5-14-2023

Rescaled Density Processes of Voter Model Perturbations on r-Regular Random Graphs converge to Feller's Branching Diffusion

Tianyue Wu
Syracuse University

Abstract

Voter model perturbations can be viewed as voter model (neutral competition) plus a small perturbation rate. Cox [C17] showed that the biased voter model, viewed as a voter model perturbation, converges to Feller's branching diffusion under mild mixing condition. We extend this result to a general class of perturbation functions on the setting of r-regular random graphs where the nearest-neighbor voting kernel has a strong mixing property, and prove a low-density diffusive limit of which the convergence of biased voter model is considered as a special case. The other special case considered is the q-voter model whose high-density ODE limit on torus for $q \approx 1$ has been proved by Agarwal, Simper and Durrett [ASD21]. We will introduce the low-density approach we use and show that a mean-field simplification occurs.

Rescaled Density Processes of Voter Model Perturbations on r-Regular Random Graphs converge to Feller's Branching Diffusion

by

Tianyue Wu

B.S., University of Washington - Seattle, 2016

M.S., Syracuse University, 2019

Dissertation

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Mathematics

Syracuse University

May 2023

Acknowledgements

I would like to thank Ted Cox for numerous discussions.

Contents

List of Figures

List of Notation

V_N	set of N vertices
V^{tr}	set of vertices of the infinite tree
R	interaction range, fixed
$d(x,y)$	graph distance
$\mathcal{N}(x)$	$\{y : 1 \leq d(x,y) \leq R\}$, $x \in V_N$
$\mathcal{N}_l(x)$	$\{y : d(x,y) = l\}$
$\overline{\mathcal{N}(x)}$	$\mathcal{N}(x) \cup \{x\}$
$\overline{\mathcal{N}_l(x)}$	$\mathcal{N}_l(x) \cup \{x\}$
$n_l^i(x,\xi)$	$\sum_{y \in \mathcal{N}_l(x)} 1_{\{\xi(y)=i\}}$, number of type i in $\mathcal{N}_l(x)$ in ξ
$S_{l,k}(x)$	$\{A \subseteq \mathcal{N}_l(x) : \lvert A\rvert = k\}$
$\widehat{B}_t^{N,x}$	rate 1 coalescing random walk starting from $x \in V_N$
$\widehat{B}_t^{e}$	rate 1 coalescing random walk starting from $e \in V^{tr}$
$\widehat{B}_t^{N,A}$	$\{\widehat{B}_t^{N,x} : x \in A\}$ for $A \subseteq V_N$
$\widehat{B}_t^{A}$	$\{\widehat{B}_t^{e} : e \in A\}$ for finite $A \subset V^{tr}$
$\tau^N(x,y)$	$\inf\{t \geq 0 : \widehat{B}_t^{N,x} = \widehat{B}_t^{N,y}\}$
$\tau^{tr}(\rho,e)$	$\inf\{t \geq 0 : \widehat{B}_t^{\rho} = \widehat{B}_t^{e}\}$
$\tau^N(A,B)$	$\inf\{t \geq 0 : \widehat{B}_t^{N,A} \cap \widehat{B}_t^{N,B} \neq \varnothing\}$

$\tau^{tr}(A,B)$	$\inf\{t \geq 0 : \widehat{B}_t^A \cap \widehat{B}_t^B \neq \varnothing\}$
$\sigma^N(A)$	$\inf\{t \geq 0 : \lvert\widehat{B}_t^{N,A}\rvert = 1\}$
$\sigma^{tr}(A)$	$\inf\{t \geq 0 : \lvert\widehat{B}_t^A\rvert = 1\}$

Chapter 1

Assumptions and main result

In [C17], Cox showed that rescaled density processes of biased voter models on large finite sets converge to Feller's branching diffusion, under the *low initial density* condition: the initial number of particles has a smaller order than the size of the space. In their work, the biased voter model is viewed as a voter model perturbation and voting kernels are assumed to have minimum mixing property: no particular site can be visited significantly more often than others, and that t^N_{meet}, the expected meeting time of two walks starting from stationary, has a larger order than the separation time t^N_{sep}. The time and mass scale are assumed to be between t^N_{sep} and t^N_{meet}.

We extend the result in Cox [C17] to a class of generalized voter model perturbations and prove a convergence for rescaled density processes to branching diffusion for the asymptotics on r-regular random graphs. The existence of the scale for getting this diffusive limit is implied by that the mixing time t^N_{mix} has a smaller order than t^N_{meet} on these graphs. Walks with this property meet either quickly, or do not meet until they "realize" that the space is finite. This property also implies that macroscopic independence will happen under suitable large time scale depending on initial number of particles.

Our introduction is structured as follows. In Section 1.1, we give the definition of voter model perturbations and state its generator. We define the rescaled density processes on r-regular graphs and state the low initial density assumption and perturbation assumptions. In Section 1.2, we define the collection of "good" r-regular graphs and show that with high probability, a simple, connected r-regular random graphs is a good graph. In particular, these graphs have strong bound on the transition probabilities. Our assumption on the time scale will be that its order is between t^N_{mix} and t^N_{meet}. We are able to relax the lower bound t^N_{sep} used in [C17] to t^N_{mix} as the mixing time for good r-regular random graphs is sufficient for providing the control on the separation distance defined in (1.5). This fact will be used frequently in later proofs.

We state the main result in Section 1.3 by using the martingale characterization of the branching diffusion with a description of the limiting drift and branching term using coalescing random walk probabilities. We will explain the implicit "double-layer" randomness in this result, and the fundamental reason for getting the branching limit.

We will describe two examples of the main result in Section 1.4. The first is the biased voter model considered in Cox [C17] and was first introduced by Williams and Bjerknes [WB79] as a model of tumor growth. This model has been studied in many literatures such as Durrett, Foo and Leder [DFL16] in which they studied the spatial Moran model that is considered as a generalization of biased voter models, and Bramson, Griffeath [BG81] where they gave an estimation on the size of the cluster of the biased voter model starting from a single particle on lattices.

The second example is the q-voter model which was introduced in Nettle [N99] as a model of language change in social networks, and in Abraams and Strogatz [AS03] as a model of language death. Under high initial density condition that initial number of particles is equal to the order of size of the space, Agarwal, Simper and Durrett [ASD21], the first paper related to the study of this model in mathematics, proved an ODE limit on torus under homogeneous mixing condition in Theorem 1.1 and 1.2 in their work. The idea followed Cox, Durrett [CD16]: as shown in Theorem 6 in Section 7.2 of [CD16], the reaction function in the limiting ODE is given by the expectation of perturbations under voter stationarity. We will give the low density diffusive limit of the q-voter model for $q < 1$ close to 1 as a corollary of our main result. The connection between our low density diffusive limit and the high density ODE limit is the following: when under low initial density, the local density of 0 at a site is very close to 1. Thus, the limiting drift in the low density diffusive limit is exactly equal to the derivative at 0 of the reaction function in the high density ODE. For complete details, see Section 1.8 in Cox, Durrett and Perkins [CDP13].

In Section 1.5, we make a comparison between the low density branching limit theorem and the Wright-Fisher limit theorem for voter models that is proved in Chen, Choi and Cox [CCC16] where a mean-field approximation for this model is used and t^N_{meet} serves as the time scale. And finally, we give an outline of the low density approach we use in Section 1.6.

1.1 Voter model perturbations

Let us begin with some definitions. Let V_N be a set of N vertices and $[N] = \{1, .., N\}$ be a numbering on it. We will assume that $V_N = [N]$. Suppose that G_N is an r-regular graph built on V_N with $r \geq 3$. Let d be the graph distance and denote $y \sim x$ if $d(x, y) = 1$. Define the nearest neighbor transition kernel q^N by

$$q^N(x, y) = 1/r \cdot 1_{\{y \sim x\}}, \quad x, y \in V_N.$$

For $\xi \in \{0, 1\}^{V_N}$, define the local densities $f^N_i = f^N_i(x, \xi)$ by

$$f^N_i(x, \xi) = \sum_{y \in V_N} q^N(x, y) \cdot 1_{\{\xi(y)=i\}}, \quad i = 0, 1.$$

For $x \in V_N$ and $l \in \mathbb{N}$, denote $\mathcal{N}_l(x) = \{y : d(x, y) = l\}$ and define

$$n^i_l(x, \xi) = \sum_{y \in \mathcal{N}_l(x)} 1\{\xi(y) = i\}, \ i = 0, 1$$

so that $n^i_l(x, \xi)$ is the count of type i in the boundary of distance-l neighborhood of x in ξ.

Let $|A|$ denote the cardinality of subset A. To simplify the story, we consider only finite interaction range which is enough for exhibition of the phenomena. Let the interaction range $R \geq 1$ be fixed. Define

$$N_R = r(r-1)^{R-1}.$$

so that for any $x \in V_N$, N_R satisfies

$$\max_{l=1,..,R} |\mathcal{N}_l(x)| \leq N_R.$$

Let $\theta^{N,i}$ be functions on $\{(l,k) : 1 \leq l \leq R,\ 1 \leq k \leq N_R\}$ for each $N \in \mathbb{N}$ and $i = 0, 1$. Define the perturbation functions

$$F_i^N(x,\xi) = \sum_{l=1}^{R} \sum_{k=1}^{N_R} \theta^{N,i}(l,k) 1_{\{n_l^i(x,\xi)=k\}}, \quad i = 0, 1.$$

Let $\tau_N \to \infty$ be a positive sequence. Define $\xi_t^{G_N}$ as the $\{0,1\}^{V_N}$-valued Markov process with rates that at each event time at site x, $\xi(x)$ makes transitions

$$\begin{aligned} 0 \to 1 \quad &\text{at rate } \tau_N f_1^N + F_1^N, \\ 1 \to 0 \quad &\text{at rate } \tau_N f_0^N + F_0^N. \end{aligned} \tag{1.1}$$

We will denote $\widehat{\xi}(x) = 1 - \xi(x)$ for convenience. Let ξ^x be the configuration

$$\xi^x(y) = \begin{cases} \xi(y), & y \neq x, \\ \widehat{\xi}(x), & y = x, \end{cases}$$

and define rate functions

$$c_N^v(x,\xi) = \widehat{\xi}(x) f_1^N(x,\xi) + \xi(x) f_0^N(x,\xi),$$
$$c_N^*(x,\xi) = \widehat{\xi}(x) F_1^N(x,\xi) + \xi(x) F_0^N(x,\xi).$$

We consider the asymptotic behavior of systems $\xi_t^{G_N}$ with time scale τ_N. Let $c_N(x,\xi)$ be the rate function of the spin-flip system $\xi_t^{G_N}$ that is defined as

$$c_N(x,\xi) = \tau_N \cdot c_N^v(x,\xi) + c_N^*(x,\xi)$$

so that the generator $\mathcal{L}_N$ of $\xi_t^{G_N}$ is

$$\mathcal{L}_N f(\xi) = \sum_{x \in V_N} c_N(x,\xi)(f(\xi^x) - f(\xi))$$

for functions $f : \{0,1\}^{V_N} \to \mathbb{R}$.

Consider the density processes

$$Z_t^N = \frac{|\xi_t^{G_N}|}{\tau_N} = \frac{1}{\tau_N} \sum_{x \in V_N} \xi_t^{G_N}(x), \quad \log N \ll \tau_N \ll N.$$

The following "low initial density" assumption is in force:

$$Z_0^N = \frac{|\xi_0^{G_N}|}{\tau_N} \to c \in [0,\infty). \tag{1.2}$$

This condition implies that there exist a constant $C_{1.3} > 0$ such that

$$Z_0^N \le C_{1.3}, \quad \text{for all } N. \tag{1.3}$$

The fact (1.3) above will be applied frequently in later proofs.

The following perturbation assumptions will be assumed through out this work. For $i = 0, 1$, $N \in \mathbb{N}$, $1 \le l \le R$ and $1 \le k \le N_R$,

(P1) $\quad \theta^{N,i}(l,k) \ge 0$,

(P2) $\quad \theta^i(l,k) = \lim_{N\to\infty} \theta^{N,i}(l,k)$ exists.

In particular, (P2) guarantees boundedness of the perturbation functions $F_i(x,\xi)$.

1.2 Bound on transition probabilities

Denote $B_l(x) = \{y \in V_N : d(x,y) \le l\}$ as the distance-l neighborhood of x. To have a quantitative description of whether $B_l(x)$ is locally a finite tree , we introduce the *tree excess* defined in Section 2.2 of [LS10], denoted as $tx(B_l(x))$:

$tx(B_l(x)) =$ the maximum number of edges that can be deleted from the induced subgraph on $B_l(x)$ while keeping it connected.

Heuristically, the tree excess represents how many loops are there in the neighborhood. In particular, $tx(B_l(x)) = 1$ means that $B_l(x)$ contains one loop, and $tx(B_l(x)) = 0$ means that $B_l(x)$ does not contain any loops, and hence is a finite tree.

Denote $l_N = (1/5)\log_{r-1} N$ and define

$$\Gamma_N = \{x \in G_N : tx(B_{l_N}(x)) = 0\},$$
$$\Gamma'_N = \{x \in G_N : tx(B_{l_N}(x)) = 1\}.$$

Let α_0 be the constant in Theorem 6.3.2 of Durrett [D07] and let $\gamma = \alpha_0^2/2$. The transition function $p_t^N(x,y)$ is defined as the probability kernel of continuous time random walk with jump kernel q^N:

$$p_t^N(x,y) = e^{-t}\sum_{k=0}^{\infty}\frac{t^k}{k!}q_k^N(x,y) \quad x,y \in V_N \tag{1.4}$$

where q_k^N is the k-th iteration of q^N. This implies that the stationary distribution π^N of p_t^N is the uniform distribution on V_N. Define

$$\Delta_t^N = \max_{x,y\in V_N}\left|\frac{p_t^N(x,y)}{\pi^N(x)} - 1\right| = N\cdot\max_{x,y\in V_N}\left|p_t^N(x,y) - \frac{1}{N}\right|. \tag{1.5}$$

We call G_N a good graph if it has the following properties:

(i) G_N is connected,

(ii) $tx(B_{l_N}(x)) \le 1$ for all $x \in V_N$,

(iii) $|\{x : tx(B_{l_N}(x)) \neq 0\}| \le r^2 N^{3/5}$,

(iv) Δ_t^N satisfies that

$$\Delta_t^N \le Ne^{-\gamma t}, \quad \text{for all } t > 0. \tag{1.6}$$

In particular, (ii) and (iii) imply

$$V_N = \Gamma_N \cup \Gamma'_N \quad \text{and} \quad |\Gamma'_N| \le r^2 N^{3/5}, \tag{1.7}$$

and (iv) implies that for each $x, y \in V_N$ and $t \ge 0$,

$$|p_t^N(x,y) - 1/N| \le e^{-\gamma t}. \tag{1.8}$$

Denote $\mathcal{G}_N(r)$ as the collection of all simple r-regular graphs on V_N, and endow $\mathcal{G}_N(r)$ with uniform measure. Define the sequence of events

$$E_N = \{G \in \mathcal{G}_N(r) : G \text{ is good}\}.$$

In the following definition and the next proposition, P denotes the uniform measure on $\mathcal{G}_N(r)$. And we say that a sequence of events $A_N \subseteq \mathcal{G}_N(r)$ occurs *with high probability*, which we will abbreviate as *w.h.p.*, if

$$P(A_N) = \frac{|A_N|}{|\mathcal{G}_N(r)|} \to 1 \quad \text{as} \quad N \to \infty.$$

Proposition 1.1. *Let G be chosen uniformly at random from $\mathcal{G}_N(r)$. Then $G \in E_N$ w.h.p.*

See Section A.1 for the proof of this result.

1.3 Main result

We use the following martingale problem characterization of Feller's branching diffusion. An adapted a.s.-continuous non-negative real valued process Z_t on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ is said to be a Feller's branching diffusion with drift θ and branching rate β started at Z_0 if its law solves the martingale problem:

$(\mathrm{MP})^{\beta,\theta}_{Z_0}$ $\quad M_t = Z_t - Z_0 - \theta \int_0^t Z_s ds$ is a continuous $(\mathcal{F}_t)$-martingale with predictable square function $\langle M \rangle_t = \beta \int_0^t Z_s ds.$

Let $C_0[0,\infty)$ be the space of continuously differentiable functions vanishing at infinity. The generator $\mathcal{L}$ of the solution of the above martingale problem is given by

$$\mathcal{L}f(z) = (\theta z) f'(z) + (\beta z) \frac{1}{2} f''(z)$$

for $f \in C_0[0,\infty)$ such that $f', f'' \in C_0[0,\infty)$.

Let G^{tr} be the r-regular infinite tree with vertex set V^{tr}. Introduce a system of coalescing random walks $\{\widehat{B}^e_t, e \in V^{tr}\}$. Each $\widehat{B}^e_t$ has rate 1 with step distribution q^{tr} that is given by

$$q^{tr}(e, e') = 1/r \cdot 1_{\{e' \sim e\}}, \quad e, e' \in V^{tr}.$$

For $A \subseteq V^{tr}$ finite, define $\widehat{B}^A_t = \{\widehat{B}^e_t, e \in A\}$. For finite disjoint $A,\ B \subseteq V^{tr}$, define the stopping times

$$\sigma^{tr}(A) = \inf\{t > 0 : |\widehat{B}^A_t| = 1\},$$
$$\tau^{tr}(A, B) = \inf\{t > 0 : \widehat{B}^A_t \cap \widehat{B}^B_t \neq \varnothing\}.$$

Denote ρ as the root of the infinite tree. Introduce the "escape probability":

$$p^{tr,e} = \sum_e q^{tr}(\rho, e) P(\tau^{tr}(\rho, e) = \infty).$$

$p^{tr,e}$ is the probability that a walk starting at ρ never returns to ρ after leaving. A good reference for it is Spitzer [S64]. Next, define

$$\mathcal{N}_l(\rho) = \{e \in V^{tr} : d(\rho, e) = l\},$$
$$S_{l,k}(\rho) = \{A \subseteq \mathcal{N}_l(\rho) : |A| = k\},$$
$$\overline{\mathcal{N}_l(\rho)} = \mathcal{N}_l(\rho) \cup \{\rho\}.$$

and define the coalescing random walk probabilities

$$p^{tr,1}(l, k) = \sum_{\substack{A \in S_{l,k}(\rho) \\ B = \overline{\mathcal{N}_l(\rho)} \cap A^c}} P(\tau^{tr}(A, B) = \infty,\ \sigma^{tr}(A) < \infty),$$
$$p^{tr,0}(l, k) = \sum_{\substack{A \in S_{l,k}(\rho) \\ B = \overline{\mathcal{N}_l(\rho)} \cap A^c}} P(\tau^{tr}(A, B) = \infty,\ \sigma^{tr}(B) < \infty).$$

We will use the notation $a_N \ll b_N$ to mean that $a_N/b_N \to 0$ as $N \to \infty$. Let $P^{\beta,\theta}_{Z_0}$ be the law of the solution of $(\mathrm{MP})^{\beta,\theta}_{Z_0}$ on $C([0,\infty), \mathbb{R}^+)$. Denote P^N as the law of $Z^N_\cdot$. Our main result is

Theorem 1.2. *Assume (1.2), (P1)-(P2) and*

$$\log N \ll \tau_N \ll N. \tag{1.9}$$

For any sequence of r-regular graphs $\{G_N\}$ *such that* $G_N \in E_N$ *for each* N,

$$P^N \Rightarrow P^{\beta,\theta}_{Z_0}$$

where

$$\begin{aligned} \beta &= 2p^{tr,e}, \\ \theta &= \sum_{l=1}^{R}\sum_{k=1}^{N_R}(\theta^1(l,k)p^{tr,1}(l,k) - \theta^0(l,k)p^{tr,0}(l,k)). \end{aligned}$$

Here $\Rightarrow$ denotes weak convergence. The upper bound in (1.9) is the order of t^N_{meet} of good r-regular random graphs which comes from the result of Oliveira [O11] who showed that t^N_{meet} is equal to $O(N)$ for most graphs. This generalized the result in Cox [C89] who showed that the order of voter model consensus time on torus for $d \geq 3$ is equal to the size of the space. The lower bound comes from the main result of Lubetzky and Sly [LS10] who proved that t^N_{mix} for good r-regular random graphs has order $\log N$.

Implicitly, there is a double-layer randomness in the law P^N: one comes from the random graph G_N, and the other is from the voter model perturbation $\xi^N_\cdot$. Less formally, the theorem above says that the rescaled density processes converge to Feller's branching diffusion *with high probability*. Therefore, Proposition 1.1 is essential for getting the theorem above.

The key for obtaining the branching limit is low initial density. As the jump kernel q^{tr} defines transient walks, the low initial density assumption implies the following consequence: as particles spread out quick and become far apart, only those at the boundary can involve in the time evolution. The occurrence of a transition from 1 to 0 at a site depends asymptotically on its local density of 0. Therefore, at the large time scale τ_N, this proportion can be approximated by the "escape" probability. This makes transitions at a site alike a branching mechanism.

1.4 Applications of Theorem 1.2

We describe two examples of Theorem 1.2. In each case we will see that the perturbation conditions (P1)-(P2) hold.

Example 1.3. *(Biased voter models).* Suppose that $b_N \to b$ is a convergent non-negative sequence. Cox [C17] considered the biased voter model which has transitions

$$0 \to 1 \quad \text{at rate } \tau_N f_1^N + b_N f_1^N,$$
$$1 \to 0 \quad \text{at rate } \tau_N f_0^N.$$

The perturbation assumption (P1) is satisfied by $b_N \geq 0$, and convergence of b_N implies (P2). If $b_N = 0$ for all N, then $\xi_\cdot^{G_N}$ is the basic voter model and one has

$$P^N \Rightarrow P_{Z_0}^{2p^{tr,e},0}.$$

If $b_N \to b \geq 0$, then

$$P^N \Rightarrow P_{Z_0}^{2p^{tr,e},bp^{tr,e}}.$$

Example 1.4. *(q-Voter models with $q < 1$ close to 1).* Let $\xi_t^{G_N}(x)$ make transitions

$$i \to 1-i \quad \text{at rate} \quad \tau_N \cdot (f_{1-i}^N)^q, \quad i = 0, 1.$$

We consider $q = q_N < 1$ close to 1. Note that these rates imply that interaction range $R = 1$ so that l has only one value $l = 1$. Thus, we will use $n_i(x, \xi) = |\{y \sim x : \xi(y) = i\}|$ instead of the notation $n_l^i(x, \xi)$ in the definition of perturbation functions.

Following Section 1.1 of Agarwal, Simper and Durrett [ASD21], we view $\xi_t^{G_N}$ as a voter model perturbation as follows. Let $q_N = 1 - \delta_N$, $\delta_N \in (0, 1)$. For $i = 0, 1$, we can write

$$(f_i^N)^{q_N} = f_i^N + ((f_i^N)^{q_N} - f_i^N) = f_i^N + (f_i^N) \cdot ((f_i^N)^{-\delta_N} - 1).$$

This implies

$$\tau_N \cdot (f_{1-i}^N)^{q_N} = \tau_N f_i^N + (f_i^N) \cdot \frac{(f_i^N)^{-\delta_N} - 1}{1/\tau_N}. \tag{1.10}$$

Define

$$c_k^N = \frac{k}{r} \cdot \frac{(r/k)^{\delta_N} - 1}{1/\tau_N}, \quad k = 1, .., r, \quad i = 0, 1.$$

It is clear that $c_k^N \geq 0$ for all $N \in \mathbb{N}$ and $k = 1, .., r$ so that (P1) is satisfied. Now (1.10) can be written as

$$\tau_N f_i^N + \sum_{k=1}^{r} c_k^N \cdot 1_{\{n_i(x,\xi)=k\}}, \quad i = 0, 1.$$

Define

$$c_k = \begin{cases} \dfrac{k}{r} \cdot \log\left(\dfrac{r}{k}\right), & k = 1, .., r-1, \\ 0, & k = r. \end{cases}$$

Note that if $\delta_N = \tau_N^{-1}$, then for each $k = 1, .., r$, $c_k^N \to c_k$ as $N \to \infty$: let $u = k/r$,

$$\begin{aligned} \lim_N c_k^N &= u \cdot \lim_N \frac{u^{-\delta_N} - 1}{\delta_N} \\ &= u \cdot \left[-(u^\alpha)' \Big|_{\alpha=0} \right] \\ &= u \cdot [u^0 \cdot (-\log(u))] = c_k. \end{aligned}$$

This implies that c_k^N satisfies (P2).

Define $\mathcal{N}(\rho) = \{e \in V^{tr} : e \sim \rho\}$ as the nearest neighborhood of the root ρ on the infinite tree, and $\overline{\mathcal{N}(\rho)} = \mathcal{N}(\rho) \cup \{\rho\}$.

Corollary 1.5. *Let $\delta_N \sim \tau_N^{-1}$. Suppose that $\xi_t^{G_N}$ is the q-voter model defined in the example above and assume (1.2). If $\log N \ll \tau_N \ll N$, then $P^N \Rightarrow P_{Z_0}^{\beta,\theta}$ as $N \to \infty$ where*

$$\beta = 2p^{tr,e},$$

$$\theta = \sum_{k=1}^{r} c_k \cdot \sum_{\substack{A \subseteq \mathcal{N}(\rho),\ |A|=k \\ B=\overline{\mathcal{N}(\rho)} \cap A^c}} \Big(P(\tau^{tr}(A, B) = \infty,\ \sigma^{tr}(A) < \infty)$$

$$- P(\tau^{tr}(A,B) = \infty,\ \sigma^{tr}(B) < \infty)\Big).$$

1.5 Comparison with high density diffusive limit theorem

Chen, Choi and Cox [CCC16] proved that density processes of voter models with high initial density on large finite sets converge to Wright-Fisher diffusion, under mild mixing condition. Convergence to either Feller's branching diffusion or Wright-Fisher diffusion reflects that mean-field simplification occurs for the asymptotics. This is guaranteed by macroscopic independence between particles. We now discuss the effects of the choices for time scale τ_N and mass scale m_N according to different initial density. To get a clear picture of the story, we will assume that models are defined on good r-regular random graphs G_N which is a special case that satisfies the mixing conditions assumed in [CCC16] and [C17].

