Glauber-Exclusion dynamics: rapid mixing regime[*]

Ryokichi Tanaka[†]   Kenkichi Tsunoda[‡]

Abstract

We show that for any attractive Glauber-Exclusion process on the one-dimensional lattice of size N with periodic boundary condition, if the corresponding hydrodynamic limit equation has a reaction term with a strictly convex potential, then the total-variation mixing time is of order $O(\log N)$. In particular, the result covers the full high-temperature regime in the original model introduced by De Masi, Ferrari and Lebowitz (1985).

1 Introduction

The Glauber-Exclusion dynamics is a stochastic spin system which is a superposition of a Glauber dynamics and a speeded-up symmetric simple exclusion process (SSEP) on lattices. The model was introduced by De Masi, Ferrari and Lebowitz in [DMFL85] and in [DMFL86], where they showed that the hydrodynamic scaling limit yields a reaction-diffusion equation (see also [KOV89] and [DMP91]). The family of limiting equations possibly interpolates dynamics of different natures in the macroscopic scale where the reaction term admits a single-well or a double-well potential. A salient feature of this model is that both types can arise even when the dynamics is constrained in the one-dimensional lattice. It seems plausible to expect, but still challenging to verify that the time needed for the stochastic model to reach an equilibrium would reflect the corresponding type in the hydrodynamic limit. The purpose of this article is to show

[†]Department of Mathematics, Kyoto University, Kyoto 606-8502 JAPAN.
  E-mail: rtanaka@math.kyoto-u.ac.jp
[‡]Faculty of Mathematics, Kyushu University, Fukuoka 819-0395, JAPAN.
  E-mail: tsunoda@math.kyushu-u.ac.jp

a rapid mixing property of the underlying stochastic model when it is hinted by the hydrodynamic equation. More precisely, we consider a general class of Glauber-Exclusion models on the one-dimensional integer lattice of size N with periodic boundary condition, and show that if the limiting reaction-diffusion equation has a strictly convex single well potential, then the total-variation mixing time is of order $O(\log N)$. In particular, this covers the full high-temperature regime in the original model of De Masi *et al*. [DMFL85, DMFL86].

For any integer $N \geq 1$, let $\mathbb{Z}_N := \mathbb{Z}/N\mathbb{Z}$ and $\Omega_N := \{-1, 1\}^{\mathbb{Z}_N}$. The generator of Glauber (spin flip) dynamics is defined by

$$L_G f(\eta) := \sum_{x \in \mathbb{Z}_N} c(x, \eta) \left(f(\eta^x) - f(\eta) \right) \quad \text{for } \eta \in \Omega_N,$$

where η^x is the configuration whose spin is $-\eta(x)$ at x and coincides with the spin of η at the other sites, and $c(x, \eta)$ is a positive value called *jump rate*. In the associated continuous-time Markov chain, the transition $\eta \to \eta^x$ occurs at rate $c(x, \eta)$. We assume that $c(x, \eta)$ depends only on $(\eta(y))_{|y-x| \leq K}$ for some $K \geq 0$ independent of N, is periodic, i.e., $c(x, \eta) = f(\eta_{\cdot + x})$ for some function f where $\eta_{\cdot + x} := (\eta(y + x))_{y \in \mathbb{Z}_N}$, and that there exists a constant $c_0 > 0$ independent of N such that $c(x, \eta) \geq c_0$ for all $(x, \eta) \in \mathbb{Z}_N \times \Omega_N$. The generator of SSEP is defined by

$$L_E f(\eta) := \frac{1}{2} \sum_{x \in \mathbb{Z}_N} \left(f(\eta^{x, x+1}) - f(\eta) \right) \quad \text{for } \eta \in \Omega_N,$$

where $\eta^{x, x+1}$ is the configuration whose spins are obtained by exchanging the spins of η at x and $x + 1$. Let us consider the sum of generators where the SSEP is speeded-up at rate N^2,

$$L_N := L_G + N^2 L_E,$$

and call the associated continuous-time Markov chain $\{\eta_t\}_{t \geq 0}$ on Ω_N a *Glauber-Exclusion* process [1]. Further if the jump rate $c(x, \eta)$ for the Glauber dynamics is attractive (see the definition in Section 2.1), then we call the corresponding processes an *attractive* Glauber-Exclusion process. Note that the Glauber-Exclusion process is irreducible on Ω_N since the jump rate in the Glauber part is positive, and thus there exists a unique stationary distribution π_N on Ω_N. We point out that except for a very special case the process is not reversible as it was observed by Gabrielli *et al*. [GJLLV97, Section 3] (see Remark 1.5), and no explicit description of the stationary distribution is available.

For any $\eta \in \Omega_N$, let $\mathbb{P}_\eta$ be the distribution of the process $\{\eta_t\}_{t \geq 0}$ starting from η. We define the *total-variation mixing time* for each $0 < \delta < 1$,

$$t_{\text{mix}}^N(\delta) := \inf \left\{ t \geq 0 \ : \ \max_{\eta \in \Omega_N} \|\mathbb{P}_\eta(\eta_t \in \cdot) - \pi_N\|_{\text{TV}} \leq \delta \right\},$$

where the total-variation distance is defined by

$$\|\mu - \nu\|_{\text{TV}} := \max_{A \subset \Omega_N} |\mu(A) - \nu(A)| = \frac{1}{2} \sum_{\eta \in \Omega_N} |\mu(\eta) - \nu(\eta)|,$$

for any pair of probability distributions μ, ν on Ω_N. For $\rho \in [-1, 1]$, let us define $R(\rho) := \mathbb{E}_{\nu_\rho}[-2\eta(0)c(0, \eta)]$ where ν_ρ is the product of Bernoulli measures on $\{-1, 1\}$ with mean ρ. The function $R : [-1, 1] \to \mathbb{R}$ is referred to as the reaction function and $R(\rho)$ as it appears in (1.2) is referred to as the reaction term. A function V is called a *potential* if $V'(\rho) = -R(\rho)$ for $\rho \in [-1, 1]$. We say that V is strictly convex if $V''(\rho) > 0$ on $[-1, 1]$.

---

[1]This is also called a Glauber+Kawasaki process in literature.

Theorem 1.1. *For any attractive Glauber-Exclusion process on $\mathbb{Z}_N$, if the corresponding reaction function R admits a strictly convex potential function V, then there exists a constant C such that for all $0 < \delta < 1$, we have*

$$t_{\mathrm{mix}}^N(\delta) \leq \frac{1}{\kappa} \log \frac{N}{\delta} + C,$$

for all large enough N, where $\kappa := \min_{\rho \in [-1,1]} V''(\rho) > 0$.

For example, let us consider the jump rate in the Glauber part introduced by De Masi *et al.* [DMFL85, DMFL86]: for $0 \leq \gamma < 1$,

$$c(x, \eta) := 1 - \gamma\eta(x)(\eta(x+1) + \eta(x-1)) + \gamma^2\eta(x+1)\eta(x-1), \tag{1.1}$$

for $(x, \eta) \in \mathbb{Z}_N \times \Omega_N$. This jump rate is positive and attractive for any $0 \leq \gamma < 1$. The reaction term $R(\rho) = \mathbb{E}_{\nu_\rho}[-2\eta(x)c(x,\eta)] = -2(1-2\gamma)\rho - 2\gamma^2\rho^3$ has a strictly convex potential function if $0 \leq \gamma < 1/2$. Therefore this range $0 \leq \gamma < 1/2$ may be considered as the high temperature regime analogous to the mean-field setting. Theorem 1.1 implies that for every $0 \leq \gamma < 1/2$ there exists a constant C such that for all $0 < \delta < 1$, we have

$$t_{\mathrm{mix}}^N(\delta) \leq \frac{1}{2(1-2\gamma)} \log \frac{N}{\delta} + C,$$

for all large enough N. However, we believe that the constant in front of $\log N$ in Theorem 1.1 would not be optimal and leave it open to find an exact constant which might lead to a cutoff particularly for the special case (1.1) by De Masi *et al.* For background and discussions on the cutoff phenomenon, see [LP17, Chapter 18].

Let us recall the result from the hydrodynamic limit which has been obtained by Kipnis, Olla and Varadhan [KOV89, Appendix] in the case of finite volume $\mathbb{Z}_N$ and by De Masi *et al.* [DMFL86] and [DMP91, Chapter VI] in the case of infinite volume $\mathbb{Z}$. For background information in the theory of hydrodynamic limit, see [KL99]; for recent progresses, see [BBP19] and references therein. For each fixed time $t \geq 0$, as $N \to \infty$ under an appropriate condition of convergence of the initial data, the density field

$$\mu_t^N(dx) = \frac{1}{N} \sum_{x \in \mathbb{Z}_N} \eta_t(x)\delta_{\frac{x}{N}}(dx) \quad \text{on } \mathbb{R}/\mathbb{Z},$$

converges weakly to a unique weak solution of the reaction-diffusion equation:

$$\frac{\partial\rho}{\partial t} = \frac{1}{2}\frac{\partial^2\rho}{\partial x^2} + R(\rho), \tag{1.2}$$

where $R(\rho)$ is the reaction term defined just before the statement of Theorem 1.1. The following is a reason why the reaction term is obtained by the average under the Bernoulli measure: for a fixed time $T > 0$ since the SSEP is speeded-up by N^2 whereas the spin-flip dynamics occurs slowly $O(N)$ in a unit time, in each (macroscopically small) box of size εN for $\varepsilon > 0$ the distribution at time T is approximately the Bernoulli measure whose parameter is the empirical average of spins in the box. There one should be able to replace the distribution by the Bernoulli at each fixed time T — the proof of deriving (1.2) is basically devoted to verify this heuristic rigorously. Let us point out that this has been verified only for a fixed time T, which is not enough to understand the full dynamics up to the equilibrium. One observes that a combination of (microscopically) slow reaction with a fast stirring creates a non-local interaction among spins even though the Glauber updates depend only on configurations in bounded ranges (e.g., nearest-neighbors). Because of this particular nature of the dynamics our analysis would not simply boil down to either the Glauber dynamics or the SSEP, where much

sharper results have been established for the total-variation mixing time in [LS16] and [Lac16, Lac17], respectively. Concerning the Glauber-Exclusion process, very little is known for the stationary distribution, which has attracted intensive studies recently (e.g., [FLT19]). Although it is not always necessary to have full understanding of the stationary distribution (e.g., a sharp result on the mixing time has been established for a large class of noisy voter models in [CPS16]), it had been unknown that the mixing time was actually $O(\log N)$ in the present setting. Given the result on the hydrodynamic limit, our result would not be valid beyond the strict convexity condition on potentials: even for the particular example by De Masi *et al.* (1.1), it is tempting to determine the exact order of total-variation mixing time in the range $1/2 \leq \gamma < 1$.

Our proof is based on an enhancement of the above replacement which is now valid in time range up to N^ε for a small enough $\varepsilon > 0$ with a uniform error control. We call it a *replacement lemma* (see the precise statement in Lemma 3.3); the proof uses a classical coupling between a SSEP and independent simple random walks and it sharpens estimates in [DMFL86]. The idea has a lot of similarities to the "v-function" method in [DMP91, Chapter IX], however our proof is self-contained and enables us to expose an explicit constant factor in the upper bound if not optimal. We use this estimate under the standard monotone coupling process in order to control error terms. After obtaining a recursive estimate with a fine error control, we find a coalescing time of the coupling process and this leads an upper bound of the total-variation mixing time (Section 2). In the entire part of this paper, we focus on the one-dimensional lattice with periodic boundary condition, but this is mainly for simplicity of notations; the methods we employ are valid in the higher dimensional lattices.

Let us give some other explicit examples to which our result applies.

Example 1.2 (The De Masi-Ferrari-Lebowitz model and its variants)**.** Brassesco *et al.* have considered the following jump rate with external field [BPSV00b] (compare with [BPSV00a]):

$$c_\mu(x, \eta) := c(x, \eta) - \frac{\mu}{2}\eta(x),$$

where $c(x, \eta)$ is the jump rate of De Masi *et al.* (1.1), and μ is a real value. If $0 \leq \mu < 2(1 - \gamma)^2$, then $c_\mu(x, \eta) > 0$. Note that introducing the parameter μ breaks the flip symmetry $\eta \leftrightarrow -\eta$ of the stationary distribution and still keeps the jump rate attractive. The reaction term has the following form: $R(\rho) = \mathbb{E}_{\nu_\rho}[-2\eta(x)c(x, \eta)] = -2(1 - 2\gamma)\rho - 2\gamma^2\rho^3 + \mu$, which has a strictly convex potential function if $0 \leq \gamma < 1/2$. (It should be noted however that the main purpose to introduce μ is to understand the dynamics when $\gamma > 1/2$ in [BPSV00b].) Concerning a more general class of processes in the one-dimensional lattice, see [Fer90, Section 4]. [2]

Remark 1.3. For any $d \geq 1$, one is able to generalize the process on the d-dimensional lattice by setting the jump rate

$$c(x, \eta) = \frac{1}{d}\sum_{i=1}^{d}\left(1 - \gamma\eta(x)\left(\eta(x - e_i) + \eta(x + e_i)\right) + \gamma^2\eta(x - e_i)\eta(x + e_i)\right),$$

where $\{e_i\}_{i=1,\dots,d}$ is the standard basis in $\mathbb{R}^d$.

Example 1.4 (The Chafee-Infante model)**.** This model has been analyzed in [FLT19, Section 8]. Let us fix $a_0, a_1, a_2 > 0$, and define

$$c(x, \eta) := a_0\mathbf{1}_{\{\eta(x-1)=\eta(x+1)\neq\eta(x)\}} + a_1\mathbf{1}_{\{\eta(x-1)=\eta(x)=\eta(x+1)\}} + a_2\mathbf{1}_{\{\eta(x-1)\neq\eta(x+1)\}}.$$

---

[2]In the same paper, there is a typographical error where the Glauber-Exclusion process with the jump rate (1.1) is analyzed; it is in fact $\gamma < 1/3$ instead of $\gamma \leq 1/2$ [Fer90, p.1528, line 22].

Note that this jump rate is positive and attractive if $a_0 \geq a_2 \geq a_1$. A computation yields

$$\mathbb{E}_{\nu_\rho}[-2\eta(x)c(x,\eta)] = \frac{1}{2}\left((a_0 - 3a_1 - 2a_2)\rho - (a_0 + a_1 - 2a_2)\rho^3\right).$$

Fix $a, b > 0$. If we define $a_1 := a$, $a_2 := a + 2b$ and $a_0 := a + 8b$, then the jump rate is positive and monotone and we have

$$a_0 - 3a_1 - 2a_2 = 4(b - a) \qquad \text{and} \qquad a_0 + a_1 - 2a_2 = 4b.$$

Hence

$$R(\rho) = \mathbb{E}_{\nu_\rho}[-2\eta(x)c(x,\eta)] = 2(b - a)\rho - 2b\rho^3,$$

where $R(\rho)$ has a strictly convex potential function if $b < a$.

Remark 1.5. In general, an explicit form of the stationary distribution π_N is not available even in the setting which we are focusing on in this article. Gabrielli *et al.* have shown that the Glauber-Exclusion process on periodic integer lattices is reversible if and only if the jump rate has the form

$$c(x,\eta) = (a_1 + a_2\eta(x))h(x,\eta),$$

where $h(x,\eta)$ is independent of $\eta(x)$ with some constants a_1, a_2, and in this particular case π_N is a Bernoulli measure [GJLLV97, Section 3].

