \infty} T_{\mathrm{mix}}^{(N)} \max_{x \in \Omega_N} \mu_N(x) = 0.$
(3) *There exists a unique* quasi-stationary distribution *associated to x, i.e., for all N large enough the sub-matrix $[Q_N]_x$ in which the row and column have been erased is* irreducible*. See also (Quattropani and Sau, 2022, Remark 3.2 and Section 4.4).*

Then there exists a sequence $(\lambda_x^{(N)})_{N \in \mathbb{N}}$ such that

$$\lim_{N \to \infty} \left| \frac{\lambda_x^{(N)}}{\frac{\mu_N(x)}{R_N(x)}} - 1 \right| = 0, \qquad R_N(x) := \sum_{s=0}^{T_{\mathrm{mix}}^{(N)}} Q_N^s(x,x),$$

and

$$\lim_{N \to \infty} \sup_{t \geq 0} \left| \frac{\mathsf{P}_{\mu_N}(\tau_x > t)}{(1 - \lambda_x^{(N)})^t} - 1 \right| = 0,$$

with P_{μ_N} the path measure of the Markov chain starting from μ_N.

In words, the FVTL states that, under the assumptions, the hitting time of a state x is well approximated by a geometric (exponential, in continuous time) random variable with mean $\frac{R(x)}{\mu(x)}$.

2.3. The geometry of regular random graphs.

In this section we collect a number of definitions and asymptotic results about the *typical geometry* of regular random graphs. We start by introducing notation and conclude by stating some basic facts regarding typical d-regular random graphs.

Given a connected graph $G = (V, E)$ and two vertices $x, y \in V$, we denote by $\mathrm{dist}(x,y)$ the distance between x and y, given by the number of edges of the shortest path joining x and y. For a vertex $x \in V$ and a radius $h \geq 1$, we denote by $\mathcal{B}_x(h)$ the ball of radius h centered at x, seen as a subgraph of G. The tree excess of a graph quantifies how far away from a tree a given connected graph is.

Definition 2.3 (Tree excess). Given a connected graph $G = (V, E)$, the *tree excess* of G, denoted by $\texttt{tx}(G)$, is the minimum number of edges that must be removed from G in order to obtain a tree. This can also be written as

$$\texttt{tx}(G) = |V| - |E| - 1.$$

Next, we introduce the notion of *locally tree-like vertices and edges*.

Definition 2.4 (Locally tree-like vertex). Given a graph $G = (V, E)$, we say that a vertex $x \in V$ is locally tree-like up to distance $\ell \geq 1$, denoted by $x \in \mathrm{LTL}(\ell)$, if the subgraph induced by the vertices at distance at most ℓ from x, i.e., $\mathcal{B}_x(\ell)$, is a tree.

Definition 2.5 (Locally tree-like edge). Given a graph $G = (V, E)$ and an edge $e \in E$, we say that $e = \{x, y\}$ is locally tree-like up to distance $\ell \geq 1$, denoted by $e \in \mathrm{LTLE}(\ell)$, if the subgraph induced by the vertices at distance at most ℓ from x or y, i.e., $\mathcal{B}_x(\ell) \cup \mathcal{B}_y(\ell)$, after removal of the edge e, is composed of two disjoint trees of depth at most ℓ.

For $d \geq 3$ and $n > d$ such that nd is even, there exists at least one graph with n vertices and constant degree equal to d. In particular, for this range of parameters the d-regular random graph of size n, drawn from the uniform distribution on the set of all graphs with n vertices and constant degree equal to d, is well defined. Denote by $m = dn/2$ the number of edges of this random graph. The following proposition collects properties of the d-regular random graph that will be needed in the rest of the paper.

Proposition 2.6 (Geometry of regular random graphs and random walks; Lubetzky and Sly (2010)). *Fix $d \geq 3$, and let $G_{d,n}(\omega)$ denote the d-regular random graph of size n. Then, whp,*

 (i) *G is connected.*
 (ii) *For every $x \in [n]$, $\texttt{tx}(\mathcal{B}_x(\frac{1}{5}\log_d n)) \leq 1$.*
 (iii) *$|\mathrm{LTLE}(\frac{1}{5}\log_d n)| = m - o(n)$.*
 (iv) *There exists a constant $C = C_d > 1$ such that for $t = C \log n$ and for every $\varepsilon > 0$,*

$$d_{\mathrm{TV}}(t) := \max_{x \in [n]} \| P^t(x, \cdot) - \pi \|_{\mathrm{TV}} \leq \varepsilon. \tag{2.3}$$

In particular, for t_n such that $\lim_{n \to \infty} \log n / t_n = 0$, then $\lim_{n \to \infty} d_{\mathrm{TV}}(t_n) = 0$.

3. COMPUTATION OF THE EXPECTATION

In this section we prove Theorem 1.1, in particular, we provide a *whp* first-order estimate of the expected density of discordant edges at time $t = t_n$, for every choice of the sequence $(t_n)_{n \in \mathbb{N}}$. In Section 3.1 we collect some basic results for the voter model on the infinite d-regular tree $\mathcal{T}_d$, the *local approximation* of the d-regular random graph. We show that the density of discordant edges at time t on $\mathcal{T}_d$ with $\mathrm{Ber}(u)$ initialisation behaves as $2u(1 - u)(1 - f(t))$. This will be done by exploiting duality: we reduce the problem to the analysis of a pair of coalescing random walks starting at the two extremes of an edge of the tree. In Section 3.2 we show that a classical coupling argument suffices to show that the approximation holds for every sequence $t = (t_n)_{n \in \mathbb{N}}$ such that $t/\log n \to 0$. When t starts to become comparable with the typical distance of the graph, i.e., $t_n = \Theta(\log n)$, the coupling argument fails and we need a refined analysis of the process to compute the expected density. This scenario will be handled in Section 3.3. The idea is to track a pair of random walks starting from the two extremes of a typical (hence, *locally tree-like*) edge up

11

to their first meeting. Using the First Visit Time Lemma, we show that, under the event that the two random walks do not meet after a short time, the time of their first meeting is well approximated by exponential random variable with rate $\theta_d^{-1} n$.

3.1. **Discordant edges on the regular tree.** Let $\mathcal{T}_d$ denote the *infinite d-regular tree* with $d \geq 3$. Our first lemma concerns the probability that two independent random walks, starting at distance one from each other, do not meet within a time t. Recall that $\mathbf{P}^{\mathcal{T}_d}$ denotes the law of two independent random walks starting from then endvertices of an edge in $\mathcal{T}_d$.

Lemma 3.1 (Meeting on a regular tree). *Let $x, y \in \mathcal{T}_d$ with $\mathrm{dist}(x, y) = 1$. Then*

$$f_d(t) = \mathbf{P}^{\mathcal{T}_d}\left(\tau_{\mathrm{meet}}^{x,y} \leq t\right) = \sum_{\kappa=0}^{\infty} e^{-2t} \frac{(2t)^\kappa}{\kappa!} \sum_{s=0}^{\lfloor \frac{\kappa-1}{2} \rfloor} \binom{2s}{s} \frac{1}{s+1} \left(\frac{1}{d}\right)^{s+1} \left(\frac{d-1}{d}\right)^s. \tag{3.1}$$

From this it follows that

$$\mathbf{P}^{\mathcal{T}_d}\left(\tau_{\mathrm{meet}}^{x,y} = \infty\right) = \frac{d-2}{d-1} = \theta_d. \tag{3.2}$$

Proof. Let $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ be two independent continuous-time random walks starting from $x \sim y$, respectively. Put $Z_t = \mathrm{dist}(X_t, Y_t)$ and note that $(Z_t)_{t\geq 0}$ is a continuous-time biased random walk on $\mathbb{N}_0$ that starts at 1, jumps at rate 2, and has jump distribution given by

$$p(z, z+1) = 1 - p(z, z-1) = \frac{d-1}{d}, \qquad z \in \mathbb{N}. \tag{3.3}$$

The claim in (3.1) follows from the construction of the random walk $(Z_t)_{t\geq 0}$ via a Poisson process of rate 2 for the jump times, together with a skeleton chain given by the discrete-time biased random walk with jump distribution given by (3.3), and explicit expressions for hitting times of discrete-time random walks based on path-counting arguments.

To get (3.2) it is enough to note that, by the standard Gambler's ruin argument,

$$\mathbf{P}^{\mathcal{T}_d}\left(\tau_{\mathrm{meet}}^{x,y} < \infty\right) = \mathbf{P}\left(Z_t = 0 \text{ for some } t \geq 0 \mid Z_0 = 1\right) = \frac{p(0,-1)}{p(0,1)} = \frac{1}{d-1},$$

from which the claim follows. $\qquad\square$

Remark 3.2. Note that (3.2) can be derived from (3.1) by direct computation. Indeed, clearly $f(0) = 0$. The generating function $c(x) = \sum_{k=0}^{\infty} C_k x^k$ for the Catalan numbers $C_k = \binom{2k}{k} \frac{1}{k+1}$ solves the quadratic equation $c(x) = 1 + xc(x)^2$. Since $C_0 = 1$, we have

$$c(x) = \frac{\left[1 - \sqrt{1 - 4x}\right]}{2x}.$$

It follows that

$$f(\infty) = \frac{1}{d} c\left(\frac{d-1}{d^2}\right) = \frac{1}{d} \frac{\left[1 - \frac{d-2}{d}\right]}{2 \frac{d-1}{d^2}} = \frac{1}{d-1} = 1 - \theta_d.$$

Lemma 3.3 (Density of discordant edges). *Consider the voter model on $\mathcal{T}_d$ with initial density u. Then, for any given edge e and any time $t \geq 0$,*

$$\mathbf{P}_u^{\mathcal{T}_d}\left(e \in D_t\right) = 2u(1-u)(1 - f_d(t)),$$

12

where $f_d(t)$ is given by (3.1).

Proof. Put $e = \{x, y\}$. Note that in order to have $e \in D_t$, by duality, the (backward) random walks starting from x and y at time t do not coalesce before time 0 and end up at vertices with distinct initial opinions. The initial opinions are independent of the trajectories of these random walks. The claim follows by noting that $1 - f_d(t)$ is the probability that the two random walks do not meet in time t, and $2u(1 - u)$ is the probability that two distinct vertices have different initial opinions. □

3.2. **Expectation for *small times*.** We start this section with a general bound that does not require the graph G to be random nor the size of G to grow to infinity.

Lemma 3.4 (Discordant LTLEs). *Let $G = (V, E)$ be any d-regular graph and let $e = \{x, y\} \in E$ be $\mathrm{LTLE}(\ell)$ for some ℓ. Consider the voter model on G with initial density u. Then, for any time $T > 0$,*

$$\sup_{t \in [0,T]} |\mathbf{P}_u(e \in D_t) - 2u(1 - u)(1 - f_d(t))| \leq \frac{4T}{\ell}.$$

Proof. Note first that the term $2u(1 - u)(1 - f_d(t))$ corresponds to the probability that a given edge in $\mathcal{T}_d$ is discordant at time t. In particular, if we consider the dual system with coalescing random walks, then the probabilities we are interested in are the same, provided the random walks do not leave the tree-like neighbourhood of e. Therefore consider the event

$\mathcal{E}_T = \{$the number of jumps the random walks perform before time T is bounded by $\ell\}$,

and observe that, by the Markov inequality,

$$\mathbf{P}(\mathcal{E}_T^c) = \mathbf{P}^{\mathcal{T}_d}(\mathcal{E}_T^c) \leq \mathbf{P}^{\mathcal{T}_d}(X \geq \ell) \leq \frac{2T}{\ell},$$

where $X \sim \mathrm{Poisson}(2T)$ stochastically dominates the number of jumps that both random walks make up to time T. By construction, for $t \leq T$ we can write

$$\mathbf{P}_u(e \in D_t, \mathcal{E}_T) = \mathbf{P}_u^{\mathcal{T}_d}(e \in D_t, \mathcal{E}_T),$$

from which we can estimate

$$|\mathbf{P}_u(e \in D_t) - 2u(1 - u)(1 - f_d(t))| = \left|\mathbf{P}_u(e \in D_t) - \mathbf{P}_u^{\mathcal{T}_d}(e \in D_t)\right|$$

$$\leq \left|\mathbf{P}_u(e \in D_t, \mathcal{E}_T) - \mathbf{P}_u^{\mathcal{T}_d}(e \in D_t, \mathcal{E}_T)\right| + 2\mathbf{P}^{\mathcal{T}_d}(\mathcal{E}_T^c) \leq \frac{4T}{\ell},$$

which settles the claim. □

Proposition 3.5 (Short time average). *Consider the voter model on a regular random graph $G_{d,n}(\omega)$ with initial density $u \in (0, 1)$. Then, for any time t_n satisfying $t_n / \log n \to 0$,*

$$|\mathbf{E}_u[\mathcal{D}_{t_n}] - 2u(1 - u)(1 - f_d(t_n))| \xrightarrow{\mathbb{P}} 0. \tag{3.4}$$

Proof. Note that for every realisation of G,

$$\mathbf{E}_u[|D_{t_n}|] = \sum_{e \in E} \mathbf{P}_u(e \in D_{t_n}^n).$$

13

Call $E_\star \subset E$ the set of $\mathrm{LTLE}\big(\frac{1}{5}\log_d n\big)$ edges. Then, by Lemma 3.4 and our assumption on t_n,

$$\Big|\mathbf{E}_u[\mathcal{D}_{t_n}] - 2u(1-u)(1 - f_d(t_n))\Big| \leq \Big|\frac{1}{m}\sum_{e\in E}\mathbf{P}_u\big(e \in D_t\big) - 2u(1-u)(1 - f_d(t_n))\Big|$$

$$\leq \frac{1}{m}\sum_{e\in E_\star}\frac{4t_n}{\frac{1}{5}\log_d n} + \frac{1}{m}|E_\star^c| \xrightarrow{\mathbb{P}} 0,$$

where the last convergence follows from Proposition 2.6(iii). $\qquad\square$

3.3. **Expectation for *long times*.** In this section we show how the behaviour in Proposition 3.5 can be extended up to the linear time scale. Clearly, we need a different argument, since we can no longer assume that the dual random walks starting at the extremes of a locally tree-like edge do not exit their neighbourhood of size $\frac{1}{5}\log_d(n)$. In order to proceed, we adapt the arguments in (Cooper et al., 2010, Lemmas 17 and 20) to our framework, which will constitute a crucial tool for our analysis. In what follows, we fix

$$\sigma_n = \lceil(\log\log n)^2\rceil. \tag{3.5}$$

The following lemma says that *whp* for every pair of starting vertices that are more than σ_n apart the probability that the two random walks meet before time $\log^3 n$ is exponentially small in σ_n.

