

Hydrodynamic Limits and Propagation of Chaos for Interacting Random Walks in Domains

Zhen-Qing Chen and Wai-Tong (Louis) Fan

January 26, 2014

Abstract

A new non-conservative stochastic reaction-diffusion system in which two families of random walks in two adjacent domains interact near the interface is introduced and studied in this paper. Such a system can be used to model the transport of positive and negative charges in a solar cell or the population dynamics of two segregated species under competition. We show that in the macroscopic limit, the particle densities converge to the solution of a coupled nonlinear heat equations. For this, we first prove that propagation of chaos holds by establishing the uniqueness of a new BBGKY hierarchy. A local central limit theorem for reflected diffusions in bounded Lipschitz domains is also established as a crucial tool.

Contents

1	Introduction	2
1.1	Main results	4
1.2	Key ideas	4
2	Notations and Preliminaries	5
2.1	Interacting random walks in domains	6
2.1.1	Discrete approximation of surface measure	7
2.1.2	Reflected diffusion and random walk approximation	8
2.1.3	Random walks with interaction	9
2.2	Discrete heat kernel	10
2.3	Hydrodynamic limit: system of nonlinear PDEs	12
2.4	Main results (rigorous statements)	16
3	Propagation of Chaos	19
3.1	Duality	19
3.2	Annihilation near the interface	21
3.3	Uniform bound and equi-continuity	22
3.4	Limiting hierarchy	25
3.5	Uniqueness of infinite hierarchy	28
3.5.1	Duhamel tree expansion	29
3.5.2	Telescoping via Chapman-Kolmogorov equation	30
3.5.3	Comparison with a ‘dominating’ tree	33
3.5.4	Estimating J_N	34

3.5.5	Proof of uniqueness	38
4	Hydrodynamic Limits	39
4.1	Constructing martingales	39
4.2	Tightness	41
4.3	Identifying the limit	43
5	Local Central Limit Theorem	43
5.1	Discrete relative isoperimetric inequality	43
5.1.1	CTRW on scaled graph aD^ε	44
5.1.2	An extension lemma	44
5.1.3	Discrete relative isoperimetric inequality	48
5.2	Nash's inequality and Poincaré inequality	49
5.3	Mixing time	50
5.4	Gaussian bound and uniform Hölder continuity of p^ε	50
5.5	Proof of local CLT	51

1 Introduction

The original motivation of this project is to study the transports of positive and negative charges in solar cells. We model a solar cell by a domain in \mathbb{R}^d that is divided into two disjoint sub-domains D_+ and D_- by an interface I , a $(d - 1)$ -dimensional hypersurface, which can be possibly disconnected. D_+ and D_- represent the hybrid medium that confine the positive and the negative charges, respectively. At microscopic level, positive and negative charges are initially modeled by N independent reflected Brownian motion (RBM) with drift on D_+ and on D_- respectively. (In this paper, they are actually modeled by N independent random walks on lattices inside D_+ and D_- that serve as discrete approximation of RBM with drifts.) These random motions model the transport of positive (respectively negative) charges under an electric potential (See Figure 1).

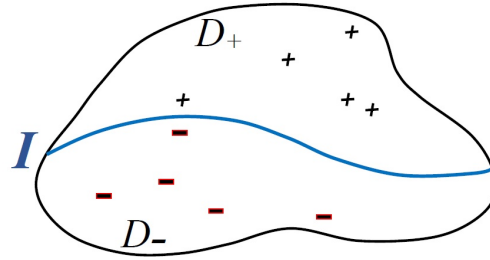


Figure 1: I is the interface of D_+ and D_-

These two types of particles annihilate each other at a certain rate when they come close to each other near the interface I . This interaction models the annihilation, trapping, recombination and separation phenomena of the charges. The interaction distance is of microscopic order ε where $N\varepsilon^d$ is comparable to 1, and the intensity of annihilation per pair is of order λ/ε where $\lambda \geq 0$ is a given parameter. This means that, intuitively and roughly speaking, according to a random time clock which runs with a speed proportional to the number of pairs (one particle of each type) of distance ε , we annihilate a pair (picked uniformly among those pairs of distance less than ε) with an exponential rate of parameter λ/ε . The above scaling guarantees that in the limit, a nontrivial proportion of particles are annihilated in any open time interval. We investigate the scaling limit of the empirical distribution of positive and negative charges;

that is, the hydrodynamic limit of this interacting diffusion system. We show that in the macroscopic level, the empirical distribution converges to a deterministic measure whose density satisfies a system of partial differential equations that has non-linear interaction at the interface.

The study of *hydrodynamic limits* of particle systems with stochastic dynamics is of fundamental importance in many areas. This study dates back to the sixth Hilbert problem in year 1900, which concerns the mathematical treatment of the axioms of physics, and to Boltzmann's work on principles of mechanics. Proving hydrodynamic limits corresponds to establishing *the law of large number* for the empirical measure of some attributes (such as position, genetic type, spin type, etc.) of the individuals in the systems. It contributes to our better understanding of the asymptotic behavior of many phenomena, such as chemical reactions [31], population dynamics [20, 32, 34], super-conductivity [40], quantum dynamics [21], fluid dynamics [24], etc. It reveals fascinating connections between the microscopic stochastic systems and deterministic partial differential equations that describe the macroscopic pictures. It also provides approximations via stochastic models to some partial differential equations that are hard or impossible to solve directly.

Since the work of Boltzmann and Hilbert, there have been many different lines of research on stochastic particle systems. Various models were constructed and different techniques were developed to establish hydrodynamic limits. Among those techniques, the entropy method [25] and the relative entropy method [40] are considered to be general methods. Unfortunately these methods do not seem to work for our model due to the singular interaction near the interface.

Many models studied in literature are conservative, for example exclusion processes [27, 28] and Fleming-Viot type systems [8, 9]. Reaction-diffusion systems (R-D system) constitute a class of models that are typically non-conservative. These are systems which have hydrodynamic limits of the form $\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u + R(u)$ (a reaction-diffusion equation, or R-D equation in short), where $R(u)$ is a function in u which is thought of as the reaction term. R-D systems arise from many different contexts and have been studied by many authors. For instance, for the case $R(u)$ is a polynomial in u , these systems contain the Schlögl's model and were studied in [18, 19, 29, 30] on a cube with Neumann boundary conditions, and in [4, 5] on a periodic lattice. Recently, perturbations of the voter models which contain the Lotka-Volterra systems are considered in [14]. In addition to results on hydrodynamic limit, [14] also established general conditions for the existence of non-trivial stationary measures and for extinction of the particles.

Our model is a non-conservative stochastic particle system which consists of two types of particles. In [9], Burdzy and Quastel studied an annihilating-branching system of two types of particles, for which the total number of particles of each type remains constant over the time. Its hydrodynamic limit is described by a linear heat equation with zero average temperature. In contrast, besides being non-conservative, our model gives rise to a system of nonlinear differential equations that seems to be new. Moreover, the interaction between two types of particles is singular near the interface of the two media, which gives rise to a boundary integral term in the hydrodynamic limit. The approach of this paper provides some new tools that are potentially useful for the study of other non-equilibrium systems.

We now give some more details on the discrete approximation of the spatial motions in our modeling. We approximate D_{\pm} by square lattices D_{\pm}^{ε} of side length ε , and then approximate reflected diffusions on D_{\pm} by continuous time random walks (CTRWs) on D_{\pm}^{ε} . The rigorous formulation of the particle system is captured by the operator $\mathfrak{L}^{\varepsilon}$ in (2.8).

Let $X_i^{\pm}(t)$ be the position of the particle with index i in \overline{D}_{\pm} at time t . We prescribe each particle a mass $1/N$ and consider the normalized empirical measures

$$\mathfrak{X}_t^{N,+}(dx) := \frac{1}{N} \sum_{\alpha:\alpha\sim t} \mathbf{1}_{X_{\alpha}^{+}(t)}(dx) \quad \text{and} \quad \mathfrak{X}_t^{N,-}(dy) := \frac{1}{N} \sum_{\beta:\beta\sim t} \mathbf{1}_{X_{\beta}^{-}(t)}(dy).$$

Here $\mathbf{1}_y(dx)$ stands for the Dirac measure concentrated at the point y , while $\alpha \sim t$ if and only if the particle X_{α}^{+} is alive at time t , and $\beta \sim t$ if and only if the particle X_{β}^{-} is alive at time t . For fixed positive integer N and $t > 0$, $\mathfrak{X}_t^{N,\pm}$ is a random measure on \overline{D}_{\pm} . We want to study the asymptotic behavior, when $N \rightarrow \infty$ (or equivalently $\varepsilon \rightarrow 0$), of the evolution in time t of the pair $(\mathfrak{X}_t^{N,+}, \mathfrak{X}_t^{N,-})$.

1.1 Main results

Our first main result (Theorem 2.20) implies the following. Suppose each particle in D_{\pm} is approximating a RBM with gradient drift $\frac{1}{2}\nabla(\log\rho_{\pm})$, where $\rho_{\pm} \in C^1(\overline{D}_{\pm})$ is strictly positive. Then under appropriate assumptions on the initial configuration $(\mathfrak{X}_0^{N,+}, \mathfrak{X}_0^{N,-})$, the normalized empirical measure $(\mathfrak{X}_t^{N,+}(dx), \mathfrak{X}_t^{N,-}(dy))$ converges in distribution to a deterministic measure

$$(u_+(t, x)\rho_+(x)dx, u_-(t, y)\rho_-(y)dy)$$

for all $t > 0$, where (u_+, u_-) is the solution of the following coupled heat equations:

$$\begin{cases} \frac{\partial u_+}{\partial t} = \frac{1}{2}\Delta u_+ + \frac{1}{2}\nabla(\log\rho_+) \cdot \nabla u_+ & \text{on } (0, \infty) \times D_+ \\ \frac{\partial u_+}{\partial \vec{n}_+} = \frac{\lambda}{\rho_+} u_+ u_- \mathbf{1}_I & \text{on } (0, \infty) \times \partial D_+ \end{cases} \quad (1.1)$$

and

$$\begin{cases} \frac{\partial u_-}{\partial t} = \frac{1}{2}\Delta u_- + \frac{1}{2}\nabla(\log\rho_-) \cdot \nabla u_- & \text{on } (0, \infty) \times D_- \\ \frac{\partial u_-}{\partial \vec{n}_-} = \frac{\lambda}{\rho_-} u_+ u_- \mathbf{1}_I & \text{on } (0, \infty) \times \partial D_- \end{cases} \quad (1.2)$$

where \vec{n}_{\pm} is the inward unit normal vector field on ∂D_{\pm} of D_{\pm} and $\mathbf{1}_I$ is the indicator function on I . Note that $\rho_{\pm} = 1$ corresponds to the particular case when there is no drift.

The above result tells us that for any fixed time $t > 0$, the probability distribution of a randomly picked particle in D_{\pm}^{ε} at time t is close to $c_{\pm}(t)u_{\pm}(t, x)$ when N is large, where $c_{\pm}(t) = (\int_{D_{\pm}^{\varepsilon}} u_{\pm}(t))^{-1}$ is a normalizing constant. In fact, the above convergence holds at the level of the path space. That is, the full trajectory (and hence the joint law at different times) of the particle profile converges to the deterministic scaling limit described by (1.1) and (1.2), not only its distribution at a given time.

Question: How about the limiting joint distribution of more than one particles ?

Our second main result (Theorem 2.22) answers this question. It asserts that **propagation of chaos** holds true for our system; that is, when the number of particles tends to infinity, their positions appear to be independent of each other. More precisely, suppose n and m unlabeled particles in D_+^{ε} and D_-^{ε} , respectively, are chosen uniformly among the living particles at time t . Then, as $N \rightarrow \infty$, the probability joint density function for their positions converges to

$$c_{(n,m)}(t) \prod_{i=1}^n u_+(t, r_i) \prod_{j=1}^m u_-(t, s_j)$$

uniformly for $(\vec{r}, \vec{s}) \in \overline{D}_+^n \times \overline{D}_-^m$ and for t in any compact time interval, where $c_{(n,m)}(t)$ is a normalizing constant.

1.2 Key ideas

A key step in our proof of propagation of chaos (Theorem 2.22) is Theorem 3.7. The latter establishes uniqueness of solution for the infinite system of equations satisfied by the correlation functions of the particles in the limit $N \rightarrow \infty$. Such infinite system of equations is sometimes called *BBGKY-hierarchy* in statistical physics. Our BBGKY hierarchy involves boundary terms on the interface, which is new to the literature. Our proof of uniqueness involves a representation and manipulations of the hierarchy in terms of trees. This technique is related to but different from that in [21] which used Feynman diagrams. It is potentially useful in the study of other stochastic models involving coupled differential equations.

To establish hydrodynamic limit result (Theorem 2.20), we employ the classical tightness plus finite dimensional distribution approach. Tightness of $(\mathfrak{X}^{N,+}, \mathfrak{X}^{N,-})$ in the Skorokhod space is proved in Theorem 4.4. This together with the propagation of chaos result (Theorem 2.22) establishes the hydrodynamic limit of the interacting random walks.

Two new tools for discrete approximation of random walks in domains are developed in this article. Namely, the local central limit theorem (local CLT) for reflected random walk on bounded Lipschitz domains (Theorem 2.12) and the ‘discrete surface measure’ (Lemma 2.4). We believe these tools are potentially useful in many discrete schemes which involve reflected Brownian motions.

Weak convergence of simple random walk on D_{\pm}^{ε} to RBM has been established for general bounded domains in [6] and [7]. However, we need more for our model; namely a local convergence result which guarantees that the convergence rate is uniform up to the boundary. For this, we establish the local CLT. We further generalize the weak convergence result and the local limit theorem to deal with RBMs with gradient drift. There are two reasons for us to consider gradient drift. First, it is physically natural to assume the particles are subject to an electric potential. Second, the maximal extension theorem, [13, Theorem 6.6.9], which is a crucial technical tool used in [6] and [7], has established only in symmetric setting. The proof of the local CLT is based on a ‘discrete relative isoperimetric inequality’ (Theorem 5.5) which leads to the Poincaré inequality and the Nash inequality. The crucial point is that these two inequalities are uniform in ε (scaling of lattice size) and is invariant under the dilation of the domain $D \mapsto aD$.

The paper is organized as follows. In section 2, we introduce the stochastic model and some preliminary facts that will be used later. We then prove the existence and uniqueness of solution for the coupled PDE. The main results, Theorem 2.20 and Theorem 2.22, will be rigorously formulated. We also mention various extensions of our main results in Remark 2.25. Section 3 and section 4 contains the proof of Theorem 2.22 and Theorem 2.20 respectively. Section 5 is devoted to the proofs of the discrete relative isoperimetric inequality and the local CLT.

2 Notations and Preliminaries

For the reader’s convenience, we list our notations here:

\mathbb{Z}	set of all integers
\mathbb{Z}_+	$\{1, 2, 3, \dots\}$ positive integers
\mathbb{N}	$\{0, 1, 2, \dots\}$ non-negative integers
\mathbb{R}	set of all real numbers
$\mathcal{B}(E)$	Borel measurable functions on E
$\mathcal{B}_b(E)$	bounded Borel measurable functions on E
$\mathcal{B}^+(E)$	non-negative Borel measurable functions on E
$C(E)$	continuous functions on E
$C_b(E)$	bounded continuous functions on E
$C^+(E)$	non-negative continuous functions on E
$C_c(E)$	continuous functions on E with compact support
$D([0, \infty), E)$	space of <i>càdlàg</i> paths from $[0, \infty)$ to E equipped with the Skorokhod metric
\mathcal{H}^m	m -dimensional Hausdorff measure
$M_+(E)$	space of finite non-negative Borel measures on E equipped with the weak topology
$M_{\leq 1}(E)$	$\{\mu \in M_+(E) : \mu(E) \leq 1\}$
$M_1(E)$ (or $\mathcal{P}(E)$)	$\{\mu \in M_+(E) : \mu(E) = 1\}$
$\eta_t^{\varepsilon, \pm}(x)$	number of living particles at $x \in D_{\pm}^{\varepsilon}$ at time t
$(\eta_t^{\varepsilon})_{t \geq 0}$	process with generator $\mathfrak{L}^{\varepsilon} = \mathfrak{L}_0^{\varepsilon} + \mathfrak{R}^{\varepsilon}$ in Definition 2.8
$(\xi_t^0)_{t \geq 0}$ and $(\eta_t^0)_{t \geq 0}$	independent processes with generator $\mathfrak{L}_0^{\varepsilon}$
E^{ε}	$\mathbb{N}^{D_+^{\varepsilon}} \times \mathbb{N}^{D_-^{\varepsilon}}$, state space of $(\eta_t^{\varepsilon})_{t \geq 0}$
$\mathfrak{X}_t^{N, \pm}(dz)$	$\frac{1}{N} \sum_{x \in D_{\pm}^{\varepsilon}} \eta_t^{\pm}(x) \mathbf{1}_x(dz)$, the normalized empirical measure in \overline{D}_{\pm}

\mathfrak{E}	$M_{\leq 1}(\overline{D}_+) \times M_{\leq 1}(\overline{D}_-)$, the state space of $(\mathfrak{X}_t^{N,+}, \mathfrak{X}_t^{N,-})_{t \geq 0}$
$\{\mathcal{F}_t^X : t \geq 0\}$	filtration induced by the process (X_t) , i.e. $\mathcal{F}_t^X = \sigma(X_s, s \leq t)$
$\mathbf{1}_x$	indicator function at x or the Dirac measure at x (depending on the context)
$\xrightarrow{\mathcal{L}}$	weak convergence of random variables (or processes)
$\langle f, \mu \rangle$	$\int f(x) \mu(dx)$
$x \vee y$	$\max\{x, y\}$
$x \wedge y$	$\min\{x, y\}$
C, C_1, C_2, \dots	positive constants
I^ε	' ε -point approximation' of I constructed in Lemma 2.4
σ_ε	'discrete surface measure' constructed in Lemma 2.4

Process	Semigroup	Heat kernel	Measure	Generator	State space
$X^\pm(t)$	P_t^\pm	$p^\pm(t, x, y)$	ρ_\pm	\mathcal{A}^\pm	\overline{D}_\pm
$X^{\varepsilon, \pm}(t)$	$P_t^{\varepsilon, \pm}$	$p^{\varepsilon, \pm}(t, x, y)$	m_ε^\pm	$\mathcal{A}_\varepsilon^\pm$	D_\pm^ε
$X_{(n,m)}(t)$	$P_t^{(n,m)}$	$p = p^{(n,m)}$	$\rho = \rho_{(n,m)}$	$\mathcal{A}^{(n,m)}$	$\overline{D}_+^n \times \overline{D}_-^m$
$X_{(n,m)}^\varepsilon(t)$	$P_t^{(n,m), \varepsilon}$	$p^\varepsilon = p^{(n,m), \varepsilon}$	m_ε	$\mathcal{A}_\varepsilon^{(n,m)}$	$(D_+^\varepsilon)^n \times (D_-^\varepsilon)^m$

where in the above,

$$p^{(n,m)}(t, (\vec{r}, \vec{s}), (\vec{r}', \vec{s}')) := \prod_{i=1}^n p^+(t, r_i, r'_i) \prod_{j=1}^m p^-(t, s_j, s'_j)$$

$$\rho_{(n,m)}(\vec{r}, \vec{s}) := \prod_{i=1}^n \rho_+(r_i) \prod_{j=1}^m \rho_-(s_j).$$

We also use the following abbreviations:

a.s.	almost surely
LDCT	Lebesgue dominated convergence theorem
CTRW	continuous time random walk
RBM	reflected Brownian motion
local CLT	local central limit theorem
LHS	left hand side
RHS	right hand side
WLOG	without loss of generality

Definition 2.1. A Borel subset E of \mathbb{R}^d is called \mathcal{H}^m -**rectifiable** if E is a countable union of Lipschitz images of bounded subsets of \mathbb{R}^m with $\mathcal{H}^m(E) < \infty$ (As usual, we ignore sets of \mathcal{H}^m measure 0). Here \mathcal{H}^m denotes the m -dimensional Hausdorff measure.

Definition 2.2. A **bounded Lipschitz domain** $D \subset \mathbb{R}^d$ is a bounded connected open set such that for any $\xi \in \partial D$, there exists $r_\xi > 0$ such that $B(\xi, r_\xi) \cap D$ is represented by $B(\xi, r_\xi) \cap \{(y', y^d) \in \mathbb{R}^d : \phi_\xi(y') < y^d\}$ for some coordinate system centered at ξ and a Lipschitz function ϕ_ξ with Lipschitz constant M , where $M = M_D > 0$ does not depend on ξ and is called the Lipschitz constant of D .

Assumption 2.3. D_\pm are given adjacent bounded Lipschitz domains in \mathbb{R}^d such that $I := \overline{D}_+ \cap \overline{D}_- = \partial D_+ \cap \partial D_-$ is a finite union of disjoint connected \mathcal{H}^{d-1} -rectifiable sets, $\rho_\pm \in W^{(1,2)}(D_\pm) \cap C^1(\overline{D}_\pm)$ are given functions which are strictly positive, $\lambda > 0$ is a fixed parameter.

2.1 Interacting random walks in domains

In this subsection, we describe the interacting random walk model. We start with some key ingredients needed in discrete approximation.

2.1.1 Discrete approximation of surface measure

To capture the boundary behavior of our processes near the interface I in the discrete scheme, we need a discrete approximation of the surface measure σ on I . The construction of I^ε and σ_ε in the following lemma is a key to our approximation scheme. For us, $\mathbb{N} := \{0, 1, 2, \dots\}$ denotes the set of non-negative integers.

Lemma 2.4. *Suppose D is a bounded Lipschitz domain of \mathbb{R}^d . Let $I \subset \partial D$ be closed, connected and \mathcal{H}^{d-1} -rectifiable. Let $\varepsilon_j = 2^{-j}$ for $j \in \mathbb{N}$. Then there exist finite subsets $I^{(j)} = I^{\varepsilon_j}$ of I and functions $\sigma_{(j)} = \sigma_{\varepsilon_j} : I^{(j)} \rightarrow [\varepsilon_j^{d-1}/C, C\varepsilon_j^{d-1}]$ such that (a) and (b) below hold simultaneously:*

(a)

$$\sup_{x \in \overline{D}} \# \left(I^{(j)} \cap B(x, s) \right) \leq C \left(\frac{s}{\varepsilon_j} \vee 1 \right)^{d-1} \quad \forall s \in (0, \infty), j \in \mathbb{N}, \quad (2.1)$$

where $\#A$ denotes the number of elements in the finite set A , $B(x, s) = \{y \in \mathbb{R}^d : |y - x| < s\}$ is the ball with radius s centered at x , and C is a constant that depends only on D .

(b) For any equi-continuous and uniformly bounded family $\mathcal{F} \subset C(I)$,

$$\lim_{j \rightarrow \infty} \sup_{f \in \mathcal{F}} \left| \sum_{I^{(j)}} f \sigma_{(j)} - \int_I f d\sigma \right| = 0. \quad (2.2)$$

Proof We can always split I into small pieces. The point is to guarantee that each piece is not too small, so that $\sigma_{(j)}/\varepsilon_j^{d-1} \geq C$ and that (2.1) holds. Since I is \mathcal{H}^{d-1} -rectifiable, we have

$$C^{-1} R^{d-1} \leq \sup_{x \in I} \mathcal{H}^{d-1}(I \cap B(x, R)) \leq C R^{d-1}$$

for $R \in (0, 1]$, where C does not depend on R . Since I is closed, it is regular with dimension $d-1$ in the terminology of section 1 of [16]. Hence by [15] or section 2 of [16], we can build ‘‘dyadic cubes’’ for I . More precisely, there exists a family of partitions $\{\Delta_j\}_{j \in \mathbb{Z}}$ of I into ‘‘cubes’’ Q such that

- (i) if $j \leq k$, $Q \in \Delta_j$ and $Q' \in \Delta_k$, then either $Q \cap Q' = \emptyset$ or $Q \subset Q'$;
- (ii) if $Q \in \Delta_j$, then

$$\begin{aligned} C^{-1} 2^j &\leq \text{diam}(Q) \leq C 2^j \quad \text{and} \\ C^{-1} 2^{j(d-1)} &\leq \mathcal{H}^{d-1}(Q) \leq C 2^{j(d-1)}; \end{aligned}$$

(iii)

$$\mathcal{H}^{d-1}(\{x \in Q : \text{dist}(x, I \setminus Q) \leq r 2^j\}) \leq C r^{1/C} 2^{j(d-1)}$$

for all $Q \in \Delta_j$ and $r > 0$.

Here the constant C is independent of j , Q , or r . Note that \mathcal{H}^{d-1} is the surface measure σ of ∂D and that property (iii) tells us that the cubes have relatively small boundary. In particular, (iii) implies $\sigma(\partial Q \cap I) = 0$ for all cube Q .

Suppose $\Delta_j = \{U_i^{(j)}\}_{i=1}^{k_j}$. We pick one point $z_i^{(j)}$ from each $U_i^{(j)}$ to form the set $I^{(j)}$. Finally, we define $\sigma_{(j)}(z_i^{(j)}) := \sigma(U_i^{(j)})$. It follows from (ii) that $\sigma_{(j)} \in [\varepsilon_j^{d-1}/C, C\varepsilon_j^{d-1}]$ for some C which depends only on D . The inequality (2.1) follows from $C^{-1} \varepsilon_j^{d-1} \leq \sigma(U_i^{(j)})$ and the Lipschitz property of ∂D . It remains to check (2.2).

Fix any $\eta > 0$. There exists $\lambda = \lambda(\eta) > 0$ such that $|f(x) - f(y)| < \eta$ whenever $|x - y| < \lambda$. Hence for j large enough (depending only on λ),

$$\left| \int_I g d\sigma - \sum_{I^{(j)}} g \sigma_{(j)} \right| = \left| \sum_i \left(\int_{U_i^{(j)}} g d\sigma - g(z_i^{(j)}) \sigma(U_i^{(j)}) \right) \right| \leq \eta \sum_i \sigma(U_i^{(j)}) = \eta \sigma(I).$$

The desired convergence (2.2) now follows. \square

Remark 2.5. (2.2) implies that we have the weak convergence $\sum_{z \in I^{(j)}} \sigma_{(j)} \delta_z \rightarrow \sigma|_I$ on the space $M_+(I)$ of positive finite measure Borel measures on I . Here δ_z is the dirac delta measure at z , and $\sigma|_I$ is the surface measure restricted to I . (2.1) is a control on the number of points locally in I^{ε_j} . We call I^ε the ‘ ε -point approximation’ of I and σ_ε the ‘discrete surface measure’ associated to I^ε . \square

Remark 2.6. The above lemma remains true if I is the finite union of disjoint closed connected and \mathcal{H}^{d-1} -rectifiable subsets of ∂D . This enables us to deal with disconnected interface I . \square

2.1.2 Reflected diffusion and random walk approximation

We now describe the motion of each underlying particle. First we fix a bounded Lipschitz domain $D \subset \mathbb{R}^d$ and any $\varepsilon > 0$. Without loss of generality, we assume that the origin $0 \in D$. Let $\varepsilon\mathbb{Z}^d$ be the union of all closed line segments joining nearest neighbors in $\varepsilon\mathbb{Z}^d$, and $(D^\varepsilon)^*$ the connected component of $D \cap \varepsilon\mathbb{Z}^d$ that contains the point 0. Set $D^\varepsilon = (D^\varepsilon)^* \cap \varepsilon\mathbb{Z}^d$. We can view D^ε as the vertices of a graph whose edges coming from $(D^\varepsilon)^*$. We also denote the graph-boundary $\partial D^\varepsilon := \{x \in D^\varepsilon : v_\varepsilon(x) < 2d\}$, where $v_\varepsilon(x)$ is the degree of x in D^ε .

Suppose $\rho \in W^{1,2}(D) \cap C^1(\bar{D})$ is strictly positive. Define

$$\mathcal{E}(f, g) := \frac{1}{2} \int_D \nabla f(x) \cdot \nabla g(x) \rho(x) dx.$$

Since D is Lipschitz, $(W^{1,2}(D), \mathcal{E})$ is a regular Dirichlet form on $L^2(D; \rho)$ and so there is a ρ -symmetric diffusion X associated with it (cf. for example [12]).

Definition 2.7. We call X the $(I_{d \times d}, \rho)$ -**reflected diffusion**, where $I_{d \times d}$ is the $d \times d$ identity matrix. When $\rho = 1$, X is called the **reflected Brownian motion (RBM)** in D . Hence a $(I_{d \times d}, \rho)$ -reflected diffusion is a RBM in D with drift $\frac{1}{2} \nabla \log \rho$.

The L^2 -infinitesimal generator of X is

$$\mathcal{A} = \frac{1}{2\rho} \nabla \cdot (\rho \nabla) = \frac{1}{2} \Delta + \frac{1}{2} \nabla(\log \rho) \cdot \nabla.$$

Moreover, X has the Skorokhod representation:

$$dX_t = dB_t + \frac{1}{2} \nabla \log \rho(X_t) dt + \bar{n}(X_t) dL_t \quad \text{for } t \geq 0, \mathbb{P}_x\text{-a.s. } x \in \bar{D}, \quad (2.3)$$

where \bar{n} is the inward unit normal of ∂D and L is the positive continuous additive function (PCAF) of X whose Revuz measure is $\frac{1}{2} \sigma$ (c.f.[12]). We call L the boundary local time of X .

Next, we define X^ε to be a continuous time random walk (CTRW) on D^ε with exponential waiting time of parameter $\frac{d}{\varepsilon^2}$ and one step transition probabilities

$$p_{xy} := \frac{\mu_{xy}}{\sum_y \mu_{xy}},$$

where $\{\mu_{xy} : x, y \in D^\varepsilon\}$ are symmetric weights (conductances) to be constructed in two steps as follows: First, for every $x \in D^\varepsilon \setminus \partial D^\varepsilon$ and $i = 1, 2, \dots, d$, define

$$\mu_{x, x+\varepsilon \bar{e}_i} := \left(1 + \frac{1}{2} \ln \frac{\rho(x + \varepsilon \bar{e}_i)}{\rho(x)}\right) \left(\frac{\rho(x) + \rho(x + \varepsilon \bar{e}_i)}{2}\right) \frac{\varepsilon^{d-2}}{2} \quad (2.4)$$

$$\mu_{x, x-\varepsilon \bar{e}_i} := \left(1 + \frac{1}{2} \ln \frac{\rho(x)}{\rho(x - \varepsilon \bar{e}_i)}\right) \left(\frac{\rho(x) + \rho(x - \varepsilon \bar{e}_i)}{2}\right) \frac{\varepsilon^{d-2}}{2} \quad (2.5)$$

Clearly, $\mu_{xy} = \mu_{yx}$ for all $x, y \in D^\varepsilon \setminus \partial D^\varepsilon$. Note that since ρ is C^1 and strictly positive on \overline{D} , when ε is sufficiently small, $\mu_{x, x+\varepsilon e_i}$ and $\mu_{x, x-\varepsilon e_i}$ are strictly positive for every $x \in D^\varepsilon \setminus \partial D^\varepsilon$ and $i = 1, 2, \dots, d$. Second, we define

$$\mu_{xy} := \begin{cases} \mu_{yx}, & \text{if } x \in \partial D^\varepsilon, y \in D^\varepsilon \setminus \partial D^\varepsilon \\ \frac{\varepsilon^{d-2}}{2}, & \text{if } x, y \in \partial D^\varepsilon \text{ are adjacent in } D^\varepsilon \end{cases}$$

Now $\mu_{xy} = \mu_{yx}$ for all $x, y \in D^\varepsilon$. A heuristic reason of the above construction can be found in [23].

