THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à l'Université Paris Dauphine

Cutoff phenomenon for some interacting particle systems

Soutenue par
Hong-Quan TRAN
Le 28 mai 2024

École doctorale n°543
SDOSE

Spécialité
Mathématiques

Préparée au
CEREMADE

Composition du jury :

Cristina TONINELLI — *Présidente du jury*
Directrice de recherche, CEREMADE

Cyril LABBÉ — *Rapporteur*
Professeur des université, Université Paris Cité

Hubert LACOIN — *Rapporteur*
Professeur, IMPA

Pietro CAPUTO — *Examinateur*
Professeur, Università Roma Tre

Nina GANTERT — *Examinatrice*
Professeur, Technische Universität München

Justin SALEZ — *Directeur de thèse*
Professeur des universités, Université Paris Dauphine

Cutoff phenomenon for some interacting particle systems

Ph.D thesis prepared by **Hong-Quan Tran**

under the supervision of **Justin Salez**

CEREMADE, PARIS DAUPHINE UNIVERSITY

To my parents

Remerciements

Je tiens en premier lieu à remercier mon directeur de thèse, Justin Salez, qui a supervisé ce travail de façon remarquable. Je suis très chanceux d'avoir été guidé par lui dans mes premiers pas de recherche, depuis le stage de M2 jusqu'à maintenant. Je lui suis reconnaissant pour sa patience et sa disponibilité, ainsi que pour l'autonomie qu'il m'a accordée. Ce serait négligent de ma part de ne pas mentionner, parmi beaucoup d'autres choses, les problèmes qu'il a bien choisis pour moi, le temps immense qu'il a consacré à relire mes brouillons, ses commentaires très pratiques et efficaces sur la rédaction et la présentation des mathématiques, et son aide pour la suite de ma carrière. Tout ce que j'apprends de lui m'accompagnera pour la suite de ma vie de chercheur. Merci pour tout, Justin !

Je tiens ensuite à remercier Hubert Lacoin et Cyril Labbé, qui ont accepté d'être les rapporteurs de cette thèse, une tâche particulièrement pénible dans un temps un peu pressé. Je suis également reconnaissant à Hubert pour son aide dans ma recherche de postdoc. J'ai aussi beaucoup apprécié les discussions diverses avec Cyril, notamment lors du groupe de travail sur le temps de mélange. Je tiens également à exprimer ma reconnaissance et ma gratitude à Cristina Toninelli, Nina Gantert et Pietro Caputo d'avoir accepté de faire partie de mon jury. Je suis heureux que mon travail puisse avoir le privilège de votre attention. De plus, merci à Cristina d'avoir organisé le groupe de travail avec Justin, ce qui m'a donné l'opportunité de discuter avec d'autres chercheurs ayant un sujet proche du mien.

Je tiens ensuite à exprimer ma gratitude à Jean-François Le Gall pour son aide dans ma recherche de postdoc, ainsi que pour ses cours lumineux qui m'ont éclairé au moment le plus inattendu dans ma thèse.

Je suis reconnaissant au CEREMADE pour l'environnement de travail stimulant et accueillant. Durant mes années de thèse, j'ai eu le plaisir d'être chargé de TD des cours donnés par Stéphane Mischler, Juliette Bouhours, Joseph Lehec, Pierre Cardaliaguet, Marc Hoffmann, François Simenhaus, et Justin lui-même. J'ai sans doute beaucoup appris sur l'enseignement et ainsi sur la responsabilité en travaillant avec vous. Un merci particulier à Pierre pour son aide dans ma recherche de postdoc. J'ai également pris beaucoup de plaisir à discuter de sujets divers avec les membres du CEREMADE, notamment avec Mathieu Lewin et Clémence Cosco. Je tiens également à exprimer ma gratitude aux secrétaires du CEREMADE: Isabelle Bellier, Anne-Laure Chagnon et César Faivre pour leurs aides rapides et efficaces chaque fois que je les sollicitais. Ça me fait toujours plaisir de vous déranger.

Un immense et chaleureux merci à tous les camarades doctorants et post-doctorants du CEREMADE, avec qui j'ai partagé beaucoup de souvenirs pendant mes années de thèse. J'ai eu un immense plaisir d'avoir partagé le bureau avec Adrien Hairault, Adrien Serguret, Claudia, Amirali, Luca et Thaddeus. Merci à Adrien Hairault de m'avoir aidé dans diverses tâches, que ce soit pour déclarer mes impôts, trouver un logement ou créer un site personnel, etc. Merci à Lucas Rey pour des belles discussions que l'on a eu, et même pour les solutions de certaines (beaucoup) questions de mathématiques. J'ai eu grand plaisir à discuter de mon thème de recherche avec Brune, Alexandre et Ons. Merci à Lorenzo, qui m'a donné goût à certains auteurs français et également à son cheesecake incroyable. Merci à Théo, qui a peut-être partagé avec moi le plus grand nombre de pauses-déjeuners. Merci à Yueh-Sheng et Adéchola pour leur accompagnement dans les démarches personnelles et professionnelles. Merci à Luke, Ruihua, Jean, Changqing et Umberto pour les séances de sport que nous avons faites ensemble. Merci à Claudia, Kathi et Giovanni pour les séances mémorables de jeux de société. Merci à Charly, Daniele, Florin, Tommy, João, Raphaël(s), Grégoire, Amal, Lucas(s), Antoine(s), César, Tommy, Richard, Jean-Paul, Hugo et Lukasz pour les moments conviviaux passés ensemble.

Je remercie également mon club de ping-pong ATT XV. Merci à eux pour les affrontements sportifs féroces mais aussi amicaux et amusants, et également pour la convivialité partagée.

Les dernières lignes sont dédiées à ma famille et mes amis. Merci à eux pour leur soutien précieux et les bons moments que nous avons passés ensemble tout au long de cette thèse. Sans eux, ce travail n'aurait pas pu être réalisé.

Résumé

Sur un espace d'états fini, une chaîne de Markov irréductible à temps continu converge vers sa mesure stationnaire unique, ou en d'autres termes, *se mélange*. La convergence est mesurée par rapport à la distance en variation totale. Dans la théorie moderne des chaînes de Markov, nous nous intéressons aux chaînes où l'espace d'états devient grand. En étudiant certains modèles de mélange de cartes, Aldous, Diaconis et Shashahani ont découvert le phénomène remarquable maintenant connu sous le nom de *cutoff*: lorsque l'espace d'états devient grand, la distance entre la chaîne et l'équilibre reste proche de sa valeur maximale pendant une longue période, puis chute soudainement vers zéro sur une échelle de temps beaucoup plus courte. Depuis, le phénomène de cutoff a été observé dans de nombreux contextes différents, tels que les chaînes de naissance et de mort, les systèmes de spin à haute température, les systèmes de particules en interaction, etc. Malgré l'accumulation de modèles, il n'existe pas encore de théorie générale permettant de prédire efficacement cutoff. Au lieu de cela, le cutoff est montré modèle par modèle.

Dans cette thèse, nous étudions trois systèmes de particules en interaction: le processus d'exclusion unidimensionnel avec réservoirs, le processus de Glauber-Exclusion dans le régime à haut température, et le processus de Zero-Range à champ-moyen avec potentiel croissant sous-linéairement. Pour chaque modèle, nous établissons cutoff et fournissons une estimation fine pour le trou spectral. Nous nous concentrons particulièrement sur le cadre de la percolation de l'information introduit par Lubetzky et Sly, qui nous permet de montrer le cutoff même sans connaître la formule explicite de la mesure invariante.

Mots clés : temps de mélange, trou spectral, phénomène de cutoff, dynamique de Glauber, processus d'Exclusion, processus de Zero-Range, percolation de l'information

Abstract

On a finite state space, an irreducible continuous-time Markov chain converges to its unique stationary measure, or in other words, *mixes*. The convergence is often measured by the total variation distance. In the modern theory of Markov Chains, we are interested in the case where the state space becomes large. When studying some models of card shuffling, Aldous, Diaconis, and Shashahani discovered a remarkable phenomenon now known as *cutoff*: as the state space becomes large, the distance between the chain and equilibrium stays close to its maximal value for a long time and then suddenly drops to near zero in a much shorter time scale. Since then, the cutoff phenomenon has been observed in many different contexts, such as birth and death chains, high-temperature spin systems, interacting particle systems, etc. Despite the accumulation of models, there is not yet a general theory to effectively predict cutoff. Instead, cutoff is proved model by model.

In this thesis, we study three models: the one-dimensional Exclusion process with reservoirs, the Glauber-Exclusion process in the high-temperature regime, and the mean-field Zero-Range process with increasing sublinear potential. These three models all fall under the category of interacting particle systems. For each model, we establish cutoff and provide a sharp estimate on the spectral gap. We particularly focus on the information percolation framework introduced by Lubetzky and Sly, which allows us to show cutoff even without knowing the explicit formula of the invariant measure.

Keywords : mixing times, spectral gap, cutoff phenomenon, Glauber dynamics, Exclusion process, Zero-Range process, information percolation

Contents

Chapter 1

Introduction

Goals

In this chapter, we briefly introduce the study of mixing times and the cutoff phenomenon. Then, we provide one of the simplest examples: the simple random walk on the hypercube. Though simple, this model illustrates the field of study very well. Furthermore, the method presented here is the base to build on the proofs for other models of interest, two of which are presented in Chapter 2 and 3.

Contents

1 Preliminaries

First, we provide some backgrounds on continuous-time Markov chains on finite state spaces. Since there are already many books about this subject, we will be very brief. Most of the materials here can be found in Chapter 2 of the book [62].

Let $\mathcal{X}$ be a finite set, which we call the *state space*. The elements of $\mathcal{X}$ are called *states* or *configurations*. Let $\mathfrak{X}$ be the set of right continuous functions $\omega : [0, \infty) \to \mathcal{X}$, with finitely many jumps in any finite interval. $\mathfrak{X}$ is given the smallest σ-algebra $\mathcal{F}$ such that the mapping $\omega \mapsto \omega(t)$ is measurable for each $t \geq 0$. Let the process $X = (X_t)_{t\geq 0}$ be defined by

$$X_t(\omega) = \omega(t).$$

Markov chains. A continuous-time Markov chain on $\mathcal{X}$ is given by the following.

1. A right continuous filtration $(\mathcal{F}_t)_{t\geq 0}$, to which X is adapted, and
2. A collection of probability measure $\{\mathbb{P}_x, x \in \mathcal{X}\}$ on $\mathfrak{X}$,

satisfying

$$\forall x \in \mathcal{X},\ \mathbb{P}_x [X_0 = x] = 1,$$

and, for any bounded measurable function φ on $\mathfrak{X}$,

$$\forall x \in \mathcal{X}, \forall s \geq 0, \mathbb{E}_x [\varphi(X_{s+\cdot})|\mathcal{F}_s] = \mathbb{E}_{X_s} [\varphi],\ \mathbb{P}_x \text{ almost surely,} \tag{1.1}$$

with $(X_{s+\cdot}(t))_{t\geq 0} = (X_{s+t})_{t\geq 0}$. Naturally, $\mathbb{P}_x$ is called the law of the process starting at configuration x. Condition (1.1) is called the *Markov property*.

A *generator* $\mathcal{L}$ on $\mathcal{X}$, sometimes known under the name Q-matrix on $\mathcal{X}$, is a square matrix indexed by elements of $\mathcal{X}$ such that

$$\forall x, y \in \mathcal{X}, x \neq y,\ \mathcal{L}(x, y) \geq 0,$$
$$\forall x \in \mathcal{X},\ \mathcal{L}(x, x) = -\sum_{y\in\mathcal{X}\setminus\{x\}} \mathcal{L}(x, y).$$

Probabilistic description. For any $\theta \in \mathbb{R}_{>0}$, we denote by $\exp(\theta)$ the exponential distribution with parameter θ, i.e. the distribution with density $\mathbb{1}_{\{s>0\}}\theta e^{-\theta s}\mathrm{d}s$. For any Markov chain $(X_t)_{t\geq 0}$, one can show that there exists a unique generator $\mathcal{L}$ such that X has the following probabilistic description: under $\mathbb{P}_x$, there exists a random time $\tau \sim \exp(|\mathcal{L}(x, x)|)$ such that

1. X stays at x in the time interval $[0, \tau)$, i.e. $X_t = x$, $\forall t \in [0, \tau)$.
2. At time τ, X jumps randomly to a new state different from x, with

$$\forall y \neq x,\ \mathbb{P}_x [X_\tau = y] = \frac{\mathcal{L}(x, y)}{|\mathcal{L}(x, x)|}.$$

A Markov process X is uniquely determined by its generator $\mathcal{L}$, and vice versa, given a generator $\mathcal{L}$, we can define a Markov process X by the above probabilistic description. Consequently, to define a Markov process X, one only need to provide its generator. We call a function $\varphi : \mathcal{X} \to \mathbb{R}$ an *observable*. Usually, the generator $\mathcal{L}$ is given by its action on the set of observables: for any observable φ, $\mathcal{L}\varphi$ is an observable defined by,

$$\forall x \in \mathcal{X},\ \mathcal{L}\varphi(x) = \sum_{y \in \mathcal{X}} \mathcal{L}(x, y) \left(\varphi(y) - \varphi(x)\right).$$

From now until the end of this section, we fix a Markov process $\mathcal{X}$ with some generator $\mathcal{L}$. For any probability distribution μ on $\mathcal{X}$, the law of the process starting from μ is

$$\mathbb{P}_\mu := \sum_{x \in \mathcal{X}} \mu(x) \mathbb{P}_x.$$

One can prove that if the chain starts with a law μ, then the law at time any time $t > 0$, i.e. $\mathbb{P}_\mu \left[X_t \in \cdot\right]$, is given by $\mu e^{t\mathcal{L}}$, where

$$e^{t\mathcal{L}} := \sum_{k=0}^{\infty} \frac{t^k \mathcal{L}^k}{k!}.$$

A probability measure π is *invariant* if $\pi\mathcal{L} = 0$, since it implies that

$$\forall t \geq 0, \mathbb{P}_\pi \left[X_t \in \cdot\right] = \pi e^{t\mathcal{L}} = \pi.$$

We say that $\mathcal{L}$ is *reversible* with respect to a probability measure π if

$$\forall x, y \in \mathcal{X}, \pi(x)\mathcal{L}(x, y) = \pi(y)\mathcal{L}(y, x).$$

If such a measure π exists, it is not hard to see that π is also invariant. We say that $\mathcal{L}$ is *irreducible* if, for any initial state x, the chain can eventually reach every state y with strictly positive probability. This means that for any $x, y \in \mathcal{X}$, there exists a sequence $(x_i)_{0 \leq i \leq k}$ for some $k \in \mathbb{Z}_+$ such that $x_0 = x, x_k = y$, and $\mathcal{L}(x_i, x_{i+1}) > 0$, $\forall 0 \leq i \leq k-1$. When the chain is irreducible, there exists a unique invariant probability distribution π, and the distribution of the process converges pointwise to π when the time tends to infinity, regardless of the initial condition:

$$\forall x, y \in \mathcal{X}, \mathbb{P}_x \left[X_t = y\right] \xrightarrow[t \to \infty]{} \pi(y).$$

The convergence is often measured with respect to the *total variation distance*, defined by, for any two probability distributions μ, ν on $\mathcal{X}$,

$$\mathrm{d}_{\mathrm{TV}}\left(\mu, \nu\right) = \max_{A \subset \mathcal{X}} |\mu(A) - \nu(A)|.$$

The speed of convergence is quantified by the *mixing times*:

$$\mathrm{t}_{\mathrm{mix}}(x; \epsilon) := \inf \left\{t \geq 0 : \mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x \left[X_t \in \cdot\right], \pi\right) \leq \epsilon\right\},\quad x \in \mathcal{X}, \epsilon \in (0, 1).$$

Let $\mathcal{D}(\cdot)$ be the worst-case distance to equilibrium:

$$\forall t \geq 0,\ \mathcal{D}(t) := \max_{x\in\mathcal{X}} \mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \pi\right).$$

The worst-case mixing times defined by:

$$\mathrm{t}_{\mathrm{mix}}(\epsilon) := \inf\left\{t \geq 0 : \mathcal{D}(t) \leq \epsilon\right\},\ \ \epsilon \in (0,1).$$

This means that $\mathrm{t}_{\mathrm{mix}}(\epsilon)$ is the time for the chain from any initial state to get within distance ϵ from the stationary distribution π. We write $\mathrm{t}_{\mathrm{mix}}$ for $\mathrm{t}_{\mathrm{mix}}(1/4)$. Understanding the mixing times is the central goal of the modern theory of Markov chains.

One can prove that if the generator $\mathcal{L}$ is irreducible, then 0 is an eigenvalue of $\mathcal{L}$ with multiplicity 1, and the other eigenvalues of $\mathcal{L}$ have negative real parts. A quantity of interest is the smallest number among the absolute values of the real parts of the non-zero eigenvalues of $\mathcal{L}$, also called the *spectral gap* of $\mathcal{L}$, denoted by $\mathtt{gap}$, which governs the asymptotic speed of convergence to equilibrium:

$$\mathtt{gap} = \lim_{t\to\infty} -\frac{1}{t}\log\max_{x\in\mathcal{X}} \mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \pi\right). \tag{1.2}$$

It is worth noting that, for any observable φ,

$$\mathtt{gap} \leq \inf_{x,y}\liminf_{t\to\infty} -\frac{1}{t}\log\left|\mathbb{E}_x\left[\varphi(X_t)\right] - \mathbb{E}_y\left[\varphi(X_t)\right]\right|. \tag{1.3}$$

Asymptotic notations. For two sequences of positive numbers $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$, we write $a_n = \mathcal{O}\left(b_n\right)$ (resp. $\Omega(b_n), \Theta(b_n), o(b_n)$) if a_n is upper bounded by (resp. lower bounded by, upper and lower bounded by, negligible compared to) b_n up to a constant factor independent of n.

The cutoff phenomenon. In the modern theory of Markov chains, we are very often interested in studying the mixing times when the state space becomes large. We are thus led to study the mixing times of a family of chains with generators $(\mathcal{L}_n)_{n\geq 1}$ on a family of state spaces $(\mathcal{X}_n)_{n\geq 1}$, where the size of $\mathcal{X}_n$ grows to infinity. For many natural families of chains, a remarkable phase transition known as *cutoff* occurs:

$$\forall \epsilon \in (0,1) \text{ fixed },\ \frac{\mathrm{t}^{(n)}_{\mathrm{mix}}(\epsilon)}{\mathrm{t}^{(n)}_{\mathrm{mix}}(1-\epsilon)} \xrightarrow[n\to\infty]{} 1.$$

In words, the asymptotic of $\mathrm{t}^{(n)}_{\mathrm{mix}}(\epsilon)$ does not depend on ϵ anymore. This implies that the distance to equilibrium stays near 1 for a long time and then suddenly drops to 0 in a much shorter time scale. We say that a family of chains exhibits cutoff with a window of size $\mathcal{O}\left(w_n\right)$

if $w_n = \mathcal{O}\left(\mathrm{t}_{\mathrm{mix}}^{(n)}\right)$ and

$$\lim_{\alpha\to\infty} \liminf_{n\to\infty} \mathcal{D}_n\left(\mathrm{t}_{\mathrm{mix}}^{(n)} - \alpha w_n\right) = 1,$$
$$\lim_{\alpha\to\infty} \limsup_{n\to\infty} \mathcal{D}_n\left(\mathrm{t}_{\mathrm{mix}}^{(n)} + \alpha w_n\right) = 0.$$

Since its discovery by Aldous, Diaconis, and Shahshahani in the context of card shuffling in [2, 21, 23], cutoff has been observed in various contexts: birth and death chains [22, 15, 27], random walks on finite groups [94, 23], high-temperature spin systems [68, 67, 69, 66, 98], interacting particle systems [50, 51, 52, 40, 87, 48, 49, 29, 13], or random walks on various models of sparse random graphs [38, 81, 64, 8, 7]. Despite the accumulating of models shown to exhibit cutoff, there is not yet a theory to effectively predict whether a family of chains exhibit cutoff (though a first step in this direction is undertaken in [85]). We refer the readers to the book [59] for more backgrounds on the mixing times and the cutoff phenomenon, and to the books [56, 62, 31] for more backgrounds on continuous-time Markov chains.

Our contribution. We add three examples to the long list of models exhibiting cutoff: the non-reversible one-dimensional Exclusion process with reservoirs (Chapter 2), the Glauber-Exclusion process (Chapter 3), and the mean-field Zero-Range process with sublinear increasing potential function (Chapter 4).

Structure of the thesis. In the rest of this chapter, we present one of the simplest examples: the simple random walk on the hypercube. It will be used to illustrate the central questions in the fields. Furthermore, the materials we presented here serve as the base to develop more complicated methods in Chapter 2 and Chapter 3.

2 The random walk on the hypercube

From now until the end of this chapter, let $\mathcal{X} = \{0,1\}^N$ be the set of all binary vectors of length N. We denote by $\leq$ the usual coordinate-wise order on $\mathcal{X}$. For any $x \in \mathcal{X}$ and $u \in [N]$, we denote by $x^{u,0}$ (resp. $x^{u,1}$) the configuration obtained by replacing the $u-$th coordinate of x by 0 (resp. 1). We consider the generator $\mathcal{L}$ which acts on any observable $\varphi : \mathcal{X} \to \mathbb{R}$ by

$$\mathcal{L}\varphi(x) = \sum_{u=1}^{N} \left(\varphi(x^{u,1-x(u)}) - \varphi(x)\right). \tag{1.4}$$

This means that each coordinate flips at rate 1, and the coordinates evolve independently. The hypercube $\mathcal{X}$ is naturally equipped with a graph structure, where two different configurations x, y are adjacent if they differ exactly at one coordinate. Then, the chain with generator $\mathcal{L}$ is a random walk on $\mathcal{X}$, where at rate N, the walk chooses a site uniformly among its neighbors and jumps there, hence the name "the random walk on the hypercube". This model is classical, and we do not claim any credit for the results presented here.

Reversibility. The generator $\mathcal{L}$ is irreducible and reversible with respect to the product law $\pi = \mathcal{B}_{1/2}^{\otimes N}$. Here $\mathcal{B}_p$ denotes the Bernoulli random variable with parameter $p \in [0,1]$.

The rest of this chapter is dedicated to proving the following results.

Theorem 1.1 (Sub-optimal upper bound)**.** *There exists a constant κ such that for any $\epsilon \in (0,1)$ fixed,*

$$\mathrm{t_{mix}}(\epsilon) \leq \frac{\log N}{2} + \kappa \log(1/\epsilon). \tag{1.5}$$

Theorem 1.2 (Spectral gap)**.** $\mathtt{gap} = 2$.

Theorem 1.3 (Cutoff for the hypercube)**.** *Fix $\epsilon \in (0,1)$. There exist constants $\kappa_1(\epsilon), \kappa_2(\epsilon)$ such that*

$$\frac{\log N}{4} - \kappa_1(\epsilon) \leq \mathrm{t_{mix}}(\epsilon) \leq \frac{\log N}{4} + \kappa_2(\epsilon).$$

In other words, the model exhibits cutoff at time $\frac{\log N}{4}$ with cutoff window $\mathcal{O}(1)$.

The three theorems above are presented in order of difficulty of their proofs. The upper bound in Theorem 1.1 is off by a factor 2 compared to the optimal bound in Theorem 1.3. However, proving the optimal upper bound is often the most challenging step to establish cutoff. In fact, for many models conjectured to exhibit cutoff, the existing upper and lower bounds are different by a factor of 2, and the lower bound is often conjectured to be optimal, see e.g. [102]. Showing cutoff then means getting rid of this extra factor 2, and some cases have been tackled only very recently, see e.g. [50, 52]. Nevertheless, though suboptimal, Theorem 1.1 already gives us the correct magnitude of the mixing times.

3 The grand coupling

Reformulation of the generator. It is more convenient to rewrite (1.4) differently as follows

$$\mathcal{L}\varphi(x) = 2\sum_{i=1}^{N}\left[\frac{1}{2}\left(\varphi(x^{u,0}) - \varphi(x)\right) + \frac{1}{2}\left(\varphi(x^{u,1}) - \varphi(x)\right)\right]. \tag{1.6}$$

Reinterpretation. With this reformulation, the model can be interpreted as follows. The value at each site, at rate 2, is replaced by an independent $\mathcal{B}_{1/2}$. The advantage of this interpretation is made more precise below.

Graphical construction. Let $\Xi = (\Xi_u)_{1\leq u\leq N}$ be N independent homogeneous Poisson processes of intensity 2, and let $(\xi_k)_{k\geq 1}$ be a sequence of i.i.d. Bernoulli variable $\mathcal{B}_{1/2}$, independent of Ξ. Given Ξ, $(\xi_k)_{k\geq 1}$, and an initial configuration $x \in \mathcal{X}$, we construct the process $(X_t^x)_{t\geq 0}$ as follows.

- The process $(X_t^x)_{t\geq 0}$ is a piecewise constant process starting at x which can only jump when Ξ jumps.

- Whenever Ξ_u jumps, say it is the $k-$th jump of Ξ, replace the u-th coordinate of X^x by ξ_k.

Then, the process $(X_t^x)_{t\geq 0}$ is a Markov process with generator $\mathcal{L}$ starting from x, or in other words, $(X_t^x)_{t\geq 0}$ follows the law $\mathbb{P}_x$.

The grand coupling. We can use the same Poisson process Ξ and the same independent Bernoulli variables $(\xi_k)_{k\geq 1}$ to construct the process from any configuration $x \in \mathcal{X}$. In other words, we construct the coupling of $(X^x)_{x\in\mathcal{X}}$, where we update the same site simultaneously using the same Bernoulli variable. One important observation is that in this coupling, once a site is updated, its values in all processes will coincide since then. Moreover, this coupling preserves order, i.e. $x \leq x' \Rightarrow \forall t \geq 0, X_t^x \leq X_t^{x'}$.

Explicit measure at time t. Thanks to the independence of the coordinates, the law of the system at any time t is a product law given by

$$\mathbb{P}_x\left[X_t \in \cdot\right] = \bigotimes_{i=1}^{N} \mathcal{B}_{1/2+e^{-2t}(x_i-1/2)}. \tag{1.7}$$

4 The suboptimal upper bound and the spectral gap

We now prove Theorem 1.1 and Theorem 1.2.

Proof. We use the grand coupling to construct the process from any configuration. Let x, y be arbitrary configurations. Recall that $(X_t^x)_{t\geq 0}$ and $(X_t^y)_{t\geq 0}$ denote the processes starting from x and y, respectively. Let $\mathbf{0}, \mathbf{1}$ denote the all-zero and all-one configurations, respectively. By monotonicity, almost surely,

$$\forall t \geq 0,\ X_t^{\mathbf{0}} \leq X_t^x, X_t^y \leq X_t^{\mathbf{1}}.$$

This means that

$$\begin{aligned}
\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}\left[X_t^x \in \cdot\right], \mathbb{P}\left[X_t^y \in \cdot\right]\right) &\leq \mathbb{P}\left[X_t^x \neq X_t^y\right] \\
&\leq \mathbb{P}\left[X_t^{\mathbf{1}} \neq X_t^{\mathbf{0}}\right] \\
&= \mathbb{P}\left[\exists u \in [N] : \Xi_u(t) = 0\right] \\
&\leq \sum_{u\in[N]} \mathbb{P}\left[\Xi_u(t) = 0\right] \\
&= Ne^{-2t}.
\end{aligned}$$

Here, we have used a union bound in the last inequality and the fact that $\Xi_u(t)$ is a Poisson variable with parameter $2t$ in the last equality. By convexity of the total variation distance, this implies

$$\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}\left[X_t^x \in \cdot\right], \pi\right) \leq Ne^{-2t}. \tag{1.8}$$

We can take $t = \dfrac{\log N + \log(1/\epsilon)}{2}$ to make the last expression equal ϵ, which finishes the proof of Theorem 1.1. Moreover, (1.8) and (1.2) imply that

$$\texttt{gap} \geq 2.$$

On the other hand, for arbitrary $u \in [N]$, by (1.3), for $\varphi : x \mapsto x(u)$,

$$\texttt{gap} \leq \lim_{t\to\infty} -\frac{1}{t} \log \left| \mathbb{E}\left[X_t^{\mathbf{1}}(u)\right] - \mathbb{E}\left[X_t^{\mathbf{0}}(u)\right] \right| = \lim_{t\to\infty} -\frac{1}{t} \log \mathbb{P}\left[\Xi_u(t) = 0\right] = 2.$$

Hence, $\texttt{gap} = 2$, which finishes the proof of Theorem 1.2. □

5 The lower bound

In this section, we prove the lower bound in Theorem 1.3. Let $X_\infty \sim \pi$ (then $X_t \xrightarrow[t\to\infty]{d} X_\infty$). We use the method known as *distinguishing statistics*, see [59], Chapter 7 for background on the method. In short, to prove that the laws of X_t and X_∞ are far apart, we only need to prove that the law of $\varphi(X_t)$ and $\varphi(X_\infty)$ are far apart, for some observable φ. The initial complicated problem of comparing some high-dimensional laws becomes comparing the law of two real random variables, which is much easier. It then remains to choose a suitable observable φ to obtain the lower bound. In this example, the comparison is carried out using the first and second moments method. Surprisingly, this simple method often gives the correct estimate of the mixing times.

Proof. We consider the process $(X_t^{\mathbf{1}})_{t\geq 0}$. It is natural to consider the observable φ given by

$$\forall x \in \mathcal{X},\ \varphi(x) = \sum_{u=1}^{N} x(u).$$

By (1.7), we have

$$\mathbb{E}\left[\varphi(X_t^{\mathbf{1}})\right] = \frac{N}{2} + \frac{N}{2} e^{-2t}, \tag{1.9}$$

$$\mathrm{Var}\left[\varphi(X_t^{\mathbf{1}})\right] = N \frac{1 - e^{-2t}}{2} \frac{1 + e^{-2t}}{2}. \tag{1.10}$$

By letting t tend to infinity,

$$\mathbb{E}\left[\varphi(X_\infty)\right] = \frac{N}{2}, \tag{1.11}$$

$$\mathrm{Var}\left[\varphi(X_\infty)\right] = \frac{N}{4}. \tag{1.12}$$

Hence, $\mathrm{Var}\left[\varphi(X_t^{\mathbf{1}})\right]$ and $\mathrm{Var}\left[\varphi(X_\infty)\right]$ are both of magnitude N. So if

$$\mathbb{E}\left[\varphi(X_t^{\mathbf{1}})\right] - \mathbb{E}\left[\varphi(X_\infty)\right] \gg \sqrt{N},$$

then the system is not mixed yet. More precisely, by Proposition 7.9 in [59],

$$\left\|\mathbb{P}\left[X_t^{\mathbf{1}} \in \cdot\right] - \pi\right\|_{TV} \geq 1 - 8\frac{\max\left\{\mathrm{Var}\left[\varphi(X_t^{\mathbf{1}})\right], \mathrm{Var}\left[\varphi(X_\infty)\right]\right\}}{\left(\mathbb{E}\left[\varphi(X_t^{\mathbf{1}})\right] - \mathbb{E}\left[\varphi(X_\infty)\right]\right)^2}.$$

By (1.10) and (1.12),

$$\max\left\{\mathrm{Var}\left[\varphi(X_t^{\mathbf{1}})\right], \mathrm{Var}\left[\varphi(X_\infty)\right]\right\} = N/4.$$

Moreover, by (1.9) and (1.11),

$$\mathbb{E}\left[\varphi(X_t^{\mathbf{1}})\right] - \mathbb{E}\left[\varphi(X_\infty)\right] = \frac{N}{2}e^{-2t}.$$

Therefore,

$$\left\|\mathbb{P}\left[X_t^{\mathbf{1}} \in \cdot\right] - \pi\right\|_{TV} \geq 1 - \frac{8e^{4t}}{N}. \tag{1.13}$$

For $t = \frac{1}{4}\left(\log N - \log\left(\frac{1}{1-\epsilon}\right) - \log 8\right)$, the right-hand side of (1.13) is equal to ϵ. Hence,

$$\mathrm{t_{mix}}(\epsilon) \geq \frac{1}{4}\left(\log N - \log\left(\frac{1}{1-\epsilon}\right) - \log 8\right),$$

which finishes our proof. □

6 The upper bound

First, we recall a result about perturbation of a product measure.

Lemma 1.4 (Perturbation of a product measure)**.** *Let $\Omega = \{0,1\}^n$ for some number n. For each subset $E \subset [n]$, let φ_E be a distribution on $\{0,1\}^E$. Let $p \in (0,1)$, and let ν be the product measure $\mathcal{B}_p^{\otimes n}$ on Ω. Let μ be the measure on Ω obtained by first sampling a subset $E \subset [n]$ via some measure $\tilde{\mu}$, and then, conditionally on E, generating independently the values on E via φ_E and the values on $[n] \setminus E$ via $\mathcal{B}_p^{\otimes [n]\setminus E}$. Then*

$$4\mathrm{d}_{TV}(\mu,\nu)^2 \leq \left\|\frac{\mu}{\nu} - 1\right\|_{L^2(\nu)}^2 \leq \mathbb{E}\left[\theta^{|E\cap E'|}\right] - 1, \tag{1.14}$$

where E, E' are i.i.d. with law $\tilde{\mu}$, and $\theta = \max\left\{\frac{1}{p}, \frac{1}{1-p}\right\}$.

Roughly speaking, Lemma 1.4 says that if we perturb the product measure by first choosing a set $E \subset [n]$ randomly and then perturbing the values of the sites in E by a prescribed law φ_E, then the perturbed measure is still close to the product measure if the perturbed region is *small* and *delocalized* (measured by the exponential moments at the end of (1.14)).

Lemma 1.4 seems to be one of the most versatile inequalities in proving cutoff. Originally introduced by Miller and Peres in [76] to prove cutoff for the lamplighter chain, it has been used to prove cutoff for the Glauber dynamics of the Ising model by Lubetzky and Sly in [67, 68, 69], for the Exclusion with reservoirs on any network by Salez in [87], for the non-reversible one-dimensional SSEP with reservoirs in [99] (see Chapter 2), and for the Glauber-Exclusion process at high temperature in [98] (see Chapter 3). We expect the list will continue to grow. For the proof of Lemma 1.4, see Lemma 1 in [87]. Now we are ready to prove the upper bound in Theorem 1.3.

Proof of the upper bound in Theorem 1.3. We consider the process starting from an arbitrary configuration $x \in \mathcal{X}$. We construct our process using the graphical construction described above. Let t be a number that we choose later. Let Red and Blue be defined by:

$$\begin{aligned} \mathsf{Red} &:= \{u \in [N] : \Xi_u(t) = 0\}, \\ \mathsf{Blue} &:= [N] \setminus \mathsf{Red}. \end{aligned}$$

Note that, conditionally on Ξ,

$$\begin{aligned} X_t^x(\mathsf{Blue}) &\sim \mathcal{B}_{1/2}^{\otimes \mathsf{Blue}}, \\ X_t^x(\mathsf{Red}) &= x(\mathsf{Red}), \end{aligned}$$

and moreover, $X_t^x(\mathsf{Red})$ and $X_t^x(\mathsf{Blue})$ are independent. This means that the law of X_t^x can be regarded as a product measure perturbed at the sites in Red. Hence, we can apply Lemma 1.4 to conclude that

$$4\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \pi\right)^2 \le \mathbb{E}\left[2^{|\mathsf{Red} \cap \mathsf{Red}'|}\right] - 1,$$

where Red' is an independent copy of Red. Note that, by definition, the variables $(\mathbb{1}_{\{u \in \mathsf{Red}\}})_{u \in [N]}$, $(\mathbb{1}_{\{u \in \mathsf{Red}'\}})_{u \in [N]}$ are independent. Therefore

$$\begin{aligned} \mathbb{E}\left[2^{|\mathsf{Red} \cap \mathsf{Red}'|}\right] &= \mathbb{E}\left[2^{\sum_{u \in [N]} \mathbb{1}_{\{u \in \mathsf{Red}\}} \mathbb{1}_{\{u \in \mathsf{Red}'\}}}\right] \\ &= \prod_{u \in [N]} \mathbb{E}\left[2^{\mathbb{1}_{\{u \in \mathsf{Red}\}} \mathbb{1}_{\{u \in \mathsf{Red}'\}}}\right] \\ &= \prod_{u \in [N]} \mathbb{E}\left[1 + \mathbb{1}_{\{u \in \mathsf{Red}\}} \mathbb{1}_{\{u \in \mathsf{Red}'\}}\right] \\ &= \prod_{u \in [N]} \left(1 + \mathbb{P}\left[u \in \mathsf{Red}\right] \mathbb{P}\left[u \in \mathsf{Red}'\right]\right) \\ &= \prod_{u \in [N]} \left(1 + \mathbb{P}\left[u \in \mathsf{Red}\right]^2\right). \end{aligned}$$

Note that, by definition,

$$\mathbb{P}\left[u \in \mathsf{Red}\right] = \mathbb{P}\left[\Xi_u(t) = 0\right] = e^{-2t}. \tag{1.15}$$

Therefore,

$$\mathbb{E}\left[2^{|\mathsf{Red} \cap \mathsf{Red}'|}\right] = \prod_{u \in [N]} \left(1 + e^{-4t}\right) = (1 + e^{-4t})^N \le e^{Ne^{-4t}}. \tag{1.16}$$

(1.15) and (1.16) together implies

$$4\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \pi\right)^2 \leq e^{Ne^{-4t}} - 1. \tag{1.17}$$

For $t = \frac{1}{4}\left(\log N + 2\log(1/\epsilon)\right)$,

$$e^{Ne^{-4t}} = e^{\epsilon^2} \leq 1 + e\epsilon^2 \leq 1 + 4\epsilon^2, \tag{1.18}$$

where we have used the inequality $e^{\theta} \leq 1 + e\theta, \forall \theta \in (0,1)$. Hence,

$$e^{Ne^{-4t}} - 1 \leq 4\epsilon^2.$$

The equations (1.17) and (1.18) together imply

$$\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \pi\right) \leq \epsilon.$$

Therefore,

$$\mathrm{t}_{\mathrm{mix}}(x;\epsilon) \leq \frac{\log N + 2\log(1/\epsilon)}{4},$$

which finishes our proof. □

7 The cutoff profile

Once one succeeds in proving cutoff, a more ambitious question is to understand how the system relaxes to equilibrium inside the cutoff windows. Recall that the worst-case total variation distance to equilibrium is defined by

$$\forall t \in \mathbb{R}_+,\ \mathcal{D}(t) := \max_{x \in \mathcal{X}} \mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \pi\right).$$

For $\mu \in \mathbb{R}, \sigma \in \mathbb{R}_{>0}$, let $\mathcal{N}\left(\mu, \sigma^2\right)$ denote the Gaussian distribution with mean μ and variance σ^2. Then one can prove the following.

Theorem 1.5 (Cutoff profile)**.** *For any $s \in \mathbb{R}$ fixed,*

$$\lim_{N\to\infty} \mathcal{D}\left(\frac{\log N}{4} + s\right) = \mathrm{d}_{TV}\left(\mathcal{N}(0,1), \mathcal{N}\left(e^{-2s}, 1\right)\right).$$

Let $t(s) := \frac{\log N}{4} + s$. The intuition behind this result is as follows. Let

$$\varphi(x) := \sum_{u \in [N]} x(u).$$

We expect that the following limits hold due to the Central Limit Theorem:

$$\frac{\varphi\left(X_\infty\right)-\mathbb{E}\left[\varphi(X_\infty)\right]}{\sqrt{\operatorname{Var}\left[\varphi(X_\infty)\right]}} \xrightarrow[N\to\infty]{d} \mathcal{N}\left(0,1\right), \tag{1.19}$$

$$\frac{\varphi\left(X^{\mathbf{1}}_{t(s)}\right)-\mathbb{E}\left[\varphi\left(X^{\mathbf{1}}_{t(s)}\right)\right]}{\sqrt{\operatorname{Var}\left[\varphi(X^{\mathbf{1}}_{t(s)})\right]}} \xrightarrow[N\to\infty]{d} \mathcal{N}\left(0,1\right). \tag{1.20}$$

Moreover, by direct computation,

$$\operatorname{Var}\left[\varphi\left(X^{\mathbf{1}}_{t(s)}\right)\right] = \operatorname{Var}\left[\varphi(X_\infty)\right](1+o(1)). \tag{1.21}$$

The equations (1.19), (1.20), and (1.21) together imply

$$\frac{\varphi(X^{\mathbf{1}}_{t(s)})-\mathbb{E}\left[\varphi(X_\infty)\right]}{\sqrt{\operatorname{Var}\left[\varphi(X_\infty)\right]}} \xrightarrow{N\to\infty} \mathcal{N}\left(e^{-2s},1\right). \tag{1.22}$$

We expect that the distribution of the system is dictated by the law of the observable φ since it is the distinguishing statistic. So, we expect that when N tends to infinity,

$$\begin{aligned} \mathrm{d_{TV}}\left(\mathbb{P}\left[X^{\mathbf{1}}_{t(s)}\in\cdot\right],\pi\right) &\approx \mathrm{d_{TV}}\left(\mathbb{P}\left[\varphi(X^{\mathbf{1}}_{t(s)})\in\cdot\right],\mathbb{P}\left[\varphi(X_\infty)\in\cdot\right]\right) \\ &\approx \mathrm{d_{TV}}\left(\mathcal{N}\left(0,1\right),\mathcal{N}\left(e^{-2s},1\right)\right), \end{aligned}$$

where the last approximation is due to (1.19) and (1.22). Usually, to prove the cutoff profile, one needs to understand in detail the law of the system at any time t close to the mixing times. The knowledge can come from either the understanding of the spectrum of the chain, as in [94, 78], or a reasonably explicit approximation of the law at time t as in [51, 37].

Though interesting, the question of obtaining the cutoff profile is not treated in this thesis. In this thesis, we only focus on showing cutoff.

8 Summary and perspectives

In many examples, the proof of cutoff is proceeded as above. For the lower bound, we say that the law of any observable mixes slower than the law of the system. On many occasions, there is a natural choice for the observable, and it often gives the correct estimate on the mixing times. For many examples, the suboptimal upper bound, such as in Theorem 1.1, is relatively easy to prove and matches the magnitude of the lower bound. However, the prefactor is often not sharp, cf. [102]. Proving the upper bound with the matching leading term is often much more challenging. Nevertheless, Lemma 1.4 proves to be quite useful for this difficult task.

In the rest of this thesis, we discuss three models: the Exclusion process with reservoirs (Chapter 2), the Glauber-Exclusion process (Chapter 3), and the Zero-Range process (Chapter 4). We will see that the proofs for cutoff of the first two models follow precisely the outline of the proof for the random walk on the hypercube, both relying on Lemma 1.4. The proof for the Zero-Range process follows a similar general strategy, though the materials we use there are different.

Chapter 2

Cutoff for the exclusion process with reservoirs

Goals

In this chapter, we prove cutoff with a diffusive window for the Symmetric Simple Exclusion Process (SSEP) with reservoirs on the segment, confirming a conjecture of Gantert, Nestoridi, and Schmid in [34]. Our proof follows exactly the outline of the one presented in Chapter 1. However, the law of the system is no longer a product law as in the case of the SRW on the hypercube. The novel ingredients needed to overcome this difficulty include the negative dependence of the system, the information percolation framework introduced by Lubetzky and Sly, and an anticoncentration inequality at the conditional level. The results presented here have been published in [99].

Contents

1 Introduction

1.1 Model

The simple exclusion process is an interacting particle system where the particles attempt to perform simple random walks on a graph, but they are not allowed to jump on top of each other (the exclusion rule). Since its introduction by Spitzer as a simplified model for a gas of interacting particles in [89] (see also [61]), it has been shown to exhibit many interesting phenomena, thus received a lot of attention from mathematicians and theoretical physicists, see, for example, [47, 60, 103]. Recently, a huge amount of work has been devoted to studying the convergence to equilibrium of the conservative (without reservoirs) model in finite volume, see [50, 51, 52, 58, 77, 102, 105]. Here, we study the non-conservative symmetric one-dimensional variant: the SSEP on the segment (the bulk) with reservoirs at the two endpoints, where the particles are allowed to enter or exit the bulk through the reservoirs. We refer the readers to the papers [5, 54] for an introduction and motivations on the model and to [34, 37, 87] for recent developments. More precisely, let $N \in \mathbb{Z}_+$ be the length of the segment, and let $p, q \in [0,1]$ be the densities of the reservoirs at the two endpoints. We consider the process $(X_t)_{t\geq 0}$ taking values in the state space $\mathcal{X} = \{0,1\}^N$, whose infinitesimal generator $\mathcal{L}$ acts on an observable $\varphi : \mathcal{X} \to \mathbb{R}$ by

$$\begin{aligned}\mathcal{L}\varphi(x) = \sum_{u=1}^{N-1} & N^2(\varphi(x^{u\leftrightarrow u+1}) - \varphi(x)) \\ & + N^2[p\varphi(x^{1,1}) + (1-p)\varphi(x^{1,0}) - \varphi(x)] \\ & + N^2[q\varphi(x^{N,1}) + (1-q)\varphi(x^{N,0}) - \varphi(x)],\end{aligned} \tag{2.1}$$

where $x^{u\leftrightarrow u+1}, x^{u,0}, x^{u,1}$ are the configurations obtained from x by swapping the u-th and $(u+1)$-th coordinates, resetting the u-th coordinate to 0, resetting the u-th coordinate to 1, respectively. Here, time is accelerated by a factor N^2 so that the process is observed on a diffusive time scale. Then $(X_t)_{t\geq 0}$ is a SSEP on the segment $[N] := \{1, \dots, N\}$ with one reservoir of density p placed at site 1 and the other reservoir of density q placed at site N. If $(p,q) \notin \{(0,0),(1,1)\}$, then $\mathcal{L}$ is irreducible, and there is a unique invariant probability distribution π. When $p = q$, the system is reversible w.r.t the product measure $\mathcal{B}_p^{\otimes N}$. When $p \neq q$, the process is no longer reversible, and the invariant measure π is given by a more complicated matrix-ansatz formula in [20] (see also [60]), which seems not easy to implement in the study of mixing times.

Previous works. The conservative SSEP (*without* reservoirs) has been thoroughly studied in [50, 51, 52, 77, 102]. In particular, Lacoin proves cutoff for the segment in [50] and for the circle in [52], and he even provides the cutoff profile for the latter model in [51]. On the contrary, only a few works have been written on the non-conservative model. We mention here some recent developments. In [34], Gantert, Nestoridi, and Schmid prove a pre-cutoff for the model: for any fixed $\epsilon \in (0,1)$,

$$\frac{1}{2\pi^2} \leq \liminf_{N\to\infty} \frac{\mathrm{t_{mix}}(\epsilon)}{\log N} \leq \limsup_{N\to\infty} \frac{\mathrm{t_{mix}}(\epsilon)}{\log N} \leq C,$$

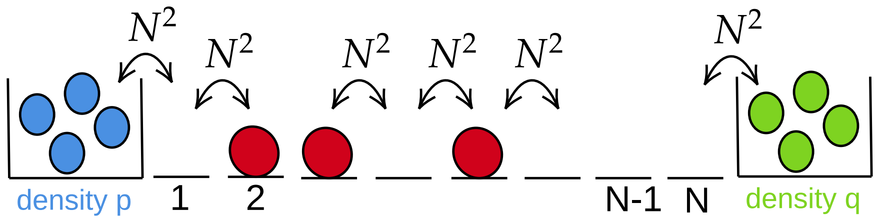


Figure 2.1: SSEP on the segment with reservoirs of densities p, q.

for some constant C independent of ϵ, and they conjecture that the system exhibits cutoff with the right estimate on the mixing times given by the lower bound. Their proof relies on an extension of the coupling used by Lacoin in [50] for the conservative model. In [37], Gonçalves, Jara, Marinho, and Menezes study the reversible case where the two reservoirs have the same density, $p = q$. Using Yau's famous relative entropy method in [104], they prove that

$$\mathrm{t}_{\mathrm{mix}}(\epsilon) = \frac{\log N + \mathcal{O}_\epsilon(1)}{2\pi^2},$$

where we recall that $\mathcal{O}(\cdot)$ means being bounded by the quantity inside the brackets up to a prefactor not depending on N, and the notation $\mathcal{O}_\epsilon(\cdot)$ emphasizes that the prefactor may depend on ϵ. Recently, Salez studies in [87] the model on general networks where reservoirs of the same density can be placed at arbitrary sites. By exploiting the negative dependence property of the system, he proves that, under some mild conditions on the network,

$$\mathrm{t}_{\mathrm{mix}}(\epsilon) = \frac{\log N + \mathcal{O}_\epsilon(1)}{2\lambda},$$

where N is the size of the network, and λ is the absolute value of the smallest eigenvalue of the Laplacian matrix of the network (see §1.3.3 below for more details). Interestingly, the Laplacian of a network does not depend on the densities but only the rates of contact with the reservoirs. In a more recent work [86], Salez studies the model on general networks where the reservoirs can have different densities, proving that λ coincides with the spectral gap of the system. In particular, in the case study in [37], $\lambda \approx \pi^2$, so all the results mentioned above are consistent. We stress that the works above provide a more comprehensive study than the results we just mentioned. The work [34] is more devoted to studying the case where the random motion of the particles is asymmetric, the work [37] provides the convergence profile for the system from any smooth initial condition, and the work [87] is more concerned with the characterization of cutoff. However, as far as we know, cutoff has been established in the works [37, 87] only for the case where every reservoir has the same density p, which subsequently implies that the system is reversible w.r.t the product measure $\mathcal{B}_p^{\otimes N}$.

Our contribution. In this chapter, we prove cutoff for the SSEP on the segment with reservoirs of arbitrary densities $p, q \in (0, 1)$, thus confirming the conjecture of Gantert, Nestoridi, and Schmid (see Conjecture 1.7 in [34]), and also making a step towards the study of irreversible models. Our proof exploits the information percolation framework introduced by Lubetzky and Sly in [68], the negative dependence of the system, and an anticoncentration inequality at the conditional level. In particular, our proof does not require the explicit formula for the invariant measure π given in [20]. We believe that this approach is applicable to other models.

1.2 Results

Our main result is that the model exhibits cutoff at time $\frac{\log N}{2\pi^2}$ with a window of order $\mathcal{O}(1)$, as conjectured by Gantert, Nestoridi, and Schmid.

Theorem 2.1 (Main theorem)**.** *For any $p, q, \epsilon \in (0,1)$ fixed, we have*

$$\mathrm{t_{mix}}(\epsilon) = \frac{\log N}{2\pi^2} + \mathcal{O}_{p,q,\epsilon}(1).$$

In fact, we will prove a stronger result, which makes precise the dependence of the lower order term on p, q, ϵ and is subsequently still valid when we allow p, q to vary with N. Without loss of generality (by the symmetry between p and q and the duality between particles and holes), we suppose that

$$q \leq \min\{p, 1-p\}. \tag{2.2}$$

We define the weight of the configuration x to be

$$S(x) = \sum_{u \in [N]} x(u).$$

We denote by $\mathbb{E}^0$ the expectation w.r.t the absorbing model where the two reservoir densities p, q are zero. Let $\mathbf{1}$ denote the configuration where every site is occupied, and let t^* be the time that the expected weight of the process starting from $\mathbf{1}$ falls under a specific threshold:

$$t^* := \inf\left\{t \geq 0 : \mathbb{E}^0_{\mathbf{1}}\left[S(X_t)\right] \leq \sqrt{Np} \vee 1\right\}. \tag{2.3}$$

In fact, we will see that t^* is the time at which the expected weight of the process (with densities of the reservoirs p, q) starting from the assumingly worst initial condition $\mathbf{1}$ becomes "close enough" to the expected weight at equilibrium. We prove that t^* is a good estimate on the mixing times for the chain associated with the generator $\mathcal{L}$ given in (2.1).

Theorem 2.2 (Non asymptotic estimates)**.** *Under assumption* (2.2)*, there exists a universal constant C such that for any $\epsilon \in]0,1[$,*

$$t^* - C\left(1 + \log\left(\frac{1}{1-\epsilon}\right)\right) \leq \mathrm{t_{mix}}(\epsilon) \leq t^* + C\left(1 + \log\left(\frac{1}{\epsilon}\right) + \log\left(\frac{1}{1-p}\right)\right). \tag{2.4}$$

Estimating t^* is a classical problem corresponding to the study of the discrete heat equation on the segment. Using the ingredients presented in Appendix A in [37], we will show that,

$$t^* = \frac{1}{\pi^2}\log\left(\frac{N}{\sqrt{Np} \vee 1}\right) \pm \mathcal{O}(1), \tag{2.5}$$

which means that the leading term is of order $\log N$, and the lower order term is bounded by some universal constant. This immediately implies the following corollary, of which Theorem 2.1 is a direct consequence.

Corollary 2.3. *Under assumption (2.2), the system exhibits cutoff at t^* when*

$$\frac{1}{1-p} = N^{o(1)}. \tag{2.6}$$

Moreover, if p is bounded away from 1, then the cutoff window is of order $\mathcal{O}(1)$, and if p is also bounded away from 0, then $t^* = \dfrac{\log N}{2\pi^2} + \mathcal{O}(1)$.

Condition (2.6) is valid for almost all pairs (p, q) satisfying (2.2), except for the case where p and q vary with N and $q \to 0,\ p \to 1$ rapidly when $N \to \infty$.

Intuition. Under assumption (2.2), the weight of configuration **1** (which is N) is the farthest from the expected weight at equilibrium (which is $N \times \dfrac{p+q}{2}$). So, we expect **1** to be the worst initial configuration. At equilibrium, we expect the sites to be "almost independent", and hence

$$\mathrm{Var}_\pi[S] \asymp \mathbb{E}_\pi[S] = \frac{N(p+q)}{2} \asymp Np,$$

where $a \asymp b$ means $a = \mathcal{O}(b)$ and $b = \mathcal{O}(a)$. It is not easy to explain in detail the intuition for the equality above. However, we mention that the above equality is clearly true when $p = q$ (since in this case, $\pi = \mathcal{B}_p^{\otimes N}$). In fact, in our proof, we only need $\mathrm{Var}_\pi[S] = \mathcal{O}(\mathbb{E}_\pi[S])$, which is immediate from the negative dependent property (see Section 2.1). In many spin systems that exhibit cutoff, cutoff is shown to occur precisely at the time when the expected weight of the system from the worst initial configuration falls within the deviation of the weight at equilibrium. Examples, among many others, include the simple random walk on the hypercube, the Glauber dynamics of the Ising model in high temperature [67, 68, 69], the reversible SSEP with reservoirs [87], the noisy voter model [16]. We believe that the time t^* proposed above is exactly that time. However, we are not able to prove the other way of the inequality mentioned above, namely $\mathbb{E}_\pi[S] = \mathcal{O}(\mathrm{Var}_\pi[S])$, to verify this rigorously.

1.3 Open problems

In this subsection, we present some open questions related to the SSEP with reservoirs.

1.3.1 Cutoff for all densities

With the intuition presented in the last subsection, we conjecture that cutoff should not be restrained by Condition (2.6).

Conjecture 1 (Cutoff for arbitrary densities)**.** *Under condition (2.2), the system still exhibits cutoff at time t^* when $N \to \infty$, even when p and q vary with N and Condition (2.6) does not hold. In particular, the system exhibits cutoff at time $\dfrac{\log N}{2\pi^2}$ if $p = 1,\ q = 0$.*

In fact, our proof relies on Lemma 1.4 about perturbation of the product measure, which only works well under Condition (2.6). We believe that this is the limit of our method but not of the problem.

1.3.2 Cutoff profile

It is also tempting to understand how the system relaxes to equilibrium inside the cutoff window. Let

$$t(b) = \frac{\log N}{2\pi^2} + \frac{b}{\pi^2}. \tag{2.7}$$

Let $X_\infty \sim \pi$. We expect that there exists some constant $f_{p,q}$ such that the following limit holds

$$\lim_{N\to\infty} \frac{\mathbb{E}_{\mathbf{1}}\left[S(X_{t(b)})\right] - \mathbb{E}\left[S(X_\infty)\right]}{\sqrt{\mathrm{Var}\left[S(X_\infty)\right]}} \xrightarrow[N\to\infty]{} f_{p,q}e^{-b}, \tag{2.8}$$

for any $b \in \mathbb{R}$, because a similar result holds in [37] when $p = q$. View the result in [37], we conjecture the following

Conjecture 2 (Cutoff profile)**.** *Let p, q be fixed and satisfy (2.2), and let $t(b)$ and $f_{p,q}$ be defined as in (2.7), (2.8). Then*

$$\lim_{N\to\infty} \mathcal{D}\left(t(b)\right) = \mathrm{d}_{TV}\left(\mathcal{N}\left(0,1\right), \mathcal{N}\left(f_{p,q}e^{-b}, 1\right)\right).$$

The intuition behind this conjecture is as presented in the previous chapter. We expect that the following limits hold due to the Central Limit Theorem:

$$\frac{S(X_\infty) - \mathbb{E}\left[S(X_\infty)\right]}{\sqrt{\mathrm{Var}\left[S(X_\infty)\right]}} \xrightarrow[N\to\infty]{d} \mathcal{N}\left(0,1\right), \tag{2.9}$$

$$\frac{S\left(X^{\mathbf{1}}_{t(b)}\right) - \mathbb{E}\left[S\left(X^{\mathbf{1}}_{t(b)}\right)\right]}{\sqrt{\mathrm{Var}\left[S(X_\infty)\right]}} \xrightarrow[N\to\infty]{d} \mathcal{N}\left(0,1\right). \tag{2.10}$$

The equations (2.8),(2.9), and (2.10) together imply

$$\frac{S(X^{\mathbf{1}}_{t(b)}) - \mathbb{E}\left[S(X_\infty)\right]}{\sqrt{\mathrm{Var}\left[S(X_\infty)\right]}} \xrightarrow[N\to\infty]{d} \mathcal{N}\left(f_{p,q}e^{-b}, 1\right). \tag{2.11}$$

So, we expect that, when N tends to infinity,

$$\begin{aligned}\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_{\mathbf{1}}\left[X_{t(b)} \in \cdot\right], \pi\right) &\approx \mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_{\mathbf{1}}\left[S(X_{t(b)}) \in \cdot\right], \mathbb{P}\left[S(X_\infty) \in \cdot\right]\right) \\ &\approx \mathrm{d}_{\mathrm{TV}}\left(\mathcal{N}\left(0,1\right), \mathcal{N}\left(f_{p,q}e^{-b}, 1\right)\right).\end{aligned}$$

The work [37] makes rigorous this intuition using Yau's relative entropy method for the case $p = q$. We also mention the analogous result for the conservative case obtained by Lacoin in [51]. However, proving the result in our settings when $p \neq q$ seems much more challenging. First, we do not even have the exact computation for $\mathrm{Var}\left[S(X_\infty)\right]$, but only an upper bound. Second, in [37] and [51], the authors manage to approximate the measure at any time near the mixing times with some fairly explicit measure, allowing them to have very fine computations to prove the cutoff profile. However, in our case, even the invariant distribution given in [20] seems complicated to implement to study mixing features.

1.3.3 In general geometry

We briefly recall the settings in [86] and [87] (with some slight adaptations).

Settings in general geometry. A finite network is a quadruple $G = (n, c, \kappa, \rho)$, consisting of

- a number $n \in \mathbb{Z}_+$, which is the size of the network. We identify the sites of the network with the set $[n]$.
- a symmetric array $c : [n] \times [n] \to \mathbb{R}_+$, whose entries are called conductances.
- a function $\kappa : [n] \to \mathbb{R}_+$, whose entries are called external rates.
- a function $\rho : \mathrm{Supp}(\kappa) \to [0, 1]$, whose entries are called densities of the external reservoirs, where $\mathrm{Supp}(\kappa) = \{i \in [n] : \kappa(i) > 0\}$.

Naturally, the support of $c(\cdot, \cdot)$ is regarded as the network's edges, and the support of $\kappa(\cdot)$ is regarded as the network boundary. The exclusion process associated with the network G is the continuous time Markov process taking value in $\mathcal{X} = \{0, 1\}^n$, with generator $\mathcal{L}$ acts on an observable $\varphi : \mathcal{X} \to \mathbb{R}$ by

$$\begin{aligned}(\mathcal{L}\varphi)(x) :=& \frac{1}{2} \sum_{i,j \in [n]} c(i, j) \left[(\varphi(x^{i \leftrightarrow j}) - \varphi(x)) \right] + \\ &+ \sum_{i \in [n]} \kappa(i) \left[\rho(i)\varphi(x^{i,1}) + (1 - \rho(i))\varphi(x^{i,0}) - \varphi(x) \right].\end{aligned}$$

In words, each pair of sites $\{i, j\}$ exchange contents at rate $c(i, j)$, and the content at the site i is resampled at rate $\kappa(i)$ by an independent Bernoulli $\mathcal{B}_{\rho(i)}$. We assume that the network is connected and that $\rho(\cdot)$ is not identically equal to 0 or 1 to ensure the irreducibility of the generator $\mathcal{L}$. In the article [87], ρ is a constant function, i.e. does not vary on $[n]$. In the settings in [86] (version 1 on arXiv), Section 4, this is equivalent to taking $\kappa(i) = r_i^+ + r_i^-$, $i \in [n]$ and $\rho(i) = \dfrac{r_i^+}{\kappa(i)}$, $i \in \mathrm{Supp}(\kappa)$. The Laplace matrix of the network G is given by $\Delta = (\Delta_{ij})_{1 \le i,j \le n}$, where

$$\Delta_{ij} := \begin{cases} c_{ij} & \text{if } i \neq j, \\ -\kappa(i) - \sum_{k \in [n] \setminus \{i\}} c_{ik} & \text{if } i = j. \end{cases}$$

We remark that the matrix Δ does not depend on the function $\rho(\cdot)$. Let λ be the smallest eigenvalue of the matrix $-\Delta$. In [86], Salez proves the following analogy of Aldous' spectral gap conjecture.

Theorem 2.4 (Theorem 2 in [86], version 1 on arXiv). *The spectral gap of the generator $\mathcal{L}$ coincides with λ.*

In [87], he proves the following theorem.

Theorem 2.5 (Characterization of cutoff, Corollary 3 in [87]). *Consider the exclusion processes with reservoirs associated with a sequence of networks $(G_n)_{n\geq 1}$, where the associated functions $\rho^{(n)}$ are constant functions and bounded away from 0 and 1, i.e. there exists $\delta \in (0,1)$ such that $\delta < \rho^{(n)}(\cdot) < 1-\delta$. Then the cutoff phenomenon occurs if and only if the so-called "product condition",*

$$\lambda_n \times \mathrm{t}_{\mathrm{mix}}^{(n)}(\epsilon) \xrightarrow[n\to\infty]{} +\infty,$$

is satisfied.

Taking into account Corollary 2.3, Theorem 2.4, and Theorem 2.5, we boldly conjecture that the condition $\rho(\cdot)$ is constant on its support is not necessary for cutoff to occur.

Conjecture 3. *Theorem 2.5 is still correct upon removing the hypothesis $\rho^{(n)}$ are constant functions.*

In particular, the exclusion process with reservoirs in higher dimensions is included in the regime above. We refer the reader to Subsection 1.3 in [87] for more details.

1.4 Structure of the proofs

In Subsection 2.1, we recall the negative dependence property of the SSEP with reservoirs and the exponential bound introduced by Miller and Peres. In Subsection 2.2, we collect some elementary but useful estimates. In Subsection 2.3, we introduce our method. Section 3 and Section 4 are devoted to the proofs of the upper and lower bound in Theorem 2.2, respectively. Finally, in Section 5, we compute t^* explicitly.

2 Preliminaries

2.1 Negative dependence property

We recall the notion of negative dependence, which is essential throughout our proof.

Definition 2.6 (Negative dependence). A random vector $Z = (Z_1, \dots, Z_n)$ taking values in $\{0,1\}^n$, for some $n \in \mathbb{Z}_+$, is said to be negatively dependent (ND) if it satisfies

$$\forall A \subset [n], \quad \mathbb{E}\left[\prod_{i\in A} Z_i\right] \leq \prod_{i\in A} \mathbb{E}\left[Z_i\right]. \tag{2.12}$$

An important property of the SSEP with reservoirs is that it preserves the negative dependence property.

