Some results in weak KPZ universality

Shalin Parekh

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2022

Abstract

Some results in half-space KPZ universality

Shalin Parekh

Stochastic partial differential equations (SPDEs) are a central object of study in the field of stochastic analysis. Their study involves a number of different tools coming from probability theory, functional analysis, harmonic analysis, statistical mechanics, and dynamical systems. Conversely SPDEs are an extremely useful paradigm to study scaling limit phenomena encountered throughout many other areas of mathematics and physics. The present thesis is concerned mainly with one particular SPDE, introduced in [99], called the Kardar-Parisi-Zhang (KPZ) equation, which appears universally as a fluctuation limit of height profiles of microscopic models such as interacting particle systems, directed polymers, and corner growth models. Such limit results are deemed instances of “weak KPZ universality," a field born from the seminal paper of Bertini and Giacomin [16]. We extend results on weak KPZ universality in a number of different directions. In one direction, we prove a version of Bertini-Giacomin’s result in a half-space by adapting their methods to this setting, thus extending a result of [45] and completing the final step towards the proof of a conjecture from [15]. In another direction, we also prove a result for the free energy for directed polymers in an octant converging to the KPZ equation in a half-space with a nontrivial normalization at the boundary. In a third direction, we return to the whole-space regime and extend the Bertini-Giacomin result to the case of several different initial data coupled together, proving joint convergence of ASEP with its basic coupling to KPZ driven by the same realization of its noise. Finally we prove a “nonlinear" version of the law of the iterated logarithm for the KPZ equation in a weak-noise but strong-nonlinearity regime. Beyond their intrinsic purpose, one application of all these extensions and generalizations is to take limits of known results and identities for discrete systems and pass them to the limit to obtain nontrivial information about the KPZ equation itself, which is a well-known methodology launched by I. Corwin and coauthors [40].

Table of Contents

List of Figures

Acknowledgements

I am grateful for many useful conversations and help from my advisor Ivan Corwin. He has supported me during my PhD and introduced me to the modern paradigms of mathematical research. I am also tremendously thankful for many useful and enjoyable mathematical interactions with the students and faculty at Columbia university. Specifically I want to thank Sayan Das, Evgeni Dimitrov, Hindy Drillick, Hugo Falconet, Georgy Gaitsgori, Yujin Kim, Ioannis Karatzas, Yier Lin, Konstantin Matetski, Mark Rychnovsky, Hao Shen, Li-Cheng Tsai, Xuan Wu, and Weitao Zhu. I am also greatly indebted to my former teachers Chris Bishop, Dmitry Chelkak, Hao Wu, and Gordan Zitkovic who introduced me to various areas of modern probability theory and encouraged me to pursue PhD studies. Finally, I would like to express my gratitude to my family for their support over the years.

Chapter 1: Introduction and Background

This thesis presents a collection of results about scaling limits of weakly asymmetric systems arising in probability theory and statistical mechanics. The central object throughout this thesis will be the Kardar-Parisi-Zhang (KPZ) equation. This is a stochastic PDE first introduced in [99] that arises naturally as a scaling limit of discrete growth models and interacting particle systems. The projects described herein involve proving convergence of some of these models to the KPZ equation and its variants, and as corollaries taking limits of discrete identities for these models in order to prove nontrivial continuum identities.

The KPZ equation can be formally written as a PDE on the domain $\mathbb{R}_+ \times \mathbb{R}$ where intuitively the $\mathbb{R}_+$ coordinate stands for the time variable and the $\mathbb{R}$ coordinate stands for the spatial variable. It is given by

$$\partial_t h(t,x) = \partial_x^2 h(t,x) + (\partial_x h(t,x))^2 + \xi(t,x),$$

where ξ is a space-time white noise, that is, a random function characterized non-rigorously by the property that its values are independent at distinct space-time values. One can formalize this by realizing the noise ξ as a Gaussian measure on some infinite dimensional vector space of Schwarz distributions with covariance structure given by the $L^2(\mathbb{R}_+ \times \mathbb{R})$ inner product.

In light of the physical relevance of the KPZ equation, one often imagines the solution h as a time-evolving height profile. The equation formally says that, starting from a given initial height profile $h(0,x)$, the time evolution of the height profile is governed by three separate phenomena: a smoothing mechanism governed by the heat operator, a “lateral growth" (also called “slope-dependent growth") mechanism governed by the nonlinear term, and a random forcing governed by the noise term.

1.1 The Kardar-Parisi-Zhang Equation

In the mathematical physics literature, the relevance of the KPZ equation is that it arises “universally" as a scaling limit of a vast collection of models, conjecturally any model that exhibits the three mechanisms described above, namely smoothing, *weakly* slope-dependent growth, and random forcing (noise). The weakly slope-dependent growth refers to the fact that one needs to tune the asymmetry of any given model to zero (at a very specific rate) to find KPZ fluctuations in the limit. This reflects the fact that the KPZ equation interpolates as a one-parameter orbit between two bona fide fixed points, a symmetric one and a strongly asymmetric one. The former is the so-called Edwards-Wilkinson fixed point (also known as the additive-noise stochastic heat equation) which is a Gaussian object that appears in the small-time limit of the KPZ equation under a 1:2:4 scaling of the exponents in fluctuations, space, and time respectively. This object appears as the limit of models exhibiting smoothing, *symmetric* growth, and random forcing. The other is the KPZ fixed point which appears as the large-time limit of the KPZ equation under a 1:2:3 scaling of the exponents in fluctuations, space, and time respectively (one also needs to subtract a divergent height shift). This object has been rigorously constructed only recently [54, 117] and conjecturally appears as the limit of models exhibiting smoothing, *strongly* slope-dependent growth, and random forcing. See [40] for more detailed information about this crossover. The recent open problems in the area have been to extend the class of models for which the convergence to the KPZ equation and KPZ fixed point are known, and in this thesis we will present a few steps toward the former.

Despite the fairly straightforward interpretation of the KPZ equation, it is classically ill-posed due to the singularity of the noise term ξ, and making sense of it (let alone proving convergence of models to it) is a tremendous challenge in and of itself. Instead one often defines the *Hopf-Cole solution* of the KPZ equation as $h := \log(Z)$ where Z solves the multiplicative-noise stochastic heat equation (SHE)

$$\partial_t Z(t,x) = \partial_x^2 Z(t,x) + Z(t,x)\xi(t,x).$$

This is still classically ill-posed because ξ is a Schwartz distribution and not a proper function, but one can nonetheless make sense of this equation more sensibly than working directly with the KPZ equation itself, exploiting the classical Itô-Walsh stochastic calculus as developed in [150]. Specifically one defines the solution to be any random function Z, adapted to the filtration of the noise ξ, and satisfying the Duhamel relation

$$Z(T,X) = \int_{\mathbb{R}} P_T(X-Y)Z_0(Y)dY + \int_0^T \int_{\mathbb{R}} P_{T-S}(X-Y)Z(S,Y)\xi(dS,dY), \tag{1.1}$$

where $P_T(X) = (2\pi T)^{-1/2}e^{-X^2/2T}$ is the heat kernel. See Chapter 2.4 below where we discuss how to rigorously perform the Picard iteration to obtain a solution to the above Duhamel relation in some Banach space of functions. A result of [122] shows that the solution of the stochastic heat equation started from any positive measure necessarily stays strictly positive for all $t > 0$ and $x \in \mathbb{R}$, hence we may indeed take the logarithm to obtain a solution of the KPZ equation. This method of constructing solutions for the KPZ equations is called the Hopf-Cole solution and is done in most of the projects described below.

One exception is Chapter 5 below, where we present some joint work with Yier Lin, in which we work directly with the KPZ equation using Hairer's theory of regularity structures [83]. This is the only chapter in which we work directly with the KPZ equation and so let us briefly describe our result there and how is a manifestation of weak universality. To motivate our result, recall that the law of the iterated logarithm for Brownian motion states that if B is a standard Brownian motion then $\limsup_{t\to 0}(2t\log\log(1/t))^{-1/2}B_t = 1$. Strassen in a seminal work [141] generalized this statement to show the functional form of this statement, namely that if we let $B^\epsilon(t) = \epsilon^{-1/2}B(\epsilon t)$ then the set of limit points as $\epsilon \to 0$ in $C[0,1]$ of the sequence $\{(2\log\log(1/\epsilon))^{-1/2}B^\epsilon\}_\epsilon$ is almost surely equal to the unit ball of its Cameron-Martin space. Our goal in Chapter 5 is to prove a “nonlinear" version of Strassen’s law in which the fluctuations respect the Burgers-like structure of the KPZ equation. We thus formulate the following.

For $\epsilon \in (0, e^{-e}]$ let $C_\epsilon := (\log\log(1/\epsilon))^{1/2}$, and let h^ϵ denote the Hopf-Cole solution to the KPZ equation

$$\partial_t h^\epsilon = \partial_x^2 h^\epsilon + C_\epsilon(\partial_x h^\epsilon)^2 + \xi,$$

with initial data $h(0,x) = 0$. Then for any $s, y \geq 0$ the set of limit points as $\epsilon \downarrow 0$ in $\mathcal{C}_{s,y} := C([0,s] \times [-y,y])$ of the sequence of functions $C_\epsilon^{-1}\epsilon^{-1/2}h^\epsilon(\epsilon^2 t, \epsilon x)$ is a.s. equal to the compact set K_{Zero} given by the closure in $\mathcal{C}_{s,y}$ of the set of smooth functions h satisfying

$$h(0,x) = 0, \qquad \left\|\partial_t h - \partial_x^2 h - (\partial_x h)^2\right\|_{L^2([0,s]\times[-y,y])} \leq 1.$$

If we instead let $h^\epsilon(0,x)$ be a two sided Brownian motion (fixed for different values of ϵ) then the same result holds but with compact limit set K_{Br} given by the closure of smooth functions $h \in \mathcal{C}_{s,y}$ satisfying

$$h(0,0) = 0, \qquad \|\partial_x h(0,\cdot)\|_{L^2[-y,y]} \leq 1, \qquad \left\|\partial_t h - \partial_x^2 h - (\partial_x h)^2\right\|_{L^2([0,s]\times[-y,y])} \leq 1.$$

If we likewise define k^ϵ to be the solution of

$$\partial_t k^\epsilon = \partial_x^2 k^\epsilon + C_\epsilon^{-1}(\partial_x k^\epsilon)^2 + \xi,$$

then the same compact limit set results hold for $C_\epsilon^{-1}\epsilon^{1/2}k^\epsilon(\epsilon^{-2}t, \epsilon^{-1}x)$. Moreover, the same results hold in stronger topologies given by parabolic Holder seminorms up to but excluding exponent 1/2 (see Definition 5.4.7).

Note that in the first result stated above (for the family h^ϵ), the nonlinearity must be scaled along with the parameter ϵ, so that it blows up in the $\epsilon \to 0$ limit. If we did not do this then the limiting compact set would simply agree with that of the linearized equation $\partial_t h_{Linear} = \partial_x^2 h_{Linear} + \xi$, namely

$$K_{Linear} := \{h \in \mathcal{C}_{s,y} : h(0,x) = 0, \big\|\partial_t h - \partial_x^2 h\big\|_{L^2([0,s]\times[-y,y])} \le 1\}.$$

Indeed this can be proved by decomposing $h_{KPZ} = h_{Linear} + v$ where h_{KPZ} is the Hopf-Cole solution to KPZ with initial data zero and v is a remainder term which has better regularity than h_{Linear} (see e.g. Theorem 3.19 of [132]). Then under the scaling necessary to obtain Strassen's law, it is easy to check that the remainder term converges a.s. to zero in the topology of $C([0,s]\times[-y,y])$ and the set of limit points for the part corresponding to h_{Linear} can be shown to be K_{Linear} by applying Proposition 5.1.2 (see Example 5.4.6 below).

Likewise in the second result stated above (for the family k^ϵ), the nonlinearity must be scaled along with the parameter ϵ so that it vanishes in the $\epsilon \to 0$ limit. If we did not do this then the asymptotics would be wrong entirely and instead one would need to apply a scaling that respects the tail behavior of the KPZ fixed point [117], namely $(\log\log(1/\epsilon))^{2/3}$, and this is done in [53].

The fact that the nonlinearity of the KPZ equation must be scaled along with the parabolic scaling of space-time to obtain a nontrivial limit set in the Strassen law can be seen as a manifestation of weak KPZ universality, which roughly states that the KPZ equation is only scale-invariant up to a one-parameter family of equations which interpolates between two *bona fide* fixed points [40]. The manner in which we prove the above theorem uses the theory of regularity structures [86] and is robust enough to prove similar theorems for other rough equations such as Φ^4_2.

1.2 Interacting Particle Systems

An interacting particle system on the integer lattice is just a time-homogeneous Feller process on the state space $\{0,1\}^{\mathbb{Z}}$. In a seminal work [16], Bertini and Giacomin proved that a certain interacting particle system called the asymmetric simple exclusion process (ASEP) converges to the KPZ equation in a certain sense. The ASEP can be interpreted as a countable collection of particles performing asymmetric random walks on the integer lattice subject to hardcore repulsion, that is, jumps are suppressed whenever one particle tries to jump onto another. The repulsion limits the number of particles at each site to at most one, and makes the system physically interesting and difficult to analyze.

One may define a time-evolving height profile $h_t(x)$ to the particle system in a natural way so that $h_t(x+1) - h_t(x)$ is 1 if there is a particle at site x at time t and is -1 if there isn't. The dynamics of $h_t(x)$ can be independently described as follows: starting from some initial height profile $h_0(x)$ each *peak* turns into a *valley* at some exponential rate p and independently each valley turns into a peak at some exponential rate q. Here by peaks we mean points $x \in \mathbb{Z}$ such that $h_t(x)$ is larger than both $h_t(x-1)$ and $h_t(x+1)$, and likewise valleys are the locations where it is smaller, see Figure 2.1. Bertini and Giacomin showed that if one scales the system diffusively in time and space, and *if one tunes the parameters of the models simultaneously* at the appropriate rate, then nonlinear (non-Gaussian) fluctuations appear. Their precise result is that if $h_t^\epsilon(x)$ denotes the height function with $p = e^{\sqrt{\epsilon}}$ and $q = e^{-\sqrt{\epsilon}}$ then $\epsilon^{1/2} h^\epsilon_{\epsilon^{-2}t}(\epsilon^{-1}x) - \epsilon^{-1}t$ converges in law to the solution of the KPZ equation.

The proof of Bertini and Giacomin is done using the following idea: let $\eta \in \{0,1\}^{\mathbb{Z}}$ denote the configuration space, i.e., $\eta_t(x) = h_t(x+1) - h_t(x)$. Bertini and Giacomin obtain a discrete Hopf-Cole transform by defining

$$Z_t^\epsilon(x) = \exp\big(-\lambda h_t(x) + \nu t\big)$$

where λ, ν will be specified later. Then

$$dZ_t^\epsilon(x) = \Omega(x) Z_t^\epsilon(x) dt + dM_t(x)$$

with M_t a martingale, and

$$\Omega(x) = \nu + (e^{2\lambda} - 1)c^R(\eta, x) + (e^{-2\lambda} - 1)c^L(\eta, x)$$

where c^R and c^L are the left and right jump rates for the configuration η. Letting $p = e^{\sqrt{\epsilon}}/2$ and $q = e^{-\sqrt{\epsilon}}/2$, and

$$\nu = p + q - 1, \qquad \lambda = \frac{1}{2}\log\left(\frac{q}{p}\right)$$

we get that

$$\frac{1}{2}\Delta Z_t^\epsilon(x) = \frac{1}{2}Z_t^\epsilon(x)\big(e^{-\lambda\eta_t(x+1)} + e^{\lambda\eta_t(x)} - 2\big) = \Omega(x)Z_t^\epsilon(x)$$

which is shown by considering the four different possible cases for $(\eta_t(x), \eta_t(x+1)) \in \{-1, 1\}^2$. For this choice of λ, ν, we thus obtain

$$dZ_t^\epsilon(x) = \frac{1}{2}\Delta Z_t^\epsilon(x)dt + dM_t(x),$$

so that $Z_t^\epsilon(x)$ satisfies a discrete stochastic heat equation. One sees that the discrete martingales satisfy

$$\frac{d}{dt}\langle M(x), M(y)\rangle_t \approx \begin{cases} 0, & x \neq y \\ \epsilon Z_t(x)^2 - \nabla^+ Z_t(x)\nabla^- Z_t(x), & x = y \end{cases},$$

so that the noise is "approximately uncorrelated and multiplicative." This discrete heat equation can then be extensively analyzed using the Duhamel formula, from which convergence to KPZ can be shown purely through discrete heat kernel estimates and martingale analysis.

Now let us discuss exclusion systems with open boundaries and their SPDE limits. Although the spatial coordinate of the KPZ equation takes values on the full real-number line, one can ask whether there exists a sensible version of the equation in which the spatial variable takes values either on a half-space $[0,\infty)$ or on a bounded interval $[0,1]$. One needs to impose boundary conditions in order to make the equation on such a domain, but it can be sensibly done as in [45, 70]. One might ask what boundary conditions may arise naturally as the limit of discrete models such as interacting particle systems, and this is the subject of Chapter 2 below.

It turns out that the Neumann boundary condition (suitably interpreted) on the KPZ equation on a half-space can be made to arise as the limit of a natural probabilistic model called the open ASEP. Open ASEP is a version of ASEP in which the particles are confined to just the non-negative integers and one imposes a boundary condition at the zero site in which particles can be created or annihilated at any specified rates α and δ respectively. Half-line open ASEP may be viewed as a Markov process on the state space $\{0,1\}^{\mathbb{Z}_+}$. Our main result, which generalized [45] to the case of arbitrary boundary conditions, states that there are tunings of the creation and annihilation rates such that one can obtain the Neumann boundary KPZ equation with any desired boundary condition. However the assumptions required for the convergence result of [45] is that the Neumann boundary conditions are positive (both for the prelimiting model and consequently the continuum equation as well). This is necessary to ensure that certain heat kernel estimates are true, which are used crucially in proving tightness and identification of the limit points.

Our improvement of [45] is primarily a technical improvement of the heat kernel estimates so that they do not depend on the positivity of the Neumann boundary parameters. Specifically we develop an entire analysis for the deterministic Robin-boundary heat kernels in the discrete regime, and our estimates are robust enough that they can be passed to the limit and used in the continuum as well. We also clarify a number of omitted details and fix various small mistakes. Consequently we extend their convergence result to all possible values of boundary conditions. This made rigorous an important identity from [15] that allows to prove the long-time fluctuation behavior of half-space KPZ at the origin for the critical value of the boundary parameter $A=-1/2$ at which a certain interesting phase transition occurs in half-space KPZ models (see Chapter 2.1).

Following [45] we also prove a version of this result where there are two boundaries and the limiting object is the KPZ equation on a bounded interval with Neumann boundary conditions at both boundaries. The discrete system is depicted in Figure 1. This result also required us to improve a separate collection of very fine heat kernel estimates. Our bounded interval result was used in [42] in order to classify invariant measures for KPZ on a bounded interval by taking a limit of the invariant measures for ASEP, which are known. All of this will be done in Chapter 2 below.

We also extend the Bertini-Giacomin result in a slightly different direction in Chapter 4. The ASEP has a natural coupling of its dynamics when one starts from several distinct initial data simultaneously. Specifically when a particle from one configuration jumps from x to y then a particle from the other configuration also jumps from x to y, assuming that there is a particle at x in the second configuration and that site y is not blocked in the second configuration. This is called the *basic coupling* for exclusion dynamics and it was introduced by Liggett in the 70s in order to classify the steady states of exclusion systems [111]. We extend the convergence result of Bertini and Giacomin to show that convergence of the height functions does hold jointly to the solution of KPZ driven by the same noise. The two sequences of initial data can be anything near stationarity.

The manner in which we prove our result is by proving a result about space-time processes coupled onto the same probability space that individually solve the KPZ equation in law. Namely suppose we have two standard space-time white noises ξ^1, ξ^2 coupled onto the same probability space. Suppose furthermore that they satisfy the following conditions:

1. $\mathbb{E}[(\xi^1, f)(\xi^2, g)] = 0$ for all $f, g \in L^2(\mathbb{R}_+ \times \mathbb{R})$ which have disjoint supports.

2. For every $t > 0$ the spatial process $h^2(t, \cdot) - h^1(t, \cdot)$ has a.s. finite p-variation for some $p < 2$, where h^i is a solution of $\partial_t h^i = \partial_x^2 h^i + \lambda_i (\partial_x h^i)^2 + \xi^i$, for $i = 1, 2$. Here $\lambda_1, \lambda_2 \in \mathbb{R}$ and furthermore we assume Hölder continuity and sub-linear growth at infinity of the initial data $h^i(0, \cdot)$.

Then $\xi^1 = \xi^2$.

The intuition for this result is as follows: if the noises are uncorrelated or only partially correlated then the spatial process in the difference of the solutions of the KPZ equations will be locally Brownian with some diffusivity. One expects a sort of contrapositive to this, and hence our result. The proof uses a result of Perkowski and Rosati [132] which developed the pathwise solution theory for KPZ on the full line, similarly to Hairer's seminal work on the circle [86].

1.3 Directed Polymers

Analogous to the Bertini-Giacomin result for interacting particle systems, in [3] it was shown that the KPZ equation can also be made to arise as the limit of the partition function in a certain model of random walks in random environments called *directed polymers.* Directed polymers are natural probabilistic objects that were first introduced in [94, 96]. They generalize directed first- and last-passage percolation and have deep connections to statistical mechanics and stochastic analysis. Specifically, we consider an environment $\{\omega_{i,j}\}_{(i,j)\in\mathbb{Z}_{\geq 0}\times\mathbb{Z}}$ consisting of i.i.d., mean-zero, finite-variance random variables. The standard deviation of the weights is referred to as the *inverse temperature*. One may define a partition function $Z^\omega(n,x)$ as a sum over all *directed* nearest-neighbor simple random walk paths $(i,\gamma_i)_{0\leq i\leq n}$ of length n starting from $(0,x)$, of the product of all weights $e^{\omega_{i,\gamma_i}}$ along the path. Similarly, there is also a natural way to define *random* Markovian transition densities associated to this environment ω, wherein a nearest-neighbor path γ has probability proportional to the product of weights $e^{\omega_{i,\gamma_i}}$ along it. As is standard practice in statistical mechanics, one may then ask questions about the existence of infinite-volume limits of these path measures and their typical fluctuation scale, as well as the typical scale and shape of the fluctuations of the partition function itself [37].

Many seminal results in these directions have been proved, perhaps most notably that there is a phase transition which becomes apparent in high dimensions. Specifically, in spatial dimensions greater than two, there is a strictly positive critical value of the inverse temperature below which weak disorder holds, meaning that the fluctuations of a typical polymer path look like Brownian motion and one may construct infinite-length path measures [39, 37]. Such polymers are said to exhibit *weak disorder.* In contrast, lower-dimensional polymers at any nonzero inverse temperature are known to be characterized by *strong disorder*, meaning that the path fluctuations are quite different and there is no sensible notion of an infinite volume Gibbs measure [37]. The results of [3, 2] examined the partition function in a regime that lies *between* strong and weak disorder. Specifically, in spatial dimension one, they scaled the inverse temperature of the model like $n^{-1/4}$ and simultaneously applied diffusive scaling to the partition function, and there they observed that the fluctuations are governed by the multiplicative-noise stochastic heat equation and that the path measures themselves have a continuum analogue. This is called *intermediate disorder* and is a manifestation of weak KPZ universality.

To briefly describe the intermediate disorder result of [3] we consider an iid random environment $\{\omega_{ij}\}_{i\geq 0, j\in\mathbb{Z}}$. For a nearest-neighbor path $S=(S_i)_{i=1}^t$ on $\mathbb{Z}$, we define the weight of the path

$$\omega(S)=\prod_{i=1}^t \omega_{i,S_i}.$$

One then defines the directed polymer partition function as

$$Z^\omega(t,x)=\sum_{\substack{\text{paths } S \text{ from}\\ (0,0) \text{ to } (t,x)}} \omega(S).$$

Let $\{\eta_{ij}\}_{i\geq 0, j\in\mathbb{Z}}$ be an iid random environment of mean-zero variables, with variance β^2. Defining $Z_n^{\omega^n}(t,x)$ be the partition function associated to the **rescaled** environment

$$\omega_{ij}^n := 1+n^{-1/4}\eta_{ij},$$

the main result of [3] essentially states that $2^{-nT}n^{1/2}Z_n^{\omega^n}(nT,n^{1/2}X)$ converges in distribution as $n\to\infty$, to the solution of the multiplicative-noise stochastic heat equation (that is, the Hopf-Cole transform of the KPZ equation):

$$\partial_T Z = \partial_X^2 Z + \beta Z \xi,$$

with Dirac initial data. This is done by first writing

$$Z_n^{\omega^n}(t,x) = \mathbb{E}_{SSRW}\left[\prod_{i=1}^{n}(1+n^{-1/4}\eta_{i,S_i})\middle|\eta\right],$$

for a simple symmetric random walk (SSRW) denoted by (S_i) within the expectation. Then by expanding the product and writing the expression in terms of transition densities of the random walk, one obtains an n^{th} degree polynomial in the variables $\eta_{t,x}$, whose coefficients consist of products of transition densities. The authors then show, under the diffusive scalings of $Z_n^{\omega^n}$ defined above, term-by-term convergence of that series to an explicit random variable which admits a chaos series expansion in terms of iterated integrals against a Gaussian space-time white noise. The latter chaos series is none other than the Picard iteration of the Duhamel relation (1.1) for the stochastic heat equation, which gives a sketch of the required convergence.

In this thesis, the main result of Chapter 3 below is to take a more complicated intermediate disorder limit (involving inhomogeneous transition densities) of a nontrivial polymer identity attributable to symmetries arising from representation theory, specifically Proposition 8.1 from [14] which is depicted in Figure 3.1 below. It is an invariance in distribution for the partition function of a certain polymer model (with inverse gamma weights) in an octant under a switching of the boundary weights. The limiting identity turns out to relate the multiplicative-noise *half-space* stochastic heat equation with Dirichlet boundary condition to the same equation with Robin boundary condition (latter corresponding to the Neumann-boundary KPZ under the Hopf-Cole transform).

Specifically we prove the following. Let $Z_{Rob}^{(A)}(T,X)$ denote the solution of the multiplicative SHE with Robin boundary parameter A and delta initial data $Z_{Rob}^{(A)}(0,X) = \delta_0(X)$ (this is defined precisely in Chapter 2.4 below). Let $Z_{Dir}^{(A)}(T,X)$ be the solution to the multiplicative SHE with

Dirichlet (zero) boundary condition and initial data $Z^{(A)}_{Dir}(0,X) = e^{B_X - (A+\frac{1}{2})X}$, where B is a standard Brownian motion independent of ξ. Then for each boundary value $A \in \mathbb{R}$ and all $T \geq 0$ we have the following equality of distributions:

$$Z^{(A)}_{Rob}(T,0) \stackrel{d}{=} \lim_{X \to 0} \frac{Z^{(A)}_{Dir}(T,X)}{X}.$$

The main challenge in proving this was to find the limit of the Dirichlet side of the equation, one may note that the normalization by X near the origin makes the result more difficult. Convergence of polymers to Neumann KPZ on the left hand side was already known thanks to [151] based on the seminal work of [2].

Our limiting identity seems to have been used in the physics literature to derive some moment formulas and conjectural asymptotics for the half-space KPZ equation [101]. The ultimate motivation in proving this identity was to try to prove the law of large numbers and the Gaussian fluctuation behavior of the supercritical half-space KPZ equation at the origin as conjectured in [10], though this still seems out of reach.

1.4 Partitions

Returning to the topic of interacting particle systems, there is a canonical choice of initial data in ASEP and TASEP called the “narrow wedge," in which one has all particles to the left of the origin and only empty sites to the right. When started from narrow-wedge initial data, one can view TASEP and ASEP as a Markov process on partitions, by viewing a partition as a Young diagram oriented in the Russian convention. For TASEP in particular, this viewpoint is useful because the evolution up to time t can be discretized in time and then studied using the theory of Markov chains on Schur processes, see [25]. Thus one may want to study other measures on partitions, such as those weighted by the number of irreducible representations of the symmetric group associated to that partition, or more generally by the number of chains of subpartitions.