We call X^ε the ε -**approximation** of X . Clearly, X^ε is symmetric with respect to the measure m_ε defined by

$$m_\varepsilon(x) := \frac{\varepsilon^2}{d} \sum_y \mu_{xy}.$$

Since $\rho \in C^1(\overline{D})$, there exists a constant $C > 0$ such that

$$C^{-1} \leq \inf_x \frac{m_\varepsilon(x)}{\varepsilon^d} \leq \sup_x \frac{m_\varepsilon(x)}{\varepsilon^d} \leq C. \quad (2.6)$$

Moreover, $\lim_{\varepsilon \rightarrow 0} \frac{m_\varepsilon(x^\varepsilon)}{\varepsilon^d} = \rho(x)$ whenever $x^\varepsilon \in D^\varepsilon$ converges to $x \in D$.

A special but important case is when $\rho \equiv 1$. In this case, X is simply the reflected Brownian motion on D , and X^ε is a simple random walk on the graph D^ε . It is proved in [6] that X^ε converges weakly to the reflected Brownian motion X as $\varepsilon \rightarrow 0$.

Recall that by Assumption 2.3, we are given $\rho_\pm \in W^{(1,2)}(D_\pm) \cap C^1(\overline{D}_\pm)$. We denote by X^\pm a $(I_{d \times d}, \rho_\pm)$ -reflected diffusion in D_\pm , and by $X^{\varepsilon, \pm}$ the ε -approximation of X^\pm .

2.1.3 Random walks with interaction

Fix $\varepsilon = \varepsilon_j = 2^{-j}$ ($j \in \mathbb{N}$) and $N = 2^{jd}$ such that $N\varepsilon^d = 1$. Assume there are N “+” particles in D_+^ε and N “-” particles in D_-^ε at $t = 0$. Each particle moves as an independent CTRW $X^{\varepsilon, \pm}$ (see the previous subsection) in its respective domain D_\pm^ε . Let I^ε be the finite subset of I defined in Lemma 2.4. For each $z \in I^\varepsilon$, pick an $z_+ \in D_+^\varepsilon$ and an $z_- \in D_-^\varepsilon$ which are closest to z (See Figure 2). A pair of particles of opposite charges at (z_+, z_-) is being killed with a certain rate to be explained. Note that for ε small enough, we have $\sup_{z \in I^\varepsilon} |z_\pm - z| \leq 2M\varepsilon$, where M is the Lipschitz constant of I .

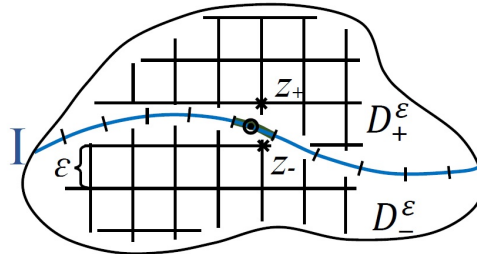


Figure 2: $z \in I^\varepsilon \subset I$, $z_\pm \in D_\pm^\varepsilon$

The state space of the particle system is the collection of configurations

$$E^\varepsilon := \{ \eta^\varepsilon = (\eta^{\varepsilon, +}, \eta^{\varepsilon, -}) : \eta^{\varepsilon, \pm} : D_\pm^\varepsilon \rightarrow \mathbb{N} \}. \quad (2.7)$$

The state of the particle system at time t is a random element $\eta_t^\varepsilon = (\eta_t^{\varepsilon, +}, \eta_t^{\varepsilon, -}) \in E^\varepsilon$. Here $\eta_t^{\varepsilon, \pm}(x)$ stands for the number of “ \pm ” particles at $x \in D_\pm^\varepsilon$ at time t . We omit ε and N for convenience when there is no ambiguity. For example, we write η_t and $m(x)$ in place of η_t^ε and $m_\varepsilon(x)$ respectively. The function ξ such that $\xi(x) = 1$ and $\xi(y) = 0$ for $y \neq x$ is denoted as $\mathbf{1}_x$.

Definition 2.8. η_t is defined to be the unique strong Markov process which has the generator $\mathfrak{L} = \mathfrak{L}^\varepsilon$ given by

$$\mathfrak{L}^\varepsilon := \mathfrak{L}_0^\varepsilon + \mathfrak{K}^\varepsilon, \quad (2.8)$$

where $\mathfrak{L}_0^\varepsilon$ is the generator of two families of independent random walks in D_+^ε and D_-^ε , respectively, with no annihilation between them, namely

$$\begin{aligned} \mathfrak{L}_0^\varepsilon f(\eta) &:= \frac{d}{\varepsilon^2} \sum_{x,y \in D_+^\varepsilon} \eta^+(x) p_{xy}^+ \{f(\eta^+ - \mathbf{1}_x + \mathbf{1}_y, \eta^-) - f(\eta)\} \\ &+ \frac{d}{\varepsilon^2} \sum_{x,y \in D_-^\varepsilon} \eta^-(x) p_{xy}^- \{f(\eta^+, \eta^- - \mathbf{1}_x + \mathbf{1}_y) - f(\eta)\} \end{aligned} \quad (2.9)$$

and \mathfrak{K}^ε is the operator corresponding to annihilation between particles of opposite signs at the interface I^ε , namely

$$\mathfrak{K}^\varepsilon f(\eta) := \frac{\lambda}{\varepsilon} \sum_{z \in I^\varepsilon} \Psi_\varepsilon(z) \eta^+(z_+) \eta^-(z_-) \{f(\eta^+ - \mathbf{1}_{z_+}, \eta^- - \mathbf{1}_{z_-}) - f(\eta)\}, \quad (2.10)$$

where p_{xy}^\pm is the one-step transition probabilities for the CTRW $X^{\varepsilon, \pm}$ on D_\pm^ε (without any interaction) and

$$\Psi_\varepsilon(z) := \frac{\sigma_\varepsilon(z)}{\varepsilon^{d-1}} \frac{\varepsilon^{2d}}{m(z_+)m(z_-)} \quad (2.11)$$

with σ_ε and I^ε being constructed by Lemma 2.4.

The expression for \mathfrak{K}^ε comes from the underlying assumptions of the model: First, the term $\eta^+(z_+) \eta^-(z_-)$ is combinatorial in nature. Since there are $\eta^+(z_+) \eta^-(z_-)$ pairs of particles at position (z_+, z_-) , the chance of killing is proportional to the number of ways of selecting a pair of particles near the interface. Second, each pair of particles near I disappears at rate $(\lambda/\varepsilon) \Psi_\varepsilon(z)$ where λ is a parameter. Intuitively, in the limit, the amount of annihilation in a neighborhood of a point is proportional to the surface area of the interface I in that neighborhood. The scaling $1/\varepsilon$ is suggested by the observation that there are about $1/\varepsilon$ "layers" starting from the interface I , so that the chance for a particle to arrive near I is of order ε . $\Psi_\varepsilon(z)$ is comparable to 1 and can be viewed as a normalizing constant with respect to the lattice. This choice (2.11) is justified in the proof of Theorem 3.5.

2.2 Discrete heat kernel

Throughout this subsection, D is a bounded Lipschitz domain in \mathbb{R}^d , $\rho \in W^{(1,2)}(D) \cap C^1(\overline{D})$ is strictly positive, X is a $(I_{d \times d}, \rho)$ -reflected diffusion. It is well known (cf. [1, 26] and the references therein) that X has a transition density $p(t, x, y)$ with respect to the symmetrizing measure $\rho(x)dx$ (i.e., $\mathbb{P}_x(X_t \in dy) = p(t, x, y) \rho(y) dy$ and $p(t, x, y) = p(t, y, x)$), that p is locally Hölder continuous and hence $p \in C((0, \infty) \times \overline{D} \times \overline{D})$, and that we have the followings: there are constants $c_1 \geq 1$ and $c_2 \geq 1$ such that

$$\frac{1}{c_1 t^{d/2}} \exp\left(\frac{-c_2 |y - x|^2}{t}\right) \leq p(t, x, y) \leq \frac{c_1}{t^{d/2}} \exp\left(\frac{-|y - x|^2}{c_2 t}\right) \quad (2.12)$$

for every $(t, x, y) \in (0, \infty) \times \overline{D} \times \overline{D}$. Using (2.12) and the Lipschitz assumption for the boundary, we can check that

$$\sup_{x \in \overline{D}} \sup_{0 < \delta \leq \delta_0} \frac{1}{\delta} \int_{D^\delta} p(t, x, y) dy \leq \frac{C_1}{t^{1/2}} + C_2 \quad \text{for } t > 0 \quad \text{and hence} \quad (2.13)$$

$$\sup_{x \in \overline{D}} \int_{\partial D} p(t, x, y) \sigma(dy) \leq \frac{C_1}{t^{1/2}} + C_2 \quad \text{for } t > 0, \quad (2.14)$$

where $C_1, C_2, \delta_0 > 0$ are constants which depends only on d, D and ρ .

On other hand, suppose $g \in \mathcal{B}_b([0, T] \times \partial D)$. Then for $t \in [0, T]$ and $x \in \bar{D}$,

$$\mathbb{E}^x \left[\int_0^t g(s, X_s) dL_s \right] = \frac{1}{2} \int_0^t \int_{\partial D} g(s, y) p(s, x, y) \sigma(dy) ds. \quad (2.15)$$

Now let X^ε be the ε -approximation of X with symmetrizing measure m_ε . The transition density p^ε of X^ε with respect to the measure m_ε is

$$p^\varepsilon(t, x, y) := \frac{\mathbb{P}^x(X_t^\varepsilon = y)}{m_\varepsilon(y)}, \quad t > 0, x, y \in D^\varepsilon. \quad (2.16)$$

Clearly, p^ε is strictly positive and is symmetric in x and y .

We will prove in Section 5 that p^ε enjoys two-sided Gaussian bound and is jointly Hölder continuous uniform in $\varepsilon \in (0, \varepsilon_0)$ for some $\varepsilon_0 > 0$, and that p^ε converges to p uniformly on compact subsets of $(0, \infty) \times \bar{D} \times \bar{D}$. In rigorous terms, we have the following three results. The important point is that the constants involved in these results do not depend on ε .

Theorem 2.9. (*Gaussian upper bound*) *There exist $C_k = C_k(d, D, \rho, T) > 0$, $k = 1, 2$, and $\varepsilon_0 = \varepsilon_0(d, D, \rho, T) \in (0, 1]$ such that for every $\varepsilon \in (0, \varepsilon_0)$ and $x, y \in D^\varepsilon$,*

$$p^\varepsilon(t, x, y) \leq \frac{C_1}{(\varepsilon \vee t^{1/2})^d} \exp\left(-C_2 \frac{|x-y|^2}{t}\right) \quad \text{for } t \in [\varepsilon, T], \quad (2.17)$$

and

$$p^\varepsilon(t, x, y) \leq \frac{C_1}{(\varepsilon \vee t^{1/2})^d} \exp\left(-C_2 \frac{|x-y|}{t^{1/2}}\right) \quad \text{for } t \in (0, \varepsilon). \quad (2.18)$$

Observe that (2.17) implies that (2.18) also holds for $t \in [\varepsilon, T]$.

Theorem 2.10. (*Gaussian lower bound*) *There exist $C_k = C_k(d, D, \rho, T) > 0$, $k = 1, 2$, and $\varepsilon_0 = \varepsilon_0(d, D, \rho, T) \in (0, 1]$ such that for every $\varepsilon \in (0, \varepsilon_0)$ and $x, y \in D^\varepsilon$,*

$$p^\varepsilon(t, x, y) \geq \frac{C_1}{(\varepsilon \vee t^{1/2})^d} \exp\left(-C_2 \frac{|x-y|^2}{t}\right) \quad \text{for } t \in (0, T]. \quad (2.19)$$

Theorem 2.11. (*Hölder continuity*) *There exist positive constants $\gamma = \gamma(d, D, \rho)$, $\varepsilon_0(d, D, \rho)$ and $C(d, D, \rho)$ such that for all $\varepsilon \in (0, \varepsilon_0)$, we have*

$$|p^\varepsilon(t, x, y) - p^\varepsilon(t', x', y')| \leq C \frac{(|t-t'|^{1/2} + |x-x'| + |y-y'|)^\gamma}{(t \wedge t')^{\sigma/2} (1 \wedge t \wedge t')^{d/2}}. \quad (2.20)$$

Theorem 2.12. (*Local CLT*)

$$\lim_{n \rightarrow \infty} \sup_{t \in [a, b]} \sup_{x, y \in \bar{D}} |p^{(2^{-n})}(t, x, y) - p(t, x, y)| = 0$$

for any compact interval $[a, b] \subset (0, \infty)$.

To establish the tightness of $\{\mathfrak{X}^{N,+}, \mathfrak{X}^{N,-}\}$, we need the following uniform estimate for the heat kernel $p^\varepsilon(t, x, y)$ of CTRW on D^ε near the boundary of D^ε . It is the discrete analog of (2.14).

Lemma 2.13. *There exist $C = C(d, D, \rho, T) > 0$ and $\varepsilon_0 = \varepsilon_0(d, D, \rho) > 0$ such that*

$$\sup_{x \in D^\varepsilon} \varepsilon^{d-1} \sum_{y \in \partial D^\varepsilon} p^\varepsilon(t, x, y) \leq \frac{C}{\varepsilon \vee t^{1/2}} \quad (2.21)$$

for all $t \in [0, T]$ and $\varepsilon \in (0, \varepsilon_0)$. Here ∂D^ε is the graph-boundary of D^ε , which is all the vertices in D^ε with degree less than $2d$.

Proof Fix $\theta \in [0, T]$. By the Gaussian upper bound in Theorem 2.9, we have

$$\begin{aligned}
& \sum_{y \in \partial D^\varepsilon} p^\varepsilon(\theta, x, y) \\
& \leq \frac{C_1}{(\varepsilon \vee \theta^{1/2})^d} \sum_{y \in \partial D^\varepsilon} \exp\left(\frac{-|y-x|}{\varepsilon \vee \theta^{1/2}}\right) \\
& = \frac{C_1}{(\varepsilon \vee \theta^{1/2})^d} \int_0^\infty |\{y \in D^\varepsilon : |f(y)| > r\}| dr \quad \text{by setting } f(y) = \mathbf{1}_{\partial D^\varepsilon}(y) \exp\left(\frac{-|y-x|}{\varepsilon \vee \theta^{1/2}}\right) \\
& = \frac{C_1}{(\varepsilon \vee \theta^{1/2})^d} \int_0^1 |\{\partial D^\varepsilon \cap B(x, (\varepsilon \vee \theta^{1/2})(-\ln r))\}| dr \quad (\text{since } f \leq 1) \\
& = \frac{C_1}{(\varepsilon \vee \theta^{1/2})^{d+1}} \int_0^\infty |\{\partial D^\varepsilon \cap B(x, s)\}| \exp\left(\frac{-s}{\varepsilon \vee \theta^{1/2}}\right) ds \quad (\text{where } s = (\varepsilon \vee \theta^{1/2})(-\ln r)) \\
& \leq \frac{C_1}{(\varepsilon \vee \theta^{1/2})^d} \vee \frac{C_2}{\varepsilon^{d-1}(\varepsilon \vee \theta^{1/2})^{d+1}} \int_0^\infty s^{d-1} \exp\left(\frac{-s}{\varepsilon \vee \theta^{1/2}}\right) ds \\
& \leq \frac{1}{\varepsilon^{d-1}} \left(\frac{C_1}{\varepsilon \vee \theta^{1/2}} \vee \frac{C_2}{\varepsilon \vee \theta^{1/2}} \int_0^\infty w^{d-1} e^{-w} dw \right) \quad (\text{where } w = \frac{s}{\varepsilon \vee \theta^{1/2}}).
\end{aligned}$$

Here C_i are all constants which depend only on d, D and T . Note that in the second last line, we used the fact that $|\{\partial D^\varepsilon \cap B(x, s)\}| \leq C((s/\varepsilon)^{d-1} \vee 1)$, which follows from Lemma 2.4. The proof is now complete. \square

In general, we use " \pm " for quantities related to X^\pm . For example, \mathcal{A}^\pm , $(P_t^\pm)_{t \geq 0}$ and $p^\pm(t, x, y)$ denote the generator, semigroup and transition density of the reflected diffusion X^\pm in D_\pm with respect to $\rho_\pm(x)dx$. In addition, we use " ε " for quantities related to the CTRWs in the discrete domains D_\pm^ε . For example, $p^{\varepsilon, \pm}(t, x, y)$ denotes the transition density of the CTRW $X^{\varepsilon, \pm}$ on D_\pm^ε with respect to the measure m_\pm^ε . We also denote $p^\varepsilon(t, (\vec{r}, \vec{s}), (\vec{r}', \vec{s}')) := \prod_{i=1}^n p^\varepsilon(t, r_i, r'_i) \prod_{j=1}^m p^\varepsilon(t, s_j, s'_j)$ for $(\vec{r}, \vec{s}) \in (D_+^\varepsilon)^n \times (D_-^\varepsilon)^m$.

By applying Lemma 2.13 to $p^{\varepsilon, \pm}(t, x, y)$, then by the boundedness of Φ_ε in (2.11), Theorem 2.12, Lemma 2.4 and LDCT, we have the following approximation for the local time of X^\pm on I .

Proposition 2.14. (*Discrete local time*)

$$\lim_{\varepsilon \rightarrow 0} \int_0^t \varepsilon^{d-1} \sum_{z \in I^\varepsilon} p^{\varepsilon, \pm}(\theta, x, z_\pm) \sigma_\varepsilon(z) d\theta = \int_0^t \int_I p^\pm(\theta, x, z) \sigma(dz) d\theta. \quad (2.22)$$

2.3 Hydrodynamic limit: system of nonlinear PDEs

In this subsection, we provide suitable notion of solutions for the coupled PDE (1.1) and (1.2), and then prove the existence and uniqueness of the solution. Throughout this subsection, D is a bounded Lipschitz domain, $\rho \in W^{(1,2)}(D) \cap C^1(\bar{D})$ is strictly positive, X is a $(I_{d \times d}, \rho)$ -reflected diffusion, $\{P_t\}$ and $p(t, x, y)$ are the semigroup and the transition density of X , with respect to the measure $\rho(x)dx$.

Observe that (1.1) is a second order parabolic equation for u_+ with Robin boundary condition, and (1.2) is a similar equation for u_- . This leads us to consider the following Robin boundary problem, where $g \in \mathcal{B}_b([0, \infty) \times \partial D)$.

$$\begin{cases} \frac{\partial u}{\partial t} = \frac{1}{2} \Delta u + \frac{1}{2} \nabla(\log \rho) \cdot \nabla u & \text{for } x \in D, t > 0, \\ \frac{\partial u}{\partial \vec{n}} = \frac{1}{\rho} g u & \text{for } x \in \partial D, t > 0, \\ u(0, x) = \varphi(x) & \text{for } x \in D. \end{cases} \quad (2.23)$$

By Itô's formula and the Skorokhod representation for X , we see that a classical solution of (2.23),

should it exists, has the probabilistic representation

$$u(t, x) := \mathbb{E}^x \left[\varphi(X_t) e^{-\int_0^t g(t-s, X_s) dL_s} \right]. \quad (2.24)$$

where L is the boundary local time of X .

Definition 2.15. u defined by (2.24) is called a probabilistic solution of (2.23).

First we show that the function u defined by (2.24) is continuous.

Lemma 2.16. *Suppose $\varphi \in \mathcal{B}_b(\overline{D})$, $g \in \mathcal{B}_b^+([0, T] \times \partial D)$ and u is defined by (2.24). Then $u \in C((0, T] \times \overline{D})$. Moreover, if $\varphi \in C(\overline{D})$, then $u \in C([0, T] \times \overline{D})$.*

Proof Observe that for any $r \in [0, t]$,

$$\begin{aligned} u(t, x) &= \mathbb{E}^x \left[\varphi(X_t) e^{-\int_r^t g(t-s, X_s) dL_s} e^{-\int_0^r g(t-s, X_s) dL_s} \right] \\ &= \mathbb{E}^x \left[\varphi(X_t) e^{-\int_r^t g(t-s, X_s) dL_s} \right] + \mathbb{E}^x \left[\varphi(X_t) e^{-\int_r^t g(t-s, X_s) dL_s} \left(e^{-\int_0^r g(t-s, X_s) dL_s} - 1 \right) \right]. \end{aligned} \quad (2.25)$$

By Markov property, the first term is

$$\mathbb{E}^x \left[\mathbb{E}^{X_r} [\varphi(X_{t-r}) e^{-\int_0^{t-r} g(t-r-s, X_s) dL_s}] \right] = \mathbb{E}^x [u(t-r, X_r)].$$

Since X has the strong Feller property (see [1]) and u is bounded, $x \mapsto \mathbb{E}^x [u(t-r, X_r)]$ is continuous on \overline{D} for any fixed $t > 0$ and $r \in (0, t)$.

The second term of (2.25) converges to 0 uniformly on $(0, T] \times \overline{D}$, as $r \rightarrow 0$. This is because its absolute value is bounded by

$$\begin{aligned} & \|\varphi\|_\infty \mathbb{E}^x \left[1 - e^{-\int_0^r g(t-s, X_s) dL_s} \right] \\ & \leq \|\varphi\|_\infty \mathbb{E}^x \left[\int_0^r g(t-s, X_s) dL_s \right] \quad \text{by mean-value theorem} \\ & \leq \|\varphi\|_\infty \|g\|_\infty \frac{1}{2} \int_0^r \int_{\partial D} p(s, x, y) \sigma(dy) ds \\ & \leq \|\varphi\|_\infty \|g\|_\infty (2C_1 \sqrt{r} + C_2 r). \end{aligned}$$

Hence, u is continuous in $x \in \overline{D}$.

By a similar calculation as in (2.25), we have

$$\begin{aligned} u(t+a, x) - u(t, x) &= \mathbb{E}^x [u(t, X_a) - u(t, x)] \\ & \quad + \mathbb{E}^x \left[\varphi(X_{t+a}) e^{-\int_a^{t+a} g(t+a-s, X_s) dL_s} \left(e^{-\int_0^a g(t+a-s, X_s) dL_s} - 1 \right) \right]. \end{aligned}$$

Hence,

$$\begin{aligned} & |u(t+a, x) - u(t, x)| \\ & \leq \mathbb{E}^x [|u(t, X_a) - u(t, x)|] + \|\varphi\|_\infty \mathbb{E}^x \left[\int_0^a g(t+a-s, X_s) dL_s \right] \quad \text{by mean-value theorem} \\ & \leq \int_D |u(t, z) - u(t, x)| p(a, x, z) dz + \|\varphi\|_\infty \|g\|_\infty \frac{1}{2} \int_0^a \int_{\partial D} p(s, x, z) \sigma(dz) ds. \end{aligned}$$

Both terms go to 0 uniformly in $x \in \overline{D}$ as a goes to 0. (In fact, the first term goes to 0 uniformly since the semigroup P_t is strongly continuous on $C(\overline{D})$. For the second term, $\int_0^a \int_{\partial D} p(s, x, z) \sigma(dz) ds \leq 2C_1 \sqrt{a} + C_2 a$ also goes to 0 uniformly in x .) Hence u is continuous in $t \in (0, T]$ uniformly in $x \in \overline{D}$. Therefore, $u \in C((0, T] \times \overline{D})$. If $\varphi \in C(\overline{D})$, we can extend the above argument to show that $u \in C([0, T] \times \overline{D})$. \square

Remark 2.17. In fact, one can allow g to be unbounded and show that the conclusion of Lemma 2.16 remains true if $g\sigma$ satisfies a Kato class condition:

$$\limsup_{a \rightarrow 0} \sup_{x \in \overline{D}} \int_0^a \int_{\partial D} p(s, x, z) |g(t + a - s, z)| \sigma(dz) ds = 0.$$

□

Proposition 2.18. Suppose $\varphi \in C(\overline{D})$ and $g \in \mathcal{B}_0^+([0, T] \times \partial D)$. Then

$$u(t, x) := \mathbb{E}^x \left[\varphi(X_t) e^{-\int_0^t g(t-s, X_s) dL_s} \right]$$

is the unique element in $C([0, T] \times \overline{D})$ that satisfies the following integral equation:

$$u(t, x) = P_t \varphi(x) - \frac{1}{2} \int_0^t \int_{\partial D} p(t-r, x, y) g(r, y) u(r, y) \sigma(dy) dr. \quad (2.26)$$

Proof By Lemma 2.16, $u(t, x) := \mathbb{E}^x[\varphi(X_t) e^{-\int_0^t g(t-s, X_s) dL_s}]$ lies in $C([0, T] \times \overline{D})$. Moreover, by Markov property and (2.15) we have

$$\begin{aligned} u(t, x) &= \mathbb{E}^x[\varphi(X_t)] - \mathbb{E}^x[\varphi(X_t) (1 - e^{-\int_0^t g(t-s, X_s) dL_s})] \\ &= P_t \varphi(x) - \mathbb{E}^x \left[\varphi(X_t) e^{-\int_r^t g(t-s, X_s) dL_s} \Big|_{r=0}^{r=t} \right] \\ &= P_t \varphi(x) - \mathbb{E}^x \left[\varphi(X_t) \int_0^t g(t-r, X_r) e^{-\int_r^t g(t-s, X_s) dL_s} dL_r \right] \\ &= P_t \varphi(x) - \mathbb{E}^x \left[\int_0^t g(t-r, X_r) \mathbb{E}^{X_r} \left[\varphi(X_{t-r}) e^{-\int_0^{t-r} g(t-r-s, X_s) dL_s} \right] dL_r \right] \\ &= P_t \varphi(x) - \mathbb{E}^x \left[\int_0^t g(t-r, X_r) u(t-r, X_r) dL_r \right] \\ &= P_t \varphi(x) - \frac{1}{2} \int_0^t \int_{\partial D} p(r, x, y) g(t-r, y) u(t-r, y) \sigma(dy) dr. \end{aligned}$$

Hence u satisfies the integral equation. It remains to prove uniqueness. Suppose $\tilde{u} \in C([0, T] \times \overline{D})$ also satisfies the integral equation. Then $w = u - \tilde{u} \in C([0, T] \times \overline{D})$ solves

$$w(t, x) = -\frac{1}{2} \int_0^t \int_{\partial D} p(t-r, x, y) g(r, y) w(r, y) \sigma(dy) dr. \quad (2.27)$$

By a Gronwall type argument and (2.14), we can show that $w = 0$. More precisely, let $\psi(s) = \sup_{x \in \overline{D}} |w(s, x)|$. Then

$$0 \leq \psi(t) \leq \int_0^t \psi(r) \left(\frac{A}{\sqrt{t-r}} + B \right) dr \quad \forall t \geq 0.$$

Note that

$$\int_0^t \psi(r) \left(\frac{A}{\sqrt{t-r}} + B \right) dr = \frac{\partial}{\partial t} \int_0^t \psi(r) (2A\sqrt{t-r} + Bt) dr - \psi(t)Bt.$$

Combining the above two inequalities, we have

$$(1 + Bt)\psi(t) \leq \frac{\partial}{\partial t} \int_0^t \psi(r) (2A\sqrt{t-r} + Bt) dr.$$

Integrating both sides with respect to t on the interval $[0, t_0]$, we have

$$0 \leq \int_0^{t_0} (1 + Bt)\psi(t) dt \leq \int_0^{t_0} \psi(r) (2A\sqrt{t_0-r} + Bt_0) dr.$$

From this we have $\psi = 0$ on $[0, t_0]$, where $t_0 > 0$ is small enough so that $2A\sqrt{t_0} + Bt_0 < 1$. Let $\tilde{\psi}(t) = \psi(t + t_0)$, we can show that

$$0 \leq \tilde{\psi}(t) \leq \int_0^t \tilde{\psi}(r) \left(\frac{A}{\sqrt{t-r}} + B \right) dr \quad \text{for all } t \geq 0$$

and repeat the argument to obtain $\tilde{\psi} = 0$ on $[0, t_0]$ (i.e. $\psi = 0$ on $[0, 2t_0]$). Inductively, we obtain $\psi = 0$ on $[0, T]$. \square

Now we come to our coupled equation.

Proposition 2.19. *For $T > 0$, consider the Banach space $\Lambda_T = C([0, T] \times \overline{D}_+) \times C([0, T] \times \overline{D}_-)$ with norm $\|(u, v)\| := \|u\|_\infty + \|v\|_\infty$. Suppose $u_+(0) = f \in C(\overline{D}_+)$ and $u_-(0) = g \in C(\overline{D}_-)$. Then there is a unique element $(u_+, u_-) \in \Lambda_T$ which satisfies the coupled integral equation*

$$\begin{cases} u_+(t, x) = P_t^+ f(x) - \frac{\lambda}{2} \int_0^t \int_I p^+(t-r, x, z) [u_+(r, z) u_-(r, z)] d\sigma(z) dr \\ u_-(t, y) = P_t^- g(y) - \frac{\lambda}{2} \int_0^t \int_I p^-(t-r, y, z) [u_+(r, z) u_-(r, z)] d\sigma(z) dr. \end{cases} \quad (2.28)$$

Moreover, (u_+, u_-) satisfies

$$\begin{cases} u_+(t, x) = \mathbb{E}^x [f(X_t^+) e^{-\lambda \int_0^t u_-(t-s, X_s^+) dL_s^+}] \\ u_-(t, y) = \mathbb{E}^y [g(X_t^-) e^{-\lambda \int_0^t u_+(t-s, X_s^-) dL_s^-}], \end{cases} \quad (2.29)$$

where L^\pm is the boundary local time of X^\pm on the interface I .

Functions (u_+, u_-) satisfying equation (2.28) will be called a weak solution of (1.1) and (1.2), as it can be shown that they are weakly differentiable and solve the equations in the distributional sense.

Proof of Proposition 2.19. Define the operator S on Λ_T by $S(u, v) = (S^+v, S^-u)$, where

$$\begin{aligned} S^+v(t, x) &= \mathbb{E}^x \left[f(X_t^+) e^{-\lambda \int_0^t v(t-s, X_s^+) dL_s^+} \right] \quad \text{for } (t, x) \in [0, T] \times \overline{D}_+, \\ S^-u(t, y) &= \mathbb{E}^y \left[g(X_t^-) e^{-\lambda \int_0^t u(t-s, X_s^-) dL_s^-} \right] \quad \text{for } (t, y) \in [0, T] \times \overline{D}_-. \end{aligned}$$

Lemma 2.16 implies that S maps into Λ_T . Moreover, for $(t, x) \in [0, T] \times \overline{D}_+$,

$$\begin{aligned} |(S^+v_1 - S^+v_2)(t, x)| &= \left| \mathbb{E}^x \left[f(X_t^+) \left(e^{-\lambda \int_0^t v_1(t-s, X_s^+) dL_s^+} - e^{-\lambda \int_0^t v_2(t-s, X_s^+) dL_s^+} \right) \right] \right| \\ &\leq \|f\|_\infty \mathbb{E}^x \left[\left| \lambda \int_0^t v_1(t-s, X_s^+) dL_s^+ - \lambda \int_0^t v_2(t-s, X_s^+) dL_s^+ \right| \right] \\ &= \|f\|_\infty \lambda \mathbb{E}^x \left[\int_0^t |v_1(t-s, X_s^+) - v_2(t-s, X_s^+)| dL_s^+ \right] \\ &\leq \lambda \|f\|_\infty \|v_1 - v_2\|_\infty \mathbb{E}^x [L_t^+] \\ &= \lambda \|f\|_\infty \|v_1 - v_2\|_\infty \frac{1}{2} \int_0^t \int_I p^+(s, x, y) \sigma(dy) ds \\ &\leq C_1 \lambda \sqrt{T} \|f\|_\infty \|v_1 - v_2\|_\infty. \end{aligned}$$

A similar result holds for $S^-u_1 - S^-u_2$. Hence,

$$\begin{aligned} \|S(u_1, v_1) - S(u_2, v_2)\|_\infty &= \|S^+v_1 - S^+v_2\|_\infty + \|S^-u_1 - S^-u_2\|_\infty \\ &\leq C_1 \lambda \sqrt{T} \|u_0\|_\infty \|v_1 - v_2\|_\infty + C_2 \lambda \sqrt{T} \|v_0\|_\infty \|u_1 - u_2\|_\infty \\ &\leq \gamma \|(u_1, v_1) - (u_2, v_2)\| \end{aligned}$$

for some $\gamma < 1$ when T is small enough.