Proposition 2.7 (The ND property is preserved by SSEP with reservoirs, Lemma 4 in [87]). *Let the generator $\mathfrak{L}$ on $\mathcal{X}$ be defined by*

$$\mathfrak{L} = \sum_{u,v=1}^{N} a_{u,v}\mathcal{L}_{u,v} + \sum_{u=1}^{N} a_u^0 \mathcal{L}_u^0 + \sum_{u=1}^{N} a_u^1 \mathcal{L}_u^1,$$

where

$$\mathcal{L}_{u,v} f(x) = f(x^{u\leftrightarrow v}) - f(x),$$
$$\mathcal{L}_u^0 f(x) = f(x^{u,0}) - f(x),$$
$$\mathcal{L}_u^1 f(x) = f(x^{u,1}) - f(x),$$

and $a_{u,v}, a_u^0, a_u^1 \in \mathbb{R}_+$, $1 \le u, v \le N$. Then $\mathfrak{L}$ preserves the negative dependence property, i.e. if $Z(0) \sim \mu$ is a negatively dependent vector for some measure μ on $\mathcal{X}$, then $Z(t) \sim \mu e^{\mathfrak{L}t}$ is also ND.

In fact, the SSEP with reservoirs preserves a much richer property called the Strongly Rayleigh property, of which negative dependence is a consequence. We refer the readers to the beautiful paper [6] for more details. This proposition, however, is available exclusively for the case where the exclusion process is symmetric and is not applicable in the case of ASEP or TASEP.

A particular case of inequality (2.12) is when $|A| = 2$, which implies that the coordinates of an ND vector Z are negatively correlated. As a consequence, the weight of the vector Z is concentrated around its mean.

Lemma 2.8 (Concentration of the weight of an ND vector). *Let $Z = (Z_1, \dots, Z_n)$ be an ND vector. Let $S = \sum_{i=1}^n Z_i$. Then*

$$\mathrm{Var}[S] \le \mathbb{E}[S]. \tag{2.13}$$

2.2 Some elementary estimates

We first give a lemma about the symmetric simple random walk on the segment.

Lemma 2.9 (Simple random walk on the segment). *Let $(U(t))_{t\ge 0}$ be a continuous-time symmetric simple random walk on $\{0, \dots, N+1\}$, which jumps to the left (or to the right) at rate N^2. For any $u \in \{0, \dots, N+1\}$, let T_u be the first time that the walk reaches u, i.e. $T_u = \inf\{t \ge 0 : U(t) = u\}$. Then there exists a constant $c > 0$ (independent of N) such that for any $u \in \{0, \dots, N+1\}$,*

$$\mathbb{P}_u[T_0 \ge 2] < e^{-c}. \tag{2.14}$$

This lemma can be proved by a classical hitting time estimate for non-negative supermartingale (see, e.g. Proposition 2.1 in [59], for the discrete version).

We identify a subset $E \subset [n]$ with the vector $(\mathbb{1}_{\{i\in E\}})_{1\le i\le n}$. We remark here that the negative dependence property comes in very handy, as it allows us to bound the exponential moment in the Lemma 1.4 by some quantity that depends only on the marginal of the random vector E, as stated in the following corollary.

Corollary 2.10 (Negative dependent perturbation of a product measure)**.** *Under the notations in Lemma 1.4, if the random set E is ND, then*

$$4\mathrm{d}_{TV}(\mu,\nu)^2 \leq e^{(\theta-1)\sum_{i=1}^{n}\mathbb{P}[i\in E]^2} - 1.$$

For the proofs of Lemma 1.4 and Corollary 2.10, see Lemma 1 in [87].

2.3 Framework and some definitions

The SSEP with reservoirs $(X_t)_{t\geq 0}$ evolves according to the following transitions:

1. $x \mapsto x^{u\leftrightarrow u+1}$ (exchange between site u and site $u+1$), which occurs at rate N^2,
2. Resampling the value at site 1 by an independent Bernoulli $\mathcal{B}_p$, which occurs at rate N^2,
3. Resampling the value at site N by an independent Bernoulli $\mathcal{B}_q$, which occurs at rate N^2.

We introduce another Markov process closely related to the SSEP with reservoirs.

The colored interchange process. Let $\mathcal{W} := \mathcal{S}_N \times \{R, B, G\}^N$, where $\mathcal{S}_N$ is the symmetric group on $[N]$, and R, B, G stand for red, blue, and green. Each element $(\sigma, b) \in \mathcal{W}$ describes a way to put N colored labelled individuals on the segment as follows. The individuals are labelled $1, 2, \ldots, N$. For any $u \in [N]$, the individual labelled u is located at site $\sigma(u)$ and colored $b(u)$. *The colored interchange process* $W = (W_t)_{t\geq 0}$ is defined as the Markov process taking values in the state space $\mathcal{W}$ which evolves according to the following transitions:

1. $(\sigma, b) \mapsto ((u, u+1)\circ\sigma, b)$ (the individuals at sites u and $u+1$ exchange their positions), which occurs at rate N^2, for $1 \leq u \leq N-1$,
2. $(\sigma, b) \mapsto (\sigma, b^{\sigma^{-1}(1),B})$ (recoloring the individual at site 1 blue), which occurs at rate N^2,
3. $(\sigma, b) \mapsto (\sigma, b^{\sigma^{-1}(N),G})$ (recoloring the individual at site N green), which occurs at rate N^2.

We denote by $\mathbb{P}_w[\cdot]$ and $\mathbb{E}_w[\cdot]$ the probability and expectation taken with respect to the law of process W starting from w. There is a natural coupling between $(X_t)_{t\geq 0}$ and $(W_t)_{t\geq 0}$ as follows.

Natural coupling between $(X_t)_{t\geq 0}$ and $(W_t)_{t\geq 0}$. A coupling of the two processes is given by making the transitions $1, 2, 3$ listed above of the two processes $(X_t)_{t\geq 0}$ and W occur at the same time.

Roughly speaking, the labels and colors are added to keep better track of the exchange of information inside the bulk and to memorize which reservoirs the resamplings come from. For an introduction to the interchange process and its relation with the exclusion process, see [59], chapter 23.

The "pushforward" function. Let the function $f_* : \mathcal{X} \times \mathcal{W} \times \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ be defined by, for any $w = (\sigma, b) \in \mathcal{W}$ and $x, v^B, v^G \in \mathcal{X}$,

$$f_*(x, w, v^B, v^G)(u) = \begin{cases} x(\sigma^{-1}(u)) & \text{if } b\left(\sigma^{-1}(u)\right) = R, \\ v^B(u) & \text{if } b\left(\sigma^{-1}(u)\right) = B, \\ v^G(u) & \text{if } b\left(\sigma^{-1}(u)\right) = G. \end{cases} \tag{2.15}$$

The interest of introducing the process W is the following lemma, whose proof is straightforward from the definition of the natural coupling.

Lemma 2.11. *Let $w_0 \in \mathcal{W}$ be the configuration where for any $u \in [N]$, the individual labelled u is located at site u and is colored red:*

$$w_0 = (\mathrm{Id}, (R, \dots, R)),$$

Let W, ξ^B, ξ^G be independent and as follows.

- *W is a colored interchange process starting at w_0.*
- *$\xi^B \sim \mathcal{B}_p^{\otimes N}$.*
- *$\xi^G \sim \mathcal{B}_q^{\otimes N}$.*

Let $(X_t)_{t\geq 0}$ be the SSEP with reservoirs started from some configuration $x \in \mathcal{X}$. Then for any $t \geq 0$,

$$X_t \overset{d}{=} f_*(x, W_t, \xi^B, \xi^G), \tag{2.16}$$

where $\overset{d}{=}$ means equal in distribution.

Red, blue, and green regions. For any $w = (\sigma, b) \in \mathcal{W}$, we denote by $\mathsf{Red}(w)$ **the red region**, i.e. the set of the sites containing the red individuals:

$$\mathsf{Red}(w) := \{u \in [N] | b(\sigma^{-1}(u)) = R\}.$$

The blue region $\mathsf{Blue}(w)$ and the green region $\mathsf{Green}(w)$ are defined similarly. As the two reservoirs recolor the particles blue or green, the red region evolves exactly as an SSEP with two reservoirs of density 0. Then t^* is the time at which the red region becomes small enough:

$$t^* = \inf\left\{t \geq 0 : \mathbb{E}_{w_0}\left[|\mathsf{Red}(W_t)|\right] \leq \sqrt{Np} \vee 1\right\}, \tag{2.17}$$

with w_0 as in Lemma 2.11. We present a graphical construction of W that allows us to reveal the green region before the red and blue regions.

Graphical construction of the colored interchange process. We can construct W in the following way.

$$W = \Psi(\sigma, b, \Xi^1, \Xi^N, (\Xi_u^G)_{1\leq u\leq N-1}, (\Xi_u^{BR})_{1\leq u\leq N-1}), \tag{2.18}$$

where $(\sigma, b) \in \mathcal{W}$ and $\Xi^1, \Xi^N, (\Xi_u^G)_{1\le u\le N-1}, (\Xi_u^{BR})_{1\le u\le N-1}$ are independent and described as follows.

- Ξ^1 and Ξ^N are homogeneous Poisson processes of intensity N^2 which indicate the times at which we recolor the individuals at site 1 and site N, respectively.
- Ξ_u^G and Ξ_u^{BR} are homogeneous Poisson processes of intensity N^2, $1 \le u \le N-1$. Each time Ξ_u^G jumps, the two individuals at sites u and $u+1$ exchange their positions if at least one of them is green, and each time Ξ_u^{BR} jumps, the two individuals at sites u and $u+1$ exchange their positions if none of them is green.

This construction gives the colored interchange process W with initial condition (σ, b).

Trajectories of a single individual. It is well known, see e.g. Chapter 23 of [59], that if we observe the trajectory of a single labelled individual, we see a continuous-time simple random walk on the segment where the conductance of any edge is N^2. As we add the colors here, we also see that the labelled individual is always recolored at rate N^2 when it is at site 1 or N.

Trajectories of the green regions. The above construction allows revealing $(\mathsf{Green}(W_t))_{t\ge0}$ before $(\mathsf{Red}(W_t))_{t\ge0}$ and $(\mathsf{Blue}(W_t))_{t\ge0}$. In fact, $(\mathsf{Green}(W_t))_{t\ge0}$ is measurable with respect to the σ-algebra $\mathcal{H}_G$ generated by Ξ^1, Ξ^N, and Ξ^G.

Number of crossings. Let $(L_t)_{t\ge0}$ be the process that counts the number of times that a blue or red individual is recolored green. Note that if an individual is recolored blue at site 1, then it needs to cross the bulk to be recolored green. Accordingly, we call $(L_t)_{t\ge0}$ the number of crossings. A simple but important observation is that $(L_t)_{t\ge0}$ is also $\mathcal{H}_G$-measurable.

3 The upper bound

Recall that

$$\mathcal{D}(t) = \max_{x\in\mathcal{X}} \mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \pi\right).$$

Our strategy is to compare directly two processes from two arbitrary configurations x and $\tilde{x}$ and use the fact that

$$\mathcal{D}(t) \le \max_{x,\tilde{x}\in\mathcal{X}} \mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \mathbb{P}_{\tilde{x}}\left[X_t \in \cdot\right]\right), \tag{2.19}$$

which is due to the convexity of the total variation distance. Our goal now is to compare the distributions of those two processes at our predicted time t^*. The idea is to match perfectly the green region of the two processes by the graphical construction above and to view the distributions on the remaining sites as a product measure perturbed by the red region to compare them using Lemma 1.4.

We will need the following lemmas and propositions.

Lemma 2.12 (Exponential decay of the red region)**.** *For any $w \in \mathcal{W}$, for any $t \geq 0$,*

$$\mathbb{E}_w\left[|\mathsf{Red}(W_{2t})|\right] \leq e^{-c\lfloor t \rfloor} |\mathsf{Red}(w)|,$$

where c is the constant in Lemma 2.9.

Lemma 2.13 (Fast increase of the number of crossings)**.** *There exists a constant C such that for any $\epsilon \in (0,1)$, for $t_2 = C(1+\log(1/\epsilon))$, for any initial configuration $w \in \mathcal{W}$,*

$$\mathbb{P}_w\left[L_{t_2} < 2N\right] \leq \epsilon/4.$$

Proposition 2.14 (Negative dependence property of conditional law)**.** *For any initial configuration $w \in \mathcal{W}$, almost surely, conditionally on $\mathcal{H}_G$, $\mathsf{Red}(W_t)$ is negatively dependent at any time $t \geq 0$.*

Lemma 2.15 (Conditional anticoncentration inequality)**.** *Let t be a positive number, and let $w = (\sigma, b) \in \mathcal{W}$ be an initial configuration. Let $W_s = (\sigma_s, b_s)$. For any individual u that is colored blue or red in w, for any site v, on the event $\{L_t \geq 2N\}$,*

$$\mathbb{P}_w\left[\sigma_t(u) = v, \sigma_s(u) \notin \mathsf{Green}(W_s), \forall 0 \leq s \leq t | \mathcal{H}_G\right] \leq \frac{1}{N}.$$

The following is a direct consequence, obtained by summing the inequality in Lemma 2.15 over all red individuals in w.

Corollary 2.16 (Conditional marginal of $\mathsf{Red}(W_t)$)**.** *Let t be a positive number and $w \in \mathcal{W}$ be an initial configuration. For any site v, on the event $\{L_t \geq 2N\}$,*

$$\mathbb{P}_w\left[v \in \mathsf{Red}(W_t) | \mathcal{H}_G\right] \leq \frac{|\mathsf{Red}(w)|}{N}.$$

Now we are ready to prove the upper bound in Theorem 2.2.

Proof of the upper bound. Let $x, \tilde{x} \in \mathcal{X}$ arbitrary, and let c be the constant in Lemma 2.9. Let w_0, W, ξ^B, ξ^G be defined as in Lemma 2.11. Let $(\zeta_t)_{t\geq 0}$ and $(\tilde{\zeta}_t)_{t\geq 0}$ be defined by

$$\begin{aligned} \zeta_t &= f_*(x, W_t, \xi^B, \xi^G), \\ \tilde{\zeta}_t &= f_*(\tilde{x}, W_t, \xi^B, \xi^G). \end{aligned}$$

By Lemma 2.11,

$$\begin{aligned} \mathbb{P}_x\left[X_t \in \cdot\right] &= \mathbb{P}_{w_0}\left[\zeta_t \in \cdot\right], \\ \mathbb{P}_{\tilde{x}}\left[X_t \in \cdot\right] &= \mathbb{P}_{w_0}\left[\tilde{\zeta}_t \in \cdot\right]. \end{aligned}$$

So now we can compare the distributions of ζ_t and $\tilde{\zeta}_t$ instead. We divide the proof into two cases: $Np \leq 1$ and $Np > 1$.

Case 1: $Np \le 1$. We see that if $|\mathsf{Red}(W_t)| = 0$, then $\mathbb{P}_{w_0}[\zeta_t \in \cdot | W_t] = \mathbb{P}_{w_0}\left[\tilde{\zeta}_t \in \cdot | W_t\right]$, by definition of f_*. Hence for $t = t^* + 2m$, for some $m \in \mathbb{Z}_+$,

$$\begin{aligned}
&\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_{w_0}[\zeta_t \in \cdot], \mathbb{P}_{w_0}\left[\tilde{\zeta}_t \in \cdot\right]\right) \\
&\le \mathbb{E}\left[\left\|\mathbb{P}_{w_0}[\zeta_t \in \cdot | W_t] - \mathbb{P}_{w_0}\left[\tilde{\zeta}_t \in \cdot | W_t\right]\right\|_{TV}\right] \\
&\le \mathbb{P}_{w_0}[|\mathsf{Red}(W_t)| > 0] \\
&\le \mathbb{E}_{w_0}[|\mathsf{Red}(W_t)|] \le e^{-cm}\mathbb{E}_{w_0}[|\mathsf{Red}(W_{t^*})|] = e^{-cm},
\end{aligned}$$

where the first inequality is due to Jensen's inequality, the second inequality is by upper bounding $\|\cdot\|_{TV}$ by 1 on the event $\{\mathsf{Red}(W_t) > 0\}$, the third inequality is because $|\mathsf{Red}(W_t)| \in \mathbb{Z}_+$, the last inequality is by Lemma 2.12, and the equality is by (2.17). We can take $m = \left\lceil \frac{-\log \epsilon}{c} \right\rceil$ to make e^{-cm} smaller than ϵ, which finishes the proof.

Case 2: $Np > 1$. Let t_1, t_2, α be some positive numbers we will choose later. Let $t = t_1 + t_2$. We see that,

$$\begin{aligned}
&\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_{w_0}[\zeta_t \in \cdot], \mathbb{P}_{w_0}\left[\tilde{\zeta}_t \in \cdot\right]\right) \\
&\le \mathbb{E}_{w_0}\left[\left\|\mathbb{P}_{w_0}[\zeta_t \in \cdot | W_{t_1}] - \mathbb{P}_{w_0}\left[\tilde{\zeta}_t \in \cdot | W_{t_1}\right]\right\|_{TV}\right] \\
&= \mathbb{E}_{w_0}\left[\left\|\mathbb{P}_{W_{t_1}}[\zeta_{t_2} \in \cdot] - \mathbb{P}_{W_{t_1}}\left[\tilde{\zeta}_{t_2} \in \cdot\right]\right\|_{TV}\right] \\
&\le \mathbb{P}_{w_0}[|\mathsf{Red}(W_{t_1})| > \alpha] + \max_{w:|\mathsf{Red}(w)| \le \alpha}\left\|\mathbb{P}_w[\zeta_{t_2} \in \cdot] - \mathbb{P}_w\left[\tilde{\zeta}_{t_2} \in \cdot\right]\right\|_{TV},
\end{aligned} \tag{2.20}$$

where the first inequality is by Jensen's inequality, the equality is due to the Markov property of the process W at time t_1, and the second inequality is by upper bounding the total variation distance by 1 on the event $\{|\mathsf{Red}(W_{t_1})| > \alpha\}$. For any $w \in \mathcal{W}$, we use the graphical construction in Section 2 to construct the process W starting from w. Then

$$\begin{aligned}
&\left\|\mathbb{P}_w[\zeta_{t_2} \in \cdot] - \mathbb{P}_w\left[\tilde{\zeta}_{t_2} \in \cdot\right]\right\|_{TV} \\
&\le \mathbb{E}\left[\left\|\mathbb{P}_w[\zeta_{t_2} \in \cdot | \mathcal{H}_G] - \mathbb{P}_w\left[\tilde{\zeta}_{t_2} \in \cdot | \mathcal{H}_G\right]\right\|_{TV}\right] \\
&\le \mathbb{P}_w[L_{t_2} < 2N] + \mathbb{E}\left[\left\|\mathbb{P}_w[\zeta_{t_2} \in \cdot | \mathcal{H}_G] - \mathbb{P}_w\left[\tilde{\zeta}_{t_2} \in \cdot | \mathcal{H}_G\right]\right\|_{TV} \mathbb{1}_{\{L_{t_2} \ge 2N\}}\right],
\end{aligned} \tag{2.21}$$

where the first inequality is by Jensen's inequality, and the second inequality is simply by upper bounding the total variation distance by 1 on the event $\{L_{t_2} < 2N\}$. Combining (2.20) and (2.21), we deduce that $\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_{w_0}[\zeta_t \in \cdot], \mathbb{P}_{w_0}\left[\tilde{\zeta}_t \in \cdot\right]\right)$ does not exceed

$$\begin{aligned}
&\mathbb{P}_{w_0}[|\mathsf{Red}(W_{t_1})| > \alpha] + \max_{w \in \mathcal{W}} \mathbb{P}_w[L_{t_2} < 2N] \\
&+ \max_{w:|\mathsf{Red}(w)| \le \alpha} \mathbb{E}\left[\left\|\mathbb{P}_w[\zeta_{t_2} \in \cdot | \mathcal{H}_G] - \mathbb{P}_w\left[\tilde{\zeta}_{t_2} \in \cdot | \mathcal{H}_G\right]\right\|_{TV} \mathbb{1}_{\{L_{t_2} \ge 2N\}}\right].
\end{aligned} \tag{2.22}$$

We separately estimate the three terms in the sum (2.22).

The first term. We choose $t_1 = t^* + 2m$ for some $m \in \mathbb{Z}_+$ that we will choose later. Then by

Lemma 2.12 and equality (2.17), $\mathbb{E}_{w_0}[|\mathsf{Red}(W_{t_1})|] \leq \mathbb{E}_{w_0}[|\mathsf{Red}(W_{t^*})|]\, e^{-cm} = \sqrt{Np}e^{-cm}$. Hence, by Markov's inequality,

$$\mathbb{P}_{w_0}[|\mathsf{Red}(W_{t_1})| \geq \alpha] \leq \frac{\sqrt{Np}e^{-cm}}{\alpha}.$$

The second term. This term is smaller than $\epsilon/4$ for t_2 as in Lemma 2.13.

The third term. Let $w \in \mathcal{W}$. Observe that, under $\mathbb{P}_w$, $\mathsf{Red}(W_{t_2}) \cup \mathsf{Blue}(W_{t_2}) = [N] \setminus \mathsf{Green}(W_{t_2})$ is $\mathcal{H}_G$-measurable. We write $\zeta_{t_2}^{RB}$ for $\zeta_{t_2}(\mathsf{Red}(X_{t_2}) \cup \mathsf{Blue}(X_{t_2}))$ and $\tilde{\zeta}_{t_2}^{RB}$ for $\tilde{\zeta}_{t_2}(\mathsf{Red}(X_{t_2}) \cup \mathsf{Blue}(X_{t_2}))$. Conditionally on $\mathcal{H}_G$, by construction, the distribution of ζ_{t_2} and $\tilde{\zeta}_{t_2}$ on $\mathsf{Green}(W_{t_2})$ is a product of $\mathcal{B}_q$, independent of the restriction of ζ_{t_2} and $\tilde{\zeta}_{t_2}$ on $\mathsf{Red}(X_{t_2}) \cup \mathsf{Blue}(X_{t_2})$, so we can safely project onto $\mathsf{Red}(X_{t_2}) \cup \mathsf{Blue}(X_{t_2})$ to obtain

$$\begin{aligned}
&\mathbb{E}\left[\left\|\mathbb{P}_w[\zeta_{t_2} \in \cdot|\mathcal{H}_G] - \mathbb{P}_w\left[\tilde{\zeta}_{t_2} \in \cdot|\mathcal{H}_G\right]\right\|_{TV} \mathbb{1}_{\{L_{t_2} \geq 2N\}}\right] \\
&= \mathbb{E}\left[\left\|\mathbb{P}_w\left[\zeta_{t_2}^{RB} \in \cdot|\mathcal{H}_G\right] - \mathbb{P}_w\left[\tilde{\zeta}_{t_2}^{RB} \in \cdot|\mathcal{H}_G\right]\right\|_{TV} \mathbb{1}_{\{L_{t_2} \geq 2N\}}\right] \\
&\leq \mathbb{E}\left[\left\|\mathbb{P}_w\left[\zeta_{t_2}^{RB} \in \cdot|\mathcal{H}_G\right] - \nu_{\mathsf{Red}(X_{t_2}) \cup \mathsf{Blue}(X_{t_2})}\right\|_{TV} \mathbb{1}_{\{L_{t_2} \geq 2N\}}\right] \\
&\quad + \mathbb{E}\left[\left\|\mathbb{P}_w\left[\tilde{\zeta}_{t_2}^{RB} \in \cdot|\mathcal{H}_G\right] - \nu_{\mathsf{Red}(X_{t_2}) \cup \mathsf{Blue}(X_{t_2})}\right\|_{TV} \mathbb{1}_{\{L_{t_2} \geq 2N\}}\right] \\
&\leq 2\sup_x \mathbb{E}\left[\left\|\mathbb{P}_w\left[\zeta_{t_2}^{RB} \in \cdot|\mathcal{H}_G\right] - \nu_{\mathsf{Red}(X_{t_2}) \cup \mathsf{Blue}(X_{t_2})}\right\|_{TV} \mathbb{1}_{\{L_{t_2} \geq 2N\}}\right], \qquad (2.23)
\end{aligned}$$

where $\nu_{\mathsf{Red}(X_{t_2}) \cup \mathsf{Blue}(X_{t_2})}$ is the product measure of $\mathcal{B}_p$ on $\mathsf{Red}(X_{t_2}) \cup \mathsf{Blue}(X_{t_2})$. Here, we have used the triangle inequality at the first inequality and then bound everything by taking the supremum over x and $\tilde{x}$. By construction of ζ, conditionally on $\mathcal{H}_G$, the variable $\zeta_{t_2}^{RB}$ can be constructed by first sampling the set $\mathsf{Red}(W_{t_2})$, then sampling the values of ζ_{t_2} on $\mathsf{Red}(W_{t_2})$ conditionally on $\mathsf{Red}(W_{t_2})$, and then sampling independently the values of ζ_{t_2} on $\mathsf{Blue}(W_{t_2})$ by a product of $\mathcal{B}_p$. By Proposition 2.14, almost surely, conditionally on $\mathcal{H}_G$, $\mathsf{Red}(W_{t_2})$ is negative dependent. So thanks to Corollary 2.10, for $a = \dfrac{1}{\min\{p, 1-p\}}$, almost surely,

$$\begin{aligned}
&2\left\|\mathbb{P}_w\left[\zeta_{t_2}^{RB} \in \cdot|\mathcal{H}_G\right] - \nu_{\mathsf{Red}(X_{t_2}) \cup \mathsf{Blue}(X_{t_2})}\right\|_{TV} \\
&\leq \sqrt{\exp\left(\sum_{i \in \mathsf{Red}(X_{t_2}) \cup \mathsf{Blue}(X_{t_2})} (a-1)\mathbb{P}_w\left[i \in \mathsf{Red}(W_{t_2})|\mathcal{H}_G\right]^2\right) - 1}, \qquad (2.24)
\end{aligned}$$

By Corollary 2.16, on the event $\{L_{t_2} \geq 2N\}$,

$$\begin{aligned}
\sum_{i \in \mathsf{Red}(X_{t_2}) \cup \mathsf{Blue}(X_{t_2})} \mathbb{P}_w\left[i \in \mathsf{Red}(W_{t_2})|\mathcal{H}_G\right]^2 &\leq |\mathsf{Red}(X_{t_2}) \cup \mathsf{Blue}(X_{t_2})| \frac{|\mathsf{Red}(w)|^2}{N^2} \\
&\leq N\frac{|\mathsf{Red}(w)|^2}{N^2} \\
&= \frac{|\mathsf{Red}(w)|^2}{N}. \qquad (2.25)
\end{aligned}$$

The equations (2.23), (2.24), and (2.25) together imply that the third term in (2.22) is upper bounded by $\sqrt{\exp\left((a-1)\alpha^2/N\right)-1}$. So altogether, with $t_1 = t^* + 2m$ and t_2 as in Lemma 2.13,

$$\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_{w_0}\left[\zeta_t \in \cdot\right], \mathbb{P}_{w_0}\left[\tilde{\zeta}_t \in \cdot\right]\right) \leq \frac{\sqrt{N}pe^{-cm}}{\alpha} + \frac{\epsilon}{4} + \sqrt{\exp\left((a-1)\alpha^2/N\right)-1}. \tag{2.26}$$

We take $\alpha = \dfrac{\epsilon\sqrt{N}}{\sqrt{8a}}$. Then

$$\frac{a-1}{N}\alpha^2 \leq \frac{\epsilon^2}{8} \leq \log\left(1+\frac{\epsilon^2}{4}\right),$$

where we have used the inequality $\log(1+x) \geq x/2$ if $0 < x < 1$. This implies that the last term on the right-hand side of (2.26) is smaller than $\epsilon/2$. The first term on the right-hand side of (2.26) now becomes $\dfrac{\sqrt{8ap}}{\epsilon}e^{-cm}$. Observe that

$$ap = p\max\left\{\frac{1}{p}, \frac{1}{1-p}\right\} = \max\left\{1, \frac{p}{1-p}\right\} \leq \frac{1}{1-p},$$

Hence

$$\frac{\sqrt{8ap}}{\epsilon}e^{-cm} \leq \frac{1}{\epsilon}\sqrt{\frac{8}{1-p}}e^{-cm}.$$

We can take

$$m = \left\lceil \frac{1}{c}\left(\frac{1}{2}\log(128) + \frac{1}{2}\log\left(\frac{1}{1-p}\right) + 2\log(1/\epsilon)\right)\right\rceil$$

to make that term smaller than $\epsilon/4$. Hence for $t = t^* + 2m + t_2$ with m and t_2 defined as above, $\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \mathbb{P}_{\tilde{x}}\left[X_t \in \cdot\right]\right) < \epsilon$, for any $x, \tilde{x} \in \mathcal{X}$, which finishes our proof. ☐

The rest of this section is dedicated to proving Lemma 2.12, Lemma 2.13, Lemma 2.15, and Proposition 2.14. First, we prove Lemma 2.12.

Proof of Lemma 2.12. As we mentioned near the end of Subsection 2.3, each labelled individual moves as a continuous-time simple random walk on the segment where every edge has conductance N^2. Each individual is also recolored at site 1 and N at rate N^2. Hence the time at which a red individual is recolored is also the time that a continuous-time random walk on $\mathbb{Z}$, whose edges are given conductance N^2, starting from the same site reaches 0 or $N+1$ (we imagine that the red individual jumps to site 0 or site $N+1$ when it is recolored). Hence, by Lemma 2.9, the probability that an individual remains red up to time 2 is smaller than e^{-c}. Summing over all red individuals, we get

$$\mathbb{E}_w\left[|\mathsf{Red}(W_2)|\right] \leq e^{-c}\left|\mathsf{Red}(w)\right|.$$

Moreover, since $|\mathsf{Red}(W_t)|$ almost surely decreases with respect to t as the individuals are only recolored blue or green, the conclusion is obtained simply by iterating the above inequality via the Markov property. ☐

To prove Lemma 2.13, Proposition 2.14, and Lemma 2.15, we need to understand the evolution of the system conditionally on $\mathcal{H}_G$. We start by describing how the system evolves once we fix a realization of Ξ^1, Ξ^N, and Ξ^G.

Observation 2.17 (Evolution of the red and blue individuals conditionally on $\mathcal{H}_G$)**.** *Let $T_0 = 0$, and let $(T_i)_{i\geq 1}$ be the times at which a point of Ξ^1, Ξ^N, or Ξ^G appears, which are $\mathcal{H}_G$-measurable and strictly increase to infinity almost surely. Conditionally on $\mathcal{H}_G$, the green region is constant on $[T_i, T_{i+1})$, $\forall i \geq 1$. From the graphical construction of W, we see that conditionally on $\mathcal{H}_G$, when Ξ^{BR} is revealed, the red and blue individuals evolve as follows. On any time interval (T_i, T_{i+1}), $i \geq 0$, two neighbor (red or blue) individuals exchange their positions at rate N^2, and at time T_i, $i \geq 1$, the system is forced to take some transitions by the environment, as follows.*

- *Two neighbor green individuals exchange their positions. Then nothing happens to the red and blue individuals.*
- *A red or blue individual, say at site u, is forced to exchange positions with a green individual, say at site $u+1$, due to a point of Ξ^G.*
- *The reservoirs recolor a green (or red) individual blue at site 1 or a blue (or red) individual green at site N, due to a point of Ξ^1 or Ξ^N.*

In short, the blue and red individuals evolve as a simple exclusion process conditionally on the environment created by the green region.

Now we prove Proposition 2.14.

Proof of Proposition 2.14. The proof is inspired by that of Proposition 2.7. We fix a realization of Ξ^1, Ξ^N, Ξ^G and let $(T_i)_{i\geq 0}$ be defined as in Observation 2.17. We will prove the following two statements, which are conditional on $\mathcal{H}_G$:

1. For any $i \geq 0$, if $\mathsf{Red}(W_{T_i})$ is ND, then $\mathsf{Red}(W_t)$ is also ND, for any $t \in [T_i, T_{i+1})$.
2. For any $i \geq 0$, if $\mathsf{Red}(W_{T_i-})$ is ND, then $\mathsf{Red}(W_{T_i})$ is also ND.

These two statements imply that if $\mathsf{Red}(W_0)$ is ND (which is the case when W_0 is deterministic), then at any time $t \geq 0$, $\mathsf{Red}(W_t)$ is still ND, which is precisely what we want. Now we prove the first statement. From Observation 2.17, we see that on the time interval $[T_i, T_{i+1})$, the red region evolves according to the generator $\tilde{\mathcal{L}}$ given by

$$\tilde{\mathcal{L}} = N^2 \sum_u \mathcal{L}_{u,u+1},$$

with $\mathcal{L}_{u,u+1}$ as in Proposition 2.7, where the sum is taken over all the site u such that $\{u, u+1\} \cap \mathsf{Green}(W_{T_i}) = \emptyset$, as Ξ^{BR} allows us to exchange the contents on any edge $(u, u+1)$ if none of them is green. We can now apply Proposition 2.7 to conclude that the ND property of the red region is preserved from time T_i to $T_{i+1}-$. Now we prove the second statement. According to Observation 2.17, at T_i, only a few following things can happen.

- Two green individuals exchange their positions. Then nothing happens to the red region, i.e. $\mathsf{Red}(W_{T_i}) = \mathsf{Red}(W_{T_i-})$. It is straightforward that ND property is preserved.
- A red or blue individual is forced to exchange positions with a green individual, says the exchange happens between two sites v and $v+1$. Then $\mathsf{Red}(W_{T_i}) = \mathsf{Red}(W_{T_i-})^{v\leftrightarrow v+1}$. This does not affect the ND property as the inequality (2.12) for $\mathsf{Red}(W_{T_i})$ is exactly that for $\mathsf{Red}(W_{T_i-})$ when we replace A by $A^{v\leftrightarrow v+1}$.
- The reservoirs recolor an individual, for example, at site 1. This corresponds to setting the first coordinate to 0. If $1 \notin A$, this operation does not affect the inequality (2.12). If $1 \in A$, the two sides of inequality (2.12) become 0. In both cases, this operation preserves the inequality (2.12) and hence the ND property.

This finishes our proof. □

Proof of Lemma 2.13. Let $(T_i)_{i\geq 0}$ be as in Observation 2.17. Let $m \in \mathbb{N}$ be a number that we will choose later. We call *the modified dynamics* the evolution of the system where we close the reservoir at site N during the time interval $[0, 2m]$ and close the reservoir at site 1 during time $]2m, 4m]$, i.e. we ignore the points of Ξ^N on the time interval $[0, 2m]$ and the points of Ξ^1 on $]2m, 4m]$.

We claim that at time $4m$, we have fewer blue and red individuals recolored green in the modified dynamics than in the original dynamics (actually, closing some reservoirs during any time only slows down $(L_t)_{t\geq 0}$). More precisely, let $(\bar{L}_t)_{t\geq 0}$ be the number of crossings in the modified dynamics. Note that $\bar{L}_t$ is the number of triple (u, T_k, T_l) such that in the modified dynamics, the individual u is green at time T_k- (if $k > 0$) and recolored blue at time T_k (or u is simply blue or red at time $T_0 = 0$), then keeps its color until being recolored green at time T_l, for some $u \in [N]$ and some $k, l \in \mathbb{Z}_+$ such that $T_l \leq t$. In fact, each time $(\bar{L}_t)_{t\geq 0}$ jumps is a time T_l of such a triple. A similar interpretation goes for $(L_t)_{t\geq 0}$ in the original dynamics. Suppose we observe the modified dynamics in the time windows $[0, 4m]$. Suppose also that (u, T_k, T_l) is a triple counted by $(\bar{L}_t)_{t\geq 0}$ in the modified dynamics. Due to the close of the reservoirs, an individual can only be recolored blue in time $[0, 2m]$ and recolored green in time $(2m, 4m]$. Hence, almost surely, $0 \leq T_k \leq 2m < T_l \leq 4m$. Now look at the original dynamics. The particle u is necessarily green at time T_k- (if $k > 0$) since opening the reservoir at site N during time $[0, 2m]$ cannot change u to blue at time T_k-. Hence the individual u is still recolored blue at time T_k. Afterward, either it is recolored green at some time $T'_l \in [T_k, 2m]$ due to the effect of Ξ_N on $[0, 2m]$, or it remains blue until time $2m$. In the latter case, necessarily, it remains blue until the time T_l- (since opening the reservoir at site 1 cannot change u to green in $[2m, T_l-]$), and hence it is recolored green at time T_l. In summary, in the original dynamics, there exists $T_{l'}$ such that the particle u is recolored blue at time T_k, then keeps its color until being recolored green at time $T_{l'}$, where $T_k < T_{l'} \leq T_l$. The association of (u, T_k, T_l) and $(u, T_k, T_{l'})$ is injective due to the identification of the time T_k in the two triples. Note that the triple $(u, T_k, T_{l'})$ is counted by $(L_t)_{t\geq 0}$. So we can conclude that $\bar{L}_t \leq L_t$.

Now we prove that in the modified dynamics, many colorings occur at site N. The point here is that when we close the reservoir at site N, the other reservoir at site 1 quickly paints almost the whole bulk blue, and vice versa, when we close the reservoir at site 1, the reservoir at site N quickly paints almost the whole bulk green, and hence there are many blue individuals recolored green. More precisely, we write $\overline{\mathbb{E}}[\cdot], \overline{\mathbb{P}}[\cdot], \overline{\mathrm{Var}}[\cdot]$ for the expectation, the probability, and the variance taken with respect to the modified dynamics. In the modified dynamics, for any initial configuration w, the set $\mathsf{Green}(W_{2m})$ is ND, by the same argument as in Proposition 2.14. Hence by Lemma 2.7,

$$\overline{\mathrm{Var}}_w[|\mathsf{Green}(W_{2m})|] \leq \overline{\mathbb{E}}_w[|\mathsf{Green}(W_{2m})|].$$

In the modified dynamics, in the time interval $[0, 2m]$, no individual is recolored green, and the green individuals evolve as simple random walks on the segment killed at site 1 when being recolored blue. Hence, by Lemma 2.9 and a proof similar to that of Lemma 2.12, for any $m \in \mathbb{Z}_+$,

$$\overline{\mathbb{E}}_w[|\mathsf{Green}(W_{2m})|] \leq |\mathsf{Green}(w)|\, e^{-cm} \leq Ne^{-cm}.$$

Then for $m = \left\lceil \frac{1}{c} \log \frac{1000}{\epsilon} \right\rceil$,

$$\begin{aligned}
\overline{\mathbb{P}}_w[|\mathsf{Green}(W_{2m})| \geq N/4] &\leq \overline{\mathbb{P}}_w\left[|\mathsf{Green}(W_{2m})| - \overline{\mathbb{E}}_w[|\mathsf{Green}(W_{2m})|] \geq N/4 - Ne^{-cm}\right] \\
&\leq \frac{\overline{\mathrm{Var}}_w[|\mathsf{Green}(W_{2m})|]}{(1/4 - e^{-cm})^2 N^2} \\
&\leq \frac{Ne^{-cm}}{(1/4 - e^{-cm})^2 N^2} \\
&\leq \frac{\epsilon}{32}.
\end{aligned}$$

By the same argument,

$$\overline{\mathbb{P}}_w[|\mathsf{Green}(W_{4m})| \leq 3N/4] = \overline{\mathbb{P}}_w[|\mathsf{Red}(W_{4m})| + |\mathsf{Blue}(W_{4m})| \geq N/4] \leq \epsilon/32.$$

We conclude that

$$\begin{aligned}
\overline{\mathbb{P}}_w[L_{4m} \geq N/2] &\geq \overline{\mathbb{P}}_w[|\mathsf{Green}(W_{2m})| < N/4, |\mathsf{Green}(W_{4m})| > 3N/4] \\
&\geq 1 - \overline{\mathbb{P}}_w[|\mathsf{Green}(W_{2m})| \geq N/4] - \overline{\mathbb{P}}_w[|\mathsf{Green}(W_{4m})| \leq 3N/4] \\
&\geq 1 - \epsilon/16.
\end{aligned}$$

This implies that, in the original dynamics, we also have

$$\mathbb{P}_w[L_{4m} \geq N/2] \geq 1 - \epsilon/16.$$

Therefore, by an argument of union bound and the Markov property,

$$\mathbb{P}_w[L_{16m} \geq 2N] \geq 1 - 4 \times \epsilon/16 = 1 - \epsilon/4.$$

We choose $t_2 = 16m$ to conclude the proof. □

Remark 1. The number t_2 above is not meant to be optimal. In fact, if we are interested only in the case of large N, we can choose $m = \frac{1+o(1)}{c} \log 4$.

The rest of this section is devoted to proving the anticoncentration inequality in Lemma 2.15. We introduce some notations that we will use in Lemma 2.15 and Proposition 2.18. Let $(T_i)_{i\geq 0}$ be defined as in Observation 2.17. We fix $t \in \mathbb{R}_+$ and denote by $T_{a_1}, \ldots, T_{a_r}$ the times at which a green individual is recolored blue at site 1, and $T_{b_1}, \ldots, T_{b_s}$ the times at which a red or blue individual is recolored green at site N, during the time interval $[0, t]$. All these times are $\mathcal{H}_G$-measurable. By abuse of notation, we relabel those individuals by $\tilde{a}_1, \ldots, \tilde{a}_r, \tilde{b}_1, \ldots, \tilde{b}_s$.

Another description of the colored interchange process. The above notations can be understood as follows. At each time T_{a_i}, the reservoir at site 1 removes the green individual at it and replaces that one with a new blue one, which we label $\tilde{a}_i$, $1 \leq i \leq r$. We have similar interpretations for $T_{b_i}, \tilde{b}_i$, $1 \leq i \leq s$. We denote the random walk of the individual $\tilde{a}_i$ by A_i. This means A_i is the trajectory of the blue individual $\tilde{a}_i$ until it is removed from the bulk by the reservoir at site N. If the individual labelled u is blue or red at time 0, we denote by $\tilde{\sigma}(u, \cdot)$ the walk of the individual u until being removed by the reservoir at site N. Observe that the walk $\tilde{\sigma}(u, \cdot)$ either survives up to time t or is removed at some time T_{b_l}. We will use the notation $\{\tilde{\sigma}(u, t) = v\}$ to indicate the event that the individual u survives up to time t and reaches the site v at time t, and $\{\tilde{\sigma}(u, t) = \tilde{b}_l\}$ to say that the individual u is removed at time T_{b_l}. Similar notations are used for the walk A_k, $1 \leq k \leq r$.
We will need the following proposition.

Proposition 2.18 (Crossing inequality)**.** *For any $w \in \mathcal{W}$, for any individual labelled u that is either blue or red in w, for any site v, for any $1 \leq k \leq r$, $1 \leq l \leq s$,*

$$\mathbb{P}_w \left[\tilde{\sigma}(u, t) = v, A_k(t) = \tilde{b}_l \middle| \mathcal{H}_G \right] \leq \mathbb{P}_w \left[\tilde{\sigma}(u, t) = \tilde{b}_l, A_k(t) = v \middle| \mathcal{H}_G \right]. \tag{2.27}$$

We show how we can prove Lemma 2.15 using Proposition 2.18.

Proof of Lemma 2.15. Let u, v be as in the statement of Lemma 2.15. Summing the inequality in Proposition 2.18 over l, we get

$$\mathbb{P}_w \left[\tilde{\sigma}(u, t) = v, A_k \text{ is killed by time } t | \mathcal{H}_G \right] \leq \mathbb{P}_w \left[\tilde{\sigma}(u, \cdot) \text{ is killed by time } t, A_k(t) = v | \mathcal{H}_G \right].$$

Now summing over k, we get

$$\begin{aligned} &\mathbb{E}_w \left[\mathbb{1}_{\{\tilde{\sigma}(u,t)=v\}} \times \#\{\text{walks among } A_1, \ldots, A_r \text{ that are killed by time } t\} \middle| \mathcal{H}_G \right] \\ &\leq \sum_{k=1}^{s} \mathbb{P}_w \left[\tilde{\sigma}(u, \cdot) \text{ is killed by time } t, A_k(t) = v | \mathcal{H}_G \right]. \end{aligned} \tag{2.28}$$

On the event $\{L_t \geq 2N\}$, there are at least $2N$ times at which a blue or red individual is recolored green by time t. Since originally there are at most N blue and red individuals, then there are always at least N times a green individual is recolored blue at site 1 and then recolored green again by time t. In other words, at least N walks are killed at site N by time t among $A_1, \dots, A_r$. Hence the left-hand side of equation (2.28) is at least $N \times \mathbb{P}\left[\tilde{\sigma}(u,t) = v \middle| \mathcal{H}_G\right]$. On the other hand, we can realize the trajectories of $\tilde{\sigma}(u,\cdot)$ and $A_1, A_2, \dots, A_r$ altogether by revealing Ξ^{BR}. Subsequently, the events on the right-hand side of (2.28) are pairwise disjoint since A_k and A_l cannot both occupy site v at time t, for any $k \neq l$. This implies that the sum on the right-hand side of (2.28) is smaller than 1. So overall, on the event $\{L_t \geq 2N\}$,

$$N \times \mathbb{P}_w\left[\tilde{\sigma}(u,t) = v \middle| \mathcal{H}_G\right] \leq 1. \tag{2.29}$$

Furthermore,

$$\mathbb{P}_w\left[\tilde{\sigma}(u,t) = v \middle| \mathcal{H}_G\right] = \mathbb{P}_w\left[\sigma_t(u) = v, \sigma_s(u) \notin \mathsf{Green}(W_s), \forall 0 \leq s \leq t \middle| \mathcal{H}_G\right]. \tag{2.30}$$

Combining (2.29) and (2.30), we get what we want. □

We now prove Proposition 2.18.

Proof of Proposition 2.18. We encourage the reader to draw a picture to follow the proof more easily. The intuition behind this is as follows. We consider two walks $\tilde{\sigma}(u,\cdot), A_k$, and we condition on the event that one of them survives and reaches the site v at time t, while the other is removed at time T_{b_l} by the reservoir at site N. If the two walks do not cross each other, then necessarily, $\tilde{\sigma}(u,t) = \tilde{b}_l$ and $A_k(t) = v$. On the other hand, if the two walks cross, then we can no longer distinguish them, and hence in this case, $\{\tilde{\sigma}(u,t) = \tilde{b}_l, A_k(t) = v\}$ and $\{\tilde{\sigma}(u,t) = v, A_k(t) = \tilde{b}_l\}$ are equiprobable. The proof is to turn this intuition into a rigorous mathematical argument.

By Observation 2.17, we see that $\tilde{\sigma}(u,\cdot)$ is a simple random walk conditionally on the environment created by the green region. If that walk is still alive at the time when $\tilde{a}_k$ is born, then it must be on the right of $\tilde{a}_k$ at that time (since $\tilde{a}_k$ is born at site 1). Those two individuals then evolve as a simple exclusion process with two individuals conditionally on the environment as we realize Ξ^{BR}, still by Observation 2.17. We propose another graphical construction of those two walks as follows. At any time t, if there are two walks alive, say at two sites u, v with $u < v$, we refer to the individual at site u as the individual on the left and to the individual at site v as the individual on the right. In case there is only one walk alive, we refer to it as the only individual. Let Ξ_R, Ξ_L, Ξ_E be 3 independent Poisson point processes with Ξ_R and Ξ_L of intensity $N^2 \mathrm{d}t \otimes \mathrm{Card}$ on $\mathbb{R}_+ \times \{L, R\}$, and Ξ_E of intensity $2N^2 \mathrm{d}t$ on $\mathbb{R}_+$. R, L, and E stand for right, left, and exchange, and Card is the counting measure. The rule of evolution is as follows.

- For each point (s, w) of Ξ_R, the individual on the right at time $s-$ (or the only individual if there is only one walk alive) attempts to make a jump at time s to the left if $w = L$ or to the right if $w = R$. It succeeds except when trying to jump on a site occupied by a green individual or the individual on the left, in which case the jump is cancelled.

- For each point (s, w) of Ξ_L, the individual on the left at time $s-$ (if there exists) attempts to make a jump at time s to the left if $w = L$ or to the right if $w = R$. It succeeds except when trying to jump on a site occupied by a green individual or by the individual on the right, in which case the jump is cancelled.

- For each point s of Ξ_E, we exchange the positions of the two individuals with probability $1/2$ if they are adjacent at time s.
- For each time T_i, the two walks make the jump forced by the environment, as explained in Observation 2.17.

This construction gives us the same distribution of the two walks as the one given by Ξ^{BR}, as the rates of transitions given by the two constructions are always the same. In short, Ξ_R and Ξ_L are used to generate the walks on the right and the left, respectively. When we realize Ξ_R, Ξ_L, we observe two random walks that are not allowed to jump on top of each other, i.e. do not cross. Ξ_E is then used to make precise where those two walks exchange their positions, i.e. to add their possible crossings. With this construction, the set $\{\tilde{\sigma}(u,s), A_k(s)\}$ is entirely determined by Ξ_R, Ξ_L, for any $0 \le s \le t$, since Ξ_E has no effect on that set. Now conditionally on the (Ξ_R, Ξ_L)-measurable event $\{\tilde{\sigma}(u,t), A_k(t)\} = \{v, \tilde{b}_l\}$ (which means that one walk reaches v at time t and the other walk was killed at b_l), we realize Ξ_E to obtain a number m of exchange edges between the two walks. We see that if $m = 0$, then we necessarily have $\tilde{\sigma}(u,t) = \tilde{b}_l, A_k(t) = v$ due to monotonicity, which means

$$0 = \mathbb{P}_w\left[\tilde{\sigma}(u,t) = v, A_k(t) = \tilde{b}_l, m = 0 \middle| \mathcal{H}_G\right] \le \mathbb{P}_w\left[\tilde{\sigma}(u,t) = \tilde{b}_l, A_k(t) = v, m = 0 \middle| \mathcal{H}_G\right]. \quad (2.31)$$

If $m > 0$, we cannot distinguish the two individuals anymore since the probability that the sum of m independent Bernoulli variables of parameter $1/2$ is even (or odd) is $1/2$, which means

$$\mathbb{P}_w\left[\tilde{\sigma}(u,t) = v, A_k(t) = \tilde{b}_l, m > 0 \middle| \mathcal{H}_G\right] = \mathbb{P}_w\left[\tilde{\sigma}(u,t) = \tilde{b}_l, A_k(t) = v, m > 0 \middle| \mathcal{H}_G\right], \quad (2.32)$$

since both are equal to $\frac{1}{2}\mathbb{P}_w\left[\{W_t(t), A_k(t)\} = \{j, \tilde{b}_l\}, m > 0 \middle| \mathcal{H}_G\right]$. Summing (2.31) and (2.32), we get what we want. □

4 The lower bound

We finish the proof of Theorem 2.2 by proving the lower bound on $t_{\text{mix}}(\epsilon)$ in Theorem 2.2.

Proof of the lower bound. We still use the distinguishing statistics method. We consider the process with the initial condition $X_0 = \mathbf{1}$. Recall that the weight of a configuration x is

$$S(x) := \sum_{u=1}^{N} x(u).$$

$S(x)$ will be our distinguishing statistic. Recall that, for any $t \ge 0$,

$$X_t \overset{d}{=} f_*(\mathbf{1}, W_t, \xi^B, \xi^G),$$

with W_t, ξ^B, ξ^G defined as in Lemma 2.11. We write S_t for $S\left(f_*(\mathbf{1}, W_t, \xi^B, \xi^G)\right)$, and S_∞ for $S(X_\infty)$, where $X_\infty \sim \pi$. By symmetry, we see that,

$$\mathbb{E}\left[|\mathsf{Blue}(W_t)|\right] = \mathbb{E}\left[|\mathsf{Green}(W_t)|\right] = \frac{N - \mathbb{E}\left[|\mathsf{Red}(W_t)|\right]}{2}.$$

Hence

$$\begin{aligned}\mathbb{E}\left[S_t\right] &= \mathbb{E}\left[\mathbb{E}\left[S_t | W_t\right]\right] \\ &= \mathbb{E}\left[|\mathsf{Red}(W_t)| + p\,|\mathsf{Blue}(W_t)| + q\,|\mathsf{Green}(W_t)|\right] \\ &= \mathbb{E}\left[|\mathsf{Red}(W_t)|\right] + \frac{N - \mathbb{E}\left[|\mathsf{Red}(W_t)|\right]}{2}(p+q),\end{aligned}$$

Letting $t \to \infty$, and observing that $\mathbb{E}\left[|\mathsf{Red}(W_t)|\right] \xrightarrow{t\to\infty} 0$ by Lemma 2.12, we get

$$\begin{aligned}\mathbb{E}\left[S_\infty\right] &= N \times \frac{p+q}{2}, \\ \mathbb{E}\left[S_t\right] - \mathbb{E}\left[S_\infty\right] &= \mathbb{E}\left[|\mathsf{Red}(W_t)|\right]\left(1 - \frac{p+q}{2}\right).\end{aligned}$$

By Proposition 2.7, X_t is ND. Hence by inequality (2.13),

$$\mathrm{Var}\left[S_t\right] \leq \mathbb{E}\left[S_t\right],$$

and by letting $t \to \infty$,

$$\mathrm{Var}\left[S_\infty\right] \leq \mathbb{E}\left[S_\infty\right].$$

Hence, under assumption (2.2),

$$\max\left\{\mathrm{Var}\left[S_t\right], \mathrm{Var}\left[S_\infty\right]\right\} \leq \frac{N(p+q)}{2} + \mathbb{E}\left[|\mathsf{Red}(W_t)|\right]\left(1 - \frac{p+q}{2}\right) \leq Np + \mathbb{E}\left[|\mathsf{Red}(W_t)|\right].$$

So by Proposition 7.9 in [59],

$$\|\mathbb{P}_{\mathbf{1}}\left[X_t \in \cdot\right] - \pi\|_{TV} \geq 1 - 8\frac{\max\left\{\mathrm{Var}\left[S_t\right], \mathrm{Var}\left[S_\infty\right]\right\}}{\left(\mathbb{E}\left[S_t\right] - \mathbb{E}\left[S_\infty\right]\right)^2} \geq 1 - 8\frac{Np + \mathbb{E}\left[|\mathsf{Red}(W_t)|\right]}{\left(1 - \frac{p+q}{2}\right)^2 \mathbb{E}\left[|\mathsf{Red}(W_t)|\right]^2}.$$

Let $t = t^* - 2m$, for some positive integer $m < t^*/2$ that we will choose later, and let c be the constant in Lemma 2.9. By Lemma 2.12 and the Markov property, $\mathbb{E}\left[|\mathsf{Red}(W_t)|\right] \geq \mathbb{E}\left[|\mathsf{Red}(W_{t^*})|\right] e^{cm} = (\sqrt{Np} \vee 1)e^{cm}$, hence the last term is bigger than

$$1 - \frac{8}{\left(1 - \frac{p+q}{2}\right)^2 e^{2cm}} - \frac{8}{\left(1 - \frac{p+q}{2}\right)^2 e^{cm}}.$$

Note that with our assumption, $1 > 1 - \frac{p+q}{2} > \frac{1}{2}$. So we conclude that the term above is greater than $1 - \frac{32}{e^{2cm}} - \frac{32}{e^{cm}}$, which is greater than ϵ for $m = \left\lceil \frac{1}{c}\log\left(\frac{64}{1-\epsilon}\right)\right\rceil$. This implies that $\mathrm{t_{mix}}(\epsilon) \geq t^* - 2m$, if $m < t^*/2$. Besides, this inequality is trivial when $m \geq t^*/2$, which finishes our proof. □

Remark 2. This method can also give a lower bound on the mixing times from any initial configuration x. In fact, we can prove that, for some constant C,

$$\mathrm{t_{mix}}(x;\epsilon) \geq t^*(x) - C\left(1 + \log\left(\frac{1}{1-\epsilon}\right)\right),$$

where

$$t^*(x) = \inf\left\{t \geq 0 : \left|\mathbb{E}_x^0\left[S_t\right] - \mathbb{E}\left[S_\infty\right]\right| \leq \sqrt{Np} \vee 1\right\}.$$

5 The computation of t^*

For the sake of completeness, we include here the proof of equality (2.5), which is a particular case of the computations presented in Appendix A of [37].

Proof of (2.5). We consider the model where the reservoirs are of density $p = q = 0$. We define $f_t : [N] \to [0,1]$ by $f_t(u) = \mathbb{E}_{\mathbf{1}}\left[X_t(u)\right]$. Then by Dynkin's formula, $\{f_t;\, t \geq 0\}$ is the unique solution of the system of equations

$$\begin{cases} \dfrac{d}{dt} f_t(u) = \Delta f_t(u) & \text{for } t \geq 0 \text{ and } u \in [N], \\ f_0(u) = 1 & \text{for all } u \in [N], \end{cases}$$

where Δ is the discrete Laplacian defined by, for any function $\varphi : [N] \to \mathbb{R}$,

$$\Delta\varphi = N^2(\varphi(u+1) + \varphi(u-1) - 2\varphi(u)),\ \forall u \in [N],$$

with the convention that $\varphi(0) = \varphi(N+1) = 0$. The eigenfunctions of Δ are given by

$$\varphi_l(u) = \sqrt{2}\sin\left(\frac{\pi l u}{N+1}\right), \quad 1 \leq l \leq N,$$

with the corresponding eigenvalues $-\lambda_l$ given by

$$\lambda_l = 2N^2\left(1 - \cos\left(\frac{\pi l}{N+1}\right)\right), \quad 1 \leq l \leq N.$$

Besides, $(\varphi_l)_{1\leq l\leq N}$ is an orthonormal basis for the scalar product given by

$$\langle \varphi, \psi \rangle = \frac{1}{N+1}\sum_{u=1}^{N} \varphi(u)\psi(u),$$

for any $\varphi, \psi : [N] \to \mathbb{R}$. Let t be an arbitrary positive number. We see that

$$\mathbb{E}_{\mathbf{1}}\left[|\mathsf{Red}(W_t)|\right] = \sum_{u=1}^{N} f_t(x) = (N+1)\left\langle f_t, f_0 \right\rangle. \tag{2.33}$$

Let

$$f_0 = \sum_{l=1}^{N} c_l \varphi_l$$

be the decomposition of f_0 in the basis $(\varphi_l)_{1\leq l\leq N}$. Then

$$f_t = \sum_{l=1}^{N} c_l e^{-\lambda_l t} \varphi_l.$$

Hence

$$\langle f_t, f_0 \rangle = \sum_{l=1}^{N} c_l^2 e^{-\lambda_l t}. \tag{2.34}$$

Note that $\lambda_1 = \inf\{\lambda_1, \ldots, \lambda_N\}$. Hence

$$c_1^2 e^{-\lambda_1 t} \leq \langle f_t, f_0 \rangle \leq \left(\sum_{l=1}^{N} c_l^2\right) e^{-\lambda_1 t} = \langle f_0, f_0 \rangle\, e^{-\lambda_1 t} = \frac{N}{N+1} e^{-\lambda_1 t}.$$

Plug in (2.33), we see that

$$(N+1) c_1^2 e^{-\lambda_1 t} \leq \mathbb{E}_{\mathbf{1}}\left[|\mathsf{Red}(W_t)|\right] \leq N e^{-\lambda_1 t},$$

which means

$$\frac{1}{\lambda_1}\left(2\log(c_1) + \log\left(\frac{N+1}{\mathbb{E}_{\mathbf{1}}\left[|\mathsf{Red}(W_t)|\right]}\right)\right) \leq t \leq \frac{1}{\lambda_1}\log\left(\frac{N}{\mathbb{E}_{\mathbf{1}}\left[|\mathsf{Red}(W_t)|\right]}\right). \tag{2.35}$$

We finish the proof by estimating λ_1 and c_1. Note that by Taylor's expansion of function cosine around 0,

$$\lambda_1 = 2N^2\left(\frac{1}{2}\frac{\pi^2}{(N+1)^2} + \mathcal{O}\left(\frac{\pi^4}{(N+1)^4}\right)\right) = \pi^2 + \mathcal{O}\left(1/N\right). \tag{2.36}$$

Besides,

$$\begin{aligned} c_1 &= \langle \mathbf{1}, \varphi_1 \rangle \\ &= \frac{1}{N+1}\sum_{l=1}^{N} \varphi_1(l) \\ &= \frac{\sqrt{2}}{N+1}\sum_{l=1}^{N} \sin\frac{\pi l}{N+1}. \end{aligned}$$

By some classical trigonometric computations,

$$c_1 = \frac{\sqrt{2}}{N+1} \times \frac{\cos\dfrac{\pi}{2(N+1)}}{\sin\dfrac{\pi}{2(N+1)}} = \frac{2\sqrt{2}}{\pi}\left(1 + \mathcal{O}\left(1/N^2\right)\right), \tag{2.37}$$

where we have used the Taylor expansion of the sine and cosine functions around 0. Note that at t^*, $\mathbb{E}_{\mathbf{1}}\left[|\mathsf{Red}(W_{t^*})|\right] = \sqrt{Np} \vee 1$, by (2.17). Then the three equations (2.35), (2.36), (2.37) together imply

$$t^* = \frac{1}{\pi^2} \log\left(\frac{N}{\sqrt{Np} \vee 1}\right) \pm \mathcal{O}\left(1\right),$$

which is precisely what we want. □

Chapter 3

The Glauber-Exclusion process

Goals

In this chapter, we discuss the Glauber-Exclusion process, a superposition of a Glauber dynamics and the Symmetric Simple Exclusion Process (SSEP) on the lattice. The model was shown to admit a reaction-diffusion equation as the hydrodynamic limit. We define a notion of temperature regimes via the reaction function in the equation and prove cutoff in the full high-temperature regime for the attractive model in dimensions 1 and 2 with periodic boundary conditions. Our results show that the equation in the hydrodynamic limit reflects the mixing behavior of the large but finite system. Besides, cutoff is proved despite the lack of reversibility and an explicit formula for the invariant measure. We also provide the spectral gap and prove pre-cutoff in all dimensions. The proof still follows the outline of the one presented in Chapter 1. However, nontrivial adaptations and novel ingredients are needed. Our proof involves a new interpretation of attractiveness, the information percolation framework introduced by Lubetzky and Sly, anti-concentration of simple random walk on the lattice, and a coupling inspired by excursion theory. We hope that this approach can find new applications in the future. The results presented here can be found in the preprint [98].

Contents

1 Introduction

The Glauber-Exclusion process, introduced in [17, 18], is a superposition of a Glauber dynamics and an appropriately accelerated symmetric simple exclusion process (SSEP) on a lattice. The model is also called Glauber-Kawasaki dynamics in [19]. Here, we use the term Glauber-Exclusion as in [93]. The model can be described as follows.

Notations. Fix a number $d \in \mathbb{Z}_+$. For any $L \in \mathbb{Z}_+$, let $\Lambda_L^d := (\mathbb{Z}/L\mathbb{Z})^d$ be the torus of side-length L in dimension d. For any finite set Ω, we call the elements of $\{-1,1\}^\Omega$ the spin configurations on Ω, and we denote by $\leq$ the usual coordinate-wise order on $\{-1,1\}^\Omega$. Our state space is $\mathcal{X} := \{-1,1\}^{\Lambda_L^d}$, the set of all spin configurations on Λ_L^d. For any $x \in \mathcal{X}$, $u, u' \in \Lambda_L^d$, we denote by $x^{u \leftrightarrow u'}$ the configuration obtained from x by exchanging the values at two sites u and u', and $x^{u,1}$ (resp. $x^{u,-1}$) the configuration obtained from x by replacing the value at site u by 1 (resp. -1). We also recall that $\exp(\theta)$ denotes the exponential distribution with parameter θ, i.e. the distribution with density $\mathbb{1}_{\{s>0\}}\theta e^{-\theta s}\mathrm{d}s$.

The generator of the symmetric simple exclusion process (SSEP) $\mathcal{L}_E$ acts on an observable $\varphi : \mathcal{X} \to \mathbb{R}$ by

$$(\mathcal{L}_E \varphi)(x) := \sum_{u \sim u'} \left(\varphi(x^{u \leftrightarrow u'}) - \varphi(x) \right), \tag{3.1}$$

where the sum is taken over all pairs of neighbors $\{u, u'\}$ on the lattice. This means that $\mathcal{L}_E$ exchanges the contents of two neighboring sites u, u' at rate 1.

To define the Glauber dynamics, we first need the following definition.

Flip-rate function. Throughout this chapter, we fix a number $m \in \mathbb{Z}_+$ and a function $c : \{-1,1\}^{B(0,m)} \to \mathbb{R}_{>0}$, where $B(0,m)$ is the subset of $\mathbb{Z}^d$ consisting of all sites of distance at most m from 0. We call this function c the *local flip-rate function*.

For any L large enough, we identify $B(0,m)$ with the subset of Λ_L^d consisting of all sites of distance at most m from site 0. The function c extends to a *global flip-rate function* $\hat{c} : \Lambda_L^d \times \mathcal{X} \to \mathbb{R}_{>0}$ as follows.

