On the convergence of densities of finite voter models to the Wright–Fisher diffusion

Yu-Ting Chen[a,1], Jihyeok Choi[b,2] and J. Theodore Cox[b,3]

[a]*Center of Mathematical Sciences and Applications, Harvard University, Cambridge, MA 02138, USA.*
E-mail: yuting_chen@cmsa.fas.harvard.edu
[b]*Department of Mathematics, Syracuse University, Syracuse, NY 13244-1150, USA*

Abstract. We study voter models defined on large finite sets. Through a perspective emphasizing the martingale property of voter density processes, we prove that in general, their convergence to the Wright–Fisher diffusion only involves certain averages of the voter models over a small number of spatial locations. This enables us to identify suitable mixing conditions on the underlying voting kernels, one of which may just depend on their eigenvalues in some contexts, to obtain the convergence of density processes. We show by examples that these conditions are satisfied by a large class of voter models on growing finite graphs.

Résumé. Nous étudions le modèle du votant sur des ensembles contenant un nombre grand mais fini de sites. En nous servant des propriétés de martingales des densités du modèle du votant nous prouvons qu' il y a convergence vers une diffusion Wright–Fisher. De plus cette preuve de convergence n'utilise que certaines moyennes sur un petit nombre de sites. Ceci nous permet d'identifier des conditions de mélange concernant le noyau du votant sous-jacent. Dans certains cas une de ces conditions nous permet de démontrer la convergence des densités en n'utilisant que les valeurs propres des noyaux. Nous donnons des exemples montrant que ces conditions de mélange sont satisfaites pour une grande classe de modèles du votant sur des ensembles croissants de graphes.

MSC: Primary 60K35; 82C22; secondary 60F05; 60J60

Keywords: Wright–Fisher diffusion; Voter model; Interacting particle system; Dual processes; Semimartingale convergence theorem

1. Introduction

The goal of this work is to investigate the convergence of density processes in finite voter models to the Wright–Fisher diffusion. This convergence gives a mean-field approximation for voter models, and is also closely related to the mean-field approximation of coalescence times for the associated dual Markov chains (cf. the recent work of Oliveira [23] and [24]). Earlier examples for such convergence of density processes are few and include the traditional mean-field models and the voter models on d-dimensional tori for $d \geq 2$ (cf. Cox [5]). In the present work, we give mixing conditions on the underlying voting kernels which hold for a large class of finite voter models, and in particular generalize the earlier results.

We introduce the class of voter models considered throughout this paper. See Chapter V of Liggett [18] or Section 4.3 of Liggett [19] for a general account of voter models. Recall that for a finite set E, a Q-matrix q is indexed by $x, y \in E$ and satisfies

$$q(x, y) \geq 0 \quad \forall x \neq y \quad \text{and} \quad q(x) \equiv -q(x, x) = \sum_{y: y \neq x} q(x, y) \tag{1.1}$$

(see Chapter 2 of [19]). For such a pair (q, E) with q irreducible, the associated continuous-time voter model (ξ_s) is the $\{0, 1\}^E$-valued Markov chain evolving according to the following rule. At independent exponential random times, the "voter" at site x replaces its "opinion," which is 0 or 1, with that of another site chosen independently according to $q(x, \cdot)$ on $E \setminus \{x\}$. More precisely, the voter model (ξ_s) is the pure-jump Markov process on $\{0, 1\}^E$ with generator

$$\mathcal{L}f(\xi) \equiv \sum_{x \in E} c(\xi, \xi^x)(f(\xi^x) - f(\xi)). \tag{1.2}$$

Here, for any configuration ξ, ξ^x is obtained by switching the opinion of ξ at x to the opposite one and differs from ξ only at this site, and the flip rate at which ξ changes to ξ^x is given by

$$c(\xi, \xi^x) = \sum_{y \in E} [\xi(x)\widehat{\xi}(y) + \widehat{\xi}(x)\xi(y)] q(x, y), \tag{1.3}$$

for $\widehat{\xi} = 1 - \xi$. Hence, the Q-matrix q can be interpreted as the *voting kernel* of (ξ_s). By allowing q to be a general Q-matrix as in (1.3), we can consider the case that the total voting rates $q(x)$ defined in (1.1) are site-dependent.

We consider in particular the density process $(p_1(\xi_s))$ of such a voter model, where

$$p_1(\xi) = \sum_{x \in E} \pi(x)\xi(x) \tag{1.4}$$

and π is the unique stationary (probability) distribution of the irreducible q-Markov chain, that is the Markov chain with semigroup $(e^{tq}; t \geq 0)$. The simplest example arises from the mean-field model in which each $q(x, \cdot)$ is the uniform distribution on the set $E \setminus \{x\}$, and it is often called the Moran model in population genetics. In this setting, π is the uniform distribution on E, and it is straightforward to apply diffusion approximation to the density processes. More precisely, these processes, after time-changes by suitable constants, converge in distribution in the Skorokhod space to the Wright–Fisher diffusion as the "population size" $|E|$ tends to infinity. Here, we recall that the Wright–Fisher diffusion, denoted by

$$(Y, (\mathbb{P}_u)_{u \in [0,1]})$$

throughout this paper, is a Markov process on $[0, 1]$ which uniquely solves the well-posed martingale problem for

$$\mathcal{G} \equiv \frac{1}{2}x(1-x)\frac{d^2}{dx^2} \tag{1.5}$$

and initial condition u for every $u \in [0, 1]$. In particular, the Wright–Fisher diffusion is a continuous martingale with predictable quadratic variation

$$\langle Y \rangle_t = \int_0^t Y_s(1 - Y_s)\, ds. \tag{1.6}$$

See Section 10.3 in Ethier and Kurtz [13] for the convergence of these density processes and Chapter 4 in the same reference for martingale problems.

For more realistic modelling, several works consider finite voter models where the voting kernels q are defined by spatial structures, or more precisely by the transition kernels of (simple) random walks on graphs (see Chapter 14 in [2,5], Section 6.9 in [11,12,23] and [26]). We note that in theoretical biology, such voter models play an important role

in the study of evolutionary dynamics where the use of general spatial structures for the underlying social networks of biological identities is essential (cf. [4,8,25] and the references there). Voter models in these contexts become harder to analyze, but the mean-field case mentioned above may still serve as an important example in their studies.

For density processes in spatial voter models, the work Cox [5] obtains a similar diffusion approximation on d-dimensional discrete tori for $d \geq 2$. It proves that if the initial laws for voter models are Bernoulli product measures with a constant density, then the density processes, again after suitable constant time-changes, converge to the Wright–Fisher diffusion. We note that the voting kernels defining the voter models in [5] are nearest-neighbor ones allowing only "local" interactions, whereas interactions in the mean-field case are defined by voters living in "well-mixed" populations and are very different in nature. Hence, the fact that the Wright–Fisher diffusion appears as the diffusion limit in both cases suggests that this type of diffusion approximation of density processes should occur in some generality. We will focus on the case as in [5] that the initial conditions are Bernoulli product measures.

To introduce our perspective on this question, we restrict our attention to the simple case that

$$q = p - \mathrm{Id}_E \tag{1.7}$$

for some symmetric probability matrix p with zero diagonal throughout this section. Here, Id_E is the identity matrix indexed by elements of E, and such a Q-matrix q arises when we consider the usual time-change of a discrete-time Markov chain with transition matrix p by an independent rate-1 Poisson process (cf. Section 20.1 of [17]). In Section 2, we will give our result for general irreducible voting kernels q, and more notation is required then. Now, the stationary distribution π for a voting kernel q of the form (1.7) is the uniform distribution, and the density process

$$\bigl(p_1(\xi_s)\bigr) \text{ is a martingale with jump size } \frac{1}{|E|}. \tag{1.8}$$

By introducing a constant time-scale factor $\gamma > 0$, the density process has predictable quadratic variation

$$\bigl\langle p_1(\xi_{\gamma \cdot})\bigr\rangle_t = \frac{2\gamma}{|E|} \int_0^t p_{10}(\xi_{\gamma s})\, \mathrm{d}s, \tag{1.9}$$

where

$$p_{10}(\xi) = \frac{1}{|E|} \sum_{x,y \in E} q(x,y)\xi(x)\widehat{\xi}(y) \tag{1.10}$$

is a weighted average of $(1,0)$ pairs in the configuration ξ. See Proposition 3.1 for these properties of density processes.

This observation should readily reveal the similarity of the density process and the Wright–Fisher diffusion in terms of martingales, under the condition that the population size $|E|$ is large and the predictable quadratic variation of the density process, a weighted average of $(1,0)$ pairs in $(\xi_{\gamma s})$ by (1.9), satisfies

$$\bigl\langle p_1(\xi_{\gamma \cdot})\bigr\rangle_t \approx \int_0^t p_1(\xi_{\gamma s})\bigl[1 - p_1(\xi_{\gamma s})\bigr]\mathrm{d}s \quad \text{as } |E| \longrightarrow \infty \tag{1.11}$$

(recall the predictable quadratic variation (1.6) of the Wright–Fisher diffusion). The mean-field case gives the simplest example satisfying this condition, since

$$p_{10}(\xi) = \frac{|E|}{|E| - 1} p_1(\xi)\bigl[1 - p_1(\xi)\bigr], \tag{1.12}$$

and hence (1.11) holds plainly with $\gamma = |E|/2$. In general, if we pass $|E|$ to infinity and $p_1(\xi_0)$ converges, then under (1.11) the density processes should converge to a continuous martingale by (1.8) which solves the well-posed martingale problem associated with the differential operator $\mathcal{G}$ in (1.5). In other words, the limiting object should be the Wright–Fisher process, and indeed, standard martingale arguments confirm this. See Section 5 for the details, and also its last two paragraphs for the use of general initial conditions.

We will formalize the condition (1.11) by the convergence *in probability* of the differences

$$\left\langle p_1(\xi_\gamma.) \right\rangle_t - \int_0^t p_1(\xi_{\gamma_s}) \left[1 - p_1(\xi_{\gamma_s}) \right] \mathrm{d}s \qquad (1.13)$$

for any $t \in (0, \infty)$. Here, we pass to the limit along a sequence of voter models, started with Bernoulli product measures with a constant density and defined by $(q^{(n)}, E_n)_{n \in \mathbb{N}}$ with $|E_n| \longrightarrow \infty$, and a sequence of constant time scales (γ_n). Our first main result in this paper shows that such convergence of the differences (1.13) is in fact an *equivalent* condition for the convergence of the voter densities toward the Wright–Fisher diffusion. See Theorem 2.1.

Let us discuss how the method of moments in Cox [5] can be applied to general finite voter models, and compare this method with the method of martingale problems stated above. In [5], the convergence of densities for voter models on discrete tori of dimension $d \geq 2$ toward the Wright–Fisher diffusion was obtained by proving that certain coalescence times of random walks are approximately sums of independent exponential variables and then appealing to the method of moments via the well-known duality between voter models and coalescing Markov chains (see [18] or (3.7) below). In fact, there is an equivalence between such almost exponentiality of coalescence times and the convergence to the Wright–Fisher diffusion of voter density processes. It takes several forms, and they hold in general (see Proposition 2.5 and Proposition 2.6). To apply these, we note that the recent work of Oliveira in [23] obtains the required asymptotic behavior of coalescence times for general Markov chains under Aldous's condition discussed below. This result can be readily used to get the mean-field behavior for one-dimensional marginals of the associated voter densities. Nonetheless, in contrast to the method of moments, we believe that the present approach by martingale problems gives greater insight into why the convergence to the Wright–Fisher diffusion should hold. It leads to an equivalent condition in terms of the lower-order densities in (1.13).

The second main result of this paper is concerned with sufficient conditions for the convergence of the differences (1.13) in terms of the underlying sequence of voting kernels $q^{(n)}$. By Proposition 5.3 below, the convergence in probability of the differences (1.13) for $q^{(n)}$ can be reinforced to convergence in L^2-norm. Hence with duality, it can be shown that this convergence is equivalent to a condition involving the coalescence times of *four* $q^{(n)}$-Markov chains (recall (1.9) and see the remark below Proposition 5.3). We give two simpler sufficient conditions for the convergence, and each involves just *two* $q^{(n)}$-Markov chains. These conditions result from the classical conditions for almost exponentiality of hitting times (see Aldous [1] and Proposition 5.23 of Aldous and Fill [2]), and carry the informal idea that the time for two independent chains to coalesce "falls far behind" the time for the chain to get close to stationarity. See Theorem 2.2 for the precise formulations. In formalizing the time to stationarity, while one of our two conditions (cf. Theorem 2.2(i)) uses mixing times and also appears in [23] for almost exponentiality of coalescence times, the other one (cf. Theorem 2.2(ii)) is based on spectral gaps and can be weaker, or more readily applied in some instances. On the other hand, by duality and our result for the convergence of voter densities, the latter condition can also serve as a weaker condition for the convergence in distribution of coalescence times to sums of independent exponential variables (Proposition 2.5). See also Section 1.1 in [23] for this issue when it comes to the stronger L_1-Wasserstein approximation of coalescence times.

As a final remark, we compare our results with the convergence of the rescaled *measure-valued* densities of voter models on $\mathbb{Z}^d$ to super-Brownian motions as in Cox, Durrett and Perkins [7] for $d \geq 2$ and to a nonnegative solution of an SPDE as in Mueller and Tribe [22] for $d = 1$. These voter models live on infinite spatial structures which, after rescaling, converge in the natural way to tractable geometric objects, namely Euclidean spaces of the same dimension, and hence allow more detailed studies of the associated voter models. In our case, the analysis relies on the martingale property of densities, and we circumvent the issue of limiting spatial structures by turning to analytic conditions for almost exponentiality of coalescence times.

The paper is organized as follows. In Section 2, we present our main results for general finite voter models. In Section 3, we study some martingales associated with a density process and use the duality equation for voter models to interpret these martingale properties in terms of coalescing Markov chains. In Section 4, we characterize the convergence of the second moment of density processes in terms of the asymptotic exponentiality of coalescence times. The results in this section are the core of our approach to obtain the convergence of density processes. In Section 5, we study tightness of densities and prove a general version (see Theorem 2.1) of the statement that the convergence of density processes to the Wright–Fisher diffusion is equivalent to the convergence in probability of the differences in (1.13). As an application of this result, we prove in Section 6 two sufficient conditions, each involving only two

independent q-Markov chains, for the convergence of voter densities (see Theorem 2.2). In Section 7, we discuss some connections between the convergence of coalescence times and the convergence of density processes, and the main results will be given below in Proposition 2.5 and Proposition 2.6. Finally, Section 8 is devoted to a few examples to illustrate our sufficient conditions (see Theorem 2.2 and Corollary 2.3) for the convergence of density processes to the Wright–Fisher diffusion.

2. Main results

From this section on, we consider voter models subject to irreducible Q-matrices (recall (1.1)) unless otherwise mentioned. We work with a sequence of irreducible Q-matrices

$$\left(q^{(n)}, E_n\right)_{n\in\mathbb{N}}$$

with stationary (probability) distributions $(\pi^{(n)})$ whenever we study voter models on large sets, and a pair (q, E) with stationary distribution π otherwise. The voter models associated with such a sequence $(q^{(n)}, E_n)$ started at Bernoulli product measures μ_u with density $\mu_u(\xi(x) = 1) = u$ are denoted by $((\xi_s), \mathbb{P}_{\mu_u}^{(n)})$. We will *always* assume that

$$|E_n| \longrightarrow \infty.$$

Whenever necessary, other quantities depending on $(q^{(n)}, E_n)$ will carry subscripts "n" or superscripts "(n)."

We start with our result for the equivalent condition of the convergence of voter densities to the Wright–Fisher diffusion. Now, for any pair (q, E), the associated density process $(p_1(\xi_{\gamma t}))$ for $\gamma > 0$ is a martingale with jump size bounded above by $\max_{x\in E} \pi(x)$, and its predictable quadratic variation takes a more general form than (1.9) which is for the simpler case (1.7). To state the formula for the general case, we set up some notation. Introduce the following measures on the product space $E \times E$ induced by π and q:

$$\nu(x, y) \equiv \pi(x)^2 q(x, y)\mathbb{1}_{x\neq y}, \tag{2.1}$$

$$\bar{\nu}(x, y) \equiv \nu(x, y)/\nu(\mathbb{1}). \tag{2.2}$$

In addition, set $p_{10}(\xi)$ and $p_{01}(\xi)$ as the $\bar{\nu}$-weighted averages of the ordered pairs $(1, 0)$ and $(0, 1)$, respectively, in the configuration ξ, given by

$$p_{10}(\xi) = \sum_{x,y\in E} \bar{\nu}(x, y)\xi(x)\widehat{\xi}(y), \tag{2.3}$$

$$p_{01}(\xi) = \sum_{x,y\in E} \bar{\nu}(x, y)\widehat{\xi}(x)\xi(y). \tag{2.4}$$

Then

$$\left\langle p_1(\xi_{\gamma\cdot})\right\rangle_t = \gamma\nu(\mathbb{1}) \int_0^t \left[p_{10}(\xi_{\gamma s}) + p_{01}(\xi_{\gamma s})\right] \mathrm{d}s \tag{2.5}$$

(see Proposition 3.1 below). Note that if q is of the particular form (1.7), then $\nu(\mathbb{1}) = 1/|E|$, both $p_{10}(\xi)$ and $p_{01}(\xi)$ agree with the right-hand side of (1.10), and the right-hand sides of (1.9) and (2.5) are equal.

Below we use $\xrightarrow[n\to\infty]{(d)}$ to denote convergence in distribution and write

$$\pi_{\mathrm{diag}} = \sum_{x\in E} \pi(x)^2.$$

Theorem 2.1. *Let $u \in (0, 1)$ and let (γ_n) be a sequence of strictly positive constants. Assume that*

$$\lim_{n\to\infty} \pi_{\mathrm{diag}}^{(n)} = 0. \tag{2.6}$$

Then the convergence of density processes

$$\left(p_1(\xi_{\gamma_n \cdot}), \mathbb{P}^{(n)}_{\mu_u}\right) \xrightarrow[n \to \infty]{(d)} (Y, \mathbb{P}_u) \tag{2.7}$$

under the Skorokhod J_1-topology for càdlàg functions holds if and only if the following mean-field condition *holds: for any $T \in (0, \infty)$,*

$$\gamma_n \nu_n(\mathbb{1}) \int_0^T \left[p_{10}(\xi_{\gamma_n s}) + p_{01}(\xi_{\gamma_n s})\right] ds$$

$$- \int_0^T p_1(\xi_{\gamma_n s})\left[1 - p_1(\xi_{\gamma_n s})\right] ds \xrightarrow[n \to \infty]{(d)} 0. \tag{2.8}$$

We will show in Section 4 below (see Theorem 4.1) that the condition (2.6) is in fact necessary for (2.7).