Lemma 3.6 (Meeting time of distant random walks). *Let $G_{d,n}(\omega)$ be a regular random graph. Then there exists a constant $c_0 > 0$ such that*

$$\mathbb{1}\left\{\max_{x,y:\ \mathrm{dist}(x,y)>\sigma_n}\mathbf{P}\big(\tau_{\mathrm{meet}}^{x,y} \leq \log^3 n\big) \leq \mathrm{e}^{-c_0\sigma_n}\right\} \xrightarrow{\mathbb{P}} 1.$$

Proof. The proof comes in several steps.

1. Consider two independent random walks $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ that start from x and y, respectively. The number of steps each random walk performs before time $\log^3 n$ has distribution $\mathrm{Poisson}(\log^3 n)$. In particular, Markov's inequality with exponential moments implies that, for $\lambda = \log\log n$,

$$\mathbf{P}\bigg(\ X_t \text{ or } Y_t \text{ performs more than } \log^4 n \text{ steps before time } \log^3 n\ \bigg)$$

$$\leq 2\mathrm{e}^{-\lambda\log^4 n}\exp\big\{(\mathrm{e}^\lambda - 1)\log^3 n\big\} = 2\exp\big\{(\mathrm{e}^\lambda - 1)\log^3 n - \lambda\log^4 n\big\}$$

$$= 2\exp\big\{(\log n - 1 - \log\log n\log n)\log^3 n\big\} \leq \mathrm{e}^{-c_0\sigma_n},$$

provided n is large enough. We can consider only the case where X and Y perform discrete-time random walks in which at every step one of the random walks is selected uniformly at random to move. We then just need to verify that *whp*

$$\mathbf{P}^{\mathrm{dt}}\big(\tau_{\mathrm{meet}}^{x,y} \leq 2\log^4 n\big) \leq \mathrm{e}^{-c_0\sigma_n},$$

uniformly over the pairs x and y that are at least σ_n apart.

2. Let Z_t denote the graph distance of X_t and Y_t at time t. Note that $Z_0 \geq \sigma_n$, and $Z_{t+1} - Z_t \in \{-1, 0, 1\}$ for every $t \geq 0$ (the case $Z_{t+1} - Z_t = 0$ corresponds to jumps like the

14
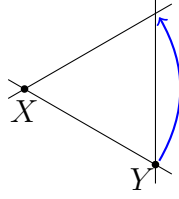
Figure 4. In a general d-regular graph the potential existence of cycles of length 3 allows the distance between the two random walks to remain constant after 1 step. In the picture, if Y moves along the blue arrow, then it remains at distance 1 to X.

one in Figure 4). This implies that the random walks cannot meet before time σ_n. Recall from Proposition 2.6 that *whp*

$$\mathtt{tx}(\mathcal{B}_{2\sigma_n}(v)) \leq 1, \qquad \text{for all } v \in V.$$

It is now enough to show that

$$q_n = \mathbf{P}^{\mathrm{dt}}\left(\tau^{x,y}_{\mathrm{meet}} \in \{\sigma_n, \dots, 2\log^4 n\}\right) \leq 4\log^8 n e^{-2c_0\sigma_n} \leq e^{-c_0\sigma_n},$$

for some positive c_0 and all n large enough. Put $\rho = \max\{t \leq \tau^{x,y}_{\mathrm{meet}} : Z_t \geq \sigma_n\}$. We rewrite the probability of interest as

$$q_n \leq 2\log^4 n \max_{r \leq 2\log^4 n} \mathbf{P}^{\mathrm{dt}}\left(\tau^{x,y}_{\mathrm{meet}} \leq 2\log^4 n, \rho = r\right).$$

We aim at providing bounds on the above probability that are uniform in $r \leq 2\log^4 n$. To that end, consider the ball of radius $2\sigma_n$ around X_ρ, which by assumption satisfies

$$\mathtt{tx}(\mathcal{B}_{2\sigma_n}(X_\rho)) \leq 1.$$

We distinguish between two cases: (1) the quantity above is 0 (and thus $\mathcal{B}_{2\sigma_n}(X_\rho)$ is a tree); (2) the tree excess of $\mathcal{B}_{2\sigma_n}(X_\rho)$ is 1.

3. First, consider the case when $\mathcal{B}_{2\sigma_n}(X_\rho)$ is a tree. Call $\mathfrak{p}$ the unique path joining X_ρ and Y_ρ within the tree, and recall that $|\mathfrak{p}| = \mathrm{dist}(X_\rho, Y_\rho) = \sigma_n$. Let us argue that at least one of the random walks does $\sigma_n/2$ steps along $\mathfrak{p}$. Indeed, suppose first that for all $t \in \{\rho, \dots, \tau^{x,y}_{\mathrm{meet}}\}$ neither X_t nor Y_t leaves the tree $\mathcal{B}_{2\sigma_n}(X_\rho)$. In order for the two random walks to meet, they need to jointly traverse the path $\mathfrak{p}$. If, on the other hand, one of the two random walks leaves $\mathcal{B}_{2\sigma_n}(X_\rho)$, then our claim is met by the definition of ρ. Here, note that when we say that the first/second random walk must have made $\sigma_n/2$ steps along $\mathfrak{p}$, we do not require that the steps are made consecutively. Nonetheless, under $\{\rho = r\}$, requiring that at least one of the two random walks does $\sigma_n/2$ steps along $\mathfrak{p}$ within time $2\log^4 n$ is equivalent to requiring that

$$\hat{\tau}_r = \inf\{u \geq r : \mathrm{dist}(X_u, Y_r) \wedge \mathrm{dist}(X_r, Y_u) \leq \tfrac{\sigma_n}{2}\}$$

satisfies $\hat{\tau}_r \leq 2\log^4 n$. We can now bound, uniformly in $r \leq 2\log^4 n$,

$$\mathbf{P}^{\mathrm{dt}}\left(\tau^{x,y}_{\mathrm{meet}} \leq 2\log^4 n, \rho = r\right) \leq \mathbf{P}^{\mathrm{dt}}\left(\hat{\tau}_r \leq 2\log^4 n, \rho = r\right) \leq 2\mathsf{P}\left(\tau_{\mathrm{abs}} \leq 2\log^4 n\right),$$
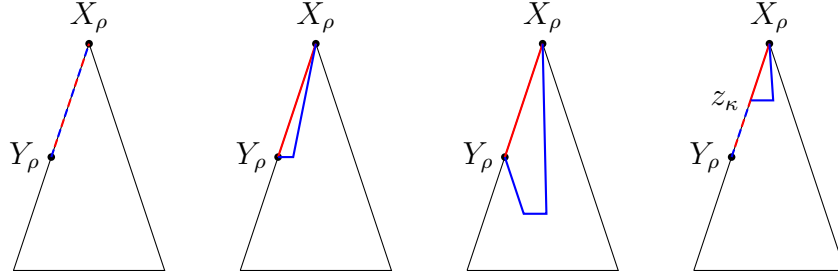
15

where τ_{abs} is the absorption time at the origin of a biased random walk on $\mathbb{Z}$, and $(W_u)_{u \geq 0}$ is such that $W_0 = \sigma_n/2$ and

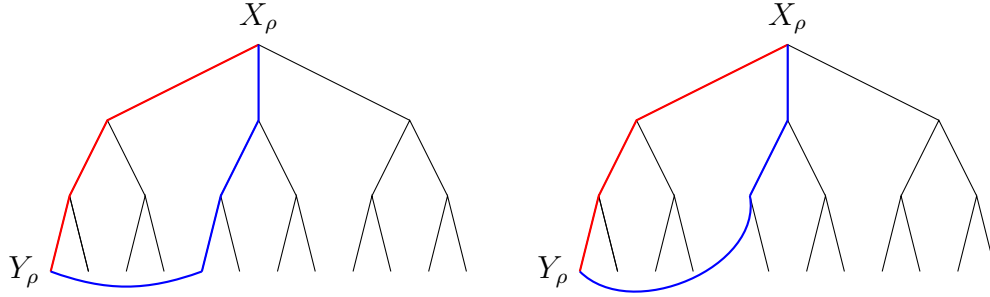$$p = \mathsf{P}(W_{t+1} - W_t = 1) = 1 - \mathsf{P}(W_{t+1} - W_t = -1) = \frac{d-1}{d}.$$

Hence, by the classical Gambler's ruin argument,

$$\mathsf{P}(\tau_{\mathrm{abs}} \leq 2 \log^4 n) \leq \mathsf{P}(\tau_{\mathrm{abs}} < \infty) = \left(\frac{1-p}{p}\right)^{\sigma_n/2} = \left(\frac{1}{d-1}\right)^{\sigma_n/2} \leq 2^{-\frac{\sigma_n}{2}}. \qquad (3.6)$$

The claim follows by choosing $c_0 \leq \frac{\log 2}{5}$.

Figure 5. A sketch of the 4 possible ways the two paths $\mathfrak{p}$ and $\tilde{\mathfrak{p}}$ described below can look like.

Figure 6. On the left $|\tilde{\mathfrak{p}}| = |\mathfrak{p}| + 1$, on the right $|\tilde{\mathfrak{p}}| = |\mathfrak{p}|$.

4. Next, consider the case when $\mathrm{tx}(\mathcal{B}_{2\sigma_n}(X_\rho)) = 1$. Then there are at most two paths joining Y_ρ to X_ρ in $\mathcal{B}_{2\sigma_n}(X_\rho)$ (see Figure 5). Call these two paths $\mathfrak{p} = (z_0 = Y_\rho, z_1, \ldots, z_{\sigma_n} = X_\rho)$ and $\tilde{\mathfrak{p}} = (\tilde{z}_0 = Y_\rho, \tilde{z}_1, \ldots, \tilde{z}_\ell = X_\rho)$, with $\sigma_n \leq \ell \leq 3\sigma_n + 1$ (the case $\ell = \sigma_n$ is exemplified in Figure 6). Let κ denote the index of the last vertex such that the paths coincide, i.e.,

$$\kappa = \sup \{i \leq \sigma_n : z_i = \tilde{z}_i\}.$$

We split the two paths as $\mathfrak{p} = (\mathfrak{p}_\kappa, \mathfrak{p}')$ and $\tilde{\mathfrak{p}} = (\mathfrak{p}_\kappa, \mathfrak{p}'')$, where $\mathfrak{p}_\kappa = (z_0, z_1, \ldots, z_\kappa)$. If $\kappa \geq \frac{1}{2}\sigma_n$, then the argument from the previous paragraph can be easily adapted as follows. We first note that, in order for the random walks to meet, either Y or X must reach $z_{\frac{1}{2}\kappa}$. If Y is the random walk that reaches this vertex, then the comparison with the birth-and-death chain is done in the same way as before. If the random walk that reaches $z_{\frac{1}{2}\kappa}$ is X,

16

then we still need to account for the time it takes X to arrive at z_κ, which can be taken care of via a union bound, yielding an additional factor $2 \log^4 n$ in the calculation. More precisely, consider the random variables

$$\tilde{\tau}_r^1 = \inf\{u \geq r\colon X_u = z_{\frac{1}{2}\kappa}\}, \qquad \tilde{\tau}_r^2 = \inf\{u \geq r\colon Y_u = z_{\frac{1}{2}\kappa}\}.$$

We can bound

$$\mathbf{P}^{\mathrm{dt}}\left(\tau_{\mathrm{meet}}^{x,y} \leq 2 \log^4 n, \rho = r\right) \leq \mathbf{P}^{\mathrm{dt}}\left(\tilde{\tau}_r^1 \wedge \tilde{\tau}_r^2 \leq 2 \log^4 n, \rho = r\right)$$

$$\leq \mathbf{P}^{\mathrm{dt}}\left(\tilde{\tau}_r^1 \leq 2 \log^4 n, \rho = r\right) + \mathbf{P}^{\mathrm{dt}}\left(\tilde{\tau}_r^2 \leq 2 \log^4 n, \rho = r\right)$$

$$\leq \mathsf{P}\left(\tau_{\mathrm{abs}} \leq 2 \log^4 n\right) + \sum_{i=1}^{2 \log^4 n} \mathsf{P}\left(\tau_{\mathrm{abs}} \leq \log^4 n - i\right)$$

$$\leq (\log^4 n + 1) 2^{-\frac{\sigma_n}{4}}.$$

For the bound in the second line, the first term follows an analogous reasoning as in the tree case. To handle the second term, we first apply union bounds on the first time the random walk X reaches z_κ, and the result follows from the strong Markov property and the same argument as in the tree case. Conversely, if $\kappa \leq \frac{1}{2}\sigma_n$, then both $\mathfrak{p}'$ and $\mathfrak{p}''$ are long paths (of size at least $\frac{1}{2}\sigma_n$). Once again, in order for the random walks to meet, at least one of them must traverse at least half of one of the paths $\mathfrak{p}'$ or $\mathfrak{p}''$. The argument follows the same lines as in the previous case, and we refrain from spelling out the details. $\square$

Proposition 3.7 (Exponential scaling of meeting time starting in stationarity). *Let $G_{d,n}(\omega)$ be a regular random graph. There exists a sequence of random variables $\Theta_d = (\Theta_d^{(n)})_{n \in \mathbb{N}}$ such that*

$$\sup_{t \geq 0} \left| \frac{\mathbf{P}(\tau_{\mathrm{meet}}^{\pi \otimes \pi} > t)}{\exp\left(-\frac{2 \Theta_d t}{n}\right)} - 1 \right| \xrightarrow{\mathbb{P}} 0, \tag{3.7}$$

and

$$\Theta_d^{(n)} \xrightarrow{\mathbb{P}} \theta_d = \frac{d-2}{d-1}.$$

Proof. We show that for the discrete-time version of the process, in which at each time step one randomly chosen walk performs a jump, we have

$$\sup_{t \geq 0} \left| \frac{\mathbf{P}^{\mathrm{dt}}(\tau_{\mathrm{meet}}^{\pi \otimes \pi} > t)}{\left(1 - \frac{\Theta_d}{n}\right)^t} - 1 \right| \xrightarrow{\mathbb{P}} 0. \tag{3.8}$$

The claim in (3.7) follows by Poissonisation. For the discrete-time process we can exploit the full power of the FVTL in Theorem 2.2. The proof comes in several steps.