This will be done in Chapter 6. In joint work with Ivan Corwin [44], our goal was to answer a question posed to us by Richard Stanley: find the limit shape (as $n \to \infty$) of any sequence λ_n of partitions of n which maximizes the number of subpartitions among all other partitions of n. We found this problem to be interesting because it is a zero-temperature limit ($\beta \to \infty$) of the measure on partitions of n in which each partition λ has weight proportional to $s(\lambda)^\beta$ where $s(\lambda)$ is the number of subpartitions of λ. We expected to find a limit shape of the Young diagrams as $n \to \infty$ and the side lengths are normalized by $n^{1/2}$. Our methods are similar in principle to the one used for large deviations of the *uniform measure* on partitions of n, studied in [56] and the limit shape we obtained is actually the same one. Our methods are purely variational and analytic despite the problem having been motivated from exactly solvable models and algebraic tools.

Our limit shape, called *Vershik's Curve* can be described as the graph of the function $f(x) = \frac{2\sqrt{3}}{\pi}\log(2\cosh(\frac{\pi}{2\sqrt{3}}x))$ lying above the graph of $|x|$. One advantage of our formulation is that we can prove that the partition of n that maximizes k-chains of subpartitions also has the same limit shape, for any fixed value of k. One possible future direction is to try to study the maximizer of k-chains where k grows with n. We expect so see a nontrivial transition from Vershik's Curve to a different curve called the Vershik-Kerov-Logan-Shepp curve (see [148] and [115]) when $k = \lfloor \alpha n^{1/2} \rfloor$ for $\alpha > 0$. The limit shape should depend on α (and interpolate smoothly between the two curves depending on whether α is close to zero or to infinity), and we expect to be able to study this by looking at large deviations of domino tilings as in [36]. This particular result is not a manifestation of KPZ universality, rather it is only peripherally related in that it was motivated by methods that are related as described above, for instance via the relation to TASEP.

Chapter 2: KPZ limit of open exclusion processes with Neumann boundary

It was recently proved in [45] that under weakly asymmetric scaling, the height functions for open ASEP on the half-line and on a bounded interval converge to the Hopf-Cole solution of the KPZ equation with Neumann boundary conditions. In their assumptions [45] chose positive values for the Neumann boundary data, and they assumed initial data which is close to stationarity. By developing more extensive heat-kernel estimates, we extend their results to negative values of the Neumann boundary parameters, and we also show how to generalize their results to narrow-wedge initial data (which is very far from stationarity). As a corollary via [15], we obtain the Laplace transform of the one-point distribution for half-line KPZ, and use this to confirm $t^{1/3}$-scale GOE Tracy-Widom long-time fluctuations.

2.1 Introduction

2.1.1 Main Results

Fix parameters $\alpha, \gamma, p, q \geq 0$. We define half-space ASEP (or ASEP-H) to be the interacting particle system on positive integers $\mathbb{Z}_{>0}$ where particles at site $x > 1$ jump to site $x-1$ at exponential rate q if site $x-1$ is unoccupied, and particles at site $x > 0$ jump to site $x+1$ at rate p if site $x+1$ is unoccupied. If a neighboring site is occupied, then a particle feels no inclination to jump there (i.e., rate 0). All jumps are independent of each other. Furthermore, particles at site $x = 1$ are created at rate α if site $x = 1$ is unoccupied, and particles at site $x = 1$ are destroyed at rate γ if site $x = 1$ is occupied. We may define a height function associated to this particle system as follows: let $h_t(0)$ denote twice the number of particles annihilated minus twice the number of particles created at site $x = 0$ up to time t. Then define $h_t(x+1)$ to be $h_t(x) + 1$ if there is a particle at site $x+1$, and define $h_t(x+1)$ to be $h_t(x) - 1$ otherwise. By linear interpolation, we view the height function as a random element of $C([0,\infty),\mathbb{R})$.

The preceding paragraph gives a notion of ASEP on the half-line $\mathbb{Z}_{\geq 0}$, but we can also define a similar particle system on the bounded interval $\{0,...,N\}$. To do this, we fix two more parameters $\beta,\delta \geq 0$, and we create and annihilate particles at site N with rates δ,β, respectively. Our main result may now be stated as follows:

Theorem 2.1.1 (Main Result)**.** *Fix some parameters $A,B \in \mathbb{R}$. For half-line ASEP, we let $I = [0,\infty)$ and for bounded-interval ASEP, we set $I=[0,1]$. For $\epsilon > 0$, we define:*

$$p = \frac{1}{2}e^{\sqrt{\epsilon}}, \qquad q = \frac{1}{2}e^{-\sqrt{\epsilon}}, \qquad \mu_A = 1 - A\epsilon, \qquad \mu_B = 1 - B\epsilon.$$

Let us also define the creation/annihilation rates:

$$\alpha = \frac{p^{3/2}(p^{1/2} - \mu_A q^{1/2})}{p-q}, \qquad \beta = \frac{p^{3/2}(p^{1/2} - \mu_B q^{1/2})}{p-q},$$

$$\gamma = \frac{q^{3/2}(q^{1/2} - \mu_A p^{1/2})}{q-p}, \qquad \delta = \frac{q^{3/2}(q^{1/2} - \mu_B p^{1/2})}{q-p}.$$

Let $h_t^\epsilon(x)$ denote the height function associated to this particle system. We then define

$$H^\epsilon(T,X) := \epsilon^{1/2} h^\epsilon_{\epsilon^{-2}T}(\epsilon^{-1}X) - \big(\frac{1}{2}\epsilon^{-1} + \frac{1}{24}\big)T.$$

Assume that $H^\epsilon(0,X)$ is near-equilibrium (see Definition 2.5.2), and converges weakly to some initial data H_0. Then $H^\epsilon(T,X)$ converges in distribution to the Hopf-Cole solution of the KPZ equation (cf. Definition 2.4.1) on the interval I with Neumann boundary parameters A (at $X=0$) and B (at $X=1$ if $I=[0,1]$), started from H_0. The convergence occurs in the Skorokhod Space $D([0,\infty),C(I))$.

This result is proved as Theorem 2.5.7 below. Before this, we will first introduce the Hopf-Cole solution of KPZ with Neumann boundary conditions, and prove some existence results about it (see Section 4). We get a similar result when the initial data is not near equilibrium, but we must subtract a logarithmically-divergent height-shift:

Theorem 2.1.2 (Extension to Narrow-Wedge Initial data)**.** *Let A, B and I be as before. For $\epsilon > 0$, let $p, q, \alpha, \beta, \gamma, \delta$, and $h_t^\epsilon(x)$ be the same as in Theorem 2.1.1. We define*

$$H^\epsilon(T, X) := \epsilon^{1/2} h^\epsilon_{\epsilon^{-2}T}(\epsilon^{-1}X) - \big(\frac{1}{2}\epsilon^{-1} + \frac{1}{24}\big)T - \frac{1}{2}\log\epsilon.$$

Assume that we start from the initial configuration of zero particles $($i.e., $H^\epsilon(0, X) = -\epsilon^{-1/2}X)$. Then we have that $H^\epsilon(T, X)$ converges in distribution to the Hopf-Cole solution of the KPZ equation on the interval I with Neumann boundary parameters A $($at $X = 0)$ and B $($at $X = 1$ if $I = [0, 1])$. The initial data is narrow-wedge, and the convergence occurs in the Skorokhod Space $D((0, \infty), C(I))$, see Definition 2.6.4.

This will be proved as Theorem 2.6.8 below. See Section 6 for the precise result.

Let us discuss for a moment why our results are more general than those of [45]. There, the authors prove convergence of open ASEP to the KPZ equation under the assumption of non-negative values for the Neumann boundary conditions. However, the non-negativity only seems to be a technical restriction which simply makes the analysis a little bit easier. In other words, the zero value for the Neumann boundary condition does not actually correspond to any meaningful phase transition in the associated particle system (see Remark 2.11 in [45]), hence one would expect that it is just a superficial restriction. Thus, the main purpose of the current work is to generalize their results to the case when the boundary parameters for the PDE are negative. This corresponds to branching in the associated diffusion process, hence the associated kernels will be super-probability measures in general. This brings about some new challenges which were not seen in [45], and we develop several new heat-kernel estimates in order to resolve these challenges. We also found a few small mistakes in the proofs of tightness in [45, 47] and related papers; therefore we will show how to fix these issues in the current work (see the discussions at the beginning of the proofs of Propositions

2.5.4 and 2.6.2).

One further restriction which the authors used in [45] was to assume near-equilibrium initial data, which roughly means that the sequence of initial conditions of the particle system are close to stationarity. We also remove that assumption in the current work, and (as stated in Theorem 2.1.2 above) we extend their results to narrow-wedge initial data for the ASEP (which corresponds to δ_0 initial data for the stochastic heat equation).

It may seem a little bit uncertain what the immediate usage of such a technical generalization might be. The first application is the following theorem. It is the main result of [15] and was proved modulo the main result of our paper, Theorem 2.1.2. Specifically, it gives us the exact one-point statistics (via the moment-generating function) for half-line KPZ. We will combine our result with theirs and reiterate it here:

Corollary 2.1.3. *Let $H(T,X)$ denote the solution to the KPZ equation on $[0,\infty)$ with Neumann boundary parameter $A=-1/2$ (this choice of A corresponds to the triple point of half-line ASEP; see Section 1 and Remark 2.11 of [45]). Then for $\xi>0$ we have that*

$$\mathbb{E}\Big[\exp\Big(-\xi\exp\big(H(T,0)+\frac{T}{24}\big)\Big)\Big]=\mathbb{E}\Big[\prod_{k=1}^{\infty}\frac{1}{\sqrt{1+4\xi\exp[(T/2)^{1/3}\mathbf{a}_k]}}\Big]$$

where $\mathbf{a}_1>\mathbf{a}_2>\dots$ forms the GOE Point process (see Definition 6.1 in [15]).

As an easy corollary of this fact, we obtain a limit theorem for the half-space KPZ equation with Neumann$(-1/2)$ boundary condition on $[0,\infty)$. As expected we get a random-matrix-type of distribution in the limit.

Corollary 2.1.4. *Let $H(T,X)$ as in the previous theorem. Then one has weak convergence*

$$\mathbb{P}\Big(\frac{H(T,0)+\frac{T}{24}}{(T/2)^{1/3}}\leq x\Big)\overset{T\to\infty}{\longrightarrow}F_{GOE}(x)$$

where $F_{GOE}(x) = \mathbb{P}(\mathbf{a}_1 \leq x)$ is the Tracy-Widom GOE distribution.

Proof. This may be derived via Proposition 2.1.3 below together with [24, Lemma 4.1.39] in a manner which is now standard, see for instance [26, Corollary 2.4]. However we also give an original proof for completeness. Fix x, and let

$$Q(T) := H(T,0) + \frac{T}{24} - (T/2)^{1/3}x.$$

Replacing ξ with $\xi e^{-(T/2)^{1/3}x}$ in Corollary 1.3, we find that

$$\mathbb{E}[\exp(-\xi \exp(Q(T)))] = \mathbb{E}\left[\prod_{k=1}^{\infty} \frac{1}{\sqrt{1 + 4\xi \exp\left[(T/2)^{1/3}\left(\mathbf{a}_k - x\right)\right]}}\right].$$

Letting $T \to \infty$ on both sides, applying the dominated convergence theorem, and noting that $\mathbf{a}_1 = \max_k \mathbf{a}_k$, we find that for all $\xi > 0$,

$$\lim_{T\to\infty} \mathbb{E}[\exp(-\xi \exp(Q(T)))] = \mathbb{E}\left[1_{[\mathbf{a}_1 - x \leq 0]}\right] = \mathbb{P}\left(\mathbf{a}_1 \leq x\right).$$

At this point in the proof, we state a general fact: If $\{X_n\}$ is any collection of non-negative random variables such that for every $\xi > 0$ we have that $\mathbb{E}\left[e^{-\xi X_n}\right] \overset{n\to\infty}{\longrightarrow} c \in [0,1]$, then we necessarily have that X_n converges in distribution (as $n \to \infty$) to a random variable which equals 0 with probability c, and equals $+\infty$ with probability $1-c$. To prove this, denote by $C([0,\infty],\mathbb{R})$ the space of functions f which have a limit at infinity, and denote by $f(\infty)$ that limit. We clearly have that $\mathbb{E}\left[f\left(X_n\right)\right] \overset{n\to\infty}{\longrightarrow} cf(0) + (1-c)f(\infty)$ for any function $f \in C([0,\infty],\mathbb{R})$ which is a finite linear combination of functions of the form $x \mapsto e^{-\xi x}$, with $\xi \geq 0$. By Stone-Weierstrass, such functions are uniformly dense in $C([0,\infty],\mathbb{R})$, therefore we conclude the same result for every $f \in C([0,\infty],\mathbb{R})$. From the previous paragraph and the preceding computations, we conclude that

$$e^{Q(T)} \overset{d}{\longrightarrow} 0 \cdot 1_{[\mathbf{a}_1 \leq x]} + \infty \cdot 1_{[\mathbf{a}_1 > x]}, \qquad \text{as } T \to \infty.$$

Consequently, we find that

$$\lim_{T\to\infty} \mathbb{P}(Q(T) \leq 0) = \lim_{T\to\infty} \mathbb{P}\left(e^{Q(T)} \leq 1\right) = \mathbb{P}\left(\mathbf{a}_1 \leq x\right),$$

as desired. $\square$

Although the primary application of our results is the aforementioned limit theorem, we will mention that there are also other applications which come to mind. The main key behind our approach is the fine heat kernel estimates which are given in Section 3 below, and these are based on the work of [57] and [45]. These heat kernels $\mathbf{p}_t^R$ may be interpreted as transition probabilities for random walks with branching or killing in discrete-space/continuous-time. Indeed, positive choices for the boundary parameter A correspond to the random walk particle being killed at the boundary with some probability depending on A, whereas negative choices for A correspond to a new particle being created at the boundary at some rate depending on A (so we get a branching process). As an indirect corollary, we get some fairly intricate information about the behavior of continuous-time random walks and branching processes.

2.1.2 Historical Background

Interacting particle systems in one spatial dimension have been extensively studied in recent years. Of particular interest is the Asymmetric Simple Exclusion Process (ASEP), in which particles on $\mathbb{Z}$ jump independently to the left and right at exponential rates in a collectively Markovian way, subject to a drift in one chosen direction and with other particles acting as a deterrent to the general movement. Such processes fall within what has come to be known as the "KPZ Universality Class," a broad collection of space-time processes coming from both mathematics and physics, which are unified by some salient features [40]. The "universality" of this class refers to the *long-time* behavior of the model: how it looks on large scales in both time and space. Objects which fall into this class will generally have temporal fluctuations on the order of $t^{1/3}$ and the statistics will be described by probability distributions which come out of random matrix theory, such as the GUE and GOE eigenvalue distributions. Moreover, the space-time fluctuations of such objects are generally seen to converge under suitable *weak* scalings to the solution of the KPZ *equation*:

$$\partial_T H = \frac{1}{2}\partial_X^2 H + \lambda(\partial_X H)^2 + \xi$$

where ξ is a Gaussian space-time white noise and $\lambda > 0$ is a constant. The first major result in this direction was [16], where the authors used a Hopf-Cole transformation at the discrete level in order to obtain the KPZ equation as a limit of ASEP on the whole line. Since then, this discrete Hopf-Cole method has proved very useful in a number of generalizations, see for instance [5, 48, 57, 45, 47, 105]. Ever since the original Bertini-Giacomin result, a number of other methods have also been developed to find the KPZ equation arising in some kind of weakly asymmetric environment. In particular, the notion of energy solutions [73, 81] has been very useful under the assumption of stationary initial data, see for instance [61]. Furthermore, Hairer in [83] has recently developed a notion of regularity structures which has also proved successful in a number of weakly asymmetric results [90] directly at the level of the KPZ equation. The theory of paracontrolled distributions [79] has also yielded results in this direction, see [73] and [80] for one example of success in this area using the notion of “energy solution."

While the aforementioned results were either on the whole line or on the circle, the subject of the current paper is how to deal with ASEP on intervals with boundary (namely, $\mathbb{Z}_{\geq 0}$ and $\{0, ..., N\}$), and to prove a small step towards KPZ universality of such a model. The features of such a model were physically analyzed as early as the seventies: in [110], Liggett considered open ASEP on the half-line in the special case that $\frac{\alpha}{p} + \frac{\gamma}{q} = 1$, and he characterized the invariant measures for the process and found the domains of attractions of each invariant measure. Convergence to the KPZ equation was first considered and proved in the primary reference [45], in the case that the Neumann boundary parameters $A, B \geq 0$.

The key feature of open ASEP is that the boundary conditions are governed by *sources* and *sinks*, where particles are created and annihilated at certain exponential rates. Under the weak scaling, the choice of rates corresponds directly to choosing Neumann boundary parameters in the KPZ equation (or equivalently, choosing Dirichlet boundary conditions for the stochastic Burgers equation, the spatial derivative of KPZ).

Derrida [59] has pointed out that on a bounded interval, open ASEP exhibits three distinct *phases* as the number N of particle sites becomes very large. These phases are described by a statistic called the current, which may be defined as the limit (as $N \to \infty$) of the expected value (under the stationary measure) of the number of particles to cross a single bond in one unit of time (usually we also normalize by $(p-q)^{-1}$, to account for the asymmetry). These three phases may be described as high-density, low-density, and maximal current (we will not give many details but see and Section 1 of [45]). In essence, the high-density phase occurs when the destruction rate β at the right boundary is very small, hence particles do not leave the system fast enough, causing jamming. In contrast, the low-density phase occurs when the creation rate α at the left boundary is very small, thus particles are created so slowly that system still lacks efficiency. The maximal-current phase is characterized by quick creation at the left boundary and quick destruction at the right boundary, so that the system operates at maximal efficiency.

To put our work into the physical context described in the preceding paragraph, our main result (Theorem 2.1.1) essentially states that the KPZ equation describes the height-function fluctuations as we zoom into an $\epsilon^{1/2}$-window around the critical triple point between the three phases on the open ASEP phase diagram, see page 4 of [45]. The boundary parameters A and B govern the "direction" from which we approach the triple point.

We note that [26] and [22] contained the first results which recognized that the Laplace transform of various one-point distributions appearing in the KPZ universality class may be written as a multiplicative functional of some determinantal point process (usually Airy or GOE/GUE), and indeed [15] builds on such types of results. We should also mention that we are not the first ones to study the heat kernels for random walks or Brownian motion with branching; [129] studies Neumann-boundary heat kernels in continuous time and space. His work is for general bounded domains in $\mathbb{R}^d$, though his estimates are not as intricate as ours.

Two weeks prior to posting this preprint, there was a related result which was posted by Gonçalves, Perkowski, and Simon in [74]. In that paper, they prove (under the assumption of stationary initial data and boundary parameters $A = B = -\frac{1}{2}$) weakly asymmetric convergence of ASEP to stochastic Burgers using very different methods than the ones we use here (the Hopf-Cole transform). Instead, they use the notion of energy solutions which was developed in [73] and proved to be unique in [81]. While their method has the advantage of avoiding the Hopf-Cole transform, it is only able to deal with stationary initial data due to the fact that they use various hydrodynamic estimates which are only available for stationary (Bernoulli) initial data with boundary parameters $A = B = -\frac{1}{2}$. We should mention that the energy solutions in [74] contain additional error terms which are not observed in our work. These additional terms may effectively be seen as Itô corrections which arise since we work at the level of the SHE and they work directly with KPZ.

Outline: The paper will be organized as follows. In Section 2 we will formally define ASEP on the half-line and on bounded intervals. Furthermore, we will define the appropriate scaling of the model under which it converges to the KPZ equation. Section 3 contains a number of heat kernel estimates for the Robin-boundary Laplacian in discrete-space and continuous-time. We also give a construction of the continuum Robin heat kernel and prove a couple of useful estimates about it. Section 3 is the main key behind our approach, hence this is the longer section which is extensively used in the remainder of the paper. In Section 4, we will define what it means to solve the KPZ equation on $[0,\infty)$ and $[0,1]$ with Neumann boundary conditions. This involves the usual notions of mild solutions and weak solutions. In Section 5, we will prove that the model defined in Section 2 converges to the KPZ equation with Neumann boundary conditions, under the assumption of "near-equilibrium" initial data. The key behind this is to prove some uniform Hölder estimates for the weakly-scaled, exponentially transformed height functions of the ASEP, and to prove a "crucial cancellation" as in Section 5 of [45]. The key novelty of our approach will be that we avoid the Green's function analysis of [45] but instead rely purely on the heat kernel estimates to prove this cancellation. In Section 6, we will try to generalize our results to the case when the initial data in the stochastic heat equation is δ_0, which is very far from equilibrium. The idea here is to note that in short time, the solution to the stochastic heat equation started from δ_0 stabilizes to equilibrium.

Acknowledgements: The author wishes to thank Ivan Corwin for suggesting the problem, for providing helpful discussions about various issues which came up during the writing of the paper, and also for thoroughly reading the preliminary drafts of this paper. We also wish to thank Hao Shen and Li-Cheng Tsai, who provided some very useful discussions.

2.2 Definition of the Model and Scalings

Much of this section is adapted directly from the primary reference [45], but we reiterate all of the definitions provided there (albeit emphasizing different points) so that the reader will have easy access to this material.

Let us first introduce some notational conventions. Firstly, we will always use lowercase letters s, t, x, y when working with our particle system on a microscopic (discrete) scale, or looking at "local" properties. In contrast, we will use capital letters S, T, X, Y when working on a macroscopic scale or dealing with objects which are continuous (or limiting to a continuous object). Hopefully this will become clearer as the reader progresses through the paper.

Let us define a few operators which we will work with throughout the paper. Firstly, if $f : \mathbb{Z} \to \mathbb{R}$ is any function then

$$\nabla^+ f(x) := f(x+1) - f(x)$$

$$\nabla^- f(x) := f(x-1) - f(x)$$

$$\Delta f(x) := f(x+1) + f(x-1) - 2f(x).$$

One should note that $\nabla^+\nabla^- = \nabla^-\nabla^+ = \Delta$.

Also, we will always use Λ to denote either the discrete interval $\{0, ..., N\}$ or the discrete half-line $\mathbb{Z}_{\geq 0}$ or both (depending on context). Similarly, we will use I to denote the continuous interval $[0, 1]$ or $[0, \infty)$.

Definition 2.2.1 (ASEP-H). *Fix $p, q, \alpha, \gamma \geq 0$. We define the half-line ASEP (or ASEP-H for brevity), to be the following continuous-time Markov process. We let $\eta(x) \in \{-1, 1\}$ denote the occupation variable at site $x \geq 1$, where the value -1 denotes an unoccupied site and $+1$ denotes an occupied site. The state space is then $\eta \in \{-1, 1\}^{\mathbb{Z}_{\geq 1}}$. The dynamics are given as follows: For $x > 0$, a particle jumps from site x to $x + 1$ at exponential rate*

$$\frac{p}{4}(1 + \eta_t(x))(1 - \eta_t(x+1)).$$

In other words, it jumps at rate p if site x is occupied and site $x + 1$ is not, otherwise it feels no inclination to jump (or the site x may be unoccupied). Similarly, for $x \geq 1$, a particle jumps from site $x + 1$ to x at rate

$$\frac{q}{4}(1 - \eta_t(x))(1 + \eta_t(x+1)).$$

Furthermore, a particle at site $x = 1$ is annihilated and created (respectively) at rates

$$\frac{\gamma}{2}(1 + \eta_t(0)) \qquad \textit{and} \qquad \frac{\alpha}{2}(1 - \eta_t(0)).$$

All jumps and annihilations/creations occur independently of each other.

Definition 2.2.2 (ASEP-B)**.** *Fix $p, q, \alpha, \gamma, \beta, \delta \geq 0$. We define the half-line ASEP (or ASEP-B for brevity), to be the following continuous-time Markov process. We let $\eta(x) \in \{-1, 1\}$ denote the occupation variable at site x, where we think of -1 denoting an unoccupied site and $+1$ denoting an occupied site. The state space is then $\eta \in \{-1, 1\}^{\{1,\dots,N\}}$. The dynamics are the same as those of the ASEP-H, except that a particle at site $x = N$ is annihilated or created (respectively) at rates*

$$\frac{\beta}{2}(1 + \eta_t(0)) \,, \qquad \frac{\delta}{2}(1 - \eta_t(0)).$$

All jumps and annihilations/creations occur independently of each other.

Definition 2.2.3 (Height Functions)**.** *Consider the model ASEP-H or ASEP-B defined above. For $t \geq 0$ we define $h_t(0)$ to be twice the net number of particles removed (i.e., twice annihilations minus creations) that have occurred up to time t. We then define*

$$h_t(x) := h_t(0) + \sum_{j=0}^{x} \eta_t(j)$$

which is an integrated version of the ASEP (in the sense that $\nabla^+ h_t(x) = \eta_t(x + 1)$).

The purpose of the above definition of $h_t(0)$ is so that the annihilation or creation of particles does not affect the value of the height function on the interior points of the interval.

We now move onto the appropriate scaling for our models. A well-known fact is that there is no way to scale time, space, and fluctuations in the KPZ equation so as to leave its solution invariant in law (the universal fixed point for the [1:2:3] scaling is highly nontrivial and was recently just proved for TASEP in [117]). However there are a couple of well-known $weak$ scalings which fix the law, in which we simultaneously scale the model parameters $together$ with the time, space, and fluctuations. The weak scaling considered for the ASEP height functions in this paper is the following:

Definition 2.2.4 (Weakly Asymmetric Scaling)**.** *Throughout this paper we will fix $A, B \in \mathbb{R}$. Let $\epsilon > 0$ small enough so that all rates defined below are positive. We define $p = \frac{1}{2}e^{\sqrt{\epsilon}}$ and $q = \frac{1}{2}e^{-\sqrt{\epsilon}}$, and we also define $\mu_A = 1 - A\epsilon$ and $\mu_B = 1 - B\epsilon$. We then define*

$$\alpha = \frac{p^{3/2}(p^{1/2} - \mu_A q^{1/2})}{p - q}, \qquad \beta = \frac{p^{3/2}(p^{1/2} - \mu_B q^{1/2})}{p - q},$$

$$\gamma = \frac{q^{3/2}(q^{1/2} - \mu_A p^{1/2})}{q - p}, \qquad \delta = \frac{q^{3/2}(q^{1/2} - \mu_B p^{1/2})}{q - p},$$

and we will always consider ASEP-H and ASEP-B with these parameters. For ASEP-B, we make the further assumption that $\epsilon = \frac{1}{N}$, where N is the length of the bounded interval.

Remark 2.2.5. *As in Remark 2.11 of [45], we note that we have the following asymptotics:*

$$p = \frac{1}{2} + \frac{1}{2}\sqrt{\epsilon} + O(\epsilon),$$

$$q = \frac{1}{2} - \frac{1}{2}\sqrt{\epsilon} + O(\epsilon).$$

For the creation/annihilation rates, we have

$$\alpha = \frac{1}{4} + \big(\frac{3}{8} + \frac{1}{4}A\big)\sqrt{\epsilon} + O(\epsilon), \qquad \beta = \frac{1}{4} + \big(\frac{3}{8} + \frac{1}{4}B\big)\sqrt{\epsilon} + O(\epsilon),$$

$$\gamma = \frac{1}{4} - \big(\frac{3}{8} + \frac{1}{4}A\big)\sqrt{\epsilon} + O(\epsilon), \qquad \delta = \frac{1}{4} - \big(\frac{3}{8} + \frac{1}{4}B\big)\sqrt{\epsilon} + O(\epsilon).$$

As stated in [45], this physically corresponds to "zooming into an $\epsilon^{1/2}$-window" around the critical triple point of the open ASEP.