The rest of this article is organized as follows: in Section 2 we prove Theorem 1.1 provided Lemma 2.2 whose full proof is deferred until subsequent sections, in Section 3 we prove Lemma 2.2 by using a key replacement lemma Lemma 3.3, in Section 4 we prove Lemma 3.3 and complete the proof of upper bound for the mixing time, and in Appendix A and B we give the proofs of technical lemmas which we use in the proof of Lemma 3.3 in Section 4.

Notation: Throughout the article, we use $C, C', C'', \ldots$ to denote numerical constants whose exact values may change from line to line, and $C_\varepsilon, C'_\varepsilon, C''_\varepsilon, \ldots$ to indicate its dependency only on ε for a parameter ε. For a set A, we denote the indicator function of A and the cardinality of A by $\mathbf{1}_A$ and $|A|$, respectively. When A is a Lebesgue measurable subset of $\mathbb{R}$, we also denote the normalized Lebesgue measure of A in $\mathbb{R}$ by $|A|$, for which we believe there is no danger of confusion from the context. For a real-valued function $f(N)$ in positive integers N, we write $f(N) = O(N)$ if there exists a constant $C \geq 0$ such that $|f(N)| \leq CN$ for all large enough N, and $f(N) = o(1)$ if $|f(N)| \to 0$ as $N \to \infty$. For any real number x, we denote by $\lfloor x \rfloor$ the largest integer at most x and by $\lceil x \rceil$ the smallest integer at least x.

2 Proof of Theorem 1.1

2.1 Monotone coupling

We define a partial order in Ω_N by setting

$$\eta \geq \zeta \iff \eta(x) \geq \zeta(x) \quad \text{for any } x \in \mathbb{Z}_N.$$

A jump rate $c(x,\eta)$ $(x \in \mathbb{Z}_N, \eta \in \Omega_N)$ is called *attractive* if $\eta \geq \zeta$ implies that

$$c(x,\eta) \leq c(x,\zeta) \text{ for } \eta(x) = \zeta(x) = 1, \quad \text{and} \quad c(x,\eta) \geq c(x,\zeta) \text{ for } \eta(x) = \zeta(x) = -1.$$

An attractive jump rate yields a monotone coupling of the associated chains (η_t^+, η_t^-), i.e., a coupling where the order of configurations are kept preserved $\eta_t^+ \geq \eta_t^-$ for all $t \geq 0$ almost surely if they have started with $\eta_0^+ \geq \eta_0^-$ [Lig10, Theorem 4.11, p.143]. We

construct a monotone coupling (η_t^+, η_t^-) for the Glauber-Exclusion process based on this coupling, for which we provide the explicit form for the sake of convenience: given $\eta \geq \zeta$ (otherwise we define the jump rate 0 from (η, ζ)), for each $x \in \mathbb{Z}_N$,

$$(\eta, \zeta) \mapsto (\eta^{x,x+1}, \zeta^{x,x+1}) \quad \text{at rate } \frac{N^2}{2},$$

and the states of (η, ζ) at x are updated according to the monotone coupling of Glauber dynamics with the same (attractive) jump rate, namely, we change the states of (η, ζ) at x by the following rates;

$$(\eta(x), \zeta(x)) = (-1, -1) \implies \begin{cases} (1, 1) & \text{with rate } c(x, \zeta) \\ (1, -1) & \text{with rate } c(x, \eta) - c(x, \zeta), \end{cases}$$

$$(\eta(x), \zeta(x)) = (1, -1) \implies \begin{cases} (-1, -1) & \text{with rate } c(x, \eta) \\ (1, 1) & \text{with rate } c(x, \zeta), \end{cases}$$

$$(\eta(x), \zeta(x)) = (1, 1) \implies \begin{cases} (-1, -1) & \text{with rate } c(x, \eta) \\ (1, -1) & \text{with rate } c(x, \zeta) - c(x, \eta). \end{cases}$$

The joint process $\{(\eta_t^+, \eta_t^-)\}_{t \geq 0}$ yields a Markovian coupling, whose distribution we denote by the same symbol $\mathbb{P}$ as long as there is no danger of confusion. If we define

$$\tau := \inf\{t \geq 0 \ : \ \eta_t^+ = \eta_t^-\},$$

then it holds that $\tau < \infty$ and $\eta_t^+ = \eta_t^-$ for all $t \geq \tau$ almost surely if $\eta_0^+ \geq \eta_0^-$ since the chain is irreducible. Let $\mathbf{1}$ and $-\mathbf{1}$ be the configurations whose states are all 1 and all -1, respectively.

Lemma 2.1. *For any attractive Glauber-Exclusion process, if the corresponding reaction term $R(\rho)$ has a potential $V(\rho)$, i.e., $R(\rho) = -V'(\rho)$ such that*

$$\kappa := \min_{\rho \in [-1,1]} V''(\rho) > 0,$$

then for the monotone coupling (η_t^+, η_t^-) with the initial configurations $(\mathbf{1}, -\mathbf{1})$, there exists a constant $C > 0$ such that for any $0 < \delta < 1$ and

$$t_{N,\delta} := \frac{1}{\kappa} \log \frac{N}{\delta} + C,$$

we have $\mathbb{P}(\tau > t_{N,\delta}) \leq \delta$ for all large enough N.

We show Lemma 2.1 in Section 2.2.

Proof of Theorem 1.1. Let $\eta, \zeta \in \Omega_N$ be two arbitrary configurations (not necessarily $\eta \geq \zeta$). We consider the monotone coupling $\{(\eta_t^+, \eta_t^-)\}_{t \geq 0}$ with initial configurations $(\mathbf{1}, -\mathbf{1})$. Note that $\mathbf{1} \geq \eta, \zeta \geq -\mathbf{1}$; enlarging the probability space, we construct a coupling $\mathbb{P}$ (denoted by the same symbol) among four copies of original chains $\eta_t^+, \eta_t^-, \eta_t, \zeta_t$ such that

$$\eta_t^+ \geq \eta_t, \zeta_t \geq \eta_t^- \quad \text{for all } t \geq 0 \text{ almost surely,}$$

with $\eta_0 = \eta$ and $\zeta_0 = \zeta$, and further any two of η_t^+, η_t and η_t^- (resp. η_t^+, ζ_t and η_t^-) are monotone coupling. Then $\eta_t^+ = \eta_t^-$ implies that $\eta_t = \zeta_t$ almost surely, and thus letting

$$\tau := \inf\{t \geq 0 \ : \ \eta_t^+ = \eta_t^-\},$$

we have that

$$\|\mathbb{P}_\eta(\eta_t \in \cdot) - \mathbb{P}_\zeta(\zeta_t \in \cdot)\|_{\mathrm{TV}} \leq \mathbb{P}(\eta_t \neq \zeta_t) \leq \mathbb{P}(\eta_t^+ \neq \eta_t^-) \leq \mathbb{P}(\tau > t).$$

By Lemma 2.1 for any $0 < \delta < 1$ if $t_{N,\delta} = (1/\kappa) \log(N/\delta) + C$ for some constant C, then $\mathbb{P}(\tau > t_{N,\delta}) \leq \delta$. Noting that

$$\max_{\eta \in \Omega_N} \|\mathbb{P}_\eta(\eta_t \in \cdot\,) - \pi_N\|_{\mathrm{TV}} \leq \max_{\eta, \zeta \in \Omega_N} \|\mathbb{P}_\eta(\eta_t \in \cdot\,) - \mathbb{P}_\zeta(\zeta_t \in \cdot\,)\|_{\mathrm{TV}},$$

we obtain

$$t_{\mathrm{mix}}^N(\delta) \leq \frac{1}{\kappa} \log \frac{N}{\delta} + C,$$

for all large enough N and conclude the proof. $\qquad\square$

2.2 The upper bound

For a continuous-time Markov chain $\{\eta_t\}_{t \geq 0}$ generated by L_N on the state space $\Omega_N = \{-1, 1\}^{\mathbb{Z}_N}$, we define the *normalized magnetization* by

$$S(\eta_t) := \frac{1}{N} \sum_{x \in \mathbb{Z}_N} \eta_t(x).$$

We say that $f : \{-1, 1\}^{\mathbb{Z}_N} \to \mathbb{R}$ is a *local function* if there exists an integer $0 \leq K < N$ independent of N and a subset I in $\mathbb{Z}_N$ with $|I| \leq K$ such that $f(\eta)$ depends only on $\{\eta_x\}_{x \in I}$ for any $\eta \in \{-1, 1\}^{\mathbb{Z}_N}$. Let us denote by $\mathrm{supp}\, f$ the smallest such subset I for a local function f. For example, $f(\eta) := \eta(-1)\eta(0)\eta(1)$ for $\eta \in \{-1, 1\}^{\mathbb{Z}_N}$ is a local function with $\mathrm{supp}\, f = \{-1, 0, 1\}$. Let $f_\emptyset(\eta) = 1$. For a non-empty subset I in $\mathbb{Z}_N$, let $f_I(\eta) := \prod_{x \in I} \eta(x)$ and call it an *elementary local function*. Note that any local function f is a linear combination of elementary local functions:

$$f(\eta) = \sum_{I \subset \mathrm{supp}\, f} a_I f_I(\eta) \quad \text{for } \eta \in \Omega_N,$$

where the summation is over all subsets in $\mathrm{supp}\, f$ and $a_I \in \mathbb{R}$ and this follows from the Fourier expansion of f on $\{-1, 1\}^K$.

For $\eta \in \Omega_N$ and for $x \in \mathbb{Z}_N$, denoting $L_N f_x(\eta)$ where $f_x(\eta) := \eta(x)$ simply by $(L_N \eta)(x)$, we have

$$(L_N \eta)(x) = -2\eta(x)c(x, \eta) + \frac{N^2}{2}\left(\eta(x+1) + \eta(x-1) - 2\eta(x)\right),$$

and thus

$$L_N S(\eta) = \frac{1}{N} \sum_{x \in \mathbb{Z}_N} -2\eta(x)c(x, \eta).$$

Since for a given jump rate $c(x, \eta)$, the function $\eta(0)c(0, \eta)$ is local, expanding it by elementary local functions in the way above, we have

$$2\eta(0)c(0, \eta) = \sum_{y \in \mathbb{Z}_N} a_y f_y(\eta) + g(\eta) \quad \text{where } g(\eta) := \sum_{|I| \neq 1} a_I f_I.$$

Note that $\eta(x)c(x, \eta) = \eta_{\cdot+x}(0)c(0, \eta_{\cdot+x})$ and that $\frac{1}{N}\sum_{y \in \mathbb{Z}_N} f_y(\eta_{\cdot+x}) = S(\eta)$ to obtain the following decomposition:

$$L_N S(\eta) = -\alpha S(\eta) - F(\eta) \quad \text{where } \alpha = \sum_{y \in \mathbb{Z}_N} a_y \text{ and } F(\eta) := \frac{1}{N} \sum_{x \in \mathbb{Z}_N} g(\eta_{\cdot+x}).$$

Note that the reaction term $R(\rho)$ is obtained from $L_N S(\eta)$ by

$$R(\rho) = \mathbb{E}_{\nu_\rho}[L_N S(\eta)].$$

Letting $G(\rho) := \mathbb{E}_{\nu_\rho} F(\eta)$, one has $R(\rho) = -\alpha\rho - G(\rho)$. If $V(\rho)$ is a potential function of $R(\rho)$, i.e., $V'(\rho) = -R(\rho)$, then we have that $V''(\rho) = \alpha + G'(\rho)$.

For a pair of configurations η_0^+, η_0^- such that $\eta_0^+ \geq \eta_0^-$, we run the monotone coupling $\{(\eta_t^+, \eta_t^-)\}_{t \geq 0}$ and denote by $\mathcal{F}_t := \sigma(\{\eta_s^+, \eta_s^-\}_{0 \leq s \leq t})$ the σ-algebra generated by the monotone coupling process up to time t for $t \geq 0$. Letting

$$\xi_t := S(\eta_t^+) - S(\eta_t^-)$$

for $t \geq 0$, we have that

$$\mathbb{E}\left[\xi_{t+T} \mid \mathcal{F}_t\right] = \xi_t - \alpha \int_t^{t+T} \mathbb{E}\left[\xi_s \mid \mathcal{F}_t\right] ds - \int_t^{t+T} \mathbb{E}\left[F(\eta_s^+) - F(\eta_s^-) \mid \mathcal{F}_t\right] ds, \qquad (2.1)$$

where $\mathbb{E}\left[\cdot \mid \mathcal{F}_t\right]$ stands for the conditional expectation with respect to the σ-algebra $\mathcal{F}_t$. For each $x \in \mathbb{Z}_N$, we define the *local average* of $\eta \in \Omega_N$ around x up to time t by letting $\{w(t)\}_{t \geq 0}$ be the simple random walk (SRW) starting from x with rate 1,

$$\Phi_x(\eta, t) := \mathbf{E}_x \eta(w(t)) \quad \text{for } \eta \in \Omega_N \text{ and } t \geq 0,$$

where $\mathbf{E}_x$ denotes the expectation with respect to the distribution of $\{w(t)\}_{t \geq 0}$ such that $w(0) = x$.

Lemma 2.2 (Main replacement lemma). *For all small enough $\varepsilon > 0$, there exist constants $C > 0$ and T_0 such that for all $T_0/N^2 \leq t \leq 1/N^{1-\varepsilon}$ and all $\widetilde{\eta} = (\eta^+, \eta^-)$ with $\eta^+ \geq \eta^-$, we have*

$$\left| \mathbb{E}_{\widetilde{\eta}}[F(\eta_t^+) - F(\eta_t^-)] - \frac{1}{N} \sum_{x \in \mathbb{Z}_N} [G(\Phi_x(\eta^+, N^2 t)) - G(\Phi_x(\eta^-, N^2 t))] \right|$$

$$\leq \frac{C}{t^{\varepsilon/4} N^{2\varepsilon}} \xi_0 + C \int_0^t \mathbb{E}\, \xi_s \, ds + C \exp(-(N^2 t)^{\varepsilon/8}).$$

We show Lemma 2.2 in Section 3 and now prove Lemma 2.1. We note that the proof of Lemma 3.3 reveals that $0 < \varepsilon < 1/28$ is sufficient for the argument. The proof of Lemma 2.1 is based on showing a strong contraction (see (2.2)) between two coupled magnetization chains. This is achieved by dealing with the function F, which is a main 'non-linear' term in this discrete setting.