1. Consider the multi-graph Γ with vertex set $V_\Gamma = \{(x,y)\colon x, y \in [n], x \neq y\} \cup \{\Delta\}$, where Δ is the merge of the vertices on the diagonal that retains the edges. More precisely,

$$((x,y),(v,w)) \in E_\Gamma \text{ if and only if } \begin{cases} x = v \text{ and } (y,w) \in E, \\ y = w \text{ and } (x,v) \in E, \end{cases}$$

and

$$(\Delta, (v,w)) \in E_\Gamma \text{ if and only if } (v,w) \in E,$$

17

with each edge adjacent to Δ having multiplicity 2. Roughly, Γ is such that every vertex except Δ has degree $2d$, while Δ has degree $4m = 2dn$. It follows that the sum of the degrees in Γ equals $2m_\Gamma = 2dn(n-1) + 2dn = 2dn^2$, and that the number of vertices is given by $N = n^2 - n + 1$. Throughout this proof, let P denote the distribution of the simple random walk on Γ, P_Γ the associated transition matrix, and V_Γ the set of vertices. The stationary distribution for this random walk π_Γ is given by

$$\pi_\Gamma(\Delta) = \frac{1}{n},$$

while, for all (x, y) with $x \neq y$,

$$\pi_\Gamma(x, y) = \frac{1}{n^2}.$$

2. Concerning the mixing time in (2.2), it is not hard to show that we can choose $T_{\mathrm{mix}} = \log^2 n$. For completeness, we provide a full proof by adapting the proof of (Cooper et al., 2010, Lemma 12) to our asynchronous setting. As shown in Friedman (2008), the second largest eigenvalue of P, denoted by $\lambda_\star$, is bounded away from 1 *whp*, from which it follows via (Levin and Peres, 2017, Corollary 12.13) that the same holds for the matrix $P_{\otimes 2} = \frac{1}{2}(P \otimes I + I \otimes P)$. Write $\lambda_{\otimes 2}$ (respectively, λ_Γ) to denote the spectral gap of the matrix $P_{\otimes 2}$ (respectively, P_Γ). By (Levin and Peres, 2017, Theorem 13.10) and with the abbreviation

$$\Phi_{\otimes 2} := \min_{S \subset V^2 \,:\, \pi_{\otimes 2}(S) \leq \frac{1}{2}} \frac{\sum_{\boldsymbol{x} \in S} \pi(\boldsymbol{x}) \sum_{\boldsymbol{y} \in S^c} P_{\otimes 2}(\boldsymbol{x}, \boldsymbol{y})}{\pi_{\otimes 2}(S)} \tag{3.9}$$

for the *conductance* of the product chain, we know that

$$\tfrac{1}{2}\Phi_{\otimes 2}^2 \leq 1 - \lambda_{\otimes 2} \leq 2\Phi_{\otimes 2}. \tag{3.10}$$

Therefore $\Phi_{\otimes 2}$ is bounded away from 0 *whp*. By the construction of Γ, we must have $\Phi_{\otimes 2} \leq \Phi_\Gamma$, where the latter is the conductance of the Markov chain P_Γ. Indeed, the graph Γ is constructed from $G \times G$ by contracting some vertices and retaining the edges. Moreover, again by (3.10) to the Markov chain P_Γ,

$$1 - \lambda_\Gamma \geq \tfrac{1}{2}\Phi_\Gamma^2 \geq \tfrac{1}{2}\Phi_{\otimes 2}^2 > 0. \tag{3.11}$$

Hence, we deduce that λ_Γ must be bounded away from 1 *whp*. Finally, by (Levin and Peres, 2017, Theorem 12.4), we conclude that $O(\log n)$ steps suffice to get (2.2).

3. Thus, Assumptions (1) and (3) in Theorem 2.2 are satisfied *whp* with $N = n^2 - n + 1$, $Q_N = P_\Gamma$, $\mu_N = \pi_\Gamma$. Moreover, choosing $x = \Delta$ we easily check the *whp* existence of a unique quasi-stationary distribution, thanks to the fact that $d \geq 3$, G is connected, and the two random walks move asynchronously. In order to use the FVTL, we are left to compute the expected number of returns to Δ before time $T_{\mathrm{mix}} \leq \log^2 n$, starting at Δ. For $t \geq 1$, put

$$R_t(\Delta) = \sum_{s=0}^{t} \mathsf{P}\big(X_s = \Delta \mid X_0 = \Delta\big).$$

We will show that

$$R_{T_{\mathrm{mix}}}(\Delta) \xrightarrow{\mathbb{P}} 1 + \frac{1 - \theta_d}{\theta_d} = \frac{1}{\theta_d} = \frac{d-1}{d-2}. \tag{3.12}$$

The term 1 comes from the fact that the random walk starts at Δ.

4. We start by giving the heuristics of the proof. Call (x, y) the vertex of Γ visited at time 1. By Proposition 2.6(ii), the probability that the edge $e = (x, y) \in E$ is $\mathrm{LTLE}(\frac{1}{5} \log_d n)$ is $1 - o(1)$. On this event, the probability that there will be a second visit to Δ before T_{mix} is upper bounded by $\frac{1}{d-1}$. A simple combination of these arguments yields an upper bound. As for the lower bound, the probability that the random walk on Γ starting at (x, y) reaches a vertex (v, w) with $\mathrm{dist}_G(v, w) \geq \sigma_n$ before reaching Δ is $\theta_d + o(1)$. Once at (v, w), the probability that the random walk visits Δ before T_{mix} is $o(1)$. Hence, up to $o(1)$ corrections, $R_{T_{\mathrm{mix}}}(\Delta)$ can be estimated by the expectation of a geometric random variable with parameter $\theta_d + o(1)$, representing the first excursion of the random walk that hits (v, w) with $\mathrm{dist}_G(v, w) \geq \sigma_n$ before returning to Δ. **5.** With this heuristics in mind, we are ready to provide the proof of (3.12). We will do so by first assuming that

$$\mathsf{P}\left(\tfrac{1}{5} \log_d n \leq \tau_\Delta^+ \leq T_{\mathrm{mix}} \mid X_0 = \Delta\right) \xrightarrow{\mathbb{P}} 0 \tag{3.13}$$

and

$$\mathsf{P}\left(\tau_\Delta^+ \leq \tfrac{1}{5} \log_d n \mid X_0 = \Delta\right) \xrightarrow{\mathbb{P}} 1 - \theta_d. \tag{3.14}$$

Write

$$R_{T_{\mathrm{mix}}}(\Delta) = 1 + \sum_{s=1}^{T_{\mathrm{mix}}} \mathsf{P}\left(X_s = \Delta \mid X_0 = \Delta\right)$$

$$= 1 + \sum_{r=1}^{T_{\mathrm{mix}}} \sum_{s=r}^{T_{\mathrm{mix}}} \mathsf{P}\left(X_s = \Delta, \tau_\Delta^+ = r \mid X_0 = \Delta\right)$$

$$= 1 + \sum_{r=1}^{T_{\mathrm{mix}}} \sum_{s=r}^{T_{\mathrm{mix}}} \mathsf{P}\left(X_s = \Delta \mid X_r = \Delta\right) \mathsf{P}\left(\tau_\Delta^+ = r \mid X_0 = \Delta\right) \tag{3.15}$$

$$= 1 + \sum_{r=1}^{T_{\mathrm{mix}}} \mathsf{P}\left(\tau_\Delta^+ = r \mid X_0 = \Delta\right) \sum_{s=0}^{T_{\mathrm{mix}}-r} \mathsf{P}\left(X_s = \Delta \mid X_0 = \Delta\right)$$

$$= 1 + \sum_{r=1}^{T_{\mathrm{mix}}} \mathsf{P}\left(\tau_\Delta^+ = r \mid X_0 = \Delta\right) R_{T_{\mathrm{mix}}-r}(\Delta).$$

To establish the upper bound in (3.12), observe that (3.15) yields

$$R_{T_{\mathrm{mix}}}(\Delta) \leq 1 + \mathsf{P}\left(\tau_\Delta^+ \leq T_{\mathrm{mix}} \mid X_0 = \Delta\right) R_{T_{\mathrm{mix}}}(\Delta),$$

which yields

$$R_{T_{\mathrm{mix}}}(\Delta) \leq \left(1 - \mathsf{P}\left(\tau_\Delta^+ \leq T_{\mathrm{mix}} \mid X_0 = \Delta\right)\right)^{-1},$$

and the claim follows by combining (3.13) and (3.14).

6. To establish the lower bound in (3.12), observe that (3.15) yields

$$R_{T_{\mathrm{mix}}}(\Delta) \geq 1 + \sum_{r=1}^{\frac{1}{5} \log_d n} \mathsf{P}\left(\tau_\Delta^+ = r \mid X_0 = \Delta\right) R_{T_{\mathrm{mix}}-r}(\Delta)$$

$$\geq 1 + \mathsf{P}\left(\tau_\Delta^+ \leq \tfrac{1}{5} \log_d n \mid X_0 = \Delta\right) R_{T_{\mathrm{mix}}-\frac{1}{5} \log_d n}(\Delta).$$

19

By iterating the estimate above, we obtain

$$R_{T_{\mathrm{mix}}}(\Delta) \geq \sum_{i=0}^{5T_{\mathrm{mix}}/\log_d n} \mathsf{P}\left(\tau_\Delta^+ \leq \tfrac{1}{5}\log_d n \mid X_0 = \Delta\right)^i = \frac{1}{\theta_d} - o_{\mathbb{P}}(1), \tag{3.16}$$

where the last equality follows from (3.14). This concludes the verification of the hypotheses of Theorem 2.2.

To conclude the proof, we still need to verify (3.13) and (3.14). We start with (3.13). By Proposition 2.6(iii), we have

$$\mathsf{P}\left(X_1 \in \mathrm{LTLE}\left(\tfrac{1}{5}\log_d n\right) \mid X_0 = \Delta\right) \xrightarrow{\mathbb{P}} 1. \tag{3.17}$$

Indeed, by construction, the random walk on Γ starting at Δ after a single step visits a $(x,y) \in E$ uniformly at random. Assume that $X_1 = (x,y) \in \mathrm{LTLE}\left(\tfrac{1}{5}\log_d n\right)$, and let $\tilde{\tau}$ denote the first time at which the process on Γ is found at some (w,z) with $\mathrm{dist}_G(w,z) = \sigma_n$. Then, up to time $\tau_\Delta \wedge \tilde{\tau}$, the random walk $(X_t)_{t \geq 1}$ can be coupled with a biased random walk on $\mathbb{Z}$ (given by the distance between the two vertices) that starts from 1, jumps to the right (respectively, left) with probability $\frac{d-1}{d}$ (respectively, $\frac{1}{d}$). Call this random walk $(W_t)_{t \geq 0}$, and let τ_j, $j \in \mathbb{Z}$, be its first hitting of j. Then a gambler's ruin argument yields

$$\mathsf{E}_1\left(\tau_0 \wedge \tau_{\sigma_n}\right) \leq \frac{d\sigma_n}{d-2},$$

which implies

$$\mathsf{P}_1\left(\tau_0 \wedge \tau_{\sigma_n} > \tfrac{1}{5}\log_d n - 1\right) \leq \mathsf{P}_1\left(\tau_0 \wedge \tau_{\sigma_n} > \sigma_n^3\right) \leq \frac{1}{\sigma_n} \tag{3.18}$$

for all n sufficiently large. Compute

$$\mathsf{P}\left(\tau_\Delta^+ \wedge \tilde{\tau} \leq \tfrac{1}{5}\log_d n \mid X_0 = \Delta\right)$$

$$\geq \sum_{(x,y) \in (V_\Gamma \setminus \{\Delta\}) \cap \mathrm{LTLE}\left(\tfrac{1}{5}\log_d n\right)} \mathsf{P}_\Delta\left(X_1 = (x,y)\right)\mathsf{P}_{(x,y)}\left(\tau_\Delta^+ \wedge \tilde{\tau} \leq \tfrac{1}{5}\log_d n - 1\right)$$

$$\overset{(3.18)}{\geq} \left(1 - \frac{1}{\sigma_n}\right)\mathsf{P}\left(X_1 \in \mathrm{LTLE}\left(\tfrac{1}{5}\log_d n\right) \mid X_0 = \Delta\right).$$

Note that the last line above converges to 1 in probability as $n \to \infty$ (see (3.17)), which concludes the verification of (3.13).