1. $\hat{c}(0,\cdot)$ is local and given by c:

$$\forall x \in \mathcal{X},\ \hat{c}(0,x) := c\left(x|_{B(0,m)}\right).$$

2. $\hat{c}$ is invariant by translation:

$$\forall x \in \mathcal{X}, u \in \Lambda_L^d,\ \hat{c}(u,x) := \hat{c}(0, x_{u+\cdot}),$$

   where

$$x_{u+\cdot}(u') := x(u+u'),\ \forall x \in \mathcal{X},\ u, u' \in \Lambda_L^d.$$

The generator $\mathcal{L}_G$ of the Glauber dynamics associated with the local flip-rate function $c(\cdot)$ acts on an observable $\varphi : \mathcal{X} \to \mathbb{R}$ by

$$(\mathcal{L}_G\varphi)(x) := \sum_{u\in\Lambda_L^d} \hat{c}(u,x)(\varphi(x^{u,-x(u)}) - \varphi(x)). \tag{3.2}$$

This means that $\mathcal{L}_G$ flips the spin on site u of configuration x at rate $\hat{c}(u,x)$, for any $(u,x) \in \Lambda_L^d \times \mathcal{X}$. We suppose that the Glauber dynamics is attractive.

Hypothesis 1 (Attractiveness)**.** *For any $x, y \in \{-1,1\}^{B(0,m)}$ such that $x \leq y$,*

$$c(x) \geq c(y) \text{ if } x(0) = y(0) = 1, \tag{3.3}$$

$$c(x) \leq c(y) \text{ if } x(0) = y(0) = -1. \tag{3.4}$$

Hypothesis 1 implies that, for any $x, y \in \mathcal{X}$ such that $x \leq y$, for any $u \in \Lambda_L^d$,

$$\hat{c}(u,x) \geq \hat{c}(u,y) \text{ if } x(u) = y(u) = 1, \tag{3.5}$$

$$\hat{c}(u,x) \leq \hat{c}(u,y) \text{ if } x(u) = y(u) = -1. \tag{3.6}$$

These conditions are necessary and sufficient to construct a Markovian coupling of two processes whose generators are $\mathcal{L}_G$ that preserves order. More precisely, suppose that $(Y_t^1)_{t\geq 0}$ and $(Y_t^2)_{t\geq 0}$ are two Markov processes with generator $\mathcal{L}_G$ starting at y_1, y_2, respectively. Then there is a Markovian coupling of $(Y_t^1)_{t\geq 0}$ and $(Y_t^2)_{t\geq 0}$ such that, almost surely,

$$y_1 \leq y_2 \Rightarrow \forall t \geq 0,\ Y_t^1 \leq Y_t^2.$$

See Theorem 4.11, p.143 in [62] for more details. When such an order-preserving coupling exists, we say that the process Y is attractive.

Although we will work with a general function $c(\cdot)$, we usually use the following example considered by De Masi *et al* in [18] to illustrate our method.

Example 3.1 (Example of De Masi *et al*)**.** *The model considered is in dimension one, i.e. $d = 1$, with the local flip-rate function c given by*

$$\forall x \in \{-1,1\}^{\{1,0,-1\}},\ c(x) := 1 - \gamma x(0)(x(1) + x(-1)) + \gamma^2 x(1)x(-1),$$

for some $\gamma \in [0,1)$. This corresponds to the global flip-rate function $\hat{c}$ given by

$$\forall x \in \mathcal{X}, u \in \Lambda_L,\ c(u,x) := 1 - \gamma x(u)(x(u+1) + x(u-1)) + \gamma^2 x(u+1)x(u-1).$$

The *Glauber-Exclusion process* is a continuous time Markov process $X = (X_t)_{t\geq 0}$, taking values in $\mathcal{X}$, whose generator is

$$\mathcal{L}_{GE} := L^2\mathcal{L}_E + \mathcal{L}_G. \tag{3.7}$$

This is a superposition of a Glauber dynamics and a SSEP accelerated by a diffusive factor L^2.

For $\rho \in [-1,1]$, we denote by $\mathfrak{R}_\rho$ the Rademacher probability measure on $\{-1,1\}$ with average ρ. More precisely,

$$\mathfrak{R}_\rho(1) = \frac{1+\rho}{2},$$
$$\mathfrak{R}_\rho(-1) = \frac{1-\rho}{2}.$$

Reaction function. We denote by ν_ρ the product measure of $\mathfrak{R}_\rho$ on $\{-1,1\}^{B(0,m)}$. We define the polynomial $R(\cdot) : [-1,1] \to \mathbb{R}$ by

$$\forall \rho \in [-1,1],\ R(\rho) := \mathbb{E}_{\nu_\rho}\left[-2\xi_0 c(\xi)\right],$$

where $\xi = (\xi_u)_{u\in B(0,m)} \sim \nu_\rho$. We call $R(\cdot)$ the *reaction function*. The reason behind this name will be explained in the next paragraph.

The hydrodynamic limit. Originally, De Masi, Ferrari, and Lebowitz introduced the model in [17, 18] for the infinite lattice $\mathbb{Z}^d$ and proved that the hydrodynamic limit of the model is a reaction-diffusion equation. We recall here the analogous result obtained by Kipnis, Olla, and Varadhan [46] for our case in a finite box. Let the density field on $(\mathbb{R}/\mathbb{Z})^d$ at time t be defined by

$$\mu_t^L(\mathrm{d}u) = \frac{1}{L^d}\sum_{u\in\Lambda_L^d} X_t(u)\delta_{u/L}(\mathrm{d}u) \ \text{ on } (\mathbb{R}/\mathbb{Z})^d.$$

Then, when L grows to infinity, under appropriate convergence conditions on the initial data, the density field converges weakly to the solution $\rho : \mathbb{R}_+ \times (\mathbb{R}/\mathbb{Z})^d \to [-1,1]$ of the following reaction-diffusion equation:

$$\frac{\partial \rho}{\partial t} = \Delta\rho + R(\rho), \tag{3.8}$$

where Δ is the Laplacian in $(\mathbb{R}/\mathbb{Z})^d$, and $R(\cdot)$ is the function defined above. Equation (3.8) explains the name reaction function of $R(\cdot)$. The intuition behind this result is as follows. Applying the generator $\mathcal{L}_{GE}$ to the function $\varphi : x \mapsto x(u)$, we see that

$$\mathcal{L}\varphi(x) = \left(L^2\sum_{u'\sim u}(x(u') - x(u))\right) - 2x(u)\hat{c}(u,x).$$

Therefore,

$$\frac{d}{dt}\mathbb{E}\left[X_t(u)\right] = \mathbb{E}\left[\mathcal{L}\varphi(X_t)\right] = \left(L^2\sum_{u'\sim u}\mathbb{E}\left[X_t(u') - X_t(u)\right]\right) - 2\mathbb{E}\left[X_t(u)\hat{c}(u,X_t)\right].$$

Suppose that $\mathbb{E}\left[X_t(u)\right]$ is well approximated by $\rho(t,u/L)$ for some smooth function ρ, for any site u and any fixed time t. Then

$$L^2\sum_{u'\sim u}\mathbb{E}\left[X_t(u') - X_t(u)\right] \approx L^2\sum_{u'\sim u}\left(\rho(t,u'/L) - \rho(t,u/L)\right) \approx \Delta\rho(t,u/L).$$

Moreover, the local law around a site u at any fixed time t is believed to be close to that of a product of $\mathfrak{R}_{\mathbb{E}[X_t(u)]}$, a phenomenon known as *local equilibrium*, see Subchapter 1.2 and Chapter 3 in [47] for more details. Then

$$\mathbb{E}\left[X_t(u)\hat{c}(u,X_t)\right] \approx \mathbb{E}_{\nu_{\mathbb{E}[X_t(u)]}}\left[\xi_0 c(\xi)\right] = R\left(\mathbb{E}\left[X_t(u)\right]\right) \approx R(\rho(t,u/L)).$$

These intuitive arguments indicate that $\rho(\cdot,\cdot)$ satisfies equation (3.8). Showing the convergence amounts to verifying the above intuition rigorously.

Based on this result, we define the *temperature regimes* as follows.

Temperature regimes.

- *High-temperature regime*: The system is at high temperature if the reaction function $R(\cdot)$ has a unique root ρ_* in $[-1,1]$, and $R'(\rho_*) < 0$.
- *Critical-temperature regime*: The system is at critical temperature if the reaction function $R(\cdot)$ has a unique root (not counting multiplicity) ρ_* in $[-1,1]$, and $R'(\rho_*) = 0$.
- *Low-temperature regime*: The system is at low temperature if the reaction function $R(\cdot)$ has at least two roots (not counting multiplicity) in $[-1,1]$.

Explanation for the temperature regimes. Let us explain why we define the temperature regimes for the model as above. Consider the equation (3.8) with the initial condition $\rho(0,\cdot)$ constant in space, say $\rho(0,\cdot) = \rho_0$ for some $\rho_0 \in [-1,1]$. We then see that $\rho(t,\cdot)$ must also be constant in space, so we can safely denote by $\rho(t)$ the value of the function ρ at time t. The equation (3.8) then simply becomes an ODE:

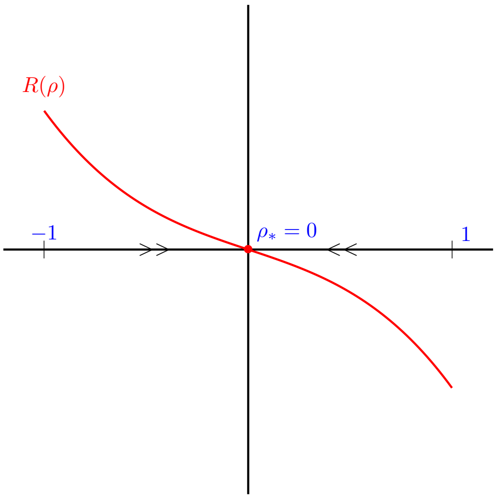
$$\begin{cases} \rho' = R(\rho), \\ \rho(0) = \rho_0. \end{cases} \tag{3.9}$$

Though simple, the solution of (3.9) can exhibit very different behaviors, depending on $R(\cdot)$. Consider the flip-rate function in Example 3.1. Then $R(\rho) = -2(1-2\gamma)\rho - 2\gamma^2\rho^3$. One can verify that the system is at high temperature if $0 \le \gamma < 1/2$, at critical temperature if $\gamma = 1/2$, and at low temperature if $1/2 < \gamma < 1$. We present in Figure 3.1 the phase diagrams of the ODE (3.9) for two different values of γ.

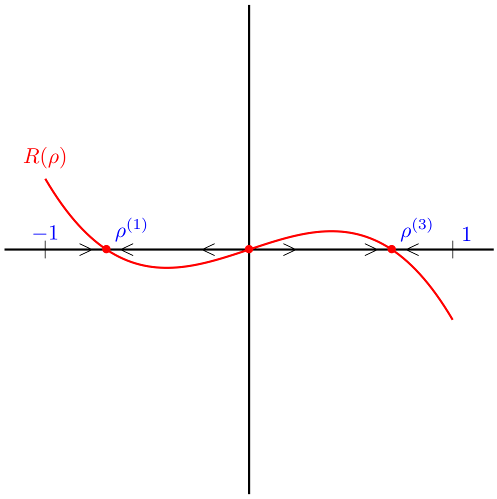
Let $\rho_+(\cdot)$ and $\rho_-(\cdot)$ denote the solutions of the ODE (3.9) with initial conditions $\rho_+(0) := 1$ and $\rho_-(0) := -1$. Viewing the diagram, we see that, in the first case, for any initial condition ρ_0, the solution of ODE (3.9) converges to the unique root ρ_* exponentially fast: $|\rho(t) - \rho_*| \lesssim e^{R'(\rho_*)t}$. In the second case, $R(\cdot)$ has 3 roots $\rho^{(1)} < \rho^{(2)} < \rho^{(3)}$, and $\rho_-(t) \xrightarrow{t\to\infty} \rho^{(1)}$, and $\rho_+(t) \xrightarrow{t\to\infty} \rho^{(3)}$. Intuitively, this implies that the system mixes rapidly at high temperatures, while at low temperatures, it exhibits the phenomenon known as *metastability*; see [10] for more details.

In this chapter, we study the model in the high-temperature regime.

Hypothesis 2 (High temperature)**.**



(a) High temperature



(b) Low temperature

Figure 3.1: Phase diagram for $\gamma = 5/12$ and $\gamma = 7/12$.

- *$R(\rho)$ admits a unique root ρ_* in $[-1, 1]$,*
- *$R'(\rho_*) < 0$.*

The primary purpose of this chapter is to study the mixing times of the system in the high-temperature regime.

Conjectured universal behaviors. A stochastic spin system, with an appropriate notion of temperature, under an appropriate time scale, is conjectured to exhibit the following behaviors.

- At high temperatures, the system mixes fast. The inverse-spectral-gap $\texttt{gap}^{-1}$ is $\mathcal{O}(1)$, the mixing times are logarithmic in the size of the lattice, and cutoff occurs.
- At critical temperatures, the inverse gap is (sub)-polynomial in the size of the lattice, and so are the mixing times.
- At low temperatures, the inverse gap is exponential in the size of the lattice, and so is the mixing time, and there is no cutoff.

See, for example, [41, 55, 101] for the conjectures.

Literature. The Glauber dynamics is arguably the most studied interacting particle system in the literature. Tons of work have been devoted to verifying the conjectured behaviors above for it; see Subsection 1.1 in [65] for a brief review of the development of the subject. The readers can look at [1, 28, 42, 43, 44, 63, 68, 71, 73, 72, 90, 91, 92, 106, 107] for results in the high-temperature regime, at [25, 26, 65, 80] for results in the critical-temperature regime, and at [14, 45, 71, 88, 95] for results in the low-temperature regime. We especially mention here the work [68], which proves cutoff in the full high-temperature regime in any dimension. The proof in [68] involves a new framework called *information percolation*, see also [33, 67, 69, 70]. For background on Glauber dynamics, see the lecture notes [74], or Chapter 15 in [59] for a quick introduction. In particular, Theorem 15.4 in [59] presents a very simple example that illustrates the difference in mixing behaviors at low and high temperatures.

As mentioned in Chapter 2, the Exclusion Process is also an emblematic interacting particle system. It describes the relaxation of a gas of interacting particles to equilibrium. The convergence is usually studied from two points of view: the macroscopic evolution of the density of particles, which is the study of hydrodynamic limits, or the microscopic evolution of the law of the particles and its total variation distance to equilibrium, which is the study of mixing times. The readers can see [47] for an introduction to the study of hydrodynamic limits and [59] for an introduction to the study of mixing times. For the results on the hydrodynamic limit of SSEP, see, for example, [5, 36, 37, 58, 46]. For the results on the mixing times, see [24, 50, 51, 52, 53, 77, 79, 102].

Cutoff for SSEP has been proved in relatively few cases: for the complete graph in [53], for dimension one in [50, 51, 52], see also [37, 99] for nonconservative variants. The only result available in general geometries for a nonconservative variant is obtained in [87]. However, it is restricted to the case where the system is reversible w.r.t a product of i.i.d. Bernoulli.

As mentioned before, the Glauber-Exclusion process was introduced in [18] and also considered in [46]. To the best of our knowledge, the only works on its mixing times, also the most related works to our work, are [93] and [100]. In [100], Tsunoda studies the model in dimension one and proves exponentially slow mixing when the primitive of the reaction function (the potential function) has two or more local minima, which implies that the system is at low temperature. In [93], Tanaka and Tsunoda prove the upper bound $\mathcal{O}(\log L)$ on the mixing times for any dimension under the condition that $\max_{\rho \in [-1,1]} R'(\rho) < 0$, which subsequently implies that the system is at high temperature. These two works together show a phase transition of the mixing times for the model in Example 3.1 between low and high temperatures, consistent with the conjectured behaviors. In particular, we expect cutoff to occur in the high-temperature regime. However, to establish cutoff, not only do we have to prove that the upper and lower bounds are of the same magnitude $\log L$, but we also need to show that the pre-factors in front of $\log L$ in the two bounds are the same. Surprisingly, not even a lower bound of matching magnitude $\log L$ is available, even though it is often "the easy part" in proving cutoff. It was not yet known whether the upper bound in [93] is sharp either.

Our contribution. We sharpen the upper bound in [93] and prove a matching lower bound, therefore showing cutoff for the full high-temperature regime for *any* attractive rate function c, in dimensions 1 and 2. Furthermore, for any dimension, in the full high-temperature regime, we prove that the inverse gap is of order 1 and provide lower and upper bounds of magnitude $\log L$ on the mixing times. Our proof involves a new interpretation of the attractiveness, the information percolation framework introduced by Lubetzky and Sly in [68], anti-concentration of simple random walk on the lattice, and what we call *the excursion coupling.* We hope that this approach can find new applications in the future.

2 Results

The first main result of the chapter is the following.

Theorem 3.2 (Cutoff in dimensions 1 and 2)**.** *Let $\epsilon \in (0,1)$ be fixed. For $d \in \{1,2\}$, there exists a constant κ such that*

$$\frac{\log|\Lambda_L^d|}{2|R'(\rho_*)|} - \kappa \log\log L \leq \mathrm{t_{mix}}(\epsilon) \leq \frac{\log|\Lambda_L^d|}{2|R'(\rho_*)|} + \kappa \log\log L. \tag{3.10}$$

Remark 3. Here, cutoff is proved when the dynamics is generally not reversible, and no explicit formula for the invariant measure π is available. In fact, in [32], Gabrielli *et al* showed that the system is reversible if and only if the function $c(\cdot)$ is of the following form

$$c(x) = (a_1 + a_2 x(0))h(x), \tag{3.11}$$

where $h(x)$ is independent of $x(0)$, and in this case π is a product of i.i.d. Rademacher. To the best of our knowledge, it is also the only case that an explicit formula for π is available. However, one can show that if a function $c(\cdot)$ satisfies the reversibility condition (3.11), then it is attractive if and only if it is a constant function.

Our second main result is about the spectral gap of the system.

Theorem 3.3 (Spectral gap)**.** *For any dimension d, there exists a constant κ such that*

$$|R'(\rho_*)| - \frac{\kappa}{\log L} \leq \mathtt{gap} \leq |R'(\rho_*)| + \frac{\kappa}{\log L}. \tag{3.12}$$

Theorem 3.2 and Theorem 3.3 together confirm the conjectured behaviors of the system in the full high-temperature regime in dimensions 1 and 2. Furthermore, they connect the mixing behaviors of the Glauber-Exclusion process with the behaviors of the solutions of the ODE (3.9).

Our third result investigates the mixing times in higher dimensions.

Theorem 3.4 (Pre-cutoff in higher dimensions)**.** *Let $\epsilon \in (0,1)$ be fixed. For any $d \geq 3$, there exists a constant κ such that*

$$\frac{1}{d}\frac{\log|\Lambda_L^d|}{|R'(\rho_*)|} - \kappa \leq \mathrm{t_{mix}}(\epsilon) \leq \frac{\log|\Lambda_L^d|}{|R'(\rho_*)|} + \kappa\sqrt{\log L}. \tag{3.13}$$

Theorem 3.4 says that the magnitude of the mixing times in all dimensions is $\log L$. Furthermore, the pre-factor in front of $\log L$ is bounded between two constants independent of the precision ϵ. This phenomenon is known as *pre-cutoff*; see Chapter 18 in [59] for more details. Establishing cutoff for the process in dimensions $d \geq 3$ remains an open problem. We believe that neither the lower nor the upper bounds in Theorem 3.4 is optimal. Furthermore, our proof of the upper bound in dimensions 1 and 2 relies on the fact that the simple random walks in dimensions 1 or 2 are recurrent, which is no longer valid for higher dimensions. We do not know if this is just a coincidence. The Interchange Process, a Markov process closely related to the Exclusion process, is conjectured to behave somehow differently in dimensions $d \geq 3$ compared to $d \in \{1,2\}$ in [96]. This fact is verified for all dimensions $d \geq 5$ in [30]. We do not know whether this phase transition in infinite volume is connected to a change of mixing behavior in finite volume.

Nevertheless, the upper bound in Theorem 3.4 is already better than that in [93]. In [93], the authors impose the assumption that the derivative R' of R is uniformly negative, i.e. $\max_{\rho\in[-1,1]} R'(\rho) < 0$, and prove the upper bound $\frac{\log|\Lambda_L^d|}{\left|\max_{\rho\in[-1,1]} R'(\rho)\right|}$ on the mixing time. Hence, our assumption is weaker, and the pre-factor in front of $\log L$ is sharper.

Example 3.5. *Let θ be a strictly positive number. In dimension one, we consider the local flip-rate function $c(\cdot) : \{-1,1\}^{\{-1,0,1\}} \to \mathbb{R}_+$ defined by*

$$\forall x \in \{-1,1\}^{\{-1,0,1\}},\ c(x) := \theta + 2 \times \mathbb{1}_{\{x(0)=1,\, x(1)=-1\}}. \tag{3.14}$$

The corresponding global flip-rate function $\hat{c}$ is given by

$$\hat{c}(u,x) := \theta + 2 \times \mathbb{1}_{\{x(u)=1,\, x(u+1)=-1\}}. \tag{3.15}$$

It is not hard to see that the function $c(\cdot)$ satisfies Hypothesis 1. The corresponding reaction function is

$$R(\rho) = \rho^2 - 2\theta\rho - 1. \tag{3.16}$$

By direct computation, one can show that $R(\rho)$ has a unique root in the interval $[-1,1]$:

$$\rho_* = \theta - \sqrt{\theta^2+1},$$

and moreover,

$$R'(\rho_*) < 0.$$

However, if $0 < \theta < 1$, then $\forall \rho \in (\theta, 1]$, $R'(\rho) > 0$. This means that $R(\rho)$ satisfies our conditions but does not satisfy the condition in [93].

Structure of the proof. In Section 3, we interpret the attractiveness in a more intuitive way. We also discuss some properties of monotone boolean functions, which will be used later in our proof. In Section 4, we construct the dual coupling, which will subsequently be used in Section 5 to prove Theorem 3.3, Theorem 3.4, and the lower bound in Theorem 3.2. Finally, in Section 6, we prove the upper bound in Theorem 3.2. For simplicity, we will assume that $d = 1$. Nevertheless, many ingredients are valid for other dimensions, and we will comment on how to adapt the proof to higher dimensions when necessary.

3 Monotone boolean functions and attractiveness

This section aims to discuss some properties of monotone boolean functions and to prove Proposition 3.9, which allows us to interpret the attractiveness more intuitively.

3.1 Monotone boolean functions

Definition 3.6 (Boolean function). A *boolean function* is a function from some hypercube $\{-1,1\}^n$, $n \in \mathbb{Z}_+$, into $\{-1,1\}$. A boolean function f is increasing if $x \le y \Rightarrow f(x) \le f(y)$, for any $x, y \in \{-1,1\}^n$.

Definition 3.7 (Pivotal set). For an increasing boolean function f on $\{-1,1\}^n$, its *pivotal set* is given by

$$\mathsf{Piv}\,(f) := \{j \in [n] : \exists x = (x_1, \dots, x_n) \in \{-1,1\}^n \text{ such that } f(x^{j,1}) = 1;\ f(x^{j,-1}) = -1\}.$$

In particular, to evaluate the value of an increasing boolean function f on a spin configuration x, we do not need to know every coordinate of x but only the coordinates in $\mathsf{Piv}\,(f)$. The following lemma is elementary but very useful in our proof.

Lemma 3.8 (Pivotal sets of monotone boolean functions)**.** *For any increasing boolean function f on $\{-1,1\}^n$,*

$$\mathsf{Piv}\,(f) \neq \emptyset \iff f(1,\dots,1) = 1 \text{ and } f(-1,\dots,-1) = -1.$$

In particular, for any increasing boolean function f on $\{-1,1\}$,

$$|\mathsf{Piv}\,(f)\,| = \frac{f(1) - f(-1)}{2}.$$

3.2 Interpretation of attractiveness

We canonically identify the functions on $\mathcal{X}$ which depend only on the coordinates in $B(0,m)$ with the functions on $\{-1,1\}^{B(0,m)}$. We now give a more intuitive interpretation of the attractiveness.

Proposition 3.9 (Interpretation of attractiveness)**.** *There exist $q \in \mathbb{Z}_+$ and increasing boolean functions $f_1, \dots, f_q$ on $\{-1,1\}^{B(0,m)}$ and positive numbers $\lambda_1, \dots, \lambda_q$, such that*

$$\forall x \in \{-1,1\}^{B(0,m)},\ c(x) = \sum_{i=1}^{q} \lambda_i \mathbb{1}_{\{f_i(x) = -x(0)\}}. \tag{3.17}$$

In particular, for any function φ on $\mathcal{X}$,

$$\forall x \in \mathcal{X},\ \mathcal{L}_G\varphi(x) = \sum_{u \in \Lambda_L} \sum_{i=1}^{q} \lambda_i \left(\varphi(x^{u, f_i(x_{u+\cdot})}) - \varphi(x)\right). \tag{3.18}$$

Moreover, we can choose $f_1 \equiv 1,\ f_2 \equiv -1$ and λ_1, λ_2 strictly positive.

Interpretation. Proposition 3.9 is of independent interest. It implies that the technical conditions (3.5), (3.6) for the attractiveness of the Glauber dynamics can be replaced by a more intuitive interpretation: for each site u, $x(u)$ is replaced by $f_i(x_{u+\cdot})$ at rate λ_i, $1 \le i \le q$. The monotonicity of $(f_i)_{1 \le i \le q}$ clearly implies that the system is attractive. In fact, for two processes

$(Y_t^1)_{t\geq 0}$ and $(Y_t^2)_{t\geq 0}$ whose generator is $\mathcal{L}_G$, one can construct a coupling of them by updating the same sites simultaneously using the same functions. This coupling preserves the order thanks to the monotonicity of $(f_i)_{1\leq i\leq q}$. Besides, the updates given by deterministic functions $f_1 \equiv 1$ or $f_2 \equiv -1$ can be carried out without looking at the spins of any sites. We say that these updates are *oblivious*. We will see that oblivious updates play a crucial role in our proof.

Illustration by Example 3.1. For the flip-rate function considered in Example 3.1, one can verify that the following functions and rates satisfy Proposition 3.9. For any $x = (x(-1), x(0), x(1)) \in \{-1,1\}^{\{-1,0,1\}}$,

$$\begin{aligned}
f_1(x) &= 1, & \lambda_1 &= (1-\gamma)^2,\\
f_2(x) &= -1, & \lambda_2 &= (1-\gamma)^2,\\
f_3(x) &= \operatorname{sgn}(x(-1)+x(0)+x(1)), & \lambda_3 &= 4\gamma^2,\\
f_4(x) &= x(-1), & \lambda_4 &= 2(\gamma-\gamma^2),\\
f_5(x) &= x(1), & \lambda_5 &= 2(\gamma-\gamma^2).
\end{aligned}$$

Here sgn denotes the sign function: $\operatorname{sgn}(s) = \begin{cases} 1 & \text{if } s>0,\\ 0 & \text{if } s=0,\\ -1 & \text{if } s<0.\end{cases}$

We first need the following lemma to prove Proposition 3.9.

Lemma 3.10 (Decomposition of monotone functions)**.** *Let Ω be a finite set with a partial order $\prec$. Let $f:\Omega\to\mathbb{R}_+$ be an increasing function, meaning that $x\prec y \Rightarrow f(x)\leq f(y)$. Then there exist $q\in\mathbb{Z}_+$ and increasing (not boolean) functions $f_1,\dots,f_q:\Omega\to\{0,1\}$ and positive numbers $\lambda_1,\dots,\lambda_q$ such that*

$$f(\cdot) = \sum_{i=1}^{q} \lambda_i f_i(\cdot). \tag{3.19}$$

Moreover, we can choose $f_1(\cdot) = \mathbb{1}_{\{f(\cdot)>0\}}$ and $\lambda_1 = \mathbb{1}_{\{f\not\equiv 0\}} \min_{\{f(\cdot)>0\}} f$.

Lemma 3.10 can be proved by induction on the size of the support of the function f. We omit the proof here. Now we prove Proposition 3.9.

Proof of Proposition 3.9. First we prove that (3.17) implies (3.18). Indeed, if (3.17) is true, then by definition of $\hat{c}(\cdot)$, for any u in Λ_L and $x\in\mathcal{X}$,

$$\hat{c}(u,x) = \sum_{i=1}^{q} \lambda_i \mathbb{1}_{\{f_i(x_{u+\cdot})=-x(u)\}}. \tag{3.20}$$

Therefore, for any function φ on $\mathcal{X}$,

$$\begin{aligned}\mathcal{L}_G\varphi(x) &= \sum_{u\in\Lambda_L} \hat{c}(u,x)(\varphi(x^{u,-x(u)}) - \varphi(x)) \\ &= \sum_{u\in\Lambda_L}\sum_{i=1}^{q} \lambda_i \mathbb{1}_{\{f_i(x_{u+\cdot})=-x(u)\}}(\varphi(x^{u,-x(u)}) - \varphi(x)) \\ &= \sum_{u\in\Lambda_L}\sum_{i=1}^{q} \lambda_i \mathbb{1}_{\{f_i(x_{u+\cdot})=-x(u)\}}(\varphi(x^{u,f_i(x_{u+\cdot})}) - \varphi(x)) \\ &= \sum_{u\in\Lambda_L}\sum_{i=1}^{q} \lambda_i(\varphi(x^{u,f_i(x_{u+\cdot})}) - \varphi(x)),\end{aligned}$$

where the second equality is due to (3.20), the third equality is because $x^{u,f_i(x_{u+\cdot})} = x^{u,-x(u)}$ whenever $f_i(x_{u+\cdot}) = -x(u)$, and the last equality is because $\varphi(x^{u,f_i(x_{u+\cdot})}) - \varphi(x) = 0$ whenever $f_i(x_{u+\cdot}) = x(u)$. It remains to show (3.17).

Let $n := |B(0,m)| - 1$. We identify the set $B(0,m)$ with $\{0,1,\dots,n\}$ such that the site 0 is identified with number 0. Then (3.17) is reformulated as follows: there exist increasing boolean functions $\tilde{f}_1,\dots,\tilde{f}_q$ on $\{-1,1\}^{n+1}$ and positive numbers $\lambda_1,\dots,\lambda_q$ such that, for any $\xi_0,\dots,\xi_n \in \{-1,1\}$,

$$c(\xi_0,\dots,\xi_n) = \sum_{i=1}^{q} \lambda_i \mathbb{1}_{\{\tilde{f}_i(\xi_0,\dots,\xi_n)=-\xi_0\}}. \tag{3.21}$$

Note that Hypothesis 1 implies that $c(-1,\xi_1,\dots,\xi_n)$ is increasing and $c(1,\xi_1,\dots,\xi_n)$ is decreasing as functions of $(\xi_1,\dots,\xi_n)$. We can now apply Lemma 3.10 to the function $c(-1,\cdot)$ to conclude that there exist increasing functions $f_1,\dots,f_{q_1}$ from $\{-1,1\}^n$ to $\{0,1\}$ and positive numbers $\lambda_1,\dots,\lambda_{q_1}$, for some $q_1 \in \mathbb{Z}_+$, such that, for any $\xi_1,\dots,\xi_n \in \{-1,1\}$,

$$c(-1,\xi_1,\dots,\xi_n) = \sum_{i=1}^{q_1} \lambda_i f_i(\xi_1,\dots,\xi_n). \tag{3.22}$$

Moreover, we can choose

$$\begin{aligned}f_1(\xi_1,\dots,\xi_n) &= \mathbb{1}_{\{c(-1,\xi_1,\dots,\xi_n)>0\}} \equiv 1, \\ \lambda_1 &= \min_{\xi_1,\dots,\xi_n\in\{-1,1\}} c(-1,\xi_1,\dots,\xi_n) > 0.\end{aligned}$$

Here, we have used the fact that $c(\cdot)$ only takes strictly positive values. We then define the boolean functions $(\tilde{f}_i)_{1\le i\le q_1}$ on $\{-1,1\}^{n+1}$ by

$$\tilde{f}_i(\xi_0,\dots\xi_n) = \begin{cases} 1 & \text{if } \xi_0 = 1, \\ 2f_i(\xi_1,\dots,\xi_n) - 1 & \text{if } \xi_0 = -1.\end{cases}$$

It is easy to see that $\tilde{f}_i$ is increasing for any i. Note also that $\tilde{f}_1 \equiv 1$.

Let $\varphi_1 : \{-1,1\}^{n+1} \to \mathbb{R}_+$ be defined by, for any $\xi_0, \dots, \xi_n \in \{-1,1\}$,

$$\varphi_1(\xi_0, \dots, \xi_n) = \sum_{i=1}^{q_1} \lambda_i \mathbb{1}_{\{\tilde{f}_i(\xi_0,\dots,\xi_n)=-\xi_0\}}.$$

Noting that

$$\begin{aligned}\mathbb{1}_{\{\tilde{f}_i(1,\xi_1\dots,\xi_n)=-1\}} &= 0,\\ \mathbb{1}_{\{\tilde{f}_i(-1,\xi_1,\dots,\xi_n)=1\}} &= \mathbb{1}_{\{f_i(\xi_1,\dots,\xi_n)=1\}} = f_i(\xi_1,\dots,\xi_n),\end{aligned}$$

we conclude that

$$\begin{aligned}\varphi_1(1, \xi_1, \dots, \xi_n) &= 0,\\ \varphi_1(-1, \xi_1, \dots, \xi_n) &= c(-1, \xi_1, \dots, \xi_n).\end{aligned}$$

Similarly $c(1, \xi_1, \dots, \xi_n)$ is a positive decreasing function of $(\xi_1, \dots, \xi_n)$. We can still apply Lemma 3.10, by replacing the relation $\leq$ by $\geq$, to conclude that there exist decreasing functions (increasing functions w.r.t relation $\geq$) $g_1, \dots, g_{q_2}$ and positive numbers $\lambda'_1, \dots, \lambda'_{q_2}$ such that

$$c(1, \xi_1, \dots, \xi_n) = \sum_{i=1}^{q_2} \lambda'_i g_i(\xi_1, \dots, \xi_n). \tag{3.23}$$

By a similar argument, we claim that the boolean functions $(\tilde{f}_{q_1+i})_{1\leq i\leq q_2}$ defined by

$$\tilde{f}_{q_1+i}(\xi_0, \dots \xi_n) = \begin{cases} 1 - 2g_i(\xi_1, \dots, \xi_n) & \text{if } \xi_0 = 1,\\ -1 & \text{if } \xi_0 = -1,\end{cases}$$

are increasing. Similarly, we can choose $\lambda'_1 > 0$ and $g_1 \equiv 1$. Hence $\tilde{f}_{q_1+1} \equiv -1$. Similarly, we define $\varphi_2 : \{-1,1\}^{n+1} \to \mathbb{R}_+$ by, for all $\xi_0, \dots, \xi_n \in \{-1,1\}$,

$$\varphi_2(\xi_0, \dots, \xi_n) = \sum_{i=1}^{q_2} \lambda'_i \mathbb{1}_{\{\tilde{f}_{q_1+i}(\xi_0,\dots,\xi_n)=-\xi_0\}}.$$

φ_2 satisfies

$$\begin{aligned}\varphi_2(1, \xi_1, \dots, \xi_n) &= c(1, \xi_1, \dots, \xi_n),\\ \varphi_2(-1, \xi_1, \dots, \xi_n) &= 0.\end{aligned}$$

Hence

$$c = \varphi_1 + \varphi_2.$$

We can set $\lambda_{q+i} = \lambda'_i$, $1 \leq i \leq q_2$ to write

$$c(\xi_0, \dots, \xi_n) = \sum_{i=1}^{q_1+q_2} \lambda_i \mathbb{1}_{\left\{\tilde{f}_i(\xi_0,\dots,\xi_n)=-\xi_0\right\}}. \tag{3.24}$$

We can renumber these functions and their coefficients such that $\tilde{f}_1 \equiv 1$, $\tilde{f}_2 \equiv -1$. So we have established (3.21). This finishes our proof. □

Remark 4. We can adapt this proposition to *any* flip rate function satisfying (3.5), (3.6), on any graph. In this case, each site u has its own update functions $f_{u,1}, \dots f_{u,q_u}$, which are increasing and applied at rate $\lambda_{u,1}, \dots, \lambda_{u,q_u}$ respectively. In our case, those functions are local and invariant by translation because the function $\hat{c}$ is.

From now on, we fix a choice of $(f_i)_{1\leq i\leq q}$ and $(\lambda_i)_{1\leq i\leq q}$ that satisfies Proposition 3.9. For our purpose, we will rewrite the reaction function R and its derivative R' in terms of $(f_i)_{1\leq i\leq q}$ and $(\lambda_i)_{1\leq i\leq q}$. To do this, let $\lambda := \sum_{i=1}^{q} \lambda_i$. For any $1 \leq i \leq q$, $j \in B(0, m)$, $\xi_{-m}, \dots, \xi_m \in \{-1, 1\}$, let us define

$$\begin{aligned}&\nabla_j f_i(\xi_{-m}, \dots, \xi_m)\\ &:= f_i(\xi_{-m}, \dots, \xi_{j-1}, 1, \xi_{j+1}, \dots, \xi_m) - f_i(\xi_{-m}, \dots, \xi_{j-1}, -1, \xi_{j+1}, \dots, \xi_m).\end{aligned}$$

Lemma 3.11 (R and R' in terms of $(f_i)_{1\leq i\leq q}$ and $(\lambda_i)_{1\leq i\leq q}$)**.** *For any* $\rho \in [-1, 1]$,

$$R(\rho) = \left(\sum_{i=1}^{q} \lambda_i \mathbb{E}_{\nu_\rho}\left[f_i\right]\right) - \lambda\rho, \tag{3.25}$$

$$R'(\rho) = \left(\sum_{i=1}^{q} \sum_{j\in B(0,m)} \lambda_i \mathbb{E}_{\nu_\rho}\left[\frac{1}{2}\nabla_j f_i\right]\right) - \lambda. \tag{3.26}$$

Proof. Recall that $R(\rho) = \mathbb{E}_{\nu_\rho}\left[-2\xi_0 c(\xi)\right]$. Hence, by (3.17).

$$\begin{aligned}
R(\rho) &= \mathbb{E}_{\nu_\rho}\left[-2\xi_0 \sum_{i=1}^{q} \lambda_i \mathbb{1}_{\{f_i(\xi)=-\xi_0\}}\right]\\
&= \sum_{i=1}^{q} \lambda_i \mathbb{E}_{\nu_\rho}\left[-2\xi_0 \mathbb{1}_{\{f_i(\xi)=-\xi_0\}}\right]\\
&= \sum_{i=1}^{q} \lambda_i \mathbb{E}_{\nu_\rho}\left[(f_i(\xi) - \xi_0)\mathbb{1}_{\{f_i(\xi)=-\xi_0\}}\right]\\
&= \sum_{i=1}^{q} \lambda_i \mathbb{E}_{\nu_\rho}\left[f_i(\xi) - \xi_0\right]\\
&= \left(\sum_{i=1}^{q} \lambda_i \mathbb{E}_{\nu_\rho}\left[f_i(\xi)\right]\right) - \lambda\rho.
\end{aligned}$$

So we have established (3.25). We also know that, for any $\rho \in [-1, 1]$,

$$\mathfrak{R}_\rho = \frac{1+\rho}{2}\delta_1 + \frac{1-\rho}{2}\delta_{-1},$$

and hence

$$\frac{d}{d\rho}\mathfrak{R}_\rho = \frac{1}{2}\delta_1 - \frac{1}{2}\delta_{-1}.$$

Therefore, for $\nu_\rho = \mathfrak{R}_\rho^{\otimes B(0,m)}$,

$$\frac{d}{d\rho}\nu_\rho = \sum_{j \in B(0,m)} \mathfrak{R}_\rho^{\otimes\{-m,\ldots,j-1\}} \otimes \left(\frac{1}{2}\delta_1 - \frac{1}{2}\delta_{-1}\right) \otimes \mathfrak{R}_\rho^{\otimes\{j+1,\ldots,m\}}.$$

This equality and (3.25) lead to (3.26). □

4 Dual coupling

4.1 Graphical Construction

We call $\Lambda_L \times \mathbb{R}_+$ the *space-time slab*, where each slice $\Lambda_L \times \{t\}$ is equipped with the graph structure of Λ_L. We identify Λ_L with the subset $\{0, \ldots, L-1\}$ of $\mathbb{Z}_+$.

We introduce a notation essential to our proof.

A collection of marks. A collection of marks $\mathcal{C}$ is a pair $(\mathcal{C}^{\text{exclusion}}, \mathcal{C}^{\text{glauber}})$, where $\mathcal{C}^{\text{exclusion}}$ and $\mathcal{C}^{\text{glauber}}$ are subsets of $\Lambda_L \times \mathbb{R}_+$ and $[q] \times \Lambda_L \times \mathbb{R}_+$, respectively, that satisfy the following conditions.

- There is at most one mark at a time:

$$\forall t > 0, \left|\mathcal{C}^{\text{exclusion}} \cap \Lambda_L \times \{t\}\right| + \left|\mathcal{C}^{\text{glauber}} \cap [q] \times \Lambda_L \times \{t\}\right| \leq 1.$$

- $\mathcal{C}$ is locally finite in time: for any $0 < t_1 < t_2$,

$$\left|\mathcal{C}^{\text{exclusion}} \cap \Lambda_L \times [t_1, t_2]\right| + \left|\mathcal{C}^{\text{glauber}} \cap [q] \times \Lambda_L \times [t_1, t_2]\right| < \infty.$$

In other words, $\mathcal{C}$ can be identified with a counting measure on $(\Lambda_L \times \mathbb{R}_+) \cup ([q] \times \Lambda_L \times \mathbb{R}_+)$ whose projection onto $\mathbb{R}_+$ has no multiple point and is locally finite. We call an element $(u, t) \in \mathcal{C}^{\text{exclusion}}$ the exclusion mark on the edge $(u, u+1)$ at time t and an element $(i, u, t) \in \mathcal{C}^{\text{glauber}}$ the Glauber mark of type i on site u at time t.

Effect of a collection of marks. Given a collection of marks $\mathcal{C}$ and a configuration $x_0 \in \mathcal{X}$, the process $(X_t^{x_0})_{t \geq 0}$ is defined recursively according to the following rules.

- $X_0^{x_0} = x_0$ (x_0 is the initial configuration).
- The process $(X_t^{x_0})_{t \geq 0}$ is piecewise constant and can only jump when a mark appears.

- When we see an exclusion mark on the edge $(u, u+1)$, we make the transition $x \mapsto x^{u \leftrightarrow u+1}$ (exchange the spins at sites u and $u+1$).
- When we see a Glauber mark of type i at site u, we make the transition $x \mapsto x^{u, f_i(x_{u+\cdot})}$ (update site u using the function f_i).

Roughly speaking, the marks tell us how the process $(X_t^{x_0})_{t\geq 0}$ is updated. We have kept the dependency on $\mathcal{C}$ implicit to lighten the notation. The Glauber-Exclusion process associated with the generator $\mathcal{L}_{GE}$ can be constructed as follows.

Graphical Construction 1. Let the background process

$$\Xi = \left((\Xi_u^{\text{exclusion}})_{0\leq u\leq L-1}, (\Xi_{i,u}^{\text{Glauber}})_{1\leq i\leq q,\, 0\leq u\leq L-1}\right) \tag{3.27}$$

be defined as follows. $(\Xi_u^{\text{exclusion}})_{0\leq u\leq L-1}, (\Xi_{i,u}^{\text{Glauber}})_{1\leq i\leq q,\, 0\leq u\leq L-1}$ are independent homogeneous Poisson processes, and

- $\Xi_u^{\text{exclusion}}$ is of intensity L^2, $0 \leq u \leq L-1$,
- $\Xi_{i,u}^{\text{Glauber}}$ is of intensity λ_i, $1 \leq i \leq q$, $0 \leq u \leq L-1$.

Almost surely, a realization of Ξ naturally defines a collection of marks $\mathcal{C}$ as follows. For any $(i, u, t) \in [q] \times \Lambda_L \times \mathbb{R}_+$,

- $(u, t) \in \mathcal{C}^{\text{exclusion}}$ iff the process $\Xi_u^{\text{exclusion}}$ jumps at time t,
- $(i, u, t) \in \mathcal{C}^{\text{glauber}}$ iff the process $\Xi_{i,u}^{\text{Glauber}}$ jumps at time t.

Then the process $(X_t^{x_0})_{t\geq 0}$ constructed by the rules above is a Markov process with generator $\mathcal{L}^{GE}$ starting from configuration x_0. An illustration is given in Figure 3.2, where each exclusion mark is naturally drawn as a left-right arrow on the corresponding edge, and each Glauber mark is drawn as $\times$ on the corresponding site, with the corresponding update function next to it.

Remark 5. In the future, we may construct the (random) collection of marks $\mathcal{C}$ differently, but as long as it has the same distribution as the one given by Ξ, the process $(X_t^{x_0})_{t\geq 0}$ constructed by the above rules is still a Markov process associated with generator $\mathcal{L}^{GE}$ starting at x_0.

The alphabet. Let $\mathcal{A} := \bigcup_{i=0}^{\infty} (\mathbb{Z}_+ \times B(0, m)^i)$ be the set of all words beginning with an element in $\mathbb{Z}_+$ followed by elements in $B(0, m)$. We define a lexicographical order on $\mathcal{A}$ using the usual orders on $\mathbb{Z}_+$ and $B(0, m)$. For any word $\mathsf{w} \in \mathcal{A}$, we call the elements of $\{\mathsf{w}\} \times B(0, m)$ the children of w. The alphabet $\mathcal{A}$ will be used to label the branching processes that naturally arise when we study the *update history* of our process.

Update history. Consider a collection of marks $\mathcal{C}$ and the resulting process $(X_t^x)_{t\geq 0}$ for some $x \in \mathcal{X}$. Let $t \geq 0$, and let $E \subset \Lambda_L$ be a set whose spins we want to know at time t. A natural way to reveal the spins of E at time t is to go backward in time, tracing the history of the information involved. At each site of E at time t, we put a particle labeled by the corresponding element in $\mathcal{A}$ (we have $E \subset \Lambda_L \subset \mathbb{Z}_+ \subset \mathcal{A}$). The history of E, going backward in time, is as follows.

1. Every time we see an exclusion mark, the particles at the endpoints of the corresponding edge, if there are any, jump to the other endpoints.

2. Every time we see a Glauber mark, say at site u, if there is a particle w there, we remove w, and at each site of $B(u, m)$, we add a new particle. We canonically label the newborn particles by the children of w. If a particle is born on a site already occupied by some other particles, all of them will move together as one particle.

We follow all the particles backward until time 0. The *update history* of the set E at time t is the sub-collection of marks $\mathcal{C}_{his}(E)$ consisting of all the marks (exclusion and Glauber) we meet on the trajectories of all the particles described above. For our purpose, it is more convenient to consider $\mathcal{C}_{his}^{bw}(E)$, the version of $\mathcal{C}_{his}(E)$ when we reverse the time interval $[0, t]$: $\mathcal{C}_{his}^{bw}(E)$ is the set of marks obtained from $\mathcal{C}_{his}(E)$ by replacing the time s in each mark by $t - s$.

We define a spin configuration $\mathfrak{s}$ on the set of all particles appearing as follows.

- For any particle w reaching time 0, say at site u,

$$\mathfrak{s}_{\mathsf{w}} := x(u).$$

- For any particle w not reaching time 0, let i be the type of the Glauber mark removing w. Then

$$\mathfrak{s}_{\mathsf{w}} := f_i((\mathfrak{s}_{\mathsf{w}'})_{\mathsf{w}' \in \{\mathsf{w}\} \times B(0,m)}).$$

These rules allow us to define recursively the spins of all particles that appear. In particular, by construction,

$$\forall u \in E, \, \mathfrak{s}_u = X_t^x(u),$$

where $\mathfrak{s}_u$ is the spin of the particle labeled u.

4.2 The branching exclusion process (BEP)

We fix a subset E of Λ_L in this subsection. We claim that the history itself, viewed backward in time, is a Markov process, called the *branching exclusion process* (BEP), defined as follows.

Definition of the BEP. We consider the following way of putting marks and particles in the space-time slab. We go forward in time.

1. At time 0, at each site of E, we put a particle labeled by the corresponding element of $\mathcal{A}$. Each particle is associated with q Poisson clocks, called the Glauber clocks of type $1, \ldots, q$, of intensity $\lambda_1, \ldots, \lambda_q$, respectively. All the Glauber clocks are independent.

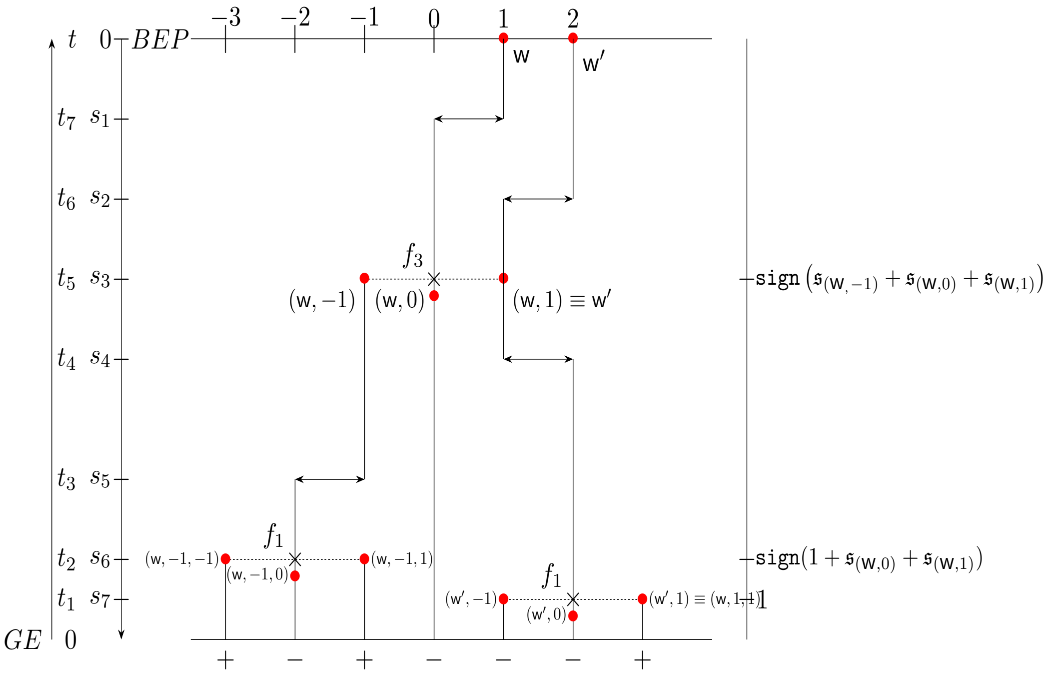


Figure 3.2: History of Glauber-Exclusion dynamics.

2. Each edge of the lattice is associated with a Poisson clock of intensity L^2, called the exclusion clock. The exclusion clocks are independent and independent of the Glauber clocks. Each time an exclusion clock rings, we put an exclusion mark on the corresponding edge.

3. The particles follow the exclusion marks until one Glauber clock rings, says the clock of type i of particle w at time s. Suppose that w is at site u at time $s-$. Then we put a Glauber mark of type i on site u at time s. After that, we remove particle w and add new particles to all sites in $B(u, m)$, canonically labeled by the children of w. We think of this as the particle w splits into its neighbor sites. Each particle born on a non-occupied site is given q independent Glauber clocks, independent of all exclusion clocks and all Glauber clocks of other particles. Each particle born on an occupied site "moves together" with the particles already there. The particles moving together are called a group of particles, and they are associated with the same q Glauber clocks. All the particles then continue to evolve with similar rules.

The BEP starting from E consists of two parts $(\Psi_{BEP}(E), \mathcal{C}_{BEP}(E))$, where

- $\Psi_{BEP}(E)$ is the collection of trajectories of all the particles, i.e. the map associating each point (u, t) of the space-time slab with the set of particles occupying the site u at time t.
- $\mathcal{C}_{BEP}(E)$ is the set consisting of all the marks on the trajectories of all the particles constructed above.

Remark 6. Knowing E, the pair $(\Psi_{BEP}(E), \mathcal{C}_{BEP}(E))$ is determined by $\mathcal{C}_{BEP}(E)$.

We say that a particle w is *alive* at time t if it was born before time t and has not been removed by time t. Let W_t be the set consisting of the particles with the smallest labels in each group of particles alive at time t, for any $t \in \mathbb{R}_+$. A spin configuration on W_t is identified with a spin configuration on the set of all particles alive at time t, where the particles in a group are given the same spin.

Spins of particles in BEP. Each spin configuration $\mathfrak{s}$ on W_t is extended to $\bigcup_{0\le s\le t} W_s$ by the following rules. For any w in $\left(\bigcup_{0\le s\le t} W_s\right) \setminus W_t$, w must be removed at some point in $[0,t]$ by a Glauber mark, say of type i. Then

$$\mathfrak{s}_{\mathsf{w}} := f_i\left((\mathfrak{s}_{\mathsf{w}'})_{\mathsf{w}'\in\{\mathsf{w}\}\times B(0,m)}\right).$$

This allows us to extend the spin configuration to all the particles alive at some point in the time interval $[0,t]$.

By the homogeneity of the Poisson processes, we can see that $\mathcal{C}_{BEP}(E)$ and $\mathcal{C}^{bw}_{his}(E)$ have the same distribution. In fact, we can show that their first marks have the same distribution, then conditionally on their first marks, their second marks have the same distribution, etc. This leads to the following result.

Proposition 3.12 (History is BEP)**.** *Let $t \in \mathbb{R}_+$, $x \in \mathcal{X}$. Consider the BEP starting from E and the history of E at time t given by a collection of marks generated by Ξ. Then the point processes $\mathcal{C}_{BEP}(E)$ and $\mathcal{C}^{bw}_{his}(E)$ have the same distribution. In particular, for some $x \in \mathcal{X}$, if we define the spin configuration $\mathfrak{s}$ on W_t by*

$$\forall \mathsf{w} \in W_t,\ \mathfrak{s}_{\mathsf{w}} := x(u),\ \textit{if } \mathsf{w} \textit{ is at site } u \textit{ at time } t,$$

then

$$X^x_t(E) \overset{d}{=} (\mathfrak{s}_u)_{u\in E}. \tag{3.28}$$

Example 3.13. *Figure 3.2 also illustrates the BEP starting from $E = \{1,2\}$. The time axis upward is the one of the Glauber-Exclusion process, and the one downward is the one of the BEP.*

The idea of looking at the history backward in time is not new. It is known as *duality* in the context of interacting particle systems. For our case, the BEP was already introduced by De Masi *et al* in the original paper [18] (for the proof of Proposition 3.12, see Theorem 3.1 in [18]). The primary motivation is that it makes the study of correlation functions easier. To study the $k-$points correlation function, we only need to follow the history of k particles backward instead of realizing the whole process forward. This idea also relates to the famous "coupling from the past" algorithm in [82] and is also an essential ingredient of the information percolation framework in [68]. Proposition 3.12 says that the history itself can be viewed as a Markov process (the BEP). This will be useful for the construction of our coupling.

We introduce some further necessary notations.

Update functions. By construction, for any particle w alive at some point in the time interval $[0,t]$, there exists an increasing boolean function $F_{\mathsf{w},t}$ on $\{-1,1\}^{W_t}$ such that for any spin configuration $\mathfrak{s}$ on W_t, its extension to $\bigcup_{0\le s\le t} W_s$ satisfies

$$\mathfrak{s}_{\mathsf{w}} = F_{\mathsf{w},t}((\mathfrak{s}_{\mathsf{w}'})_{\mathsf{w}'\in W_t}). \tag{3.29}$$

The function $F_{\mathsf{w},t}$ describes the dependence of the spin of w on the spins of the particles alive at time t.

Pivotal sets of the update functions. For any $t \ge 0$, for any particle w that is alive at some time $s < t$, we denote by

$$\mathsf{Piv}\,(\mathsf{w},t) := \mathsf{Piv}\,(F_{\mathsf{w},t})\,. \tag{3.30}$$

We write $\mathsf{Piv}\,(E,t)$ for $\bigcup_{\mathsf{w}\in E}\mathsf{Piv}\,(\mathsf{w},t)$.

Observation 3.14. *For any $s < t$,*

$$\mathsf{Piv}\,(\mathsf{w},t) \subset \bigcup_{\mathsf{w}'\in \mathsf{Piv}(\mathsf{w},s)} \mathsf{Piv}\,(\mathsf{w}',t)\,. \tag{3.31}$$

Update times of the pivotal set. Let the sequence $(T_j)_{j\ge 0}$ be defined recursively, as follows.

$$\begin{aligned} T_0 &= 0,\\ T_{j+1} &= \inf\{t \ge T_j : \text{a Glauber clock of a particle in } \mathsf{Piv}\,(E,T_j) \text{ rings at time } t\}\,. \end{aligned}$$

Observation 3.15. $(\mathsf{Piv}\,(E,t))_{t\ge 0}$ *is piece-wise constant and only jumps at the times* $(T_j)_{j>0}$*. The same conclusion is true for* $(F_{\mathsf{w},t})_{t\ge 0}$*, for any* $\mathsf{w}\in E$*.*

Information percolation. We now briefly explain the idea that we borrow from [68], which will be made more precise later. First, to determine the spin of $\mathsf{w}\in E$, we do not need to follow the whole history, but only the part that is "essential", say $(\mathsf{Piv}\,(\mathsf{w},s))_{0\le s\le t}$. In particular, if $\mathsf{Piv}\,(\mathsf{w},s) = \emptyset$ for some $s < t$, then the update function $F_{\mathsf{w},s}$ degenerates, i.e. becomes a constant function. Then, we can determine $\mathfrak{s}_{\mathsf{w}}$ without tracking further the history in the past. Second, we regard $(\mathsf{Piv}\,(\mathsf{w},t))_{t\ge 0}$ as a percolation cluster in the space-time slab. Surprisingly, we prove in Lemma 3.20 below that the information percolation cluster defined this way is "subcritical" in the full high-temperature regime. This subcriticality means that the update functions quickly degenerate, so the spins at time t quickly become independent of those at time 0 as t grows to infinity, and the system mixes quickly.

4.3 The idealized branching process (IBP)

We will compare the BEP with its idealized version, the *idealized branching process* (IBP). The IBP starting from a particle $\mathsf{w}\in\mathcal{A}$ is defined as follows.

- Initially, w has q independent Poisson clocks, of intensities $\lambda_1, \dots, \lambda_q$, called the Glauber clocks of type $1, \dots, q$ respectively.
- Each time a clock of a particle w' rings, we remove w' and add the children of w' into the set of particles. Each of them is given q independent Glauber clocks, and the clocks of different particles are independent.

The IBP starting from w is a pair $(\Psi_{IBP}, \mathcal{C}_{IBP})$, where

- Ψ_{IBP} is the process associating each time $t \in \mathbb{R}_+$ with the set of particles alive at time t, i.e. the set of particles born before time t and not having been removed by time t.
- $\mathcal{C}_{IBP}$ is the counting measure on $[q] \times \mathbb{R}_+$ that tracks the type of the Glauber clocks that ring.

Remark 7. In the IBP, the particles no longer live on the lattice. All the particles are different, and all their clocks are independent. In particular, it is a dimension-free object, i.e. it does not depend on the size L of the system.

For any finite subset E of $\mathcal{A}$ such that for all $\mathsf{w}_1, \mathsf{w}_2 \in E$, w_1 is not a descendant of w_2, the IBP starting from E consists of $|E|$ independent IBPs starting from particles labeled by the elements of E.

Spins and update functions of the IBP. Let $\widetilde{W}_t$ be the set of particles of the IBP alive at time t. Similarly to what we have done for the BEP, for any spin configuration $\tilde{\mathfrak{s}}$ on $\widetilde{W}_t$, we extend $\tilde{\mathfrak{s}}$ to $\bigcup_{0 \le s \le t} \widetilde{W}_s$ by defining for each $\mathsf{w} \notin \widetilde{W}_t$,

$$\tilde{\mathfrak{s}}_{\mathsf{w}} := f_i\big((\tilde{\mathfrak{s}}_{\mathsf{w}'})_{\mathsf{w}' \in \{\mathsf{w}\} \times B(0,m)}\big).$$

We can extend $\tilde{\mathfrak{s}}$ to all particles alive at some time $s < t$ with this procedure. In particular, we can define the spin of w. We define the update functions $\tilde{F}_{\mathsf{w},t}$, the pivotal sets $(\widetilde{\mathsf{Piv}}(\tilde{F}_{\mathsf{w},t}))_{t \ge 0}$, and the update times $(\tilde{T}_j)_{j \in \mathbb{Z}_+}$ analogously as for the BEP.

4.4 The coupling

Let E be a subset of Λ_L. We will construct a coupling between the BEP and the IBP starting from E. Let there be one exclusion clock at each edge of the lattice and an infinite number of Glauber clocks of types $1, \dots, q$. All of them are independent. We will use these clocks to construct the two processes. We say that the coupling is *succesful* until time t if $\mathsf{Piv}\,(E,t) = \widetilde{\mathsf{Piv}}(E,t)$, and for each $\mathsf{w} \in \mathsf{Piv}\,(E,t)$, the Glauber clocks associated with it at time t are the same in the BEP and the IBP. We now construct the coupling, which should be thought of as embedding the IBP in the space-time slab $\Lambda_L \times \mathbb{R}_+$.

1. Whenever an exclusion clock rings, we put an exclusion mark at the corresponding edge. When a particle of the BEP meets an exclusion mark, it jumps to the other endpoint of the edge.

2. At time 0, each particle $\mathsf{w} \in E$ is associated with the same q Glauber clocks in the BEP and the IBP.

3. For each particle, the set of clocks associated with it stays unchanged in the interval $(T_j, T_{j+1}]$, where $(T_j)_{j\geq 0}$ are the update times of $\mathsf{Piv}\,(E, \cdot)$.

4. Suppose that the coupling is still successful until time T_j. Then T_{j+1} corresponds to the ring of a Glauber clock of some particle $\mathsf{w} \in \mathsf{Piv}\,(E, T_j)$. Note that $T_{j+1} = \tilde{T}_{j+1}$ because the same clock of w rings in the IBP. Suppose that w is at site u at time $T_{j+1}-$. Then one of two cases can happen.

   (a) **Case 1**: At time $T_{j+1}-$, there is no particle in $B(u, m)$ except w. This means that the children of w are born on sites not occupied by other particles in the pivotal set. Then we update $\mathsf{Piv}\,(E, T_{j+1})$ and $\widetilde{\mathsf{Piv}}(E, \tilde{T}_{j+1})$: we let the particles in $\mathsf{Piv}\,(E, T_{j+1}-) \setminus \{\mathsf{w}\}$ keep their clocks (in IBP and BEP), and we remove from $\mathsf{Piv}\,(E, T_{j+1}-)$ and $\widetilde{\mathsf{Piv}}(E, \tilde{T}_{j+1}-)$ the particles whose spins are no longer necessary to determine the spins atop, and then we associate the children of w in BEP and IBP with the same Glauber clocks. Note that $\mathsf{Piv}\,(E, T_{j+1}) \subset (\mathsf{Piv}\,(E, T_{j+1}-) \setminus \{\mathsf{w}\}) \cup (\{\mathsf{w}\} \times B(0, m))$. In this case, the coupling is still successful until time T_{j+1}.

   (b) **Case 2**: At time $T_{j+1}-$, there is some particles in $B(u, m)$ other than w. Then from time T_{j+1}, all the particles in the two pivotal sets are given different Poisson clocks. The coupling becomes unsuccessful since then.

We say that the coupling is successful until infinity if it is successful until t for any finite t. The interest in the above coupling is the following.

Observation 3.16 (Spins atop coincide)**.** *If the coupling is successful until time t, and if $\mathfrak{s}_{\mathsf{w}} = \tilde{\mathfrak{s}}_{\mathsf{w}},\ \forall \mathsf{w} \in \mathsf{Piv}\,(E, t)$, then $\mathfrak{s}_{\mathsf{w}} = \tilde{\mathfrak{s}}_{\mathsf{w}},\ \forall \mathsf{w} \in E$.*

We will show that the coupling is successful until infinity with high probability.

Proposition 3.17 (The coupling is successful w.h.p)**.** *Let E be an arbitrary subset of Λ_L. Consider the coupling of the BEP and the IBP starting from E described above. Then*

$$\mathbb{P}\left[\textit{The coupling is not successful until infinity}\right] \leq \beta \frac{|E|^2}{L}, \tag{3.32}$$

for some constant β.