Proof of Lemma 2.1. Fix $0 < \varepsilon < 1/28$ and let $T_* := 1/N^{1-\varepsilon}$. We first claim that for all $t \geq 0$,

$$\mathbb{E}\, \xi_{t+T_*} \leq \left(1 - \frac{\kappa}{N^{1-\varepsilon}} + \frac{C_\varepsilon}{N^{1+\varepsilon^2}}\right) \mathbb{E}\, \xi_t + \frac{C_\varepsilon'}{N^2}. \qquad (2.2)$$

We prove this inequality later and let us conclude Lemma 2.1 by invoking this estimate. For any $0 < \delta < 1$, if we define

$$t_{N,\delta} := \frac{1}{\kappa}\left(\log N + \log\frac{4}{\delta} + C_\varepsilon'\right),$$

then applying (2.2) inductively $\lfloor t_{N,\delta}/T_* \rfloor$-times and using the monotone coupling $\mathbb{E}\, \xi_t \leq \mathbb{E}\, \xi_s$ for $t \geq s$ yield

$$\mathbb{E}\, \xi_{t_{N,\delta}} \leq \left(1 - \frac{\kappa}{N^{1-\varepsilon}} + \frac{C_\varepsilon}{N^{1+\varepsilon^2}}\right)^{\lfloor t_{N,\delta}/T_* \rfloor} \mathbb{E}\, \xi_0 + \frac{C_\varepsilon'}{N^2} \sum_{n=0}^{\lfloor t_{N,\delta}/T_* \rfloor - 1} \left(1 - \frac{\kappa}{N^{1-\varepsilon}} + \frac{C_\varepsilon}{N^{1+\varepsilon^2}}\right)^n$$

$$\leq \xi_0 \exp\left(-\kappa t_{N,\delta} + C_\varepsilon'\right) + N^{1-\varepsilon}\frac{C_\varepsilon''}{N^2} \leq \frac{\delta}{2N} + \frac{C_\varepsilon''}{N^{1+\varepsilon}},$$

for all large enough N, where we have used $\xi_0 \leq 2$ in the last inequality. Recalling that $\tau := \inf\{t \geq 0 : \xi_t = 0\}$, we obtain by the Markov inequality,

$$\mathbb{P}(\tau > t_{N,\delta}) = \mathbb{P}(\xi_{t_{N,\delta}} \geq 1/N) \leq N \cdot \mathbb{E}\,\xi_{t_{N,\delta}} \leq \frac{\delta}{2} + \frac{C_\varepsilon''}{N^\varepsilon} < \delta,$$

for all large enough N, as required.

Let us turn to the claim (2.2). We run the monotone coupling with the initial state $(\mathbf{1}, -\mathbf{1})$. Note that for all $s, t > 0$,

$$\left|\int_s^{s+t} \mathbb{E}\left[\xi_u \mid \mathcal{F}_s\right] du - t\xi_s\right| \leq Ct^2. \tag{2.3}$$

Indeed, since the SSEP dynamics given by L_E preserves $S(\eta)$, the difference $|\xi_u - \xi_s|$ is at most $2/N$ times the number of Glauber updates within the time interval $[s, u]$ for $0 \leq s < u$, which is stochastically dominated by the Poisson random variable with intensity $O(N(u - s))$. Integrating in u,

$$\mathbb{E}\left[\xi_u - \xi_s \mid \mathcal{F}_s\right] \leq \mathbb{E}\left[|\xi_u - \xi_s| \mid \mathcal{F}_s\right] \leq C(u - s),$$

we obtain (2.3). For each $t \geq 0$, we integrate

$$\mathbb{E}\left[F(\eta_s^+) - F(\eta_s^-) \mid \mathcal{F}_t\right] \quad \text{for } t \leq s \leq t + T_*,$$

where we decompose the time interval for the integration into two parts; for the first part $t \leq s \leq t + T_0/N^2$, we use the following elementary estimate

$$\left|\int_t^{t + \frac{T_0}{N^2}} \mathbb{E}\left[F(\eta_s^+) - F(\eta_s^-) \mid \mathcal{F}_t\right] ds\right| \leq \frac{2\|g\|_\infty T_0}{N^2},$$

(for the definition of g, see the third paragraph in Section 2.2) and for the second part $t + T_0/N^2 \leq s \leq T_*$, we will apply to the Markov property and Lemma 2.2. By (2.3), it follows that

$$\int_t^{t+s} \mathbb{E}\left[\xi_u \mid \mathcal{F}_t\right] du \leq s\xi_t + Cs^2,$$

which we integrate over $t \leq s \leq t + T_*$ (and we note that this factor comes from the second term in the right hand side in the inequality in Lemma 2.2). Direct computations yield

$$\int_0^{T_*} \left(s + \frac{1}{s^{\varepsilon/4} N^{2\varepsilon}}\right) ds = \frac{1}{2} T_*^2 + \frac{1}{(1 - \varepsilon/4) N^{2\varepsilon}} T_*^{1 - \varepsilon/4} \leq \frac{C_\varepsilon}{N^{1 + \varepsilon^2}},$$

(which is a factor in front of ξ_t; we have used $(1 - \varepsilon)(1 - \varepsilon/4) + 2\varepsilon \geq 1 + \varepsilon^2$), and

$$\int_0^{T_*} \left(s^2 + \exp\left(-(N^2 s)^{\varepsilon/8}\right)\right) ds \leq \frac{1}{3} T_*^3 + \frac{1}{N^2} \int_0^\infty \exp\left(-s^{\varepsilon/8}\right) ds \leq \frac{C_\varepsilon'}{N^2},$$

(which is a remaining error term), combining above estimates, we have that

$$\int_t^{t+T_*} \mathbb{E}\left[F(\eta_s^+) - F(\eta_s^-) \mid \mathcal{F}_t\right] ds \geq -\frac{C_\varepsilon''}{N^2} - \frac{C_\varepsilon}{N^{1+\varepsilon^2}} \xi_t$$
$$+ \int_0^{T_*} \frac{1}{N} \sum_{x \in \mathbb{Z}_N} \left[G(\Phi_x(\eta_t^+, N^2 s)) - G(\Phi_x(\eta_t^-, N^2 s))\right] ds,$$

$$\tag{2.4}$$

where we have used $\xi_t \geq 0$, the Markov property and Lemma 2.2. For any pairs of configurations with $\eta^+ \geq \eta^-$, it holds that for $s \geq 0$,

$$\Phi_x(\eta^+, N^2 s) \geq \Phi_x(\eta^-, N^2 s).$$

Recall that a potential function $V(\rho)$ of $R(\rho)$ satisfies that $V''(\rho) = \alpha + G'(\rho)$. Letting $\kappa := \min_{\rho \in [-1,1]} V''(\rho) > 0$, we have that by the Taylor theorem for $\rho^+ \geq \rho^-$,

$$G(\rho^+) - G(\rho^-) = -\alpha(\rho^+ - \rho^-) + (V'(\rho^+) - V'(\rho^-)) \geq (-\alpha + \kappa)(\rho^+ - \rho^-).$$

Plugging into this inequality $\Phi_x(\eta_t, N^2 s)$ and noting that

$$\frac{1}{N} \sum_{x \in \mathbb{Z}_N} \Phi_x(\eta_t, N^2 s) = \frac{1}{N} \sum_{x \in \mathbb{Z}_N} \mathbf{E}_x \eta_t(w(N^2 s)) = S(\eta_t),$$

we obtain by averaging over all sites $x \in \mathbb{Z}_N$, for all $s \geq 0$,

$$\frac{1}{N} \sum_{x \in \mathbb{Z}_N} [G(\Phi_x(\eta_t^+, N^2 s)) - G(\Phi_x(\eta_t^-, N^2 s))] \geq (-\alpha + \kappa)(S(\eta_t^+) - S(\eta_t^-)). \tag{2.5}$$

For all large enough N, taking expectation in (2.4), we obtain from (2.1) and (2.5) for all $t \geq 0$,

$$\mathbb{E}[\xi_{t+T_*} \mid \mathcal{F}_t] \leq \left(1 + \frac{C_\varepsilon}{N^{1+\varepsilon^2}}\right) \xi_t - \alpha \int_t^{t+T_*} \mathbb{E}[\xi_s \mid \mathcal{F}_t] \, ds + (\alpha - \kappa) T_* \xi_t + \frac{C_\varepsilon'}{N^2}. \tag{2.6}$$

By the monotone coupling, if the initial state is $(\mathbf{1}, -\mathbf{1})$, then $\mathbb{E}\, \xi_t \leq \mathbb{E}\, \xi_s$ for $t \geq s$, whence taking expectation in (2.6) yields

$$\left(1 + \frac{\alpha}{N^{1-\varepsilon}}\right) \mathbb{E}\, \xi_{t+T_*} \leq \left(1 + \frac{C}{N^{1+\varepsilon^2}}\right) \mathbb{E}\, \xi_t + (\alpha - \kappa) T_* \mathbb{E}\, \xi_t + \frac{C_\varepsilon'}{N^2}.$$

Since $T_* = 1/N^{1-\varepsilon}$, multiplying the inverse of $1 + \alpha/N^{1-\varepsilon}$ in both sides yields (2.2), which completes the proof of Lemma 2.1. $\qquad\square$

3 Proof of Lemma 2.2

For any positive integer $K \geq 1$ independent of N such that $K \leq N$, let us fix K-distinct sites $\{x_1, \ldots, x_K\}$ in $\mathbb{Z}_N$. We consider an SSEP generated by L_E with particles marked by labels $\{1, \ldots, K\}$ as a stirring dynamics, i.e., a K-marked SSEP $\{z(t)\}_{t \geq 0}$ with rate 1 starting from $z_i(0) = x_i$ where $z(t) = (z_i(t))_{i=1,\ldots,K} \in (\mathbb{Z}_N)^K$, and for each bond $\{x, x+1\}$ in $\mathbb{Z}_N$, an exchange occurs at random times which are independent Poisson point processes with intensity $1/2$. Note that for the N-marked SSEP $\widetilde{z}(t) = (\widetilde{z}_x(t))_{x \in \mathbb{Z}_N}$ with rate 1 starting from $\widetilde{z}_x(0) = x$ for $x \in \mathbb{Z}_N$, we naturally identify $z(t)$ with a subset of $\widetilde{z}(t)$ such that $z_i(t) = \widetilde{z}_{x_i}(t)$ for all $i = 1, \ldots, K$ and for all $t \geq 0$ almost surely. Let us define

$$\eta^{z(t)}(x) := \begin{cases} \eta(z_i(t)) & \text{if } x = x_i \text{ for } i = 1, \ldots, K, \\ \eta(x) & \text{if } x \notin \{x_1, \ldots, x_K\}, \end{cases}$$

for $x \in \mathbb{Z}_N$ and $\eta \in \Omega_N$. This process is constant for all $x \notin \{x_1, \ldots, x_K\}$ and records the values observed by $z_i(t)$ on η at time t and each x_i for $i = 1, \ldots, K$.

Lemma 3.1. *For any local function* $f : \Omega_N \to \mathbb{R}$, *let*

$$F(\eta) := \frac{1}{N} \sum_{x \in \mathbb{Z}_N} f(\eta_{\cdot + x}) \quad \text{for } \eta \in \Omega_N,$$

and $\{z(t)\}_{t\geq 0}$ be the N-marked SSEP with rate 1 starting from each site in $\mathbb{Z}_N$. Then there exists a constant $C > 0$ depending only on $\|f\|_{\infty} = \max_{\eta \in \Omega_N} |f(\eta)|$, the size of the support $|\mathrm{supp}\, f|$ and the jump rate such that the following holds: for any $t > 0$ and for any monotone coupling (η_t^+, η_t^-) with the initial configurations $\widetilde{\eta} := (\eta^+, \eta^-)$ where $\eta^+ \geq \eta^-$,

$$\left| \mathbb{E}_{\widetilde{\eta}}[F(\eta_t^+) - F(\eta_t^-)] - \mathbf{E}\left[F(\eta^{+z(N^2 t)}) - F(\eta^{-z(N^2 t)}) \right] \right| \leq C \int_0^t \mathbb{E}\left(S(\eta_s^+) - S(\eta_s^-) \right) ds,$$

where $\eta^{+z(N^2 t)}$ and $\eta^{-z(N^2 t)}$ are $\eta^{z(N^2 t)}$ for $\eta = \eta^+$ and η^-, respectively, and $\mathbf{E}$ stands for the expectation with respect to $\{z(t)\}_{t\geq 0}$.

We defer the proof for a moment and state a technical lemma needed for it. For any $t > 0$, if we define

$$\widetilde{f}_{t-s}(\eta) := \mathbf{E}\, f(\eta^{z(N^2 (t-s))}) \quad \text{for } s \in [0, t] \text{ and } \eta \in \Omega_N,$$

then

$$\mathbb{E}\, \widetilde{f}_0(\eta_t) - \mathbb{E}\, \widetilde{f}_t(\eta) = \int_0^t \mathbb{E}\, (L_N + \partial_s)\, \widetilde{f}_{t-s}(\eta_s)\, ds,$$

where $\partial_s := \partial/\partial s$. Noting that

$$\partial_s \widetilde{f}_{t-s}(\eta) = -N^2 L_E \widetilde{f}_{t-s}(\eta),$$

by the definition of generator $L_N = L_G + N^2 L_E$, we have

$$(L_N + \partial_s) \widetilde{f}_{t-s}(\eta_s) = L_N \widetilde{f}_{t-s}(\eta_s) - N^2 L_E \widetilde{f}_{t-s}(\eta_s) = L_G \widetilde{f}_{t-s}(\eta_s),$$

and $\mathbb{E}\, \widetilde{f}_0(\eta_t) = \mathbb{E}\, f(\eta_t)$ and $\mathbb{E}\, \widetilde{f}_t(\eta) = \mathbf{E}\, f(\eta^{z(N^2 t)})$, we obtain

$$\mathbb{E}\, f(\eta_t) - \mathbf{E}\, f(\eta^{z(N^2 t)}) = \int_0^t \mathbb{E}\, L_G \widetilde{f}_{t-s}(\eta_s)\, ds \quad \text{for } t > 0. \tag{3.1}$$

We use the following lemma to show Lemma 3.1.