Definition 2.2.6 (Gärtner Transform)**.** *For $x \in \mathbb{Z}_{\geq 0}$ we define the ASEP-H Gärtner transformed height function as*

$$Z_t(x) := e^{-\lambda h_t(x) + \nu t},$$

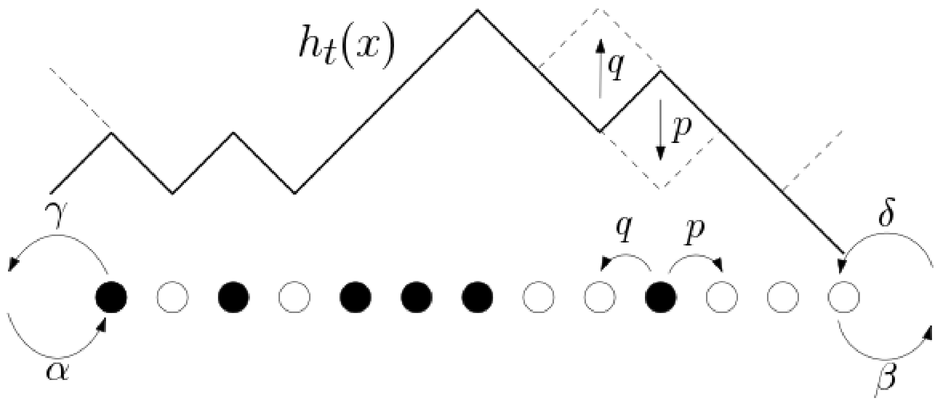

Figure 2.1: A depiction of the height function at a time t, in bold. We also show the jump/creation/annihilation rates and the way the height function changes when jumps of this rate occur.

where

$$\lambda = \frac{1}{2}\log\frac{q}{p}, \qquad \nu = p + q - 2\sqrt{pq}.$$

The definition is the same for ASEP-B, but we only consider $x \in \{0, ..., N\}$.

We remark that although the above Z depends on ϵ via the parameters p, q, we chose not to make this explicit because of notational convenience. The choice for these specific values of λ, ν is not immediately obvious, but becomes clear in the following theorem, which is really the key behind proving weak convergence to SHE in the limit as $\epsilon \to 0$. We also remark that

$$\lambda = -\sqrt{\epsilon}, \qquad \nu = \frac{1}{2}\epsilon + \frac{1}{24}\epsilon^2 + O(\epsilon^3)$$

which is the reason behind the temporal drift in Theorem 2.1.1.

Lemma 2.2.7 (Hopf-Cole-Gärtner Transform)**.** *The Gärtner-transformed height functions* $Z_t(x)$ *for ASEP-H satisfy the following discrete SHE:*

$$dZ_t(x) = \frac{1}{2}\Delta Z_t(x)dt + dM_t(x)$$

where $M_t(x)$ *is a pure-jump martingale (i.e., a sum of compensated Poisson processes) whose predictable bracket* $\langle M(x), M(y)\rangle_t$ *satisfies the following asymptotics as* $\epsilon \to 0$*:*

$$\frac{d}{dt}\langle M(x), M(y)\rangle_t = \begin{cases} 0 & x \neq y \\ \epsilon Z_t(x)^2 - \nabla^+ Z_t(x)\nabla^- Z_t(x) + o(\epsilon) Z_t(x)^2 & x = y > 0 \\ \epsilon Z_t(x)^2 + o(\epsilon) Z_t(x)^2 & x = y = 0 \end{cases} \tag{2.1}$$

for all $x \in \mathbb{Z}_{\geq 0}$. Moreover, Z_t satisfies a discrete Robin boundary condition for all t

$$Z(-1) = \mu_A Z_t(0).$$

For ASEP-B, the same holds true for $x \in \{0, ..., N-1\}$, but we have that the third asymptotic in (2.1) holds for $x = y = N$ and moreover

$$Z_t(N+1) = \mu_B Z_t(N).$$

Proof. The complete proof is found in Lemmas 3.1 and 3.3 of [45]. We also note that the original paper [69] was the first to recognize this transformation in the whole-line case, while [16] used it to prove the first of these KPZ-type convergence results. We further remark that (for fixed choices of p, q, μ_A, μ_B) this lemma *only* holds true for the exact values of $\alpha, \gamma, \beta, \delta$ chosen in Definition 2.2.4 as well as the values of λ, ν chosen in Definition 2.2.6. See the proof in [45]. □

2.3 Heat Kernel Estimates

In this (fairly lengthy) section, we will provide all of the technical estimates which will be used in the analysis used to prove convergence of the Gärtner-transformed height functions to the solutions of the SHE. There will not be very much motivation, but the purpose of the estimates will become clear as the reader progresses through Sections 5 and 6. Therefore, the reader may wish to skip some of the more technical results of this section, and come back as needed while progressing through the remainder of the paper.

Before moving onto the estimates, one should remark that the Robin heat kernels $\mathbf{p}_t^R(x,y)$ which are introduced below may be interpreted as transition probabilities for a continuous-time, discrete-space random walk with killing and branching at the boundary with certain probability, according the the corresponding parameters A, B. The case when $A > 0$ corresponds to killing, while the case $A < 0$ corresponds to branching (and similarly for B). Likewise, the continuous-space kernels P_T which are introduced in Section 3.3 may be interpreted as transition densities for Brownian motion with branching/killing. See for instance Section 4.1.1 of [45], [133], or [97] for more on this subject.

As a notational convention, we will usually write C for constants, and we will not generally specify when irrelevant terms are being absorbed into the constants. We will also write $C(A), C(A,T)$, or $C(a,b,A,B,T)$ whenever we want to specify exactly which parameters the constant depends on. This will not always be specified, though.

We also mention that we will occasionally use estimates from Appendix A of [57], where they derive very useful estimates for the (standard) whole-line heat kernel p_t. Hence the reader may wish to briefly look at that appendix while proceeding.

2.3.1 Half-Line Estimates

Unless otherwise specified, $A \in \mathbb{R}$ will be fixed. For $\epsilon > 0$, we set $\mu_A = 1 - A\epsilon$, and we let $\mathbf{p}_t^R(x,y)$ be the semi-discrete heat kernel on $\mathbb{Z}_{\geq -1}$ with Robin boundary parameter μ_A: this is defined as the fundamental solution to the discrete-space, continuous-time heat equation:

$$\partial_t \mathbf{p}_t^R(x,y) = \frac{1}{2}\Delta \mathbf{p}_t^R(x,y), \qquad \mathbf{p}_0^R(x,y) = 1_{\{x=y\}}, \qquad x, y \in \mathbb{Z}_{\geq 0}$$

with the boundary condition

$$\mathbf{p}_t^R(-1,y) = \mu_A \mathbf{p}_t^R(0,y).$$

Here, the discrete Laplacian Δ is taken in the second spatial coordinate. The "generalized image method" of Section 4.1 of [45] says that

$$\mathbf{p}_t^R(x,y) = p_t(x-y) + \mu_A p_t(x+y+1) + (1-\mu_A^{-2})\sum_{z=2}^{\infty} p_t(x+y+z)\mu_A^z \tag{2.2}$$

where $p_t(x)$ is the standard heat kernel on the whole line $\mathbb{Z}$ (i.e., the unique solution to the continuous-time, discrete-space equation $\partial_t p_t(x) = \frac{1}{2}\Delta p_t(x)$ with $p_0(x) = 1_{\{x=0\}}$). One may also directly check that (2.2) holds true.

Proposition 2.3.1. *Fix $A \in \mathbb{R}$ and $T > 0$. For $b \geq 0$, there is a constant $C(A,b,T)$ (not depending on ϵ) such that for all $t \in [0, \epsilon^{-2}T]$, and all $x, y \in \mathbb{Z}_{\geq 0}$ we have that*

$$\mathbf{p}_t^R(x,y) \leq C(A,b,T)(1 \wedge t^{-1/2})e^{-b|x-y|(1\wedge t^{-1/2})}.$$

Proof. Note that the first two terms appearing in the right-hand-side of Equation (2.2) already satisfy a bound of the desired form, by the standard (whole-line) heat kernel estimates (A.12) of [57]. Therefore it suffices to show that the third term appearing in the right side of (2.2) also satisfies a bound of the desired type. Since $\mu_A = 1 - A\epsilon$ it follows that $\log(\mu_A) \leq C(A)\epsilon$. For $t \in [0, \epsilon^{-2}T]$ we then have $\epsilon \leq C(T)(1 \wedge t^{-1/2})$ and therefore $\log(\mu_A) \leq C(A,T)(1 \wedge t^{-1/2})$. Summarizing, there exists $C(A,T)$ such that for all $t \in [0, \epsilon^{-2}T]$ we have that

$$\mu_A \leq e^{C(A,T)(1\wedge t^{-1/2})}.$$

If $A \geq 0$ then we can just take $C(A,T) = 0$, otherwise C will be positive. Using the standard heat kernel bound (A.12) of [57] it follows that for $b > C(A,T)$ we have

$$\begin{aligned}\sum_{z=2}^{\infty} p_t(x+y+z)\mu_A^z &\le C(b)\sum_{z=2}^{\infty}(1\wedge t^{-1/2})e^{-b(x+y+z)(1\wedge t^{-1/2})}\cdot e^{C(A,T)(1\wedge t^{-1/2})z}\\ &= C(b)e^{-b(x+y)(1\wedge t^{-1/2})}(1\wedge t^{-1/2})\frac{e^{2(C(A,T)-b)(1\wedge t^{-1/2})}}{1-e^{(C(A,T)-b)(1\wedge t^{-1/2})}}\\ &\le C(A,b,T)e^{-b(x+y)(1\wedge t^{-1/2})}\end{aligned}$$

where we used the fact that $e^{-2q}/(1-e^{-q}) \le 1/q$ for any $q \ge 0$. Since $t \le \epsilon^{-2}T$ we have that $|1-\mu_A^{-2}| \le C(A)\epsilon \le C(A,T)(1\wedge t^{-1/2})$ and the result follows. $\square$

Proposition 2.3.2. *Fix $A \in \mathbb{R}$ and $T > 0$. For $b \ge 0$, there is a constant $C(A,b,T)$ (not depending on ϵ) such that for all $t \in [0,\epsilon^{-2}T]$, all $x,y \in \mathbb{Z}_{\ge 0}$, all $|n| \le \lceil t^{1/2}\rceil$, and all $v \in [0,1]$ we have that*

$$|\mathbf{p}_t^R(x+n,y)-\mathbf{p}_t^R(x,y)| \le C(A,b,T)(1\wedge t^{-(1+v)/2})|n|^v e^{-b|x-y|(1\wedge t^{-1/2})}.$$

When $b=0$ this bound holds for all n, not just for $n \le t^{1/2}$.

Proof. The corresponding bounds already hold for the whole-line kernel $p_t(x)$, by (A.13) of [57]. So we proceed exactly as in Proposition 2.3.1, using (2.2) with p_t replaced by its n-point gradient $\nabla_n p_t$. The fact that the bound holds for all n when $b=0$ is a consequence of the triangle inequality applied to the case when $n=1$ and $v=1$. $\square$

Corollary 2.3.3. *For any $T \ge 0$ and $a_1,a_2 \ge 0$, there exists some constant $C = C(a_1,a_2,A,T)$ such that for all $x \ge 0$ and $t \le \epsilon^{-2}T$ we have*

$$\begin{aligned}\sum_{y\ge 0}\mathbf{p}_t^R(x,y)e^{a_1|x-y|(1\wedge t^{-1/2})}e^{a_2\epsilon y} &\le Ce^{a_2\epsilon x},\\ \sum_{y\ge 0}|\nabla^{\pm}\mathbf{p}_t^R(x,y)|e^{a_1|x-y|(1\wedge t^{-1/2})}e^{a_2\epsilon y} &\le C(1\wedge t^{-1/2})e^{a_2\epsilon x}.\end{aligned}$$

If we replace $e^{a_2\epsilon y}$ with $e^{a_2\epsilon|x-y|}$ in the first two expressions, then these bounds hold without the factor $e^{a_2\epsilon x}$ on the RHS.

Proof. Using Proposition 2.3.1, we find that $\mathbf{p}_t^R(x,y) \le C(b,T,A)(1\wedge t^{-1/2})e^{-b|x-y|(1\wedge t^{-1/2})}$, and moreover

$$e^{a_2\epsilon y} \le e^{a_2\epsilon x}e^{a_2\epsilon|x-y|} \le e^{a_2\epsilon x}e^{a_2 T^{1/2}(1\wedge t^{-1/2})|x-y|}$$

since $t \le \epsilon^{-2}T$. Consequently,

$$\sum_{y\ge 0}\mathbf{p}_t^R(x,y)e^{a_1|x-y|(1\wedge t^{-1/2})}e^{a_2\epsilon y} \le Ce^{a_2\epsilon x}\sum_{y\in\mathbb{Z}}(1\wedge t^{-1/2})e^{(-b+a_1+a_2T^{1/2})|x-y|(1\wedge t^{-1/2})}.$$

Letting $b := 1+a_1+a_2T^{1/2}$ we compute the sum on the RHS:

$$\sum_{y\in\mathbb{Z}}(1\wedge t^{-1/2})e^{-|x-y|(1\wedge t^{-1/2})} = (1\wedge t^{-1/2})\frac{1+e^{-(1\wedge t^{-1/2})}}{1-e^{-(1\wedge t^{-1/2})}} \le C$$

where we used the fact that $(1+e^{-q})/(1-e^{-q}) \le 1+2/q$ for $q \ge 0$. This proves the first inequality, and the second one is proved similarly using Proposition 2.3.2 instead of 2.3.1. The final statement is proved in a similar way. $\square$

We now turn to proving temporal bounds for the semi-discrete heat kernel, which will be useful in proving tightness. This involves different methods that the ones used to prove the spatial bounds above.

Lemma 2.3.4. *Let $p_t(x)$ denote the standard heat kernel on the whole line $\mathbb{Z}$ (as defined below Equation* (2.2)*). Then for $\mu > 0$, $t \ge 0$ and $x \in \mathbb{Z}_{\ge 0}$ we have the equality*

$$\sum_{z=-\infty}^{\infty} p_t(x+z)\mu^z = \mu^{-x}\exp\left[\frac{1}{2}\big(\mu+\mu^{-1}-2\big)t\right].$$

Proof. Fixing μ, let $F(t,x)$ denote the sum on the left side. Note that F is defined by a convolution involving p_t, therefore

$$\partial_t F(t,x) = \frac{1}{2}\Delta F(t,x) = \frac{1}{2}(\mu+\mu^{-1}-2)F(t,x).$$

Furthermore $F(0,x) = \mu^{-x}$, because $p_0(x) = 1_{[x=0]}$. This proves the given formula. Another way of putting this is that F is defined by convolving the semigroup p_t with an eigenfunction of Δ, whose eigenvalue is $\mu + \mu^{-1} - 2$. $\square$

Proposition 2.3.5. *Fix $A \in \mathbb{R}$ and $T > 0$. There is a constant $C(A,T)$ (not depending on ϵ) such that for all $s < t \in [0, \epsilon^{-2}T]$, all $x, y \in \mathbb{Z}_{\geq 0}$, and all $v \in [0,1]$ we have that*

$$|\mathbf{p}_t^R(x,y) - \mathbf{p}_s^R(x,y)| \leq C(A,T)(1 \wedge s^{-1/2-v})(t-s)^v.$$

Proof. First note that it suffices to prove these formulas when $v = 0$ and $v = 1$ since the middle cases follow by a straightforward interpolation. The $v = 0$ case follows from Proposition 2.3.1, and thus we only need to prove the $v = 1$ case. We will only consider the case when $A \leq 0$ (so $\mu_A \geq 1$), because the $A > 0$ case easier and involves similar methods (see Proposition 4.11 of [45]). To this end, we define a function

$$F(t,x) = \sum_{z=-\infty}^{\infty} p_t(x+z)\mu_A^z = \mu_A^{-x} \exp\left[\frac{1}{2}(\mu_A + \mu_A^{-1} - 2)t\right].$$

We rewrite the right side of Equation (2.2) as

$$\begin{aligned} &p_t(x-y) + \mu_A p_t(x+y+1) + (1-\mu_A^{-2})F(t,x+y) - (1-\mu_A^{-2}) \sum_{z=-1}^{\infty} p_t(x+y-z)\mu_A^{-z} \\ &=: J_1(t,x,y) \quad + \quad J_2(t,x,y) \quad + \quad J_3(t,x,y) \quad - \quad J_4(t,x,y). \end{aligned}$$

Now we only need to bound each of the differences $|J_i(t,x,y) - J_i(s,x,y)|$ for $1 \leq i \leq 4$. When $i = 1, 2$, the desired bounds follow directly from equation (A.10) of [57]. For the $i = 3$ bound, note that

$$
\begin{aligned}
|F(t,x)-F(s,x)| &= \mu_A^{-x} e^{\frac{1}{2}(\mu_A+\mu_A^{-1}-2)s}\left(e^{\frac{1}{2}(\mu_A+\mu_A^{-1}-2)(t-s)}-1\right) \\
&\le 1\cdot e^{\frac{1}{2}(\mu_A+\mu_A^{-1}-2)s}\cdot\left(\frac{1}{2}(\mu_A+\mu_A^{-1}-2)(t-s)\, e^{\frac{1}{2}(\mu_A+\mu_A^{-1}-2)(t-s)}\right) \\
&= e^{\frac{1}{2}(\mu_A+\mu_A^{-1}-2)t}\cdot\frac{1}{2}(\mu_A+\mu_A^{-1}-2)(t-s)
\end{aligned}
$$

where we used the bound $e^q - 1 \le qe^q$ with $q = \frac{1}{2}(\mu_A+\mu_A^{-1}-2)(t-s)$. Now we again recall that $\mu_A = 1 - A\epsilon$ so that $\mu_A^{-1} = 1 + A\epsilon + A^2\epsilon^2 + O(\epsilon^3)$, and therefore $\mu_A + \mu_A^{-1} - 2 \le CA^2\epsilon^2$ for small enough ϵ. Hence when $s,t \le \epsilon^{-2}T$, the previous bound gives

$$
|F(t,x)-F(s,x)| \le e^{\frac{1}{2}CA^2T}\cdot\frac{1}{2}CA^2\epsilon^2(t-s) = C(A,T)\,\epsilon^2(t-s).
$$

Now we recall that $1-\mu_A^{-2} \le C(A)\epsilon$ and $\epsilon < C(T)s^{-1/2}$ so that

$$
\begin{aligned}
|J_3(t,x,y)-J_3(s,x,y)| &= (1-\mu_A^{-2})|F(t,x+y)-F(s,x+y)| \\
&\le C(A,T)\,\epsilon^3(t-s) \\
&\le C(A,T)s^{-3/2}(t-s)
\end{aligned}
$$

which proves the desired bound for J_3.

To prove the bound for J_4, we again apply the whole-line estimate (A.10) of [57]

$$
\begin{aligned}
|J_4(t,x,y)-J_4(s,x,y)| &\le (1-\mu_A^{-2})\sum_{z=-1}^{\infty}\big|p_t(x+y-z)-p_s(x+y-z)\big|\mu_A^{-z} \\
&\le (1-\mu_A^{-2})\sum_{z=-1}^{\infty} Cs^{-3/2}(t-s)\mu_A^{-z} \\
&= Cs^{-3/2}(t-s)\mu_A(1+\mu_A^{-1})
\end{aligned}
$$

and we implicitly assume that ϵ is small enough so that $\mu_A(1+\mu_A^{-1}) \leq C(A)$. This proves the claim. $\square$

Proposition 2.3.6 (Long-time Estimate). *There exist constants $C = C(A,B)$ and $K = K(A,B)$ such that for every $t \geq 0$ and $x, y \geq 0$ we have that*

$$\mathbf{p}_t^R(x,y) \leq C(t^{-1/2}+\epsilon)e^{K\epsilon^2 t}.$$

We remark that these are "long-time" estimates because they are true uniformly over all $t > 0$, i.e., constants don't depend on any terminal time $\epsilon^{-2}T$ (as opposed to most results here).

Proof. In Equation (6.2.2), the first two terms are clearly bounded by $Ct^{-1/2}$ by equation (A.10) of [57] and the third term is bounded in absolute value by $|1-\mu_A^{-2}|\exp\big[(\mu_A+\mu_A^{-1}-2)t\big]$ by Lemma 2.3.4. Since $\mu_A = 1 - A\epsilon$, it follows that $|1-\mu_A^{-2}| < C\epsilon$ and $\mu_A + \mu_A^{-1} - 2 \leq K\epsilon^2$. This completes the proof. $\square$

Proposition 2.3.7. *For all A and $T > 0$, there exists $C = C(A,T)$ such that for $t \in [0, \epsilon^{-2}T]$ and $v \in [0,1]$ we have*

$$\sup_{x\in\mathbb{Z}_{\geq 0}} \left| \sum_{y\geq 0} \boldsymbol{p}_t^R(x,y) \ - \ 1 \right| \leq C\epsilon^v t^{v/2}.$$

Proof. One may use Lemma 2.3.8 (below) in order to prove this claim. However, we choose to give a different proof which will generalize easily to the bounded-interval case. Let

$$f(t,x) = \sum_{y\geq 0} \mathbf{p}_t^R(x,y).$$

By (1) we know $\mathbf{p}_t^R(x,y) = \mathbf{p}_t^R(y,x)$ and therefore

$$\begin{aligned}\partial_t f(t,x) = \frac{1}{2}\Delta f(t,x) &= \frac{1}{2}\sum_{y\geq 0}\big(\mathbf{p}_t^R(y+1,x) + \mathbf{p}_t^R(y-1,x) - 2\mathbf{p}_t^R(y,x)\big) \\ &= \frac{1}{2}\big(\mathbf{p}_t^R(-1,x) - \mathbf{p}_t^R(0,x)\big) = \frac{1}{2}(\mu_A - 1)\mathbf{p}_t^R(0,x)\end{aligned}$$

where we canceled out many terms in the first equality of the second line. Note that $f(0,x)=1$, and $|\mu_A - 1| = |A|\epsilon$. Finally (by Proposition 1.2) $\mathbf{p}_t^R(0,x) \le C(A,T)t^{-1/2}$, therefore when $v\in[0,1]$ we have

$$
\begin{aligned}
|f(t,x)-1| &\le \int_0^t |\partial_s f(s,x)|ds = \frac{1}{2}|A|\epsilon \int_0^t \mathbf{p}_s^R(0,x)ds \\
&\le C(A,T)\;\epsilon \int_0^t s^{-1/2}ds = C(A,T)\;\epsilon\; t^{1/2} \le C(A,T)\epsilon^v t^{v/2}
\end{aligned}
$$

where in the last inequality we used the fact that $\epsilon = \epsilon^v\epsilon^{1-v} \le C(T)\epsilon^v t^{(v-1)/2}$ since $t\in[0,\epsilon^{-2}T]$. This proves the claim. □

We now turn to proving certain “cancellation estimates" which will be used in identifying the limiting measure on $C([0,T],C([0,\infty))$ as the solution to the stochastic heat equation.

Lemma 2.3.8. *For the next few estimates we will distinguish between different values of A by writing $\boldsymbol{p}_t^R(x,y;A)$ for the (ϵ-dependent) Robin heat kernel of parameter A. For all A, all $T>0$, and all $b\ge 0$, there exists $C(A,b,T)$ such that for all $x,y\in\mathbb{Z}_{\ge 0}$ and $t\in[0,\epsilon^{-2}T]$ we have*

$$
\begin{aligned}
|\mathbf{p}_t^R(x,y;A)-\mathbf{p}_t^R(x,y;0)| &\le C(A,b,T)\;\epsilon\; e^{-b(x+y)(1\wedge t^{-1/2})}, \\
|\nabla^{\pm}\mathbf{p}_t^R(x,y;A)-\nabla^{\pm}\mathbf{p}_t^R(x,y;0)| &\le C(A,b,T)\;\epsilon\;(1\wedge t^{-1/2})e^{-b(x+y)(1\wedge t^{-1/2})}.
\end{aligned}
$$

where $\nabla^{\pm}$ denotes the discrete gradient in the first spatial coordinate.