The motivation for the coupling is as follows. We are inspired by the local equilibrium phenomenon, so we try to couple the BEP and the IBP so that when we generate the spins of the particles at time t in the two processes according to the same distribution (assuming that those sets are the same), then the spins atop are the same. If this is the case, as the spins atop of the IBP are independent, the spins atop of the BEP are also independent. The naive coupling is to match every particle in the two processes. However, the total number of particles increases exponentially fast, so we can not expect the naive coupling to work until the mixing times. Nevertheless, it is *not necessary* to match the evolution of every particle, but only the ones that truly influence the spins atop, say $\mathsf{Piv}(E, \cdot)$.

The idea to look only at $(\mathsf{Piv}(E,t))_{t\geq 0}$ ultimately comes from [68] (in this paper, the authors call it the update support function, see Subsection 2.2 in [68]). In Subsection 4.5, we analyze the IBP, and in particular, we prove that $\widetilde{\mathsf{Piv}}(\cdot)$ behaves like a subcritical Galton Watson process. In Subsection 4.6, we use this to prove Proposition 3.17.

4.5 Analysis of the IBP

In this subsection, we consider the IBP starting from a particle $\mathsf{w} \in \mathcal{A}$. For convenience, we identify the set of children of w with $B(0,m)$. Let the functions $\phi, \psi : \mathbb{R}_+ \to \mathbb{R}$ be defined by

$$\phi(t) := \mathbb{P}\left[\widetilde{\mathsf{Piv}}(\mathsf{w},t) \neq \emptyset\right], \tag{3.33}$$

$$\psi(t) := \mathbb{E}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t)|\right], \tag{3.34}$$

which means that $\phi(t)$ is the probability that the process $\widetilde{\mathsf{Piv}}(\mathsf{w},\cdot)$ still survives at time t and $\psi(t)$ is the average size of $\widetilde{\mathsf{Piv}}(\mathsf{w},t)$. The main purpose of this subsection is to analyze several aspects of the IBP and, in particular, to prove that $\psi(t) \sim e^{R(\rho_*)t}$.

The average size of $\widetilde{\mathsf{Piv}}(\mathsf{w},t)$ conditionally on survival is

$$\mathbb{E}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t)|\,|\widetilde{\mathsf{Piv}}(\mathsf{w},t) \neq \emptyset\right] = \frac{\psi(t)}{\phi(t)}. \tag{3.35}$$

Note that if $\widetilde{\mathsf{Piv}}(\mathsf{w},t) = \emptyset$, then the spin of w does not depend on the spins at time t anymore, so we can safely write

$$\vartheta(t) = \mathbb{E}\left[\tilde{\mathfrak{s}}_{\mathsf{w}}|\widetilde{\mathsf{Piv}}(\mathsf{w},t) = \emptyset\right] \tag{3.36}$$

for the average spin of w conditionally on the extinction of $\widetilde{\mathsf{Piv}}(\mathsf{w},t)$. Clearly, the functions ψ, ϕ, ϑ do not depend on our choice of $\mathsf{w} \in \mathcal{A}$.

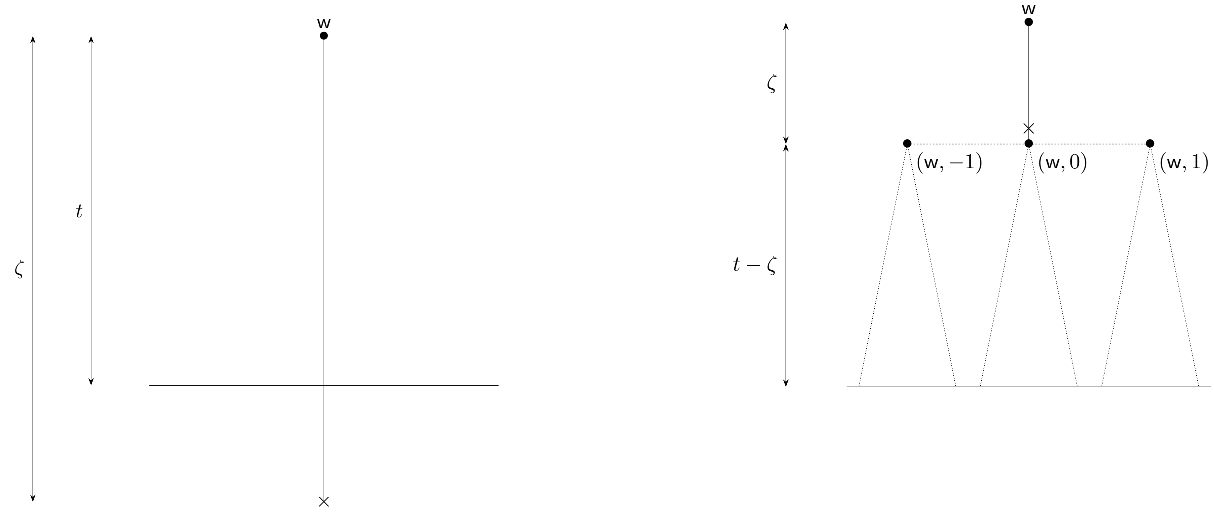
First, we have a result linking the average of $\tilde{\mathfrak{s}}_{\mathsf{w}}$ with the ODE (3.9).

Lemma 3.18 (Spin atop of an IBP). *Let $\rho_0 \in [-1,1]$ and $t \geq 0$. Consider the IBP starting from w. Suppose that, conditionally on the IBP, the spins at time t are generated independently by a product of Rademacher: $\tilde{\mathfrak{s}}(\widetilde{W}_t) \sim \mathfrak{R}_{\rho_0}^{\otimes \widetilde{W}_t}$. Then*

$$\mathbb{E}\left[\tilde{\mathfrak{s}}_{\mathsf{w}}\right] = \rho(t), \tag{3.37}$$

where ρ is the solution of equation (3.9).

Proof. Let $g(t) := \mathbb{E}\left[\tilde{\mathfrak{s}}_{\mathsf{w}}\right]$. Let ζ be the first time that a Glauber clock of w rings. Let $\mathcal{E}(i)$ be the event that the Glauber clock ringing at time ζ is of type i. Then ζ and $\mathcal{E}(i)$ are independent, and $\zeta \sim \exp(\lambda)$, $\mathbb{P}\left[\mathcal{E}(i)\right] = \lambda_i/\lambda$. Moreover, the branches starting from the children of w are



(a) $\zeta > t$ (b) $\zeta < t$

Figure 3.3: Conditionally on the first ring

independent and independent of ζ and $\mathcal{E}(i)$, each behaving as an IBP. We see that

$$g(t) = \mathbb{E}\left[\mathbb{1}_{\{\zeta>t\}}\tilde{\mathfrak{s}}_{\mathsf{w}}\right] + \mathbb{E}\left[\mathbb{1}_{\{\zeta\le t\}}\tilde{\mathfrak{s}}_{\mathsf{w}}\right]. \tag{3.38}$$

We estimate the two terms on the right-hand side of (3.38) separately.

- **First term**: conditionally on ζ (see Figure 3.3), on the event $\{\zeta > t\}$, $\tilde{\mathfrak{s}}_{\mathsf{w}} \sim \mathfrak{R}_{\rho_0}$, and hence

$$\mathbb{E}\left[\mathbb{1}_{\{\zeta>t\}}\tilde{\mathfrak{s}}_{\mathsf{w}}\right] = \mathbb{E}\left[\mathbb{1}_{\{\zeta>t\}}\mathbb{E}\left[\tilde{\mathfrak{s}}_{\mathsf{w}}|\zeta\right]\right] = \mathbb{E}\left[\mathbb{1}_{\{\zeta>t\}} \times \rho_0\right] = \rho_0\mathbb{P}\left[\zeta > t\right] = \rho_0 e^{-\lambda t}.$$

- **Second term:** we see that

$$\begin{aligned}\mathbb{E}\left[\mathbb{1}_{\{\zeta\le t\}}\tilde{\mathfrak{s}}_{\mathsf{w}}\right] &= \sum_{i=1}^{q}\mathbb{E}\left[\mathbb{1}_{\mathcal{E}(i)}\mathbb{1}_{\{\zeta\le t\}}f_i((\tilde{\mathfrak{s}}_{\mathsf{w}'})_{\mathsf{w}'\in\{\mathsf{w}\}\times B(0,m)})\right] \\ &= \sum_{i=1}^{q}\mathbb{P}\left[\mathcal{E}(i)\right]\mathbb{E}\left[\mathbb{1}_{\{\zeta\le t\}}f_i\left((\tilde{\mathfrak{s}}_{\mathsf{w}'})_{\mathsf{w}'\in\{\mathsf{w}\}\times B(0,m)}\right)\right],\end{aligned}$$

since $\mathcal{E}(i)$ is independent of ζ and of the IBPs starting from the children of w. Conditionally on ζ, on the event $\{\zeta < t\}$, by construction, $(\tilde{\mathfrak{s}}_{\mathsf{w}'})_{\mathsf{w}'\in\{\mathsf{w}\}\times B(0,m)}$ are i.i.d. $\mathfrak{R}_{g(t-\zeta)}$. Hence

$$\mathbb{E}\left[\mathbb{1}_{\{\zeta<t\}}f_i\left((\tilde{\mathfrak{s}}_{\mathsf{w}'})_{\mathsf{w}'\in\{\mathsf{w}\}\times B(0,m)}\right)\right] = \mathbb{E}\left[\mathbb{1}_{\{\zeta<t\}}\mathbb{E}_{\nu_{g(t-\zeta)}}\left[f_i\right]\right].$$

Hence

$$\begin{aligned}\mathbb{E}\left[\mathbb{1}_{\{\zeta\le t\}}\tilde{\mathfrak{s}}_{\mathsf{w}}\right] &= \sum_{i=1}^{q}\mathbb{P}\left[\mathcal{E}(i)\right]\mathbb{E}\left[\mathbb{1}_{\{\zeta<t\}}\mathbb{E}_{\nu_{g(t-\zeta)}}\left[f_i\right]\right]\\ &= \sum_{i=1}^{q}\frac{\lambda_i}{\lambda}\mathbb{E}\left[\mathbb{1}_{\{\zeta<t\}}\mathbb{E}_{\nu_{g(t-\zeta)}}\left[f_i\right]\right]\\ &= \mathbb{E}\left[\mathbb{1}_{\{\zeta<t\}}\sum_{i=1}^{q}\frac{\lambda_i}{\lambda}\mathbb{E}_{\nu_{g(t-\zeta)}}\left[f_i\right]\right]\\ &= \mathbb{E}\left[\mathbb{1}_{\{\zeta<t\}}\times\frac{1}{\lambda}\times\left[R(g(t-\zeta))+\lambda g(t-\zeta)\right]\right],\end{aligned}$$

where the last equality comes from Lemma 3.11.

This implies that the first term of (3.38) equals $e^{-\lambda t}\rho_0$ and the second term of (3.38) equals

$$\begin{aligned}\int_0^t \lambda e^{-\lambda s}\frac{1}{\lambda}\left(R(g(t-s))+\lambda g(t-s)\right)\mathrm{d}s &= \int_0^t e^{-\lambda s}\left(R(g(t-s))+\lambda g(t-s)\right)\mathrm{d}s\\ &= \int_0^t e^{-\lambda(t-s)}\left(R(g(s))+\lambda g(s)\right)\mathrm{d}s,\end{aligned}$$

where the last inequality follows from the change of variable $s' = t - s$. This implies that

$$\begin{aligned}g(t) &= e^{-\lambda t}\rho_0 + \int_0^t e^{-\lambda(t-s)}\left(R(g(s))+\lambda g(s)\right)\mathrm{d}s\\ &= e^{-\lambda t}\left(\rho_0 + \int_0^t e^{\lambda s}\left(R(g(s))+\lambda g(s)\right)\mathrm{d}s\right).\end{aligned}$$

Now we can differentiate both sides with respect to t to conclude that

$$\begin{aligned}g'(t) &= -\lambda e^{-\lambda t}\left(\rho_0 + \int_0^t e^{\lambda s)}\left(R(g(s))+\lambda g(s)\right)\mathrm{d}s\right) + e^{-\lambda t}e^{\lambda t}\left(R(g(t))+\lambda g(t)\right)\\ &= -\lambda g(t) + R(g(t)) + \lambda g(t)\\ &= R(g(t)).\end{aligned}$$

This means that g solves equation (3.9) with initial condition $g(0) = \rho_0$. This finishes our proof. □

Recall that ρ_+ and ρ_- are solutions of equation (3.8) with initial conditions 1 and -1. We have the following lemma.

Lemma 3.19.

$$\phi(t) = \frac{\rho_+(t) - \rho_-(t)}{2}, \tag{3.39}$$

$$\vartheta(t) = \frac{\rho_+(t) + \rho_-(t)}{2(1-\phi(t))}. \tag{3.40}$$

Proof. Let us generate the IBP up to time t. We denote by $\mathbb{E}_+[\cdot]$ the probability taken with respect to the IBP when we assign the spins at time t to be all-plus, i.e. $\tilde{\mathfrak{s}}_{\mathsf{w}'} := 1$, $\forall \mathsf{w}' \in \widetilde{W}_t$

and $\mathbb{E}_-[\cdot]$ the probability taken when we assign the spins at time t to be all-minus, i.e. $\tilde{\mathfrak{s}}_{\mathsf{w}'} := -1,\ \forall \mathsf{w}' \in \widetilde{W}_t$. Note that

$$\mathbb{E}_+[\tilde{\mathfrak{s}}_{\mathsf{w}}] = \mathbb{E}_+\left[\tilde{\mathfrak{s}}_{\mathsf{w}} \mathbb{1}_{\left\{\widetilde{\mathsf{Piv}}(\mathsf{w},t)\neq\emptyset\right\}}\right] + \mathbb{E}_+\left[\tilde{\mathfrak{s}}_{\mathsf{w}} \mathbb{1}_{\left\{\widetilde{\mathsf{Piv}}(\mathsf{w},t)=\emptyset\right\}}\right]. \tag{3.41}$$

Recall that $\widetilde{\mathsf{Piv}}(\mathsf{w},t) = \mathsf{Piv}\left(\tilde{F}_{\mathsf{w},t}\right)$, where $\tilde{\mathfrak{s}}_{\mathsf{w}} = \tilde{F}_{\mathsf{w},t}\left((\tilde{\mathfrak{s}}_{\mathsf{w}'})_{\mathsf{w}'\in\widetilde{W}_t}\right)$ for some increasing boolean function $\tilde{F}_{\mathsf{w},t}$. Hence

$$\mathbb{E}_+\left[\tilde{\mathfrak{s}}_{\mathsf{w}} \mathbb{1}_{\left\{\widetilde{\mathsf{Piv}}(\mathsf{w},t)\neq\emptyset\right\}}\right] = \mathbb{E}_+\left[\mathbb{1}_{\left\{\widetilde{\mathsf{Piv}}(\mathsf{w},t)\neq\emptyset\right\}}\right] = \mathbb{P}\left[\widetilde{\mathsf{Piv}}(\mathsf{w},t) \neq \emptyset\right],$$

where the first equality is due to Lemma 3.8. Moreover, when $\widetilde{\mathsf{Piv}}(\mathsf{w},t) = \emptyset$, $\tilde{\mathfrak{s}}_{\mathsf{w}}$ no longer depends on the spins of particles in $\widetilde{W}_t$, so the second term in (3.41) is equal to $(1-\phi(t))\vartheta(t)$ by definition of the functions ϕ, ϑ. This implies that

$$\rho_+(t) = \phi(t) + (1-\phi(t))\vartheta(t).$$

Similarly

$$\rho_-(t) = -\phi(t) + (1-\phi(t))\vartheta(t).$$

These two equalities imply what we want. ☐

In the following proposition, we prove that ψ satisfies a particular differential equation, implying that the pivotal set is subcritical, one of the most important results in this chapter.

Proposition 3.20 (Differential equation for ψ)**.** *There exist a univariate polynomial Q_1 and a bivariate polynomial Q_2 such that*

$$\psi' = \psi\left(R'(\rho_*) + (\vartheta - \rho_*)Q_1(\vartheta) + \phi Q_2(\phi, \vartheta)\right). \tag{3.42}$$

Proof. The idea is to proceed as in the proof of Lemma 3.18. See Figure 3.3 for intuition. Let us generate the IBP starting from one particle w up to time t. Let ζ be the first time that a Glauber clock of w rings. Then $\zeta \sim \exp(\lambda)$. We see that

$$\begin{aligned}\psi(t) &= \mathbb{E}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t)|\right] \\ &= \mathbb{E}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t)|\mathbb{1}_{\{\zeta>t\}}\right] + \mathbb{E}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t)|\mathbb{1}_{\{\zeta<t\}}\right].\end{aligned} \tag{3.43}$$

We estimate the two terms in (3.43) separately.

- **First term:** when $\zeta > t$, $\widetilde{\mathsf{Piv}}(\mathsf{w},t) = \{\mathsf{w}\}$ by construction, and hence

$$\mathbb{E}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t)|\mathbb{1}_{\{\zeta>t\}}\right] = \mathbb{E}\left[|\{\mathsf{w}\}|\mathbb{1}_{\{\zeta>t\}}\right] = \mathbb{P}[\zeta > t]. \tag{3.44}$$

- **Second term:** For any subset $A \subset B(0,m)$, any $i \in [q]$, any spin configuration η, ξ on

$B(0,m)$, let us define the spin configuration $\xi^{A,\eta}$ on $B(0,m)$, the boolean function $f_i^{A,\eta}$ on $\{-1,1\}^{B(0,m)}$, and the events $\mathcal{E}(i), \mathcal{E}(A,\eta)$ by

$$\begin{aligned}
\xi^{A,\eta}(j) &:= \eta(j) \times \mathbb{1}_{\{j\in A\}} + \xi(j)\mathbb{1}_{\{j\in A^C\}}, \\
f_i^{A,\eta}(\xi) &:= f_i(\xi^{A,\eta}), \\
\mathcal{E}(i) &:= \{\text{The Glauber clock of } \mathsf{w} \text{ ringing at time } \zeta \text{ is of type } i\}, \\
\mathcal{E}(A,\eta) &:= \{\widetilde{\mathsf{Piv}}(w',t) \neq \emptyset, \forall \mathsf{w}' \in A, \text{ and } \widetilde{\mathsf{Piv}}(w',t) = \emptyset, \forall \mathsf{w}' \in A^C\} \cap \{(\tilde{\mathfrak{s}}_{\mathsf{w}'})_{\mathsf{w}'\in A^C} = \eta|_{A^C}\}.
\end{aligned}$$

Note that, on the event $\mathcal{E}(i) \cap \mathcal{E}(A,\eta)$,

$$\widetilde{\mathsf{Piv}}(\mathsf{w},t) = \bigcup_{\mathsf{w}'\in \mathsf{Piv}\left(f_i^{A^C,\eta}\right)} \widetilde{\mathsf{Piv}}(\mathsf{w}',t).$$

Hence

$$\mathbb{E}\left[\left|\widetilde{\mathsf{Piv}}(\mathsf{w},t)\right| \mathbb{1}_{\{\zeta<t\}}\mathbb{1}_{\mathcal{E}(i)}\mathbb{1}_{\mathcal{E}(A,\eta)}\right] = \sum_{\mathsf{w}'\in \mathsf{Piv}\left(f_i^{A^C,\eta}\right)} \mathbb{E}\left[\left|\widetilde{\mathsf{Piv}}(\mathsf{w}',t)\right| \mathbb{1}_{\{\zeta<t\}}\mathbb{1}_{\mathcal{E}(i)}\mathbb{1}_{\mathcal{E}(A,\eta)}\right]. \tag{3.45}$$

By construction, $\zeta, \mathcal{E}(i)$ and the branches of the IBP starting at the children of w are independent. Hence, for any non empty $A \subset B(0,m)$, $\mathsf{w}' \in A, \eta \in \{-1,1\}^{B(0,m)}$,

$$\begin{aligned}
&\mathbb{E}\left[\left|\widetilde{\mathsf{Piv}}(\mathsf{w}',t)\right| \mathbb{1}_{\{\zeta<t\}}\mathbb{1}_{\mathcal{E}(i)}\mathbb{1}_{\mathcal{E}(A,\eta)}\right] \\
&= \mathbb{P}\left[\mathcal{E}(i)\right]\mathbb{E}\left[\left|\widetilde{\mathsf{Piv}}(\mathsf{w}',t)\right| \mathbb{1}_{\{\zeta<t\}}\mathbb{1}_{\mathcal{E}(A,\eta)}\right]. \\
&= \frac{\lambda_i}{\lambda}\mathbb{E}\left[\mathbb{1}_{\{\zeta<t\}}\left|\widetilde{\mathsf{Piv}}(\mathsf{w}',t)\right| \prod_{\mathsf{w}''\in A\setminus\{\mathsf{w}'\}} \mathbb{1}_{\left\{\widetilde{\mathsf{Piv}}(\mathsf{w}'',t)\neq\emptyset\right\}} \prod_{\mathsf{w}''\in A^C} \mathbb{1}_{\left\{\widetilde{\mathsf{Piv}}(\mathsf{w}'',t)=\emptyset, \tilde{\mathfrak{s}}_{\mathsf{w}''}=\eta_{\mathsf{w}''}\right\}}\right].
\end{aligned}$$

Note that the branches starting at the children of w at time ζ are independent IBPs independent of ζ, shifted by a time ζ. Then we can integrate out the randomness of these branches to conclude that

$$\mathbb{E}\left[\left|\widetilde{\mathsf{Piv}}(\mathsf{w}',t)\right| \mathbb{1}_{\{\zeta<t\}}\mathbb{1}_{\mathcal{E}(i)}\mathbb{1}_{\mathcal{E}(A,\eta)}\right] = \frac{\lambda_i}{\lambda}\mathbb{E}\left[\mathbb{1}_{\{\zeta<t\}}\psi\phi^{|A|-1}\prod_{\mathsf{w}''\in A^C}(1-\phi)\mathfrak{R}_\vartheta(\eta_{\mathsf{w}''})\right],$$

where we have used the definition of ψ, ϕ, and ϑ. In the formula above, the functions ψ, ϕ, ϑ inside the expectation are evaluated at $t-\zeta$. Hence

$$\mathbb{E}\left[\left|\widetilde{\mathsf{Piv}}(\mathsf{w}',t)\right| \mathbb{1}_{\{\zeta<t\}}\mathbb{1}_{\mathcal{E}(i)}\mathbb{1}_{\mathcal{E}(A,\eta)}\right] = \mathbb{E}\left[\mathbb{1}_{\{\zeta<t\}}\mathcal{M}(i,A,\eta)\right],$$

with

$$\mathcal{M}(i,A,\eta) := \frac{\lambda_i}{\lambda}\psi\phi^{|A|-1}(1-\phi)^{2m+1-|A|}\prod_{\mathsf{w}''\in A^C}\mathfrak{R}_\vartheta(\eta_{\mathsf{w}''}),$$

where the functions ψ, ϕ, ϑ are evaluated at $t-\zeta$. For convention, $\mathcal{M}(i,A,\eta) := 0$ if $|A| = 0$. Hence

$$\begin{aligned}\mathbb{E}\left[\left|\widetilde{\mathsf{Piv}}(\mathsf{w},t)\right| \mathbb{1}_{\{\zeta<t\}}\right] &= \sum_{i,A,\eta} \mathbb{E}\left[\mathbb{1}_{\mathcal{E}(i)}\mathbb{1}_{\mathcal{E}(A,\eta)}\mathbb{1}_{\{\zeta<t\}}|\widetilde{\mathsf{Piv}}(\mathsf{w},t)|\right] \\ &= \sum_{i,A,\eta} \sum_{\mathsf{w}'\in\mathsf{Piv}\left(f_i^{A^C,\eta}\right)} \mathbb{E}\left[\mathbb{1}_{\mathcal{E}(i)}\mathbb{1}_{\mathcal{E}(A,\eta)}\mathbb{1}_{\{\zeta<t\}}|\widetilde{\mathsf{Piv}}(\mathsf{w}',t)|\right] \\ &= \mathbb{E}\left[\mathbb{1}_{\{\zeta<t\}} \sum_{i,A,\eta} \left|\mathsf{Piv}\left(f_i^{A^C,\eta}\right)\right| \mathcal{M}(i,A,\eta)\right], \end{aligned} \tag{3.46}$$

where the sums are taken on all $i \in [q], A \subset B(0,m)$, and $\eta \in \{-1,1\}^{B(0,m)}$. Note that,

$$\sum_{i,A,\eta} \mathbb{1}_{\{|A|=0\}} \left|\mathsf{Piv}\left(f_i^{A,\eta}\right)\right| \mathcal{M}(i,A,\eta) = 0, \tag{3.47}$$

and

$$\sum_{i,A,\eta} \mathbb{1}_{\{|A|\geq 2\}} \left|\mathsf{Piv}\left(f_i^{A^C,\eta}\right)\right| \mathcal{M}(i,A,\eta) = \psi\phi g, \tag{3.48}$$

where g is a polynomial of ϕ and ϑ, and the functions ψ, ϕ, ϑ are evaluated at $t-\zeta$. We now estimate the sum on the subsets A such that $|A| = 1$. Note that if $A = \{j\}$, then

$$\left|\mathsf{Piv}\left(f_i^{\{j\}^C,\eta}\right)\right| = \frac{1}{2}\nabla_j f_i(\eta),$$

due to Lemma 3.8. Hence

$$\begin{aligned}&\sum_{i,A,\eta} \mathbb{1}_{\{|A|=1\}} \left|\mathsf{Piv}\left(f_i^{A^C,\eta}\right)\right| \mathcal{M}(i,A,\eta) \\ &= \sum_{i,j,\eta} \frac{1}{2}\nabla_j f_i(\eta)\mathcal{M}(i,\{j\},\eta) \\ &= \sum_{i,j,\eta} \frac{1}{2}\nabla_j f_i(\eta) \times \frac{\lambda_i}{\lambda}\psi(1-\phi)^{2m} \prod_{\mathsf{w}''\in B(0,m)\setminus\{j\}} \mathfrak{R}_\vartheta(\eta_{\mathsf{w}''}) \\ &= \psi(1-\phi)^{2m} \sum_{i=1}^{q} \frac{\lambda_i}{\lambda} \sum_{j\in B(0,m)} \mathbb{E}_{\nu_\vartheta}\left[\frac{1}{2}\nabla_j f_i\right] \\ &= \psi(1-\phi)^{2m}\frac{1}{\lambda}(R'(\vartheta)+\lambda), \end{aligned} \tag{3.49}$$

where we have used Lemma 3.11 in the last equality. The equations (3.46), (3.47), (3.48), (3.49) together imply that

$$\mathbb{E}\left[\widetilde{\mathsf{Piv}}(\mathsf{w},t)\mathbb{1}_{\{\zeta<t\}}\right] = \mathbb{E}\left[\mathbb{1}_{\{\zeta<t\}} \times \left[\psi(1-\phi)^{2m}\left(\frac{R'(\vartheta)}{\lambda}+1\right) + \psi\phi g\right]\right], \tag{3.50}$$

where the functions ψ, ϕ, ϑ inside the expectation are evaluated at $t-\zeta$. The equations (3.44), (3.50), (3.43) together imply that

$$\begin{aligned}\psi(t) &= \mathbb{P}\left[\zeta > t\right] + \mathbb{E}\left[\mathbb{1}_{\{\zeta<t\}} \times \left[\psi(1-\phi)^{2m}\left(\frac{R'(\vartheta)}{\lambda}+1\right)+\psi\phi g\right]\right] \\ &= e^{-\lambda t} + \int_0^t \lambda e^{-\lambda s}\left[\psi(1-\phi)^{2m}\left(\frac{R'(\vartheta)}{\lambda}+1\right)+\psi\phi g\right]\mathrm{d}s\end{aligned}$$

where the functions ψ, ϕ, ϑ inside the integral sign are evaluated at $t-s$. We make a change of variable $s' = t - s$ to conclude that

$$\begin{aligned}\psi(t) &= e^{-\lambda t} + \int_0^t \lambda e^{-\lambda(t-s)}\left[\psi(1-\phi)^{2m}\left(\frac{R'(\vartheta)}{\lambda}+1\right)+\psi\phi g\right]\mathrm{d}s \\ &= e^{-\lambda t}\left(1 + \int_0^t \lambda e^{\lambda s}\left[\psi(1-\phi)^{2m}\left(\frac{R'(\vartheta)}{\lambda}+1\right)+\psi\phi g\right]\mathrm{d}s\right),\end{aligned}$$

where the functions ψ, ϕ, ϑ are now evaluated at s. Now, we can differentiate both sides with respect to t to conclude that

$$\psi' = -\lambda\psi + e^{-\lambda t}\lambda e^{\lambda t}\left[\psi(1-\phi)^{2m}\left(\frac{R'(\vartheta)}{\lambda}+1\right)+\psi\phi g\right],$$

where all functions $\psi', \phi, \psi, \vartheta$ are evaluated at t. Simplifying that formula, we obtain

$$\psi' = \psi\left(-\lambda + (1-2\phi)^{2m}\lambda + (1-2\phi)^{2m}R'(\vartheta) + \phi g\right).$$

Recall that R is a univariate polynomial. Hence, we can find a univariate polynomial Q_1 and a bivariate polynomial Q_2 such that

$$\begin{aligned}(\vartheta-\rho_*)Q_1(\vartheta) &= R'(\vartheta) - R'(\rho_*), \\ \phi Q_2(\phi,\vartheta) &= \left((1-2\phi)^{2m}-1\right)(R'(\vartheta)+\lambda) + \phi g.\end{aligned}$$

This finishes the proof. □

Remark 8. In the proof above, the monotonicity of the functions $(f_i)_{1\le i\le q}$ is crucial, and so is Hypothesis 1.

We derive the asymptotic of ψ using Hypothesis 2.

Lemma 3.21 (Asymptotic of some important functions)**.** *There exists a constant κ such that, for any $t \ge 0$, all the numbers $\log\psi(t), \log\phi(t), \log(\rho_+(t)-\rho_*), \log(\rho_*-\rho_-(t))$ lie in the interval $[R'(\rho_*)t-\kappa, R'(\rho_*)t+\kappa]$.*

Proof of Lemma 3.21. We prove the result for $\psi(t)$. The results for other functions are proved similarly. By Hypothesis 2, we easily see that

$$\begin{aligned}R(\rho) &> 0,\ \forall\rho\in[-1,\rho_*), \\ R(\rho) &< 0,\ \forall\rho\in(\rho_*,1].\end{aligned}$$

Hence $\rho_+(t) \searrow \rho_*$. Moreover,

$$\lim_{t\to\infty} \frac{\rho'_+(t)}{\rho_+(t)-\rho} = \lim_{t\to\infty} \frac{R(\rho_+(t)) - R(\rho_*)}{\rho_+(t)-\rho_*} = R'(\rho_*).$$

Then for t large enough,

$$\frac{\rho'_+(t)}{\rho_+(t)-\rho_*} \leq \frac{R'(\rho_*)}{2}.$$

Then, by Gronwall's lemma, there exists a positive constant κ_1 such that

$$\forall t \geq 0,\ \rho_+(t) - \rho_* \leq \kappa_1 e^{\frac{1}{2}R'(\rho_*)t}.$$

One can prove that the inequality above is still true if we replace $\rho_+(t) - \rho_*$ by the positive functions $\rho_* - \rho_-(t)$, $\phi(t)$, $|\vartheta(t) - \rho_*|$. Therefore,

1. $\phi(t)$, $|\vartheta(t) - \rho_*|$ are absolutely integrable with respect to t.
2. $Q_1(\vartheta)$ and $Q_2(\phi, \vartheta)$ are bounded uniformly in t.

Hence $(\vartheta - \rho_*)Q_1(\vartheta) + \phi Q_2(\phi, \vartheta)$ is absolutely integrable as a function of t. Recall that Proposition 3.20 implies that

$$\frac{\psi'}{\psi} - R'(\rho_*) = (\vartheta - \rho_*)Q_1(\vartheta) + \phi Q_2(\phi, \vartheta).$$

Hence, taking integration from 0 to t, and using the inequality $|\int f - \int g| \leq \int |f - g|$, and noting that $\psi(0) = 1$, we obtain

$$\left|\log \psi(t) - R'(\rho_*)t\right| \leq \int_0^t |(\vartheta - \rho_*)Q_1(\vartheta) + \phi Q_2(\phi, \vartheta)| \leq \kappa,$$

where

$$\kappa = \int_0^\infty |(\vartheta - \rho_*)Q_1(\vartheta) + \phi Q_2(\phi, \vartheta)|\,.$$

This finishes our proof. □

Remark 9. Hypothesis 2 is crucial for the proof above.

We present some elementary but useful results on $\widetilde{W}_t$.

Lemma 3.22. *Consider the IBP starting from a set E. Then there exists a constant κ such that, for any number $t > 0$,*

$$\mathbb{E}\left[|\widetilde{W}_t|^2\right] \leq e^{\kappa t}|E|^2.$$

Proof of Lemma 3.22. Note that $(|\widetilde{W}_t|)_{t\geq 0}$ is a Markov process itself. Let $\mathcal{L}$ be the generator associated with $(|\widetilde{W}_t|)_{t\geq 0}$. Then, for any function $\varphi : \mathbb{Z}_+ \to \mathbb{R}$, for any $n \in \mathbb{Z}_+$,

$$\mathcal{L}\varphi(n) = \lambda n\left[\varphi(n + 2m + 1 - 1) - \varphi(n)\right] = \lambda n\left[\varphi(n + 2m) - \varphi(n)\right].$$

We also know that, see e.g. [31],

$$\frac{d}{dt}\mathbb{E}\left[\varphi(|\widetilde{W}_t|)\right] = \mathbb{E}\left[\mathcal{L}\varphi(|\widetilde{W}_t|)\right].$$

Applying the formula above for $\varphi : n \mapsto n^2$, we obtain

$$\begin{aligned}\frac{d}{dt}\mathbb{E}\left[|\widetilde{W}_t|^2\right] &= \mathbb{E}\left[\lambda|\widetilde{W}_t|(4m|\widetilde{W}_t| + 4m^2)\right] \\ &\leq \lambda(4m + 4m^2)\mathbb{E}\left[|\widetilde{W}_t|^2\right].\end{aligned}$$

Then, we can apply Gronwall's lemma to conclude that

$$\mathbb{E}\left[|\widetilde{W}_t|^2\right] \leq e^{(4m+4m^2)\lambda t}\mathbb{E}\left[|\widetilde{W}_0|^2\right],$$

which finishes the proof. □

Lemma 3.23. *Consider the BEP and IBP starting from $E \subset \Lambda_L$. Then, $|\widetilde{W}_t|$ stochastically dominates $|W_t|$ for any $t \geq 0$.*

Proof of Lemma 3.23. The reason is that the particles in the two processes are given q Poisson clocks with the same rates. For the BEP, each time a clock of a particle rings, the particle is removed, and *at most* $2m+1$ new particles are added to W_t because some new particles are born on the sites already occupied. In contrast, for the process $|\widetilde{W}_t|$, each time a clock of a particle rings, *exactly* $2m + 1$ new particles are added. □

4.6 Success of the coupling

Now we can prove Proposition 3.17. We will need a result about anti-concentration of the Interchange Process IP(2).

Interchange process IP(2). The interchange process IP(2) on Λ_L^d, whose edges have conductance L^2, is a couple of random walk $(\mathcal{U}_1,\mathcal{U}_2)$ that has the following description. The state space is $(\Lambda_L^d \times \Lambda_L^d) \setminus \{(u, u)|u \in \Lambda_L^d\}$. For any initial condition, the evolution is as follows. Each edge of Λ_L^d is associated with an independent Poisson clock of intensity L^2. Whenever a clock rings, if $\mathcal{U}_1$ (or $\mathcal{U}_2$) is at one endpoint of the corresponding edge, it jumps to the other endpoint.

From now on, we denote by $\mathsf{dist}(\cdot, \cdot)$ the shortest-path distance on the lattice. We have the following result.

Lemma 3.24 (Anticoncentration)**.** *For any $d \in \mathbb{Z}_{>0}$. Let $(\mathcal{U}_1,\mathcal{U}_2)$ be an IP(2) on Λ_L^d, whose edges have conductance L^2. Let θ be a strictly positive number and $\zeta \sim \exp(\theta)$, independent of $(\mathcal{U}_1,\mathcal{U}_2)$. Then, for any $k \in \mathbb{Z}_+$,*

$$\max_{u_1,u_2\in\Lambda_L^d;\, u_1\neq u_2} \mathbb{P}_{u_1,u_2}\left[\mathsf{dist}(\mathcal{U}_1(\zeta),\mathcal{U}_2(\zeta)) \leq k\right] = \begin{cases}\mathcal{O}_{\theta,k}(1/L) & \text{if } d = 1,\\ \mathcal{O}_{\theta,k}(\log L/L^2) & \text{if } d = 2,\\ \mathcal{O}_{\theta,k}(1/L^2) & \text{if } d \geq 3.\end{cases}$$

The proof of Lemma 3.24 is postponed to Appendix. To lighten the notation, we write $\mathsf{Piv}(t)$ for $\mathsf{Piv}(E,t)$ and $\widetilde{\mathsf{Piv}}(t)$ for $\widetilde{\mathsf{Piv}}(E,t)$. The proof of Proposition 3.17 is divided into two following lemmas.

Lemma 3.25. *With the same notations as in Proposition 3.17,*

$$\mathbb{P}\left[\textit{The coupling is not successful until infinity}\right] = \mathcal{O}\left(\frac{\sum_{j=0}^{\infty}\mathbb{E}\left[\widetilde{\mathsf{Piv}}(\tilde{T}_j)\right]}{L}\right).$$

Lemma 3.26.

$$\sum_{j=0}^{\infty}\mathbb{E}\left[\left|\widetilde{\mathsf{Piv}}(\tilde{T}_j)\right|\right] = \mathcal{O}\left(|E|^2\right).$$

Clearly, those two lemmas imply Proposition 3.17. Those two lemmas separate the effect of the exclusion process and the IBP.

Proof of Lemma 3.25. We say that two particles w_1 and w_2 have an interaction if there is a time t such that:

1. Both are in the pivotal set at time $t-$, i.e. $\mathsf{w}_1, \mathsf{w}_2 \in \mathsf{Piv}(t-)$, and they do not occupy the same site.
2. At time t, a clock of one of the two particles rings, and then one of the children is born on the site occupied by the other particle.

In short, two particles in $\mathsf{Piv}(\cdot)$ have an interaction if one splits into the site occupied by the other. For example, in Figure 3.2, w and w' have an interaction at time s_3. By definition, the coupling is successful up to time t if there is no interaction up to time t. Hence

$$\mathbb{P}\left[\text{The coupling is not successful until infinity}\right] \leq \mathbb{P}\left[\text{There is an interaction}\right].$$

We only need to prove

$$\mathbb{P}\left[\text{There is an interaction}\right] \leq \mathcal{O}\left(\frac{\sum_{j=0}^{\infty}\mathbb{E}\left[\widetilde{\mathsf{Piv}}(\tilde{T}_j)\right]}{L}\right). \tag{3.51}$$

Note that

$$\mathbb{P}\left[\text{There is an interaction}\right] = \sum_{\mathsf{w}_1,\mathsf{w}_2}\mathbb{P}\left[\text{The first interaction is between } \mathsf{w}_1,\mathsf{w}_2\right].$$

Recall the definition of the update times $(T_j)_{j\in\mathbb{Z}_+}$ of the pivotal set. Let $\mathcal{E}(j,\mathsf{w}_1,\mathsf{w}_2)$ be the event that T_j is the first time such that both w_1 and w_2 are in the pivotal set. Let $\mathcal{E}(j)$ be the event that there is no interaction up to time T_j, and let $\mathcal{E}(\mathsf{w}_1,\mathsf{w}_2)$ be the event that there is an interaction between $\mathsf{w}_1,\mathsf{w}_2$. Then

$$\mathbb{P}\left[\text{The first interaction is of } \mathsf{w}_1,\mathsf{w}_2\right] \leq \sum_{j=0}^{\infty}\mathbb{P}\left[\mathcal{E}(j,\mathsf{w}_1,\mathsf{w}_2)\cap\mathcal{E}(j)\cap\mathcal{E}(\mathsf{w}_1,\mathsf{w}_2)\right].$$

Let $\mathcal{F}_j$ be the σ-algebra generated by the Poisson processes up to time T_j. Then

$$\mathbb{P}\left[\mathcal{E}(j,\mathsf{w}_1,\mathsf{w}_2)\cap\mathcal{E}(j)\cap\mathcal{E}(\mathsf{w}_1,\mathsf{w}_2)\right]=\mathbb{E}\left[\mathbb{1}_{\mathcal{E}(j,\mathsf{w}_1,\mathsf{w}_2)}\mathbb{1}_{\mathcal{E}(j)}\mathbb{P}\left[\mathcal{E}(\mathsf{w}_1,\mathsf{w}_2)|\mathcal{F}_j\right]\right]$$

Let τ be the time such that a Glauber clock of w_1 or w_2 rings counted from T_j:

$$\tau:=\inf\left\{t>0:\text{a Glauber clock of }\mathsf{w}_1\text{ or }\mathsf{w}_2\text{ rings at time }T_j+t\right\}.$$

Conditionally on $\mathcal{F}_j$, on the event $\mathcal{E}(j,\mathsf{w}_1,\mathsf{w}_2)\cap\mathcal{E}(j)$, from time T_j onwards, w_1 and w_2 move as an IP(2) on the lattice, until one of their Glauber clocks rings at time $\tau\sim\exp(2\lambda)$. To have an interaction, the distance between w_1 and w_2 at time τ must be smaller than m. Therefore, by Lemma 3.24, conditionally on $\mathcal{F}_j$, on the event $\mathcal{E}(j,\mathsf{w}_1,\mathsf{w}_2)\cap\mathcal{E}(j)$,

$$\mathbb{P}\left[\mathcal{E}(\mathsf{w}_1,\mathsf{w}_2)|\mathcal{F}_j\right]\leq\beta/L,$$

for some constant β. We deduce that,

$$\mathbb{P}\left[\text{There is an interaction}\right]\leq\sum_{\mathsf{w}_1,\mathsf{w}_2}\sum_{j=0}^{\infty}\frac{\beta}{L}\mathbb{P}\left[\mathcal{E}(j,\mathsf{w}_1,\mathsf{w}_2)\cap\mathcal{E}(j)\right].$$

Note that on the event $\mathcal{E}(j,\mathsf{w}_1,\mathsf{w}_2)$, either w_1 or w_2 (or both of them) is born at time T_j. We denote by $\mathcal{E}(j,\mathsf{w}_i)$ the event that w_i is born at time T_j, $i\in\{1,2\}$. Hence

$$\sum_{\mathsf{w}_1,\mathsf{w}_2}\mathbb{1}_{\mathcal{E}(j,\mathsf{w}_1,\mathsf{w}_2)}\mathbb{1}_{\mathcal{E}(j)}\leq\sum_{i\in\{1,2\}}\sum_{\mathsf{w}_1,\mathsf{w}_2}\mathbb{1}_{\mathcal{E}(j,\mathsf{w}_1,\mathsf{w}_2)}\mathbb{1}_{\mathcal{E}(j)}\mathbb{1}_{\mathcal{E}(j,\mathsf{w}_i)}.$$

Moreover, note also that on the event $\mathcal{E}(j,\mathsf{w}_1,\mathsf{w}_2)$, $\mathsf{w}_2\in\mathsf{Piv}\left(T_j\right)$. Hence

$$\begin{aligned}\sum_{\mathsf{w}_1,\mathsf{w}_2}\mathbb{1}_{\mathcal{E}(j,\mathsf{w}_1,\mathsf{w}_2)}\mathbb{1}_{\mathcal{E}(j)}\mathbb{1}_{\mathcal{E}(j,\mathsf{w}_1)}&\leq\sum_{\mathsf{w}_1,\mathsf{w}_2}\mathbb{1}_{\{\mathsf{w}_2\in\mathsf{Piv}(T_j)\}}\mathbb{1}_{\mathcal{E}(j,\mathsf{w}_1)}\mathbb{1}_{\mathcal{E}(j)}\\&\leq\sum_{\mathsf{w}_1}\left|\mathsf{Piv}\left(T_j\right)\right|\mathbb{1}_{\mathcal{E}(j,\mathsf{w}_1)}\mathbb{1}_{\mathcal{E}(j)}\\&\leq(2m+1)\left|\mathsf{Piv}\left(T_j\right)\right|\mathbb{1}_{\mathcal{E}(j)},\end{aligned}$$

where all these inequalities are true almost surely. The last inequality is because at most $(2m+1)$ particles are born at time T_j. Note that, on the event $\mathcal{E}(j)$, $\mathsf{Piv}\left(T_j\right)=\widetilde{\mathsf{Piv}}(\tilde{T}_j)$. This implies that

$$\sum_{\mathsf{w}_1,\mathsf{w}_2}\mathbb{1}_{\mathcal{E}(j,\mathsf{w}_1,\mathsf{w}_2)}\mathbb{1}_{\mathcal{E}(j)}\leq2(2m+1)\left|\widetilde{\mathsf{Piv}}(\tilde{T}_j)\right|.$$

Taking expectation, we conclude that

$$\mathbb{P}\left[\text{There is an interaction}\right]\leq2(2m+1)\frac{\beta}{L}\sum_{j=0}^{\infty}\mathbb{E}\left[\widetilde{\mathsf{Piv}}(\tilde{T}_j)\right],$$

which finishes the proof. □

We finish this section by proving Lemma 3.26. The left-hand side is the sum of the expected sizes of the process $\left|\widetilde{\mathsf{Piv}}(\cdot)\right|$ at its update times. The intuition is that the inequality is true if we replace $\widetilde{\mathsf{Piv}}(\cdot)$ by a subcritical Galton-Watson process. The only problem is that $\left|\widetilde{\mathsf{Piv}}(\cdot)\right|$ is not Markovian. Nevertheless, it is "subcritical", and the idea is to enlarge the process $\left|\widetilde{\mathsf{Piv}}(\cdot)\right|$ to make it Markovian and then adapt the proof of the subcritical Galton-Watson case.

Proof of Lemma 3.26. Let t_0 be a number such that $\psi(t_0) < 1$, which exists thanks to Lemma 3.21. We rewrite the left-hand side as

$$\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\mathbb{E}\left[\left|\widetilde{\mathsf{Piv}}(\tilde{T}_j)\right|\mathbb{1}_{\left\{it_0\leq \tilde{T}_j<(i+1)t_0\right\}}\right].$$

Let the sequence $(B_i)_{i\in\mathbb{Z}_+}$ be defined recursively as follows.

$$\begin{aligned}B_0 &:= \widetilde{\mathsf{Piv}}(0),\\ B_i &:= \bigcup_{\mathsf{w}\in B_{i-1}}\widetilde{\mathsf{Piv}}(\mathsf{w}, it_0),\ \forall i\geq 1.\end{aligned}$$

It is not hard to see that $\widetilde{\mathsf{Piv}}(it_0)\subset B_i, \forall u\in\mathbb{Z}_+$, because given the IBP up to time it_0 and the spins of the particles in B_i, by definition of B_i, we can recursively determine the spins of the particles in B_{i-1}, and so on, up to time 0. It is also not hard to see that the sequence $(|B_i|)_{i\in\mathbb{Z}_+}$ is a subcritical Galton-Watson process, by its definition and the definition of t_0.

Let κ be the constant in Lemma 3.22. We claim that

$$\mathbb{E}\left[\sum_{j=0}^{\infty}\left|\widetilde{\mathsf{Piv}}(\tilde{T}_j)\right|\mathbb{1}_{\left\{it_0\leq \tilde{T}_j<(i+1)t_0\right\}}\right]\leq \frac{e^{\kappa t_0}}{2m}\mathbb{E}\left[|B_i|^2\right].$$

To see this, consider the IBP starting from B_i (at time it_0) up to time $(i+1)t_0$. We denote this process by $\mathtt{Gw}_i$. Let $(\mathcal{T}_j)_{j\in\mathbb{Z}_+}$ be the update times of $\mathtt{Gw}_i$. Then $\mathtt{Gw}_i(it_0) = B_i$, and

$$\mathbb{E}\left[\sum_{j=0}^{\infty}\left|\widetilde{\mathsf{Piv}}(\tilde{T}_j)\right|\mathbb{1}_{\left\{it_0\leq \tilde{T}_j<(i+1)t_0\right\}}\right]\leq \mathbb{E}\left[\sum_{j=0}^{\infty}\left|\mathtt{Gw}_i(\mathcal{T}_j)\mathbb{1}_{\{it_0\leq \mathcal{T}_j<(i+1)t_0\}}\right|\right].$$

This is because $\mathtt{Gw}_i$ can be thought of as an enlargement of $\widetilde{\mathsf{Piv}}(\cdot)$ in the time interval $[it_0, (i+1)t_0]$, where we first enlarge the set $\widetilde{\mathsf{Piv}}(it_0)$ to B_i, and then we do not remove any particle of the branching process starting from B_i.

Note that the size of $\mathtt{Gw}_i(\cdot)$ increases almost surely, hence

$$\left|\mathtt{Gw}_i(\mathcal{T}_j)\mathbb{1}_{\{it_0\leq \mathcal{T}_j<(i+1)t_0\}}\right|\leq \left|\mathtt{Gw}_i((i+1)t_0)\right|.$$

Moreover, at each update time $\mathcal{T}_j$, the size of $\mathtt{Gw}_i$ increases by 2m, and hence

$$\sum_{j=0}^{\infty} \mathbb{1}_{\{it_0 \le \mathcal{T}_j < (i+1)t_0\}} \le \frac{|\mathtt{Gw}_i((i+1)t_0)|}{2m}.$$

This implies

$$\sum_{j=0}^{\infty} \left|\mathtt{Gw}_i(\mathcal{T}_j)\mathbb{1}_{\{it_0 \le \mathcal{T}_j < (i+1)t_0\}}\right| \le \frac{|\mathtt{Gw}_i((i+1)t_0)|^2}{2m}.$$

By Lemma 3.22,

$$\mathbb{E}\left[|\mathtt{Gw}_i((i+1)t_0)|^2\right] \le e^{\kappa t_0}\mathbb{E}\left[|B_i|^2\right].$$

This proves the claim. Now, we can take the sum over i to conclude that

$$\mathbb{E}\left[\sum_{j=0}^{\infty}\left|\widetilde{\mathsf{Piv}}(\tilde{T}_j)\right|\right] \le \frac{e^{\kappa t_0}}{2m}\sum_{i=0}^{\infty}\mathbb{E}\left[|B_i|^2\right].$$

Since $(|B_i|)_{i\in\mathbb{Z}_+}$ is a subcritical Galton-Watson process, we have

$$\sum_{i=0}^{\infty}\mathbb{E}\left[|B_i|^2\right] = \mathcal{O}\left(\mathbb{E}\left[|B_0|^2\right]\right) = \mathcal{O}\left(|E|^2\right).$$

This finishes the proof. □

5 Application of the dual coupling

Now we can prove Theorem 3.3, Theorem 3.4, and the lower bound in Theorem 3.2.

Proposition 3.27 (Replacement lemma)**.** *Suppose that $X_0 \sim \mathfrak{R}_{\rho_0}^{\otimes\Lambda_L}$, for some $\rho_0 \in (-1,1)$. Then, for any subset $E \subset \Lambda_L$,*

$$\sup_{t\ge 0} \mathrm{d}_{TV}\left(Law\left(X_t(E)\right), \mathfrak{R}_{\rho(t)}^{\otimes|E|}\right) \le \frac{\beta|E|^2}{L}, \tag{3.52}$$

where $\rho(\cdot)$ is the solution of equation (3.9), and β is the constant in Proposition 3.17.

Remark 10. The proposition above is of independent interest. It shows that the joint distribution of $o\left(\sqrt{L}\right)$ arbitrary sites is close to that of a product of i.i.d. Rademacher. This provides sharp estimates on the correlation functions, allowing us to derive the correct lower bound on the mixing times using distinguishing statistics.

Proposition 3.27 is just a direct corollary of Proposition 3.17.

Proof. Consider the coupling of BEP and IBP starting from E in Subsection 4.4. We generate the spins at time t by $\mathfrak{s}(W_t) \sim \mathfrak{R}_{\rho_0}^{\otimes W_t}$ and $\tilde{\mathfrak{s}}(\widetilde{W}_t) \sim \mathfrak{R}_{\rho_0}^{\otimes\widetilde{W}_t}$. Conditionally on the success of the coupling until time t, we couple the spins on $\mathsf{Piv}(E,t)$ and $\widetilde{\mathsf{Piv}}(E,t)$ such that $\mathfrak{s}_{\mathsf{w}} = \tilde{\mathfrak{s}}_{\mathsf{w}},\ \forall \mathsf{w} \in \mathsf{Piv}(E,t)$. Then Observation 3.16 implies that

$$\mathfrak{s}_{\mathsf{w}} = \tilde{\mathfrak{s}}_{\mathsf{w}},\ \forall \mathsf{w} \in E.$$

Recall that, by Proposition 3.12,

$$X_t(E) \overset{d}{=} (\mathfrak{s}_{\mathsf{w}})_{\mathsf{w}\in E},$$

and by Lemma 3.18,

$$(\tilde{\mathfrak{s}}_{\mathsf{w}})_{\mathsf{w}\in E} \overset{d}{=} \mathfrak{R}_{\rho(t)}^{\otimes E}.$$

Hence

$$\begin{aligned} \mathrm{d_{TV}}\left(\mathrm{Law}\left(X_t(E)\right), \mathfrak{R}_{\rho(t)}^{\otimes|E|}\right) &\le \mathbb{P}\left[(\mathfrak{s}_{\mathsf{w}})_{\mathsf{w}\in E} \neq (\tilde{\mathfrak{s}}_{\mathsf{w}})_{\mathsf{w}\in E}\right] \\ &\le \mathbb{P}\left[\text{The coupling is not successful until time } t\right] \\ &\le \frac{\beta|E|^2}{L}. \end{aligned}$$

This finishes our proof. □

Corollary 3.28 (1-point and 2-point correlation functions)**.** *Under the hypothesis of Proposition 3.27, for any $u, v \in \Lambda_L$, for any $t \ge 0$,*

$$|\mathbb{E}\left[X_t(u)\right] - \rho(t)| \le 2\beta/L, \tag{3.53}$$

$$|\mathrm{Cov}\left(X_t(u), X_t(v)\right)| \le 12\beta/L. \tag{3.54}$$

Proof. Let $u, v \in \Lambda_L$, and let ξ_1, ξ_2 be i.i.d. Rademacher $\mathfrak{R}_{\rho(t)}$. Let β be the constant in Proposition 3.17.

1. Applying Proposition 3.27 with $E := \{u\}$, we deduce that there exists a coupling of $X_t(u)$ and ξ_1 such that

$$\mathbb{P}\left[X_t(u) \neq \xi_1\right] \le \frac{\beta}{L}.$$

Hence,

$$\begin{aligned} |\mathbb{E}\left[X_t(u)\right] - \rho(t)| &= |\mathbb{E}\left[X_t(u) - \xi_1\right]| \\ &= \left|\mathbb{E}\left[(X_t(u) - \xi_1)\mathbb{1}_{\{X_t(u)\neq\xi_1\}}\right]\right| \\ &\le \mathbb{E}\left[|X_t(u) - \xi_1|\,\mathbb{1}_{\{X_t(u)\neq\xi_1\}}\right] \\ &\le 2\mathbb{P}\left[X_t(u) \neq \xi_1\right] \\ &\le 2\beta/L. \end{aligned}$$

2. Similarly, applying Proposition 3.27 with $E := \{u, v\}$, we deduce that there exists a coupling of $(X_t(u), X_t(v))$ and (ξ_1, ξ_2) such that if we let $\mathcal{E} := \{(X_t(u), X_t(v)) \neq (\xi_1, \xi_2)\}$, then $\mathbb{P}[\mathcal{E}] \leq \frac{4\beta}{L}$. Hence

$$\begin{aligned}|\mathrm{Cov}(X_t(u), X_t(v))| &= |\mathbb{E}[X_t(u)X_t(v)] - \mathbb{E}[X_t(u)]\,\mathbb{E}[X_t(v)]| \\ &\leq |\mathbb{E}[X_t(u)X_t(v)] - \mathbb{E}[\xi_1\xi_2]| + |\mathbb{E}[\xi_1\xi_2] - \mathbb{E}[X_t(u)]\,\mathbb{E}[X_t(v)]|\,.\end{aligned}$$

The first term in the last expression is

$$|\mathbb{E}[(X_t(u)X_t(v) - \xi_1\xi_2)\mathbb{1}_{\mathcal{E}}]| \leq 2\mathbb{P}[\mathcal{E}] \leq 8\beta/L.$$

The second term does not exceed

$$|\mathbb{E}[\xi_1]|\,|\mathbb{E}[\xi_2] - \mathbb{E}[X_t(v)]| + |\mathbb{E}[X_t(v)]|\,|\mathbb{E}[\xi_1] - \mathbb{E}[X_t(u)]| \leq 2\beta/L + 2\beta/L = 4\beta/L.$$

This finishes our proof. □

We are ready to prove the lower bound on the mixing times. We denote by $(X_t^+)_{t\geq 0}$ the Glauber-Exclusion process starting from all-plus and recall that $\rho_+(\cdot)$ is the solution of (3.9) with the initial condition $\rho_+(0) = 1$.

Proof of the lower bound. The proof uses the classical distinguishing statistics method (see Section 7.3 in [59]). Our distinguishing statistic is the total magnetization $\sum_{u\in\Lambda_L} x(u)$. By abuse of notation, let $X_\infty \sim \pi$ (so that for any initial configuration, $X_t \xrightarrow[t\to\infty]{d} X_\infty$). Let β be defined as in Proposition 3.17. From (3.53), we deduce that

$$\left|\mathbb{E}\left[\sum_{u\in\Lambda_L} X_t^+(u)\right] - L\rho_+(t)\right| \leq 2\beta,$$

and by letting $t \to \infty$,

$$\left|\mathbb{E}\left[\sum_{u\in\Lambda_L} X_\infty(u)\right] - L\rho_*\right| \leq 2\beta.$$

From (3.54), we see that

$$\begin{aligned}\mathrm{Var}\left[\sum_{u\in\Lambda_L} X_t^+(u)\right] &= \sum_{u\in\Lambda_L} \mathrm{Var}\left[X_t^+(u)\right] + \sum_{u\neq v} \mathrm{Cov}\left(X_t^+(u), X_t^+(v)\right) \\ &\leq L + L(L-1) \times 12\beta/L \\ &< (12\beta + 1)L.\end{aligned}$$

By letting t tend to infinity, we get

$$\mathrm{Var}\left[\sum_{u\in\Lambda_L} X_\infty(u)\right] \le (12\beta+1)L.$$

So by Proposition 7.9 in [59],

$$\mathrm{d_{TV}}\left(\mathbb{P}\left[X_t^+\in\cdot\right],\pi\right) \ge 1-8\frac{\max\left\{\mathrm{Var}\left[\sum_{u\in\Lambda_L}X_t^+(u)\right],\mathrm{Var}\left[\sum_{u\in\Lambda_L}X_\infty(u)\right]\right\}}{\left(\mathbb{E}\left[\sum_{u\in\Lambda_L}X_t^+(u)\right]-\mathbb{E}\left[\sum_{u\in\Lambda_L}X_\infty(u)\right]\right)^2}.$$

Let $t := \dfrac{\log L}{-2R(\rho_*)} - \kappa$ for some constant κ that we will choose later. Let κ_1 be a constant satisfying Lemma 3.21. Then

$$\begin{aligned}\rho_+(t)-\rho_* &\ge e^{R(\rho_*)t-\kappa_1}\\ &= \frac{e^{-R(\rho_*)\kappa-\kappa_1}}{\sqrt{L}}.\end{aligned}$$

Therefore, when L is large enough,

$$\begin{aligned}\mathbb{E}\left[\sum_{u\in\Lambda_L}X_t^+(u)\right]-\mathbb{E}\left[\sum_{u\in\Lambda_L}X_\infty(u)\right] &\ge L(\rho_+(t)-\rho_*)-4\beta\\ &= \sqrt{L}e^{-R(\rho_*)\kappa-\kappa_1}-4\beta.\\ &> \frac{1}{2}\sqrt{L}e^{-R(\rho_*)\kappa-\kappa_1}.\end{aligned}$$

Therefore,

$$8\frac{\max\left\{\mathrm{Var}\left[\sum_{u\in\Lambda_L}X_t^+(u)\right],\mathrm{Var}\left[\sum_{u\in\Lambda_L}X_\infty(u)\right]\right\}}{\left(\mathbb{E}\left[\sum_{u\in\Lambda_L}X_t^+(u)\right]-\mathbb{E}\left[\sum_{u\in\Lambda_L}X_\infty(u)\right]\right)^2} < \frac{32(12\beta+1)}{e^{-2R(\rho_*)\kappa-2\kappa_1}}.$$

Note that as $R'(\rho_*)<0$, for κ large enough, the last expression is smaller than $1-\epsilon$. Hence

$$\mathrm{d_{TV}}\left(\mathbb{P}\left[X_t^+\in\cdot\right],\pi\right)\ge\epsilon,$$

and therefore,

$$\mathrm{t_{mix}}(\epsilon) > t = \frac{\log L}{-2R(\rho_*)}-\kappa,$$

which finishes our proof. □

Remark 11. The result in higher dimensions is proved similarly. One can first show a version of Proposition 3.17 for higher dimensions, where the bound on the right-hand side is replaced by $\mathcal{O}\left(\dfrac{|E|^2\log L}{L^2}\right)$ for dimension $d=2$, and by $\mathcal{O}\left(\dfrac{|E|^2}{L^2}\right)$ for $d\ge3$, and then propagates these changes. The only modification needed in the proof is to use the anticoncentration in Lemma 3.24 for the appropriate dimension.

Now we prove the lower bound in Theorem 3.3 and the upper bound in Theorem 3.4. First, we need the following lemma.

Lemma 3.29. *Consider a BEP starting with one particle* w. *Then, there exist constants* κ *and* β *such that for any* $t \in \mathbb{R}_+$,

$$\mathbb{E}\left[|\mathsf{Piv}(\mathsf{w}, t)|\right] \le \psi(t) + e^{\kappa t}\frac{\beta}{\sqrt{L}}.$$

Proof of Theorem 3.3 and the lower bound in Theorem 3.4. Let $x \in \mathcal{X}$ be an arbitrary configuration. We want to compare $\mathbb{P}_x[X_t \in \cdot]$ and $\mathbb{P}_\pi[X_t \in \cdot]$ for some $t > 0$. We use the same BEP starting from $E = \Lambda_L$ to generate X_t^x and X_t^π, the processes starting from x and π. Let $\mathsf{w} \in \Lambda_L$ be arbitrary. We see that

$$\begin{aligned}
\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x[X_t \in \cdot], \mathbb{P}_\pi[X_t \in \cdot]\right) &\le \mathbb{P}\left[X_t^x \ne X_t^\pi\right] \\
&\le \sum_{u \in \Lambda_L} \mathbb{P}\left[X_t^x(u) \ne X_t^\pi(u)\right] \\
&\le \sum_{u \in \Lambda_L} \mathbb{P}\left[\mathsf{Piv}(u, t) \ne \emptyset\right] \\
&= L\mathbb{P}\left[\mathsf{Piv}(\mathsf{w}, t) \ne \emptyset\right] \\
&\le L\mathbb{E}\left[|\mathsf{Piv}(\mathsf{w}, t)|\right].
\end{aligned}$$

By Observation 3.14, it is not hard to see that,

$$\mathbb{E}\left[|\mathsf{Piv}(\mathsf{w}, mt)|\right] \le \mathbb{E}\left[|\mathsf{Piv}(\mathsf{w}, t)|\right]^m.$$

Hence, with the constant κ and β as in Lemma 3.29,

$$\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x[X_{mt} \in \cdot], \mathbb{P}_\pi[X_{mt} \in \cdot]\right) \le L\mathbb{E}\left[|\mathsf{Piv}(\mathsf{w}, mt)|\right] \le L\left(\psi(t) + e^{\kappa t}\frac{\beta}{\sqrt{L}}\right)^m. \tag{3.55}$$

Provided that $\left(\psi(t) + e^{\kappa t}\frac{\beta}{\sqrt{L}}\right) < 1$, we can take $m = \left\lceil \frac{\log L - \log \epsilon}{-\log\left(\psi(t) + e^{\kappa t}\frac{\beta}{\sqrt{L}}\right)} \right\rceil$ to make the right-hand side smaller than ϵ. This implies that

$$\mathrm{t_{mix}}(x; \epsilon) \le mt \le \left(\frac{\log L - \log \epsilon}{-\log\left(\psi(t) + e^{\kappa t}\frac{\beta}{\sqrt{L}}\right)} + 1\right) t,$$

for any t such that $\left(\psi(t) + e^{\kappa t}\frac{\beta}{\sqrt{L}}\right) < 1$. It remains to choose a suitable t that leads to our results. We will choose t such that $e^{\kappa t}\frac{\beta}{\sqrt{L}} \le \psi(t)$. For such a number t, with κ_1 the constant in Lemma 3.21, we have

$$\log\left(\psi(t)+e^{\kappa t}\frac{\beta}{\sqrt{L}}\right)\le \log\psi(t)+\log 2\le R'(\rho_*)t+\kappa_1+\log 2. \tag{3.56}$$

Therefore, if $R'(\rho_*)t+\kappa_1+\log 2<0$,

$$\begin{aligned}\mathrm{t_{mix}}(\epsilon) &\le \frac{\log L-\log\epsilon}{-(R'(\rho_*)t+\kappa_1+\log 2)}t+t\\ &= \frac{\log L-\log\epsilon}{-R'(\rho_*)+(\kappa_1+\log 2)/t}+t\\ &= \frac{\log L}{-R'(\rho_*)}+\mathcal{O}\left(\frac{\log L}{t}\right)+t.\end{aligned}$$

We can take $t=\sqrt{\log L}$, which satisfies all the conditions we impose on t, to finish the proof of the upper bound in Theorem 3.4 for $d=1$. The proof for higher dimensions is similar.