Lemma 3.2. *For any local function $f : \{-1, 1\}^{\mathbb{Z}_N} \to \mathbb{R}$, it holds that*

$$f(\eta^+) - f(\eta^-) = \sum_{x \in \mathrm{supp}\, f} \left(\eta^+(x) - \eta^-(x) \right) \varphi_{x,f}(\eta^+, \eta^-),$$

for any $(\eta^+, \eta^-) \in \Omega_N^2$, where $\varphi_{x,f}(\eta^+, \eta^-)$ is a polynomial with respect to $\eta^+(y), \eta^-(y)$ for $y \in \mathrm{supp}\, f \setminus \{x\}$ for each $x \in \mathrm{supp}\, f$. Moreover, we have

$$\max_{\eta^+, \eta^- \in \Omega_N} |\varphi_{x,f}(\eta^+, \eta^-)| \leq 2^{|\mathrm{supp}\, f|} \max_{\eta \in \Omega_N} |f(\eta)|.$$

Proof. For any nonempty subset I in $\mathbb{Z}_N$, let $\varphi_I(\eta) := \prod_{x \in I} \eta(x)$ and $\varphi_{\emptyset}(\eta) \equiv 1$ for $\eta \in \Omega_N$, and recall that any local function $f : \Omega_N \to \mathbb{R}$ is a linear combination of $\{\varphi_I\}_{I \subset \mathrm{supp}\, f}$ with coefficients in $\mathbb{R}$. Hence it suffices to show the claim for each φ_I with $I \neq \emptyset$. This follows from the induction on the size of I; indeed, if $|I| = 1$, then $\varphi_I(\eta^+) - \varphi_I(\eta^-) = \eta^+(x) - \eta^-(x)$ with $I = \{x\}$. For $y \notin I$,

$$\varphi_{I \cup \{y\}}(\eta^+) - \varphi_{I \cup \{y\}}(\eta^-) = \eta^+(y)\varphi_I(\eta^+) - \eta^-(y)\varphi_I(\eta^-)$$
$$= (\eta^+(y) - \eta^-(y))\varphi_I(\eta^+) + \eta^-(y)\left(\varphi_I(\eta^+) - \varphi_I(\eta^-) \right).$$

This shows that for the base of elementary local functions φ_I,

$$\varphi_I(\eta^+) - \varphi_I(\eta^-) = \sum_{x \in I} (\eta^+(x) - \eta^-(x))\varphi_{x,I}(\eta^+, \eta^-),$$

where $\varphi_{x,I}(\eta^+, \eta^-)$ is a polynomial with respect to $\eta^+(y), \eta^-(y)$ for $y \in I \setminus \{x\}$ for each $x \in I$, and further $|\varphi_{x,I}(\eta^+, \eta^-)| \leq 1$. Since $f = \sum_{I \subset \mathrm{supp}\, f} a_I \varphi_I$, where $a_I := E_{\mathrm{supp}f}[f\varphi_I]$, with $E_{\mathrm{supp}f}$ the expectation with respect to the product of the $1/2$-Bernoulli distribution on $\{-1, 1\}^{\mathrm{supp}f}$, $|a_I| \leq \|f\|_\infty := \max_{\eta \in \Omega_N} |f(\eta)|$ and we have that

$$
\begin{aligned}
f(\eta^+) - f(\eta^-) &= \sum_{I \subset \mathrm{supp}\, f} a_I (\varphi_I(\eta^+) - \varphi_I(\eta^-)) \\
&= \sum_{I \subset \mathrm{supp}\, f} a_I \sum_{x \in \mathrm{supp}\, f} (\eta^+(x) - \eta^-(x)) \varphi_{x,I}(\eta^+, \eta^-) 1_{\{x \in I\}} \\
&= \sum_{x \in \mathrm{supp}\, f} (\eta^+(x) - \eta^-(x)) \sum_{I \subset \mathrm{supp}\, f} a_I \varphi_{x,I}(\eta^+, \eta^-) 1_{\{x \in I\}}.
\end{aligned}
$$

Letting $\varphi_{x,f}(\eta^+, \eta^-) := \sum_{I \subset \mathrm{supp}\, f} a_I \varphi_{x,I}(\eta^+, \eta^-) 1_{\{x \in I\}}$, we obtain $|\varphi_{x,f}(\eta^+, \eta^-)| \leq 2^{|\mathrm{supp}\, f|} \|f\|_\infty$, as required. $\qquad\square$

Proof of Lemma 3.1. For any local function f, let $I := \mathrm{supp}\, f = \{x_1, \ldots, x_{|I|}\}$, and we consider $|I|$-marked SSEP $\{z(t)\}_{t \geq 0}$ with rate 1 and the initial state $z_i(0) = x_i$ for $i = 1, \ldots, |I|$. For any fixed $\eta \in \Omega_N$, let us write $\widetilde{f}_{t-s}(\eta) := \mathbf{E}\, f(\eta^{z(N^2(t-s))})$, where $\mathbf{E}$ stands for the expectation with respect to $z(N^2(t-s))$. Since for now we fix t, s and N, we just write for the sake of brevity with no danger of confusion

$$
\widetilde{f}_{t-s}(\eta) := \mathbf{E}\, f(\eta^{\mathcal{Z}}) \quad \text{where } \mathcal{Z} := z(N^2(t-s)),
$$

and $\mathcal{Z}$ is also considered as a subset in $\mathbb{Z}_N$. Then we have

$$
\begin{aligned}
L_G \widetilde{f}_{t-s}(\eta) &= \sum_{x \in \mathbb{Z}_N} c(x, \eta) \left(\widetilde{f}_{t-s}(\eta^x) - \widetilde{f}_{t-s}(\eta) \right) = \sum_{x \in \mathbb{Z}_N} c(x, \eta) \left(\mathbf{E}\, f((\eta^x)^{\mathcal{Z}}) - \mathbf{E}\, f(\eta^{\mathcal{Z}}) \right) \\
&= \mathbf{E} \sum_{x \in \mathbb{Z}_N} c(x, \eta) \left(f((\eta^x)^{\mathcal{Z}}) - f(\eta^{\mathcal{Z}}) \right).
\end{aligned}
$$

Since $f(\eta^{\mathcal{Z}})$ depends only on $\eta(y)$ for $y \in \mathcal{Z}$, if $x \notin \mathcal{Z}$, then $f((\eta^x)^{\mathcal{Z}}) - f(\eta^{\mathcal{Z}}) = 0$, and thus

$$
\sum_{x \in \mathbb{Z}_N} c(x, \eta) \left(f((\eta^x)^{\mathcal{Z}}) - f(\eta^{\mathcal{Z}}) \right) = \sum_{x \in \mathcal{Z}} c(x, \eta) \left(f((\eta^x)^{\mathcal{Z}}) - f(\eta^{\mathcal{Z}}) \right).
$$

Given $\mathcal{Z}$ and $x \in \mathcal{Z}$, the term $c(x, \eta) f(\eta^{\mathcal{Z}})$ is a local function depending only on η_y for $y \in \mathcal{Z} \cup I_x$, where $I_x := \mathrm{supp}\, c(x, \cdot)$, i.e., the support of $c(x, \cdot)$. Hence Lemma 3.2 implies that for $\eta^+, \eta^- \in \Omega_N$, we have

$$
c(x, \eta^+) f(\eta^{+\mathcal{Z}}) - c(x, \eta^-) f(\eta^{-\mathcal{Z}}) = \sum_{y \in \mathcal{Z} \cup I_x} (\eta^+(y) - \eta^-(y)) \varphi_{y,\mathcal{Z}}(\eta^+, \eta^-),
$$

where $\varphi_{y,\mathcal{Z}}(\eta^+, \eta^-)$ is a polynomial with respect to $\eta^+(z), \eta^-(z)$ for $z \in \mathcal{Z} \cup I_x$, and moreover, there exists a constant C depending only on the size of the support for $c(x, \eta) f(\eta^{\mathcal{Z}})$ (which is at most $|\mathrm{supp}\, f| + |\mathrm{supp}\, c(x, \cdot)|$), and $\max_{\eta \in \Omega_N} |c(x, \eta) f(\eta^{\mathcal{Z}})| \leq \|c(0, \cdot)\|_\infty \|f\|_\infty$, such that

$$
\max_{\eta^+, \eta^- \in \Omega_N} |\varphi_{y,\mathcal{Z}}(\eta^+, \eta^-)| \leq C.
$$

This shows that

$$
|c(x, \eta^+) f(\eta^{+\mathcal{Z}}) - c(x, \eta^-) f(\eta^{-\mathcal{Z}})| \leq C \sum_{y \in \mathcal{Z} \cup I_x} |\eta^+(y) - \eta^-(y)|,
$$

and we note that the same bound holds for $|c(x, \eta^+) f((\eta^{+x})^{\mathcal{Z}}) - c(x, \eta^-) f((\eta^{-x})^{\mathcal{Z}})|$:

$$
|c(x, \eta^+) f((\eta^{+x})^{\mathcal{Z}}) - c(x, \eta^-) f((\eta^{-x})^{\mathcal{Z}})| \leq C \sum_{y \in \mathcal{Z} \cup I_x} |\eta^+(y) - \eta^-(y)|.
$$

Summarizing the above discussion, we have that for any $\eta^+, \eta^- \in \Omega_N$,

$$|L_G \widetilde{f}_{t-s}(\eta^+) - L_G \widetilde{f}_{t-s}(\eta^-)| \le \mathbf{E} \sum_{x \in \mathcal{Z}} 2C \sum_{y \in \mathcal{Z} \cup I_x} |\eta^+(y) - \eta^-(y)|.$$

Therefore applying the above inequality to η_s^+, η_s^-, we obtain

$$\mathbb{E} |L_G \widetilde{f}_{t-s}(\eta_s^+) - L_G \widetilde{f}_{t-s}(\eta_s^-)| \le 2C \mathbb{E} \, \mathbf{E} \sum_{x \in \mathcal{Z}} \sum_{y \in \mathcal{Z} \cup I_x} |\eta_s^+(y) - \eta_s^-(y)|$$

$$= 2C \mathbf{E} \sum_{x \in \mathcal{Z}} \sum_{y \in \mathcal{Z} \cup I_x} \mathbb{E} |\eta_s^+(y) - \eta_s^-(y)|,$$

where we have used the Fubini theorem in the last equality. Since $\eta_s^+ \ge \eta_s^-$ under the monotone coupling, taking the average $\eta_s^+ - \eta_s^-$ by the translation on $\mathbb{Z}_N$ in the last term, namely, plugging $\eta^+_{\cdot + z, s} - \eta^-_{\cdot + z, s}$ into the above inequality and taking the arithmetic mean over $z \in \mathbb{Z}_N$, we obtain

$$2C \mathbf{E} \sum_{x \in \mathcal{Z}} \sum_{y \in \mathcal{Z} \cup I_x} \mathbb{E} \left(S(\eta_s^+) - S(\eta_s^-) \right) \le 2C |I| (|I| + |I_0|) \mathbb{E} \left(S(\eta_s^+) - S(\eta_s^-) \right),$$

and thus

$$\frac{1}{N} \sum_{x \in \mathbb{Z}_N} \mathbb{E} |L_G \widetilde{f}_{t-s}(\eta^+_{\cdot + x, s}) - L_G \widetilde{f}_{t-s}(\eta^-_{\cdot + x, s})| \le C' \mathbb{E} \left(S(\eta_s^+) - S(\eta_s^-) \right),$$

for a constant $C' > 0$, where $\eta^+_{\cdot + x, s} := (\eta_s^+(y + x))_{y \in \mathbb{Z}_N}$ and $\eta^-_{\cdot + x, s}$ is defined similarly. We recall that $F(\eta) = (1/N) \sum_{x \in \mathbb{Z}_N} f(\eta_{\cdot + x})$ and note that $(L_G F)(\eta) = \sum_{x \in \mathbb{Z}_N} (L_G f)(\eta_{\cdot + x})$. Hence integrating in s from 0 to t, we have that by (3.1),

$$\left| \mathbb{E} \left[F(\eta_t^+) - F(\eta_t^-) \right] - \mathbf{E} \left[F(\eta^{+z(N^2 t)}) - F(\eta^{-z(N^2 t)}) \right] \right| \le C' \int_0^t \mathbb{E} \left(S(\eta_s^+) - S(\eta_s^-) \right) ds.$$

Note that the constant C' depends only on $\|f\|_\infty$ and $|\operatorname{supp} f|$ as well as $\|c(0, \cdot)\|_\infty$ and $|\operatorname{supp} c(0, \cdot)|$. This shows the claim. $\qquad \square$

The proof of Lemma 2.2 is based on the following lemma, which we prove in Section 4.

Lemma 3.3. *For K-distinct sites $x_1, \ldots, x_K$ in $\mathbb{Z}_N$ where $K \ge 1$, let $\psi : \Omega_N \mapsto \mathbb{R}$ be a local function of the form*

$$\psi(\eta) := \prod_{i=1}^K \eta(x_i), \quad \text{for } \eta \in \Omega_N.$$

Let $z(t) = \{z_i(t)\}_{i=1,\ldots,K}$ for $t \ge 0$ be a K-marked SSEP on $\mathbb{Z}_N$ starting from $\{x_i\}_{i=1,\ldots,K}$. If we define

$$\Delta_T(\eta) := \mathbf{E} \, \psi(\eta^{z(T)}) - \prod_{i=1}^K \mathbf{E} \, \eta(z_i(T)) \quad \text{for } \eta \in \Omega_N \text{ and } T \ge 0,$$

then for all $0 < \varepsilon < 1/28$ there exist constants $C > 0$ and T_0 such that for all $T_0 \le T \le N^{2+\varepsilon}$ the following holds: for all $\eta^+, \eta^- \in \Omega_N$,

$$|\Delta_T(\eta^+) - \Delta_T(\eta^-)| \le \frac{C}{T^{\varepsilon/4}} \sum_{i=1}^K \mathbf{E} |\eta^+(z_i(T)) - \eta^-(z_i(T))| + C \exp(-T^{\varepsilon/8}),$$

where the implied constant C depends only on K and ε.

Proof of Lemma 2.2. Fix an arbitrary $\widetilde{\eta} = (\eta^+, \eta^-)$ with $\eta^+ \geq \eta^-$. For each $x \in \mathbb{Z}_N$, let $z(t) = (z_i(t))_{i=1,\ldots,K}$ for $t \geq 0$ be a K-marked SSEP starting from $\{x_i + x\}_{i=1,\ldots,K}$ and $\mathbf{E}_x$ be the corresponding expectation. Then for all $T \geq T_0$ and for all configurations $\eta^+, \eta^- \in \Omega_N$, we have

$$\left| \sum_{x \in \mathbb{Z}_N} \left(\prod_{i=1}^{K} \mathbf{E}_x \eta^+(z_i(T)) - \prod_{i=1}^{K} \mathbf{E}_x \eta^-(z_i(T)) \right) - \sum_{x \in \mathbb{Z}_N} \left(\Phi_x(\eta^+, T)^K - \Phi_x(\eta^-, T)^K \right) \right|$$

$$\leq \frac{C}{\sqrt{T}} \sum_{x \in \mathbb{Z}_N} \sum_{i=1}^{K} \left| \mathbf{E}_x \eta^+(z_i(T)) - \mathbf{E}_x \eta^-(z_i(T)) \right|. \tag{3.2}$$

Indeed, recall that $\Phi_{z_i(0)}(\eta, T) = \mathbf{E}\,\eta(z_i(T))$, and a "smoothing effect" of SRW (Lemma B.1) implies that for $i \neq j$,

$$|\Phi_{z_i(0)}(\eta, T) - \Phi_{z_j(0)}(\eta, T)| \leq 2\|\mathbb{P}_{z_i(0)}(z_i(T) \in \cdot) - \mathbb{P}_{z_j(0)}(z_j(T) \in \cdot)\|_{\mathrm{TV}}$$

$$\leq \frac{2C|x_i - x_j|}{\sqrt{T}} \leq \frac{2CD}{\sqrt{T}} \quad \text{for all } T \geq T_0 \text{ and } \eta \in \Omega_N, \tag{3.3}$$

where $D := \max_{i,j=1,\ldots,K} |x_i - x_j|$. For a short hand notation, we write $\Phi_i^+ := \Phi_{z_i(0)}(\eta^+, T)$ for $i = 1, \ldots, K$ and similarly for Φ_i^-. Let us use the following decomposition:

$$\prod_{i=1}^{K} \Phi_i^+ - \prod_{i=1}^{K} \Phi_i^- = \left(\Phi_1^+ - \Phi_1^- \right) \prod_{i \neq 1} \Phi_i^+ + \Phi_1^- \left(\Phi_2^+ - \Phi_2^- \right) \prod_{i \neq 1,2} \Phi_i^+ + \cdots + \left(\prod_{i \neq K} \Phi_i^- \right) \left(\Phi_K^+ - \Phi_K^- \right). \tag{3.4}$$

Expanding the product in a similar way, we obtain by (3.3) for $1 \leq l < K$,

$$\left| \prod_{i=i_1,\ldots,i_l} \Phi_i^+ - (\Phi_j^+)^l \right| \leq \frac{2CDl}{\sqrt{T}} \leq \frac{2CDK}{\sqrt{T}} \quad \text{for } j, i_1, \ldots, i_l = 1, \ldots, K,$$

for all $T \geq T_0$ and the same for Φ_i^-. We combine those estimates and see that by triangle inequalities the difference between $\prod_{i=1}^{K} \Phi_i^+ - \prod_{i=1}^{K} \Phi_i^-$ and

$$\left(\Phi_1^+ - \Phi_1^- \right) (\Phi_1^+)^{K-1} + \Phi_2^- \left(\Phi_2^+ - \Phi_2^- \right) (\Phi_2^+)^{K-2} + \cdots + (\Phi_K^-)^{K-1} \left(\Phi_K^+ - \Phi_K^- \right)$$

is at most

$$\frac{2CDK}{\sqrt{T}} \sum_{i=1}^{K} (\Phi_i^+ - \Phi_i^-).$$

Note that the second sum of the left-hand side in (3.2) can be written as the sum of the penultimate display by the translation on $\mathbb{Z}_N$. Therefore summing up over all $x \in \mathbb{Z}_N$, and abbreviating the constant, we obtain (3.2).