Proof. Note by (6.2.2) that

$$
|\mathbf{p}_t^R(x,y;A)-\mathbf{p}_t^R(x,y;0)| = (\mu_A-1)p_t(x+y+1) + (1-\mu_A^{-2})\sum_{z=2}^{\infty} p_t(x+y+z)\mu_A^z
$$

where p_t is the standard (whole-line) heat kernel. Since $\mu_A = 1-A\epsilon$, it follows that $|\mu_A-1|$ and $|1-\mu_A^{-2}|$ are both bounded by $C(A)\epsilon$. Moreover $p_t(x+y+1)\le C(b)e^{-b(x+y)(1\wedge t^{-1/2})}$ by the standard heat kernel bound (A.12) of [57]. Hence we just need to show that

$$\sum_{z=2}^{\infty} p_t(x+y+z)\mu_A^z \le C(A,b,T)e^{-b(x+y)(1\wedge t^{-1/2})}.$$

But this was done during Proposition 2.3.1. The proof for the gradient estimates is similar, but we get left with an extra $1 \wedge t^{-1/2}$ by setting $v = 1$ in Proposition 2.3.2. □

Lemma 2.3.9. *For $A \in \mathbb{R}$, $t \ge 0$, and $x, y \in \mathbb{Z}_{\ge 0}$ we define*

$$K_t(x,y;A) := \nabla^+ \mathbf{p}_t^R(x,y;A)\nabla^- \mathbf{p}_t^R(x,y;A).$$

For $a, T \ge 0$ there exists a constant $C = C(a,A,T)$ such that

$$\sum_{y\ge 1}\int_0^{\epsilon^{-2}T} |K_t(x,y;A) - K_t(x,y;0)|e^{a\epsilon|x-y|}dt \le C\epsilon^{1/2}.$$

Proof. Write

$$\begin{aligned}
|K_t(x,y;A) - K_t(x,y;0)| &\le |\nabla^+ \mathbf{p}_t^R(x,y;A) - \nabla^+ \mathbf{p}_t^R(x,y;0)||\nabla^- \mathbf{p}_t^R(x,y;A)| \\
&\quad + |\nabla^+ \mathbf{p}_t^R(x,y;0)||\nabla^- \mathbf{p}_t^R(x,y;A) - \nabla^- \mathbf{p}_t^R(x,y;0)| \\
&=: J_1(t,x,y) + J_2(t,x,y).
\end{aligned}$$

By Lemma 2.3.8 we have that

$$\begin{aligned}
J_1(t,x,y) &= |\nabla^+ \mathbf{p}_t^R(x,y;A) - \nabla^+ \mathbf{p}_t^R(x,y;0)||\nabla^- \mathbf{p}_t^R(x,y;A)| \\
&\le C(A,T)\,\epsilon\,(1\wedge t^{-1/2})|\nabla^- \mathbf{p}_t^R(x,y;A)|.
\end{aligned}$$

Since $t \in [0, \epsilon^{-2}T]$ we have that $\epsilon = \epsilon^{1/2}\epsilon^{1/2} \le C(T)\epsilon^{1/2}(1\wedge t^{-1/4})$ and thus we see that

$$J_1(t,x,y) \le C(A,T)\,\epsilon^{1/2}(1\wedge t^{-3/4})|\nabla^- \mathbf{p}_t^R(x,y;A)|.$$

Consequently

$$\sum_{y\geq 1} J_1(t,x,y)e^{a\epsilon|x-y|} \leq C(A,T)\epsilon^{1/2}(1\wedge t^{-3/4})\sum_{y\geq 1}|\nabla^- \mathbf{p}_t^R(x,y;A)|e^{a\epsilon|x-y|}$$
$$\leq C(a,A,T)\epsilon^{1/2}(1\wedge t^{-5/4}).$$

In the last inequality we used the second bound in Corollary 2.3.3. Now integrating both sides of this inequality from 0 to $\epsilon^{-2}T$ we find that

$$\begin{aligned}\sum_{y\geq 1}\int_0^{\epsilon^{-2}T} J_1(t,x,y)e^{a\epsilon|x-y|}dt &= \int_0^{\epsilon^{-2}T}\Big(\sum_{y\geq 1} J_1(t,x,y)e^{a\epsilon|x-y|}\Big)dt \\ &\leq \int_0^\infty C(A,T,a)\epsilon^{1/2}(1\wedge t^{-5/4})dt \\ &= C(A,T,a)\epsilon^{1/2}.\end{aligned}$$

A similar argument shows that J_2 satisfies a similar bound. This proves the claim. $\square$

Lemma 2.3.10. *Let us write $\mathbf{p}_t^R(x,y;A)$ as in the preceding lemma. For $x,\bar{x}\in\mathbb{Z}_{\geq 0}$, we have that*

$$\sum_{y\geq 0}\int_0^\infty \nabla^+\mathbf{p}_t^R(x,y;0)\nabla^+\mathbf{p}_t^R(\bar{x},y;0)dt = 1_{\{x=\bar{x}\}}.$$

Proof. Let us recall the summation-by-parts identity: if $u,v:\mathbb{Z}_{\geq -1}\to\mathbb{R}$ are absolutely summable, then

$$\sum_{y=0}^\infty u(y)\Delta v(y) = u(-1)\nabla^- v(0) - \sum_{y=-1}^\infty \nabla^+ u(y)\nabla^+ v(y).$$

Letting $u=\mathbf{p}_t^R(x,\cdot;0)$ and $v=\mathbf{p}_t^R(\bar{x},\cdot;0)$, the boundary terms will vanish and we get

$$-\sum_{y=0}^\infty \nabla^+\mathbf{p}_t^R(x,y;0)\nabla^+\mathbf{p}_t^R(\bar{x},y;0) = \sum_{y=0}^\infty \mathbf{p}_t^R(x,y;0)\Delta\mathbf{p}_t^R(\bar{x},y;0) = \sum_{y=0}^\infty \Delta\mathbf{p}_t^R(x,y;0)\mathbf{p}_t^R(\bar{x},y;0).$$

Since $\Delta \mathbf{p}_t^R = 2\partial_t \mathbf{p}_t^R$, this implies that

$$-\sum_{y=0}^{\infty} \nabla^+ \mathbf{p}_t^R(x,y;0)\nabla^+ \mathbf{p}_t^R(\bar{x},y;0) = \sum_{y=0}^{\infty} \mathbf{p}_t^R(x,y;0)\partial_t \mathbf{p}_t^R(\bar{x},y;0) + \partial_t \mathbf{p}_t^R(x,y;0)\mathbf{p}_t^R(\bar{x},y;0).$$

Integrating both sides from $t=0$ to ∞ and using the semigroup property, we find that

$$\begin{aligned}
-\sum_{y=0}^{\infty} \int_0^{\infty} \nabla^+ \mathbf{p}_t^R(x,y;0)\nabla^+ \mathbf{p}_t^R(\bar{x},y;0)dt &= \sum_{y=0}^{\infty} \mathbf{p}_t^R(x,y;0)\mathbf{p}_t^R(\bar{x},y,0)\Big|_{t=0}^{\infty} \\
&= \mathbf{p}_{2t}^R(x,\bar{x};0)\Big|_{t=0}^{\infty} \\
&= 0 - 1_{\{x=\bar{x}\}}
\end{aligned}$$

which proves the claim. $\square$

Proposition 2.3.11. *Let $A \in \mathbb{R}$, $T > 0$, and $a > 0$. Let K_t be as in Lemma 2.3.9. There exists some $\epsilon_0 = \epsilon_0(A,T,a)$ and some $c_* = c_*(A,T,a) < 1$ such that for $\epsilon < \epsilon_0$ and $x \geq 1$*

$$\sum_{y\geq 1} \int_0^{\epsilon^{-2}T} |K_t(x,y,A)|e^{a\epsilon|x-y|}dt \leq c_*.$$

Proof. For the moment being, let us suppose that we have already proved the claim when $A = 0$: there exists some $c_* < 1$ such that for small enough ϵ we have

$$\sum_{y\geq 1} \int_0^{\epsilon^{-2}T} |K_t(x,y;0)|e^{a\epsilon|x-y|}dt \leq c_*.$$

By Proposition A.8, we also have that

$$\sum_{y\geq 1} \int_0^{\epsilon^{-2}T} |K_t(x,y;A) - K_t(x,y;0)|e^{a\epsilon|x-y|}dt \leq C(a,A,T)\epsilon^{1/2}.$$

Putting together both these bounds and using the triangle inequality, we find that

$$\sum_{y\geq 1}\int_0^{\epsilon^{-2}T} |K_t(x,y;A)|\, e^{a\epsilon|x-y|}dt \leq c_* + C(a,A,T)\epsilon^{1/2}.$$

Now consider ϵ small enough so that $C(a,A,T)\epsilon^{1/2} \leq (1-c_*)/2$, and we see that the RHS is smaller than $(1+c_*)/2 =: c_*' < 1$ which proves the claim for arbitrary A.

It remains to prove the claim when $A = 0$. From now on, we will implicitly assume that $A = 0$ and we will just write $K_t(x,y)$ and $\mathbf{p}_t^R(x,y)$ with the understanding that $A = 0$. Let us first consider the case when $a = 0$. For this, we imitate the proof of Proposition 5.4 in [45]. Using Cauchy-Schwarz, it is true that

$$\sum_{y\geq 1}\int_0^{\epsilon^{-2}T} |K_t(x,y)|dt < \left(\sum_{y\geq 1}\int_0^{\infty}(\nabla^+\mathbf{p}_t^R(x,y))^2 dt\right)^{1/2}\left(\sum_{y\geq 1}\int_0^{\infty}(\nabla^-\mathbf{p}_t^R(x,y))^2 dt\right)^{1/2}.$$

Using Lemma 2.3.10, it is easy to see that the RHS of this expression is equal to 1. Moreover, the inequality is strict since $\nabla^+\mathbf{p}_t^R \neq \nabla^-\mathbf{p}_t^R$. This proves that, for each fixed $x \in \mathbb{Z}_{\geq 0}$ and $\epsilon > 0$, the LHS is strictly less than 1. However, this strict inequality may no longer be true after taking the supremum over all x and ϵ. Thus, a stronger argument is needed. Recall the Lagrange identity, which says

$$\left(\sum_i a_i^2\right)\left(\sum_i b_i^2\right) - \left(\sum_i a_i b_i\right)^2 = \sum_{i<j}(a_i b_j - a_j b_i)^2.$$

This means that

$$\left(\sum_{y\geq 1}(\nabla^+\mathbf{p}_t^R(x,y))^2\right)\left(\sum_{y\geq 1}(\nabla^-\mathbf{p}_t^R(x,y))^2\right) - \left(\sum_{y\geq 1}|K_t(x,y)|\right)^2$$

$$= \sum_{1\leq \bar{y}<y}\Big(|\nabla^+\mathbf{p}_t^R(x,y)\nabla^-\mathbf{p}_t^R(x,\bar{y})| - |\nabla^-\mathbf{p}_t^R(x,y)\nabla^+\mathbf{p}_t^R(x,\bar{y})|\Big)^2. \tag{2.3}$$

Now, just as in Corollary 5.4 of [45], we claim that there exists some $t_0 > 0$ such that for every

$\epsilon > 0$, $x \in \mathbb{Z}_{\geq 0}$, and $t \leq t_0$ we have that

$$\nabla^{\pm}\mathbf{p}_t^R(x,x) \leq -\frac{4}{5}, \quad \text{and} \quad |\nabla^{-}\mathbf{p}_t^R(x,x+1)| \leq \frac{1}{5}. \tag{2.4}$$

Indeed, the corresponding bound may be seen to be true for the standard heat kernel on the whole line $\mathbb{Z}$: for $t \leq t_0$ we have

$$\nabla^{\pm}p_t(0) \leq -\frac{9}{10}, \quad \text{and} \quad |\nabla^{-}p_t(-1)| \leq \frac{1}{10}. \tag{2.5}$$

In fact when $t = 0$ the left quantity is simply -1 and the right quantity is 0. Moreover, there is no dependence on ϵ and both quantities are continuous in t, which shows that (2.5) is indeed true. Now we use the simple relation (see Equation (6.2.2)) $\mathbf{p}_t^R(x,y) = p_t(x-y) + p_t(x+y+1)$ in order to deduce Equation (2.4) from (2.5).

Now, given that (2.4) is true, this implies that for $t \leq t_0$, $x \in \mathbb{Z}_{\geq 0}$, and $\epsilon > 0$:

$$\Big(|\nabla^{+}\mathbf{p}_t^R(x,x+1)\nabla^{-}\mathbf{p}_t^R(x,x)| - |\nabla^{-}\mathbf{p}_t^R(x,x+1)\nabla^{+}\mathbf{p}_t^R(x,x)|\Big)^2 \geq \Big(\frac{4}{5}\cdot\frac{4}{5} - \frac{1}{5}\cdot 1\Big)^2 > \frac{1}{6}.$$

This in turn implies that the expression in (2.3) is bounded below by $1/6$, uniformly over $t \in [0,t_0]$, $x \in \mathbb{Z}_{\geq 0}$, and $\epsilon > 0$, i.e.,

$$\Big(\sum_{y\geq 1}(\nabla^{+}\mathbf{p}_t^R(x,y))^2\Big)\Big(\sum_{y\geq 1}(\nabla^{-}\mathbf{p}_t^R(x,y))^2\Big) - \Big(\sum_{y\geq 1}|K_t(x,y)|\Big)^2 > 1/6$$

Now this expression is of the form $f^2g^2 - h^2 > 1/6$, where f, g, h are functions of (x,ϵ), defined by the previous expression. We can rewrite this as $(fg-h)(fg+h) > 1/6$. Now, Cauchy-Schwarz implies that $h \leq fg$, so that $fg + h \leq 2fg$. Moreover, the heat kernel estimates (Propositions 2.3.1, 2.3.2, and Corollary 2.3.3) imply that for $t \leq \epsilon^{-2}T$, we have that $f, g \leq C$ for some absolute constant C. Hence $(fg-h)\cdot 2C^2 \geq (fg-h)(fg+h) > 1/6$, so that $fg-h > 1/(12C^2) =: c > 0$.

Summarizing, there exists $c > 0$ such that for all $t \le t_0$, $x \in \mathbb{Z}_{\ge 0}$, and $\epsilon > 0$, we have that

$$\Big(\sum_{y\ge 1}(\nabla^+\mathbf{p}_t^R(x,y))^2\Big)^{1/2}\Big(\sum_{y\ge 1}(\nabla^-\mathbf{p}_t^R(x,y))^2\Big)^{1/2} - \sum_{y\ge 1}|K_t(x,y)| \;\;>\;\; c.$$

Using the Cauchy-Schwarz inequality and Lemma 2.3.10, the time-integral of the first term on the LHS (from $t = 0$ to ∞) is bounded above by 1, therefore we integrate both sides of the above expression and we get

$$1 - \int_0^\infty \sum_{y\ge 1}|K_t(x,y)|dt > ct_0$$

which completes the proof when $a = 0$, with $c_* = 1 - ct_0$.

To prove the general case with $a > 0$, Proposition 2.3.2 gives $|K_t(x,y)| \le C(1\wedge t^{-2})e^{-b|x-y|(1\wedge t^{-1/2})}$, so that

$$\sup_{x\ge 0}\sum_{y\ge 1}|K_t(x,y)|(e^{a\epsilon|x-y|}-1) \le C(1\wedge t^{-2})\sum_{z\in\mathbb{Z}} e^{-b(1\wedge t^{-1/2})|z|}(e^{a\epsilon|z|}-1).$$

By the dominated convergence theorem, this in turn implies that:

$$\lim_{\epsilon\to 0}\int_0^\infty \sup_{x\ge 0}\sum_{y\ge 1}|K_t(x,y)|(e^{a\epsilon|x-y|}-1)dt = 0.$$

Note that by Equation (6.2.2), $\mathbf{p}_t^R$ does not actually depend on ϵ when $A = 0$, therefore the preceding expression implies that

$$\lim_{\epsilon\to 0}\sup_{x\ge 0}\int_0^\infty \sum_{y\ge 1}|K_t(x,y)|e^{a\epsilon|x-y|}dt = \sup_{x\ge 0}\int_0^\infty\sum_{y\ge 1}|K_t(x,y)|dt \le c_* < 1.$$

Hence for small enough $\epsilon > 0$, the LHS is smaller than $(1+c_*)/2 =: c_*'$, completing the proof. □

Corollary 2.3.12. *In the same setting as Proposition 2.3.11, for any $S \in [0,T]$ there is a $C =$*

$C(A,T,S,a)$ *such that for* $s=\epsilon^{-2}S$ *we have*

$$\sum_{y\geq 1}\int_0^s |\nabla^+\boldsymbol{p}_t^R(x,y)\nabla^-\boldsymbol{p}_t^R(x,y)|e^{a\epsilon|x-y|}(s-t)^{-1/2}dt\leq C\epsilon.$$

Proof. We mimic the proof given in Proposition 5.4 of [45]. We split the integral into two pieces, one from 0 to $s/2$ and he other from $s/2$ to s. For the first integral, we note that $(s-t)^{-1/2}\leq \sqrt{2}s^{-1/2}=\sqrt{2}S^{-1/2}\epsilon$ when $t<s/2$, and therefore

$$\begin{aligned}&\sum_{y\geq 1}\int_0^{s/2} |\nabla^+\mathbf{p}_t^R(x,y)\nabla^-\mathbf{p}_t^R(x,y)|e^{a\epsilon|x-y|}(s-t)^{-1/2}dt\\ &\leq C(S)\ \epsilon\sum_{y\geq 1}\int_0^{s/2} |\nabla^+\mathbf{p}_t^R(x,y)\nabla^-\mathbf{p}_t^R(x,y)|e^{a\epsilon|x-y|}dt\\ &< C(S)\ \epsilon\end{aligned}$$

where we used Proposition 2.3.11 in the final line. For the second part of the integral, we note from Proposition 2.3.2 that $|\nabla^+\mathbf{p}_t^R(x,y)\nabla^-\mathbf{p}_t^R(x,y)|\leq Ct^{-1}|\nabla^-\mathbf{p}_t^R(x,y)|$ and thus Corollary 2.3.3 shows

$$\sum_{y\geq 1}|\nabla^+\mathbf{p}_t^R(x,y)\nabla^-\mathbf{p}_t^R(x,y)|e^{a\epsilon|x-y|}\leq C(a,A,T)(1\wedge t^{-3/2}).$$

Integrating both sides from $s/2$ to s we see that

$$\begin{aligned}&\sum_{y\geq 1}\int_{s/2}^{s} |\nabla^+\mathbf{p}_t^R(x,y)\nabla^-\mathbf{p}_t^R(x,y)|e^{a\epsilon|x-y|}(s-t)^{-1/2}dt\\ &\leq C(a,A,T)\int_{s/2}^s t^{-3/2}(s-t)^{-1/2}dt\\ &=C(a,A,T)s^{-1}\int_{1/2}^1 u^{-3/2}(1-u)^{-1/2}du\\ &\leq C(a,A,S,T)\epsilon^2.\end{aligned}$$

We made a substitution $t=su$ in the third line, and we used $s^{-1}=\epsilon^2S^{-1}$ in the final line. This proves the claim. □

Finally we conclude this section with an estimate which falls into none of the above categories:

Proposition 2.3.13. *For every $t \geq s \geq 0$ and $x, y \geq 0$ we have that*

$$\mathbf{p}_s^R(x,y) \leq e^{t-s}\mathbf{p}_t^R(x,y).$$

Proof. One may directly verify that

$$\mathbf{p}_s^R(x,y) = \sum_{n=0}^{\infty} e^{-s}\frac{s^n}{n!}\mathbf{p}^n(x,y)$$

where $\mathbf{p}^n(x,y)$ is the fundamental solution to the discrete-time, discrete-space equation $\mathbf{p}^{n+1}(x,y) - \mathbf{p}^n(x,y) = \frac{1}{2}\Delta\mathbf{p}^n(x,y)$, with Robin boundary conditions $\mathbf{p}^n(-1,y) = \mu_A\mathbf{p}^n(0,y)$. Therefore

$$\mathbf{p}_s^R(x,y) = \sum_{n=0}^{\infty} e^{-s}\frac{s^n}{n!}\mathbf{p}^n(x,y) \leq e^{t-s}\sum_{n=0}^{\infty} e^{-t}\frac{t^n}{n!}\mathbf{p}^n(x,y) = e^{t-s}\mathbf{p}_t^R(x,y)$$

which proves the claim. □

2.3.2 Bounded Interval Estimates

In this section, we prove all of the estimates of Section 3.1 for the Robin heat kernel on the bounded interval $\{0,...,N\}$. We will fix $A, B \in \mathbb{R}$, and we set $\mu_A := 1 - A\epsilon$ and $\mu_B = 1 - B\epsilon$ for $\epsilon := \frac{1}{N}$. Then we let $\mathbf{p}_t^R(x,y)$ be the Robin heat kernel satisfying $\mathbf{p}_t^R(-1,y) = \mu_A\mathbf{p}_t^R(0,y)$ and $\mathbf{p}_t^R(N+1,y) = \mu_B\mathbf{p}_t^R(N,y)$. We will need a formula similar to (2.2), though unfortunately we have to rely on a more complicated inductive formula.

In order to derive such a formula, [45] used the following procedure: Start with an arbitrary “test" function $\varphi : \{0,...,N\} \to \mathbb{R}$. Then extend φ to a function $\tilde{\varphi}$ on all of $\mathbb{Z}$ such that $\tilde{\varphi}(z-1) - \mu_A\tilde{\varphi}(z)$ is an odd function, and such that $\tilde{\varphi}(N+1+z) - \mu_B\tilde{\varphi}(N+z)$ is also an odd function. This may be done inductively, first defining $\tilde{\varphi}$ on $\{-N,...,-1\}$ and on $\{N+1,...,2N\}$, then on $\{-2N,...,-N-1\}$ and $\{2N+1,...,3N\}$, and so forth. Then it would necessarily hold true that for any $x \in \{0, ..., N\}$:

$$\sum_{y=0}^{N} \mathbf{p}_t^R(x,y)\varphi(y) = \sum_{y\in\mathbb{Z}} p_t(x-y)\tilde{\varphi}(y)$$

where the p_t on the RHS is the standard (whole-line) heat kernel. Indeed, both the LHS and RHS solve the heat equation with Robin boundary conditions and initial data φ. Using this fact and rearranging terms, Lemmas 4.6 and 4.7 in [45] proved the following semi-explicit formula

$$\mathbf{p}_t^R(x,y) = \sum_{k\in\mathbb{Z}} I_k p_t(x - i(y,k)) + \epsilon \sum_{\substack{k\in\mathbb{Z}\\k\neq 0}} \sum_{z=k(N+1)}^{k(N+1)+N} p_t(x-z)E_k(z,y) \tag{2.6}$$

where

$$i(y,k) = \begin{cases} (k+1)(N+1) - y - 1, & k \equiv 1 \pmod 2 \\ y + k(N+1), & k \equiv 0 \pmod 2 \end{cases}$$

and the factors I_k and $E_k(z,y)$ (which depend on μ_A and μ_B) satisfy the following inductive relations for $m \geq 0$:

$I_0 = 1$ and $E_0(x,y) = 0$ for $x, y \in \{0, ..., N\}$. Then

$$\begin{aligned} I_{-(m+1)} &= \mu_A I_m, \\ I_{m+1} &= \mu_B I_{-m}. \end{aligned}$$

For $x \in \{(-m-1)(N+1), ..., (-m-1)(N+1)+N\}$, and $y \in \{0, ..., N\}$

$$E_{-(m+1)}(x,y) := \mu_A E_m(-x-1,y) + \epsilon^{-1}(\mu_A^2 - 1) \sum_{k=0}^{m} \mu_A^{-x-2-i(y,k)} I_k 1_{\{i(y,k)\leq -x-2\}}$$
$$+ (\mu_A^2 - 1) \sum_{k=0}^{m} \sum_{z=k(N+1)}^{k(N+1)+N} \mu_A^{-x-2-z} 1_{\{z\leq -x-2\}} E_k(z,y).$$

For $x \in \{(m+1)(N+1), ..., (m+1)(N+1)+N\}$, and $y \in \{0, ..., N\}$

$$E_{m+1}(x,y) := \mu_B E_{-m}(2N+1-x,y) + \epsilon^{-1}(\mu_B^2 - 1) \sum_{k=-m}^{0} \mu_B^{x-2(N+1)+i(y,k)} I_k 1_{\{i(y,k) \geq 2(N+1)-x\}}$$
$$+ (\mu_B^2 - 1) \sum_{k=-m}^{0} \sum_{z=k(N+1)}^{k(N+1)+N} \mu_B^{x-2(N+1)+z} 1_{\{z \geq 2(N+1)-x\}} E_k(z,y).$$

We will repeatedly use the fact that $\epsilon = N^{-1}$ throughout this section, so we strongly emphasize this relation.

Lemma 2.3.14. *There exists a constant $C_0 := C_0(A,B)$ such that for all $k \in \mathbb{Z}$,*

$$\sup_{\substack{z \in \{k(N+1), ..., k(N+1)+N\} \\ y \in \{0,...,N\}}} |E_k(z,y)| \leq C_0^{|k|}.$$

Proof. Assuming C_0 has already been fixed, we we prove the claim using induction on k (and we will find an explicit lower bound for C_0 later). It is clearly true for $k = 0$, since $E_0 = 0$. So suppose the result is true for all k such that $|k| \leq m$. Using the recursive relation for $E_{-(m+1)}$ as well as the fact that $|\mu_A^2 - 1| \leq 3A\epsilon$ for small enough ϵ, we find that

$$|E_{-m-1}(x,y)| \leq \mu_A C_0^m + 3A(m+1)\mu_A^{(m+1)N} + 3A\epsilon \bigg(\sum_{k=0}^{m} N \mu_A^{(m+1-k)N} C_0^k \bigg).$$

Now notice that $\mu_A^{(m+1)N} = (1 - A/N)^{N(m+1)} \leq e^{|A|(m+1)}$. Similarly, and using $\epsilon = N^{-1}$ we get

$$3A\epsilon \bigg(\sum_{k=0}^{m} N \mu_A^{(m+1-k)N} C_0^k \bigg) \leq 3Ae^{|A|} \sum_{k=0}^{m} e^{|A|(m-k)} C_0^k = 3Ae^{|A|} \frac{C_0^{m+1} - e^{|A|(m+1)}}{C_0 - e^{|A|}}.$$

Therefore if ϵ is small enough so that $\mu_A < 2$ then

$$|E_{-(m+1)}(x,y)| \leq 2C_0^m + 3A(m+1)e^{|A|(m+1)} + 3Ae^{|A|} \frac{C_0^{m+1} - e^{|A|(m+1)}}{C_0 - e^{|A|}}$$
$$\leq K(A) \left(e^{2|A|m} + \frac{C_0^{m+1}}{C_0 - e^{|A|}} \right)$$

where $K(A)$ is a constant depending only on A and not m. We claim that for large enough choices of C_0 the last expression can be made smaller than C_0^{m+1}. Indeed, take $C_0 \geq e^{2|A|}$ such that $2K(A)/(C_0 - e^{|A|}) \leq 1$. Then

$$\begin{aligned} C_0^{-(m+1)}|E_{-(m+1)}(x,y)| &\leq K(A)\left[\frac{1}{C_0}\left(\frac{e^{2|A|}}{C_0}\right)^m + \frac{1}{C_0 - e^{|A|}}\right] \\ &\leq K(A)\left[\ \frac{1}{C_0}\quad + \quad \frac{1}{C_0 - e^{|A|}}\right] \\ &\leq 1. \end{aligned}$$

Thus the inductive hypothesis is satisfied for $k = -(m+1)$. The case when $k = m+1$ is similar, but one needs to replace A with B throughout the proof. $\square$

As a consequence of these explicit formulas, we have the same spatial bounds as those of the half-line.

Proposition 2.3.15. *Fix $A, B \in \mathbb{R}$ and $T > 0$. For $b \geq 0$, there is a constant $C(A,B,b,T)$ (not depending on $\epsilon = N^{-1}$) such that for all $t \in [0, \epsilon^{-2}T]$, and all $x, y \in \{0, ..., N\}$ we have that*

$$\mathbf{p}_t^R(x,y) \leq C(A,B,b,T)(1 \wedge t^{-1/2})e^{-b|x-y|(1\wedge t^{-1/2})}.$$

Proof. We will prove the desired bound for both terms in the RHS of equation (2). For the first term, we only need to consider terms in the sum with $|k| > 2$ since the other terms easily satisfy the desired bound. Note by induction that $I_k \leq (\mu_A \vee \mu_B)^{|k|}$ for all $k \in \mathbb{Z}$. Using the same argument as in Proposition 2.3.1, we have

$$\mu_A \vee \mu_B \leq e^{C(A,B)\epsilon} \leq e^{C(A,B,T)(1\wedge t^{-1/2})}$$

for all $t \in [0, \epsilon^{-2}T]$. By using the standard estimates [57] as well as the fact that $|x - i(y;k)| \geq$

$(|k|-1)N$ for all $x,y\in\{0,...,N\}$ we find that if $bN>C(A,B,T)$ then

$$\begin{aligned}\sum_{|k|>2} I_k p_t(x-i(y;k)) &\leq \sum_{|k|>2} e^{C(A,B,T)(1\wedge t^{-1/2})|k|}\cdot C(b)(1\wedge t^{-1/2})e^{-b(|k|-1)N(1\wedge t^{-1/2})}\\ &= C(b)(1\wedge t^{-1/2})\frac{e^{2(C(A,B,T)-bN)(1\wedge t^{-1/2})}}{1-e^{(C(A,B,T)-bN)(1\wedge t^{-1/2})}}\\ &\leq C(b)\frac{e^{(C(A,B,T)-bN)(1\wedge t^{-1/2})}}{bN-C(A,B,T)}\end{aligned}$$

where we used $e^{-2q}/(1-e^{-q})\leq e^{-q}/q$ in the last line. Now, we may as well assume that b (or N) is large enough so that $bN-C(A,B,T)>bN/2$. Using this together with the fact that $N\geq|x-y|$ we get that the last expression is bounded above by

$$C(b)N^{-1}e^{-\frac{1}{2}b|x-y|(1\wedge t^{-1/2})},$$

then using the fact that $N^{-1}=\epsilon\leq C(T)(1\wedge t^{-1/2})$ gives the desired bound on the second term in equation (2), after replacing $\frac{1}{2}b$ with b.

Now for the second term in (2.6). As before we only consider terms in the sum with $|k|>2$. Note by Lemma 2.3.14 that

$$\begin{aligned}\epsilon\sum_{|k|>2}\sum_{z=k(N+1)}^{k(N+1)+N} p_t(x-z)E_k(z,y) &\leq C(b)(1\wedge t^{-1/2})\sum_{|k|>2} e^{-b(|k|-1)N(1\wedge t^{-1/2})}C_0^{|k|}\\ &= C(b)(1\wedge t^{-1/2})e^{-bN(1\wedge t^{-1/2})}\sum_{k\in\mathbb{Z}} e^{\left(\log C_0-bN(1\wedge t^{-1/2})\right)|k|}\end{aligned}$$

where we used $\epsilon=N^{-1}$ and $|x-z|\geq(|k|-1)N$ in the first inequality. Noting that $N(1\wedge t^{-1/2})>T^{-1/2}$ we may assume that $b>2T^{1/2}\log C_0$ so that $\log C_0-bN(1\wedge t^{-1/2})<-\log C_0$ and thus the last expression is bounded by

$$C(b)(1\wedge t^{-1/2})e^{-bN(1\wedge t^{-1/2})}\frac{1}{1-C_0^{-1}}\leq C(A,B,b,T)(1\wedge t^{-1/2})e^{-b|x-y|(1\wedge t^{-1/2})}$$

where we used $|x-y| \le N$ and assumed wlog that $C_0 > 1$. This proves the claim. $\square$

Proposition 2.3.16. *Fix $A, B \in \mathbb{R}$ and $T > 0$. For $b \ge 0$, there is a constant $C(A,B,b,T)$ (not depending on $\epsilon = N^{-1}$) such that for all $t \in [0, \epsilon^{-2}T]$, all $v \in [0,1]$, all $|n| \le \lceil t^{1/2} \rceil$ and all $x, y \in \{0, ..., N\}$ we have that*

$$|\mathbf{p}_t^R(x+n,y) - \mathbf{p}_t^R(x,y)| \le C(A,B,b,T)(1 \wedge t^{-1/2-v})|n|^v e^{-b|x-y|(1\wedge t^{-1/2})}.$$

When $b = 0$ this bound holds for all n, not just for $n \le t^{1/2}$.