We are left with proving (3.14). In this case, we can assume once again that $X_1 \in \mathrm{LTLE}\left(\tfrac{1}{5}\log_d n\right)$, which in particular implies that, up to time $\tfrac{1}{5}\log_d n$, the random walk can be perfectly coupled with a biased random walk on $\mathbb{Z}$ that starts at 1 and jumps to the right (respectively, left) with probability $\frac{d-1}{d}$ (respectively, $\frac{1}{d}$). The claim follows by noting that

$$\mathsf{P}\left(\tau_\Delta^+ \leq \tfrac{1}{5}\log_d n \mid X_0 = \Delta\right) = o_{\mathbb{P}}(1) + \mathsf{P}_1\left(\tau_0 \leq \tfrac{1}{5}\log_d n - 1\right)$$

$$= o_{\mathbb{P}}(1) + \mathsf{P}_1\left(\tau_0 < \infty\right) = o(1) + \frac{1}{d-1},$$

which concludes the verification of (3.14). $\qquad\square$

Lemma 3.8 (Tail of meeting time of distant random walks). *Let $G_{d,n}(\omega)$ be a regular random graph and consider two independent random walks. Then*

$$\max_{x,y:\ \mathrm{dist}(x,y)\geq\sigma_n}\ \sup_{t\geq0}\left|\mathbf{P}\big(\tau_{\mathrm{meet}}^{x,y}>t\big)-\mathrm{e}^{-\frac{2t\theta_d}{n}}\right|\xrightarrow{\ \mathbb{P}\ }0.\tag{3.19}$$

Proof. We need to deduce from Proposition 3.7 an (additive) analogous result when the two random walks start far away. For that we use Lemma 3.6.

For simplicity we work again in discrete time, so that what we actually need to show is

$$\max_{x,y:\ \mathrm{dist}(x,y)\geq\sigma_n}\ \max_{t\geq0}\left|\mathbf{P}^{\mathrm{dt}}(\tau_{\mathrm{meet}}^{x,y}>t)-\left(1-\frac{\theta_d}{n}\right)^t\right|\xrightarrow{\ \mathbb{P}\ }0.\tag{3.20}$$

Fix x and y such that $\mathrm{dist}(x,y)\geq\sigma_n$, and let $\mathcal{E}$ be the event that the two random walks meet before the mixing time T_{mix}. Since $T_{\mathrm{mix}}\leq\log^3 n$, Lemma 3.6 states that $\mathsf{P}(\mathcal{E})\leq\mathrm{e}^{-c_0\sigma_n}$ *whp* for every pair (x,y) as above. Hence, if $t\leq T_{\mathrm{mix}}$, then

$$\min_{x,y:\ \mathrm{dist}(x,y)\geq\sigma_n}\mathbf{P}^{\mathrm{dt}}\big(\tau_{\mathrm{meet}}^{x,y}>t\big)\geq 1-\mathrm{e}^{-c_0\sigma_n}.\tag{3.21}$$

On the other hand, if $t>T_{\mathrm{mix}}$, then by definition of T_{mix},

$$\max_{(x,y)\in V_\Gamma}\big|\mathsf{P}\big(X_t^x=v,Y_t^y=w\big)-\pi_\Gamma(v,w)\big|\leq\frac{1}{n^6}.\tag{3.22}$$

Therefore

$$\mathbf{P}^{\mathrm{dt}}\big(\tau_{\mathrm{meet}}^{x,y}>t\big)=\sum_{(v,w)\in V_\Gamma\setminus\{\Delta\}}\mathsf{P}\big(\tau_\Delta^{(x,y)}>T_{\mathrm{mix}},X_{T_{\mathrm{mix}}}^x=v,Y_{T_{\mathrm{mix}}}^y=w\big)\mathsf{P}\big(\tau_\Delta^{(v,w)}>t-T_{\mathrm{mix}}\big)$$

$$\leq\sum_{(v,w)\in V_\Gamma\setminus\{\Delta\}}\mathsf{P}\big(X_{T_{\mathrm{mix}}}^x=v,Y_{T_{\mathrm{mix}}}^y=w\big)\mathsf{P}\big(\tau_\Delta^{(v,w)}>t-T_{\mathrm{mix}}\big)$$

$$\leq\sum_{(v,w)\in V_\Gamma\setminus\{\Delta\}}\pi_\Gamma(v,w)\mathsf{P}\big(\tau_\Delta^{(v,w)}>t-T_{\mathrm{mix}}\big)+\frac{1}{n^6}$$

$$\leq\mathsf{P}\big(\tau_\Delta^{\pi_\Gamma}>t-T_{\mathrm{mix}}\big)+\frac{1}{n^4}.$$

On the other hand,

$$\mathbf{P}^{\mathrm{dt}}\big(\tau_{\mathrm{meet}}^{x,y}>t\big)$$

$$=\sum_{(v,w)\in V_\Gamma\setminus\Delta}\mathsf{P}\big(\tau_\Delta^{(x,y)}>T_{\mathrm{mix}},X_{T_{\mathrm{mix}}}^x=v,Y_{T_{\mathrm{mix}}}^y=w\big)\mathsf{P}\big(\tau_\Delta^{(v,w)}>t-T_{\mathrm{mix}}\big)$$

$$\geq\left(\sum_{(v,w)\in V_\Gamma\setminus\{\Delta\}}\mathsf{P}\big(X_{T_{\mathrm{mix}}}^x=v,Y_{T_{\mathrm{mix}}}^y=w\big)\mathsf{P}\big(\tau_\Delta^{(v,w)}>t-T_{\mathrm{mix}}\big)\right)-\mathsf{P}\big(\tau_{\mathrm{meet}}^{x,y}\leq T_{\mathrm{mix}}\big)$$

$$\overset{(3.22)}{\geq}\sum_{(v,w)\in V_\Gamma\setminus\{\Delta\}}\pi_\Gamma(v,w)\mathsf{P}\big(\tau_\Delta^{(v,w)}>t-T_{\mathrm{mix}}\big)-\frac{1}{n^6}-\mathsf{P}\big(\tau_{\mathrm{meet}}^{x,y}\leq T_{\mathrm{mix}}\big)$$

$$\overset{(3.21)}{=}\mathsf{P}\big(\tau_\Delta^{\pi_\Gamma}>t-T_{\mathrm{mix}}\big)-\frac{1}{n^4}-\mathrm{e}^{-c_0\sigma_n}.$$

21

This, when combined with Propostition 3.7, implies the claim because

$$\left|\left(1-\frac{\theta_d}{n}\right)^t - \left(1-\frac{\theta_d}{n}\right)^{t-T_{\text{mix}}}\right| = \left(1-\frac{\theta_d}{n}\right)^t\left|1-\left(1-\frac{\theta_d}{n}\right)^{-T_{\text{mix}}}\right| \le e^{-\frac{t\theta_d}{n}}\left|1-\left(1-\frac{\theta_d}{n}\right)^{-T_{\text{mix}}}\right|$$

converges to zero as $n \to \infty$, uniformly in $t \ge 0$. $\qquad\square$

Proposition 3.9 (Tail of meeting time of adjacent random walks on a LTLE). *Let $G_{d,n}(\omega)$ be a regular random graph, let $e = (x, y)$ be in $\mathrm{LTLE}\left(\frac{1}{5}\log_d n\right)$, and consider two independent random walks starting from the endvertices of e. Then, for all t_n such that $\lim_{n\to\infty} t_n = \infty$ and $\lim_{n\to\infty} t_n/n = s \ge 0$,*

$$\left|\mathbf{P}\left(\tau_{\text{meet}}^{x,y} > t_n\right) - \theta_d\, e^{-2s\theta_d}\right| \xrightarrow{\mathbb{P}} 0.$$

Proof. Note that it is enough to assume that $t \ge \sigma_n^3$, since otherwise the claim follows from Lemma 3.4. Let $\tau_{\text{far}}^{x,y}$ be the first time at which the two random walks are at distance at least σ_n, and $\tau_{\text{exit}}^{x,y}$ the first time at which one of the two random walks exits the locally tree-like neighbourhood of radius $\frac{1}{5}\log_d n$ of e. Put $\tau_0^{x,y} = \tau_{\text{meet}}^{x,y} \wedge \tau_{\text{far}}^{x,y} \wedge \tau_{\text{exit}}^{x,y}$. For simplicity, we start by analysing the discrete-time process. Up to time $\tau_0^{x,y}$, the process can be coupled to a biased random walk $(W_t)_{t\ge0}$ with $W_0 = 1$ and

$$p = \mathsf{P}\left(W_{t+1} - W_t = 1\right) = 1 - \mathsf{P}\left(W_{t+1} - W_t = -1\right) = \frac{d-1}{d}.$$

In particular,

$$\mathbf{E}^{\text{dt}}\left[\tau_0^{x,y}\right] \le \frac{d\sigma_n}{d-2},$$

and therefore

$$\begin{aligned}
\mathbf{P}^{\text{dt}}\left(\tau_0^{x,y} = \tau_{\text{exit}}^{x,y}\right) &\le \mathbf{P}^{\text{dt}}\left(\tau_{\text{exit}}^{x,y} \le \sigma_n^3\right) + \mathbf{P}^{\text{dt}}\left(\tau_0^{x,y} > \sigma_n^3\right) \\
&\le 0 + \frac{1}{\sigma_n^3}\frac{d\sigma_n}{d-2} = o(1),
\end{aligned} \tag{3.23}$$

where the last term follows from the Markov inequality. Similarly, by a gambler ruin argument,

$$\mathbf{P}^{\text{dt}}\left(\tau_{\text{far}}^{x,y} > \sigma_n^3\right) = o(1),$$

and therefore

$$\mathbf{P}\left(\tau_{\text{meet}}^{x,y} > t_n\right) = \mathbf{P}\left(\tau_{\text{meet}}^{x,y} > t_n,\, \tau_0^{x,y} = \tau_{\text{far}}^{x,y} \le \sigma_n^3\right) + o(1). \tag{3.24}$$

The probability in the right-hand side of (3.24) can be written, in the discrete-time setting, as

$$\mathbf{P}^{\mathrm{dt}}\left(\tau_{\mathrm{meet}}^{x,y} > t_n, \, \tau_0^{x,y} = \tau_{\mathrm{far}}^{x,y} \le \sigma_n^3\right)$$

$$= \sum_{u=0}^{\sigma_n^3} \sum_{(v,w): \, \mathrm{dist}(v,w) > \sigma_n} \mathbf{P}^{\mathrm{dt}}\left(\tau_0^{x,y} = \tau_{\mathrm{far}}^{x,y} = u, \, (X_{\tau_0^{x,y}}, Y_{\tau_0^{x,y}}) = (v,w)\right) \mathbf{P}\left(\tau_{\mathrm{meet}}^{v,w} > t_n - u\right)$$

$$\sim \sum_{u=0}^{\sigma_n^3} \sum_{(v,w): \, \mathrm{dist}(v,w) > \sigma_n} \mathbf{P}^{\mathrm{dt}}\left(\tau_0^{x,y} = \tau_{\mathrm{far}}^{x,y} = u, \, (X_{\tau_0^{x,y}}, Y_{\tau_0^{x,y}}) = (v,w)\right) \left(1 - \frac{\theta_d}{n}\right)^{t_n - u}$$

$$= \left(1 - \frac{\theta_d}{n}\right)^{t_n} \sum_{u=0}^{\sigma_n^3} \mathbf{P}^{\mathrm{dt}}\left(\tau_0^{x,y} = \tau_{\mathrm{far}}^{x,y} = u\right) \left(1 - \frac{\theta_d}{n}\right)^{-u}$$

$$\sim \left(1 - \frac{\theta_d}{n}\right)^{t_n} \sum_{u=0}^{\sigma_n^3} \mathbf{P}^{\mathrm{dt}}\left(\tau_0^{x,y} = \tau_{\mathrm{far}}^{x,y} = u\right) \sim \mathrm{e}^{-s\theta_d} \mathbf{P}^{\mathrm{dt}}\left(\tau_0^{x,y} = \tau_{\mathrm{far}}^{x,y} \le \sigma_n^3\right) \sim \theta_d \, \mathrm{e}^{-s\theta_d} \qquad whp.$$

In order to pass to the continuous-time setting, it is enough to realise that, for all $\varepsilon \in (0, s)$,

$$\mathbf{P}\left(\tau_{\mathrm{meet}}^{x,y} > t_n\right) = \sum_{u=0}^{\infty} \mathsf{P}\left(\mathrm{Poisson}(2t_n) = u\right) \mathbf{P}^{\mathrm{dt}}\left(\tau_{\mathrm{meet}}^{x,y} > u\right)$$

$$= \sum_{u=2(s-\varepsilon)n \vee 0}^{2(s+\varepsilon)n} \mathsf{P}\left(\mathrm{Poisson}(2sn) = u\right) \mathbf{P}^{\mathrm{dt}}\left(\tau_{\mathrm{meet}}^{x,y} > u\right) + o_{\mathbb{P}}(1)$$

$$= \theta_d \, \mathrm{e}^{-2s\theta_d} + O_{\mathbb{P}}(\varepsilon),$$

after which the claim follows by taking the limit $\varepsilon \downarrow 0$. $\qquad \square$

The next two statements turn the result in Proposition 3.9 into a statement about the expected density of discordant edges at time $O(n)$ for the voter model.