Next, we discuss our second main result which gives sufficient conditions for the mean-field condition (2.8). We need some notation concerning the mixing of the q-Markov chain. Let $(q_t) = (e^{tq})$ be the semigroup of the q-Markov chain on E, and d_E be the maximal total variation distance

$$d_E(t) = \max_{x \in E} \left\| q_t(x, \cdot) - \pi(\cdot) \right\|_{\mathrm{TV}}, \tag{2.9}$$

where $\|\cdot\|_{\mathrm{TV}}$ refers to the total variation distance. Note that $d_E(t)$ is always finite. We recall that the mixing time

$$\mathbf{t}_{\mathrm{mix}} = \inf\left\{t \geq 0 : d_E(t) \leq \frac{1}{2e}\right\} < \infty \tag{2.10}$$

provides, informally speaking, one measurement of the time for the one-dimensional marginals to get close to the equilibrium distribution π. An alternative for this purpose for the q-Markov chain is the associated relaxation time $\mathbf{g}^{-1}$, where $\mathbf{g} \in (0, \infty)$ is the spectral gap and is the second smallest eigenvalue of $-q$. We refer to [2] and [17] for standard properties of spectral gaps and their connections with mixing times (the arguments there can be adapted in a straightforward manner to the context of Markov chains defined by general Q-matrices according to the setup in Section 1.1 of [3]). In particular, we note that $\mathbf{g}^{-1} \leq \mathbf{t}_{\mathrm{mix}}$.

Next, let $M_{U,U'}$ be the meeting time of two independent q-Markov chains with semigroup (q_t) started at spatial locations (U, U'), where the sites U and U' are independent and distributed according to π. We define the expected meeting time to be

$$\mathbf{t}_{\mathrm{meet}} = \mathbf{E}[M_{U,U'}]. \tag{2.11}$$

Theorem 2.2. *For each $n \in \mathbb{N}$, let $\mathbf{g}_n$, $\mathbf{t}^{(n)}_{\mathrm{mix}}$ and $\mathbf{t}^{(n)}_{\mathrm{meet}}$ be the spectral gap, mixing time and expected meeting time of the $q^{(n)}$-Markov chain, respectively. In addition, we put*

$$\pi^{(n)}_{\max} = \max\left\{\pi^{(n)}(x); x \in E_n\right\},$$

$$q^{(n)}_{\max} = \max\left\{q^{(n)}(x); x \in E_n\right\}.$$

(*Recall that the voting rates $q^{(n)}(x)$ are defined in* (1.1).) *Suppose that either of the following conditions is satisfied*:

(i) $\quad \lim_{n \to \infty} \pi^{(n)}_{\mathrm{diag}} = 0 \quad$ *and* $\quad \lim_{n \to \infty} \dfrac{\mathbf{t}^{(n)}_{\mathrm{mix}}}{\mathbf{t}^{(n)}_{\mathrm{meet}}} = 0,$

(ii) *the $q^{(n)}$-Markov chains are reversible and satisfy,*

$$\lim_{n \to \infty} \pi^{(n)}_{\mathrm{diag}} = 0 \quad \text{and} \quad \lim_{n \to \infty} \frac{\log(e \vee \mathbf{t}^{(n)}_{\mathrm{meet}} \sqrt{\pi^{(n)}_{\max} q^{(n)}_{\max} \nu_n(\mathbb{1})})}{\mathbf{g}_n \mathbf{t}^{(n)}_{\mathrm{meet}}} = 0. \tag{2.12}$$

Then for all $u \in [0, 1]$, (2.8) holds with $\gamma_n = \mathbf{t}^{(n)}_{\mathrm{meet}}$, and consequently, (2.7) holds.

Let us make some observation for the condition (ii) of Theorem 2.2. From an inequality (see (3.21)) proved later on, we have

$$\mathbf{t}_{\text{meet}}^{(n)}\pi_{\max}^{(n)}q_{\max}^{(n)} \geq \mathbf{t}_{\text{meet}}^{(n)}\sqrt{\pi_{\max}^{(n)}q_{\max}^{(n)}\nu_n(\mathbb{1})} \geq \frac{(1-\pi_{\text{diag}}^{(n)})^2}{4}. \tag{2.13}$$

Also, it is plain that

$$\lim_{n\to\infty}\pi_{\text{diag}}^{(n)} = 0 \quad \Longleftrightarrow \quad \lim_{n\to\infty}\pi_{\max}^{(n)} = 0. \tag{2.14}$$

Hence if the voting rates $(q^{(n)}(x); x \in E_n)$ are uniformly bounded and $\lim_{n\to\infty}\pi_{\text{diag}}^{(n)} = 0$, then $\lim_{n\to\infty}\mathbf{t}_{\text{meet}}^{(n)} = \infty$, and moreover, $\mathbf{t}_{\text{meet}}^{(n)}$ has order at least $(\pi_{\max}^{(n)})^{-1}$. This, applied to the second part of (2.12), gives the following.

Corollary 2.3. *If the Markov chains defined by $(q^{(n)}, E_n)$ are reversible and satisfy $\lim_{n\to\infty}\pi_{\text{diag}}^{(n)} = 0$,*

$$\limsup_{n\to\infty}\max_{x\in E_n}q^{(n)}(x) < \infty \quad and \quad \liminf_{n\to\infty}\mathbf{g}_n > 0,$$

then the same conclusions of Theorem 2.2 hold. In particular, these conditions hold when $q^{(n)} = p^{(n)} - \text{Id}_{E_n}$ for symmetric probability matrices $p^{(n)}$ (not necessarily with zero diagonals), and the Markov chains defined by $(q^{(n)}, E_n)$ satisfy $\liminf_{n\to\infty}\mathbf{g}_n > 0$.

If the sequence $(\mathbf{t}_{\text{meet}}^{(n)}\pi_{\max}^{(n)}q_{\max}^{(n)})_{n\in\mathbb{N}}$ is bounded above, then plainly the second condition in (2.12) reduces to

$$\lim_{n\to\infty}\mathbf{g}_n\mathbf{t}_{\text{meet}}^{(n)} = \infty. \tag{2.15}$$

This is the condition suggested by Aldous and Fill on almost exponentiality of hitting times in [2], for the particular case of the first meeting time of two independent q-Markov chains (see also Section 1.1 in [23]). Moreover, if $q^{(n)} = p^{(n)} - \text{Id}_{E_n}$ for a probability matrix $p^{(n)}$ and the matrices $p^{(n)}$ satisfy sufficient symmetry (see Chapter 7 in [2] for the notion of *symmetric chains* and note that it is stronger than requiring $p^{(n)}(x, y) = p^{(n)}(y, x)$ for any x, y), then $2\mathbf{t}_{\text{meet}}^{(n)}$ is equal to the so-called *random target time* and so can be expressed explicitly in terms of the eigenvalues of $p^{(n)}$ (Section 4.2 in [2]). In this case, the condition (2.15) only involves the eigenvalues of $-q^{(n)}$.

Remark 2.4. *One notion of "transience" (respectively, "recurrence") for a sequence of finite Markov chains (see Section 15.2.3 in [2]) is essentially that the sequence $(\mathbf{t}_{\text{meet}}^{(n)}\pi_{\max}^{(n)}q_{\max}^{(n)})_{n\in\mathbb{N}}$ be bounded above (respectively, tend to infinity). See Remark 8.1 for more details on this terminology. Theorem 2.2 applies in both cases. In fact, we use considerably more delicate arguments in the present proof of Theorem 2.2, in order to take into account the recurrent case as well.*

Our last results concern coalescence times of Markov chains. Suppose again that we have a sequence of irreducible Q-matrices $(q^{(n)}, E_n)$, with stationary distributions $(\pi^{(n)})$. For a given n, let $U_1, U_2, \ldots$ be i.i.d. with distribution $\pi^{(n)}$. Let $(\hat{X}_t^x, x \in E_n)$ be a system of coalescing $q^{(n)}$-Markov chains, with $\hat{X}_0^x = x$, independent of the U_i's. This means that the $q^{(n)}$-Markov chains $\hat{X}^x$ move independently until they meet, at which time they coalesce and move together. Define the coalescence times

$$\mathsf{C}_{k,j}^{(n)} = \inf\{t > 0; |\{\hat{X}_t^{U_1}, \ldots, \hat{X}_t^{U_k}\}| = j\}, \quad 1 \leq j \leq k \leq |E_n|,$$

and let $Z_2, Z_3, \ldots$ be independent exponential random variables with $\mathbf{E}[Z_j] = 1/\binom{j}{2}$. In the mean-field case, it is well-known and easy to check that with $\gamma_n = |E_n|/2$,

$$\frac{\mathsf{C}_{k,j}^{(n)}}{\gamma_n} \xrightarrow[n\to\infty]{(d)} \sum_{i=j+1}^{k} Z_i, \quad 1 \leq j < k < \infty. \tag{2.16}$$

(See Chapter 14 in [2].) In fact, this convergence is an easy consequence of the convergence of voter model densities to the Wright–Fisher diffusion.

Proposition 2.5. *If* (2.7) *holds, then so does* (2.16). *In particular, if either of the conditions of Theorem 2.2 hold, then so does* (2.16) *with* $\gamma_n = \mathbf{t}_{\text{meet}}^{(n)}$.

We refer the readers to [23] and [24] for recent results on the almost exponentiality of Markov chain hitting times of general sets, and in particular, of Markov chain coalescence times. These results give the convergence in (2.16) with explicit convergence rates under slightly different conditions than the ones we give here. Remarkably, the convergence of the "full" coalescence times $\hat{\mathsf{C}}_1^{(n)}$ of $\{\hat{X}^x; x \in E_n\}$ is also obtained in [23], where

$$\hat{\mathsf{C}}_j^{(n)} = \inf\{t \geq 0; |\{\hat{X}_t^x; x \in E_n\}| = j\}, \quad 1 \leq j \leq |E_n|.$$

In this direction, we also have Proposition 2.6 below, which interprets the convergence of full coalescence times in terms of the convergence of voter densities to the Wright–Fisher diffusion.

Proposition 2.6. *Let* $\tau_1^{(n)}$ *denote the first hitting time of* 1 *by the density process* $(p_1(\xi_{\gamma_n t}))$, *and* τ_1^Y *the first hitting time of* 1 *by the Wright–Fisher diffusion* (Y_t). *Then the following convergences are equivalent*:

$$\left(\frac{\tau_1^{(n)}}{\gamma_n}, \mathbb{P}_{\mu_u}^{(n)}\right) \xrightarrow[n \to \infty]{(d)} (\tau_1^Y, \mathbb{P}_u), \quad \forall u \in [0, 1], \tag{2.17}$$

$$\frac{\hat{\mathsf{C}}_j^{(n)}}{\gamma_n} \xrightarrow[n \to \infty]{(d)} \sum_{i=j+1}^{\infty} Z_i, \quad \forall j \in \mathbb{N}. \tag{2.18}$$

We note that the convergence (2.17) does not follow immediately from the weak convergence of density processes since first hitting times are in general not continuous with respect to the Skorokhod J_1-topology. To see this, we may reinforce the convergence (2.7) to almost-sure convergence in the Skorokhod J_1-topology by the Skorokhod representation (see [13]). Then, for example, the approximating density processes $(p_1(\xi_{\gamma_n \cdot}), \mathbb{P}_{\mu_u}^{(n)})$ may "linger" very close to the absorbing state 1 for long periods of time before getting absorbed at 1, while the limiting process (Y_t) has already reached 1. Hence, (2.17) rules out this lingering behavior of the density process $(p_1(\xi_{\gamma_n \cdot}), \mathbb{P}_{\mu_u}^{(n)})$ for all large n in particular.

3. Martingale property and duality

Fix a Markov chain defined by (q, E) with stationary distribution π, and consider the corresponding voter model (ξ_t). Recall the definition (1.4) of p_1, and set

$$p_0 \equiv 1 - p_1.$$

In this section, we identify some martingales associated with the density process $(p_1(\xi_t))$ and then resort to the duality equation for voter models (see (3.7) below) for their interpretations in terms of coalescing Markov chains.

Proposition 3.1. *For any initial configuration* $\xi \in \{0, 1\}^E$, *all of the following three processes are* $\mathbb{P}_\xi$-*martingales*:

  (i) $(p_1(\xi_t))$.
  (ii) $(p_1(\xi_t) p_0(\xi_t) + \nu(\mathbb{1}) \int_0^t [p_{10}(\xi_s) + p_{01}(\xi_s)] \, ds)$.
  (iii) $(p_1^2(\xi_t) - \nu(\mathbb{1}) \int_0^t [p_{10}(\xi_s) + p_{01}(\xi_s)] \, ds)$.

Proof. Recall that the generator $\mathcal{L}$ and the flip rates of the voter model (ξ_t) are given by (1.2) and (1.3), respectively. In the following, we will show

$$\mathcal{L}p_1 \equiv 0, \tag{3.1}$$

$$\mathcal{L}(p_1 p_0) \equiv -\nu(\mathbb{1})(p_{10} + p_{01}). \tag{3.2}$$

Then our assertions for the processes in (i) and (ii) follow from these and a standard result of Markov processes. The fact that the process in (iii) is a martingale then follows from the analogous properties of the processes in (i) and (ii), since $p_1^2 = p_1 - p_1 p_0$.

We first show (3.1). Plainly

$$p_1(\xi^x) - p_1(\xi) = \pi(x)[\widehat{\xi}(x) - \xi(x)], \tag{3.3}$$

and thus by (1.2) we get

$$\begin{aligned}
\mathcal{L}p_1(\xi) &= \sum_{x \in E} \widehat{\xi}(x) \sum_{y \in E} \xi(y)q(x, y)\pi(x) - \sum_{x \in E} \xi(x) \sum_{y \in E} \widehat{\xi}(y)q(x, y)\pi(x) \\
&= \sum_{x, y \in E} \xi(y)\pi(x)q(x, y) - \sum_{x, y \in E} \xi(x)\xi(y)\pi(x)q(x, y) \\
&\quad - \sum_{x, y \in E} \xi(x)\pi(x)q(x, y) + \sum_{x, y \in E} \xi(x)\xi(y)\pi(x)q(x, y) = 0,
\end{aligned}$$

because $\sum_{y \in E} q(x, y) = 0$ and $\sum_{x \in E} \pi(x)q(x, y) = 0$ for all $y \in E$. Hence, (3.1) follows, and the density process $(p_1(\xi_t))$ is a martingale.

Next, to show (3.2), we note that for any $x \in E$,

$$\begin{aligned}
p_1(\xi^x)&p_0(\xi^x) - p_1(\xi)p_0(\xi) \\
&= [p_1(\xi^x) - p_1(\xi)] \cdot [p_0(\xi^x) - p_0(\xi)] + p_0(\xi)[p_1(\xi^x) - p_1(\xi)] \\
&\quad + p_1(\xi)[p_0(\xi^x) - p_0(\xi)] \\
&= -\pi(x)^2 + p_0(\xi)[p_1(\xi^x) - p_1(\xi)] + p_1(\xi)[p_0(\xi^x) - p_0(\xi)],
\end{aligned}$$

where we have used (3.3) and the analogue $p_0(\xi^x) - p_0(\xi) = \pi(x)[\xi(x) - \widehat{\xi}(x)]$ in the last line. Since $\mathcal{L}p_1 = \mathcal{L}p_0 = 0$, the last equality implies that

$$\begin{aligned}
\sum_{x \in E} c(x, \xi)&[p_1(\xi^x)p_0(\xi^x) - p_1(\xi)p_0(\xi)] \\
&= -\sum_{x, y \in E} [\xi(x)\widehat{\xi}(y) + \widehat{\xi}(x)\xi(y)]q(x, y)\pi(x)^2 \\
&= -\nu(\mathbb{1})[p_{10}(\xi) + p_{01}(\xi)],
\end{aligned}$$

where the last equality follows from the definitions (2.1)–(2.4). This gives (3.2), and our assertion for (ii) is proved. The proof is complete. $\qquad\square$

Recall that μ_u denotes the Bernoulli product measure on E with density $\mu_u(\xi(x) = 1) = u$.

Corollary 3.2. *For any $\gamma, t \in (0, \infty)$ and initial configuration $\xi \in \{0, 1\}^E$, the martingale $(p_1(\xi_{\gamma t}))$ under $\mathbb{P}_\xi$ has predictable quadratic variation process*

$$\langle p_1(\xi_{\gamma \cdot}) \rangle_t = \gamma \nu(\mathbb{1}) \int_0^t [p_{10}(\xi_{\gamma s}) + p_{01}(\xi_{\gamma s})] \, ds. \tag{3.4}$$

Also for any $u \in [0, 1]$, we have

$$\mathbb{E}_{\mu_u}\big[p_1(\xi_{\gamma t})p_0(\xi_{\gamma t})\big] = u(1-u)(1-\pi_{\mathrm{diag}}) - \gamma\nu(\mathbb{1})\int_0^t \mathbb{E}_{\mu_u}\big[p_{10}(\xi_{\gamma s}) + p_{01}(\xi_{\gamma s})\big]\,ds. \tag{3.5}$$

Proof. The equation (3.4) follows readily from Proposition 3.1 for the process in (iii) and the standard characterization of predictable quadratic variations (cf. [15]). Similarly, by (ii) in the same proposition, we have

$$\mathbb{E}_{\xi}\big[p_1(\xi_{\gamma t})p_0(\xi_{\gamma t})\big] = p_1(\xi)p_0(\xi) - \gamma\nu(\mathbb{1})\int_0^t \mathbb{E}_{\xi}\big[p_{10}(\xi_{\gamma s}) + p_{01}(\xi_{\gamma s})\big]\,ds, \tag{3.6}$$

and so a randomization of the initial configuration ξ by μ_u leads to (3.5). $\qquad\square$

The rest of this section is devoted to interpreting the above results by coalescing Markov chains, and now we recall duality. Using the coalescing Markov chains $(\hat{X}^x, x \in E)$ introduced in Section 2, we can formulate the duality equation for voter models (see Chapter V of [18] or Section 4.3 of [19]) as

$$\mathbb{E}_{\eta}\bigg[\prod_{x \in F} \xi_t(x)\bigg] = \mathbf{E}\bigg[\prod_{x \in F} \eta\big(\hat{X}^x_t\big)\bigg] \quad \forall \eta \in \{0, 1\}^E, t \in \mathbb{R}_+ \tag{3.7}$$

for any nonempty subset F of E. The readers will see later on that the duality formula becomes particularly tractable for a voter model with initial law μ_u.

We will make frequent use of a special case of (3.7) stated as follows. For convenience, let $(X^x_t, x \in E)$ be another system of q-Markov chains with Q-matrix q and $X^x_0 = x$, but now consisting of independent chains. We define the first meeting times of X^x and X^y by

$$M_{x,y} = \inf\big\{t \ge 0 \colon X^x_t = X^y_t\big\}, \quad x, y \in E.$$

Then (3.7) implies

$$\mathbb{E}_{\xi}\big[\xi_t(x)\widehat{\xi}_t(y)\big] = \mathbf{E}\big[\xi\big(X^x_t\big)\widehat{\xi}\big(X^y_t\big); M_{x,y} > t\big]. \tag{3.8}$$

Next, we recall that (U, U') has law $\pi \otimes \pi$, and now introduce (V, V') with law

$$\mathbf{P}\big(V = a, V' = b\big) \equiv \bar{\nu}(a, b), \quad a, b \in E \tag{3.9}$$

(recall the definition of $\bar{\nu}$ from (2.2)). We assume, in addition, that these random elements (U, U') and (V, V') are independent of the system $(X^x; x \in E)$.