We will denote ξ_t^N as the rate τ_N voter model on V_N in this section: that is ξ_t^N has rate $c_N(x,\xi)$ defined in Section 1.1 with $c_N^*(x,\xi) = 0$. We will write the density process Z_t^N defined in that section as X_t^N below:

$$X_t^N = \frac{1}{m_N}\sum_{x\in V_N}\xi_t^N(x) = \frac{N}{m_N}\sum_{x\in V_N}\pi^N(x)\xi_t^N(x)$$

where $m_N \to \infty$ is the mass scale and π^N is the uniform distribution on V_N

$$\pi^N(x) = \frac{1}{N}, \qquad \text{for } x \in V_N,$$

which is the stationary distribution of p_t^N defined in (1.4). Define

$$\begin{aligned} p_{10}^N(\xi) &= \sum_x \pi^N(x)\xi(x)\Big(\sum_y q^N(x,y)\widehat{\xi}(y)\Big) \\ &= \sum_{x,y}\pi^N(x)q^N(x,y)\xi(x)\widehat{\xi}(y), \quad \xi \in \{0,1\}^{V_N}. \end{aligned}$$

Let us recall a few definitions. The Feller's branching diffusion Z_t with zero drift is a continuous martingale with quadratic variation

$$\langle Z \rangle_t = \beta \int_0^t Z_s \; ds,$$

and Z_t has generator

$$\mathcal{L}f(z) = (\beta z)\frac{1}{2}f''(z), \quad z \in [0, \infty).$$

The Wright-Fisher diffusion Y_t is a continuous martingale that has quadratic variation

$$\langle Y \rangle_t = \int_0^t Y_s(1 - Y_s) \; ds,$$

and Y_t has generator

$$\mathcal{G}f(z) = (z(1-z))\frac{1}{2}f''(z), \quad \text{for } f \in C^2([0,1]).$$

where $C^2([0,1])$ denotes second order continuously differentiable functions on $[0,1]$.

Using the decomposition which we will introduce in Chapter 2, one obtains the quadratic variation process for X_t^N as

$$\langle X^N \rangle_t = \frac{1}{\tau_N} \int_0^t 2p_{10}^N(\xi_s^N) \; ds.$$

Introduce

$$P(U = x, V' = y) = \pi^N(x)q^N(x,y),$$
$$P(U = x, U' = y) = \pi^N(x)\pi^N(y),$$

and let $M_{x,y}$ be the first meeting time of two independent rate 1 random walks starting from $x, y \in V_N$ respectively. Recall that t_{mix}^N is the mixing time that has order

$$t_{mix}^N \sim \log N \text{ as } N \to \infty,$$

and t_{meet}^N is the expected meeting time of two independent walks starting from stationary

$$t_{meet}^N = E(M_{U,U'}) \sim N \text{ as } N \to \infty.$$

The crucial fact is that if $\tau_N \sim t^N_{meet}$, one gets the following exponential decay of the tail probability

$$P(M_{U,V'} > \tau_N t) \approx \text{constant} \cdot e^{-t} \text{ for } N \text{ large} \tag{1.11}$$

while if $t^N_{mix} \ll \tau_N \ll t^N_{meet}$,

$$P(M_{U,V'} > \tau_N t) \approx \text{constant independent of } t \text{ for } N \text{ large.} \tag{1.12}$$

(1.11) is from Corollary 4.2 in [CCC16] and (1.12) is from Proposition 4.2 in [C17].

Under low initial density, τ_N and m_N satisfy

$$\begin{aligned} &t^N_{mix} \ll \tau_N = m_N \ll t^N_{meet}, \\ &X^N_0 = \frac{|\xi^N_0|}{m_N} \to x \in [0, \infty). \end{aligned} \tag{1.13}$$

Note that $N/m_N \to \infty$. In fact, if ξ^N_t is the voter model perturbation defined in (1.1) and, in addition, its drift term is positive, then X^N_t has exponential growth asymptotically and hence is an unbounded process.

By the duality equation for voter models,

$$E^{\xi^N_0}(p^N_{10}(\xi^N_{2t})) = E\Big[\xi^N_0(\widehat{B}^{U,N}_{2t\tau_N})\widehat{\xi}^N_0(\widehat{B}^{V',N}_{2t\tau_N}); M_{U,V'} > 2t\tau_N\Big].$$

The time and mass scale in (1.13) imply the following two kernel properties: first, for any $x, y \in V_N$,

$$m_N p^N_{t\tau_N}(x,y) \le \frac{2m_N}{N} \to 0$$

This implies that with large probability walks do not hit ξ^N_0 by time at the scale so that

$$P(\widehat{B}^{N,x}_{2t\tau_N} \in \xi^N_0, \widehat{B}^{N,y}_{2t\tau_N} \notin \xi^N_0) \approx P(\widehat{B}^{N,x}_{2t\tau_N} \in \xi^N_0, \tau^N(x,y) > t\tau_N).$$

We will prove the fact above in Chapter 4. This implies

$$E^{\xi^N_0}(p^N_{10}(\xi^N_{2t})) \approx E(\xi^N_0(\widehat{B}^{U,N}_{2t\tau_N}); M_{U,V'} > t\tau_N).$$

Secondly, as $\tau_N \gg t^N_{mix}$, then macroscopic independence is implied by the bound

$$|Np^N_{t\tau_N}(x,y) - 1| \leq Ne^{-\gamma(t\tau_N)} \to 0 \tag{1.14}$$

so that

$$E(\xi^N_0(\widehat{B}^{U,N}_{2t\tau_N}); M_{U,V'} > t\tau_N) \approx E(\xi^N_0(\widehat{B}^{U,N}_{2t\tau_N})) \cdot P(M_{U,V'} > t\tau_N).$$

Together with (1.12) leads to the mean-field simplification

$$\langle X^N \rangle_t \approx 2P(M_{U,V'} > \tau_N t) \int_0^t X^N_s \, ds.$$

Note that on G_N, the limit of the probabilities $P(M_{U,V'} > \tau_N t)$ is exactly the escape probability $p^{tr,e}$ defined in Section 1.3.

Under high initial density, the time and mass scale are

$$\begin{aligned} \tau_N &= m_N \sim t^N_{meet}, \\ X^N_0 &= \frac{|\xi^N_0|}{m_N} \to x \in [0,1]. \end{aligned} \tag{1.15}$$

Under (1.15),

$$X^N_\cdot \Rightarrow Y_\cdot \quad \text{as } N \to \infty$$

according to the main result of [CCC16]. Note that $X^N_\cdot$ is a bounded process.

For simplicity, let us assume $m_N = \tau_N = N$. For $u \in [0,1]$, let μ_u be the product measure on $\{0,1\}^{V_N}$,

$$\mu_u(\xi = 1 \text{ on } A) = u^{|A|}$$

for $A \subseteq V_N$ finite. The duality equation for voter models implies that for $x, y \in V_N$,

$$P^{\mu_u}(\xi^N_t(x) = 1, \xi^N_t(y) = 0) = u(1-u)P(M_{x,y} > t\tau_N)$$

so that

$$E^{\mu_u}(p^N_{10}(\xi^N_t)) = u(1-u)P(M_{U,V'} > t\tau_N).$$

Furthermore, the duality equation also implies

$$E^{\mu_u}(X_t^N(1-X_t^N)) = E^{\mu_u}\left(\frac{1}{N^2}\sum_{x,y}\xi_t^N(x)\widehat{\xi}_t^N(y)\right)$$
$$= u(1-u)P(M_{U,U'} > t\tau_N).$$

As $\tau_N = N \gg t_{mix}^N$, a bound similar to (1.14) can be obtained. This implies that particles must be separated apart by a macroscopic distance so that successive meetings are roughly independent. Also, note the fact that

$$P(M_{U,U'} > \tau_N t) \to e^{-t} \quad \text{ as } N \to \infty$$

which can be implied by [O11], and (1.11) says that the meeting time $M_{U,V'}$ is almost exponential. In symbol, this mean that

$$\langle X^N \rangle_t \approx \int_0^t X_s^N(1-X_s^N)\, ds.$$

For details of the last line, see Lemma 6.2 in [CCC16].

We make a final remark. The Wright-Fisher diffusion originally can be viewed as the limiting average of allele densities. Thus, the macroscopic independence plays the major role for obtaining this diffusive limit. For details regarding this point, see the proof of (3.1), (6.6) and the discussion in Section 1 of Cox [C89] on Kingman's coalescent lurking behind under the large scale.

1.6 Outline

The low density approach we use is structured as follows. In Section 2 we obtain a semimartingale decomposition. In Section 3, we give an argument of a comparison with the biased voter model and derive some bounds on Z_t^N. In Chapter 4, we provide some technical estimates of the drift term and branching rate using duality for voter models. Convergence of these estimates to coalescing random walk probabilities on the infinite tree will be proved in Chapter 5. Finally in Chapter 6 we establish tightness by verifying Aldou's criterion and identification of the limit is done by L^1-estimation.

From now on, $\{G_N\}$ is a sequence of good graphs. We will use the notation ξ_t^N to denote the processes $\xi_t^{G_N}$. Until further notice, $\sum_x$ will denote $\sum_{x \in V_N}$.

Chapter 2

Semimartingale decomposition

Define

$$\mathcal{L}_N^v f(\xi) = \tau_N \cdot \sum_x c_N^v(x,\xi)(f(\xi^x) - f(\xi)),$$
$$\mathcal{L}_N^* f(\xi) = \sum_x c_N^*(x,\xi)(f(\xi^x) - f(\xi)).$$

so that the generator $\mathcal{L}_N$ of ξ_t^N can be written as

$$\mathcal{L}_N f(\xi) = (\mathcal{L}_N^v + \mathcal{L}_N^*) f(\xi).$$

For $\xi \in \{0,1\}^{V_N}$, define

$$v_N(\xi) = \sum_x (\widehat{\xi}(x) f_1^N(x,\xi) + \xi(x) f_0^N(x,\xi)),$$
$$d^{N,1}(\xi) = \sum_x \widehat{\xi}(x) F_1^N(x,\xi),$$
$$d^{N,0}(\xi) = \sum_x \xi(x) F_0^N(x,\xi).$$

The next proposition provide the decomposition for $|\xi_t^N|$.

Proposition 2.1. *For all $t \geq 0$,*

$$|\xi_t^N| = |\xi_0^N| + D_t^N + M_t^N$$

where M_t^N is a $P^{\xi_0^N}$-martingale with quadratic variation

$$\langle M^N \rangle_t = \int_0^t [\tau_N v_N(\xi_s^N) + d^{N,1}(\xi_s^N) + d^{N,0}(\xi_s^N)] \; ds$$

and

$$D_t^N = \int_0^t [d^{N,1}(\xi_s^N) - d^{N,0}(\xi_s^N)] \; ds.$$

Denote

$$\begin{aligned} m_1^N(\xi) &= d^{N,1}(\xi) - d^{N,0}(\xi), \\ m_2^N(\xi) &= \tau_N v_N(\xi) + [d^{N,1}(\xi) + d^{N,0}(\xi)]. \end{aligned}$$

An immediate consequence of Proposition 2.1 is the decomposition of $Z_\cdot^N$:

Corollary 2.2. *Let $\mathcal{M}_t^N = \dfrac{1}{\tau_N} M_t^N$. Then $\mathcal{M}_t^N$ is a martingale with quadratic variation*

$$\langle \mathcal{M}^N \rangle_t = \frac{1}{\tau_N^2} \int_0^t m_2^N(\xi_s^N) \; ds.$$

Moreover, define

$$\mathcal{D}_t^N = \frac{1}{\tau_N} \int_0^t m_1^N(\xi_s^N) \; ds.$$

Then for all $t \geq 0$, Z_t^N has the semimartingale decomposition

$$Z_t^N = Z_0^N + \mathcal{M}_t^N + \mathcal{D}_t^N.$$

To prove Proposition 2.1 we will need several preliminary results. In the next lemma we give some basic facts which will be frequently used in the proofs later. Define functions

$$f(\xi) = |\xi|, \quad g(\xi) = |\xi|^2.$$

Lemma 2.3.

(a) $f(\xi^x) - f(\xi) = \widehat{\xi}(x) - \xi(x)$.

(b) $g(\xi^x) - g(\xi) = 1 + 2|\xi|(\widehat{\xi}(x) - \xi(x))$.

(c) $\displaystyle\sum_{x,y} q^N(x,y)\widehat{\xi}(x)\xi(y) = \sum_{x,y} q^N(x,y)\xi(x)\widehat{\xi}(y)$.

Proof. (a) is immediate by a simple calculation. For (b), we have

$$|\xi^x| = \Big(\sum_{y\neq x}\xi(y)\Big) + \widehat{\xi}(x) = |\xi| + (\widehat{\xi}(x) - \xi(x))$$

so that

$$\begin{aligned} g(\xi^x) - g(\xi) &= (|\xi^x| - |\xi|)(|\xi^x| + |\xi|) \\ &= (\widehat{\xi}(x) - \xi(x))(2|\xi| + (\widehat{\xi}(x) - \xi(x)) \\ &= 2|\xi|(\widehat{\xi}(x) - \xi(x)) + 1. \end{aligned}$$

And lastly, (c) is by symmetry of the kernel $q^N(x,y)$.

∎

Lemma 2.4.

(a) $\mathcal{L}_N^v f(\xi) = 0$.

(b) $\mathcal{L}_N^v g(\xi) = \tau_N v_N(\xi)$.

Proof. For (a), by Lemma 2.3 (a), we have

$$\begin{aligned} \mathcal{L}_N^v f(\xi) &= \sum_x \tau_N c_N^v(x,\xi)(f(\xi^x) - f(\xi)) \\ &= \tau_N \sum_{x,y} q^N(x,y)(\xi(x)\widehat{\xi}(y) + \widehat{\xi}(x)\xi(y)) \cdot (\widehat{\xi}(x) - \xi(x)) \\ &= \tau_N\Big(\sum_{x,y} q^N(x,y)(\xi(x)\widehat{\xi}(x)\widehat{\xi}(y) - \widehat{\xi}(x)\xi(x)\xi(y)) \qquad (2.1) \\ &\quad + \sum_{x,y} q^N(x,y)(\widehat{\xi}(x)\xi(y) - \xi(x)\widehat{\xi}(y))\Big). \end{aligned}$$

By the fact that $\xi(x)\widehat{\xi}(x) = 0$, (2.1) is equal to

$$\tau_N \sum_{x,y} q^N(x,y)(\widehat{\xi}(x)\xi(y) - \xi(x)\widehat{\xi}(y)) = 0$$

where the last equality is by Lemma 2.3 (c).

For (b), by Lemma 2.3 (a) and (b)

$$\mathcal{L}_N^v g(\xi) = \sum_x \tau_N c_N^v(x,\xi)(g(\xi^x) - g(\xi))$$

$$
\begin{aligned}
&= \sum_x \tau_N c_N^v(x,\xi)(1 + 2|\xi|(\widehat{\xi}(x) - \xi(x))) \\
&= \tau_N \Big[\sum_x c_N^v(x,\xi) + 2|\xi| \sum_x c_N^v(x,\xi)(\widehat{\xi}(x) - \xi(x)) \Big] \\
&= \tau_N \sum_x c_N^v(x,\xi) = \tau_N v_N(\xi) \qquad (2.2)
\end{aligned}
$$

where (2.2) is by part (a). ■

Lemma 2.5.

(a) $\mathcal{L}_N^* f(\xi) = d^{N,1}(\xi) - d^{N,0}(\xi)$.

(b) $\mathcal{L}_N^* g(\xi) = d^{N,1}(\xi) + d^{N,0}(\xi) + 2|\xi| \cdot (d^{N,1}(\xi) - d^{N,0}(\xi))$.

Proof. Direct calculations show that

$$
\begin{aligned}
\mathcal{L}_N^* f(\xi) &= \sum_x c_N^*(x,\xi)(f(\xi^x) - f(\xi)) \\
&= \sum_x (\widehat{\xi}(x) \cdot F_1^N(x,\xi) + \xi(x) \cdot F_0^N(x,\xi)) \cdot (\widehat{\xi}(x) - \xi(x)) \\
&= \sum_x \widehat{\xi}(x) F_1^N(x,\xi) - \sum_x \xi(x) F_0^N(x,\xi) \\
&= d^{N,1}(\xi) - d^{N,0}(\xi)
\end{aligned}
$$

and by Lemma 2.3 (b),

$$
\begin{aligned}
\mathcal{L}_N^* g(\xi) &= \sum_x c_N^*(x,\xi)(g(\xi^x) - g(\xi)) \\
&= \sum_x c_N^*(x,\xi)(1 + 2|\xi|(\widehat{\xi}(x) - \xi(x))) \\
&= (d^{N,1}(\xi) + d^{N,0}(\xi)) + 2|\xi|(d^{N,1}(\xi) - d^{N,0}(\xi)).
\end{aligned}
$$

■

Proof of Proposition 2.1. Recall that $f(\xi) = |\xi|$, and $g(\xi) = |\xi|^2$. Define

$$
\begin{aligned}
M_t^N &= f(\xi_t^N) - f(\xi_0^N) - \int_0^t \mathcal{L}_N f(\xi_s^N) ds, \\
Q_t^N &= g(\xi_t^N) - g(\xi_0^N) - \int_0^t \mathcal{L}_N g(\xi_s^N) ds.
\end{aligned}
$$

By Theorem I.5.2 in [L85], M_t^N and Q_t^N are martingales. Combining Lemma 2.4 and 2.5, M_t^N equals to

$$\begin{aligned} M_t^N &= |\xi_t^N| - |\xi_0^N| - \int_0^t [d^{N,1}(\xi_s^N) - d^{N,0}(\xi_s^N)]\ ds \\ &= |\xi_t^N| - |\xi_0^N| - D_t^N, \end{aligned} \tag{2.3}$$

and Q_t^N equals to

$$\begin{aligned} Q_t^N = |\xi_t^N|^2 - |\xi_0^N|^2 - \int_0^t &[\tau_N \upsilon_N(\xi_s^N) + (d^{N,1}(\xi_s^N) + d^{N,0}(\xi_s^N)) \\ &+ 2|\xi|(d^{N,1}(\xi_s^N) - d^{N,0}(\xi_s^N))]\ ds. \end{aligned}$$

Apply Exercise 2.9.29 in [EK86] to M_t^N and Q_t^N so that one gets the martingale

$$(M_t^N)^2 - \int_0^t [\tau_N \upsilon_N(\xi_s^N) + d^{N,1}(\xi_s^N) + d^{N,0}(\xi_s^N)]\ ds.$$

Now the integral part of the last line gives $\langle M^N \rangle_t$. ■

Chapter 3

Comparison with biased voter model

In this chapter, we show that ξ_t^N is close to the voter model over a short time period. This will be done by a comparison with the biased voter model $\xi_t^{N,b}$ which we define now.

For $\xi \in \{0,1\}^{V_N}$, recall the definition of local densities

$$f_i^N = f_i^N(x,\xi) = \sum_y q^N(x,y) 1_{\{\xi(y)=i\}}, \quad i = 0,1$$

Define the bias parameter

$$b = \sup_N \sum_{l=1}^{R} \sum_{k=1}^{N_R} (\theta^{N,0}(l,k) + \theta^{N,1}(l,k)).$$

For $x \in V_N$, define

$$\begin{aligned}
&\mathcal{N}(x) = \{y \in V_N : 1 \leq d(x,y) \leq R\}, \\
&\overline{\mathcal{N}(x)} = \mathcal{N}(x) \cup \{x\}, \\
&\mathcal{K} = 1 + R \cdot N_R.
\end{aligned}$$

A few facts from the definitions above are

$$\begin{aligned}
\mathcal{N}(x) &= \bigcup_{l=1}^{R} \mathcal{N}_l(x), \\
\mathcal{K} \geq |\overline{\mathcal{N}(x)}| &\geq \Big(\sum_{y \in \mathcal{N}(x)} \xi(y) \Big) \vee \Big(\sum_{y \in \mathcal{N}(x)} \widehat{\xi}(y) \Big).
\end{aligned}$$

Recall that $n_l^i(x,\xi)$ is the count in ξ of type i in $\mathcal{N}_l(x)$:

$$n_l^i(x,\xi) = \sum_{y\in\mathcal{N}_l(x)} 1\{\xi(y) = i\},\ \ i = 0,1 \tag{3.1}$$

and define the total count of type 1 in $\mathcal{N}(x)$

$$n_1(x,\xi) = \sum_{l=1}^{R} n_l^1(x,\xi).$$

The biased voter model $\xi_t^{N,b}$ has the following dynamics: at $x \in V_N$, $\xi_t^{N,b}(x)$ makes transitions

$$\begin{aligned} 0 \to 1 &\quad \text{at rate } \tau_N f_1^N + b\cdot n_1(x,\xi), \\ 1 \to 0 &\quad \text{at rate } \tau_N f_0^N \end{aligned}$$

so that the rate function for $\xi_t^{N,b}(x)$ is

$$c_N^b(x,\xi) = \tau_N c_N^v(x,\xi) + \widehat{\xi}(x)\cdot(b\cdot n_1(x,\xi)).$$

We first provide the bounds on biased voter models in the next proposition.

Proposition 3.1. *For $t \geq 0$ and $\xi_0^N \in \{0,1\}^{V_N}$,*

(a) $E^{\xi_0^N}|\xi_t^{N,b}| \leq |\xi_0^N|\cdot e^{(b\mathcal{K})t}$.

(b) $E^{\xi_0^N}(|\xi_t^{N,b}|^2) \leq \left[|\xi_0^N|^2 + |\xi_0^N|(2\tau_N + b\mathcal{K})\cdot te^{t(b\mathcal{K})}\right]\cdot e^{(2b\mathcal{K})t}$.

Proof. We follow the idea of Corollary 2.1 in [C17]. Notice that

$$\begin{aligned} \sum_x \widehat{\xi}(x) n_1(x,\xi) &= \sum_x \sum_{y\in\mathcal{N}(x)} \widehat{\xi}(x)\xi(y) \\ &= \sum_y \xi(y)\Big(\sum_{x\in\mathcal{N}(y)} \widehat{\xi}(x)\Big) \\ &\leq \mathcal{K}|\xi|. \end{aligned} \tag{3.2}$$

By (3.2) and applying Proposition 2.1 to $|\xi_t^{N,b}|$, we have

$$E^{\xi_0^N}|\xi_t^{N,b}| = |\xi_0^N| + \int_0^t b\cdot\sum_x E^{\xi_0^N}(\widehat{\xi}_s^{N,b}(x)\cdot n_1(x,\xi_s^{N,b}))\ ds$$

$$\leq |\xi_0^N| + (b\mathcal{K}) \int_0^t E^{\xi_0^N} |\xi_s^{N,b}| \; ds.$$

Thus, apply Grownwall's inequality to the function $f(s) = E^{\xi_0^N}(|\xi_s^{N,b}|)$ and the proof of (a) is complete.

For (b), using Theorem I.5.2 in [L85] and Lemma 2.3(b), we get

$$\begin{aligned} E^{\xi_0^N}(|\xi_t^{N,b}|^2) = |\xi_0^N|^2 + \int_0^t \tau_N E^{\xi_0^N}[v_N(\xi_s^{N,b})] \\ + b \cdot E^{\xi_0^N}\left[\sum_x \widehat{\xi}_s^{N,b}(x) n_1(x, \xi_s^{N,b}) \cdot \Big(1 + 2|\xi_s^{N,b}| \cdot (\widehat{\xi}_s^{N,b}(x) - \xi_s^{N,b}(x))\Big)\right] \; ds. \end{aligned}$$

By (3.2),

$$\begin{aligned} &\sum_x \widehat{\xi}(x) n_1(x,\xi) \cdot (1 + 2|\xi| \cdot (\widehat{\xi}(x) - \xi(x))) \\ &\qquad = (1 + 2|\xi|) \sum_x n_1(x,\xi)\widehat{\xi}(x) \leq (1+2|\xi|) \cdot \mathcal{K}|\xi|. \end{aligned}$$

This implies

$$E^{\xi_0^N}(|\xi_t^{N,b}|^2) \leq |\xi_0^N|^2 + \int_0^t 2\tau_N E^{\xi_0^N} |\xi_s^{N,b}| + E^{\xi_0^N}[(1 + 2|\xi_s^{N,b}|)(b\mathcal{K} \cdot |\xi_s^{N,b}|)] \; ds.$$

Rearrange the terms and apply part (a), we obtain

$$\begin{aligned} E^{\xi_0^N}(|\xi_t^{N,b}|^2) &\leq |\xi_0^N|^2 + (2\tau_N + b\mathcal{K}) \int_0^t E^{\xi_0^N} |\xi_s^{N,b}| \; ds + 2b\mathcal{K} \int_0^t E^{\xi_0^N}(|\xi_s^{N,b}|^2) \; ds \\ &\leq \left[|\xi_0^N|^2 + |\xi_0^N|(2\tau_N + b\mathcal{K}) \cdot t e^{t(b\mathcal{K})} \right] + 2b\mathcal{K} \int_0^t E^{\xi_0^N}(|\xi_s^{N,b}|^2) \; ds. \end{aligned}$$

The result follows by applying Gronwall's inequality to the function $g(s) = E^{\xi_0^N}(|\xi_s^{N,b}|^2)$. ∎

We now construct two couplings which make a comparison between particle systems. The existence of both will be implied by Theorem III.1.5 in [L85].