Denote by $\psi_x(\eta) = \prod_{i=1,\ldots,K} \eta(x + x_i)$. Combining Lemma 3.3 with (3.2) for any $\varepsilon < 1/28$, we have that for all $T_0 \leq T \leq N^{2+\varepsilon}$ (this restriction on the range for T is

required to apply Lemma 3.3),

$$
\left| \frac{1}{N} \sum_{x \in \mathbb{Z}_N} \mathbf{E}_x[\psi_x(\eta^{+z(T)}) - \psi_x(\eta^{-z(T)})] - \frac{1}{N} \sum_{x \in \mathbb{Z}_N} [\Phi_x(\eta^+, T)^K - \Phi_x(\eta^-, T)^K] \right|
$$

$$
\leq \left| \frac{1}{N} \sum_{x \in \mathbb{Z}_N} \mathbf{E}_x[\psi_x(\eta^{+z(T)}) - \psi_x(\eta^{-z(T)})] - \frac{1}{N} \sum_{x \in \mathbb{Z}_N} \left(\prod_{i=1}^{K} \mathbf{E}_x \eta^+(z_i(T)) - \prod_{i=1}^{K} \mathbf{E}_x \eta^-(z_i(T)) \right) \right|
$$

$$
+ \frac{C}{\sqrt{T} N} \sum_{x \in \mathbb{Z}_N} \sum_{i=1}^{K} \left| \mathbf{E}_x \eta^+(z_i(T)) - \mathbf{E}_x \eta^-(z_i(T)) \right|
$$

$$
\leq \frac{C}{T^{\varepsilon/4} N} \sum_{x \in \mathbb{Z}_N} \sum_{i=1}^{K} \mathbf{E}_x \left(\eta^+(z_i(T)) - \eta^-(z_i(T)) \right) + C e^{-T^{\varepsilon/8}}
$$

$$
= \frac{CK}{T^{\varepsilon/4} N} \sum_{x \in \mathbb{Z}_N} \left(\Phi_x(\eta^+, T) - \Phi_x(\eta^-, T) \right),
$$

where we have used (3.2) in the first inequality and Lemma 3.3 and ε is less than $1/2$ in the second inequality. If we define $F(\eta) = (1/N) \sum_{x \in \mathbb{Z}_N} \psi_x(\eta)$, then by Lemma 3.1 for $t = T/N^2$ and for $\widetilde{\eta} = (\eta^+, \eta^-)$ where $\eta^+ \geq \eta^-$,

$$
\left| \mathbb{E}_{\widetilde{\eta}}[F(\eta_t^+) - F(\eta_t^-)] - \mathbb{E} \left[F(\eta^{+z(T)}) - F(\eta^{-z(T)}) \right] \right| \leq C \int_0^t \mathbb{E} \left(S(\eta_s^+) - S(\eta_s^-) \right) ds,
$$

and plugging $T = tN^2$ in the penultimate display, we conclude that for all $T_0/N^2 \leq t \leq N^\varepsilon$,

$$
\left| \mathbb{E}_{\widetilde{\eta}}[F(\eta_t^+) - F(\eta_t^-)] - \frac{1}{N} \sum_{x \in \mathbb{Z}_N} [\Phi_x(\eta^+, N^2 t)^K - \Phi_x(\eta^-, N^2 t)^K] \right|
$$

$$
\leq \frac{CK}{t^{\varepsilon/4} N^{1+2\varepsilon}} \sum_{x \in \mathbb{Z}_N} \left(\Phi_x(\eta^+, N^2 t) - \Phi_x(\eta^-, N^2 t) \right) \tag{3.5}
$$

$$
+ C \int_0^t \mathbb{E} \left(S(\eta_s^+) - S(\eta_s^-) \right) ds + C \exp(-(N^2 t)^{\varepsilon/8}).
$$

Note that for any $s \geq 0$,

$$
\xi_0 = S(\eta^+) - S(\eta^-) = \frac{1}{N} \sum_{x \in \mathbb{Z}_N} \left(\Phi_x(\eta^+, s) - \Phi_x(\eta^-, s) \right).
$$

Recall that in the general case $F(\eta)$ has the form $F(\eta) = (1/N) \sum_{x \in \mathbb{Z}_N} \sum_{|I| \neq 1} a_I f_I(\eta_{\cdot+x})$ and $G(\rho) = \sum_{|I| \neq 1} a_I \rho^{|I|}$. Taking a linear combination and applying to the above inequality (3.5), we conclude the proof of Lemma 2.2. $\qquad\square$

4   Replacement lemma: Proof of Lemma 3.3

For an integer $K \geq 1$ we consider the K-marked SSEP $z(t) := \{z_i(t)\}_{i=1,\dots,K}$ on $\mathbb{Z}_N$, and the K-marked independent SRWs $z^0(t) := \{z_i^0(t)\}_{i=1,\dots,K}$ with rate 1, i.e., each particle runs as a SRW having an independent Poisson clock with intensity 1 (see the explicit generator in Appendix A). We use the following coupling between these two processes; the result is known (e.g. [DMP91]) and has been used extensively in [DMFL86]. Since we need to clarify the dependence on every constant involved in the processes and we also need to refer to the construction of the coupling, we include the proof of the following claim.

Lemma 4.1. *For any integers $N \geq 2$ and $K \geq 1$ with $N \geq K$, fix any initial condition $z(0) = z^0(0) = \{x_1, \ldots, x_K\}$ in $\mathbb{Z}_N$ where $x_i \neq x_j$ for $i \neq j$, $i, j = 1, \ldots, K$. Then there exists a coupling $\mathbf{P}$ between a K-marked SSEP $\{z(t)\}_{t \geq 0}$ and K-marked independent SRWs $\{z^0(t)\}_{t \geq 0}$ with rate 1 such that the following holds: for any $\varepsilon > 0$ there exists T_0 such that for all $T_0 \leq T \leq N^{2+\varepsilon}$,*

$$\mathbf{P}\left(\max_{j=1,\ldots,K} \max_{0 \leq t \leq T} |z_j(t) - z_j^0(t)| \geq KT^{\frac{1}{4}+3\varepsilon}\right) \leq 4K^2 \exp\left(-T^{\varepsilon/4}\right),$$

and $z_1(t) = z_1^0(t)$ for all $t \geq 0$ almost surely.

In the following, we will use several results on the SRW on $\mathbb{Z}_N$; all of them are classical — we include the proofs for the sake of completeness in Appendix A and B.

Proof of Lemma 4.1. We construct a coupling from the graphical representation on $\mathbb{Z}_N \times [0, \infty)$. For each bond $\{x, x+1\}$ in $\mathbb{Z}_N$, we define random exclusion times $\mathcal{E}^{x,x+1} = \{\mathcal{E}_n^{x,x+1}; n \geq 1\}$ where $\mathcal{E}^{x,x+1}$ are independent Poisson point processes with intensity 1. Furthermore, for each bond $\{x, x+1\}$, we define a sequence of independent $1/2$-Bernoulli random variables $\{U_n^{x,x+1}; n \geq 1\}$ where $U_n^{x,x+1} \in \{0, 1\}$ for $n \geq 1$ and these sequences are independent of each other and of everything else. For each $t = \mathcal{E}_n^{x,x+1}$, we draw a double arrow between (x, t) and $(x+1, t)$ with mark $U_n^{x,x+1}$.

Given the initial condition $z_i(0) = x_i$ for $i = 1, \ldots, K$, the K-particles evolve as follows: for $t > s \geq 0$, given $z_i(s)$, $i = 1, \ldots, K$, we put $z_i(t) = z_i(s)$ for all $i = 1, \ldots, K$ as far as there is no double arrow in time $(s, t]$. Starting from time $t = 0$, for each $i = 1, \ldots, K$, if there is a double arrow with mark 1 between $(z_i(t-), t)$ and (x, t), then we let $z_i(t) = x$, and do nothing otherwise. Then $\{(z_i(t))_{i=1,\ldots,K}\}_{t \geq 0}$ is almost surely defined and has the distribution of K-marked SSEP starting from $z_i(0) = x_i$ for $i = 1, \ldots, K$.

Concerning the K-marked independent SRWs, let us define $\{(z_i^0(t))_{i=1,\ldots,K}\}_{t \geq 0}$ depending on the number $i = 1, \ldots, K$. Suppose that there is a double arrow between x and $x+1$ at time t for $x \in \mathbb{Z}_N$ and $t \geq 0$. There are three cases according to the possible configurations on the sites $x, x+1$:

(1) There are no particles from z on the sites $x, x+1$. In this case we do nothing.

(2) There is a single particle from z on the sites $x, x+1$. Suppose that $z_i(t-) = x$ or $x+1$. If the double arrow has the mark 1, then we define

$$z_i^0(t) := z_i^0(t-) + (z_i(t) - z_i(t-)),$$

and keep everything else, and if the double arrow has the mark 0, then we do nothing.

(3) There are two particles $z_i(t-), z_j(t-)$ on the sites $x, x+1$. Assume that $i < j$ and the double arrow has the mark $U \in \{0, 1\}$. Then we define

$$U = 1 \implies z_i^0(t) := z_i^0(t-) + (z_i(t) - z_i(t-)) \quad \text{and} \quad z_j^0(t) := z_j^0(t-),$$

and

$$U = 0 \implies z_j^0(t) := z_j^0(t-) + (z_i(t) - z_j(t-)) \quad \text{and} \quad z_i^0(t) := z_i^0(t-),$$

and we keep everything else.

The construction gives K-independent SRWs $\{(z_i^0(t))_{i=1,\ldots,K}\}_{t \geq 0}$ with rate 1 (cf. Figure 1).

Let us show that the resulting coupling $\mathbf{P}$ is the desired one. We analyze $z_i(t) - z_i^0(t)$ for $t \geq 0$ and for each $i = 1, \ldots, K$. Note that in this construction we have $z_1^0(t) = z_1(t)$ for all $t \geq 0$. For $j > 1$, $z_j(t) - z_j^0(t)$ can differ only during the period when $|z_j(t) - z_i(t)| = 1$
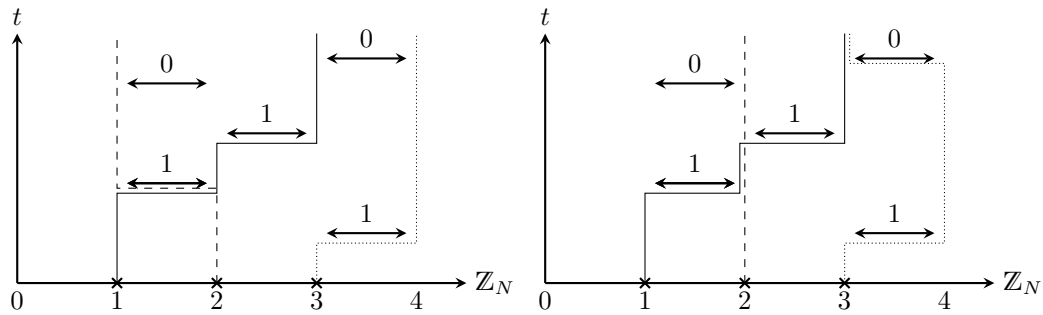
Figure 1: Left: 3-marked SSEP starting from $x_1 = 1, x_2 = 2, x_3 = 3$. Right: 3-independent SRWs starting from $x_1 = 1, x_2 = 2, x_3 = 3$. Three paths (which are solid, dotted, and densely dotted, respectively) correspond to trajectories of three marked particles in the processes, respectively.

for some $i < j$. For each $i < j$, denoting by $w_{i,j}(t)$ the difference created in that period up to time t, we have

$$z_j(t) - z_j^0(t) = \sum_{i<j} w_{i,j}(t) \quad \text{for } j = 2, \ldots, K.$$

For $i < j$, if we define $\tau_\star := \inf\{t \geq 0 \ : \ |z_i(t) - z_j(t)| = 1\}$, then $\tau_\star < \infty$ almost surely, and given $\tau_\star$ and $z_i(\tau_\star), z_j(\tau_\star)$, let

$$T_\star := \inf\{s \geq 0 \ : \ |z_i(s + \tau_\star) - z_j(s + \tau_\star)| > 1\}.$$

Then the strong Markov property shows that $T_\star$ conditioned on the σ-algebra $\mathcal{G}_{\tau_\star}$ associated with the filtration $\mathcal{G}_t := \sigma(\{z(s)\}_{0 \leq s \leq t})$ for $t \geq 0$ and the stopping time $\tau_\star$ has the exponential distribution with rate 1, and is independent of the random exchange times between $z_i(\tau_\star)$ and $z_j(\tau_\star)$ since $T_\star$ depends only on the random exchange times after $\tau_\star$ on $(x-1, x)$ and $(x+1, x+2)$ for $\{z_i(\tau_\star), z_j(\tau_\star)\} = \{x, x+1\}$. Let us define for $t \geq 0$ and $i < j$,

$$\theta_{i,j}(t) := |\{s \in [0, t] \ : \ |z_i(s) - z_j(s)| = 1\}|,$$

where $|\cdot|$ is the normalized Lebesgue measure. Conditioned on the random variable $\theta_{i,j}(t)$, each $w_{i,j}(t)$ has the distribution of a SRW with rate 1 starting from 0 on $\mathbb{Z}_N$ at time $\theta_{i,j}(t)$. (Note however that for each $j > 2$, $\{w_{i,j}(t)\}_{i<j}$ are not independent.)