Proposition 3.3. *For any $\gamma, t > 0$ and initial configuration $\xi \in \{0, 1\}^E$,*

$$\mathbb{E}_{\xi}\big[p_1(\xi_{\gamma t})p_0(\xi_{\gamma t})\big] = \mathbf{E}\big[\xi\big(X^U_{\gamma t}\big)\widehat{\xi}\big(X^{U'}_{\gamma t}\big); M_{U,U'} > \gamma t\big], \tag{3.10}$$

$$\mathbb{E}_{\xi}\big[p_{10}(\xi_{\gamma t})\big] = \mathbf{E}\big[\xi\big(X^V_{\gamma t}\big)\widehat{\xi}\big(X^{V'}_{\gamma t}\big); M_{V,V'} > \gamma t\big], \tag{3.11}$$

$$\mathbb{E}_{\xi}\big[p_{01}(\xi_{\gamma t})\big] = \mathbf{E}\big[\widehat{\xi}\big(X^V_{\gamma t}\big)\xi\big(X^{V'}_{\gamma t}\big); M_{V,V'} > \gamma t\big]. \tag{3.12}$$

Proof. By the duality equation (3.8) and the definitions of p_1, p_0 and (U, U'), we have

$$\mathbb{E}_{\xi}\big[p_1(\xi_{\gamma t})p_0(\xi_{\gamma t})\big] = \sum_{x,y \in E} \pi(x)\pi(y)\mathbf{E}\big[\xi\big(X^x_{\gamma t}\big)\widehat{\xi}\big(X^y_{\gamma t}\big); M_{x,y} > t\big]$$

$$= \mathbf{E}\big[\xi\big(X^U_{\gamma t}\big)\widehat{\xi}\big(X^{U'}_{\gamma t}\big); M_{U,U'} > \gamma t\big],$$

which proves (3.10). The equations (3.11) and (3.12) can be derived in the same fashion by using the definition (3.9) of (V, V'). $\qquad\square$

We point out that (3.10)–(3.12) are closely related to the tail distributions of some particular meeting times. By (3.11) and (3.12), we have

$$\sup_{\xi \in \{0,1\}^E} \mathbb{E}_\xi \big[p_{10}(\xi_{\gamma t}) + p_{01}(\xi_{\gamma t}) \big] \le 2\mathbf{P}(M_{V,V'} > \gamma t). \tag{3.13}$$

Moreover, if we start the voter model with the product measure μ_u for $u \in [0, 1]$, then Proposition 3.3 implies

$$\mathbb{E}_{\mu_u} \big[p_1(\xi_{\gamma t}) p_0(\xi_{\gamma t}) \big] = u(1 - u)\mathbf{P}(M_{U,U'} > \gamma t) \tag{3.14}$$

and

$$\mathbb{E}_{\mu_u} \big[p_{10}(\xi_{\gamma t}) \big] = \mathbb{E}_{\mu_u} \big[p_{01}(\xi_{\gamma t}) \big] = u(1 - u)\mathbf{P}(M_{V,V'} > \gamma t). \tag{3.15}$$

As a particular application of (3.14) and (3.15), we give simple proofs for some known results in Markov chain theory in Corollary 3.4 below (see Section 5.3 of Chapter 3 in [2]).

Corollary 3.4. *The tail distributions of $M_{U,U'}$ and $M_{V,V'}$ are related by the formula: for any $\gamma, t > 0$,*

$$\mathbf{P}(M_{U,U'} > \gamma t) = 1 - \pi_{\mathrm{diag}} - 2\gamma \nu(\mathbb{1}) \int_0^t \mathbf{P}(M_{V,V'} > \gamma s)\, ds. \tag{3.16}$$

Moreover, we have

$$\mathbf{E}[M_{V,V'}] = \frac{1 - \pi_{\mathrm{diag}}}{2\nu(\mathbb{1})}, \tag{3.17}$$

$$\mathbf{E}[M_{U,U'}] = \nu(\mathbb{1})\mathbf{E}\big[M_{V,V'}^2\big]. \tag{3.18}$$

Proof. We start with (3.16). If we fix $u \in (0, 1)$, and plug (3.14) and (3.15) into (3.5), then cancelling the factor $u(1 - u)$ gives (3.16). We remark that (3.16) can be alternatively derived by a standard Markov chain "last time" decomposition (see Section A.2 of [7]), and leave the details to the readers.

We then consider the two equalities (3.17) and (3.18). Since q is irreducible, the meeting time $M_{x,y}$ is finite a.s. for any $x, y \in E$. Thus, by setting $\gamma = 1$ and passing $t \longrightarrow \infty$ in the identity (3.16), we deduce (3.17). To obtain the second equality (3.18), we set $\gamma = 1$ and integrate both sides of (3.16):

$$\begin{aligned}
\mathbf{E}[M_{U,U'}] &= 2\nu(\mathbb{1}) \int_0^\infty \left(\frac{1 - \pi_{\mathrm{diag}}}{2\nu(\mathbb{1})} - \int_0^t \mathbf{P}(M_{V,V'} > s)\, ds \right) dt \\
&= 2\nu(\mathbb{1}) \int_0^\infty \int_t^\infty \mathbf{P}(M_{V,V'} > s)\, ds\, dt \\
&= 2\nu(\mathbb{1}) \int_0^\infty s\mathbf{P}(M_{V,V'} > s)\, ds \\
&= 2\nu(\mathbb{1}) \frac{\mathbf{E}[M_{V,V'}^2]}{2} = \nu(\mathbb{1})\mathbf{E}\big[M_{V,V'}^2\big],
\end{aligned}$$

where (3.17) is used in the second equality below. We have proved (3.18). The proof is complete. □

Remark 3.5. (1) *Some useful consequences of Corollary 3.4 are the following. First, (3.17) and Markov's inequality imply that for any $\gamma, t > 0$,*

$$2\gamma \nu(\mathbb{1})\mathbf{P}(M_{V,V'} > \gamma t) \le \frac{1 - \pi_{\mathrm{diag}}}{t}. \tag{3.19}$$

Second, passing $t \longrightarrow \infty$ in (3.16), we obtain

$$2\gamma \nu(\mathbb{1}) \int_0^\infty \mathbf{P}(M_{V,V'} > \gamma s)\, \mathrm{d}s \leq 1. \tag{3.20}$$

Finally, from (3.17), (3.18) and the Cauchy–Schwarz inequality we obtain a useful lower bound of $\mathbf{t}_{\mathrm{meet}} = \mathbf{E}[M_{U,U'}]$:

$$\mathbf{t}_{\mathrm{meet}} \geq \frac{1}{\nu(\mathbb{1})} \left(\frac{1 - \pi_{\mathrm{diag}}}{2} \right)^2. \tag{3.21}$$

See also Section 5.1 of [2] for a similar inequality.

(2) *If q is of the form (1.7) for a symmetric probability matrix p with zero diagonal, then $\nu(\mathbb{1}) = 1/|E|$ and $\mathbf{P}((V, V') = (a, b)) = \pi(a)q(a, b)$. In this case, (3.17) and (3.18) reduce to*

$$\mathbf{E}[M_{V,V'}] = \frac{|E| - 1}{2} \quad \text{and} \quad \mathbf{E}\big[M_{V,V'}^2\big] = |E| \cdot \mathbf{E}[M_{U,U'}],$$

respectively.

4. Pairwise coalescence times

Throughout this section we take an arbitrary sequence of irreducible Markov chains defined by Q-matrices $(q^{(n)}, E_n)_{n \in \mathbb{N}}$. With $\pi^{(n)}$ being the stationary distribution of $q^{(n)}$, we write

$$\pi_{\mathrm{diag}}^{(n)} = \sum_{x \in E_n} \pi^{(n)}(x)^2,$$

$$\nu_n(x, y) \equiv \pi^{(n)}(x)^2 q^{(n)}(x, y)\mathbb{1}_{x \neq y}, \qquad \bar{\nu}_n = \frac{\nu_n}{\nu_n(\mathbb{1})}$$

as before. Let (ξ_s) with law $\mathbb{P}_\lambda^{(n)}$ denote the voter model defined by the voting kernel $q^{(n)}$ with initial distribution λ. By convention, $\mathbb{P}_\xi^{(n)} = \mathbb{P}_{\delta_\xi}^{(n)}$ for delta measures δ_ξ.

In this section, we consider the density processes of these voter models and study the necessary and sufficient conditions for the convergence of their second moments to the second moment of the Wright–Fisher diffusion $(Y, (\mathbb{P}_u)_{u \in [0,1]})$ which is defined by the differential operator $\mathcal{G}$ in (1.5). Our main result in this section is Theorem 4.1 below. In the following, let $\mathbf{e}$ denote the exponential random variable with mean 1, and $\mathscr{L}(X)$ denote the law of a random element X.

Theorem 4.1. *Assume that*

$$\lim_{n \to \infty} \pi_{\mathrm{diag}}^{(n)} = \Delta \in [0, 1), \tag{4.1}$$

and let (γ_n) be a sequence of constants in $(0, \infty)$. Then the following conditions are equivalent.

(1) *For some $u \in (0, 1)$,*

$$\lim_{n \to \infty} \mathbb{E}_{\mu_u}^{(n)}\big[p_1(\xi_{\gamma_n t}) p_0(\xi_{\gamma_n t})\big] = (1 - \Delta)\mathbb{E}_u\big[Y_t(1 - Y_t)\big] \quad \forall t \in \mathbb{R}_+. \tag{4.2}$$

(2) *For all $t \in \mathbb{R}_+$,*

$$\lim_{n \to \infty} 2\gamma_n \nu_n(\mathbb{1}) \int_0^t \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\, \mathrm{d}s = (1 - \Delta)\big(1 - \mathrm{e}^{-t}\big).$$

(3) *For all $\mu \in \mathbb{R}_+$,*

$$\lim_{n \to \infty} 2\gamma_n \nu_n(\mathbb{1}) \mathbf{E}^{(n)}\big[1 - e^{-\mu M_{V,V'}/\gamma_n}\big] = (1 - \Delta)\mathbb{E}\big[1 - e^{-\mu \mathbf{e}}\big].$$

(4) $\quad \mathscr{L}\left(\dfrac{M_{U,U'}}{\gamma_n}\right) \xrightarrow[n \to \infty]{\text{(d)}} (1 - \Delta) \cdot \mathscr{L}(\mathbf{e}) + \Delta \cdot \delta_0.$

Moreover, if any of these four conditions holds, then (4.2) *holds for any $u \in [0, 1]$.*

Proof. We will prove this theorem in the order: $(2) \Longleftrightarrow (4)$, $(1) \Longleftrightarrow (2)$, and finally $(2) \Longleftrightarrow (3)$.

Step 1: $(2) \Longleftrightarrow (4)$. Note that (4) is equivalent to

$$\mathbf{P}^{(n)}(M_{U,U'} > \gamma_n t) \longrightarrow (1 - \Delta)e^{-t}, \quad \forall t > 0,$$

and so it follows immediately from (3.16) and (4.1) that (2) and (4) are equivalent.

Step 2: $(1) \Longleftrightarrow (2)$. Suppose that (1) holds for some $u \in (0, 1)$. Note that

$$\mathbb{E}_u\big[Y_t(1 - Y_t)\big] = u(1 - u)e^{-t}, \quad t \in \mathbb{R}_+.$$

Using the foregoing equality, (3.14) and (3.16), we see that (1) implies

$$(1 - \Delta)u(1 - u)e^{-t} = \lim_{n \to \infty} \mathbb{E}^{(n)}_{\mu_u}\big[p_1(\xi_{\gamma_n t}) p_0(\xi_{\gamma_n t})\big]$$

$$= u(1 - u)\bigg[(1 - \Delta) - \lim_{n \to \infty} 2\gamma_n \nu_n(\mathbb{1}) \int_0^t \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\,ds\bigg].$$

Cancelling out the factor $u(1 - u)$ on both sides of the foregoing equality, we obtain (2). For the converse, we take any $u \in (0, 1)$ and then reverse this argument. In particular, (4.2) holds for all $u \in [0, 1]$ whenever (2) applies.

Step 3: $(2) \Longleftrightarrow (3)$. Let us make some elementary observations. First, for any $(0, \infty)$-valued random variable X and any $\mu > 0$, it is elementary to obtain

$$\mathbf{E}\big[1 - e^{-\mu X}\big] = \mu \int_0^\infty e^{-\mu s}\mathbf{P}(X > s)\,ds \tag{4.3}$$

$$= \mu^2 \int_0^\infty e^{-\mu t} \int_0^t \mathbf{P}(X > s)\,ds\,dt. \tag{4.4}$$

In addition, (3.20) gives

$$2\gamma_n \nu_n(\mathbb{1}) \int_0^\infty \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\,ds \le 1, \quad \forall n \in \mathbb{N}. \tag{4.5}$$

Now assume that (2) holds. Taking $X = M_{V,V'}/\gamma_n$ in (4.4), we have for any $\mu > 0$

$$2\gamma_n \nu_n(\mathbb{1})\mathbf{E}^{(n)}\big[1 - e^{-\mu M_{V,V'}/\gamma_n}\big] = \mu^2 \int_0^\infty e^{-\mu t} \int_0^t 2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\,ds\,dt. \tag{4.6}$$

We pass $n \to \infty$ for both sides of the foregoing equality. The bound (4.5) justifying the use of the dominated convergence theorem, the limit of the right-hand side of (4.6) equals

$$(1 - \Delta)\mu^2 \int_0^\infty e^{-\mu t}\big(1 - e^{-t}\big)\,dt = (1 - \Delta)\mathbb{E}\big[1 - e^{-\mu \mathbf{e}}\big],$$

where the last equality follows from (4.4) with $X = \mathbf{e}$. We have proved (3).

The proof that (3) implies (2) is more involved. Employing (4.3) again, we see that (3) implies that for all $\lambda > 0$,

$$2\gamma_n \nu_n(\mathbb{1}) \int_0^\infty e^{-\lambda s}\mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\,ds \to (1 - \Delta) \int_0^\infty e^{-\lambda s}\mathbb{P}(\mathbf{e} > s)\,ds \tag{4.7}$$

as $n \to \infty$. For any $\mu > 0$ and $t > 0$, define

$$f_{n,\mu}(t) = \frac{\mathrm{e}^{-\mu t}\mathbf{P}^{(n)}(M_{V,V'} > \gamma_n t)}{\int_0^\infty \mathrm{e}^{-\mu s}\mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\,\mathrm{d}s},$$

$$f_\mu(t) = \frac{\mathrm{e}^{-\mu t}\mathbb{P}(\mathbf{e} > t)}{\int_0^\infty \mathrm{e}^{-\mu s}\mathbb{P}(\mathbf{e} > s)\,\mathrm{d}s}.$$

Applying (4.7) twice, we obtain for any $\lambda > 0$,

$$\int_0^\infty \mathrm{e}^{-\lambda t} f_{n,\mu}(t)\,\mathrm{d}t \; = \; \frac{2\gamma_n \nu_n(\mathbb{1})\int_0^\infty \mathrm{e}^{-(\lambda+\mu)t}\mathbf{P}^{(n)}(M_{V,V'} > \gamma_n t)\,\mathrm{d}t}{2\gamma_n \nu_n(\mathbb{1})\int_0^\infty \mathrm{e}^{-\mu s}\mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\,\mathrm{d}s}$$

$$\longrightarrow \int_0^\infty \mathrm{e}^{-\lambda t} f_\mu(t)\,\mathrm{d}t$$

as $n \to \infty$. Hence, we deduce from Lévy's continuity theorem for Laplace transforms of distributions on $\mathbb{R}_+$ (cf. Theorem 4.3 in [16]) and (4.7) with λ replaced by $\mu > 0$ that for all $t \geq 0$,

$$2\gamma_n \nu_n(\mathbb{1}) \int_0^t \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\,\mathrm{d}s \; \geq \; 2\gamma_n \nu_n(\mathbb{1}) \int_0^t \mathrm{e}^{-\mu s}\mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\,\mathrm{d}s$$

$$\to (1 - \Delta) \int_0^t \mathrm{e}^{-\mu s}\mathbb{P}(\mathbf{e} > s)\,\mathrm{d}s \tag{4.8}$$

as $n \to \infty$. Since $\mu > 0$ is arbitrary, this implies

$$\liminf_{n\to\infty} 2\gamma_n \nu_n(\mathbb{1}) \int_0^t \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\,\mathrm{d}s \geq (1 - \Delta)\int_0^t \mathbb{P}(\mathbf{e} > s)\,\mathrm{d}s. \tag{4.9}$$

To prove the converse inequality, we start with the decomposition

$$2\gamma_n \nu_n(\mathbb{1}) \int_0^t \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\,\mathrm{d}s = 2\gamma_n \nu_n(\mathbb{1}) \int_0^t \left(1 - \mathrm{e}^{-\mu s}\right)\mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\,\mathrm{d}s$$

$$+ 2\gamma_n \nu_n(\mathbb{1}) \int_0^t \mathrm{e}^{-\mu s}\mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\,\mathrm{d}s. \tag{4.10}$$

Fix any $\mu > 0$. By Markov's inequality and the elementary fact that $1 - \mathrm{e}^{-\mu s} \leq \mu s$ if $\mu s \geq 0$, the first integral on the right-hand side above is bounded by

$$2\gamma_n \nu_n(\mathbb{1}) \int_0^t \mu \frac{\mathbf{E}^{(n)}[M_{V,V'}]}{\gamma_n}\,\mathrm{d}s \leq \mu t,$$

where the last inequality is a consequence of (3.17). Applying the foregoing inequality to (4.10) and using (4.8), we obtain

$$\limsup_{n\to\infty} 2\gamma_n \nu_n(\mathbb{1}) \int_0^t \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\,\mathrm{d}s \leq \mu t + (1 - \Delta)\int_0^t \mathrm{e}^{-\mu s}\mathbb{P}(\mathbf{e} > s)\,\mathrm{d}s.$$

If we let $\mu \longrightarrow 0$ in the above inequality and then combine the result with (4.9), we obtain

$$2\gamma_n \nu_n(\mathbb{1}) \int_0^t \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\,\mathrm{d}s \longrightarrow (1 - \Delta)\int_0^t \mathbb{P}(\mathbf{e} > s)\,\mathrm{d}s, \tag{4.11}$$

which is (2). The proof of the theorem is now complete. $\qquad\square$

Corollary 4.2. *Under the assumption* (4.1), *any of* (1)–(4) *of Theorem* 4.1 *implies that*

$$\lim_{n\to\infty} 2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}(M_{V,V'} > \gamma_n t) = (1-\Delta)\mathrm{e}^{-t}, \quad \forall t > 0. \tag{4.12}$$

If in addition the limit $m_\Delta = \lim_{n\to\infty} 2\gamma_n \nu_n(\mathbb{1})$ *exists, then* $m_\Delta \in [1-\Delta, +\infty]$ *and*

$$\mathscr{L}\left(\frac{M_{V,V'}}{\gamma_n}\right) \xrightarrow[n\to\infty]{\mathrm{(d)}} \frac{1-\Delta}{m_\Delta}\mathscr{L}(\mathbf{e}) + \left(1 - \frac{1-\Delta}{m_\Delta}\right)\delta_0$$

with the convention that $\frac{1}{+\infty} = 0$.