We say that $\xi \leq \eta$ if $\xi(x) \leq \eta(x)$ for every $x \in V_N$. The first coupling is between the voter model perturbation and biased voter model. To verify the assumptions of Theorem III.1.5, suppose that $\xi \leq \eta$. If $\eta(x) = \xi(x) = 0$, then

$$c_N(x,\xi) = \tau_N f_1^N(x,\xi) + F_1^N(x,\xi)$$

$$
\begin{aligned}
&= \tau_N f_1^N(x,\xi) + \sum_{l=1}^{R}\sum_{k=1}^{N_R} \theta^{N,i}(l,k) 1_{\{n_l^1(x,\xi)=k\}} \\
&\leq \tau_N f_1^N(x,\eta) + b \sum_{l=1}^{R}\sum_{k=1}^{N_R} 1_{\{n_l^1(x,\eta)=k\}} \\
&\leq \tau_N f_1^N(x,\eta) + b n_1(x,\eta) = c_N^b(x,\eta)
\end{aligned}
$$

and similarly, if $\eta(x) = \xi(x) = 1$,

$$
\begin{aligned}
c_N(x,\xi) &= \tau_N f_0^N(x,\xi) + F_0^N(x,\xi) \\
&\geq \tau_N f_0^N(x,\eta) = c_N^b(x,\eta).
\end{aligned}
$$

Thus by Theorem III.1.5 in [L85], given that $\xi_0^N = \xi_0^{N,b}$, there is a common probability space such that with probability 1,

$$
\xi_t^N \leq \xi_t^{N,b} \quad \text{for all } t \geq 0. \tag{3.3}
$$

The coupling between the voter model and the biased voter model is constructed in a similar way. Let $\xi_t^{N,v}$ be the voter model with generator $\mathcal{L}_N^v$ and suppose that $\xi \leq \eta$. If $\eta(x) = \xi(x) = 0$,

$$
\tau_N f_1^N(x,\eta) \leq \tau_N f_1^N(x,\eta) + b n_1(x,\eta) = c_N^b(x,\eta)
$$

and if $\eta(x) = \xi(x) = 1$,

$$
\tau_N f_0^N(x,\xi) \geq \tau_N f_0^N(x,\eta) = c_N^b(x,\eta).
$$

Thus, the assumptions of Theorem III.1.5 in [L85] are verified. Given that $\xi_0^{N,v} = \xi_0^N$, there is a common probability space such that with probability 1,

$$
\xi_t^{N,v} \leq \xi_t^{N,b} \quad \text{for all } t \geq 0. \tag{3.4}
$$

By Corollary III.1.7 in [L85], (3.3) and (3.4) imply

$$
E^{\xi_0^N}(h(\xi_t^N)) \leq E^{\xi_0^N}(h(\xi_t^{N,b})) \tag{3.5}
$$

$$
E^{\xi_0^N}(h(\xi_t^{N,v})) \leq E^{\xi_0^N}(h(\xi_t^{N,b})) \tag{3.6}
$$

for any function $h = h(\xi)$ such that $h(\xi) \le h(\eta)$ given that $\xi \le \eta$.

Corollary 3.2. *For any $\xi_0^N \in \{0,1\}^{V_N}$ and $t \ge 0$,*

(a) $E^{\xi_0^N}|\xi_t^N| \le |\xi_0^N| \cdot e^{(b\mathcal{K})t}$.

(b) $E^{\xi_0^N}[|\xi_t^N|^2] \le \left[|\xi_0^N|^2 + |\xi_0^N|(2\tau_N + b\mathcal{K}) \cdot te^{t(b\mathcal{K})}\right] \cdot e^{(2b\mathcal{K})t}$.

Proof. Part (a) is immediate by (3.5) by taking $h(\xi) = |\xi|$ so that

$$E^{\xi_0^N}|\xi_t^N| \le E^{\xi_0^N}|\xi_t^{N,b}|.$$

Similarly for part (b), take $h(\xi) = |\xi|^2$ and obtain

$$E^{\xi_0^N}(|\xi_t^N|^2) \le E^{\xi_0^N}(|\xi_t^{N,b}|^2).$$

■

Recall the definitions v_N and $d^{N,i}$

$$\begin{aligned} v_N(\xi) &= \sum_x (\widehat{\xi}(x) f_1^N(x,\xi) + \xi(x) f_0^N(x,\xi)) \\ d^{N,1}(\xi) &= \sum_x \widehat{\xi}(x) F_1^N(x,\xi) \\ d^{N,0}(\xi) &= \sum_x \xi(x) F_0^N(x,\xi). \end{aligned}$$

The next proposition will be used in Corollary 3.7 to show that the voter model perturbation is close to the voter model over a short time period.

Proposition 3.3. *There exists a constant $C_{3.3}$ such that for all $t \ge 0$,*

(a) $E^{\xi_0^N}|v_N(\xi_t^N) - v_N(\xi_t^{N,v})| \le 4(2 \vee (2bC_{3.3})) \cdot |\xi_0^N| \left[(e^{(b\mathcal{K})t} - 1) + te^{(b\mathcal{K})t}\right]$.

(b) $E^{\xi_0^N}|d^{N,i}(\xi_t^N) - d^{N,i}(\xi_t^{N,v})| \le C_{3.3}(2 \vee (2bC_{3.3})) \cdot |\xi_0^N| \left[(e^{(b\mathcal{K})t} - 1) + te^{(b\mathcal{K})t}\right]$, *for $i = 0, 1$.*

To prepare for the proof, we first give some preliminary bounds. For $x \in V_N$, $\xi \in \{0,1\}^{V_N}$, $A \subseteq V_N$, define

$$\chi(x,\xi,A) = \prod_{a \in A} \xi(a).$$

Lemma 3.4. *For $\xi, \eta \in \{0,1\}^{V_N}$, and $A, B \subseteq V_N$ disjoint,*

$$|\chi(x,\xi,A)\chi(x,\widehat{\xi},B) - \chi(x,\eta,A)\chi(x,\widehat{\eta},B)| \le \sum_{y \in A\cup B} |\xi(y) - \eta(y)|.$$

Proof. By twice using the fact that $|\prod z_i - \prod w_i| \le \sum |z_i - w_i|$ for z_i, w_i such that $|z_i|, |w_i| \le 1$ (Lemma 3.4.3 in [D19]), we have

$$\begin{aligned}
&|\chi(x,\xi,A)\chi(x,\widehat{\xi},B) - \chi(x,\eta,A)\chi(x,\widehat{\eta},B)| \\
&\quad \le |\chi(x,\xi,A) - \chi(x,\eta,A)| + |\chi(x,\widehat{\xi},B) - \chi(x,\widehat{\eta},B)| \\
&\quad \le \sum_{a\in A} |\xi(a) - \eta(a)| + \sum_{b \in B} |\widehat{\xi}(b) - \widehat{\eta}(b)| \\
&\quad = \sum_{y \in A\cup B} |\xi(y) - \eta(y)|.
\end{aligned}$$

■

For $i = 0, 1$, define $d^{N,i}_{l,k}$

$$d^{N,1}_{l,k}(\xi) = \sum_x \widehat{\xi}(x) 1_{\{n^1_l(x,\xi)=k\}},$$

$$d^{N,0}_{l,k}(\xi) = \sum_x \xi(x) 1_{\{n^0_l(x,\xi)=k\}},$$

then we can write $d^{N,i}$ as

$$\begin{aligned}
d^{N,1}(\xi) &= \sum_x \widehat{\xi}(x) F^N_1(x,\xi) \\
&= \sum_x \widehat{\xi}(x) \Big(\sum_{l=1}^{R} \sum_{k=1}^{N_R} \theta^{N,1}(l,k) 1_{\{n^1_l(x,\xi)=k\}} \Big) \\
&= \sum_{l=1}^{R} \sum_{k=1}^{N_R} \theta^{N,1}(l,k) \Big(\sum_x \widehat{\xi}(x) 1_{\{n^1_l(x,\xi)=k\}} \Big) \\
&= \sum_{l=1}^{R} \sum_{k=1}^{N_R} \theta^{N,1}(l,k) d^{N,1}_{l,k}(\xi), \\
d^{N,0}(\xi) &= \sum_x \xi(x) F^N_0(x,\xi) = \sum_{l=1}^{R} \sum_{k=1}^{N_R} \theta^{N,0}(l,k) d^{N,0}_{l,k}(\xi).
\end{aligned}$$

Lemma 3.5. *For $\xi \in \{0,1\}^{V_N}$,* *(a)* $v_N(\xi) \le 2|\xi|$.

(b) There is a constant $C_{3.5}$ *such that* $d^{N,i}_{l,k}(\xi) \le C_{3.5} \cdot |\xi|$, *for* $i = 0,1$ *and* $1 \le l \le R$, $1 \le k \le N_R$.

Proof. Part (a) is by

$$\begin{aligned} v_N(\xi) &= \sum_{x,y} q^N(x,y)(\widehat{\xi}(x)\xi(y) + \xi(x)\widehat{\xi}(y)) \\ &\le \sum_{x,y} q^N(x,y)(\xi(x) + \xi(y)) = 2|\xi|. \end{aligned}$$

For (b), for $i = 1$, by definition we have

$$\begin{aligned} d^{N,1}_{l,k}(\xi) &= \sum_x \widehat{\xi}(x) 1_{\{n^1_l(x,\xi)=k\}} \qquad (3.7) \\ &= \sum_x \sum_{\substack{A\in S_{l,k}(x) \\ B=\overline{\mathcal{N}_l(x)}\cap A^c}} \chi(x,\xi,A)\cdot \chi(x,\widehat{\xi},B) \\ &\le \sum_x \sum_{A\in S_{l,k}(x)} \chi(x,\xi,A) \\ &\le \sum_x \sum_{A\subseteq \mathcal{N}(x)} \sum_{a\in\mathcal{N}(x)} \xi(a) \le \mathcal{K}2^{\mathcal{K}}\cdot|\xi| \qquad (3.8) \end{aligned}$$

so that we could choose $C_{3.5} = \mathcal{K}2^{\mathcal{K}}$.

Similarly for $i = 0$,

$$\begin{aligned} d^{N,0}_{l,k}(\xi) &= \sum_x \xi(x) 1_{\{n^0_l(x,\xi)=k\}} \\ &= \sum_x \sum_{\substack{A\in S_{l,k}(x) \\ B=\overline{\mathcal{N}_l(x)}\cap A^c}} \chi(x,\widehat{\xi},A)\cdot \chi(x,\xi,B) \\ &\le \sum_x \sum_{B\subseteq \overline{\mathcal{N}(x)}} \chi(x,\xi,B) \\ &\le \sum_x \sum_{B\subseteq \overline{\mathcal{N}(x)}} \sum_{b\in B} \xi(b) \le \mathcal{K}2^{\mathcal{K}}|\xi|. \end{aligned}$$

∎

Lemma 3.6. *For any* $\eta,\ \xi \in \{0,1\}^{V_N}$ *with* $\xi \le \eta$, *(a)* $|v_N(\xi) - v_N(\eta)| \le 4(|\eta| - |\xi|)$.

(b) $|d^{N,i}_{l,k}(\xi) - d^{N,i}_{l,k}(\eta)| \le C_{3.5}(|\eta| - |\xi|)$, $i = 0, 1$.

Proof. For (a), direct calculation with applying Lemma 3.4 gives

$$\begin{aligned}
&|v_N(\xi) - v_N(\eta)| \\
&\quad \le \sum_{x,y} q^N(x,y)\Big(|\widehat{\xi}(x)\xi(y) - \widehat{\eta}(x)\eta(y)| + |\xi(x)\widehat{\xi}(y) - \eta(x)\widehat{\eta}(y)|\Big) \\
&\quad \le \sum_{x,y} q^N(x,y)\cdot 2\Big(\sum_{a\in\{x,y\}} |\xi(a) - \eta(a)|\Big) \\
&\quad \le \sum_{x,y} q^N(x,y)\cdot 2\Big(|\xi(x) - \eta(x)| + |\xi(y) - \eta(y)|\Big) \\
&\quad = 2\Big(\sum_{x,y} q^N(x,y)(\eta(y) - \xi(y)) + \sum_{x,y} q^N(x,y)(\eta(x) - \xi(x))\Big) \\
&\quad = 4(|\eta| - |\xi|).
\end{aligned}$$

For (b), by Lemma 3.4 we have

$$\begin{aligned}
&|d^{N,1}_{l,k}(\xi) - d^{N,1}_{l,k}(\eta)| \\
&\quad \le \sum_x \sum_{\substack{A\in S_{l,k}(x)\\ B=\overline{\mathcal{N}_l(x)}\cap A^c}} |\chi(x,\xi,A)\chi(x,\widehat{\xi},B) - \chi(x,\eta,A)\chi(x,\widehat{\eta},B)| \\
&\quad \le \sum_x \sum_{\substack{A\in S_{l,k}(x)\\ B=\overline{\mathcal{N}_l(x)}\cap A^c}} \sum_{y\in A\cup B} |\xi(y) - \eta(y)| \\
&\quad \le \sum_x \sum_{y\in\overline{\mathcal{N}(x)}} |\xi(y) - \eta(y)| \cdot \sum_{\substack{A\in S_{l,k}(x)\\ B=\overline{\mathcal{N}_l(x)}\cap A^c}} 1.
\end{aligned}$$

Since $\xi \le \eta$, the last line is bounded by

$$2^{\mathcal{K}} \cdot \sum_x \sum_{y\in\overline{\mathcal{N}(x)}} (\eta(y) - \xi(y)) = 2^{\mathcal{K}}\mathcal{K} \cdot (|\eta| - |\xi|).$$

The calculation for $i = 0$ is similar. ∎

Now we prove Proposition 3.3.

Proof of Proposition 3.3. We first bound the difference of the total masses. By Lemma 2.4

(a) and Lemma A.1, $|\xi_t^{N,v}|$ is a martingale and $|\xi_t^{N,b}|$ is a submartingale. This implies

$$0 \le E^{\xi_0^N}|\xi_t^{N,b}| - E^{\xi_0^N}|\xi_t^N| \le \left(E^{\xi_0^N}|\xi_t^{N,b}| - |\xi_0^N|\right) + \left|E^{\xi_0^N}|\xi_t^N| - |\xi_0^N|\right|, \tag{3.9}$$

$$0 \le E^{\xi_0^N}|\xi_t^{N,b}| - E^{\xi_0^N}|\xi_t^{N,v}| = E^{\xi_0^N}|\xi_t^{N,b}| - |\xi_0^N|. \tag{3.10}$$

Using Proposition 3.1(a),

$$E^{\xi_0^N}|\xi_t^{N,b}| - |\xi_0^N| \le |\xi_0^N|(e^{(b\mathcal{K})t} - 1). \tag{3.11}$$

And by Proposition 2.1,

$$\begin{aligned}\left|E^{\xi_0^N}|\xi_t^N| - |\xi_0^N|\right| &= \left|E^{\xi_0^N}(M_t^N) + E^{\xi_0^N}(D_t^N)\right| \\ &= \left|E^{\xi_0^N}(D_t^N)\right| \le \int_0^t E^{\xi_0^N}|m_1^N(\xi_s^N)|\, ds\end{aligned}$$

To bound $E^{\xi_0^N}|m_1^N(\xi_s^N)|$, we have

$$\begin{aligned}|m_1^N(\xi)| &= \left|\sum_{l=1}^R \sum_{k=1}^{N_R} (\theta^{N,1}(l,k) d_{l,k}^{N,1}(\xi) - \theta^{N,0}(l,k) d_{l,k}^{N,0}(\xi))\right| \\ &\le b \cdot \sum_{l=1}^R \sum_{k=1}^{N_R} (|d_{l,k}^{N,1}(\xi)| + |d_{l,k}^{N,0}(\xi)|).\end{aligned} \tag{3.12}$$

Recall that $\mathcal{K} = 1 + R \cdot N_R$. By Lemma 3.5 (b), (3.12) is bounded by

$$b \cdot 2C_{3.5}|\xi| \cdot \sum_{l=1}^R \sum_{k=1}^{N_R} 1 \le 2b\mathcal{K} \cdot C_{3.5}|\xi|.$$

Define $C_{3.3} = \mathcal{K} \cdot C_{3.5}$. Therefore, by Corollary 3.2(a),

$$\begin{aligned}\left|E^{\xi_0^N}|\xi_t^N| - |\xi_0^N|\right| &\le (2bC_{3.3}) \cdot \int_0^t E^{\xi_0^N}|\xi_s^N|\, ds \\ &\le (2bC_{3.3}) \cdot |\xi_0^N| t e^{(b\mathcal{K})t}.\end{aligned} \tag{3.13}$$

We now prove part (a). By Lemma 3.6(a), the coupling (3.3) and (3.4), (3.9) and (3.10),

$$\begin{aligned}&E^{\xi_0^N}|v_N(\xi_t^N) - v_N(\xi_t^{N,v})| \\ &\le E^{\xi_0^N}|v_N(\xi_t^N) - v_N(\xi_t^{N,b})| + E^{\xi_0^N}|v_N(\xi_t^{N,v}) - v_N(\xi_t^{N,b})| \\ &\le 4\left(E^{\xi_0^N}\left||\xi_t^N| - |\xi_t^{N,b}|\right| + E^{\xi_0^N}\left||\xi_t^{N,v}| - |\xi_t^{N,b}|\right|\right)\end{aligned}$$

$$= 4\Big(E^{\xi_0^N}(|\xi_t^{N,b}| - |\xi_t^N|) + E^{\xi_0^N}(|\xi_t^{N,b}| - |\xi_t^{N,v}|) \Big)$$
$$\le 4\Big[2\Big(E^{\xi_0^N}(|\xi_t^{N,b}| - |\xi_0^N|) \Big) + \Big| E^{\xi_0^N}(|\xi_t^N| - |\xi_0^N|) \Big| \Big].$$

By (3.11) and (3.13), the last line is bounded by

$$4\Big[2|\xi_0^N|(e^{(b\mathcal{K})t} - 1) + (2bC_{3.3})|\xi_0^N| t e^{(b\mathcal{K})t} \Big] \le 4(2 \vee (2bC_{3.3}))|\xi_0^N| \Big[(e^{(b\mathcal{K})t} - 1) + t e^{(b\mathcal{K})t} \Big].$$

For (b), we have that for $i = 0, 1$,

$$|d^{N,i}(\xi_t^N) - d^{N,i}(\xi_t^{N,v})| \le |d^{N,i}(\xi_t^N) - d^{N,i}(\xi_t^{N,b})| + |d^{N,i}(\xi_t^{N,v}) - d^{N,i}(\xi_t^{N,b})|$$
$$\le \sum_{l=1}^{R} \sum_{k=1}^{N_R} \Big(|d_{l,k}^{N,i}(\xi_t^N) - d_{l,k}^{N,i}(\xi_t^{N,b})| + |d_{l,k}^{N,i}(\xi_t^{N,b}) - d_{l,k}^{N,i}(\xi_t^{N,v})| \Big).$$

By Lemma 3.6, the coupling result (3.9),

$$E^{\xi_0^N} |d_{l,k}^{N,i}(\xi_t^N) - d_{l,k}^{N,i}(\xi_t^{N,b})| \le C_{3.5}\Big(E^{\xi_0^N}|\xi_t^{N,b}| - E^{\xi_0^N}|\xi_t^N| \Big)$$
$$\le C_{3.5}\Big[\Big(E^{\xi_0^N}|\xi_t^{N,b}| - |\xi_0^N| \Big) + \Big| E^{\xi_0^N}|\xi_t^N| - |\xi_0^N| \Big| \Big] \tag{3.14}$$

and similarly by (3.4) and (3.10),

$$E^{\xi_0^N} |d_{l,k}^{N,i}(\xi_t^{N,b}) - d_{l,k}^{N,i}(\xi_t^{N,v})| \le C_{3.5}\Big(E^{\xi_0^N}|\xi_t^{N,b}| - E^{\xi_0^N}|\xi_t^{N,v}| \Big)$$
$$\le C_{3.5}\Big[E^{\xi_0^N}|\xi_t^{N,b}| - |\xi_0^N| \Big]. \tag{3.15}$$

Combining (3.14) and (3.15), applying (3.11) and (3.13) we have

$$E^{\xi_0^N}|d^{N,i}(\xi_t^N) - d^{N,i}(\xi_t^{N,v})|$$
$$\le \mathcal{K} \cdot C_{3.5} \cdot \Big(2\Big[E^{\xi_0^N}|\xi_t^{N,b}| - |\xi_0^N| \Big] + \Big| E^{\xi_0^N}|\xi_t^N| - |\xi_0^N| \Big| \Big)$$
$$\le C_{3.3} \cdot \Big(2\Big[E^{\xi_0^N}|\xi_t^{N,b}| - |\xi_0^N| \Big] + \Big| E^{\xi_0^N}|\xi_t^N| - |\xi_0^N| \Big| \Big)$$
$$\le C_{3.3}(2 \vee (2bC_{3.3})) \cdot |\xi_0^N| \Big[(e^{(b\mathcal{K})t} - 1) + t e^{(b\mathcal{K})t} \Big].$$

■

Recall that

$$m_1^N(\xi) = d^{N,1}(\xi) - d^{N,0}(\xi),$$
$$m_2^N(\xi) = \tau_N v_N(\xi) + [d^{N,1}(\xi) + d^{N,0}(\xi)].$$

Corollary 3.7. *There is a constant $C_{3.7}$ such that for all $t \geq 0$,*

(a) $\dfrac{1}{\tau_N} E^{\xi_0^N} |m_1^N(\xi_t^N) - m_1^N(\xi_t^{N,v})| \leq C_{3.7} \cdot Z_0^N \Big((e^{(b\mathcal{K})t} - 1) + te^{(b\mathcal{K})t} \Big).$

(b) $\dfrac{1}{\tau_N^2} E^{\xi_0^N} |m_2^N(\xi_t^N) - m_2^N(\xi_t^{N,v})| \leq C_{3.7} \Big(1 + \dfrac{1}{\tau_N} \Big) \cdot Z_0^N \Big((e^{(b\mathcal{K})t} - 1) + te^{(b\mathcal{K})t} \Big).$

Proof. For (a), apply Proposition 3.3 (b) we have

$$\begin{aligned}
&\frac{1}{\tau_N} E^{\xi_0^N} |m_1^N(\xi_t^N) - m_1^N(\xi_t^{N,v})| \\
&\quad \leq \frac{1}{\tau_N} \Big(E^{\xi_0^N} |d^{N,1}(\xi_t^N) - d^{N,1}(\xi_t^{N,v})| + E^{\xi_0^N} |d^{N,0}(\xi_t^N) - d^{N,0}(\xi_t^{N,v})| \Big) \\
&\quad \leq 2C_{3.3}(2 \vee 2bC_{3.3}) \cdot Z_0^N \Big[(e^{(b\mathcal{K})t} - 1) + te^{(b\mathcal{K})t} \Big]. \qquad (3.16)
\end{aligned}$$

And for (b), let N be large. Apply both (a) and (b) of Proposition 3.3 and obtain

$$\begin{aligned}
&\frac{1}{\tau_N^2} E^{\xi_0^N} |m_2^N(\xi_t^N) - m_2^N(\xi_t^{N,v})| \\
&\quad \leq \frac{1}{\tau_N^2} \Big(\tau_N \cdot E^{\xi_0^N} |v_N(\xi_t^N) - v_N(\xi_t^{N,v})| + E^{\xi_0^N} |d^{N,1}(\xi_t^N) - d^{N,1}(\xi_t^{N,v})| \\
&\qquad + E^{\xi_0^N} |d^{N,0}(\xi_t^N) - d^{N,0}(\xi_t^{N,v})| \Big) \\
&\quad \leq \frac{1}{\tau_N} \Big[\tau_N \cdot 4(2 \vee (2bC_{3.3})) + 2C_{3.3}(2 \vee (2bC_{3.3})) \Big] \cdot \frac{|\xi_0^N|}{\tau_N} \Big[(e^{(b\mathcal{K})t} - 1) + te^{(b\mathcal{K})t} \Big] \\
&\quad \leq \Big[4 + \frac{2C_{3.3}}{\tau_N} \Big] \cdot (2 \vee (2bC_{3.3})) \cdot Z_0^N \Big[(e^{(b\mathcal{K})t} - 1) + te^{(b\mathcal{K})t} \Big] \qquad (3.17) \\
&\quad \leq (4 + 2C_{3.3}) \cdot (2 \vee (2bC_{3.3})) \cdot \Big[1 + \frac{1}{\tau_N} \Big] \cdot Z_0^N \Big[(e^{(b\mathcal{K})t} - 1) + te^{(b\mathcal{K})t} \Big]. \qquad (3.18)
\end{aligned}$$

The proof for the corollary is completed by taking

$$C_{3.7} = \Big[(2C_{3.3}) \vee (4 + 2C_{3.3})) \Big] \cdot (2 \vee 2bC_{3.3}).$$

■

Chapter 4

Estimation by the voter model

Given the results in the previous chapter, the main work in this chapter will be in proving Proposition 4.1 to obtain the estimate of the drift term and branching rate of the voter perturbation. We shall see that this estimation is almost trivial by the comparison bound in Chapter 3 and Proposition 4.2 which gives estimates for the voter model.

The key of this estimation is that we choose a correct time scale t_N for averaging the macroscopic density of 1's. In particular, this time scale should provide sufficient mixing condition so that uniformity of local densities can be established, meanwhile it keeps the motion of particles at time points of the scale relatively (comparing to τ_N) local so that if two particles coalesce, then they must meet early. Recall from (1.9) that

$$\log N \ll \tau_N \ll N.$$

We choose t_N such that

$$t_N = \delta_N \tau_N \quad \text{where} \quad \delta_N \to 0 \quad \text{and} \quad t_N \gg \log N.$$

This definition gives the following consequences:

$$\frac{t_N}{\tau_N} \to 0 \quad \text{as } N \to \infty, \tag{4.1}$$

$$Ne^{-\gamma t_N} \leq \frac{1}{N} \quad \text{for large } N, \tag{4.2}$$

and by (1.8)

$$\max_{x,y\in V_N}\left|p^N_{t_N}(x,y)-\frac{1}{N}\right|\le\frac{1}{N}\quad\text{for large } N. \tag{4.3}$$

We will see later in the proofs that several error terms along the estimation are defined according to the consequences above.