Corollary 3.10 (Discordant edges at small distances). *Let $G_{d,n}(\omega)$ be a d-regular random graph, let $e = (x, y)$ be in $\mathrm{LTLE}(\frac{1}{5}\log_d n)$, and consider the voter model with initial density $u \in (0, 1)$. Then, for all t_n such that $\lim_{n\to\infty} t_n = \infty$ and $\lim_{n\to\infty} t_n/n = s \in [0, \infty)$,*

$$\left|\mathbf{P}_u\left(e \in D_{t_n}\right) - 2u(1-u)\,\theta_d\,\mathrm{e}^{-2s\theta_d}\right| \xrightarrow{\mathbb{P}} 0.$$

Proof. The claim follows from Proposition 3.9 via duality. $\qquad \square$

Proposition 3.11 (Expected density of discordant edges). *Let $G_{d,n}(\omega)$ be a d-regular random graph, and consider the voter model with initial density $u \in (0, 1)$. Then, for all t_n such that $\lim_{n\to\infty} t_n/n = s \in [0, \infty)$,*

$$\left|\mathbf{E}_u[\mathcal{D}_{t_n}] - 2u(1-u)\,\theta_d\,\mathrm{e}^{-2s\theta_d}\right| \xrightarrow{\mathbb{P}} 0.$$

Proof. The claim follows from Corollary 3.10 and the same argument as in the proof of Proposition 3.5. $\qquad \square$

4. Beyond the expectation

In this section we prove Theorem 1.2. We start with a proposition that implies the first statement of the theorem.

Proposition 4.1 (Concentration for worst case initialization). *Consider times t_n such that $t_n \to \infty$ and $t_n/n \to 0$. Then, for any $\varepsilon > 0$,*

$$\sup_{\xi \in \{0,1\}^V} \mathbf{P}_\xi \left(|\mathcal{D}_{t_n} - \mathbf{E}_\xi [\mathcal{D}_{t_n}]| > \varepsilon \right) \xrightarrow{\mathbb{P}} 0.$$

Proof. Putting $E_\star = \mathrm{LTLE}(\frac{1}{5} \log_d n) \subset E$ and recalling that $|E_\star| = m - o(n)$ *whp* by Proposition 2.6 (ii), we have

$$
\begin{aligned}
\mathbf{E}[|D_t|^2] &= \sum_{e,e' \in E} \mathbf{P}\left(e, e' \in D_t\right) = \sum_{e,e' \in E_\star} \mathbf{P}\left(e, e' \in D_t\right) + o(n^2) \\
&= \sum_{e \in E_\star} \left[\sum_{e' \in E_\star,\, \mathrm{dist}(e,e')>\sigma_\mathrm{n}} \mathbf{P}\left(e, e' \in D_t\right) + O(d^{\sigma_n}) \right] + o(n^2),
\end{aligned}
\tag{4.1}
$$

with σ_n as in (3.5). Fix $e = (x,y), e' = (x',y') \in E_\star$ such that $\mathrm{dist}(e,e') > \sigma_n$, and consider the event

$$\mathcal{E} = \mathcal{E}_{e,e'} := \left\{ \tau_\mathrm{meet}^{x,x'} \wedge \tau_\mathrm{meet}^{x,y'} \wedge \tau_\mathrm{meet}^{y,x'} \wedge \tau_\mathrm{meet}^{y,y'} > t \right\}.$$

We observe that, on the event $\mathcal{E}$, the events $\{e \in D_t\}$ and $\{e' \in D_t\}$ are negatively correlated. Indeed, denote by σ_e (respectively, $\sigma_{e'}$) a realisation of length t of the two independent random walk trajectories starting at x and y (respectively, x' and y'). Let $\mathcal{H}(\sigma_{e'})$ denote the set of possible realisations of $\sigma_{e'}$ that never meet the trajectory σ_e. We then have

$$
\begin{aligned}
\mathbf{P}_\xi(e, e' \in D_t, \mathcal{E}) &= \sum_{\sigma_e \in \{e \in D_t\}} \sum_{\sigma_{e'} \in \{e' \in D_t\} \cap \mathcal{H}(\sigma_e)} \mathbf{P}_\xi(\sigma_e) \mathbf{P}_\xi\left(\sigma_{e'} \mid \sigma_e\right) \\
&= \sum_{\sigma_e \in \{e \in D_t\}} \sum_{\sigma_{e'} \in \{e' \in D_t\} \cap \mathcal{H}(\sigma_e)} \mathbf{P}_\xi(\sigma_e) \mathbf{P}_\xi\left(\sigma_{e'}\right) \\
&\leq \sum_{\sigma_e \in \{e \in D_t\}} \sum_{\sigma_{e'} \in \{e' \in D_t\}} \mathbf{P}_\xi(\sigma_e) \mathbf{P}_\xi\left(\sigma_{e'}\right) = \mathbf{P}_\xi(e \in D_t)\, \mathbf{P}_\xi(e' \in D_t),
\end{aligned}
\tag{4.2}
$$

which gives the claimed negative dependence.

Furthermore, by Lemma 3.8 , we have $\mathbf{P}(\mathcal{E}^c) = o_\mathbb{P}(1)$, which together with (4.2) guarantees that

$$\mathbf{P}_\xi\left(e, e' \in D_t\right) \leq \mathbf{P}_\xi\left(e, e' \in D_t, \mathcal{E}\right) + \mathbf{P}_\xi(\mathcal{E}^c) \leq \mathbf{P}_\xi(e \in D_t)\mathbf{P}_\xi\left(e' \in D_t\right) + o_\mathbb{P}(1).$$

Inserting the latter inequality into (4.1), we get

$$\mathbf{E}_\xi[|D_t|^2] = \sum_{e\in E_\star}\left[\sum_{e'\in E_\star,\,\mathrm{dist}(e,e')>\sigma_n}\mathbf{P}_\xi\left(e,e'\in D_t\right)+O(d^{\sigma_n})\right]+o(n^2)$$

$$\leq \sum_{e\in E_\star}\left[\sum_{e'\in E_\star,\,\mathrm{dist}(e,e')>\sigma_n}\left[\mathbf{P}_\xi(e\in D_t)\mathbf{P}_\xi\left(e'\in D_t\right)+o_\mathbb{P}(1)\right]+O(d^{\sigma_n})\right]+o(n^2)$$

$$\leq \mathbf{E}_\xi[|D_t|]^2+o_\mathbb{P}(n^2)+o(n^2)=(1+o_\mathbb{P}(1))\mathbf{E}_\xi[|D_t|]^2.$$

Therefore, by Chebyshev's inequality,

$$\mathbf{P}_\xi\left(\left||D_t|-\mathbf{E}_\xi|D_t|\right|>\varepsilon m\right)\leq\mathbf{P}_\xi\left(\left||D_t|-\mathbf{E}_\xi|D_t|\right|>\varepsilon\mathbf{E}_\xi|D_t|\right)\leq\frac{\mathrm{Var}_\xi(|D_t|)}{\varepsilon^2\mathbf{E}_\xi[|D_t|]^2}=o_\mathbb{P}(1).$$

$\square$

We next consider times t_n such that $t_n/n\to s\in(0,\infty)$ and use results from Chen et al. (2016). Let

$$\gamma_n=\mathbf{E}[\tau_{\mathrm{meet}}^{\pi\otimes\pi}] \tag{4.3}$$

be the expected meeting time of two independent continuous-time simple random walks with uniformly chosen initial positions. Chen et al. (2016) (see Theorems 2.1 and 2.2 therein) provide conditions on the underlying sequence of graphs under which the proportion of type-1 vertices $(\mathcal{B}_{\gamma_n t}^n)_{t\geq0}$ converges as a process in the space of càdlàg paths to the Fisher-Wright diffusion $(\bar{\mathcal{B}}_t)_{t\geq0}$ for γ_n as in Eq. (4.3). In the case of d-regular random graphs,

$$\gamma_n=\frac{1}{2}\theta_d^{-1}\,n+o_\mathbb{P}(1), \tag{4.4}$$

and the requirements in Chen et al. (2016) are easily verified *whp*.

The following lemma is the central estimate that we need in order to establish convergence of the proportion of discordant edges.

Lemma 4.2 (Linking density of discordant edges to density of type-1)**.** *Let s_n converge to a finite value. Then*

$$\mathbf{E}_u\left[\left|\mathcal{D}_{s_n\gamma_n}-2\theta_d\mathcal{B}_{s_n\gamma_n}(1-\mathcal{B}_{s_n\gamma_n})\right|\right]\xrightarrow{\mathbb{P}}0.$$

Before proving this lemma, we conclude the proof of Theorem 1.2.

Proof of Theorem 1.2 (ii). Recall that we are assuming that $t_n/n\to s$. Consider $s_n=t_n/\gamma_n$ and note that $\lim_{n\to\infty}s_n=2s\,\theta_d$, due to (4.4). Hence Lemma 4.2 yields

$$\mathbf{E}_u\left[\left|\mathcal{D}_{t_n}-2\theta_d\mathcal{B}_{t_n}(1-\mathcal{B}_{t_n})\right|\right]\xrightarrow{\mathbb{P}}0. \tag{4.5}$$

Moreover, since $x\mapsto x(1-x)$ is continuous and so is the Fisher-Wright diffusion $(\bar{\mathcal{B}}_t)_{t\geq0}$, Proposition A.2 implies that

$$2\theta_d\mathcal{B}_{t_n}(1-\mathcal{B}_{t_n})\quad\text{converges in distribution to}\quad 2\theta_d\bar{\mathcal{B}}_{2s\,\theta_d}(1-\bar{\mathcal{B}}_{2s\,\theta_d}).$$

In view of (4.5), the same holds *whp* for $\mathcal{D}_{t_n}$, which concludes the proof. $\square$

To prove Lemma 4.2, we use the following concentration estimate derived in Chen et al. (2016).

Lemma 4.3 (Concentration on arbitrary connected graphs). *Fix a finite connected graph G on n vertices and m edges and let $(\eta_t)_{t\geq 0}$ denote the voter model evolving on G. For any $0 \leq s \leq t$,*

$$\sup_{\eta_0} \left| \mathbf{E}_{\eta_0}[\mathcal{D}_t] - 2\mathbf{P}[\tau^e > s] \frac{1}{n^2} \sum_x \eta_0(x) \sum_x (1 - \eta_0(x)) \right| \leq \mathbf{P}[\tau^e \in (s,t]] + 4\mathbf{P}[\tau^e > s] d_{\mathrm{TV}}(t-s),$$

where τ^e denotes the meeting time of two random walks starting from the two vertices at the end of a uniformly chosen edge of G, and d_{TV} is defined as in (2.3).

The proof of Lemma 4.3 is provided in (Chen et al., 2016, Lemma 6.1). Note that it applies to any deterministic connected graph on which the voter dynamics takes place.

We will also need the asymptotic estimate

$$\mathbf{P}[\tau^e \geq k \log^2 n] \xrightarrow{\mathbb{P}} \theta_d, \qquad k \in \mathbb{N}, \tag{4.6}$$

which follows from Proposition 3.9.

Proof of Lemma 4.2. Fix $\delta_n = 2\log^2 n$, and split

$$\mathbf{E}_u\left[\left|\mathcal{D}_{s_n\gamma_n} - 2\theta_d \mathcal{B}_{s_n\gamma_n}(1 - \mathcal{B}_{s_n\gamma_n})\right|\right] \leq \mathbf{E}_u\left[\left|\mathcal{D}_{s_n\gamma_n} - 2\theta_d \mathcal{B}_{s_n\gamma_n-\delta_n}(1 - \mathcal{B}_{s_n\gamma_n-\delta_n})\right|\right]$$
$$+ 2\theta_d \mathbf{E}_u\left[\left|\mathcal{B}_{s_n\gamma_n-\delta_n}(1 - \mathcal{B}_{s_n\gamma_n-\delta_n}) - \mathcal{B}_{s_n\gamma_n}(1 - \mathcal{B}_{s_n\gamma_n})\right|\right]. \tag{4.7}$$

We will show that each of the expectations in the right-hand side above converges to zero.