Proof. We prove (4.12), from which the second assertion immediately follows. We may assume that (2) of Theorem 4.1 holds. For each $n \geq 1$, define

$$f_n(t) = 2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}(M_{V,V'} > \gamma_n t), \quad t \in (0, \infty).$$

Then each f_n is continuous and decreasing. Moreover, by (3.19),

$$0 \leq f_n(t) \leq \frac{1}{t}.$$

Now fix $a > 0$ and define $G_n(t) = 1 - af_n(t)$, $t \in [a, \infty)$. By the above inequality, (G_n) is a sequence of distribution functions on $[a, \infty)$. Hence by Helly's selection principle, there exist a subsequence (G_{n_k}) and some (sub-)distribution function G such that $G_{n_k}(t) \longrightarrow G(t)$ for every continuity point $t \in (a, \infty)$ of G. Since G is monotone, it can have only countably many discontinuity points, and hence for any $a < s < t$,

$$\lim_{k\to\infty}\int_s^t G_{n_k}(u)\,\mathrm{d}u = \int_s^t G(u)\,\mathrm{d}u$$

by dominated convergence. It then follows from (2) of Theorem 4.1 that

$$\int_s^t G(u)\,\mathrm{d}u = (t - s) + a(1-\Delta)\big(\mathrm{e}^{-t} - \mathrm{e}^{-s}\big),$$

which implies that $G(t) = 1 - a(1-\Delta)\mathrm{e}^{-t}$ for every continuity point $t \in (a, \infty)$ of G. Since G is increasing, this equality holds for any $t \in (a, \infty)$. Therefore, G is continuous on (a, ∞) and we have

$$\lim_{k\to\infty} G_{n_k}(t) = 1 - a(1-\Delta)\mathrm{e}^{-t} \quad \text{for any } t \in (a, \infty).$$

As the limit does not depend on the subsequence, this proves that

$$\lim_{n\to\infty} f_n(t) = (1-\Delta)\mathrm{e}^{-t} \quad \text{for any } t \in (a, \infty).$$

Since $a > 0$ is arbitrary, we have proved (4.12).   □

We now study what is left out in the conclusion of Corollary 4.2 and consider, informally, the instant s_n after which the tail of $2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}(M_{V,V'}/\gamma_n \in \cdot)$ starts to behave like the $(1-\Delta)$ multiple of the standard exponential distribution. The following result will play a crucial role in the proof of Theorem 2.2.

Proposition 4.3. *Suppose that* (4.1) *and any of* (1)–(4) *of Theorem* 4.1 *holds.*

(1) *Let* $(s_n) \subseteq \mathbb{R}_+$ *be any sequence such that*

$$\liminf_{n\to\infty} 2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}(M_{V,V'} > s_n) \geq 1 - \Delta. \tag{4.13}$$

Then $s_n = \mathrm{o}(\gamma_n)$ *as* $n \longrightarrow \infty$.

(2) *Let $(s_n) \subseteq \mathbb{R}_+$ be any sequence such that*

$$\lim_{n \to \infty} 2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}(M_{V,V'} > s_n) = 1 - \Delta. \tag{4.14}$$

If (s_n') is a sequence in $\mathbb{R}_+$ such that $s_n' \geq s_n$ and $s_n' = \mathrm{o}(\gamma_n)$, then (4.14) holds with (s_n') in place of (s_n).

Proof. Consider (1) first, and we may assume that (4) of Theorem 4.1 holds. Assume the converse that s_n/γ_n does not converge to zero. By passing to a subsequence if necessary, we may assume without loss of generality that (s_n) satisfies $s_n/\gamma_n \longrightarrow \delta$ for some $\delta \in (0, \infty]$ as $n \longrightarrow \infty$. By assumption,

$$1 - \Delta \leq \liminf_{n \to \infty} 2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}(M_{V,V'} > s_n)$$

$$\leq \lim_{n \to \infty} 2\gamma_n \nu_n(\mathbb{1}) \frac{1 - \pi_{\mathrm{diag}}^{(n)}}{2s_n \nu_n(\mathbb{1})} = (1 - \Delta) \lim_{n \to \infty} \frac{\gamma_n}{s_n},$$

where the second inequality is due to (3.19). Hence, we must have $\delta \leq 1$. On the other hand, by (3.16) of Corollary 3.4,

$$\mathbf{P}^{(n)}\left(\frac{M_{U,U'}}{\gamma_n} > \frac{s_n}{\gamma_n}\right) = 1 - \pi_{\mathrm{diag}}^{(n)} - 2\gamma_n \nu_n(\mathbb{1}) \int_0^{s_n/\gamma_n} \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s) \, \mathrm{d}s. \tag{4.15}$$

Using (4) of Theorem 4.1, we get

$$\lim_{n \to \infty} \mathbf{P}^{(n)}\left(\frac{M_{U,U'}}{\gamma_n} > \frac{s_n}{\gamma_n}\right) = (1 - \Delta)\mathrm{e}^{-\delta}.$$

Apply this to (4.15), and we obtain

$$(1 - \Delta)\left(1 - \mathrm{e}^{-\delta}\right) = \lim_{n \to \infty} 2\gamma_n \nu_n(\mathbb{1}) \int_0^{s_n/\gamma_n} \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s) \, \mathrm{d}s$$

$$\geq \liminf_{n \to \infty} 2\gamma_n \nu_n(\mathbb{1})\left(\frac{s_n}{\gamma_n}\right)\mathbf{P}^{(n)}(M_{V,V'} > s_n)$$

$$\geq (1 - \Delta)\delta$$

by the definition of δ and (4.13). As a consequence, $1 - \mathrm{e}^{-\delta} \geq \delta$, whereas it is easy to see that $1 - \mathrm{e}^{-\delta'} < \delta'$ as long as $\delta' > 0$. This proves that δ must be 0, so (1) follows.

To prove (2), we let $(s_n') \subset \mathbb{R}_+$ satisfy $s_n' \geq s_n$ and $s_n' = \mathrm{o}(\gamma_n)$. It is immediate that

$$\limsup_{n \to \infty} 2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}\left(M_{V,V'} > s_n'\right)$$

$$\leq \lim_{n \to \infty} 2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}(M_{V,V'} > s_n) = 1 - \Delta \tag{4.16}$$

by the present assumption (4.14). To obtain the converse inequality, we fix $\varepsilon > 0$. Since $\frac{s_n'}{\gamma_n} < \varepsilon$ for all large enough n,

$$2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}\left(M_{V,V'} > s_n'\right)$$

$$\geq 2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}(M_{V,V'} > \varepsilon\gamma_n) \longrightarrow (1 - \Delta)\mathrm{e}^{-\varepsilon}$$

as $n \longrightarrow \infty$ by our assumption on the validity of any of (1)–(4) in Theorem 4.1 and Corollary 4.2. Since $\varepsilon > 0$ is arbitrary, we deduce that

$$\liminf_{n \to \infty} 2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}\left(M_{V,V'} > s_n'\right) \geq 1 - \Delta. \tag{4.17}$$

We now get the asserted equality (4.14) for (s_n') from (4.16) and (4.17). The proof is complete. $\qquad \square$

5. Proof of Theorem 2.1

In this section, we prove limit theorems for density processes. We will focus on the martingale property of the density processes and use semimartingale limit theorems for our purpose. As before, we take a sequence of irreducible Q-matrices $(q^{(n)}, E_n)$ with stationary distributions $(\pi^{(n)})$ and a sequence of strictly positive constants (γ_n).

We first introduce some notation for density processes used throughout this section. For each n, we write $Y_n = (Y_n(t))$ for the density processes $(p_1(\xi_{\gamma_n t}))$ of the voter model defined by $q^{(n)}$. By Proposition 3.1, each Y_n is a càdlàg $(\mathscr{F}_t^n)$-martingale, where

$$\mathscr{F}_t^n = \sigma(\xi_{\gamma_n s}; s \le t). \tag{5.1}$$

We recall from (3.4) that the predictable quadratic variation process of Y_n is given by the continuous process

$$\langle Y_n \rangle_t = \gamma_n \nu_n(\mathbb{1}) \int_0^t \left[p_{10}(\xi_{\gamma_n s}) + p_{01}(\xi_{\gamma_n s}) \right] ds. \tag{5.2}$$

Note that the process in (5.2) is different from the quadratic variation process, which is given by

$$[Y_n]_t = \sum_{s:s \le t} \left(\Delta Y_n(s) \right)^2$$

(see [15]).

In the following theorem, we refer to [15] for the necessary stochastic calcuus definitions and theorems.

Theorem 5.1. *Assume that* (4.1) *holds with* $\Delta = 0$ *and any of* (1)–(4) *of Theorem* 4.1 *holds. Then we have the following.*

(1) *For any* $u \in [0, 1]$, *the sequence of laws of the càdlàg martingales*

$$\left(Y_n, \mathbb{P}_{\mu_u}^{(n)} \right), \quad n \in \mathbb{N} \tag{5.3}$$

   is C-tight.
(2) *For any* $u \in [0, 1]$, *every subsequential limit of the laws of the martingales in* (5.3) *is the law of a continuous nonnegative martingale bounded by* 1.
(3) *Suppose that, by choosing a subsequence if necessary, the sequence of laws of the martingales in* (5.3) *converges to the law of a continuous martingale* Z. *Then a stronger convergence holds*:

$$\left(Y_n, [Y_n], \langle Y_n \rangle \right) \xrightarrow[n \to \infty]{(d)} \left(Z, [Z], [Z] \right). \tag{5.4}$$

Proof. We begin with the proof of the assertion:

the sequence of laws of $\left(\langle Y_n \rangle, \mathbb{P}_{\mu_u}^{(n)} \right), \quad n \in \mathbb{N}, \text{ is } C\text{-tight.} \tag{5.5}$

This follows by verifying (5.6) and (5.9) below for the sequence of laws of $\langle Y_n \rangle$ (cf. Theorem VI.4.5 of [15]). First, we check the compact containment condition:

$$\forall \varepsilon > 0, T > 0, \exists K > 0 \text{ such that } \sup_{n \in \mathbb{N}} \mathbb{P}_{\mu_u}^{(n)} \left(\sup_{s \le T} \langle Y_n \rangle_s \ge K \right) \le \varepsilon. \tag{5.6}$$

We make use of the monotonicity of $\langle Y_n \rangle$, which gives, for every $\varepsilon > 0$, $T > 0$ and $n \in \mathbb{N}$,

$$\mathbb{P}_{\mu_u}^{(n)} \left(\sup_{s \le T} \langle Y_n \rangle_s \ge K \right) \le \frac{\mathbb{E}_{\mu_u}^{(n)}[\langle Y_n \rangle_T]}{K}$$

$$= \frac{\gamma_n \nu_n(\mathbb{1})}{K} \int_0^T \mathbb{E}_{\mu_u}^{(n)} \left[p_{10}(\xi_{\gamma_n s}) + p_{01}(\xi_{\gamma_n s}) \right] ds$$

$$\leq \frac{2\gamma_n \nu_n(\mathbb{1})}{K} \int_0^T \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\, \mathrm{d}s \tag{5.7}$$

$$\leq \frac{1}{K}, \tag{5.8}$$

where (5.7) follows from (3.15), and (5.8) from (3.20). We have proved (5.6).

The second condition which we need to check is Aldous's criterion:

$$\forall \varepsilon > 0, \quad \lim_{\theta \to 0} \limsup_{n \to \infty} \sup_{S,T:S \leq T \leq S+\theta} \mathbb{P}_{\mu_u}^{(n)}\big(\big|\langle Y_n \rangle_T - \langle Y_n \rangle_S\big| \geq \varepsilon\big) = 0, \tag{5.9}$$

where S and T range over all finite $(\mathscr{F}_t^n)$-stopping times. For any $\theta > 0$ and finite $(\mathscr{F}_t^n)$-stopping times S and T satisfying $S \leq T \leq S + \theta$, we have

$$\mathbb{E}_{\mu_u}^{(n)}\big[\big|\langle Y_n \rangle_T - \langle Y_n \rangle_S\big|\big] = \gamma_n \nu_n(\mathbb{1}) \mathbb{E}_{\mu_u}^{(n)}\bigg[\int_S^T \big[p_{10}(\xi_{\gamma_n s}) + p_{01}(\xi_{\gamma_n s})\big]\, \mathrm{d}s\bigg]$$

$$\leq \gamma_n \nu_n(\mathbb{1}) \mathbb{E}_{\mu_u}^{(n)}\bigg[\int_S^{S+\theta} \big[p_{10}(\xi_{\gamma_n s}) + p_{01}(\xi_{\gamma_n s})\big]\, \mathrm{d}s\bigg]$$

$$= \gamma_n \nu_n(\mathbb{1}) \mathbb{E}_{\mu_u}^{(n)}\bigg[\mathbb{E}_{\xi_{\gamma_n S}}^{(n)}\bigg[\int_0^\theta \big[p_{10}(\xi_{\gamma_n s}) + p_{01}(\xi_{\gamma_n s})\big]\, \mathrm{d}s\bigg]\bigg]$$

$$\leq 2\gamma_n \nu_n(\mathbb{1}) \int_0^\theta \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\, \mathrm{d}s, \tag{5.10}$$

where the last inequality follows from (3.13). Note that the right-hand side of the last inequality is independent of the stopping times S and T. By assumption, condition (2) of Theorem 4.1 holds, and thus

$$\lim_{n \to \infty} 2\gamma_n \nu_n(\mathbb{1}) \int_0^\theta \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n r)\, \mathrm{d}r = 1 - \mathrm{e}^{-\theta}.$$

Our claim (5.9) now follows by applying this equality to the right-hand side of (5.10). We have proved that the sequence of laws of the continuous processes $\langle Y_n \rangle$ is tight, or more precisely, C-tight.

The next step is to prove the desired properties (1)–(3) of the sequence of laws of the càdlàg martingales Y_n, given that we have obtained the C-tightness of the sequence of laws of $\langle Y_n \rangle$. Since the sequence of laws of the initial conditions $(Y_n(0))$ is clearly tight, we may apply Theorem VI.4.13 of [15] and conclude that the sequence $(Y_n, \mathbb{P}_{\mu_u}^{(n)})$ is tight. Since the jumps of Y_n are uniformly bounded by $\pi_{\max}^{(n)}$, and $\pi_{\max}^{(n)} \longrightarrow 0$ as $n \longrightarrow \infty$ on account of the assumption that $\Delta = 0$, it follows from Proposition VI.3.26 of [15] the sequence of laws of (Y_n) is C-tight, so we have proved (1) of our theorem.

We now consider (2). Suppose that $(Z, \mathbb{Q}_u)$ is a subsequential limit of the sequence of laws of $(Y_n, \mathbb{P}_{\mu_u}^{(n)})$. For convenience, we may assume that

$$\big(Y_n, \mathbb{P}_{\mu_u}^{(n)}\big) \xrightarrow[n \to \infty]{(\mathrm{d})} (Z, \mathbb{Q}_u).$$

Since $(Y_n, \mathbb{P}_{\mu_u}^{(n)})$ is C-tight and each member is a nonnegative martingale uniformly bounded by 1, it follows that the limiting object $(Z, \mathbb{Q}_u)$ is a continuous martingale bounded by 1 by Proposition IX.1.1 in [15], and (2) follows. Moreover, Corollary VI.6.30 of [15] implies that

$$\big(Y_n, [Y_n]\big) \xrightarrow[n \to \infty]{(\mathrm{d})} \big(Z, [Z]\big). \tag{5.11}$$

It remains to prove (3), and we need to reinforce the convergence in (5.11) to (5.4). To this end, it suffices to show that the sequence of laws of $\langle Y_n \rangle$ converge to the law of $[Z] = \langle Z \rangle$ as well. We have shown the C-tightness of

the sequence of laws of $\langle Y_n \rangle$ in the proof of (1). Hence, by taking a subsequence if necessary, we may assume that the sequence of laws of $(Y_n, \langle Y_n \rangle)$ converges to the law of (Z, B) for some continuous increasing process B. The sequence (Y_n) is obviously uniformly integrable. We will show in the last paragraph of this proof that $\{\langle Y_n \rangle_T\}_{n \in \mathbb{N}}$ is L_2-bounded for any $T > 0$, and hence uniformly integrable. It then follows that both Z and $Z^2 - B$ are continuous martingales with respect to the filtration generated by Z and B. The standard characterization of $\langle Z \rangle$ implies that $\langle Z \rangle = B$, and we can reinforce the convergence (5.11) to

$$\big(Y_n, [Y_n], \langle Y_n \rangle\big) \xrightarrow[n \to \infty]{(d)} \big(Z, [Z], [Z]\big). \tag{5.12}$$

To complete the proof, we verify that for any fixed $T > 0$,

$$\sup_{n \in \mathbb{N}} \mathbb{E}^{(n)}_{\mu_u}\big[\langle Y_n \rangle_T^2\big] < \infty. \tag{5.13}$$

With (5.2) as our starting point, we expand and use the Markov property at time $s < u$ to obtain

$$\begin{aligned}
\mathbb{E}^{(n)}_{\mu_u}\big[\langle Y_n \rangle_T^2\big] &= 2\big(\gamma_n \nu_n(\mathbb{1})\big)^2 \int_0^T \mathrm{d}s \int_s^T \mathrm{d}u \mathbb{E}^{(n)}_{\mu_u}\big[\big[p_{10}(\xi_{\gamma_n s}) + p_{01}(\xi_{\gamma_n s})\big] \\
&\qquad \times \mathbb{E}^{(n)}_{\xi_{\gamma_n s}}\big[p_{10}(\xi_{\gamma_n(u-s)}) + p_{01}(\xi_{\gamma_n(u-s)})\big]\big] \\
&\le 2\big(\gamma_n \nu_n(\mathbb{1})\big)^2 \int_0^T \mathbb{E}^{(n)}_{\mu_u}\big[p_{10}(\xi_{\gamma_n s}) + p_{01}(\xi_{\gamma_n s})\big]\mathrm{d}s \\
&\qquad \times \int_0^T 2\mathbf{P}^{(n)}(M_{V,V'} > \gamma_n u)\,\mathrm{d}u,
\end{aligned}$$

where the last inequality is due to (3.13). Applying (3.13) again, we obtain

$$\mathbb{E}^{(n)}_{\mu_u}\big[\langle Y_n \rangle_T^2\big] \le 2\bigg[2\gamma_n \nu_n(\mathbb{1}) \int_0^T \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n u)\,\mathrm{d}u\bigg]^2 \le 2$$

by (3.20). This gives (5.13), and the proof of (5.12) is complete. □

Corollary 5.2. *Suppose that* (2.6) *holds (i.e.,* (4.1) *holds with* $\Delta = 0$*). Then the convergence* (2.7) *implies the mean-field condition* (2.8).

Proof. Suppose that the sequence of laws of $(Y_n, \mathbb{P}^{(n)}_{\mu_u})$ converges to the distribution of the Wright–Fisher diffusion. This implies that condition (1) of Theorem 4.1 holds. As a consequence, Theorem 5.1 applies, and thus (5.4) must hold with the limit Z distributed as the Wright–Fisher diffusion Y and hence

$$[Z]_t = \int_0^t Z_s(1 - Z_s)\,\mathrm{d}s. \tag{5.14}$$

Since

$$w \longmapsto \bigg(\int_0^t w(s)\big[1 - w(s)\big]\mathrm{d}s;\, t \in \mathbb{R}_+\bigg) : D\big(\mathbb{R}_+, [0, 1]\big) \longrightarrow D(\mathbb{R}_+, \mathbb{R}) \tag{5.15}$$

defines a continuous function (cf. the proof of Proposition 3.7.1 in [13]) for any $T \in (0, \infty)$, the equation (5.14) and the convergence (5.4) imply

$$\langle Y_n \rangle_T - \int_0^T Y_n(s)\big[1 - Y_n(s)\big]\mathrm{d}s \xrightarrow[n \to \infty]{(d)} 0,$$

which is exactly the mean-field condition (2.8). □

Our strategy to complete the proof of Theorem 2.1 is to argue that if (2.6) and the mean-field condition (2.8) hold then the conditions of Theorem 4.1 must hold, so that Theorem 5.1 applies. To do this, we first show that the mean-field condition is itself a statement of local convergence in $L^p(\mathbb{P})$ for any $p \in [1, \infty)$.