Recall that $\xi^{N,v}_t$ is the voter model with rates c^v_N. Define

$$\begin{aligned}
V^{N,v}_t &= \frac{1}{\tau_N^2}E^{\xi^N_0}(\tau_N\cdot v_N(\xi^{N,v}_t))\\
&= \frac{1}{\tau_N}\sum_x E^{\xi^N_0}\left[\widehat{\xi}^{N,v}_t(x)f^N_1(x,\xi^{N,v}_t)+\xi^{N,v}_t(x)f^N_0(x,\xi^{N,v}_t)\right]\\
&= \frac{1}{\tau_N}\sum_x\sum_y q^N(x,y)\cdot E^{\xi^N_0}\left[\widehat{\xi}^{N,v}_t(x)\xi^{N,v}_t(y)+\xi^{N,v}_t(x)\widehat{\xi}^{N,v}_t(y)\right]\\
&= \frac{2}{\tau_N}\sum_x\sum_y q^N(x,y)\cdot E^{\xi^N_0}\left[\widehat{\xi}^{N,v}_t(x)\xi^{N,v}_t(y)\right].
\end{aligned}$$

Recall the definitions

$$\begin{aligned}
\mathcal{N}_l(x) &= \{y\in V_N : d(x,y)=l\},\\
\overline{\mathcal{N}_l(x)} &= \mathcal{N}_l(x)\cup\{x\},\\
S_{l,k}(x) &= \{A\subseteq\mathcal{N}_l(x) : |A|=k\},\\
\chi(x,\xi,A) &= \prod_{a\in A}\xi(a),
\end{aligned}$$

and define

$$\begin{aligned}
V^{N,1}_t(l,k) &= \frac{1}{\tau_N}E^{\xi^N_0}(d^{N,1}_{l,k}(\xi^{N,v}_t))\\
&= \frac{1}{\tau_N}\sum_x E^{\xi^N_0}\left(\widehat{\xi}^{N,v}_t(x)\cdot 1\{n^1_l(x,\xi^{N,v}_t)=k\}\right)\\
&= \frac{1}{\tau_N}\sum_x\sum_{\substack{A\in S_{l,k}(x)\\ B=\overline{\mathcal{N}_l(x)}\cap A^c}} E^{\xi^N_0}(\chi(x,\xi^{N,v}_t,A)\cdot\chi(x,\widehat{\xi}^{N,v}_t,B)),
\end{aligned}$$

$$\begin{aligned}
V^{N,0}_t(l,k) &= \frac{1}{\tau_N}E^{\xi^N_0}(d^{N,0}_{l,k}(\xi^{N,v}_t))\\
&= \frac{1}{\tau_N}\sum_x E^{\xi^N_0}\left(\xi^{N,v}_t(x)\cdot 1\{n^0_l(x,\xi^{N,v}_t)=k\}\right)
\end{aligned}$$

$$= \frac{1}{\tau_N} \sum_x \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} E^{\xi_0^N}(\chi(x, \widehat{\xi}_t^{N,v}, A) \cdot \chi(x, \xi_t^{N,v}, B)).$$

Let $\{\widehat{B}_t^{N,x} : x \in V_N\}$ be a family of rate 1 coalescing random walks on G_N with step distribution q^N such that $\widehat{B}_0^{N,x} = x$. For $A \subseteq V_N$, define $\widehat{B}_t^{N,A} = \{\widehat{B}_t^{N,x} : x \in A\}$. The duality equation for voter models is

$$P^{\xi_0^N}(\xi_t^{N,v} = 1 \text{ on } A) = P(\widehat{B}_{t\tau_N}^{N,A} \in \xi_0^N) \quad \text{for } A \subseteq V_N. \tag{4.4}$$

See [HL75] for more details about the duality equation above. Note that any $A \subseteq V_N$ is finite since V_N is finite. By (4.4), $V_{2\delta_N}^{N,v}$, $V_{2\delta_N}^{N,1}(l,k)$ and $V_{2\delta_N}^{N,0}(l,k)$ can be written as

$$V_{2\delta_N}^{N,v} = \frac{2}{\tau_N} \sum_{x,y} q^N(x,y) P(\widehat{B}_{2t_N}^{N,x} \notin \xi_0^N,\ \widehat{B}_{2t_N}^{N,y} \in \xi_0^N),$$

$$V_{2\delta_N}^{N,1}(l,k) = \frac{1}{\tau_N} \sum_x \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} P(\widehat{B}_{2t_N}^{N,A} \subseteq \xi_0^N,\ \widehat{B}_{2t_N}^{N,B} \subseteq \widehat{\xi}_0^N),$$

$$V_{2\delta_N}^{N,0}(l,k) = \frac{1}{\tau_N} \sum_x \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} P(\widehat{B}_{2t_N}^{N,A} \subseteq \widehat{\xi}_0^N,\ \widehat{B}_{2t_N}^{N,B} \subseteq \xi_0^N).$$

Define the first meeting times

$$\tau^N(x,y) = \inf\{t \geq 0 : \widehat{B}_t^{N,x} = \widehat{B}_t^{N,y}\},$$

$$\tau^N(A,B) = \inf\{t \geq 0 : \widehat{B}_t^{N,A} \cap \widehat{B}_t^{N,B} \neq \varnothing\},$$

and let $\sigma^N(A)$ denote the time at which walks starting from sites in A have coalesced into a single particle,

$$\sigma^N(A) = \inf\{t \geq 0 : |\widehat{B}_t^{N,A}| = 1\}.$$

Define the coalescing walk probabilities on G_N

$$p^{N,e} = \frac{1}{N} \sum_{x,y} q^N(x,y) P(\tau^N(x,y) > t_N),$$

$$p^{N,1}(l,k) = \frac{1}{N} \sum_x \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} P(\tau^N(A,B) > t_N,\ \sigma^N(A) \leq t_N),$$

$$p^{N,0}(l,k) = \frac{1}{N}\sum_x \sum_{\substack{A\in S_{l,k}(x)\\ B=\overline{\mathcal{N}_l(x)}\cap A^c}} P(\tau^N(A,B) > t_N,\ \sigma^N(B) \le t_N).$$

The main result of this chapter is

Proposition 4.1. *There is a sequence $\varepsilon^N_{4.1} \to 0$ as $N \to \infty$ such that*

(a) $\left|\frac{1}{\tau_N} E^{\xi^N_0}[m^N_1(\xi^N_{2\delta_N})] - \theta^N Z^N_0\right| \le \varepsilon^N_{4.1}(1+Z^N_0)^2,$

(b) $\left|\frac{1}{\tau_N^2} E^{\xi^N_0}[m^N_2(\xi^N_{2\delta_N})] - \beta^N Z^N_0\right| \le \varepsilon^N_{4.1}(1+Z^N_0)^2$

where

$$\theta^N = \sum_{l=1}^{R}\sum_{k=1}^{N_R}(\theta^{N,1}(l,k)p^{N,1}(l,k) - \theta^{N,0}(l,k)p^{N,0}(l,k)), \tag{4.5}$$

$$\beta^N = 2p^{N,e}. \tag{4.6}$$

We need to first prove the follow two propositions for the voter model.

Proposition 4.2. *There is a sequence $\varepsilon^N_{4.2} \to 0$ as $N \to \infty$ such that*

(a) $\left|\frac{1}{\tau_N} E^{\xi^N_0}[m^N_1(\xi^{N,v}_{2\delta_N})] - \theta^N Z^N_0\right| \le \varepsilon^N_{4.2}(1+(Z^N_0)^2),$

(b) $\left|\frac{1}{\tau_N^2} E^{\xi^N_0}[m^N_2(\xi^{N,v}_{2\delta_N})] - \beta^N Z^N_0\right| \le \varepsilon^N_{4.2}(1+Z^N_0+(Z^N_0)^2).$

where θ^N and β^N are as in (4.5) and (4.6).

Proposition 4.3. *There are sequences $\varepsilon^{N,v},\ \varepsilon^{N,i} \to 0$ such that*

(a) $\left|V^{N,v}_{2\delta_N} - 2p^{N,e}\cdot Z^N_0\right| \le \varepsilon^{N,v}\cdot\left[1+(Z^N_0)^2\right].$

(b) $\left|V^{N,i}_{2\delta_N}(l,k) - p^{N,i}(l,k)\cdot Z^N_0\right| \le \varepsilon^{N,i}\cdot\left[1+(Z^N_0)^2\right]$*, for $i = 0, 1,\ 1 \le l \le R$ and $1 \le k \le N_R$.*

We first give some preliminary results. Let $\{B^{N,x}_t : x \in V_N\}$ be a family of rate 1 independent walks with step distribution q^N such that $B^{N,x}_0 = x$ so that $B^{N,x}_t$ has transition function p^N_t as defined in (1.4). Lemma 4.4 gives bounds on probabilities of independent walks, and Lemma 4.5 gives a bound on meeting time probabilities.

Lemma 4.4. *For any $x, y \in V_N$, $x \ne y$,*

(a) $P(B^{N,x}_t \in \xi^N_0) \le \frac{2|\xi^N_0|}{N}$ *, for $t \ge \frac{1}{\gamma}\log N$.*

(b) $\left| P(B_t^{N,x} \in \xi_0^N) - \frac{|\xi_0^N|}{N} \right| \le |\xi_0^N| \cdot e^{-\gamma t}$*, for all* $t \ge 0$.

Proof. For (a), by (1.8) we have

$$p_t^N(x,y) \le \frac{1}{N} + e^{-\gamma t}, \quad \text{for all } t > 0.$$

Thus if $t \ge \frac{1}{\gamma} \log N$,

$$p_t^N(x,y) \le \frac{2}{N} \quad \text{for all } x, y \in V_N.$$

This implies

$$P(B_t^{N,x} \in \xi_0^N) = \sum_w p_t^N(x,w)\xi_0^N(w) \le |\xi_0^N| \cdot \frac{2}{N}. \tag{4.7}$$

For (b),

$$\begin{aligned} \left| P(B_t^{N,x} \in \xi_0^N) - \frac{|\xi_0^N|}{N} \right| &= \left| \sum_w \xi_0^N(w)(p_t^N(x,w) - 1/N) \right| \\ &\le \sum_w \xi_0^N(w)|p_t^N(x,w) - 1/N| \\ &\le |\xi_0^N| \cdot e^{-\gamma t}. \end{aligned}$$

where the last inequality is by (1.8).

Lemma 4.5. $P(\tau^N(x,y) \in (t_N, 2t_N]) \le \frac{2e^2(1+t_N)}{N}$.

Proof. For $t \in (t_N, 2t_N]$ and large N, by (4.3) we have

$$P(B_t^{N,x} = B_t^{N,y}) = \sum_z p_t^N(x,z)p_t^N(y,z) \le \frac{2}{N} \sum_z p_t^N(y,z) = \frac{2}{N}.$$

Then by Lemma A.2,

$$\begin{aligned} P(\tau^N(x,y) \in (t_N, 2t_N]) &\le e^2 \int_{t_N}^{2t_N+1} P(B_t^{N,x} = B_t^{N,y})\, dt \\ &\le \frac{2e^2(t_N+1)}{N}. \end{aligned}$$

■

We now prove Proposition 4.3(a).

Proof of Proposition 4.3 (a). We can write the probability $P(\widehat{B}^{N,x}_{2t_N} \notin \xi^N_0,\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0)$ as a sum of three terms:

$$P(\widehat{B}^{N,x}_{2t_N} \notin \xi^N_0,\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0) \tag{4.8}$$

$$= \left[P(\widehat{B}^{N,x}_{2t_N} \notin \xi^N_0,\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0) - P(\tau^N(x,y) > 2t_N,\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0)\right] \tag{4.9}$$

$$+ \left[P(\tau^N(x,y) > 2t_N,\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0) - P(\tau^N(x,y) > t_N,\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0)\right] \tag{4.10}$$

$$+ P(\tau^N(x,y) > t_N,\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0). \tag{4.11}$$

To bound (4.9), we have

$$\begin{aligned}
&\left|P(\widehat{B}^{N,x}_{2t_N} \notin \xi^N_0,\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0) - P(\tau^N(x,y) > 2t_N,\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0)\right| \\
&\quad = P(\tau^N(x,y) > 2t_N,\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0) - P(\widehat{B}^{N,x}_{2t_N} \notin \xi^N_0,\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0) \\
&\quad = P(\widehat{B}^{N,x}_{2t_N} \in \xi^N_0,\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0,\ \tau^N(x,y) > 2t_N) \\
&\quad \leq P(B^{N,x}_{2t_N} \in \xi^N_0,\ B^{N,y}_{2t_N} \in \xi^N_0) && (4.12) \\
&\quad \leq \frac{4|\xi^N_0|^2}{N^2}. && (4.13)
\end{aligned}$$

Note that (4.12) is implied by the fact that coalescing walks are independent until their first meeting. And (4.13) is by Lemma 4.4(a).

To bound (4.10), notice that

$$\begin{aligned}
&\left|P(\tau^N(x,y) > 2t_N,\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0) - P(\tau^N(x,y) > t_N,\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0)\right| \\
&\quad = P(\tau^N(x,y) \in (t_N, 2t_N],\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0) \\
&\quad \leq P(\tau^N(x,y) \in (t_N, 2t_N]) \\
&\quad \leq \frac{2e^2(1+t_N)}{N}. && (4.14)
\end{aligned}$$

and (4.14) is by Lemma 4.5.

Lastly, to estimate (4.11), use the Markov property and Lemma 4.4(b) to obtain

$$\left|P(\tau^N(x,y) > t_N,\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0) - \frac{|\xi^N_0|}{N} P(\tau^N(x,y) > t_N)\right|$$

$$
\begin{aligned}
&= \Bigg| \sum_{w \in V_N} \Big(P(\tau^N(x,y) > t_N,\ \widehat{B}^{N,y}_{t_N} = w) \cdot P(\widehat{B}^{w}_{t_N} \in \xi^N_0) \\
&\quad - P(\tau^N(x,y) > t_N,\ \widehat{B}^{N,y}_{t_N} = w) \cdot \frac{|\xi^N_0|}{N} \Big) \Bigg| \\
&\le \sum_{w \in V_N} P(\tau^N(x,y) > t_N,\ \widehat{B}^{N,y}_{t_N} = w) \cdot \Big| P(\widehat{B}^{w}_{t_N} \in \xi^N_0) - \frac{|\xi^N_0|}{N} \Big| \\
&\le 1 \cdot |\xi^N_0| e^{-\gamma t_N}. \qquad (4.15)
\end{aligned}
$$

Therefore, since we can write $p^{N,e} \cdot Z^N_0$ as

$$
p^{N,e} \cdot Z^N_0 = \frac{1}{\tau_N} \sum_{x,y} q^N(x,y) \Big[P(\tau^N(x,y) > t_N) \cdot \frac{|\xi^N_0|}{N} \Big],
$$

then (4.13), (4.14) and (4.15) together imply

$$
\begin{aligned}
&\Big| V^{N,v}_{2\delta_N} - 2p^{N,e} \cdot Z^N_0 \Big| \\
&\quad \le \frac{2}{\tau_N} \sum_{x,y} q^N(x,y) \Big| P(\widehat{B}^{N,x}_{2t_N} \notin \xi^N_0,\ \widehat{B}^{N,y}_{2t_N} \in \xi^N_0) - \frac{|\xi^N_0|}{N} P(\tau^N(x,y) > t_N) \Big| \\
&\quad \le \frac{2N}{\tau_N} \Big(\frac{4|\xi^N_0|^2}{N^2} + \frac{2e^2(1+t_N)}{N} + |\xi^N_0| e^{-\gamma t_N} \Big) \\
&\quad \le 2 \Big(\frac{4\tau_N}{N} (Z^N_0)^2 + \frac{2e^2(1+t_N)}{\tau_N} + C_{1.3} \cdot N e^{-\gamma t_N} \Big) \\
&\quad \le 2 \Big(\frac{4\tau_N}{N} (Z^N_0)^2 + \frac{2e^2(1+t_N)}{\tau_N} + C_{1.3} \cdot N^{-1} \Big)
\end{aligned}
$$

where the last line is by (4.2). Take

$$
\varepsilon^{N,v} = 2 \max \Big(\frac{4\tau_N}{N}, \frac{2e^2(1+t_N)}{\tau_N} + C_{1.3} \cdot N^{-1} \Big)
$$

and $\varepsilon^{N,v} \to 0$ by (4.1) and (4.2). ■

For Proposition 4.3(b), we will prove case $i = 1$ and the proof of $i = 0$ is exactly similar. This will be done in three steps, as summarized in Lemma 4.6 - 4.9. Recall that

$$
\begin{aligned}
&\mathcal{N}(x) = \{y : 1 \le d(x,y) \le R\}, \quad \overline{\mathcal{N}(x)} = \mathcal{N}(x) \cup \{x\} \\
&\overline{\mathcal{N}_l(x)} = \mathcal{N}_l(x) \cup \{x\}, \\
&S_{l,k}(x) = \{A \subseteq \mathcal{N}_l(x) : |A| = k\},
\end{aligned}
$$

$$\mathcal{K} = 1 + R \cdot N_R.$$

Let $C_R = \mathcal{K}^2 \cdot 2^{\mathcal{K}}$. The following simple fact will be used frequently in later proofs:

$$\sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} |A||B| \le \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} \mathcal{K}^2 \le C_R. \tag{4.16}$$

Fix $1 \le l \le R$ and $1 \le k \le N_R$ and let $A, B \in S_{l,k}(x)$. In the figure of each of the following three steps, the time goes up.

Step 1. Define events

$$E_{11}(A,B) = \{\widehat{B}^{N,A}_{2t_N} \subseteq \xi^N_0,\ \widehat{B}^{N,B}_{2t_N} \cap \xi^N_0 \neq \varnothing,\ \tau^N(A,B) > 2t_N\},$$
$$E_{12}(A,B) = \{\widehat{B}^{N,A}_{2t_N} \subseteq \xi^N_0,\ \tau^N(A,B) > 2t_N\}.$$

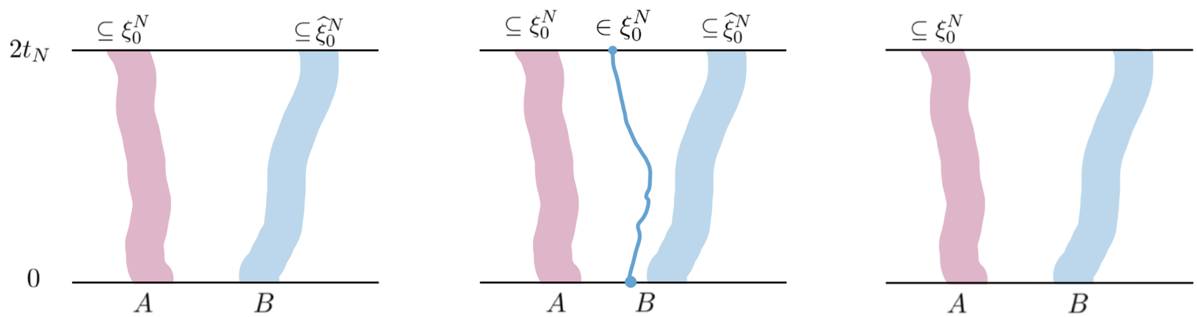

Figure 4.1: Event in the definition of $V^{N,1}_{2\delta_N}$, $E_{11}(A,B)$ and $E_{12}(A,B)$ from left to right.

The relation between $E_{11}(A,B)$ and $E_{12}(A,B)$ is

$$\{\widehat{B}^{N,A}_{2t_N} \subseteq \xi^N_0,\ \widehat{B}^{N,B}_{2t_N} \subseteq \widehat{\xi}^N_0\} \cup E_{11}(A,B) = E_{12}(A,B).$$

Let

$$\mathcal{E}_{1j} = \frac{1}{\tau_N} \sum_x \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} P(E_{1j}(A,B)), \quad j = 1, 2.$$

Thus, $V^{N,1}_{2\delta_N}(l,k) = \mathcal{E}_{12} - \mathcal{E}_{11}$. Lemma 4.6 below shows that $\mathcal{E}_{11}$ is negligible.

Lemma 4.6. $\mathcal{E}_{11} \le 4C_R \dfrac{\tau_N}{N} \cdot (Z^N_0)^2$.

Proof. By Lemma 4.4(a) and using independence between walks,

$$\begin{aligned}
P(E_{11}(A,B)) &= P(\widehat{B}^{N,A}_{2t_N} \subseteq \xi^N_0,\ \widehat{B}^{N,B}_{2t_N} \cap \xi^N_0 \neq \varnothing,\ \tau^N(A,B) > 2t_N) \\
&\le \sum_{a\in A, b\in B} P(\widehat{B}^{N,A}_{2t_N} \in \xi^N_0,\ \widehat{B}^{N,B}_{2t_N} \in \xi^N_0\ ,\ \tau^N(A,B) > 2t_N) \\
&\le \sum_{a\in A, b\in B} P(B^a_{2t_N} \in \xi^N_0,\ B^b_{2t_N} \in \xi^N_0) \\
&\le |A||B|\frac{4|\xi^N_0|^2}{N^2}. \qquad (4.17)
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathcal{E}_{11} &\le \frac{1}{\tau_N}\cdot\frac{4|\xi^N_0|^2}{N^2}\sum_x \sum_{\substack{A\in S_{l,k}(x)\\ B=\overline{\mathcal{N}_l(x)}\cap A^c}} |A||B| \\
&\le \frac{1}{\tau_N}\cdot\frac{4|\xi^N_0|^2}{N^2}\cdot NC_R = 4C_R\frac{\tau_N}{N}\cdot(Z^N_0)^2 \qquad (4.18)
\end{aligned}$$

where the last inequality is by (4.16). ■

Step 2. Define

$$\begin{aligned}
E_{21}(A,B) &= \{\widehat{B}^{N,A}_{2t_N} \subseteq \xi^N_0,\ \tau^N(A,B) \in (t_N, 2t_N]\}, \\
E_{22}(A,B) &= \{\widehat{B}^{N,A}_{2t_N} \subseteq \xi^N_0,\ \tau^N(A,B) > t_N\}.
\end{aligned}$$

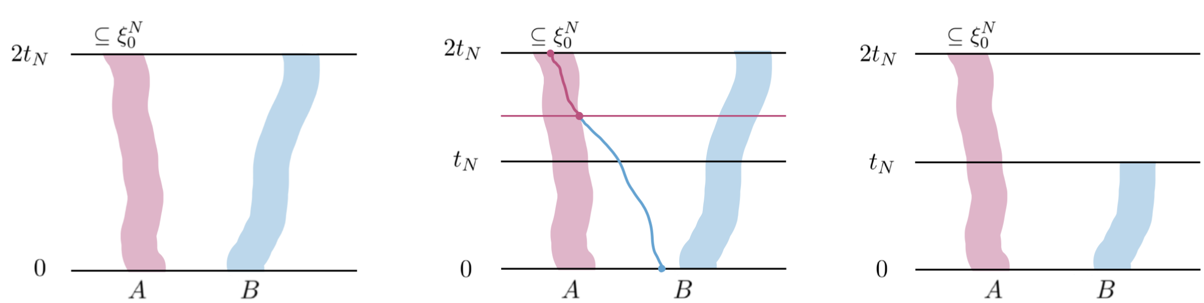

Figure 4.2: $E_{12}(A,B)$, $E_{21}(A,B)$, $E_{22}(A,B)$ from left to right.

By the definition above, $E_{12}(A,B) \cup E_{21}(A,B) = E_{22}(A,B)$. Let

$$\mathcal{E}_{2j} = \frac{1}{\tau_N}\sum_x \sum_{\substack{A\in S_{l,k}(x)\\ B=\overline{\mathcal{N}_l(x)}\cap A^c}} P(E_{2j}(A,B)), \quad j=1,2.$$

so that $\mathcal{E}_{12} = \mathcal{E}_{22} - \mathcal{E}_{21}$. In Lemma 4.7 we show that $\mathcal{E}_{21}$ is negligible.

Lemma 4.7. $\mathcal{E}_{21} \le C_R \dfrac{2e^2(1+t_N)}{\tau_N}$.

Proof. By Lemma 4.5,

$$\begin{aligned} P(E_{21}(A,B)) &= P(\widehat{B}^{N,A}_{2t_N} \subseteq \xi^N_0,\ \tau^N(A,B) \in (t_N, 2t_N]) \\ &\le \sum_{a\in A, b\in B} P(\tau^N(a,b) \in (t_N, 2t_N]) \\ &\le |A||B| \frac{2e^2(1+t_N)}{N}. \end{aligned}$$

Therefore, by (4.16)

$$\begin{aligned} \mathcal{E}_{21} &\le \frac{1}{\tau_N} \cdot \frac{2e^2(1+t_N)}{N} \sum_x \sum_{\substack{A\in S_{l,k}(x) \\ B=\overline{\mathcal{N}_l(x)}\cap A^c}} |A||B| \\ &\le C_R \frac{2e^2(1+t_N)}{\tau_N}. \end{aligned}$$

■

Step 3. Define

$$\begin{aligned} E_{31}(A,B) &= \{\widehat{B}^{N,A}_{2t_N} \subseteq \xi^N_0,\ \tau^N(A,B) > t_N,\ \sigma^N(A) > 2t_N\}, \\ E_{32}(A,B) &= \{\widehat{B}^{N,A}_{2t_N} \subseteq \xi^N_0,\ \tau^N(A,B) > t_N,\ ,\ \sigma^N(A) \in (t_N, 2t_N]\}, \\ E_{33}(A,B) &= \{\widehat{B}^{N,A}_{2t_N} \in \xi^N_0,\ \tau^N(A,B) > t_N,\ \sigma^N(A) \le t_N\}. \end{aligned}$$

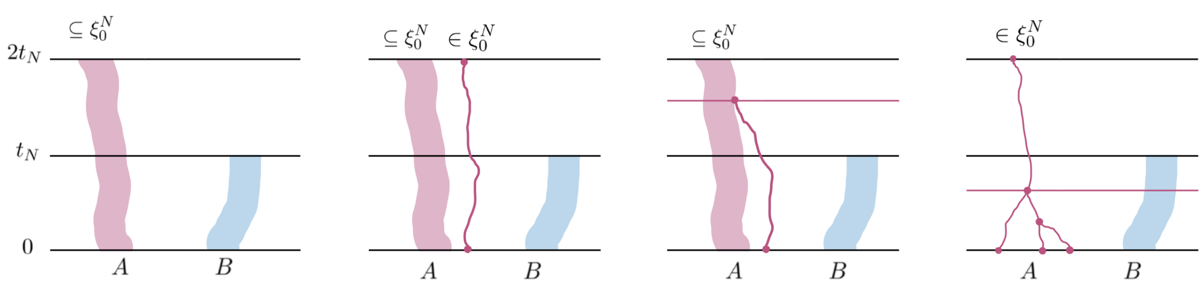

Figure 4.3: Illustration of $E_{22}(A,B)$, $E_{31}(A,B)$, $E_{32}(A,B)$ and $E_{33}(A,B)$.