We observe that for each i, j with $i < j$, the random variable $\theta_{i,j}(t)$ is stochastically dominated by the occupation time of SRW with rate 2 on the 3-sites $\{-1, 0, 1\}$ on $\mathbb{Z}_N$ starting from $z_0 := z_i(0) - z_j(0)$. Hence by Lemma A.2, letting $\theta(T)$ be the occupation time at 0 up to time T for the continuous-time SRW on $\mathbb{Z}_N$ with rate 2 starting from 0, we have that for any $\varepsilon > 0$, there exists T_0 such that for all $T_0 \leq T \leq N^{2+\varepsilon}$ and for each pair i, j with $i < j$,

$$\mathbb{P}\left(\theta_{i,j}(T) \geq 6T^{\frac{1}{2}+2\varepsilon}\right) \leq 3\mathbb{P}\left(\theta(T) \geq 2T^{\frac{1}{2}+2\varepsilon}\right) \leq 3\exp\left(-T^{\varepsilon/4}\right). \tag{4.1}$$

Then, it holds that for all large enough T satisfying that $T \leq N^{2+\varepsilon}$ for each pair $i < j$,

$$\mathbb{P}\left(\max_{0 \leq t \leq \theta_{i,j}(T)} |w_{i,j}(t)| \geq T^{\frac{1}{4}+3\varepsilon}\right)$$

$$\leq \mathbb{P}\left(\max_{0 \leq t \leq 6T^{\frac{1}{2}+2\varepsilon}} |w_{i,j}(t)| \geq T^{\frac{1}{4}+3\varepsilon}\right) + \mathbb{P}\left(\theta_{i,j}(T) \geq 6T^{\frac{1}{2}+2\varepsilon}\right)$$

$$\leq 3\exp\left(-\frac{1}{6}T^{2\varepsilon}\right) + 3\exp\left(-T^{\varepsilon/4}\right) \leq 4\exp\left(-T^{\varepsilon/4}\right),$$

where we have used Lemma A.1 with $(6T^{\frac{1}{2}+2\varepsilon})^{\frac{1}{2}+\varepsilon} < T^{\frac{1}{4}+3\varepsilon}$ for $T \geq T_0$ (a re-chosen T_0 if necessary) and (4.1) in the second inequality.

Noting that for each $j > 1$ and for $t, T > 0$,

$$\mathbf{P}\left(\max_{0 \leq t \leq T} |z_j(t) - z_j^0(t)| \geq KT\right) \leq \sum_{i<j} \mathbf{P}\left(\max_{0 \leq t \leq T} |w_{i,j}(t)| \geq T\right),$$

we have that for any $\varepsilon > 0$ and for all large enough T with $T \leq N^{2+\varepsilon}$,

$$\mathbf{P}\left(\max_{0 \leq t \leq T} |z_j(t) - z_j^0(t)| \geq KT^{\frac{1}{4}+3\varepsilon}\right) \leq 4K \exp\left(-T^{\varepsilon/4}\right).$$

Therefore the union bound over $j = 1, \ldots, K$ implies that for any $\varepsilon > 0$ there exists T_0' such that for all $T_0' \leq T \leq N^{2+\varepsilon}$,

$$\mathbf{P}\left(\max_{j=1,\ldots,K} \max_{0 \leq t \leq T} |z_j(t) - z_j^0(t)| \geq KT^{\frac{1}{4}+3\varepsilon}\right) \leq 4K^2 \exp\left(-T^{\varepsilon/4}\right),$$

as required. $\qquad\square$

In order to prove Lemma 3.3, we use a coupling between a K-marked SSEP and K-marked independent SRWs as it is constructed in Lemma 4.1. A heuristic explanation for the proof of Lemma 3.3 is as follows. Let $s_T := T - T^{\frac{1}{2}+7\varepsilon}$. Decomposing the time interval $[0, T]$ as $[0, s_T]$ and $[s_T, T]$, we construct some events (which will be denoted by $\mathcal{G}, \mathcal{H}_1$ and $\mathcal{H}_2$) to ensure that K-particles in the K-marked SSEP are fairly well-separated at time s_T and behave like independent SRWs during the time range $[s_T, T]$. Since in the first stage $[0, s_T]$ those particles are already fairly separated, they remain sufficiently separated during $[s_T, T]$. Moreover, the 'good' event $\mathcal{G} \cap \mathcal{H}_1 \cap \mathcal{H}_2$ happens with high probability. The construction of these events is a key ingredient to prove Lemma 3.3.

Proof of Lemma 3.3. Recall that we consider a local function $\psi(\eta) = \prod_{i=1}^{K} \eta(x_i)$ for K-distinct sites $x_1, \ldots, x_K$ in $\mathbb{Z}_N$. We consider a coupling $\mathbf{P}$ between $\{z(t)\}_{t \geq 0}$ and K-marked independent SRWs $\{z^0(t)\}_{t \geq 0}$ with rate 1 such that $z_i(0) = z_i^0(0)$ for $i = 1, \ldots, K$ as it is constructed in Lemma 4.1. For all $0 < \varepsilon < 1/28$, let $s_T := T - T^{\frac{1}{2}+7\varepsilon}$. First we define the event $\mathcal{G}$ where it holds that for all $i = 1, \ldots, K$, if $i \neq j$, then

$$|z_i^0(s_T) - z_j^0(s_T)| \geq T^{\frac{1}{2}-\varepsilon}.$$

Note that $z_1^0(t) = z_1(t)$ holds for all $t \geq 0$ almost surely in the coupling $\mathbf{P}$. By a classical estimate for the SRW (Lemma B.2) applied to

$$z_i^0(t) - z_j^0(t) \quad \text{for } i, j = 1, \ldots, K \text{ with } i \neq j \text{ conditioned on } \{z_1^0(t)\}_{0 \leq t \leq T},$$

there exists T_0 such that for all $T_0 \leq T \leq N^{2+\varepsilon}$, we have almost surely in $\mathbf{P}$,

$$\mathbf{P}\left(\mathcal{G}^c \mid \{\sigma(z_1^0(t))\}_{0 \leq t \leq T}\right) \leq \frac{CK^2}{T^{\varepsilon/4}},$$

in particular, since $z_1^0(t) = z_t(t)$ for all $t \geq 0$ almost surely in $\mathbf{P}$, we have that

$$\mathbf{P}\left(\mathcal{G}^c \mid \sigma(z_1(T))\right) \leq \frac{CK^2}{T^{\varepsilon/4}} \qquad \text{almost surely in } \mathbf{P}, \tag{4.2}$$

where $\mathcal{G}^c$ stands for the complement of the event $\mathcal{G}$. Next we define the event $\mathcal{H}_1$ where it holds that for all $i = 1, \ldots, K$,

$$|z_i(s_T) - z_i^0(s_T)| \leq KT^{\frac{1}{4}+3\varepsilon}.$$

Then Lemma 4.1 shows that

$$\mathbf{P}\left(\mathcal{H}_1^c\right) \leq 4K^2 \exp(-T^{\varepsilon/4}) \qquad \text{for all } T_0 \leq T \leq N^{2+\varepsilon}, \tag{4.3}$$

where we take another T_0 if necessary and denote it by the same symbol. Based on the coupling $\mathbf{P}$, we couple $\{z(t)\}_{s_T \leq t < T}$ and K-marked independent SRWs $\{w^0(t)\}_{t \geq 0}$ with rate 1 such that $w_i^0(0) = z_i(s_T)$ for each $i = 1, \ldots, K$ in the same way as in Lemma 4.1——we denote the resulting coupling by the same symbol $\mathbf{P}$. Let us define the event $\mathcal{H}_2$ where it holds that for all $i = 1, \ldots, K$,

$$|z_i(t + s_T) - w_i^0(t)| \leq K T^{\frac{1}{8}+4\varepsilon} \quad \text{for all } 0 \leq t \leq T^{\frac{1}{2}+7\varepsilon}$$

and for all $i \neq j$,

$$|w_i^0(t) - w_j^0(t)| \geq \frac{1}{2} T^{\frac{1}{4}+\varepsilon} \quad \text{for all } 0 \leq t \leq T^{\frac{1}{2}+7\varepsilon}.$$

Lemma 4.1 and an application of the maximal inequality (the second claim of Lemma A.1 with $\lambda = 2$ applied to $\{w_i^0(t) - w_j^0(t)\}$ for $0 \leq t \leq T^{\frac{1}{2}+7\varepsilon}$ with $i \neq j$) imply that by the Markov property of $\{z(t)\}_{t \geq 0}$,

$$\mathbf{P}\left(\mathcal{H}_2^c \mid \sigma(z(s_T), z^0(s_T))\right) \mathbf{1}_{\mathcal{G}} \leq 4K^2 \exp(-T^{\varepsilon/8}) + 3K^2 \exp(-\frac{1}{12}T^{\varepsilon})$$

$$\leq 5K^2 \exp(-T^{\varepsilon/8}) \quad \text{for all } T_0 \leq T \leq N^{2+\varepsilon}, \tag{4.4}$$

where we have used

$$\left(T^{\frac{1}{2}+7\varepsilon}\right)^{\frac{1}{4}+3\varepsilon} = T^{\frac{1}{8}+\frac{13}{4}\varepsilon+21\varepsilon^2} < T^{\frac{1}{8}+4\varepsilon} \quad \text{for } 0 < \varepsilon < \frac{1}{28}, \quad \text{and}$$

$$T^{\left(\frac{1}{2}+7\varepsilon\right)\left(\frac{1}{2}+\varepsilon\right)} > T^{\frac{1}{4}+\varepsilon} \quad \text{for } \varepsilon > 0,$$

and further we choose another T_0 depending only on ε if necessary. Note that given the event $\mathcal{G} \cap \mathcal{H}_1 \cap \mathcal{H}_2$, we have by the triangle inequality

$$|z_i(t + s_T) - z_j(t + s_T)| \geq |w_i^0(t) - w_j^0(t)| - |z_i(t + s_T) - w_i^0(t)| - |z_j(t + s_T) - w_j^0(t)|$$

$$\geq \frac{1}{2} T^{\frac{1}{4}+\varepsilon} - 2K T^{\frac{1}{8}+4\varepsilon} > 1$$

for all $i \neq j$, for all $0 \leq t \leq T - s_T$ and for all $T_0 \leq T \leq N^{2+\varepsilon}$, choosing another T_0 depending only on ε and K if necessary, and thus by the construction of coupling between $\{z(t)\}_{s_T \leq t \leq T}$ and $\{w^0(t)\}_{t \geq 0}$, we have that for all $i = 1, \ldots, K$,

$$z_i(t + s_T) = w_i^0(t) \quad \text{for all } 0 \leq t \leq T - s_T.$$

Let us decompose

$$\mathbf{E}\,\psi(\eta^{z(T)}) = \mathbf{E}\,\psi(\eta^{z(T)})\mathbf{1}_{\mathcal{G}} + \mathbf{E}\,\psi(\eta^{z(T)})\mathbf{1}_{\mathcal{G}^c}, \tag{4.5}$$

and estimate the second term of (4.5). For each $i = 1, \ldots, K$, if we define $g_i(\eta) := \prod_{j \neq i} \eta(x_j)$ for $\eta \in \Omega_N$, then $g_i(\eta^{z(T)}) = \prod_{j \neq i} \eta(z_j(T))$, and we have

$$\mathbf{E}\,\psi(\eta^{z(T)})\mathbf{1}_{\mathcal{G}^c} = \mathbf{E}\left[\mathbf{E}\left[\psi(\eta^{z(T)})\mathbf{1}_{\mathcal{G}^c} \mid \sigma(z_i(T))\right]\right]$$

$$= \mathbf{E}\left[\eta(z_i(T))\mathbf{E}\left[g_i(\eta^{z(T)})\mathbf{1}_{\mathcal{G}^c} \mid \sigma(z_i(T))\right]\right].$$

Letting $r(i, \eta, T) := \mathbf{E}\left[g_i(\eta^{z(T)})\mathbf{1}_{\mathcal{G}^c} \mid \sigma(z_i(T))\right]$, by (4.2) we obtain $|r(i, \eta, T)| \leq C/T^{\varepsilon/4}$ for all $T_0 \leq T \leq N^{2+\varepsilon}$. Note that for all $\eta^+, \eta^- \in \Omega_N$,

$$|\mathbf{E}\left(\eta^+(z_i(T)) - \eta^-(z_i(T))\right)g_i(\eta^{+z(T)})\mathbf{1}_{\mathcal{G}^c}| = |\mathbf{E}\,r(i, \eta^+, T)\left(\eta^+(z_i(T)) - \eta^-(z_i(T))\right)|$$

$$\leq \frac{C}{T^{\varepsilon/4}}\mathbf{E}\,|\eta^+(z_i(T)) - \eta^-(z_i(T))|.$$

For $\eta^+, \eta^- \in \Omega_N$, using the decomposition

$$\prod_{i=1}^{K} \eta^+(x_i) - \prod_{i=1}^{K} \eta^-(x_i) = (\eta^+(x_1) - \eta^-(x_1)) \prod_{j \neq 1} \eta^+(x_j)$$

$$+ \eta^-(x_1)(\eta^+(x_2) - \eta^-(x_2)) \prod_{j \neq 1,2} \eta^+(x_j) + \cdots + \left(\prod_{i \neq K} \eta^-(x_i) \right) (\eta^+(x_K) - \eta^-(x_K)),$$

and permuting the role of x_i for $i = 1, \ldots, K$, we obtain

$$\left| \mathbf{E} \, \psi(\eta^{+z(T)}) \mathbf{1}_{\mathcal{G}^c} - \mathbf{E} \, \psi(\eta^{-z(T)}) \mathbf{1}_{\mathcal{G}^c} \right| \leq \frac{C}{T^{\varepsilon/4}} \sum_{i=1}^{K} \mathbf{E} \, |\eta^+(z_i(T)) - \eta^-(z_i(T))|. \tag{4.6}$$

Concerning the first term of (4.5), we have by (4.3) and (4.4) for all $T_0 \leq T \leq N^{2+\varepsilon}$,

$$\mathbf{E} \, \psi(\eta^{z(T)}) \mathbf{1}_{\mathcal{G}} = \mathbf{E} \, \psi(\eta^{z(T)}) \mathbf{1}_{\mathcal{G} \cap \mathcal{H}_1 \cap \mathcal{H}_2} + O\left(e^{-T^{\varepsilon/8}} \right),$$

and $\mathbf{E} \, \psi(\eta^{z(T)}) \mathbf{1}_{\mathcal{G} \cap \mathcal{H}_1 \cap \mathcal{H}_2} = \mathbf{E} \left[\mathbf{E} \left[\psi(\eta^{z(T)}) \mathbf{1}_{\mathcal{H}_2} \mid \mathcal{F}_{s_T} \right] \mathbf{1}_{\mathcal{G} \cap \mathcal{H}_1} \right]$. Given $\mathcal{G} \cap \mathcal{H}_1$, one has

$$\mathbf{E} \left[\psi(\eta^{z(T)}) \mathbf{1}_{\mathcal{H}_2} \mid \mathcal{F}_{s_T} \right] = \mathbf{E} \prod_{i=1}^{K} \eta(w_i^0(T - s_T)) \mathbf{1}_{\mathcal{H}_2} = \mathbf{E} \prod_{i=1}^{K} \eta(w_i^0(T - s_T)) + O(e^{-T^{\varepsilon/8}}),$$

and

$$\mathbf{E} \prod_{i=1}^{K} \eta(w_i^0(T - s_T)) = \prod_{i=1}^{K} \mathbf{E} \, \eta(w_i^0(T - s_T)) = \prod_{i=1}^{K} \Phi_{z_i(s_T)}(\eta, T - s_T),$$

where the first equality follows since $\{w^0(t)\}_{t \geq 0}$ are independent SRWs and we recall $\Phi_x(\eta, t) = \mathbf{E}_x \eta(w(t))$ a local average of η by the SRW with rate 1 around x. Then we use the following estimate on SRW: for all $t > 0$,

$$\|\mathbb{P}_{z_i(s_T)}(w(t) \in \cdot) - \mathbb{P}_{z_i^0(s_T)}(w(t) \in \cdot)\|_{\mathrm{TV}} \leq \frac{C|z_i(s_T) - z_i^0(s_T)|}{\sqrt{t}},$$