Proof. Again we use equation (2), replacing p_t with $\nabla_n p_t$ and noting that the corresponding bound already holds for the whole-line kernel (A.13) of [57]. Then we proceed exactly as in Proposition 2.3.15. The fact that the bound holds for all n when $b = 0$ is a consequence of the triangle inequality applied to the case when $n = 1$. $\square$

Corollary 2.3.17. *Fix $A, B \in \mathbb{R}$. For any $T \ge 0$ and $a \ge 0$, there exists some constant $C = C(a, A, B, T)$ such that for all $x \in \{0, ..., N\}$ and $t \le \epsilon^{-2}T$ we have*

$$\sum_{y\ge 0} \mathbf{p}_t^R(x,y) e^{a|x-y|(1\wedge t^{-1/2})} \le C,$$

$$\sum_{y\ge 0} |\nabla^+ \mathbf{p}_t^R(x,y)| e^{a|x-y|(1\wedge t^{-1/2})} \le C(1 \wedge t^{-1/2}).$$

Proof. One may mimic the proof of Corollary 2.3.3 with $a_2 = 0$, but use Propositions 2.3.15 and 2.3.16 instead of 2.3.1 and 2.3.2. $\square$

We will now prove the temporal estimate (analog of Proposition 2.3.5 on bounded intervals). For this, the above spatial methods do not work well, so we need information on the spectrum of the Laplacian with Robin boundary conditions.

Lemma 2.3.18. *Consider the operator $\frac{1}{2}\Delta$ on $\mathbb{R}^{\{0,...,N\}}$ with Robin boundary conditions $f(-1) := \mu_A f(0)$ and $f(N+1) := \mu_B f(N)$. This is a symmetric operator (and hence orthonormally diagonalizable with real eigenvalues). Consider N large. For $1 \le k \le N-1$, there is an* *eigenvalue of the form* $\lambda_k = 1 - \cos\omega_k$ *with* $\omega_k \in [k\pi/(N+1), (k+1)\pi/(N+1)]$. *In particular there are at most two positive eigenvalues* λ_N, λ_{N+1}, *and moreover these satisfy the bound* $\lambda_{N+1} \vee \lambda_N \leq C(A,B)\epsilon^2$.

Proof. The fact that $\frac{1}{2}\Delta$ is symmetric follows from the matrix representation using the standard basis $\{e_x\}_{0\leq x\leq N}$ where $e_x(y) = 1_{[x=y]}$:

$$\Delta = \begin{bmatrix} \mu_A - 2 & 1 & 0 & \dots & 0 & 0 \\ 1 & -2 & 1 & \dots & 0 & 0 \\ 0 & 1 & -2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -2 & 1 \\ 0 & 0 & 0 & \dots & 1 & \mu_B - 2 \end{bmatrix}.$$

If $A = B = 0$ (which means that $(\mu_A, \mu_B) = (1,1)$) then one may check directly that $N+1$ independent eigenvectors are given by $\cos(\omega_k x)$ where $\omega_k = k\pi/(N+1)$ for $k \in \{0, ..., N\}$. In this case the eigenvalues are precisely $\lambda_k = 1 - \cos\omega_k$.

If $(\mu_A, \mu_B) \neq (1,1)$ then by equation (4.27) in [45], all negative eigenvalues of this operator are of the form $\lambda = 1 - \cos\omega$, where $\omega \in (0, \pi)$ solves the following equation

$$f(\omega) := \sin(\omega(N+2)) - (\mu_A + \mu_B)\sin(\omega(N+1)) + \mu_A\mu_B \sin(\omega N) = 0.$$

Letting $\omega = k\pi/(N+1)$, we see that

$$\begin{aligned} f\left(\frac{k\pi}{N+1}\right) &= \sin\left(k\pi + \frac{k\pi}{N+1}\right) + \mu_A\mu_B \sin\left(k\pi - \frac{k\pi}{N+1}\right) \\ &= (1-\mu_A\mu_B)\sin\left(k\pi + \frac{k\pi}{N+1}\right) \\ &= (-1)^k(1-\mu_A\mu_B)\sin\left(\frac{k\pi}{N+1}\right) \\ &= (-1)^k\left(\frac{A+B}{N} - \frac{AB}{N^2}\right)\sin\left(\frac{k\pi}{N+1}\right). \end{aligned}$$

Now notice that unless $A = B = 0$, the middle term in this last expression cannot vanish for arbitrarily large N. We already ruled out this case, thus when N is large enough this last expression will always alternate sign as a function of $k \in \{1, ..., N-1\}$. By the intermediate value theorem applied to f, we find $N-1$ solutions to the eigenvalue equation above, one in each of the intervals $\big(k\pi/(N+1), (k+1)\pi/(N+1)\big)$ for $k \in \{1, ..., N-1\}$ (this does not work for $k \in \{0, N\}$ since $\omega = 0, \pi$ are solutions which do not correspond to nontrivial eigenvectors). Thus the first claim is proved.

Now let us consider the (at most two) positive eigenvalues. By the basic methods of recursive sequences, there must be an associated eigenfunction of the form

$$\psi(x) = \mu^{-x} + c\mu^{x-N}$$

for some $\mu > 0$ and $c \in \mathbb{R}$ which satisfy the relations

$$\mu_A = \frac{\mu + c\mu^{-(N+1)}}{1 + c\mu^{-N}} =: h_1(c,\mu) \quad , \quad \mu_B = \frac{\mu^{-(N+1)} + c\mu}{\mu^{-N} + c} =: h_2(c,\mu).$$

Note that the functions h_1, h_2 just defined satisfy the relation $h_2(c,\mu) = h_1(c,\mu^{-1}) = h_1(c^{-1},\mu)$ which means that we may assume $\mu > 1$ and $|c| \le 1$, after possibly interchanging the roles of μ_A and μ_B a couple of times (this does not change the eigenvalues or eigenfunctions).

If the above relations are satisfied with $\mu > 1$ and $c \in [-1, 0]$ then $\mu < h_1(c, \mu)$ and thus $1 < \mu < \mu_A$ so that the eigenvalue associated with μ satisfies

$$\lambda = \mu + \mu^{-1} - 2 \leq \mu_A + \mu_A^{-1} - 2 \leq C(A)\epsilon^2.$$

In the last inequality we used the fact that $\mu_A = 1 - A\epsilon$ so that $\mu_A^{-1} = 1 + A\epsilon + A^2\epsilon^2 + O(\epsilon^3)$.

On the other hand, if the above relations are satisfied with $\mu > 1$ and $c \in (0, 1]$, then $\mu_A\mu_B > 1$ (because $h_1(c, \mu)h_2(c, \mu) > 1$ for $c > 0$ by direct computation) and also $\mu_A < \mu$ (which is because $h_1(c, \mu) < \mu$ for $c > 0$). The fact that $\mu_A\mu_B > 1$ implies that $A < 0$ or $B < 0$, let's say $A < 0$. Note that

$$|\mu_A - 1| = \frac{|\mu - 1||1 - c\mu^{-(N+1)}|}{1 + c\mu^{-N}}.$$

Since $\mu^{-1} < \mu_A^{-1}$ and $c \in [0, 1]$, it follows that $1 - c\mu^{-(N+1)} \geq 1 - \mu_A^{-(N+1)}$. But $\mu_A^{-(N+1)} = (1 - A/N)^{-(N+1)} \leq e^{-|A|}$ for large enough N. Using these bounds together with the last expression shows

$$|\mu - 1| = \frac{1 + c\mu^{-N}}{1 - c\mu^{-(N+1)}}|\mu - 1| \leq \frac{2}{1 - e^{-|A|}}|\mu_A - 1| = C(A)|\mu_A - 1|$$

and thus $\mu < 1 + C(A)|\mu_A - 1| = 1 + C(A)\epsilon$, so that the associated eigenvalue satisfies

$$\lambda = \mu + \mu^{-1} - 2 \leq C(A)\epsilon^2.$$

This completes the proof. $\square$

Remark 2.3.19. *It is worth mentioning that positive eigenvalues will exist if and only if $A + B + AB < 0$, though we do not need this stronger claim. Similarly, zero will be an eigenvalue iff $A + B + AB = 0$. The key idea in proving these statements is to note that whenever 0 is an eigenvalue, the corresponding eigenfunction must be of the form $c + dx$ for some $c, d \in \mathbb{R}$, and then checking the boundary conditions necessarily forces $A + B + AB = 0$.*

Lemma 2.3.20. *In the same setting as Lemma B.4, let ψ_k denote the L^2-normalized eigenfunction associated to the eigenvalue λ_k. There exists some $C = C(A,B)$ such that for large enough N we have that $|\psi_k(x)| \leq CN^{-1/2}$ for all $0 \leq x, k \leq N$.*

Proof. If $\lambda_k < 0$ then this is proved in Lemma 4.10 of [45].

If $\lambda_k > 0$, then we can write the non-normalized eigenfunction as

$$\psi_k(x) = \mu^{-x} + c\mu^{x-N}$$

where $\mu > 1$ and $|c| \leq 1$ as in the proof of Lemma 2.3.18. Then

$$\begin{aligned}\sum_{k=0}^{N} \psi_k(x)^2 &= (1+c^2)\frac{\mu^{2(N+1)} - 1}{\mu - 1} + 2c(N+1) \\ &\leq 4(N+1)\mu^{2(N+1)} + 2(N+1)\end{aligned}$$

where we used the fact that $(x^k - 1)/(x-1) \leq kx^k$ for $x > 1$. From the proof of Lemma 2.3.18, we know that $\mu \leq 1 + C(A,B)\epsilon$, hence we see that $\mu^{2(N+1)} \leq e^{2C(A,B)} = C'(A,B)$. Hence the expression above is bounded by $(4C(A,B)+2)(N+1) \leq C'(A,B)N$. So by renormalizing, we find that $\psi_k(x) \leq C(A,B)N^{-1/2}$.

If $\lambda_k = 0$ then the associated eigenfunction is

$$\psi_k(x) = c + dx.$$

This can happen for arbitrarily large N iff $A + B + AB = 0$ and $d/c = A/N$ (as one may check by writing out the boundary conditions, for instance $1 - d/c = \psi(-1)/\psi(0) = \mu_A = 1 - A/N$, etc). The condition for the eigenfunction to be normalized simplifies to

$$c^2(N+1) + cdN(N+1) + d^2N(N+1)(2N+1)/6 = 1.$$

Then by writing $d = Ac/N$ or $c = dN/A$ we can check that $c^2 \leq C_1(A)N^{-1}$ and $d^2 \leq C_2(A)N^{-3}$. Consequently $\psi_k(x) = c + dx \leq C(A)(N^{-1/2} + (N^{-3/2})N) = C(A)N^{-1/2}$ for $x \in \{0, ..., N\}$. This proves the claim. $\square$

Proposition 2.3.21. *Fix $A, B \in \mathbb{R}$ and $T > 0$. There is a constant $C(A, B, T)$ (not depending on $\epsilon = N^{-1}$) such that for all $s < t \in [0, \epsilon^{-2}T]$ and all $v \in [0, 1]$ we have*

$$|\mathbf{p}_t^R(x, y) - \mathbf{p}_s^R(x, y)| \leq C(A, B, T)(1 \wedge s^{-1/2-v})(t - s)^v.$$

Proof. Just like Proposition 2.3.5, we only need to prove this when $v = 0$ and $v = 1$. The $v = 0$ case follows from Proposition 2.3.15. So let us prove the $v = 1$ case.

Let $S(t) := e^{\frac{1}{2}t\Delta}$ and let $\{e_x\}$ denote the standard basis functions on $\mathbb{R}^{\{0,...,N\}}$. Let ψ_k $(1 \leq k \leq N + 1)$ denote the orthonormal eigenunctions of $\frac{1}{2}\Delta$. Then $e_x = \sum_k \langle e_x, \psi_k \rangle \psi_k = \sum_k \psi_k(x)\psi_k$. So letting $\langle \cdot, \cdot \rangle$ denote the inner product on $L^2(\{0, ..., N\})$ we find

$$\mathbf{p}_t^R(x, y) = \langle S(t)e_x, e_y \rangle = \sum_{k,\ell} \psi_k(x)\psi_\ell(y)\langle S(t)\psi_k, \psi_\ell \rangle = \sum_k \psi_k(x)\psi_k(y)e^{\lambda_k t}$$

which implies that

$$\mathbf{p}_t^R(x, y) - \mathbf{p}_s^R(x, y) = \sum_k \psi_k(x)\psi_k(y)e^{\lambda_k s}(e^{\lambda_k(t-s)} - 1).$$

We will split this last sum into two pieces based on the sign of the eigenvalues. Let us first consider negative eigenvalues. Using Lemmas 2.3.18 and 2.3.20, together with the following

$$|e^{-q} - 1| \leq q, \quad q \geq 0$$

$$c_1 u^2 \leq |1 - \cos u| \leq c_2 u^2, \quad u \in [0, \pi]$$

we obtain the following bound

$$
\begin{aligned}
\sum_{k:\lambda_k<0} \psi_k(x)\psi_k(y)e^{\lambda_k s}(e^{\lambda_k(t-s)}-1) &\le C(A,B) \sum_{k:\lambda_k<0} N^{-1}e^{\lambda_k s}|\lambda_k|(t-s) \\
&\le \frac{C(A,B)}{N}(t-s)\sum_{k:\lambda_k<0} e^{\lambda_k s}\left|1-\cos\left(\frac{k\pi}{N+1}\right)\right| \\
&\le \frac{C(A,B)}{N}(t-s)\sum_{k=0}^{N} e^{-c_1 s k^2/(N+1)^2}\frac{c_2 k^2}{(N+1)^2}.
\end{aligned}
$$

This last expression can be interpreted as a Riemann sum for the integral

$$
\begin{aligned}
C(A,B)(t-s)\int_0^1 x^2 e^{-c_1 s x^2}dx &\le C(A,B)(t-s)\int_0^\infty x^2 e^{-c_1 s x^2}dx \\
&= C(A,B)(t-s)s^{-3/2}\int_0^\infty u^2 e^{-c_1 u^2}du \\
&= C(A,B)s^{-3/2}(t-s) \qquad (2.7)
\end{aligned}
$$

where we made the substitution $x\sqrt{s}=u$ in the second line. So when N is large, we are close enough to this integral that the same bound holds.

Next we consider the terms with positive eigenvalues. By Lemmas 2.3.18 and 2.3.20, and the fact that $e^q - 1 \le qe^q$ for $q \ge 0$ we see that

$$
\begin{aligned}
\sum_{k:\lambda_k>0} \psi_k(x)\psi_k(y)e^{\lambda_k s}(e^{\lambda_k(t-s)}-1) &\le 2\cdot\frac{C}{N}\cdot e^{C(A,B)\epsilon^2 s}\cdot \epsilon^2(t-s)e^{C(A,B)\epsilon^2(t-s)} \\
&= 2C\epsilon^3(t-s)e^{C(A,B)\epsilon^2 t} \qquad (2.8) \\
&\le C(A,B,T)s^{-3/2}(t-s)
\end{aligned}
$$

where in the last inequality we used the fact that $s<t<\epsilon^{-2}T$ so that $\epsilon^2 t \le T$ and $\epsilon^3 \le C(T)s^{-3/2}$. This proves the claim. $\square$

Proposition 2.3.22 (Long-time Estimate)**.** *There exist constants* $C=C(A,B)$ *and* $K=K(A,B)$

such that for every $t \geq 0$ and $x, y \geq 0$ we have that

$$\mathbf{p}_t^R(x,y) \leq C(t^{-1/2} + \epsilon)e^{K\epsilon^2 t}.$$

Just like Proposition 2.3.6 this is a "long-time" estimate because it is true uniformly over all $t > 0$, i.e., constants don't depend on any terminal time $\epsilon^{-2}T$.

Proof. In equations (2.7) and (2.8) above, note that the constants do not depend on the terminal time T. This means that there are C and K such that for any $s < t$ we have that

$$|\mathbf{p}_t^R(x,y) - \mathbf{p}_s^R(x,y)| \leq C\epsilon^3(t-s)e^{K\epsilon^2 t} + Cs^{-3/2}(t-s).$$

Dividing by $t - s$ and letting $s \to t$, we find

$$|\partial_t \mathbf{p}_t^R(x,y)| \leq C\epsilon^3 e^{K\epsilon^2 t} + Ct^{-3/2}$$

By Proposition 2.3.15, the desired bound already holds when $t \leq \epsilon^{-2}$, thus we only consider the case when $t > \epsilon^{-2}$. Using the above expression and Proposition 2.3.15,

$$\begin{aligned}
\mathbf{p}_t^R(x,y) &\leq \mathbf{p}_{\epsilon^{-2}}^R(x,y) + \int_{\epsilon^{-2}}^t |\partial_s \mathbf{p}_s^R(x,y)| ds \\
&\leq C(\epsilon^{-2})^{-1/2} + C\int_{\epsilon^{-2}}^t (\epsilon^3 e^{K\epsilon^2 s} + Cs^{-3/2}) ds \\
&= C\epsilon + \frac{C}{K}\epsilon(e^{K\epsilon^2 t} - e^K) + 2C(\epsilon - t^{-1/2}) \\
&\leq C'\epsilon e^{K\epsilon^2 t}
\end{aligned}$$

which proves the claim. □

Proposition 2.3.23. *For all $A, B \in \mathbb{R}$ and $T > 0$, there exists $C = C(A,B,T)$ such that for*

$t \in [0, \epsilon^{-2}T]$ *and* $v \in [0,1]$ *we have*

$$\sup_{0 \le x \le N} \left| \sum_{y=0}^{N} \mathbf{p}_t^R(x,y) \; - \; 1 \right| \le C\epsilon^v t^{v/2}.$$

Proof. The proof is basically the same as Proposition 2.3.7. We define

$$f(t,x) := \sum_{y=0}^{N} \mathbf{p}_t^R(x,y).$$

The fact that $\mathbf{p}_t^R$ is symmetric in x and y follows easily by symmetry of the operator $\frac{1}{2}\Delta$ with Robin boundary conditions. The same arguments used in Lemma 2.3.7 then show that

$$\partial_t f(t,x) = \frac{1}{2}(\mu_A - 1)\mathbf{p}_t^R(0,x) + \frac{1}{2}(\mu_B - 1)\mathbf{p}_t^R(N,x)$$

so that $|\partial_t f(t,x)| \le C(A,B,T) \; \epsilon \; t^{-1/2}$ by Proposition 2.3.15. Now we proceed exactly as in Lemma 2.3.7. □

We now turn to proving the cancellation estimates (the analogs of Propositions 2.3.11 and 2.3.12 for bounded-interval Robin heat kernels). For this we would like a result like Lemma 2.3.8, but we are not able to prove the corresponding result on bounded intervals, so we opt for a weaker result which will suffice for us (Lemma 2.3.25).

Lemma 2.3.24. *Let* M *and* N *be symmetric* $n \times n$ *matrices (or more generally, self-adjoint operators on some Hilbert space) whose set of eigenvalues (or spectral values) is bounded above by* $\alpha \in \mathbb{R}$. *Then*

$$\|e^M - e^N\| \le e^{\alpha}\|M - N\|$$

where e^M *denotes the matrix exponential, and* $\|\cdot\|$ *is the operator norm with respect to the underlying Hilbert space norm.*

Proof. When M and N commute the proof is easy by simultaneous diagonalization. For the general case, it is tempting to use $\|e^M - e^N\| \leq \|M - N\| e^{\max\{\|M\|, \|N\|\}}$, however this crude bound fails since it only takes into account the magnitude of the eigenvalues and not their sign (note that α might be negative). Instead we use Frechet calculus. If $\mathcal{B}$ is a Banach space, $f : \mathcal{B} \to \mathcal{B}$ is Frechet differentiable, and $\gamma : [0, 1] \to \mathcal{B}$ is a smooth curve,

$$f(\gamma(1)) - f(\gamma(0)) = \int_0^1 Df(\gamma(t))\gamma'(t)dt$$

where $Df(x) \in L(\mathcal{B}, \mathcal{B})$ is the Frechet derivative. As an immediate corollary, we have for all $a, b \in \mathcal{B}$

$$\|f(a) - f(b)\| \leq \left(\sup_{x \in [a,b]} \|Df(x)\| \right) \|a - b\|$$

where $[a, b] := \{(1 - t)a + tb : t \in [0, 1]\}$. We now specialize this bound to the case when $f(X) = e^X$ with $\mathcal{B} = L(\mathcal{H}, \mathcal{H})$ for a Hilbert space $\mathcal{H}$. In this case, there is a well-known formula for the Frechet derivative

$$Df(X)H = \int_0^1 e^{sX} H e^{(1-s)X} ds$$

which immediately implies that the operator norm of $Df(X)$ satisfies

$$\|Df(X)\| \leq \int_0^1 \|e^{sX}\| \cdot \|e^{(1-s)X}\| ds.$$

Now suppose that M, N are self-adjoint operators on $\mathcal{H}$ which satisfy $\langle Mu, u\rangle \leq \alpha\|u\|^2$ and $\langle Nu, u\rangle \leq \alpha\|u\|^2$ (i.e., the largest eigenvalue of M and of N is bounded above by α, which can be negative). In this case we easily have that $\langle Xu, u\rangle \leq \alpha\|u\|^2$ for all X in the interval $[M, N]$ (as defined above). Consequently $\|e^{sX}\| \leq e^{\alpha s}$ for all s. Thus if $X \in [M, N]$

$$\|Df(X)\| \leq \int_0^1 e^{s\alpha} e^{(1-s)\alpha} ds = e^{\alpha}$$

so that $\|f(M) - f(N)\| \leq e^{\alpha}\|M - N\|$ as desired. □

Lemma 2.3.25. *For the next few estimates we will distinguish between different values of* (A, B) *by writing* $\mathbf{p}_t^R(x, y; A, B)$ *for the (*ϵ*-dependent) Robin heat kernel of parameters* A, B*. For all* $A, B \in \mathbb{R}$*, all* $T > 0$*, and all* $b \geq 0$*, there exists* $C = C(A, B, b, T)$ *such that for all* $x, y \in \mathbb{Z}_{\geq 0}$*, all* $t \in [0, \epsilon^{-2}T]$*, and all* $v \in [0, 1]$ *we have*

$$|\mathbf{p}_t^R(x, y; A, B) - \mathbf{p}_t^R(x, y; 0, 0)| \leq C\, \epsilon^v\, (t^v \wedge t^{(3v-1)/2})\, e^{-(1-v)b|x-y|(1\wedge t^{-1/2})},$$
$$|\nabla^{\pm}\mathbf{p}_t^R(x, y; A, B) - \nabla^{\pm}\mathbf{p}_t^R(x, y; 0, 0)| \leq C\, \epsilon^v\, (t^v \wedge t^{2v-1}) e^{-(1-v)b|x-y|(1\wedge t^{-1/2})},$$

where $\nabla^{\pm}$ *denotes the discrete gradient in the first spatial coordinate.*

Proof. It suffices to prove the claim $v = 0$ and $v = 1$. The middle cases then follow from interpolation. The $v = 0$ case follows easily from Propositions 2.3.15 and 2.3.16, thus we only consider the $v = 1$ case. So we will show that

$$|\mathbf{p}_t^R(x, y; A, B) - \mathbf{p}_t^R(x, y; 0, 0)| \leq C\, \epsilon\, t,$$

$$|\nabla^{\pm}\mathbf{p}_t^R(x, y; A, B) - \nabla^{\pm}\mathbf{p}_t^R(x, y; 0, 0)| \leq C\, \epsilon\, t,$$

for some constant $C = C(A, B, T)$ (we postulate that these bounds are not optimal, but they suffice to prove Proposition 2.3.26 below). To prove these bounds, we will use Lemma 2.3.24 with the two matrices $\frac{t}{2}\Delta_{A,B}$ and $\frac{t}{2}\Delta_{0,0}$. Here $\Delta_{A,B}$ denotes the Laplacian on $\{0, ..., N\}$ with Robin boundary parameters μ_A, μ_B. Then Lemma 2.3.24 and Lemma 2.3.18 show that

$$\|e^{\frac{1}{2}\Delta_{A,B}t} - e^{\frac{1}{2}\Delta_{0,0}t}\| \leq e^{C(A,B)\epsilon^2 t} \cdot \frac{t}{2}\|\Delta_{A,B} - \Delta_{0,0}\|.$$

But $t \leq \epsilon^{-2}T$ so that $\epsilon^2 t \leq T$ and thus $e^{C(A,B)\epsilon^2 t} \leq C(A,B,T)$. Now notice that

$$\Delta_{A,B} - \Delta_{0,0} = \begin{bmatrix} \mu_A - 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \mu_B - 1 \end{bmatrix}.$$

This is a diagonal matrix with eigenvalues $\{-A\epsilon, 0, ..., 0, -B\epsilon\}$ and thus we see that $\|\Delta_{A,B} - \Delta_{0,0}\| \leq C(A,B)\epsilon$. Summarizing, we have shown that

$$\|e^{\frac{1}{2}\Delta_{A,B}t} - e^{\frac{1}{2}\Delta_{0,0}t}\| \leq C(A,B,T)\ \epsilon\ t.$$

Therefore

$$\begin{aligned} |\mathbf{p}_t^R(x,y;A,B) - \mathbf{p}_t^R(x,y;0,0)| &= \left| \left\langle (e^{\frac{1}{2}\Delta_{A,B}t} - e^{\frac{1}{2}\Delta_{0,0}t})\mathbf{1}_x, \mathbf{1}_y \right\rangle \right| \\ &\leq \|e^{\frac{1}{2}\Delta_{A,B}t} - e^{\frac{1}{2}\Delta_{0,0}t}\| \\ &\leq C(A,B,T)\epsilon t. \end{aligned}$$

Similarly,

$$\begin{aligned} |\nabla^{\pm}\mathbf{p}_t^R(x,y;A,B) - \nabla^{\pm}\mathbf{p}_t^R(x,y;0,0)| &= \left| \left\langle (e^{\frac{1}{2}\Delta_{A,B}t} - e^{\frac{1}{2}\Delta_{0,0}t})(\mathbf{1}_x - \mathbf{1}_{x\pm1}), \mathbf{1}_y \right\rangle \right| \\ &\leq 2\|e^{\frac{1}{2}\Delta_{A,B}t} - e^{\frac{1}{2}\Delta_{0,0}t}\| \\ &\leq C(A,B,T)\epsilon t \end{aligned}$$

which proves the claim. □

Lemma 2.3.26. *For $A, B \in \mathbb{R}$, $t \geq 0$, and $x, y \in \mathbb{Z}_{\geq 0}$ we define*

$$K_t(x, y; A, B) := \nabla^+ \mathbf{p}_t^R(x, y; A, B) \nabla^- \mathbf{p}_t^R(x, y; A, B).$$

For $T \geq 0$ there exists a constant $C = C(A, B, T)$ such that

$$\sum_{y=1}^{N-1} \int_0^{\epsilon^{-2}T} |K_t(x, y; A, B) - K_t(x, y; 0, 0)| dt \leq C\epsilon^{1/8}.$$

Proof. The proof is very similar to that of Lemma 2.3.9 with $a = 0$. The only difference is that instead of using Lemma 2.3.8, we now use the second bound in Lemma 2.3.25 with $v = 1/8$, and consequently we get a factor of $\epsilon^{1/8}$ instead of $\epsilon^{1/2}$. $\square$

Lemma 2.3.27. *Let us write $\mathbf{p}_t^R(x, y; A, B)$ as in the preceding lemma. For $x, \bar{x} \in \{0, ..., N\}$, we have*

$$\sum_{y \geq 0} \int_0^\infty \nabla^+ \mathbf{p}_t^R(x, y; 0, 0) \nabla^+ \mathbf{p}_t^R(\bar{x}, y; 0, 0) dt = -\frac{1}{N+1} 1_{\{x \neq \bar{x}\}} + \frac{N}{N+1} 1_{\{x = \bar{x}\}}.$$

Proof. Let us recall the summation-by-parts identity: if $u, v : \{-1, ..., N+1\} \to \mathbb{R}$, then

$$\sum_{y=0}^N u(y) \Delta v(y) = u(N+1) \nabla^+ v(N) + u(-1) \nabla^- v(0) - \sum_{y=-1}^N \nabla^+ u(y) \nabla^+ v(y).$$

Letting $u = \mathbf{p}_t^R(x, \cdot; 0, 0)$ and $v = \mathbf{p}_t^R(\bar{x}, \cdot; 0, 0)$, the boundary terms will all vanish and we get

$$-\sum_{y=0}^N \nabla^+ \mathbf{p}_t^R(x, y) \nabla^+ \mathbf{p}_t^R(\bar{x}, y) = \sum_{y=0}^N \mathbf{p}_t^R(x, y) \Delta \mathbf{p}_t^R(\bar{x}, y) = \sum_{y=0}^N \Delta \mathbf{p}_t^R(x, y) \mathbf{p}_t^R(\bar{x}, y).$$

Since $\Delta \mathbf{p}_t^R = 2\partial_t \mathbf{p}_t^R$, this implies that

$$-\sum_{y=0}^N \nabla^+ \mathbf{p}_t^R(x, y) \nabla^+ \mathbf{p}_t^R(\bar{x}, y) = \sum_{y=0}^N \mathbf{p}_t^R(x, y) \partial_t \mathbf{p}_t^R(\bar{x}, y) + \partial_t \mathbf{p}_t^R(x, y) \mathbf{p}_t^R(\bar{x}, y).$$

Integrating both sides from $t = 0$ to ∞ and using the semigroup property, we find that

$$-\sum_{y=0}^{N}\int_0^\infty \nabla^+\mathbf{p}_t^R(x,y)\nabla^+\mathbf{p}_t^R(\bar{x},y)dt = \sum_{y=0}^{N}\mathbf{p}_t^R(x,y)\mathbf{p}_t^R(\bar{x},y)\Big|_{t=0}^{\infty}$$
$$= \mathbf{p}_{2t}^R(x,\bar{x})\Big|_{t=0}^{\infty}$$
$$= \frac{1}{N+1} - 1_{\{x=\bar{x}\}}$$

which proves the claim. In the final line, we used the fact that $\lim_{t\to\infty}\mathbf{p}_t^R(x,\cdot;0,0)$ is the uniform measure on $\{0,...,N\}$, which follows from the basic theory of finite-state Markov Chains. $\square$

Proposition 2.3.28. *Let $A, B \in \mathbb{R}$, and $T > 0$. Let K_t be as in Lemma 2.3.26. There exists some $\epsilon_0 = \epsilon_0(A,B,T)$ and some $c_* = c_*(A,B,T) < 1$ such that for $\epsilon < \epsilon_0$ and $x \in \{1,...,N-1\}$*

$$\sum_{y=1}^{N-1}\int_0^{\epsilon^{-2}T}|K_t(x,y;A,B)|dt \le c_*.$$

Moreover, for any $S \in [0,T]$ there is a $C = C(A,B,T,S)$ such that for $s = \epsilon^{-2}S$ we have

$$\sum_{y=1}^{N-1}\int_0^{s}|K_t(x,y;A,B)|(s-t)^{-1/2}dt \le C\epsilon.$$

Proof. The proof of the first bound is the same as that of Proposition 2.3.11 with $a = 0$, the only difference is the factor of $\epsilon^{1/8}$ rather than $\epsilon^{1/2}$. Obviously one should use Lemmas 2.3.26 and 2.3.27 instead of the analogous half-line estimates 2.3.9 and 2.3.10.