The relation among the events above is

$$E_{22}(A,B) = \bigcup_{j=1}^{3} E_{3j}(A,B).$$

Let

$$\mathcal{E}_{3j} = \frac{1}{\tau_N} \sum_x \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} P(E_{3j}(A,B)), \quad j = 1,2,3.$$

so that $\mathcal{E}_{22} = \mathcal{E}_{31} + \mathcal{E}_{32} + \mathcal{E}_{33}$. We will show that $\mathcal{E}_{31}$ and $\mathcal{E}_{32}$ are negligible in Lemma 4.8, and that $\mathcal{E}_{33}$ is close to $p^{N,1}(l,k) \cdot Z_0^N$ in Lemma 4.9.

Lemma 4.8.

(a) $\mathcal{E}_{31} \le 4C_R \dfrac{\tau_N}{N} \cdot (Z_0^N)^2$.

(b) $\mathcal{E}_{32} \le C_R \dfrac{2e^2(1+t_N)}{\tau_N}$.

Proof. For (a), apply Lemma 4.4(a) and we have

$$\begin{aligned} P(E_{31}(A,B)) &= P(\widehat{B}^{N,A}_{2t_N} \subseteq \xi_0^N,\ \tau^N(A,B) > t_N,\ \sigma^N(A) > 2t_N) \\ &\le P(\widehat{B}^{N,A}_{2t_N} \subseteq \xi_0^N,\ \sigma^N(A) > 2t_N) \\ &\le \sum_{a,a' \in A} P(\widehat{B}^{N,A}_{2t_N} \in \xi_0^N,\ \widehat{B}^{a'}_{2t_N} \in \xi_0^N\ ,\ \tau^N(a,a') > 2t_N) \\ &\le |A|^2 \frac{4|\xi_0^N|^2}{N^2} \end{aligned} \tag{4.19}$$

where (4.19) follows similarly as (4.17). Therefore,

$$\mathcal{E}_{31} \le \frac{1}{\tau_N} \cdot \frac{4|\xi_0^N|^2}{N^2} \sum_x \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} |A|^2 \le 4C_R \frac{\tau_N}{N} \cdot (Z_0^N)^2$$

where the last inequality is from (4.16).

For (b), apply Lemma 4.5 and we have

$$\begin{aligned} P(E_{32}(A,B)) &= P(\widehat{B}^{N,A}_{2t_N} \subseteq \xi_0^N,\ \tau^N(A,B) > t_N,\ \sigma^N(A) \in (t_N, 2t_N]) \\ &\le P(\sigma^N(A) \in (t_N, 2t_N]) \end{aligned}$$

$$\leq \sum_{a,a' \in A} P(\tau^N(a,a') \in (t_N, 2t_N])$$
$$\leq |A|^2 \frac{2e^2(1+t_N)}{N}.$$

Thus, by (4.16),

$$\mathcal{E}_{32} \leq \frac{1}{\tau_N} \cdot \frac{2e^2(1+t_N)}{N} \sum_x \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} |A|^2 \leq C_R \frac{2e^2(1+t_N)}{\tau_N}.$$

∎

Recall that $p^{N,1}(l,k)$ is defined as

$$p^{N,1}(l,k) = \frac{1}{N} \sum_x \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} P(\tau^N(A,B) > t_N,\ \sigma^N(A) \leq t_N),$$

and the definition of $\mathcal{E}_{33}$ is

$$\mathcal{E}_{33} = \frac{1}{\tau_N} \sum_x \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} P(E_{33}(A,B))$$

where

$$E_{33}(A,B) = \{\widehat{B}^{N,A}_{2t_N} \in \xi^N_0,\ \tau^N(A,B) > t_N,\ \sigma^N(A) \leq t_N\}.$$

Lemma 4.9. *There is a constant $C_{4.9}$ such that $|\mathcal{E}_{33} - p^{N,1}(l,k) \cdot Z^N_0| \leq C_{4.9} \cdot N^{-1}$.*

Proof. We will write $\widehat{B}^{N,A}_t = \{a\}$ as $\widehat{B}^{N,A}_t = a$. We first decompose the probability $P(E_{33}(A,B))$ using the Markov property at t_N:

$$\begin{aligned} &P(E_{33}(A,B)) \\ &= P(\widehat{B}^{N,A}_{2t_N} \in \xi^N_0,\ \tau^N(A,B) > t_N,\ \sigma^N(A) \leq t_N) \\ &= \sum_{\substack{a' \in V_N \\ B' \subseteq V_N, a' \notin B'}} P(\widehat{B}^{N,A}_{t_N} = a',\ \widehat{B}^{N,B}_{t_N} = B',\ \tau^N(A,B) > t_N,\ \sigma^N(A) \leq t_N) \qquad (4.20) \\ &\qquad \cdot P(\widehat{B}^{a'}_{t_N} \in \xi^N_0). \end{aligned}$$

Write $P(\widehat{B}^{a'}_{t_N} \in \xi^N_0)$ in (4.20) as

$$P(\widehat{B}^{a'}_{t_N} \in \xi^N_0) = \frac{|\xi^N_0|}{N} + \Big(P(\widehat{B}^{a'}_{t_N} \in \xi^N_0) - \frac{|\xi^N_0|}{N}\Big).$$

Hence $P(E_{33}(A,B))$ is equal to the sum of $\Sigma_1(A,B)$ and $\Sigma_2(A,B)$, defined as

$$\begin{aligned}
\Sigma_1(A,B) &= \sum_{\substack{a' \in V_N \\ B' \subseteq V_N, a' \notin B'}} P(\widehat{B}^{N,A}_{t_N} = a',\ \widehat{B}^{N,B}_{t_N} = B',\ \tau^N(A,B) > t_N,\ \sigma^N(A) \le t_N) \cdot \frac{|\xi^N_0|}{N} \\
&= P(\tau^N(A,B) > t_N,\ \sigma^N(A) \le t_N) \cdot \frac{|\xi^N_0|}{N}, \\
\Sigma_2(A,B) &= \sum_{\substack{a' \in V_N \\ B' \subseteq V_N, a' \notin B'}} P(\widehat{B}^{N,A}_{t_N} = a',\ \widehat{B}^{N,B}_{t_N} = B',\ \tau^N(A,B) > t_N,\ \sigma^N(A) \le t_N) \\
&\qquad \cdot \Big(P(\widehat{B}^{a'}_{t_N} \in \xi^N_0) - \frac{|\xi^N_0|}{N}\Big).
\end{aligned}$$

Therefore, $\mathcal{E}_{33}$ is the sum of Σ_1 and Σ_2 which are defined as

$$\begin{aligned}
\Sigma_1 &= \frac{1}{\tau_N} \sum_x \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} \Sigma_1(A,B) \\
&= \frac{1}{\tau_N} \cdot \frac{|\xi^N_0|}{N} \sum_x \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} P(\tau^N(A,B) > t_N,\ \sigma^N(A) \le t_N) \\
&= p^{N,1}(l,k) \cdot Z^N_0, \\
\Sigma_2 &= \frac{1}{\tau_N} \sum_x \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} \Sigma_2(A,B).
\end{aligned}$$

We sill see that Σ_2 is negligible. Applying Lemma 4.4(b), $\Sigma_2(A,B)$ is bounded by

$$\begin{aligned}
&|\Sigma_2(A,B)| \\
&\quad \le \sum_{\substack{a' \in V_N \\ B' \subseteq V_N, a' \notin B'}} P(\widehat{B}^{N,A}_{t_N} = a',\ \widehat{B}^{N,B}_{t_N} = B',\ \tau^N(A,B) > t_N,\ \sigma^N(A) \le t_N) \\
&\qquad \cdot \left| P(\widehat{B}^{a'}_{t_N} \in \xi^N_0) - \frac{|\xi^N_0|}{N} \right| \\
&\quad \le \sum_{\substack{a' \in V_N \\ B' \subseteq V_N, a' \notin B'}} P(\widehat{B}^{N,A}_{t_N} = a',\ \widehat{B}^{N,B}_{t_N} = B',\ \tau^N(A,B) > t_N,\ \sigma^N(A) \le t_N) \cdot \Big(|\xi^N_0| e^{-\gamma t_N}\Big)
\end{aligned}$$

$$\leq 1 \cdot |\xi_0^N| e^{-\gamma t_N}.$$

Consequently, by (4.2) and (1.3),

$$|\Sigma_2| \leq \frac{1}{\tau_N} \sum_x \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} |\Sigma_2(A, B)| \leq C_R \cdot Z_0^N \cdot N e^{-\gamma t_N} \leq C_R \cdot C_{1.3} \cdot N^{-1}.$$

Therefore,

$$|\mathcal{E}_{33} - p^{N,1}(l,k) \cdot Z_0^N| = |\Sigma_2| \leq C_R \cdot C_{1.3} \cdot N^{-1}.$$

Thus, we can choose $C_{4.9} = C_{1.3} \cdot C_R$. ∎

Proof of Proposition 4.3(b). Define

$$\varepsilon^{N,1} = \max\left(8C_R \frac{\tau_N}{N},\ 2C_R \frac{2e^2(1+t_N)}{\tau_N} + C_{4.9} \cdot N^{-1}\right).$$

We have $\varepsilon^{N,1} \to 0$ since $t_N \gg \log N$. By combining the bounds in Lemma 4.6, 4.7, 4.8 and 4.9, we have

$$\begin{aligned}
&|V_{2\delta_N}^{N,1}(l,k) - p^{N,1}(l,k) \cdot Z_0^N| \\
&\quad \leq |V_{2\delta_N}^{N,1}(l,k) - \mathcal{E}_{12}| + |\mathcal{E}_{12} - \mathcal{E}_{22}| + |\mathcal{E}_{22} - \mathcal{E}_{33}| + |\mathcal{E}_{33} - p^{N,1}(l,k) \cdot Z_0^N| \\
&\quad = \mathcal{E}_{11} + \mathcal{E}_{21} + (\mathcal{E}_{31} + \mathcal{E}_{32}) + |\mathcal{E}_{33} - p^{N,1}(l,k) \cdot Z_0^N| \\
&\quad \leq 8C_R \frac{\tau_N}{N} \cdot (Z_0^N)^2 + 2C_R \frac{2e^2(1+t_N)}{\tau_N} + C_{4.9} \cdot N^{-1} \\
&\quad \leq \varepsilon^{N,1}(1 + (Z_0^N)^2).
\end{aligned}$$

The proof for showing that

$$|V_{2\delta_N}^{N,0}(l,k) - p^{N,0}(l,k) \cdot Z_0^N| \leq \varepsilon^{N,0}(1 + (Z_0^N)^2)$$

is exactly similar, and one can choose $\varepsilon^{N,0}$ to be equal to $\varepsilon^{N,1}$. ∎

Proof of Proposition 4.2. We have

$$\frac{1}{\tau_N} E^{\xi_0^N}[m_1^N(\xi_{2\delta_N}^{N,v})] = \sum_{l=1}^{R} \sum_{k=1}^{N_R} (\theta^{N,1}(l,k) \cdot V_{2\delta_N}^{N,1}(l,k) - \theta^{N,0}(l,k) \cdot V_{2\delta_N}^{N,0}(l,k)).$$

By Proposition 4.3,

$$\begin{aligned}
&\left|\frac{1}{\tau_N}E^{\xi_0^N}[m_1^N(\xi_{2\delta_N}^{N,v})] - \theta^N Z_0^N\right| \\
&\quad \le \sum_{l=1}^{R}\sum_{k=1}^{N_R} |\theta^{N,1}(l,k)| \cdot \left|V_{2\delta_N}^{N,1}(l,k) - p^{N,1}(l,k)Z_0^N\right| \\
&\quad + |\theta^{N,0}(l,k)| \cdot \left|V_{2\delta_N}^{N,0}(l,k) - p^{N,0}(l,k)Z_0^N\right| \\
&\quad \le b\mathcal{K}(\varepsilon^{N,1} + \varepsilon^{N,0}) \cdot (1 + (Z_0^N)^2).
\end{aligned}$$

The proof for part (b) is similar. By definition,

$$\begin{aligned}
&\frac{1}{\tau_N^2}E^{\xi_0^N}[m_2^N(\xi_{2\delta_N}^{N,v})] \\
&= \frac{1}{\tau_N^2}E^{\xi_0^N}\left[\tau_N v_N(\xi_{2\delta_N}^{N,v})\right] + \frac{1}{\tau_N} \cdot \left(\left[\frac{1}{\tau_N}E^{\xi_0^N}\left[d^{N,1}(\xi_{2\delta_N}^{N,v})\right]\right]\right. \\
&\quad \left. + \left[\frac{1}{\tau_N}E^{\xi_0^N}\left[d^{N,0}(\xi_{2\delta_N}^{N,v})\right]\right]\right) \\
&= V_{2\delta_N}^{N,v} + \frac{1}{\tau_N}\sum_{l=1}^{R}\sum_{k=1}^{N_R}\left(\theta^{N,1}(l,k)\left[\frac{1}{\tau_N}E^{\xi_0^N}(d_{l,k}^{N,1}(\xi_t^{N,v}))\right]\right. \\
&\quad \left. + \theta^{N,0}(l,k)\left[\frac{1}{\tau_N}E^{\xi_0^N}(d_{l,k}^{N,0}(\xi_t^{N,v}))\right]\right) \\
&= V_{2\delta_N}^{N,v} + \frac{1}{\tau_N}\sum_{l=1}^{R}\sum_{k=1}^{N_R}\left(\theta^{N,1}(l,k) \cdot V_{2\delta_N}^{N,1}(l,k) + \theta^{N,0}(l,k) \cdot V_{2\delta_N}^{N,0}(l,k)\right).
\end{aligned}$$

By Lemma 3.5(b) and the fact that $|\xi_t^{N,v}|$ is a martingale, for $i = 0, 1$

$$|V_{2\delta_N}^{N,i}(l,k)| = \left|\frac{1}{\tau_N}E^{\xi_0^N}(d_{l,k}^{N,i}(\xi_{2\delta_N}^{N,v}))\right| \le C_{3.5} \cdot \frac{1}{\tau_N}E^{\xi_0^N}(|\xi_{2\delta_N}^{N,v}|) = C_{3.5}Z_0^N.$$

Therefore,

$$\begin{aligned}
&\left|\frac{1}{\tau_N^2}E^{\xi_0^N}[m_2^N(\xi_{2\delta_N}^{N,v})] - \beta^N Z_0^N\right| \\
&\quad \le \left|V_{2\delta_N}^{N,v} - 2p^{N,e}Z_0^N\right| + \frac{b}{\tau_N}\sum_{l=1}^{R}\sum_{k=1}^{N_R}\left(|V_{2\delta_N}^{N,1}(l,k)| + |V_{2\delta_N}^{N,0}(l,k)|\right) \\
&\quad \le \varepsilon^{N,v} \cdot \left[1 + (Z_0^N)^2\right] + \frac{2b\mathcal{K}}{\tau_N} \cdot C_{3.5}Z_0^N.
\end{aligned}$$

Consequently, we could choose

$$\varepsilon^N_{4.2} = \max\left(b\mathcal{K}(\varepsilon^{N,1} + \varepsilon^{N,0}),\ \varepsilon^{N,v},\ \frac{2b\mathcal{K}}{\tau_N} \cdot C_{3.5}\right).$$

■

Proof of Proposition 4.1. For (a), apply Corollary 3.7 (a) and Proposition 4.2 (a),

$$\begin{aligned}
&\left|\frac{1}{\tau_N} E^{\xi^N_0}[m^N_1(\xi^N_{2\delta_N})] - \theta^N Z^N_0\right| \\
&\quad \le \left|\frac{1}{\tau_N} E^{\xi^N_0}[m^N_1(\xi^{N,v}_{2\delta_N})] - \theta^N Z^N_0\right| + \frac{1}{\tau_N} E^{\xi^N_0}\left|m^N_1(\xi^N_{2\delta_N}) - m^N_1(\xi^{N,v}_{2\delta_N})\right| \\
&\quad \le \varepsilon^N_{4.2}(1 + (Z^N_0)^2) + C_{3.7} \cdot Z^N_0\left((e^{(b\mathcal{K})(2\delta_N)} - 1) + (2\delta_N)e^{(b\mathcal{K})(2\delta_N)}\right).
\end{aligned}$$

Similarly for (b), apply Corollary 3.7 (b) and Proposition 4.2 (b),

$$\begin{aligned}
&\left|\frac{1}{\tau_N} E^{\xi^N_0}[m^N_2(\xi^N_{2\delta_N})] - \beta^N Z^N_0\right| \\
&\quad \le \left|\frac{1}{\tau_N} E^{\xi^N_0}[m^N_2(\xi^{N,v}_{2\delta_N})] - \beta^N Z^N_0\right| + \frac{1}{\tau_N} E^{\xi^N_0}\left|m^N_2(\xi^N_{2\delta_N}) - m^N_2(\xi^{N,v}_{2\delta_N})\right| \\
&\quad \le \varepsilon^N_{4.2}(1 + Z^N_0 + (Z^N_0)^2) + C_{3.7}\left(1 + \frac{1}{\tau_N}\right) \cdot Z^N_0\left((e^{(b\mathcal{K})(2\delta_N)} - 1)\right. \\
&\qquad \left. + (2\delta_N)e^{(b\mathcal{K})(2\delta_N)}\right).
\end{aligned}$$

Note that $e^{(b\mathcal{K})(2\delta_N)} - 1 \to 0$ as $\delta_N \to 0$. Since by (1.3), $Z^N_0 \le C_{1.3}$, then we can choose

$$\varepsilon^N_{4.1} = \max\left\{\varepsilon^N_{4.2}(1 + C_{1.3}),\ 2C_{3.7}\left((e^{(b\mathcal{K})(2\delta_N)} - 1) + (2\delta_N)e^{(b\mathcal{K})(2\delta_N)}\right)\right\}$$

so that $\varepsilon^N_{4.1} \to 0$. ■

Chapter 5

Convergence of coalescing random walk probabilities

In this chapter, we will show in Proposition 5.1 that the coalescing random walk probabilities defined on G_N in the previous chapter converge and the limits as $N \to \infty$ are the corresponding coalescing walk probabilities on the infinite tree.

Recall the definitions

$$q^N(x,y) = 1/r \cdot 1_{\{x \sim y\}}, \quad x, y \in V_N$$

and

$$q^{tr}(a,b) = 1/r \cdot 1_{\{a \sim b\}}, \quad a, b \in V^{tr}.$$

The sequence t_N satisfies $\log N \ll t_N \ll \tau_N$ and the coalescing random walk probabilities defined in Chapter 4 are

$$p^{N,e} = \frac{1}{N} \sum_{x,y \in V_N} q^N(x,y) P(\tau^N(x,y) > t_N),$$

$$p^{N,1}(l,k) = \frac{1}{N} \sum_{x} \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} P(\tau^N(A,B) > t_N,\ \sigma^N(A) \leq t_N),$$

$$p^{N,0}(l,k) = \frac{1}{N}\sum_{x}\sum_{\substack{A\in S_{l,k}(x)\\ B=\overline{\mathcal{N}_l(x)}\cap A^c}} P(\tau^N(A,B) > t_N,\ \sigma^N(B) \le t_N).$$

Recall that $\rho \in V^{tr}$ is the root of the tree, and the system of rate 1 coalescing walk system is denoted as $\{\widehat{B}_t^e : e \in V^{tr}\}$. For $A,\ B \subset V^{tr}$ disjoint, the stopping times defined in Section 1.3 are

$$\sigma^{tr}(A) = \inf\{t > 0 : |\widehat{B}_t^A| = 1\},$$
$$\tau^{tr}(A,B) = \inf\{t > 0 : \widehat{B}_t^A \cap \widehat{B}_t^B \ne \varnothing\}.$$

The coalescing walk probabilities on the infinite tree defined in Section 1.3 are

$$p^{tr,e} = \sum_{e} q^{tr}(\rho,e)P(\tau^{tr}(\rho,e) = \infty),$$
$$p^{tr,1}(l,k) = \sum_{\substack{A\in S_{l,k}(\rho)\\ B=\overline{\mathcal{N}_l(\rho)}\cap A^c}} P(\tau^{tr}(A,B) = \infty,\ \sigma^{tr}(A) < \infty),$$
$$p^{tr,0}(l,k) = \sum_{\substack{A\in S_{l,k}(\rho)\\ B=\overline{\mathcal{N}_l(\rho)}\cap A^c}} P(\tau^{tr}(A,B) = \infty,\ \sigma^{tr}(B) < \infty)$$

where $S_{l,k}(\rho) = \{A \subseteq \mathcal{N}_l(\rho) : |A| = k\}$ for $\rho \in V^{tr}$.

We now state the main result of this chapter.

Proposition 5.1. *As $N \to \infty$,*

(a) $p^{N,e} \to p^{tr,e}$.

(b) $p^{N,i}(l,k) \to p^{tr,i}.(l,k),\ i = 0,1$

We begin with proving Lemma 5.2-5.6 to prepare for the proof.

Lemma 5.2. *For any $s_N \to \infty$, $s_N < t_N$, there is a sequence $\varepsilon^N_{5.2} \to 0$ such that*

$$P(\tau^N(x,y) \in (s_N,t_N]) \le \varepsilon^N_{5.2}, \quad \text{for all } x,y \in V_N.$$

Proof. Let $x,y \in V_N$ be arbitrary. By Lemma A.2

$$P(\tau^N(x,y) \in (s_N,t_N]) \le e^2 \int_{s_N}^{t_N+1} P(B_s^{N,x} = B_s^{N,y})\, ds. \tag{5.1}$$

By (1.8),

$$\begin{aligned} P(B_s^{N,x} = B_s^{N,y}) &= \sum_z P(B_s^{N,x} = z) P(B_s^{N,y} = z) \\ &\le \sum_z p_s^N(x,z) \cdot \left(\left| p_s^N(y,z) - \frac{1}{N} \right| + \frac{1}{N} \right) \\ &\le e^{-\gamma s} + \frac{1}{N}. \end{aligned}$$

Therefore,

$$\int_{s_N}^{t_N+1} P(B_s^{N,x} = B_s^{N,y}) \, ds \le \int_{s_N}^{t_N+1} \left[e^{-\gamma s} + \frac{1}{N} \right] ds \le \frac{t_N+1}{N} + \frac{1}{\gamma} e^{-\gamma s_N}.$$

Thus, choose $\varepsilon_{5.2}^N$ to be

$$\varepsilon_{5.2}^N = e^2 \left(\frac{t_N+1}{N} + \frac{1}{\gamma} e^{-\gamma s_N} \right)$$

and $\varepsilon_{5.2}^N \to 0$ since $t_N \ll N$ and $s_N \to \infty$. ■

Let s_N be a positive sequence such that $s_N \to \infty$ and $s_N \ll \log N$. Note that $t_N \gg \log N$ implies $s_N \ll t_N$. Now define

$$\begin{aligned} p_1^{N,e} &= \frac{1}{N} \sum_x \sum_y q^N(x,y) P(\tau^N(x,y) > s_N), \\ p_1^{N,1}(l,k) &= \frac{1}{N} \sum_x \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} P(\tau^N(A,B) > s_N, \ \sigma^N(A) \le s_N), \\ p_1^{N,0}(l,k) &= \frac{1}{N} \sum_x \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} P(\tau^N(A,B) > s_N, \ \sigma^N(B) \le s_N). \end{aligned}$$

Recall that $\mathcal{K} = 1 + R \cdot N_R$ and $C_R = 2^{\mathcal{K}} \mathcal{K}^2$.

Lemma 5.3.

(a) $|p_1^{N,e} - p^{N,e}| \le \varepsilon_{5.2}^N$.

(b) $|p_1^{N,i}(l,k) - p^{N,i}(l,k)| \le 2 C_R \varepsilon_{5.2}^N$, $i = 0, 1$.

Proof. For part (a), by Lemma 5.2 we have

$$|p_1^{N,e} - p^{N,e}| \le \frac{1}{N} \sum_{x,y} q^N(x,y) |P(\tau^N(x,y) > s_N) - P(\tau^N(x,y) > t_N)|$$

$$= \frac{1}{N}\sum_{x,y} q^N(x,y)P(\tau^N(x,y) \in (s_N,t_N])$$
$$\leq \left(\frac{1}{N}\sum_{x,y} q^N(x,y)\right) \cdot \varepsilon^N_{5.2} = \varepsilon^N_{5.2}.$$

For (b), we have that for $i = 1$,

$$|P(\tau^N(A,B) > t_N,\ \sigma^N(A) \leq t_N) - P(\tau^N(A,B) > s_N,\ \sigma^N(A) \leq s_N)|$$
$$\leq P(\tau^N(A,B) \in (s_N,t_N]) + P(\sigma^N(A) \in (s_N,t_N])$$
$$\leq \sum_{\substack{a\in A\\ b\in B}} P(\tau^N(A,B) \in (s_N,t_N]) + \sum_{a',a''\in A} P(\tau^N(a',a'') \in (s_N,t_N])$$

so that by Lemma 5.2,

$$|p^{N,1}_1(l,k) - p^{N,1}(l,k)| \leq \frac{1}{N}\sum_x \sum_{\substack{A\in S_{l,k}(x)\\ B=\overline{\mathcal{N}_l(x)}\cap A^c}} (|A||B| + |A|^2)\varepsilon^N_{5.2} \leq 2C_R\varepsilon^N_{5.2}.$$

The calculation for $i = 0$ is similar. ■

Recall from Section 1.2 that for $l_N = \frac{1}{5}\log_{r-1} N$,

$$\Gamma_N = \{x \in V_N : tx(B_{l_N}(x)) = 0\},$$
$$\Gamma'_N = \{x \in V_N : tx(B_{l_N}(x)) = 1\},$$
$$V_N = \Gamma_N \cup \Gamma'_N.$$

We now define a group of walk probabilities that are averages over only sites in Γ_N:

$$p^{N,e}_2 = \frac{1}{|\Gamma_N|}\sum_{x\in\Gamma_N}\sum_y q^N(x,y)P(\tau^N(x,y) > s_N),$$
$$p^{N,1}_2(l,k) = \frac{1}{|\Gamma_N|}\sum_{x\in\Gamma_N}\sum_{\substack{A\in S_{l,k}(x)\\ B=\overline{\mathcal{N}_l(x)}\cap A^c}} P(\tau^N(A,B) > s_N,\ \sigma^N(A) \leq s_N),$$
$$p^{N,0}_2(l,k) = \frac{1}{|\Gamma_N|}\sum_{x\in\Gamma_N}\sum_{\substack{A\in S_{l,k}(x)\\ B=\overline{\mathcal{N}_l(x)}\cap A^c}} P(\tau^N(A,B) > s_N,\ \sigma^N(B) \leq s_N).$$

Eventually, we will see that the above converge to probabilities on the infinite tree, as G_N are good graphs which means that with high probability there is locally a finite tree at most of the sites. The next lemma shows that $p_2^{N,e}$ and $p_2^{N,i}(l,k)$ are close to $p_1^{N,e}$ and $p_1^{N,i}(l,k)$, respectively.