(Lemma B.1), and we have that given $\mathcal{H}_1$ for each $i = 1, \ldots, K$,

$$|\Phi_{z_i(s_T)}(\eta, T - s_T) - \Phi_{z_i^0(s_T)}(\eta, T - s_T)|$$
$$\leq 2\|\mathbb{P}_{z_i(s_T)}(w(T - s_T) \in \cdot) - \mathbb{P}_{z_i^0(s_T)}(w(T - s_T) \in \cdot)\|_{\mathrm{TV}}$$
$$\leq \frac{2C|z_i(s_T) - z_i^0(s_T)|}{\sqrt{T - s_T}} \leq \frac{2CKT^{\frac{1}{4}+3\varepsilon}}{\sqrt{T^{\frac{1}{2}+7\varepsilon}}} = \frac{2CK}{T^{\varepsilon/2}},$$

for all $T_0 \leq T \leq N^{2+\varepsilon}$. This implies that given $\mathcal{H}_1$, for each $i = 1, \ldots, K$,

$$\prod_{j \neq i} \Phi_{z_j(s_T)}(\eta, T - s_T) = \prod_{j \neq i} \Phi_{z_j^0(s_T)}(\eta, T - s_T) + O\left(\frac{1}{T^{\varepsilon/2}} \right). \tag{4.7}$$

For the simplicity of notation, let

$$\Phi_i^+ := \Phi_{z_i(s_T)}(\eta^+, T - s_T) \quad \text{and} \quad \Phi_i^{+0} := \Phi_{z_i^0(s_T)}(\eta^+, T - s_T),$$

and similarly Φ_i^- and Φ_i^{-0}. Since $z_1(t) = z_1^0(t)$ under the coupling, we have $\Phi_1^{\pm} = \Phi_1^{\pm 0}$. Hence by (4.7) we obtain given $\mathcal{H}_1$,

$$(\Phi_1^+ - \Phi_1^-) \prod_{j \neq 1} \Phi_j^+ = (\Phi_1^{+0} - \Phi_1^{-0}) \prod_{j \neq 1} \Phi_j^+$$

$$= (\Phi_1^{+0} - \Phi_1^{-0}) \prod_{j \neq 1} \Phi_j^{+0} + O\left(\frac{1}{T^{\varepsilon/2}} \right) (\Phi_1^{+0} - \Phi_1^{-0}).$$

Changing the role of $i = 1, \ldots, K$, we obtain given $\mathcal{H}_1$

$$\left(\prod_{j<i} \Phi_j^- \right) (\Phi_i^+ - \Phi_i^-) \left(\prod_{j>i} \Phi_j^+ \right)$$

$$= \left(\prod_{j<i} \Phi_j^{-0} \right) (\Phi_i^{+0} - \Phi_i^{-0}) \left(\prod_{j>i} \Phi_j^{+0} \right) + O\left(\frac{1}{T^{\varepsilon/2}} \right) (\Phi_i^{+0} - \Phi_i^{-0}).$$

Then using (3.4) and the same decomposition for Φ_i^{+0} and Φ_i^{-0} we obtain given $\mathcal{H}_1$,

$$\prod_{i=1}^K \Phi_i^+ - \prod_{i=1}^K \Phi_i^- = \prod_{i=1}^K \Phi_i^{+0} - \prod_{i=1}^K \Phi_i^{-0} + O\left(\frac{1}{T^{\varepsilon/2}} \right) \sum_{i=1}^K \left(\Phi_i^{+0} - \Phi_i^{-0} \right).$$

Therefore

$$\mathbf{E} \left(\psi(\eta^{+z(T)}) - \psi(\eta^{-z(T)}) \right) \mathbf{1}_\mathcal{G} = \mathbf{E} \left(\prod_{i=1}^K \Phi_i^{+0} - \prod_{i=1}^K \Phi_i^{-0} \right) \mathbf{1}_{\mathcal{G} \cap \mathcal{H}_1}$$

$$+ O\left(\frac{1}{T^{\varepsilon/2}} \right) \sum_{i=1}^K \mathbf{E} \left(\Phi_i^{+0} - \Phi_i^{-0} \right) \mathbf{1}_{\mathcal{G} \cap \mathcal{H}_1} + O(e^{-T^{\varepsilon/8}}).$$

$$(4.8)$$

By a similar discussion as we did for (4.6), we have

$$\mathbf{E} \left| \prod_{i=1}^K \Phi_i^{+0} - \prod_{i=1}^K \Phi_i^{-0} \right| \mathbf{1}_{\mathcal{G}^c} = O\left(\frac{1}{T^{\varepsilon/4}} \right) \sum_{i=1}^K \mathbf{E} \left| \Phi_i^{+0} - \Phi_i^{-0} \right|. \qquad (4.9)$$

Finally, we note that

$$\mathbf{E} \prod_{i=1}^K \Phi_{z_i^0(s_T)}(\eta, T - s_T) = \prod_{i=1}^K \mathbf{E} \, \Phi_{z_i^0(s_T)}(\eta, T - s_T) = \prod_{i=1}^K \Phi_{z_i(0)}(\eta, T), \qquad (4.10)$$

where the first equality holds since $\{z^0(t)\}_{t \geq 0}$ are independent SRWs and the second equality follows from the Markov property of each $z_i^0(t)$. Thus, by (4.5), (4.6), (4.8), (4.9) and (4.10), we conclude that

$$|\Delta_T(\eta^+) - \Delta_T(\eta^-)| \leq \mathbf{E} \left| \prod_{i=1}^K \Phi_i^{+0} - \prod_{i=1}^K \Phi_i^{-0} \right| \mathbf{1}_{\mathcal{G}^c \cup \mathcal{H}_1^c} + O\left(\frac{1}{T^{\varepsilon/2}} \right) \sum_{i=1}^K \mathbf{E} \left| \Phi_i^{+0} - \Phi_i^{-0} \right|$$

$$+ \frac{C}{T^{\varepsilon/4}} \sum_{i=1}^K \mathbf{E} \left| \eta^+(z_i(T)) - \eta^-(z_i(T)) \right| + O(e^{-T^{\varepsilon/8}})$$

$$\leq \frac{C}{T^{\varepsilon/4}} \sum_{i=1}^K \mathbf{E} \left| \eta^+(z_i(T)) - \eta^-(z_i(T)) \right| + O(e^{-T^{\varepsilon/8}}),$$

for all $T_0 \leq T \leq N^{2+\varepsilon}$, and this concludes the proof of Lemma 3.3. $\qquad \square$

A Classical results on SRW

Fix an integer $N \geq 2$, let $\mathbb{Z}_N := \mathbb{Z}/N\mathbb{Z}$ be the state space, where we identify $\mathbb{Z}_N$ with natural representatives $\{0, 1, \ldots, N-1\}$ by abusing notations. We consider the SSEP $\{\eta_t\}_{t \geq 0}$ given by the generator

$$L_E f(\eta) = \frac{1}{2} \sum_{x \in \mathbb{Z}_N} (f(\eta^{x,x+1}) - f(\eta)), \qquad \eta \in \{0, 1\}^{\mathbb{Z}_N},$$

where $f : \{0,1\}^{\mathbb{Z}_N} \to \mathbb{R}$, and the continuous-time SRW $\{w(t)\}_{t\geq 0}$ with rate $\lambda > 0$ defined by the generator

$$L_N^{0,\lambda} f(x) = \frac{\lambda}{2} \sum_{y=x\pm 1} (f(y) - f(x)), \qquad x \in \mathbb{Z}_N,$$

where $f : \mathbb{Z}_N \to \mathbb{R}$. Note that the SRW with rate λ runs as follows: a random particle has a Poisson clock with intensity λ, and it moves to one of two neighbors with equal probability at each time when the clock rings.

Lemma A.1. *Fix $\lambda > 0$. Let $\{w(t)\}_{t\geq 0}$ be a continuous-time SRW on $\mathbb{Z}_N$ with rate λ and $w(0) = 0$. Then for any $\varepsilon > 0$ and for any $T > 0$, we have*

$$\mathbb{P}\left(\max_{0 \leq t \leq T} |w(t)| \geq T^{\frac{1}{2}+\varepsilon} \right) \leq 3\exp\left(-\frac{\lambda}{6} T^{2\varepsilon} \right).$$

Moreover, if $|w(0)| = \lceil T^{\frac{1}{2}+\varepsilon} \rceil < N/2$, then for any $\varepsilon > 0$ and for any $T > 0$, we have

$$\mathbb{P}\left(\min_{0 \leq t \leq T} |w(t)| \leq \frac{1}{2}\lceil T^{\frac{1}{2}+\varepsilon} \rceil \right) \leq 3\exp\left(-\frac{\lambda}{24} T^{2\varepsilon} \right).$$

Proof. Let us consider a discrete-time SRW $\{\widetilde{S}_t\}_{t=0,1,\dots}$ on $\mathbb{Z}$ with $\widetilde{S}_0 = 0$. If we define $S_t := \widetilde{S}_t \mod N$, then $\{S_t\}_{t=0,1,\dots}$ is a discrete-time SRW on $\mathbb{Z}_N$ with $S_0 = 0$. Then the maximal inequality yields for any positive integers $M \geq 1$ and $D \geq 1$,

$$\mathbb{P}\left(\max_{0 \leq k \leq M} |S_k| \geq D \right) \leq \mathbb{P}\left(\max_{0 \leq k \leq M} |\widetilde{S}_k| \geq D \right) \leq 2e^{-\frac{D^2}{2M}}.$$

For each $T > 0$, the number of jumps n_T in time $[0,T]$ in the continuous-time SRW $\{w(t)\}_{t\geq 0}$ with rate λ has the Poisson distribution with intensity λT, and thus we have

$$\mathbb{P}(n_T \geq 3\lambda T) \leq e^{-3\lambda T} \mathbb{E}\, e^{n_T} = e^{-3\lambda T} e^{(e-1)\lambda T} \leq e^{-\lambda T}. \tag{A.1}$$

Therefore we obtain

$$\mathbb{P}\left(\max_{0 \leq t \leq T} |w(t)| \geq D \right) \leq \mathbb{P}\left(\max_{0 \leq k \leq n_T} |\widetilde{S}_k| \geq D, n_T < 3\lambda T \right) + \mathbb{P}(n_T \geq 3\lambda T)$$

$$\leq \mathbb{P}\left(\max_{0 \leq k \leq \lfloor 3\lambda T \rfloor} |\widetilde{S}_k| \geq D \right) + e^{-\lambda T} \leq 2e^{-\frac{D^2}{6\lambda T}} + e^{-\lambda T},$$

and letting $D = \lambda T^{\frac{1}{2}+\varepsilon}$ for any $T > 0$ and for any $\varepsilon > 0$ yields the first claim.

The second claim follows from the above discussion applying to a continuous-time SRW $w(t) - w(0)$ for $t \geq 0$ starting from 0; we omit the details. $\qquad\square$

For an integer $K \geq 1$ we consider the K-marked SSEP $z(t) := \{z_i(t)\}_{i=1,\dots,K}$ on $\mathbb{Z}_N$ running according to the SSEP, and the K-marked independent SRWs $z^0(t) := \{z_i^0(t)\}_{i=1,\dots,K}$ with rate λ, where each particle runs as a SRW having an independent Poisson clock with intensity λ.

We have used the following lemma in the proof of Lemma 4.1 in Section 4.

Lemma A.2. *Let $\{w(t)\}_{t\geq 0}$ be a continuous-time SRW on $\mathbb{Z}_N$ with $w(0) = 0$ and rate 2. For any $T > 0$, we define the occupation time at 0 for $\{w(t)\}_{t\geq 0}$ up to time T by*

$$\theta(T) := |\{t \in [0,T] \; : \; w(t) = 0\}|,$$

where $|A|$ stands for the normalized Lebesgue measure of a measurable set $A \subset \mathbb{R}$. Then for any $\varepsilon > 0$ there exists T_0 such that for all $T_0 \leq T \leq N^{2+\varepsilon}$,

$$\mathbb{P}\left(\theta(T) \geq 2T^{\frac{1}{2}+2\varepsilon} \right) \leq \exp\left(-T^{\varepsilon/4} \right).$$

Proof. For any positive real $\lambda > 0$, let $\{w(t)\}_{t \geq 0}$ be a continuous-time SRW on $\mathbb{Z}_N$ with $w(0) = 0$ and rate λ to which we will apply $\lambda = 2$. Since for the continuous-time SRW $\{w(t)\}_{t \geq 0}$ each waiting time for the next jump has the exponential distribution with parameter λ, the occupation time $\theta(T)$ has the distribution $\sum_{i=1}^{V_T} T_i$ where V_T is the number of visits of $\{w(t)\}_{t \geq 0}$ at 0 up to time T and T_i, $i = 1, 2, \ldots$ are independent exponential random variable with parameter λ. We use a discrete-time SRW $\{S_t\}_{t=0,1,\ldots}$ on $\mathbb{Z}_N$ with $S_0 = 0$ to obtain a bound for V_T and for the occupation time $\theta(T)$.