Proposition 5.3. *For any voter model defined by an irreducible Q-matrix and initial configuration ξ,*

$$\mathbb{E}_\xi\left[\left(\int_0^\infty \nu(\mathbb{1})\big[p_{10}(\xi_s) + p_{01}(\xi_s)\big]\,\mathrm{d}s\right)^m\right] \le m! \quad \forall m \in \mathbb{N}. \tag{5.16}$$

Hence, the mean-field condition (2.8) *holds if and only if for all $T \in (0, \infty)$,*

$$\lim_{n\to\infty} \mathbb{E}^{(n)}_{\mu_u}\left[\left|\int_0^T \big(\gamma_n \nu_n(\mathbb{1})\big[p_{10}(\xi_{\gamma_n s}) + p_{01}(\xi_{\gamma_n s})\big] - p_1(\xi_{\gamma_n s})p_0(\xi_{\gamma_n s})\big)\,\mathrm{d}s\right|^p\right] = 0$$
$$\forall p \in [1, \infty). \tag{5.17}$$

If we set $m = 2$ and use duality, then the convergence condition (5.17) is equivalent to a condition that can be expressed in terms of two pairs of coalescing Markov chains started at different times. We show in the next section that an argument using only a single pair of Markov chains is sufficient to obtain this convergence.

Proof of Proposition 5.3. By (3.13) and (3.20), for any initial configuration ξ,

$$\nu(\mathbb{1})\mathbb{E}_\xi\left[\int_0^\infty \big[p_{10}(\xi_s) + p_{01}(\xi_s)\big]\,\mathrm{d}s\right] \le 2\nu(\mathbb{1})\int_0^\infty \mathbf{P}(M_{V,V'} > s)\,\mathrm{d}s \le 1. \tag{5.18}$$

For $m \in \mathbb{N}$, if we expand the left-hand side of (5.16), and then use the Markov property at time $s_{m-1} < s_m$, we obtain

$$m!\big(\nu(\mathbb{1})\big)^m \int_0^\infty \mathrm{d}s_1 \int_{s_1}^\infty \mathrm{d}s_2 \cdots \int_{s_{m-1}}^\infty \mathrm{d}s_m \mathbb{E}_\xi\left[\prod_{i=1}^m \big[p_{10}(\xi_{s_i}) + p_{01}(\xi_{s_i})\big]\right]$$

$$= m!\big(\nu(\mathbb{1})\big)^m \int_0^\infty \mathrm{d}s_1 \int_{s_1}^\infty \mathrm{d}s_2 \cdots \int_{s_{m-2}}^\infty \mathrm{d}s_{m-1} \mathbb{E}_\xi\left[\prod_{i=1}^{m-1} \big[p_{10}(\xi_{s_i}) + p_{01}(\xi_{s_i})\big]\right.$$

$$\left. \times \mathbb{E}_{\xi_{s_{m-1}}} \int_{s_{m-1}}^\infty \big[p_{10}(\xi_{s_m - s_{m-1}}) + p_{01}(\xi_{s_m - s_{m-1}})\big]\,\mathrm{d}s_m\right].$$

By applying the bound (5.18) and iteration, we obtain (5.16).

For the second assertion, we only need to show that the mean-field condition implies (5.17), because the converse follows immediately from Markov's inequality. Moreover, given the mean-field condition, by Skorokhod's representation and a standard result of uniform integrability, it is enough to derive a uniform bound on the mth moment of

$$\gamma_n \nu_n(\mathbb{1}) \int_0^\infty \big[p_{10}(\xi_{\gamma_n s}) + p_{01}(\xi_{\gamma_n s})\big]\,\mathrm{d}s$$

for any $m \in \mathbb{N}$, which is precisely the content of the first assertion. Hence, (5.17) holds, and the proof is complete.
□

The following result connects the mean-field condition and the various equivalent conditions in Theorem 4.1. It completes the proof of Theorem 2.1.

Theorem 5.4. *Suppose that* (2.6) *holds (i.e.,* (4.1) *holds with $\Delta = 0$). Then the mean-field condition* (2.8) *implies all of the conditions of Theorem* 4.1 *hold, as well as the convergence* (2.7) *for all $u \in [0, 1]$.*

Proof. Let (γ_n) be a sequence of strictly positive constants so that the mean-field condition (2.8) holds. For the first assertion, it is enough to show that (1) of Theorem 4.1 holds. By taking a subsequence if necessary, we may assume that

$$I(t) = \lim_{n\to\infty} \int_0^t \mathbb{E}_{\mu_u}^{(n)}\big[p_1(\xi_{\gamma_n s}) p_0(\xi_{\gamma_n s})\big] \, ds \tag{5.19}$$

exists in $\mathbb{R}_+$ for all $t \in \mathbb{Q}_+$. Since $s \longmapsto \mathbb{E}_{\mu_u}^{(n)}[p_1(\xi_{\gamma_n s}) p_0(\xi_{\gamma_n s})]$ is uniformly bounded, a monotonicity argument implies that the foregoing limit exists for all $t \in \mathbb{R}_+$ and defines a continuous function I on $\mathbb{R}_+$. Moreover, given that the mean-field condition holds, we can write the function I as

$$I(t) = \lim_{n\to\infty} \gamma_n \nu_n(\mathbb{1}) \int_0^t \mathbb{E}_{\mu_u}^{(n)}\big[p_{10}(\xi_{\gamma_n s}) + p_{01}(\xi_{\gamma_n s})\big] \, ds \tag{5.20}$$

by Proposition 5.3. In view of (3.5) and the last display, we obtain

$$\lim_{n\to\infty} \mathbb{E}_{\mu_u}^{(n)}\big[p_1(\xi_{\gamma_n t}) p_0(\xi_{\gamma_n t})\big] = u(1-u) - I(t) \quad \forall t \in \mathbb{R}_+. \tag{5.21}$$

By the bounded convergence theorem and the definition of $I(t)$, this implies

$$I(t) = \int_0^t \big[u(1-u) - I(s)\big] \, ds$$

by (5.19). Solving this integral equation gives $I(t) = u(1-u)(1 - e^{-t})$. By plugging this solution into the right-hand side of (5.21), we find that

$$\lim_{n\to\infty} \mathbb{E}_{\mu_u}^{(n)}\big[p_1(\xi_{\gamma_n t}) p_0(\xi_{\gamma_n t})\big] = u(1-u)e^{-t},$$

which is (1) of Theorem 4.1. This proves the first assertion.

Having proved that the conditions of Theorem 4.1 hold, we may now apply Theorem 5.1. By (i) of Theorem 5.1, the family $(Y_n, \mathbb{P}_{\mu_u}^{(n)})$ is C-tight for any $u \in [0, 1]$. If $(Y_{n_k}, \mathbb{P}_{\mu_u}^{(n_k)})$ is any weakly convergent subsequence, then (ii) and (iii) of Theorem 5.1 imply that

$$\big(Y_{n_k}, \langle Y_{n_k}\rangle\big) \xrightarrow[n\to\infty]{\text{(d)}} \big(Z, \langle Z\rangle\big)$$

for a continuous martingale Z. Thanks to the continuity of the map (5.15), we deduce from the mean-field condition (2.8) that

$$\langle Z\rangle_T = \int_0^T Z_s(1 - Z_s) \, ds \quad \forall T \in \mathbb{R}_+$$

almost surely. Hence, Z is a Wright–Fisher diffusion, and the proof is complete. $\square$

Although we only consider Bernoulli initial conditions throughout this section, the readers may notice that most of the proofs do apply to the context where for each n, the initial condition of the voter model defined by $(q^{(n)}, E_n)$ is a general probability measure λ_n on $\{0, 1\}^{E_n}$.

More precisely, the same proofs of Theorem 5.1, Corollary 5.2, and Proposition 5.3 still apply, if we consider such a generalization. For the extension of Theorem 5.4, we consider general initial conditions λ_n for which the sequence of laws $\lambda_n(p_1(\xi) \in \cdot)$ converges weakly to a probability measure, say, $\widehat{\lambda}_\infty$ on $[0, 1]$, and use (3.4) instead of (3.5) to obtain an analogue of (5.21). This leads to the conclusion that, whenever the mean-field condition (2.8), with μ_u replaced by λ_n for each voter model defined $q^{(n)}$, holds, we have the weak convergence of the associated density processes to the Wright–Fisher diffusion with initial condition $\widehat{\lambda}_\infty$.

6. Proof of Theorem 2.2

For the convenience of readers, we give an informal outline of the proof of Theorem 2.2 first. We take a generic voter model as usual and a constant $\gamma > 0$. Falling back in time by a small amount δ and using the Markov property of voter models, we get for any instant s

$$\gamma v(\mathbb{1})\big[p_{10}(\xi_{\gamma s}) + p_{01}(\xi_{\gamma s})\big] - p_1(\xi_{\gamma s})p_0(\xi_{\gamma s})$$
$$\simeq \gamma v(\mathbb{1})\big(\mathbb{E}_{\xi_{\gamma(s-\delta)}}\big[p_{10}(\xi_{\gamma\delta})\big] + \mathbb{E}_{\xi_{\gamma(s-\delta)}}\big[p_{01}(\xi_{\gamma\delta})\big]\big) - p_1(\xi_{\gamma(s-\delta)})p_0(\xi_{\gamma(s-\delta)})$$

on $\sigma(\xi_{\gamma u}; u \leq s)$. We then resort to duality and interpret the right-hand side, or more generally the term

$$\gamma v(\mathbb{1})\big(\mathbb{E}_\xi\big[p_{10}(\xi_{\gamma\delta})\big] + \mathbb{E}_\xi\big[p_{01}(\xi_{\gamma\delta})\big]\big) - p_1(\xi)p_0(\xi) \tag{6.1}$$

for arbitrary ξ, by moving forward in time from the point of view of q-Markov chains. For the first two terms $\mathbb{E}_\xi[p_{10}(\xi_{\gamma\delta})]$ and $\mathbb{E}_\xi[p_{01}(\xi_{\gamma\delta})]$, we use Proposition 3.3 and read them as expectations of the function

$$(x, y) \longmapsto \xi(x)\widehat{\xi}(y)$$

of some pairs of q-Markov chains before they meet. On the other hand, $p_1(\xi)$ and $p_0(\xi)$ are the π-expectations of configurations ξ and $\widehat{\xi}$, respectively, where π is the stationary distribution of the q-Markov chain. Applying these observations to the quantity (6.1), we can regard (2.8) as a result that, informally speaking, the time that two q-Markov chains meet "falls far behind" the time that a q-Markov chain gets close to its equilibrium distribution π. See also [9] for an application of this "falling-back-moving-forward" argument.

Some additional notation will be useful in the first step of making the above precise. Recall the system (X_t^x) of independent q-Markov chains on E with semigroup (q_t) and stationary distribution π. For any real function f on E define $\pi(f) = \sum_{x \in E} f(x)\pi(x)$, $q_t f(x) = \sum_{y \in E} q_t(x, y)f(y)$, and

$$\mathrm{Var}_\pi(f) = \sum_{x \in E}\big(f(x) - \pi(f)\big)^2\pi(x).$$

The following bounds will be useful. First, we have two bounds on the difference between $q_t\xi(x)$ and $p_1(\xi)$. Recall the definition of d_E in (2.9). Since $p_1(\xi) = \pi(\xi)$ and ξ is bounded by 1, it follows that

$$\big|q_t\xi(x) - p_1(\xi)\big| \leq 2d_E(t), \quad \forall x \in E, \xi \in \{0, 1\}^E \tag{6.2}$$

(see Proposition 4.5 in [17]). A second bound (see, e.g., Lemma 2.4 of [10]) is available when (q_t) is reversible and has spectral gap **g**. In this case, for any f,

$$\mathrm{Var}_\pi(q_t f) \leq \mathrm{Var}_\pi(f)e^{-2\mathbf{g}t}. \tag{6.3}$$

Second, it follows from the definition of $\mathbf{t}_{\mathrm{mix}}$ in (2.10) that

$$d_E(k\mathbf{t}_{\mathrm{mix}}) \leq e^{-k}, \quad \forall k \in \mathbb{N} \tag{6.4}$$

(see Section 4.5 of [17]).

Proposition 6.1. *Let (q, E) be an irreducible Q-matrix. For any $0 < s < t < \infty$, we have the following estimates.*

(1) *If d_E denotes the maximal total variation distance defined by* (2.9), *then*

$$\sup_{\xi \in \{0,1\}^E} \big|\mathbb{E}_\xi\big[p_{10}(\xi_t)\big] - \mathbf{P}(M_{V,V'} > s)p_1(\xi)p_0(\xi)\big|$$
$$\leq \mathbf{P}\big(M_{V,V'} \in (s, t]\big) + 4\mathbf{P}(M_{V,V'} > s)d_E(t - s). \tag{6.5}$$

The same inequality holds if p_{10} is replaced by p_{01}.

(2) *If the q-Markov chain is reversible and* **g** *is the associated spectral gap*, *then*

$$\sup_{\xi \in \{0,1\}^E} \left| \mathbb{E}_\xi \left[p_{10}(\xi_t) \right] - \mathbf{P}(M_{V,V'} > s) p_1(\xi) p_0(\xi) \right|$$

$$\leq \mathbf{P}\big(M_{V,V'} \in (s,t]\big) + 2\sqrt{\frac{\pi_{\max} q_{\max}}{\nu(\mathbb{1})}} e^{-\mathbf{g}(t-s)}, \tag{6.6}$$

where $\pi_{\max} = \max\{\pi(x); x \in E\}$ *and* $q_{\max} = \max\{q(x); x \in E\}$. *The same inequality holds if* p_{10} *is replaced by* p_{01}.

Proof. The proofs of (1) and (2) are based on the preliminary bound

$$\left| \mathbb{E}_\xi \left[p_{10}(\xi_t) \right] - p_1(\xi) p_0(\xi) \mathbf{P}(M_{V,V'} > s) \right|$$

$$\leq \mathbf{P}\big(M_{V,V'} \in (s,t]\big) + \left| \mathbf{E}\big[q_{t-s}\xi\big(X_s^V\big) q_{t-s}\widehat{\xi}\big(X_s^{V'}\big) - p_1(\xi) p_0(\xi); M_{V,V'} > s \big] \right|. \tag{6.7}$$

To get this bound, we first use Proposition 3.3 and write for any configuration ξ,

$$\mathbb{E}_\xi \left[p_{10}(\xi_t) \right] = \mathbf{E}\big[\xi\big(X_t^V\big) \widehat{\xi}\big(X_t^{V'}\big); M_{V,V'} > t \big]$$

$$= \mathbf{E}\big[\xi\big(X_t^V\big) \widehat{\xi}\big(X_t^{V'}\big); M_{V,V'} > s \big] + \varepsilon_1(s,t;\xi), \tag{6.8}$$

where

$$\left| \varepsilon_1(s,t;\xi) \right| \leq \mathbf{P}\big(M_{V,V'} \in (s,t]\big) \tag{6.9}$$

uniformly in ξ. Applying the Markov property of the two-dimensional process $(X^V, X^{V'})$ at time s, we have

$$\mathbf{E}\big[\xi\big(X_t^V\big) \widehat{\xi}\big(X_t^{V'}\big); M_{V,V'} > s \big] = \mathbf{E}\big[\mathbf{E}\big[\xi\big(X_{t-s}^v\big) \widehat{\xi}\big(X_{t-s}^{v'}\big) \big] \big|_{(v,v')=(X_s^V, X_s^{V'})}; M_{V,V'} > s \big]$$

$$= \mathbf{E}\big[q_{t-s}\xi\big(X_s^V\big) q_{t-s}\widehat{\xi}\big(X_s^{V'}\big); M_{V,V'} > s \big]$$

$$= p_1(\xi) p_0(\xi) \mathbf{P}(M_{V,V'} > s)$$

$$\quad + \mathbf{E}\big[q_{t-s}\xi\big(X_s^V\big) q_{t-s}\widehat{\xi}\big(X_s^{V'}\big) - p_1(\xi) p_0(\xi); M_{V,V'} > s \big].$$

Combining this equality and the bound (6.9) on ε_1 with (6.8) gives (6.7).

We now consider the proof of (1). The last term in (6.7) is bounded above by

$$\mathbf{E}\big[\big| q_{t-s}\xi\big(X_s^V\big) q_{t-s}\widehat{\xi}\big(X_s^{V'}\big) - p_1(\xi) p_0(\xi) \big|; M_{V,V'} > s \big]$$

$$\leq \mathbf{E}\big[p_0(\xi) \big| p_1(\xi) - q_{t-s}\xi\big(X_s^V\big) \big|; M_{V,V'} > s \big]$$

$$\quad + \mathbf{E}\big[q_{t-s}\xi\big(X_s^V\big) \big| p_0(\xi) - q_{t-s}\widehat{\xi}\big(X_s^{V'}\big) \big|; M_{V,V'} > s \big]$$

$$\leq 4 d_E(t-s) \mathbf{P}(M_{V,V'} > s),$$

where we have used (6.2). Plugging this bound into (6.7) gives (6.5).

Next, we turn to the proof of (2). In this case, we bound the last term in (6.7) in the following way:

$$\mathbf{E}\big[\big| q_{t-s}\xi\big(X_s^V\big) q_{t-s}\widehat{\xi}\big(X_s^{V'}\big) - \pi(\xi)\pi(\widehat{\xi}) \big| \big]$$

$$\leq \mathbf{E}\big[\big| q_{t-s}\xi\big(X_s^V\big) - \pi(\xi) \big| \big] + \mathbf{E}\big[\big| q_{t-s}\widehat{\xi}\big(X_s^{V'}\big) - \pi(\widehat{\xi}) \big| \big]$$

$$\leq \mathbf{E}\big[\big(q_{t-s}\xi\big(X_s^V\big) - \pi(\xi) \big)^2 \big]^{1/2} + \mathbf{E}\big[\big(q_{t-s}\widehat{\xi}\big(X_s^{V'}\big) - \pi(\widehat{\xi}) \big)^2 \big]^{1/2}. \tag{6.10}$$

Recall the distribution of (V, V') in (3.9). For all $x \in E$ and $s \geq 0$, we have

$$\mathbf{P}(X_s^V = x) = \bar{\nu}(\{x\} \times E) = \frac{\pi(x)^2 q(x)}{\nu(\mathbb{1})} \leq \frac{\pi_{\max} q_{\max}}{\nu(\mathbb{1})} \pi(x).$$

Since (X^x) is independent of V, it follows from the foregoing inequality that

$$\mathbf{E}\big[\big(q_{t-s}\xi\big(X_s^V\big) - \pi(\xi)\big)^2\big] \leq \frac{\pi_{\max} q_{\max}}{\nu(\mathbb{1})} \operatorname{Var}_\pi(q_{t-s}\xi) \leq \frac{\pi_{\max} q_{\max}}{\nu(\mathbb{1})} e^{-2\mathbf{g}(t-s)}, \tag{6.11}$$

where we have used (6.3) and the fact that $\operatorname{Var}(f) \leq 1$ if $|f|$ is bounded by 1. The same bound holds if we replace $\xi(X_s^V)$ with $\widehat{\xi}(X_s^{V'})$. Indeed, we still have

$$\mathbf{P}(X_s^{V'} = x) \leq \frac{\pi_{\max} q_{\max}}{\nu(\mathbb{1})} \pi(x) \quad \forall x \in E \text{ and } s \geq 0,$$

since for any $x \in E$, reversibility implies

$$\mathbf{P}(V' = x) = \frac{\sum_{a:a\neq x} \pi(a)^2 q(a, x)}{\nu(\mathbb{1})}$$

$$= \frac{\pi(x) \sum_{a:a\neq x} \pi(a)q(x, a)}{\nu(\mathbb{1})} \leq \frac{\pi_{\max} q_{\max}}{\nu(\mathbb{1})} \pi(x).$$

Hence by (6.11) and its analogue when V is replaced by V', we obtain from (6.10) that

$$\mathbf{E}\big[\big|q_{t-s}\xi\big(X_s^V\big) q_{t-s}\widehat{\xi}\big(X_s^{V'}\big) - \pi(\xi)\pi(\widehat{\xi})\big|\big] \leq 2\sqrt{\frac{\pi_{\max} q_{\max}}{\nu(\mathbb{1})}} e^{-\mathbf{g}(t-s)}.$$

Plugging this bound into (6.7) completes the proof of (6.6). □

Lemma 6.2. *If* $\gamma_n = \mathbf{t}_{\mathrm{meet}}^{(n)}$, *then under either condition of Theorem* 2.2, *any of the conditions in Theorem* 4.1 *holds with* $\Delta = 0$. *Moreover, we can choose* (s_n') *satisfying* (4.14) *with* $s_n' = \mathrm{o}(\gamma_n)$ *such that with* $\delta_n = s_n'/\gamma_n$,

$$\varepsilon_n = \sup_{\xi \in \{0,1\}^{E_n}} \big|\gamma_n \nu_n(\mathbb{1})\mathbb{E}_\xi^{(n)}\big[p_{10}(\xi_{\gamma_n \cdot 2\delta_n}) + p_{01}(\xi_{\gamma_n \cdot 2\delta_n})\big] - p_1(\xi)p_0(\xi)\big| \underset{n\to\infty}{\longrightarrow} 0 \tag{6.12}$$

(*cf. the informal discussion at the beginning of this section*).