Now we prove the lower bound in Theorem 3.3. Recall that

$$\begin{aligned}\mathsf{gap} &= \lim_{t\to\infty} -\frac{1}{t}\log\max_{x\in\Omega} \mathrm{d_{TV}}\left(\mathbb{P}_x\left[X_t\in\cdot\right],\pi\right).\\ &= \lim_{m\to\infty} -\frac{1}{mt}\log\max_{x\in\Omega} \mathrm{d_{TV}}\left(\mathbb{P}_x\left[X_{mt}\in\cdot\right],\pi\right).\end{aligned}$$

On the other hand, equation (3.55) implies that, for any $t\ge 0$, $m\in\mathbb{Z}_+$,

$$\log\max_{x\in\Omega}\mathrm{d_{TV}}\left(\mathbb{P}_x\left[X_{mt}\in\cdot\right],\pi\right)\le \log L+m\log\left(\psi(t)+e^{\kappa t}\frac{\beta}{\sqrt{L}}\right).$$

Therefore,

$$\mathsf{gap}\ge \lim_{m\to\infty}-\frac{1}{mt}\left(\log L+m\log\left(\psi(t)+e^{\kappa t}\frac{\beta}{\sqrt{L}}\right)\right)=-\frac{1}{t}\log\left(\psi(t)+e^{\kappa t}\frac{\beta}{\sqrt{L}}\right).$$

Combining with (3.56), we deduce that, for any t such that $e^{\kappa t}\frac{\beta}{\sqrt{L}}\le\psi(t)$,

$$\mathsf{gap}\ge -R'(\rho_*)-\frac{\kappa_1+\log 2}{t}.$$

For $t=\dfrac{(\log L)/2-\log\beta-\kappa_1}{\kappa-R'(\rho_*)}$, $e^{\kappa t}\dfrac{\beta}{\sqrt{L}}=e^{R'(\rho_*t)-\kappa_1}\le\psi(t)$, thanks to Lemma 3.21. This choice of t gives us what we want to prove. □

Now we prove Lemma 3.29.

Proof of Lemma 3.29. We use the coupling of BEP and IBP starting from one particle w. We write $\mathsf{Piv}\,(t)$ for $\mathsf{Piv}\,(\mathsf{w},t)$ and $\widetilde{\mathsf{Piv}}(t)$ for $\widetilde{\mathsf{Piv}}(\mathsf{w},t)$. Let $\mathcal{E}$ denote the event that the coupling is successful until time t. Then, on $\mathcal{E}$, we have

$$|\mathsf{Piv}\,(t)|=\left|\widetilde{\mathsf{Piv}}(t)\right|.$$

Let β_1 be as in Proposition 3.17. Then $\mathbb{P}\left[\mathcal{E}^C\right] \leq \beta_1/L$, where $\mathcal{E}^C$ denotes the complement of $\mathcal{E}$. Hence

$$\begin{aligned}\mathbb{E}\left[|\mathsf{Piv}\,(t)\,|\right] &= \mathbb{E}\left[|\mathsf{Piv}\,(t)\,|(\mathbb{1}_{\mathcal{E}} + \mathbb{1}_{\mathcal{E}^C})\right] \\ &= \mathbb{E}\left[|\mathsf{Piv}\,(t)\,|\mathbb{1}_{\mathcal{E}}\right] + \mathbb{E}\left[|\mathsf{Piv}\,(t)\,|\mathbb{1}_{\mathcal{E}^C}\right].\end{aligned}$$

Note that

$$\mathbb{E}\left[|\mathsf{Piv}\,(t)\,|\mathbb{1}_{\mathcal{E}}\right] = \mathbb{E}\left[|\widetilde{\mathsf{Piv}}(t)|\mathbb{1}_{\mathcal{E}}\right] \leq \mathbb{E}\left[|\widetilde{\mathsf{Piv}}(t)|\right] = \psi(t).$$

Let κ_1 be the constant in Lemma 3.22. Note also that

$$\mathbb{E}\left[|\mathsf{Piv}\,(t)\,|\mathbb{1}_{\mathcal{E}^C}\right]^2 \leq \mathbb{E}\left[|W_t|\mathbb{1}_{\mathcal{E}^C}\right]^2 \leq \mathbb{E}\left[|W_t|^2\right]\mathbb{P}\left[\mathcal{E}^C\right] \leq \mathbb{E}\left[|\widetilde{W}_t|^2\right]\mathbb{P}\left[\mathcal{E}^C\right] \leq e^{\kappa_1 t} \times \frac{\beta_1}{L},$$

where we have used $|\mathsf{Piv}\,(t)\,| \leq |W_t|$ for the first inequality, the Cauchy-Schwarz inequality for the second inequality, Lemma 3.23 for the third inequality, and Lemma 3.22 for the last inequality. Therefore,

$$\mathbb{E}\left[|\mathsf{Piv}\,(t)\,|\right] \leq \psi(t) + e^{\kappa_1 t/2} \times \frac{\sqrt{\beta_1}}{\sqrt{L}}.$$

We can take $\kappa = \kappa_1/2$ and $\beta = \sqrt{\beta_1}$ to finish the proof. □

Now we prove the upper bound in Theorem 3.3. We will need the following results.

Lemma 3.30. *Consider a BEP starting with one particle* w. *Then*

$$\forall t > 0,\ \mathsf{gap} \leq -\frac{1}{t}\log \mathbb{P}\left[|\mathsf{Piv}\,(\mathsf{w}, t)\,| = 1\right]. \tag{3.57}$$

Lemma 3.31. *Consider a BEP starting with one particle* w. *Let* β *be as in Proposition 3.17. Then, there exists a constant* $\kappa > 0$ *such that*

$$\mathbb{P}\left[|\mathsf{Piv}\,(\mathsf{w}, t)\,| = 1\right] \geq \kappa e^{R'(\rho_*)t} - \frac{\beta}{L}. \tag{3.58}$$

With these results, we can prove the upper bound in Theorem 3.3.

Proof of the upper bound in Theorem 3.3. Lemma 3.30 and Lemma 3.31 together implies that, for any $t > 0$,

$$\mathsf{gap} \leq -\frac{1}{t}\log\left(\kappa e^{R'(\rho_*)t} - \frac{\beta}{L}\right),$$

for some positive constant κ. It remains to choose a suitable number t. We will choose t such that

$$\frac{\beta}{L} \leq \frac{\kappa}{2} e^{R'(\rho_*)t}. \tag{3.59}$$

For such a number t, we have

$$\begin{aligned}\mathsf{gap} &\leq -\frac{1}{t}\log\left(\frac{\kappa}{2}e^{R'(\rho_*)t}\right)\\ &= -R'(\rho_*) + \frac{\log 2 - \log\kappa}{t}.\end{aligned}$$

We can take $t = \dfrac{\log L + \log\kappa - \log 2 - \log\beta}{|R'(\rho_*)|}$, which satisfies (3.59), to finish the proof. □

We finish this section with the proofs of Lemma 3.30 and Lemma 3.31.

Proof of Lemma 3.30. The idea is that for a subcritical branching process, conditionally on the event of survival at time t, it is likely that the pivotal set contains exactly one particle. By a proof similar to that of Lemma 3.19, we can show that

$$\forall t \geq 0,\ \mathbb{E}\left[\frac{X_t^+(u) - X_t^-(u)}{2}\right] = \mathbb{P}\left[\mathsf{Piv}\left(\mathsf{w}, t\right) \neq \emptyset\right].$$

Therefore, by (1.3),

$$\begin{aligned}\mathsf{gap} &\leq \lim_{t\to\infty} -\frac{1}{t}\log\mathbb{E}\left[\frac{X_t^+(u) - X_t^-(u)}{2}\right]\\ &= \lim_{t\to\infty} -\frac{1}{t}\log\mathbb{P}\left[\mathsf{Piv}\left(\mathsf{w}, t\right) \neq \emptyset\right]\\ &= \lim_{m\to\infty} -\frac{1}{mt}\log\mathbb{P}\left[\mathsf{Piv}\left(\mathsf{w}, mt\right) \neq \emptyset\right].\end{aligned} \tag{3.60}$$

On the other hand, by recurrence, it is easy to prove that

$$\mathbb{P}\left[\mathsf{Piv}\left(\mathsf{w}, mt\right) \neq \emptyset\right] \geq \mathbb{P}\left[|\mathsf{Piv}\left(\mathsf{w}, t\right)| = 1\right]^m. \tag{3.61}$$

This is because, conditionally on time $(m-1)t$, on the event $|\mathsf{Piv}\left(\mathsf{w}, (m-1)t\right)| = 1$, the BEP starting from the unique element of $\mathsf{Piv}\left(\mathsf{w}, (m-1)t\right)$ has the same law as the BEP starting from w. The equations (3.60) and (3.61) together lead to what we want. □

Proof of Lemma 3.31. We only need to consider the case $t > 1$. Consider the coupling of the BEP and the IBP starting from w. Note that,

$$\begin{aligned}\mathbb{P}\left[|\mathsf{Piv}\left(\mathsf{w}, t\right)| = 1\right] &\geq \mathbb{P}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w}, t)| = 1, \text{the coupling is successful until infinity}\right]\\ &\geq \mathbb{P}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w}, t)| = 1\right] - \mathbb{P}\left[\text{the coupling is not successful until infinity}\right]\\ &\geq \mathbb{P}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w}, t)| = 1\right] - \frac{\beta}{L}.\end{aligned} \tag{3.62}$$

For $s \in \mathbb{R}_+$, let $\mathcal{G}_s$ be the σ-algebra generated by the IBP starting from w up to time s. Then, for any positive integer m,

$$\begin{aligned}\mathbb{P}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t)|=1\right] &\geq \mathbb{P}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t-1)|\neq\emptyset, |\widetilde{\mathsf{Piv}}(\mathsf{w},t-1)|\leq m, |\widetilde{\mathsf{Piv}}(\mathsf{w},t)|=1\right]\\ &= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{|\widetilde{\mathsf{Piv}}(\mathsf{w},t-1)|\neq\emptyset, |\widetilde{\mathsf{Piv}}(\mathsf{w},t-1)|\leq m, |\widetilde{\mathsf{Piv}}(\mathsf{w},t)|=1}\middle|\mathcal{G}_{t-1}\right]\right]\\ &= \mathbb{E}\left[\mathbb{1}_{|\widetilde{\mathsf{Piv}}(\mathsf{w},t-1)|\neq\emptyset, |\widetilde{\mathsf{Piv}}(\mathsf{w},t-1)|\leq m}\mathbb{P}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t)|=1\middle|\mathcal{G}_{t-1}\right]\right]. \end{aligned}\tag{3.63}$$

We claim that almost surely,

$$\mathbb{P}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t)|=1\middle|\mathcal{G}_{t-1}\right] \geq \kappa\alpha^{|\widetilde{\mathsf{Piv}}(\mathsf{w},t)|-1},\tag{3.64}$$

where

$$\begin{aligned}\kappa &= \mathbb{P}\left[\mathsf{w}\text{ does not receive any update in }[0,1]\right],\\ \alpha &= \min_{\xi\in\{-1,1\}}\mathbb{P}\left[\mathsf{w}\text{ receive an oblivious update in }[0,1], \tilde{\mathfrak{s}}_{\mathsf{w}}=\xi\right].\end{aligned}$$

We first prove the result using the claim. The equations (3.63) and (3.64) together imply

$$\mathbb{P}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t)|=1\right] \geq \kappa\alpha^{m-1}\mathbb{P}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t-1)|\neq\emptyset, |\widetilde{\mathsf{Piv}}(\mathsf{w},t-1)|\leq m\right].$$

Moreover,

$$\begin{aligned}\mathbb{P}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t-1)|\neq\emptyset, |\widetilde{\mathsf{Piv}}(\mathsf{w},t-1)|\leq m\right] &= \mathbb{P}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t-1)|\neq\emptyset\right]-\mathbb{P}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t-1)|>m\right]\\ &\geq \mathbb{P}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t-1)|\neq\emptyset\right]-\frac{\mathbb{E}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t-1)|\right]}{m}\\ &= \phi(t-1)-\frac{\psi(t-1)}{m}.\end{aligned}\tag{3.65}$$

Therefore,

$$\mathbb{P}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t)|=1\right] \geq \kappa\alpha^{m}\left(\phi(t-1)-\frac{\psi(t-1)}{m}\right).$$

By choosing m large enough and combining with Lemma 3.21, we get what we want. It remains to show the claim. Suppose that $\widetilde{\mathsf{Piv}}(\mathsf{w},t-1)=\{\mathsf{w}_1,\ldots,\mathsf{w}_k\}$. By the definition of pivotal set, there exist $\xi_2,\ldots,\xi_k\in\{-1,1\}$ such that $\tilde{F}_{\mathsf{w},t-1}(\tilde{\mathfrak{s}}_{\mathsf{w}_1},\xi_2,\ldots,\xi_k)$ is not a constant function. Let $\mathcal{E}_1$ be the event that w_1 does not receive any update in $[t-1,t]$, and for $2\leq i\leq k$, and let $\mathcal{E}_i$ be the event that w_i receive an oblivious update in $[t-1,t]$ and $\tilde{\mathfrak{s}}_{\mathsf{w}_i}=\xi_i$. Note that,

$$\mathbb{P}\left[|\widetilde{\mathsf{Piv}}(\mathsf{w},t)|=1\middle|\mathcal{G}_{t-1}\right] \geq \mathbb{P}\left[\bigcap_{i=1}^{k}\mathcal{E}_i\middle|\mathcal{G}_{t-1}\right].$$

Conditionally on $\mathcal{G}_{t-1}$, the branches starting from $\mathsf{w}_1,\ldots,\mathsf{w}_k$ are independent IBPs shifted by a time $t-1$. Therefore,

$$\mathbb{P}\left[\bigcap_{i=1}^{k}\mathcal{E}_i\middle|\mathcal{G}_{t-1}\right]=\prod_{i=1}^{k}\mathbb{P}\left[\mathcal{E}_i\middle|\mathcal{G}_{t-1}\right].$$

Note that

$$\mathbb{P}\left[\mathcal{E}_1\middle|\mathcal{G}_{t-1}\right]=\kappa,$$
$$\mathbb{P}\left[\mathcal{E}_i\middle|\mathcal{G}_{t-1}\right]\geq\alpha,\ \forall 2\leq i\leq k.$$

These equations imply the claim. □

6 The upper bound

In this section, we prove the upper bound in Theorem 3.2. For simplicity, we only give the proof for $d=1$, but we will comment on how to adapt the proof for $d=2$.

Recall that in the decomposition of the local flip-rate function $c(\cdot)$ in Proposition 3.9, $f_1\equiv 1$, $f_2\equiv -1$, so the updates given by f_1, f_2 are called *oblivious* because they can be carried out without looking at the spins of any sites. These updates will play a crucial role in our proof. It is more convenient to merge two oblivious deterministic updates into a single oblivious random update as follows. We allow the collection of marks to contain another type of mark: the refresh marks, which can appear on the sites in the space-time slab. We adapt the definition of the collection of marks accordingly. From now on, let

$$\bar{\rho}:=\frac{\lambda_1-\lambda_2}{\lambda_1+\lambda_2}. \tag{3.66}$$

We will use the following graphical construction.

Graphical Construction 2. Let the background process Ξ be defined as follows.

$$\Xi=\left((\Xi_u^{\text{exclusion}})_{0\leq u\leq L-1},(\Xi_u^{\text{refresh}})_{0\leq u\leq L-1},(\Xi_{i,u}^{\text{Glauber}})_{3\leq i\leq q,\,0\leq u\leq L-1}\right), \tag{3.67}$$

where $(\Xi_u^{\text{exclusion}})_{0\leq u\leq L-1},(\Xi_u^{\text{refresh}})_{0\leq u\leq L-1},(\Xi_{i,u}^{\text{Glauber}})_{3\leq i\leq q,\,0\leq u\leq L-1}$ are independent homogeneous Poisson processes, and

- $\Xi_u^{\text{exclusion}}$ is of intensity L^2, $0\leq u\leq L-1$,
- Ξ_u^{refresh} is of intensity $\lambda_1+\lambda_2$, $0\leq u\leq L-1$,
- $\Xi_{i,u}^{\text{Glauber}}$ is of intensity λ_i, $3\leq i\leq q$, $0\leq u\leq L-1$.

The process Ξ naturally defines a collection of marks $\mathcal{C}$ consisting of exclusion, refresh, and Glauber marks as follows.

- Whenever $\Xi_u^{\text{exclusion}}$ jumps, we put an exclusion mark on the edge $(u,u+1)$.

- Whenever Ξ_u^{refresh} jumps, we put a refresh mark at site u.
- Whenever $\Xi_{i,u}^{\text{Glauber}}$ jumps, we put a Glauber mark of type i at site u.

Given the collection of marks $\mathcal{C}$, an infinite sequence of independent Rademacher variables $\mathfrak{R}_{\bar{\rho}}$, and an initial condition $x_0 \in \mathcal{X}$, we construct the process, which we always denote by $(X_t^{x_0})_{t\geq 0}$, as follows.

- The process $(X_t^{x_0})_{t\geq 0}$ is a piecewise constant process starting at x_0 which can only jump when a mark appears.
- When we see an exclusion mark, say on the edge $(u, u+1)$, we make the transition $x \mapsto x^{u\leftrightarrow u+1}$ (exchange the spins of sites u and $u+1$).
- When we see a refresh mark, say at site u, we replace the $u-th$ coordinate of X by an independent $\mathfrak{R}_{\bar{\rho}}$ (randomize the spin at site u).
- When we see a Glauber mark, say of type i at site u, we make the transition $x \mapsto x^{u, f_i(x_{u+\cdot})}$ (update site u using the function f_i).

Then, when the collection of marks $\mathcal{C}$ is generated by Ξ and the Rademacher variables $\mathfrak{R}_{\bar{\rho}}$ are independent of Ξ, the process $(X_t^{x_0})_{t\geq 0}$ is a Markov process with the generator $\mathcal{L}_{GE}$ starting from x_0.

The grand coupling. We can use the same Poisson process Ξ and the same independent Rademacher variables $\mathfrak{R}_{\bar{\rho}}$ to construct the Glauber-Exclusion process from any configuration $x \in \mathcal{X}$. It is not hard to see that this coupling preserves order, i.e. , $x \leq x' \Rightarrow \forall t \geq 0,\ X_t^x \leq X_t^{x'}$.

Correlation marks associated with Glauber marks. It is convenient to associate each Glauber mark with a set of points as follows. Whenever we see a Glauber mark at site u, we put on each site in $B(u, m)$ a mark, and we call those marks the correlation marks associated with the aforementioned Glauber mark. These correlation marks indicate that the spins of those sites are (possibly) used in the corresponding Glauber updates.

We introduce another Markov process Z, which we call the *independent-site process*, constructed as follows.

Independent-site process. From the collection of marks generated by Ξ, let $(Z_t)_{t\geq 0}$ be the process taking values in $\mathcal{Z} := \{0,1\}^{\Lambda_L}$ constructed as follows. For any initial configuration $z_0 \in \mathcal{Z}$,

- The process $(Z_t^{z_0})_{t\geq 0}$ is a piecewise constant process starting at z_0 which can only jump when a mark appears.
- When we see an exclusion mark, say on the edge $(u, u+1)$, we make the transition $z \mapsto z^{u\leftrightarrow u+1}$.

- When we see a refresh mark, say at site u, we make the transition $z \mapsto z^{u,1}$.
- When we see a Glauber mark, says at site u, we remove all the sites with the associated correlation marks from Z, i.e. we make the transition $z \mapsto z \setminus B(u, m)$.

Here we have identified the elements of $\{0,1\}^{\Lambda_L}$ with the subsets of Λ_L. As its name suggests, the process Z tracks the positions of independent sites in X conditionally on Ξ. Roughly speaking, when a site is resampled according to an independent Rademacher $\mathfrak{R}_{\bar{\rho}}$ due to a refresh mark, it becomes independent of all other sites and hence is added to Z. On the other hand, when a site u is updated due to a Glauber mark, it creates a correlation with its neighbors, and all these neighbors are removed from Z. By abuse of notation, for any $z \subset \Lambda_L$, we denote by $\mathbb{P}_z[\cdot]$ the distribution of the process Z starting from z.

The grand coupling for Z. We can use the same background process Ξ to construct the process Z from any configuration $z \subset \Lambda_L$.

Remark 12 (Attractiveness of the process Z). For any $z_1, z_2 \subset \Lambda_L$ such that $z_1 \leq z_2$, if Z^{z_1} and Z^{z_2} are constructed using the grand coupling, then almost surely,

$$\forall t \geq 0,\ Z_t^{z_1} \leq Z_t^{z_2}.$$

Red, blue, and green regions. Let the red, blue, and green regions be defined as follows:

$$\begin{aligned} \mathsf{Blue}(t) &= Z_t^{\emptyset}, \\ \mathsf{Red}(t) &= \{u \in \Lambda_L | X_t^+(u) \neq X_t^-(u)\}, \\ \mathsf{Green}(t) &= \Lambda_L \setminus (\mathsf{Red}(t) \cup \mathsf{Blue}(t)). \end{aligned}$$

Here, $(X_t^+)_{t\geq 0}$ and $(X_t^-)_{t\geq 0}$ denote the process X starting from configurations all-plus and all-minus, respectively. To lighten the notation, we write $X_t^x(\mathsf{Blue})$ for $X_t^x(\mathsf{Blue}(t))$, and similarly for Red, and Green. Let $\mathcal{F}_t$ be the σ-algebra generated by $\Xi([0,t])$ and $(X_s^x(\Lambda_L \setminus \mathsf{Blue}(s)))_{x\in\mathcal{X}, 0\leq s\leq t}$. The interest in introducing red, blue, and green regions is the following.

Observation 3.32. *By construction, almost surely, conditionally on $\mathcal{F}_t$, for any $x \in \mathcal{X}$, $X_t^x(\mathsf{Blue}) \sim \mathfrak{R}_{\bar{\rho}}^{\otimes \mathsf{Blue}(t)}$ and is independent of $X_t^x(\mathsf{Red} \cup \mathsf{Green})$.*

From now on, for any subset $E \subset \Lambda_L$, we will use the shorthand notation $\nu_E := \mathfrak{R}_{\bar{\rho}}^{\otimes E}$. The meanings of the red, blue, and green regions above are as follows: the red region Red is the set of disagreements of X^+ and X^-. Conditionally on $\mathcal{F}_t$, the spins in the blue region Blue have the product measure $\nu_{\mathsf{Blue}(t)}$, independent of the spins on $\mathsf{Red} \cup \mathsf{Green}$, and the green region Green is the set of sites whose spins may present correlation but independent of the initial configuration. More precisely, notice that, as $X_t^+(\mathsf{Green}) = X_t^-(\mathsf{Green})$, by monotonicity, in the grand coupling, the spins on Green do not depend on the initial configuration anymore:

$$X_t^+(\mathsf{Green}) = X_t^-(\mathsf{Green}) = X_t^x(\mathsf{Green}),\ \forall x \in \mathcal{X}.$$

Therefore, we can safely denote it by $X_t(\mathsf{Green})$. Note also that $X_t(\mathsf{Green})$ and $X_t^x(\mathsf{Red})$, $x \in \mathcal{X}$, are $\mathcal{F}_t$ measurable.

The following lemma summarizes the discussion above.

Lemma 3.33 (The conditional law)**.** *Almost surely, for any $t \geq 0$, conditionally on $\mathcal{F}_t$,*

$$X_t^x \sim \nu_{\mathsf{Blue}(t)} \otimes \delta_{X_t^x(\mathsf{Red})} \otimes \delta_{X_t(\mathsf{Green})}, \tag{3.68}$$

where $\delta_{X_t(\mathsf{Green})}$ and $\delta_{X_t^x(\mathsf{Red})}$ are the Dirac measures on $\{-1,1\}^{\mathsf{Green}(t)}$ and $\{-1,1\}^{\mathsf{Red}(t)}$ which give all mass to $X_t(\mathsf{Green})$ and $X_t^x(\mathsf{Red})$, respectively.

6.1 Proof of the upper bound

First, note that, by the definition of Red,

$$|\mathsf{Red}(t)| = \frac{1}{2}\mathbb{E}\left[\sum_{u\in\Lambda_L}\Big(X_t^+(u) - X_t^-(u)\Big)\Big|\mathcal{F}_t\right].$$

This and Corollary 3.28 lead to the following lemma.

Lemma 3.34 (Decay of the red region)**.**

$$\mathbb{E}\left[|\mathsf{Red}(t)|\right] = \frac{L}{2}(\rho^+(t) - \rho^-(t)) + \mathcal{O}(1). \tag{3.69}$$

For any $\beta \in \mathbb{R}_{>0}$, let

$$\mathrm{BAD}_\beta = \left\{z \in \mathcal{Z} : \exists\, 0 \leq l \leq \left\lceil \frac{L}{\lceil \beta \log L\rceil}\right\rceil,\ z \cap [\![l\lceil \beta\log L\rceil, (l+1)\lceil\beta\log L\rceil - 1]\!] = \emptyset\right\},$$

where $[\![i,j]\!] := [i,j]\cap\mathbb{Z}$.

We will need the following results.

Lemma 3.35 (BAD is rarely visited)**.** *There exist constants $\beta_2, \beta_3 \in \mathbb{R}_{>0}$ such that, for any $t > \beta_3$,*

$$\max_{z\in\mathcal{Z}} \mathbb{P}_z\left[\exists s \in [t, t+1] : Z_s \in \mathrm{BAD}_{\beta_2}\right] = \mathcal{O}\left(1/L^6\right). \tag{3.70}$$

Proposition 3.36 ("Adjacent" measures quickly become close)**.** *Let $z \subset \Lambda_L$, let $u \in \Lambda_L \setminus z$, and let y be an arbitrary spin configuration on $\Lambda_L \setminus (z \cup \{u\})$. Let π_1, π_2 be the two probability distributions on $\mathcal{X}$ given by*

$$\pi_1 = \delta_1^u \otimes \nu_z \otimes \delta_y;\ \pi_2 = \nu_{z\cup\{u\}} \otimes \delta_y, \tag{3.71}$$

where δ_1^u denotes the Dirac measure on $\{-1,1\}^{\{u\}}$ which gives all mass to 1. Let β_2 be as in Lemma 3.35. Then there exists a constant κ such that

$$\|\mathbb{P}_{\pi_1}[X_1 \in \cdot] - \mathbb{P}_{\pi_2}[X_1\in\cdot]\|_{TV} \leq \kappa\left(\frac{\log L}{\sqrt{L}} + \sqrt{\mathbb{P}_{z\cup\{u\}}\left[\exists s\in[0,1] : Z_s \in \mathrm{BAD}_{\beta_2}\right]}\right). \tag{3.72}$$

The same result holds if we replace π_1 *by* $\delta^{u}_{-1} \otimes \nu_z \otimes \delta_y$.

Proposition 3.36 says that very quickly, one cannot distinguish two processes starting from the measure of the form π_2 and its perturbation π_1: the total variation distance drop from 1 to $\frac{\log L}{\sqrt{L}}$ in a time $\mathcal{O}(1)$. This is the result whose proof involves the information percolation framework and the idea from excursion theory.

With Lemma 3.35 and Proposition 3.36, we can now prove the upper bound in Theorem 3.2.

Proof of the upper bound. Let $t = t_1 + 1$, for some number t_1 that we will choose later. Let β_2, β_3 be as in Lemma 3.35. To lighten the notation, we write BAD for BAD_{β_2}. We will estimate directly $\mathrm{d_{TV}}\left(\mathbb{P}_{x_1}\left[X_t \in \cdot\right], \mathbb{P}_{x_2}\left[X_t \in \cdot\right]\right)$ for x_1, x_2 arbitrary in $\mathcal{X}$, as we do not have an explicit formula for the invariant measure π. We use the grand coupling to construct the Glauber-Exclusion process from any initial configuration. To lighten the notation, here we write Red, Blue, and Green for $\mathsf{Red}(t_1)$, $\mathsf{Blue}(t_1)$, and $\mathsf{Green}(t_1)$. Let $|\mathsf{Red}| = \mathcal{R}$ and $\mathsf{Red} = \{u_1, \ldots, u_{\mathcal{R}}\}$, which are all measurable w.r.t $\mathcal{F}_{t_1}$. Let the (random) distribution $\mathfrak{L}(x, j)$, for $x \in \mathcal{X}$, $0 \le j \le \mathcal{R}$, be defined by

$$\mathfrak{L}(x,j) = \nu_{\mathsf{Blue} \cup \{u_1,\ldots,u_j\}} \otimes \delta_{X^x_{t_1}(\mathsf{Red}\setminus\{u_1,\ldots,u_j\})} \otimes \delta_{X_{t_1}(\mathsf{Green})},$$

which means that $\mathfrak{L}(x,j)$ is the law of the random configuration obtained by resampling the spins on the sites in $\{u_1, \ldots, u_j\}$ of $X^x_{t_1}$ according to i.i.d. $\mathfrak{R}_{\bar{\rho}}$ independent of $X^x_{t_1}$. By Lemma 3.33, almost surely, $\mathfrak{L}(x,0)$ is exactly the conditional distribution of $X^x_{t_1}$ given $\mathcal{F}_{t_1}$, and

$$\mathfrak{L}(x,\mathcal{R}) = \nu_{\mathsf{Blue}\cup\mathsf{Red}} \otimes \delta_{X_{t_1}(\mathsf{Green})},$$

which does not depend on x. So we can write $\mathfrak{L}(\mathcal{R})$ for $\mathfrak{L}(x,\mathcal{R})$. The point here is that the sequence $(\mathfrak{L}(x,j))_{j=0,\ldots,\mathcal{R}}$ forms a "path" from $\mathfrak{L}(x,0)$ to $\mathfrak{L}(\mathcal{R})$. We see that, for any $x_1, x_2 \in \mathcal{X}$,

$$\begin{aligned}
\mathrm{d_{TV}}\left(\mathbb{P}_{x_1}\left[X_t \in \cdot\right], \mathbb{P}_{x_2}\left[X_t \in \cdot\right]\right) &\le \mathbb{E}\left[\|\mathbb{P}_{x_1}\left[X_t \in \cdot | \mathcal{F}_{t_1}\right] - \mathbb{P}_{x_2}\left[X_t \in \cdot | \mathcal{F}_{t_1}\right]\|_{TV}\right] \\
&= \mathbb{E}\left[\left\|\mathbb{P}_{\mathfrak{L}(x_1,0)}\left[X_1 \in \cdot\right] - \mathbb{P}_{\mathfrak{L}(x_2,0)}\left[X_1 \in \cdot\right]\right\|_{TV}\right] \\
&\le 2 \max_{x\in\mathcal{X}} \mathbb{E}\left[\left\|\mathbb{P}_{\mathfrak{L}(x,0)}\left[X_1 \in \cdot\right] - \mathbb{P}_{\mathfrak{L}(\mathcal{R})}\left[X_1 \in \cdot\right]\right\|_{TV}\right]. \qquad (3.73)
\end{aligned}$$

By the triangle inequality, for any $x \in \mathcal{X}$,

$$\left\|\mathbb{P}_{\mathfrak{L}(x,0)}\left[X_1 \in \cdot\right] - \mathbb{P}_{\mathfrak{L}(\mathcal{R})}\left[X_1 \in \cdot\right]\right\|_{TV} \le \sum_{j=0}^{\mathcal{R}-1} \left\|\mathbb{P}_{\mathfrak{L}(x,j)}\left[X_1 \in \cdot\right] - \mathbb{P}_{\mathfrak{L}(x,j+1)}\left[X_1 \in \cdot\right]\right\|_{TV}. \qquad (3.74)$$

By Proposition 3.36, there exists a constant κ such that

$$\begin{aligned}
&\left\|\mathbb{P}_{\mathfrak{L}(x,j)}\left[X_1 \in \cdot\right] - \mathbb{P}_{\mathfrak{L}(x,j+1)}\left[X_1 \in \cdot\right]\right\|_{TV} \\
&\le \kappa\left(\frac{\log L}{\sqrt{L}} + \sqrt{\mathbb{P}_{\mathsf{Blue}\cup\{u_1,\ldots,u_{j+1}\}}\left[\exists s \in [0,1] : Z_s \in \mathrm{BAD}\right]}\right) \\
&\le \kappa\left(\frac{\log L}{\sqrt{L}} + \sqrt{\mathbb{P}_{\mathsf{Blue}}\left[\exists s \in [0,1] : Z_s \in \mathrm{BAD}\right]}\right). \qquad (3.75)
\end{aligned}$$

We have used the attractiveness of Z and the fact that BAD is a decreasing subset of Λ_L in the last inequality. The equations (3.73), (3.74), (3.75) together imply

$$\begin{aligned} \mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_{x_1}\left[X_t \in \cdot\right], \mathbb{P}_{x_2}\left[X_t \in \cdot\right]\right) &\leq 2\mathbb{E}\left[\mathcal{R}\,\kappa\left(\frac{\log L}{\sqrt{L}} + \sqrt{\mathbb{P}_{\mathsf{Blue}}\left[\exists s \in [0,1] : Z_s \in \mathrm{BAD}\right]}\right)\right] \\ &\leq 2\frac{\kappa \log L}{\sqrt{L}}\mathbb{E}\left[\mathcal{R}\right] + 2\kappa\mathbb{E}\left[\mathcal{R}\sqrt{\mathbb{P}_{\mathsf{Blue}}\left[\exists s \in [0,1] : Z_s \in \mathrm{BAD}\right]}\right]. \end{aligned} \tag{3.76}$$

The latter term in the last expression above can be bounded by the Cauchy-Schwarz inequality:

$$\begin{aligned} \mathbb{E}\left[\mathcal{R}\sqrt{\mathbb{P}_{\mathsf{Blue}}\left[\exists s \in [0,1] : Z_s \in \mathrm{BAD}\right]}\right] &\leq \sqrt{\mathbb{E}\left[\mathcal{R}^2\right]\mathbb{E}\left[\mathbb{P}_{\mathsf{Blue}}\left[\exists s \in [0,1] : Z_s \in \mathrm{BAD}\right]\right]} \\ &\leq \sqrt{L^2\mathbb{P}_{\emptyset}\left[\exists s \in [t_1, t_1+1] : Z_s \in \mathrm{BAD}\right]}. \end{aligned}$$

Here, we have used the facts that $\mathcal{R} \leq L$, $\mathsf{Blue} = Z^{\emptyset}_{t_1}$ (by definition) and the Markov property of the process $Z^{\emptyset}$ at time t_1. Then, by Lemma 3.35, provided that $t_1 \geq \beta_3$,

$$\begin{aligned} \mathbb{E}\left[\mathcal{R}\sqrt{\mathbb{P}_{\mathsf{Blue}}\left[\exists s \in [0,1] : Z_s \in \mathrm{BAD}\right]}\right] &\leq \sqrt{L^2\mathcal{O}\left(1/L^6\right)} \\ &= \mathcal{O}\left(1/L^2\right). \end{aligned}$$

It remains to estimate the first term in (3.76). Thanks to Lemma 3.34 and Lemma 3.21, we can take $t_1 = \frac{\log L}{2|R'(\rho_*)|} + \frac{2\log\log L}{|R'(\rho_*)|}$ to make $\frac{\log L}{\sqrt{L}}\mathbb{E}\left[\mathcal{R}\right] = \mathcal{O}\left(1/\log L\right)$. Clearly, $t_1 \geq \beta_3$ when L is large enough. This implies that

$$\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_{x_1}\left[X_t \in \cdot\right], \mathbb{P}_{x_2}\left[X_t \in \cdot\right]\right) = \mathcal{O}\left(1/\log L\right) + \mathcal{O}\left(1/L^2\right) = \mathcal{O}\left(1/\log L\right),$$

for any $x_1, x_2 \in \mathcal{X}$. This finishes our proof.

□

The rest of the chapter is devoted to proving Lemma 3.35 and Proposition 3.36. In Subsection 6.2, we prove Lemma 3.35, and finally, in Subsection 6.3 and Subsection 6.4, we prove Proposition 3.36.

6.2 Proof of Lemma 3.35: the bad set is rarely visited

First, we need the following lemmas.

Lemma 3.37 (the bad set has small mass)**.** *There exist positive constants β_2 and β_3 such that for any $t \geq \beta_3$, and for any deterministic subset E of Λ_L of cardinality at least $\lceil \beta_2 \log L \rceil$,*

$$\max_{z \in \mathcal{Z}} \mathbb{P}_z\left[Z_t \cap E = \emptyset\right] = \mathcal{O}\left(1/L^{10}\right). \tag{3.77}$$

Lemma 3.38 (Inertia in a short time)**.** *Let τ_{out} be the time at which the process Z jumps out of the initial state Z_0:*

$$\tau_{out} = \inf\left\{t \geq 0 : Z_t \neq Z_0\right\}. \tag{3.78}$$

Then there is a constant $\beta > 0$ such that,

$$\min_{z\in\mathcal{Z}} \mathbb{P}_z\left[\tau_{out} > \frac{1}{L^3}\right] \geq \beta. \tag{3.79}$$

Now we can prove Lemma 3.35.

Proof of Lemma 3.35. Let z be an element of $\mathcal{Z}$. Let β_2, β_3 be as in Lemma 3.37. We write BAD for BAD_{β_2} to lighten the notation. We apply the union bound in Lemma 3.37 for the subsets $E_j = \llbracket j\left\lceil \beta_2 \log L\right\rceil, (j+1)\left\lceil \beta_2 \log L\right\rceil - 1\rrbracket$, $j \in \left\llbracket 0, \left\lceil \frac{L}{\lceil \beta_2 \log L\rceil}\right\rceil\right\rrbracket$, to obtain:

$$\forall t \geq \beta_3,\ \mathbb{P}_z\left[Z_t \in \mathrm{BAD}\right] = \mathcal{O}\left(1/L^9\right).$$

This implies that for any $t \geq \beta_3$,

$$\mathbb{E}_z\left[\int_t^{t+1+1/L^3} \mathbb{1}_{\{Z_s\in \mathrm{BAD}\}}\mathrm{d}s\right] = \int_t^{t+1+1/L^3} \mathbb{P}_z\left[Z_s \in \mathrm{BAD}\right]\mathrm{d}s = \mathcal{O}\left(1/L^9\right). \tag{3.80}$$

Let

$$\tau_{\mathrm{BAD}} := \inf\left\{t \geq 0 : Z_t \in \mathrm{BAD}\right\}.$$

By the Markov property,

$$\begin{aligned}
&\mathbb{E}_z\left[\int_t^{t+1+1/L^3} \mathbb{1}_{\{Z_s\in \mathrm{BAD}\}}\mathrm{d}s\right] \\
&= \mathbb{E}_z\left[\mathbb{E}_{Z_t}\left[\int_0^{1+1/L^3} \mathbb{1}_{\{Z_s\in \mathrm{BAD}\}}\mathrm{d}s\right]\right] \\
&\geq \mathbb{E}_z\left[\mathbb{E}_{Z_t}\left[\mathbb{1}_{\{\tau_{\mathrm{BAD}}\leq 1\}}\int_0^{1+1/L^3} \mathbb{1}_{\{Z_s\in \mathrm{BAD}\}}\mathrm{d}s\right]\right] \\
&\geq \mathbb{E}_z\left[\mathbb{E}_{Z_t}\left[\mathbb{1}_{\{\tau_{\mathrm{BAD}}\leq 1\}}\mathbb{E}_{Z_{\tau_{\mathrm{BAD}}}}\left[\int_0^{1+1/L^3-\tau_{\mathrm{BAD}}} \mathbb{1}_{\{Z_s\in \mathrm{BAD}\}}\mathrm{d}s\right]\right]\right].
\end{aligned} \tag{3.81}$$

Note that if $\tau_{\mathrm{BAD}} \leq 1$, then $\int_0^{1+1/L^3-\tau_{\mathrm{BAD}}} \mathbb{1}_{\{Z_s\in \mathrm{BAD}\}}\mathrm{d}s \geq \int_0^{1/L^3} \mathbb{1}_{\{Z_s\in \mathrm{BAD}\}}\mathrm{d}s$. On the other hand, with β as in Lemma 3.38, we have

$$\begin{aligned}\mathbb{E}_{Z_{\tau_{\mathrm{BAD}}}}\left[\int_0^{1/L^3}\mathbb{1}_{\{Z_s\in\mathrm{BAD}\}}\mathrm{d}s\right] &\geq \mathbb{E}_{Z_{\tau_{\mathrm{BAD}}}}\left[\frac{1}{L^3}\mathbb{1}_{\left\{\tau_{out}>\frac{1}{L^3}\right\}}\right] \\ &\geq \min_{z\in\mathcal{Z}}\frac{1}{L^3}\mathbb{P}_z\left[\tau_{out}>\frac{1}{L^3}\right] \\ &\geq \frac{\beta}{L^3}.\end{aligned}$$

This and (3.81) together imply

$$\mathbb{E}_z\left[\int_t^{t+1+1/L^3}\mathbb{1}_{\{Z_s\in\mathrm{BAD}\}}\mathrm{d}s\right] \geq \frac{\beta}{L^3}\mathbb{E}_z\left[\mathbb{E}_{Z_t}\left[\mathbb{1}_{\{\tau_{\mathrm{BAD}}\leq 1\}}\right]\right]. \tag{3.82}$$

This and (3.80) together imply

$$\mathbb{E}_z\left[\mathbb{P}_{Z_t}\left[\tau_{\mathrm{BAD}}\leq 1\right]\right] = \mathcal{O}\left(1/L^6\right). \tag{3.83}$$

Note that the left-hand side of the equation above is exactly $\mathbb{P}_z\left[\exists s\in[t,t+1]:Z_s\in\mathrm{BAD}\right]$. So this leads to what we want. □

We now prove Lemma 3.38 and Lemma 3.37.

Proof of Lemma 3.38. τ_{out} is at least the time that the first point of Ξ appears, which has distribution $\exp(L^2\times L+\lambda L)$, where $\lambda=\sum_{i=1}^q\lambda_i$. Hence

$$\begin{aligned}\mathbb{P}_z\left[\tau_{out}>\frac{1}{L^3}\right] &\geq \mathbb{P}\left[\exp(L^3+\lambda L)>\frac{1}{L^3}\right] \\ &= \exp\left(-\frac{L^3+\lambda L}{L^3}\right) \\ &= \exp(-1-\frac{\lambda}{L^2}).\end{aligned}$$

We can choose, for example, $\beta=e^{-1-\lambda}$. □

To finish this subsection, we prove Lemma 3.37.

Proof of Lemma 3.37. Recall that the process Z is attractive. Note also that for any $E\subset\Lambda_L$, the set $\{z\subset\Lambda_L|z\cap E=\emptyset\}$ is decreasing. So, we only need to prove the statement for the initial condition $z:=\emptyset$. To know whether $E\cap Z_t^{\emptyset}=\emptyset$, a natural way, again, is to trace the history of the set E backward in time. We do it by placing at each site in E at time t a particle labeled by the corresponding element in $\mathcal{A}$ and trace them back to time 0, given the collection of marks generated by Ξ.

- When a particle meets an exclusion mark, it jumps to the other endpoint of the corresponding edge.

- When a particle, say w, meet a correlation mark, it is removed, and $Z_t^{\emptyset}(\mathsf{w}) := 0$.
- When a particle, say w, meet a refresh mark, it is removed, and $Z_t^{\emptyset}(\mathsf{w}) := 1$.
- If a particle w reaches time 0, then $Z_t^{\emptyset}(\mathsf{w}) := 0$.

With those rules, we can construct $Z_t^{\emptyset}(E)$.

It is more convenient to generate the history in the forward direction. Let

$$\Xi^* = (\Xi^{\text{exclusion},*}, \Xi^{\text{Glauber},*}, \Xi^{\text{refresh},*})$$

be an independent copy of Ξ. In another copy of the space-time slab, we construct the process $(E_s)_{0\le s\le t}$ takings values in $\mathcal{Z}$ and the collection of marks $\mathcal{C}^*$ as follows.

- $E_0 := E$.
- Whenever $\Xi_u^{\text{exclusion},*}$ jumps, say at time s, we perform the following operations:
  1. Put an exclusion mark on the edge $(u, u+1)$,
  2. $E_s := E_{s-}^{u\leftrightarrow u+1}$.
- Whenever $\Xi_u^{\text{refresh},*}$ jumps, $0 \le u \le L-1$, say at time s,
  - If $u \le |E_{s-}|$, put a refresh mark on the $u-$th site in E_{s-} , say site u'. Then $E_s := E_{s-} \setminus \{u'\}$.
  - If $u > |E_{s-}|$, put a refresh mark on the $(u - |E_{s-}|)$-th site in $\Lambda_L \setminus E_{s-}$. Then $E_s := E_{s-}$.
- Whenever $\Xi_{i,u}^{\text{Glauber},*}$ jumps, say at time s,
  - if $u \le |B(E_{s-}, m)|$, put a Glauber mark on the $u-$th site in $B(E_{s-}, m)$, say site u', and put correlation marks on all the sites in $B(u', m)$. Then $E_s := E_{s-} \setminus B(u', m)$.
  - if $u > |B(E_{s-}, m)|$, put a Glauber mark on the $(u - |B(E_{s-}, m)|)-$th site in $\Lambda_L \setminus B(E_{s-}, m)$, say site u', and put correlation marks on all the sites in $B(u', m)$. Then $E_s := E_{s-}$.

By the homogeneity of the Poisson processes Ξ and Ξ^*, the collection of marks $\mathcal{C}^*|_{[0,t]}$ has the same distribution as $\mathcal{C}|_{[0,t]}$ viewed backward in time. In particular, the process $(E_s)_{0\le s\le t}$ has the same distribution as the set of particles alive when we track the history of the set E backward from time t. In particular, $\mathbb{P}\left[Z_t^{\emptyset} \cap E = \emptyset\right]$ is the same as the probability that no site is removed from $(E_s)_{0\le s\le t}$ due to the appearance of a refresh mark.

Here, we have rearranged the Poisson clocks so that the first clocks in $\Xi^{\text{refresh},*}$ indicate the refresh of the sites in $(E_s)_{0\le s\le t}$, and the first clocks in $\Xi^{\text{Glauber},*}$ indicate the Glauber updates of the sites in the neighbors of the process $(E_s)_{0\le s\le t}$. This makes our computation easier. More precisely, let $\mathcal{G}_s$ be the σ-algebra generated by Ξ^* up to time s. Let $\mathcal{T}_0 = 0$ and $(\tau_i, \mathcal{T}_i)_{i\geq 1}$ be defined recursively as follows.

$$\begin{aligned}\mathcal{T}_{i+1} = \inf \Big\{ t > \mathcal{T}_i : & \text{ a clock among } (\Xi_u^{\text{refresh},*})_{1\leq u \leq |E_{\mathcal{T}_i}|} \\ & \text{or } (\Xi_{u,i}^{\text{Glauber},*})_{3\leq i\leq q, 1\leq u\leq (2m+1)|E_{\mathcal{T}_i}|} \text{ rings at time } t \Big\} \\ \tau_{i+1} = \mathcal{T}_{i+1} - \mathcal{T}_i. &\end{aligned}$$

Then we have the following:

1. Conditionally on $\mathcal{G}_{\mathcal{T}_i}$, $\tau_{i+1} \sim \exp\left((\lambda_1+\lambda_2)\left|E_{\mathcal{T}_i}\right| + (\lambda-\lambda_1-\lambda_2)(2m+1)\left|E_{\mathcal{T}_i}\right|\right)$.

2. $$|E_s| = \left|E_{\mathcal{T}_i}\right|, \forall s \in [\mathcal{T}_i, \mathcal{T}_{i+1}); \ \left|E_{\mathcal{T}_{i+1}}\right| \geq \left|E_{\mathcal{T}_i}\right| - (2m+1),$$

   and hence

   $$\left|E_{\mathcal{T}_i}\right| \geq |E| - i(2m+1).$$

   This is because at most $(2m+1)$ sites are removed each time a Glauber mark appears.

3. $$\begin{aligned}&\mathbb{P}\left[\text{the ring at } \mathcal{T}_{i+1} \text{ is among } \Xi_1^{\text{refresh},*}, \ldots, \Xi_{|E_{\mathcal{T}_i}|}^{\text{refresh},*} | \mathcal{G}_{\mathcal{T}_i}\right] \\ &= \frac{\lambda_1+\lambda_2}{(\lambda_1+\lambda_2)+(\lambda-\lambda_1-\lambda_2)(2m+1)} \\ &=: 1-\alpha. \end{aligned} \tag{3.84}$$

Based on the arguments above,

$$\mathbb{P}\left[Z_t^{\emptyset} \cap E = \emptyset\right] = \mathbb{P}\left[\text{For any } \mathcal{T}_i \leq t, \text{ the ring at } \mathcal{T}_i \text{ is among } \Xi^{\text{Glauber},*}\right].$$

Note that, for any $l \in \mathbb{Z}_+$,

$$\begin{aligned}&\mathbb{P}\left[\text{For any } \mathcal{T}_i \leq t, \text{ the ring at } \mathcal{T}_i \text{ is among } \Xi^{\text{Glauber},*}\right] \\ &\leq \mathbb{P}\left[\mathcal{T}_l > t\right] + \mathbb{P}\left[\mathcal{T}_l \leq t, \text{the ring at } \mathcal{T}_1, \ldots, \mathcal{T}_l \text{ are among } \Xi^{\text{Glauber},*}\right].\end{aligned} \tag{3.85}$$

By equation (3.84),

$$\mathbb{P}\left[\mathcal{T}_l \leq t, \text{the ring at } \mathcal{T}_1, \ldots, \mathcal{T}_l \text{ are among } \Xi^{\text{Glauber},*}\right] \leq \alpha^l. \tag{3.86}$$

We can choose $l = 10\left\lceil \frac{\log L}{-\log\alpha}\right\rceil$ to make this term smaller than $\frac{1}{L^{10}}$. The bound for the first term on the right side of (3.85) is a consequence of the concentration of the sum $\tau_1 + \cdots + \tau_l$. Let $\beta := (\lambda_1+\lambda_2)+(\lambda-\lambda_1-\lambda_2)(2m+1)$. By Claims 1 and 2, conditionally on $\mathcal{G}_{\mathcal{T}_i}$, τ_{i+1} is stochastically dominated by $\exp\left(\beta(|E| - i(2m+1)\right)$. Set $\theta := (2m+1)l$, then provided that

$|E| > \theta + \frac{2\theta}{\beta}$,

$$\frac{\beta(|E|-(i-1)(2m+1))}{\beta(|E|-(i-1)(2m+1))-\theta} \leq 2,\ \forall 1 \leq i \leq l.$$

Therefore,

$$\begin{aligned}\mathbb{E}\left[e^{\theta \mathcal{T}_l}\right] &= \mathbb{E}\left[\mathbb{E}\left[e^{\theta(\mathcal{T}_{l-1}+\tau_l)} | \mathcal{G}_{\mathcal{T}_{l-1}}\right]\right] \\ &= \mathbb{E}\left[e^{\theta \mathcal{T}_{l-1}} \mathbb{E}\left[e^{\tau_l} | \mathcal{G}_{\mathcal{T}_{l-1}}\right]\right] \\ &\leq \mathbb{E}\left[e^{\theta \mathcal{T}_{l-1}} \times \frac{\beta(|E|-(l-1)(2m+1))}{\beta(|E|-(l-1)(2m+1))-\theta}\right] \\ &\leq 2\mathbb{E}\left[e^{\theta \mathcal{T}_{l-1}}\right].\end{aligned}$$

By induction, we deduce that

$$\mathbb{E}\left[e^{\theta \mathcal{T}_l}\right] \leq 2^l.$$

Hence, by Chernoff's bound,

$$\mathbb{P}\left[\mathcal{T}_l > t\right] \leq e^{-\theta t} 2^l = (2e^{-(2m+1)t})^l.$$

With l chosen as above, there exists β_3 such that when $t \geq \beta_3$, the last expression is smaller than $1/L^{10}$. We finish the proof by choosing β_2 such that $\beta_2 \log L \geq \theta + \theta/\beta$, which clearly exists when $\theta = (2m+1)l,\ l = 10\left\lceil \frac{\log L}{-\log q} \right\rceil$. □

6.3 First step of the proof of Proposition 3.36: reformulation

The rest of this chapter is devoted to proving Proposition 3.36. This is done in two steps. In this subsection, we do the first step: reduce our task to estimating the local time that a certain Markov process spends in a certain subset of its state space. In the next subsection, we use techniques from excursion theory to estimate this local time.

Without loss of generality, we can suppose $u = 0$. Throughout this subsection, let z, π_1, π_2 be as in the statement of Proposition 3.36. We recall Lemma 1.4 with the necessary adaptation for our state space.

Lemma 3.39 (Perturbation of a product measure)**.** *Let $\Omega = \{-1,1\}^n$. For each subset $S \subset [n]$, let φ_S be a distribution on $\{-1,1\}^S$. Let $\rho \in (-1,1)$, and let ν be the product measure $\mathfrak{R}_\rho^{\otimes n}$ on Ω. Let μ be the measure on Ω obtained by first sampling a subset $S \subset [n]$ via some measure $\tilde{\mu}$, and then, conditionally on S, generating independently the values on S via φ_S and the values on $[n] \setminus S$ via $\mathfrak{R}_p^{\otimes [n]\setminus S}$. Then*

$$4\mathrm{d}_{TV}(\mu,\nu)^2 \leq \left\| \frac{\mu}{\nu} - 1 \right\|_{L^2(\nu)}^2 \leq \mathbb{E}\left[\theta^{|S \cap S'|}\right] - 1,$$

where S, S' are i.i.d. with law $\tilde{\mu}$, and $\theta = \max\left\{\frac{2}{1+\rho}, \frac{2}{1-\rho}\right\}$.

We will use the following graphical construction. We will construct a coupling (X, Z) where X starts from π_1 or π_2, and Z starts from $z \cup \{0\}$. Recall that a site in Z is viewed as colored blue. From now on, we will color all sites not in Z black. The content of a site u at time t is its spin $X_t(u) \in \{-1, 1\}$ and its color $Z_t(u) \in \{0, 1\}$. From now on, we allow a collection of marks to contain new types of marks: blue and black marks, which can appear on the edges of Λ_L, and whose effects are explained later.

Graphical Construction 3. Let Ξ^{refresh} and Ξ^{Glauber} be defined as in Graphical Construction 2. They generate the refresh and Glauber marks as before. However, we replace the effect of the exclusion marks with that of the blue and black marks, as follows. Let $(\Xi_u^{\text{blue}})_{0 \le u \le L-1}$ and $(\Xi_u^{\text{black}})_{0 \le u \le L-1}$ be independent of $\Xi^{\text{refresh}}, \Xi^{\text{Glauber}}$ and defined as follows.

- $(\Xi_u^{\text{black}})_{0 \le u \le L-1}$ are independent Poisson clocks of intensity L^2 on the edges. When Ξ_u^{black} jumps, we put a black mark on the edge $(u, u+1)$.
- $(\Xi_u^{\text{blue}})_{0 \le u \le L-1}$ are independent Poisson clocks of intensity $2L^2$ on the edges. When Ξ_u^{blue} jumps, we put a blue mark on the edge $(u, u+1)$.

From now on, let $\Xi := (\Xi^{\text{refresh}}, \Xi^{\text{Glauber}}, \Xi^{\text{black}}, \Xi^{\text{blue}})$. Each realization of Ξ defines a collection of marks consisting of refresh, Glauber, black, and blue marks. We construct the process X from this collection of marks as follows.

- The effect of the refresh marks and Glauber marks are as before.
- Each time we see a black mark, say on the edge $(u, u+1)$, if at least one of the endpoints of the edge $(u, u+1)$ is black, we exchange the contents of those endpoints: $x \mapsto x^{u \leftrightarrow u+1}, z \mapsto z^{u \leftrightarrow u+1}$.
- Each time we see a blue mark, say on the edge $(u, u+1)$, if both endpoints are blue, we exchange their spins with probability $1/2$ (we decide by sampling an independent Bernoulli variable with parameter $1/2$).

With this construction, X is still a Glauber-Exclusion process, and Z still tracks the independent sites of X if X starts from π_2. The idea is to reveal Ξ but not the Rademacher variables $\mathfrak{R}_{\bar{\rho}}$ used to refresh the sites at the refresh marks and the Bernoulli variables used to make decisions at the blue marks.

The conditional distribution of X. Conditionally on $\Xi([0, 1])$, let $(u_1, \tau_1), (u_2, \tau_2), \ldots, (u_k, \tau_k)$ be the positions of the correlation marks that appear on the blue sites, i.e. $\forall i,\ (u_i, \tau_i -) \in Z_{\tau_i -}$. In other words, they are the correlation marks that turn some blue sites black. We claim that, conditionally on Ξ, almost surely, there is a function Φ independent of the initial configuration of X such that

$$X_1 = \Phi(y, X_{\tau_1 -}(u_1), \ldots, X_{\tau_k -}(u_k), X_1(Z_1)). \tag{3.87}$$

The explanation for the formula above is as follows. At time 1, almost surely, each site is either black or blue. The spins of the blue sites are given by $X_1(Z_1)$. The spins of the black sites can be constructed from

1. The spins of the black sites at the beginning, i.e. y,
2. The trajectories of the black sites given by Ξ^{black} and the update functions given by Ξ^{Glauber}. All that information is encoded in the function Φ,
3. The spins of the blue sites used in the update functions at the Glauber marks. Those spins are exactly $X_{\tau_1-}(u_1), \ldots, X_{\tau_k-}(u_k)$.

Let $\mathfrak{L}_1$ (resp. $\mathfrak{L}_2$) be the distribution of $(X_{\tau_1-}(u_1), \ldots, X_{\tau_k-}(u_k), X_1(Z_1))$ when the Glauber-Exclusion process starts from π_1 (resp. π_2). We remark here that if we start with the distribution π_2, then $\mathfrak{L}_2 = \mathfrak{R}_{\bar{\rho}}^{\otimes k} \otimes \mathfrak{R}_{\bar{\rho}}^{\otimes Z_1}$, and $\mathfrak{L}_1$ is a perturbation of $\mathfrak{L}_2$. Now, we can take advantage of the information-percolation framework in [68]. Thanks to partially revealing the randomness of the graphical construction, we transform the problem into comparing a product law with its perturbation, which can be done using Lemma 3.39. More formally, we have the following lemma.

Lemma 3.40 (Projection)**.**

$$\|\mathbb{P}_{\pi_1}[X_1 \in \cdot] - \mathbb{P}_{\pi_2}[X_1 \in \cdot]\|_{TV} \leq \mathbb{E}\left[\|\mathfrak{L}_1 - \mathfrak{L}_2\|_{TV}\right]. \tag{3.88}$$

Proof. We have

$$\begin{aligned}
&\|\mathbb{P}_{\pi_1}[X_1 \in \cdot] - \mathbb{P}_{\pi_2}[X_1 \in \cdot]\|_{TV} \\
&\leq \mathbb{E}\left[\|\mathbb{P}_{\pi_1}[X_1 \in \cdot|\Xi([0,1])] - \mathbb{P}_{\pi_2}[X_1 \in \cdot|\Xi([0,1])]\|_{TV}\right] \\
&\leq \mathbb{E}\left[\|\mathfrak{L}_1 - \mathfrak{L}_2\|_{TV}\right],
\end{aligned}$$

where in the last inequality, we have use (3.87) and the definitions of $\mathfrak{L}_1, \mathfrak{L}_2$. □

It remains to estimate the right-hand side in equation (3.88). First, we introduce a random walk used to track the position of the perturbation.

The perturbation position. Let the random walk $(U_1(t))_{t\geq 0}$ taking values in $\Lambda_L \cup \{\dagger\}$ be constructed as follows.

- $U_1(0) = 0$ (the walk U_1 starts at site 0).
- Whenever a refresh mark or a correlation mark appears on the site occupied by U_1, the walk U_1 is killed (it becomes the cemetery state $\dagger$).
- Whenever a black mark appears, if U_1 is at one of the two endpoints of the corresponding edge and the other endpoint is black, then U_1 takes the edge (U_1 jumps to the other endpoint).

- Whenever a blue mark appears, if U_1 is at one of the endpoints of the mark, with probability 1/2 (decided by an independent Bernoulli variable with parameter 1/2), it attempts to take the edge, and it succeeds if the other endpoint is blue and fails if the other endpoint is black.

Let U_2 be an independent copy of U_1 conditionally on Ξ, i.e. the decisions of the two walks each time a blue mark appears are independent. The law of U_1 (or U_2) conditionally on Ξ is the averaging process, introduced by Aldous in [3], defined as follows.

The averaging process. Let $H = (H_t)_{t\geq 0}$ be the piece-wise constant process taking values in the space of function from Λ_L to $[0,1]$, constructed as follows:

- $H_0(0) = 1;\ H_0(u) = 0,\ \forall u \in \Lambda_L \setminus \{0\}$.
- When we see a refresh mark, say at site u, we make the transition $h \mapsto h^{u,0}$, where $h^{u,0}$ is obtained from h by replacing the values at u by 0, i.e.

$$\forall u' \in \Lambda_L,\ h^{u,0}(u') = h(u') \times \mathbb{1}_{\{u' \neq u\}}.$$

- When we see a Glauber mark, say at site u, we make the transition $h \mapsto h^{B(u,m),0}$, where $h^{B(u,m),0}$ is obtained from h by replacing the values in $B(u,m)$ by 0, i.e.

$$\forall u' \in \Lambda_L,\ h^{B(u,m),0}(u') = h(u') \times \mathbb{1}_{\{u' \notin B(0,m)\}}.$$

- When a black mark appears, say on the edge $(u, u+1)$, if the process Z makes the transition $z \mapsto z^{u\leftrightarrow u+1}$, we make the transition $h \mapsto h^{u\leftrightarrow u+1}$, where $h^{u\leftrightarrow u+1}$ is obtained from h by exchanging the values at sites u and $u+1$, i.e.

$$h^{u\leftrightarrow u+1}(u') = h(u') \times \mathbb{1}_{\{u' \notin \{u,u+1\}\}} + h(u) \times \mathbb{1}_{\{u'=u+1\}} + h(u+1) \times \mathbb{1}_{\{u'=u\}}.$$

- When a blue mark appears, say at site u, if both sites u and $u+1$ are blue, we replace $h(u)$ and $h(u+1)$ by their average $\dfrac{h(u)+h(u+1)}{2}$, i.e. $h \mapsto \dfrac{h + h^{u\leftrightarrow u+1}}{2}$.

The averaging process is an interesting Markov process by itself. We refer the readers to [84, 13] for recent development. By construction, $(H_s)_{s\in[0,t]}$ is measurable w.r.t $\Xi([0,t])$. Moreover,

$$H_t(u) = \mathbb{P}\left[U_1(t) = u | \Xi\right],\ \forall i \in \Lambda_L, \forall t \geq 0,$$

and in a particular

$$H_t(u)^2 = \mathbb{P}\left[U_1(t) = U_2(t) = u | \Xi\right]. \tag{3.89}$$

Recall that $\mathsf{dist}(\cdot,\cdot)$ denotes the shortest-path distance on Λ_L. By convention, we define

$$\mathsf{dist}(u, \dagger) := \mathsf{dist}(\dagger, u) := \infty,\ \forall u \in \Lambda_L \cup \{\dagger\}.$$

The purpose of this subsection is to prove the following result.