The proof of the second bound can be copied verbatim from the proof of Proposition 2.3.12, with $a = 0$. $\square$

Finally we conclude this section with the analogue of Proposition 2.3.13:

Proposition 2.3.29. *For every $t \geq s \geq 0$ and $x, y \geq 0$ we have that*

$$\mathbf{p}_s^R(x,y) \leq e^{t-s}\mathbf{p}_t^R(x,y).$$

Proof. Similar to the proof of Proposition 2.3.13. $\square$

2.3.3 Continuum Estimates

Throughout this section, I will denote either $[0,\infty)$ or $[0,1]$. Similarly, Λ will denote either $\mathbb{Z}_{\geq 0}$ or $\{0,...,N\}$. We will fix Robin boundary parameters A and B. As before, $\mathbf{p}_t^R$ will denote the (discrete-space, continuous-time, ϵ-dependent) Robin heat kernel on Λ, with boundary parameters μ_A, μ_B. By an abuse of notation, we will write $\mathbf{p}_t(x,y)$ even when $x, y \in \mathbb{R}$, and this quantity is meant to be understood as a linear interpolation of the values of $\mathbf{p}_t^R$ from nearby integer-coordinate points.

Theorem 2.3.30 (Existence of the Continuum Robin Heat Kernel)**.** *For $T \geq 0$ and $X, Y \in I$ let*

$$P_T^\epsilon(X,Y) := \epsilon^{-1}\mathbf{p}^R_{\epsilon^{-2}T}(\epsilon^{-1}X, \epsilon^{-1}Y).$$

Then $P_T^\epsilon(X,Y)$ converges to a limit $P_T(X,Y)$ as $\epsilon \to 0$. For any $0 < \delta < \tau$, the convergence is uniform over $(T,X,Y) \in [\delta,\tau] \times I \times I$. Moreover, the limit $P_T(X,Y)$ satisfies

$$\partial_T P_T(X,Y) = \frac{1}{2}\partial_X^2 P_T(X,Y),$$

$$\partial_X P_T(0,Y) = AP_T(0,Y),$$

$$\partial_X P_T(1,Y) = BP_T(1,Y), \qquad \text{if } I = [0,1].$$

Furthermore, for every $X \in I$, $P_T(X,\cdot)$ converges weakly to δ_X as $T \to 0$.

Proof. Let $\tau > \delta > 0$. Using Propositions 2.3.1, 2.3.2, and 2.3.5 (or 2.3.15, 2.3.16, and 2.3.21) with $b = 0$ and $v = 1$, we easily verify that for every $S < T \in [\delta,\tau]$ and $X, Y, Z \in I$ we have (say, for $\epsilon \leq 1$)

$$P^\epsilon_T(X,Y) \leq CT^{-1/2},$$

$$|P^\epsilon_T(X,Z) - P^\epsilon_T(Y,Z)| \leq CT^{-1}|X-Y|,$$

$$|P^\epsilon_T(X,Y) - P^\epsilon_S(X,Y)| \leq CS^{-3/2}|T-S|,$$

where C is a constant depending only on A, B, and the terminal time τ. Since $S, T > \delta$, we find that $T^{-1/2} \leq \delta^{-1/2}$, also $T^{-1} \leq \delta^{-1}$, and similarly $S^{-3/2} \leq \delta^{-3/2}$. Therefore, we have proved that the collection $\{P^\epsilon_\cdot(\cdot,\cdot)\}_{\epsilon\in(0,1]}$ is uniformly bounded and uniformly Lipchitz (in both the time variable and in both spatial variables by symmetry) on $[\delta,\tau]\times I\times I$. By the Arzela-Ascoli theorem, we conclude that the family $\{P^\epsilon_\cdot(\cdot,\cdot)\}_{\epsilon\in(0,1]}$ is precompact in $C([\delta,\tau]\times I\times I)$, so there is at least one limit point as $\epsilon \to 0$.

We will now show that any limit point of the $\{P^\epsilon_\cdot(\cdot,\cdot)\}$ coincides in a *weak* sense with the fundamental solution of the Robin-boundary heat equation on I (this weak formulation is good enough, because any continuous weak solution is automatically a strong solution by the standard methods of PDE). In other words, if $P_T(X,Y)$ is a limit point, we will show that for any $\phi \in C_c^\infty((0,\infty)\times\mathbb{R})$ satisfying $\partial_X\phi(T,0) = A\phi(T,0)$ (and $\partial_X\phi(T,1) = B\phi(T,1)$ if $I = [0,1]$),

$$-\int_I\int_0^\infty P_T(X,Y)\partial_T\phi(T,X)dTdX = \int_I P_T(X,Y)\partial_X^2\phi(T,X)dTdX. \tag{2.9}$$

To prove this, first note that for any $\epsilon > 0$ and any $X, Y \in \epsilon\mathbb{Z}$, we have (by definition) that $\partial_T P^\epsilon_T(X,Y) = \epsilon^{-2}(P^\epsilon_T(X+\epsilon,Y) + P^\epsilon_T(X-\epsilon,Y) - 2P^\epsilon_T(X,Y))$. By the linear interpolation, this is still true for $X, Y \in \mathbb{R}$. Therefore we find that

$$-\int_I\int_0^\infty P^\epsilon_T(X,Y)\partial_T\phi(T,X)dTdX = \int_I\int_0^\infty \partial_T P^\epsilon_T(X,Y)\phi(T,X)dTdX \tag{2.10}$$

$$= \int_I \int_0^\infty \epsilon^{-2}\big[P_T^\epsilon(X+\epsilon,Y) + P_T^\epsilon(X-\epsilon,Y) - 2P_T^\epsilon(X,Y)\big]\phi(T,X)dTdX. \tag{2.11}$$

We can separate the expression (2.11) as the sum of three separate integrals based on the three terms appearing in the square parentheses, then we make the substitution $X \mapsto X-\epsilon$ in the first integral, we make the substitution $X \mapsto X+\epsilon$ in the second integral, and we leave the third integral as is. After those calculations we obtain the following expression:

$$\int_I \int_0^\infty P_T^\epsilon(X,Y)\cdot \epsilon^{-2}\big[\phi(T,X+\epsilon) + \phi(T,X-\epsilon) - 2\phi(T,X)\big]dTdX \;+\; O(\epsilon) \tag{2.12}$$

where the error term is due to the boundary correction near the endpoints of I. The fact that this boundary correction is $O(\epsilon)$ is a consequence of the boundary conditions: since $\mathbf{p}_t(-1,y) = \mu_A \mathbf{p}_t(0,y)$, it follows that $\partial_X^\pm P_T^\epsilon(0,Y) = AP_T(0,Y) + O(\epsilon)$, where $\partial_X^\pm$ denotes left/right derivatives. Similarly if $I=[0,1]$, then we also have $\partial_X^\pm P_T^\epsilon(1,Y) = BP_T^\epsilon(1,Y) + O(\epsilon)$. Also recall that $\partial_X \phi(T,0) = A\phi(T,0)$ (and $\partial_X\phi(T,1) = B\phi(T,1)$ if $I=[0,1]$). Using these facts and performing a first-order Taylor expansion of the integrand of the boundary correction gives the $O(\epsilon)$ error.

Since $\partial_X^2\phi$ is continuous and compactly supported, it follows that

$$\lim_{\epsilon\to 0} \sup_{\substack{T>0\\ X\in\mathbb{R}}} \Big| \partial_X^2\phi(T,X) - \epsilon^{-2}\big[\phi(T,X+\epsilon) + \phi(T,X-\epsilon) - 2\phi(T,X)\big]\Big| = 0.$$

Indeed, this can be seen by using the fundamental theorem of calculus to rewrite the parenthetical term: $\phi(T,X+\epsilon) + \phi(T,X-\epsilon) - 2\phi(T,X) = \int_X^{X+\epsilon}\int_{Z-\epsilon}^Z \partial_X^2\phi(T,W)dWdZ$, and then using uniform continuity of $\partial_X^2\phi$ on $(0,\infty)\times\mathbb{R}$.

So if $P_T(X,Y)$ is a limit point of $\{P_\cdot^\epsilon(\cdot,\cdot)\}_{\epsilon\in(0,1]}$, then there is a subsequence of $\epsilon \to 0$ such that $P_T^\epsilon(X,Y) \to P_T(X,Y)$ along that subsequence. Letting $\epsilon \to 0$ along the same subsequence in the LHS of (2.10) and in the equivalent expression (2.12) proves the desired identity (2.9), thus completing the proof that $P_T^\epsilon(X,Y)$ converges to a function $P_T(X,Y)$ which satisfies the heat equation with Robin boundary.

All that is left to show is that $P_T(X,\cdot)$ converges weakly to δ_X as $T \to 0$. In other words, if $\varphi \in C_c^\infty(\mathbb{R})$, we wish to show that $\int_I P_T^R(X,Y)\varphi(Y)dY \overset{T\to 0}{\longrightarrow} \varphi(X)$. This will require the usage of Propositions 2.3.31 and 2.3.33 proved below, hence the reader may wish to take a look at those estimates and return to this proof a bit later (note that the proofs of those estimates do not use this property of P_T, hence there is no circular logic here). First note by Proposition 2.3.33 that there exists some $C > 0$ such that for $T \le 1$ we have

$$\left| \int_I P_T(X,Y)dY - 1 \right| \le CT^{1/2}.$$

Therefore by the triangle inequality

$$\begin{aligned}\left| \int_I P_T^R(X,Y)\varphi(Y)dY - \varphi(X) \right| &\le \int_I P_T(X,Y)\big|\varphi(Y) - \varphi(X)\big|dY + \left| \int_I P_T(X,Y)dY - 1 \right| \varphi(X) \\ &\le \int_I P_T(X,Y)\big|\varphi(Y) - \varphi(X)\big|dY + CT^{1/2}\varphi(X).\end{aligned}$$

Hence it suffices to show that

$$\int_I P_T(X,Y)\big|\varphi(Y) - \varphi(X)\big|dY \overset{T\to 0}{\longrightarrow} 0.$$

Since $\varphi \in C_c^\infty(\mathbb{R})$ it follows that φ is Lipchitz so that $|\varphi(Y) - \varphi(X)| \le C|Y - X|$. Moreover by applying the first estimate in Proposition 2.3.31 with $b = 1$ we find that there exists C such that for

$T \leq 1$, we have $P_T(X,Y) \leq CT^{-1/2}e^{-|X-Y|/\sqrt{T}}$. Thus for $T \leq 1$,

$$\begin{aligned}
\int_I P_T(X,Y)|\varphi(Y)-\varphi(X)|dY &\leq CT^{-1/2}\int_I e^{-|X-Y|/\sqrt{T}}|Y-X|dY \\
&\leq CT^{-1/2}\int_{\mathbb{R}} e^{-|Z|/\sqrt{T}}|Z|dZ \\
&\leq CT^{-1/2}\int_{\mathbb{R}} e^{-|W|}|T^{1/2}W|(T^{1/2}dW) \\
&= CT^{1/2}
\end{aligned}$$

where we made the substitution $Z = Y - X$ in the second line, and another substitution $W = Z/\sqrt{T}$ in the third line. Now we let $T \to 0$ which proves the claim. $\square$

From now onward, $P_T(X,Y)$ will denote the continuum Robin heat kernel which has been constructed in Theorem 2.3.30.

Proposition 2.3.31. *Fix a terminal time $\tau > 0$. For any $b \geq 0$, there exists some constant $C = C(A,B,b,\tau)$ such that for all $S < T \leq \tau$ and $X,Y,Z \in I$ we have that*

$$P_T(X,Y) \leq CT^{-1/2}e^{-b|X-Y|/\sqrt{T}}, \tag{2.13}$$

$$|P_T(X,Z) - P_T(Y,Z)| \leq CT^{-1}|Y-Z|, \tag{2.14}$$

$$|P_T(X,Y) - P_S(X,Y)| \leq CS^{-3/2}|T-S|. \tag{2.15}$$

Proof. For $\epsilon > 0$, let $P_T^{\epsilon}(X,Y)$ be as in Theorem 2.3.30. Note that all of these estimates already hold for $P_T^{\epsilon}(X,Y)$, by Propositions 2.3.1, 2.3.2, and 2.3.5 (or by 2.3.15, 2.3.16, and 2.3.21). Moreover the constant C does not depend on ϵ. Letting $\epsilon \to 0$, it follows that these estimates still hold in the limit. $\square$

Proposition 2.3.32. *Fix a terminal time $\tau > 0$. For any $a \geq 0$, there exists a constant $C =$*

$C(a, \tau, A, B)$ *such that for* $T \leq \tau$ *and* $X \in I$,

$$\int_I P_T(X,Y) e^{aY} dY \leq C e^{aX}.$$

Proof. Using (2.13) with $b = 1 + a\tau^{1/2}$, and the fact that $e^{aY} \leq e^{aX} e^{a|X-Y|}$ for $X \in I$, one sees that there exists $C = C(a, A, B, \tau)$ such that for all $T \leq \tau$ and $X \in I$

$$\begin{aligned}
\int_I P_T(X,Y) e^{aY} dY &\leq C e^{aX} T^{-1/2} \int_I e^{-b|X-Y|/\sqrt{T}} e^{a|X-Y|} dY \\
&\leq C e^{aX} T^{-1/2} \int_{\mathbb{R}} e^{-b|Z|/\sqrt{T}} e^{a|Z|} dZ \\
&= C e^{aX} \int_{\mathbb{R}} e^{-b|W|} e^{aT^{1/2}|W|} dW \\
&\leq C e^{aX} \int_{\mathbb{R}} e^{-|W|} dW
\end{aligned}$$

where we made a substitution $Z = Y - X$ in the second line, and another substitution $W = T^{-1/2} Z$ in the next line. In the final line we used the fact that $aT^{1/2} - b \leq a\tau^{1/2} - b = -1$. $\square$

Proposition 2.3.33. *Fix a terminal time* $\tau \geq 0$. *Then there is a constant* $C = C(A, B, \tau)$ *such that for* $T \leq \tau$ *and* $X \in I$ *we have*

$$\left| \int_I P_T(X,Y) dY - 1 \right| \leq C T^{1/2}.$$

Proof. Let $P_T^\epsilon(X,Y)$ be as in Theorem 2.3.30. Using Proposition 2.3.7 or 2.3.23 with $v = 1$, there exists $C > 0$ such that for all $T \leq \tau$, $X \in I$, and (small enough) $\epsilon > 0$,

$$\left| \epsilon \sum_{Y \in \epsilon\Lambda} P_T^\epsilon(X,Y) - 1 \right| \leq C T^{1/2}. \tag{2.16}$$

Notice that

$$\epsilon \sum_{Y \in \epsilon\Lambda} P_T^\epsilon(X,Y) = \int_I P_T^\epsilon(X, \epsilon\lfloor \epsilon^{-1} Y \rfloor) dY.$$

Moreover, $P_T^\epsilon(X, \epsilon\lfloor \epsilon^{-1} Y \rfloor) \overset{\epsilon \to 0}{\longrightarrow} P_T(X,Y)$, by uniform convergence of $P_T^\epsilon(X, \cdot)$ to $P_T(X, \cdot)$.

Using Proposition 2.3.1 or 2.3.15 with $b = 1$, we have the following bound, uniformly over all small enough $\epsilon > 0$:

$$|P_T^\epsilon(X,Y)| \le CT^{-1/2}e^{-|Y-X|/\sqrt{T}}.$$

For each fixed X and T, the RHS is an integrable function of Y, so it follows from the dominated convergence theorem (together with the preceding observations) that

$$\lim_{\epsilon\to 0}\epsilon\sum_{Y\in\epsilon\Lambda}P_T^\epsilon(X,Y) = \lim_{\epsilon\to 0}\int_I P_T^\epsilon(X,\epsilon\lfloor\epsilon^{-1}Y\rfloor)dY = \int_I P_T(X,Y)dY.$$

Letting $\epsilon \to 0$ in (2.16) gives the result. □

2.4 The SHE with Robin Boundary Conditions

Next we want to describe the continuum version of the height functions in our particle system, which we expect will (in a sense) solve the KPZ Equation on the spatial domain I. Recall that this equation is formally given by

$$\partial_T H = \frac{1}{2}\partial_X^2 H + \frac{1}{2}(\partial_X H)^2 + \xi$$

where ξ is a Gaussian space-time white noise, meaning informally that $\mathbb{E}[\xi(S,X)\xi(T,Y)] = \delta(S-T)\delta(X-Y)$.

In order to solve this equation, let us first make precise how to rigorously define the noise term. One may construct ξ as the distributional time-derivative $\xi = \partial_T W$ of a cylindrical Wiener process $W = (W_T)_{T\ge 0}$ over $L^2(I)$, in which case each individual W_T may be viewed as a random element of the Sobolev space $H^s_{loc}(I)$ for $s < -1/2$. This viewpoint will be very useful to us because it allows one to define stochastic integrals against ξ, which in turn allows us to construct strong solutions to parabolic PDEs which are driven by ξ. See for instance [85, 52, 150] for the general theory of space-time stochastic integrals.

Given the regularity (or lack thereof) of the noise ξ, one may then heuristically compute (using Schauder estimates or the Kolmogorov continuity theorem) that the solution to the KPZ equation should be locally Hölder $1/2-$ in space and Hölder $1/4-$ in time, but no better. In particular, we are faced with two serious problems:

1. The nonlinear term $(\partial_X H)^2$ is undefined, since it is not possible to square the derivative of a function which is Hölder $1/2-$. So the PDE is ill-posed.

2. To make matters worse, the boundary parameters for ASEP should somehow correspond to Neumann boundary conditions for the PDE so that $\partial_X H(T,0) = A$ in the half-line case (and also $\partial_X H(T,1) = -B$ for the bounded interval case). But once again these quantities are ill-posed.

The first of the two problems described may be fixed by the well-known Hopf-Cole transform, in which we define $\mathcal{Z} := \exp H$ and then formally $\mathcal{Z}$ solves the multiplicative Stochastic Heat Equation (SHE):

$$\partial_T \mathcal{Z} = \frac{1}{2}\Delta \mathcal{Z} + \mathcal{Z}\xi$$

As it turns out, this transformation now makes the equation well-posed, and by considering the noise as a cylindrical Wiener process, one can hope to make sense of solutions in a Duhamel form, see Theorem 2.4.2 below. Once we construct this solution and prove it is positive, we can *formally* just define $H := \log \mathcal{Z}$. One may wish to look at [73, 70, 74] for methods which avoid the Hopf-Cole transformation.

So now we only need to make sense of the Neumann Boundary conditions described above for the KPZ equation. Under the Hopf-Cole transform just discussed, the Neumann boundary conditions for the KPZ equation formally become $Robin$ boundary conditions for the SHE:

$$\partial_X \mathcal{Z}(T,X)|_{X=0} = A\mathcal{Z}(T,0)$$

(and also $\partial_X \mathcal{Z}(T,X)|_{X=1} = -B\mathcal{Z}(T,1)$ for the bounded-interval case). If we were to forget about the noise term in the SHE for the moment being (thus we consider the equation $\partial_T \mathcal{Z} = \frac{1}{2}\Delta\mathcal{Z}$), then by theorem 2.3.30 the solution is given by a semigroup kernel:

$$\mathcal{Z}(T,X) = P_T * \mathcal{Z}_0(X) = \int_I P_T(X,Y)\mathcal{Z}(0,Y)dY.$$

This fact motivates the following definition of solutions in mild (Duhamel) form:

Definition 2.4.1 (Hopf-Cole Solution with Neumann Boundary Conditions)**.** *Let P_T denote the continuum Robin heat kernel constructed in Theorem 2.3.30, and let ξ denote a space-time white noise on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathcal{Z}_0$ denote some (random) Borel measure which is $\mathbb{P}$-almost surely supported on I. We say that a space-time process $\mathcal{Z} = (\mathcal{Z}(T,X))_{T>0,X\in I}$ is a mild solution to the SHE satisfying Robin boundary conditions if $\mathbb{P}$-almost surely, for every $T>0$ and $X \in I$ we have*

$$\mathcal{Z}(T,X) = \int_I P_T(X,Y)\mathcal{Z}_0(dY) + \int_0^T \int_I P_{T-S}(X,Y)\mathcal{Z}(S,Y)\xi(S,Y)dSdY$$

where the integral against the white noise is meant to be interpreted in the Itô-Walsh sense [150, 85, 52]. If $\xi = \partial_T W$ for a cylindrical Wiener Process W, then we may abbreviate the above expression as

$$\mathcal{Z}_T = P_T * \mathcal{Z}_0 + \int_0^T P_{T-S} * (\mathcal{Z}_S \cdot dW_S)$$

where the $*$ always denotes a spatial convolution.

The following proposition gives us conditions for the existence, uniqueness, and positivity of solutions, starting from initial data which is defined pointwise at each $X \in I$, and bounded in $L^2(\mathbb{P})$

under an exponential weight function.

Proposition 2.4.2 (Existence/Uniqueness/Positivity of Mild Solutions with L^2-bounded initial data)**.** *Let ξ be a space-time white noise. Suppose that we have some (random, function-valued) initial data $\mathcal{Z}_0$ which satisfies the following condition for some $a > 0$:*

$$\sup_{X \in I} e^{-aX} \mathbb{E}[\mathcal{Z}_0(X)^2] < \infty.$$

If $I = [0,1]$ we can just assume $a = 0$. Then there exists a mild solution to the SHE which is adapted to the natural filtration generated by the noise, $\mathcal{F}_T := \sigma(\{\xi(S,\cdot)\}_{S \le T})$. This mild solution is unique in the class of adapted processes $\mathcal{Z}$ satisfying

$$\sup_{\substack{X \in I \\ S \in [0,T]}} e^{-aX} \mathbb{E}[\mathcal{Z}_S(X)^2] < \infty.$$

Furthermore, if $\mathcal{Z}_0$ is a.s. positive, then $\mathcal{Z}(S,\cdot) > 0, \ \forall S$ a.s.

Proof. The argument here is adapted from [150]. The informal argument is as follows: We fix a terminal time $\tau > 0$, and we define the following sequence of iterates for $T \le \tau$ and $X \in I$:

$$\begin{aligned} u_0(T,X) &:= \int_I P_T(X,Y) \mathcal{Z}_0(Y) dY, \\ u_{n+1}(T,X) &:= \int_0^T \int_I P_{T-S}(X,Y) u_n(S,Y) \xi(S,Y) dY dS. \end{aligned}$$

The fact that the stochastic integral defining u_{n+1} actually exists will follow from Equations (2.17) and (2.18) below, but we will leave the formalism for later. If we let $z_N := \sum_0^N u_n$ then we have the relation that

$$z_{N+1}(T,X) = \int_I P_T(X,Y)\mathcal{Z}_0(Y) dY + \int_0^T \int_I P_{T-S}(X,Y) z_N(S,Y) \xi(S,Y) dY dS.$$

By letting $N \to \infty$, it follows that our mild solution should heuristically be given by $\mathcal{Z} =$

$\lim_{N\to\infty} z_N = \sum_0^\infty u_n$. Therefore, the aim is now to prove that the series $\sum_n u_n$ converges absolutely in the appropriate Banach Space.

Let us now formalize the argument. Consider the Banach space $\mathcal{B}$ consisting of $C(I)$-valued, adapted processes $u = (u(T,\cdot))_{T\in[0,\tau]}$ which satisfy the condition

$$\|u\|_{\mathcal{B}}^2 := \sup_{\substack{X\in I\\ T\in[0,\tau]}} e^{-aX}\mathbb{E}[u(T,X)^2] < \infty.$$

Let u_n be given as above. Define a sequence of functions $(f_n)_{n\geq 0}$ from $[0,\tau] \to \mathbb{R}_+$ by

$$f_n(T) := \sup_{\substack{X\in I\\ S\in[0,T]}} e^{-aX}\mathbb{E}[u_n(S,X)^2]$$

where the RHS is defined to be $+\infty$ if the stochastic integral defining u_n fails to exist (which will happen iff $\int_{[0,T]\times I} P_{T-S}(X,Y)^2\mathbb{E}[u_n(S,Y)^2]dYdS = +\infty$).