Since the times of jumps for $w(t)$ are given by a Poisson point process on $[0, \infty)$ with intensity λ, the number of jumps n_T up to time T satisfies that for all $T \geq 0$, by (A.1),

$$\mathbb{P}\left(n_T \geq 3\lambda T\right) \leq e^{-\lambda T}.$$

If we denote by U_n the number of visits at 0 up to time n for discrete-time SRW $\{S_t\}_{t=0,1,\ldots}$, then by (A.1),

$$\mathbb{P}\left(V_T \geq M\right) \leq \mathbb{P}\left(V_T \geq M \cap \{n_T < 3\lambda T\}\right) + \mathbb{P}\left(n_T \geq 3\lambda T\right) \leq \mathbb{P}\left(U_{\lfloor 3\lambda T \rfloor} \geq M\right) + e^{-\lambda T}.$$

Let $\{\widetilde{S}_t\}_{t=0,1,\ldots}$ be a discrete-time SRW on $\mathbb{Z}$ with $\widetilde{S}_0 = 0$ and for $x \in \mathbb{Z}$, let $\widetilde{U}_n(x)$ be the number of visits at x up to time n for $\{\widetilde{S}_t\}_{t=0,1,\ldots}$. Noting that $\widetilde{S}_t \mod N$ has the same distribution as S_t, we observe that for any positive integer $L > 0$,

$$\mathbb{P}\left(U_n \geq M\right) \leq \sum_{x \in (-L, L] \cap N\mathbb{Z}} \mathbb{P}\left(\widetilde{U}_n(x) \geq M \left\lceil \frac{2L}{N} \right\rceil^{-1}\right) + \mathbb{P}\left(\max_{1 \leq k \leq n} |\widetilde{S}_k| > L\right), \qquad \text{(A.2)}$$

where we have taken the union bound over $\lceil 2L/N \rceil$ points in $(-L, L] \cap N\mathbb{Z}$. For each $x \in \mathbb{Z}$, it holds that for any integer $m > 0$,

$$\mathbb{P}\left(\widetilde{U}_n(x) \geq m\right) \leq \mathbb{P}\left(\widetilde{U}_n(0) \geq m\right) \leq \left(1 - \frac{1}{2\sqrt{n}}\right)^{m-1} \leq 2\exp\left(-\frac{m}{2\sqrt{n}}\right), \qquad \text{(A.3)}$$

for all large enough n independent of x; indeed the first inequality holds by the stochastic domination as $\widetilde{S}_t$ starts at 0, and the second inequality follows since the first return time R to 0 after $\widetilde{S}_t$ leaves 0 satisfies that $\mathbb{P}\left(R > 2n\right) \geq (2\sqrt{n})^{-1}$ for all large enough n. Noting that the maximal inequality yields

$$\mathbb{P}\left(\max_{1 \leq k \leq n} |\widetilde{S}_k| \geq L\right) \leq 2\exp\left(-\frac{L^2}{2n}\right),$$

we have by (A.2) and (A.3) there exists n_0 such that for all integers M, N and $L > 0$ and for all large enough $n \geq n_0$,

$$\mathbb{P}\left(U_n \geq M\right) \leq \left\lceil \frac{2L}{N} \right\rceil \exp\left(-\frac{MN}{4L\sqrt{n}}\right) + 2\exp\left(-\frac{L^2}{2n}\right).$$

Summarizing the above estimates, we have that there exists T_0 (depending on n_0) such that for all $T \geq T_0$ and all integers $M, L > 0$,

$$\mathbb{P}\left(V_T \geq M\right) \leq \left\lceil \frac{2L}{N} \right\rceil \exp\left(-\frac{MN}{4L\sqrt{\lfloor 3\lambda T \rfloor}}\right) + 2\exp\left(-\frac{L^2}{2\lfloor 3\lambda T \rfloor}\right) + e^{-\lambda T}.$$

Hence fixing the parameter $\lambda = 2$, and for any $\varepsilon > 0$ letting $M = \lfloor T^{\frac{1}{2}+2\varepsilon} \rfloor$ and $L = \lfloor T^{\frac{1}{2}+\varepsilon} \rfloor$, we obtain

$$\mathbb{P}\left(V_T \geq T^{\frac{1}{2}+2\varepsilon}\right) \leq \frac{2T^{\frac{1}{2}+\varepsilon}}{N} \exp\left(-\frac{T^\varepsilon N}{4\sqrt{6T}}\right) + 2\exp\left(-\frac{T^{2\varepsilon}}{12}\right) + e^{-2T}.$$

If $T \leq N^{2+\varepsilon}$, then $T^{1-\varepsilon} \leq N^2$ and thus the first term in the right hand side is at most

$$2T^{3\varepsilon/2} \exp\left(-\frac{T^{\varepsilon/2}}{4\sqrt{6}}\right) \leq \exp\left(-T^{\varepsilon/3}\right),$$

for all large enough T. Therefore for any $\varepsilon > 0$ there exists T_0 such that for all $T_0 \leq T \leq N^{2+\varepsilon}$,

$$\mathbb{P}\left(V_T \geq T^{\frac{1}{2}+2\varepsilon}\right) \leq 3\exp\left(-T^{\varepsilon/3}\right). \tag{A.4}$$

Returning to the estimate on the occupation time $\theta(T)$, we have that for a sum of independent exponential random variables with parameter 2, for any positive integer $V \geq 1$,

$$\mathbb{P}\left(\sum_{i=1}^{V} T_i \geq 2V\right) \leq e^{-2V} \prod_{i=1}^{V} \mathbb{E}\, e^{T_i} = e^{-2V} \cdot 2^V \leq e^{-V},$$

where we have used $\mathbb{E}\, e^{T_i} = 2$ for each i. Therefore combining with (A.4), we have that for any $\varepsilon > 0$, there exists T_0' such that for all $T_0' \leq T \leq N^{2+\varepsilon}$,

$$\mathbb{P}\left(\theta(T) \geq 2T^{\frac{1}{2}+2\varepsilon}\right) \leq \mathbb{P}\left(\sum_{i=1}^{V_T} T_i \geq 2\lfloor T^{\frac{1}{2}+2\varepsilon}\rfloor, V_T < T^{\frac{1}{2}+2\varepsilon}\right) + \mathbb{P}\left(V_T \geq T^{\frac{1}{2}+2\varepsilon}\right)$$

$$\leq \exp\left(-\lfloor T^{\frac{1}{2}+2\varepsilon}\rfloor\right) + 3\exp\left(-T^{\varepsilon/3}\right) \leq \exp\left(-T^{\varepsilon/4}\right),$$

and this shows the claim. $\qquad\square$

B Proofs of miscellaneous lemmas

Lemma B.1. *Fix $\lambda > 0$. There exists a constant $C > 0$ such that the following holds: for $x, y \in \mathbb{Z}_N$, let $\{w_1(t)\}_{t\geq 0}$ and $\{w_2(t)\}_{t\geq 0}$ be continuous-time SRWs on $\mathbb{Z}_N$ with rate λ such that $w_1(0) = x$ and $w_2(0) = y$. Then for all $t > 0$,*

$$\|\mathbb{P}_x\left(w_1(t) \in \cdot\right) - \mathbb{P}_y\left(w_2(t) \in \cdot\right)\|_{\mathrm{TV}} \leq \frac{C|x-y|}{\sqrt{t}}.$$

Proof. Let us run $\{w_1(t)\}_{t\geq 0}$ and $\{w_2(t)\}_{t\geq 0}$ independently; then $w_1(t) - w_2(t)$ is a continuous-time SRW on $\mathbb{Z}_N$ with rate 2λ starting from $x - y$. Letting τ_0 be the first time $w_1(t) - w_2(t)$ to hit 0, we have that for any $t \geq 0$,

$$\|\mathbb{P}_x\left(w_1(t) \in \cdot\right) - \mathbb{P}_y\left(w_2(t) \in \cdot\right)\|_{\mathrm{TV}} \leq \mathbb{P}\left(\tau_0 > t\right).$$

If we consider a discrete-time SRW $\{S_t\}_{t=0,1,\ldots}$ on $\mathbb{Z}$ starting from $|x-y|$ and let τ_* be the first time when it hits 0, then there exists a constant $n_0 > 0$ such that for any $x, y \in \mathbb{Z}$ for $n > n_0$,

$$\mathbb{P}\left(\tau_* > n\right) \leq \frac{6|x-y|}{\sqrt{n}},$$

[LP17, Theorem 2.17]. Since for a continuous-time chain the number of jumps happening in $[0, t]$ has the Poisson distribution with intensity $2\lambda t$, and

$$\sum_{n=1}^{\infty} e^{-2\lambda} \frac{(2\lambda t)^n}{n!} \frac{1}{\sqrt{n}} = \sum_{|n-2\lambda t|\leq \lambda t} e^{-2\lambda} \frac{(2\lambda t)^n}{n!} \frac{1}{\sqrt{n}} + \sum_{|n-2\lambda t|>\lambda t} e^{-2\lambda} \frac{(2\lambda t)^n}{n!} \frac{1}{\sqrt{n}} \leq \frac{1}{\sqrt{\lambda t}} + \frac{2}{\lambda t},$$

where we have used the Chebyshev inequality in the second term, we obtain

$$\mathbb{P}\left(\tau_0 > t\right) \leq \sum_{n=1}^{\infty} e^{-2\lambda} \frac{(2\lambda t)^n}{n!} \mathbb{P}\left(\tau_* > n\right) \leq \frac{18|x-y|}{\sqrt{\lambda t}},$$

for $\lambda t \geq n_0$, and upon replacing the constant factor $18/\sqrt{\lambda}$ by a positive constant C, we have the claim as stated. □

Lemma B.2. *Fix $\lambda > 0$. Let $\{w(t)\}_{t \geq 0}$ be a continuous-time SRW on $\mathbb{Z}_N$ with rate λ such that $w(0) = 0$. Suppose that we have a sequence s_T such that $s_T/T \to 1$ as $T \to \infty$. Then for any $\varepsilon > 0$, there exists a constant C such that for all large enough T and N with $T \leq N^{2+\varepsilon}$ and for all $y \in \mathbb{Z}_N$,*

$$\mathbb{P}\left(|w(s_T) - y| \leq T^{\frac{1}{2}-2\varepsilon}\right) \leq \frac{C}{T^{\varepsilon/2}}.$$

Proof. Let $\{S_t\}_{t=0,1,\dots}$ be a discrete-time SRW on $\mathbb{Z}_N$ with $S_0 = 0$. For any subset I in $\mathbb{Z}_N$ of size $|I| \leq T^{\frac{1}{2}-2\varepsilon}$, let us fix a subset I_0 of size $|I|$ in $\{0, 1, \dots, N-1\}$ such that $I = I_0 \bmod N$, and define $\widetilde{I} := [-\lceil t^{1/2+\varepsilon}\rceil, \lceil t^{1/2+\varepsilon}\rceil] \cap (I_0 + N\mathbb{Z})$ in $\mathbb{Z}$ for $t \geq 0$. Taking a SRW $\{\widetilde{S}_t\}_{t=0,1,\dots}$ on $\mathbb{Z}$ such that $S_t = \widetilde{S}_t \bmod N$, we apply to $\widetilde{S}_t$ on $\mathbb{Z}$ the local limit theorem and the maximal inequality:

$$\mathbb{P}(S_t \in I) \leq \mathbb{P}\left(\widetilde{S}_t \in \widetilde{I}\right) + \mathbb{P}\left(\max_{0 \leq k \leq t} |\widetilde{S}_k| \geq t^{1/2+\varepsilon}\right)$$

$$\leq |I|\frac{t^{\frac{1}{2}+\varepsilon}}{N}(1+o(1))\frac{1}{\sqrt{\pi t}} + 2\exp\left(-\frac{t^{2\varepsilon}}{2}\right) \quad \text{as } t \to \infty,$$

where we have used $|\widetilde{I}| \leq |I|t^{\frac{1}{2}+\varepsilon}/N$ in the second inequality. Note that the number of jumps up to time s_T in a continuous-time SRW is λs_T with an additive error at most $\lambda s_T/2$ with probability at least $1 - 4/\lambda s_T$. Hence if $s_T/T \to 1$, then for any $\varepsilon > 0$ and for all large enough T and N with $T \leq N^{2+\varepsilon}$ (where we recall that $|I| \leq T^{\frac{1}{2}-2\varepsilon}$),

$$\mathbb{P}\left(w(s_T) \in I\right) \leq T^{\frac{1}{2}-2\varepsilon}\frac{T^{\frac{1}{2}+\varepsilon}}{N}(1+o(1))\frac{1}{\sqrt{\pi T}} + 2\exp\left(-(1-o(1))\frac{T^{2\varepsilon}}{2}\right) + O\left(\frac{1}{T}\right)$$

$$= O\left(\frac{T^{\frac{1}{2}-\varepsilon}}{N}\right) + O\left(\exp\left(-\frac{T^{2\varepsilon}}{3}\right)\right) + O\left(\frac{1}{T}\right) = O\left(\frac{1}{T^{\varepsilon/2}}\right).$$

Therefore for any $\varepsilon > 0$ for all large enough T and N with $T \leq N^{2+\varepsilon}$ and for any $y \in \mathbb{Z}_N$,

$$\mathbb{P}\left(|w(s_T) - y| \leq T^{\frac{1}{2}-2\varepsilon}\right) = O\left(\frac{1}{T^{\varepsilon/2}}\right)$$

as desired. □

References

[BBP19] Lorenzo Bertini, Paolo Buttà, and Adriano Pisante. On large deviations of interface motions for statistical mechanics models. *Ann. Henri Poincaré*, 20(6):1785–1821, 2019. MR3956160

[BPSV00a] S. Brassesco, E. Presutti, V. Sidoravicius, and M. E. Vares. Ergodicity and exponential convergence of a Glauber + Kawasaki process. In *On Dobrushin's way. From probability theory to statistical physics*, volume 198 of *Amer. Math. Soc. Transl. Ser. 2*, pages 37–49. Amer. Math. Soc., Providence, RI, 2000. MR1766341

[BPSV00b] S. Brassesco, E. Presutti, V. Sidoravicius, and M. E. Vares. Ergodicity of a Glauber + Kawasaki process with metastable states. *Markov Process. Related Fields*, 6(2):181–203, 2000. MR1778754

[CPS16] J. Theodore Cox, Yuval Peres, and Jeffrey E. Steif. Cutoff for the noisy voter model. *Ann. Appl. Probab.*, 26(2):917–932, 2016. MR3476629

[DMFL85] A. De Masi, P. A. Ferrari, and J. L. Lebowitz. Rigorous derivation of reaction-diffusion equations with fluctuations. *Phys. Rev. Lett.*, 55(19):1947–1949, 1985. MR0811605

[DMFL86] A. De Masi, P. A. Ferrari, and J. L. Lebowitz. Reaction-diffusion equations for interacting particle systems. *J. Statist. Phys.*, 44(3-4):589–644, 1986. MR0857069

[DMP91] Anna De Masi and Errico Presutti. *Mathematical Methods for Hydrodynamic Limits*, volume 1501 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1991. MR1175626

[Fer90] Pablo A. Ferrari. Ergodicity for spin systems with stirrings. *Ann. Probab.*, 18(4):1523–1538, 1990. MR1071806

[FLT19] J. Farfán, C. Landim, and K. Tsunoda. Static large deviations for a reaction-diffusion model. *Probab. Theory Related Fields*, 174(1-2):49–101, 2019. MR3947320

[GJLLV97] Davide Gabrielli, Giovanni Jona-Lasinio, Claudio Landim, and Maria Eulalia Vares. Microscopic reversibility and thermodynamic fluctuations. In *Boltzmann's legacy 150 years after his birth (Rome, 1994)*, volume 131 of *Atti Convegni Lincei*, pages 79–87. Accad. Naz. Lincei, Rome, 1997. MR1800477

[KL99] Claude Kipnis and Claudio Landim. *Scaling Limits of Interacting Particle Systems*, volume 320 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1999. MR1707314

[KOV89] C. Kipnis, S. Olla, and S. R. S. Varadhan. Hydrodynamics and large deviation for simple exclusion processes. *Comm. Pure Appl. Math.*, 42(2):115–137, 1989. MR0978701

[Lac16] Hubert Lacoin. The cutoff profile for the simple exclusion process on the circle. *Ann. Probab.*, 44(5):3399–3430, 09 2016. MR3551201

[Lac17] Hubert Lacoin. The simple exclusion process on the circle has a diffusive cutoff window. *Ann. Inst. H. Poincaré Probab. Statist.*, 53(3):1402–1437, 08 2017. MR3689972

[Lig10] Thomas M. Liggett. *Continuous Time Markov Processes: An Introduction*, volume 113 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. MR2574430

[LP17] David A. Levin and Yuval Peres. *Markov Chains and Mixing Times*. American Mathematical Society, Providence, RI, 2017. Second edition, With contributions by Elizabeth L. Wilmer, With a chapter on "Coupling from the past" by James G. Propp and David B. Wilson. MR3726904

[LS16] Eyal Lubetzky and Allan Sly. Information percolation and cutoff for the stochastic Ising model. *J. Amer. Math. Soc.*, 29(3):729–774, 2016. MR3486171