Hence there is a $T_0 > 0$ such that $S : \Lambda_{T_0} \rightarrow \Lambda_{T_0}$ is a contraction map. By Banach fixed point theorem, there is a unique element $(u^*, v^*) \in \Lambda_{T_0}$ such that $(u^*, v^*) = S(u^*, v^*)$. By Proposition 2.18, (u^*, v^*) is the unique weak solution to the coupled PDE on $[0, T_0]$.

Repeat the above argument, with $u_0(\cdot)$ replaced by $u^*(T_0, \cdot)$, and $v_0(\cdot)$ replaced by $v^*(T_0, \cdot)$. We see that, since $\|u^*(T_0, \cdot)\|_\infty \leq \|u_0\|_\infty$, $\|v^*(T_0, \cdot)\|_\infty \leq \|v_0\|_\infty$ and C_i ($i = 1, 2$) are the same, we can extend the solution of the coupled PDE uniquely to $[T_0, 2T_0]$. Iterating the argument, we have for any $T > 0$, the coupled PDE has a unique weak solution in Λ_T . Invoke Proposition 2.18 once more, we obtain the desired implicit probabilistic representation (2.29).

Finally, by using Markov property as in the proof of Proposition 2.18, we see that (2.29) and (2.28) are equivalent. \square

2.4 Main results (rigorous statements)

In this paper, we always assume the scaling $N\varepsilon^d = 1$ holds, so that the interacting random walk model is parameterized by a single parameter N which is the initial number of particles in each of D_+^ε and D_-^ε . More precisely, for each fixed N , we set $\varepsilon = N^{-1/d}$ and let $(\eta_t^\varepsilon)_{t \geq 0}$ be a Markov process having generator \mathfrak{L}^ε defined in (2.8) and having initial distribution satisfies $\sum_{x \in D_+^\varepsilon} \eta_0^{\varepsilon,+}(x) = \sum_{y \in D_-^\varepsilon} \eta_0^{\varepsilon,-}(y) = N$. We define the empirical measures

$$\mathfrak{X}_t^{N,\pm}(dz) := \frac{1}{N} \sum_{x \in D_\pm^\varepsilon} \eta_t^{\varepsilon,\pm}(x) \mathbf{1}_x(dz).$$

It is clear that $(\mathfrak{X}_t^{N,+}, \mathfrak{X}_t^{N,-})_{t \geq 0}$ is a continuous time Markov process (inheriting from that of η_t) with state space

$$\mathfrak{E} := M_{\leq 1}(\overline{D}_+) \times M_{\leq 1}(\overline{D}_-),$$

where $M_{\leq 1}(E)$ denotes the space of non-negative Borel measures on E with mass at most 1. $M_{\leq 1}(E)$ is a closed subset of $M_+(E)$, where the latter denotes the space of finite non-negative Borel measures on E equipped with the following metric:

$$\|\mu - \nu\| := \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{|\langle \mu, \phi_k \rangle - \langle \nu, \phi_k \rangle|}{1 + |\langle \mu, \phi_k \rangle - \langle \nu, \phi_k \rangle|}, \quad (2.30)$$

where $\{\phi_k : k \geq 1\}$ is any countable dense subset of $C(E)$. The topology induced by this metric is equivalent to the weak topology (i.e. $\|\mu_n - \mu\| \rightarrow 0$ if and only if $\langle \mu_n, f \rangle \rightarrow \langle \mu, f \rangle$ for all $f \in C(E)$). Under this metric, $M_+(\overline{D})$ is a complete separable metric space, hence so are \mathfrak{E} and the Skorokhod space $D([0, T], \mathfrak{E})$ (see e.g. Theorem 3.5.6 of [22]). Here is our first main result.

Theorem 2.20. (*Hydrodynamic Limit*) *Suppose Assumption 2.3 holds and the sequence of initial configurations η_0^ε satisfies the following conditions:*

- (i) $\mathfrak{X}_0^{N,\pm} \xrightarrow{\mathcal{L}} u_0^\pm(z) dz$ in $M_{\leq 1}(\overline{D}_\pm)$, where $u_0^\pm \in C(\overline{D}_\pm)$. (Note that $\mathfrak{X}_0^{N,\pm}$ has unit mass for all N .)
- (ii) $\overline{\lim}_{N \rightarrow \infty} \sup_{z \in D_\pm^\varepsilon} \mathbb{E} \left[(\eta_0^{\varepsilon,\pm}(z))^2 \right] < \infty$.

Then for any $T > 0$, as $\varepsilon \rightarrow 0$ along the sequence $\varepsilon_j = 2^{-j}$, we have

$$(\mathfrak{X}^{N,+}, \mathfrak{X}^{N,-}) \xrightarrow{\mathcal{L}} (\nu^+, \nu^-) \in D([0, T], \mathfrak{E}),$$

where (ν^+, ν^-) is the deterministic element in $C([0, T], \mathfrak{E})$ such that

$$(\nu_t^+(dx), \nu_t^-(dy)) = (u_+(t, x) \rho_+(x) dx, u_-(t, y) \rho_-(y) dy)$$

for all $t \in [0, T]$, and (u_+, u_-) is the unique weak solution of the coupled PDEs (1.1) and (1.2) with initial value (u_0^+, u_0^-) .

Theorem 2.20 gives the limiting probability distribution of one particle randomly picked in D_{\pm}^{ε} at time t . This is the 1-particle distribution in the terminology of statistical physics.

Question: What is the limiting joint distribution of more than one particles?

Before stating the answer, we need to introduce a standard tool in the study of stochastic particle systems: the notion of correlation functions¹. Recall that the state space of $\eta^{\varepsilon} = (\eta_t^{\varepsilon})_{t \geq 0}$ is E^{ε} defined in (2.7). We denote by

$$\Omega_{n,m}^{\varepsilon} := \left\{ \xi = (\xi^+, \xi^-) \in E^{\varepsilon} : |\xi^+| := \sum_x \xi^+(x) = n, |\xi^-| := \sum_y \xi^-(y) = m \right\}$$

the set of configurations with n and m particles in D_+^{ε} and D_-^{ε} respectively. We then define $A : E^{\varepsilon} \times E^{\varepsilon} \rightarrow \mathbb{R}$ in such a way that whenever $\xi \in \Omega_{n,m}^{\varepsilon}$,

$$A(\xi, \eta) := A^+(\xi^+, \eta^+) A^-(\xi^-, \eta^-) := \prod_{x \in D_+} A_{\xi^+(x)}^{\eta^+(x)} \prod_{x \in D_-} A_{\xi^-(x)}^{\eta^-(x)}, \quad (2.31)$$

where $n \mapsto A_k^n$ is the Poisson polynomial² of order k , namely $A_0^n := 1$ and $A_k^n := n(n-1) \cdots (n-k+1)$ for $k \geq 1$ (in particular, $A_k^n = 0$ for $k > n$). Note that A_k^n is the number of permutations of k objects chosen from n distinct objects. So $A(\xi, \eta)$ is the total number of possible site to site pairings between labeled particles having configuration ξ with a subset of labeled particles having configuration η . An alternative representation of (2.31) will be given in (3.5).

Convention: For $(\vec{r}, \vec{s}) \in (D_+^{\varepsilon})^n \times (D_-^{\varepsilon})^m$ and $\eta \in \Omega_{N,M}^{\varepsilon}$, we define $A((\vec{r}, \vec{s}), \eta)$ to be $A(\xi, \eta)$ with $\xi = (\sum_i \delta_{r_i}, \sum_j \delta_{s_j})$.

Definition 2.21. Let \mathbb{P}^{η} is the law of a process with generator $\mathcal{L}^{\varepsilon}$ and initial distribution η satisfying

$$\sum_{x \in D_+^{\varepsilon}} \eta^+(x) = \sum_{y \in D_-^{\varepsilon}} \eta^-(y) = \varepsilon^{-d}. \quad (2.32)$$

For all $t \geq 0$, we define

$$\gamma^{\varepsilon}(\xi, t) := \gamma^{\varepsilon, (n,m)}(\xi, t) := \frac{\varepsilon^{d(n+m)}}{\alpha_{\varepsilon}(\xi)} \mathbb{E}^{\eta}[A(\xi, \eta_t)] \quad (2.33)$$

for all $\xi \in \Omega_{n,m}^{\varepsilon}$, where

$$\alpha_{\varepsilon}(\xi) := m_{\varepsilon}(\vec{r}, \vec{s}) := \prod_{i=1}^n m_{\varepsilon}^+(r_i) \prod_{j=1}^m m_{\varepsilon}^-(s_j). \quad (2.34)$$

when $\xi = (\sum_i \delta_{r_i}, \sum_j \delta_{s_j})$. By convention, we also have $\gamma^{\varepsilon}((\vec{r}, \vec{s}), t) := \gamma^{\varepsilon}(\xi, t)$. Note that γ^{ε} depends on the initial configuration of η .

Intuitively, suppose we randomly pick n and m living particles in D_+ and D_- respectively at time t , then $(\vec{r}, \vec{s}) \mapsto \gamma^{\varepsilon, (n,m)}((\vec{r}, \vec{s}), t)$ is the joint probability density function for their positions, up to a normalizing constant. Therefore, it is natural that $\gamma^{\varepsilon, (n,m)}$ defined by (2.33) is called the (n, m) -**particle correlation function**.

The next is our second main result, **propagation of chaos**, for our system. It says that when the number of particles tends to infinity, they appears to be independent of each other. Mathematically, the correlation function factors out in the limit $N \rightarrow \infty$.

Theorem 2.22. (Propagation of Chaos) Under the same condition as in Theorem 2.20, for all $n, m \in \mathbb{N}$ and any compact interval $[a, b] \subset (0, \infty)$,

$$\lim_{\varepsilon \rightarrow 0} \sup_{(\vec{r}, \vec{s}) \in \bar{D}_+^n \times \bar{D}_-^m} \sup_{t \in [a, b]} \left| \gamma^{\varepsilon}((\vec{r}, \vec{s}), t) - \prod_{i=1}^n u_+(t, r_i) \prod_{j=1}^m u_-(t, s_j) \right| = 0,$$

¹More precisely, we will be using correlation functions for *unlabeled* particles. We refer the readers to [33] for the relation between labeled and unlabeled correlation functions.

²The notation A_k^n is suggested by the fact that $\mathbb{E}[A_k^{\varrho}] = \theta^k$ when ϱ is a Poisson random variable with mean θ .

where (u_+, u_-) is the weak solution of the coupled PDE.

To investigate the intensity of killing near the interface, we define $J_t^{N,\pm} \in D([0, \infty), M_+(\overline{D}_\pm))$ by

$$J_t^{N,+}(A) := \varepsilon^{d-1} \sum_{z \in I^\varepsilon} \Psi(z) \eta_t^+(z_+) \eta_t^-(z_-) \mathbf{1}_A(z_+) \quad \text{for } A \subset \overline{D}_+, \quad (2.35)$$

$$J_t^{N,-}(B) := \varepsilon^{d-1} \sum_{z \in I^\varepsilon} \Psi(z) \eta_t^+(z_+) \eta_t^-(z_-) \mathbf{1}_B(z_-) \quad \text{for } B \subset \overline{D}_-. \quad (2.36)$$

Clearly, $\langle J_t^{N,+}, 1 \rangle = \langle J_t^{N,-}, 1 \rangle$, which measures the number of encounters of the two types of particles near I . An immediately corollary of Theorem 2.22 is the following, which is what we need to identify the limit of $(\mathfrak{X}^{N,+}, \mathfrak{X}^{N,-})$.

Corollary 2.23. *For any fixed $t \in (0, \infty)$ and $\phi \in C(\overline{D}_\pm)$,*

$$\begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E}[\langle J_t^{N,\pm}, \phi \rangle] &= \frac{1}{2} \int_I u_+(t, y) u_-(t, y) \phi(y) \sigma(dy), \\ \lim_{N \rightarrow \infty} \mathbb{E}[(\langle J_t^{N,\pm}, \phi \rangle)^2] &= \left(\frac{1}{2} \int_I u_+(t, y) u_-(t, y) \phi(y) \sigma(dy) \right)^2, \\ \lim_{N \rightarrow \infty} \mathbb{E}[\langle \mathfrak{X}_t^{\pm, N}, \phi \rangle] &= \int_{D_\pm} u_\pm(t, y) \phi(y) \rho_\pm(y) dy, \\ \lim_{N \rightarrow \infty} \mathbb{E}[(\langle \mathfrak{X}_t^{\pm, N}, \phi \rangle)^2] &= \left(\int_{D_\pm} u_\pm(t, y) \phi(y) \rho_\pm(y) dy \right)^2. \end{aligned}$$

Proof We only need to apply Theorem 2.22 for the cases $(n, m) = (1, 1)$ and $(n, m) = (1, 0)$. By definition,

$$\gamma^\varepsilon(\mathbf{1}_r, t) = \frac{\varepsilon^d}{m^+(r)} \mathbb{E}^\eta[\eta_t^+(r)] \quad \text{and} \quad \gamma^\varepsilon(\mathbf{1}_r + \mathbf{1}_s, t) = \frac{\varepsilon^{2d}}{m^+(r)m^-(s)} \mathbb{E}^\eta[\eta_t^+(r)\eta_t^-(s)].$$

Using (2.6) and Lemma 2.4, we get the first two equations via Theorem 2.22. Using (2.6) and the assumption that $\rho_\pm \in C(\overline{D}_\pm)$, we have the last two equations again by Theorem 2.22. \square

Remark 2.24. (Conditions on η_0) The two conditions for the initial configuration η_0 in Theorem 2.20 are mild and natural. They are satisfied, for example, when each particle has the same random initial distribution $\frac{u_0^\pm(z)}{\sum_{D_\pm} u_0^\pm}$. Condition (ii) guarantees that, asymptotically, there is no "blow up" of number of particles at any site. More precisely, this technical condition is imposed so that we have

$$\sup_{t \geq 0} \mathbb{E} \left[\langle 1, J_t^{N,+} \rangle^2 \right] \leq C < \infty \quad \text{for sufficiently large } N. \quad (2.37)$$

The above can be easily checked by comparing with the process $\bar{\eta}$ that has no annihilation (i.e. $\bar{\eta}$ has generator $\mathfrak{L}_0^\varepsilon$). Alternatively, we can use the comparison result (3.11) to prove (2.37). \square

Remark 2.25. (Generalization) We can generalize our results in a number of ways by the same argument. For example, the initial number of particles in D_+ and D_- can be different, the condition $N\varepsilon^d = 1$ can be relaxed to $\lim_{N \rightarrow \infty} N\varepsilon^d \rightarrow 1$ where ε depends on N . The annihilation constant λ can be replaced by a space and time dependent function $\lambda(t, x) \in C([0, \infty) \times I)$. The diffusion coefficients in D_+ and D_- can be different. The condition " $\mathfrak{X}_0^{N,\pm}$ has mass one for all N " can be replaced by "the mass of $\mathfrak{X}_0^{N,\pm}$ is uniformly bounded in N ". More generally, the same method can be extended to deal with similar models with more than two types of particles. \square

The remaining part of this paper is devoted to the proof of Theorem 2.20 and Theorem 2.22. We first prove Theorem 2.22 because the proof of Theorem 2.20 relies on Theorem 2.22.

3 Propagation of Chaos

3.1 Duality

The starting point of our analysis is the discrete integral equation for γ^ε in Lemma 3.2. At the heart of its proof is the dual relation in Lemma 3.1, which says that the two independent processes $\xi^0 = (\xi_t^0)_{t \geq 0}$ and $\eta^0 = (\eta_t^0)_{t \geq 0}$ of independent random walks *with no interaction* are dual to each other with respect to the function $\frac{A(\xi, \eta)}{\alpha_\varepsilon(\xi)}$, where $\xi, \eta \in E^\varepsilon$. Such kind of dual formula for the whole grid \mathbb{Z}^d appeared in [5] and in Chapter 15 of [11].

Lemma 3.1. (*Duality for independent processes*) *Let $\xi^0 = (\xi_t^0)_{t \geq 0}$ and $\eta^0 = (\eta_t^0)_{t \geq 0}$ be independent continuous time Markov processes on E^ε with generator $\mathfrak{L}_0^\varepsilon$ defined in Definition 2.8. Then we have*

$$\mathbb{E} \left[\frac{A(\xi_t^0, \eta_0^0)}{\alpha_\varepsilon(\xi_t^0)} \right] = \mathbb{E} \left[\frac{A(\xi_0^0, \eta_t^0)}{\alpha_\varepsilon(\xi_0^0)} \right] \quad \text{for every } t \geq 0. \quad (3.1)$$

Proof Assume $\xi_0^0 \in \Omega_{n,m}^\varepsilon$ and $\eta_0^0 \in \Omega_{N,M}^\varepsilon$. Then we have $\xi_t^0 \in \Omega_{n,m}^\varepsilon$ and $\eta_t^0 \in \Omega_{N,M}^\varepsilon$ for all $t \geq 0$. Without loss of generality, we may assume $N \geq n \geq 1$ and $M \geq m \geq 1$ as otherwise both sides inside expectations of (3.1) are zero by the definition of $A(\xi, \eta)$.

Denote U the map that sends $(\vec{r}, \vec{s}) \in (D_+^\varepsilon)^n \times (D_-^\varepsilon)^m$ to $(\sum_i \delta_{r_i}, \sum_j \delta_{s_j}) \in \Omega_{n,m}^\varepsilon$ for any (n, m) . We first focus on D_+ in Step 1 and Step 2 below.

Step 1. For any $\vec{r} \in (D_+^\varepsilon)^n$ and $\eta^+ \in \Omega_{N,0}^\varepsilon$, fix some $\vec{x}^+ = (x_1^+, \dots, x_N^+) \in U^{-1}(\eta^+)$. Then by the definition (2.31) of A ,

$$A^+(\vec{r}, \eta^+) = \# \left\{ \vec{i} : \vec{x}_i^+ = \vec{r} \right\}, \quad (3.2)$$

where n -tuples $\vec{i} := (i_1, \dots, i_n)$ consist of distinct positive integers in the set $\{1, 2, \dots, N\}$, $\vec{x}_i^+ := (x_{i_1}^+, \dots, x_{i_n}^+)$ and $\#S$ denotes the number of elements in the finite set S .

Step 2. Denote by $\mathbb{P}_0^{\eta^+}$ the law of the *unlabeled* process $(\eta_t^0)_{t \geq 0}$ starting from $\eta^+ \in \Omega_{N,0}^\varepsilon$ and has generator $\mathfrak{L}_0^\varepsilon$. Let $\vec{x}^+ = (x_1^+, \dots, x_N^+) \in U^{-1}(\eta^+)$, and $\vec{X}_t^{+, \varepsilon} := (X_1^{+, \varepsilon}(t), \dots, X_N^{+, \varepsilon}(t))$ be independent CTRWs in D_+^ε starting from \vec{x} , whose law will be denoted as $\mathbb{P}^{\vec{x}^+}$. Then by (3.2), we have

$$\mathbb{E}_0^{\eta^+} [A(\vec{r}, \eta_t^0)] = \mathbb{E}[\# \{n\text{-tuples } \vec{i} : \vec{X}_i^{+, \varepsilon}(t) = \vec{r}\}] = \sum_{\vec{i}: n\text{-tuples}} \mathbb{P}^{\vec{x}_i^+} (\vec{X}_i^{+, \varepsilon}(t) = \vec{r}). \quad (3.3)$$

where $\mathbb{P}^{\vec{x}_i^+}$ is the law of $\{\vec{X}_i^{+, \varepsilon}(t); t \geq 0\}$. Denote by $p^\varepsilon(\theta, \vec{z}, \vec{w})$ the transition density of n independent CTRWs in D_+^ε . By Chapman-Kolmogorov equation, we have for any $\theta \in [0, t]$,

$$\mathbb{P}^{\vec{x}_i^+} (\vec{X}_i^{+, \varepsilon}(t) = \vec{r}) = \sum_{\vec{z} \in (D_+^\varepsilon)^n} p^\varepsilon(\theta, \vec{x}_i^+, \vec{z}) p^\varepsilon(t - \theta, \vec{z}, \vec{r}) m(\vec{z}) m(\vec{r}).$$

Putting this into (3.3), we have

$$\begin{aligned} \mathbb{E}_0^{\eta^+} [A(\vec{r}, \eta_t^0)] &= m_\varepsilon(\vec{r}) \sum_{\vec{z}} \sum_{\vec{i}} \mathbb{P}^{\vec{x}_i^+} (\vec{X}_i^{+, \varepsilon}(\theta) = \vec{z}) p^\varepsilon(t - \theta, \vec{z}, \vec{r}) \\ &= m_\varepsilon(\vec{r}) \sum_{\vec{z}} \mathbb{E}_0^{\eta^+} [A(\vec{z}, \eta_\theta^0)] p^\varepsilon(t - \theta, \vec{z}, \vec{r}) \quad \text{by (3.3) again} \\ &= m_\varepsilon(\vec{r}) \sum_{\vec{z}} \mathbb{E}_0^{\eta^+} [A(\vec{z}, \eta_\theta^0)] p^\varepsilon(t - \theta, \vec{r}, \vec{z}) \quad \text{by symmetry of } p^\varepsilon \\ &= m_\varepsilon(\vec{r}) \mathbb{E} \left[\frac{A(\vec{Y}_{t-\theta}^{+, \varepsilon}, \eta_\theta^0)}{m_\varepsilon(\vec{Y}_{t-\theta}^{+, \varepsilon})} \right], \end{aligned} \quad (3.4)$$

where \mathbb{E} is the expectation corresponding the probability measure under which each coordinate processes of $\{\vec{Y}_t^{+, \varepsilon}; t \geq 0\}$ are independent CTRWs with $\vec{Y}_0^{+, \varepsilon} = \vec{r}$ and are independent of $(\eta_t^0)_{t \geq 0}$.

Step 3. Now we work on $D_+ \times D_-$. For any $(\vec{r}, \vec{s}) \in (D_+^\varepsilon)^n \times (D_-^\varepsilon)^m$ and $\eta = (\eta^+, \eta^-) \in \Omega_{N, M}^\varepsilon$, take $\vec{x} = (x_1^+, \dots, x_N^+, x_1^-, \dots, x_M^-) \in U^{-1}(\eta)$. As in step 1, we have

$$A^+((\vec{r}, \vec{s}), \eta) = \sharp \left\{ (\vec{i}, \vec{j}) : (\vec{x}_{\vec{i}}^+, \vec{x}_{\vec{j}}^-) = (\vec{r}, \vec{s}) \right\}, \quad (3.5)$$

where \vec{i} runs over all n -tuples $\vec{i} := (i_1, \dots, i_n)$ consisting of distinct positive integers in the set $\{1, 2, \dots, N\}$, and \vec{j} over all m -tuples $\vec{j} := (j_1, \dots, j_m)$ consisting of distinct positive integers in the set $\{1, 2, \dots, M\}$.

Denote by \mathbb{P}_0^η the law of the *unlabeled* process $(\eta_t^0)_{t \geq 0}$ starting from $\eta \in \Omega_{N, M}^\varepsilon$ and has generator $\mathfrak{L}_0^\varepsilon$. Since all processes on D_+^ε are independent of those on D_-^ε , we can proceed as in step 2 (via (3.5)) to obtain

$$\mathbb{E}_0^\eta[A((\vec{r}, \vec{s}), \eta_t^0)] = m_\varepsilon(\vec{r}, \vec{s}) \mathbb{E} \left[\frac{A(\vec{Y}_{t-\theta}^\varepsilon, \eta_\theta^0)}{m_\varepsilon(\vec{Y}_{t-\theta}^\varepsilon)} \right] \quad \text{for } \theta \in [0, t], \quad (3.6)$$

where $\vec{Y}^\varepsilon := (Y_1^{+, \varepsilon}, \dots, Y_n^{+, \varepsilon}, Y_1^{-, \varepsilon}, \dots, Y_m^{-, \varepsilon})$ is independent of η^0 with $\vec{Y}_0^\varepsilon = (\vec{r}, \vec{s})$, and its components are mutually independent CTRWs on D_\pm^ε , respectively. The proof is now complete by taking $\theta = 0$. \square

We now formulate the discrete integral equations that we need. Recall the definition of \mathfrak{K}^ε from (2.10) and the definition of $\gamma^\varepsilon((\vec{r}, \vec{s}), t)$ from (2.33).

Lemma 3.2. (*Discrete integral equation for γ^ε*) For any $\varepsilon > 0$, $t > 0$, $(\vec{r}, \vec{s}) \in (D_+^\varepsilon)^n \times (D_-^\varepsilon)^m$, non-negative integers n, m and initial distribution η_0 , we have

$$\begin{aligned} \gamma^\varepsilon((\vec{r}, \vec{s}), t) &= \sum_{(\vec{r}', \vec{s}')} \gamma^\varepsilon((\vec{r}', \vec{s}'), 0) p^\varepsilon(t, (\vec{r}, \vec{s}), (\vec{r}', \vec{s}')) m(\vec{r}', \vec{s}') \\ &\quad + \int_0^t \sum_{(\vec{r}', \vec{s}')} p^\varepsilon(t-s, (\vec{r}, \vec{s}), (\vec{r}', \vec{s}')) \mathbb{E}[\mathfrak{K}^\varepsilon A((\vec{r}', \vec{s}'), \eta_s)] \varepsilon^{d(n+m)} ds, \end{aligned} \quad (3.7)$$

where \mathfrak{K}^ε acts on the η -variable of $A((\vec{r}, \vec{s}), \eta)$.

Proof Starting from (3.1), we can obtain Lemma 3.2 by ‘integration by parts’ as follows.

Let $\mathbb{P}_{(\xi^0)}$ and $\mathbb{P}_{(\eta^0)}$ be the laws of ξ^0 and η^0 respectively. (3.1) is equivalent to saying that for any ξ and η , we have

$$\mathbb{E}_{(\xi^0)} \left[\frac{A(\xi_w^0, \eta)}{\alpha_\varepsilon(\xi_w^0)} \middle| \xi_0^0 = \xi \right] = \mathbb{E}_{(\eta^0)} \left[\frac{A(\xi, \eta_w^0)}{\alpha_\varepsilon(\xi)} \middle| \eta_0^0 = \eta \right] \quad \text{for every } w \geq 0. \quad (3.8)$$

Taking $w = t - s$, we see that (3.8) is in turn equivalent to

$$F_s^{(\xi)}(\eta) := P_{t-s}^{(\xi^0)} \left(\frac{A(\cdot, \eta)}{\alpha_\varepsilon(\cdot)} \right) (\xi) = P_{t-s}^{(\eta^0)} \left(\frac{A(\xi, \cdot)}{\alpha_\varepsilon(\xi)} \right) (\eta) =: G_s^{(\eta)}(\xi) \quad \text{for every } s \in [0, t] \text{ and } t \geq 0, \quad (3.9)$$

where $P_t^{(\xi^0)}$ and $P_t^{(\eta^0)}$ are the transition semigroup of ξ^0 and η^0 , respectively, and they act on the ξ and η variables in $\frac{A(\xi, \eta)}{\alpha_\varepsilon(\xi)}$, respectively. Therefore, with $\mathfrak{L}_0^\varepsilon$ acting on the η variable, we have

$$\frac{\partial}{\partial s} F_s^{(\xi)}(\eta) = \frac{\partial}{\partial s} G_s^{(\eta)}(\xi) = -\mathfrak{L}_0^\varepsilon P_{t-s}^{(\eta^0)} \left(\frac{A(\xi, \cdot)}{\alpha_\varepsilon(\xi)} \right) (\eta) = -\mathfrak{L}_0^\varepsilon F_s^{(\xi)}(\eta). \quad (3.10)$$

Recall that η_t is the configuration process of our interacting system with generator $\mathfrak{L}_0^\varepsilon + \mathfrak{K}^\varepsilon$ (see Definition 2.8). Fix ξ and consider the function $(s, \eta) \mapsto F_s(\eta) := F_s^{(\xi)}(\eta)$. We have

$$M_s := F_s(\eta_s) - F_0(\eta_0) - \int_0^s \left(\frac{\partial F_r}{\partial r} + \mathfrak{L}_0^\varepsilon F_r + \mathfrak{K}^\varepsilon F_r \right) (\eta_r) dr$$

is a \mathcal{F}_s^η -martingale for $s \in [0, t]$. By (3.10) and the fact that $E^\eta[M_t] = E^\eta[M_0] = 0$, where \mathbb{P}^η is the law of $(\eta_t)_{t \geq 0}$ starting from η , we have

$$0 = \mathbb{E}^\eta \left[\frac{A(\xi, \eta_t)}{\alpha_\varepsilon(\xi)} \right] - P_t^{(\xi^0)} \left(\frac{A(\cdot, \eta)}{\alpha_\varepsilon(\cdot)} \right) (\xi) - \int_0^t \mathbb{E}^\eta \left[\mathfrak{K}^\varepsilon P_{t-r}^{(\xi^0)} \left(\frac{A(\cdot, \eta_r)}{\alpha_\varepsilon(\cdot)} \right) (\xi) \right] dr$$

for all ξ and η . This is equivalent to the stated equation in the lemma. \square

It is clear that $\mathfrak{K}^\varepsilon A(\xi, \eta) \leq 0$. Hence, as an immediate consequence of Lemma 3.2, we have the following comparison result:

$$\gamma^\varepsilon((\vec{r}, \vec{s}), t) \leq \sum_{(\vec{r}', \vec{s}')} \gamma^\varepsilon((\vec{r}', \vec{s}'), 0) p^\varepsilon(t, (\vec{r}, \vec{s}), (\vec{r}', \vec{s}')) m(\vec{r}', \vec{s}') \quad (3.11)$$

for all $t > 0$ and $(\vec{r}, \vec{s}) \in (D_+^\varepsilon)^n \times (D_-^\varepsilon)^m$.

3.2 Annihilation near the interface

For any $\xi = (\xi^+, \xi^-) \in E^\varepsilon$, we let $\xi_{(x)}^+ = \xi^+(x) \mathbf{1}_x$, the element that has only $\xi^+(x)$ number of particles at x , and none elsewhere. Similarly, we denote $\xi^-(y) \mathbf{1}_y$ by $\xi_{(y)}^-$. Set $\xi_{(x,y)} = (\xi^+(x) \mathbf{1}_x, \xi^-(y) \mathbf{1}_y)$, the element that has only $\xi^+(x)$ number of particles at x , $\xi^-(y)$ number of particles at y , and none elsewhere.