Lemma 3.41 (Comparision with local time)**.** *There exists a constant β such that*

$$\|\mathbb{P}_{\pi_1}[X_1 \in \cdot] - \mathbb{P}_{\pi_2}[X_1 \in \cdot]\|_{TV}^2 \le \beta \mathbb{E}\left[\int_0^1 \mathbb{1}_{\{\mathsf{dist}(U_1(s),U_2(s))=0\}} \mathrm{d}s\right]. \tag{3.90}$$

Thanks to Lemma 3.40, we only need to show the following results to prove Lemma 3.41.

Lemma 3.42 (Conditional upper bound by the Averaging process)**.**

$$\|\mathfrak{L}_1 - \mathfrak{L}_2\|_{TV}^2 \le \left(\max\left\{\frac{2}{1+\bar\rho}, \frac{2}{1-\bar\rho}\right\} - 1\right)\left(\sum_{j=1}^k H_{\tau_j-}^2(i_j) + \|H_1\|_2^2\right). \tag{3.91}$$

Lemma 3.43. *There exist a constant β such that*

$$\mathbb{E}\left[\sum_{j=1}^k H_{\tau_j-}^2(i_j) + \|H_1\|_2^2\right] \le \beta \mathbb{E}\left[\int_0^1 \mathbb{1}_{\{\mathsf{dist}(U_1(s),U_2(s))=0\}} \mathrm{d}s\right]. \tag{3.92}$$

It remains to show Lemma 3.42 and Lemma 3.43. We first show Lemma 3.42.

Proof of Lemma 3.42. By construction, $\mathfrak{L}_2$ is a product of $\mathfrak{R}_{\bar\rho}$, and $\mathfrak{L}_1$ is a perturbed version of $\mathfrak{L}_2$, and the walk U_1 tracks the perturbation position. Let

$$S_1 := \{(u_j, \tau_j-) | j \in [k], U_1(\tau_j-) = u_j\} \cup \{(u,1) | u \in Z_1, U_1(1) = u\}.$$

In words, S_1 contains places perturbed in the vector $(X_{\tau_1-}(u_1), \dots, X_{\tau_k-}(u_k), X_1(Z_1))$. Note that the walk U_1 is killed when it meets a correlation or refresh mark. Hence $|S_1| \le 1$. Let S_2 be defined as S_1 when we replace U_1 by U_2. Then, by Lemma 3.39,

$$\|\mathfrak{L}_1 - \mathfrak{L}_2\|_{TV}^2 \le \mathbb{E}\left[\theta^{|S_1 \cap S_2|} - 1 | \Xi\right], \tag{3.93}$$

where $\theta = \max\left\{\frac{2}{1+\bar\rho}, \frac{2}{1-\bar\rho}\right\}$. Note that, as $|S_1 \cap S_2| \le 1$,

$$\begin{aligned}
&\mathbb{E}\left[\theta^{|S_1 \cap S_2|} - 1 | \Xi\right] \\
&= \mathbb{E}\left[(\theta - 1)\,|S_1 \cap S_2| \,\middle|\, \Xi\right] \\
&= (\theta - 1)\left(\left(\sum_{j=1}^k \mathbb{P}\left[U_1(\tau_j-) = U_2(\tau_j-) = u_j | \Xi\right]\right) + \left(\sum_{u \in \Lambda_L} \mathbb{P}\left[U_1(1) = U_2(1) = u | \Xi\right]\right)\right) \\
&= (\theta - 1)\left(\sum_{j=1}^k H_{\tau_j-}^2(u_j) \times + \|H_1\|_2^2\right),
\end{aligned}$$

which is what we want. □

Now we show Lemma 3.43.

Proof of Lemma 3.43. Recall that $(u_1, \tau_1), \ldots, (u_k, \tau_k)$ are the correlation marks that turn some blue sites green up to time 1. Let $(u_i, \tau_i)_{i \geq k+1}$ be the correlation marks that turn some blue sites green from time 1 onwards. Let us define

$$F_t := \sum_{j=1}^{\infty} H^2_{\tau_j -}(u_j) \mathbb{1}_{\{\tau_j \leq t\}}. \tag{3.94}$$

Then $(Z_t, H_t, F_t)_{t \geq 0}$ is a Markov process. In particular, let $\mathcal{L}$ be the generator corresponding to this Markov process. Then for any function φ on the state space of this process,

$$\begin{aligned}
&\mathcal{L}\varphi(z,h,f) \\
&= \sum_{u=0}^{L-1} L^2 \left(\varphi(z^{u\leftrightarrow u+1}, h^{u\leftrightarrow u+1}, f) - \varphi(z,h,f)\right)\left(\mathbb{1}_{\{z(u)=0\}} \vee \mathbb{1}_{\{z(u+1)=0\}}\right) \\
&\quad + \sum_{u=0}^{L-1} 2L^2 \left(\varphi\left(z, \frac{h + h^{u\leftrightarrow u+1}}{2}, f\right) - \varphi(z,h,f)\right) \mathbb{1}_{\{z(u)=z(u+1)=1\}} \\
&\quad + \sum_{u=0}^{L-1} (\lambda_1 + \lambda_2) \left(\varphi(z^{u,1}, h^{i,0}, f) - \varphi(z,h,f)\right) \\
&\quad + \sum_{i=3}^{q} \sum_{u=0}^{L-1} \lambda_i \left(\varphi\left(z \setminus B(u,m), h^{B(u,m),0}, f + \sum_{u' \in B(0,m)} h^2(u') \mathbb{1}_{\{z(u')=1\}}\right) - \varphi(z,h,f)\right).
\end{aligned}$$

Let φ be the projection onto the third coordinate. Then

$$\begin{aligned}
\mathcal{L}\varphi(z,h,f) &= \sum_{i=3}^{q} \sum_{u=0}^{L-1} \lambda_i \sum_{u' \in B(u,m)} h^2(u') \mathbb{1}_{\{z(u')=1\}} \\
&= (\lambda - \lambda_1 - \lambda_2) \sum_{u=0}^{L-1} \sum_{u'=0}^{L-1} h^2(u') \mathbb{1}_{\{z(u')=1\}} \mathbb{1}_{\{u' \in B(u,m)\}} \\
&= (\lambda - \lambda_1 - \lambda_2) \sum_{u'=0}^{L-1} \sum_{u=0}^{L-1} h^2(u') \mathbb{1}_{\{z(u')=1\}} \mathbb{1}_{\{u \in B(u',m)\}} \\
&\leq (\lambda - \lambda_1 - \lambda_2)(2m+1) \sum_{u'=0}^{L-1} h^2(u') \mathbb{1}_{\{z(u')=1\}} \\
&= (\lambda - \lambda_1 - \lambda_2)(2m+1) \left\|h\right\|_2^2.
\end{aligned}$$

A classical result (see e.g. [31]) says that

$$F_t - F_0 - \int_0^t \mathcal{L}\varphi(Z_s, H_s, F_s) ds$$

is a zero-mean martingale, and hence

$$\mathbb{E}\left[F_t\right] = \mathbb{E}\left[F_0\right] + \mathbb{E}\left[\int_0^t \mathcal{L}\varphi(Z_s, H_s, F_s) ds\right]$$
$$\leq (\lambda - \lambda_1 - \lambda_2)(2m+1)\mathbb{E}\left[\int_0^t \|H_s\|_2^2 \, ds\right]. \tag{3.95}$$

Moreover, $\|H_t\|_2^2$ almost surely decreases, hence

$$\|H_1\|_2^2 \leq \int_0^1 \|H_s\|_2^2 \, \mathrm{d}s. \tag{3.96}$$

Let

$$\beta := (\lambda - \lambda_1 - \lambda_2)(2m+1) + 1.$$

Two equations (3.95), (3.96), and Fubini's theorem together imply that

$$\mathbb{E}\left[\sum_{j=1}^k H_{\tau_j -}^2(i_j) + \|H_1\|_2^2\right] \leq \beta \int_0^1 \mathbb{E}\left[\|H_s\|_2^2\right] \mathrm{d}s. \tag{3.97}$$

Note that for any $s \in [0,1]$,

$$\mathbb{E}\left[\|H_s\|_2^2\right] = \mathbb{E}\left[\sum_{u \in \Lambda_L} \mathbb{P}\left[U_1(s) = U_2(s) = u | \Xi\right]\right] = \mathbb{E}\left[\mathbb{1}_{\{\mathsf{dist}(U_1(s), U_2(s)) = 0\}}\right].$$

Therefore

$$\int_0^1 \mathbb{E}\left[\|H_s\|_2^2\right] \mathrm{d}s = \int_0^1 \mathbb{E}\left[\mathbb{1}_{\{\mathsf{dist}(U_1(s), U_2(s)) = 0\}}\right] \mathrm{d}s$$
$$= \mathbb{E}\left[\int_0^1 \mathbb{1}_{\{\mathsf{dist}(U_1(s), U_2(s)) = 0\}} \mathrm{d}s\right]. \tag{3.98}$$

Two equations (3.97), (3.98) lead to what we want. □

6.4 Second step of the proof of Proposition 3.36: the excursion coupling

This subsection is devoted to estimating the right-hand side of (3.92). It should be more convenient to consider a modified version $(\bar{Z}, \bar{U}_1, \bar{U}_2)$ of (Z, U_1, U_2), where we do not kill the walks when we see refresh marks or correlation marks, and we always color the sites occupied by $\bar{U}_1, \bar{U}_2$ blue. Note that before U_1 or U_2 are killed, $(\bar{Z}, \bar{U}_1, \bar{U}_2) = (Z, U_1, U_2)$. Therefore, almost surely,

$$\int_0^1 \mathbb{1}_{\{\mathsf{dist}(U_1(s), U_2(s)) = 0\}} \mathrm{d}s \leq \int_0^1 \mathbb{1}_{\{\mathsf{dist}(\bar{U}_1(s), \bar{U}_2(s)) = 0\}} \mathrm{d}s. \tag{3.99}$$

Besides, not killing the walks only increases the blue sites, i.e. , almost surely,

$$\bar{Z}_s \geq Z_s, \forall s \in \mathbb{R}_+. \tag{3.100}$$

An important fact here is the observation of Aldous in [3]: $(\bar{Z}, \bar{U}_1, \bar{U}_2)$ itself is a Markov process (without conditional probability), constructed as follows. We still use the collection of marks given in Graphical Construction 3. Suppose that we are at state (z, u_1, u_2).

- When a refresh mark appears, say at site u, make the transition $(z, u_1, u_2) \mapsto (z^{u,1}, u_1, u_2)$.
- When a Glauber mark appears, say at site u, make the transition $(z, u_1, u_2) \mapsto ((z \setminus b(u,m)) \cup \{u_1, u_2\}, u_1, u_2)$.
- When a black mark appears, say on the edge $(u, u+1)$, if $z(u) = 0$ or $z(u+1) = 0$, make the transition $(z, u_1, u_2) \mapsto (z^{u \leftrightarrow u+1}, u_1^{u \leftrightarrow u+1}, u_2^{u \leftrightarrow u+1})$, where $u_1^{u \leftrightarrow u+1}$ is given by

$$u_1^{u \leftrightarrow u+1} = \begin{cases} u+1 & \text{if } u_1 = u, \\ u & \text{if } u_1 = u+1, \\ u_1 & \text{if } u_1 \notin \{u, u+1\}, \end{cases}$$

  and $u_2^{u \leftrightarrow u+1}$ is defined similarly.
- When a blue mark appears, say on edge $(u, u+1)$, if $z(u) = z(u+1) = 1$,

$$\begin{aligned}
&(z, u_1, u_2) \mapsto (z, u_1, u_2) && \text{with probability } 1/4, \\
&(z, u_1, u_2) \mapsto (z, u_1^{u \leftrightarrow u+1}, u_2) && \text{with probability } 1/4, \\
&(z, u_1, u_2) \mapsto (z, u_1, u_2^{u \leftrightarrow u+1}) && \text{with probability } 1/4, \\
&(z, u_1, u_2) \mapsto (z, u_1^{u \leftrightarrow u+1}, u_2^{u \leftrightarrow u+1}) && \text{with probability } 1/4.
\end{aligned}$$

By abuse of notation, we write $\mathbb{P}_{z,u_1,u_2}[\cdot], \mathbb{E}_{z,u_1,u_2}[\cdot]$ for the probability and expectation taken with respect to the process $(\bar{Z}, \bar{U}_1, \bar{U}_2)$ starting from (z, u_1, u_2). We will compare $(\bar{U}_1, \bar{U}_2)$ with its idealized version $(\widetilde{U}_1, \widetilde{U}_2)$, defined§ as follows.

Idealized version of $(\bar{U}_1, \bar{U}_2)$. Let two random walks $(\widetilde{U}_1, \widetilde{U}_2)$ on Λ_L be constructed as follows. Each edge of the lattice is associated with a Poisson clock of intensity $2L^2$. Whenever a clock rings, if $\widetilde{U}_1$ (or $\widetilde{U}_2$) is at one endpoint of the edge, it decides to take the edge with probability 1/2. If both can take the edge, their decisions are made independently. Informally, $(\widetilde{U}_1, \widetilde{U}_2)$ is the version of $(\bar{U}_1, \bar{U}_2)$ where we assume that all the sites are always blue.

The excursion. Let $\tau_i^{0\to1}$ be the time for i-th passage from 0 to 1, and let $\tau_i^{1\to0}$ be the time for i-th passage from 1 to 0 of $\left(\mathsf{dist}\left(\bar{U}_1(t), \bar{U}_2(t)\right)\right)_{t\geq0}$. More precisely, let $\tau_i^{0\to1}, \tau_i^{1\to0}$, and $\mathcal{T}_i$

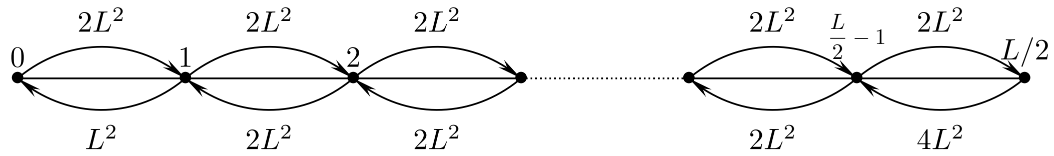


Figure 3.4: Transition rates of $\mathsf{dist}\left(\widetilde{U}_1, \widetilde{U}_2\right)$ when L is even.

be defined recursively as follows.

$$\mathcal{T}_j = \sum_{i=1}^{j-1} (\tau_i^{0\to 1} + \tau_i^{1\to 0}), \tag{3.101}$$

$$\tau_i^{0\to 1} = \inf\left\{t \geq 0 : \mathsf{dist}\left(\bar{U}_1(\mathcal{T}_i + t), \bar{U}_2(\mathcal{T}_i + t)\right) = 1\right\}, \tag{3.102}$$

$$\tau_i^{1\to 0} = \inf\left\{t \geq 0 : \mathsf{dist}\left(\bar{U}_1(\mathcal{T}_i + \tau_i^{0\to 1} + t), \bar{U}_2(\mathcal{T}_i + \tau_i^{0\to 1} + t)\right) = 0\right\}. \tag{3.103}$$

The sequence $(\tilde{\tau}_i^{0\to 1}, \tilde{\tau}_i^{1\to 0}, \tilde{\mathcal{T}}_i)_{i\geq 1}$ is defined similarly by replacing $(\bar{U}_1, \bar{U}_2)$ with $(\widetilde{U}_1, \widetilde{U}_2)$. The idea is to express the local time in terms of the excursions:

$$\int_0^1 \mathbb{1}_{\left\{\mathsf{dist}\left(\bar{U}_1(s), \bar{U}_2(s)\right)=0\right\}} \mathrm{d}s = \sum_{i\geq 1} \mathbb{1}_{\{\mathcal{T}_i < 1\}} \times (\tau_i^{0\to 1} \wedge [1 - \mathcal{T}_i]), \tag{3.104}$$

$$\int_0^1 \mathbb{1}_{\left\{\mathsf{dist}\left(\widetilde{U}_1(s), \widetilde{U}_2(s)\right)=0\right\}} \mathrm{d}s = \sum_{i\geq 1} \mathbb{1}_{\left\{\tilde{\mathcal{T}}_i < 1\right\}} \times (\tilde{\tau}_i^{0\to 1} \wedge [1 - \tilde{\mathcal{T}}_i]). \tag{3.105}$$

Let us define Υ by:

$$\Upsilon = \sum_{i\geq 1} \mathbb{1}_{\left\{\tilde{\mathcal{T}}_i < 1\right\}} \times \tilde{\tau}_i^{0\to 1}. \tag{3.106}$$

From now on, we fix a choice of β_2, β_3 as in Lemma 3.37, and we alway write BAD for BAD_{β_2}. We have the following results.

Lemma 3.44. $\mathsf{dist}\left(\widetilde{U}_1, \widetilde{U}_2\right)$ *is a Markov process with transition rates given in Figure 3.4.*

Proposition 3.45 (Excursion coupling)**.** *There is a coupling of* $(\bar{U}_1, \bar{U}_2)$ *and* $(\widetilde{U}_1, \widetilde{U}_2)$ *where*

$$\mathsf{dist}\left(\bar{U}_1(0), \bar{U}_2(0)\right) = \mathsf{dist}\left(\widetilde{U}_1(0), \widetilde{U}_2(0)\right) = 0$$

such that, almost surely, for all $i \geq 1$*,*

$$\tau_i^{0\to 1} \geq \tilde{\tau}_i^{0\to 1}, \tag{3.107}$$

$$\tau_i^{1\to 0} = \tilde{\tau}_i^{1\to 0}. \tag{3.108}$$

Lemma 3.46 (Fast passage to 1 from good environment)**.** *There exists a constant* $\beta > 0$ *such that*

$$\max_{t\geq 0} \max_{z\notin \mathrm{BAD}, u\in z} \mathbb{E}_{z,u,u}\left[(\tau_1^{0\to 1} \wedge t) \times \mathbb{1}_{\left\{\bar{Z}_s \notin \mathrm{BAD}, \forall s\in[0,t]\right\}}\right] \leq \beta \frac{\log^2 L}{L^2}.$$

Lemma 3.47 (Anti-concentration, integral form)**.**

$$\mathbb{E}\left[\Upsilon\right] = \mathcal{O}\left(1/L\right). \tag{3.109}$$

Remark 13. In dimension 2, the bound on the right-hand side of (3.109) is replaced by $\mathcal{O}\left(\frac{\log L}{L^2}\right)$.

Now we can prove Proposition 3.36.

Proof of Proposition 3.36. Thanks to Lemma 3.41 and equation (3.99), we now only have to estimate $\mathbb{E}\left[\int_0^1 \mathbb{1}_{\left\{\mathsf{dist}\left(\bar{U}_1(s),\bar{U}_2(s)\right)=0\right\}}\mathrm{d}s\right]$. We have

$$\begin{aligned}&\mathbb{E}\left[\int_0^1 \mathbb{1}_{\left\{\mathsf{dist}\left(\bar{U}_1(s),\bar{U}_2(s)\right)=0\right\}}\mathrm{d}s\right]\\&=\mathbb{E}\left[\int_0^1 \mathbb{1}_{\left\{\mathsf{dist}\left(\bar{U}_1(s),\bar{U}_2(s)\right)=0\right\}}\mathrm{d}s\left(\mathbb{1}_{\left\{\exists s\in[0,1]:\bar{Z}_s\in\mathrm{BAD}\right\}}+\mathbb{1}_{\left\{\forall s\in[0,1]:\bar{Z}_s\notin\mathrm{BAD}\right\}}\right)\right].\end{aligned} \tag{3.110}$$

For the first term, we write

$$\begin{aligned}\mathbb{E}\left[\int_0^1 \mathbb{1}_{\left\{\mathsf{dist}\left(\bar{U}_1(s),\bar{U}_2(s)\right)=0\right\}}\mathrm{d}s\mathbb{1}_{\left\{\exists s\in[0,1]:\bar{Z}_s\in\mathrm{BAD}\right\}}\right]&\le\mathbb{P}\left[\exists s\in[0,1]:\bar{Z}_s\in\mathrm{BAD}\right]\\&\le\mathbb{P}\left[\exists s\in[0,1]:Z_s\in\mathrm{BAD}\right],\end{aligned} \tag{3.111}$$

where we have used (3.100) in the last inequality. Now, we estimate the second term. We will prove that

$$\mathbb{E}\left[\int_0^1 \mathbb{1}_{\left\{\mathsf{dist}\left(\bar{U}_1(s),\bar{U}_2(s)\right)=0\right\}}\mathrm{d}s\mathbb{1}_{\left\{\forall s\in[0,1]:\bar{Z}_s\notin\mathrm{BAD}\right\}}\right]=\mathcal{O}\left(\log^2 L\,\mathbb{E}\left[\Upsilon\right]\right). \tag{3.112}$$

Thanks to (3.104), (3.105), and Lemma 3.47, we only need to prove that there exists a constant β such that

$$\mathbb{E}\left[\mathbb{1}_{\left\{\forall s\in[0,1]:\bar{Z}_s\notin\mathrm{BAD}\right\}}\mathbb{1}_{\{\mathcal{T}_i<1\}}\times(\tau_i^{0\to1}\wedge[1-\mathcal{T}_i])\right]\le\beta\log^2 L\times\mathbb{E}\left[\mathbb{1}_{\{\tilde{\mathcal{T}}_i<1\}}\times\tilde{\tau}_i^{0\to1}\right]. \tag{3.113}$$

Let β_1 be the constant in Lemma 3.46. By the Markov property at time $\mathcal{T}_i$,

$$\begin{aligned}&\mathbb{E}\left[\mathbb{1}_{\left\{\forall s\in[0,1]:\bar{Z}_s\notin\mathrm{BAD}\right\}}\mathbb{1}_{\{\mathcal{T}_i<1\}}\times(\tau_i^{0\to1}\wedge[1-\mathcal{T}_i])\right]\\&=\mathbb{E}\Big[\mathbb{1}_{\left\{\forall s\in[0,\mathcal{T}_i]:\bar{Z}_s\notin\mathrm{BAD}\right\}}\mathbb{1}_{\{\mathcal{T}_i<1\}}\\&\qquad\times\mathbb{E}_{\bar{Z}(\mathcal{T}_i),\bar{U}_1(\mathcal{T}_i),\bar{U}_2(\mathcal{T}_i)}\left[\mathbb{1}_{\left\{\forall s\in[0,1-\mathcal{T}_i]:\bar{Z}_s\notin\mathrm{BAD}\right\}}\times\left(\tau_1^{0\to1}\wedge(1-\mathcal{T}_i)\right)\right]\Big]\\&\le\mathbb{E}\left[\mathbb{1}_{\left\{\forall s\in[0,\mathcal{T}_i]:\bar{Z}_s\notin\mathrm{BAD}\right\}}\mathbb{1}_{\{\mathcal{T}_i<1\}}\frac{\beta_1\log^2 L}{L^2}\right]\\&\le\frac{\beta_1\log^2 L}{L^2}\mathbb{P}\left[\mathcal{T}_i<1\right]\\&\le\frac{\beta_1\log^2 L}{L^2}\mathbb{P}\left[\tilde{\mathcal{T}}_i<1\right].\end{aligned} \tag{3.114}$$

Here we have used Lemma 3.46 in the first inequality and the fact that $\mathcal{T}_i \geq \tilde{\mathcal{T}}_i$ (due to Proposition 3.45) in the last inequality. On the other hand,

$$\begin{aligned}\mathbb{E}\left[\mathbb{1}_{\{\tilde{\mathcal{T}}_i<1\}} \times \tilde{\tau}_i^{0\to 1}\right] &= \mathbb{E}\left[\mathbb{1}_{\{\tilde{\mathcal{T}}_i<1\}}\mathbb{E}_{\widetilde{U}_1(\tilde{\mathcal{T}}_i),\widetilde{U}_1(\tilde{\mathcal{T}}_i)}\left[\tilde{\tau}_1^{0\to 1}\right]\right] \\ &= \mathbb{E}\left[\mathbb{1}_{\{\tilde{\mathcal{T}}_i<1\}}\frac{1}{2L^2}\right]. \end{aligned} \tag{3.115}$$

Here we have used the fact that $\mathsf{dist}\left(\widetilde{U}_1(\tilde{\mathcal{T}}_i), \widetilde{U}_2(\tilde{\mathcal{T}}_i)\right) = 0$ by construction, and that the expected time needed for $\mathsf{dist}\left(\widetilde{U}_1, \widetilde{U}_2\right)$ to reach 1 from this time onwards is $\frac{1}{2L^2}$, due to Lemma 3.44. So (3.114) and (3.115) together imply (3.113). The equations (3.90), (3.99), (3.110), (3.112), (3.111) together imply that

$$\|\mathbb{P}_{\pi_1}\left[X_t \in \cdot\right] - \mathbb{P}_{\pi_2}\left[X_t \in \cdot\right]\|_{TV}^2 = \mathcal{O}\left(\frac{\log^2 L}{L} + \mathbb{P}_{z\cup\{0\}}\left[\exists s \in [0,1] : Z_s \in \text{BAD}\right]\right).$$

Together with the inequality $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$, $a, b > 0$, the inequality above gives us what we want. □

It remains to prove Lemma 3.44, Proposition 3.45, Lemma 3.46, and Lemma 3.47. We will need the following result to prove Lemma 3.44 and Proposition 3.45.

Lemma 3.48 (Strong lumpability of Markov process)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be two finite sets, and let $f : \mathcal{X} \to \mathcal{Y}$ be a surjective function. Let $(X_t)_{t\geq 0}$ be a Markov process with a generator $\mathcal{L}$ on $\mathcal{X}$. Then $(f(X_t))_{t\geq 0}$ is a Markov process for any initial condition of $(X_t)_{t\geq 0}$ if and only if*

$$\forall y_1, y_2 \in \mathcal{Y}, y_1 \neq y_2, \forall x_1, x_1' \in f^{-1}(y_1), \sum_{x\in f^{-1}(y_2)} \mathcal{L}(x_1, x) = \sum_{x\in f^{-1}(y_2)} \mathcal{L}(x_1', x). \tag{3.116}$$

Moreover, if (3.116) is satisfied, then the generator $\mathcal{L}'$ of $(f(X_t))_{t\geq 0}$ is determined by

$$\forall y_1, y_2 \in \mathcal{Y}, y_1 \neq y_2, \forall x_1 \in f^{-1}(y_1), \mathcal{L}'(y_1, y_2) = \sum_{x\in f^{-1}(y_2)} \mathcal{L}(x_1, x). \tag{3.117}$$

Lemma 3.48 is a classical result. For its proof, the readers can see Theorem 2.4 in [4] and Lemma 2.5 in [59] for the analogous result for discrete-time Markov chains. Now we prove Lemma 3.44 and Proposition 3.45.

Proof of Lemma 3.44. This is just a direct application of Lemma 3.48 for the process $(\widetilde{U}_1, \widetilde{U}_2)$ and the function $\mathsf{dist}\,(\cdot,\cdot)$. □

Proof of Proposition 3.45. The intuition is that $(\bar{U}_1, \bar{U}_2)$ evolves exactly like $(\widetilde{U}_1, \widetilde{U}_2)$ when $\bar{U}_1$ and $\bar{U}_2$ are not on a same site. This explains (3.108). The only difference is when $\bar{U}_1$ and $\bar{U}_2$ are on a same site since they might be forced to move together when their neighbors are not blue. This explains (3.107).

More precisely, for any initial configuration of $(\bar{Z}, \bar{U}_1, \bar{U}_2)$, consider the modified version $(\bar{Z}^{mod}, \bar{U}_1^{mod}, \bar{U}_2^{mod})$ of $(\bar{Z}, \bar{U}_1, \bar{U}_2)$ which is stopped at $E := \{(z, u, u) | z \in \mathcal{Z}, u \in \Lambda_L\}$. Then

$(\bar{Z}^{mod}, \bar{U}_1^{mod}, \bar{U}_2^{mod})$ is also a Markov process, and up to the time that $(\bar{Z}, \bar{U}_1, \bar{U}_2)$ reaches E, those two processes are the same. The modified version $(\widetilde{U}_1^{mod}, \widetilde{U}_2^{mod})$ of $(\widetilde{U}_1, \widetilde{U}_2)$ is defined similarly.

We apply Lemma 3.48 for the function

$$f : (z, u_1, u_2) \mapsto \mathsf{dist}\,(u_1, u_2)$$

and the chain $(\bar{Z}^{mod}, \bar{U}_1^{mod}, \bar{U}_2^{mod})$ to see that $f(\bar{Z}^{mod}, \bar{U}_1^{mod}, \bar{U}_2^{mod})$ is also a Markov process. Let $\mathcal{L}$ be the generator associated with this process. By direct computation, the process with generator $\mathcal{L}$ can be described as in Figure 3.4, except that the transition from 0 to 1 is suppressed.

Similarly, we apply Lemma 3.48 for the function

$$\tilde{f} : (u_1, u_2) \mapsto \mathsf{dist}\,(u_1, u_2)$$

and the process $(\widetilde{U}_1^{mod}, \widetilde{U}_2^{mod})$ to see that $\tilde{f}(\widetilde{U}_1^{mod}, \widetilde{U}_2^{mod})$ is also a Markov process. By direct computation, we can see that its generator is also $\mathcal{L}$.

Our coupling is as follows. Let $\mathcal{G}_t$ be the σ-algebra generated by $\Xi([0,t])$. Recall the definition of $(\mathcal{T}_i)$ in (3.101).

- Conditionally on $\mathcal{G}_{\mathcal{T}_i}$, let τ_i be the first time counted from $\mathcal{T}_i$ onwards that $\bar{U}_1$ meets a blue mark and the two walks $\bar{U}_1, \bar{U}_2$ take different decisions. Then,

  $$\tau_i \le \tau_i^{0\to 1},$$

  where the inequality is strict if the other endpoint of the blue mark is a black site.

- Conditionally on $\mathcal{G}_{\mathcal{T}_i+\tau_i^{0\to 1}}$, $\mathsf{dist}\left(\bar{U}_1, \bar{U}_2\right)$ evolves as a Markov process with generator $\mathcal{L}$ until it reaches 0. This gives us the excursion $\tau_i^{1\to 0}$.

Note that, by construction,

- τ_i is independent of $\mathcal{G}_{\mathcal{T}_i}$, and hence of $(\tau_j, \tau_j^{1\to 0})_{1\le j\le i-1}$.
- $\tau_i^{1\to 0}$ is independent of $\mathcal{G}_{\mathcal{T}_i+\tau_i^{0\to 1}}$ and hence of $(\tau_j, \tau_j^{1\to 0})_{1\le j\le i-1}$.
- For any i, $\tau_i \sim \exp(2L^2)$.
- For any i, $\tau_i^{1\to 0}$ is the time of an excursion of a Markov process with generator $\mathcal{L}$ starting at state 1 until reaching state 0.

Hence $(\tau_i, \tau_i^{1\to 0})_{i\ge 1}$ are i.i.d., and they have the same distribution as the $(\tilde{\tau}_i^{0\to 1}, \tilde{\tau}_i^{1\to 0})_{i\ge 1}$. This finishes our proof. □

Now we prove Lemma 3.46. We will need the following results.

Lemma 3.49 (Fast transition in good environment)**.** *There exist constants $\beta, \beta' > 0$ such that*

$$\min_{z\notin \mathrm{BAD}, u\in z} \mathbb{P}_{z,u,u}\left[\tau_1^{0\to 1} < \frac{\beta' \log^2 L}{L^2}\right] \ge \beta.$$

Proof of Lemma 3.46. Let $z \in \mathrm{BAD}$, $u \in z$ arbitrary. Let β, β' be as in Lemma 3.49. Then

$$\begin{aligned}
&\mathbb{E}_{z,u,u}\left[(\tau_1^{0\to1} \wedge t) \times \mathbb{1}_{\{\bar{Z}_s \notin \mathrm{BAD}, \forall s \in [0,t]\}}\right] \\
&\leq \sum_{i=0}^{\lfloor tL^2/(\beta' \log^2 L)\rfloor} \mathbb{E}_{z,u,u}\left[\tau_1^{0\to1} \times \mathbb{1}_{\{\bar{Z}_s \notin \mathrm{BAD}, \forall s \in [0,t]\}} \mathbb{1}_{\left\{\frac{i\beta' \log^2 L}{L^2} \leq \tau_1^{0\to1} < \frac{(i+1)\beta' \log^2 L}{L^2}\right\}}\right] \\
&\quad + \mathbb{E}_{z,u,u}\left[t \times \mathbb{1}_{\{\bar{Z}_s \notin \mathrm{BAD}, \forall s \in [0,t]\}} \mathbb{1}_{\{\tau_1^{0\to1} > t\}}\right] \\
&\leq \left(\sum_{i=0}^{\lfloor tL^2/(\beta' \log^2 L)\rfloor} \frac{(i+1)\beta' \log^2 L}{L^2} \mathbb{E}_{z,u,u}\left[\mathbb{1}_{\{\bar{Z}_s \notin \mathrm{BAD}, \forall s \in [0,t]\}} \mathbb{1}_{\left\{\frac{i\beta' \log^2 L}{L^2} \leq \tau_1^{0\to1}\right\}}\right]\right) \\
&\quad + t\mathbb{E}_{z,u,u}\left[\mathbb{1}_{\{\bar{Z}_s \notin \mathrm{BAD}, \forall s \in [0,t]\}} \mathbb{1}_{\{\tau_1^{0\to1} > \lfloor tL^2/(\beta' \log^2 L)\rfloor (\beta' \log^2 L)/(L^2)\}}\right] \\
&\leq \left(\sum_{i=0}^{\lfloor tL^2/(\beta' \log^2 L)\rfloor} \frac{(i+1)\beta' \log^2 L}{L^2}(1-\beta)^i\right) + t(1-\beta)^{\lfloor tL^2/(\beta' \log^2 L)\rfloor} \\
&= \mathcal{O}\left(\log^2 L/L^2\right).
\end{aligned}$$

□

Now we prove Lemma 3.49. We recall the following classical result about the hitting time of a simple random walk.

Lemma 3.50 (Hitting time of a symmetric simple random walk)**.** *Let U be a simple random walk on $\mathbb{Z}$ where edges have conductance 1. Let $\tau := \inf\{t \geq 0 : U(t) = 0\}$ be the hitting time of the site 0. Then there exists a constant $\beta > 0$ such that*

$$\max_{u \in \mathbb{Z}} \mathbb{P}_u\left[\tau \leq 2|u|^2\right] > \beta. \tag{3.118}$$

In dimension 2, with similar notations and assumptions, there exists a constant $\beta >$ such that

$$\max_{u \in \mathbb{Z}^2} \mathbb{P}_u\left[\tau \leq 2\,\|u\|^3\right] > \beta.$$

This lemma can be proved by a classical hitting time estimate for non-negative supermartingale. For completeness, we include the proof for dimension 2 in Appendix.

Proof of Lemma 3.49. First, consider the case where one of the neighbors of u, say $u+1$, is blue. Then there is a probability bounded away from 0 that a blue mark appears on edge $(u, u+1)$ before time $1/L^2$ and before any blue or black mark appears on edges $(u+1, u+2)$ and $(u-1, u)$. When this happens, with probability $1/2$, the two walks $\bar{U}_1, \bar{U}_2$ can make different decisions. Therefore, $\tau_1^{0\to1} < 1/L^2$ with a probability greater than some constant $\beta_4 > 0$.

Now we come back to the general case. Since $z \notin \mathrm{BAD}$, there exists a blue site u' at a distance at most $2\beta_2 \log L$ from u. Let U_3 be the walk starting at u', which behaves like U_1 (see paragraph "the perturbation position") when it meets the marks.

Let β_1 be a constant that we will choose later. The number of Glauber marks appearing in the interval $\left[0, \frac{\beta_1 \log^2 L}{L^2}\right]$ is a Poisson variable with average $\frac{\lambda \beta_1 \log^2 L}{L}$. Hence, for L large enough,

$$\mathbb{P}\left[\text{no Glauber mark appears in } \left[0, \frac{\beta_1 \log^2 L}{L^2}\right]\right] \geq 1/2.$$

Conditionally on this event, U_3 and $\bar{U}_1$ evolve as two independent random walks on Λ_L in $\left[0, \frac{\beta_1 \log^2 L}{L^2}\right]$, until they are at distance 1 from each other. Let

$$\tau := \inf \left\{ t \geq 0 | \mathsf{dist}\left(U_3(t), \bar{U}_1(t)\right) = 1 \right\}.$$

Note that $U_3 - \bar{U}_1$ is a simple random walk on Λ_L where edges have conductance $2L^2$ until it reaches 1 or -1. So for β as in Lemma 3.50, there exists a constant β_1 such that

$$\mathbb{P}\left[\tau < \frac{\beta_1 \log^2 L}{L^2} \middle| \text{no Glauber mark appears in } \left[0, \frac{\beta_1 \log^2 L}{L^2}\right]\right] \geq \beta.$$

Hence,

$$\mathbb{P}\left[\tau < \frac{\beta_1 \log^2 L}{L^2}, \text{no Glauber mark appears in } \left[0, \frac{\beta_1 \log^2 L}{L^2}\right]\right] \geq \beta/2.$$

By the strong Markov property at time τ and the first case, we can deduce that

$$\mathbb{P}\left[\tau_1^{0\to 1} \leq \tau + 1/L^2, \tau < \frac{\beta_1 \log^2 L}{L^2}, \text{no Glauber mark appears in } \left[0, \frac{\beta_1 \log^2 L}{L^2}\right]\right] \geq \beta_4 \beta/2.$$

This finishes our proof. □

Remark 14. In dimension $d = 2$, there is a similar version of Lemma 3.49. This time, the proof involves the 2-dimensional version of Lemma 3.50, so in the statement of the result, we replace $\log^2 L$ by $\log^3 L$. We see that, with the same notations as in the proof above, U_3 and $\bar{U}_1$ still evolve as two independent walks until they are at a distance 1 from each other or until U_3 is killed when a Glauber mark appears. The number of Glauber marks appearing in the time interval $\left[0, \frac{\beta_1 \log^3 L}{L^2}\right]$ is $\mathcal{O}\left(\log^3 L\right)$. Each time there is a ring among the Glauber clocks, the chance that the mark appearing kills U_3 is $\mathcal{O}\left(1/L^2\right)$. So the chance that U_3 is killed in the time interval $\left[0, \frac{\beta_1 \log^3 L}{L^2}\right]$ is $\mathcal{O}\left(\frac{\log^3 L}{L^2}\right) = o(1)$. So if initially, $\mathsf{dist}\left(U_3, \bar{U}_1\right) = \mathcal{O}\left(\log L\right)$, then U_3 can still meet $\bar{U}_1$ in time interval $\left[0, \frac{\beta_1 \log^3 L}{L^2}\right]$ with a chance bounded away from 0. The proof is then adapted accordingly.

It remains to prove Lemma 3.47. It is not hard to prove the result if we replace $(\widetilde{U}_1, \widetilde{U}_2)$

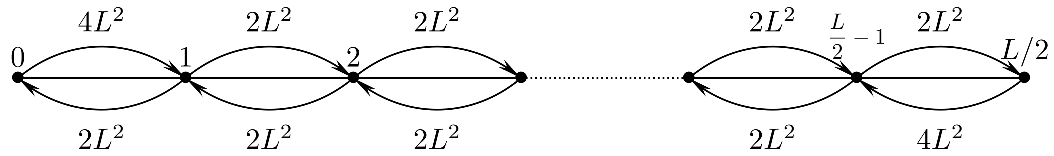


Figure 3.5: Transition rate of $\mathsf{dist}\left(\check{\mathcal{U}}_1, \check{\mathcal{U}}_2\right)$ when L is even.

with two independent SRWs. We will prove the results for two independent SRWs and use a comparison argument to transfer the result back to the pair $(\widetilde{U}_1, \widetilde{U}_2)$.

Let $\check{\mathcal{U}}_1$ be a SRW starting from 0 on Λ_L, where every edge has conductance L^2. Let $\check{\mathcal{U}}_2$ be an independent copy of $\check{\mathcal{U}}_1$. Let $(\check{\tau}_i^{0\to 1}, \check{\tau}_i^{1\to 0}, \check{\mathcal{T}}_i)_{i\geq 1}$ be defined as in (3.101), (3.102), (3.103) when we replace $(\bar{U}_1, \bar{U}_2)$ by $(\check{\mathcal{U}}_1, \check{\mathcal{U}}_2)$. For two real random variables ζ_1, ζ_2, we write $\zeta_1 \overset{d}{\leq} \zeta_2$ if ζ_2 dominates stochastically ζ_1. We will need the following results.

Lemma 3.51.

$$\check{\tau}_1^{1\to 0} \overset{d}{\leq} \tilde{\tau}_1^{1\to 0}, \tag{3.119}$$

$$\check{\tau}_1^{0\to 1} \overset{d}{\leq} \tilde{\tau}_1^{0\to 1}. \tag{3.120}$$

Lemma 3.52.

$$\sum_{i=1}^{\infty} \mathbb{P}\left[\check{\mathcal{T}}_i < 1\right] = \mathcal{O}\left(L\right). \tag{3.121}$$

Now we prove Lemma 3.47.

Proof of Lemma 3.47. In this proof, $\mathbb{E}_0\left[\cdot\right]$ means the expectation taken with respect to the law of the process $\mathsf{dist}\left(\widetilde{U}_1, \widetilde{U}_2\right)$ starting from 0. We have

$$\begin{aligned}
\mathbb{E}\left[\Upsilon\right] &= \sum_{i\geq 1} \mathbb{E}\left[\mathbb{1}_{\left\{\tilde{\mathcal{T}}_i < 1\right\}} \times \check{\tau}_i^{0\to 1}\right] \\
&= \sum_{i\geq 1} \mathbb{E}\left[\mathbb{1}_{\left\{\tilde{\mathcal{T}}_i < 1\right\}} \times \mathbb{E}_0\left[\tilde{\tau}_1^{0\to 1}\right]\right] \\
&= \sum_{i\geq 1} \frac{1}{2L^2} \mathbb{P}\left[\tilde{\mathcal{T}}_i < 1\right],
\end{aligned}$$

where we have used the Markov property for the process $\mathsf{dist}\left(\widetilde{U}_1, \widetilde{U}_2\right)$ in the second equality, and the formula of the generator of $\mathsf{dist}\left(\widetilde{U}_1, \widetilde{U}_2\right)$ in the third equality. Note that, by Lemma

3.51, $\check{\mathcal{T}}_i \overset{d}{\le} \tilde{\mathcal{T}}_i$, $\forall i \ge 1$. Therefore,

$$\begin{aligned}\sum_{i\ge1} \frac{1}{2L^2}\mathbb{P}\left[\tilde{\mathcal{T}}_i < 1\right] &\le \frac{1}{2L^2}\sum_{i\ge1}\mathbb{P}\left[\check{\mathcal{T}}_i < 1\right]\\ &= \frac{1}{2L^2}\mathcal{O}\left(L\right)\\ &= \mathcal{O}\left(L\right).\end{aligned}$$

This finishes our proof. □

It remains to prove Lemma 3.51 and Lemma 3.52. We first prove Lemma 3.51.

Proof of Lemma 3.51. See Figure 3.4 and 3.5 for intuition. A direct application of Lemma 3.48 shows that $\mathsf{dist}\left(\check{\mathcal{U}}_1, \check{\mathcal{U}}_2\right)$ is a Markov process with transition rates as in Figure 3.5. Then we see that

$$\begin{aligned}\tilde{\tau}_1^{0\to1} &\sim \exp(2L^2),\\ \check{\tau}_1^{0\to1} &\sim \exp(4L^2).\end{aligned}$$

This proves (3.120). Now let $\tau^{2\to1}$ be the (random) time it takes for $\mathsf{dist}\left(\widetilde{U}_1, \widetilde{U}_2\right)$ to reach state 1 from state 2. By Figure 3.4 and Figure 3.5, it is also the time it takes for $\mathsf{dist}\left(\check{\mathcal{U}}_1, \check{\mathcal{U}}_2\right)$ to reach state 1 from state 2. Let $(\zeta_i)_{i\ge1}, (\tau_i^{2\to1})_{i\ge1}$, and $(\xi_i)_{i\ge1}$ be three independent sequences of i.i.d. random variables such that

- $(\zeta_i)_{i\ge1}$ is a sequence of i.i.d. $\exp(L^2)$.
- $(\tau_i^{2\to1})_{i\ge1}$ is a sequence of independent copies of $\tau^{2\to1}$.
- $(\xi_i)_{i\ge1}$ is a sequence of i.i.d. Bernoulli random variables with parameter 2/3.

Then we see that

$$\begin{aligned}\tilde{\tau}_1^{1\to0} &\overset{d}{=} \frac{\zeta_1}{3} + \xi_1\tau_1^{2\to1} + \xi_1\frac{\zeta_2}{3} + \xi_1\xi_2\tau_2^{2\to1} + \dots\\ &= \sum_{i=1}^{\infty}\left[\left(\frac{\zeta_i}{3}\prod_{j=1}^{i-1}\xi_i\right) + \left(\tau_i^{2\to1}\prod_{j=1}^{i}\xi_i\right)\right].\end{aligned}$$

This is because to go to state 0 from state 1, first, we need to jump out of state 1, which takes a time $\exp(3L^2)$. Then, with probability 1/3, we succeed to jump to state 0, and with probability 2/3, we jump to state 2. Then we have to come back to state 1 and try again. Similarly, let $(\xi_i')_{i\ge1}$ be a sequence of i.i.d. Bernoulli random variables with parameter 1/2, independent of $(\zeta_i)_{i\ge1}$ and $(\tau_i^{2\to1})_{i\ge1}$. Then

$$\check{\tau}_1^{1\to0} \overset{d}{=} \sum_{i=1}^{\infty}\left[\left(\frac{\zeta_i}{4}\prod_{j=1}^{i-1}\xi_i'\right) + \left(\tau_i^{2\to1}\prod_{j=1}^{i}\xi_i'\right)\right].$$

Since $\xi_i' \overset{d}{\leq} \xi_i$, we conclude that $\check{\tau}_1^{1\to 0} \overset{d}{\leq} \tilde{\tau}_1^{1\to 0}$, which finishes our proof. □

Finally, we show Lemma 3.52. We will need the following classical result about anticoncentration of SRW on the lattice.

Lemma 3.53 (Local limit theorem)**.** *Let $\check{\mathcal{U}}$ be a SRW on Λ_L starting from 0, where every edge has conductance L^2. Then,*

$$\max_{u\in\Lambda_L} \mathbb{P}_0\left[\check{\mathcal{U}}(t) = u\right] = \mathcal{O}\left(\frac{1}{L}\left(\frac{1}{\sqrt{t}}+1\right)\right).$$

Proof of Lemma 3.52. Let us define

$$\check{\Upsilon} := \sum_{i\geq 1} \mathbb{1}_{\{\check{\mathcal{T}}_i<1\}} \times \check{\tau}_i^{0\to 1}.$$

We see that

$$\begin{aligned}
&\check{\Upsilon} - \int_0^1 \mathbb{1}_{\{\mathsf{dist}(\check{\mathcal{U}}_1(s),\check{\mathcal{U}}_2(s))=0\}}\mathrm{d}s \\
&= \mathbb{1}_{\{\mathsf{dist}(\check{\mathcal{U}}_1(1),\check{\mathcal{U}}_2(1))=0\}} \times \left(\inf\left\{t\geq 1 | \mathsf{dist}\left(\check{\mathcal{U}}_1(t),\check{\mathcal{U}}_2(t)\right)=1\right\} - 1\right)
\end{aligned}$$

Taking the expectation and using the Markov property at time 1, we get

$$\begin{aligned}
\mathbb{E}\left[\check{\Upsilon} - \int_0^1 \mathbb{1}_{\{\mathsf{dist}(\check{\mathcal{U}}_1(s),\check{\mathcal{U}}_2(s))=0\}}\mathrm{d}s\right] &= \mathbb{P}\left[\mathsf{dist}\left(\check{\mathcal{U}}_1(1),\check{\mathcal{U}}_2(1)\right)=0\right]\mathbb{E}\left[\check{\tau}_1^{0\to 1}\right] \\
&\leq 1\times \frac{1}{4L^2} \\
&= \frac{1}{4L^2}.
\end{aligned}$$

Note that $\mathbb{1}_{\{\mathsf{dist}(\check{\mathcal{U}}_1(s),\check{\mathcal{U}}_2(s))=0\}} = \mathbb{1}_{\{\check{\mathcal{U}}_1(s)-\check{\mathcal{U}}_2(s)=0\}}$, and note also that $(\check{\mathcal{U}}_1(s)-\check{\mathcal{U}}_2(s))_{s\geq 0}$ is a SRW on the lattice, where edges have conductance $2L^2$. Hence by Lemma 3.53,

$$\mathbb{P}\left[\check{\mathcal{U}}_1(s)-\check{\mathcal{U}}_2(s)=0\right] = \mathcal{O}\left(\frac{1}{L}\left(\frac{1}{\sqrt{s}}+1\right)\right).$$

Therefore,

$$\begin{aligned}
\mathbb{E}\left[\check{\Upsilon}\right] &\leq \frac{1}{4L^2} + \mathbb{E}\left[\int_0^1 \mathbb{1}_{\{\mathsf{dist}(\check{\mathcal{U}}_1(s),\check{\mathcal{U}}_2(s))=0\}}\mathrm{d}s\right] \\
&= \frac{1}{4L^2} + \int_0^1 \mathbb{P}\left[\check{\mathcal{U}}_1(s)-\check{\mathcal{U}}_2(s)=0\right]\mathrm{d}s \\
&= \frac{1}{4L^2} + \mathcal{O}\left(\int_0^1 \frac{1}{L}\left(\frac{1}{\sqrt{s}}+1\right)\mathrm{d}s\right) \\
&= \mathcal{O}\left(\frac{1}{L}\right).
\end{aligned}$$

On the other hand,

$$\mathbb{E}\left[\check{\Upsilon}\right] = \sum_{i\geq 1} \mathbb{P}\left[\check{\mathcal{T}}_i < 1\right] \times \frac{1}{4L^2}.$$

The two equations above imply that

$$\sum_{i\geq 1} \mathbb{P}\left[\check{\mathcal{T}}_i < 1\right] = \mathcal{O}\left(L\right),$$

which finishes our proof. □

Appendix

In this appendix, we prove Lemma 3.24 and Lemma 3.50. First, we prove Lemma 3.24. We will need the following result.

Lemma 3.54 (Anticoncentration of SRW, integral form)**.** *Let $\check{\mathcal{U}}_1, \check{\mathcal{U}}_2$ be two i.i.d. SRWs on Λ_L^d where edges have conductance L^2. Let θ be a strictly positive number and $\zeta \sim \exp(\theta)$, independent of $(\check{\mathcal{U}}_1, \check{\mathcal{U}}_2)$. Then, for any $k \in \mathbb{Z}_+$,*

$$\max_{u_1, u_2 \in \Lambda_L^d} \mathbb{P}_{u_1, u_2}\left[\mathsf{dist}\left(\check{\mathcal{U}}_1(\zeta), \check{\mathcal{U}}_2(\zeta)\right) \leq k\right] = \begin{cases} \mathcal{O}_{\theta,k}(1/L) & \text{if } d = 1, \\ \mathcal{O}_{\theta,k}(\log L/L^2) & \text{if } d = 2, \\ \mathcal{O}_{\theta,k}(1/L^2) & \text{if } d \geq 3. \end{cases}$$

For $u_1, u_2 \in \Lambda_L$, we write $u_1 \preceq u_2$ if $\mathsf{dist}\,(u_1, 0) \leq \mathsf{dist}\,(u_2, 0)$. In higher dimensions, for $u_1, u_2 \in \Lambda_L^d$, we write $u_1 \preceq u_2$ if there exists a permutation σ of $[d]$ such that $\forall i \in [d], u_1(i) \preceq u_2(\sigma(i))$. We will also need the following result.

Lemma 3.55 (Distance between 2 walks in an $IP(2)$)**.** *Consider the lattice Λ_L^d where every edge has conductance L^2. Let $(\mathcal{U}_1, \mathcal{U}_2)$ be an IP(2) and $(\check{\mathcal{U}}_1, \check{\mathcal{U}}_2)$ be a pair of independent SRWs on Λ_L^d. There is a Markovian coupling of $(\mathcal{U}_1, \mathcal{U}_2)$ and $(\check{\mathcal{U}}_1, \check{\mathcal{U}}_2)$ such that if the initial condition $(u_1, u_2, \check{u}_1, \check{u}_2)$ satisfies $(\check{u}_2 - \check{u}_1) \preceq (u_2 - u_1)$, then almost surely,*

$$\forall t \geq 0,\ (\check{\mathcal{U}}_2(t) - \check{\mathcal{U}}_1(t)) \preceq (\mathcal{U}_2(t) - \mathcal{U}_1(t)),$$

and in particular,

$$\mathsf{dist}\left(\check{\mathcal{U}}_2(t), \check{\mathcal{U}}_1(t)\right) \leq \mathsf{dist}\,(\mathcal{U}_2(t), \mathcal{U}_1(t))\,.$$

Clearly, Lemma 3.54 and Lemma 3.55 together imply Lemma 3.24.

Proof of Lemma 3.24. Let $u_1, u_2 \in \Lambda_L$. Consider the coupling of the IP(2) $(\mathcal{U}_1, \mathcal{U}_2)$ and a pair of independent SRWs $(\check{\mathcal{U}}_1, \check{\mathcal{U}}_2)$, both starting from u_1, u_2, independent of ζ, that satisfies Lemma 3.55. Then, by Lemma 3.55,

$$\mathbb{P}\left[\mathsf{dist}\,(\mathcal{U}_1(\zeta), \mathcal{U}_2(\zeta)) \leq k\right] \leq \mathbb{P}\left[\mathsf{dist}\left(\check{\mathcal{U}}_1(\zeta), \check{\mathcal{U}}_2(\zeta)\right) \leq k\right].$$

This and Lemma 3.54 lead to what we want. □

Lemma 3.54 is just the integral form of Lemma 3.53.

Proof of Lemma 3.54. We denote by $\check{\mathcal{U}}_1(t, i)$ the i-th coordinate of $\check{\mathcal{U}}_1(t)$, and similarly for $\check{\mathcal{U}}_2$. We see that $(\check{\mathcal{U}}_1(\cdot, i) - \check{\mathcal{U}}_2(\cdot, i))_{1 \leq i \leq d}$ are d independent SRWs on Λ_L, where every edge has conductance $2L^2$. Hence, for any $t \geq 0$,

$$\begin{aligned}\mathbb{P}\left[\mathsf{dist}\left(\mathcal{U}_1(t), \mathcal{U}_2(t)\right) \leq k\right] &\leq \prod_{i=1}^{d} \mathbb{P}\left[\left(\check{\mathcal{U}}_1(t,i) - \check{\mathcal{U}}_2(t,i)\right) \in \{-k, \dots, k\}\right] \\ &\leq \prod_{i=1}^{d} \left((2k+1) \max_{u \in \Lambda_L} \mathbb{P}\left[\check{\mathcal{U}}_1(t,i) - \check{\mathcal{U}}_2(t,i) = u\right]\right) \\ &= \mathcal{O}_k\left(\frac{1}{L^d}\left(\frac{1}{t^{d/2}} + 1\right)\right).\end{aligned}$$

Note that

$$\mathbb{P}\left[\mathsf{dist}\left(\check{\mathcal{U}}_1(\zeta), \check{\mathcal{U}}_2(\zeta)\right) \leq k\right] = \int_0^\infty \mathbb{P}\left[\mathsf{dist}\left(\check{\mathcal{U}}_1(t), \check{\mathcal{U}}_2(t)\right) \leq k\right] \theta e^{-\theta t} \mathrm{d}t.$$

We split the integral above into three intervals $[0, 1/L^2]$, $[1/L^2, 1]$, $[1, \infty)$. First, we estimate the integral on $[0, 1/L^2]$:

$$\begin{aligned}&\int_0^{1/L^2} \mathbb{P}\left[\mathsf{dist}\left(\check{\mathcal{U}}_1(t), \check{\mathcal{U}}_2(t)\right) \leq k\right] \theta e^{-\theta t} \mathrm{d}t \\ &\leq \int_0^{1/L^2} \theta e^{-\theta t} \mathrm{d}t \\ &\leq \theta / L^2.\end{aligned}$$

Now we estimate the integral on $[1, \infty)$:

$$\begin{aligned}&\int_1^\infty \mathbb{P}\left[\mathsf{dist}\left(\check{\mathcal{U}}_1(t), \check{\mathcal{U}}_2(t)\right) \leq k\right] \theta e^{-\theta t} \mathrm{d}t \\ &= \mathcal{O}_k\left(\int_1^\infty \frac{1}{L^d}\left(\frac{1}{t^{d/2}} + 1\right) \theta e^{-\theta t} \mathrm{d}t\right) \\ &= \mathcal{O}_k\left(\frac{1}{L^d} \int_1^\infty \theta e^{-\theta t} \mathrm{d}t\right) \\ &= \mathcal{O}_k\left(\frac{1}{L^d}\right).\end{aligned}$$

Finally, we estimate the integral on $[1/L^2, 1]$:

$$\begin{aligned}
&\int_{1/L^2}^{1} \mathbb{P}\left[\mathsf{dist}\left(\check{\mathcal{U}}_1(t), \check{\mathcal{U}}_2(t)\right) \le k\right] \theta e^{-\theta t} \mathrm{d}t \\
&= \mathcal{O}_k\left(\int_{1/L^2}^{1} \frac{1}{L^d}\left(\frac{1}{t^{d/2}} + 1\right)\theta e^{-\theta t}\mathrm{d}t\right) \\
&= \mathcal{O}_k\left(\frac{\theta}{L^d}\int_{1/L^2}^{1} \frac{1}{t^{d/2}}\mathrm{d}t\right) \\
&= \begin{cases}
\mathcal{O}_k\left(\frac{\theta}{L}\left(2\sqrt{t}\Big|_{1/L^2}^{1}\right)\right) = \mathcal{O}_k\left(\frac{\theta}{L}\right) & \text{if } d = 1, \\
\mathcal{O}_k\left(\frac{\theta}{L^2}\left(\log t\Big|_{1/L^2}^{1}\right)\right) = \mathcal{O}_k\left(\frac{\theta \log L}{L^2}\right) & \text{if } d = 2, \\
\mathcal{O}_k\left(\frac{\theta}{L^d}\left(\frac{t^{1-d/2}}{1-d/2}\Big|_{1/L^2}^{1}\right)\right) = \mathcal{O}_k\left(\frac{\theta}{L^2}\right) & \text{if } d \ge 3.
\end{cases}
\end{aligned}$$

The estimates above lead to what we want. □

Now we prove Lemma 3.55.

Proof of Lemma 3.55. The intuition is simple: an IP(2) evolves exactly like 2 independent SRWs when the two particles are not next to each other. The distance between 2 walks in IP(2) is forbidden to jump to 0, while the distance between two independent SRWs can. This makes the distance between 2 particles in $IP(2)$ always bigger than that of 2 independent SRWs in distribution.

More formally, note that $Y := (\mathcal{U}_1 - \mathcal{U}_2)$ and $\tilde{Y} := (\check{\mathcal{U}}_1 - \check{\mathcal{U}}_2)$ are also Markov processes. We only have to construct a coupling of Y and $\tilde{Y}$.

Let $\mathcal{L}$ (resp. $\tilde{\mathcal{L}}$) be the generator of Y (resp $\tilde{Y}$). Let $\mathfrak{e}_1, \ldots, \mathfrak{e}_d$ be the unit vectors in Λ_L^d. Then

$$\tilde{\mathcal{L}}(u, u') = \begin{cases} 2L^2 & \text{if } u' = u \pm \mathfrak{e}_i \text{ for some } i \in [d], \\ 0 & \text{if not,} \end{cases}$$

and

$$\mathcal{L}(u, u') = \begin{cases} 2L^2 & \text{if } u' = u \pm \mathfrak{e}_i \text{ for some } i \in [d] \text{ and } u' \neq 0, \\ L^2 & \text{if } (u, u') \in \{(\pm\mathfrak{e}_i, \mp\mathfrak{e}_i) | i \in [d]\}, \\ 0 & \text{if not,} \end{cases}$$

To construct the coupling kernel $\mathcal{L}^{coupling}$, we only need to construct the jump rates from each configuration $(u, \check{u})$ such that $u \preceq \check{u}$. Without loss of generality, we can suppose that $\mathsf{dist}\,(u(i), 0) \le \mathsf{dist}\,(\check{u}(i), 0)\,, \forall i \in [d]$. The kernel is as follows. For any coordinate i,

- If $\mathsf{dist}\,(u(i), 0) < \mathsf{dist}\,(\check{u}(i), 0)$,

$$\mathcal{L}^{coupling}((u, \check{u}), (u \pm \mathfrak{e}_i, \check{u})) = \mathcal{L}^{coupling}((u, \check{u}), (u, \check{u} \pm \mathfrak{e}_i)) = 2L^2.$$

- If $\mathsf{dist}\left(u(i),0\right)=\mathsf{dist}\left(\check{u}(i),0\right)$, then $u(i)=\pm\check{u}(i)$, and therefore there exists $\xi\in\{-1,1\}$ such that $\mathsf{dist}\left(u(i)\pm 1,0\right)=\mathsf{dist}\left(\check{u}(i)\pm\xi,0\right)$. Then

$$\mathcal{L}^{coupling}((u,\check{u}),(u\pm\mathfrak{e}_i,\check{u}\pm\xi\mathfrak{e}_i))=2L^2\mathbb{1}_{\{u\pm e_i\neq 0\}}.$$

If $u\in\{\pm\mathfrak{e}_i\}$, then we necessarily have $\check{u}\in\{\pm\mathfrak{e}_i\}$, and in this case,

$$\mathcal{L}^{coupling}((u,\check{u}),(-u,\check{u}))=L^2,$$
$$\mathcal{L}^{coupling}((u,\check{u}),(u,0))=2L^2.$$

For any $(u',\check{u}')\neq(u,\check{u})$ not listed in the cases above,

$$\mathcal{L}^{coupling}((u,\check{u}),(u',\check{u}'))=0.$$

We can verify that this gives us a coupling kernel for $(Y,\tilde{Y})$ that preserves the relation $\preceq$. This finishes our proof. □

Now we prove Lemma 3.50.