Lemma 5.4.

(a) $|p_1^{N,e} - p_2^{N,e}| \leq 2(r^2 N^{-2/5})$.

(b) $|p_1^{N,i}(l,k) - p_2^{N,i}(l,k)| \leq 2C_R(r^2 N^{-2/5})$, $i = 0, 1$.

Proof. For (a), by (1.7) we have

$$\begin{aligned} |p_1^{N,e} - p_2^{N,e}| &\leq \left| \frac{1}{N} - \frac{1}{|\Gamma_N|} \right| \cdot \sum_{x \in \Gamma_N} P(\tau^N(x,y) > s_N) + \frac{1}{N} \sum_{x \in \Gamma'_N} P(\tau^N(x,y) > s_N) \\ &\leq |\Gamma_N| \cdot \left| \frac{1}{N} - \frac{1}{|\Gamma_N|} \right| + \frac{|\Gamma'_N|}{N} \\ &= \frac{2|\Gamma'_N|}{N} \\ &\leq 2(r^2 N^{-2/5}). \end{aligned}$$

For (b) for $i = 1$, similarly we have that

$$\begin{aligned} &|p_1^{N,1}(l,k) - p_2^{N,1}(l,k)| \\ &\leq \left| \frac{1}{N} - \frac{1}{|\Gamma_N|} \right| \sum_{x \in \Gamma_N} \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} P(\tau^N(A,B) > s_N,\ \sigma^N(A) \leq s_N) \\ &\quad + \frac{1}{N} \sum_{x \in \Gamma'_N} \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} P(\tau^N(A,B) > s_N,\ \sigma^N(A) \leq s_N) \\ &\leq C_R \left(|\Gamma_N| \cdot \left| \frac{1}{N} - \frac{1}{|\Gamma_N|} \right| + \frac{|\Gamma'_N|}{N} \right) \\ &\leq 2C_R(r^2 N^{-2/5}) \end{aligned}$$

and for $i = 0$ the calculation is similar. ∎

We need to introduce a coupling between walks on V_N and the infinite tree in preparation of proving Lemma 5.6 which gives the convergence to coalescing random walk probabilities on the tree. Fix $x \in \Gamma_N$. Recall that $\rho \in V^{tr}$ is the root of the infinite tree and the interaction range $R > 0$ is fixed and finite. Define the exit times

$$T^N(x) = \inf\{t > 0 : \exists y \in B_R(x),\ \widehat{B}_t^{N,y} \notin B_{(1/2)l_N}(x)\},$$
$$T^{tr} = \inf\{t > 0 : \exists e \in B_R(\rho),\ B_t^e \notin B_{(1/2)l_N}(\rho)\}.$$

Note that since $x \in \Gamma_N$, then $B_{(1/2)l_N}(x)$ is a finite tree as it is loop-free. Thus, we can couple the coalescing walks started at sites in $B_R(x)$, $\widehat{B}_t^{N,B_R(x)}$, and the walks started at sites in $B_R(\rho)$, $\widehat{B}_t^{B_R(\rho)}$ up until $T^N(x)$ as follows: first, introduce a graph isomorphism ψ such that $B_R(x) = \psi^{-1}(B_R(\rho))$. Next, for $y \in B_R(x)$, define $\widetilde{B}_t^{N,y}$ as follows:

$$\widetilde{B}_t^{N,y} = \begin{cases} \psi(\widehat{B}_t^{N,y}), & t < T^N(x) \\ \widehat{B}_{t-T^N(x)}^{\psi(\widehat{B}_{T^N(x)}^{N,y})}, & t \geq T^N(x). \end{cases} \tag{5.2}$$

This definition says the following: before exiting $B_R(\rho)$, the walk $\widetilde{B}_t^{N,y}$ "duplicates" the realization of $\widehat{B}_t^{N,y}$ via the isomorphism ψ on the tree. At $t = T^N(x)$, the state of $\widetilde{B}_t^{N,y}$ is $\psi(\widehat{B}_{T^N(x)}^{N,y})$. And it performs as an usual coalescing random walk after the exit time $T^N(x)$. It is not hard to see that $\widetilde{B}_t^{N,y}$ has the same law as $\widehat{B}_t^{N,y}$ for any $y \in B_R(x)$. Thus, later when we compare the walk probabilities on G_N with those on the infinite tree in the proof of Lemma 5.6, we will keep using the notation $\widehat{B}_t^{N,y}$ instead of $\widetilde{B}_t^{N,y}$.

We now give a probability bound on $T^N(x)$ and T^{tr} that will be used in the proof of Lemma 5.6.

Lemma 5.5. $P(T^N(x) \leq 2s_N) \vee P(T^{tr} \leq 2s_N) \leq \dfrac{2s_N}{(1/2)l_N - R}$.

Proof. Observe that

$$P(T^N(x) \leq 2s_N) \leq \sum_{y \in B_R(x)} P(\exists t \leq 2s_N,\ \widehat{B}_t^{N,y} \notin B_{(1/2)l_N}(x)).$$

Notice that for $y \in B_R(x)$,

$$\begin{aligned} &P(\exists t \leq 2s_N,\ \widehat{B}_t^{N,y} \notin B_{(1/2)l_N}(x)) \\ &\quad = P(\exists t \leq 2s_N,\ d(x, \widehat{B}_t^{N,y}) \geq (1/2)l_N) \\ &\quad \leq P(\exists t \leq 2s_N,\ d(x, \widehat{B}_t^{N,y}) + d(x,y) \geq (1/2)l_N) \end{aligned}$$

$$\le P(\widehat{B}_t^{N,y} \text{ makes at least } [(1/2)l_N - R] \text{ jumps by time } 2s_N)$$

$$\le E(\text{number of jumps made by } \widehat{B}_{2s_N}^{N,y})/((1/2)l_N - R)$$

$$= \frac{2s_N}{(1/2)l_N - R},$$

and similarly,

$$P(\exists t \le 2s_N,\ \widehat{B}_t^e \notin B_{(1/2)l_N}(\rho)) \le \frac{2s_N}{(1/2)l_N - R}.$$

■

Lemma 5.6 compares the probabilities on G_N and the tree probabilities in the limit $N \to \infty$.

Lemma 5.6. *There exists a sequence $\varepsilon_{5.6}^N \to 0$ such that*

(a) $|p_2^{N,e} - p^{tr,e}| \le \varepsilon_{5.6}^N$.

(b) $|p_2^{N,i}(l,k) - p^{tr,i}(l,k)| \le 2C_R \cdot \varepsilon_{5.6}^N$, $i = 0, 1$.

Proof. For part (a), define $p_{s_N}^{N,e}(x)$

$$\begin{aligned} p_{s_N}^{N,e}(x) &= \sum_y q^N(x,y)P(\tau^N(x,y) > s_N) \\ &= \sum_y q^N(x,y)P(\tau^N(x,y) > s_N, T^N(x) > 2s_N) \\ &+ \sum_y q^N(x,y)P(\tau^N(x,y) > s_N, T^N(x) \le 2s_N) \end{aligned}$$

so that we can write

$$p_2^{N,e} = \frac{1}{|\Gamma_N|} \sum_{x \in \Gamma_N} \sum_y p_{s_N}^{N,e}(x).$$

By coupling the walks in G_N started at sites in $B_R(x)$ and the walks in G^{tr} started at sites in $B_R(\rho)$ as defined in (5.2),

$$\begin{aligned} &\sum_y q^N(x,y)P(\tau^N(x,y) > s_N, T^N(x) > 2s_N) \qquad (5.3) \\ &= \sum_e q^{tr}(\rho,e)P(\tau^{tr}(\rho,e) > s_N, T^{tr} > 2s_N). \end{aligned}$$

Define

$$\begin{aligned} p_{s_N}^{tr,e} &= \sum_e q^{tr}(\rho,e)P(\tau^{tr}(\rho,e) > s_N) \\ &= \sum_e q^{tr}(\rho,e)P(\tau^{tr}(\rho,e) > s_N, T^{tr} > 2s_N) \\ &+ \sum_e q^{tr}(\rho,e)P(\tau^{tr}(\rho,e) > s_N, T^{tr} \leq 2s_N). \end{aligned}$$

By (5.3),

$$\begin{aligned} |p_{s_N}^{N,e}(x) - p_{s_N}^{tr,e}| &\leq \sum_y q^N(x,y)P(T^N(x) \leq 2s_N) + \sum_e q^{tr}(\rho,e)P(T^{tr} \leq 2s_N) \\ &= P(T^N(x) \leq 2s_N) + P(T^{tr} \leq 2s_N) \end{aligned}$$

so that Lemma 5.5 implies

$$|p_{s_N}^{N,e}(x) - p_{s_N}^{tr,e}| \leq \frac{4s_N}{(1/2)l_N - R} \quad \text{for all } x \in \Gamma_N. \tag{5.4}$$

Since

$$\begin{aligned} |p_{s_N}^{tr,e} - p^{tr,e}| &\leq \sum_e q^{tr}(\rho,e)|P(\tau^{tr}(\rho,e) > s_N) - P(\tau^{tr}(\rho,e) = \infty)| \\ &= \sum_e q^{tr}(\rho,e)P(\tau^{tr}(\rho,e) \in (s_N,\infty)) \\ &\leq \Big[\max_{a,b\in B_R(\rho)} P(\tau^{tr}(a,b) \in (s_N,\infty))\Big] \sum_e q^{tr}(\rho,e) \\ &= \max_{a,b\in B_R(\rho)} P(\tau^{tr}(a,b) \in (s_N,\infty)). \end{aligned} \tag{5.5}$$

Note that walks on the infinite tree are transient, then for any $a, b \in B_R(\rho)$,

$$P(\tau^{tr}(a,b) \in (s_N,\infty)) \to 0 \quad \text{as } s_N \to \infty. \tag{5.6}$$

Therefore, we can define

$$\varepsilon_{5.6}^N = \frac{4s_N}{(1/2)l_N - 2R} + \max_{a,b\in B_R(\rho)} P(\tau^{tr}(a,b) \in (s_N,\infty))$$

and $\varepsilon_{5.6}^N \to 0$ because $s_N \ll \log_N$ and $s_N \to \infty$. Therefore, combining (5.4) and (5.5) , we have

$$\begin{aligned}|p_2^{N,e} - p^{tr,e}| &\le \left| \frac{1}{|\Gamma_N|} \sum_{x \in \Gamma_N} (p_{s_N}^{N,e}(x) - p_{s_N}^{tr,e}) \right| + |p_{s_N}^{tr,e} - p^{tr,e}| \\ &\le \left(\frac{1}{|\Gamma_N|} \sum_{x \in \Gamma_N} |p_{s_N}^{N,e}(x) - p_{s_N}^{tr,e}| \right) + |p_{s_N}^{tr,e} - p^{tr,e}| \\ &\le \varepsilon_{5.6}^N.\end{aligned}$$

For part (b) for $i = 1$ and $x \in \Gamma_N$, define

$$\begin{aligned}a_N(x) &= \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} P(\tau^N(A,B) > s_N,\ \sigma^N(A) \le s_N, T^N(x) > 2s_N), \\ b_N(x) &= \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} P(\tau^N(A,B) > s_N,\ \sigma^N(A) \le s_N, T^N(x) \le 2s_N).\end{aligned}$$

Write $p_{s_N}^{N,1}(x)$ as

$$\begin{aligned}p_{s_N}^{N,1}(x) &= \sum_{\substack{A \in S_{l,k}(x) \\ B = \overline{\mathcal{N}_l(x)} \cap A^c}} P(\tau^N(A,B) > s_N,\ \sigma^N(A) \le s_N) \\ &= a_N(x) + b_N(x).\end{aligned}$$

Similarly, define

$$\begin{aligned}a_N^{tr} &= \sum_{\substack{A \in S_{k,l}(\rho) \\ B = \overline{\mathcal{N}_l(\rho)} \cap A^c}} P(\tau^{tr}(A,B) > s_N,\ \sigma^{tr}(A) \le s_N,\ T^{tr} > 2s_N), \\ b_N^{tr} &= \sum_{\substack{A \in S_{k,l}(\rho) \\ B = \overline{\mathcal{N}_l(\rho)} \cap A^c}} P(\tau^{tr}(A,B) > s_N,\ \sigma^{tr}(A) \le s_N,\ T^{tr} \le 2s_N)\end{aligned}$$

and write $p_{s_N}^{tr,1}$ as

$$\begin{aligned}p_{s_N}^{tr,1} &= \sum_{\substack{A \in S_{k,l}(\rho) \\ B = \overline{\mathcal{N}_l(\rho)} \cap A^c}} P(\tau^{tr}(A,B) > s_N,\ \sigma^{tr}(A) \le s_N) \\ &= a_N^{tr} + b_N^{tr}.\end{aligned}$$

Note that the value of a_N^{tr} and b_N^{tr} does not depend on ρ.

By definition we have

$$p_2^{N,1}(l,k) = \frac{1}{|\Gamma_N|}\sum_{x\in\Gamma_N} p_{s_N}^{N,1}(x) = \frac{1}{|\Gamma_N|}\sum_{x\in\Gamma_N}(a_N(x)+b_N(x)).$$

Decompose $a_N(x)$ by

$$a_N(x) = \sum_{\substack{A\in S_{l,k}(x)\\ B=\overline{\mathcal{N}_l(x)}\cap A^c}}\ \sum_{\substack{a\in A\\ b\in B}}\ \sum_{a',a''\in A} P(\tau^N(A,B)=\tau^N(A,B),\ \sigma^N(A)=\tau^N(a',a''), \tag{5.7}$$

$$\tau^N(A,B)>s_N,\ \tau^N(a',a'')\le s_N, T^N(x)>2s_N).$$

By the coupling (5.2) between the walks on G_N started at sites in $B_R(x)$ and the walks in G^{tr} started at sites in $B_R(\rho)$, the right side of (5.7) is equal to

$$\sum_{\substack{A\in S_{k,l}(\rho)\\ B=\overline{\mathcal{N}_l(\rho)}\cap A^c}}\ \sum_{\substack{a\in A\\ b\in B}}\ \sum_{a',a''\in A} P(\tau^{tr}(A,B)=\tau^{tr}(a,b),\ \sigma^{tr}(A)=\tau^{tr}(a',a''),$$

$$\tau^{tr}(a,b)>s_N,\ \tau^{tr}(a',a'')\le s_N, T^{tr}>2s_N)$$

$$= \sum_{\substack{A\in S_{k,l}(\rho)\\ B=\overline{\mathcal{N}_l(\rho)}\cap A^c}} P(\tau^{tr}(A,B)>s_N,\ \sigma^{tr}(A)\le s_N,\ T^{tr}>2s_N),$$

hence for every $x\in\Gamma_N$,

$$a_N(x) = a_N^{tr}.$$

This implies

$$p_{s_N}^{tr,1} = a_N^{tr}+b_N^{tr} = a_N(x)+b_N^{tr} = [p_{s_N}^{N,1}(x)-b_N(x)]+b_N^{tr}$$

and we have

$$\begin{aligned}
|p_{s_N}^{N,1}(x)-p_{s_N}^{tr,1}| &\le b_N(x)+b_N^{tr}\\
&\le \sum_{\substack{A\in S_{k,l}(\rho)\\ B=\overline{\mathcal{N}_l(\rho)}\cap A^c}} P(T^{tr}\le 2s_N) + \sum_{\substack{A\in S_{l,k}(x)\\ B=\overline{\mathcal{N}_l(x)}\cap A^c}} P(T^N(x)\le 2s_N)\\
&\le 2C_R\cdot\frac{2s_N}{(1/2)l_N-R}.
\end{aligned} \tag{5.8}$$

Moreover, we have that

$$
\begin{aligned}
&|p^{tr,1}_{s_N} - p^{tr,1}(l,k)| \\
&\quad \le \sum_{\substack{A\in S_{k,l}(\rho)\\ B=\overline{\mathcal{N}_l(\rho)}\cap A^c}} |P(\tau^{tr}(A,B) > s_N,\ \sigma^{tr}(A)\le s_N) - P(\tau^{tr}(A,B)=\infty,\ \sigma^{tr}(A)<\infty)| \\
&\quad \le \sum_{\substack{A\in S_{k,l}(\rho)\\ B=\overline{\mathcal{N}_l(\rho)}\cap A^c}} P(\tau^{tr}(A,B)\in(s_N,\infty)) + P(\sigma^{tr}(A)\in(s_N,\infty)) \\
&\quad \le \sum_{\substack{A\in S_{k,l}(\rho)\\ B=\overline{\mathcal{N}_l(\rho)}\cap A^c}} \left[\sum_{\substack{a\in A\\ b\in B}} P(\tau^{tr}(a,b)\in(s_N,\infty)) + \sum_{a',a''\in A} P(\tau^{tr}(a',a'')\in(s_N,\infty))\right] \\
&\quad \le 2C_R\cdot \max_{a,b\in B_R(\rho)} P(\tau^{tr}(a,b)\in(s_N,\infty)). \qquad (5.9)
\end{aligned}
$$

Therefore, combining (5.8) and (5.9),

$$
\begin{aligned}
|p^{N,1}_2(l,k) - p^{tr,1}(l,k)| &\le \left|\frac{1}{|\Gamma_N|}\sum_{x\in\Gamma_N} p^{N,1}_{s_N}(x) - p^{tr,1}_{s_N}\right| + |p^{tr,1}_{s_N} - p^{tr,1}(l,k)| \\
&\le \left(\frac{1}{|\Gamma_N|}\sum_{x\in\Gamma_N} |p^{N,1}_{s_N}(x) - p^{tr,1}_{s_N}|\right) + |p^{tr,1}_{s_N} - p^{tr,1}(l,k)| \\
&\le 2C_R\cdot\left(\frac{2s_N}{(1/2)l_N - R} + \max_{a,b\in B_R(\rho)} P(\tau^{tr}(a,b)\in(s_N,\infty))\right) \\
&\le 2C_R\cdot\varepsilon^N_{5.6}.
\end{aligned}
$$

The proof for $i=0$ is similar. ■

Proof of Proposition 5.1. Combining Lemma 5.3, 5.4 and 5.6, we have

$$
\begin{aligned}
|p^{N,e} - p^{tr,e}| &\le |p^{N,e} - p^{N,e}_1| + |p^{N,e}_1 - p^{N,e}_2| + |p^{N,e}_2 - p^{tr,e}| \\
&\le \varepsilon^N_{5.2} + 2(r^2N^{-2/5}) + \varepsilon^N_{5.6},
\end{aligned}
$$

and for $i=0,1$,

$$
\begin{aligned}
|p^{N,i} - p^{tr,i}(l,k)| &\le |p^{N,i} - p^{N,i}_1(l,k)| + |p^{N,i}_1(l,k) - p^{N,i}_2(l,k)| + |p^{N,i}_2(l,k) - p^{tr,i}(l,k)| \\
&\le 2C_R(\varepsilon^N_{5.2} + r^2N^{-2/5} + \varepsilon^N_{5.6}).
\end{aligned}
$$

This completes the proof. ■

Chapter 6

Tightness; identification of the weak limit

In this chapter, we complete the proof of Theorem 1.2 by proving Proposition 6.1 and 6.2. Identification of the limit is done in Lemma 6.7 which shows that mean-field simplification occurs. Note that this result was based on Proposition 4.3 which relies on estimation using duality for the voter model only.

Recall that a family of laws on $D([0,\infty),\mathbb{R}^+)$ is called C-tight if it is tight and every limit point is supported by $C([0,\infty),\mathbb{R}^+)$. The drift θ and branching rate β defined in Section 1.3 are

$$\theta = \sum_{l=1}^{R}\sum_{k=1}^{N_R}(\theta^1(l,k)p^{tr,1}(l,k) - \theta^0(l,k)p^{tr,0}(l,k)),$$

$$\beta = 2p^{tr,e}.$$

Recall that G_N is a sequence of good graphs, and P^N is the law of $Z^N_\cdot$ defined in Section 1.3.

Proposition 6.1. *$\{P^N, N \in \mathbb{N}\}$ is C-tight.*

Proposition 6.2. *If P^* is any weak limit point of $\{P^N\}$, then $P^* = P^{\beta,\theta}$.*

Note that the laws $\{P^N\}$ in the two propositions above is defined over fixed sequence of good graphs.

We first give some preliminary results. Recall from Chapter 3 that

$$d_{l,k}^{N,1}(\xi) = \sum_x \widehat{\xi}(x) 1_{\{n_l^1(x,\xi)=k\}},$$
$$d_{l,k}^{N,0}(\xi) = \sum_x \xi(x) 1_{\{n_l^0(x,\xi)=k\}},$$

and $m_1^N(\xi)$ and $m_2^N(\xi)$ are written as

$$\begin{aligned} m_1^N(\xi) &= d^{N,1}(\xi) - d^{N,0}(\xi) \\ &= \sum_{l=1}^{R}\sum_{k=1}^{N_R}(\theta^{N,1}(l,k)d_{l,k}^{N,1}(\xi) - \theta^{N,0}(l,k)d_{l,k}^{N,0}(\xi)), \\ m_2^N(\xi) &= \tau_N v_N(\xi) + [d^{N,1}(\xi) + d^{N,0}(\xi)] \\ &= \tau_N v_N(\xi) + \sum_{l=1}^{R}\sum_{k=1}^{N_R}(\theta^{N,1}(l,k)d_{l,k}^{N,1}(\xi) + \theta^{N,0}(l,k)d_{l,k}^{N,0}(\xi)). \end{aligned}$$

The bias parameter b and the constant $\mathcal{K}$ defined in Chapter 3 are

$$b = \sup_N \sum_{l=1}^{R}\sum_{k=1}^{N_R}(\theta^{N,0}(l,k) + \theta^{N,1}(l,k)),$$
$$\mathcal{K} = 1 + R \cdot N_R.$$

Lemma 6.3. *There is a constant $C_{6.3}$ such that for $\xi \in \{0,1\}^{V_N}$,*

(a) $|m_1^N(\xi)| \le C_{6.3}|\xi|$.

(b) $|m_2^N(\xi)| \le (2\tau_N + C_{6.3})|\xi|$.

Proof. By Lemma 3.5, we have

$$\begin{aligned} |m_1^N(\xi)| &= |d^{N,1}(\xi) - d^{N,0}(\xi)| \\ &\le \sup_{N,l,k}(\theta^{N,1}(l,k) + \theta^{N,0}(l,k)) \cdot \sum_{l=1}^{R}\sum_{k=1}^{N_R}(|d_{l,k}^{N,1}(\xi)| + |d_{l,k}^{N,0}(\xi)|). \end{aligned}$$

By Lemma 3.5(b), $d_{l,k}^{N,i}(\xi) \le C_{3.5} \cdot |\xi|$ for $i = 0, 1$. Together with (P2), we have that

$$|m_1^N(\xi)| \le b \cdot \Big[\sum_{l=1}^{R} \sum_{k=1}^{N_R} (|d_{l,k}^{N,1}(\xi)| + |d_{l,k}^{N,0}(\xi)|) \Big]$$
$$\le 2b\mathcal{K}C_{3.5} \cdot |\xi|. \tag{6.1}$$

Similarly, we have that by Lemma 3.5 (a) and (b),

$$|m_2^N(\xi)| \le \tau_N |v_N(\xi)| + |d^{N,0}(\xi)| + |d^{N,1}(\xi)|$$
$$\le (2\tau_N + 2b\mathcal{K}C_{3.5})|\xi|$$

hence one can choose $C_{6.3} = 2b\mathcal{K}C_{3.5}$. ∎

Recall that

$$\mathcal{D}_t^N = \frac{1}{\tau_N} \int_0^t m_1^N(\xi_s^N)\, ds,$$
$$\langle \mathcal{M}^N \rangle_t = \frac{1}{\tau_N^2} \int_0^t m_2^N(\xi_s^N)\, ds.$$

The next lemma gives bound on $\mathcal{D}_\cdot^N$ and $\langle \mathcal{M}^N \rangle_\cdot$.

Lemma 6.4. *Let $K > 0$, $T > 0$ be fixed. There is a constant $C_{6.4}(K,T)$ such that for $\sup_N Z_0^N \le K$ and for any $s \le T$,*

$$\frac{1}{\tau_N} E^{\xi_0^N} |m_1^N(\xi_s^N)| \le C_{6.4}(K,T),$$
$$\frac{1}{\tau_N^2} E^{\xi_0^N} |m_2^N(\xi_s^N)| \le C_{6.4}(K,T).$$

Proof. By Lemma 6.3 and Corollary 3.2 (a), we have

$$\frac{1}{\tau_N} E^{\xi_0^N} |m_1^N(\xi_s^N)| \le C_{6.3} \cdot E^{\xi_0^N}[Z_s^N] \le C_{6.3} \cdot K e^{(b\mathcal{K})T}$$

and similarly

$$\frac{1}{\tau_N^2} E^{\xi_0^N} |m_2^N(\xi_s^N)| \le \frac{2\tau_N + C_{6.3}}{\tau_N} \cdot E^{\xi_0^N}[Z_s^N] \le 3K e^{(b\mathcal{K})T}.$$

Consequently, we could choose $C_{6.4}(K,T) = (C_{6.3} \vee 3) K e^{(b\mathcal{K})T}$ and this completes the proof. ∎

Denote $\mathcal{T}_T^N = \{$all stopping times bounded by T that are relative to $\mathcal{F}_\cdot^N\}$. We will show that $\mathcal{D}_\cdot^N$ and $\mathcal{M}^N_\cdot$ satisfy Aldou's criterion for tightness in Lemma 6.5 - 6.6. The bounds in Proposition 3.1 will be used in the proof of Lemma 6.5.