1. Apply the Markov property at time $s_n\gamma_n - \delta_n$, to obtain

$$\mathbf{E}_u\left[\left|\mathcal{D}_{s_n\gamma_n} - 2\theta_d \mathcal{B}_{s_n\gamma_n-\delta_n}(1 - \mathcal{B}_{s_n\gamma_n-\delta_n})\right|\right] \leq \sup_{\eta_0} \mathbf{E}_{\eta_0}\left[\left|\mathcal{D}_{\delta_n} - 2\theta_d \mathcal{B}_0(1 - \mathcal{B}_0)\right|\right]$$
$$\leq \sup_{\eta_0} \mathbf{E}_{\eta_0}\left[\left|\mathcal{D}_{\delta_n} - \mathbf{E}_{\eta_0}[\mathcal{D}_{\delta_n}]\right|\right] + \sup_{\eta_0}\left|\mathbf{E}_{\eta_0}[\mathcal{D}_{\delta_n}] - 2\theta_d\frac{1}{n^2}\sum_x \eta_0(x)\sum_x(1-\eta_0(x))\right|. \tag{4.8}$$

The first expectation in the right-hand side of (4.8) converges to zero by Proposition 4.1 and the fact that $|\mathcal{D}_{\delta_n} - \mathbf{E}_{\eta_0}[\mathcal{D}_{\delta_n}]|$ is bounded by one. To control the second expectation in (4.8), apply Lemma 4.3 with $t = \delta_n = 2\log^2 n$ and $s = \log^2 n$, to obtain

$$\sup_{\eta_0}\left|\mathbf{E}_{\eta_0}[\mathcal{D}_{\delta_n}] - 2\theta_d\frac{1}{n^2}\sum_x \eta_0(x)\sum_x(1-\eta_0(x))\right|$$
$$\leq \sup_{\eta_0}\left|\mathbf{E}_{\eta_0}[\mathcal{D}_{\delta_n}] - 2\mathbf{P}[\tau^e \geq \log^2 n]\frac{1}{n^2}\sum_x \eta_0(x)\sum_x(1-\eta_0(x))\right| + 2\left|\mathbf{P}[\tau^e \geq \log^2 n] - \theta_d\right|$$
$$\leq \mathbf{P}[\tau^e \in (\log^2 n, 2\log^2 n]] + 4\mathbf{P}[\tau^e > \log^2 n]d_{TV}(\log^2 n) + 2\left|\mathbf{P}[\tau^e \geq \log^2 n] - \theta_d\right|.$$

Note that both

$$\mathbf{P}[\tau^e \in (\log^2 n, 2\log^2 n]] = \mathbf{P}[\tau^e > 2\log^2 n] - \mathbf{P}[\tau^e > \log^2 n]$$

and

$$\left|\mathbf{P}[\tau^e \geq \log^2 n] - \theta_d\right|$$

26

converge to zero *whp* by (4.6), while $d_{\mathrm{TV}}(\log^2 n)$ goes to zero thanks to Proposition 2.6 (iv). Hence also the second expectation in (4.8) vanishes, and so we have

$$\mathbf{E}_u\left[\left|\mathcal{D}_{t_n\gamma_n} - 2\theta_d \mathcal{B}_{t_n\gamma_n-\delta_n}(1 - \mathcal{B}_{t_n\gamma_n-\delta_n})\right|\right] \xrightarrow{\mathbb{P}} 0. \tag{4.9}$$

2. It remains to show that the second term in the right-hand side of Eq. (4.7) vanishes *whp*, for which we argue by continuity in the proper topology. For an arbitrary $T > 0$, let $[0, T] \times \mathcal{D}[0, T]$ denote the product of $[0, T]$ with the càdlàg space $\mathcal{D}[0, T]$ endowed with the J_1-Skorohod topology. For a given $s > 0$ and an arbitrary evaluation function $h\colon [0, T] \times \mathcal{D}[0, T] \to \mathbb{R}$ defined via $h(t, \phi) = \phi(t)$, we consider the incremental function $\tilde{h}\colon [0, T]^2 \times \mathcal{D}[0, T] \to \mathbb{R}$ given by

$$\tilde{h}(t, u, \phi) = |h(t, \phi) - h(u, \phi)|,$$

and note that, by continuity of the modulus function and Lemma A.1, all points in $[0, T]^2 \times \mathcal{C}[0, T]$ are continuity points of $\tilde{h}$. In particular, since $\frac{\delta_n}{\gamma_n} \to 0$ and with $s = \lim_{n\to\infty} s_n < T$, we have that

$$\mathbf{E}_u\left[\left|\mathcal{B}_{s_n\gamma_n}(1 - \mathcal{B}_{s_n\gamma_n}) - \mathcal{B}_{s_n\gamma_n-\delta_n}(1 - \mathcal{B}_{s_n\gamma_n-\delta_n})\right|\right] = \mathbf{E}_u\left[\left|h\left(s_n, (\mathcal{B}_{u\gamma_n}(1 - \mathcal{B}_{u\gamma_n}))_{u\in[0,T]}\right)\right.\right.$$

$$\left.\left. - h\left(s_n - \frac{\delta_n}{\gamma_n}, (\mathcal{B}_{u\gamma_n}(1 - \mathcal{B}_{u\gamma_n}))_{u\in[0,T]}\right)\right|\right] = \mathbf{E}_u\left[\tilde{h}\left(s_n, s_n - \frac{\delta_n}{\gamma_n}, (\mathcal{B}_{u\gamma_n}(1 - \mathcal{B}_{u\gamma_n}))_{u\in[0,T]}\right)\right]$$

$$\xrightarrow{\mathbb{P}} \mathbf{E}_u\left[\tilde{h}\left(s, s, (\bar{\mathcal{B}}_u(1 - \bar{\mathcal{B}}_u))_{u\in[0,T]}\right)\right] = 0, \tag{4.10}$$

where the limit is justified by (1.6) and Proposition A.2 .

3. Combine (4.7), (4.9), and (4.10), to conclude the proof. □

5. Uniform concentration

In this section we prove Theorem 1.3, i.e., we sharpen the result in Theorem 1.2*(i)* to a uniform bound over sublinear times up to the scale $n^{1-o(1)}$. Note that, up to a union bound on the set of jump times for the process $\mathcal{D}_t^n$, the proof amounts to showing the following.

Proposition 5.1 (Strengthening of the pointwise concentration). *Consider the voter model with initial density $u \in (0, 1)$ on a regular random graph $G_{d,n}(\omega)$. For any fixed $\varepsilon, \delta, a > 0$,*

$$\mathbb{1}\left\{\max_{t\leq n^{1-\varepsilon}} \mathbf{P}_u\left(|\mathcal{D}_t - \mathbf{E}_u[\mathcal{D}_t]| > \delta\right) \leq n^{-a}\right\} \xrightarrow{\mathbb{P}} 1.$$

Proof. The proof comes in several steps. To ease the reading, we drop the subindex u from the notation of the voter measure $\mathbf{P}_u$ and abbreviate $p_t = \mathbf{E}[\mathcal{D}_t]$.

1. Fix

$$t \leq n^{1-\varepsilon}, \qquad K_n = \log^2 n,$$

and write A to denote a collection of K edges of the graph sampled uniformly at random, independently of the voter dynamics. Call

$$\mathcal{D}_t^A := \frac{1}{K_n} \sum_{e \in A} \mathbb{1}_{\{e \in D_t\}}$$

the fraction of edges in the random set A that are discordant at time t. By the triangle inequality, we have

$$\mathbf{P}\left(|\mathcal{D}_t - p_t| > \delta\right) \le \mathbf{P}\left(|\mathcal{D}_t - \mathcal{D}_t^A| > \tfrac{1}{2}\delta\right) + \mathbf{P}\left(|\mathcal{D}_t^A - p_t| > \tfrac{1}{2}\delta\right). \tag{5.1}$$

Note that, given $\mathcal{D}_t = q$, the quantity $\mathcal{D}_t^A$ is distributed as Z/K_n, where $Z \sim \mathrm{Bin}(K_n, q)$. Therefore, conditioning on $\mathcal{D}_t = q \in [0, 1]$, we can bound the first term on the right-hand side of (5.1) by means of the Chernoff bound

$$\mathbf{P}\left(|\mathcal{D}_t - \mathcal{D}_t^A| > \tfrac{1}{2}\delta \;\middle|\; \frac{|D_t|}{m} = q\right) \le 2\exp\left(-\frac{\delta^2}{12q}K_n\right) \le 2\exp\left(-\frac{\delta^2}{12}K_n\right). \tag{5.2}$$

Since the bound in (5.2) is uniform over $q \in [0, 1]$ and $t \le n^{1-\varepsilon}$, we conclude that, for every $a > 0$, $\mathbb{P}$-a.s.,

$$\max_{t \le n^{1-\varepsilon}} \mathbf{P}\left(|\mathcal{D}_t - \mathcal{D}_t^A| > \tfrac{1}{2}\delta\right) \le 2\exp\left(-\frac{\delta^2}{12}K_n\right) \le n^{-a},$$

for all n sufficiently large.

2. We next show that a similar bound holds for the second term on the right-hand side of (5.1). The latter will be proved by means of duality. Consider a system of n independent random walks starting from the different vertices of G and evolving independently. We consider a (multi)sub-set of $2K_n$ random walks starting at the extremes of the edges in the random set A. Note that these are distributed at time zero as $\otimes_{K_n}\nu$, where ν is the probability distribution on $[n]^2$ defined as

$$\nu(x, y) = \pi(x)\frac{1}{d}\mathbb{1}_{x \sim y}, \qquad (x, y) \in [n]^2.$$

Moreover, since $\pi \equiv \frac{1}{n}$, the two marginal distributions of ν coincide with the stationary distribution, i.e., for every $x \in [n]$,

$$\sum_{z \in [n]} \nu(x, z) = \sum_{z \in [n]} \nu(z, x) = \frac{1}{n}.$$

Observe that

$$p_t = \mathbf{E}[\mathcal{D}_t] = \frac{1}{m} \sum_{e \in E} \mathbf{P}(e \in D_t) = \mathbf{P}_\nu(e \in D_t). \tag{5.3}$$

In order to simplify the notation, when considering a system of $2K$ independent random walks starting at $\otimes_{K_n}\nu$, for each of the K_n random edges $A = \{e_1, \ldots, e_{K_n}\}$, we label the extremes as e_j^- and e_j^+, for all $j \in \{1, \ldots, K_n\}$. For $i, j \le K_n$, define the quantities

$$\tau^{e_i, e_j} := \tau_{\mathrm{meet}}^{e_i^-, e_j^-} \wedge \tau_{\mathrm{meet}}^{e_i^+, e_j^-} \wedge \tau_{\mathrm{meet}}^{e_i^-, e_j^+} \wedge \tau_{\mathrm{meet}}^{e_i^+, e_j^+}. \tag{5.4}$$

When $\tau^{e_i, e_j} \le t$ we say that *the edges i and j interact before t*. We say that $\tau^{e_i, e_j} = 0$ when $\{e_i^-, e_i^+\} \cap \{e_j^-, e_j^+\} \ne \emptyset$. Note that, for all $i, j \le K_n$ and $t \ge 0$,

$$\mathbf{P}_{\otimes_{K_n}\nu}(\tau^{e_i, e_j} \le t) = \mathbf{P}_{\nu \otimes \nu}(\tau^{e_1, e_2} \le t) \le 4\mathbf{P}_{\pi \otimes \pi}(\tau_{\mathrm{meet}}^{x, y} \le t). \tag{5.5}$$

Indeed, the inequality in (5.5) follows from the definition of ν and a union bound.

We next argue that for all $t = o(n)$ the probability on the right-hand side of Eq. (5.5) can be bounded above *whp* by

$$\mathbf{P}_{\pi \otimes \pi}(\tau_{\text{meet}}^{x,y} \leq t) = O\left(\frac{t}{n}\right). \tag{5.6}$$

Let us start by pointing out that the FVTL is not enough to deduce (5.6). Indeed, the FVTL is particularly suited to have sharp estimates of the right tail of the hitting time of a set on time scales that are *at least* as big as the mean hitting time. For this reason, we will exploit a result in Aldous and Brown (1992) that in a sense complements the FVTL on short time scales (see also (Hermon et al., 2021, Lemma 2.11)).

Lemma 5.2. *For an irreducible and reversible continuous-time Markov chain with stationary distribution μ and any subset of states A,*

$$\left| \mathbf{P}(\tau_A^\mu \leq t) - \frac{t}{\mathbf{E}[\tau_A^\mu]} \right| \leq \left(\frac{t}{\mathbf{E}[\tau_A^\mu]} \right)^2 + \frac{1}{\lambda_\star \, \mathbf{E}[\tau_A^\mu]}, \qquad 0 \leq t \leq \mathbf{E}[\tau_A^\mu], \tag{5.7}$$

where τ_A^μ is the hitting time of the set A when starting with distribution μ, and $\lambda_\star$ denotes the minimal non-trivial eigenvalue of the infinitesimal generator of the process.

In our setting we can think of the Markov chain as the simple random walk on $G \otimes G$ and of A as the diagonal, so that $\mu = \pi \otimes \pi$, $\lambda_\star \in (0,1)$ and $\mathbf{E}[\tau_A^\mu] = \mathbf{E}[\tau_\Delta^{\pi \otimes \pi}] = \Theta(n)$, where the latter follows by Proposition 3.7. Note that, as soon as $t/n \to 0$, we have that the term on the right-hand side of (5.7) is $O(t/n)$, as for $t/\mathbf{E}[\tau_A^\mu]$, so that (5.6) immediately follows.

The following lemma states that it is unlikely to have within A a density of edges interacting before time t.

Lemma 5.3. *Consider a system of n independent random walks, each starting from a different vertex of a regular random graph $G_{d,n}(\omega)$. Select a sub(multi)set A of E of size $2K_n$ with joint distribution $\otimes_{K_n} \nu$. Fix $\varepsilon, \gamma > 0$, $t \leq n^{1-\varepsilon}$ and $K_n = \log^2 n$. Let*

$$\mathcal{E}_{t,\gamma} := \left\{ \exists I \subset A \text{ s.t. } \max_{e \in I} \min_{f \in A \setminus \{e\}} \tau^{e,f} \leq t, \text{ and } |I| \geq \gamma K_n \right\},$$

where $\tau^{e,f}$ is defined as in (5.4). Then, for all $a > 0$,

$$\mathbf{P}_{\otimes_{K_n} \nu}(\mathcal{E}_{t,\gamma}) \leq n^{-a} \quad \text{whp}.$$

Before proving Lemma 5.3, we show how the latter can be exploited to conclude the proof of Proposition 5.1.

3. To conclude the proof of Proposition 5.1, we argue as follows. Let

$$\mathcal{H}_t = \{ |\mathcal{D}_t^A - p_t| > \delta \} \tag{5.8}$$

be the event of interest. Consider a system of $2K_n$ independent random walks starting at the extremes of the vertices in A, and let these evolve up to time t. Let $I \subset K_n$ denote the set of edges in A interacting within time t with other vertices of A. Consider the event $\mathcal{E}_t$ in Lemma 5.3 for which we know that, for every constants $a, \gamma > 0$, *whp*

$$\mathbf{P}(\mathcal{E}_{t,\gamma}) \leq n^{-a}.$$

Therefore
$$\mathbf{P}\left(\mathcal{H}_t\right) \le \mathbf{P}(\mathcal{H}_t \cap \mathcal{E}_{t,\gamma}^c) + n^{-a}.$$

In order to conclude the lemma, it suffices to bound the probability on the right-hand side of the above display.