Lemma 3.3. *Let \mathfrak{K}^ε be the operator defined in (2.10) and acts on the η -variable of $A(\xi, \eta)$. Then*

$$\mathfrak{K}^\varepsilon A(\xi, \eta) = \sum_{z \in I^\varepsilon} A(\xi - \xi_{(z_+, z_-)}, \eta) \cdot \mathfrak{K}^\varepsilon A(\xi_{(z_+, z_-)}, \eta). \quad (3.12)$$

Moreover, if $\xi \in \Xi := \{\xi : \xi^\pm(z_\pm) \leq 1 \text{ for every } z \in I^\varepsilon\}$, then

$$\mathfrak{K}^\varepsilon A(\xi, \eta) = -\frac{\lambda}{\varepsilon} \sum_{z \in I^\varepsilon : \xi^+(z_+) = 1} \Psi_\varepsilon(z) A(\xi + \mathbf{1}_{(0, z_-)}, \eta) \quad (3.13)$$

$$-\frac{\lambda}{\varepsilon} \sum_{z \in I^\varepsilon : \xi^-(z_-) = 1} \Psi_\varepsilon(z) A(\xi + \mathbf{1}_{(z_+, 0)}, \eta) \quad (3.14)$$

$$-\frac{\lambda}{\varepsilon} \sum_{z \in I^\varepsilon : \xi_{(z_+, z_-)} = (1, 1)} \Psi_\varepsilon(z) A(\xi, \eta). \quad (3.15)$$

Proof Observe that $A(\xi - \xi_{(x,y)}, \eta) A(\xi_{(x,y)}, \eta) = A(\xi, \eta)$. Consequently

$$\begin{aligned} & \frac{\lambda}{\varepsilon} \Psi_\varepsilon(z) \eta^+(z_+) \eta^-(z_-) (A(\xi, \eta - \mathbf{1}_{(z_+, z_-)}) - A(\xi, \eta)) \\ &= \frac{\lambda}{\varepsilon} \Psi_\varepsilon(z) \eta^+(z_+) \eta^-(z_-) A(\xi - \xi_{(z_+, z_-)}, \eta) (A(\xi_{(z_+, z_-)}, \eta - \mathbf{1}_{(z_+, z_-)}) - A(\xi_{(z_+, z_-)}, \eta)) \\ &= A(\xi - \xi_{(z_+, z_-)}, \eta) \mathfrak{K}^\varepsilon A(\xi_{(z_+, z_-)}, \eta). \end{aligned}$$

Thus (3.12) holds. On other hand,

$$\mathfrak{K}^\varepsilon A(\xi_{(z_+, z_-)}, \eta) = \Psi_\varepsilon(z) \frac{\lambda}{\varepsilon} \eta^+(z_+) \eta^-(z_-) \times \begin{cases} -1, & \text{if } \xi_{(z_+, z_-)} = (1, 0) \text{ or } (0, 1) \\ 1 - \eta^+(z_+) - \eta^-(z_-), & \text{if } \xi_{(z_+, z_-)} = (1, 1) \end{cases}.$$

Observe also that for $x \in D_+^\varepsilon$ and $y \in D_-^\varepsilon$,

$$A(\xi - \xi_{(x,y)}, \eta) \eta^+(x) \eta^-(y) = A(\xi - \xi_{(x,y)} + \mathbf{1}_{(x,y)}, \eta)$$

and

$$\begin{aligned}
& A(\xi - \xi_{(x,y)}) \eta^+(x)^2 \eta^-(y) \\
&= A(\xi - \xi_{(x,y)}, \eta) (\eta^+(x)^2 - \eta^+(x) + \eta^+(x)) \eta^-(y) \\
&= A(\xi - \xi_{(x,y)}, \eta) A(2\mathbf{1}_x, \eta^+) \eta^-(y) + A(\xi - \xi_{(x,y)}, \eta) \eta^+(x) \eta^-(y) \\
&= A(\xi - \xi_{(x,y)} + 2\mathbf{1}_{(x,0)} + \mathbf{1}_{(0,y)}, \eta) + A(\xi, \eta).
\end{aligned}$$

Similarly,

$$A(\xi - \xi_{(x,y)}, \eta) \eta^+(x) \eta^-(y)^2 = A(\xi - \xi_{(x,y)} + \mathbf{1}_{(x,0)} + 2\mathbf{1}_{(0,y)}, \eta) + A(\xi, \eta).$$

From the above calculations and (3.12), we see that for $\xi \in \Xi$,

$$\begin{aligned}
\mathfrak{K}^\varepsilon A(\xi, \eta) &= -\frac{\lambda}{\varepsilon} \sum_{z \in I^\varepsilon: \xi(z_+, z_-) = (1,0)} \Psi_\varepsilon(z) A(\xi + \mathbf{1}_{(0, z_-)}, \eta) \\
&\quad -\frac{\lambda}{\varepsilon} \sum_{z \in I^\varepsilon: \xi(z_+, z_-) = (0,1)} \Psi_\varepsilon(z) A(\xi + \mathbf{1}_{(z_+, 0)}, \eta) \\
&\quad -\frac{\lambda}{\varepsilon} \sum_{z \in I^\varepsilon: \xi(z_+, z_-) = (1,1)} \Psi_\varepsilon(z) (A(\xi, \eta) + A(\xi + \mathbf{1}_{(0, z_-)}, \eta) + A(\xi + \mathbf{1}_{(z_+, 0)}, \eta)),
\end{aligned}$$

which gives the desired formula. \square

3.3 Uniform bound and equi-continuity

We extend to define $\gamma^{\varepsilon, (n,m)}(\cdot, t)$ continuously on $\overline{D}_+^n \times \overline{D}_-^m$ while preserving the supremum and the infimum in each small ε -cube. We can accomplish this by the interpolation described in [2] or [39], or by a sequence of harmonic extensions along simplexes with increasing dimensions (described in [23]). Recall that the definition of $\gamma^{\varepsilon, (n,m)}(\cdot, t)$ depends on the initial configuration η_0 of the interacting random walks (see Definition 2.21), which has the normalization (2.32).

Theorem 3.4. *There exists $\varepsilon_0 > 0$ such that for any $(n, m) \in \mathbb{N} \times \mathbb{N}$, the family of functions $\{\gamma^\varepsilon((\vec{r}, \vec{s}), t)\}_{\varepsilon \in (0, \varepsilon_0)}$ is uniformly bounded and equi-continuous on $\overline{D}_+^n \times \overline{D}_-^m \times (0, \infty)$, which is uniform in the initial configuration η_0 that satisfies (2.32).*

Proof We first prove uniform boundedness. By (3.11) and the Gaussian upper bound in Theorem 2.9, we have

$$\begin{aligned}
\gamma^\varepsilon((\vec{r}, \vec{s}), t) &\leq \sum_{(\vec{r}', \vec{s}')} \gamma^\varepsilon((\vec{r}', \vec{s}'), 0) p^\varepsilon(t, (\vec{r}, \vec{s}), (\vec{r}', \vec{s}')) m(\vec{r}', \vec{s}') \\
&\leq \left(\frac{C}{t^{d/2}}\right)^{n+m} \sum_{(\vec{r}', \vec{s}') \in \overline{D}_+^n \times \overline{D}_-^m} A((\vec{r}', \vec{s}'), \eta_0) \varepsilon^{d(n+m)} \quad \text{whenever } \varepsilon \in (0, \varepsilon_0).
\end{aligned}$$

Since the initial distribution $\eta_0 = (\eta_0^+, \eta_0^-)$ has the property that $\sum_{x \in D_+^\varepsilon} \eta_0^+(x) = \sum_{y \in D_-^\varepsilon} \eta_0^-(y) = \varepsilon^{-d}$, we have

$$\begin{aligned}
&\sum_{(\vec{r}, \vec{s}) \in \overline{D}_+^n \times \overline{D}_-^m} A((\vec{r}, \vec{s}), \eta_0) \varepsilon^{d(n+m)} \\
&= \left(\sum_{\vec{r} \in \overline{D}_+^n} A^+(\vec{r}, \eta_0^+) \right) \left(\sum_{\vec{s} \in \overline{D}_-^m} A^-(\vec{s}, \eta_0^-) \right) \varepsilon^{d(n+m)}
\end{aligned}$$

$$\begin{aligned}
&\leq \left(\sum_{\vec{r} \in \overline{D}_+^n} \prod_{i=1}^n \eta_0^+(r_i) \right) \left(\sum_{\vec{s} \in \overline{D}_-^m} \prod_{j=1}^m \eta_0^-(s_j) \right) \varepsilon^{d(n+m)} \quad \text{since } A_k^n \leq n^k \\
&\leq \prod_{i=1}^n \left(\sum_{r_i \in \overline{D}_+} \eta_0^+(r_i) \varepsilon^d \right) \prod_{j=1}^m \left(\sum_{s_j \in \overline{D}_-} \eta_0^-(s_j) \varepsilon^d \right) = 1.
\end{aligned} \tag{3.16}$$

Thus there exist $\varepsilon_0 = \varepsilon_0(d, D, \rho)$ and $C = C(d, D, \rho) > 0$ such that for all $t \in (0, \infty)$, $(n, m) \in \mathbb{N} \times \mathbb{N}$ and $\varepsilon \in (0, \varepsilon_0)$,

$$\sup_{\varepsilon \in (0, \varepsilon_0)} \sup_{\xi \in \Omega_{n,m}^\varepsilon} \gamma^\varepsilon(\xi, t) \leq \left(\frac{C}{t^{d/2}} \right)^{n+m}. \tag{3.17}$$

We next show that both terms on the right hand side of (3.7) are equi-continuous. Recall that we can rewrite the equation (3.7) as

$$\gamma^\varepsilon((\vec{r}, \vec{s}), t) = F^\varepsilon((\vec{r}, \vec{s}), t) + G^\varepsilon((\vec{r}, \vec{s}), t),$$

where

$$\begin{aligned}
F^\varepsilon((\vec{r}, \vec{s}), t) &:= \sum_{(\vec{r}', \vec{s}')} \gamma^\varepsilon((\vec{r}', \vec{s}'), 0) p^\varepsilon(t, (\vec{r}, \vec{s}), (\vec{r}', \vec{s}')) m(\vec{r}', \vec{s}'), \\
G^\varepsilon((\vec{r}, \vec{s}), t) &:= \int_0^t \sum_{(\vec{r}', \vec{s}')} p^\varepsilon(t-s, (\vec{r}, \vec{s}), (\vec{r}', \vec{s}')) \mathbb{E}[\mathfrak{R}^\varepsilon A((\vec{r}', \vec{s}'), \eta_s)] \varepsilon^{d(n+m)} ds.
\end{aligned}$$

Now let $(\vec{r}, \vec{s}), (\vec{p}, \vec{q}) \in (D_+^\varepsilon)^n \times (D_-^\varepsilon)^m$ and $0 < t < \ell \leq \infty$. For the first term,

$$\begin{aligned}
&\left| F^\varepsilon((\vec{r}, \vec{s}), t) - F^\varepsilon((\vec{p}, \vec{q}), \ell) \right| \\
&= \left| \sum_{(\vec{r}', \vec{s}')} \left(p^\varepsilon(t, (\vec{r}, \vec{s}), (\vec{r}', \vec{s}')) - p^\varepsilon(\ell, (\vec{p}, \vec{q}), (\vec{r}', \vec{s}')) \right) \mathbb{E}[A((\vec{r}', \vec{s}'), \eta_0)] \varepsilon^{d(n+m)} \right| \\
&\leq \left(\sup_{(\vec{r}', \vec{s}')} \left| p^\varepsilon(t, (\vec{r}, \vec{s}), (\vec{r}', \vec{s}')) - p^\varepsilon(\ell, (\vec{p}, \vec{q}), (\vec{r}', \vec{s}')) \right| \right) \mathbb{E}_0 \left[\sum_{(\vec{r}', \vec{s}')} A((\vec{r}', \vec{s}'), \eta_0) \varepsilon^{d(n+m)} \right] \\
&\leq \sup_{(\vec{r}', \vec{s}')} \left| p^\varepsilon(t, (\vec{r}, \vec{s}), (\vec{r}', \vec{s}')) - p^\varepsilon(\ell, (\vec{p}, \vec{q}), (\vec{r}', \vec{s}')) \right|,
\end{aligned}$$

where we have used (3.16) in the last line. By the uniform Hölder continuity of $p^\varepsilon(t, (\vec{r}, \vec{s}), (\vec{r}', \vec{s}'))$ (see Theorem 5.12 below) and the fact that $p^\pm(t, x, y) \in C((0, \infty) \times \overline{D}_\pm \times \overline{D}_\pm)$, we see that $\{F^\varepsilon\}$ is equi-continuous at $((\vec{r}, \vec{s}), t)$. For the second term, note that

$$G^\varepsilon((\vec{p}, \vec{q}), \ell) - G^\varepsilon((\vec{r}, \vec{s}), t) = \int_t^\ell H^{(1)}(s) ds + \int_0^t H^{(2)}(s) ds, \tag{3.18}$$

where

$$\begin{aligned}
H^{(1)}(s) &:= \sum_{(\vec{r}', \vec{s}')} p^\varepsilon(\ell-s, (\vec{p}, \vec{q}), (\vec{r}', \vec{s}')) \mathbb{E}[\mathfrak{R}^\varepsilon A((\vec{r}', \vec{s}'), \eta_s)] \varepsilon^{d(n+m)} \text{ and} \\
H^{(2)}(s) &:= \sum_{(\vec{r}', \vec{s}')} \left[p^\varepsilon(\ell-s, (\vec{r}, \vec{s}), (\vec{r}', \vec{s}')) - p^\varepsilon(t-s, (\vec{p}, \vec{q}), (\vec{r}', \vec{s}')) \right] \mathbb{E}[\mathfrak{R}^\varepsilon A((\vec{r}', \vec{s}'), \eta_s)] \varepsilon^{d(n+m)}.
\end{aligned}$$

In the remaining, we will show that G^ε is equi-continuous. We first deal with $H^{(1)}$ in (3.18).

As in (3.16), we have

$$\begin{aligned} & \sum_{(\vec{r}', \vec{s}')} p^\varepsilon(\theta_1, (\vec{p}, \vec{q}), (\vec{r}', \vec{s}')) A((\vec{r}', \vec{s}'), \eta_{\theta_2}) \varepsilon^{d(n+m)} \\ & \leq \prod_{i=1}^n \left(\sum_{r'_i} p^\varepsilon(\theta_1, p_i, r'_i) \eta_{\theta_2}^+(r'_i) \varepsilon^d \right) \prod_{j=1}^m \left(\sum_{s'_j} p^\varepsilon(\theta_1, q_j, s'_j) \eta_{\theta_2}^-(s'_j) \varepsilon^d \right). \end{aligned} \quad (3.19)$$

On other hand, using (3.11), the Chapman Kolmogorov equation and assumption (ii) for η_0 , in this order, we have

$$\sup_{\theta_1, \theta_2 > 0} \sup_{a \in D_+^\varepsilon} \mathbb{E} \left[\left(\sum_{x \in D_+^\varepsilon} p^\varepsilon(\theta_1, a, x) \eta_{\theta_2}^+(x) \varepsilon^d \right)^2 \right] \leq C \quad (3.20)$$

for large enough N , where $C > 0$ is a constant.

$$\begin{aligned} & \sum_{(\vec{r}', \vec{s}')} p^\varepsilon(\theta_1, (\vec{p}, \vec{q}), (\vec{r}', \vec{s}')) \left| \mathbb{E} \left[\mathcal{K}^\varepsilon A((\vec{r}', \vec{s}'), \eta_{\theta_2}) \right] \right| \varepsilon^{d(n+m)} \\ & \leq \mathbb{E} \left[\frac{\lambda}{\varepsilon} \sum_{z \in I} \eta_{\theta_2}^+(z_+) \eta_{\theta_2}^-(z_-) \sum_{(\vec{r}', \vec{s}')} p^\varepsilon(\theta_1, (\vec{p}, \vec{q}), (\vec{r}', \vec{s}')) 2A((\vec{r}', \vec{s}'), \eta_{\theta_2}) \varepsilon^{d(n+m)} \right] \\ & \leq 2\mathbb{E} \left[\langle 1, J_{\theta_2}^N \rangle \prod_{i=1}^n \left(\sum_{r'_i} p^\varepsilon(\theta_1, p_i, r'_i) \eta_{\theta_2}^+(r'_i) \varepsilon^d \right) \prod_{j=1}^m \left(\sum_{s'_j} p^\varepsilon(\theta_1, q_j, s'_j) \eta_{\theta_2}^-(s'_j) \varepsilon^d \right) \right] \\ & \leq C \quad \text{uniformly for } \theta_1 > 0, \theta_2 > 0, (\vec{p}, \vec{q}) \in (D_+^\varepsilon)^n \times (D_-^\varepsilon)^m \text{ and } \varepsilon > 0 \text{ small enough.} \end{aligned}$$

We have used (3.19) for the second inequality. The last inequality follows from Hölder's inequality, (2.37) and (3.20). Therefore for any (n, m) ,

$$\sup_{\theta_1, \theta_2 > 0} \sup_{(\vec{p}, \vec{q}) \in (D_+^\varepsilon)^n \times (D_-^\varepsilon)^m} \sum_{(\vec{r}', \vec{s}')} p^\varepsilon(\theta_1, (\vec{p}, \vec{q}), (\vec{r}', \vec{s}')) \left| \mathbb{E} \left[\mathcal{K}^\varepsilon A((\vec{r}', \vec{s}'), \eta_{\theta_2}) \right] \right| \varepsilon^{d(n+m)} \leq C \quad (3.21)$$

for large enough N , where $C > 0$ is a constant. Hence $\int_t^\ell |H^{(1)}(s)| ds \leq C(\ell - t) \rightarrow 0$ as $\ell \rightarrow t$, uniformly for $(\vec{p}, \vec{q}), s \in (t, \ell)$ and ε small enough. Finally, we deal with $H^{(2)}$. For any $h \in (0, t)$, we have

$$\left| \int_0^t H^{(2)}(s) ds \right| \leq \int_0^{t-h} |H^{(2)}(s)| ds + \int_{t-h}^t |H^{(2)}(s)| ds.$$

By (3.21), we have $\int_{t-h}^t |H^{(2)}(s)| ds \leq Ch$. By the Hölder continuity of p^ε (cf. Theorem 5.12),

$$\begin{aligned} \int_0^{t-h} |H^{(2)}(s)| ds & \leq \int_0^{t-h} \sup_{(\vec{r}', \vec{s}')} \left| p^\varepsilon(\ell - s, (\vec{r}, \vec{s}), (\vec{r}', \vec{s}')) - p^\varepsilon(t - s, (\vec{p}, \vec{q}), (\vec{r}', \vec{s}')) \right| \\ & \quad \mathbb{E} \left[\frac{\lambda}{\varepsilon} \sum_{z \in I} \eta_s^+(z_+) \eta_s^-(z_-) \sum_{(\vec{r}', \vec{s}')} 2A((\vec{r}', \vec{s}'), \eta_s) \varepsilon^{d(n+m)} \right] ds \\ & \leq (t-h) C \frac{|\ell - t|^{\sigma_1} + \|(\vec{r}, \vec{s}) - (\vec{p}, \vec{q})\|^{\sigma_2}}{h^{\sigma_3}} \quad \text{for sufficiently small } \varepsilon > 0, \end{aligned}$$

where σ_i ($i = 1, 2, 3$) are positive constants. Since $h \in (0, t)$ is arbitrary, we see that $\left| \int_0^t H^{(2)}(s) ds \right| \rightarrow 0$ as $|\ell - t| + \|(\vec{r}, \vec{s}) - (\vec{p}, \vec{q})\| \rightarrow 0$, uniformly for small enough $\varepsilon > 0$. Hence G^ε is equi-continuous at an arbitrary $((\vec{r}, \vec{s}), t) \in (D_+^\varepsilon)^n \times (D_-^\varepsilon)^m \times (0, \infty)$. \square

From Theorem 3.4 and a diagonal selection argument, it follows that for any sequence $\varepsilon_k \rightarrow 0$ there is a subsequence along which γ^ε converges on $\overline{D}_+^n \times \overline{D}_-^m \times (0, T)$, uniformly on the compacts, to some $\gamma^{(n,m)} \in C(\overline{D}_+^n \times \overline{D}_-^m \times (0, T))$, for every $(n, m) \in \mathbb{N} \times \mathbb{N}$. Our goal is to show that

$$\gamma_t^{(n,m)}(\vec{r}, \vec{s}) = \prod_{i=1}^n u_+(t, r_i) \prod_{j=1}^m u_-(t, s_j).$$

We will achieve this by first showing that both $\Gamma = \{\gamma^{(n,m)}\}$ and $\Phi_t^{(n,m)} := \prod_{i=1}^n u_+(t, r_i) \prod_{j=1}^m u_-(t, s_j)$ satisfy the same an infinite hierarchy of equations, and then establishing uniqueness of the hierarchy.

3.4 Limiting hierarchy

Note that $D_+^n \times D_-^m$ is a bounded Lipschitz domain in $\mathbb{R}^{(n+m)d}$, and that the boundary $\partial(D_+^n \times D_-^m)$ contains the disjoint union $\cup_{i=1}^n \partial_+^i \cup \cup_{j=1}^m \partial_-^j$ where

$$\partial_+^i := \left(D_+ \times \cdots \times (\partial D_+ \cap I) \times \cdots \times D_+ \right) \times D_-^m, \quad (3.22)$$

$$\partial_-^j := D_+^n \times \left(D_- \times \cdots \times (\partial D_- \cap I) \times \cdots \times D_- \right). \quad (3.23)$$

We define the function $\rho = \rho_{(n,m)} : D_+^n \times D_-^m \rightarrow \mathbb{R}$ by $\rho(\vec{r}, \vec{s}) := \prod_{i=1}^n \rho_+(r_i) \prod_{j=1}^m \rho_-(s_j)$. We also denote $p(t, (\vec{r}, \vec{s}), (\vec{r}', \vec{s}')) := \prod_{i=1}^n p^+(t, r_i, r'_i) \prod_{j=1}^m p^-(t, s_j, s'_j)$, where p^\pm is the transition density of the reflected diffusion X^\pm on \overline{D}_\pm with respect to the measure $\rho_\pm(x)dx$. We now characterize the subsequential limits of $\{\gamma^\varepsilon\}_{\varepsilon>0}$:

Theorem 3.5. *Let η_0^ε be a sequence of initial configurations that satisfy (2.32) with $\varepsilon = N^{-1/d}$, and the conditions of Theorem 2.20; that is, their corresponding empirical measures $\mathfrak{X}_0^{N,\pm}$ converges weakly to $u_0^\pm(z)dz$ in $M_{\leq 1}(\overline{D}_\pm)$ for some $u_0^\pm \in C(\overline{D}_\pm)$ and*

$$\overline{\lim}_{\varepsilon \rightarrow 0} \sup_{z \in D_\pm^\varepsilon} \mathbb{E} [\eta_0^{\varepsilon,\pm}(z)^2] < \infty. \quad (3.24)$$

Denote by $\Gamma^\varepsilon = \{\gamma^{\varepsilon,(n,m)}; t \geq 0, n, m \in \mathbb{N}\}$ the correlation functions for the interacting random walks with initial configuration η_0^ε . Let $\Gamma = \{\gamma_t^{(n,m)}; t \geq 0, n, m \in \mathbb{N}\}$ be any subsequential limit (as $\varepsilon \rightarrow 0$) of $\Gamma^\varepsilon = \{\gamma^{\varepsilon,(n,m)}; t \geq 0, n, m \in \mathbb{N}\}$. Then the following infinite system of hierarchical equations holds:

$$\begin{aligned} \gamma_t^{(n,m)}(\vec{r}, \vec{s}) &= \int_{D_+^n \times D_-^m} \Phi^{(n,m)}(\vec{a}, \vec{b}) p(t, (\vec{r}, \vec{s}), (\vec{a}, \vec{b})) \rho(\vec{a}, \vec{b}) d(\vec{a}, \vec{b}) \\ &\quad - \frac{\lambda}{2} \int_0^t \left(\sum_{i=1}^n \int_{\partial_+^i} \gamma_\theta^{(n,m+1)}(\vec{a}, (\vec{b}, a_i)) p(t-\theta, (\vec{r}, \vec{s}), (\vec{a}, \vec{b})) \frac{\rho(\vec{a}, \vec{b})}{\rho_+(a_i)} d\sigma_{(n,m)}(\vec{a}, \vec{b}) \right. \\ &\quad \left. + \sum_{j=1}^m \int_{\partial_-^j} \gamma_\theta^{(n+1,m)}((\vec{a}, b_j), \vec{b}) p(t-\theta, (\vec{r}, \vec{s}), (\vec{a}, \vec{b})) \frac{\rho(\vec{a}, \vec{b})}{\rho_-(b_j)} d\sigma_{(n,m)}(\vec{a}, \vec{b}) \right) d\theta, \end{aligned} \quad (3.25)$$

where $d(\vec{a}, \vec{b})$ is the Lebesgue measure on \mathbb{R}^{n+m} , $\sigma_{(n,m)}$ is the surface measure of $\partial(D_+^n \times D_-^m)$ and $\Phi^{(n,m)}(\vec{a}, \vec{b}) := \prod_{i=1}^n u_0^+(a_i) \prod_{j=1}^m u_0^-(b_j)$.

Remark 3.6. (i) The equation expresses $\gamma_t^{(n,m)}$ as an integral in time involving $\gamma^{(n,m+1)}$ and $\gamma^{(n+1,m)}$, thus forming a coupled chain of equations. In statistical physics, it is sometimes called the BBGKY

hierarchy³. It describes the evolution of the limiting (n, m) -particle correlation functions and hence the dynamics of the particles.

(ii) By Proposition 2.15, (3.25) is equivalent to

$$\gamma_t^{(n,m)}(\vec{r}, \vec{s}) = \mathbb{E}^{(\vec{r}, \vec{s})} \left[\Phi^{(n,m)}(X_{(n,m)}(t)) - \lambda \int_0^t (\Upsilon \gamma_s)^{(n,m)}(X_{(n,m)}(t-s)) dL_s^{(n,m)} \right]. \quad (3.26)$$

Here $L^{(n,m)}$ is boundary local time of $X_{(n,m)}$, the symmetric reflected diffusion on $D_+^n \times D_-^m$ corresponding to $(I_{(n+m)d \times (n+m)d}, \rho_{(n,m)})$, and $(\Upsilon v)^{(n,m)}$ is a function on $\partial(D_+^n \times D_-^m)$ defined as

$$(\Upsilon v)^{(n,m)}(\vec{r}, \vec{s}) := \begin{cases} v^{(n,m+1)}(\vec{r}, (\vec{s}, r_i)) \frac{\rho_{(n,m)}(\vec{r}, \vec{s})}{\rho_+(r_i)}, & \text{if } (\vec{r}, \vec{s}) \in \partial_+^i; \\ v^{(n+1,m)}((\vec{r}, s_j), \vec{s}) \frac{\rho_{(n,m)}(\vec{r}, \vec{s})}{\rho_-(s_j)}, & \text{if } (\vec{r}, \vec{s}) \in \partial_-^j; \\ 0, & \text{otherwise.} \end{cases}$$

Observe that the coordinate processes of $X_{(n,m)}$ consist of n independent copies of reflected diffusions in \bar{D}_+ and m independent copies of reflected diffusions in \bar{D}_- .

(iii) It is easy to check by using (ii) and Proposition 2.18 that

$$\tilde{\gamma}_t^{(n,m)}(\vec{r}, \vec{s}) := \prod_{i=1}^n u_+(t, r_i) \prod_{j=1}^m u_-(t, s_j)$$

is a solution of (3.25), where (u_+, u_-) is the weak solution of the coupled PDEs (1.1)-(1.2) with initial value (u_0^+, u_0^-) .

Proof of Theorem 3.5. Recall that $\Xi := \{\xi : \xi_{\pm}(z_{\pm}) \leq 1 \text{ for every } z \in I^\varepsilon\}$. We can rewrite (3.7) as

$$\begin{aligned} \gamma^\varepsilon((\vec{r}, \vec{s}), t) &= \sum_{(\vec{a}, \vec{b})} \gamma^\varepsilon((\vec{a}, \vec{b}), 0) p^\varepsilon(t, (\vec{r}, \vec{s}), (\vec{a}, \vec{b})) m(\vec{a}, \vec{b}) \\ &\quad + \int_0^t \sum_{(\vec{a}, \vec{b}) \notin \Xi} p^\varepsilon(t-s, (\vec{r}, \vec{s}), (\vec{a}, \vec{b})) \mathbb{E}[\mathfrak{K}^\varepsilon A((\vec{a}, \vec{b}), \eta_s)] \varepsilon^{d(n+m)} ds \\ &\quad + \int_0^t \sum_{(\vec{a}, \vec{b}) \in \Xi} p^\varepsilon(t-s, (\vec{r}, \vec{s}), (\vec{a}, \vec{b})) \mathbb{E}[\mathfrak{K}^\varepsilon A((\vec{a}, \vec{b}), \eta_s)] \varepsilon^{d(n+m)} ds. \end{aligned} \quad (3.27)$$

Fix any $(n, m) \in \mathbb{N} \times \mathbb{N}$, $t > 0$ and $(\vec{r}, \vec{s}) \in (D_+^\varepsilon)^n \times (D_-^\varepsilon)^m$. By a simple counting argument and condition (3.24) for η_0^ε , we see that the first term in (3.27) equals

$$\begin{aligned} &\mathbb{E}^{\eta_0^\varepsilon} \left[\sum_{(\vec{a}, \vec{b}) \in (D_+^\varepsilon)^n \times (D_-^\varepsilon)^m} p^\varepsilon(t, (\vec{r}, \vec{s}), (\vec{a}, \vec{b})) \prod_{i=1}^n \eta^+(a_i) \prod_{j=1}^m \eta^+(b_j) \right] + o(N) \\ &= \mathbb{E}^{\eta_0^\varepsilon} \left[\prod_{i=1}^n \langle \mathfrak{X}_0^{N,+}, p^\varepsilon(t, r_i, \cdot) \rangle \prod_{j=1}^m \langle \mathfrak{X}_0^{N,-}, p^\varepsilon(t, s_j, \cdot) \rangle \right] + o(N), \end{aligned}$$

which converges to $\mathbb{E}^{(\vec{r}, \vec{s})}[\Phi^{(n,m)}(X_{(n,m)}(t))]$ by Theorem 2.12 and assumption (i) for the initial distributions. Here $\mathbb{P}^{(\vec{r}, \vec{s})}$ is the probability measure for $X_{(n,m)}$ starting at (\vec{r}, \vec{s}) .

³BBGKY stand for N. N. Bogoliubov, Max Born, H. S. Green, J. G. Kirkwood, and J. Yvon, who derived this type of hierarchy of equations in the 1930s and 1940s in a series papers.