Proof. Let $\gamma_n = \mathbf{t}_{\mathrm{meet}}^{(n)} = \mathbf{E}^{(n)}[M_{U,U'}]$. The strategy is to first prove that (4) of Theorem 4.1 holds, i.e.,

$$\frac{M_{U,U'}}{\gamma_n} \xrightarrow[n\to\infty]{(\mathrm{d})} \mathbf{e}, \tag{6.13}$$

and then use Proposition 4.3 and the bounds in Proposition 6.1 to choose a sequence (s_n') satisfying (6.12).

Suppose first that (i) of Theorem 2.2 holds, and consider the product chain comprised of two independent copies of $q^{(n)}$-Markov chains. For the product chain started at its stationary distribution $\pi \otimes \pi$, the first hitting time of the diagonal D_n has the same law as the meeting time $M_{U,U'}$. Letting $(\tilde{q}_t^{(n)})$ denote the product chain semigroup, we have the obvious inequality

$$\big\|\tilde{q}_t^{(n)}(\cdot) - \pi^{(n)} \otimes \pi^{(n)}(\cdot)\big\|_{\mathrm{TV}} \leq 2d_{E_n}(t).$$

By this inequality and our assumption that $\mathbf{t}_{\mathrm{mix}}^{(n)}/\mathbf{t}_{\mathrm{meet}}^{(n)} \longrightarrow 0$, Theorem 1.4 of [1] on the convergence of Markov chain hitting times to exponential laws applies to the product chain $\tilde{q}_t$, giving (6.13).

Now let (s_n) be a sequence with $s_n = o(\gamma_n)$ satisfying (4.14) (here $\Delta = 0$ by assumption). Note that the existence of such a sequence (s_n) is guaranteed by (4.12). Define (s_n') by

$$s_n' = s_n \vee u_n \mathbf{t}_{\text{mix}}^{(n)},$$

where (u_n) satisfies

$$\lim_{n\to\infty} u_n = \infty \quad \text{and} \quad \lim_{n\to\infty} \frac{u_n \mathbf{t}_{\text{mix}}^{(n)}}{\gamma_n} = 0.$$

Observe that $\delta_n = s_n'/\gamma_n \to 0$ as $n \to \infty$, and also that $s_n'/\mathbf{t}_{\text{mix}}^{(n)} \to \infty$ implies $d_{E_n}(s_n') \to 0$ by (6.4). Furthermore, applying (2) of Proposition 4.3 to both (s_n') and $(2s_n')$, we have

$$\lim_{n\to\infty} 2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}\big(M_{V,V'} > s_n'\big) = 1 \quad \text{and} \quad \lim_{n\to\infty} 2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}\big(M_{V,V'} \in \big(s_n', 2s_n'\big]\big) = 0. \tag{6.14}$$

By (1) of Proposition 6.1, taking $s = s_n'$ and $t = 2s_n'$, we have for any initial configuration ξ,

$$\begin{aligned}
&\Big|\gamma_n \nu_n(\mathbb{1})\mathbb{E}_\xi^{(n)}\big[p_{10}(\xi_{\gamma_n \cdot 2\delta_n}) + p_{01}(\xi_{\gamma_n \cdot 2\delta_n})\big] - p_1(\xi)p_0(\xi)\Big| \\
&\quad \leq 2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}\big(M_{V,V'} \in \big(s_n', 2s_n'\big]\big) + 8\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}\big(M_{V,V'} > s_n'\big)d_{E_n}\big(s_n'\big) \\
&\qquad + \Big|2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}\big(M_{V,V'} > s_n'\big) - 1\Big|p_1(\xi)p_0(\xi).
\end{aligned} \tag{6.15}$$

Therefore, (6.12) follows from (6.14) and (6.15).

Next, suppose that (ii) of Theorem 2.2 holds, so $\mathbf{g}_n \mathbf{t}_{\text{meet}}^{(n)} \to \infty$ as $n \to \infty$. We consider again the product chain, the hitting time of the diagonal and the meeting time $M_{U,U'}$. The product chain is reversible, and has spectral gap $\tilde{\mathbf{g}}_n = \mathbf{g}_n/2$ by Lemma 3.2 in [10]. It follows from Proposition 3.23 in [2] that the hitting time for the diagonal D_n is approximately exponentially distributed in the sense that (6.13) holds.

We again select a sequence (s_n) such that $s_n = o(\gamma_n)$ and (4.14) holds. The existence of (s_n) is due to the same reason as in the case (i). Now we choose (u_n) such that

$$\lim_{n\to\infty} u_n = \infty \quad \text{and} \quad \lim_{n\to\infty} u_n \frac{\log(e \vee \gamma_n \sqrt{\pi_{\max}^{(n)} q_{\max}^{(n)} \nu_n(\mathbb{1})})}{\mathbf{g}_n \gamma_n} = 0, \tag{6.16}$$

and define (s_n') by

$$s_n' = s_n \vee \frac{\gamma_n}{u_n}.$$

Clearly $\delta_n = s_n'/\gamma_n \to 0$, and (6.14) holds by (2) of Proposition 4.3. By (2) of Proposition 6.1 with $s = s_n'$ and $t = 2s_n'$, we get for any initial configuration ξ,

$$\begin{aligned}
&\Big|\gamma_n \nu_n(\mathbb{1})\mathbb{E}_\xi^{(n)}\big[p_{10}(\xi_{\gamma_n \cdot 2\delta_n}) + p_{01}(\xi_{\gamma_n \cdot 2\delta_n})\big] - p_1(\xi)p_0(\xi)\Big| \\
&\quad \leq 2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}\big(M_{V,V'} \in \big(s_n', 2s_n'\big]\big) + 4\sqrt{\pi_{\max}^{(n)} q_{\max}^{(n)} \nu_n(\mathbb{1})}\gamma_n e^{-\mathbf{g}_n s_n'} \\
&\qquad + \Big|2\gamma_n \nu_n(\mathbb{1})\mathbf{P}^{(n)}\big(M_{V,V'} > s_n'\big) - 1\Big|p_1(\xi)p_0(\xi).
\end{aligned} \tag{6.17}$$

As before, by our choice of (s_n') and Proposition 4.3, the first term and the third one on the right-hand side above tend to 0 as $n \to \infty$.

To show that the second term on the right-hand side of (6.17) also tends to zero, we make some observations for the condition (ii) of Theorem 2.2. Now, $\pi_{\text{diag}}^{(n)} \to 0$, and so the inequality (2.13) implies that

$$\liminf_{n\to\infty} \gamma_n \sqrt{\pi_{\max}^{(n)} q_{\max}^{(n)} \nu_n(\mathbb{1})} = \liminf_{n\to\infty} \mathbf{t}_{\text{meet}}^{(n)} \sqrt{\pi_{\max}^{(n)} q_{\max}^{(n)} \nu_n(\mathbb{1})} > 0.$$

On the other hand,

$$\frac{\mathbf{g}_n \gamma_n}{u_n} - \log\big(e \vee \gamma_n \sqrt{\pi_{\max}^{(n)} q_{\max}^{(n)} \nu_n(\mathbb{1})}\big) = \log\big(e \vee \gamma_n \sqrt{\pi_{\max}^{(n)} q_{\max}^{(n)} \nu_n(\mathbb{1})}\big)\left(\frac{\mathbf{g}_n \gamma_n}{u_n \log(e \vee \gamma_n \sqrt{\pi_{\max}^{(n)} q_{\max}^{(n)} \nu_n(\mathbb{1})})} - 1\right)$$

$$\geq \frac{\mathbf{g}_n \gamma_n}{u_n \log(e \vee \gamma_n \sqrt{\pi_{\max}^{(n)} q_{\max}^{(n)} \nu_n(\mathbb{1})})} - 1 \longrightarrow \infty,$$

where the convergence follows from the choice of u_n in (6.16). We deduce from the last two displays that

$$\mathbf{g}_n s_n' - \log\big(e \vee \gamma_n \sqrt{\pi_{\max}^{(n)} q_{\max}^{(n)} \nu_n(\mathbb{1})}\big) \geq \frac{\mathbf{g}_n \gamma_n}{u_n} - \log\big(e \vee \gamma_n \sqrt{\pi_{\max}^{(n)} q_{\max}^{(n)} \nu_n(\mathbb{1})}\big) \longrightarrow \infty,$$

which is enough for the desired convergence. The proof is complete. $\qquad\square$

We are now ready to prove Theorem 2.2.

Proof of Theorem 2.2. We have shown in the proof of Lemma 6.2 that (6.13) holds, and hence all of the equivalent conditions of Theorem 4.1 hold. Also, the sequences (δ_n) and (ε_n) defined in Lemma 6.2 satisfy $\delta_n \longrightarrow 0$ and $\varepsilon_n \longrightarrow 0$ as $n \longrightarrow \infty$.

Our goal in this proof is to prove the L^1-norm version of the mean-field condition, namely (5.17) with $p = 1$ for any $T > 0$. For this, we first note that (3.13) gives

$$\mathbb{E}_{\mu_u}^{(n)}\left[\int_0^{2\delta_n} \big|\gamma_n \nu_n(\mathbb{1})\big[p_{10}(\xi_{\gamma_n s}) + p_{01}(\xi_{\gamma_n s})\big] - p_1(\xi_{\gamma_n s}) p_0(\xi_{\gamma_n s})\big| \, \mathrm{d}s\right]$$

$$\leq 2\gamma_n \nu_n(\mathbb{1}) \int_0^{2\delta_n} \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s) \, \mathrm{d}s + 2\delta_n,$$

and the right-hand side tends to 0 as $n \longrightarrow \infty$ by (2) of Theorem 4.1 and the fact that $\delta_n \longrightarrow 0$. Hence, it remains to show that

$$\lim_{n \to \infty} \mathbb{E}_{\mu_u}^{(n)}\left[\left|\int_{2\delta_n}^T \big(\gamma_n \nu_n(\mathbb{1})\big[p_{10}(\xi_{\gamma_n s}) + p_{01}(\xi_{\gamma_n s})\big] - p_1(\xi_{\gamma_n s}) p_0(\xi_{\gamma_n s})\big) \, \mathrm{d}s\right|\right] = 0, \tag{6.18}$$

for any $T > 0$.

For convenience, we write from now on

$$\bar{p}(\xi) \equiv p_{10}(\xi) + p_{01}(\xi),$$

and for any $s \geq 2\delta_n$,

$$H_n(s) \equiv \gamma_n \nu_n(\mathbb{1}) \bar{p}(\xi_{\gamma_n s}) - \mathbb{E}_{\mu_u}^{(n)}\big[\gamma_n \nu_n(\mathbb{1}) \bar{p}(\xi_{\gamma_n s}) \big| \mathscr{F}_{s-2\delta_n}^n\big]$$

(recall the definition of $\mathscr{F}_t^n$ from (5.1)). Note that $H_n(s) \in \mathscr{F}_{s-2\delta_n}^n$. Then

$$\mathbb{E}_{\mu_u}^{(n)}\left[\left|\int_{2\delta_n}^T \gamma_n \nu_n(\mathbb{1}) \bar{p}(\xi_{\gamma_n s}) - p_1(\xi_{\gamma_n s}) p_0(\xi_{\gamma_n s}) \, \mathrm{d}s\right|\right]$$

$$\leq \mathbb{E}_{\mu_u}^{(n)}\left[\left(\int_{2\delta_n}^T H_n(s) \, \mathrm{d}s\right)^2\right]^{1/2}$$

$$+ \mathbb{E}_{\mu_u}^{(n)}\left[\int_{2\delta_n}^T \big|\mathbb{E}_{\mu_u}^{(n)}\big[\gamma_n \nu_n(\mathbb{1}) \bar{p}(\xi_{\gamma_n s}) \big| \mathscr{F}_{s-2\delta_n}^n\big] - p_1(\xi_{\gamma_n(s-2\delta_n)}) p_0(\xi_{\gamma_n(s-2\delta_n)})\big| \, \mathrm{d}s\right]$$

$$+ \mathbb{E}_{\mu_u}^{(n)}\left[\left|\int_{2\delta_n}^T p_1(\xi_{\gamma_n(s-2\delta_n)}) p_0(\xi_{\gamma_n(s-2\delta_n)}) - p_1(\xi_{\gamma_n s}) p_0(\xi_{\gamma_n s}) \, \mathrm{d}s\right|\right], \tag{6.19}$$

and so to verify (6.18) it suffices to prove that each term on the right-hand side of the above tends to 0 as $n \longrightarrow \infty$.

We first prove that the first term on the right-hand side of (6.19) tends to zero. Note that

$$\mathbb{E}_{\mu_u}^{(n)}\left[\left(\int_{2\delta_n}^{T} H_n(s)\,\mathrm{d}s\right)^2\right] = 2\mathbb{E}_{\mu_u}^{(n)}\left[\int\int_{2\delta_n \le r \le s \le T} H_n(s)H_n(r)\mathbb{1}_{r>s-2\delta_n}\,\mathrm{d}s\,\mathrm{d}r\right]. \tag{6.20}$$

To justify the restriction "$\mathbb{1}_{r>s-2\delta_n}$" for the right-hand side, we note that for $2\delta_n \le r < s - 2\delta_n$,

$$\mathbb{E}_{\mu_u}^{(n)}\big[H_n(s)|\mathscr{F}_r^n\big] = 0,$$

and hence, we obtain by conditioning on $\mathscr{F}_r^n$ that

$$\mathbb{E}_{\mu_u}^{(n)}\big[H_n(s)H_n(r)\big] = 0, \qquad 2\delta_n \le r < s - 2\delta_n.$$

Now expanding $H_n(r)H_n(s)$, we obtain

$$\mathbb{E}_{\mu_u}^{(n)}\left[\int\int_{2\delta_n \le r \le s \le T} \mathrm{d}r\,\mathrm{d}s\,H_n(s)H_n(r)\mathbb{1}_{r>s-2\delta_n}\right]$$

$$= \mathbb{E}_{\mu_u}^{(n)}\left[\int_{2\delta_n}^{T}\mathrm{d}r\int_r^{T\wedge(r+2\delta_n)}\mathrm{d}s\big(\gamma_n v_n(\mathbb{1})\big)^2\bar{p}(\xi_{\gamma_n r})\bar{p}(\xi_{\gamma_n s})\right]$$

$$-\int_{2\delta_n}^{T}\mathrm{d}r\int_r^{T\wedge(r+2\delta_n)}\mathrm{d}s\,\mathbb{E}_{\mu_u}^{(n)}\big[\big(\gamma_n v_n(\mathbb{1})\big)^2\mathbb{E}_{\mu_u}^{(n)}\big[\bar{p}(\xi_{\gamma_n r})|\mathscr{F}_{r-2\delta_n}^n\big]\bar{p}(\xi_{\gamma_n s})\big]$$

$$-\int_{2\delta_n}^{T}\mathrm{d}r\int_r^{T\wedge(r+2\delta_n)}\mathrm{d}s\,\mathbb{E}_{\mu_u}^{(n)}\big[\big(\gamma_n v_n(\mathbb{1})\big)^2\bar{p}(\xi_{\gamma_n r})\mathbb{E}_{\mu_u}^{(n)}\big[\bar{p}(\xi_{\gamma_n s})|\mathscr{F}_{s-2\delta_n}^n\big]\big]$$

$$+\int_{2\delta_n}^{T}\mathrm{d}r\int_r^{T\wedge(r+2\delta_n)}\mathrm{d}s\,\mathbb{E}_{\mu_u}^{(n)}\big[\big(\gamma_n v_n(\mathbb{1})\big)^2\mathbb{E}_{\mu_u}^{(n)}\big[\bar{p}(\xi_{\gamma_n r})|\mathscr{F}_{r-2\delta_n}^n\big]$$

$$\times \mathbb{E}_{\mu_u}^{(n)}\big[\bar{p}(\xi_{\gamma_n s})|\mathscr{F}_{s-2\delta_n}^n\big]\big]. \tag{6.21}$$

We will show that each of the four terms on the right-hand side of the last equality tends to zero as $n \to \infty$. To do this we first state three facts which we will use repeatedly. By our choice of s_n' and $\delta_n = s_n'/\gamma_n$ in Lemma 6.2, and by Proposition 4.3,

$$\lim_{n\to\infty} 2\gamma_n v_n(\mathbb{1})\mathbf{P}^{(n)}(M_{V,V'} > 2\gamma_n\delta_n) = 1. \tag{6.22}$$

By (2) of Theorem 4.1, for each $t > 0$,

$$K_t = \sup_{n\in\mathbb{N}} 2\gamma_n v_n(\mathbb{1})\int_0^t \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)\,\mathrm{d}s < \infty. \tag{6.23}$$

Finally, by Markov property and (3.13), we have for $r < s$,

$$\mathbb{E}_{\mu_u}^{(n)}\big[\bar{p}(\xi_{\gamma_n s})|\mathscr{F}_r^n\big] = \mathbb{E}_{\xi_{\gamma_n r}}^{(n)}\big[\bar{p}(\xi_{\gamma_n(s-r)})\big] \le 2\mathbf{P}^{(n)}\big(M_{V,V'} > \gamma_n(s-r)\big). \tag{6.24}$$

We start with the first term on the right-hand side of (6.21), arguing in more detail than we will for the other terms. By conditioning at time $r < s$ and using (6.24) repeatedly, we obtain

$$\mathbb{E}_{\mu_u}^{(n)}\left[\int_{2\delta_n}^{T}\mathrm{d}r\int_r^{T\wedge(r+2\delta_n)}\mathrm{d}s\big(\gamma_n v_n(\mathbb{1})\big)^2\bar{p}(\xi_{\gamma_n r})\bar{p}(\xi_{\gamma_n s})\right]$$

$$= \int_{2\delta_n}^{T}\mathrm{d}r\int_r^{T\wedge(r+2\delta_n)}\mathrm{d}s\,\mathbb{E}_{\mu_u}^{(n)}\big[\gamma_n v_n(\mathbb{1})\bar{p}(\xi_{\gamma_n r})\mathbb{E}_{\mu_u}^{(n)}\big[\gamma_n v_n(\mathbb{1})\bar{p}(\xi_{\gamma_n s})|\mathscr{F}_r^n\big]\big]$$

$$\leq \int_{2\delta_n}^{T} dr \int_{r}^{T \wedge (r+2\delta_n)} ds \, \mathbb{E}_{\mu_u}^{(n)} \big[\gamma_n \nu_n(\mathbb{1}) \bar{p}(\xi_{\gamma_n r}) \big] 2\gamma_n \nu_n(\mathbb{1}) \mathbf{P}^{(n)} \big(M_{V,V'} > \gamma_n(s-r) \big)$$

$$\leq 2\gamma_n \nu_n(\mathbb{1}) \int_{2\delta_n}^{T} \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n r) \, dr \times 2\gamma_n \nu_n(\mathbb{1}) \int_{0}^{2\delta_n} \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s) \, ds$$

$$\leq K_T \times 2\gamma_n \nu_n(\mathbb{1}) \int_{0}^{2\delta_n} \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s) \, ds \longrightarrow 0 \quad \text{as } n \longrightarrow \infty,$$

where we have used (3.13), (6.23) and (2) of Theorem 4.1.