4. In order to get the desired bound we next show negative correlation for the discordancy of a collection of edges, in the spirit of the proof of Proposition 4.1.

Lemma 5.4 (Negative correlation under no-meeting events). *Consider a sub(multi)set of edges $A_k = \{e_1, \ldots, e_k\}$ sampled according to $\otimes_k \nu$, and let $\mathcal{G}_{t,k}$ be the event in which the $2k$ independent random walks starting at the extremes of the k edges never meet before time t. Then, for all $j \in \{0, 1, 2, \ldots, k\}$,*

$$\mathbf{P}_{\otimes_k \nu} \left(\text{Exactly } j \text{ of the } k \text{ edges are discordant at time } t, \ \mathcal{G}_{t,k} \right) \le \mathbf{P}\left(\mathrm{Bin}(k, p_t) = j\right).$$

Proof. For $e = (x, y) \in A_k$, call σ_e a realisation of length t of the two independent trajectories starting at x and y. For $2 \le j \le k$, call $\mathcal{H}(\sigma_1, \ldots, \sigma_{j-1})$ the set of the possible realisations of σ_j that never meet the $2(j-1)$ trajectories $\sigma_1, \ldots, \sigma_{j-1}$. Observe that

$$\mathbf{P}_{\otimes_k \nu}(e_i \in D_t, \forall i \le j, \ e_i \notin D_t, \forall i > j, \ \mathcal{G}_{t,k})$$

$$= \sum_{\sigma_1 \in \{e_1 \in D_t\}} \sum_{\sigma_2 \in \{e_2 \in D_t\} \cap \mathcal{H}(\sigma_1)} \cdots \sum_{\sigma_k \in \{e_k \notin D_t\} \cap \mathcal{H}(\sigma_1, \ldots, \sigma_{k-1})} \prod_{j=1}^{k} \mathbf{P}\left(\sigma_i \mid \cap_{j<i} \sigma_j\right)$$

$$= \sum_{\sigma_1 \in \{e_1 \in D_t\}} \sum_{\sigma_2 \in \{e_2 \in D_t\} \cap \mathcal{H}(\sigma_1)} \cdots \sum_{\sigma_k \in \{e_k \notin D_t\} \cap \mathcal{H}(\sigma_1, \ldots, \sigma_{k-1})} \prod_{j=1}^{k} \mathbf{P}\left(\sigma_i\right)$$

$$\le \sum_{\sigma_1 \in \{e_1 \in D_t\}} \sum_{\sigma_2 \in \{e_2 \in D_t\}} \cdots \sum_{\sigma_k \in \{e_k \notin D_t\}} \prod_{j=1}^{k} \mathbf{P}\left(\sigma_i\right)$$

$$= \prod_{i=1}^{j} \mathbf{P}_{\otimes_k \nu}(e_i \in D_t) \prod_{i=j+1}^{k} \mathbf{P}_{\otimes_k \nu}(e_i \notin D_t)$$

$$= (p_t)^j (1 - p_t)^{k-j},$$

where in the last equality we use the identity in (5.3). By the product form of $\otimes_k \nu$ the result immediately follows. $\square$

5. Recall (5.8) and that I is the set of edges that are involved in some interaction before time t. Under $\mathcal{E}_{t,\gamma}^c$, we have $|I| \le \gamma K_n$. Therefore

$$\mathbf{P}\left(\mathcal{H}_t \cap \mathcal{E}_{t,\gamma}^c\right) \le \sum_{w=0}^{\gamma K_n} \mathbf{P}\left(\mathcal{H}_t; \ |I| = w\right) = \sum_{w=0}^{\gamma K_n} \sum_{B \subset A : \ |B| = w} \frac{1}{\binom{K_n}{w}} \mathbf{P}\left(\mathcal{H}_t; \ I = B\right).$$

Setting $k(w) = K_n - w$, exchangeability implies

$$\mathbf{P}\left(\mathcal{H}_t \cap \mathcal{E}_{t,\gamma}^c\right) \le \sum_{w=0}^{\gamma K_n} \mathbf{P}\left(\mathcal{H}_t; \ A \setminus I = e_1, \ldots, e_{k(w)}\right).$$

30

Observe that the event $\{A \setminus I = e_1, \ldots, e_{k(w)}\}$ implies that the first $k(w)$ edges do not inter-act with any edge in A *and* that each of the last w edges has at least one interaction. This is clearly contained in the event $\mathcal{G}_{t,k(w)} = \{\min_{i,j \le k(w)} \tau^{e_i, e_j} > t\}$ introduced in Lemma 5.4. Hence

$$\mathbf{P}\big(\mathcal{H}_t \cap \mathcal{E}_{t,\gamma}^c\big) \le \sum_{w=0}^{\gamma K_n} \mathbf{P}\big(\mathcal{H}_t \cap \mathcal{G}_{t,k(w)}\big).$$

Observe that, under the event $\mathcal{H}_t \cap \mathcal{G}_{t,k(w)}$,

$$\delta K_n < \Big| \sum_{e \in A} \mathbb{1}_{\{e \in D_t\}} - p_t K_n \Big| \le \Big| \sum_{i=1}^{K_n - w} \mathbb{1}_{\{e_i \in D_t\}} - p_t |K_n - w| \Big| + w + p_t w$$

$$\le \Big| \sum_{i=1}^{K_n - w} \mathbb{1}_{\{e_i \in D_t\}} - p_t |K_n - w| \Big| + 2\gamma K_n,$$

which implies, with

$$\mathcal{R}_{t,k(w)} := \left\{ \Big| \sum_{i=1}^{k(w)} \mathbb{1}_{\{e_i \in D_t\}} - p_t k(w) \Big| > (\delta - 2\gamma) k(w) \right\},$$

that

$$\mathcal{R}_{t,k(w)} \cap \mathcal{G}_{t,k(w)} \supset \mathcal{H}_t \cap \mathcal{G}_{t,k(w)}.$$

It follows that

$$\mathbf{P}\big(\mathcal{H}_t \cap \mathcal{E}_{t,\gamma}^c\big) \le \sum_{w=0}^{\gamma K_n} \mathbf{P}\big(\mathcal{H}_t \cap \mathcal{G}_{t,k(w)}\big) \le \sum_{w=0}^{\gamma K_n} \mathbf{P}\big(\mathcal{G}_{t,k(w)} \cap \mathcal{R}_{t,k(w)}\big), \tag{5.9}$$

from which, thanks to Lemma 5.4, for every $a > 0$ *whp*,

$$\mathbf{P}\big(\mathcal{H}_t \cap \mathcal{E}_{t,\gamma}^c\big) \le \sum_{w=0}^{\gamma K_n} \mathbf{P}\big(\mathcal{G}_{t,k(w)} \cap \mathcal{R}_{t,k(w)}\big)$$

$$\le \sum_{w=0}^{\gamma K_n} \mathsf{P}\big(\big|\text{Bin}(K_n - w, p_t) - p_t |K_n - w|\big| > (\delta - 2\gamma)(K_n - w)\big) \le n^{-a}.$$

$\square$

Proof of Lemma 5.3. The proof comes in several steps. For simplicity, we suppress the de-pendence on γ from the notation.

1. For a vertex $x \in [n]$, we call $\mathcal{P}_x$ the subset of times $[0, t]$ in which the random walk starting at x moves. Put

$$\mathcal{P}^{\text{tot}} = \{\forall x \in [n], \ |\mathcal{P}_x| \le n^{\varepsilon/2} t\}$$

and note that

$$\mathbf{P}\big(\mathcal{P}^{\text{tot}}\big) \ge 1 - n\mathsf{P}\big(\text{Poisson}(t) > n^{\varepsilon/2} t\big) \ge 1 - n^{-a}, \qquad \text{for all } a > 0,$$

where the last inequality holds for all n large enough. Therefore it is enough to show that, for all $a > 0$,

$$\mathbf{P}_{\otimes K_n \nu}\big(\mathcal{E}_{t,\gamma} \cap \mathcal{P}^{\text{tot}}\big) \le n^{-a} \qquad whp.$$

Clearly,

$$\mathbf{P}_{\otimes_{K_n}\nu}\left(\mathcal{E}_{t,\gamma}\cap\mathcal{P}^{\text{tot}}\right) \le \sum_{I\subset A:\ |I|\ge\gamma K_n}\mathbf{P}_{\otimes_{K_n}\nu}\left(\max_{e\in I}\min_{f\in A\backslash\{e\}}\tau^{e,f}\le t,\ \mathcal{P}^{\text{tot}}\right)$$

$$\le \sum_{C=\gamma K_n}^{K_n}\binom{K_n}{C}\mathbf{P}_{\otimes_C\nu}\left(\max_{e\in I}\min_{f\in I\backslash\{e\}}\tau^{e,f}\le t,\ \mathcal{P}^{\text{tot}}\right),$$

(5.10)

where in the second line, due to the exchangeability of the edges in A and with a slight abuse of notation, we can take for I the set of the first C edges of A. Given I, associate to it one possible collection of first interactions between its vertices. Call this set

$$B_I = \{(1,j_1),\ldots,(C,j_C)\},$$

where (ℓ,j_ℓ) stands for the event that the first edge with which the edge e_ℓ interacts is e_{j_ℓ}.

2. For every choice of B_I, we denote by $\mathcal{E}_t(I,B_I)$ the event $\mathcal{E}_{t,\gamma}$ in which the set of interacting edges is I and $B_I = \{(1,j_1),\ldots,(C,j_C)\}$ is the set of first interactions. Then

$$\mathbf{P}_{\otimes_C\nu}\left(\mathcal{E}_t(I,B_I)\cap\mathcal{P}^{\text{tot}}\right) \le \mathbf{P}_{\otimes_C\nu}\left(\tau^{e_1,e_{j_1}}\le t,\mathcal{P}^{\text{tot}}(1)\right)\prod_{\ell=2}^C\mathbf{P}_{\otimes_C\nu}\left(\tau^{e_\ell,e_{j_\ell}}\le t,\ \mathcal{P}^{\text{tot}}(\ell)\mid\mathcal{F}_{\ell-1}\right),$$

(5.11)

where, for $\ell\in\{1,\ldots,C\}$, we put

$$\mathcal{P}^{\text{tot}}(\ell) := \{|\mathcal{P}_{e_\ell^-}|\le n^{\varepsilon/2}t\}\cap\{|\mathcal{P}_{e_\ell^+}|\le n^{\varepsilon/2}t\}\cap\{|\mathcal{P}_{e_{j_\ell}^-}|\le n^{\varepsilon/2}t\}\cap\{|\mathcal{P}_{e_{j_\ell}^+}|\le n^{\varepsilon/2}t\}$$

and

$$\mathcal{F}_\ell := \cap_{\iota\le\ell}\{\tau^{e_\iota,e_{j_\iota}}\le t\}\cap\mathcal{P}^{\text{tot}}(\iota),\qquad \ell=\{2,\ldots,C\}.$$

By construction, regardless of the specific ordering of the elements in the set B_I, there is a collection of at least $C/2$ indices $\ell\in\{1,\ldots,C\}$ such that

$$|\Xi_\ell| := \left|\{e_\ell,e_{j_\ell}\}\cap\{e_1,e_{j_1},\ldots,e_{\ell-1},e_{j_{\ell-1}}\}\right|\le 1.$$

We claim that, for every $\ell\in\{1,\ldots,C\}$ for which $|\Xi_\ell|=0$,

$$\mathbf{P}_{\otimes_C\nu}\left(\tau^{e_\ell,e_{j_\ell}}\le t\mid\mathcal{F}_{\ell-1}\right) \le \mathbf{P}_{\nu\otimes\nu}\left(\tau^{e,f}\le t\right) + \frac{4C}{n} = O\left(\frac{C+t}{n}\right).$$

(5.12)

Indeed,

$$\mathbf{P}_{\otimes_C\nu}\left(\tau^{e_\ell,e_{j_\ell}}\le t\mid\mathcal{F}_{\ell-1}\right)$$

$$\le \mathbf{P}_{\otimes_C\nu}\left(\tau^{e_\ell,e_{j_\ell}}\le t, \{e_\ell^-,e_\ell^+,e_{j_\ell}^-,e_{j_\ell}^+\}\cap\cup_{\iota<\ell}\{e_\iota^-,e_\iota^+,e_{j_\iota}^-,e_{j_\iota}^+\}=\emptyset\mid\mathcal{F}_{\ell-1}\right) + \frac{4\ell}{n}$$

$$= \mathbf{P}_{\otimes_{K_n}\nu}\left(\tau^{e_{i\ell},e_{j_\ell}}\le t, \{e_\ell^-,e_\ell^+,e_{j_\ell}^-,e_{j_\ell}^+\}\cap\cup_{\iota<\ell}\{e_\iota^-,e_\iota^+,e_{j_\iota}^-,e_{j_\iota}^+\}=\emptyset\right) + \frac{4\ell}{n}$$

$$\le \mathbf{P}_{\otimes_{K_n}\nu}\left(\tau^{e_{i\ell},e_{j_\ell}}\le t\right) + \frac{4C}{n}.$$

Therefore, by (5.5), we conclude that, under $|\Xi_\ell|=0$,

$$\mathbf{P}_{\otimes_C\nu}\left(\tau^{e_\ell,e_{j_\ell}}\le t\mid\mathcal{F}_{\ell-1}\right) \le 4\mathbf{P}_{\pi\otimes\pi}(\tau_{\text{meet}}^{x,y}\le t) + \frac{4C}{n},$$

(5.13)

and (5.12) follows by inserting (5.6) into (5.13).