Proof of Lemma 3.50. We only need to prove the result for $\|u\|$ large enough. Let us consider a random walk $(U(t))_{t\geq 0}$ starting from u. Let k and K be two parameters such that $k\leq\|u\|\leq K$ that we will choose later. Let τ_k,τ_K be defined by

$$\tau_k=\inf\left\{t\geq 0:\|U(t)\|\leq k\right\},$$
$$\tau_K=\inf\left\{t\geq 0:\|U(t)\|\geq K\right\}.$$

Let $a(\cdot)$ be the potential kernel defined as in §4.4 in [57]. Note that $a(\cdot)$ is harmonic at every point except 0, and moreover, by Theorem 4.4.4 in [57],

$$\forall v\in\mathbb{Z}^2,\ a(v)=\frac{2}{\pi}\log\|v\|+\mathcal{O}\left(1\right). \tag{3.122}$$

Hence, $(a(U(t\wedge\tau_k\wedge\tau_K)))_{t\geq 0}$ is a bounded martingale. Note also that τ_k and τ_K are finite almost surely, thanks to the recurrence of the simple random walk in dimension two. So, we can apply the optional stopping theorem to conclude that

$$a(u)=\mathbb{E}\left[a(U(\tau_k\wedge\tau_K))\right].$$

Note that by (3.122), almost surely,

$$a(U_{\tau_k})=\frac{2}{\pi}\log k+\mathcal{O}\left(1\right),$$
$$a(U_{\tau_K})=\frac{2}{\pi}\log K+\mathcal{O}\left(1\right).$$

Hence

$$\begin{aligned} a(u) &= \mathbb{E}\left[\mathbb{1}_{\{\tau_k<\tau_K\}} a(U_{\tau_k})\right] + \mathbb{E}\left[\mathbb{1}_{\{\tau_K<\tau_k\}} a(U_{\tau_K})\right] \\ &= \mathbb{P}\left[\tau_k < \tau_K\right] \times \left(\frac{2}{\pi}\log k + \mathcal{O}\left(1\right)\right) + \mathbb{P}\left[\tau_K < \tau_k\right] \times \left(\frac{2}{\pi}\log K + \mathcal{O}\left(1\right)\right). \end{aligned}$$

Therefore,

$$\log \|u\| = \mathbb{P}\left[\tau_k < \tau_K\right] \times \left(\log k + \mathcal{O}\left(1\right)\right) + \left(1 - \mathbb{P}\left[\tau_k < \tau_K\right]\right) \times \left(\log K + \mathcal{O}\left(1\right)\right).$$

So

$$\mathbb{P}\left[\tau_k < \tau_K\right] = \frac{\log K - \log \|u\| + \mathcal{O}\left(1\right)}{\log K - \log k + \mathcal{O}\left(1\right)}.$$

We choose $k = 2$ and $K = \|u\|^{5/4}$. Then

$$\mathbb{P}\left[\tau_k < \tau_K\right] = \frac{\frac{\log \|u\|}{4} + \mathcal{O}\left(1\right)}{\frac{5\log \|u\|}{4} + \mathcal{O}\left(1\right)} \gtrsim \frac{1}{5}, \tag{3.123}$$

when $\|u\|$ is large enough. Let $U(t,1)$ and $U(t,2)$ (resp. $u(1), u(2)$) be two coordinates of $U(t)$ (resp. u). By the anticoncentration inequality, see e.g. Proposition 2.4.4 in [57] for the discrete-time version,

$$\max_{l\in\mathbb{Z}} \mathbb{P}\left[U(t,2) - u(2) = l\right] = \mathcal{O}\left(\frac{1}{\sqrt{t}}\right).$$

Therefore, by an union bound, for $t = \|u\|^3$,

$$\mathbb{P}\left[|U(t,2) - u(2)| \le K\right] = \mathcal{O}\left(\frac{K}{\sqrt{t}}\right) = \mathcal{O}\left(\frac{\|u\|^{5/4}}{\|u\|^{3/2}}\right) = \mathcal{O}\left(\|u\|^{-1/4}\right) < \frac{1}{10},$$

when $\|u\|$ is large enough. In this case,

$$\mathbb{P}\left[\left\|U(\|u\|^3, 2)\right\| \le K\right] \le \mathbb{P}\left[|U(t,2) - u(2)| \le K\right] \le 1/10,$$

and therefore,

$$\mathbb{P}\left[\tau_K \le \|u\|^3\right] \ge \mathbb{P}\left[\left\|U(\|u\|^3, 2)\right\| > K\right] \ge \frac{9}{10}. \tag{3.124}$$

The equations (3.123), (3.124) together imply that

$$\mathbb{P}\left[\tau_k \le \|u\|^3\right] \ge \frac{9}{10} + \frac{1}{5} - 1 = \frac{1}{10}.$$

Therefore

$$\mathbb{P}_u\left[\tau \leq 2\left\|u\right\|^3\right] \geq \mathbb{E}_u\left[\mathbb{1}_{\left\{\tau_k \leq \|u\|^3\right\}}\mathbb{P}_{U(\tau_k)}\left[\tau < \left\|u\right\|^3\right]\right]$$

Note that as k is a constant,

$$\mathbb{P}_{U(\tau_k)}\left[\tau < \left\|u\right\|^3\right] \geq \min_{\{v:\|v\|\leq k\}} \mathbb{P}_v\left[\tau < \left\|u\right\|^3\right] = 1 + o(1)$$

when $\|u\| \to \infty$. Therefore, when $\|u\|$ is large enough,

$$\mathbb{P}_u\left[\tau \leq 2\left\|u\right\|^3\right] \geq \beta,$$

for certain constant β. This finishes our proof. □

Chapter 4

The mean-field Zero-Range process

Goals

In this chapter, we study the mean-field Zero-Range process where the potential function r is increasing to infinity at sublinear speed, and the density of particles is bounded. We determine the mixing times of the system and establish cutoff. We also prove that the spectral gap is bounded away from zero and infinity. Our proof uses the path coupling method of Bubley and Dyer and stochastic calculus. The results presented here have been published in [97].

Contents

1 Introduction

1.1 Model

The Zero-Range process, introduced by Spitzer, is a model of interacting particle systems in continuous time. It describes the evolution of $m \geq 1$ indistinguishable particles jumping randomly across $n \geq 1$ sites, where the speed of a particle only depends on the number of its cooccupants (hence the name Zero-Range). More precisely, the interaction is represented by a function $r : \{1, 2, \dots\} \to (0, \infty)$, called *the potential function*, where $r(k)$ is the rate at which a site with k particles expels a particle. For convenience, we let $r(0) = 0$ (no jump from empty sites). In this chapter, we focus on the *mean-field* version of the model, where a jumping particle chooses its destination uniformly among all sites. More precisely, we consider a continuous-time Markov chain $X := (X_t)_{t\geq 0} = (X_t(1), X_t(2), \dots, X_t(n))_{t\geq 0}$ taking values in the state space

$$\mathcal{X} := \left\{ x \in \mathbb{Z}_+^n : \sum_{i=1}^{n} x(i) = m \right\}, \tag{4.1}$$

whose Markov generator $\mathcal{L}$ acts on an observable $\varphi : \mathcal{X} \to \mathbb{R}$ as follows:

$$(\mathcal{L}\varphi)(x) = \frac{1}{n} \sum_{1\leq i,j\leq n} r(x(i))(\varphi(x - \delta_i + \delta_j) - \varphi(x)). \tag{4.2}$$

Here $(\delta_i)_{1\leq i\leq n}$ denotes the canonical basis of $\mathbb{Z}_+^n$. The generator $\mathcal{L}$ is irreducible and reversible with respect to the following law:

$$\pi(x) \propto \prod_{i=1}^{n} \prod_{k=1}^{x(i)} \frac{1}{r(k)}, \tag{4.3}$$

with the convention that an empty product is 1. This implies that the process mixes. We recall that the *Dirichlet form* associated with our process is defined by:

$$\mathcal{E}(\varphi, \psi) := -\langle \varphi, \mathcal{L}\psi \rangle_\pi ,$$

where $\langle \varphi, \psi \rangle_\pi := \sum_{x\in\mathcal{X}} \pi(x)\varphi(x)\psi(x)$ denotes the usual inner-product in $L^2(\mathcal{X}, \pi)$. Then the *Poincaré constant*, denoted by λ_*, is defined by:

$$\lambda_* := \min \left\{ \frac{\mathcal{E}(\varphi, \varphi)}{\mathrm{Var}\,[\varphi]} \right\},$$

where the minimum is taken over all non-constant observables, and $\mathrm{Var}\,[\varphi]$ denotes the variance of φ under π. In our case, due to reversibility, λ_* coincides with the spectral gap defined in (1.2):

$$\lambda_* = \mathsf{gap}.$$

We favor the term Poincaré constant in this chapter because sometimes we discuss Poincaré constants of non-reversible models. The purpose of the present chapter is to estimate λ_* and

$t_{mix}(x;\epsilon)$, under certain assumptions on $r(\cdot)$.

Previous works. To the best of our knowledge, the total-variation mixing times of the Zero-Range process have only been studied in a few cases: the case where r is constant in [50], [51], [75], the case where r is non-decreasing and bounded in [40], and the somehow-trivial case of independent walkers where r is linear. Regarding the Poincaré constant, a notable result is given in [77], where Morris determines the order of magnitude of λ_* in the case where r is constant. Another result is obtained by Caputo in [12] for the case where r is homogeneously Lipschitz and increasing at infinity, i.e.

$$\sup_{k\geq 1} |r(k+1)-r(k)| < \infty, \tag{4.4}$$

$$\inf_{k-l\geq \delta} r(k)-r(l) > 0, \tag{4.5}$$

for some $\delta \in \mathbb{Z}_+$, where he proves that the Poincaré constant is bounded away from zero. In [39], Salez and Hermon prove a comparison principle that allows us to compare the Poincaré constant of many models with that of the mean-field model. We will use this principle below.

1.2 Main results

We consider the "intermediate" regime where the function r is non-decreasing and unbounded but grows slower than a linear function. More precisely, throughout the chapter, we assume that r satisfies:

$$r(k+1) \geq r(k),\, \forall k \in \mathbb{Z}_+, \tag{4.6}$$

$$\lim_{k\to\infty} r(k) = \infty, \tag{4.7}$$

$$\sup_{k\in\mathbb{Z}_+} \frac{r(k)}{k} < \infty. \tag{4.8}$$

We study the regime where the number of sites diverges while the density of particles per site remains bounded. More precisely, we always suppose that $m = m(n)$ and $x = x^{(n)}$, and all asymptotic statements refer to the regime:

$$n \to \infty, \qquad \frac{m}{n} \leq \rho, \tag{4.9}$$

where ρ is a positive constant. To lighten the notation, we keep the dependency upon n implicit as much as possible. By a *dimension-free* constant, we mean a number that depends only on r and ρ. Recall that our notation $\mathcal{O}(\cdot)$ (resp. $\Omega(\cdot)$, $\Theta(\cdot)$, $o(\cdot)$) means being upper bounded by (resp. lower bounded by, upper and lower bounded by, negligible compared to) the quantity inside the brackets up to a dimension-free prefactor. We define a function $R : \{1,2,\dots\} \to \mathbb{R}$ as follows:

$$\forall k \in \mathbb{Z}_+,\, R(k) = \sum_{i=1}^{k} \frac{1}{r(i)}. \tag{4.10}$$

Under condition (4.8), we easily see that

$$\lim_{k\to\infty} R(k) = \infty. \tag{4.11}$$

Our main result says that $R(\|x\|_\infty)$ is a good estimate for $\mathrm{t}_{\mathrm{mix}}(x;\epsilon)$.

Theorem 4.1 (Main result)**.** *For $\epsilon \in (0,1)$ fixed, for any initial state x,*

$$\mathrm{t}_{\mathrm{mix}}(x;\epsilon) \le (1+o(1))R(\|x\|_\infty) + \mathcal{O}(\log n). \tag{4.12}$$

In addition, if the initial state $x = x^{(n)}$ satisfies $\left\|x^{(n)}\right\|_\infty \xrightarrow[n\to\infty]{} \infty$, then

$$\mathrm{t}_{\mathrm{mix}}(x;\epsilon) \ge (1-o(1))R(\|x\|_\infty).$$

Maximizing over all initial states x, we obtain the following corollary.

Corollary 4.2 (Cutoff)**.** *Suppose additionally that $R(m) \gg \log n$. Then for $\epsilon \in (0,1)$ fixed,*

$$\frac{\mathrm{t}_{\mathrm{mix}}(\epsilon)}{R(m)} = 1 + o(1). \tag{4.13}$$

In other words, the system exhibits cutoff at time $R(m)$.

The class of functions r that satisfy conditions (4.6), (4.7), (4.8) is quite large. A natural example is when r is of the form $r(k) = k^\alpha$, $\forall k \in \mathbb{Z}_+$, for some $\alpha \in (0,1)$. In this case,

$$R(k) = (1+o(1))\frac{k^{1-\alpha}}{1-\alpha}, \text{ as } k \to \infty,$$

by the Stolz-Cesàro Theorem. Thereupon, a direct application of our result gives the following.

Example 4.3. *Suppose that $r(k) = k^\alpha$, $\forall k \in \mathbb{Z}$, for some $\alpha \in (0,1)$, and suppose that $m \gg (\log n)^{1/(1-\alpha)}$. Then the system exhibits cutoff at time $\dfrac{m^{1-\alpha}}{1-\alpha}$.*

Cutoff for the Zero-Range process was obtained in [75] for the case $r(k) = 1$ and more generally in [40] for the case where r is non-decreasing and bounded. Our work complements these results by investigating the case where $r \to \infty$. We also prove that the Poincaré constant is bounded away from zero and infinity:

Theorem 4.4 (Poincaré constant)**.** $\lambda_* = \Theta(1)$.

Thanks to the comparisons in the paper [39] of Hermon and Salez, we can extend this result to the more general case where a jumping particle chooses its destination according to a doubly stochastic matrix P rather than uniformly among all sites (for example, take P to be the transition matrix of random walk on a regular graph). More precisely, let P be an irreducible doubly stochastic transition matrix on $[n] := \{1, 2, \ldots, n\}$, and let $\mathcal{L}^P$ be the generator on $\mathcal{X}$ that acts on an observable $\varphi : \mathcal{X} \to \mathbb{R}$ by:

$$(\mathcal{L}^P\varphi)(x) = \sum_{1\le i,j\le n} r(x(i))P(i,j)(\varphi(x-\delta_i+\delta_j) - \varphi(x)).$$

Similarly, we can define the Poincaré constants $\lambda_*(P)$ and $\lambda_*(\mathcal{L}^P)$ of P and $\mathcal{L}^P$ via their associated Dirichlet forms and their stationary laws. Then we have the following.

Corollary 4.5 (Poincaré constant in arbitrary geometry)**.** $\lambda_*(\mathcal{L}^P) = \Theta(\lambda_*(P))$.

We give an example where we can compute $\lambda_*(P)$ explicitly to obtain an explicit estimate on $\lambda_*(\mathcal{L}^P)$.

Example 4.6 (Poincaré constant of torus model)**.** *Let P be the transition matrix of the simple random walk on the lattice $\mathbb{Z}^d/p\mathbb{Z}^d$, for some $p, d \in \mathbb{Z}_+$. Then $\lambda_*(\mathcal{L}^P) = \Theta(1/(dp^2))$, as $p^d \to \infty$.*

The calculation of $\lambda_*(P)$ is deferred to the end of the chapter.

Heuristics. If we ignore arrivals and only consider departures of particles, then $R(\|x\|_\infty)$ is exactly the expectation of the time it takes for the initially highest site to be emptied. In the true system, due to the conditions imposed on r, the arrival rate at each site is uniformly bounded. Consequently, $R(\|x\|_\infty)$ remains a good approximation for the emptying time. For the lower bound, we prove that before time $R(\|x\|_\infty)$, the initially highest site still has too many particles, and hence the system has not yet reached equilibrium. For the upper bound, we will see that at time $t = (1 + o(1))R(\|x\|_\infty) + \mathcal{O}(\log n)$, $\|X_t\|_\infty = \mathcal{O}(\log n)$. Afterwards, the system quickly reaches equilibrium.

2 The lower bound

2.1 Preliminaries

We will use the following two graphical constructions of the process X.

Graphical construction 1. Let Ξ be a Poisson point process of intensity $\frac{1}{n}\mathrm{d}t \otimes \mathrm{d}u \otimes \mathrm{Card} \otimes \mathrm{Card}$ on $[0,\infty) \times [0,\infty) \times [n] \times [n]$, where Card denotes the counting measure. Define the piece-wise constant process $X = (X_t)_{t\geq 0}$ taking values in $\mathcal{X}$ as follows: $X_0 = x$, and for each point (t, u, i, j) of Ξ,

$$X_t := \begin{cases} X_{t-} - \delta_i + \delta_j, & \text{if } u \leq r(X_{t-}(i)) \\ X_{t-} & \text{otherwise.} \end{cases} \tag{4.14}$$

Then X is a càdlàg Markov process starting from x with generator $\mathcal{L}$.

Graphical construction 2. Let Ψ be a Poisson point process of intensity $\mathrm{d}t \otimes \mathrm{d}u \otimes \mathrm{Card}$ on $[0,\infty) \times [0,\infty) \times [n]$. Consider the piece-wise constant process $X = (X_t)_{t\geq 0}$ which starts at $X_0 = x$ and has the following jumps: for each point (t, u, j) of Ψ,

$$X_t := \begin{cases} X_{t-} - \delta_i + \delta_j, & \text{if } \frac{1}{n}\sum_{k=1}^{i-1} r\left(X_{t-}(k)\right) < u \leq \frac{1}{n}\sum_{k=1}^{i} r(X_{t-}(k)), \\ & \text{for some } i \in [n], \\ X_{t-} & \text{otherwise.} \end{cases} \tag{4.15}$$

Then X is also a Markov process starting from x with generator $\mathcal{L}$. We can view the Poisson process in Graphical Construction 2 as the repartition of the Poisson process in Graphical Construction 1 according to the destination of the jumps.

Filtration. We always note $(\mathcal{F}_t)_{t\geq 0}$ *the filtration* generated by the Poisson processes in the graphical construction we are using, where $\mathcal{F}_t$ is the $\sigma - algebra$ generated by these processes up to time t. It is immediate that the process X is adapted to the filtration and has càdlàg trajectories.

Mean-field jump rate. At any time t, as the model is mean-field, the arrival rate at each site is the same. We denote this quantity by ζ_t:

$$\zeta_t := \frac{1}{n}\sum_{j=1}^{n} r(X_t(j)).$$

Condition (4.8) implies that

$$\zeta_t \leq \left(\frac{1}{n}\sum_{j=1}^{n} X_t(j)\right) \sup_{k\in\mathbb{Z}_+} \frac{r(k)}{k} \leq \rho \sup_{k\in\mathbb{Z}_+} \frac{r(k)}{k} =: \kappa. \tag{4.16}$$

Hence, the number of particles arriving at each site is stochastically dominated by a Poisson process of dimension-free intensity κ.

Martingale associated with an observable. For any observable $\varphi : \mathcal{X} \to \mathbb{R}$, under $\mathbb{P}_x$, the process $M = (M_t)_{t\geq 0}$ given by

$$M_t := \varphi(X_t) - \varphi(x) - \int_0^t \mathcal{L}\varphi(X_u)\mathrm{d}u \tag{4.17}$$

is a zero-mean martingale, see e.g Theorem 3.32 in [62]. Let φ_1, φ_2 be two observables, and let $M^{(1)}$, $M^{(2)}$ be the associated martingales. Then the *predictable covariation* of $M^{(1)}$ and $M^{(2)}$ is given by

$$\left\langle M^{(1)}, M^{(2)}\right\rangle_t = \int_0^t \sum_{y\in\mathcal{X}} \mathcal{L}(X_u, y)\left(\varphi_1(y) - \varphi_1(X_u)\right)\left(\varphi_2(y) - \varphi_2(X_u)\right) du. \tag{4.18}$$

We recall a lemma on the concentration of martingales with jumps (see [35]):

Lemma 4.7 (Concentration of martingale)**.** *Let $(M_t)_{t\geq 0}$ be a zero-mean càdlàg martingale w.r.t a filtration that satisfies the usual conditions. Suppose that $M_0 = 0$ and $M_t - M_{t-} \leq K$ for all $t > 0$ and some $0 \leq K < \infty$. Then for each $a > 0, b > 0$,*

$$\mathbb{P}\left[\exists t \geq 0 : M_t \geq a, \langle M, M\rangle_t \leq b^2\right] \leq \exp\left[-\frac{a^2}{2(aK + b^2)}\right]. \tag{4.19}$$

Gain/loss at a site. We will need the following quantities:

1. For $i \in [n]$, let $G_i(t)$ be the counting process that counts the number of particles arriving at site i up to time t. We call G_i *the gain* at site i.

2. Let $L_i(t)$ be the counting process that counts the number of particles jumping out of site i up to time t. We call L_i *the loss* at site i.

Obviously,

$$X_t(i) = X_0(i) + G_i(t) - L_i(t).$$

2.2 Elementary concentration inequalities

We list here some useful inequalities, whose proofs are simple applications of Chernoff's bound (see, e.g. [9] for more details on Chernoff's bound).

Lemma 4.8 (Poisson concentration). *Let Z be a Poisson variable. Then for any $B > 1$,*

$$\mathbb{P}\left[Z \geq B\mathbb{E}Z\right] \leq e^{-(1+B\ln B-B)\mathbb{E}Z}.$$

Lemma 4.9 (Concentration of sum of independent exponential variables). *Let $\lambda_1, \ldots, \lambda_k$ be positive numbers, and let $\xi_1, \ldots, \xi_k$ be independent random variables such that $\xi_i \sim \exp(\lambda_i)$, $\forall 1 \leq i \leq k$. Let*

$$S := \sum_{i=1}^{k} \xi_i.$$

For $B > 0$ arbitrary, we have the following inequalities:

$$\mathbb{P}\left[S - \mathbb{E}S \leq -\sqrt{\mathrm{Var}\left[S\right]}B\right] \leq e^{-B^2/4}, \tag{4.20}$$

and

$$\mathbb{P}\left[S - \mathbb{E}S \geq \lambda \mathrm{Var}\left[S\right] + \frac{B}{\lambda}\right] \leq e^{-B/2}, \tag{4.21}$$

where $\lambda = \min_{1 \leq i \leq k} \{\lambda_i\}$.

2.3 Proof of the lower bound

In this subsection, we prove the lower bound on $\mathrm{t}_{\mathrm{mix}}(x; \epsilon)$ in Theorem 4.1. First, we analyze the law of a single site at equilibrium. We still denote by X_∞ a random variable following the law π.

Proposition 4.10 (Single site marginals at equilibrium). *There exists a dimension-free constant $q > 0$ such that*

$$\forall k \in \mathbb{Z}_+, \quad \frac{\mathbb{P}\left[X_\infty(1) = k\right]}{\mathbb{P}\left[X_\infty(1) = k-1\right]} < \frac{q}{r(k)}.$$

Proof. Let x be an arbitrary configuration in $\mathcal{X}$ such that $x(1) \geq 1$. Thanks to (4.9), the number of sites that have at least 2ρ particles is at most $\frac{n}{2}$, and hence the number of sites that have less than 2ρ particles is at least $\frac{n}{2}$. For any $l \in [n] \setminus \{1\}$ such that $x(l) < 2\rho$, thanks to (4.6),

$$\frac{\pi(x)}{\pi(x-\delta_1+\delta_l)} = \frac{r(x(l)+1)}{r(x(1))} \leq \frac{r(\lceil 2\rho \rceil)}{r(x(1))}.$$

Taking average over all sites l such that $x(l) < 2\rho$, we obtain

$$\begin{aligned}\pi(x) &\leq \frac{r(\lceil 2\rho \rceil)}{r(x(1))}\frac{2}{n} \sum_{\substack{l\neq 1\\ x(l)<2\rho}} \pi(x-\delta_1+\delta_l) \\ &\leq \frac{r(\lceil 2\rho \rceil)}{r(x(1))}\frac{2}{n} \sum_{l\neq 1} \pi(x-\delta_1+\delta_l).\end{aligned}$$

Now we take the sum over all x such that $x(1) = k$ to get:

$$\begin{aligned}\mathbb{P}\left[X_\infty(1)=k\right] &\leq \frac{r(\lceil 2\rho \rceil)}{r(k)}\frac{2}{n} \sum_{x(1)=k}\sum_{l\neq 1} \pi(x-\delta_1+\delta_l) \\ &= \frac{r(\lceil 2\rho \rceil)}{r(k)}\frac{2}{n} \sum_{l\neq 1}\sum_{x(1)=k} \pi(x-\delta_1+\delta_l) \\ &= \frac{r(\lceil 2\rho \rceil)}{r(k)}\frac{2}{n} \sum_{l\neq 1}\sum_{\substack{z(1)=k-1,\\ z(l)\geq 1}} \pi(z) \qquad \text{(change of variable: } z = x-\delta_1+\delta_l) \\ &\leq \frac{r(\lceil 2\rho \rceil)}{r(k)}\frac{2}{n} \sum_{l\neq 1}\sum_{z(1)=k-1} \pi(z) \\ &= \frac{r(\lceil 2\rho \rceil)}{r(k)}\frac{2}{n} \sum_{l\neq 1} \mathbb{P}\left[X_\infty(1)=k-1\right] \\ &\leq \frac{r(\lceil 2\rho \rceil)}{r(k)}\frac{2n}{n} \mathbb{P}\left[X_\infty(1)=k-1\right] \\ &= \frac{q}{r(k)} \mathbb{P}\left[X_\infty(1)=k-1\right],\end{aligned}$$

where $q = 2r(\lceil 2\rho \rceil)$. ☐

Now we study the effect of arrivals at a particular site. Recall that L_i denotes the loss at site i, as defined at the end of subsection 2.1.

Lemma 4.11 (Effect of arrivals)**.** *Let x be an arbitrary initial configuration. For $i \in [n]$ and $h \in \mathbb{Z}_+$ such that $x(i) \geq h$, there exist independent variables $U_k \sim \exp\left(\frac{n-1}{n} r(k)\right)$, $h \geq k \geq 1$, such that for any $0 \leq k \leq h-1$,*

1. *$T_k := U_h + U_{h-1} + \cdots + U_{h-k}$ is a stopping time,*
2. *almost surely, $L_i(T_k) \geq k+1$,*
3. *almost surely, $X_t(i) \geq h-k$, $\forall t \in [0, T_k)$.*

The intuition is as follows: if we ignore arrivals and consider only departures at site i, then T_k is the time to have $k+1$ particles depart from i. Arrivals, on the one hand, slow down a site from being emptied, but on the other hand, accelerate the rate of expelling and hence increase the loss. So the inequalities at points 2 and 3 follow.

Proof. We use the graphical construction 1. We first prove for the case $k = 0$. Let U_h be defined by

$$U_h = \inf\left\{t \geq 0 : \Xi\Big([0,t] \times [0, r(h)] \times \{i\} \times ([n] \setminus \{i\})\Big) > 0\right\}.$$

It is clear that U_h is a stopping time and $U_h \sim \exp\left(\frac{n-1}{n} r(h)\right)$. Moreover, by definition of U_h,

$$\Xi\Big([0, U_h) \times [0, r(h)] \times \{i\} \times ([n] \setminus \{i\})\Big) = 0,$$

so before time U_h, $X(i)$ cannot fall from h to $h-1$. In other words, $\forall t \in [0, U_h)$, $X_t(i) \geq h$. In particular, $X_{U_h-}(i) \geq h$, so there should be a jump from site i to $[n] \setminus \{i\}$ at U_h. Consequently, $L_i(U_h) \geq 1$ and $X_{U_h}(i) = X_{U_h-}(i) - 1 \geq h - 1$, almost surely, which finishes the case $k = 0$. Now we define U_k inductively by:

$$U_k = \inf\left\{t \geq 0 : \Xi\Big((T_{h-k-1}, T_{h-k-1} + t] \times [0, r(k)] \times \{i\} \times ([n] \setminus \{i\})\Big) > 0\right\},$$

where $T_{h-k-1} = U_h + U_{h-1} + \cdots + U_{k+1}$. The variables $(U_k)_{h \geq k \geq 1}$ are independent by the stationary and independent increments of Poisson processes. In addition,

$$U_k \sim \exp\left(\frac{n-1}{n} r(k)\right).$$

The claim is simply obtained by induction and by the strong Markov property. □

Useful variables. Lemma 4.11 allows us to compare certain random times with the random variables $(S_k)_{k \geq 1}$ defined by

$$S_k = \sum_{i=1}^{k} \xi_i, \tag{4.22}$$

where $(\xi_i)_{i \geq 1}$ is a sequence of independent random variables such that $\xi_i \sim \exp(r(i))$, $\forall i \in \mathbb{Z}_+$. Obviously, $\mathbb{E}[S_k] = \sum_{i=1}^{k} \frac{1}{r(i)} = R(k)$, $\mathrm{Var}[S_k] = \sum_{i=1}^{k} \frac{1}{r(i)^2}$. Due to (4.6), (4.7), (4.8), the functions r, R diverge, so we easily see that

$$\lim_{k \to \infty} \frac{\mathrm{Var}[S_k]}{\mathbb{E} S_k} = 0. \tag{4.23}$$

Lemma 4.11 makes precise the fact that arrivals can only slow down a site from being emptied, while Proposition 4.10 together with condition (4.7) say that at equilibrium, the typical height of a site cannot be very large. This leads to the lower bound on $\mathrm{t_{mix}}(x;\epsilon)$ in Theorem 4.1:

Proof of the lower bound in Theorem 4.1. Let $\delta \in (0,1)$ be fixed, and let $x \in \mathcal{X}$ be arbitrary.

We only need to prove that for $\|x\|_\infty$ sufficiently large,

$$\frac{\mathrm{t}_{\mathrm{mix}}(x;\epsilon)}{R(\|x\|_\infty)} > 1-\delta. \tag{4.24}$$

Without loss of generality, suppose that site 1 is originally the highest, i.e. $\|x\|_\infty = x(1)$. Our distinguishing statistic is $\varphi : x \mapsto x(1)$. We know that for any $A \subset \mathcal{X}$,

$$\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \pi\right) \geq \mathbb{P}_x\left[X_t \in A\right] - \pi(A).$$

We choose

$$A = \{y \in \mathcal{X} : y(1) \geq k\},$$

where

$$k = \sup\left\{l \in \mathbb{Z}_+ : |R(l)| \leq \frac{\delta}{2}R(\|x\|_\infty)\right\}.$$

We only need to show that $\pi(A) = o(1)$, and for $t = (1-\delta)R(x(1))$, $\mathbb{P}_x\left[X_t \in A\right] = 1 - o(1)$. Thanks to (4.11), $\|x\|_\infty \gg 1$ ensures that $R(\|x\|_\infty) \gg 1$ and hence $k \gg 1$, so by Proposition 4.10 and (4.7), $\mathbb{P}\left[X_\infty \in A\right] = o(1)$. On the other hand, we apply Lemma 4.11 with $i = 1, h = x(1)$ to conclude that there exists a stopping time

$$T_{x(1)-k-1} \overset{(d)}{=} \frac{n}{n-1}(S_{x(1)} - S_k),$$

where the sequence $(S_i)_{i\geq 1}$ is defined in (4.22), such that $X_t(1) \geq k+1$, $\forall t \in [0, T_{x(1)-k-1})$. We define $S_{k,x(1)} = S_{x(1)} - S_k$, for any $k < x(1)$. Then

$$\begin{aligned}
\mathbb{P}_x\left[X_{(1-\delta)R(x(1))} \in A\right] &\geq \mathbb{P}_x\left[T_{x(1)-k-1} > (1-\delta)R(x(1))\right] \\
&\geq \mathbb{P}\left[S_{k,x(1)} > (1-\delta)R(x(1))\right] \\
&= 1 - \mathbb{P}\left[S_{k,x(1)} \leq (1-\delta)R(x(1))\right].
\end{aligned}$$

Moreover, $R(k) \leq \frac{\delta}{2}R(x(1))$ by definition of k, and hence $\mathbb{E}\left[S_{k,x(1)}\right] = R(x(1)) - R(k) \geq (1-\delta/2)R(x(1))$. So by the concentration inequality (4.20),

$$\begin{aligned}
\mathbb{P}\left[S_{k,x(1)} \leq (1-\delta)R(x(1))\right] &\leq \mathbb{P}\left[S_{k,x(1)} - \mathbb{E}S_{k,x(1)} \leq -\frac{\delta}{2}R(x(1))\right] \\
&\leq \exp\left(-\frac{1}{4}\cdot\frac{\delta^2}{4}R(x(1))^2\mathrm{Var}\left[S_{k,x(1)}\right]^{-1}\right) \\
&\leq \exp\left(-\frac{\delta^2}{16}R(x(1))\mathrm{Var}\left[S_{x(1)}\right]^{-1}\right). \\
&= o(1),
\end{aligned}$$

where in the third inequality we have used the fact that $\mathrm{Var}\left[S_{x(1)}\right] > \mathrm{Var}\left[S_{k,x(1)}\right]$, and in the last equality we have used (4.23) and the fact that $x(1) = \|x\|_\infty \xrightarrow{n\to\infty} \infty$. Hence $\mathbb{P}_x\left[X_t \in A\right] =$

$1 - o(1)$, which finishes our proof. $\square$

3 The upper bound

We will prove the following statements:

Proposition 4.12 (Dissolution)**.** *There exist dimension-free constants σ, α_1 such that for any $\delta \in (0,1)$ fixed, for any $x \in \mathcal{X}$, for any $t \geq (1+\delta)R(\|x\|_\infty) + \sigma \log n$,*

$$\mathbb{P}_x\left[\|X_t\|_\infty \geq \alpha_1 \log n\right] = \mathcal{O}\left(n^{-2}\right).$$

Proposition 4.13 (Quick convergence to equilibrium)**.** *Let α_1 be defined as in Proposition 4.12. Then there exists a dimension-free constant α so that for any configuration x such that $\|x\|_\infty \leq \alpha_1 \log n$,*

$$\mathrm{d}_{TV}\left(\mathbb{P}_x\left[X_{\alpha \log n} \in \cdot\right], \pi\right) = \mathcal{O}\left(n^{-2}\right).$$

First we see how these propositions lead to the upper bound on $\mathrm{t_{mix}}(x;\epsilon)$ in Theorem 4.1:

Proof of the upper bound in Theorem 4.1: Let σ, α_1, α be defined as in Proposition 4.12 and Proposition 4.13. Let x be an arbitrary configuration; let $t_1 = (1+\delta)R(\|x\|_\infty) + \sigma \log n$, for some δ, $t_2 = \alpha \log n$, and $t = t_1 + t_2$. We only need to prove that for arbitrary $\delta \in (0,1)$ fixed, for n large enough, for any $x \in \mathcal{X}$,

$$\mathrm{t_{mix}}(x;\epsilon) \leq t.$$

By the convexity of the total variation distance,

$$\begin{aligned}\mathrm{d_{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \pi\right) &\leq \sum_{y \in \mathcal{X}} \mathbb{P}_x\left[X_{t_1} = y\right] \mathrm{d_{TV}}\left(\mathbb{P}_y\left[X_{t_2} \in \cdot\right], \pi\right) \\ &\leq \mathbb{P}_x\left[\|X_{t_1}\|_\infty \geq \alpha_1 \log n\right] + \max_{\|y\|_\infty \leq \alpha_1 \log n} \mathrm{d_{TV}}\left(\mathbb{P}_y\left[X_{t_2} \in \cdot\right], \pi\right),\end{aligned}$$

which is $\mathcal{O}\left(n^{-2}\right)$ by Proposition 4.12 and Proposition 4.13, hence smaller than ϵ when n is large enough. $\square$

The structure of the rest of this section is as follows. In Subsection 3.1, we prove Proposition 4.12 and provide some analysis on the trajectory of the system which will be used in the proof of Proposition 4.13. In Subsection 3.2, we prove Proposition 4.13 by the path coupling method of Bubley and Dyer.

3.1 Dissolution

Recall that G_i denotes the gain at site i, as defined at the end of subsection 2.1. First we give an estimate on G_i at the predicted time.

Lemma 4.14 (Estimating the gain at a site)**.** *For any dimension-free constant d, there exists a dimension-free constant c_0 such that for $t = (1+\delta/4)(R(\|x\|_\infty) + d\log n)$, for any $\delta \in (0,1)$ fixed,*

$$\mathbb{P}_x\left[G_i(t) \geq c_0(R(\|x\|_\infty) \vee \log n)\right] = \mathcal{O}\left(n^{-5}\right). \tag{4.25}$$

Proof. We use the graphical construction 2. Since $\zeta_t < \kappa$ at all time (see (4.16)), G_i is simply dominated by the Poisson process $\Psi|_{\cdot\times[0,\kappa]\times\{i\}}$. Consequently, at time $t = (1+\delta/4)(R(\|x\|_\infty) + d\log n)$, $G_i(t)$ is stochastically dominated by a random variable Y, with

$$Y \sim \text{Poisson}\Big(\kappa(1+\delta/4)(R(\|x\|_\infty) + d\log n)\Big).$$

Then the result is simply obtained by Lemma 4.8, for c_0 large enough. □

For any site i, Lemma 4.11 says that arrivals can only accelerate the loss L_i, while Lemma 4.14 gives us a (random) upper bound on G_i. They together lead to the following proposition:

Proposition 4.15 (First phase dissolution)**.** *Let $d = \dfrac{7}{r(1)}$, and let c_0 be defined as in Lemma 4.14. Then for any $x \in \mathcal{X}$, for any $\delta \in (0,1)$ fixed, for $t = (1+\delta/4)(R(\|x\|_\infty) + d\log n)$,*

$$\mathbb{P}_x\left[\|X_t\|_\infty \leq c_0(R(\|x\|_\infty) \vee \log n)\right] = 1 - \mathcal{O}\left(n^{-2}\right). \tag{4.26}$$

Proof. Let $i \in [n]$. We apply Lemma 4.11 with $h = x(i)$ to conclude that there exists a stopping time $T_{x(i)-1} \overset{(d)}{=} \dfrac{n}{n-1} S_{x(i)}$ such that $L_i(T_{x(i)-1}) \geq x(i)$. Note that $S_{x(i)}$ is dominated stochastically by $S_{\|x\|_\infty}$, hence by using (4.21), we deduce that

$$\begin{aligned}
&\mathbb{P}\left[T_{x(i)-1} \geq \frac{n}{n-1}\left(\mathbb{E}\left[S_{\|x\|_\infty}\right] + r(1)\text{Var}\left[S_{\|x\|_\infty}\right] + \frac{6\log n}{r(1)}\right)\right] \\
&= \mathbb{P}\left[S_{x(i)} \geq \left(\mathbb{E}\left[S_{\|x\|_\infty}\right] + r(1)\text{Var}\left[S_{\|x\|_\infty}\right] + \frac{6\log n}{r(1)}\right)\right] \\
&\leq \mathbb{P}\left[S_{\|x\|_\infty} \geq \left(\mathbb{E}\left[S_{\|x\|_\infty}\right] + r(1)\text{Var}\left[S_{\|x\|_\infty}\right] + \frac{6\log n}{r(1)}\right)\right] \\
&\leq n^{-3}.
\end{aligned}$$

Note that $\text{Var}\left[S_{\|x\|_\infty}\right] = o(R(\|x\|_\infty) \vee \log n)$ by (4.23), and hence

$$\frac{n}{n-1}\left(\mathbb{E}\left[S_{\|x\|_\infty}\right] + r(1)\text{Var}\left[S_{\|x\|_\infty}\right] + \frac{6\log n}{r(1)}\right) < t,$$

when n is large enough. Consequently, for n large enough,

$$\mathbb{P}_x\left[L_i(t) < x(i)\right] \leq \mathbb{P}_x\left[T_{x(i)-1} > t\right] \leq n^{-3}.$$

We take a union bound of this and the event in (4.25) over all sites to conclude that

$$\mathbb{P}\left[\exists i,\, G_i(t) \geq c_0(R(\|x\|_\infty) \vee \log n)\right] + \mathbb{P}\left[\exists i,\, L_i(t) < x(i)\right] = \mathcal{O}\left(n^{-2}\right).$$

The claim follows. □

We now recall a simple version of Gronwall's lemma that we will use a lot:

Lemma 4.16 (Gronwall's lemma)**.** *Let α, β be some positive numbers. Let $u : [0, \infty) \to \mathbb{R}^+$ be a continuously differentiable function such that $\frac{d}{dt}u(t) \leq -\beta u(t) + \alpha$. Then*

$$u(t) < \frac{\alpha}{\beta} + \left(u(0) - \frac{\alpha}{\beta}\right) e^{-\beta t}.$$

In particular, if $t \geq \frac{\log u(0)}{\beta}$, then $u(t) < \frac{\alpha}{\beta} + 1$.

For θ a positive number, and for $i \in [n]$, we define the observable $\varphi_i^\theta : \mathcal{X} \to \mathbb{R}$ by

$$\varphi_i^\theta(x) = e^{\theta x(i)}, \tag{4.27}$$

and we define the observable φ^θ by

$$\varphi^\theta(x) = \frac{1}{n}\sum_{i=1}^{n} \varphi_i^\theta(x).$$

Lemma 4.17 (Estimate on $\mathcal{L}\varphi^\theta$)**.** *For any dimension-free constants $\theta, \beta > 0$, there exists a number $L = L(\theta, \beta)$ such that, for any configuration x,*

$$\mathcal{L}\varphi^\theta(x) \leq -\beta\varphi^\theta(x)\mathbb{1}_{\{\varphi^\theta(x)>L\}} + (e^\theta - 1)\kappa\varphi^\theta(x)\mathbb{1}_{\{\varphi^\theta(x)\leq L\}}. \tag{4.28}$$

In particular,

$$\mathcal{L}\varphi^\theta(x) \leq -\beta\varphi^\theta(x) + ((e^\theta - 1)\kappa + \beta)L. \tag{4.29}$$

Proof. For simplicity, we write φ instead of φ^θ and φ_i instead of φ_i^θ. It is not difficult to see that

$$\begin{aligned}
\frac{\mathcal{L}\varphi_i(x)}{\varphi_i(x)} &= \frac{e^\theta - 1}{n}\sum_{j\in[n]\setminus\{i\}} r(x(j)) - \frac{1 - e^{-\theta}}{n}\sum_{j\in[n]\setminus\{i\}} r(x(i)) \\
&= \frac{e^\theta - 1}{n}\sum_{j\in[n]} r(x(j)) - \left((1 - e^{-\theta})\frac{n-1}{n} + \frac{e^\theta - 1}{n}\right) r(x(i)).
\end{aligned}$$

Hence, by (4.16),

$$\mathcal{L}\varphi_i(x) \leq (e^\theta - 1)\kappa\varphi_i(x) - (1 - e^{-\theta})\varphi_i(x) r(x(i)).$$

Taking the average over all sites i we get

$$\mathcal{L}\varphi(x) \le (e^\theta - 1)\kappa\varphi(x) - (1-e^{-\theta})\frac{\sum_{i\in[n]} r(x(i))\varphi_i(x)}{n}.$$

The claim follows when $\varphi(x) \le L$. It remains to consider the case $\varphi(x) > L$. For any $c \in \mathbb{Z}_+$, $r(x(i))\varphi_i(x) \ge r(c)(\varphi_i(x) - e^{\theta c})$ due to the monotonicity of r, hence

$$\begin{aligned}\sum_{i\in[n]} r(x(i))\varphi_i(x) &\ge r(c)\sum_{i\in[n]}(\varphi_i(x) - e^{\theta c})\\ &\ge r(c)(n\varphi(x) - ne^{\theta c}).\end{aligned}$$

Consequently,

$$\mathcal{L}\varphi(x) \le (e^\theta - 1)\kappa\varphi(x) - (1-e^{-\theta})r(c)(\varphi(x) - e^{\theta c}).$$

Let $L = L(c) = 2e^{\theta c}$. If $\varphi(x) > L$, then $\varphi(x) - e^{\theta c} > \dfrac{\varphi(x)}{2}$, which implies:

$$\mathcal{L}\varphi(x) \le (e^\theta - 1)\kappa\varphi(x) - (1-e^{-\theta})r(c)\frac{\varphi(x)}{2}. \tag{4.30}$$

We can take c large enough to make the right-hand side of the inequality above smaller than $-\beta\varphi(x)$, which finishes the proof of (4.28). (4.29) is obtained by rewriting (4.28) as follows,

$$\begin{aligned}\mathcal{L}\varphi(x) &\le -\beta\varphi(x)(1-\mathbb{1}_{\{\varphi(x)\le L\}}) + (e^\theta-1)\kappa\varphi(x)\mathbb{1}_{\{\varphi(x)\le L\}}\\ &= -\beta\varphi(x) + ((e^\theta-1)\kappa+\beta)\varphi(x)\mathbb{1}_{\{\varphi(x)\le L\}}\\ &\le -\beta\varphi(x) + ((e^\theta-1)\kappa+\beta)L,\end{aligned}$$

which is what we want. □

A good estimate on $\mathcal{L}\varphi^\theta$ will guarantee the good behavior of the trajectories of X, as stated in the following proposition.

Proposition 4.18 (Dissolution). *Let $\theta, \beta > 0$ be some dimension-free constants. Then for any $x \in \mathcal{X}$, for any $t \ge \dfrac{\theta}{\beta}\|x\|_\infty$,*

$$\mathbb{P}_x\left[\|X_t\|_\infty \ge \frac{6}{\theta}\log n\right] = \mathcal{O}\left(n^{-5}\right). \tag{4.31}$$

Proof. We still write φ instead of φ^θ, for simplicity. Let $L = L(\theta,\beta)$ be defined as in Lemma 4.17. Let $u(t) = \mathbb{E}_x[\varphi(X_t)]$. By (4.17),

$$\frac{d}{dt}u(t) = \mathbb{E}_x[\mathcal{L}\varphi(X_t)].$$

This and (4.29) imply:

$$\frac{d}{dt}u(t) \leq -\beta u(t) + ((e^{\theta} - 1)\kappa + \beta)L.$$

Therefore, by Lemma 4.16,

$$u(t) \leq \frac{((e^{\theta} - 1)\kappa + \beta)L}{\beta} + u(0)e^{-\beta t}.$$

Hence for $t \geq \frac{\theta}{\beta}\|x\|_{\infty}$, $u(t) \leq \frac{((e^{\theta}-1)\kappa+\beta)L}{\beta} + 1$. Note that $e^{\theta\|x\|_{\infty}} \leq n\varphi(x)$, hence $\mathbb{E}\left[e^{\theta\|X_t\|_{\infty}}\right] \leq nu(t)$. Then the claim is a simple consequence of Chernoff's bound. □

We now prove Proposition 4.12:

Proof of Proposition 4.12. We fix $\delta \in (0,1)$. Let c_0 be defined as in Lemma 4.14. Let $\theta_1 > 0$ be fixed, and let β_1 be such that $\frac{\theta_1 c_0}{\beta_1} \leq \frac{\delta}{2}$. Let $L_1 = L(\theta_1, \beta_1)$ as in Lemma 4.17. Let $t_1 = (1+\delta/4)(R(\|x\|_{\infty}) + d\log n)$, $t_2 = \frac{\delta}{2}(R(\|x\|_{\infty}) \vee \log n)$. By definition of β_1 and Proposition 4.15,

$$\mathbb{P}_x\left[\frac{\theta_1\|X_{t_1}\|_{\infty}}{\beta_1} \leq t_2\right] = 1 - \mathcal{O}\left(n^{-2}\right).$$

By Markov property at time t_1 and inequality (4.31) , we conclude that for any $t \geq t_1 + t_2$,

$$\mathbb{P}_x\left[\|X_t\|_{\infty} \geq \frac{6}{\theta_1}\log n\right] = \mathcal{O}\left(n^{-2}\right).$$

We choose $\alpha_1 = \frac{6}{\theta_1}$ and $\sigma = \frac{5}{4}d + \frac{1}{2}$ to conclude the proof. □

The estimate on $\mathcal{L}\varphi^{\theta}$ in Lemma 4.17 also ensures that the system quickly reaches the set where φ^{θ} is small.

Proposition 4.19 (Exponential moment of hitting time)**.** *Let $\theta, \beta > 0$ be some dimension-free constants, and let $L = L(\theta, \beta)$ be defined as in Lemma 4.17. Let T be the hitting time of the set $\{x \in \mathcal{X} : \varphi^{\theta}(x) \leq L\}$. Then for any $x \in \mathcal{X}$,*

$$\mathbb{E}_x\left[e^{\beta T}\right] \leq e^{\theta}\varphi^{\theta}(x)/L. \tag{4.32}$$

Proof. We write φ instead of φ^{θ}. The idea is that if φ is large, then the drift $\mathcal{L}\varphi$ is negative, and its magnitude is of the same order as φ. Hence the system will quickly reach the set where φ is small, which will be made precise by stochastic calculus. Consider the function $F : \mathbb{R}^2 \to \mathbb{R}$ defined by $F(u,v) = uv$, which is twice continuously differentiable. By (4.17),

$$\varphi(X_t) = \varphi(X_0) + \int_0^t \mathcal{L}\varphi(X_u)\mathrm{d}u + M_t,$$

where $(M_t)_{t\geq 0}$ is a martingale. Moreover, $\varphi(X)$ is a pure-jump process since X is piece-wise constant. For a càdlàg process Y, we denote by ΔY its jumps: $\Delta Y(s) = Y(s) - Y(s-)$. We define $G(t) = F(e^{\beta t}, \varphi(X_t))$. Note that the function $t \mapsto e^{\beta t}$ has bounded variation. Consequently, applying Itô's formula (for example, see Theorem 33 in chapter 2 of [83]) to the function F and the semi-martingales $t \mapsto \varphi(X_t)$ and $t \mapsto e^{\beta t}$, we get:

$$\begin{aligned} G(t) =& \varphi(X_0) + \int_0^t e^{\beta u}\mathcal{L}\varphi(X_u)\mathrm{d}u + \int_0^t e^{\beta u}\mathrm{d}M_u + \int_0^t \varphi(X_u)\beta e^{\beta u}\mathrm{d}u \\ &+ \sum_{0\leq s\leq t}\left[\Delta G(s) - \frac{\partial F}{\partial v}\left(e^{\beta s-}, \varphi(X_{s-})\right)\cdot \Delta\varphi(X_s)\right]. \end{aligned} \tag{4.33}$$

On the other hand, as $\dfrac{\partial F}{\partial v}(u,v) = u$ and the function $t \mapsto e^{\beta t}$ is continuous,

$$\begin{aligned} \Delta G(s) &= e^{\beta s}\varphi(X_s) - e^{\beta s}\varphi(X_{s-}) \\ &= e^{\beta s}\Delta\varphi(X_s) \\ &= \frac{\partial F}{\partial v}\left(e^{\beta s-}, \varphi(X_{s-})\right)\cdot\Delta\varphi(X_s). \end{aligned}$$

Hence the last term in the right-hand side of (4.33) is zero. Moreover, the term $\int_0^t e^{\beta u}\mathrm{d}M_u$ is a martingale. Applying the formula at time $t \wedge T$, we get:

$$\begin{aligned} &e^{\beta(t\wedge T)}\varphi(X_{t\wedge T}) \\ &= \varphi(X_0) + \int_0^{t\wedge T} e^{\beta u}\Big(\mathcal{L}\varphi(X_u) + \beta\varphi(X_u)\Big)\mathrm{d}u + \int_0^{t\wedge T} e^{\beta u}\mathrm{d}M_u. \end{aligned} \tag{4.34}$$

By Lemma 4.17, $\mathcal{L}\varphi(X_u) + \beta\varphi(X_u) \leq 0$ when $u < T$. It follows that $\left(e^{\beta(t\wedge T)}\varphi(X_{t\wedge T})\right)_{t\geq 0}$ is a supermartingale. Thus,

$$\mathbb{E}_x\left[e^{\beta(t\wedge T)}\varphi(X_{t\wedge T})\right] \leq \varphi(x). \tag{4.35}$$

Clearly, if $x \in \arg\min\varphi$, then $\mathcal{L}\varphi(x) \geq 0$, and hence by (4.28), $\varphi(x) \leq L$. In particular, $\{\varphi \leq L\} \neq \varnothing$, so $T < \infty$ a.s. as the process is irreducible. In (4.35), letting $t \to \infty$ and using Fatou's lemma, we obtain

$$\mathbb{E}_x\left[e^{\beta T}\varphi(X_T)\right] \leq \varphi(x).$$

It is easy to see that $\varphi(X_T) \geq \varphi(X_{T-})/e^{\theta} \geq Le^{-\theta}$, which leads to our claim. □

The next lemma says that if we start from a configuration x such that $\|x\|_\infty = \mathcal{O}(\log n)$, then this remains true for a long time.

Lemma 4.20 (Stability of trajectories)**.** *Let α_1 be defined as in Proposition 4.12. There is a* *dimension-free constant α_2 such that*

$$\sup_{\|x\|_\infty \leq \alpha_1 \log n} \mathbb{P}_x \left[\exists t \in [0, (\log n)^2], \|X_t\|_\infty > \alpha_2 \log n\right] = \mathcal{O}\left(n^{-3}\right). \tag{4.36}$$

Proof. Suppose that $\|x\|_\infty \leq \alpha_1 \log n$. Let $\theta_1 = \frac{6}{\alpha_1}$, and let β_1 be a constant, and let $t = \frac{\theta_1}{\beta_1}\alpha_1 \log n \geq \frac{\theta_1}{\beta_1}\|x\|_\infty$. Then by (4.31),

$$\mathbb{P}_x \left[\|X_t\|_\infty > \alpha_1 \log n\right] = \mathcal{O}\left(n^{-5}\right).$$

Moreover, by Lemma 4.8, for a dimension-free constant α_2' large enough,

$$\mathbb{P}\left[G_i(t) \geq \alpha_2' \log n\right] = \mathcal{O}\left(n^{-5}\right).$$

Taking a union bound, we deduce that

$$\mathbb{P}_x \left[\{\|X_t\|_\infty > \alpha_1 \log n\} \cup \{\exists i : G_i(t) \geq \alpha_2' \log n\}\right] = \mathcal{O}\left(n^{-4}\right).$$

This implies

$$\mathbb{P}_x \left[\|X_t\|_\infty \leq \alpha_1 \log n, \sup_{s\in[0,t]} \|X_s\|_\infty \leq (\alpha_1 + \alpha_2') \log n\right] \geq 1 - \mathcal{O}\left(n^{-4}\right).$$

The inequality remains true when we take the supremum over all x such that $\|x\|_\infty \leq \alpha_1 \log n$. Iterating, and using the Markov property, we deduce that, for any $k \in \mathbb{Z}_+$,

$$\begin{aligned}\mathbb{P}_x \left[\|X_{kt}\|_\infty \leq \alpha_1 \log n, \sup_{s\in[0,kt]} \|X_s\|_\infty \leq (\alpha_1 + \alpha_2') \log n\right] &\geq (1 - \mathcal{O}\left(n^{-4}\right))^k \\ &\geq 1 - \mathcal{O}\left(kn^{-4}\right).\end{aligned}$$

We finish the proof simply by taking $k = \lfloor(\log n)^2\rfloor$, and $\alpha_2 = \alpha_1 + \alpha_2'$. □

In the next proposition, we prove that the bound on $\|X\|_\infty$ above leads to a strong bound on $\varphi^\theta(X)$ for some $\theta > 0$.

Proposition 4.21 (Strong concentration of trajectories)**.** *Let θ_2, β_2 be two positive dimension-free constants and $L_2 = L(\theta_2, \beta_2)$ as in Lemma 4.17. Let T be the hitting time of the set $\{\varphi^{\theta_2} \leq L_2\}$. Let α_1 be defined as in Proposition 4.12. Then, provided that θ_2 is small enough, for any x such that $\|x\|_\infty \leq \alpha_1 \log n$,*

$$\mathbb{P}_x \left[\sup_{s\in[T,(\log n)^2]} \varphi^{\theta_2}(X_s) > L_2 + 4\right] = \mathcal{O}\left(n^{-3}\right). \tag{4.37}$$

We will need the following lemma:

Lemma 4.22 (Martingale estimate)**.** *Let α_1 be defined as in Proposition 4.12. Let θ_2 be a positive dimension-free constant. Suppose that $\|x\|_\infty \leq \alpha_1 \log n$. Let $(M_t)_{t\geq 0}$ be defined by*

$$M_t = \varphi^{\theta_2}(X_t) - \varphi^{\theta_2}(X_0) - \int_0^t \mathcal{L}\varphi^{\theta_2}(X_u)\mathrm{d}u,$$

which is a martingale according to (4.17). Then, for θ_2 small enough,

$$\mathbb{P}_x\left[\sup_{s\in[0,(\log n)^2]} |M_s| \geq 1\right] = \mathcal{O}\left(n^{-3}\right).$$

For simplicity, in the proofs of Lemma 4.22 and Proposition 4.21, we still write φ instead of φ^{θ_2} and φ_i instead of $\varphi_i^{\theta_2}$. First we see how Lemma 4.22 leads to Proposition 4.21:

Proof of Proposition 4.21. Let θ_2 and M be as in Lemma 4.22. We will prove that

$$\left\{\sup_{s\in[T,(\log n)^2]} \varphi(X_s) > L_2 + 4\right\} \subset \left\{\sup_{s\in[0,(\log n)^2]} |M_s| \geq 1\right\} \bigcup \{T \geq (\log n)^2\},$$

and then we show that the probabilities of the events on the right-hand side are $\mathcal{O}\left(n^{-3}\right)$. By contrapositivity, suppose that for a realization of X which is càdlàg almost surely, we have $\sup_{s\in[0,(\log n)^2]} |M_s| < 1$ and $T < (\log n)^2$. We prove that

$$\sup_{s\in[T,(\log n)^2]} \varphi(X_s) \leq L_2 + 4.$$

For $h \in [T, (\log n)^2]$ arbitrary, let $s_h = \sup\{s \in [0, h] : \varphi(X_{s-}) \leq L_2\}$. Note that for any $s \in [0, (\log n)^2]$,

$$|\Delta\varphi(X_s)| = |\Delta M_s| \leq 2 \sup_{s\in[0,(\log n)^2]} |M_s| \leq 2.$$

Moreover, by definition of s_h, $\varphi(X_{s_h-}) \leq L_2$, and hence $\varphi(X_{s_h}) \leq L_2 + \Delta\varphi(X_{s_h}) \leq L_2 + 2$. Also by definition of s_h, $\varphi(X_u) > L_2$ when $s_h \leq u < h$, and hence $\mathcal{L}\varphi(X_u) < 0$ by Lemma 4.17. Consequently,

$$\varphi(X_h) = \varphi(X_{s_h}) - M_{s_h} + M_h + \int_{s_h}^h \mathcal{L}\varphi(X_u)\mathrm{d}u \leq L_2 + 4, \tag{4.38}$$

which proves the inclusion. Besides, Proposition 4.19 gives us $\mathbb{P}_x\left[T \geq (\log n)^2\right] = \mathcal{O}\left(n^{-3}\right)$ by Chernoff's bound. Combining this and Lemma 4.22, we deduce the claim. □

Now we prove Lemma 4.22:

Proof of Lemma 4.22. We will provide good control on ΔM and $\langle M, M\rangle$, and afterward we use Lemma 4.7. By (4.17), the process $(M_t^{(i)})_{t\geq 0}$ defined by

$$M_t^{(i)} = \varphi_i(X_t) - \varphi_i(X_0) - \int_0^t \mathcal{L}\varphi_i(X_u)\mathrm{d}u$$

is a zero-mean martingale. It is clear that M is the average of $M^{(i)}$:

$$M_t = \frac{1}{n}\sum_{i=1}^{n} M_t^{(i)}.$$

Due to the conservation of the number of particles, the martingales $(M^{(i)})_{i\in[n]}$ have negative covariations. More precisely, according to (4.18), we have

$$d\left\langle M^{(i)}, M^{(j)}\right\rangle_t = \sum_{1\le k,l\le n} \frac{r(X_t(k))}{n}\Big(\varphi_i(X_t-\delta_k+\delta_l)-\varphi_i(X_t)\Big)\times \\ \times\Big(\varphi_j(X_t-\delta_k+\delta_l)-\varphi_j(X_t)\Big).$$

Note that when $i\neq j$, $(\varphi_i(x-\delta_k+\delta_l)-\varphi_i(x))\,(\varphi_j(x-\delta_k+\delta_l)-\varphi_j(x))$ is negative if $\{k,l\} = \{i,j\}$ and is zero otherwise. Hence $\left\langle M^{(i)}, M^{(j)}\right\rangle_t \le 0$, $\forall i\neq j$. Consequently, for all t positive,

$$\langle M, M\rangle_t \le \frac{1}{n^2}\sum_{i=1}^{n}\left\langle M^{(i)}, M^{(i)}\right\rangle_t.$$

Let α_2 be defined as in Lemma 4.20, and let U be the exit time from $\{\|\cdot\|_\infty \le \alpha_2\log n\}$. Note that, almost surely, for any $u\ge 0$, $\Delta M_u = 0$ or there exist $k,l\in[n]$ such that

$$\Delta M_u = \varphi(X_u-\delta_k+\delta_l)-\varphi(X_u) = \frac{1}{n}(e^{\theta_2}-1)(e^{\theta_2 X_{u-}(l)}-e^{\theta_2(X_{u-}(k)-1)}).$$

In either case, before time U, almost surely,

$$|\Delta M_u| \le \frac{1}{n}(e^{\theta_2}-1)e^{\theta_2\|X_{u-}\|_\infty} \le (e^{\theta_2}-1)n^{\theta_2\alpha_2-1}.$$

Similarly,

$$|\varphi_i(x-\delta_k+\delta_l)-\varphi_i(x)| \le (e^{\theta_2}-1)e^{\theta_2\|x\|_\infty}(\mathbb{1}_{\{k=i\}}\vee\mathbb{1}_{\{l=i\}}).$$

Hence for any $u<U$,

$$|\varphi_i(x-\delta_k+\delta_l)-\varphi_i(x)|^2 \le (e^{\theta_2}-1)^2 n^{2\theta_2\alpha_2}(\mathbb{1}_{\{k=i\}}\vee\mathbb{1}_{\{l=i\}}).$$

For any $i\in[n]$, let $M^{(i),U}$ be the martingale $M^{(i)}$ stopped at time U, i.e. ,

$$\forall t\ge 0,\ M_t^{(i),U} := M_{t\wedge U}^{(i)}.$$

Similarly, let M^U be the martingale M stopped at time U. We have

$$\left\langle M^{(i),U}, M^{(i),U}\right\rangle_t = \int_0^{t\wedge U}\sum_{1\le k,l\le n}\frac{r(X_u(k))}{n}\left(\varphi_i(X_u-\delta_k+\delta_l)-\varphi_i(X_u)\right)^2 \mathrm{d}u.$$

Then, by dividing the double sum to the sum where $k = i$ or $l = i$ or $k \neq i \neq l$, we get

$$\left\langle M^{(i),U}, M^{(i),U} \right\rangle_t \leq \int_0^{t \wedge U} (r(X_u(i)) + \zeta_u (e^{\theta_2} - 1)^2 n^{2\theta_2 \alpha_2} \mathrm{d}u,$$

where we recall that ζ_u is the mean-field jump rate. Taking the sum over $i \in [n]$, we get

$$\begin{aligned} \left\langle M^U, M^U \right\rangle_t &\leq \frac{1}{n^2} \int_0^{t \wedge U} 2n\zeta_u n^{2\theta_2\alpha_2} (e^{\theta_2} - 1)^2 \mathrm{d}u \\ &\leq \frac{1}{n^2} \int_0^{t \wedge U} 2n\kappa n^{2\theta_2\alpha_2} (e^{\theta_2} - 1)^2 \mathrm{d}u \\ &= \mathcal{O}\left(tn^{2\theta_2\alpha_2 - 1}\right), \end{aligned}$$

where we have used (4.16) in the second inequality. Now for θ_2 small enough, $2\theta_2\alpha_2 - 1 < -1/2$, which implies $\Delta M_u = \mathcal{O}(1/\sqrt{n})$ and $\left\langle M^U, M^U \right\rangle_t = \mathcal{O}(1/\sqrt{n})$ if $u < U$ and $t = (\log n)^2$. Then we apply Lemma 4.7 to the martingale $(M_t^{U \wedge (\log n)^2})_{t \geq 0} := (M_{U \wedge (\log n)^2 \wedge t})_{t \geq 0}$, with $a = 1$, $K = b^2 = \mathcal{O}(1/\sqrt{n})$, to obtain

$$\mathbb{P}_x \left[\sup_{s \geq 0} M_s^{U \wedge (\log n)^2} \geq 1 \right] \leq e^{-\Omega(\sqrt{n})}.$$

Using the same argument for $-M$, and taking a union bound, we deduce that

$$\mathbb{P}_x \left[\sup_{s \geq 0} \left| M_s^{U \wedge (\log n)^2} \right| \geq 1 \right] \leq 2e^{-\Omega(\sqrt{n})}.$$

Taking a union bound with the event in Lemma 4.20, we get

$$\begin{aligned} \mathbb{P}_x \left[\exists s \in [0, (\log n)^2] : |M_s| \geq 1 \right] &\leq \mathbb{P}_x \left[\sup_{s \geq 0} \left| M_s^{U \wedge (\log n)^2} \right| \geq 1 \right] + \mathbb{P}_x \left[U \leq (\log n)^2 \right] \\ &= \mathcal{O}\left(n^{-3}\right), \end{aligned}$$

which finishes our proof. □

3.2 Path coupling via tagged particles

For $k \in \mathbb{Z}_+$, define

$$\Delta_r(k) := r(k+1) - r(k) \geq 0. \tag{4.39}$$

Let Θ be a Poisson point process of intensity $\frac{1}{n} \mathrm{d}t \otimes \mathrm{d}u \otimes \mathrm{Card}$ on $\mathbb{R}_+ \times \mathbb{R}_+ \times [n]$, independent of the Poisson processes used in the graphical construction of X. For a site $i \in [n]$, define an $[n]$- valued process $I = (I_t)_{t \geq 0}$ by setting $I_0 = i$ and for each (t, u, k) in Θ,

$$I_t := \begin{cases} k & \text{if } u \leq \Delta_r(X_{t-}(I)) \\ I_{t-} & \text{otherwise,} \end{cases} \tag{4.40}$$

where $X_t(I) := X_t(I_t)$. This definition means that conditionally on X, the process I will jump with the time-varying rate $\Delta(X_t(I))$, and the destination is uniformly chosen among all sites. A simple but important observation is that $(X_t + \delta_{I_t})_{t\geq 0}$ has the same distribution as a Zero-Range process starting at $x + \delta_i$. We call I a tagged particle, and we stress here that the construction of I relies strictly on condition (4.6). For another site j, similarly, we can construct a second tagged particle J starting from $J_0 = j$ using the same process Θ. Thus we have a coupling $(X_t + \delta_{I_t}, X_t + \delta_{J_t})_{t\geq 0}$ of two Zero-Range processes starting at $x+\delta_i$ and $x+\delta_j$ respectively. We call the particles of X *non-tagged particles*. We denote by $\mathbb{P}_{x,i,j}$ the law of the process (X, I, J) starting from (x, i, j) and $\mathbb{E}_{x,i,j}$ the expectation taken w.r.t $\mathbb{P}_{x,i,j}$. Let τ be the coalescence time of I and J:

$$\tau := \inf\{t \geq 0 : I_t = J_t\}. \tag{4.41}$$

By the classical relation between $\mathrm{d}_{\mathrm{TV}}(\cdot,\cdot)$ and coupling, we have:

$$\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_{x+\delta_i}\left[X_t \in \cdot\right], \mathbb{P}_{x+\delta_j}\left[X_t \in \cdot\right]\right) \leq \mathbb{P}_{x,i,j}\left[\tau > t\right]. \tag{4.42}$$

By construction, if the two tagged particles manage to jump at the same time, then coalescence occurs. However, the jump rates of the tagged particles depend on their number of cooccupants, which complicates our task. We will need to analyze carefully the trajectories of (X, I, J) to obtain a good estimate on τ.