For the second term on the right-hand side of (6.21), again applying (6.24) repeatedly, we obtain

$$0 \leq \int_{2\delta_n}^{T} dr \int_{r}^{T \wedge (r+2\delta_n)} ds \, \mathbb{E}_{\mu_u}^{(n)} \big[\big(\gamma_n \nu_n(\mathbb{1}) \big)^2 \mathbb{E}_{\mu_u}^{(n)} \big[\bar{p}(\xi_{\gamma_n r}) | \mathscr{F}_{r-2\delta_n}^{n} \big] \bar{p}(\xi_{\gamma_n s}) \big]$$

$$\leq \int_{2\delta_n}^{T} dr \int_{r}^{T \wedge (r+2\delta_n)} ds \, 2\gamma_n \nu_n(\mathbb{1}) \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n 2\delta_n) 2\gamma_n \nu_n(\mathbb{1}) \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n s)$$

$$\leq \int_{2\delta_n}^{T} dr \int_{r}^{T \wedge (r+2\delta_n)} ds \, \big[2\gamma_n \nu_n(\mathbb{1}) \mathbf{P}^{(n)}(M_{V,V'} > 2\gamma_n \delta_n) \big]^2$$

$$\leq 2\delta_n T \times \big(2\gamma_n \nu_n(\mathbb{1}) \mathbf{P}^{(n)}(M_{V,V'} > 2\gamma_n \delta_n) \big)^2 \longrightarrow 0 \quad \text{as } n \longrightarrow \infty,$$

where we have made use of the fact that $s \geq r \geq 2\delta_n$ above, (6.22) and the fact that $\delta_n \to 0$.

The third term on the right-hand side of (6.21) is slightly different from the previous one. Now, we use (6.24) in the following way:

$$0 \leq \int_{2\delta_n}^{T} dr \int_{r}^{T \wedge (r+2\delta_n)} ds \, \mathbb{E}_{\mu_u}^{(n)} \big[\big(\gamma_n \nu_n(\mathbb{1}) \big)^2 \mathbb{E}_{\mu_u}^{(n)} \big[\bar{p}(\xi_{\gamma_n s}) | \mathscr{F}_{s-2\delta_n}^{n} \big] \bar{p}(\xi_{\gamma_n r}) \big]$$

$$\leq \int_{2\delta_n}^{T} dr \int_{r}^{T \wedge (r+2\delta_n)} ds \, 2\gamma_n \nu_n(\mathbb{1}) \mathbf{P}^{(n)}(M_{V,V'} > 2\gamma_n \delta_n) 2\gamma_n \nu_n(\mathbb{1}) \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n r)$$

$$\leq 2\delta_n \times 2\gamma_n \nu_n(\mathbb{1}) \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n 2\delta_n) \int_{0}^{T} dr \, 2\gamma_n \nu_n(\mathbb{1}) \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n r)$$

$$\leq 2\delta_n \times 2\gamma_n \nu_n(\mathbb{1}) \mathbf{P}^{(n)}(M_{V,V'} > \gamma_n 2\delta_n) \times K_T \longrightarrow 0 \quad \text{as } n \longrightarrow \infty,$$

which follows from (6.22), (6.23), and the fact that $\delta_n \to 0$.

Finally, for the last term on the right-hand side of (6.21), the bound (6.24) remains useful and we get

$$0 \leq \int_{2\delta_n}^{T} dr \int_{r}^{T \wedge (r+2\delta_n)} ds \, \mathbb{E}_{\mu_u}^{(n)} \big[\mathbb{E}_{\mu_u}^{(n)} \big[\gamma_n \nu_n(\mathbb{1}) \bar{p}(\xi_{\gamma_n r}) | \mathscr{F}_{r-2\delta_n}^{n} \big]$$

$$\times \mathbb{E}_{\mu_u}^{(n)} \big[\gamma_n \nu_n(\mathbb{1}) \bar{p}(\xi_{\gamma_n s}) | \mathscr{F}_{s-2\delta_n}^{n} \big] \big]$$

$$\leq \int_{2\delta_n}^{T} dr \int_{r}^{T \wedge (r+2\delta_n)} ds \, \big(2\gamma_n \nu_n(\mathbb{1}) \mathbf{P}^{(n)}(M_{V,V'} > 2\gamma_n \delta_n) \big)^2$$

$$\leq 2\delta_n T \times \big(2\gamma_n \nu_n(\mathbb{1}) \mathbf{P}^{(n)}(M_{V,V'} > 2\gamma_n \delta_n) \big)^2 \longrightarrow 0 \quad \text{as } n \longrightarrow \infty$$

since $\delta_n \to 0$ and we have (6.22). We have thus verified the desired convergence for the first term of (6.19).

We now make some observations for the other two terms in (6.19). To handle the second term, we apply the Markov property of the voter model to the integrand at time $s - 2\delta_n$. It follows from Lemma 6.2 that the integrand

$$\big| \mathbb{E}_{\mu_u}^{(n)} \big[\gamma_n \nu_n(\mathbb{1}) \bar{p}(\xi_{\gamma_n s}) | \mathscr{F}_{s-2\delta_n}^{n} \big] - p_1(\xi_{\gamma_n(s-2\delta_n)}) p_0(\xi_{\gamma_n(s-2\delta_n)}) \big|$$

is uniformly bounded by ε_n, so that the second term is no larger than $\varepsilon_n T$. By a simple change-of-variable argument, the third term above is easily seen to be bounded by $4\delta_n$. Since both of the sequences (ε_n) and (δ_n) tend to zero, the last two terms in (6.19) both tend to zero. This completes the proof of Theorem 2.2. $\square$

7. Coalescence times and density processes

Let (D_t) be the pure-death process on $\mathbb{N}$ which jumps from k to $k-1$ at rate $\binom{k}{2}$, $k \geq 2$. Set $Z_1 = \infty$, and recall that we let $Z_2, Z_3, \ldots$ be independent exponential variables with mean $\mathbf{E}[Z_j] = 1/\binom{j}{2}$. For any integer $k \geq 2$, it is easy to see from independence of Z_j and $\sum_{i=j+1}^{k} Z_i$ that

$$\mathbf{P}_k(D_t = j) = \mathbf{P}\left(\sum_{i=j+1}^{k} Z_i \leq t < \sum_{i=j}^{k} Z_i \right), \quad 1 \leq j \leq k.$$

Furthermore, (D_t) and the Wright–Fisher diffusion (Y_t) are linked by the following duality equation (see equation (7.21) of [27]):

$$\mathbb{E}_u[Y_t^k] = \mathbf{E}_k[u^{D_t}], \quad \forall u \in [0, 1], k \in \mathbb{N}, t \in \mathbb{R}_+. \tag{7.1}$$

The proofs of Proposition 2.5 and Proposition 2.6 are both based on this simple equality.

Proof of Proposition 2.5. Let us fix $t > 0$ and $k \geq 2$. By the duality equation (3.7), and the fact that the initial law of ξ_0 is μ_u,

$$\mathbb{E}_{\mu_u}^{(n)}\big[\big(p_1(\xi_{\gamma_n t})\big)^k\big] = \sum_{j=1}^{k} u^j \mathbf{P}^{(n)}\big(\big|\{\hat{X}_{t\gamma_n}^{U_1}, \ldots, \hat{X}_{t\gamma_n}^{U_k}\}\big| = j\big)$$

$$= \sum_{j=1}^{k} u^j \mathbf{P}^{(n)}\big(\mathsf{C}_{k,j}^{(n)} \leq \gamma_n t < \mathsf{C}_{k,j-1}^{(n)}\big), \tag{7.2}$$

with the convention that $\mathsf{C}_{k,k}^{(n)} = 0$. On the other hand, by assumption and the duality equation (7.1),

$$\lim_{n \to \infty} \mathbb{E}_{\mu_u}^{(n)}\big[\big(p_1(\xi_{\gamma_n t})\big)^k\big] = \mathbb{E}_u[Y_t^k]$$

$$= \mathbf{E}_k[u^{D_t}] = \sum_{j=1}^{k} u^j \mathbf{P}_k[D_t = j]$$

$$= \sum_{j=1}^{k} u^j \mathbf{P}\left(\sum_{i=j+1}^{k} Z_i \leq t < \sum_{i=j}^{k} Z_i \right). \tag{7.3}$$

Combining (7.2) and (7.3) we see that

$$\lim_{n \to \infty} \sum_{j=1}^{k} u^j \mathbf{P}^{(n)}\big(\mathsf{C}_{k,j}^{(n)} \leq \gamma_n t < \mathsf{C}_{k,j-1}^{(n)}\big) = \sum_{j=1}^{k} u^j \mathbf{P}\left(\sum_{i=j+1}^{k} Z_i \leq t < \sum_{i=j}^{k} Z_i \right).$$

The foregoing equality holds for all $u \in [0, 1]$, so it must be the case that

$$\lim_{n \to \infty} \mathbf{P}^{(n)}\big(\mathsf{C}_{k,j}^{(n)} \leq \gamma_n t < \mathsf{C}_{k,j-1}^{(n)}\big) = \mathbf{P}\left(\sum_{i=j+1}^{k} Z_i \leq t < \sum_{i=j}^{k} Z_i \right), \quad \forall 1 \leq j \leq k.$$

It follows by dominated convergence that for any $\lambda > 0$

$$\lim_{n\to\infty} \int_0^\infty \lambda e^{-\lambda t} \mathbf{P}^{(n)}\big(\mathsf{C}_{k,j}^{(n)} \le \gamma_n t < \mathsf{C}_{k,j-1}^{(n)}\big)\,dt = \int_0^\infty \lambda e^{-\lambda t} \mathbf{P}\bigg(\sum_{i=j+1}^k Z_i \le t < \sum_{i=j}^k Z_i\bigg)\,dt$$

and hence

$$\lim_{n\to\infty}\big(\mathbf{E}^{(n)}\big[e^{-\lambda \mathsf{C}_{k,j-1}^{(n)}/\gamma_n}\big] - \mathbf{E}^{(n)}\big[e^{-\lambda \mathsf{C}_{k,j}^{(n)}/\gamma_n}\big]\big) = \mathbf{E}\big[e^{-\lambda \sum_{i=j}^k Z_i}\big] - \mathbf{E}\big[e^{-\lambda \sum_{i=j+1}^k Z_i}\big]$$

for any $1 \le j \le k$ and our assertion follows plainly. $\square$

Proof of Proposition 2.6. The proof of Proposition 2.6 is a slight generalization of Proposition 2.5, so we will skip some details. We start with two equalities. First, as in (7.2), we have

$$\mathbb{P}_{\mu_u}^{(n)}\big(\tau_1^{(n)} \le \gamma_n t\big) = \mathbb{E}_{\mu_u}^{(n)}\bigg[\prod_{x\in E_n} \xi_{\gamma_n t}(x)\bigg]$$

$$= \sum_{j=1}^{|E_n|} u^j \mathbf{P}^{(n)}\big(\hat{\mathsf{C}}_j^{(n)} \le \gamma_n t < \hat{\mathsf{C}}_{j-1}^{(n)}\big). \tag{7.4}$$

Also by (7.1), we have

$$\mathbb{P}_u\big(\tau_1^Y \le t\big) = \lim_{k\to\infty} \mathbb{E}_u\big[Y_t^k\big] = \sum_{j=1}^\infty u^j \mathbf{P}\bigg(\sum_{i=j+1}^\infty Z_i \le t < \sum_{i=j}^\infty Z_i\bigg). \tag{7.5}$$

That (2.18) implies (2.17) now follows from the two displays (7.4) and (7.5) and dominated convergence.

The converse also uses the same two displays, but now we need another elementary result: For any nonnegative a_j^n, for $n, j \in \mathbb{N}$, with $\sum_j a_j^n \le 1$, the condition that

$$\lim_{n\to\infty} \sum_j a_j^n u^j \quad \text{exists for every } u \in (0, 1) \tag{7.6}$$

is enough to obtain that $\lim_{n\to\infty} a_j^n$ exists for every $j \in \mathbb{N}$. Indeed, if (n_k) and (n_k') are two subsequences such that $\lim_{k\to\infty} a_j^{n_k}$ and $\lim_{k\to\infty} a_j^{n_k'}$ exist for all $j \in \mathbb{N}$, then the limits are all in $[0, 1]$, and so by dominated convergence (7.6) implies

$$\sum_j \Big(\lim_{k\to\infty} a_j^{n_k}\Big) u^j = \sum_j \Big(\lim_{k\to\infty} a_j^{n_k'}\Big) u^j, \quad u \in [0, 1).$$

We deduce from these that $\lim_{k\to\infty} a_j^{n_k} = \lim_{k\to\infty} a_j^{n_k'}$ for all $j \in \mathbb{N}$, which, by a diagonal argument on selecting convergent subsequences of $(a_j^n)_{n\in\mathbb{N}}$ each $j \in \mathbb{N}$, is enough for our claim that $\lim_{n\to\infty} a_j^n$ exists for every $j \in \mathbb{N}$.

Using this elementary result, and assuming (2.17) so that the right-hand side of (7.4) converges to the right-hand side of (7.5), we obtain that for every $j \in \mathbb{N}$,

$$\lim_{n\to\infty} \mathbf{P}^{(n)}\big(\hat{\mathsf{C}}_j^{(n)} \le \gamma_n t < \hat{\mathsf{C}}_{j-1}^{(n)}\big)$$

must exist and this limit must be

$$\mathbf{P}\bigg(\sum_{i=j+1}^\infty Z_i \le t < \sum_{i=j}^\infty Z_i\bigg).$$

This establishes (2.18), and the proof is complete. $\square$

8. Examples

In this section, we consider various sequences of $(q^{(n)}, E_n)$-Markov chains for which one of the conditions of Theorem 2.2 and Corollary 2.3 applies, and hence the convergence of the corresponding voter model densities in (2.7) holds.

The Q-matrices $q^{(n)}$ considered below are of the form $q^{(n)} = p^{(n)} - \mathrm{Id}_{E_n}$, where $p^{(n)}$ is a symmetric probability matrix but not necessarily has zero diagonal. In this case, $1 - \lambda$ is an eigenvalue of $-q^{(n)}$ if and only if λ is an eigenvalue of $p^{(n)}$. If in addition $p^{(n)}$ has zero diagonal, the inequality (3.21) for such a particular Q-matrix $q^{(n)}$ becomes

$$\mathbf{t}_{\text{meet}}^{(n)} \geq \frac{(|E_n| - 1)^2}{4|E_n|}. \tag{8.1}$$

All our examples below can be viewed as random walks on graphs, although we do not use this language for the examples in Section 8.1 which include and generalize Theorem 2 of [5].

8.1. *Discrete tori*

For $n, d \in \mathbb{N}$, we consider irreducible $(q^{(n)}, E_n)$-Markov chains where for $d, n \in \mathbb{N}$,

$$E_n = \big((-n/2, n/2] \cap \mathbb{Z}\big)^d$$

and $q^{(n)}(x, y) = q^{(n)}(0, y - x)$ for $x \neq y$. Here, the difference $y - x$ is read coordinate-wise mod n. By the assumed symmetry of $q^{(n)}$, the bound (8.1) applies.

8.1.1. *Nearest-neighbor walk*
Assume $d \geq 2$ and $q^{(n)}(x, y) = (2d)^{-1}$ if $|x - y| = 1$ (the difference is computed mod n coordinate-wise). Then as $n \longrightarrow \infty$, $\mathbf{t}_{\text{mix}}^{(n)} = \mathrm{O}(n^2)$ in all dimensions d (see Theorem 5.5 in [17]) and

$$\mathbf{t}_{\text{meet}}^{(n)} \sim \begin{cases} \frac{1}{2\pi} |E_n| \log |E_n| & \text{if } d = 2, \\ G_d |E_n| & \text{if } d \geq 3, \end{cases} \tag{8.2}$$

where the constant G_d is the expected number of visits to the origin by a simple symmetric random walk in $\mathbb{Z}^d$ starting at the origin (see [5]). Hence, (i) of Theorem 2.2 holds, and we have the convergence of voter model densities to the Wright–Fisher diffusion in (2.7). This result was first obtained in Theorem 2 of [5].

Remark 8.1. *As in Section* 13.2.3 *in* [2], *we say that the sequence* $(q^{(n)}, E_n)_{n \in \mathbb{N}}$ *is transient if* $\sup_{n \in \mathbb{N}} \mathbf{t}_{\text{meet}}^{(n)}/|E_n|$ *is finite, and is recurrent otherwise. We note that the asymptotic behavior in* (8.2) *indicates recurrence for* $d = 2$ *and transience for* $d \geq 3$. *This is consistent with the fact that simple symmetric random walk on* $\mathbb{Z}^d$ *is recurrent if* $d = 2$ *and is transient if* $d \geq 3$.

With this notion in mind, we note that (8.1) *gives the correct asymptotic rate of growth for* $\mathbf{t}_{\text{meet}}^{(n)}$ *for the transient case* $d \geq 3$, *but not for the recurrent case* $d = 2$.

8.1.2. *Intermediate-range random walk*
We consider the random walks studied in [6], which have range tending to infinity. Let (m_n) be a sequence of positive integers such that $m_n < n/2$ for all n and $m_n \to \infty$. For any $d \geq 1$, let

$$\Lambda_n = \Lambda_n^d = \big([-m_n, m_n] \cap \mathbb{Z}\big)^d \setminus \{0\}, \tag{8.3}$$

and put

$$q^{(n)}(x, y) = |\Lambda_n|^{-1} \quad \text{if } y - x \in \Lambda_n \tag{8.4}$$

(again the difference $y - x$ is read mod n coordinate-wise).

Proposition 8.2. *Assume $d = 2$ and*

$$\lim_{n \to \infty} \frac{m_n^2}{\log n} = 0. \tag{8.5}$$

Then

$$\mathbf{t}_{\mathrm{mix}}^{(n)} = \mathrm{O}\big(n^2/m_n^2\big) \quad \text{as } n \to \infty, \tag{8.6}$$

$$\liminf_{n \to \infty} \frac{\mathbf{t}_{\mathrm{meet}}^{(n)}}{n^2 \log n / m_n^2} > 0. \tag{8.7}$$

Taken together, (8.6) and (8.7) imply condition (i) of Theorem 2.2, and so we have the convergence of voter model densities in (2.7).