3. Assume now that for some $\ell \in \{1, \ldots, C\}$ we have $|\Xi_\ell| = 1$. More precisely, under $\mathcal{F}_{\ell-1} \cap \{|\Xi_\ell| = 1\}$, one among e_ℓ and e_{j_ℓ} is still unsampled at step ℓ. Without loss of generality, assume that e_ℓ is such an unsampled random edge. By the same argument used to derive (5.13), and putting

$$\mathcal{J}_\ell := \left\{ \{e_\ell^-, e_\ell^+\} \bigcap \cup_{\iota < \ell} \{e_\iota^-, e_\iota^+, e_{j_\iota}^-, e_{j_\iota}^+\} = \emptyset \right\},$$

we bound

$$\mathbf{P}_{\otimes C \nu} \left(\tau^{e_\ell, e_{j_\ell}} \leq t, \mathcal{P}^{\text{tot}}(\ell) \,|\, \mathcal{F}_{\ell-1} \right) \leq 2 \, \mathbf{P}_{\otimes C \nu} \left(\tau^{e_\ell, e_{j_\ell}} \leq t, \mathcal{P}^{\text{tot}}(\ell), \mathcal{J}_\ell \,|\, \mathcal{F}_{\ell-1} \right) + \frac{8C}{n}.$$

Arguing as in (Oliveira, 2013, Proposition 4.5), we get that, under $|\Xi_\ell| = 1$ with e_ℓ unsampled at step ℓ,

$$
\begin{aligned}
&\mathbf{P}_{\otimes C \nu} \left(\tau^{e_\ell, e_{j_\ell}} \leq t, \mathcal{P}^{\text{tot}}(\ell), \mathcal{J}_\ell \,\big|\, \mathcal{F}_{\ell-1} \right) \\
&\leq \mathbf{P}_{\otimes C \nu} \left(\tau^{e_\ell, e_{j_\ell}} \leq t \,\big|\, \mathcal{F}_{\ell-1}, \mathcal{J}_\ell, \mathcal{P}^{\text{tot}}(\ell) \right) \\
&= \mathbf{P}_{\otimes C \nu} \left(\tau_{\text{meet}}^{e_\ell^-, e_{j_\ell}^-} \wedge \tau_{\text{meet}}^{e_\ell^+, e_{j_\ell}^-} \wedge \tau_{\text{meet}}^{e_\ell^-, e_{j_\ell}^+} \wedge \tau_{\text{meet}}^{e_\ell^+, e_{j_\ell}^+} \leq t \,\big|\, \mathcal{F}_{\ell-1}, \mathcal{J}_\ell, \mathcal{P}^{\text{tot}}(\ell) \right) \\
&\leq 4 \mathbf{P}_{\otimes C \nu} \left(\tau_{\text{meet}}^{e_\ell^-, e_{j_\ell}^-} \leq t \,\big|\, \mathcal{F}_{\ell-1}, \mathcal{J}_\ell, \mathcal{P}^{\text{tot}}(\ell) \right) \\
&= 4 \mathbf{E} \left[\mathbf{P}_{\otimes C \nu} \left(\tau_{\text{meet}}^{e_\ell^-, e_{j_\ell}^-} \leq t \,\big|\, \mathcal{F}_{\ell-1}, \mathcal{J}_\ell, \mathcal{P}^{\text{tot}}(\ell), \mathcal{P}_{e_\ell^-}, \mathcal{P}_{e_{j_\ell}^-} \right) \Big|\, \mathcal{F}_{\ell-1}, \mathcal{J}_\ell, \mathcal{P}^{\text{tot}}(\ell) \right],
\end{aligned}
$$

(5.14)

where the second inequality follows from the equality of the marginals. Note that, conditionally on $\mathcal{F}_{\ell-1}$ and $\mathcal{J}_\ell$, for all $s \leq t$ the position of the random walk starting at e_ℓ^- is distributed according to π. On the other hand, the position of the random walk starting at $e_{j_\ell}^-$ at time s is distributed in a non-trivial way. Moreover, the positions of the two random walks at time s are independent. Hence

$$
\begin{aligned}
&\mathbf{P}_{\otimes C \nu} \left(\tau_{\text{meet}}^{e_\ell^-, e_{j_\ell}^-} \leq t \,\big|\, \mathcal{F}_{\ell-1}, \mathcal{J}_\ell, \mathcal{P}^{\text{tot}}(\ell), \mathcal{P}_{e_\ell^-}, \mathcal{P}_{e_{j_\ell}^-} \right) \\
&\leq \sum_{s \in \mathcal{P}_{e_\ell^-} \cup \mathcal{P}_{e_{j_\ell}^-}} \mathbf{P}_{\otimes C \nu} \left(X_s^{(e_\ell^-)} = X_s^{(e_{j_\ell}^-)} \,\big|\, \mathcal{F}_{\ell-1}, \mathcal{J}_\ell, \mathcal{P}^{\text{tot}}(\ell), \mathcal{P}_{e_\ell^-}, \mathcal{P}_{e_{j_\ell}^-} \right) \\
&\leq \sum_{s \in \mathcal{P}_{e_\ell^-} \cup \mathcal{P}_{e_{j_\ell}^-}} \max_\lambda \sum_{x \in [n]} \lambda(x) \pi(x) \leq \frac{|\mathcal{P}_{e_\ell^-} \cup \mathcal{P}_{e_{j_\ell}^-}|}{n}.
\end{aligned}
$$

(5.15)

Inserting (5.15) into the expectation in (5.14), we obtain

$$\mathbf{P}_{\otimes C \nu} \left(\tau^{e_{i_\ell}, e_{j_\ell}} \leq t, \mathcal{P}^{\text{tot}}(\ell), \mathcal{J}_\ell \,\big|\, \mathcal{F}_{\ell-1}, |\Xi_\ell| = 1 \right) \leq \frac{8 n^{\varepsilon/2} t}{n}. \tag{5.16}$$

Inserting (5.13)–(5.16) into (5.11), we obtain

$$\mathbf{P}_{\otimes C \nu} \left(\mathcal{E}_t(I, B_I) \cap \mathcal{P}^{\text{tot}} \right) \leq \left(\frac{8(n^{\varepsilon/2} t + C)}{n} \right)^{C/2} \leq \left(\sqrt{\frac{8(n^{\varepsilon/2} t + K)}{n}} \right)^C =: \Gamma^C. \tag{5.17}$$

4. By (5.10), (5.17) and the fact that there are $(C-1)^C$ different ways to select the set B_I, we are now ready to conclude that *whp*

$$\mathbf{P}_{\otimes_{K_n}\nu}\left(\mathcal{E}_{t,\gamma}\cap\mathcal{P}^{\mathrm{tot}}\right) \leq \sum_{C=\gamma K_n}^{K_n}\binom{K_n}{C}C^C\Gamma^C \leq \sum_{C=\gamma K_n}^{K_n}(K_n^2\Gamma)^C.$$

Since, by our choice of t and K_n,

$$K_n^2\Gamma \leq n^{-\varepsilon/5}$$

for all n large enough, we obtain

$$\mathbf{P}_{\otimes_{K_n}\nu}\left(\mathcal{E}_{t,\gamma}\cap\mathcal{P}^{\mathrm{tot}}\right) \leq K_n n^{-\frac{\varepsilon\gamma}{5}K_n},$$

concluding the proof. $\qquad\square$

APPENDIX A. AUXILLIARY FACTS FOR CÀDLÀG PROCESSES

In the next lemma, for an arbitrary $T > 0$, we let $[0,T]\times\mathcal{D}[0,T]$ denote the product of $[0,T]$ with the càdlàg space $\mathcal{D}[0,T]$ endowed with the J_1-Skorohod topology.

Lemma A.1. *Consider the function $h\colon [0,T]\times\mathcal{D}[0,T]\to\mathbb{R}$ defined via $h(s,\phi)=\phi(s)$. Every point in $[0,T]\times\mathcal{C}[0,T]$ is a continuity point of h.*

Proof. Assume that (s_n,ϕ_n) is a sequence that converges to (s,ϕ), with $\phi\in\mathcal{C}[0,T]$. We need to verify that $h(s_n,\phi_n)$ converges to $h(s,\phi)$. Indeed, if ϕ_n converges to ϕ, then there exists a sequence of increasing functions $\lambda_n : [0,T]\to[0,T]$ with $\lambda_n(0)=0$ and $\lambda_n(T)=T$ such that $\lambda_n(u)\to u$ and $\phi_n(u)-\phi(\lambda_n(u))\to 0$ uniformly over $u\in[0,T]$. From this we obtain

$$||\phi_n-\phi||_\infty \leq ||\phi_n-\phi\circ\lambda_n||_\infty + ||\phi\circ\lambda_n-\phi||_\infty \to 0,$$

since any function in $\mathcal{C}[0,T]$ is uniformly continuous. In particular, ϕ_n converges to ϕ uniformly, which readily implies that $\phi_n(s_n)\to\phi(s)$ and concludes the proof of the lemma. $\qquad\square$

Proposition A.2. *Assume that $(X_u^n)_{u\in[0,T]}$ is a sequence of càdlàg processes that converge in distribution to a process $(X_u)_{u\in[0,T]}$ that has almost surely continuous trajectories. For any sequence $t_n\to t\in[0,T]$, $X_{t_n}^n$ converges in distribution to X_t.*

Proof. From the assumptions of the proposition, it follows that $((X_u^n)_{u\in[0,T]},t_n)_{n\in\mathbb{N}}$ converges in distribution to $((X_u)_{u\in[0,T]},t)$. Let $h\colon [0,T]\times\mathcal{D}[0,T]\to\mathbb{R}$ as in Lemma A.1. Note that, since $(X_u)_{u\in[0,T]}$ is almost surely continuous, it follows that $X_{t_n}^n = h(t_n,(X_u^n)_{u\in[0,T]})$ converges in distribution to $X_t = h(t,(X_u)_{u\in[0,T]})$. $\qquad\square$

REFERENCES

Aldous, D. J. and Brown, M. (1992). Inequalities for rare events in time-reversible Markov chains I. *Lecture Notes-Monograph Series*, pages 1–16.

Chen, Y.-T. (2021). Precise asymptotics of some meeting times arising from the voter model on large random regular graphs. *Electronic Communications in Probability*, 26:1–13.

Chen, Y.-T., Choi, J., and Cox, J. T. (2016). On the convergence of densities of finite voter models to the Wright–Fisher diffusion. *Annales de l'Institut Henri Poincaré*, 52(1):286–322.

Clifford, P. and Sudbury, A. (1973). A model for spatial conflict. *Biometrika*, 60(3):581–588.

Cooper, C., Frieze, A., and Radzik, T. (2010). Multiple random walks in random regular graphs. *SIAM Journal on Discrete Mathematics*, 23(4):1738–1761.

Cooper, C. and Frieze, A. M. (2005). The cover time of random regular graphs. *SIAM Journal on Discrete Mathematics*, 18(4):728–740.

Cox, J. (1989). Coalescing random walks and voter model consensus times on the torus in zd. *Annals of Probabability*, 17(4):1333–1366.

Cox, J. and Greven, A. (1990). On the long term behaviour of some finite particle systems. *Probababilty Theory and Related Fields*, 85:195–237.

Cox, J. and Greven, A. (1994). The finite systems scheme: An abstract theorem and a new example. *CRM Proceedings and Lecture Notes*, 5:55–66.

Donnelly, P. and Welsh, D. (1983). Finite particle systems and infection models. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 94, pages 167–182. Cambridge University Press.

Durrett, R. (2008). *Probability Models for DNA Sequence Evolution (2nd. ed.)*. Springer, New York.

Fernley, J. and Ortgiese, M. (2019). Voter models on subcritical inhomogeneous random graphs. *arXiv preprint arXiv:1911.13187*.

Friedman, J. (2008). *A Proof of Alon's Second Eigenvalue Conjecture and Related Problems*. American Mathematical Soc.

Hermon, J., Li, S., Yao, D., and Zhang, L. (2021). Mean field behavior during the big bang for coalescing random walk. *arXiv preprint arXiv:2105.11585*.

Holley, R. A. and Liggett, T. M. (1975). Ergodic theorems for weakly interacting infinite systems and the voter model. *Annals of Probability*, pages 643–663.

Levin, D. A. and Peres, Y. (2017). *Markov Chains and Mixing Times*. American Mathematical Society, Providence, RI.

Lubetzky, E. and Sly, A. (2010). Cutoff phenomena for random walks on random regular graphs. *Duke Mathematical Journal*, 153(3):475–510.

Manzo, F., Quattropani, M., and Scoppola, E. (2021). A probabilistic proof of Cooper & Frieze's "first visit time lemma". *ALEA Lat. Am. J. Probab. Math. Stat.*, 18(2):1739–1758.

Oliveira, R. I. (2013). Mean field conditions for coalescing random walks. *Annals of Probability*, 41(5):3420–3461.

Quattropani, M. and Sau, F. (2022). On the meeting of random walks on random DFA. *arXiv preprint arXiv:2204.02827*.

Sood, V., Antal, T., and Redner, S. (2008). Voter models on heterogeneous networks. *Physical Review E*, 77(4):041121.

MATHEMATICAL INSTITUTE, LEIDEN UNIVERISITY, P.O. BOX 9521, 2300RA LEIDEN, THE NETHERLANDS

Email address: {l.avena,r.baldasso,r.s.hazra,denholla,m.quattropani}@math.leidenuniv.nl