We now prove that the second term in (3.27) tends to 0 as $\varepsilon \rightarrow 0$. The integrand with respect to ds is at most

$$\mathbb{E} \left[\frac{\lambda}{\varepsilon} \sum_{z \in I} \eta_{\theta}^+(z_+) \eta_{\theta}^-(z_-) \sum_{(\vec{a}, \vec{b}) \notin \Xi} p^{\varepsilon}(t - \theta, (\vec{r}, \vec{s}), (\vec{a}, \vec{b})) 2A((\vec{a}, \vec{b}), \eta_{\theta}) \varepsilon^{d(n+m)} \right]. \quad (3.28)$$

Note that $\{(\vec{a}, \vec{b}) \notin \Xi\}$ is a subset of

$$\bigcup_{w \in I} \left[\left(\bigcup_{k=2}^n \{(\vec{a}, \vec{b}) : \vec{a}(w_+) = k\} \right) \cup \left(\bigcup_{\ell=2}^m \{(\vec{a}, \vec{b}) : \vec{b}(w_-) = \ell\} \right) \right], \quad (3.29)$$

and that for fixed $w \in I$ and $k \in \{2, \dots, n\}$, we further have

$$\{(\vec{a}, \vec{b}) : \vec{a}(w_+) = k\} = \bigcup_{\substack{i_1, \dots, i_k \\ \text{distinct}}} \{(\vec{a}, \vec{b}) : a_{i_1} = \dots = a_{i_k} = w_+\}.$$

Now we restrict the sum over $\{(\vec{a}, \vec{b}) \notin \Xi\}$ in (3.28) to the subset $\{(\vec{a}, \vec{b}) : a_{i_1} = \dots = a_{i_k} = w_+\}$, where $w \in I$, $k \in \{2, \dots, n\}$ and (i_1, \dots, i_k) are fixed. Moreover, we denote (a_1, \dots, a_k) by \vec{a}_k and (a_{k+1}, \dots, a_n) by $\vec{a} \setminus \vec{a}_k$. Then

$$\begin{aligned} & \mathbb{E} \left[\frac{\lambda}{\varepsilon} \sum_{z \in I} \eta_{\theta}^+(z_+) \eta_{\theta}^-(z_-) \sum_{\{(\vec{a}, \vec{b}) : a_{i_1} = \dots = a_{i_k} = w_+\}} p^{\varepsilon}(t - \theta, (\vec{r}, \vec{s}), (\vec{a}, \vec{b})) 2A((\vec{a}, \vec{b}), \eta_{\theta}) \varepsilon^{d(n+m)} \right] \\ & \leq p^{\varepsilon}(t - \theta, (r_1, \dots, r_k), (w_+, \dots, w_+)) \varepsilon^{kd} \sum_{(\vec{a} \setminus \vec{a}_k, \vec{b})} p^{\varepsilon}(t - \theta, \vec{r} \setminus \vec{r}_k, \vec{a} \setminus \vec{a}_k) p(t - \theta, \vec{s}, \vec{b}) \varepsilon^{d(n+m-k)} \\ & \quad \cdot \frac{\lambda}{\varepsilon} \mathbb{E} \left[\sum_{z \in I} \eta_{\theta}^+(z_+) \eta_{\theta}^-(z_-) 2A((\vec{a}, \vec{b}), \eta_{\theta}) \right] \\ & \leq \frac{C \varepsilon^{kd}}{(t - \theta)^{kd/2}} \frac{\lambda}{\varepsilon} \#|I^{\varepsilon}| \sup_{(\vec{a}, \vec{b})} \mathbb{E}[\eta_{\theta}^+(z_+) \eta_{\theta}^-(z_-) 2A((\vec{a}, \vec{b}), \eta_{\theta})] \\ & \leq \frac{\lambda \varepsilon^{(k-1)d}}{(t - \theta)^{kd/2}} C \quad \text{where } C = C(n, m, \theta, d, D_{\pm}) \\ & \leq \frac{\lambda \varepsilon^d}{(t - \theta)^{kd/2}} C = O(\varepsilon^d) \quad \text{since } k \geq 2. \end{aligned}$$

The second to the last inequality above follows from the bound $\#|I^{\varepsilon}| \leq C \varepsilon^{-(d-1)}$ (see Lemma 2.4) and the uniform upper bound (3.17). Repeat the above argument for the other subsets of $\{(\vec{a}, \vec{b}) \notin \Xi\}$ and use the fact $\#|I^{\varepsilon}| \leq C \varepsilon^{-(d-1)}$ again (for $w \in I$ in (3.29)), we have, for any $\theta \in (0, t)$, (3.28) is of order ε and hence converges to 0 uniformly for (\vec{r}, \vec{s}) , as $\varepsilon \rightarrow 0$. The second term in (3.27) then converges to 0, by (3.21) and LDCT.

For the third term in (3.27), we split the integrand with respect to $d\theta$ into three terms corresponding to (3.13), (3.14) and (3.15) respectively. The term corresponding to (3.13) equals

$$\begin{aligned} & -\frac{\lambda}{\varepsilon} \sum_{(\vec{a}, \vec{b}) \in \Xi} p^{\varepsilon}(t - s, (\vec{r}, \vec{s}), (\vec{a}, \vec{b})) \sum_{\substack{z \in I^{\varepsilon} \\ \vec{a}(z_+) = 1}} \Psi(z) A((\vec{a}, (\vec{b}, z_-)), \eta_{\theta}) \varepsilon^{d(n+m)} \\ & = -\frac{\lambda}{\varepsilon} \sum_{z \in I^{\varepsilon}} \sum_{\substack{(\vec{a}, \vec{b}) \in \Xi \\ \vec{a}(z_+) = 1}} p^{\varepsilon}(t - s, (\vec{r}, \vec{s}), (\vec{a}, \vec{b})) \Psi(z) A((\vec{a}, (\vec{b}, z_-)), \eta_{\theta}) \varepsilon^{d(n+m)} \end{aligned}$$

$$\begin{aligned}
&= -\frac{\lambda}{\varepsilon} \sum_{z \in I^\varepsilon} \Psi(z) \sum_{i=1}^n \sum_{\substack{(\vec{a}, \vec{b}) \in \Xi \\ a_i = z_+}} p^\varepsilon(t-s, (\vec{r}, \vec{s}), (\vec{a}, \vec{b})) A((\vec{a}, (\vec{b}, z_-)), \eta_\theta) \varepsilon^{d(n+m)} \\
&= -\frac{\lambda}{\varepsilon} \sum_{z \in I^\varepsilon} \Psi(z) \sum_{i=1}^n p^\varepsilon(t-s, r_i, z_+) \sum_{(\vec{a} \setminus a_i, \vec{b}) \in \Xi} p^\varepsilon(t-s, (\vec{r} \setminus r_i, \vec{s}), (\vec{a} \setminus a_i, \vec{b})) \\
&\quad \times \frac{m((\vec{a}, (\vec{b}, z_-)))}{\varepsilon^{d(n+m+1)}} \gamma^\varepsilon((\vec{a}, (\vec{b}, z_-)), \theta) \varepsilon^{d(n+m)} \\
&= -\lambda \sum_{i=1}^n \sum_{(\vec{a} \setminus a_i, \vec{b}) \in \Xi} p^\varepsilon(t-s, (\vec{r} \setminus r_i, \vec{s}), (\vec{a} \setminus a_i, \vec{b})) m(\vec{a} \setminus a_i, \vec{b}) \\
&\quad \times \sum_{z \in I^\varepsilon} \sigma_\varepsilon(z) p^\varepsilon(t-s, r_i, z_+) \gamma^\varepsilon((\vec{a}, (\vec{b}, z_-)), \theta).
\end{aligned}$$

By Theorem 2.12 and Lemma 2.4,

$$\lim_{\varepsilon \rightarrow 0} \sum_{z \in I^\varepsilon} \sigma_\varepsilon(z) p^\varepsilon(t-s, r_i, z_+) \gamma^\varepsilon((\vec{a}, (\vec{b}, z_-)), \theta) = \int_I p(t-s, r_i, z) \gamma_\theta(\vec{a}, (\vec{b}, z)) d\sigma(z)$$

and the convergence is uniform for $r_i \in D_+^\varepsilon$. Therefore, by applying Theorem 2.12 again, the term corresponding to (3.13) converges to

$$-\lambda \sum_{i=1}^n \int_{\partial_+^i} \gamma_\theta(\vec{a}, (\vec{b}, a_i)) p(t-\theta, (\vec{r}, \vec{s}), (\vec{a}, \vec{b})) \frac{\rho_{(n,m)}(\vec{r}, \vec{s})}{\rho_+(r_i)} d\vec{b} da_1 \cdots d\sigma(a_i) \cdots da_n.$$

We repeat the same argument for the term corresponding to (3.14). Moreover, note that the term corresponding to (3.15) will not contribute to the limit as $\varepsilon \rightarrow 0$, by the same argument we used for the second term in (3.27). Therefore, the integrand of the second term in (3.27) converges to

$$\begin{aligned}
&-\lambda \sum_{i=1}^n \int_{\partial_+^i} \gamma_\theta(\vec{a}, (\vec{b}, a_i)) p(t-\theta, (\vec{r}, \vec{s}), (\vec{a}, \vec{b})) \frac{\rho_{(n,m)}(\vec{a}, \vec{b})}{\rho_+(a_i)} d\vec{b} da_1 \cdots d\sigma(a_i) \cdots da_n \\
&-\lambda \sum_{j=1}^m \int_{\partial_-^j} \gamma_\theta((\vec{a}, b_j), \vec{b}) p(t-\theta, (\vec{r}, \vec{s}), (\vec{a}, \vec{b})) \frac{\rho_{(n,m)}(\vec{a}, \vec{b})}{\rho_-(b_j)} d\vec{a} db_1 \cdots d\sigma(b_j) \cdots db_m.
\end{aligned}$$

The integral for $\theta \in (0, t)$ in the third term in (3.27) then converges to the desired quantity, by (3.21) and LDCT. The proof is complete. \square

In view of Remark 3.6(iii), the proof of Theorem 2.22 (Propagation of Chaos) will be complete once we establish the uniqueness of the solution of the limiting hierarchy (3.25). This will be accomplished in Theorem 3.7 in the next subsection.

3.5 Uniqueness of infinite hierarchy

Uniqueness of BBGKY hierarchy is an important issue in statistical physics. For example, it is a key step in the derivation of the cubic non-linear Schrödinger equation from the quantum dynamics of many body systems obtained in [21]. Our BBGKY hierarchy (3.25) is new to the literature and the proof of its uniqueness involves a representation and manipulations of the hierarchy in terms of trees. The technique is related but different from that in [21], which used the Feynman diagrams.

Note that, by Theorem 3.5, $\gamma_t^{(n,m)}$ can be extended continuously to $t = 0$. Uniqueness of solution for the hierarchy will be established on a subset of the space

$$C([0, T], \mathcal{D}) := \bigoplus_{(n,m) \in \mathbb{N} \times \mathbb{N}} C([0, T], \overline{D}_+^n \times \overline{D}_-^m)$$

equipped with the product topologies induced by the uniform norm $\|\cdot\|_{(T,n,m)}$ on $[0, T] \times \overline{D}_+^n \times \overline{D}_-^m$.

Theorem 3.7. (*Uniqueness of the infinite hierarchy*) *Given any $T > 0$. Suppose $\beta_t = \{\beta_t^{(n,m)}\} \in C([0, T], \mathcal{D})$ is a solution to the infinite hierarchy (3.25) with zero initial condition (i.e. $\beta_0 = \Phi = 0$) and satisfies $\|\beta_t^{(n,m)}\|_{(T,n,m)} \leq C^{n+m}$ for some $C \geq 0$. Then we have $\|\beta_t^{(n,m)}\|_{(T,n,m)} = 0$ for every $n, m \in \mathbb{N}$.*

The remaining of this subsection is devoted to give a proof of this theorem.

Convention in this subsection: $\beta = \{\beta^{(n,m)}\}$ will always denote the functions stated in Theorem 3.7. For notational simplicity, we will also assume $\lambda = 2$ and $\rho_{\pm} = 1$. The proof for the general case is the same. we will also drop T from the notation $\|\beta_t^{(n,m)}\|_{(T,n,m)}$.

It is convenient to rewrite the infinite hierarchy (3.25) in a more compact form as

$$\gamma_t^{(n,m)} = P_t^{(n,m)} \Phi^{(n,m)} - \int_0^t P_{t-s}^{(n,m)} \left(\sum_{i=1}^n V_i^+ \gamma_s^{(n,m+1)} + \sum_{j=1}^m V_j^- \gamma_s^{(n+1,m)} \right) ds, \quad (3.30)$$

where $V_i^+ \gamma^{(n,m+1)}$ is a measure concentrated on ∂_+^i defined as

$$\begin{aligned} V_i^+ \gamma^{(n,m+1)} &:= \gamma^{(n,m+1)}(\vec{a}, (\vec{b}, a_i)) d\sigma_{(n,m)} \Big|_{\partial_+^i}(\vec{a}, \vec{b}) \\ &= \gamma^{(n,m+1)}(\vec{a}, (\vec{b}, a_i)) d\sigma \Big|_I(a_i) d(\vec{a} \setminus a_i) d\vec{b}. \end{aligned}$$

Here $\sigma_{(n,m)} \Big|_{\partial_+^i}$ is the surface measure of $\partial(D_+^n \times D_-^m)$ restricted to ∂_+^i . Similarly, $V_j^- \gamma^{(n+1,m)}$ is a measure concentrating on ∂_-^j defined as

$$\begin{aligned} V_j^- \gamma^{(n+1,m)} &:= \gamma^{(n+1,m)}((\vec{a}, b_j), \vec{b}) d\sigma_{(n,m)} \Big|_{\partial_-^j}(\vec{a}, \vec{b}) \\ &= \gamma^{(n+1,m)}((\vec{a}, b_j), \vec{b}) d\sigma \Big|_I(b_j) d\vec{a} d(\vec{b} \setminus b_j). \end{aligned}$$

3.5.1 Duhamel tree expansion

We now describe the infinite hierarchy in detail. It is natural and illustrative to represent the infinite hierarchy in terms of a tree structure, with the ‘root’ at the top and the ‘leaves’ at the bottom. Fix two positive integers n and m . We construct a sequence of finite trees $\{\mathbb{T}_N^{(n,m)} : N = 0, 1, 2, \dots\}$ recursively as follows.

1. $\mathbb{T}_0^{(n,m)}$ is the root, with label (n, m) .
2. $\mathbb{T}_1^{(n,m)}$ is constructed from $\mathbb{T}_0^{(n,m)}$ by attaching $n + m$ new vertices (call them leaves of $\mathbb{T}_1^{(n,m)}$) to it. More precisely, we attach $n + m$ new vertices to the root by drawing n ‘+’ edges and m ‘-’ edges from the root. Those new leaves drawn by the ‘+’ edges are labeled $(n, m + 1)$, while those drawn by the ‘-’ edges are labeled $(n + 1, m)$. We also label the edges as $\{+i\}_{i=1}^n$ and $\{-j\}_{j=1}^m$ (See Figure 3).
3. When $N = 2$, we view each of the $n + m$ leaves of $\mathbb{T}_1^{(n,m)}$ as a ‘root’ (with a new label, being either $(n, m + 1)$ or $(n + 1, m)$), and then attach new leaves (leaves of $\mathbb{T}_2^{(n,m)}$) to it by drawing ‘ \pm ’ edges. Hence there are $(m + n)(m + n + 1)$ new leaves, coming from $n^2 + m(n + 1)$ new ‘+’ edges and $n(m + 1) + m^2$ new ‘-’ edges.
4. Having drawn $\mathbb{T}_{N-1}^{(n,m)}$, we construct $\mathbb{T}_N^{(n,m)}$ by attaching new edges and new leaves from each leaf of $\mathbb{T}_{N-1}^{(n,m)}$ by the same construction, viewing a leaf of $\mathbb{T}_{N-1}^{(n,m)}$ as a ‘root’.

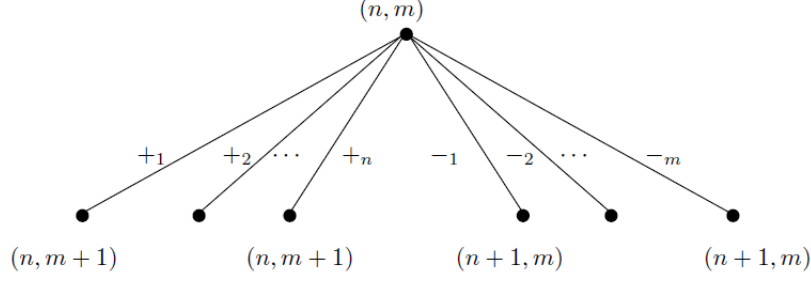


Figure 3: $\mathbb{T}_1^{(n,m)}$

In $\mathbb{T}_N^{(n,m)}$, the root is connected to each leaf by a unique path $\vec{\theta} = (\theta_1, \theta_2, \dots, \theta_N)$ formed by the ‘ \pm ’ edges. Moreover, such a path passes through a sequence of labels formed by the leaves of $\mathbb{T}_k^{(n,m)}$ ($k = 1, 2, \dots, N$). We denote these labels by $\vec{l}(\vec{\theta}) = (l_1(\vec{\theta}), l_2(\vec{\theta}), \dots, l_N(\vec{\theta}))$. For example, when $(n, m) = (2, 5)$, $N = 3$ and the path is $\vec{\theta} = (+1, -6, -5)$. Then $\vec{l}(\vec{\theta}) = ((2, 6), (3, 6), (4, 6))$ and the path connects the root to a leaf of $\mathbb{T}_3^{(2,5)}$ with label $(4, 6)$. Note that the label is not one-to-one. For example, $\vec{l}(+1, -6, -5) = \vec{l}(+2, -6, -4)$.

For mnemonic reason, we use the same notation $\mathbb{T}_N^{(n,m)}$ to denote the collection of paths that connect the root to a leaf in $\mathbb{T}_N^{(n,m)}$. By induction, the total number of paths (or the total number of leaves) is

$$(n+m)(n+m+1)\cdots(n+m+N-1) = \frac{(n+m+N-1)!}{(n+m-1)!}. \quad (3.31)$$

Iterating (3.30) N times gives

$$\begin{aligned} \beta_t^{(n,m)} &= - \int_{t_2=0}^t P_{t-t_2}^{(n,m)} \left(\sum_{i=1}^n V_i^+ \beta_{t_2}^{(n,m+1)} + \sum_{j=1}^m V_j^- \beta_{t_2}^{(n+1,m)} \right) dt_2 \\ &= \dots \\ &= (-1)^N \int_{t_2=0}^t \int_{t_3=0}^{t_2} \dots \int_{t_{N+1}=0}^{t_N} dt_2 \dots dt_{N+1} \\ &\quad \sum_{\vec{\theta} \in \mathbb{T}_N^{(n,m)}} P_{t-t_2}^{(n,m)} V_{\theta_1} P_{t_2-t_3}^{l_1(\vec{\theta})} V_{\theta_2} P_{t_3-t_4}^{l_2(\vec{\theta})} V_{\theta_3} \dots P_{t_N-t_{N+1}}^{l_{N-1}(\vec{\theta})} V_{\theta_N} \beta_{t_{N+1}}^{l_N(\vec{\theta})}, \end{aligned} \quad (3.32)$$

where V_{θ_i} (for $i = 1, 2, \dots, N$) is defined by $V_{+i} = V_i^+$ and $V_{-j} = V_j^-$. For example, when $(n, m) = (2, 5)$, $N = 3$ and the path is $\vec{\theta} = (+1, -6, -5)$. Then

$$P^{(n,m)} V_{\theta_1} P^{l_1(\vec{\theta})} V_{\theta_2} P^{l_2(\vec{\theta})} V_{\theta_3} \beta^{l_3(\vec{\theta})} = P^{(2,5)} V_1^+ P^{(2,6)} V_6^- P^{(3,6)} V_5^- \beta^{(4,6)}.$$

3.5.2 Telescoping via Chapman-Kolmogorov equation

By (3.31), the right hand side of (3.32) is a sum of $(n+m)(n+m+1)\cdots(n+m+N-1)$ terms of multiple integrals. We will apply the bound $\|\beta_t^{(p,q)}\|_{(p,q)} \leq C^{p+q}$ to each term, and then simplify the integrand using Chapman-Kolmogorov equation.

We demonstrate this for the twelve terms for the case $(n, m, N) = (1, 2, 2)$. The twelve terms on the right hand side of (3.32) are

$$P_{t-t_2}^{(1,2)} \left\{ V_1^+ P_{t_2-t_3}^{(1,3)} \left(V_1^+ \beta_{t_3}^{(1,4)} + (V_1^- + V_2^- + V_3^-) \beta_{t_3}^{(2,3)} \right) \right.$$

$$\begin{aligned}
& + V_1^- P_{t_2-t_3}^{(2,2)} \left((V_1^+ + V_2^+) \beta_{t_3}^{(2,3)} + (V_1^- + V_2^-) \beta_{t_3}^{(3,2)} \right) \\
& + V_2^- P_{t_2-t_3}^{(2,2)} \left((V_1^+ + V_2^+) \beta_{t_3}^{(2,3)} + (V_1^- + V_2^-) \beta_{t_3}^{(3,2)} \right) \}. \tag{3.33}
\end{aligned}$$

The first four terms came from the leftmost leaf of the previous level, we group them together to obtain, for $(x, y_1, y_2) \in \overline{D}_+ \times \overline{D}_-^2$,

$$\begin{aligned}
& \left| P_{t-t_2}^{(1,2)} V_1^+ P_{t_2-t_3}^{(1,3)} \left(V_1^+ \beta_{t_3}^{(1,4)} + (V_1^- + V_2^- + V_3^-) \beta_{t_3}^{(2,3)} \right) (x, y_1, y_2) \right| \\
& \leq C^5 \int d\sigma(x') dy'_1 dy'_2 p^{(1,2)}(t-t_2, (x, y_1, y_2), (x', y'_1, y'_2)) \\
& \quad \left(\int d\sigma(a) db_1 db_2 db_3 + \int da d\sigma(b_1) db_2 db_3 + \int da db_1 d\sigma(b_2) db_3 + \int da db_1 db_2 d\sigma(b_3) \right) \\
& \quad p^{(1,3)}(t_2-t_3, (x', y'_1, y'_2, x'), (a, b_1, b_2, b_3)) \\
& = C^5 \int d\sigma(x') p^+(t-t_2, x, x') \left(\int d\sigma(a) p^+(t_2-t_3, x', a) + \int d\sigma(b_1) p^-(t-t_3, y_1, b_1) \right. \\
& \quad \left. + \int d\sigma(b_2) p^-(t-t_3, y_2, b_2) + \int d\sigma(b_3) p^-(t-t_3, x', b_3) \right).
\end{aligned}$$

Note the telescoping effect upon using the Chapman-Kolmogorov equation for the middle two terms in the last equality above gives rise to $t-t_3$ rather than t_2-t_3 .

We apply (2.14) to obtain

$$\begin{aligned}
& \left\| P_{t-t_2}^{(1,2)} V_1^+ P_{t_2-t_3}^{(1,3)} \left(V_1^+ \beta_{t_3}^{(1,4)} + (V_1^- + V_2^- + V_3^-) \beta_{t_3}^{(2,3)} \right) \right\|_{(1,2)} \\
& \leq C^5 \frac{C_+}{\sqrt{t-t_2}} \left(\frac{C_+}{\sqrt{t_2-t_3}} + \frac{2C_-}{\sqrt{t-t_3}} + \frac{C_-}{\sqrt{t_2-t_3}} \right),
\end{aligned}$$

where $C_{\pm} = C(D_{\pm}, T)$ are positive constants. Repeat the above argument for the remaining eight terms of (3.33), we obtain

$$\begin{aligned}
\|\beta_t^{(1,2)}\|_{(1,2)} & \leq C^5 \int_{t_2=0}^t \int_{t_3=0}^{t_2} \frac{C_+}{\sqrt{t-t_2}} \left(\frac{C_+}{\sqrt{t_2-t_3}} + \frac{2C_-}{\sqrt{t-t_3}} + \frac{C_-}{\sqrt{t_2-t_3}} \right) \\
& \quad + \frac{2C_-}{\sqrt{t-t_2}} \left(\frac{C_+}{\sqrt{t-t_3}} + \frac{C_+}{\sqrt{t_2-t_3}} + \frac{C_-}{\sqrt{t-t_3}} + \frac{C_-}{\sqrt{t_2-t_3}} \right). \tag{3.34}
\end{aligned}$$

The key is to visualize the twelve terms on the right as 12 paths of $\mathbb{T}_2^{(1,2)}$ *with the edges relabeled*. We denote this relabeled tree by $\mathbb{S}_2^{(1,2)}$ (See Figure 4, ignoring the five leaves in $\mathbb{S}_3^{(1,2)}$ at the moment). More precisely, since all twelve terms on the right are of the form $\frac{C_{\pm}}{\sqrt{t_p-t_2}} \frac{C_{\pm}}{\sqrt{t_q-t_3}}$, we only need to record the indexes (p, q) and the \pm sign. For example, the first four terms can be represented by

$$(+1+2, +1-1, +1-1, +1-2).$$

Each $+1-1$ corresponds to $\frac{C_+}{\sqrt{t-t_2}} \frac{C_-}{\sqrt{t-t_3}}$ and hence it appears twice. In $\mathbb{S}_2^{(1,2)}$, these four paths are formed by a $+1$ edge followed by four edges with labels $\{+2, -1, -1, -2\}$.

In general, we obtain $\mathbb{S}_N^{(n,m)}$ by relabeling the edges of $\mathbb{T}_N^{(n,m)}$, while keeping the labels for the vertices and the \pm sign for the edges. The relabeling of edges are performed as follows:

1. At level 1, we assign the number ‘1’ to all the edges connected to the root. Hence we have n ‘+1’ edges and m ‘-1’ edges, rather than the labels $\{+i\}_{i=1}^n$ and $\{-j\}_{j=1}^m$ (See Figure 5 for $\mathbb{S}_1^{(n,m)}$).

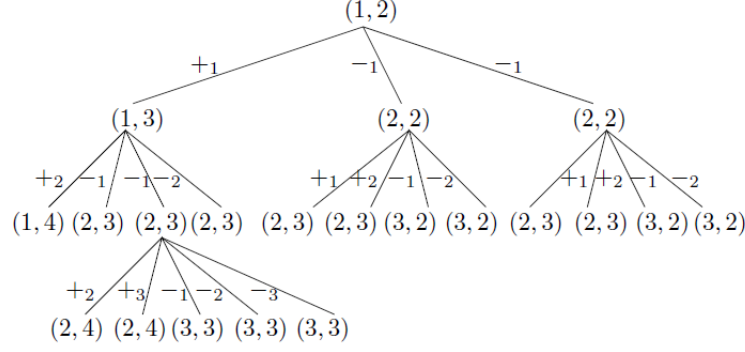


Figure 4: $\mathbb{S}_2^{(1,2)}$ together with 5 leaves in $\mathbb{S}_3^{(1,2)}$

2. At level $k \geq 2$, consider the set $\Lambda^+ := \{+1, \dots, +1, +2, +3, \dots, +k\}$ in which we have n copies of $+1$ (hence there are $n + k - 1$ elements in Λ^+ , in which n of them are $+1$). Similarly, let $\Lambda^- := \{-1, \dots, -1, -2, -3, \dots, -k\}$ in which we have m copies of -1 . For an arbitrary leaf ξ of $\mathbb{T}_{k-1}^{(n,m)}$, let R^ξ be the labels of (the edges of) the path from the root to ξ in $\mathbb{S}_{k-1}^{(n,m)}$, counting with multiplicity. Finally, the collection of new labels of the edges below ξ , denoted by L^ξ , is chosen in such a way that

$$\Lambda^+ \cup \Lambda^- = R^\xi \cup L^\xi \quad (\text{counting multiplicity}).$$

Since $|R^\xi| = k - 1$ and $|L^\xi| = n + m + k - 1$ (again, counting multiplicity), the cardinalities of the two sides match:

$$(n + k - 1) + (m + k - 1) = (k - 1) + (n + m + k - 1).$$

Induction shows that $R^\xi \subset \Lambda^+ \cup \Lambda^-$ and the choice for L^ξ is unique. For example, for leaf $\xi = (1, 3)$ of $\mathbb{T}_1^{(1,2)}$, $R^\xi = \{+1\}$, $\Lambda^+ := \{+1, +2\}$ and $\Lambda^- := \{-1, -2\}$. So $L^\xi = \{+2, -1, -1, -2\}$, which gives the new labels to the edges below ξ ; see Figure 4.

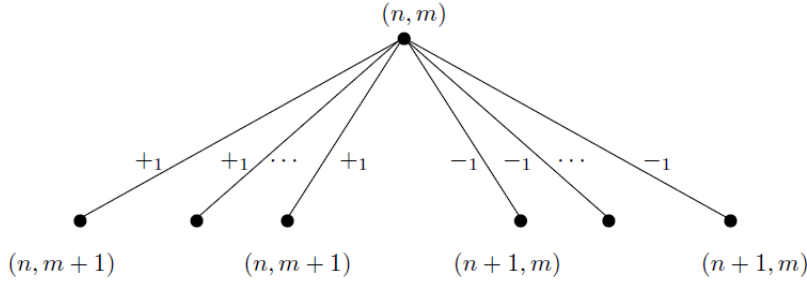


Figure 5: $\mathbb{S}_1^{(n,m)}$

As a further illustration, we continue to ‘grow’ $\mathbb{S}_2^{(1,2)}$ (see Figure 4) by adding suitably labeled edges to leaves of $\mathbb{S}_2^{(1,2)}$. Precisely, let ξ be a leaf of $\mathbb{S}_2^{(1,2)}$.

- If $R^\xi = \{+1, +2\}$, then $L^\xi = \{+3, -1, -1, -2, -3\}$ (this is the case for the leftmost leaf, which has label $(1, 4)$)
- If $R^\xi = \{+1, -1\}$, then $L^\xi = \{+2, +3, -1, -2, -3\}$ (shown in Figure 4).
- If $R^\xi = \{+1, -2\}$, then $L^\xi = \{+2, +3, -1, -1, -3\}$.
- If $R^\xi = \{-1, +1\}$, then $L^\xi = \{+2, +3, -1, -2, -3\}$.

- If $R^\xi = \{-1, +2\}$, then $L^\xi = \{+1, +3, -1, -2, -3\}$.
- If $R^\xi = \{-1, -1\}$, then $L^\xi = \{+1, +2, +3, -2, -3\}$.
- If $R^\xi = \{-1, -2\}$, then $L^\xi = \{+1, +2, +3, -1, -3\}$.

For mnemonic reason, we use the same notation $\mathbb{S}_N^{(n,m)}$ to denote the collection of paths from the root to the leaves of $\mathbb{S}_N^{(n,m)}$. Any such path is represented by the *ordered* (new) labels of the edges. We now ‘forget’ the sign of the edges and only record the integer labels. For example, the path $(-1, +1, -2, +3)$ is replaced by $(1, 1, 2, 3)$.

Using the hypothesis $\|\beta_t^{(n,m)}\|_{(n,m)} \leq C^{n+m}$ (of Theorem 3.7) and applying Chapman-Kolmogorov equation to (3.32), and then applying (2.14), we obtain the following lemma by the same argument that we used to obtain (3.34).

Lemma 3.8.

$$\|\beta_t^{(n,m)}\|_{(n,m)} \leq C^{n+m+N} (C_+ \vee C_-)^N I_N^{(n,m)}(t), \quad (3.35)$$

where

$$I_N^{(n,m)}(t) := \int_{t_2=0}^t \cdots \int_{t_{N+1}=0}^{t_N} \sum_{\vec{v} \in \mathbb{S}_N^{(n,m)}} \frac{1}{\sqrt{(t_{v_1} - t_2)(t_{v_2} - t_3) \cdots (t_{v_N} - t_{N+1})}} dt_2 \cdots dt_{N+1}. \quad (3.36)$$

Our goal is to show that $I_N^{(n,m)}(t) \leq (Ct)^{N/2}$ for some $C = C(n, m) > 0$. This will imply $\|\beta_t^{(n,m)}\|_{(n,m)} = 0$ for $t > 0$ small enough. Clearly we have

$$\begin{aligned} I_N^{(n,m)}(t) &\leq \frac{(n+m+N-1)!}{(n+m-1)!} \int_{t_2=0}^t \cdots \int_{t_{N+1}=0}^{t_N} \frac{1}{\sqrt{(t-t_2)(t_2-t_3) \cdots (t_N-t_{N+1})}} dt_2 \cdots dt_{N+1} \\ &= \frac{(n+m+N-1)!}{(n+m-1)!} \frac{(\pi t)^{N/2}}{\Gamma(\frac{N+2}{2})}. \end{aligned}$$

Unfortunately, this crude bound is asymptotically larger than $(Ct)^{N/2}$ for any $C > 0$.

3.5.3 Comparison with a ‘dominating’ tree

Note that

$$\begin{aligned} I_3^{(1,2)}(t) &\leq \int_{t_2=0}^t \int_{t_3=0}^{t_2} \int_{t_4=0}^{t_3} \left(\frac{3}{\sqrt{t-t_2}} \right) \left(\frac{2}{\sqrt{t_2-t_3}} + \frac{2}{\sqrt{t-t_3}} \right) \\ &\quad \left(\frac{2}{\sqrt{t_3-t_4}} + \frac{2}{\sqrt{t_2-t_4}} + \frac{1}{\sqrt{t-t_4}} \right) dt_2 dt_3 dt_4. \end{aligned}$$

This is obtained by comparing the labels in $\mathbb{S}_3^{(1,2)}$ with a ‘dominating’ labeling, in which the labels of the edges below *every* leaf of $\mathbb{S}_2^{(1,2)}$ is $\{3, 3, 2, 2, 1\}$ (the \pm sign is discarded). This trick enables us to group the terms at each level.