Lemma 6.5. *Assume that* $\sup_N Z_0^N \le K$. *For every* $\varepsilon > 0,\ T > 0,$

$$\sup_N \sup_{\substack{S_1,S_2\in\mathcal{T}_T^N\\ S_1\le S_2\le S_1+\delta}} P^{\xi_0^N}(|\mathcal{D}_{S_2}^N - \mathcal{D}_{S_1}^N| > \varepsilon) \to 0, \tag{6.2}$$

$$\sup_N \sup_{\substack{S_1,S_2\in\mathcal{T}_T^N\\ S_1\le S_2\le S_1+\delta}} P^{\xi_0^N}(|\langle\mathcal{M}^N\rangle_{S_2} - \langle\mathcal{M}^N\rangle_{S_1}| > \varepsilon) \to 0 \tag{6.3}$$

as $\delta \downarrow 0$.

Proof. Fix $T > 0$ and $\delta > 0$. Let $S_1,\ S_2 \in \mathcal{T}_T^N$ such that $S_1 \le S_2 \le S_1 + \delta$. We first prove (6.2). Apply Lemma 6.3(a) and we have

$$\begin{aligned} |\mathcal{D}_{S_2}^N - \mathcal{D}_{S_1}^N| &= \left|\int_{S_1}^{S_2} \frac{1}{\tau_N} m_1^N(\xi_s^N)\, ds\right| \\ &\le \frac{1}{\tau_N}\int_{S_1}^{S_1+\delta} |m_1^N(\xi_s^N)|\, ds \\ &\le C_{6.3}\cdot \frac{1}{\tau_N}\int_{S_1}^{S_1+\delta} |\xi_s^N|\, ds. \end{aligned} \tag{6.4}$$

Recall that $\xi_s^{N,b}$ is the biased voter model defined in Chapter 3. Define

$$Z_s^{N,b} = \frac{|\xi_s^{N,b}|}{\tau_N}.$$

By the coupling (3.3), (6.4) is bounded by

$$C_{6.3}\cdot\frac{1}{\tau_N}\int_{S_1}^{S_1+\delta} |\xi_s^{N,b}|\, ds \le C_{6.3}\delta\cdot\Big(\sup_{s\le T+1} Z_s^{N,b}\Big)$$

Since $|\xi_s^{N,b}|$ is a submartingale, we can apply Doob's inequality and together with Proposition 3.1(a),

$$\begin{aligned} P^{\xi_0^N}(|\mathcal{D}_{S_2}^N - \mathcal{D}_{S_1}^N| > \varepsilon) &\le P^{\xi_0^N}\Big(C_{6.3}\delta\cdot\Big(\sup_{s\le T+1} Z_s^{N,b}\Big) > \varepsilon\Big) \\ &\le (C_{6.3}\delta/\varepsilon)\cdot E^{\xi_0^N}\Big(Z_{T+1}^{N,b}\Big) \\ &\le \delta\cdot(C_{6.3}K/\varepsilon)e^{(b\mathcal{K})(T+1)}. \end{aligned}$$

The proof of (6.3) is similar. Apply Lemma 6.3(b) and by (3.3),

$$\begin{aligned}
|\langle \mathcal{M}^N\rangle_{S_2} - \langle \mathcal{M}^N\rangle_{S_1}| &= \left| \int_{S_1}^{S_2} \frac{1}{\tau_N^2} m_2^N(\xi_s^N)\ ds \right| \\
&\le \frac{1}{\tau_N^2} \int_{S_1}^{S_1+\delta} |m_2^N(\xi_s^N)|\ ds \\
&\le (2\tau_N + C_{6.3}) \cdot \frac{1}{\tau_N^2} \int_{S_1}^{S_1+\delta} |\xi_s^N|\ ds \\
&\le 3 \int_{S_1}^{S_1+\delta} Z_s^{N,b}\ ds \\
&\le 3\delta \cdot \Big(\sup_{s\le T+1} Z_s^{N,b} \Big).
\end{aligned}$$

And by Doob's inequality and Proposition 3.1(a),

$$\begin{aligned}
P^{\xi_0^N}(|\langle \mathcal{M}^N\rangle_{S_2} - \langle \mathcal{M}^N\rangle_{S_1}| > \varepsilon) &\le P^{\xi_0^N}\Big(3\delta \cdot \Big(\sup_{s\le T+1} Z_s^{N,b} \Big) > \varepsilon \Big) \\
&\le (3\delta/\varepsilon) \cdot E^{\xi_0^N}\Big(Z_{T+1}^{N,b} \Big) \\
&\le \delta(3K/\varepsilon) e^{(b\mathcal{K})(T+1)}.
\end{aligned}$$

This completes the proof. ■

Lemma 6.6. *For every $T, K > 0$,*

$$\sup_N \sup_{\xi_0^N : Z_0^N \le K} P^{\xi_0^N}\Big(\sup_{s\le T} |\mathcal{D}_s^N| > L \Big) \to 0, \tag{6.5}$$

$$\sup_N \sup_{\xi_0^N : Z_0^N \le K} P^{\xi_0^N}\Big(\sup_{s\le T} |\langle \mathcal{M}^N\rangle_s| > L \Big) \to 0 \tag{6.6}$$

as $L \to \infty$.

Proof. We have that by Markov inequality and Lemma 6.4,

$$\begin{aligned}
P^{\xi_0^N}\Big(\sup_{s\le T} |\mathcal{D}_s^N| > L \Big) &\le P^{\xi_0^N}\Big(\frac{1}{\tau_N} \int_0^T |m_1^N(\xi_s^N)|\ ds > L \Big) \\
&\le \frac{1}{L} \cdot \frac{1}{\tau_N} \int_0^T E^{\xi_0^N} |m_1^N(\xi_s^N)|\ ds \\
&\le \frac{T \cdot C_{6.4}(K,T)}{L}.
\end{aligned}$$

The proof for $\langle \mathcal{M}^N \rangle$ is similar.

■

We now give the estimate on drift and quadratic variation which will be used to identify the weak limit.

Lemma 6.7. *(L^1-estimates) For any $T \geq 0$,*

$$E^{\xi_0^N}\left|\mathcal{D}_T^N - \theta\int_0^T Z_t^N dt\right| \to 0, \tag{6.7}$$

$$E^{\xi_0^N}\left|\langle\mathcal{M}^N\rangle_T - \beta\int_0^T Z_t^N dt\right| \to 0 \tag{6.8}$$

as $N \to \infty$.

Proof. We follow the idea of Proposition 5.1 in Cox [C17]. Fix $T > 0$ and suppose that $\sup_N Z_0^N \leq K$. Recall that the sequence $\delta_N \to 0$ satisfies $\log N \ll \delta_N\tau_N \ll \tau_N \ll N$. For convenience, denote

$$d_t^N = \frac{1}{\tau_N} m_1^N(\xi_t^N)$$

and define

$$I_1^N = \int_0^{2\delta_N} [d_t^N - \theta \cdot Z_t^N]\, dt,$$
$$I_2^N = \int_{T-2\delta_N}^{T} [d_t^N - \theta \cdot Z_t^N]\, dt,$$
$$I_3^N = \int_{2\delta_N}^{T-2\delta_N} [d_t^N - \theta \cdot Z_t^N]\, dt.$$

Thus we have

$$\mathcal{D}_T^N - \theta\int_0^T Z_s^N ds = I_1^N + I_3^N + I_3^N.$$

It is enough to show that $E^{\xi_0^N}|I_i^N| \to 0$ for $i = 1, 2, 3$ as $N \to \infty$. By Lemma 6.4(a) and Corollary 3.7(a) ,

$$\begin{aligned} E^{\xi_0^N}|I_1^N| &\leq \int_0^{2\delta_N} E^{\xi_0^N}|d_t^N - \theta \cdot Z_s^N|\, dt \\ &\leq \int_0^{2\delta_N} [E^{\xi_0^N}|d_t^N| + bE^{\xi_0^N}|Z_s^N|]\, dt \\ &\leq (2\delta_N)\cdot\left[C_{6.4}(K,T) + b\cdot Ke^{(b\mathcal{K})T}\right] \to 0 \quad \text{ as } \delta_N \to 0. \end{aligned}$$

And similarly,

$$E^{\xi_0^N}|I_2^N| \le (2\delta_N) \cdot \Big[C_{6.4}(K,T) + b \cdot Ke^{(b\mathcal{K})T}\Big] \to 0$$

as $\delta_N \to 0$.

Next we bound I_3^N. Since $I_2^N \to 0$ as $\delta_N \to 0$, for convenience we redefine I_3^N as

$$I_3^N = \int_{2\delta_N}^T [d_t^N - \theta \cdot Z_t^N]\, dt.$$

Let $(\mathcal{F}_t^N)$ be the canonical filtration generated by (ξ_t^N). Define

$$\begin{aligned} h_{1,t}^N &= d_t^N - E(d_t^N|\mathcal{F}_{t-2\delta_N}^N), \\ h_{2,t}^N &= E(d_t^N|\mathcal{F}_{t-2\delta_N}^N) - \theta^N Z_{t-2\delta_N}^N, \\ h_{3,t}^N &= \theta^N Z_{t-2\delta_N}^N - \theta Z_t^N. \end{aligned}$$

Write

$$\begin{aligned} &d_t^N - \theta \cdot Z_t^N \\ &\quad = \Big[d_t^N - E(d_t^N|\mathcal{F}_{t-2\delta_N}^N)\Big] + \Big[E(d_t^N|\mathcal{F}_{t-2\delta_N}^N) - \theta^N Z_{t-2\delta_N}^N\Big] + \Big[\theta^N Z_{t-2\delta_N}^N - \theta Z_t^N\Big] \\ &\quad = h_{1,t}^N + h_{2,t}^N + h_{3,t}^N. \end{aligned}$$

We first bound the intergal of $h_{2,t}^N$ and $h_{3,t}^N$. For $h_{2,t}^N$, notice that by Markov property,

$$h_{2,t}^N = E^{\xi_{t-2\delta_N}^N}(d_{2\delta_N}^N) - \theta^N Z_{t-2\delta_N}^N.$$

By Proposition 4.1(a) and Corollary 3.2 (a)-(b),

$$\begin{aligned} E^{\xi_0^N}\left|\int_{2\delta_N}^T h_{2,t}^N\, dt\right| &\le \int_{2\delta_N}^T E^{\xi_0^N}|E^{\xi_{t-2\delta_N}^N}(d_{2\delta_N}^N) - \theta^N Z_{t-2\delta_N}^N|\, dt \\ &\le \varepsilon_{4.1}^N \cdot \int_{2\delta_N}^T E^{\xi_0^N}[(1 + Z_{t-2\delta_N}^N)^2]\, dt \\ &\le \varepsilon_{4.1}^N \cdot T\Big(1 + 2Ke^{(b\mathcal{K})T} + \Big[K^2 + 3KTe^{(b\mathcal{K})T}\Big] \cdot e^{(2b\mathcal{K})T}\Big). \end{aligned} \tag{6.9}$$

For $h_{3,t}^N$, we have

$$h_{3,t}^N = \theta^N\Big[Z_{t-2\delta_N}^N - Z_t^N\Big] + [\theta^N - \theta]Z_t^N.$$

Since

$$\int_{2\delta_N}^{T} [Z^N_{t-2\delta_N} - Z^N_t]\; dt = \int_0^{T-2\delta_N} Z^N_t\; dt - \int_{2\delta_N}^{T} Z^N_t\; dt$$
$$= \int_0^{2\delta_N} Z^N_t\; dt - \int_{T-2\delta_N}^{T} Z^N_t\; dt,$$

so that by Corollary 3.2 (a),

$$E^{\xi_0^N}\left|\int_{2\delta_N}^{T} h^N_{3,t}\; dt\right| \le b\cdot\Big[\int_0^{2\delta_N} E^{\xi_0^N}(Z^N_t)\; dt + \int_{T-2\delta_N}^{T} E^{\xi_0^N}(Z^N_t)\; dt\Big]$$
$$+ |\theta^N - \theta|\int_{2\delta_N}^{T} E^{\xi_0^N}(Z^N_t)\; dt$$
$$\le \Big[4b\delta_N + |\theta^N - \theta|\cdot T\Big]\cdot K e^{(b\mathcal{K})T}. \tag{6.10}$$

Lastly to bound $h^N_{1,t}$, we prove the following two claims.

Claim 1: For each $t > 2\delta_N$, $E^{\xi_0^N}\left[h^N_{1,t}\cdot h^N_{1,s}\right] = 0$ if $s < t - 2\delta_N$.

Proof of Claim 1. $\mathcal{F}^N_s \subseteq \mathcal{F}^N_{t-2\delta_N}$ if $s < t - 2\delta_N$, hence tower rule can be applied:

$$E^{\xi_0^N}(h^N_{1,t}|\mathcal{F}^N_s) = E^{\xi_0^N}(d^N_t - E^{\xi_0^N}(d^N_t|\mathcal{F}^N_{t-2\delta_N})|\mathcal{F}^N_s)$$
$$= E^{\xi_0^N}(d^N_t|\mathcal{F}^N_s) - E^{\xi_0^N}\left[E^{\xi_0^N}(d^N_t|\mathcal{F}^N_{t-2\delta_N})|\mathcal{F}^N_s\right]$$
$$= E^{\xi_0^N}(d^N_t|\mathcal{F}^N_s) - E^{\xi_0^N}(d^N_t|\mathcal{F}^N_s) = 0.$$

Therefore,

$$E^{\xi_0^N}[h^N_{1,t}\cdot h^N_{1,s}] = E^{\xi_0^N}\left[E^{\xi_0^N}(h^N_{1,t}\cdot h^N_{1,s}|\mathcal{F}^N_s)\right]$$
$$= E^{\xi_0^N}\left[h^N_{1,s}\cdot E^{\xi_0^N}(h^N_{1,t}|\mathcal{F}^N_s)\right] = 0.$$

□

Claim 2: There is a constant $C(K,T)$ such that $E^{\xi_0^N}[(h^N_{1,t})^2] \le C(K,T)$ for every $t \le T$.

Proof of Claim 2. By applying conditional Jensen, we have

$$E^{\xi_0^N}[(h^N_{1,t})^2] = E^{\xi_0^N}\left[(d^N_t - E^{\xi_0^N}(d^N_t|\mathcal{F}^N_{t-2\delta_N}))^2\right]$$
$$\le E^{\xi_0^N}\left[(d^N_t)^2 + E^{\xi_0^N}((d^N_t)^2|\mathcal{F}^N_{t-2\delta_N})\right] + 2E^{\xi_0^N}[|d^N_t|\cdot|E^{\xi_0^N}(d^N_t|\mathcal{F}^N_{t-2\delta_N})|]$$

$$\leq 2E^{\xi_0^N}[(d_t^N)^2] + 2(E^{\xi_0^N}[(d_t^N)^2])^{1/2} \cdot (E^{\xi_0^N}[(E^{\xi_0^N}((d_t^N)^2|\mathcal{F}^N_{t-2\delta_N})])^{1/2}$$
$$= 4E^{\xi_0^N}[(d_t^N)^2].$$

By Lemma 6.3 and Lemma 3.2 (b),

$$\begin{aligned} 4E^{\xi_0^N}[(d_t^N)^2] &= 4E^{\xi_0^N}\left[\left|\frac{1}{\tau_N} m_1^N(\xi_t^N)\right|^2\right] \\ &\leq 4C_{6.3}^2 E^{\xi_0^N}[(Z_t^N)^2] \\ &\leq 4C_{6.3}^2 \cdot \left[K^2 + 3K \cdot Te^{T(b\mathcal{K})}\right] \cdot e^{(2b\mathcal{K})T} \equiv C(K,T). \end{aligned}$$

□

Let $C(K,T)$ be from Claim 2. Apply Claim 1 and Claim 2 and we get

$$\begin{aligned} &E^{\xi_0^N}\left[\left(\int_{2\delta_N}^T h_{1,t}^N \, dt\right)^2\right] \\ &= E^{\xi_0^N}\left[\int_{2\delta_N}^T \int_{2\delta_N}^T h_{1,t}^N \cdot h_{1,s}^N \, ds \, dt\right] \\ &\leq \int_{2\delta_N}^T \int_{2\delta_N}^T E^{\xi_0^N}(h_{1,t}^N \cdot h_{1,s}^N) \, ds \, dt \\ &= 2\int_{2\delta_N}^T \left(\int_{\{t-2\delta_N \leq s < t\}} + \int_{\{2\delta_N < s < t-2\delta_N\}}\right) E^{\xi_0^N}(h_{1,t}^N \cdot h_{1,s}^N) \, ds \, dt \qquad (6.11) \\ &= 2\int_{2\delta_N}^T \int_{t-2\delta_N}^t E^{\xi_0^N}(h_{1,t}^N \cdot h_{1,s}^N) \, ds \, dt \\ &\leq 2\int_{2\delta_N}^T \int_{t-2\delta_N}^t E^{\xi_0^N}[(h_{1,t}^N)^2]^{1/2} \cdot E^{\xi_0^N}[(h_{1,s}^N)^2]^{1/2} \, ds \, dt \\ &\leq \delta_N \cdot 2T \cdot C(K,T). \end{aligned}$$

Consequently, combine (6.9), (6.10) and (6.11), so that

$$\begin{aligned} E^{\xi_0^N}|I_3^N| &\leq \sum_{i=1}^3 E^{\xi_0^N}\left|\int_{2\delta_N}^T h_{i,t}^N \, dt\right| \\ &\leq \varepsilon_{4.1}^N \cdot T\Big(1 + 2Ke^{(b\mathcal{K})T} + \left[K^2 + 3KTe^{(b\mathcal{K})T}\right] \cdot e^{(2b\mathcal{K})T}\Big) \\ &\quad + \left[4b\delta_N + |\theta^N - \theta| \cdot T\right] \cdot Ke^{(b\mathcal{K})T} \\ &\quad + \left[\delta_N \cdot 2T \cdot C(K,T)\right]^{1/2} \to 0, \qquad \text{as } N \to \infty. \end{aligned}$$

Note that $|\theta^N - \theta| \to 0$ by Proposition 5.1 and (P2).

For the proof of (6.8), define

$$J^N_{1,T} = \int_0^T \left[\frac{1}{\tau_N} \upsilon_N(\xi^N_s) - \beta Z^N_s\right] ds,$$
$$J^N_{2,T} = \int_0^T \frac{1}{\tau_N^2}(d^{N,1}(\xi^N_s) + d^{N,0}(\xi^N_s))\ ds$$

so that we can write

$$\langle \mathcal{M}^N \rangle_T - \beta \int_0^T Z^N_s ds = J^N_{1,T} + J^N_{2,T}.$$

The way of showing that

$$E^{\xi^N_0}|J^N_{1,T}| \to 0 \quad \text{as } N \to \infty$$

is similar to the proof of (6.7). For showing that $E^{\xi^N_0}|J^N_{2,T}| \to 0$, we can follow the proof in (6.1) and apply Lemma 3.2(a) so that

$$\begin{aligned} E^{\xi^N_0}|J^N_{2,T}| &\leq \frac{2b\mathcal{K}C_{3.5}}{\tau_N} \cdot \int_0^T E^{\xi^N_0}(Z^N_s)\ ds \\ &\leq \frac{2b\mathcal{K}C_{3.5}}{\tau_N} \cdot T \cdot [Ke^{(2b\mathcal{K})T}] \to 0 \quad \text{as } N \to \infty. \end{aligned}$$

■

The next lemma says that $\{\mathcal{M}^N_t, t \leq T\}$ is uniformly bounded in L^2 which is implied by its bounded fourth moment.

Lemma 6.8. *Suppose that* $\sup_N Z^N_0 \leq K$. *Let* $T > 0$. *There is a constant* $C_{6.8}(K,T)$ *such that*

$$\sup_N E^{\xi^N_0}\left(\sup_{t\leq T} \langle \mathcal{M}^N \rangle_t^2\right) \leq C_{6.8}(K,T) \cdot T^2. \tag{6.12}$$

so that

$$\sup_N E^{\xi^N_0}\left(\sup_{t\leq T} |\mathcal{M}^N_t|^4\right) < \infty. \tag{6.13}$$

Proof. Fix $T > 0$. To see (6.12), we have that for any $t \le T$,

$$E^{\xi_0^N}[\langle \mathcal{M}^N \rangle_t^2] = E^{\xi_0^N}\left[\left(\int_0^t \frac{1}{\tau_N^2} m_2^N(\xi_s^N) ds\right)^2\right]$$
$$\le T^2 \cdot E^{\xi_0^N}\left[\sup_{s \le T}\left(\frac{1}{\tau_N^2} m_2^N(\xi_s^N)\right)^2\right].$$

By Lemma 6.3(b) and the coupling (3.3), for large N we have

$$\left(\frac{1}{\tau_N^2} m_2^N(\xi_s^N)\right)^2 \le 9(Z_s^N)^2 \le 9(Z_s^{N,b})^2.$$

Let $C_{6.8}(K,T) = 9 \cdot 4[K^2 + 3K(T \cdot e^{T(b\mathcal{K})})] \cdot e^{(2b\mathcal{K})T}$ so that by Corollary 3.2 (b) and Doob's inequality,

$$E^{\xi_0^N}\left[\sup_{s \le T}\left(\frac{1}{\tau_N^2} m_2^N(\xi_s^N)\right)^2\right] \le 9E^{\xi_0^N}\left[\sup_{s \le T}(Z_s^{N,b})^2\right] \le 9 \cdot 2^2 E^{\xi_0^N}((Z_T^{N,b})^2) \le C_{6.8}(K,T).$$

This implies

$$\sup_N E^{\xi_0^N}\left(\sup_{t \le T}\langle \mathcal{M}^N \rangle_t^2\right) \le T^2 \cdot \sup_N E^{\xi_0^N}\left[\sup_{s \le T}\left(\frac{1}{\tau_N^2} m_2^N(\xi_s^N)\right)^2\right] \le T^2 \cdot C_{6.8}(K,T).$$

We now prove (6.13). The jumps of $\mathcal{M}_\cdot^N$ are bounded by

$$|\mathcal{M}_t^N - \mathcal{M}_{t-}^N| \le |Z_t^N - Z_{t-}^N| + |\mathcal{D}_t^N - \mathcal{D}_{t-}^N|$$
$$= \frac{1}{\tau_N}\Big||\xi_t^N| - |\xi_{t-}^N|\Big| \le \frac{1}{\tau_N}, \; P^{\xi_0^N}\text{-a.s.} \tag{6.14}$$

Denote

$$\Delta \mathcal{M}_T^N = \mathcal{M}_T^N - \mathcal{M}_{T-}^N,$$
$$(\mathcal{M}^N)_T^* = \sup_{t \le T}|\mathcal{M}_t^N|,$$
$$(\Delta\mathcal{M}^N)_T^* = \sup_{t \le T}|\Delta\mathcal{M}_t^N|.$$

(6.14) implies

$$|(\Delta\mathcal{M}^N)_T^*| = \sup_{t \le T}|\mathcal{M}_t^N - \mathcal{M}_{t-}^N| \le \frac{1}{\tau_N}.$$

Therefore, by Theorem A.3 with $\phi(x) = x^4$, there is a constant $C = C_\phi$ such that

$$E^{\xi_0^N}\left(\sup_{t\le T}|\mathcal{M}_t^N|^4\right) \le C_\phi\left[E^{\xi_0^N}[\langle\mathcal{M}^N\rangle_T^2] + E^{\xi_0^N}\left(\left[(\Delta\mathcal{M}^N)_T^*\right]^4\right)\right]$$
$$\le C_\phi\left[T^2\cdot C_{6.8}(K,T)+1\right] < \infty.$$

∎

Proof of Proposition 6.1. It is enough to show that the quadruple $(Z_\cdot^N, \mathcal{D}_\cdot^N, \langle\mathcal{M}^N\rangle_\cdot, \mathcal{M}_\cdot^N)$ is C-tight.

(i) $\mathcal{D}_\cdot^N$ and $\langle\mathcal{M}^N\rangle_\cdot$: By Lemma 6.5 and Lemma 6.6, assumptions of Theorem VI.4.5 in [JS87] are satisfied so that by this theorem, $\mathcal{D}_\cdot^N$ and $\langle\mathcal{M}^N\rangle_\cdot$ are tight, and this implies that they are in fact C-tight since both of them are integral.

(ii) $\mathcal{M}_\cdot^N$: Since $\langle\mathcal{M}^N\rangle_\cdot$ is C-tight, then by Theorem VI.4.13 in [JS87], $\mathcal{M}_\cdot^N$ is tight. By Proposition VI.3.26 in [JS87] and (6.14), $\mathcal{M}_\cdot^N$ is C-tight.

(iii) $Z_\cdot^N$ is C-tight by (i) and (ii) following Corollary 2.2. ∎

Proof of Proposition 6.2. By Skorokhod's Theorem, Proposition 6.1 implies that there is a subsequence of laws P^{N_k} such that $P^{N_k} \Rightarrow P$ in $D([0,\infty),\mathbb{R}^+)$ (choose a further subsequence, if needed) and we may assume that on a common probability space,

$$(Z_\cdot^{N_k}, \mathcal{D}_\cdot^{N_k}, \langle\mathcal{M}^{N_k}\rangle_\cdot) \to (Z_\cdot, \mathcal{D}_\cdot, \mathcal{L}_\cdot) \quad \text{a.s.}$$

where $Z_\cdot, \mathcal{D}_\cdot, \langle\mathcal{M}\rangle_\cdot$ are continuous.

Lemma 6.7 implies that for any $\varepsilon > 0$,

$$P^{\xi_0^N}\left[\left|\mathcal{D}_t^N - \theta\int_0^t Z_s^N ds\right| > \varepsilon\right] \to 0,$$
$$P^{\xi_0^N}\left[\left|\langle\mathcal{M}^N\rangle_t - \beta\int_0^t Z_s^N ds\right| > \varepsilon\right] \to 0$$

as $N \to 0$, by Chebychev's inequality. By these probability estimates, it follows that

$$\mathcal{D}_t = \theta\int_0^t Z_s ds,$$
$$\mathcal{L}_t = \beta\int_0^t Z_s ds.$$

Proposition 2.1 implies that $\mathcal{M}^{N_k}_\cdot = Z^{N_k}_\cdot - \mathcal{D}^{N_k}_\cdot - Z^N_0$. By $Z^{N_k}_\cdot \to Z_\cdot$ a.s. and $\mathcal{D}^{N_k}_\cdot \to \mathcal{D}_\cdot$ a.s, we have $\mathcal{M}^{N_k}_\cdot \to \mathcal{M}_\cdot$ a.s. and $(Z_\cdot, \mathcal{M}_\cdot, \mathcal{D}_\cdot)$ satisfies

$$\begin{aligned} Z_t &= Z_0 + \mathcal{M}_t + \mathcal{D}_t \\ &= Z_0 + \mathcal{M}_t + \theta \int_0^t Z_s ds. \end{aligned}$$

Moreover, $\mathcal{M}_\cdot$ is continuous as $\mathcal{M}^N_\cdot$ is C-tight by Proposition 6.1.