The inequality (4.42) gives us the estimate on the total variation distance of two processes starting from two adjacent configurations. We can extend the comparison to two arbitrary configurations by choosing a path between them and then using triangle inequality. This simple but powerful idea, originally due to Bubley and Dyer in [11], is the strategy that we will implement.

Throughout this subsection, α_1 will be as in Proposition 4.12, and θ_2, β_2, L_2 will be as in Proposition 4.21. Proposition 4.21 and (4.32) imply that starting from any configuration x such that $\|x\|_\infty \leq \alpha_1 \log n$, the system will reach the set $\{\varphi^{\theta_2} \leq L_2\}$ quickly (in time $\mathcal{O}(\log n)$) and then remains in $\{\varphi^{\theta_2} \leq L_2 + 4\}$ for a long time, namely $\Omega((\log n)^2)$. We will prove that if this is the case, then the coalescence time τ is likely to be $\mathcal{O}(\log n)$.

From now on, let $\mathsf{Good} := \{\varphi^{\theta_2} \leq L_2 + 4\}$. We say that a configuration x is *good* if $x \in \mathsf{Good}$ and x is *bad* otherwise. We introduce the process (X^*, I^*) taking values in $\mathsf{Good} \times [n]$ whose infinitesimal generator $\mathcal{L}^*$ acts on an observable $\varphi : \mathsf{Good} \times [n] \to \mathbb{R}$ by:

$$\begin{aligned} \mathcal{L}^*\varphi(x,i) = &\sum_{k,l} \frac{r(x(k))}{n} \left(\varphi(x - \delta_k + \delta_l, i) - \varphi(x,i)\right) \mathbb{1}_{\{x-\delta_k+\delta_l \in \mathsf{Good}\}} + \\ &+ \sum_{j=1}^{n} \frac{\Delta_r(x(i))}{n} (\varphi(x,j) - \varphi(x,i)). \end{aligned} \tag{4.43}$$

This definition means that X^* is a Zero-Range process constrained to staying in Good, and I^* jumps with a time-varying rate $\Delta_r(X_t^*(I^*))$ and chooses its destination uniformly among all sites, where $X_t^*(I^*) := X_t^*(I_t^*), \forall t \geq 0$. We can use the same Poisson processes in the graphical construction of (X, I) and (X^*, I^*) to obtain a coupling of them, and we can construct the second tagged particle J^* analogously. By their construction, if the processes (X, I, J) and (X^*, I^*, J^*)

start from the same configuration (x, i, j), then they coincide up to the time when X turns bad. For any $\theta > 0$, we define the observable $\varphi_*^\theta : \mathsf{Good} \times [n] \to \mathbb{R}_+$ by

$$\varphi_*^\theta(x, i) := \varphi_i^\theta(x) = e^{\theta x(i)}.$$

Lemma 4.23 (Exponential moment of the number of cooccupants)**.** *There exist dimension-free constants c_2 and K such that for any $(x, i, j) \in \mathsf{Good} \times [n] \times [n]$, when $t \geq c_2(x(i) \vee x(j))$,*

$$\mathbb{E}_{x,i,j}\left[e^{\theta_2(X_t^*(I^*) \vee (X_t^*(J^*))}\right] \leq K. \tag{4.44}$$

Proof. We will prove that there exist positive dimension-free constants a_2, b_2 such that

$$\mathcal{L}^* \varphi_*^{\theta_2}(x, i) \leq -a_2 \varphi_*^{\theta_2}(x, i) + b_2, \tag{4.45}$$

for any $(x, i) \in \mathsf{Good} \times [n]$. Let us see how (4.45) leads to the claim:

Let $u(t) = \mathbb{E}_{x,i}\left[\varphi_*^{\theta_2}(X_t^*, I_t^*)\right]$. Then

$$u'(t) = \mathbb{E}_{x,i}\left[\mathcal{L}^* \varphi_*^{\theta_2}(X_t^*, I_t^*)\right] \leq -a_2 u(t) + b_2.$$

Hence, by Lemma 4.16,

$$\mathbb{E}_{x,i}\left[e^{\theta_2 X_t^*(I^*)}\right] \leq \frac{b_2}{a_2} + 1,$$

for any $t \geq \dfrac{\theta_2}{a_2} x(i)$. We take $c_2 = \dfrac{\theta_2}{a_2}$, $K = 2\left(\dfrac{b_2}{a_2} + 1\right)$, so the claim follows from the inequality

$$e^{\theta_2(X_t^*(I^*) \vee X_t^*(J^*))} \leq e^{\theta_2 X_t^*(J^*)} + e^{\theta_2 X_t^*(I^*)}.$$

It remains to prove (4.45). It is similar to Lemma 4.17 except now we have an extra term corresponding to the jump of the tagged particle I, which is controlled by the fact that the system is constrained to staying in Good. More precisely,

$$\begin{aligned}
\mathcal{L}^* \varphi_*^{\theta_2}(x, i) = &\sum_{k \neq i} \frac{r(x(k))}{n}\left(e^{\theta_2} - 1\right)\varphi_*^{\theta_2}(x, i)\mathbb{1}_{\{x - \delta_k + \delta_i \in \mathsf{Good}\}} \\
&+ \sum_{k \neq i} \frac{r(x(i))}{n}\left(e^{-\theta_2} - 1\right)\varphi_*^{\theta_2}(x, i)\mathbb{1}_{\{x - \delta_i + \delta_k \in \mathsf{Good}\}} \\
&+ \sum_{j=1}^{n} \frac{\Delta_r(x(i))}{n}(\varphi_*^{\theta_2}(x, j) - \varphi_*^{\theta_2}(x, i)).
\end{aligned}$$

We will bound the three terms above to obtain an upper bound on $\mathcal{L}^* \varphi_*^{\theta_2}$:

The first term: since $\frac{1}{n}\sum_{k=1}^{n} r(x(k)) \leq \kappa$ by (4.16), hence,

$$\sum_{k \neq i} \frac{r(x(k))}{n}\left(e^{\theta_2} - 1\right)\varphi_*^{\theta_2}(x, i)\mathbb{1}_{\{x - \delta_k + \delta_i \in \mathsf{Good}\}} \leq \kappa(e^{\theta_2} - 1)\varphi_*^{\theta_2}(x, i).$$

The third term: it is negative when $x(i) > \frac{1}{\theta_2}\log(L_2+4)$ since $x \in \mathsf{Good}$, hence,

$$\Delta_r(x(i))\left(\frac{\sum_{j=1}^n e^{\theta_2 x(j)}}{n} - e^{\theta_2 x(i)}\right) < \max_{k \le \frac{1}{\theta_2}\log(L_2+4)} \Delta_r(k)(L_2+4) =: c,$$

for some dimension-free constant c.

The second term: we first observe that there are at most $\frac{n}{N}$ sites l such that $x(l) > N\rho$, for any constant $N > 0$, thanks to (4.9). On the other hand, if $x(i) > N\rho$ and $x(l) \le N\rho$, then $\varphi^{\theta_2}(x - \delta_i + \delta_l) \le \varphi^{\theta_2}(x)$, so if x is good, then so is $x - \delta_i + \delta_l$. Consequently, as $e^{-\theta_2} - 1$ is negative, we have

$$\begin{aligned}
&\sum_{k \ne i} \frac{r(x(i))}{n}(e^{-\theta_2}-1)\varphi_*^{\theta_2}(x,i)\mathbb{1}_{\{x-\delta_i+\delta_k \in \mathsf{Good}\}} \\
\le &\sum_{k \ne i} \frac{r(x(i))}{n}(e^{-\theta_2}-1)\varphi_*^{\theta_2}(x,i)\mathbb{1}_{\{x-\delta_i+\delta_k \in \mathsf{Good}\}}\mathbb{1}_{\{x(i)>N\rho\}} \\
\le &\sum_{k \ne i} \frac{r(\lceil N\rho \rceil)}{n}(e^{-\theta_2}-1)\varphi_*^{\theta_2}(x,i)\mathbb{1}_{\{x(i)>N\rho\}}\mathbb{1}_{\{x(k)\le N\rho\}} \\
\le & r(\lceil N\rho \rceil)(e^{-\theta_2}-1)\varphi_*^{\theta_2}(x,i)\mathbb{1}_{\{x(i)>N\rho\}}(1-1/N) \\
= & r(\lceil N\rho \rceil)(e^{-\theta_2}-1)\varphi_*^{\theta_2}(x,i)(1-1/N) + \\
&\quad + r(\lceil N\rho \rceil)(1-e^{-\theta_2})\varphi_*^{\theta_2}(x,i)(1-1/N)\mathbb{1}_{\{x(i)\le N\rho\}}.
\end{aligned}$$

We sum the three inequalities, and afterwards we take

$$a_2 = -\kappa(e^{\theta_2}-1) + (1-1/N)r(\lceil N\rho \rceil)(1-e^{-\theta_2})$$

and

$$b_2 = c + r(\lceil N\rho \rceil)e^{\theta_2 N\rho}.$$

We choose N large enough to make $a_2 > 0$, which is what we needed. □

We fix a constant c_2 which satisfies Lemma 4.23. For any initial configuration $(x,i,j) \in \mathsf{Good} \times [n] \times [n]$, we define successively the stopping times $(T_k)_{k\ge 1}$ as follows:

$$\begin{aligned}
T_1 &= c_2(x(i) \vee x(j)) + 1 = c_2(X_0^*(I^*) \vee X_0^*(J^*)) + 1, \\
T_k &= T_{k-1} + c_2\left(X_{T_{k-1}}^*(I^*) \vee X_{T_{k-1}}^*(J^*)\right) + 1.
\end{aligned} \tag{4.46}$$

Lemma 4.24 (Bound of τ by T_k)**.** *Let $(T_k)_{k\ge1}$ be defined as in (4.46). Then there exists a dimension-free constant c_3 such that for any $(x,i,j) \in \mathsf{Good} \times [n] \times [n]$, for any $k \ge 1$,*

$$\mathbb{P}_{x,i,j}\left[\tau \ge T_k; X\big|_{[0,T_k]} = X^*\big|_{[0,T_k]}\right] \le (1-c_3)^k. \tag{4.47}$$

Proof. We only need to prove for $k = 1$, then use induction and the strong Markov property.

Let c_2 be the constant used in the definition of $(T_k)_{k\geq 1}$, and let K be the corresponding constant in Lemma 4.23, and let $t = c_2(x(i) \vee x(j)) = T_1 - 1$. By (4.44) and Chernoff's bound,

$$\mathbb{P}_{x,i,j}\left[X_t^*(I^*) \vee X_t^*(J^*) \geq a\right] \leq \frac{K}{e^{\theta_2 a}},$$

for any $a > 0$. We choose a large enough to make the right-hand side less than 1/2. We will prove that for any $(x, i, j) \in \mathsf{Good} \times [n] \times [n]$ such that $x(i) \vee x(j) < a$, there exists a dimension-free constant $c > 0$ such that

$$\mathbb{P}_{x,i,j}\left[\tau < 1\right] > c. \tag{4.48}$$

Assuming for the moment that we have (4.48), let us prove the lemma. It is not hard to see that

$$\begin{aligned}
&\mathbb{P}_{x,i,j}\left[\tau > T_1, X\big|_{[0,T_1]} = X^*\big|_{[0,T_1]}\right] \\
&\leq \mathbb{P}_{x,i,j}\left[X_t^*(I^*) \vee X_t^*(J^*) \geq a\right] + \mathbb{P}_{x,i,j}\left[X_t^*(I^*) \vee X_t^*(J^*) < a, X\big|_{[0,t]} = X^*\big|_{[0,t]}, \tau > t+1\right] \\
&\leq \mathbb{P}_{x,i,j}\left[X_t^*(I^*) \vee X_t^*(J^*) \geq a\right] + (1-c)\mathbb{P}_{x,i,j}\left[X_t^*(I^*) \vee X_t^*(J^*) < a\right] \\
&\leq 1 - c\mathbb{P}_{x,i,j}\left[X_t^*(I^*) \vee X_t^*(J^*) < a\right] \\
&\leq 1 - c/2.
\end{aligned}$$

In the second inequality, we have used (4.48) and the Markov property at time t. We deduce the claim simply by taking $c_3 = c/2$. It remains to prove (4.48).

Suppose that x is good and $x(i) \vee x(j) < a$. The scenario is that in a finite time, there is no particle arriving at i and j, and the tagged particles wait for two sites i, j to be completely emptied, and afterwards they jump at the same time. More precisely, we use the mixed graphical construction for the process X as follows: let Ξ and Ψ be two independent Poisson processes defined as in Graphical Construction 1 and Graphical Construction 2. Consider the process X which starts from x and has the following jumps: for each $(t, u, e) \in \Psi$ where $e \in \{i, j\}$,

$$X_t := \begin{cases} X_{t-} - \delta_l + \delta_e, & \text{if } \frac{1}{n}\sum_{k=1}^{l-1} r\left(X_{t-}(k)\right) < u \leq \frac{1}{n}\sum_{k=1}^{l} r(X_{t-}(k)), \\ & \quad \text{for some } l \in [n], \\ X_{t-} & \text{otherwise,} \end{cases} \tag{4.49}$$

and for each $(t, u, k, l) \in \Xi$ where $l \in [n] \setminus \{i, j\}$,

$$X_t := \begin{cases} X_{t-} - \delta_k + \delta_l, & \text{if } r(X_{t-}(k)) \geq u, \\ X_{t-} & \text{otherwise.} \end{cases} \tag{4.50}$$

Then X is a Markov process with generator $\mathcal{L}$ on $\mathcal{X}$. Here, we use Ψ to indicate the jumps to two sites i, j and Ξ to indicate other jumps. Let

- $A = \left\{\Psi\Big([0,1] \times [0,\kappa] \times \{i,j\}\Big) = 0\right\}$,

- $B_i = \left\{ \Xi\Big([0,1/2] \times [0, r(1)] \times \{i\} \times ([n] \setminus \{i,j\}) \Big) \geq a \right\}$,
- $B_j = \left\{ \Xi\Big([0,1/2] \times [0, r(1)] \times \{j\} \times ([n] \setminus \{i,j\}) \Big) \geq a \right\}$,
- $C = \left\{ \Theta\Big([0,1/2] \times [0, \max_{k \leq a} \Delta_r(k)] \times [n] \Big) = 0 \right\} \cap \left\{ \Theta\Big([1/2,1] \times [0, r(1)] \times [n] \Big) \geq 1 \right\}$.

In fact, A is the event that there is no non-tagged particle arriving at two sites i,j up to time 1. If A happens, then B_i and B_j ensure that all non-tagged particles of two sites i,j jump to $[n] \setminus \{i,j\}$ in $[0,1/2]$. If A, B_i, B_j happen, then two sites i,j are empty in $[1/2,1]$, and event C ensures that the two tagged particles stay at $\{i,j\}$ in $[0,1/2]$ then jump at the same time (hence coalescence) in $[1/2,1]$. Moreover, the Poisson random variables used in the definitions of these events are independent and have parameters $\Theta(1)$. We conclude that the events above are independent and their probabilities are $\Theta(1)$. It follows that

$$\mathbb{P}_{x,i,j}[\tau < 1] \geq \mathbb{P}_{x,i,j}[A \cap B_i \cap B_j \cap C] = \Theta(1),$$

which finishes our proof. □

Lemma 4.25 (Exponential moment of T_k)**.** *Let c_2 and $(T_k)_{k \geq 1}$ be as in (4.46), and let K be the corresponding constant in Lemma 4.23. Let $\theta_3 = \theta_2 / c_2$. Then for any $(x,i,j) \in \mathsf{Good} \times [n] \times [n]$, for any $k \geq 1$,*

$$\mathbb{E}_{x,i,j}\left[e^{\theta_3 T_k}\right] \leq n(Ke^{\theta_3})^k (L_2 + 4).$$

Proof. By convention, let $T_0 = 0$. For any $k \geq 2$, note that T_{k-1} is $\mathcal{F}_{T_{k-2}}$-measurable by its definition. Conditionally on $\mathcal{F}_{T_{k-2}}$, we have

$$\begin{aligned}
\mathbb{E}_{x,i,j}\left[e^{\theta_3 T_k}\right] &= \mathbb{E}_{x,i,j}\left[\mathbb{E}_{x,i,j}\left[e^{\theta_3 T_k} | \mathcal{F}_{T_{k-2}}\right]\right] \\
&= \mathbb{E}_{x,i,j}\left[e^{\theta_3 (T_{k-1}+1)} \mathbb{E}_{x,i,j}\left[e^{\theta_3 c_2 (X^*_{T_{k-1}}(I^*) \vee X^*_{T_{k-1}}(J^*)} | \mathcal{F}_{T_{k-2}}\right]\right] \\
&= \mathbb{E}_{x,i,j}\left[e^{\theta_3 (T_{k-1}+1)} \mathbb{E}_{X^*_{T_{k-2}}, I^*_{T_{k-2}}, J^*_{T_{k-2}}}\left[e^{\theta_2 (X^*_I(T_1) \vee X^*_J(T_1))}\right]\right] \\
&\leq K e^{\theta_3} \mathbb{E}\left[e^{\theta_3 T_{k-1}}\right],
\end{aligned}$$

where the inequality is due to (4.44). Moreover,

$$\mathbb{E}_{x,i,j}\left[e^{\theta_3 T_1}\right] = \mathbb{E}_{x,i,j}\left[e^{\theta_2 (x(i) \vee x(j)) + \theta_3}\right] \leq n(L_2 + 4) e^{\theta_3},$$

where the last inequality is due to the fact that $x \in \mathsf{Good}$. The claim is then obtained by induction. □

Corollary 4.26 (Quick coalescence while staying good)**.** *There exists a dimension-free constant α_3 such that for any $(x,i,j) \in \mathsf{Good} \times [n] \times [n]$,*

$$\mathbb{P}_{x,i,j}[\tau > \alpha_3 \log n; X_t = X^*_t, \forall t \leq \alpha_3 \log n] = \mathcal{O}\left(n^{-3}\right). \tag{4.51}$$

Proof. For any $\alpha_3 > 0$ and $k \in \mathbb{Z}_+$, the left-hand side is upper bounded by

$$\begin{aligned}&\mathbb{P}_{x,i,j}\left[T_k > \alpha_3 \log n\right] + \mathbb{P}_{x,i,j}\left[\tau \geq T_k; X|_{[0,T_k]} = X^*|_{[0,T_k]}\right] \\ &\leq n(Ke^{\theta_3})^k (L_2+4) n^{-\theta_3\alpha_3} + (1-c_3)^k,\end{aligned}$$

for some dimension-free constants K, θ_3, c_3, due to Lemma 4.24, Lemma 4.25, and Chernoff's bound. We choose $k = \mathcal{O}(\log n)$ such that $(1-c_3)^k = \mathcal{O}(n^{-3})$ and α_3 large enough such that $(Ke^{\theta_3})^k n^{1-\theta_3\alpha_3} = \mathcal{O}(n^{-3})$ to get what we wanted. □

Now we can finally prove the quick coalescence for a configuration x such that $\|x\|_\infty \leq \alpha_1 \log n$.

Proposition 4.27 (Quick coalescence)**.** *Recall that α_1 is fixed in this subsection. There exists a dimension-free constant α such that for any x such that $\|x\|_\infty \leq \alpha_1 \log n$, for any $i, j \in [n]$,*

$$\mathbb{P}_{x,i,j}\left[\tau \geq \alpha \log n\right] = \mathcal{O}\left(n^{-3}\right). \tag{4.52}$$

Proof. Let $\alpha = \alpha_4 + \alpha_3$, where α_3 is as in Corollary 4.26, and α_4 is a dimension-free constant that we will choose later. Let T be the hitting time of the set $\{\varphi^{\theta_2} \leq L_2\}$. The probability that we want to estimate does not exceed the following sum:

$$\begin{aligned}&\mathbb{P}_x\left[T \geq \alpha_4 \log n\right] \\ &+ \mathbb{P}_x\left[\sup_{s\in[T,(\log n)^2]} \varphi^{\theta_2}(X_s) > L_2 + 4\right] \\ &+ \mathbb{P}_{x,i,j}\left[T < \alpha_4 \log n; \sup_{s\in[T;(\log n)^2]} \varphi^{\theta_2}(X_s) \leq L_2 + 4; \tau \geq T + \alpha_3 \log n\right].\end{aligned} \tag{4.53}$$

We simply prove that all the terms are $\mathcal{O}(n^{-3})$:

1. **The first term:** By (4.32) and Chernoff's bound, it is upper bounded by $\mathcal{O}\left(n^{\theta_2\alpha_1 - \beta_2\alpha_4}\right)$, which is $\mathcal{O}(n^{-3})$ when α_4 is large enough.

2. **The second term** is $\mathcal{O}(n^{-3})$ by Proposition 4.21.

3. **The last term:** From the time T onward, we couple the processes (X, I, J) and (X^*, I^*, J^*) starting from (X_T, I_T, J_T) by using the same Poisson processes for their graphical constructions. We observe that up to time $(\log n)^2$, (X, I, J) and (X^*, I^*, J^*) coincide. Therefore, this term is $\mathcal{O}(n^{-3})$ by the strong Markov property at time T and (4.51), which finishes our proof.

□

Now we can prove Proposition 4.13.

Proof of Proposition 4.13. Let $t = \alpha \log n$, where α is as in Proposition 4.27. We say that two configurations are adjacent if they differ only by one jump. (4.52) implies that for any x, y such that x, y are adjacent and $\|x\|_\infty \vee \|y\|_\infty \leq \alpha_1 \log n$, $\mathrm{d_{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \mathbb{P}_y\left[X_t \in \cdot\right]\right) = \mathcal{O}\left(n^{-3}\right)$. Now for x, y arbitrary such that $\|x\|_\infty \vee \|y\|_\infty \leq \alpha_1 \log n$, we can always connect x and y by a path, i.e. a sequence $(\omega_0, \omega_1, \dots, \omega_k)$ in $\mathcal{X}$ such that $\omega_0 = x, \omega_k = y$, and ω_{l-1} is adjacent to ω_l for $1 \leq l \leq k$. Furthermore, we can pick one of the shortest paths to make sure that $k \leq m$ and

$$\max_{1 \leq l \leq k} \|\omega_l\|_\infty \leq \|x\|_\infty \vee \|y\|_\infty \leq \alpha_1 \log n.$$

Then by triangle inequality,

$$\mathrm{d_{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \mathbb{P}_y\left[X_t \in \cdot\right]\right) \leq \sum_{u=1}^{k} \mathrm{d_{TV}}\left(\mathbb{P}_{\omega_{u-1}}\left[X_t \in \cdot\right], \mathbb{P}_{\omega_u}\left[X_t \in \cdot\right]\right) \leq m\mathcal{O}\left(n^{-3}\right) = \mathcal{O}\left(n^{-2}\right),$$

where the last equality is due to (4.9). By stationarity of π and convexity of $\mathrm{d_{TV}}(\cdot, \cdot)$,

$$\begin{aligned}
\mathrm{d_{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \pi\right) &\leq \sum_{y \in \mathcal{X}} \pi(y) \mathrm{d_{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \mathbb{P}_y\left[X_t \in \cdot\right]\right) \\
&= \sum_{\{y : \|y\|_\infty > \alpha_1 \log n\}} \pi(y) \mathrm{d_{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \mathbb{P}_y\left[X_t \in \cdot\right]\right) \\
&\quad + \sum_{\{y : \|y\|_\infty \leq \alpha_1 \log n\}} \pi(y) \mathrm{d_{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \mathbb{P}_y\left[X_t \in \cdot\right]\right) \\
&\leq \pi(\|y\|_\infty > \alpha_1 \log n) + \pi(\|y\|_\infty \leq \alpha_1 \log n) \mathcal{O}\left(n^{-2}\right).
\end{aligned}$$

Moreover, by letting $t \to \infty$ in Proposition 4.12, we obtain $\pi(\|y\|_\infty > \alpha_1 \log n) = \mathcal{O}\left(n^{-2}\right)$. Combining it with the above inequality, we deduce the claim. □

4 The Poincaré constant

This section is devoted to proving Theorem 4.4 and Corollary 4.5. We first recall a classical lemma for general Markov processes:

Lemma 4.28 (Lower bound on Poincaré constant)**.** *Let Ω be a finite state space, and let $\mathcal{L}$ be an irreducible reversible Markov generator on Ω. Fix $\gamma > 0$, and suppose that for any $(x, y) \in \Omega \times \Omega$ such that $\mathcal{L}(x, y) > 0$, there exists a coupling $\mathbb{P}_{x,y}$ of two processes with generator $\mathcal{L}$ starting from x and y such that*

$$\mathbb{E}_{x,y}\left[e^{\gamma \tau}\right] < \infty,$$

where τ is the coalescence time of the two processes. Then $\lambda_*(\mathcal{L}) > \gamma$.

Proof. Let $A = \max_{x,y : \mathcal{L}(x,y) > 0} \mathbb{E}_{x,y}\left[e^{\gamma \tau}\right]$. Then for any x, y such that $\mathcal{L}(x, y) > 0$,

$$\mathrm{d_{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \mathbb{P}_y\left[X_t \in \cdot\right]\right) \leq \mathbb{P}_{x,y}\left[\tau > t\right] \leq e^{-\gamma t} \mathbb{E}_{x,y}\left[e^{\gamma \tau}\right] \leq A e^{-\gamma t}.$$

Now for $(x, y) \in \Omega \times \Omega$ arbitrary, as $\mathcal{L}$ is irreducible, we can connect x and y by a path, i.e. a sequence $(\omega_0, \omega_1, \dots, \omega_k)$ in Ω such that $\omega_0 = x, \omega_k = y$, and $\mathcal{L}(\omega_{l-1}, \omega_l) > 0$, for $1 \leq l \leq k$.

Picking one of the shortest paths ensures that $k \le |\Omega|$. Hence by the triangle inequality,

$$\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \mathbb{P}_y\left[X_t \in \cdot\right]\right) \le \sum_{l=1}^{k} \mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_{\omega_{l-1}}\left[X_t \in \cdot\right], \mathbb{P}_{\omega_l}\left[X_t \in \cdot\right]\right) \le A|\Omega|e^{-\gamma t}.$$

By stationarity of π and convexity of $\mathrm{d}_{\mathrm{TV}}(\cdot,\cdot)$,

$$\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \pi\right) \le \sum_{y\in\Omega} \pi(y)\mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \mathbb{P}_y\left[X_t \in \cdot\right]\right) \le A|\Omega|e^{-\gamma t}.$$

We deduce that

$$-\frac{1}{t}\log\max_{x\in\Omega} \mathrm{d}_{\mathrm{TV}}\left(\mathbb{P}_x\left[X_t \in \cdot\right], \pi\right) \ge \gamma - \frac{1}{t}(\log A + \log|\Omega|).$$

The claim is obtained simply by letting $t \to \infty$. □

Thanks to Lemma 4.28, in order to prove $\lambda_* = \Omega(1)$, we just need to prove the following:

Proposition 4.29 (Exponential moment of coalescence time)**.** *There exists a dimension-free constant γ such that for all $(x,i,j) \in \mathcal{X} \times [n] \times [n]$, for the coupling of two Zero-Range processes starting from $x+\delta_i$, $x+\delta_j$ using tagged particles as described above,*

$$\mathbb{E}_{x,i,j}\left[e^{\gamma\tau}\right] < \infty. \tag{4.54}$$

Proof. We only need to prove the result for big enough n. Let α_1 be as in Proposition 4.12. Let $\theta_1 = 6/\alpha_1, \beta_1$ be a dimension-free constant, and $L_1 = L(\theta_1,\beta_1)$ as in Lemma 4.17. Let $\theta_2 < \theta_1$ be a constant that satisfies Proposition 4.21, and let $L_2 > L_1$ and Good be defined as in subsection 3.2. Let T_1 be the hitting time of the set $\{\varphi^{\theta_1} \le L_1\}$. (4.32) says that $\mathbb{E}_x\left[e^{\beta_1 T_1}\right]$ is finite for all $x \in \mathcal{X}$, so we only need to prove the result for $x \in \{\varphi^{\theta_1} \le L_1\}$. By abuse of notation, we will note $\mathbb{P}_*(\cdot)$ (resp. $\mathbb{E}_*(\cdot)$) for the maximum of $\mathbb{P}_{x,i,j}(\cdot)$ (resp. $\mathbb{E}_{x,i,j}(\cdot)$) taken over all x such that $\varphi^{\theta_1}(x) \le L_1$ and all $(i,j) \in [n]\times[n]$. We will prove that there exists γ such that $\mathbb{E}_*\left[e^{\gamma(\tau\wedge k)}\right]$ is bounded uniformly in k. Then the claim is proved simply by letting k tend to infinity using the Monotone Convergence Theorem.

Let $A = \mathsf{Good} \cup \{y : \exists x \in \mathsf{Good}, \mathcal{L}(x,y) > 0\}$. Let T_{bad} be the exit time from Good, and let $T_2 := T_{bad} \wedge \alpha_3 \log n$, where α_3 is defined in Corollary 4.26. In fact, A is the set of all possible values of X up to time T_{bad}, and in particular, $X_{T_2} \in A$.By our definitions of the dimension free constants, when n is large enough, $\{\varphi^{\theta_1} \le L_1\} \subset \{\varphi^{\theta_2} \le L_2\} \cap \{\|\cdot\|_\infty \le \alpha_1 \log n\}$. Consequently, by Proposition 4.21 and Proposition 4.27,

$$\mathbb{P}_*\left[\tau \ge T_2\right] = \mathcal{O}\left(n^{-3}\right).$$

Note that $T_2 \le \alpha_3 \log n$, and hence $\mathbb{E}_*\left[e^{\gamma(\tau\wedge k)}\mathbb{1}_{\{\tau<T_2\}}\right] \le e^{\gamma\alpha_3\log n} = n^{\gamma\alpha_3}$. We deduce that

$$\mathbb{E}_*\left[e^{\gamma(\tau\wedge k)}\right] \le n^{\gamma\alpha_3} + \mathbb{E}_*\left[e^{\gamma(\tau\wedge k)}\mathbb{1}_{\{\tau\ge T_2\}}\right].$$

Conditionally on $\mathcal{F}_{T_2}$, by the strong Markov property, we have

$$\mathbb{E}_*\left[e^{\gamma(\tau\wedge k)}\mathbb{1}_{\{\tau\geq T_2\}}\right] \leq \mathbb{E}_*\left[e^{\gamma(T_2\wedge k)}\mathbb{1}_{\{\tau\geq T_2\}}\mathbb{E}_{X_{T_2}}\left[e^{\gamma(\tau\wedge k)}\right]\right]$$
$$\leq n^{\gamma\alpha_3}\mathbb{P}_*\left[\tau\geq T_2\right]\max_{y\in A}\mathbb{E}_y\left[e^{\gamma T_1}\right]\mathbb{E}_*\left[e^{\gamma(\tau\wedge k)}\right],$$

where in the last inequality we use the fact that $X_{T_2}\in A$, almost surely. Putting things together, we obtain

$$\mathbb{E}_*\left[e^{\gamma(\tau\wedge k)}\right] \leq n^{\gamma\alpha_3} + \mathcal{O}\left(n^{\gamma\alpha_3-3}\right)\max_{y\in A}\mathbb{E}_y\left[e^{\gamma T_1}\right]\mathbb{E}_*\left[e^{\gamma(\tau\wedge k)}\right]. \tag{4.55}$$

Consequently,

$$\mathbb{E}_*\left[e^{\gamma(\tau\wedge k)}\right] \leq \frac{n^{\gamma\alpha_3}}{1-\mathcal{O}\left(n^{\gamma\alpha_3-3}\right)\max_{y\in A}\mathbb{E}_y\left[e^{\gamma T_1}\right]},$$

provided that the denominator of the right-hand side is positive. Note that for $y\in A$, $\|y\|_\infty = \mathcal{O}\left(\log n\right)$, and hence $\varphi^{\theta_1}(y) < n^p$, for some dimension-free constant p. Then by (4.32) and Jensen's inequality, for all $y\in A$ and $\gamma<\beta_1$,

$$\mathbb{E}_y\left[e^{\gamma T_1}\right] \leq \mathbb{E}_y\left[e^{\beta_1 T_1}\right]^{\gamma/\beta_1} = \mathcal{O}\left(n^{\gamma p/\beta_1}\right).$$

Then the denominator above is $1-\mathcal{O}\left(n^{\gamma\alpha_3+\gamma p/\beta_1-3}\right)$, which is positive when γ is small enough. This finishes our proof. □

We now prove Theorem 4.4 and Corollary 4.5.

Proof of Theorem 4.4 and Corollary 4.5. In [39] (more precisely, in Corollary 3 and Lemma 13), the authors prove that for P a doubly stochastic transition matrix on $[n]$,

$$\lambda_*(\mathcal{L}) \leq \frac{\lambda_*(\mathcal{L}^P)}{\lambda_*(P)} \leq (1-1/n)\frac{\mathbb{E}\left[r(X_\infty(1))\right]}{\mathrm{Var}\left[X_\infty(1)\right]}.$$

Lemma 4.28 and Proposition 4.29 readily imply that $\lambda_*(\mathcal{L}) = \Omega(1)$, and hence so is $\dfrac{\lambda_*(\mathcal{L}^P)}{\lambda_*(P)}$. It remains to prove that

$$\frac{\mathbb{E}\left[r(X_\infty(1))\right]}{\mathrm{Var}\left[X_\infty(1)\right]} = \mathcal{O}\left(1\right). \tag{4.56}$$

We consider two cases: the case where the density is bounded away from zero : $\dfrac{1}{2}\leq\mathbb{E}\left[X_\infty(1)\right] = \dfrac{m}{n}\leq\rho$, and the case of low density: $\mathbb{E}\left[X_\infty(1)\right] = \dfrac{m}{n} < \dfrac{1}{2}$.

In case the density is bounded away from zero, it is easy to deduce from Proposition 4.10 that $\mathbb{E}\left[r(X_\infty(1))\right] = \Theta(1)$ and $\mathrm{Var}\left[X_\infty(1)\right] = \Theta(1)$, and hence the claim.

In the case of low density, we have $\mathrm{Var}\left[X_\infty(1)\right] \geq \mathbb{E}\left[X_\infty(1)\right] - \mathbb{E}\left[X_\infty(1)\right]^2$ and $\mathbb{E}\left[r(X_\infty(1))\right] \leq$

$\sup_{k\in\mathbb{Z}_+} \frac{r(k)}{k}\mathbb{E}\left[X_\infty(1)\right]$. We deduce that

$$\frac{\mathbb{E}\left[r(X_\infty(1))\right]}{\mathrm{Var}\left[X_\infty(1)\right]} \leq \frac{\sup_{k\in\mathbb{Z}_+} \frac{r(k)}{k}}{1-\mathbb{E}\left[X_\infty(1)\right]} < 2 \sup_{k\in\mathbb{Z}_+} \frac{r(k)}{k},$$

which finishes the proof. □

Finally, we compute the Poincaré constant of the transition matrix P in Example 4.6.

Proof of example 4.6. Let P_1 be the transition matrix of the simple random walk on $\mathbb{Z}/p\mathbb{Z}$. It is not hard to see that both P and P_1 are reversible w.r.t the uniform measures on their domains. Note that, for $f_1, \ldots, f_d$ some eigenfunctions of P_1 with eigenvalues $\lambda_1, \ldots, \lambda_d$ respectively, the function $f : (\mathbb{Z}/p\mathbb{Z})^d \to \mathbb{R}$ defined by

$$f(x(1), \ldots, x(d)) = \prod_{i=1}^{d} f_i(x(i))$$

is an eigenfunction of P with eigenvalue $\lambda := \frac{\sum_{i=1}^d \lambda_i}{d}$. Moreover, as the eigenfunctions of P_1 generate the space of functions from $\mathbb{Z}/p\mathbb{Z}$ to $\mathbb{R}$ due to reversibility, the functions f of the form above also generate the space of functions from $(\mathbb{Z}/p\mathbb{Z})^d$ to $\mathbb{R}$. Hence every eigenvalue of P is of the form $\lambda = \frac{\sum_{i=1}^d \lambda_i}{d}$, where $\lambda_i,\ 1 \leq i \leq d$, are some eigenvalues of P_1. It is well-known that the eigenvalues of P_1 are $\cos\left(\frac{2\pi k}{p}\right),\ 0 \leq k \leq p$ (see e.g. Chapter 12 of [59]). Due to reversibility of P, $\lambda_*(P)$ is the smallest eigenvalue of $I - P$ (see e.g. Chapter 12 of [59]), so it is given by

$$\lambda_*(P) = 1 - \frac{1}{d}\left(\cos\left(\frac{2\pi}{p}\right) + d - 1\right) = \frac{1}{d}\left(1 - \cos\left(\frac{2\pi}{p}\right)\right) \approx \frac{1}{d} \cdot \frac{2\pi^2}{p^2},$$

which finishes our proof. □

Bibliographie

[1] M. Aizenman and R. Holley. "Rapid convergence to equilibrium of stochastic Ising models in the Dobrushin Shlosman regime". *Percolation theory and ergodic theory of infinite particle systems (Minneapolis, Minn., 1984–1985).* Vol. 8. IMA Vol. Math. Appl. Springer, New York, 1987, pp. 1–11 (cit. on p. 47).

[2] David Aldous and Persi Diaconis. "Shuffling cards and stopping times". *American Mathematical Monthly* 93.5 (1986), pp. 333–348 (cit. on p. 5).

[3] David Aldous and Daniel Lanoue. "A lecture on the averaging process". *Probability Surveys* 9 (2012), pp. 90–102 (cit. on pp. 98, 102).

[4] Frank Ball and Geoffrey F. Yeo. "Lumpability and marginalisability for continuous-time Markov chains". *Journal of Applied Probability* 30.3 (1993), pp. 518–528 (cit. on p. 105).

[5] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim. "Large deviations for the boundary driven symmetric simple exclusion process". *Mathematical Physics, Analysis and Geometry. An International Journal Devoted to the Theory and Applications of Analysis and Geometry to Physics* 6.3 (2003), pp. 231–267 (cit. on pp. 16, 48).

[6] Julius Borcea, Petter Brändén, and Thomas M. Liggett. "Negative dependence and the geometry of polynomials". *Journal of the American Mathematical Society* 22.2 (2009), pp. 521–567 (cit. on p. 23).

[7] Charles Bordenave, Pietro Caputo, and Justin Salez. "Cutoff at the "entropic time" for sparse Markov chains". *Probability Theory and Related Fields* 173.1-2 (2019), pp. 261–292 (cit. on p. 5).

[8] Charles Bordenave and Hubert Lacoin. "Cutoff at the entropic time for random walks on covered expander graphs". *Journal of the Institute of Mathematics of Jussieu. JIMJ. Journal de l'Institut de Mathématiques de Jussieu* 21.5 (2022), pp. 1571–1616 (cit. on p. 5).

[9] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence.* Oxford University Press, Oxford, 2013, pp. x+481 (cit. on p. 125).

[10] Anton Bovier and Frank den Hollander. *Metastability.* Vol. 351. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Cham, 2015, pp. xxi+581 (cit. on p. 46).

[11] R. Bubley and M. Dyer. "Path coupling: A technique for proving rapid mixing in Markov chains". Miami Beach, FL, USA: IEEE, 1997, pp. 223–231 (cit. on p. 139).

[12] Pietro Caputo. "Spectral gap inequalities in product spaces with conservation laws". *Stochastic analysis on large scale interacting systems.* Vol. 39. Adv. Stud. Pure Math. Math. Soc. Japan, Tokyo, 2004, pp. 53–88 (cit. on p. 121).

[13] Pietro Caputo, Matteo Quattropani, and Federico Sau. "Cutoff for the averaging process on the hypercube and complete bipartite graphs". *Electronic Journal of Probability* 28 (2023), Paper No. 100, 31 (cit. on pp. 5, 98).

[14] J. Chayes, L. Chayes, and R. Schonmann. "Exponential decay of connectivities in the two-dimensional ising model". *Journal of Statistical Physics* (1987) (cit. on p. 47).

[15] Guan-Yu Chen and Laurent Saloff-Coste. "Computing cutoff times of birth and death chains". *Electronic Journal of Probability* 20 (2015), no. 76, 47 (cit. on p. 5).

[16] J. Theodore Cox, Yuval Peres, and Jeffrey E. Steif. "Cutoff for the noisy voter model". *The Annals of Applied Probability* 26.2 (2016), pp. 917–932 (cit. on p. 19).

[17] A. De Masi, P. A. Ferrari, and J. L. Lebowitz. "Rigorous derivation of reaction-diffusion equations with fluctuations". *Physical Review Letters* 55.19 (1985), pp. 1947–1949 (cit. on pp. 43, 45).

[18] A. De Masi, P. A. Ferrari, and J. L. Lebowitz. "Reaction-diffusion equations for interacting particle systems". *Journal of Statistical Physics* 44.3-4 (1986), pp. 589–644 (cit. on pp. 43–45, 48, 60).

[19] A. De Masi, T. Funaki, E. Presutti, and M. E. Vares. "Fast-reaction limit for Glauber-Kawasaki dynamics with two components". *ALEA. Latin American Journal of Probability and Mathematical Statistics* 16.2 (2019), pp. 957–976 (cit. on p. 43).

[20] B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier. "Exact solution of a 1D asymmetric exclusion model using a matrix formulation". *Journal of Physics A: Mathematical and General* 26 (1993), pp. 1493–1517 (cit. on pp. 16, 18, 21).

[21] Persi Diaconis. "The cutoff phenomenon in finite Markov chains". *Proceedings of the National Academy of Sciences of the United States of America* 93.4 (1996), pp. 1659–1664 (cit. on p. 5).

[22] Persi Diaconis and Laurent Saloff-Coste. "Separation cut-offs for birth and death chains". *The Annals of Applied Probability* 16.4 (2006), pp. 2098–2122 (cit. on p. 5).

[23] Persi Diaconis and Mehrdad Shahshahani. "Generating a random permutation with random transpositions". *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 57.2 (1981), pp. 159–179 (cit. on p. 5).

[24] Persi Diaconis and Mehrdad Shahshahani. "Time to reach stationarity in the Bernoulli-Laplace diffusion model". *SIAM Journal on Mathematical Analysis* 18.1 (1987), pp. 208–218 (cit. on p. 48).

[25] Jian Ding, Eyal Lubetzky, and Yuval Peres. "The mixing time evolution of Glauber dynamics for the mean-field Ising model". *Communications in Mathematical Physics* 289.2 (2009), pp. 725–764 (cit. on p. 47).

[26] Jian Ding, Eyal Lubetzky, and Yuval Peres. "Mixing time of critical Ising model on trees is polynomial in the height". *Communications in Mathematical Physics* 295.1 (2010), pp. 161–207 (cit. on p. 47).

[27] Jian Ding, Eyal Lubetzky, and Yuval Peres. "Total variation cutoff in birth-and-death chains". *Probability Theory and Related Fields* 146.1-2 (2010), pp. 61–85 (cit. on p. 5).

[28] R. L. Dobrushin and S. B. Shlosman. "Completely analytical interactions: constructive description". *Journal of Statistical Physics* 46.5-6 (1987), pp. 983–1014 (cit. on p. 47).

[29] Dor Elboim and Dominik Schmid. "Mixing times and cutoff for the TASEP in the high and low density phase" (2022). eprint: 2208.08306 (math.PR) (cit. on p. 5).

[30] Dor Elboim and Allan Sly. *Infinite cycles in the interchange process in five dimensions*. 2023. arXiv: 2211.17023 [math.PR] (cit. on p. 50).

[31] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes – characterization and convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. New York: John Wiley & Sons Inc., 1986, pp. x+534 (cit. on pp. 5, 72, 100).

[32] Davide Gabrielli, Giovanni Jona-Lasinio, Claudio Landim, and Maria Eulalia Vares. "Microscopic reversibility and thermodynamic fluctuations". *Boltzmann's legacy 150 years after his birth (Rome, 1994)*. Vol. 131. Atti Convegni Lincei. Accad. Naz. Lincei, Rome, 1997, pp. 79–87 (cit. on p. 49).

[33] Shirshendu Ganguly and Insuk Seo. "Information percolation and cutoff for the random-cluster model". *Random Structures & Algorithms* 57.3 (2020), pp. 770–822 (cit. on p. 47).

[34] Nina Gantert, Evita Nestoridi, and Dominik Schmid. "Mixing times for the simple exclusion process with open boundaries". *The Annals of Applied Probability* 33.2 (2023), pp. 972–1012 (cit. on pp. 15–17).

[35] Sara van de Geer. "Exponential inequalities for martingales, with application to maximum likelihood estimation for counting processes". *The Annals of Statistics* 23.5 (1995), pp. 1779–1801 (cit. on p. 124).

[36] P. Gonçalves, M. Jara, O. Menezes, and A. Neumann. "Non-equilibrium and stationary fluctuations for the SSEP with slow boundary". *Stochastic Processes and their Applications* 130.7 (2020), pp. 4326–4357 (cit. on p. 48).

[37] Patrícia Gonçalves, Milton Jara, Rodrigo Marinho, and Otávio Menezes. "Sharp Convergence to Equilibrium for the SSEP with Reservoirs" (2021). arXiv: 2110.06353 [math.PR] (cit. on pp. 12, 16–18, 20, 38, 48).

[38] Jonathan Hermon. "Cutoff for Ramanujan graphs via degree inflation". *Electronic Communications in Probability* 22 (2017), Paper No. 45, 10 (cit. on p. 5).

[39] Jonathan Hermon and Justin Salez. "A version of Aldous' spectral-gap conjecture for the zero range process". *The Annals of Applied Probability* 29.4 (2019), pp. 2217–2229 (cit. on pp. 121, 122, 147).

[40] Jonathan Hermon and Justin Salez. “Cutoff for the mean-field zero-range process with bounded monotone rates”. *The Annals of Probability* 48.2 (2020), pp. 742–759 (cit. on pp. 5, 121, 122).

[41] P. Hohenberg and B. Halperin. “Theory of dynamic critical phenomena” (1977) (cit. on p. 47).

[42] Richard Holley. “On the asymptotics of the spin-spin autocorrelation function in stochastic Ising models near the critical temperature”. *Spatial stochastic processes*. Vol. 19. Progr. Probab. Birkhäuser Boston, Boston, MA, 1991, pp. 89–104 (cit. on p. 47).

[43] Richard Holley and Daniel Stroock. “Logarithmic Sobolev inequalities and stochastic Ising models”. *Journal of Statistical Physics* 46.5-6 (1987), pp. 1159–1194 (cit. on p. 47).

[44] Richard A. Holley and Daniel W. Stroock. “Uniform and L^2 convergence in one-dimensional stochastic Ising models”. *Communications in Mathematical Physics* 123.1 (1989), pp. 85–93 (cit. on p. 47).

[45] Dmitry Ioffe. “Exact large deviation bounds up to T_c for the Ising model in two dimensions”. *Probability Theory and Related Fields* 102.3 (1995), pp. 313–330 (cit. on p. 47).

[46] C. Kipnis, S. Olla, and S. R. S. Varadhan. “Hydrodynamics and large deviation for simple exclusion processes”. *Communications on Pure and Applied Mathematics* 42.2 (1989), pp. 115–137 (cit. on pp. 45, 48).

[47] Claude Kipnis and Claudio Landim. *Scaling limits of interacting particle systems*. Vol. 320. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1999, pp. xvi+442 (cit. on pp. 16, 46, 48).

[48] Cyril Labbé and Hubert Lacoin. “Cutoff phenomenon for the asymmetric simple exclusion process and the biased card shuffling”. *The Annals of Probability* 47.3 (2019), pp. 1541–1586 (cit. on p. 5).

[49] Cyril Labbé and Hubert Lacoin. “Mixing time and cutoff for the weakly asymmetric simple exclusion process”. *The Annals of Applied Probability* 30.4 (2020), pp. 1847–1883 (cit. on p. 5).

[50] Hubert Lacoin. “Mixing time and cutoff for the adjacent transposition shuffle and the simple exclusion”. *The Annals of Probability* 44.2 (2016), pp. 1426–1487 (cit. on pp. 5, 6, 16, 17, 48, 121).

[51] Hubert Lacoin. “The cutoff profile for the simple exclusion process on the circle”. *The Annals of Probability* 44.5 (2016), pp. 3399–3430 (cit. on pp. 5, 12, 16, 20, 48, 121).

[52] Hubert Lacoin. “The simple exclusion process on the circle has a diffusive cutoff window”. *Annales de l’Institut Henri Poincaré Probabilités et Statistiques* 53.3 (2017), pp. 1402–1437 (cit. on pp. 5, 6, 16, 48).

[53] Hubert Lacoin and Rémi Leblond. “Cutoff phenomenon for the simple exclusion process on the complete graph”. *ALEA. Latin American Journal of Probability and Mathematical Statistics* 8 (2011), pp. 285–301 (cit. on p. 48).

[54] C. Landim, A. Milanés, and S. Olla. “Stationary and nonequilibrium fluctuations in boundary driven exclusion processes”. *Markov Processes and Related Fields* 14.2 (2008), pp. 165–184 (cit. on p. 16).

[55] Kent Bækgaard Lauritsen and Hans C. Fogedby. “Critical exponents from power spectra”. *Journal of Statistical Physics* 72.1-2 (1993), pp. 189–205 (cit. on p. 47).

[56] Gregory F. Lawler. *Introduction to stochastic processes*. Second. Chapman & Hall/CRC, Boca Raton, FL, 2006, pp. xiv+234 (cit. on p. 5).

[57] Gregory F. Lawler and Vlada Limic. *Random Walk: A Modern Introduction*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2010 (cit. on pp. 116, 117).

[58] Tzong-Yow Lee and Horng-Tzer Yau. “Logarithmic Sobolev inequality for some models of random walks”. *The Annals of Probability* 26.4 (1998), pp. 1855–1873 (cit. on pp. 16, 48).

[59] David A. Levin and Yuval Peres. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2017, pp. xvi+447 (cit. on pp. 5, 8, 9, 23, 24, 26, 37, 48, 49, 78, 79, 105, 148).

[60] Thomas M. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*. Vol. 324. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1999, pp. xii+332 (cit. on p. 16).

[61] Thomas M. Liggett. *Interacting particle systems*. Classics in Mathematics. Springer-Verlag, Berlin, 2005, pp. xvi+496 (cit. on p. 16).

[62] Thomas M. Liggett. *Continuous time Markov processes*. Vol. 113. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2010, pp. xii+271 (cit. on pp. 2, 5, 44, 124).

[63] Sheng Lin Lu and Horng-Tzer Yau. “Spectral gap and logarithmic Sobolev inequality for Kawasaki and Glauber dynamics”. *Communications in Mathematical Physics* 156.2 (1993), pp. 399–433 (cit. on p. 47).

[64] Eyal Lubetzky and Yuval Peres. "Cutoff on all Ramanujan graphs". *Geometric and Functional Analysis* 26.4 (2016), pp. 1190–1216 (cit. on p. 5).

[65] Eyal Lubetzky and Allan Sly. "Critical Ising on the square lattice mixes in polynomial time". *Communications in Mathematical Physics* 313.3 (2012), pp. 815–836 (cit. on p. 47).

[66] Eyal Lubetzky and Allan Sly. "Cutoff for the Ising model on the lattice". *Inventiones Mathematicae* 191.3 (2013), pp. 719–755 (cit. on p. 5).

[67] Eyal Lubetzky and Allan Sly. "An exposition to information percolation for the Ising model". *Annales de la Faculté des Sciences de Toulouse. Mathématiques. Série 6* 24.4 (2015), pp. 745–761 (cit. on pp. 5, 9, 19, 47).

[68] Eyal Lubetzky and Allan Sly. "Information percolation and cutoff for the stochastic Ising model". *Journal of the American Mathematical Society* 29.3 (2016), pp. 729–774 (cit. on pp. 5, 9, 18, 19, 47, 48, 60, 61, 64, 97).

[69] Eyal Lubetzky and Allan Sly. "Universality of cutoff for the Ising model". *The Annals of Probability* 45.6A (2017), pp. 3664–3696 (cit. on pp. 5, 9, 19, 47).

[70] Eyal Lubetzky and Allan Sly. "Fast initial conditions for Glauber dynamics". *Probability Theory and Related Fields* 181.1-3 (2021), pp. 647–667 (cit. on p. 47).

[71] F. Martinelli. "On the two-dimensional dynamical Ising model in the phase coexistence region". *Journal of Statistical Physics* 76.5-6 (1994), pp. 1179–1246 (cit. on p. 47).

[72] F. Martinelli and E. Olivieri. "Approach to equilibrium of Glauber dynamics in the one phase region. I. The attractive case". *Communications in Mathematical Physics* 161.3 (1994), pp. 447–486 (cit. on p. 47).

[73] F. Martinelli and E. Olivieri. "Approach to equilibrium of Glauber dynamics in the one phase region. II. The general case". *Communications in Mathematical Physics* 161.3 (1994), pp. 487–514 (cit. on p. 47).

[74] Fabio Martinelli. "Lectures on Glauber dynamics for discrete spin models". *Lectures on probability theory and statistics (Saint-Flour, 1997)*. Vol. 1717. Lecture Notes in Math. Springer, Berlin, 1999, pp. 93–191 (cit. on p. 48).

[75] Mathieu Merle and Justin Salez. "Cutoff for the mean-field zero-range process". *The Annals of Probability* 47.5 (2019), pp. 3170–3201 (cit. on pp. 121, 122).

[76] Jason Miller and Yuval Peres. "Uniformity of the uncovered set of random walk and cutoff for lamplighter chains". *The Annals of Probability* 40.2 (2012), pp. 535–577 (cit. on p. 9).

[77] Ben Morris. "The mixing time for simple exclusion". *The Annals of Applied Probability* 16.2 (2006), pp. 615–635 (cit. on pp. 16, 48, 121).

[78] Evita Nestoridi and Sam Olesker-Taylor. "Limit profiles for reversible Markov chains". *Probability Theory and Related Fields* 182.1-2 (2022), pp. 157–188 (cit. on p. 12).

[79] Roberto Imbuzeiro Oliveira. "Mixing of the symmetric exclusion processes in terms of the corresponding single-particle random walk". *The Annals of Probability* 41.2 (2013), pp. 871–913 (cit. on p. 48).

[80] Lars Onsager. "Crystal statistics. I. A two-dimensional model with an order-disorder transition". *Physical Review. Series II* 65 (1944), pp. 117–149 (cit. on p. 47).

[81] Narutaka Ozawa. "An entropic proof of cutoff on Ramanujan graphs". *Electronic Communications in Probability* 25 (2020), Paper No. 77, 8 (cit. on p. 5).

[82] James Gary Propp and David Bruce Wilson. "Exact sampling with coupled Markov chains and applications to statistical mechanics". *Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995)*. Vol. 9. 1-2. 1996, pp. 223–252 (cit. on p. 60).

[83] Philip E. Protter. *Stochastic integration and differential equations*. Vol. 21. Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin, 2005, pp. xiv+419 (cit. on p. 134).

[84] Matteo Quattropani and Federico Sau. "Mixing of the averaging process and its discrete dual on finite-dimensional geometries". *The Annals of Applied Probability* 33.2 (2023), pp. 936–971 (cit. on p. 98).

[85] Justin Salez. "Cutoff for non-negatively curved Markov chains". *J. Eur. Math. Soc* (2023) (cit. on p. 5).

[86] Justin Salez. "Spectral gap and curvature of monotone Markov chains" (2023). arXiv: `2305.04688 [math.PR]` (cit. on pp. 17, 21, 22).

[87] Justin Salez. "Universality of cutoff for exclusion with reservoirs". *The Annals of Probability* 51.2 (2023), pp. 478–494 (cit. on pp. 5, 9, 10, 16, 17, 19, 21–24, 48).

[88] R. Schonmann. “Second order large deviation estimates for ferromagnetic systems in the phase coexistence region”. *Communications in Mathematical Physics* (1987) (cit. on p. 47).

[89] Frank Spitzer. “Interaction of Markov processes”. *Advances in Mathematics* 5 (1970), 246–290 (1970) (cit. on p. 16).

[90] Daniel W. Stroock and Bogusław Zegarliński. “The equivalence of the logarithmic Sobolev inequality and the Dobrushin-Shlosman mixing condition”. *Communications in Mathematical Physics* 144.2 (1992), pp. 303–323 (cit. on p. 47).

[91] Daniel W. Stroock and Bogusław Zegarliński. “The logarithmic Sobolev inequality for continuous spin systems on a lattice”. *Journal of Functional Analysis* 104.2 (1992), pp. 299–326 (cit. on p. 47).

[92] Daniel W. Stroock and Bogusław Zegarliński. “The logarithmic Sobolev inequality for discrete spin systems on a lattice”. *Communications in Mathematical Physics* 149.1 (1992), pp. 175–193 (cit. on p. 47).

[93] Ryokichi Tanaka and Kenkichi Tsunoda. “Glauber-exclusion dynamics: rapid mixing regime”. *Electronic Journal of Probability* 27 (2022), Paper No. 141, 26 (cit. on pp. 43, 48, 50).

[94] Lucas Teyssier. “Limit profile for random transpositions”. *The Annals of Probability* 48.5 (2020), pp. 2323–2343 (cit. on pp. 5, 12).

[95] L. Thomas. “Bound on the mass gap for finite volume stochastic ising models at low temperature”. *Communications in Mathematical Physics* (1989) (cit. on p. 47).

[96] Bálint Tóth. “Improved lower bound on the thermodynamic pressure of the spin 1/2 Heisenberg ferromagnet”. *Letters in Mathematical Physics* 28.1 (1993), pp. 75–84 (cit. on p. 50).

[97] Hong Quan Tran. “The mean-field zero-range process with unbounded monotone rates: mixing time, cutoff, and Poincaré constant”. *The Annals of Applied Probability* 33.3 (2023), pp. 1732–1757 (cit. on p. 119).

[98] Hong-Quan Tran. *Cutoff for the Glauber-Exclusion process in the full high-temperature regime: an information percolation approach.* 2023. arXiv: `2310.15810 [math.PR]` (cit. on pp. 5, 10, 41).

[99] Hong-Quan Tran. “Cutoff for the non reversible SSEP with reservoirs”. *Electronic Journal of Probability* 28 (2023), Paper No. 1 (cit. on pp. 10, 15, 48).

[100] Kenkichi Tsunoda. “Exponentially slow mixing and hitting times of rare events for a reaction-diffusion model”. *ALEA. Latin American Journal of Probability and Mathematical Statistics* 19.2 (2022), pp. 1161–1184 (cit. on p. 48).

[101] Fu-Gao Wang and Chin-Kun Hu. “Universality in dynamic critical phenomena”. *Journal of Physics A: Mathematical and General* 56.2 (1997), pp. 2310–2313 (cit. on p. 47).

[102] David Bruce Wilson. “Mixing times of Lozenge tiling and card shuffling Markov chains”. *The Annals of Applied Probability* 14.1 (2004), pp. 274–325 (cit. on pp. 6, 12, 16, 48).

[103] Anthony J. Wood, Richard A. Blythe, and Martin R. Evans. “Combinatorial mappings of exclusion processes”. *Journal of Physics. A. Mathematical and Theoretical* 53.12 (2020), pp. 123001, 51 (cit. on p. 16).

[104] Horng-Tzer Yau. “Relative entropy and hydrodynamics of Ginzburg-Landau models”. *Letters in Mathematical Physics* 22.1 (1991), pp. 63–80 (cit. on p. 17).

[105] Horng-Tzer Yau. “Logarithmic Sobolev inequality for generalized simple exclusion processes”. *Probability Theory and Related Fields* 109.4 (1997), pp. 507–538 (cit. on p. 16).

[106] Bogusław Zegarliński. “On log-Sobolev inequalities for infinite lattice systems”. *Letters in Mathematical Physics* 20.3 (1990), pp. 173–182 (cit. on p. 47).

[107] Bogusław Zegarliński. “Dobrushin uniqueness theorem and logarithmic Sobolev inequalities”. *Journal of Functional Analysis* 105.1 (1992), pp. 77–111 (cit. on p. 47).

RÉSUMÉ

Sur un espace d'états fini, une chaîne de Markov irréductible à temps continu converge vers sa mesure stationnaire unique, ou en d'autres termes, *se mélange*. La convergence est mesurée par rapport à la distance en variation totale. Dans la théorie moderne des chaînes de Markov, nous nous intéressons aux chaînes où l'espace d'états devient grand. En étudiant certains modèles de mélange de cartes, Aldous, Diaconis et Shashahani ont découvert le phénomène remarquable maintenant connu sous le nom de *cutoff*: lorsque l'espace d'états devient grand, la distance entre la chaîne et l'équilibre reste proche de sa valeur maximale pendant une longue période, puis chute soudainement vers zéro sur une échelle de temps beaucoup plus courte. Depuis, le phénomène de cutoff a été observé dans de nombreux contextes différents, tels que les chaînes de naissance et de mort, les systèmes de spin à haute température, les systèmes de particules en interaction, etc. Malgré l'accumulation de modèles, il n'existe pas encore de théorie générale permettant de prédire efficacement cutoff. Au lieu de cela, le cutoff est montré modèle par modèle.

Dans cette thèse, nous étudions trois systèmes de particules en interaction: le processus d'exclusion unidimensionnel avec réservoirs, le processus de Glauber-Exclusion dans le régime à haut température, et le processus de Zero-Range à champ-moyen avec potentiel croissant sous-linéairement. Pour chaque modèle, nous établissons cutoff et fournissons une estimation fine pour le trou spectral. Nous nous concentrons particulièrement sur le cadre de la percolation de l'information introduit par Lubetzky et Sly, qui nous permet de montrer le cutoff même sans connaître la formule explicite de la mesure invariante.

MOTS CLÉS

temps de mélange, trou spectral, phénomène de cutoff, dynamique de Glauber, processus d'Exclusion, processus de Zero-Range, percolation de l'information

ABSTRACT

On a finite state space, an irreducible continuous-time Markov chain converges to its unique stationary measure, or in other words, *mixes*. The convergence is often measured by the total variation distance. In the modern theory of Markov Chains, we are interested in the case where the state space becomes large. When studying some models of card shuffling, Aldous, Diaconis, and Shashahani discovered a remarkable phenomenon now known as *cutoff*: as the state space becomes large, the distance between the chain and equilibrium stays close to its maximal value for a long time and then suddenly drops to near zero in a much shorter time scale. Since then, the cutoff phenomenon has been observed in many different contexts, such as birth and death chains, high-temperature spin systems, interacting particle systems, etc. Despite the accumulation of models, there is not yet a general theory to effectively predict cutoff. Instead, cutoff is proved model by model.

In this thesis, we study three models: the one-dimensional Exclusion process with reservoirs, the Glauber-Exclusion process in the high-temperature regime, and the mean-field Zero-Range process with increasing sublinear potential. These three models all fall under the category of interacting particle systems. For each model, we establish cutoff and provide a sharp estimate on the spectral gap. We particularly focus on the information percolation framework introduced by Lubetzky and Sly, which allows us to show cutoff even without knowing the explicit formula of the invariant measure.

KEYWORDS

mixing times, spectral gap, cutoff phenomenon, Glauber dynamics, Exclusion process, Zero-Range process, information percolation