For the general case, let ξ be an arbitrary leaf of $\mathbb{S}_{k-1}^{(n,m)}$. Note that in L^ξ , each of the integers $2, 3, \dots, k$ appears at most twice and the integer 1 appears at most $n+m$ times. We compare L^ξ with the ‘dominating’ label \tilde{L}^ξ defined below:

Level k	$ \tilde{L}^k = L^k $	\tilde{L}^k
1	$n + m$	$1, 1, 1, \dots, 1$
2	$n + m + 1$	$2, 2, 1, 1, 1, \dots, 1$
3	$n + m + 2$	$3, 3, 2, 2, 1, 1, 1, \dots, 1$
\dots	\dots	\dots
$n + m - 1$	$2(n + m) - 2$	$n + m - 1, n + m - 1, n + m - 2, \dots, 3, 3, 2, 2, 1, 1$
$n + m$	$2(n + m) - 1$	$n + m, n + m, n + m - 1, n + m - 1, \dots, 3, 3, 2, 2, 1$
$n + m + 1$	$2(n + m)$	$n + m + 1, n + m + 1, n + m, n + m, \dots, 3, 3, 2, 2$
$n + m + 2$	$2(n + m) + 1$	$n + m + 2, n + m + 2, n + m + 1, \dots, 3, 3, 2$
\dots	\dots	\dots
N	$n + m + N - 1$	$N, N, N - 1, N - 1, \dots, c, b, a$

In the last row, if $n + m + N - 1$ is even, then $a = b = (N - n - m + 3)/2$ and $c = b + 1$; if $n + m + N - 1$ is odd, then $a = (N - n - m + 2)/2$ and $b = c = a + 1$.

We can now group the terms in each level k as a sum of k terms to obtain

$$\begin{aligned}
I_N^{(n,m)}(t) &\leq \int_{t_2=0}^t \int_{t_3=0}^{t_2} \dots \int_{t_{N+1}=0}^{t_N} \left(\frac{n+m}{\sqrt{t-t_2}} \right) \left(\frac{2}{\sqrt{t_2-t_3}} + \frac{n+m-1}{\sqrt{t-t_3}} \right) \\
&\quad \left(\frac{2}{\sqrt{t_3-t_4}} + \frac{2}{\sqrt{t_2-t_4}} + \frac{n+m-2}{\sqrt{t-t_4}} \right) \dots \\
&\quad \left(\frac{2}{\sqrt{t_{n+m}-t_{n+m+1}}} + \dots + \frac{2}{\sqrt{t_2-t_{n+m+1}}} + \frac{1}{\sqrt{t-t_{n+m+1}}} \right) \\
&\quad \prod_{k=n+m+1}^N \left(\frac{2}{\sqrt{t_k-t_{k+1}}} + \frac{2}{\sqrt{t_{k-1}-t_{k+1}}} + \dots + \frac{2}{\sqrt{t_2-t_{k+1}}} \right) dt_2 \dots dt_{N+1}.
\end{aligned}$$

In the last term, we have used the observation that when $k > n + m$, the smallest element in \tilde{L}^k is at least 2 and so the sum stops before reaching $1/\sqrt{t-t_{k+1}}$. From this and the simple estimates like

$$\frac{2}{\sqrt{t_3-t_4}} + \frac{2}{\sqrt{t_2-t_4}} + \frac{n+m-2}{\sqrt{t-t_4}} \leq \frac{2(n+m)}{3} \left(\frac{1}{\sqrt{t_3-t_4}} + \frac{1}{\sqrt{t_2-t_4}} + \frac{1}{\sqrt{t-t_4}} \right),$$

we have derived the following

Lemma 3.9.

$$I_N^{(n,m)}(t) \leq \frac{(n+m)^{(n+m)}}{(n+m)!} 2^N J_N(t), \tag{3.37}$$

where

$$J_N(t) := \int_0^t \int_0^{t_2} \dots \int_0^{t_N} \prod_{i=2}^{N+1} \left(\sum_{j=1}^{i-1} \frac{dt_i}{\sqrt{t_j-t_i}} \right). \tag{3.38}$$

3.5.4 Estimating J_N

Our goal in this section is show that $J_N(t) \leq (Ct)^{N/2}$ for some $C > 0$. Our proof relies on the following recursion formula pointed out to us by David Speyer:

$$J_N(t) = \sum_{k=1}^N \frac{1}{k!} \sum_{\substack{n_2+n_3+\dots+n_{k+1}=N \\ n_2, n_3, \dots, n_{k+1} \geq 1}} \prod_{j=2}^{k+1} \int_0^t \frac{J_{n_{j-1}}(t_j)}{\sqrt{t-t_j}} dt_j. \tag{3.39}$$

We assume (3.39) for now and use it to establish the following lemma. The proof of (3.39) will be given immediately after it.

Lemma 3.10. $J_N(t)$ is homogeneous in the sense that

$$J_N(t) = J_N \cdot t^{N/2} \quad \text{where } J_N := J_N(1). \quad (3.40)$$

Moreover,

$$2^N \leq J_N \leq \frac{(N+1)^N \pi^{N/2}}{(N+1)!}. \quad (3.41)$$

Proof $J_N(t) = J_N \cdot t^{N/2}$ is obvious from (3.39) after a change of variable. Let \mathcal{M}_N be the collection of functions $f : \{2, 3, \dots, N+1\} \rightarrow \{1, 2, \dots, N\}$ satisfying $f(i) < i$. We can rewrite $J_N(t_1)$ as

$$J_N(t_1) = \sum_{f \in \mathcal{M}_N} \int_{t_2=0}^{t_1} \int_{t_3=0}^{t_2} \cdots \int_{t_{N+1}=0}^{t_N} \prod_{i=2}^{N+1} \frac{dt_i}{\sqrt{t_{f(i)} - t_i}}. \quad (3.42)$$

This is a sum of $N!$ terms. When we put $t_1 = 1$, the smallest term is

$$\int_{t_2=0}^1 \int_{t_3=0}^{t_2} \cdots \int_{t_{N+1}=0}^{t_N} \prod_{j=2}^{N+1} \frac{dt_j}{\sqrt{1 - t_j}} = \frac{1}{N!} \left(\int_0^1 \frac{1}{\sqrt{1-s}} ds \right)^N = \frac{2^N}{N!}.$$

Hence we have the lower bound $2^N \leq J_N$. Unfortunately, the largest term is exactly

$$\int_{t_2=0}^1 \int_{t_3=0}^{t_2} \cdots \int_{t_{N+1}=0}^{t_N} \prod_{i=1}^N \frac{dt_j}{\sqrt{t_{j-1} - t_j}} = \frac{\pi^{N/2}}{\Gamma(\frac{N+2}{2})}.$$

which grows faster than C^N for any $C \in (0, \infty)$. Hence for the upper bound, we will employ the recursion formula (3.39). We apply the homogeneity to the right hand side of (3.39) to obtain

$$J_N = \sum_{k=1}^N \frac{1}{k!} \sum_{\substack{n_2+n_3+\dots+n_{k+1}=N \\ n_2, \dots, n_{k+1} \geq 1}} \prod_{j=2}^{k+1} J_{n_j-1} \int_0^1 \frac{t_j^{(n_j-1)/2}}{\sqrt{1-t_j}} dt_j.$$

The integrals are now simple one dimensional and can be evaluated:

$$J_N = \sum_{k=1}^N \frac{1}{k!} \sum_{\substack{n_2+n_3+\dots+n_{k+1}=N \\ n_2, \dots, n_{k+1} \geq 1}} \prod_{j=2}^{k+1} \left(J_{n_j-1} \cdot \frac{\sqrt{\pi} \Gamma((n_j+1)/2)}{\Gamma((n_j+2)/2)} \right). \quad (3.43)$$

Since $\frac{\sqrt{\pi} \Gamma((n_j+1)/2)}{\Gamma((n_j+2)/2)} \leq \sqrt{\pi}$, we have $J_N \leq K_N$, where K_N is defined by the recursion

$$K_N = \sum_{k=1}^N \frac{1}{k!} \sum_{\substack{n_2+n_3+\dots+n_{k+1}=N \\ n_2, \dots, n_{k+1} \geq 1}} \prod_{j=2}^{k+1} (K_{n_j-1} \cdot \sqrt{\pi})$$

with $K_0 = 1$. The generating function $\phi(x) := \sum_{N=0}^{\infty} K_N x^N$ of K_N clearly satisfies $\phi(x) = \exp(\sqrt{\pi} x \phi(x))$. We thus see that $\phi(x) = W(-\sqrt{\pi} x) / (-\sqrt{\pi} x)$, where W is the Lambert W function. By Lagrange inversion Theorem (see Theorem 5.4.2 of [37]), $W(z) = \sum_{k=1}^{\infty} (-k)^{k-1} z^k / k!$ (for $|z| < 1/e$). Hence by comparing coefficients in the series expansion of $\phi(x)$, we have $K_N = \frac{(N+1)^N \pi^{N/2}}{(N+1)!}$ as desired. \square

Remark 3.11. By Stirling's formula, $\frac{(N+1)^N \pi^{N/2}}{(N+1)!} \sim (\sqrt{\pi} e)^N$ (where $a(N) \sim b(N)$ means $\lim_{N \rightarrow \infty} \frac{a(N)}{b(N)} = 1$). Hence $J_N \leq C^N$ for some $C > 0$. Monte Carlo simulations suggests that $J_N \sim \pi^N$. The recursion (3.43) also makes it clear that J_N 's are all in $\mathbb{Q}[\pi]$ (polynomials in π with rational coefficients) and makes it easy to compute them recursively. This is because $\frac{\sqrt{\pi} \Gamma((n_j+1)/2)}{\Gamma((n_j+2)/2)}$ is rational if n_j is odd and is a rational multiple of π if n_j is even. For example, $J_1 = 2$, $J_2 = 2 + \pi$ and $J_3 = 4 + \frac{10\pi}{3}$. \square

We now turn to the proof the recursion formula (3.39) which is restated in the following lemma.

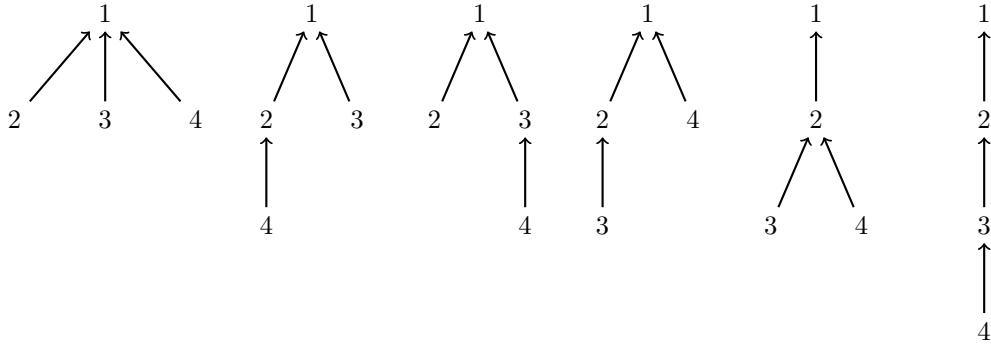
Lemma 3.12.

$$J_N(t) = \sum_{k=1}^N \frac{1}{k!} \sum_{\substack{n_2+n_3+\dots+n_{k+1}=N \\ n_2, n_3, \dots, n_{k+1} \geq 1}} \prod_{j=2}^{k+1} \int_0^t \frac{J_{n_j-1}(t_j)}{\sqrt{t-t_j}} dt_j \quad (3.44)$$

provided that we set $J_0(t) = 1$.

Proof The proof is based on standard combinatorial methods for working on sums over planar rooted trees (see [37]).

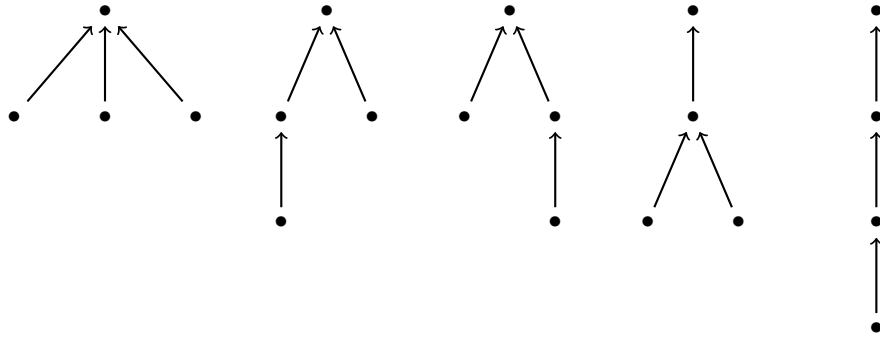
Step 1: Summing over labeled trees. Recall (3.42). There are $N!$ elements in \mathcal{M}_N . We can visualize each of them as a rooted tree with vertex set $\{1, 2, \dots, N+1\}$ and a directed edge from i to $f(i)$ for each i . For example, the 6 elements of \mathcal{M}_3 can be represented by



The trees are drawn so that arrows point upwards and the children of a given vertex are listed from left to right. Note that the second and fourth tree of the list are the same up to relabeling the vertices. The idea is to group terms in (3.42) like this together. First, we rewrite (3.42) in terms of trees. Let \mathcal{D}_N be the set of ‘decreasing trees’, which are trees whose vertices are labeled by $\{1, 2, \dots, N+1\}$ and such that $i < j$ whenever there is an edge $i \leftarrow j$. Then

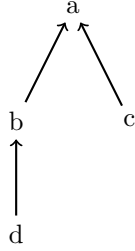
$$J_N(t_1) = \sum_{T \in \mathcal{D}_N} \int_{t_1 \geq t_2 \geq \dots \geq t_{N+1} \geq 0} \prod_{(i \leftarrow j) \in \text{Edge}(T)} \frac{dt_j}{\sqrt{t_i - t_j}}. \quad (3.45)$$

Step 2: Summing over unlabeled trees. A planar tree is a rooted unlabeled tree where, for each vertex, the children of that vertex are ordered. We draw a planar tree so that its children are ordered from left to right. Here are the 5 planar rooted trees on 3+1 vertices:



Let \mathcal{T}_k be the set of planar rooted trees with k vertices. In general, there are $\frac{(2N)!}{N!(N+1)!}$ (the Catalan number) elements in \mathcal{T}_{N+1} , see exercise 6.19 in [37]. We now group all the integrals in (3.45) with the

same planar tree. For example, two different labeled trees (the second and forth in our list of labeled trees) both give the same unlabeled planar tree (which is the second in the above list). We redraw this unlabeled planar tree T_0 below and attach letters $\{a, b, c, d\}$ to T_0 for later use.



The integrands corresponding to the second and forth labeled trees are

$$\frac{dt_2 dt_3 dt_4}{\sqrt{(t_1 - t_2)(t_2 - t_4)(t_1 - t_3)}} \quad \text{and} \quad \frac{dt_2 dt_3 dt_4}{\sqrt{(t_1 - t_2)(t_2 - t_3)(t_1 - t_4)}}.$$

They are the same to

$$\frac{dt_b dt_c dt_d}{\sqrt{(t_a - t_b)(t_a - t_c)(t_b - t_d)}},$$

once we relabel the variables by the vertices of T_0 . That is, $(1, 2, 3, 4) \rightarrow (a, b, c, d)$ for the first term and $(1, 2, 3, 4) \rightarrow (a, b, d, c)$ for the second.

We now go back and keep track of the bounds of integration. In the first integral, they are $t_a \geq t_b \geq t_c \geq t_d$ and, in the second integral, they are $t_a \geq t_b \geq t_d \geq t_c$. We can group these together as

$$t_a \geq t_b, \quad t_a \geq t_c, \quad t_b \geq t_d, \quad t_b > t_c,$$

which is the same as $t_a \geq t_b \geq t_c$ and $t_b \geq t_d$.

In general, the inequality constraints we have are of two types. First, whenever we have an edge $u \leftarrow v$, we get the inequality $t_u \geq t_v$. Second, if v and w are children of u with v to the left of w , then $t_v \geq t_w$. Let $P(T, t_1)$ be the polytope cut out by these inequalities where t_1 is the variable at the root. We have proved

$$J_N(t_1) = \sum_{T \in \mathcal{T}_{N+1}} \int_{P(T, t_1)} \prod_{(u \leftarrow v) \in \text{Edge}(T)} \frac{dt_v}{\sqrt{t_u - t_v}}, \quad (3.46)$$

Step 3: Grouping terms for which the root has degree k . We abbreviate

$$\omega(T, t_1) := \prod_{(u \leftarrow v) \in \text{Edge}(T)} \frac{dt_v}{\sqrt{t_u - t_v}}$$

if T has more than one vertex (otherwise $\omega(T, t_1) := 1$). Then (3.46) translates into

$$J_N(t_1) = \sum_{T \in \mathcal{T}_{N+1}} \int_{P(T, t_1)} \omega(T, t_1). \quad (3.47)$$

Fix an integer k and let T be a tree whose root has degree k . Removing the root leaves behind k children, denoted in chronological order by t_2, \dots, t_{k+1} , and k planar subtrees T_j having t_j as its root for $2 \leq j \leq k+1$. Then

$$\int_{P(T, t_1)} \omega(T, t_1) = \int_{t_1 \geq t_2 \geq \dots \geq t_{k+1} \geq 0} \prod_{j=2}^{k+1} \frac{dt_j}{\sqrt{t_1 - t_j}} \int_{P(T_j, t_j)} \omega(T_j, t_j).$$

Hence, group together the terms where the root has degree k , we have

$$J_N(t_1) = \sum_{k=1}^N \sum_{T_2, \dots, T_{k+1}} \int_{t_1 \geq t_2 \geq \dots \geq t_{k+1} \geq 0} \prod_{j=2}^{k+1} \frac{dt_j}{\sqrt{t_1 - t_j}} \int_{P(T_j, t_j)} \omega(T_j, t_j). \quad (3.48)$$

Here the summation conditions include that $\sum_{j=2}^{k+1} |T_j| = N$ and each different ordering of $(T_2, T_3, \dots, T_{k+1})$ are considered to be different, where $|T_j|$ is the number of vertices in T_j . This abbreviation applies whenever $\sum_{T_2, \dots, T_{k+1}}$ appears.

On other hand, we have by applying (3.47) to each $J_{n_j-1}(t_j)$ below that

$$\begin{aligned} & \sum_{k=1}^N \frac{1}{k!} \sum_{\substack{n_2+n_3+\dots+n_{k+1}=N \\ n_2, \dots, n_{k+1} \geq 1}} \prod_{j=2}^{k+1} \int_0^{t_1} \frac{J_{n_j-1}(t_j)}{\sqrt{t_1 - t_j}} dt_j \\ &= \sum_{k=1}^N \frac{1}{k!} \sum_{\substack{n_2+n_3+\dots+n_{k+1}=N \\ n_2, \dots, n_{k+1} \geq 1}} \prod_{j=2}^{k+1} \sum_{|T_j|=n_j} \int_0^{t_1} \frac{dt_j}{\sqrt{t_1 - t_j}} \int_{P(T_j, t_j)} \omega(T_j, t_j) \\ &= \sum_{k=1}^N \frac{1}{k!} \sum_{T_2, \dots, T_{k+1}} \prod_{j=2}^{k+1} \int_0^{t_1} \frac{dt_j}{\sqrt{t_1 - t_j}} \int_{P(T_j, t_j)} \omega(T_j, t_j) \\ &= \sum_{k=1}^N \frac{1}{k!} \sum_{T_2, \dots, T_{k+1}} \int_{[0, t_1]^{k+1}} \prod_{j=2}^{k+1} \frac{dt_j}{\sqrt{t_1 - t_j}} \int_{P(T_j, t_j)} \omega(T_j, t_j). \end{aligned} \quad (3.49)$$

Step 4: Identifying the integrals. It remains to show that (3.48) is equal to (3.49). Let \mathcal{S}_k denote the space of permutations of $\{2, 3, \dots, k+1\}$. Then the right hand side of (3.49) is equal to

$$\begin{aligned} & \sum_{k=1}^N \frac{1}{k!} \sum_{T_2, \dots, T_{k+1}} \sum_{\sigma \in \mathcal{S}_k} \int_{t_1 \geq t_{\sigma(2)} \geq \dots \geq t_{\sigma(k+1)} \geq 0} \prod_{j=2}^{k+1} \frac{dt_{\sigma(j)}}{\sqrt{t_1 - t_{\sigma(j)}}} \int_{P(T_{\sigma(j)}, t_{\sigma(j)})} \omega(T_{\sigma(j)}, t_{\sigma(j)}) \\ &= \sum_{k=1}^N \frac{1}{k!} \sum_{\sigma \in \mathcal{S}_k} \sum_{T_2, \dots, T_{k+1}} \int_{t_1 \geq s_2 \geq \dots \geq s_{k+1} \geq 0} \prod_{j=2}^{k+1} \frac{ds_j}{\sqrt{t_1 - s_j}} \int_{P(T_{\sigma(j)}, s_j)} \omega(T_{\sigma(j)}, s_j) \\ &= \sum_{k=1}^N \frac{1}{k!} \sum_{\sigma \in \mathcal{S}_k} \sum_{T_2, \dots, T_{k+1}} \int_{t_1 \geq s_2 \geq \dots \geq s_{k+1} \geq 0} \prod_{j=2}^{k+1} \frac{ds_j}{\sqrt{t_1 - s_j}} \int_{P(T_j, s_j)} \omega(T_j, s_j) \\ &= \sum_{k=1}^N \sum_{T_2, \dots, T_{k+1}} \int_{t_1 \geq t_2 \geq \dots \geq t_{k+1} \geq 0} \prod_{j=2}^{k+1} \frac{dt_j}{\sqrt{t_1 - t_j}} \int_{P(T_j, t_j)} \omega(T_j, t_j), \end{aligned}$$

which is $J_N(t)$ by (3.48). This completes the proof of the lemma. \square

3.5.5 Proof of uniqueness

Proof of Theorem 3.7. By Lemma 3.8, Lemma 3.9 and Lemma 3.10, we have

$$\|\beta_t^{(n,m)}\|_{(n,m)} \leq C_1(n,m) C_2(D_+, D_-, T)^N t^{N/2} \quad (3.50)$$

for all $t \in [0, T]$ and $N \in \mathbb{N}$. This implies that there is a constant $\tau > 0$ so that $\|\beta_t^{(n,m)}\|_{(n,m)} = 0$ for $t \leq \tau$ and for all $(n, m) \in \mathbb{N} \times \mathbb{N}$. Note that $\tilde{\beta}_t := \beta_{\tau+t}$ also satisfies the hierarchy (3.30), and that $\tilde{\beta}_0 = 0$. Using the hypothesis $\|\beta_t^{(n,m)}\|_{(T,n,m)} \leq C^{n+m}$, we can extend to obtain $\|\beta_t^{(n,m)}\|_{(n,m)} = 0$ for $t \in [0, T]$. \square

4 Hydrodynamic Limits

This section is devoted to the proof of Theorem 2.20. Throughout this section, we assume the conditions of Theorem 2.20 hold.

4.1 Constructing martingales

Since $\eta_t = \eta_t^\varepsilon$ has a finite state space, we know that for all bounded function $F : \mathbb{R}_+ \times E^\varepsilon \rightarrow \mathbb{R}$ that is smooth in the first coordinate with $\sup_{(s,x)} \left| \frac{\partial F}{\partial s}(s,x) \right| < C < \infty$, we have two \mathfrak{F}_t^η -martingales below:

$$M(t) := F(t, \eta_t) - F(0, \eta_0) - \int_0^t \frac{\partial F}{\partial s}(s, \eta_s) + \mathfrak{L}F(s, \cdot)(\eta_s) ds \quad (4.1)$$

and

$$N(t) := M(t)^2 - \int_0^t \mathfrak{L}(F^2(s, \cdot))(\eta_s) - 2F(s, \eta_s) \mathfrak{L}F(s, \cdot)(\eta_s) ds, \quad (4.2)$$

where $\mathfrak{L} = \mathfrak{L}^\varepsilon$ is the generator defined in (2.8). See Lemma 5.1 (p.330) of [27] or Proposition 4.1.7 of [22] for a proof. We will use this fact to construct some important martingales in Lemma 4.1 below.

In general, suppose $X = (X_t)_{t \geq 0}$ is a CTRW in a finite state space E , whose one step transition probability is p_{xy} and mean holding time at x is $h(x)$. Its infinitesimal generator of X is the discrete operator

$$\mathcal{A}f(x) := \frac{1}{h(x)} \sum_{y \in E} p_{xy}(f(y) - f(x)).$$

The formal adjoint \mathcal{A}^* of \mathcal{A} is defined by

$$\mathcal{A}^*f(x) := \sum_{y \in E} \left(\frac{1}{h(y)} p_{yx}f(y) - \frac{1}{h(x)} p_{xy}f(x) \right).$$

It can be easily checked that

$$\langle f, \mathcal{A}^*g \rangle = \langle \mathcal{A}f, g \rangle, \quad \text{where } \langle f, g \rangle := \frac{1}{N} \sum_{x \in E} f(x)g(x). \quad (4.3)$$

We denote by $\mathcal{A}_\varepsilon^\pm$ the generator of the CTRW $X^{\pm, \varepsilon}$ on D_\pm^ε , respectively, and by $\mathcal{A}_\varepsilon^{*, \pm}$ the corresponding formal adjoint. In our case, $h(x) = h_\varepsilon(x) = \varepsilon^2/d$ for all x . We can check that if $f \in C^2(D_\pm)$, then

$$\lim_{\varepsilon \rightarrow 0} \mathcal{A}_\varepsilon^\pm f(x^\varepsilon) = \mathcal{A}^\pm f(x) \text{ whenever } x^\varepsilon \in D_\pm^\varepsilon \text{ converges to } x \in D_\pm. \quad (4.4)$$

Lemma 4.1. *For any $\phi \in \mathcal{B}_b(D_+)$,*

$$M(t) := M_\phi^{+, N}(t) := \langle \phi, \mathfrak{X}_t^{N, +} \rangle - \langle \phi, \mathfrak{X}_0^{N, +} \rangle - \int_0^t \langle \mathcal{A}_\varepsilon^+ \phi, \mathfrak{X}_s^{N, +} \rangle ds + \lambda \int_0^t \langle J_s^{N, +}, \phi \rangle ds \quad (4.5)$$

is an \mathfrak{F}_t^η -martingale for $t \geq 0$, where $J^{N, +}$ is the measure-valued process defined by (2.35). Moreover, if $\phi \in C^1(\overline{D}_+)$, then there is a constant $C > 0$ independent of N so that for every $T > 0$,

$$\mathbb{E} \left[\sup_{t \in [0, T]} M^2(t) \right] \leq \frac{CT}{N}. \quad (4.6)$$

Similar statements hold for $\mathfrak{X}^{N, -}$.

Proof The lemma follows by applying (4.1) and (4.2) to the function

$$F(s, \eta) := f(\eta) := \frac{1}{N} \sum_{x \in D_+} \phi^+(x) \eta^+(x) \quad , (s, \eta) \in [0, \infty) \times E^\varepsilon.$$

We spell out the details here for completeness. Observe that $f(\eta_t) = \langle \phi, \mathfrak{X}_t^{N,+} \rangle$. Fix $x_0 \in D_+^\varepsilon$, and define $\eta_{x_0}^+$ to be the function from E^ε to \mathbb{R} which maps η to $\eta^+(x_0)$. Then by the definition of $\mathfrak{L} = \mathfrak{L}^\varepsilon$ in (2.8),

$$\mathfrak{L}\eta_{x_0}^+(\eta) = (\mathcal{A}_\varepsilon^{*,+} \eta^+)(x_0) - \sum_{\{z \in I^\varepsilon: z_+ = x_0\}} \frac{\lambda}{\varepsilon} \Psi(z) \eta_t^+(z_+) \eta_t^-(z_-). \quad (4.7)$$

Similarly, for $y_0 \in D_-^\varepsilon$, we have

$$\mathfrak{L}\eta_{y_0}^-(\eta) = (\mathcal{A}_\varepsilon^{*,-} \eta^-)(y_0) - \sum_{\{z \in I^\varepsilon: z_- = y_0\}} \frac{\lambda}{\varepsilon} \Psi(z) \eta_t^+(z_+) \eta_t^-(z_-). \quad (4.8)$$

Hence, by linearity of \mathfrak{L} , (4.7) and then (4.3), we have

$$\begin{aligned} \mathfrak{L}f(\eta) &= \frac{1}{N} \sum_{x \in D_+} \phi(x) (\mathfrak{L}\eta_x^+)(\eta) \\ &= \frac{1}{N} \sum_{x \in D_+} \phi(x) (\mathcal{A}_\varepsilon^{*,+} \eta^+(x)) - \frac{\lambda}{N\varepsilon} \sum_{z \in I^\varepsilon} \Psi(z) \eta^+(z_+) \eta^-(z_-) \phi(z_+) \\ &= \frac{1}{N} \sum_{x \in D_+} \eta^+(x) (\mathcal{A}_\varepsilon^+ \phi(x)) - \frac{\lambda}{N\varepsilon} \sum_{z \in I^\varepsilon} \Psi(z) \eta^+(z_+) \eta^-(z_-) \phi(z_+). \end{aligned}$$

Hence

$$\mathfrak{L}f(\eta_s) = \langle \mathcal{A}_\varepsilon^+ \phi, \mathfrak{X}_s^{N,+} \rangle - \frac{\lambda}{N\varepsilon^d} \langle J_s^{N,+}, \phi \rangle \quad (4.9)$$

and $M(t)$ is an \mathfrak{F}_t^η -martingale by (4.1). Next, we compute $\mathbb{E}[\langle M \rangle_t]$. Note that

$$\mathfrak{L}(f^2)(\eta) = \frac{1}{N^2} \sum_{a \in D_+^\varepsilon} \sum_{b \in D_+^\varepsilon} \phi^+(a) \phi^+(b) \mathfrak{L}(\eta_a \eta_b)(\eta),$$

where $\mathfrak{L}(\eta_a \eta_b)$ can be computed explicitly using (2.8). Hence from (4.2), we can check that

$$\mathbb{E}[M^2(t)] = \mathbb{E}[\langle M \rangle_t] = \mathbb{E} \left[\int_0^t \mathfrak{L}(f^2)(\eta_s) - 2f(\eta_s) \mathfrak{L}f(\eta_s) ds \right] = \int_0^t \mathbb{E}[g(\eta_r)] dr,$$

where

$$\begin{aligned} g(\eta) &= \frac{1}{N^2} \left(\sum_{y, z \in D_+} \eta^+(z) h^{-1}(z) p_{zy} (\phi(y) - \phi(z))^2 + \frac{\lambda}{\varepsilon} \sum_{z \in I^\varepsilon} \Psi(z) \eta(z_+) \eta^-(z_-) (-\phi(z_+))^2 \right) \\ &\leq \frac{1}{N^2} \left(\varepsilon^2 \|\nabla \phi\|_\infty^2 \sum_{x \in D_+} \eta(x) h^{-1}(x) + \frac{\lambda}{\varepsilon} \|\phi\|_\infty^2 \sum_{z \in I^\varepsilon} \Psi(z) \eta^+(z_+) \eta^-(z_-) \right) \\ &\leq \frac{d \|\nabla \phi\|_\infty^2}{N} + \frac{\lambda \|\phi\|_\infty^2}{N^2} \left(\frac{1}{\varepsilon} \sum_{z \in I^\varepsilon} \Psi(z) \eta^+(z_+) \eta^-(z_-) \right). \end{aligned}$$

After taking expectation for $g(\eta_r)$, the first term is the last display is of order at most $1/N$ since $\phi \in C^1(\bar{D}_+)$, while the second term inside the bracket is of order at most $1/N$, uniformly in $r \in [0, t]$, by (2.37). Hence $\mathbb{E}[M^2(t)] \leq \frac{C}{N}$ for some $C = C(\phi, d, D_\pm, \lambda)$. Doob's maximal inequality then gives (4.6). \square

Remark 4.2. From the second term of (4.9), we see that if the parameter of the killing time is of order λ/ε , then we need $N\varepsilon^d$ to be comparable to 1. \square

4.2 Tightness

The following simple observation is useful for proving tightness when the transition kernel of the process has a singularity at $t = 0$. It says that we can break down the analysis of the fluctuation of functionals of a process on $[0, T]$ into two parts. One part is near $t = 0$, and the other is away from $t = 0$ where we have a bound for a higher moment. Its proof, which is based on the Prohorov's theorem, is simple and is omitted (detail can be found in [23]).