By Itô's isometry and the definition of f_n, we see that

$$\begin{aligned}\mathbb{E}[u_{n+1}(T,X)^2] &= \int_0^T \int_I P_{T-S}(X,Y)^2\mathbb{E}[u_n(S,Y)^2]dYdS \\ &\leq \int_0^T \left(\int_I P_{T-S}(X,Y)^2 \cdot e^{aY}dY\right) f_n(S)dS. \qquad (2.17)\end{aligned}$$

By Propositions 2.3.31 and 2.3.32 above, we easily obtain the bound:

$$\int_I P_T(X,Y)^2 e^{aY}dY \leq CT^{-1/2}e^{aX}, \qquad \forall X \in I,\ T \leq \tau. \qquad (2.18)$$

Here C is a constant depending only on A, B and the time horizon τ. Note that f_n is an increasing function, which implies that $T \mapsto \int_0^T (T-S)^{-1/2} f_n(S)dS$ is also increasing in T (which can be proved by substituting $S = TU$). Together with equations (2.17) and (2.18), this implies that

$$f_{n+1}(T) \leq C \int_0^T (T-S)^{-1/2} f_n(S) dS$$

which we can iterate twice to obtain that

$$f_{n+2}(T) \leq C' \int_0^T f_n(S) dS$$

By assumption, we have that $\sup_{X \in I} e^{-aX} \mathbb{E}[\mathcal{Z}_0(X)^2] < \infty$ which implies (for instance by (2.18)) that $f_0(T) < \infty$ for all $T \in [0, \tau]$. So by iterating this recursion, one obtains the result that $f_n(T) \leq CT^{n/2}/(n/2)!$, which implies that the stochastic integral defining u_n always exists, and moreover that $\sum_n \|u_n\|_{\mathcal{B}} < \infty$, as desired.

Uniqueness is proved in a similar way: If $\mathcal{Z}, \mathcal{Z}'$ are two different solutions then

$$\mathcal{Z}(T,X) - \mathcal{Z}'(T,X) = \int_0^T \int_I P_{T-S}(X,Y) \big[\mathcal{Z}(S,Y) - \mathcal{Z}'(S,Y)\big] \xi(S,Y) dY dS$$

Squaring both sides and taking expectations,

$$\mathbb{E}\Big[\big(\mathcal{Z}(T,X) - \mathcal{Z}'(T,X)\big)^2\Big] = \int_0^T \int_I P_{T-S}(X,Y)^2 \mathbb{E}\Big[\big(\mathcal{Z}(S,Y) - \mathcal{Z}'(S,Y)\big)^2\Big] dY dS$$

so now Gronwall's lemma (or direct iteration) shows that both sides must be zero.

The positivity result can be adapted from [122]. $\square$

Although the above proposition gives us existence and uniqueness results for a wide class of initial data, there are still some natural choices of initial data which are not covered. In section 6, we will especially need the case of δ_0 initial data, which is clearly not covered by the above proposition (since it is not function-valued).

Proposition 2.4.3 (Existence/Uniqueness/Positivity of Mild Solutions with δ_0-initial data)**.** *Fix a space-time white noise ξ. There exists a $C(I)$-valued process $(\mathcal{Z}_T)_{T\geq 0}$ which is adapted to $\mathcal{F}_T := \sigma(\{\xi(S,\cdot)\}_{S\leq T})$, and satisfies the conditions of Definition 2.4.1 with $\mathcal{Z}_0 = \delta_0$:*

$$\mathcal{Z}_T(X) = P_T(X,0) + \int_0^T \int_I P_{T-S}(X,Y)\mathcal{Z}_S(Y)\xi(S,Y)dYdS.$$

This mild solution is unique in the class of adapted processes $\mathcal{Z}$ satisfying

$$\sup_{\substack{X\in I\\ S\in[0,T]}} S\cdot \mathbb{E}[\mathcal{Z}_S(X)^2] < \infty.$$

Furthermore, $\mathcal{Z}_S > 0$, $\forall S$ a.s.

Proof. Similar to before, let us fix a terminal time τ and then define a Banach space $\mathcal{B}$ which consists of $C(I)$-valued, adapted processes $u = (u(T,\cdot))_{T\in[0,\tau]}$ satisfying

$$\|u\|_{\mathcal{B}}^2 := \sup_{\substack{X\in I\\ T\in[0,\tau]}} T\cdot \mathbb{E}[u(T,X)^2] < \infty.$$

As before, define a sequence of iterates $T\leq \tau$ and $X\in I$:

$$\begin{aligned} u_0(T,X) &:= P_T(X,0),\\ u_{n+1}(T,X) &:= \int_0^T \int_I P_{T-S}(X,Y)u_n(S,Y)\xi(S,Y)dYdS. \end{aligned}$$

As before, we just need to show that $\sum_n \|u_n\|_{\mathcal{B}} < \infty$.

Similar to the proof of Proposition 2.4.2, we define

$$f_n(T) := \sup_{\substack{X\in I\\ S\in[0,T]}} S^{1/2} P_S(X,0)^{-1}\mathbb{E}[u_n(S,X)^2]$$

where the RHS is defined as $+\infty$ if the stochastic integral defining u_n fails to exist.

Using the Itô isometry, the definition of f_n, and the fact that (by Proposition 2.3.31) $P_{T-S} \lesssim (T-S)^{-1/2}$, we compute

$$\begin{aligned}
\mathbb{E}[u_{n+1}(T,X)^2] &= \int_0^T \int_I P_{T-S}(X,Y)^2 \mathbb{E}[u_n(S,Y)^2] dY dS \\
&\leq \int_0^T \int_I P_{T-S}(X,Y)^2 \cdot S^{-1/2} P_S(Y,0) f_n(S) dY dS \\
&\leq C \int_0^T (T-S)^{-1/2} S^{-1/2} \bigg[\int_I P_{T-S}(X,Y) P_S(Y,0) dY \bigg] f_n(S) dS \\
&= C P_T(X,0) \int_0^T (T-S)^{-1/2} S^{-1/2} f_n(S) dS
\end{aligned}$$

where we used the semigroup property in the final line. Multiplying both sides by $T^{1/2} P_T(X,0)^{-1}$, we find that

$$T^{1/2} P_T(X,0)^{-1} \mathbb{E}[u_{n+1}(T,X)] \leq C T^{1/2} \int_0^T (T-S)^{-1/2} S^{-1/2} f_n(S) dS.$$

Just like before, f_n is an increasing function, therefore (by making a substitution $S = TU$) one may see that the RHS of the last expression is an increasing function of T. It follows that

$$f_{n+1}(T) \leq C T^{1/2} \int_0^T (T-S)^{-1/2} S^{-1/2} f_n(S) dS$$

which we can iterate twice to obtain

$$f_{n+2}(T) \leq C' T^{1/2} \int_0^T S^{-1/2} f_n(S) dS.$$

Using the fact that (by Prop. 2.3.31) $\sup_{T \in [0,\tau]} f_0(T) \leq C$, we can iterate this recursion to obtain

$$f_n(T) \lesssim T^{n/2} / (n/2)!$$

We have just proved that

$$\mathbb{E}[u_n(T,X)^2] \leq CP_T(X,0)T^{(n-1)/2}/(n/2)!$$

for a constant C not depending on $T \in [0,\tau]$, $X \in I$, or $n \in \mathbb{N}$. By Proposition 2.3.31 we know that $P_T(X,0) \leq CT^{-1/2}$, and therefore it follows that $\sum_n \|u_n\|_{\mathcal{B}} < \infty$, which proves existence of the mild solution.

Uniqueness and positivity are proved in the same manner as in Proposition 2.4.2. $\square$

Although the mild solution is one notion of what it means to solve the SHE, it is not the only natural definition of a solution. In particular, the notion of a weak solution (which uses the idea of pairing against test functions) is also very useful, and is actually equivalent to the notion of a mild solution:

Proposition 2.4.4 (Equivalence of Weak Solutions and Mild Solutions)**.** *Let W be a cylindrical Wiener process defined on some $(\Omega, \mathcal{F}, \mathbb{P})$. Denote by $\mathscr{T}$ the collection of all $\varphi \in \mathcal{S}(\mathbb{R})$ such that $\varphi'(0) = A\varphi(0)$ and also $\varphi'(1) = B\varphi(1)$ if $I = [0,1]$. Then a $C(I)$-valued process $(\mathcal{Z}_T)_{T\geq 0}$ is a mild solution to the SHE if and only if for every $\varphi \in \mathscr{T}$ we have that*

$$(\mathcal{Z}_T, \varphi) = (\mathcal{Z}_0, \varphi) + \frac{1}{2}\int_0^T (\mathcal{Z}_S, \varphi'')dS + \int_0^T (\mathcal{Z}_S\varphi, dW_S)$$

where $(\varphi, \psi) := \int_I \varphi\psi$ denotes the $L^2(I)$-pairing. If the latter relation holds, we call $(\mathcal{Z}_T)$ a weak solution of the SHE.

Proof. If $\mathcal{Z}_0$ is a mild solution as defined in 2.4.1, then

$$\mathcal{Z}_T = P_T * \mathcal{Z}_0 + \int_0^T P_{T-S} * (\mathcal{Z}_S dW_S).$$

Pairing both sides against a φ and using self-adjointness of $\frac{1}{2}\Delta$ with Robin boundary conditions,

$$(\mathcal{Z}_T, \varphi) = (P_T * \mathcal{Z}_0, \varphi) + \int_0^T \big(\mathcal{Z}_S \cdot (P_{T-S} * \varphi), dW_S\big). \tag{2.19}$$

Next, noting that $\partial_T (P_T * \mathcal{Z}_0, \varphi) = \frac{1}{2}(P_T * \mathcal{Z}_0, \varphi'')$ we have that

$$\begin{aligned}
(P_T * \mathcal{Z}_0, \varphi) &= (\mathcal{Z}_0, \varphi) + \frac{1}{2}\int_0^T (P_S * \mathcal{Z}_0, \varphi'') dS \\
&= (\mathcal{Z}_0, \varphi) + \frac{1}{2}\int_0^T \left(\mathcal{Z}_S - \int_0^S P_{S-U} * (\mathcal{Z}_U dW_U)\ \ ,\ \ \varphi''\right) dS \\
&= (\mathcal{Z}_0, \varphi) + \frac{1}{2}\int_0^T (\mathcal{Z}_S, \varphi'') dS - \frac{1}{2}\int_0^T \left(\mathcal{Z}_U \cdot \left[\int_U^T P_{S-U} * \varphi'' dS\right]\ \ , dW_U\right) \\
&= (\mathcal{Z}_0, \varphi) + \frac{1}{2}\int_0^T (\mathcal{Z}_S, \varphi'') dS - \int_0^T \left(\mathcal{Z}_U \cdot \Big[P_{T-U} * \varphi - \varphi\Big]\ \ , dW_U\right).
\end{aligned} \tag{2.20}$$

In the third line we distributed terms and switched the order of integration (cf. stochastic Fubini's theorem). In the last line we merely applied the fundamental theorem of calculus to the term within the square bracket. Now Equations (2.19) and (2.20) together imply that $(\mathcal{Z}_T)$ is a weak solution. Thus mild solutions are weak.

To prove the converse, note that if

$$(\mathcal{Z}_T, \varphi) = (\mathcal{Z}_0, \varphi) + \frac{1}{2}\int_0^T (\mathcal{Z}_S, \varphi'') dS + \int_0^T (\mathcal{Z}_S \varphi, dW_S).$$

Using Itô's formula for Hilbert-space valued processes, we see that if $\varphi \in C^\infty([0,\tau] \times \mathbb{R})$ such that $\varphi_T := \varphi(T, \cdot) \in \mathscr{T}$ for all T, then

$$d(\mathcal{Z}_T, \varphi_T) \ \ = \ \ (d\mathcal{Z}_T, \varphi_T) + (\mathcal{Z}_T, d\varphi_T) \ \ = \ \ \big(\mathcal{Z}_T, \frac{1}{2}\varphi_T'' + \partial_T \varphi_T\big) dT + (\mathcal{Z}_T \varphi_T, dW_T).$$

Then fixing a time $T > 0$, and setting $\varphi_S := P_{T-S} * \varphi$, we can integrate from 0 to T to obtain:

$$
\begin{aligned}
(\mathcal{Z}_T, \varphi) - (P_T * \mathcal{Z}_0, \varphi) &= (\mathcal{Z}_T, \varphi_T) - (\mathcal{Z}_0, \varphi_0) \\
&= \int_0^T \big(\mathcal{Z}_S, \frac{1}{2}\varphi_S'' + \partial_S \varphi_S\big) dS + \int_0^T (\mathcal{Z}_S \varphi_S, dW_S) \\
&= 0 + \int_0^T \big(\mathcal{Z}_S \cdot (P_{T-S} * \varphi), dW_S\big)
\end{aligned}
$$

which (after rearranging terms) is equivalent to

$$
(\mathcal{Z}_T, \varphi) = \left(P_T * \mathcal{Z}_0 + \int_0^T P_{T-S} * (\mathcal{Z}_S dW_S) \ \ , \ \varphi \right)
$$

so that $(\mathcal{Z}_T)$ is a mild solution. $\square$

2.5 Proof of Tightness and Identification of the Limit

Let us first establish a few topological conventions. Throughout this section, I will denote either the interval $[0, \infty)$ or $[0, 1]$. We will endow $C(I)$ with the topology of uniform convergence on compact sets if $I = [0, \infty)$, and the topology of uniform convergence when $I = [0, 1]$. The space $D([0, \tau], C(I))$ will denote the space of all right-continuous functions from $[0, \tau] \to C(I)$ which have left limits. We will endow $D([0, \tau], C(I))$ with the Skorokhod topology (see [17]), which may be metrized by

$$
\sigma(\Phi, \Psi) = \inf_{\lambda \in \Lambda} \max \left\{ \|\lambda - id\|_{L^\infty[0,\tau]} \ , \ \sup_{T \in [0,\tau]} d_{C(I)}\big(\Phi(T), \Psi(\lambda(T))\big) \right\}
$$

where Λ is the space of increasing homeomorphisms from $[0, \tau]$ to itself, and $d_{C(I)}$ is a metric inducing the topology on $C(I)$.

Definition 2.5.1 (Parabolic Scaling)**.** *Let $\Lambda = \mathbb{Z}_{\geq 0}$ for ASEP-H and $\Lambda = \{0, ..., N\}$ for ASEP-B. Let $Z_t(x)$ denote the Gärtner-transformed height functions from Section 2. For $\epsilon > 0$, $X \in \epsilon^{-1}\Lambda$,* *and $T \geq 0$, we define the parabolically scaled process*

$$\mathcal{Z}^\epsilon(T, X) = Z_{\epsilon^{-2}T}(\epsilon^{-1}X)$$

where $Z_t(x)$ is the Gärtner-transformed height function from Definition 2.7. We extend $\mathcal{Z}^\epsilon(T, \cdot)$ from $\epsilon\Lambda$ to the whole interval I by linear interpolation. Throughout this section, we will fix a terminal time τ, and consider (the law of) $\mathcal{Z}^\epsilon$ as a measure on the Skorokhod space $D([0, \tau], C(I))$.

We remark that $\mathcal{Z}^\epsilon$ depends on ϵ in two completely different ways, first due to the space-time scaling in Definition 2.5.1, but secondly also because the non-scaled discrete-space process Z_t depends on ϵ via the ϵ-scaled parameters $p, q, \alpha, \beta, \gamma, \delta$. This is what we meant by a weak scaling: we scale the model parameters along with the height functions.

Assumption 2.5.2 (Near-Equilibrium Initial Conditions)**.** *For ASEP-H, we will always assume that the initial data $\mathcal{Z}^\epsilon$ satisfies the following bounds: There exists $a \geq 0$ such that for each terminal time $\tau > 0$, each $p \geq 1$, and each $\alpha \in [0, \frac{1}{2})$ there exists some constant $C = C(p, \tau, \alpha)$ such that for all $T \in [0, \tau]$, $X, Y \in I$, and $\epsilon > 0$ small enough:*

$$\|\mathcal{Z}_0^\epsilon(X)\|_p \leq Ce^{aX},$$
$$\|\mathcal{Z}_0^\epsilon(X) - \mathcal{Z}_0^\epsilon(Y)\|_p \leq C|X - Y|^\alpha e^{a(X+Y)}.$$

Here $\|Z\|_p := \mathbb{E}[|Z|^p]^{1/p}$ denotes the L^p-norm with respect to the probability measure. For ASEP-B, we will assume the same bounds but with $a = 0$.

As a consequence of Kolmogorov's continuity criterion, Assumption 2.5.2 ensures that with large probability, the random functions $\{\mathcal{Z}_0^\epsilon\}_{\epsilon>0}$ are pointwise bounded and equi-Hölder $1/2-$ (locally). In turn, Arzela-Ascoli's theorem and Prohorov's theorem ensure that as $\epsilon \to 0$ the $\mathcal{Z}_0^\epsilon$ converge weakly (along a subsequence) to some random function which has the same regularity as Brownian motion. Hence, Assumption 2.5.2 is merely a technical restriction which ensures that the initial data have limit points which are not too wildly behaved on small scales, and whose moments grow exponentially in X (at worst). For instance a product of Bernoulli's will satisfy the assumptions in 3.2, as the associated height function converges to a Brownian motion. On the other hand, narrow-wedge (zero-particle) initial data will not satisfy Assumption 2.5.2. We will find a way to deal with such data in Section 6 below.

We are almost ready to move onto the main results, however we will need a technical result which will be a very useful black box for obtaining L^p estimates. This result was originally obtained in Lemma 3.1 of [57] and was further elaborated in Lemma 4.18 of [45], so we will not give the proof.

Lemma 2.5.3. *Let $\|X\|_p = \mathbb{E}[|X|^p]^{1/p}$ denote the L^p norm with respect to the probability measure, and let M denote the martingale appearing in Proposition 2.2.7. Let $F : [0,\infty) \times \mathbb{Z}_{\geq 0} \to \mathbb{R}$ be any bounded function. There exists a constant C (not depending on the function F) such that that for any $t > 1$ we have that*

$$\bigg\| \int_0^t \sum_{y \in \Lambda} F(s,y) dM_s(y) \bigg\|_p^2 \leq C\epsilon \int_0^t \sum_y \overline{F}(s,y)^2 \|Z_s(y)\|_p^2 ds$$

where the bar denotes a local supremum:

$$\overline{F}(s,y) := \sup_{|u-s| \leq 1} F(u,y).$$

In practice, we will almost always use the above lemma when $F(s,y) = \mathbf{p}_s^R(x,y)$ for some $x \in \Lambda$. In this case, by proposition 2.3.29 or 2.3.13, it always holds true that

$$\sup_{|u-s| \leq 1} \mathbf{p}_u^R(x,y) \leq e^1 \mathbf{p}_{s+1}^R(x,y)$$

and therefore by Proposition 2.3.1 or 2.3.15 (with $b = 0$) it then follows that

$$\overline{\mathbf{p}}_s^R(x,y)^2 \leq e^2 \mathbf{p}_{s+1}^R(x,y)^2 \leq C s^{-1/2} \mathbf{p}_{s+1}^R(x,y). \tag{2.21}$$

This fact will be used repeatedly during the technical estimates below.

We now move onto the main results.

Proposition 2.5.4 (Tightness)**.** *Fix a terminal time $\tau \geq 0$, and assume that the sequence $\{\mathcal{Z}_0^\epsilon\}_{\epsilon>0}$ of initial data satisfies Assumption 2.5.2. For all $p \geq 1$ and $\alpha \in [0, 1/2)$, there exists a constant $C = C(\alpha, p, \tau)$ such that*

$$\|\mathcal{Z}^\epsilon(T, X)\|_p \leq Ce^{aX} \tag{2.22}$$

$$\|\mathcal{Z}^\epsilon(T, Y) - \mathcal{Z}^\epsilon(T, X)\|_p \leq C|X - Y|^\alpha e^{a(X+Y)} \tag{2.23}$$

$$\|\mathcal{Z}^\epsilon(T, X) - \mathcal{Z}^\epsilon(S, X)\|_p \leq Ce^{2aX}(|T - S|^{\alpha/2} \vee \epsilon^\alpha) \tag{2.24}$$

uniformly over all $X, Y \in [0, \infty)$ and $S, T \in [0, \tau]$. For ASEP-B, the same estimates are satisfied for $X, Y \in [0, 1]$ and $a = 0$. Consequently, the laws of the $\mathcal{Z}^\epsilon$ are tight in the Skorokhod space $D([0, \tau], C(I))$, and moreover the limit point lies in $C([0, \tau], C(I))$.

Proof. This was originally proved as Proposition 4.15 of [45]. However, after a thorough discussion with the authors, we actually found a small mistake in equation (4.57) of that paper where $\mathbf{p}_{t-s}^R(x, y)$ should be replaced by $\mathbf{p}_{t-s+1}^R(x, y)$, and this messes up the iteration. This is not so obviously fixed; therefore we will give a new variant of the proof here, which is loosely based on the argument in Proposition 3.2 of [57].

For the first estimate, note by Proposition 2.2.7 that Z_t will satisfy the Duhamel-form equation

$$Z_t(x) = \mathbf{p}_t^R * Z_0(x) + \int_0^t \sum_{y \in \Lambda} \mathbf{p}_{t-s}^R(x, y) dM_s(y).$$

Using this identity and the fact that $(x+y)^2 \leq 2x^2 + 2y^2$, we find that

$$\|Z_t(x)\|_p^2 \leq 2\big\|\mathbf{p}_t^R * Z_0(x)\big\|_p^2 + 2\bigg\| \int_0^t \sum_{y\in\Lambda} \mathbf{p}_{t-s}^R(x,y) dM_s(y) \bigg\|_p^2. \tag{2.25}$$

The first term on the RHS of (2.25) can be bounded as follows:

$$\bigg\| \sum_{y\in\Lambda} \mathbf{p}_t^R(x,y) Z_0(y) \bigg\|_p \leq \sum_y \mathbf{p}_t^R(x,y) \|Z_0(y)\|_p \leq \sum_y \mathbf{p}_t^R(x,y) e^{a\epsilon y} \leq C e^{a\epsilon x} \tag{2.26}$$

where we applied Assumption 2.5.2 in the second inequality and Corollary 2.3.3 in the next one. For $t \geq 1$, the second term on the RHS of (2.25) can be bounded using Lemma 2.5.3 with (2.21):

$$\bigg\| \int_0^t \sum_{y\in\Lambda} \mathbf{p}_{t-s}^R(x,y) dM_s(y) \bigg\|_p^2 \leq C\epsilon \int_0^t (t-s)^{-1/2} \sum_y \mathbf{p}_{t-s+1}^R(x,y) \|Z_s(y)\|_p^2 ds. \tag{2.27}$$

Let us define the following quantity:

$$[Z_t]_p := \sup_{x\in\Lambda} e^{-a\epsilon x} \|Z_t(x)\|_p.$$

Multiplying both sides of (2.25) by $e^{-2a\epsilon x}$ and then using (2.26) and (2.27), we see that

$$e^{-2a\epsilon x} \|Z_t(x)\|_p^2 \leq C + C\epsilon e^{-2a\epsilon x} \int_0^t (t-s)^{-1/2} \bigg(\sum_y \mathbf{p}_{t-s+1}^R(x,y) e^{2a\epsilon y} \bigg) [Z_s]_p^2 \, ds.$$

Now, by Corollary 2.3.3 we know that

$$\sup_{x\geq 0} \; e^{-2a\epsilon x} \sum_y \mathbf{p}_{t-s+1}^R(x,y) e^{2a\epsilon y} < \infty$$

so by taking the sup over x on both sides of the previous expression, we find that

$$[Z_t]_p^2 \leq C + C\epsilon \int_0^t (t-s)^{-1/2} [Z_s]_p^2 ds. \tag{2.28}$$

Now we would like to iterate the inequality in (2.28), however the problem is that we have only proved this bound for $t \geq 1$ (see the statement of Lemma 2.5.3), but we also need to prove it for $t \in [0,1]$ in order to apply the iteration. For $t \leq 1$, we have (by the definition of Z_t) that $Z_t(x) \leq e^{2\sqrt{\epsilon}N(t)}Z_0(x)$, where $N(t)$ denotes the net number of jumps which occur at site x up to time t. Since $t \leq 1$, each $N(t)$ is stochastically dominated by a Poisson random variable of rate 1. Then by Assumption 2.5.2, $\|Z_t(x)\|_p \leq \mathbb{E}[e^{pN(t)}]^{1/p} \cdot \|Z_0\|_p \leq Ce^{a\epsilon x}$, and thus $\sup_{t\leq 1}[Z_t]_p < \infty$, which proves that (2.28) holds for $t \leq 1$ as well.

Iterating (2.28) twice, we find that

$$[Z_t]_p^2 \leq C + C^2 \epsilon t^{1/2} + \alpha C^2 \epsilon^2 \int_0^t [Z_s]_p^2 ds \tag{2.29}$$

where $\alpha = \int_u^t (s-u)^{-1/2}(t-s)^{-1/2}ds$, which is a constant not depending on u or t, as can be verified by making the substitution $v = (t-s)/(s-u)$. Since $\epsilon t^{1/2} \leq T^{1/2}$, the second term on the RHS of (2.29) may be absorbed into the first one, only changing the constant C. Then an easy application of Gronwall's lemma shows that

$$[Z_t]_p \leq [Z_0]_p e^{\alpha C^2 \epsilon^2 t} \leq C$$

where we used the fact that $\epsilon^2 t \leq T$ which is a constant not depending on ϵ or t or x. We have shown that

$$\|Z_t(x)\|_p \leq Ce^{a\epsilon x}$$

which (after changing to macroscopic variables $X = \epsilon x$ and $T = \epsilon^2 t$) proves the first bound (2.22).

Let us now move onto the second bound (2.23). Note that

$$Z_t(x) - Z_t(y) = \sum_z \big(\mathbf{p}_t^R(x,z) - \mathbf{p}_t^R(y,z)\big) Z_0(z) + \int_0^t \sum_z \big(\mathbf{p}_t^R(x,z) - \mathbf{p}_t^R(y,z)\big) dM_s(z).$$

Let us name the terms on the right side as I_1, I_2 respectively.

In order to bound I_1, let us start by extending Z_0 to a function $\tilde{Z}_0$, defined on all $\mathbb{Z}$, such that $\tilde{Z}_0(z-1) - \mu_A \tilde{Z}_0(z)$ is an odd function. In the bounded interval case, we also require that $\tilde{Z}_0(N+1+z) - \mu_B \tilde{Z}_0(N+z)$ is an odd function. This extended $\tilde{Z}_0$ can be constructed inductively: first let $\tilde{Z}_0(-1) = \mu_A Z(0)$, then define $\tilde{Z}_0(-2) = \mu_A \tilde{Z}_0(-1) - \big(Z_0(0) - \mu_A Z_0(1)\big)$, and so on. In the bounded interval case, one would first construct $\tilde{Z}_0$ on $\{-N, ..., -1\}$ and $\{N+1, ..., 2N\}$, then on $\{-2N, ..., -N-1\}$ and $\{2N+1, ..., 3N\}$, etc.