Proof of Proposition 8.2. To obtain (8.6) and (8.7), we make use of results from of [6]. Since conditions (P1)–(P3) in [6] hold by Proposition 1.1 there, we deduce from Theorem 1.7 of [6] that if $\lim_{n \to \infty} s_n/(n^2/m_n^2) = \infty$, then

$$\sum_{x \in E_n} \big| q_{s_n}^{(n)}(0, x) - \pi_n(x) \big| \leq n^2 \sup_{x \in E_n} \left| q_{s_n}^{(n)}(0, x) - \frac{1}{n^2} \right| \to 0,$$

which implies (8.6).

Next, to get (8.7), we first reduce $\mathbf{t}_{\mathrm{meet}}^{(n)}$ to a simpler time. Let (X_t) be the Markov chain on E_n given by $q^{(n)}$, and

$$H_0 = \inf\{t \geq 0 : X_t = 0\}$$

be the hitting time of 0. Since the difference of two rate-one random walks is a rate-two random walk (see also Proposition 7.1 and Proposition 14.5 of [2] for a more general fact), we have

$$\mathbf{E}^{(n)}[M_{U,U'}] = \frac{\mathbf{E}_{\pi^{(n)}}^{(n)}[H_0]}{2}. \tag{8.8}$$

The limit (8.7) can then be derived from the estimates on the expectations $\mathbf{E}_x^{(n)}[H_0]$ given in Theorem 1.3 in [6], but we will use instead the following simpler argument, which relies on only (6.1) from [6].

We now claim

$$\liminf_{n \to \infty} \frac{\mathbf{E}_{\pi^{(n)}}^{(n)}[H_0]}{n^2 \log n / m_n^2} \geq \frac{12}{\pi}, \tag{8.9}$$

as entails (8.7) by (8.8). For $x \in E_n$ and $\lambda > 0$, let

$$G_n(x, \lambda) = \int_0^\infty \mathrm{e}^{-\lambda s} q_s^{(n)}(0, x) \, \mathrm{d}s.$$

By standard Markov chain arguments,

$$\mathbf{E}_x^{(n)}\big[\exp(-\lambda H_0)\big] = \frac{G_n(x, \lambda)}{G_n(0, \lambda)}. \tag{8.10}$$

Clearly $\sum_{x \in E_n} \pi^{(n)}(x) G_n(x, \lambda) = \frac{1}{n^2 \lambda}$, and thus by (8.10),

$$\mathbf{E}_{\pi^{(n)}}^{(n)}\big[\mathrm{e}^{-\lambda H_0}\big] = \frac{(n^2 \lambda)^{-1}}{G_n(0, \lambda)}. \tag{8.11}$$

According to (6.1) of [6] with $t_n = \log n / m_n^2$, we have

$$\lim_{n \to \infty} \frac{G_n(0, \lambda/(n^2 t_n))}{t_n} = \lambda^{-1} + \frac{12}{\pi}.$$

Applying this fact to (8.11), we get

$$\lim_{n \to \infty} \mathbf{E}_{\pi^{(n)}}^{(n)} \left[e^{-\lambda H_0 / n^2 t_n} \right] = \frac{1}{1 + (12/\pi)\lambda}.$$

We deduce (8.9) from the Skorokhod representation and Fatou's lemma. The proof is complete. $\square$

In view of (8.7), the condition $m_n^2 / \log n \to 0$ implies $\mathbf{t}_{\text{meet}}^{(n)} / |E_n| \to \infty$. This means the Markov chain sequences considered in this example are, like the nearest-neighbor $d = 2$ case, recurrent. Also, although we will not give the details here, Theorem 2.2 still holds if instead of (8.5) we consider the (transient) case in which (8.5) is replaced with $\lim_{n \to \infty} m_n^2 / \log n = \infty$.

Proposition 8.3. *Assume $d = 1$ and*

$$\lim_{n \to \infty} \frac{m_n}{\sqrt{n}} = \infty. \tag{8.12}$$

For $q^{(n)}$ given as in (8.4), let $\mathbf{g}_n$ denote the spectral gap of $q^{(n)}$. Then

$$\lim_{n \to \infty} n \mathbf{g}_n = \infty. \tag{8.13}$$

Assuming Proposition 8.3 for now, we may write the second condition in (ii) of (2.2) in the form

$$\frac{\log(e \vee \mathbf{t}_{\text{meet}}^{(n)} \sqrt{\pi_{\max}^{(n)} q_{\max}^{(n)} \nu_n(\mathbb{1})})}{\mathbf{g}_n \mathbf{t}_{\text{meet}}^{(n)}} \leq \frac{\log[e \vee (\mathbf{t}_{\text{meet}}^{(n)} / n)]}{(n \mathbf{g}_n)(\mathbf{t}_{\text{meet}}^{(n)} / n)}.$$

By the meeting time bound (8.1), $\mathbf{t}_{\text{meet}}^{(n)} / n$ is bounded away from 0, and thus (8.13) implies that the right-hand side above tends to 0. That is, condition (ii) of Theorem 2.2 holds and we obtain convergence of voter model densities to the Wright–Fisher diffusion.

For the proof of Proposition 8.3, we recall the definition of the bottleneck ratio Φ_* here. For a reversible Markov chain (q, E) with $q = p - I$ for a probability matrix p with zero diagonal, define

$$\Phi(S) = \frac{\sum_{x \in S, y \in S^\complement} \pi(x) q(x, y)}{\pi(S)}, \quad S \subset E, \tag{8.14}$$

and

$$\Phi_*(q) = \min \left\{ \Phi(S); S \subset E, \pi(S) \leq \frac{1}{2} \right\}. \tag{8.15}$$

The inequality we need is

$$\mathbf{g} \geq \frac{1}{2} \left(\Phi_*(q) \right)^2. \tag{8.16}$$

See Section 13.3.2 in [17] for this inequality, and note that $\mathbf{g}$ is equal to $1 - \lambda$ for λ being the second largest eigenvalue of p.

Proof of Proposition 8.3. It is easy to see from the definition (8.14) that

$$\Phi(S) = \frac{|\partial S|}{2m_n|S|}, \tag{8.17}$$

where $\partial S = \{(x, y); x \in S, y \in S^{\complement}, 1 \le |x - y| \le m_n\}$. For $1 \le k \le n/2$, let I_k be an "interval" of k elements in E_n:

$$I_k = \{0, 1, \ldots, k - 1\}.$$

A little thought shows that the minimum of $|\partial S|$ among all S with $|S| = k$ is obtained by taking $S = I_k$, which implies that

$$\Phi_* = \min\{\Phi(I_k); 1 \le k \le \lfloor n/2 \rfloor\}.$$

It is easy to check that if $m_n < k \le n/2$, then

$$|\partial I_k| = 2 \sum_{j=1}^{m_n} j = m_n(m_n + 1).$$

Similarly, if $1 \le k \le m_n$, then $|\partial I_k|$ is

$$2 \sum_{j=m_n-k+1}^{m_n} j = m_n(m_n + 1) - (m_n - k)(m_n - k + 1) = k(2m_n - k + 1).$$

It follows from (8.17) that

$$\Phi(I_k) = \begin{cases} \frac{m_n+1}{2k} & \text{if } m_n < k \le n/2, \\ \frac{2m_n-k+1}{2m_n} & \text{if } 1 \le k \le m_n. \end{cases}$$

Taking $k = n/2$, we see that

$$\Phi_*(q^{(n)}) = \Phi(I_{\lfloor n/2 \rfloor}) = \frac{m_n + 1}{2\lfloor n/2 \rfloor}.$$

It is now immediate from (8.12) and the inequality (8.16) that

$$\lim_{n \to \infty} n\mathbf{g}_n \ge \lim_{n \to \infty} \frac{n}{2} \cdot \frac{m_n^2}{4\lfloor n/2 \rfloor^2} = \infty,$$

which completes the proof. $\qquad \square$

Although we will not prove it here, the condition (8.12) implies that $\mathbf{t}_{\text{meet}}^{(n)} = O(n)$, which means that the chain considered in Proposition 8.3 is transient.

8.2. *Random walk on simple graphs*

We consider in this section graphs which are simple, that is have no loops or multiple edges, and are connected. The simple random walk on such a graph $G = (\mathsf{V}, \mathsf{E})$ with vertex set V and edge set E is the Markov chain (q, V) with $q(x, y) = 1/\deg(x)$ if (x, y) is an edge for $x \ne y$. Note that q is reversible with stationary distribution

$$\pi(x) = \frac{\deg(x)}{2|\mathsf{E}|}.$$

See [20] for a survey and the standard terminology of random walks on graphs.

8.2.1. *Hypercubes*

For $n \in \mathbb{N}$, take $\mathsf{V}_n = \{0, 1\}^n$ and for $x, y \in \mathsf{V}_n$ let $|x - y| = \sum_{j=1}^n |x_i - y_i|$. We draw an edge between any $x, y \in \mathsf{V}_n$ with $|x - y| = 1$, and obtain the n-dimensional hypercube, a connected n-regular graph. The random walk Q-matrix $q^{(n)}$ on this graph is given by $q^{(n)}(x, y) = 1/n$ if $|x - y| = 1$, and is irreducible and symmetric with $\pi(x) \equiv 2^{-n}$. Furthermore, it is known (see Example 5.15 in [2]) that

$$\mathbf{g}_n = \left(\frac{n}{2}\right)^{-1} \quad \text{and} \quad \mathbf{t}_{\text{meet}}^{(n)} \sim 2^{n-1} \quad \text{as } n \longrightarrow \infty.$$

It is easy to see from these facts that (ii) of Theorem 2.2 is satisfied.

8.2.2. *Expander graphs*

Fix $\alpha \in (0, \infty)$ and $k \in \mathbb{N}$ with $k \geq 3$, and take a (k, α)-expander family of graphs (G_n) with corresponding random walk Q-matrices $q^{(n)}$. Here as in Section 13.6 of [17], (G_n) is a graph sequence such that the number of vertices of G_n tends to infinity, each G_n is connected and k-regular, and satisfies

$$\Phi_*\big(q^{(n)}\big) \geq \alpha, \quad \forall n \in \mathbb{N}$$

(see (8.15) for notation). By (8.16), $\liminf \mathbf{g}_n \geq \frac{1}{2}\alpha^2$, and thus the conditions of Corollary 2.3 apply.

8.3. *Random walk on general graphs*

We now consider finite graphs without the simplicity condition, nor the connectivity condition. For such a graph G with vertex set V, its edge set E is now defined by using an adjacency matrix

$$A : \mathsf{V} \times \mathsf{V} \longrightarrow \mathbb{Z}_+$$

with $A(x, x) \in \{0, 1, 2\}$, so that $A(x, y)$ gives the number of edges joining x and y. For $x \neq y$, $A(x, y)$ simply gives the number of edges between x and y. In Section 8.3.1 below, we consider several models of random graphs due to Friedman [14] in which the convention is that $A(x, x) = 1$ means a "half-loop" at x, and $A(x, x) = 2$ means a "whole-loop" at x.

If we take a sequence (G_n) of such general graphs with $G_n = (\mathsf{V}_n, \mathsf{E}_n)$ and E_n being encoded by A_n, then the $q^{(n)}$-random walk on G_n has Q-matrix defined by

$$q^{(n)}(x, y) = \frac{A_n(x, y)}{A_n(x)}, \quad x \neq y,$$

where $A_n(x) = \sum_{y \in \mathsf{V}_n} A_n(x, y)$. Hence, $q^{(n)} = p^{(n)} - I$, where the xth row of $p^{(n)}$ is obtained by dividing the xth row of A_n by $A_n(x)$. In this case, the second largest eigenvalue of the transition matrix $p^{(n)}$ is different from 1 if and only if the graph G_n is connected, and so the second smallest eigenvalue of $-q^{(n)}$ is different from 0 if and only if the graph G_n is connected.

8.3.1. *Random regular graphs*

The work [14] considers various models of growing random k-regular graphs (G_n) on n vertices (see the models $\mathcal{G}_{n,k}$, $\mathcal{H}_{n,k}$, $\mathcal{I}_{n,k}$, and $\mathcal{J}_{n,k}$ there), and each is defined for a large set of admissible degrees k. For simplicity, we only consider the model $\mathcal{G}_{n,k}$ below, although the following discussion applies to other models $\mathcal{H}_{n,k}$, $\mathcal{I}_{n,k}$, and $\mathcal{J}_{n,k}$ in [14] for moderately large admissible degrees k as well.

The random regular graphs $\mathcal{G}_{n,k}$ are defined for even integers k with $\mathsf{V}_n = \{1, \dots, n\}$, and for each n the edge set is given by

$$\mathsf{E}_n = \big\{\big(x, \rho_j(x)\big), \big(x, \rho_j^{-1}(x)\big); j = 1, \dots, k/2, x \in \mathsf{V}_n\big\},$$

where $\rho_1, \ldots, \rho_{k/2}$ are i.i.d. permutations of $\{1, \ldots, n\}$ and each ρ_i is chosen uniformly from the set of $n!$ permutations. Then for any even integer k, we have

$$\lim_{n \to \infty} \mathbb{P}\left(\max_{2 \leq i \leq n} \left| \lambda_i(\mathcal{G}_{n,k}) \right| \leq \frac{2\sqrt{k-1}}{k} + \varepsilon \right) = 1, \quad \forall \varepsilon > 0,$$

where

$$1 = \lambda_1(\mathcal{G}_{n,k}) \geq \lambda_2(\mathcal{G}_{n,k}) \geq \cdots \geq \lambda_n(\mathcal{G}_{n,k})$$

are the ordered eigenvalues associated with the normalized adjacency matrix $p^{(n)}$ on $\mathcal{G}_{n,k}$. More precisely, we have

$$\mathbb{P}\left(\max_{2 \leq i \leq n} \left| \lambda_i(\mathcal{G}_{n,k}) \right| > \frac{2\sqrt{k-1}}{k} + \varepsilon \right) \leq \frac{c}{n^{\lceil (\sqrt{k-1}+1)/2 \rceil - 1}}, \tag{8.18}$$

where c is a constant. See Theorem 1.1 in [14], and also [21] for estimates of mixing times on other random regular graphs.

If we assume in addition that $\mathcal{G}_{n,k}$, for $n \in \mathbb{N}$, are independent random graphs, then it follows from (8.18) and the Borel–Cantelli lemma that for each even $k \geq 12$,

$$\liminf_{n \to \infty} \mathbf{g}(\mathcal{G}_{n,k}) > 0 \quad \text{a.s.}$$

Since the stationary distribution of $q^{(n)}$ is always uniform, the sequence of $q^{(n)}$-Markov chains now satisfies the conditions of Corollary 2.3 with probability one (with respect to the randomness of $q^{(n)}$). We obtain the convergence of voter model densities (2.7) along $(\mathcal{G}_{n,k})_{n \in \mathbb{N}}$ with probability one.

References

[1] D. J. Aldous. Markov chains with almost exponential hitting times. *Stochastic Process. Appl.* **13** (1982) 305–310. MR0671039

[2] D. J. Aldous and J. A. Fill. *Reversible Markov Chains and Random walks on Graphs*. Unfinished monograph, recompiled 2014. Available at http://www.stat.berkeley.edu/~aldous/RWG/book.html.

[3] P. Caputo, T. M. Liggett and T. Richthammer. Proof of Aldous' spectral gap conjecture. *J. Amer. Math. Soc.* **23** (2010) 831–851. MR2629990

[4] Y.-T. Chen. Sharp benefit-to-cost rules for the evolution of cooperation on regular graphs. *Ann. Appl. Probab.* **23** (2013) 637–664. MR3059271

[5] J. T. Cox. Coalescing random walks and voter model consensus times on the torus in $\mathbb{Z}^d$. *Ann. Probab.* **17** (1989) 1333–1366. MR1048930

[6] J. T. Cox. Intermediate range migration in the two-dimensional stepping stone model. *Ann. Appl. Probab.* **20** (2010) 785–205. MR2680548

[7] J. T. Cox, R. Durrett and E. A. Perkins. Rescaled voter models converge to super-Brownian motion. *Ann. Probab.* **28** (2000) 185–234. MR1756003

[8] J. T. Cox, R. Durrett and E. A. Perkins. Voter model perturbations and reaction diffusion equations. *Astérisque* **349** (2013) 1–113. MR3075759

[9] J. T. Cox, M. Merle and E. A. Perkins. Coexistence in a two-dimensional Lotka–Volterra model. *Electron. J. Probab.* **15** (2010) 1190–1266. MR2678390

[10] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.* **6** (1996) 695–750. MR1410112

[11] R. Durrett. *Random Graph Dynamics. Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge Univ. Press, Cambridge, 2010. MR2656427

[12] P. Donnelly and D. Welsh. Finite particle systems and infection models. *Math. Proc. Cambridge Philos. Soc.* **94** (1983) 167–182. MR0704809

[13] S. N. Ethier and T. G. Kurtz. *Markov Processes, Characterization and Convergence. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. Wiley, New York, 1986. MR0838085

[14] J. Friedman. A proof of Alon's second eigenvalue conjecture and related problems. *Mem. Amer. Math. Soc.* **195** (910) (2008) 1–100. MR2437174

[15] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*, 2nd edition. *Grundlehren der Mathematischen Wissenschaften* **288**. Springer, Berlin, 2003. MR1943877

[16] O. Kallenberg. *Foundations of Modern Probability*, 2nd edition. *Probability and Its Applications*. Springer, New York, 2002. MR1876169

[17] D. A. Levin, Y. Peres and E. L. Wilmer. *Markov Chains and Mixing Times*. Amer. Math. Soc., Providence, RI, 2009. MR2466937

[18] T. M. Liggett. *Interacting Particle Systems. Grundlehren der Mathematischen Wissenschaften* **276**. Springer, New York, 1985. MR0776231

[19] T. M. Liggett. *Continuous Time Markov Processes. An Introduction. Graduate Studies in Mathematics* **113**. Amer. Math. Soc., Providence, RI, 2010. MR2574430

[20] L. Lovász. Random walks on graphs: A survey. In *Combinatorics, Paul Erdös Is Eighty, Vol. 2* (*Keszthely, 1993*). *Bolyai Soc. Math. Stud.* **2** 353–397. János Bolyai Math. Soc., Budapest, 1996. MR1395866

[21] E. Lubetzky and A. Sly. Cutoff phenomena for random walks on random regular graphs. *Duke Math. J.* **153** (2010) 475–510. MR2667423

[22] C. Mueller and R. Tribe. Stochastic p.d.e.'s arising from the long range contact and long range voter processes. *Probab. Theory Related Fields* **102** (1995) 519–545. MR1346264

[23] R. I. Oliveira. Mean field conditions for coalescing random walks. *Ann. Probab.* **41** (2013) 3420–3461. MR3127887

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

[25] H. Ohtsuki, C. Hauert, E. Lieberman and M. A. Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441** (2006) 502–505.

[26] V. Sood and S. Redner. Voter model on heterogeneous graphs. *Phys. Rev. Lett.* **94** (2005) 178701.

[27] S. Tavaré. Line-of-descent and genealogical processes, and their applications in population genetics models. *Theoret. Population Biol.* **26** (1984) 119–164. MR0770050