Lastly, we show that $\mathcal{L}_\cdot$ is equal to $\langle \mathcal{M} \rangle_\cdot$. By Lemma 6.8, the sequence of martingales

$$\{(\mathcal{M}^{N_k}_t)^2 - \langle \mathcal{M}^{N_k} \rangle_t, t \leq T\}$$

is uniformly bounded in L^2 for every $T > 0$. This implies that it is an uniformly integrable family. Since for every $T > 0$ and $t \leq T$,

$$(\mathcal{M}^{N_k}_t)^2 - \langle \mathcal{M}^{N_k} \rangle_t \to \mathcal{M}^2_t - \mathcal{L}_t \text{ a.s.,} \quad \text{as } N_k \to \infty,$$

then it implies that $\{\mathcal{M}^2_t - \mathcal{L}_t,\ t \leq T\}$ is an L^2-martingale for every $T > 0$ and thus $\mathcal{L}_\cdot$ is the quadratic variation of $\mathcal{M}_\cdot$, in symbol,

$$\langle \mathcal{M} \rangle_t = \lim_k \langle \mathcal{M}^{N_k} \rangle_t = \beta \int_0^t Z_s ds.$$

■

Appendix

A.1 Proof of Proposition 1.1

Property (ii) and (iv) is implied by Lemma 2.1 in [LS10] and Theorem 6.3.2 in Durrett [D07], respectively. Property (i) is implied by Theorem II.4.24 in [H17].

For (iii), we prove the following simple fact. For any binomial random variable $X \sim Bin(m, p)$ with $m \in \mathbb{N}$ and $p \in (0, 1)$, $P(X \geq 1)$ can be bounded by

$$\begin{aligned} P(X \geq 1) &= \sum_{k=1}^{m} \binom{m}{k} p^k (1-p)^{m-k} \\ &= \sum_{k=0}^{m-1} \binom{m}{k+1} p^{k+1} (1-p)^{m-(k+1)} \\ &= \binom{m}{1} p \cdot \sum_{k=0}^{m-1} \binom{m}{1} p^k (1-p)^{m-k-1} \Big/ \binom{k+1}{k} \leq \binom{m}{1} p. \end{aligned} \tag{A.1}$$

We note the fact in [LS10] that $tx(B_{l_N}(x))$ is stochastically bounded above by the binomial random variable $R_{l_N} \sim Bin(r(r-1)^{l_N}, r(r-1)^{l_N}/N)$. Thus, by (A.1),

$$P(tx(B_{l_N}(x)) \neq 0) \leq P(R_{l_N} \geq 1) \leq \binom{r(r-1)^{l_N}}{1} \frac{r(r-1)^{l_N}}{N} \leq \frac{r^2}{N} \cdot N^{2/5}.$$

Define

$$Y(x) = \begin{cases} 1, & \text{if } tx(B_{l_N}(x)) \geq 1 \\ 0, & \text{otherwise} \end{cases}, \quad Y = \sum_x Y(x).$$

Notice that

$$EY = \sum_x E(Y(x)) = \sum_x P(tx(B_{l_N}(x)) \neq 0) \leq N \cdot \left(\frac{r^2}{N} \cdot N^{2/5}\right) = r^2 N^{2/5},$$

Thus by Markov's inequality we have

$$P(Y \geq r^2 N^{3/5}) \leq \frac{EY}{r^2 N^{3/5}} \leq \frac{N^{2/5}}{N^{3/5}} \to 0.$$

■

A.2 An elementary lemma for submartingales

Let S be countable. Suppose that $\mathcal{L}$ is a Markov generator and let the process

$$(X_t, (\mathcal{F}_t), (P^x)_{x \in S})$$

be defined by $\mathcal{L}$. Let f be a bounded function in the domain of $\mathcal{L}$ so that $Ef(X_t) < \infty$.

Lemma A.1. *If $\mathcal{L}f \geq 0$, then $f(X_t)$ is a submartingale. In particular, if $\mathcal{L}f = 0$, then $f(X_t)$ is a martingale.*

Proof. By Theorem I.5.2 in [L85],

$$M_t = f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)ds$$

is a P^{X_0}-martingale. For any $h > 0$, We have

$$\begin{aligned} E^x(f(X_{t+h}) - f(X_t)|\mathcal{F}_t) &= E^x(M_{t+h} - M_t|\mathcal{F}_t) + E^x\Big(\int_t^{t+h} \mathcal{L}f(X_s)ds\Big|\mathcal{F}_t\Big) \\ &= E^x\Big(\int_t^{t+h} \mathcal{L}f(X_s)ds\Big|\mathcal{F}_t\Big) \\ &= E^{X_t}\Big(\int_0^h \mathcal{L}f(X_s)ds\Big). \end{aligned}$$

Since $\mathcal{L}f \geq 0$, the last line is non-negative P^x-a.s. for any $x \in S$. ■

A.3 Meeting time lemma

The proof of Lemma 5.2 is based on the following frequently used result. The particular proof we provide here is due to Le Gall. The idea is that the walks have to stay at a place for a while before one of them leaves.

Lemma A.2. *Let P^{x_1,x_2} be the law of two independent rate 1 Markov chains X_t^1, X_t^2 with starting points x_1, x_2, and let M be the first meeting time of the chains. Then for any x,*

$$E^{x,x}\Big[\int_0^1 1_{\{X_s^1=X_s^2\}} ds\Big] \geq 1/e^2, \tag{A.2}$$

and for $t_1 < t_2$ and y_1, y_2,

$$P^{y_1,y_2}P(M \in (t_1,t_2]) \leq e^2 \int_{t_1}^{t_2+1} P^{y_1,y_2}(X_s^1 = X_s^2)\ ds. \tag{A.3}$$

Proof. Assume (A.2). Let $\mathcal{F}_M$ be the σ-algebra generated by M. We have

$$E^{y_1,y_2}\Big[\int_{t_1}^{t_2+1} 1_{\{X_s^1=X_s^2\}} ds \Big| \mathcal{F}_M\Big] \geq 1_{\{M\in(t_1,t_2]\}} \cdot E^{y_1,y_2}\Big[\int_{t_1}^{t_2+1} 1_{\{X_s^1=X_s^2\}} ds \Big| \mathcal{F}_M\Big]. \tag{A.4}$$

Notice that the right side of (A.4) is bounded below by

$$1_{\{M\in(t_1,t_2]\}} \cdot E^{y_1,y_2}\Big[\int_{M}^{M+1} 1_{\{X_s^1=X_s^2\}} ds \Big| \mathcal{F}_M\Big] = 1_{\{M\in(t_1,t_2]\}} \cdot E^{X_M^1,X_M^2}\Big[\int_0^1 1_{\{X_s^1=X_s^2\}} ds\Big]$$

by the strong Markov property. Therefore,

$$\begin{aligned}
&\int_{t_1}^{t_2+1} P^{y_1,y_2}(X_s^1 = X_s^2)\ ds \\
&\quad = E^{y_1,y_2}\Big(E^{y_1,y_2}\Big[\int_{t_1}^{t_2+1} 1_{\{X_s^1=X_s^2\}} ds \Big| \mathcal{F}_M\Big]\Big) \\
&\quad \geq E^{y_1,y_2}\Big(1_{\{M\in(t_1,t_2]\}} \cdot E^{X_M^1,X_M^2}\Big[\int_0^1 1_{\{X_s^1=X_s^2\}} ds\Big]\Big) \\
&\quad = \sum_x P(M \in (t_1,t_2], X_M^1 = X_M^2 = x) \cdot E^{x,x}\Big[\int_0^1 1_{\{X_s^1=X_s^2\}} ds\Big] \\
&\quad \geq P(M \in (t_1,t_2]) \cdot \frac{1}{e^2}.
\end{aligned}$$

To see that (A.2) is true, define

$$T_i = \inf\{t > 0 : X_t^i \neq x\},\ i = 1, 2.$$

As X_t^1 and X_t^2 are independent rate 1, then T_1 and T_2 are independent rate 1 exponential random variables. Since for any $s \in [0, 1)$,

$$\begin{aligned} P^{x,x}(X_s^1 = X_s^2,\ T_1 \wedge T_2 > 1) &= P^{x,x}(X_s^1 = X_s^2 | T_1 \wedge T_2 > 1) \cdot P^{x,x}(T_1 \wedge T_2 > 1) \\ &= 1 \cdot P^{x,x}(T_1 \wedge T_2 > 1). \end{aligned}$$

This implies

$$\begin{aligned} E^{x,x}\left[\int_0^1 1_{\{X_s^1 = X_s^2\}} ds\right] &\geq \int_0^1 P^{x,x}(X_s^1 = X_s^2,\ T_1 \wedge T_2 > 1) ds \\ &= P^{x,x}(T_1 \wedge T_2 > 1) \\ &= P^{x,x}(T_1 > 1,\ T_2 > 1) = 1/e^2. \end{aligned}$$

■

A.4 A continuous version of Burkholder's inequality

The next theorem states a continuous time version of Burkholder's inequality (Theorem 21.1 in [B73]). We follow Burkholder's proof in the original paper which adapted to this continuous time version.

Suppose that X_t is an L^2-cadlag martingale with $X_0 = 0$ and predictable square function $\langle X \rangle_t$. Furthermore, assume $\langle X \rangle_t$ is continuous. Denote

$$\Delta X_t = X_t - X_{t-},$$

$$X_t^* = \sup_{s \leq t} |X_s|,$$

$$\Delta X_t^* = \sup_{s \leq t} |\Delta X_s|.$$

Let $\phi : [0, \infty) \to \mathbb{R}$, continuous and non-decreasing, satisfies $\phi(0) = 0$ and the following growth condition: let $c = c_{6.1} > 0$ be the constant of (6.1) in [B73] such that

$$\phi(2x) \leq c\phi(x) \quad \text{for all } x > 0. \tag{A.5}$$

Two immediate facts are implied. First, since $\phi \geq 0$ and is non-decreasing,

$$\phi(a \vee b) \leq \phi(a) + \phi(b). \tag{A.6}$$

Second, for non-negative integers k,

$$\phi(2^k x) \leq c^k \phi(x). \tag{A.7}$$

Theorem A.3. *There is a constant $c = c_\phi > 0$ such that*

$$E\phi(X_t^*) \leq c_\phi \Big[E\phi(\langle X \rangle_t^{1/2}) + E\phi(\Delta X_t^*) \Big] \quad \textit{for any } t > 0. \tag{A.8}$$

We need the following two results to prove the theorem above.

Lemma A.4. *(Lemma 7.1 in [B73]) Suppose that f and g are non-negative random variables and $\beta > 1$, $\delta > 0$ and $\varepsilon > 0$ satisfy*

$$P(g > \beta\lambda,\ f \geq \delta\lambda) \leq \varepsilon P(g > \lambda) \textit{ for all } \lambda > 0.$$

Suppose $\gamma = \gamma(\beta) > 0$ and $\eta = \eta(\delta) > 0$ satisfy

$$\phi(\beta\lambda) \leq \gamma\phi(\lambda) \quad \textit{and} \quad \phi(\delta^{-1}\lambda) \leq \eta\phi(\lambda) \quad \textit{for all } \lambda > 0. \tag{A.9}$$

If in addition $\gamma\varepsilon < 1$, then

$$E\phi(g) \leq \frac{\gamma\eta}{1 - \gamma\varepsilon} E\phi(f).$$

Suppose that β and δ are given. Let k be a positive integer satisfying $2^{k-1} < \beta \leq 2^k$. Then one could choose $\gamma = c_{6.1}^k$, and η can be chosen according to δ in the same way.

Lemma A.5. *Assume $\beta > 1$ and $0 \le \delta < \beta - 1$. Then for every $T > 0$,*

$$P(X_T^* > \beta\lambda, \langle X\rangle_T^{1/2} \vee \Delta X_T^* \le \delta\lambda) \le \frac{\delta^2}{(\beta - 1 - \delta)^2} P(M_T^* > \lambda)$$

for all $\lambda > 0$.

Proof of Theorem A.3. In Lemma A.4, let $\beta = 2$. Then (A.5) is satisfied by taking $\gamma = c_{6.1}$. Choose $0 < \delta < (1/4) \wedge \gamma^{-1}$ so that $4\gamma\delta < 1$ and $\delta^{-1} > \gamma \vee 4 = c_{6.1} \vee 4$. Thus, one could take $\eta = 2^{j\vee 2}$ where j satisfies $2^{j-1} < c_{6.1} \le 2^j$. Now let

$$g = X_T^*, \quad f = \langle X\rangle_T^{1/2} \vee \Delta X_T^*,$$

Note that $\delta < 1/2$ so that $2\delta < 1$, thus

$$\varepsilon = \frac{\delta^2}{(\beta - 1 - \delta)^2} = \frac{\delta^2}{(1-\delta)^2} < 4\delta^2.$$

It follows that $\gamma\varepsilon < (4\gamma\delta)\delta < \delta < 1$. By Lemma A.4 and (A.6),

$$\begin{aligned} E\phi(X_T^*) &\le \frac{c_{6.1}\eta}{1 - c_{6.1}\varepsilon} E\phi(\langle X\rangle_T^{1/2} \vee \Delta X_T^*) \\ &\le \frac{c_{6.1}\eta}{1 - c_{6.1}\varepsilon}\Big[E\phi(\langle X\rangle_T^{1/2}) + E\phi(\Delta X_T^*)\Big]. \end{aligned}$$

Now we can choose

$$c_\phi = \frac{c_{6.1}\eta}{1 - c_{6.1}\varepsilon}$$

and this completes the proof of Theorem A.3. ■

Suppose $\phi(x) = x^4$. Then $\phi(2x) = 16\phi(x)$ so we can take $c_{6.1} = 16$ in (A.5). Now let

$$\beta = 2,$$

$$\gamma = c_{6.1} = 16,$$

$$\delta = (1/2)\gamma^{-1} = 2^{-5}$$

so that

$$\gamma\varepsilon = \frac{\gamma\delta^2}{(\beta - 1 - \delta)^2} = (1/2)\frac{\delta}{(1-\delta)^2} = \frac{2^{-6}}{(1 - 2^{-5})^2} = \frac{16}{31^2}.$$

By (A.7), we can take $\eta = c_{6.1}^5 = 16^5$ so that (A.9) is satisfied. Thus, we may take

$$c_\phi = \frac{\gamma\eta}{1-\gamma\varepsilon} = \frac{16^6}{1-16/31^2} = (16)^6\left(\frac{961}{945}\right)$$

in (A.8).

To prove Lemma A.5, we will make use of the following result.

Lemma A.6. *Let τ_1, τ_2 be stopping times such that $\tau_1 \leq \tau_2$ a.s., and let*

$$H_t = X_{t\wedge\tau_1} - X_{t\wedge\tau_2}.$$

Then H_t is an L^2-martingale with predictable square function $\langle X\rangle_{t\wedge\tau_1} - \langle X\rangle_{t\wedge\tau_2}$.

The lemma above in discrete time is an easy consequence of the fact that the sequence $1\{\tau \geq n\}$ is bounded and predictable for any stopping time τ, and hence a martingale transformation by it is again a martingale.

For continuous time setting, we use Proposition II.2.2 in [IW81]: in their notation, the stochastic integral

$$I^X(\psi)(t) = \int_0^t \psi \; dX_s$$

is defined for all processes $\psi = \psi(\omega, t) \in \mathcal{L}^2$ where $\mathcal{L}^2$ is defined in (1.1) of Section II.1, with respect to the right continuous martingale X that is square-integrable. Simply take the constant function $\psi = 1$ so that Lemma A.6 follows (2.3) of Proposition II.2.2 in [IW81]. ■

Proof of Lemma A.5. Let $T > 0, \beta > 1$ and $0 < \delta < \beta - 1$. For $\lambda > 0$, define

$$\tau(a) = \inf\{t \geq 0 : X_t^* > a\}, \qquad \text{for } a > 0,$$

$$\sigma = \inf\{t \geq 0 : \langle X\rangle_t^{1/2} \vee \Delta X_t^* > \delta\lambda\},$$

$$H_t^\lambda = X_{t\wedge\tau(\beta\lambda)\wedge\sigma} - X_{t\wedge\tau(\lambda)\wedge\sigma}.$$

By Lemma A.6, H_t^λ is an L^2-martingale with predictable square function

$$\langle H^\lambda\rangle_t = \langle X\rangle_{t\wedge\tau(\beta\lambda)\wedge\sigma} - \langle X\rangle_{t\wedge\tau(\lambda)\wedge\sigma}.$$

Claim 1. $P(X_T^* > \beta\lambda, \langle X\rangle_t^{1/2} \vee \Delta X_t^* \leq \delta\lambda) \leq P(\tau(\beta\lambda) \leq T, \sigma \geq T)$.

Proof. This follows directly from the definition of $\tau(\cdot)$: Notice that

$$\{X_T^* > \beta\lambda\} \subseteq \{\tau(\beta\lambda) \leq T\},$$

$$\{\langle X\rangle_T^{1/2} \vee \Delta X_T^* \leq \delta\lambda\} \subseteq \{\sigma \geq T\}.$$

Claim 2. $P(\tau(\beta\lambda) \leq T, \sigma \geq T) \leq P(H_T^* \geq \lambda(\beta - 1 - \delta))$.

Proof. On the event $\Gamma = \{\tau(\beta\lambda) \leq T, \sigma \geq T\}$,

$$H_t^\lambda = X_{t\wedge\tau(\beta\lambda)} - X_{t\wedge\tau(\lambda)}$$

and thus

$$|H_t^\lambda| \geq |X_{t\wedge\tau(\beta\lambda)}| - |X_{t\wedge\tau(\lambda)}| \geq |X_{t\wedge\tau(\beta\lambda)}| - \sup_{t\in[0,T]} |X_{t\wedge\tau(\lambda)}|.$$

On Γ, $\tau(\lambda) \leq T$ and $\Delta X_T^* \leq \delta\lambda$ which implies

$$\sup_{t\in[0,T]} |X_{t\wedge\tau(\lambda)}| \leq |X_{\tau(\lambda)}| \leq |X_{\tau(\lambda)-}| + |\Delta X_{\tau(\lambda)}^*| \leq \lambda + \delta\lambda.$$

Similarly since $\tau(\beta\lambda) \leq T$ on Γ,

$$\sup_{t\in[0,T]} |X_{t\wedge\tau(\beta\lambda)}| = |X_{\tau(\beta\lambda)}| \geq \beta\lambda.$$

Consequently,

$$|H_t^\lambda| \geq \lambda\beta - (\lambda + \delta\lambda) = \lambda(\beta - 1 - \delta).$$

Claim 3. $P((H^\lambda)_T^* > \lambda(\beta - 1 - \delta)) \leq \dfrac{1}{\lambda^2(\beta - 1 - \delta)^2} E(\langle H^\lambda\rangle_T)$.

Proof. Since H_T^λ is a martingale, then $(H_T^\lambda)^2$ is a submartingale and $E[(H_T^\lambda)^2] = E(\langle H^\lambda\rangle_T)$. Thus, the above claim follows from Doob's inequality (Theorem 1.4 in [CW90].)

Claim 4. $E(\langle H^\lambda\rangle_T) \leq (\delta\lambda)^2 P(X_T^* > \lambda)$.

Proof. Since $\tau(\lambda) \leq \tau(\beta\lambda)$,

$$\langle H^\lambda\rangle_T = \langle X\rangle_{T\wedge\tau(\beta\lambda)\wedge\sigma} - \langle X\rangle_{T\wedge\tau(\lambda)\wedge\sigma}$$

$$
\begin{aligned}
&= (\langle X\rangle_{T\wedge\tau(\beta\lambda)\wedge\sigma} - \langle X\rangle_{T\wedge\tau(\lambda)\wedge\sigma}) \cdot (1_{\{\tau(\lambda)\le T\}} + 1_{\{\tau(\lambda)>T\}}) \\
&= (\langle X\rangle_{T\wedge\tau(\beta\lambda)\wedge\sigma} - \langle X\rangle_{T\wedge\tau(\lambda)\wedge\sigma}) \cdot 1_{\{\tau(\lambda)\le T\}} \\
&\le \langle X\rangle_{T\wedge\tau(\beta\lambda)\wedge\sigma} \cdot 1_{\{\tau(\lambda)\le T\}} \\
&\le \langle X\rangle_{T\wedge\tau(\beta\lambda)\wedge\sigma} \cdot 1_{\{X_T^*\ge\lambda\}} \\
&\le (\delta\lambda)^2 \cdot 1_{\{X_T^*\ge\lambda\}}
\end{aligned}
$$

where the last inequality follows from the assumed continuity of $\langle X\rangle_t$. Consequently,

$$E(\langle H^\lambda\rangle_T) \le \delta^2\lambda^2 \cdot P(X_T^* \ge \lambda).$$

Since λ was arbitrary, this completes the proof of Lemma A.5. ■

Bibliography

[AS03] Abraams, D. M., Strogatz, S. H. (2003) *Modelling the dynamics of language death*. Nature 424, 900.

[ASD21] Agarwal, P., Simper, M., Durrett, R. (2021) *The q-voter model on the torus*. Electronic Journal of Probability 26.

[BG80] Bramson, M., Griffeath, D. (1980) *On the Willimas-Bjerknes tumor growth model, II.* Proc. Camb. Phil Soc. 88, 339-357.

[BG81] Bramson, M., Griffeath, D. (1981) *On the Willimas-Bjerknes tumor growth model, I.* Ann. Prob. 9, 173-185.

[B73] Burkholder, D. L. (1973) *Distribution Function Inequalities for Martingales*. Ann. Probab. 1(1): 19-42.

[CCC16] Chen, Y. T., Choi, J., Cox, J. T. (2016) *On the convergence of densities of finite voter models to the Wright-Fisher diffusion*. Ann. Inst. Henri Poincaré Probab. Stat. 52, 286–322.

[CW90] Chung, K. L., Williams, R. J. (1990) *Introduction to Stochastic Integration.* Birkhäuser, New York, NY.

[CS73] Clifford, P., Sudbury, A. (1973) *A model for spatial conflict*. Biometrika, 60, 581-588.

[C89] Cox, J. T. (1989) *Coalescing random walks and voter model consensus times on the torus in* $\mathbb{Z}^d$. The Annals of Probability, Vol. 17, No. 4, 1333-1366.

[C17] Cox, J. T. (2017) *Densities of Biased Voter Models on Finite Sets Converge to Feller's Branching Diffusion*. Markov Processes Relat. Fields 23, 421–444.

[CD16] Cox, J.T., and Durrett, R. (2016) *Evolutionary games on the torus with weak selection*. Stoch. Proc. Appl. 126, 2388–2409.

[CDP13] Cox, J. T., Durrett, R., Perkins, E. A. (2013) *Voter Model Perturbations and Reaction Diffusion Equations*. Astérisque, no. 349.

[D07] Durrett, R. (2007) *Random Graph Dynamics.* Cambridge University Press.

[D19] Durrett, R. (2019) *Probability: Theory and Examples.* 5th ed. Cambridge University Press.

[DFL16] Durrett, R., Foo, J., Leder, K. (2016) *Spatial Moran models, II: Cancer initiation in spatially structured tissue*. J. Math. Biol. 72 (5), 1369–1400.

[EK86] Ethier, S. N., Kurtz, T. G. (1986) *Markov Processes: Characterization and Convergence*. Wiley, New York.

[F51] Feller, W. (1951) *Diffusion processes in genetics*. Proc. Second Berkeley Symp. Math. Statist. Prob. 227-246. Univ. of California Press.

[H17] van der Hofstad, R. (2017) *Random Graphs and Complex Networks*. Cambridge University Press.

[HL75] Holley, R. A., Liggett, T. M. (1975) *Ergodic theorens for weakly interacting infinite systems and the voter model.* Ann. Probab. 3(4): 643-663,1985.

[IW81] Ikeda, N. , Watanabe, S. (1981) *Stochastic Differential Equations and Diffusion Processes*. 1st Ed.

[J69] Jiřina, M. (1969) *On Feller's branching diffusion processes.* Časopis pro pěstování matematiky, Vol. 94 , No. 1, 84–90.

[JS87] Jacod, J., Shiryaev, A. N. (1987) *Limit Theorems for Stochastic Processes.* Springer, New York.

[L85] Liggett, T. M. (1985) *Interacting Particle Systems*. Springer, New York.

[L99] Liggett, T. M. (1999) *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Grundlehren der mathematischen Wissenschaften (GL, volume 324).

[LS10] Lubetzky, E., Sly, A. (2010) *Cutoff phenomena for random walks on random regular graphs*. Duke Mathematical Journal.

[N99] Nettle, D. (1999) *Using Social Impact Theory to simulate language change*. Lingua, Vol. 108, Issues 2–3, 95-117.

[O11] Oliveira, R. I.(2011) *Mean field conditions for coalescing random walks*. Ann. Probab. 41(5): 3420-3461.

[P88a] Perkins, E. A. (1988) *A Space-Time Property of a Class of Measure-Valued Branching Diffusions*. Trans. Amer. Math. Soc. Vol.305, No. 2 pp. 743-795.

[S64] Spitzer, F. (1976) *Principles of Random Walk*. 1st ed. Springer New York, NY.

[S71] Spitzer, F. (1971) *Random Fields and Interacting Particle Systems*. Mathematical Association of America.

[SV79] Stroock, D. W., Varadhan, S. R. (1979) *Multidimensional Diffusion Processes*. Springer-Verlag Berlin Heidelberg.

[WB79] Williams, T., Bjerknes, R. (1972) *Stochastic model for abnormal clone spread through epithelial basal layer*. Nature 236, 19-21.

Vita

NAME

Tianyue Wu

EDUCATION

B.S., Mathematics, University of Washington - Seattle, May 2016.

M.S., Applied Statistics, Syracuse University, May 2019.