Lemma 4.3. *Let $\{Y_N\}$ be a sequence of real valued processes such that $t \mapsto \int_0^t Y_N(r) dr$ is continuous on $[0, T]$ a.s., where $T \in [0, \infty)$. Suppose the following holds:*

(i) *There exists $q > 1$ such that $\overline{\lim}_{N \rightarrow \infty} \mathbb{E}[\int_h^T |Y_N(r)|^q dr] < \infty$ for any $h > 0$,*

(ii) *$\lim_{\alpha \searrow 0} \overline{\lim}_{N \rightarrow \infty} \mathbb{P}(\int_0^\alpha |Y_N(r)| dr > \varepsilon_0) = 0$ for any $\varepsilon_0 > 0$.*

Then $\{\int_0^t Y_N(r) dr; t \in [0, T]\}_{N \in \mathbb{N}}$ is tight in $C([0, T], \mathbb{R})$.

Here is our tightness result for $\{\langle \mathfrak{X}^{N,+}, \mathfrak{X}^{N,-} \rangle\}$. We need Lemma 4.3 in the proof mainly because we do not know if $\overline{\lim}_{N \rightarrow \infty} \mathbb{E} \int_0^T \langle \mathcal{A}_\varepsilon^+ \varphi_+, \mathfrak{X}_s^{N,+} \rangle^2 ds$ is finite or not.

Theorem 4.4. *The sequence $\{\langle \mathfrak{X}^{N,+}, \mathfrak{X}^{N,-} \rangle\}$ is relatively compact in $D([0, T], \mathfrak{E})$ and any subsequential limit of the laws of $\{\langle \mathfrak{X}^{N,+}, \mathfrak{X}^{N,-} \rangle\}$ carries on $C([0, T], \mathfrak{E})$. Moreover, for all $\varphi_\pm \in C^2(\overline{D}_\pm)$,*

$$\left\{ \int_0^t \langle J_s^{N,+}, \varphi_\pm \rangle ds \right\}, \quad \left\{ \int_0^t \langle \mathcal{A}_\varepsilon^+ \varphi_+, \mathfrak{X}_s^{N,+} \rangle ds \right\} \quad \text{and} \quad \left\{ \int_0^t \langle \mathcal{A}_\varepsilon^- \varphi_-, \mathfrak{X}_s^{N,-} \rangle ds \right\}$$

are all tight in $C([0, T], \mathbb{R})$.

Proof We write \mathfrak{X}^\pm in place of $\mathfrak{X}^{N,\pm}$ for convenience. By Stone-Weierstrass Theorem, $C^2(\overline{D}_\pm)$ is dense in $C(\overline{D}_\pm)$ in uniform topology. It suffices to check that $\{(\langle \mathfrak{X}^+, \phi^+ \rangle, \langle \mathfrak{X}^-, \phi^- \rangle)\}$ is relatively compact in $D([0, T], \mathbb{R}^2)$ for all $\phi^\pm \in C^2(\overline{D}_\pm)$ (cf. Proposition 1.7 (p.54) of [27]) for this weak tightness criterion). By Prohorov's theorem (see Theorem 1.3 and remark 1.4 of [27]), $\{(\langle \mathfrak{X}^+, \phi^+ \rangle, \langle \mathfrak{X}^-, \phi^- \rangle)\}$ is relatively compact in $D([0, T], \mathbb{R}^2)$ if (1) and (2) below holds:

(1) For all $t \in [0, T]$ and $\varepsilon_0 > 0$, there exists a compact set $K(t, \varepsilon_0) \subset \mathbb{R}^2$ such that

$$\sup_N \mathbb{P} \left((\langle \mathfrak{X}_t^+, \phi^+ \rangle, \langle \mathfrak{X}_t^-, \phi^- \rangle) \notin K(t, \varepsilon_0) \right) < \varepsilon_0;$$

(2) For all $\varepsilon_0 > 0$,

$$\lim_{\delta \rightarrow 0} \overline{\lim}_{N \rightarrow \infty} \mathbb{P} \left(\sup_{\substack{|t-s| < \delta \\ 0 \leq s, t \leq T}} \left| (\langle \mathfrak{X}_t^+, \phi^+ \rangle, \langle \mathfrak{X}_t^-, \phi^- \rangle) - (\langle \mathfrak{X}_s^+, \phi^+ \rangle, \langle \mathfrak{X}_s^-, \phi^- \rangle) \right|_{\mathbb{R}^2} > \varepsilon_0 \right) = 0.$$

We first check (1). Since ϕ^\pm is bounded on \overline{D}_\pm and $|\langle \mathfrak{X}_t^\pm, 1 \rangle| \leq 1$ for all $t \in [0, \infty)$, we can always take $K = [-\|\phi^+\|_\infty, \|\phi^+\|_\infty] \times [-\|\phi^-\|_\infty, \|\phi^-\|_\infty]$.

To verify (2), since $|(x_1, y_1) - (x_2, y_2)|_{\mathbb{R}^2} \leq |x_1 - x_2| + |y_1 - y_2|$, we only need to focus on \mathfrak{X}^+ . By Lemma 4.1,

$$\left| \langle \phi, \mathfrak{X}_t^+ \rangle - \langle \phi, \mathfrak{X}_s^+ \rangle \right| = \left| \int_s^t \langle \mathcal{A}_\varepsilon^+ \phi, \mathfrak{X}_r^+ \rangle dr - \int_s^t \frac{\lambda}{N \varepsilon^d} \langle J_r^{N,+}, \phi \rangle dr + (M_\phi(t) - M_\phi(s)) \right|. \quad (4.10)$$

So we only need to verify (2) with $\langle \phi, \mathfrak{X}_t^+ \rangle - \langle \phi, \mathfrak{X}_s^+ \rangle$ replaced by each of the 3 terms on RHS of the above equation (4.10).

For the first term of (4.10), we apply Lemma 4.3 for the case $q = 2$ and $Y_N(r) = \langle \mathcal{A}_\varepsilon^+ \phi, \mathfrak{X}_r^+ \rangle$. Since $\phi \in C^2(\overline{D}_+)$, we have

$$\sup_{x \in D^\varepsilon \setminus \partial D^\varepsilon} |\mathcal{A}_\varepsilon^+ \phi(x)| \leq C(\phi) \quad \text{and} \quad \sup_{x \in \partial D^\varepsilon} |\varepsilon \mathcal{A}_\varepsilon^+ \phi(x)| \leq C(\phi)$$

for some constant $C(\phi)$ which only depends only on ϕ . Using Lemma 2.13, we have

$$\mathbb{E} [\langle |\mathcal{A}_\varepsilon^+ \phi|, \mathfrak{X}_r^{N,+} \rangle] \leq \frac{1}{N} \sum_{i=1}^N \sum_{D^\varepsilon} |\mathcal{A}_\varepsilon^+ \phi(\cdot)| p^{\varepsilon,+}(r, x_i, \cdot) m_\varepsilon(\cdot) \leq C_1(d, D_+, \phi) + \frac{C_2(d, D_+, \phi)}{\varepsilon \vee r^{\frac{1}{2}}},$$

which is in $L^1[0, T]$ as a function in r . This implies hypothesis (ii) of Lemma 4.3, via the Chebyshev's inequality. Hypothesis (i) of Lemma 4.3 can be verified easily using the upper bound (3.17) for the correlation function, or by direct comparison to the process without annihilation:

$$\begin{aligned} & \mathbb{E} \left[\langle |\mathcal{A}_\varepsilon \phi|, \mathfrak{X}_r^+ \rangle^2 \right] \\ & \leq \left(\frac{1}{N} \sum_{i=1}^N \mathbb{E} [\mathcal{A}_\varepsilon \phi(X_r^i)] \right)^2 + \frac{1}{N^2} \sum_{i=1}^N \mathbb{E} [(\mathcal{A}_\varepsilon \phi)^2(X_r^i)] - \frac{1}{N^2} \sum_{i=1}^N (\mathbb{E} [\mathcal{A}_\varepsilon \phi(X_r^i)])^2 \\ & \leq C(d, D, \phi) \left(1 + \frac{1}{\sqrt{r}} + \frac{1}{r} \right). \end{aligned}$$

For the second term of (4.10), by (2.37) we have $\overline{\lim}_{N \rightarrow \infty} \mathbb{E} \left[\int_0^T \langle 1, J_r^N \rangle^2 dr \right] < \infty$. Hence (2) holds for this term by Lemma 4.3.

For the third term of (4.10), by Chebyshev's inequality, Doob's maximal inequality and Lemma 4.1, we have

$$\begin{aligned} \mathbb{P} \left(\sup_{|t-s| < \delta} |M_\phi(t) - M_\phi(s)| > \varepsilon_0 \right) & \leq \frac{1}{\varepsilon_0^2} \mathbb{E} \left[\left(\sup_{|t-s| < \delta} |M_\phi(t) - M_\phi(s)| \right)^2 \right] \\ & \leq \frac{1}{\varepsilon_0^2} \mathbb{E} \left[\left(2 \sup_{t \in [0, T]} |M_\phi(t)| \right)^2 \right] \\ & \leq 16 \mathbb{E} [M_\phi(T)^2] \leq \frac{C}{N}. \end{aligned}$$

We have proved that (2) is satisfied. Hence $\{(\mathfrak{X}^{N,+}, \mathfrak{X}^{N,-})\}$ is relatively compact. Using (2) and the metric of \mathfrak{E} , we can check that any subsequential limit L^* of the laws of $\{(\mathfrak{X}^{N,+}, \mathfrak{X}^{N,-})\}$ concentrates on $C([0, \infty), \mathfrak{E})$. \square

Remark 4.5. In general, to prove tightness for (X_n, Y_n) in $D([0, T], A \times B)$, it is NOT enough to prove tightness separately for (X_n) and (Y_n) in $D([0, T], A)$ and $D([0, T], B)$ respectively. (However, the latter condition implies tightness in $D([0, T], A) \times D([0, T], B)$ trivially). See Exercise 22(a) in Chapter 3 of [22]. For example, $(\mathbf{1}_{[1+\frac{1}{n}, \infty)}, \mathbf{1}_{[1, \infty)})$ converges in $D_{\mathbb{R}}[0, \infty) \times D_{\mathbb{R}}[0, \infty)$ but not in $D_{\mathbb{R}^2}[0, \infty)$. The reason is that the two processes can jump at different times ($t = 1$ and $t = 1 + \frac{1}{n}$) that become identified in the limit (only one jump at $t = 1$); this can be avoided if one of the two processes is C -tight (i.e. has only continuous limiting values), which is satisfied in our case since $\mathfrak{X}^{N,+}$ and $\mathfrak{X}^{N,-}$ turns out to be both C -tight. \square

Remark 4.6. Even without condition (ii) of Theorem 2.20 for η_0 , we can still verify hypothesis (i) of Lemma 4.3. Actually, applying (4.5) to suitable test functions, we have

$$\lim_{\alpha \rightarrow 0} \overline{\lim}_{N \rightarrow \infty} \mathbb{E} \left[\int_0^\alpha \langle J_s^{N,+}, 1 \rangle ds \right] = 0. \quad (4.11)$$

\square

4.3 Identifying the limit

Suppose $(\mathfrak{X}^{\infty,+}, \mathfrak{X}^{\infty,-})$ is a subsequential limit of $(\mathfrak{X}^{N,+}, \mathfrak{X}^{N,-})$, say the convergence is along the subsequence $\{N'\}$. By the Skorokhod representation Theorem, the continuity of the limit in t and [22, Theorem 3.10.2], there exists a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ such that

$$\lim_{N' \rightarrow \infty} \sup_{t \in [0, T]} \left\| (\mathfrak{X}_t^{+,N'}, \mathfrak{X}_t^{-,N'}) - (\mathfrak{X}_t^{\infty,+}, \mathfrak{X}_t^{\infty,-}) \right\|_{\mathfrak{E}} = 0 \quad \mathbb{P}\text{-a.s.}, \quad (4.12)$$

Hence we have for any $t > 0$ and $\phi \in C(\overline{D}_+)$,

$$\lim_{N' \rightarrow \infty} \mathbb{E}[\langle \mathfrak{X}_t^+, \phi \rangle] = \mathbb{E}[\langle \mathfrak{X}_t^{+, \infty}, \phi \rangle] \quad \text{and} \quad \lim_{N' \rightarrow \infty} \mathbb{E}[\langle (\mathfrak{X}_t^+)^2, \phi \rangle] = \mathbb{E}[\langle (\mathfrak{X}_t^{+, \infty})^2, \phi \rangle].$$

Combining with Corollary 2.23, we have

$$\langle \mathfrak{X}_t^{+, \infty}, \phi \rangle = \langle u_+(t), \phi \rangle_{\rho_+} \quad \mathbb{P}\text{-a.s. for every } t \geq 0 \text{ and for } \phi.$$

Here we have used the simple fact that if $\mathbb{E}[X] = (\mathbb{E}[X^2])^{1/2} = a$, then $X = a$ a.s.

Suppose $\{\phi_k\}$ is a countable dense subset of $C(\overline{D}_+)$. Then for every $t \geq 0$,

$$\langle \mathfrak{X}_t^{\infty,+}, \phi_k \rangle = \langle u_+(t), \phi_k \rangle_{\rho_+} \quad \text{for every } k \geq 1 \text{ } \mathbb{P}\text{-a.s.}$$

Since $\mathfrak{X}^{\infty,+} \in C((0, \infty), M_+(\overline{D}_+))$, we can pass to rational numbers to obtain

$$\langle \mathfrak{X}_t^{\infty,+}, \phi_k \rangle = \langle u_+(t), \phi_k \rangle_{\rho_+} \quad \text{for every } t \geq 0 \text{ and } k \geq 1 \text{ } \mathbb{P}\text{-a.s.}$$

Hence,

$$\mathfrak{X}_t^{\infty,+}(dx) = u_+(t, x) \rho_+(x) dx \quad \text{for every } t \geq 0 \quad \mathbb{P}\text{-a.s.}$$

Similarly,

$$\mathfrak{X}_t^{\infty,-}(dy) = u_-(t, y) \rho_-(y) dy \quad \text{for every } t \geq 0 \quad \mathbb{P}\text{-a.s.}$$

In conclusion, any subsequential limit is the dirac delta measure

$$\delta_{u_+(t,x)dx, u_-(t,y)dy} \in M_1(D([0, \infty), \mathfrak{E})).$$

This together with Theorem 4.4 completes the proof of Theorem 2.20.

5 Local Central Limit Theorem

Suppose $D \subset \mathbb{R}^d$ is a bounded Lipschitz domain and $\rho \in W^{1,2}(D) \cap C^1(\overline{D})$ is strictly positive. Suppose X is a $(I_{d \times d}, \rho)$ -reflected diffusion and X^ε be an ε -approximation of X , described in the subsection 2.1.2. In this section, we prove the local central limit theorem (Theorem 2.12), the Gaussian upper bound (Theorem 2.9) and the Hölder continuity (Theorem 5.12) for p^ε . The proofs are standard once we establish a discrete analogue of a relative isoperimetric inequality (Theorem 5.5) for bounded Lipschitz domains.

5.1 Discrete relative isoperimetric inequality

Note that any Lipschitz domain enjoys the uniform cone property and any bounded $(d-1)$ -dimensional manifold of class $C^{0,1}$ has finite perimeter. Hence, by Corollary 3.2.3 (p.165) and Theorem 6.1.3 (p.300) of [35], we have the following relative isoperimetric inequality.

Proposition 5.1 (Relative isoperimetric inequality). *Let $D \subset \mathbb{R}^d$ be a bounded Lipschitz domain and $r \in (0, 1]$. Then*

$$S(r, D) := \sup_{U \in \mathfrak{G}} \frac{|U|^{\frac{d-1}{d}}}{\sigma(\partial U \cap D)} < \infty, \quad (5.1)$$

where \mathfrak{G} is the collection of open subsets $U \subset D$ such that $|U| \leq r|D|$ and $\partial U \cap D$ is a manifold of class $C^{0,1}$ (i.e. each point $x \in \partial U \cap D$ has a neighborhood which can be represented by the graph of a Lipschitz function). Moreover, $S(r, D) = S(r, aD)$ for all $a > 0$.

In this subsection, we establish a discrete analogue for the relative isoperimetric inequality (Theorem 5.5).

First, we study the scaled graph aD^ε and gather some basic properties of a continuous time random walk on a finite set.

5.1.1 CTRW on scaled graph aD^ε

Recall the m_ε -symmetric CTRW X^ε on D^ε defined in the Subsection 2.1.2. The Dirichlet form $(\mathcal{E}^{(\varepsilon)}, l^2(m_\varepsilon))$ of X^ε in $l^2(m_\varepsilon)$ is given by

$$\mathcal{E}^{(\varepsilon)}(f, g) := \frac{1}{2} \sum_{x, y \in D^\varepsilon} (f(y) - f(x))(g(y) - g(x)) \mu_{xy}, \quad (5.2)$$

where $\mu_{xy} = \mu_{x,y}^{D^\varepsilon}$ are the conductance on the graph D^ε defined in the Subsection 2.1.2. The stationary measure $\pi = \pi^{D^\varepsilon}$ of X^ε is given by $\pi(x) = m_\varepsilon(x)/m(D^\varepsilon)$, where $m(D^\varepsilon) := \sum_{x \in D^\varepsilon} m_\varepsilon(x)$.

We now consider the scaled graph $aD^\varepsilon = (aD)^{a\varepsilon}$, which is an approximation to the bounded Lipschitz domain aD by square lattice $a\varepsilon\mathbb{Z}^d$. Clearly the degrees of vertices are given by $v^{aD^\varepsilon}(ax) = v^{D^\varepsilon}(x)$. Define the function $\rho_{(aD)}$ on aD by $\rho_{(aD)}(ax) := \rho(x)$. Then define the CTRW X^{aD^ε} using $\rho_{(aD)}$ as we have done for X^ε using ρ .

The mean holding time of X^{aD^ε} is $(a\varepsilon)^2/d$. Clearly, the symmetrizing measure m^{aD^ε} and the stationary probability measure π^{aD^ε} have the scaling property $m^{aD^\varepsilon}(ax) = a^d m^{D^\varepsilon}(x)$ and $\pi^{aD^\varepsilon}(ax) = \pi^{D^\varepsilon}(x)$. Let $p_{aD}^{a\varepsilon}$ be the transition density of X^{aD^ε} with respect to the symmetrizing measure m^{aD^ε} . Then

$$a^d p_{aD}^{a\varepsilon}(a^2 t, ax, ay) = p_D^\varepsilon(t, x, y) \quad (5.3)$$

for every $t > 0$, $\varepsilon > 0$, $a > 0$ and $x, y \in D^\varepsilon$.

We will simply write m and π for the symmetrizing measure and the stationary probability measure when there is no ambiguity for the underlying graph.

5.1.2 An extension lemma

Following the notation of [36], we let G be a finite set, $K(x, y)$ be a Markov kernel on G and π the stationary measure of K . Note that a Markov chain on a finite set induces a natural graph structure as follows. Let $Q(e) := \frac{1}{2}(K(x, y)\pi(x) + K(y, x)\pi(y))$ for any $e = (x, y) \in G \times G$. Define the set of directed edges $E := \{e = (x, y) \in G \times G : Q(e) > 0\}$.

We use the following 2 different notions for the ‘‘boundary’’ of $A \subset G$:

$$\begin{aligned} \partial_e A &:= \{e = (x, y) \in E : x \in A, y \in G \setminus A \text{ or } y \in A, x \in G \setminus A\}, \\ \partial A &:= \{x \in A : \exists y \in G \setminus A \text{ such that } (x, y) \in E\}. \end{aligned}$$

Observe that each edge in $\partial_e A$ is counted twice. Set

$$Q(\partial_e A) := \frac{1}{2} \sum_{e \in \partial_e A} q(e) = \frac{1}{2} \sum_{x \in A, y \in G \setminus A} (K(x, y)\pi(x) + K(y, x)\pi(y)).$$

Definition 5.2. For any $r \in (0, 1)$, define

$$S_\pi(r, G) := \sup_{\{A \subset G: \pi(A) \leq r\}} \frac{2|A|^{(d-1)/d}}{|\partial_e A|} \quad \text{and} \quad \tilde{S}_\pi(r, G) := \sup_{\{A \subset G: \pi(A) \leq r\}} \frac{\pi(A)^{(d-1)/d}}{Q(\partial_e A)}. \quad (5.4)$$

We call $1/\tilde{S}_\pi(r, G)$ an **isoperimetric constant** of the chain (K, π) . It provides rich information about the geometric properties of G and the behavior of the chain (cf. [36]).

In our case, $G = aD^\varepsilon$, $\pi(x) = \frac{m(x)}{m(aD^\varepsilon)}$ and $K(x, y) = p_{xy}$ in aD^ε , where $p_{x,y}$ is the one-step transition probabilities of X^{aD^ε} defined in subsection 2.1.2. For $a = 1$ and $A \subset D^\varepsilon$, we have

$$\begin{aligned} \partial_e A &= \{(x, y) \in A \times D^\varepsilon \setminus A \cup D^\varepsilon \setminus A \times A : \text{the line segment } (x, y] \subset D\}, \\ \partial A &= \{x \in A : \exists y \in D^\varepsilon \setminus A \text{ such that } |x - y| = \varepsilon \text{ and the line segment } [x, y] \subset D\}, \\ \tilde{\partial} A &:= \{x \in A : \exists y \in \varepsilon\mathbb{Z}^d \text{ such that } |x - y| = \varepsilon \text{ and } (x, y] \cap \partial D \neq \emptyset\}, \\ \Delta A &:= \tilde{\partial} A \setminus \partial A. \end{aligned}$$

In this notation, we have $\partial D^\varepsilon = \emptyset$, $\tilde{\partial} D^\varepsilon = \{x \in D^\varepsilon : v(x) < 2d\}$, $A \cap \tilde{\partial} D^\varepsilon = \tilde{\partial} A$ and $\tilde{\partial} D^\varepsilon = \Delta A \cup (\partial A \cap \tilde{\partial} D^\varepsilon) \cup (\tilde{\partial} D^\varepsilon \setminus A)$. See Figure 6 for an illustration.

Definition 5.3. We say that $A \subset D^\varepsilon$ is *grid-connected* if $\partial_e A_1 \cap \partial_e A_2 \neq \emptyset$ whenever $A = A_1 \cup A_2$.

It is easy to check that A is grid-connected if and only if for every $x, y \in A$, there exists $\{x_1 = x, x_2, \dots, x_{m-1}, x_m = y\} \subset A$ such that each line segments $[x_j, x_{j+1}] \subset D$ and $|x_j - x_{j+1}| = \varepsilon$.

The following is a key lemma which allows us to derive the relative isoperimetric inequality for the discrete setting from that in the continuous setting, and hence leads us to Theorem 5.5.

Lemma 5.4. (*Extension of sub-domains*) Let π_{srw} be the stationary measure of the simple random walk (SRW) on D^ε . For any $r \in (0, 1)$, there exist positive constants $\varepsilon_1(d, D, r)$, $M_1(d, D, r)$ and $M_2(d, D, r)$ such that if $\varepsilon \in (0, \varepsilon_1)$, then for all grid-connected $A \subset D^\varepsilon$ with $\pi_{srw}(A) \leq r$, we can find a connected open subset $U \subset D$ which contains A and satisfies:

- (a) $\partial U \cap D$ is \mathcal{H}^{d-1} -rectifiable,
- (b) $|U| \leq \frac{49r+1}{50}|D|$,
- (c) $\varepsilon^d |A| \leq M_1 |U|$,
- (d) $M_2 \varepsilon^{d-1} |\partial A| \geq \sigma(\partial U \cap D)$.

Proof Since the proof for each $r \in (0, 1)$ is the same, we just give a proof for the case $r = 1/2$.

For $x \in \varepsilon\mathbb{Z}^d$, let $U_x := \prod_{i=1}^d (x_i - \frac{\varepsilon}{2}, x_i + \frac{\varepsilon}{2})$ be the cube in the dual lattice which contains x . Since A is grid-connected, we have $(W_1)^\circ$ is connected in \mathbb{R}^d , where $(W_1)^\circ$ is the interior of $W_1 := \cup_{x \in A} (\overline{U_x} \cap D)$. (See Figure 6 for an illustration.)

Note that we cannot simply take $U = (W_1)^\circ$ because (d) may fail, for example when ΔA contributes too much to $\partial U \cap D$, i.e., when $(W_1)_\Delta := \partial W_1 \cap (\bigcup_{x \in \Delta A} \partial U_x)$ is large. However, $\Delta A \subset \tilde{\partial} D^\varepsilon$ is close to ∂D and so we can fill in the gaps between ΔA and ∂D to eliminate those contributions. In this process, we may create some extra pieces for $\partial U \cap D$, but we will show that those pieces are small enough. Following this observation, we will eventually take $U = (W_1 \cup W_2)^\circ$ where $W_2 \subset D_h$ for some small enough $h > 0$.

Since D is a bounded Lipschitz domains, we can choose $h > 0$ small enough so that $|D_h| < |D|/200$. Moreover, $\pi_{srw}(A) < 1/2$ implies $\varepsilon^d |A| \leq \varepsilon^d |\partial(D^\varepsilon)| + m_{srw}(D^\varepsilon)/2$. So we can choose ε small enough so that $|W_1| \leq \varepsilon^d |A| \leq \frac{101}{200}|D|$. Hence U satisfies (b). By Lipschitz property again, there exists $M_1 > 0$ such that $|U_x \cap D| \geq |U_x|/M_1 = \varepsilon^d/M_1$ for any $x \in D^\varepsilon$. Hence (c) is satisfied.

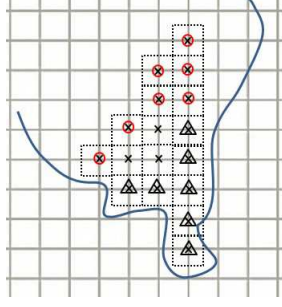


Figure 6: $W_1 := \cup_{x \in A} (\overline{U_x} \cap D)$

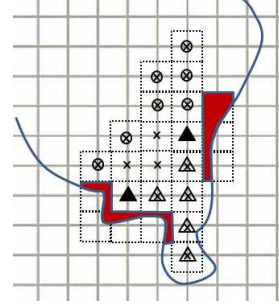


Figure 7: W_3 is the shaded part

It remains to construct W_2 in such a way that $W_2 \subset D_h$ for some small enough $h > 0$ (more precisely, for h small enough so that $|D_h| < |D|/200$) and that (a) and (d) are satisfied. We will construct W_2 in 3 steps:

Step 1: (Construct W_3 to seal the opening between ∂D and the subset of $(W_1)_\partial$ which are close to ∂D . See Figure 7.) Write $\Delta A = \Delta_1 A \cup \Delta_2 A$ where $\Delta_2 A := \Delta A \setminus \Delta_1 A$ and

$$\Delta_1 A := \{x \in \Delta A : \exists y \in \partial A \text{ such that } \max\{|x_i - y_i| : 1 \leq i \leq d\} = 1\}.$$

Points in $\Delta_1 A$ are marked in solid black in Figure 7. For $x \in \Delta_1 A$, consider the following closed cube centered at x :

$$T_x := \bigcup_{y \in \widehat{B}(x, 10R\varepsilon)} \overline{U_y}, \text{ where } R = \sqrt{d}(M+1)$$

Let Θ_x be the union of all connected components of $T_x \cap D$ whose closure intersects $\overline{U_x}$ and define

$$W_3 := \bigcup_{x \in \Delta_1 A} \Theta_x.$$

Step 2: (Fill in the gaps between ∂D and $(W_1)_\Delta$ near Δ_2 . See Figure 8) Note that $\cup_{x \in \Delta_1 A} \partial U_x$ does not contribute to $\partial(W_1 \cup W_3) \cap D$. Let W_4 be the union of all connected components of $D \setminus (W_1 \cup W_3)$ whose closure intersects $\overline{U_x}$ for some $x \in \Delta_2 A$.

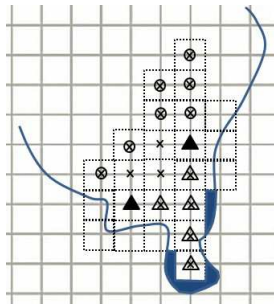


Figure 8: W_4 is the shaded part

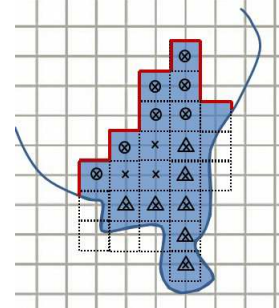


Figure 9: U is the shaded part

Step 3: Finally, take $W_2 := W_3 \cup W_4$, and set $U := (W_1 \cup W_3 \cup W_4)^\circ$. (See Figure 9.)

It is clear that U is connected and $\partial U \cap D \subset \bigcup_{x \in \mathbb{Z}^d} \partial U_x$ is piecewise linear, so (a) is satisfied. For any $W \subset D$, we have $\partial W \cap D = W_\partial \cup W_\Delta \cup W_\nabla$, where

$$W_\partial := \partial W \cap \left(\bigcup_{x \in \partial A} \partial U_x \right), \quad W_\Delta := \partial W \cap \left(\bigcup_{x \in \Delta A} \partial U_x \right) \text{ and } W_\nabla := \partial W \setminus \left(\bigcup_{x \in \partial A \cup \Delta A} \partial U_x \right).$$

Therefore, $\sigma(\partial W \cap D) \leq \sigma(W_\partial) + \sigma(W_\Delta) + \sigma(W_\nabla)$ whenever the corresponding surface measures are defined. It is clear that by construction we have

- $(W_1)_\nabla = \emptyset$,
- $(W_1 \cup W_3)_\partial \subset (W_1)_\partial$, $(W_1 \cup W_3)_\Delta \subset (W_1)_{\Delta_2}$, $(W_1 \cup W_3)_\nabla \subset \bigcup_{x \in \Delta_1} \bigcup_{y \in \widehat{B}(x, 10R\varepsilon)} \partial U_y$
where $(W_1)_{\Delta_2}$ is defined analogously as $(W_1)_\Delta$, with Δ replaced by Δ_2 ,
- $U_\partial \subset (W_1 \cup W_3)_\partial$, $U_\Delta = \emptyset$, $U_\nabla \subset (W_1 \cup W_3)_\nabla$.

Now $\sigma(U_\partial) \leq \sigma((W_1)_\partial) \leq |\partial A| 2d\varepsilon^{d-1}$. Moreover, each $x \in \partial A$ is adjacent to at most $3^d - 1$ points in $\Delta_1 A \cup \Delta A$, and for each $x \in \Delta_1 A$, there are at most $|\widehat{B}(10R\varepsilon)| \leq (20R + 1)^d$ cubes in T_x . So we have

$$\sigma(U_\nabla) \leq \sigma((W_1 \cup W_3)_\nabla) \leq (3^d - 1)|\partial A| (20R + 1)^d 2d\varepsilon^{d-1}.$$

Hence (d) is satisfied.

Since $\text{diam}(T_x) < 20R\sqrt{d}\varepsilon$, we have $W_3 \subset D_{(20R\sqrt{d}+1)\varepsilon}$. To complete the proof, it suffices to show that $W_4 \subset D_{(10R)\varepsilon}$. This is equivalent to show that any curve in $D \setminus \overline{W_1 \cup W_3}$ starting from any point in $(W_1)_{\Delta_2}$ must lie in $D_{(10R)\varepsilon}$.

Let $\gamma[0, 1]$ be an arbitrary continuous curve starting at an arbitrary point $p \in (W_1)_{\Delta_2}$ such that $\gamma(0, 1) \subset D \setminus \overline{W_1}$ and $\text{dist}(\gamma(t), \partial D) > (10R)\varepsilon$ for some $t \in (0, 1]$. Define $\Omega_{D^\varepsilon} := \left(\bigcup_{x \in D^\varepsilon} \overline{U_x} \right) \cap D$. Since $(W_1)_\Delta \subset \partial(\Omega_{D^\varepsilon}) \cap D \subset \bigcup_{z \in \partial D^\varepsilon} \partial U_z$ and $\sup_{z \in \partial D^\varepsilon} \text{dist}(z, \partial D) < \varepsilon$, the time when γ first exits $D \setminus \overline{\Omega_{D^\varepsilon}}$ must be less than t by continuity of γ . That is,

$$\tau := \inf \left\{ s > 0 : \gamma(s) \in \left(\bigcup_{z \in \partial D^\varepsilon \setminus A} \partial U_z \right) \cap D \right\} < t.$$

It suffices to show that $\gamma(0, \tau] \cap \Theta_x \neq \emptyset$ for some $x \in \Delta_1 A$. We do so by constructing a continuous curve $\tilde{\gamma}$ which is close to γ and passes through ∂U_x for some $x \in \Delta_1 A$.

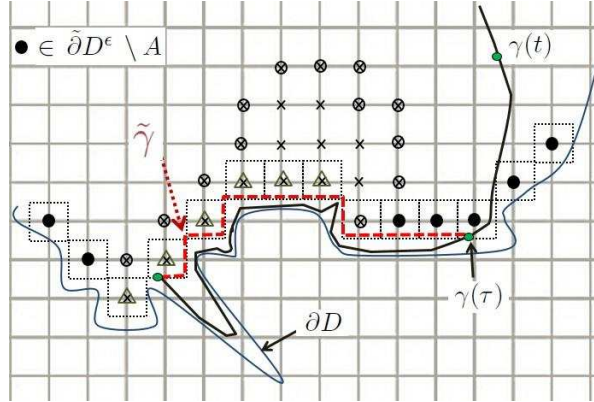


Figure 10: γ and a corresponding continuous $\tilde{\gamma} \subset \partial(\Omega_{D^\varepsilon}) \cap D$

Since $\sup_{s \in [0, \tau]} \text{dist}(\gamma(s), \partial D) < 2R\varepsilon$, we can choose ε small enough (depending only on D) and split $[0, \tau]$ into finitely many disjoint intervals I 's so that the $4R\varepsilon$ -tube of each $\gamma(I)$ lies in a coordinate ball $B(I)$ of D . For $s \in I$, project $\gamma(s)$ vertically upward (along the d -th coordinate of $B(I)$) onto $\partial(\Omega_{D^\varepsilon}) \cap D$ to obtain $\hat{\gamma}(s)$. Note that $\hat{\gamma}$ maybe discontinuous even in the interior of I . However, it is continuous on $[0, \tau]$ except possibly for finitely many points. Let $\{0 \leq s_1 < s_2 < \dots < s_m \leq \tau\}$ be the collection of discontinuities for $\hat{\gamma}([0, \tau])$. Then $0 < |\hat{\gamma}(s_j-) - \hat{\gamma}(s_j+)| \leq 2R\varepsilon$ and we can connect $\hat{\gamma}(s_j-)$ to $\hat{\gamma}(s_j+)$ by a continuous curve $\beta_j : [0, 1] \rightarrow \partial(\Omega_{D^\varepsilon}) \cap D \cap B(\hat{\gamma}(s_j-), 8R\varepsilon) \cap B(\hat{\gamma}(s_j+), 8R\varepsilon)$.

Define $\tilde{\gamma} : [0, \tau + m] \rightarrow \partial(\Omega_{D^\varepsilon}) \cap D$ to be the continuous curve obtained by concatenating $\hat{\gamma}$ and $\{\beta_j : j = 1, 2, \dots, m\}$ (See Figure 10). Then $\tilde{\gamma}(0) = \hat{\gamma}(0) = p \in (W_1)_{\Delta_2}$ and $\tilde{\gamma}(\tau + m) = \hat{\gamma}(\tau) = \gamma(\tau) \in \partial(\Omega_{D^\varepsilon}) \cap D \setminus (\partial W_1 \cap D)$. By the continuity of $\tilde{\gamma}$, there is some $t_* \in (0, \tau + m)$ such that $\tilde{\gamma}(t_*) \in (W_1)_{\Delta_1}$. (Roughly speaking, on ∂D^ε , $\Delta_2 A$ is separated from $\partial D^\varepsilon \setminus A$ by $\Delta_1 A$.)

Now for some $1 \leq j \leq m$, we have $\tilde{\gamma}(t_*)$ and $\gamma(s_j)$ are connected in $D \setminus \Omega_{D^\varepsilon} \subset D \setminus \overline{W_1}$, and

$$|\tilde{\gamma}(t_*) - \gamma(s_j)| \leq |\tilde{\gamma}(t_*) - \hat{\gamma}(s_j)| + |\hat{\gamma}(s_j) - \gamma(s_j)| \leq 8R\varepsilon + R\varepsilon.$$

Hence $\tilde{\gamma}(t_*) \in \partial U_x$ for some $x \in \Delta_1 A$. We therefore have $\gamma(s_j) \in \Theta_x$. The proof is now complete. \square

5.1.3 Discrete relative isoperimetric inequality

Let π_{srw} be the stationary measure of the simple random walk (SRW) on the graph under consideration and recall Definition 5.2.

Theorem 5.5 (Discrete relative isoperimetric inequality). *For every $r \in (0, 1)$, there exists $\hat{S}_{srw} = \hat{S}_{srw}(d, D, r) \in (0, \infty)$ and $\varepsilon_1 = \varepsilon_1(d, D, r) \in (0, \infty)$ such that*

$$\begin{aligned} \sup_{\varepsilon \in (0, \varepsilon_1)} S_{srw}(r, D^\varepsilon) &\leq \hat{S}_{srw}, \text{ and} \\ \tilde{S}_{srw}(r, D^\varepsilon) &\leq \frac{2d(m_{srw}(D^\varepsilon))^{1/d}}{\varepsilon} \hat{S}_{srw} \text{ for every } \varepsilon \in (0, \varepsilon_1). \end{aligned}$$

Proof We can also assume that A is grid-connected. This is because

$$\frac{|A|^{(d-1)/d}}{|\partial_e A|} \leq \frac{|A_1|^{(d-1)/d}}{|\partial_e A_1|} \vee \frac{|A_2|^{(d-1)/d}}{|\partial_e A_2|}$$

whenever $A = A_1 \cup A_2$ with $\partial_e A_1 \cap \partial_e A_2 = \emptyset$. From Lemma 5.4 and Proposition 5.1, we have

$$\sup_{\varepsilon \in (0, \varepsilon_1)} \sup_{\{A \subset D^\varepsilon : \pi(A) \leq r\}} \frac{|A|^{(d-1)/d}}{|\partial A|} \leq M_2 M_1^{(d-1)/d} S\left(\frac{49r+1}{50}, D\right).$$

We thus have the first inequality since $4d|\partial A| \geq |\partial_e A| \geq 2|\partial A|$. The second inequality follows from the first since $q(e) = \frac{(\varepsilon)^d}{2dm(D^\varepsilon)}$. \square

For the CTRW X^{aD^ε} on aD^ε , we let π be the stationary measure. Observe that, because $\pi^{aD^\varepsilon}(aA) = \pi^{D^\varepsilon}(A)$ and $m(aD^\varepsilon) = a^d m(D^\varepsilon)$, we have

$$S_\pi(r, aD^\varepsilon) = S_\pi(r, D^\varepsilon) \quad \text{and} \quad \tilde{S}_\pi(r, aD^\varepsilon) = \tilde{S}_\pi(r, D^\varepsilon) \quad (5.5)$$

for all $a > 0$ and $r > 0$. Hence we only need to consider the case $a = 1$. In view of Theorem 5.5 and (2.6), we have (taking $r = 1/2$)

Corollary 5.6. *There exist positive constants $\hat{S} = \hat{S}(d, D, \rho)$, $\varepsilon_1 = \varepsilon_1(d, D, \rho)$ and $\hat{C} = \hat{C}(d, D, \rho)$ such that*

$$\sup_{\varepsilon \in (0, \varepsilon_1)} S_\pi(1/2, D^\varepsilon) \leq \hat{S}, \text{ and} \quad (5.6)$$

$$\tilde{S}_\pi(1/2, D^\varepsilon) \leq \frac{\hat{C}}{\varepsilon} \hat{S} \text{ for every } \varepsilon \in (0, \varepsilon_1). \quad (5.7)$$

As an immediate consequence of Corollary 5.6 and (5.5), we have the following Poincaré inequality.

Corollary 5.7. (Poincaré inequality) *There exist $\varepsilon_1 = \varepsilon_1(d, D, \rho) > 0$ such that*

$$\frac{|D| a^{d-2}}{16 \widehat{C}^2 \widehat{S}^2} \|f - \langle f \rangle_\pi\|_{l^2(\pi)}^2 \leq \mathcal{E}_{aD}^{a\varepsilon}(f)$$

for all $f \in l^2(aD^\varepsilon, \pi)$, $\varepsilon \in (0, \varepsilon_1)$, $a > 0$. Here $\langle f \rangle_\pi := \sum f \pi$, \widehat{C} and \widehat{S} are the same constants in Corollary 5.6, and $\mathcal{E}_{aD}^{a\varepsilon}$ is the Dirichlet form in $l^2(m^{aD^\varepsilon})$ of the CTRW X^{aD^ε} (see (5.2)).

Proof By Corollary 5.6, the isoperimetric constant

$$\mathcal{I} := \inf_{\pi(A) \leq 1/2} \frac{Q(\partial A)}{\pi(A)} \geq 2^{1/d} \frac{1}{\widehat{S}} \geq 2^{1/d} \frac{\varepsilon}{\widehat{C} \widehat{S}}.$$

Hence, by the Cheeger's inequality (see [36, Lemma 3.3.7]),

$$\inf_f \frac{\mathcal{E}_{aD}^{a\varepsilon}(f)}{\|f - \langle f \rangle_\pi\|_{l^2(\pi)}^2} \geq \frac{d m(aD^\varepsilon)}{(a\varepsilon)^2} \frac{\mathcal{I}^2}{8} \geq \frac{|D|}{16} \frac{a^{d-2}}{\widehat{C}^2 \widehat{S}^2}$$

when $\varepsilon > 0$ is small enough. \square

The above Poincaré inequality already tells us a positive lower bound for the spectral gap of X^{aD^ε} and hence gives us an estimate for the mixing time. However, we will state a stronger result in Proposition 5.9 in the next subsection.

5.2 Nash's inequality and Poincaré inequality

The discrete relative isoperimetric inequality leads to the following two functional inequalities; namely, a Poincaré inequality and a Nash inequality that are uniform in ε and in scaling $D \mapsto aD$. The uniformity in scaling helps proving the near diagonal lower bound for p^ε .

Theorem 5.8. (Nash's inequality and Poincaré inequality uniform in ε and in scaling) *There exist $\varepsilon_1 = \varepsilon_1(d, D, \rho) > 0$ and $C = C(d, D, \rho) > 0$ such that*

$$\|f - \langle f \rangle_\pi\|_{l^2(\pi)}^{2(1+2/d)} \leq 8 \tilde{S}_\pi(1/2, D^\varepsilon)^2 \left(\frac{(a\varepsilon)^2}{d m(aD^\varepsilon)} \mathcal{E}_{aD}^{a\varepsilon}(f) \right) \|f\|_{l^1(\pi)}^{4/d} \quad (5.8)$$

$$\|f\|_{l^2(m)}^{2(1+2/d)} \leq C \left(\mathcal{E}_{aD}^{a\varepsilon}(f) + (\widehat{C} \widehat{S} a)^{-2} \|f\|_{l^2(m)}^2 \right) \|f\|_{l^1(m)}^{4/d} \quad (5.9)$$

for every $f \in l^2(aD^\varepsilon)$, $\varepsilon \in (0, \varepsilon_1)$ and $a \in (0, \infty)$, where \widehat{C} and \widehat{S} are the same constants in Corollary 5.6; $\langle f \rangle_\pi := \sum f \pi$ and $\mathcal{E}_{aD}^{a\varepsilon}$ is the Dirichlet form in $l^2(m^{aD^\varepsilon})$ of the CTRW X^{aD^ε} (see (5.2)).

Proof Note that $\frac{(a\varepsilon)^2}{d m(aD^\varepsilon)} \mathcal{E}_{aD}^{a\varepsilon}(f)$ is the Dirichlet form of the unit speed CTRW with the same one-step transition probabilities as that of X^{aD^ε} . Hence (5.8) follows directly from [36, Theorem 3.3.11] and (5.7). For (5.9), let $R = (2^{1/d} \delta a \varepsilon)^{-1}$ with $\delta \geq (\widehat{C} \widehat{S} a)^{-1}$. For any nonempty subset $A \subset aD^\varepsilon$,

$$\frac{Q(\partial A) + \frac{1}{R} \pi(A)}{\pi(A)^{\frac{d-1}{d}}} \geq \frac{1}{\widehat{S}} \wedge \frac{1}{R} \left(\frac{1}{2} \right)^{\frac{1}{d}} \geq \left((\widehat{C} \widehat{S})^{-1} \wedge a \delta \right) \varepsilon = (\widehat{C} \widehat{S})^{-1} \varepsilon.$$

Hence,

$$\sup_{A \subset aD^\varepsilon} \frac{\pi(A)^{\frac{d-1}{d}}}{Q(\partial A) + \frac{1}{R} \pi(A)} \leq \frac{\widehat{C} \widehat{S}}{\varepsilon}. \quad (5.10)$$

By [36, Theorem 3.3.10],

$$\|f\|_{l^2(\pi)}^{2(1+2/d)} \leq 16 \left(\frac{C}{\varepsilon} \right)^2 \left(\frac{(a\varepsilon)^2}{d m(aD^\varepsilon)} \mathcal{E}_{aD}^{a\varepsilon}(f) + \frac{1}{8R^2} \|f\|_{l^2(\pi)}^2 \right) \|f\|_{l^1(\pi)}^{4/d}.$$

Using the relations $\|f\|_{l^2(\pi)}^2 = (m(aD^\varepsilon))^{-1} \|f\|_{l^2(m)}^2$, $\|g\|_{l^1(\pi)} = (m(aD^\varepsilon))^{-1} \|g\|_{l^1(m)}$ and (2.6), we get the desired inequality (5.9). \square

5.3 Mixing time

By the Poincaré inequalities in (5.8) and [36, Corollary 2.3.2], we obtain an estimate on the time needed to reach stationarity.

Proposition 5.9. *(Mixing time estimate) There exists $C > 0$ which depends only on d such that*

$$\left| p_{aD}^{a\varepsilon}(t, x, y) - \frac{1}{m(aD^\varepsilon)} \right| \leq C \min \left\{ (a\widehat{C}\widehat{S})^d t^{-d/2}, \frac{1}{(a\varepsilon)^d} \exp\left(\frac{-dt}{8(a\widehat{C}\widehat{S})^2}\right) \right\}$$

for every $t > 0$, $x, y \in aD^\varepsilon$, $\varepsilon \in (0, \varepsilon_1)$ and $a > 0$. Here \widehat{C} and \widehat{S} are the constants in Corollary 5.6.

Proof By (5.8) and Theorem 2.3.1 of [36], we have

$$\left| m(aD^\varepsilon) p_{aD}^{a\varepsilon}\left(\frac{(a\varepsilon)^2}{d}t, x, y\right) - 1 \right| \leq \left(\frac{d(8\widehat{S}^2)}{2t}\right)^{d/2}$$

After simplification and using (5.7), we obtain the upper bound which is of order $t^{-d/2}$. On other hand, by Corollary 5.7 and [36, Lemma 2.1.4], we obtain the exponential term on the right hand side. \square

5.4 Gaussian bound and uniform Hölder continuity of p^ε

Equipped with the Nash inequality (5.9) and the Poincaré inequality (5.8), one can follow a now standard procedure (see, for example, [10] or [17]) to obtain two sided Gaussian estimates for p^ε . In the following, C_1 , C_2 and ε_0 are positive constants which depends only on d , D , ρ and T .

More precisely, we only need the Nash inequality (5.9) and Davies' method to obtain the following Gaussian upper bound.

Theorem 5.10. *There exist constants $C_i = C_i(d, D, \rho, T) > 0$, $i = 1, 2$, and $\varepsilon_0 = \varepsilon_0(d, D, \rho, T) \in (0, 1]$ such that*

$$p_{aD}^{a\varepsilon}(t, x, y) \leq \frac{C_1}{(a\varepsilon \vee t^{1/2})^d} \exp\left(\frac{C_2}{a^2}t - \frac{|y-x|^2}{(a\varepsilon)^2 \vee t}\right)$$

for every $t \geq a\varepsilon$, $\varepsilon \in (0, \varepsilon_0)$, $a > 0$ and $x, y \in aD^\varepsilon$. Moreover, the following weaker bound holds for all $t > 0$:

$$p_{aD}^{a\varepsilon}(t, x, y) \leq \frac{C_1}{(a\varepsilon \vee t^{1/2})^d} \exp\left(\frac{C_2}{a^2}t - \frac{|y-x|}{a\varepsilon \vee t^{1/2}}\right).$$

In particular, this implies the upper bound in Theorem 2.9 which is the case when $a = 1$.

We can then apply the Poincaré inequality (5.8) and argue as in section 3 of [17] to obtain the near diagonal lower bound. A more comprehensive proof is given in [23].

Lemma 5.11.

$$p^\varepsilon(t, x, y) \geq \frac{C_2}{(\varepsilon \vee t^{1/2})^d}$$

for every $(t, x, y) \in (0, \infty) \times D^\varepsilon \times D^\varepsilon$ with $|x - y| \leq C_1 t^{1/2}$ and $\varepsilon \in (0, \varepsilon_0)$.

The Gaussian lower bound for p^ε in Theorem 2.10 then follows from the Lipschitz property of D and a well-known chaining argument (see, for example, page 329 of [38]). Therefore, we have the two-sided Gaussian bound for p^ε as stated in Theorem 2.9 and Theorem 2.10. It then follows from a standard ‘oscillation’ argument (cf. Theorem 1.31 in [39] or Theorem II.1.8 in [38]) that p^ε is Hölder continuous in (t, x, y) , uniformly in ε . More precisely,

Theorem 5.12. *There exist positive constants $\gamma = \gamma(d, D, \rho)$, $\varepsilon_0(d, D, \rho)$ and $C(d, D, \rho)$ such that for all $\varepsilon \in (0, \varepsilon_0)$, we have*

$$|p^\varepsilon(t, x, y) - p^\varepsilon(t', x', y')| \leq C \frac{(|t - t'|^{1/2} + \|x - x'\| + \|y - y'\|)^\gamma}{(t \wedge t')^{\sigma/2} [1 \wedge (t \wedge t')^{d/2}]}. \quad (5.11)$$

5.5 Proof of local CLT

The following weak convergence result for RBM with drift is a natural generalization of [6, Theorem 3.3].

Theorem 5.13. *Let $D \subset \mathbb{R}^d$ be a bounded domain whose boundary ∂D has zero Lebesgue measure. Suppose D also satisfies:*

$$C^1(\overline{D}) \text{ is dense in } W^{1,2}(D).$$

Suppose $\rho \in W^{1,2}(D) \cap C^1(\overline{D})$ is strictly positive. Then for every $T > 0$, as $k \rightarrow \infty$,

- (i) $(X^{2^{-k}}, \mathbb{P}_m)$ converges weakly to the stationary process (X, \mathbb{P}_ρ) in the Skorokhod space $D([0, T], \overline{D})$.
- (ii) $(X^{2^{-k}}, \mathbb{P}_{x_k})$ converges weakly to (X, \mathbb{P}_x) in the Skorokhod space $D([0, T], \overline{D})$ whenever x_k converges to $x \in D$.

Proof For (i), the proof follows from a direct modification of the proof of [6, Theorem 3.3]. Recall the definition of the one-step transition probabilities p_{xy} , defined in the paragraph that contains (2.4) and (2.5). Observe that, since $\rho \in C^1(\overline{D})$, approximations using Taylor's expansions in the proofs of [6, Lemma 3.1 and Lemma 2.2] continue to work with the current definition of p_{xy} . Thus we have

$$\begin{aligned} \lim_{k \rightarrow \infty} \mathcal{E}^{2^{-k}}(f, f) &= \frac{1}{2} \int_D |\nabla f(x)|^2 \rho(x) dx, \quad \forall f \in C^1(\overline{D}), \quad \text{and} \\ \lim_{k \rightarrow \infty} L^{(2^{-k})} f &= \frac{1}{2} \Delta f + \frac{1}{2} \nabla(\log \rho) \cdot \nabla f \quad \text{uniformly in } D, \quad \forall f \in C_c^\infty(D). \end{aligned}$$

The process X^ε has a Lévy system $(N^\varepsilon(x, dy), t)$, where for $x \in D^\varepsilon$,

$$N^\varepsilon(x, dy) = \frac{d}{\varepsilon^2} \sum_{z: z \leftrightarrow x} p_{xz} \delta_{\{z\}}(dy).$$

Following the same calculations as in the proof of [6, Theorem 3.3], while noting that [13, Theorem 6.6.9] (in place of [6, Theorem 1.1]) can be applied to handle general symmetric reflected diffusions as in our present case, we get part (i). Part (ii) follows from part (i) by a localization argument (cf. [7, Remark 3.7]). \square

We can now present the proof of the local CLT.

Proof of Theorem 2.12. For each $\varepsilon > 0$ and $t > 0$, we extend $p^\varepsilon(t, \cdot, \cdot)$ to $\overline{D} \times \overline{D}$ in such a way that p^ε is nonnegative and continuous on $(0, \infty) \times \overline{D} \times \overline{D}$, and that both the maximum and the minimum values are preserved on each cell in the grid $\varepsilon \mathbb{Z}^d$. This can be done in many ways, say by the interpolation described in [2], or a sequence of harmonic extensions along the simplexes (described in [23]).

Consider the family $\{t^{d/2} p^\varepsilon\}_\varepsilon$ of continuous functions on $(0, \infty) \times \overline{D} \times \overline{D}$. Theorem 2.9 and Theorem 5.12 give us uniform pointwise bound and equi-continuity respectively. By Arzela-Ascoli Theorem, it is relatively compact. i.e. for any sequence $\{\varepsilon_n\} \subset (0, 1]$ which decreases to 0, there is a subsequence $\{\varepsilon_{n'}\}$ and a continuous $q : (0, \infty) \times \overline{D} \times \overline{D} \rightarrow [0, \infty)$ such that $p^{\varepsilon_{n'}}$ converges to q locally uniformly.

On other hand, by part (ii) of Theorem 5.13, if the original sequence $\{\varepsilon_k\}$ is a subsequence of $\{2^{-k}\}$, then $q = p$. More precisely, the weak convergence implies that for all $t > 0$,

$$\int_D \phi(y) p(t, x, y) dy = \int_D \phi(y) q(t, x, y) dy \quad \text{for all } \phi \in C_c(D) \text{ and } x \in D.$$

Then by the continuity of both p and q in the second coordinate, we have $q = p$ on $(0, \infty) \times D \times \overline{D}$. Since $p(t, \cdot, \cdot)$ and $q(t, \cdot, \cdot)$ are continuous on $\overline{D} \times \overline{D}$ (cf. [1]), we obtain $p = q$ on $(0, \infty) \times \overline{D} \times \overline{D}$. In conclusion, we have p^ε converges to p locally uniformly through the sequence $\{\varepsilon_n = 2^{-n}; n \geq 1\}$. \square

Acknowledgements: We thank Krzysztof Burdzy, Rekhe Thomas and Tatiana Toro for helpful discussions. In particular, we are grateful to David Speyer for pointing out the recurrence relation (3.39)

to us. We also thank Guozhong Cao, Samson A. Jenekhe, Christine Luscombe, Oleg Prezhdo and Rudy Schlaf for discussions on solar cells. Financial support from NSF Solar Energy Initiative grant DMR-1035196 as well as NSF grant DMS-1206276 is gratefully acknowledged.

References

- [1] R.F. Bass and P. Hsu. Some potential theory for reflecting Brownian motion in Hölder and Lipschitz domains. *Ann. Probab.* **19** (1991), 486-508.
- [2] R.F. Bass and T. Kumagai. Symmetric Markov chains on Z^d with unbounded range. *Trans. Amer. Math. Soc.* **360** (2008), 2041-2075.
- [3] D.J. Blount. Comparison of stochastic and deterministic models of a linear chemical reaction with diffusion. *Ann. Probab.* **19** (1991), 1440-1462.
- [4] C. Boldrighini, A. De Masi and A. Pellegrinotti. Nonequilibrium fluctuations in particle systems modelling reaction-diffusion equations. *Stoch. Proc. Appl.* **42** (1992), 1-30.
- [5] C. Boldrighini, A. De Masi, A. Pellegrinotti and E. Presutti. Collective phenomena in interacting particle systems. *Stoch. Proc. Appl.* **25** (1987), 137-152.
- [6] K. Burdzy and Z.-Q. Chen. Discrete approximations to reflected Brownian motion. *Ann. Probab.* **36** (2008), 698-727.
- [7] K. Burdzy and Z.-Q. Chen. Reflected random walk in fractal domains. *Ann. Probab.* **41** (2011), 2791-2819.
- [8] K. Burdzy, R. Holyst and P. March. A Fleming-Viot particle representation of Dirichlet Laplacian. *Comm. Math. Phys.* **214** (2000), 679-703.
- [9] K. Burdzy and J. Quastel. An annihilating-branching particle model for the heat equation with average temperature zero. *Ann. Probab.* **34** (2006), 2382-2405.
- [10] E.A. Carlen, S. Kusuoka and D.W. Stroock. Upper bounds for symmetric Markov transition functions. *Ann. Inst. Henri Poincaré-Probab. Statist.* **23** (1986), 245-287. MR 898496.
- [11] M.-F. Chen. *From Markov Chains to Non-Equilibrium Particle Systems. 2nd ed.* World Scientific, 2003.
- [12] Z.-Q. Chen. On reflecting diffusion processes and Skorokhod decompositions. *Probab. Theory Relat. Fields* **94** (1993), 281-316.
- [13] Z.-Q. Chen and M. Fukushima. *Symmetric Markov Processes, Time Change and Boundary Theory.* Princeton. University Press, 2012.
- [14] J.T. Cox, R. Durrett and E. Perkins. Voter model perturbations and reaction diffusion equations. *Astérisque* **349**, 2013.
- [15] G. David. Morceaux de graphes Lipschitziens et intégrales singulières sur une surface. *Rev. Mat. Iberoamericana.* **4** (1988), 73114 (French). MR 1009120.
- [16] G. David and S. Semmes. Singular integrals and rectifiable sets in \mathbb{R}^n : Beyond Lipschitz graphs. *Astérisque.* **193** (1991), 152. MR 1113517 (92j:42016)9120.
- [17] T. Delmotte. Parabolic Harnack inequality and estimates of Markov chains on graphs. *Revista Matemática Iberoamericana.* **15**(1) (1999), 181-232.

- [18] P. Dittrich. A stochastic model of a chemical reaction with diffusion. *Probab. Theory Relat. Fields.* **79** (1988), 115-128.
- [19] P. Dittrich. A stochastic partial system: Fluctuations around a nonlinear reaction-diffusion equation. *Stochastic Processes. Appl.* **30** (1988), 149-164.
- [20] R. Durrett and S. Levin. The importance of being discrete (and spatial). *Theoretical population biology* **46.3** (1994), 363-394.
- [21] L. Erdős, B. Schlein and H. T. Yau. Derivation of the cubic non-linear Schrödinger equation from quantum dynamics of many-body systems. *Inventiones Mathematicae* **167**(3) (2007), 515-614.
- [22] S.N. Ethier and T.G. Kurtz. *Markov processes. Characterization and Convergence.* Wiley, New York, 1986. MR0838085.
- [23] W.-T. Fan. Systems of reflected diffusions with interactions through membranes. PhD thesis. In preparation, 2014.
- [24] F. Golse. Hydrodynamic limits. *Euro. Math. Soc.* **1** (2005), 699-717.
- [25] M.Z. Guo, G.C. Papanicolaou and S.R.S. Varadhan. Nonlinear diffusion limit for a system with nearest neighbor interactions. *Comm. Math. Phys.* **118** (1988), 31-59.
- [26] P. Gyrya and L. Saloff-Coste. *Neumann and Dirichlet Heat Kernels in Inner Uniform Domains.* Paris: Socit mathmatique de France, 2011.
- [27] C. Kipnis and C. Landim. *Scaling Limits of Interacting Particle Systems.* Springer, 1998.
- [28] C. Kipnis, S. Olla and S.R.S. Varadhan. Hydrodynamics and large deviations for simple exclusion process. *Comm. Pure Appl. Math.* **42** (1989), 115-137.
- [29] P. Kotelenez. Law of large numbers and central limit theorem for linear chemical reactions with diffusion. *Ann. Probab.* **14** (1986), 173-193
- [30] P. Kotelenez. High density limit theorems for nonlinear chemical reactions with diffusion. *Probab. Theory Relat. Fields* **78** (1988), 11-37.
- [31] T. Kurtz. Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J. Appl. Probab.* **8** (1971), 344-356.
- [32] T. Kurtz. *Approximation of Population Processes.* SIAM, Philadelphia, 1981.
- [33] R. Lang and N.X. Xanh. Smoluchowski's theory of coagulation in colloids holds rigorously in the Boltzmann-Grad-limit. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete.* **54** (1980), 227-280.
- [34] R. M. May and M. A. Nowak. Evolutionary games and spatial chaos. *Nature* **359.6398** (1992), 826-829.
- [35] V.G. Mazja. *Sobolev Spaces.* Springer Verlag, 1985
- [36] L. Saloff-Coste. *Lectures on Finite Markov Chains.* Lecture Notes in Math., Springer, 1997.
- [37] R.P. Stanley. *Enumerative Combinatorics II.* Cambridge University Press, Cambridge, 1999.
- [38] D.W. Stroock. Diffusion semigroups corresponding to uniformly elliptic divergence form operators. *Séminaire de Probabilités XXII.* Lecture Notes in Math., vol. 1321, pp. 316-347, Springer, Berlin, 1988.
- [39] D.W. Stroock and W. Zheng. Markov chain approximations to symmetric diffusions. *Ann. Inst. Henri. Poincar-Probab. Statist.* **33** (1997), 619-649. MR 1473568.

- [40] H.T. Yau. Relative entropy and hydrodynamics of Ginzburg-Landau models. *Lett. Math. Phys.* **22** (1991), 63-80.

Zhen-Qing Chen

Department of Mathematics, University of Washington, Seattle, WA 98195, USA

Email: zqchen@uw.edu

Wai-Tong (Louis) Fan

Department of Mathematics, University of Washington, Seattle, WA 98195, USA

Email: louisfan@math.washington.edu