By Assumption 2.5.2, we know that $\|Z_0(z) - Z_0(w)\|_p \leq Ce^{a\epsilon(x+y)}(\epsilon|z-w|)^\alpha$ for $\alpha < 1/2$. In the half-line case, one may check (using the defining property of $\tilde{Z}_0$) that $\tilde{Z}_0$ will satisfy the same property with the same constant a. In the bounded interval case, $\tilde{Z}_0$ will satisfy this property on $\mathbb{Z}$, for *some* large enough $a > 0$ which does not depend on ϵ (see Lemma 2.3.14). By construction, it is true that

$$\sum_{z \in \Lambda} \mathbf{p}_t^R(x,z) Z_0(z) = \sum_{z \in \mathbb{Z}} p_t(x-z) \tilde{Z}_0(z)$$

where the p_t on the RHS is the standard (whole-line) heat kernel on $\mathbb{Z}$. Consequently,

$$\begin{aligned}
\|I_1\|_p &\leq \sum_{z \in \mathbb{Z}} p_t(x-z) \big\| \tilde{Z}_0(z) - \tilde{Z}_0(z+y-x) \big\|_p \\
&\leq C \sum_{z \in \mathbb{Z}} p_t(x-z) (\epsilon|x-y|)^\alpha e^{a\epsilon(2z+y-x)} \\
&\leq C(\epsilon|x-y|)^\alpha e^{a\epsilon(x+y)}.
\end{aligned}$$

In the final line, we used the fact that $\sum_z p_t(x-z)e^{2a\epsilon z} = e^{2a\epsilon x} e^{(\cosh(2a\epsilon)-1)t} \sim e^{2a\epsilon x} e^{2a^2\epsilon^2 t} \leq Ce^{2a\epsilon x}$, which is true because $e^{2a\epsilon z}$ is an eigenfunction of $\frac{1}{2}\Delta$ on $\mathbb{Z}$ with eigenvalue $\cosh(2a\epsilon) - 1$. This proves that I_1 satisfies the desired bound.

Next we need to bound I_2, so we are going to apply Lemma 2.5.3 with $F(s,z) = \mathbf{p}^R_{t-s}(x,z) - \mathbf{p}^R_{t-s}(y,z)$. Note that if $|t-s-u| \leq 1$, then by the triangle inequality, Proposition 2.3.2 (or 2.3.15), and Proposition 2.3.13 (or 2.3.29), we have

$$\begin{aligned}(\mathbf{p}^R_u(x,z) - \mathbf{p}^R_u(y,z))^2 &\leq |\mathbf{p}^R_u(x,z) - \mathbf{p}^R_u(y,z)|(\mathbf{p}^R_u(x,z) + \mathbf{p}^R_u(y,z)) \\ &\leq C(1 \wedge u^{-(1+\alpha)/2})|x-y|^\alpha \cdot e^1(\mathbf{p}^R_{t-s+1}(x,z) + \mathbf{p}^R_{t-s+1}(y,z)) \\ &\leq C(t-s)^{-(1+\alpha)/2}|x-y|^\alpha \cdot (\mathbf{p}^R_{t-s+1}(x,z) + \mathbf{p}^R_{t-s+1}(y,z)).\end{aligned}$$

Hence $\sup_{|s-u|\leq 1} F_u(s,z)^2$ is bounded by the last expression. Thus by Lemma 2.5.3 and Equation (2.22) above, we find that for $t \geq 1$ and $\alpha \in [0,1)$ we have:

$$\begin{aligned}\|I_2\|_p^2 &\leq C\epsilon|x-y|^\alpha \int_0^t (t-s)^{-(1+\alpha)/2} \sum_z \big(\mathbf{p}^R_{t-s+1}(x,z) + \mathbf{p}^R_{t-s+1}(y,z)\big)\|Z_s(z)\|_p^2 ds \\ &\leq C\epsilon|x-y|^\alpha \int_0^t (t-s)^{-(1+\alpha)/2} \sum_z \big(\mathbf{p}^R_{t-s+1}(x,z) + \mathbf{p}^R_{t-s+1}(y,z)\big)e^{2a\epsilon z} ds \\ &= C\epsilon|x-y|^\alpha t^{(1-\alpha)/2}\big(e^{2a\epsilon x} + e^{2a\epsilon y}\big) \\ &\leq C(\epsilon|x-y|)^\alpha e^{2a\epsilon(x+y)}.\end{aligned}$$

Let us justify each of the inequalities above. In the second line we used (2.22) to bound $\|Z_s(z)\|_p^2$ by $Ce^{2a\epsilon z}$. In the third line, we applied Corollaries 2.3.3 and 2.3.17 in order to bound the sum over z, and then we used the fact that $\int_0^t (t-s)^{-(1+\alpha)/2} ds = Ct^{(1-\alpha)/2}$. In the final line, we used the fact that $t^{(1-\alpha)/2} \leq C\epsilon^{\alpha-1}$ since $t \leq \epsilon^{-2}T$, and we also noted that $e^{2a\epsilon x} + e^{2a\epsilon y} \leq 2e^{2a\epsilon(x+y)}$. Taking square roots in the above expression, we find that for $\alpha < 1$,

$$\|I_2\|_p \leq C(\epsilon|x-y|)^{\alpha/2} e^{a\epsilon(x+y)}$$

which (after changing to macroscopic variables) completes the proof of the second bound (2.44).

Let us move onto the third estimate. Using Lemma 2.2.7 and the semigroup property of the heat kernel, we can write

$$Z_t(x) = \sum_{y\geq 0} \mathbf{p}^R_{t-s}(x,y) Z_s(y) + \int_s^t \sum_{y\geq 0} \mathbf{p}^R_{t-u}(x,y) dM_u(y)$$

where $\mathbf{p}^R_t$ is the Robin heat kernel on $\mathbb{Z}_{\geq 0}$ with Robin boundary conditions. Therefore

$$\begin{aligned}
Z_t(x) - Z_s(x) = &\sum_{y\geq 0} \mathbf{p}^R_{t-s}(x,y)\big(Z_s(y) - Z_s(x)\big) \\
&+ \Bigg[\sum_{y\geq 0} \mathbf{p}^R_{t-s}(x,y) \; - \; 1\Bigg] Z_s(x) + \int_s^t \sum_{y\geq 0} \mathbf{p}^R_{t-u}(x,y) dM_u(y).
\end{aligned}$$

Naming the terms on the RHS J_1, J_2, J_3 (in that order), we will prove the desired bound for each of the terms J_i.

For J_1 we use the spatial L^p bound (2.23) at time $S = \epsilon s$, and the fact that $r^\alpha \leq e^r$ (since $\alpha < 1$) to obtain

$$\begin{aligned}
\|J_1\|_p &\leq \sum_{y\geq 0} \mathbf{p}^R_{t-s}(x,y) \|Z_s(y) - Z_s(x)\|_p \\
&\leq C \sum_{y\geq 0} \mathbf{p}^R_{t-s}(x,y) \cdot |\epsilon x - \epsilon y|^\alpha e^{a(\epsilon x + \epsilon y)} \\
&\leq C \sum_{y\geq 0} \mathbf{p}^R_{t-s}(x,y) \cdot \epsilon^\alpha [1 \vee (t-s)^{\alpha/2}] e^{|x-y|[1\wedge (t-s)^{-1/2}]} e^{a(\epsilon x + \epsilon y)} \\
&\leq C\epsilon^\alpha [1 \vee (t-s)^{\alpha/2}] e^{2a\epsilon x}. \qquad (2.30)
\end{aligned}$$

where we used Corollary 2.3.3 in the last inequality. Now bounding J_2, we can simply use Proposition 2.3.7 together with (2.22) to note that

$$\begin{aligned}
\|J_2\|_p &\leq \Bigg|\sum_{y\geq 0} \mathbf{p}^R_{t-s}(x,y) \; - \; 1\Bigg| \cdot \|Z_s(x)\|_p \\
&\leq C\epsilon^\alpha (t-s)^{\alpha/2} \cdot e^{a\epsilon x}
\end{aligned}$$

where C does not depend on x. As for J_3, we use Lemma 2.5.3 and (2.21) to obtain for $t-s \geq 1$,

$$\left\| \int_s^t \sum_{y\geq 0} \mathbf{p}^R_{t-u}(x,y) dM_u(y) \right\|_p^2 \leq C\epsilon \int_s^t (t-u)^{-1/2} \sum_{y\geq 0} \mathbf{p}^R_{t-u+1}(x,y) \|Z_u(y)\|_p^2 du.$$

Using (2.22) and then Corollary 2.3.3 we find that

$$\sum_{y\geq 0} \mathbf{p}^R_{t-u+1}(x,y) \|Z_u(y)\|_p^2 \leq C \sum_{y\geq 0} \mathbf{p}^R_{t-u+1}(x,y) \cdot e^{2a\epsilon y} \leq C e^{2a\epsilon x}.$$

Consequently

$$\begin{aligned} C\epsilon \int_s^t (t-u)^{-1/2} \sum_{y\geq 0} \mathbf{p}^R_{t-u+1}(x,y) \|Z_u(y)\|_p^2 du &\leq C\epsilon e^{2a\epsilon x} \int_s^t (t-u)^{-1/2} du \\ &= C\epsilon (t-s)^{1/2} e^{2a\epsilon x} \\ &\leq C\epsilon^{2\alpha} (t-s)^{\alpha} e^{2a\epsilon x}. \end{aligned}$$

In the last line we used the fact that $2\alpha < 1$ and that $t-s \leq \epsilon^{-2}\tau$, so that $\epsilon = \epsilon^{2\alpha}\epsilon^{1-2\alpha} \leq C\epsilon^{2\alpha}(t-s)^{\alpha-1/2}$.

To show tightness on $D([0,\tau], C(I))$ for $0 < \delta \leq \tau$, these three estimates imply (respectively) that the $\{\mathcal{Z}^\epsilon\}_{\epsilon\in(0,1]}$ are uniformly bounded, uniformly spatially Hölder, and uniformly temporally Hölder (**except** for jumps of order ϵ^α) with large probability. Now we apply the version of Arzela-Ascoli for Skorokhod spaces (together with Prohorov's theorem) to obtain tightness, see [Bil97, Chapter 3] for the precise formulation of compactness in D.

The fact that any limit point lies in $C([\delta,\tau], C(I))$ is a straightforward consequence of Kolmogorov's continuity criterion. □

Definition 2.5.5 (Martingale Problem for the SHE)**.** *Fix a terminal time $\tau \geq 0$. Let $\mathbb{P}$ be a probability measure on $\Omega := C([0,\tau], C(I))$, and for $T \in [0,\tau]$, denote by $\mathcal{L}_T : \Omega \to C(I)$ the evaluation* *map at time T. Define $\mathscr{T}$ to be the set of all test functions $\varphi \in C_c^\infty(\mathbb{R})$ such that $\varphi'(0) = A\varphi(0)$, and also $\varphi'(1) = B\varphi(1)$ if $I = [0,1]$. We say that $\mathbb{P}$ solves the martingale problem for the SHE with Robin boundary parameters A and B on I if for every $\varphi \in \mathscr{T}$, the processes*

$$Y_T(\varphi) := (\mathcal{L}_T, \varphi) - (\mathcal{L}_0, \varphi) - \int_0^T (\mathcal{L}_S, \varphi'')dS,$$
$$Q_T(\varphi) := Y_T(\varphi)^2 - \int_0^T \|\mathcal{L}_S\varphi\|_{L^2(I)}^2 dS$$

are $\mathbb{P}$-local martingales (with respect to the filtration $\mathcal{F}_T := \sigma(\{\mathcal{L}_S : S \leq T\})$). Here $(\psi, \varphi) := \int_I \psi\varphi$ denotes the $L^2(I)$ pairing.

Here is the motivation behind the preceding definition. If $\mathcal{L}$ is the solution of the SHE, then formally, we have that

$$\mathcal{L}_T - \mathcal{L}_0 - \frac{1}{2}\int_0^T \Delta\mathcal{L}_S dS = \int_0^T \mathcal{L}_S dW_S$$

for some cylindrical Wiener Process $(W_S)_{S\geq 0}$ over $L^2(I)$. Consequently, if we integrate both sides in the above expression against any test function $\varphi \in \mathscr{T}$, and use self-adjointness of the operator Δ with Robin boundary conditions on I, we should have that

$$(\mathcal{L}_T, \varphi) - (\mathcal{L}_0, \varphi) - \int_0^T (\mathcal{L}_S, \varphi'')dS = \int_0^T (\mathcal{L}_S\varphi, dW_S) \tag{2.31}$$

One may verify directly from the definition of mild solutions that this is indeed true. The RHS in this expression is clearly a martingale for any φ (being an Itô integral against a cylindrical Wiener process). Furthermore by the Itô isometry, the quadratic variation of the martingale appearing on the RHS of (2.31) is given by

$$\int_0^T \|\mathcal{L}_S\varphi\|_{L^2(I)}^2 dS$$

which shows that the expressions $Y_T(\varphi)$ and $Q_T(\varphi)$ appearing in Definition 2.5.5 are indeed martingales when $\mathcal{Z}$ is a solution to the SHE.

Conversely, if $\mathcal{L}$ is any $C(I)$-valued process (not necessarily the solution to the SHE) such that $Y_T(\varphi)$ and $Q_T(\varphi)$ are always local martingales, one might hope that (by some version of Levy's characterization theorem) it should be possible to construct a cylindrical Wiener Process W such that (2.31) holds true for every φ, and thus $\mathcal{L}$ solves the SHE in some weak (PDE) sense. This is indeed true, as proved in Proposition 5.9 of [45]. We will reiterate the proof here because it will be needed later.

Proposition 2.5.6 (Uniqueness of Solution to Martingale Problem)**.** *Fix a terminal time τ, and let Ω and $\mathcal{L}_T$ be as in Definition 2.5.5. Suppose that $\mathbb{P}$ is a solution to the martingale problem for the SHE. Then we may enlarge the probability space $\big(\Omega,\ \mathcal{B}orel,\ \mathbb{P}\big)$ so as to admit a space-time white noise W such that $\mathbb{P}$-almost surely, for any $T \in [0,\tau]$ we have that*

$$\mathcal{L}_T(X) = P_T * \mathcal{L}_0(X) + \int_0^T P_{T-S} * (\mathcal{L}_S dW_S)$$

Proof. Let us first illustrate this same principle in a one-dimensional setting. Let $b, \sigma : \mathbb{R} \to \mathbb{R}$, and suppose that L_t is a process for which

$$Y_t := L_t - \int_0^t b(L_s)ds,$$

$$Q_t := Y_t^2 - \int_0^t \sigma(L_s)^2 ds$$

are both local martingales. Now suppose that we want to construct a Brownian motion W_t such that $dL_t = b(L_t)dt + \sigma(L_t)dW_t$. The solution is then to define

$$W_t := \int_0^t \sigma^{-1}(L_s) 1_{\{\sigma(L_s)\neq 0\}} dY_s + \int_0^t 1_{\{\sigma(L_s)=0\}} d\overline{W}_s \tag{2.32}$$

for an independent BM $\overline{W}$. It is then easily checked that $\langle W\rangle_t = t$ so that W is a Brownian motion, and W clearly satisfies the desired relation.

The same general idea will apply in the infinite-dimensional setting, where $b(z) = \frac{1}{2}\Delta z$ and $\sigma(z) = z$. However, the formalisms needed to define integrals against the (now infinite-dimensional) process Y_s become much more subtle, therefore we will use the notion of martingale measures as developed by Walsh, see for instance [150, Chapter 2].

Returning to the original statement, we define $Y_T(\varphi)$ and $Q_T(\varphi)$ to be the local martingales appearing in Definition 2.5.5. By localization, we are just going to assume that Y and Q are themselves martingales, with the understanding that we should technically choose a sequence of stopping times $\tau_n \to \infty$ and apply the usual tricks used in localization.

By the assumption that $Q_T(\varphi)$ is a martingale, we have that $\langle Y(\varphi)\rangle_T = Y_T(\varphi)^2 - Q_T(\varphi) = \int_0^T \|\mathcal{L}_S\varphi\|_{L^2(I)}^2 dS$ and thus by polarization we see that

$$\langle Y(\varphi), Y(\psi)\rangle_T = \int_0^T (\mathcal{L}_S\varphi, \mathcal{L}_S\psi)dS$$

so if φ and ψ have disjoint supports, then $\langle Y(\varphi), Y(\psi)\rangle = 0$. Thus Y is actually an *orthogonal* martingale measure [150, page 288], so we can define stochastic integrals of predictable space-time processes against it. In particular, we define

$$W_T(\varphi) := \int_0^T \int_I \mathcal{L}_S(X)^{-1}\varphi(X)1_{\{L_S(X)\neq 0\}}Y(dX\ dS) + \int_0^T \int_I 1_{\{\mathcal{L}_S(X)=0\}}\varphi(X)\overline{W}(dX\ dS)$$

where $\overline{W}$ is an independent space-time white noise over $[0,\tau) \times I$. Note the analogy between this formula and equation (2.32) in the one-dimensional setting. Then one may check that

$$\langle W(\varphi), W(\psi)\rangle_T = (\varphi, \psi)T$$

so by Levy's characterization theorem, it follows that W is a space-time white noise over $[0,\tau] \times I$, and moreover, it is true by construction that for every $\varphi \in \mathscr{T}$,

$$Y_T(\varphi) = \int_0^T (\mathcal{L}_S \varphi, dW_S)$$

which implies that

$$\begin{aligned}(\mathcal{L}_T, \varphi) &= (\mathcal{L}_0, \varphi) + \frac{1}{2}\int_0^T (\mathcal{L}_S, \varphi'')dS + Y_T(\varphi) \\ &= (\mathcal{L}_0, \varphi) + \frac{1}{2}\int_0^T (\mathcal{L}_S, \varphi'')dS + \int_0^T (\mathcal{L}_S \varphi, dW_S).\end{aligned}$$

Thus $(\mathcal{L}_T)$ is a weak solution, which (by Proposition 2.4.4) coincides with the mild solution. □

Theorem 2.5.7 (Main Theorem: Identification of Limit)**.** *Assume that as $\epsilon \to 0$, the sequence of initial data $\mathcal{Z}_0^\epsilon$ converges weakly in $C(I)$ to some initial data $\mathcal{Z}_0$. Then, as $\epsilon \to 0$, the laws of the $\mathcal{Z}^\epsilon$ converge weakly in $D([0,T], C(I))$ to the law of the mild solution of the SHE with Robin boundary conditions, and initial data $\mathcal{Z}_0$.*

Proof. The previous proposition shows that the laws of the $\mathcal{Z}^\epsilon$ do indeed converge (subsequentially) to some limiting measure as $\epsilon \to 0$. When $A, B \geq 0$ the identification of the limit is shown in [45] via Lemma 5.8, Lemma 5.7, Proposition 5.6, and Proposition 5.9 (in that logical order). When $A < 0$ or $B < 0$ the proof is very similar, but a small modification is needed in Lemma 5.8 (because it relies on a previous result, Proposition 5.1, which is only true for $A, B \geq 0$). Rather than trying to pinpoint the numerous little modifications which are needed, we will reproduce the whole proof here in generality, for the sake of completeness and clarity.

By Proposition 2.5.6, it suffices to show that if $\mathbb{P}$ is a probability measure on $C([0,\tau], C(I))$ which is a limit point of the laws of the $\mathcal{Z}^\epsilon$, then the processes

$$Y_T(\varphi) := (\mathcal{L}_T, \varphi) - \frac{1}{2}\int_0^T (\mathcal{L}_S, \varphi'')dS$$

$$Q_T(\varphi) := Y_T(\varphi)^2 - \int_0^T (\mathcal{L}_T^2, \varphi^2) dS$$

are both $\mathbb{P}$-local martingales, where $\mathcal{L}_T : C([0,\tau], C(I)) \to C(I)$ denotes the canonical projection (e.g., evaluation at T). The main idea is to note that for any (fixed) $\epsilon > 0$, if we define

$$(\varphi, \psi)_\epsilon := \epsilon \sum_{x \in \Lambda} \varphi(\epsilon x) \psi(\epsilon x)$$

then by Theorem 2.2.7, the process

$$Z_t(x) - \frac{1}{2} \int_0^t \Delta Z_t(x) dt$$

is a martingale, so after summing against a test function and changing to macroscopic variables, the process

$$Y_T^\epsilon(\varphi) := (\mathcal{Z}_T^\epsilon, \varphi)_\epsilon - \frac{1}{2} \int_0^T (\Delta_\epsilon \mathcal{Z}_S^\epsilon, \varphi)_\epsilon dS$$

is also martingale, where

$$\Delta_\epsilon \varphi(X) := \epsilon^{-2}(\varphi(X+\epsilon) + \varphi(X-\epsilon) - 2\varphi(X)).$$

Recall that $\varphi \in \mathscr{T}$, which means that $\varphi'(0) = A\varphi(0)$. Using summation-by-parts and a Taylor series expansion of φ near $X = 0$, we have:

$$\begin{aligned}
(\Delta_\epsilon \mathcal{Z}_S^\epsilon, \varphi)_\epsilon - (\mathcal{Z}_S^\epsilon, \Delta_\epsilon \varphi)_\epsilon &= \epsilon^{-2} \cdot \epsilon \Big[\mathcal{Z}_S^\epsilon(-\epsilon)\big(\varphi(0) - \varphi(-\epsilon)\big) - \varphi(-\epsilon)\big(\mathcal{Z}_S^\epsilon(0) - \mathcal{Z}_S^\epsilon(-\epsilon)\big) \Big] \\
&= \epsilon^{-1} \Big[\mathcal{Z}_S^\epsilon(-\epsilon) \cdot \big(\epsilon A \varphi(0) + O(\epsilon^2)\big) - \varphi(-\epsilon) \cdot \epsilon A \mathcal{Z}_S^\epsilon(0) \Big] \\
&= \epsilon^{-1} \Big[(1 - \epsilon A) \mathcal{Z}_S^\epsilon(0) \cdot \big(\epsilon A \varphi(0) + O(\epsilon^2)\big) - \big(\varphi(0) - \epsilon A \varphi(0) + O(\epsilon^2)\big) \cdot \epsilon A \mathcal{Z}_S^\epsilon(0) \Big] \\
&= \epsilon^{-1} \Big[\mathcal{Z}_S^\epsilon(0) \cdot O(\epsilon^2) \Big]
\end{aligned}$$

where the $O(\epsilon^2)$ term is a non-random quantity depending only on the test function φ. This computation was specific to the half-line case, but the bounded-interval case is similar (except that the number of boundary terms doubles since there are two boundary points instead of just one) and we get the same bound (except that there will be a $\mathcal{Z}_S^\epsilon(1)$ term appearing next to the $O(\epsilon^2)$ term as well).

Summarizing, we have shown that

$$\big|(\Delta_\epsilon \mathcal{Z}_S^\epsilon, \varphi)_\epsilon - (\mathcal{Z}_S^\epsilon, \Delta_\epsilon \varphi)_\epsilon\big| \leq C\epsilon \big|\mathcal{Z}_S^\epsilon(0)\big| \tag{2.33}$$

and by (2.22) the RHS tends to zero in $L^2(\Omega)$ as $\epsilon \to 0$. Now, by the fundamental theorem of calculus, one may easily check that $\Delta_\epsilon \varphi(X) = \int_X^{X+\epsilon} \int_{Z-\epsilon}^{Z} \varphi''(W) dW dZ$, and hence by uniform continuity of φ'', it follows that

$$\lim_{\epsilon \to 0} \sup_{X \in \mathbb{R}} \big|\varphi''(X) - \Delta_\epsilon \varphi(X)\big| = 0. \tag{2.34}$$

Letting $F(\epsilon) := \sup_{X\in\mathbb{R}} \big|\varphi''(X) - \Delta_\epsilon \varphi(X)\big|$, we find that

$$\begin{aligned}
\big\|(\mathcal{Z}_S^\epsilon, \Delta_\epsilon \varphi)_\epsilon - (\mathcal{Z}_S^\epsilon, \varphi'')_\epsilon\big\|_2 &\leq \epsilon \sum_{X \in \epsilon\Lambda} \|Z_S^\epsilon(X)\|_2 \cdot F(\epsilon) \cdot 1_{\{\varphi(X)>0\}} \\
&\leq \left(\epsilon \sum_{X\in\epsilon\Lambda} Ce^{aX} \cdot 1_{\{\varphi(X)>0\}}\right) \cdot F(\epsilon) \\
&\leq C|supp(\varphi)|e^{a|supp(\varphi)|} F(\epsilon) \\
&= C \cdot F(\epsilon)
\end{aligned} \tag{2.35}$$

where $|supp(\varphi)|$ denotes the supremum of the support of φ. In the second line we used (2.22) to bound $\|Z_S^\epsilon(X)\|_2$ by Ce^{aX}, and in the third line we used compact support of φ.

Using (2.33), (2.34), and (2.35), it follows that

$$\left\|(\Delta_\epsilon \mathcal{Z}_S^\epsilon, \varphi)_\epsilon - (\mathcal{Z}_S^\epsilon, \varphi'')_\epsilon\right\|_2 \stackrel{\epsilon\to 0}{\longrightarrow} 0.$$

Here, as usual, we are denoting $\|X\|_p := \mathbb{E}[|X|^p]^{1/p}$. Therefore, we can write for $T \le \tau$:

$$Y_T^\epsilon(\varphi) = (\mathcal{Z}_T^\epsilon, \varphi)_\epsilon - \frac{1}{2}\int_0^T (\mathcal{Z}_S^\epsilon, \varphi'')_\epsilon dS + R_T^\epsilon(\varphi) \tag{2.36}$$

where $\lim_{\epsilon\to 0}\|R_T^\epsilon\|_2 = 0$. Letting $\epsilon \to 0$ in this last expression, it follows that $Y_T(\varphi)$ is a $\mathbb{P}$-martingale for any limit point $\mathbb{P}$ of the law of $\mathcal{Z}^\epsilon$. Here we are using the fact that $(\mathcal{Z}_T^\epsilon, \varphi)_\epsilon$ converges weakly along some subsequence to $(\mathcal{L}_T, \varphi)$ under $\mathbb{P}$, which can be seen using a Riemann sum approximation.

Thus we only need to show that $Q_T(\varphi)$ is a $\mathbb{P}$-martingale (with respect to the canonical filtration on $C([0,\tau], C(I))$) for any limit point $\mathbb{P}$ of the law of $\mathcal{Z}^\epsilon$. This will be the the more difficult part of the proof (due to the extra term appearing in (2.1)), and it is really where the brunt of our efforts will go. The basic idea is as follows: We define

$$Q_T^\epsilon(\varphi) := Y_T^\epsilon(\varphi)^2 - \langle Y^\epsilon(\varphi)\rangle_T$$

where $\langle Y^\epsilon(\varphi)\rangle$ denotes the predictable bracket of the martingale $Y_T^\epsilon(\varphi)$. Our goal will be to write Q_T^ϵ in a form similar to (2.36), i.e.,

$$Q_T^\epsilon(\varphi) = Y_T^\epsilon(\varphi)^2 - \int_0^T ((\mathcal{Z}_S^\epsilon)^2, \varphi^2)_\epsilon dS + R_T^\epsilon$$

where R_T^ϵ is some error term whose L^2 norm vanishes as $\epsilon \to 0$. So the remainder of the proof will be devoted to this goal. Indeed, using Equation (2.1), it holds that

$$\langle M\rangle_t = \int_0^t \Big((\epsilon + o(\epsilon))Z_s(x)^2 - \nabla^+ Z_s(x)\nabla^- Z_s(x)\Big)ds.$$

Consequently, we find that

$$\langle Y^\epsilon(\varphi)\rangle_T = \epsilon^{-2}\int_0^T \Big((\epsilon + o(\epsilon))^2((\mathcal{Z}_S^\epsilon)^2, \varphi^2)_\epsilon + \epsilon\cdot(\nabla_\epsilon^+ \mathcal{Z}_S^\epsilon \cdot \nabla_\epsilon^- \mathcal{Z}_S^\epsilon\ ,\ \varphi^2)_\epsilon\Big)dS$$

where $\nabla_\epsilon^\pm \varphi(X) = \varphi(X) - \varphi(X \pm \epsilon)$. Since $\epsilon^{-2}(\epsilon + o(\epsilon))^2 = 1 + o(1)$, it is clear that

$$\epsilon^{-2}\int_0^T (\epsilon + o(\epsilon))^2((\mathcal{Z}_S^\epsilon)^2, \varphi^2)_\epsilon dS$$

converges weakly to $\int_0^T (\mathcal{L}_S^2, \varphi^2)dS$ (under $\mathbb{P}$) as $\epsilon \to 0$ (subsequentially). Thus defining

$$R_T^\epsilon(\varphi) := \epsilon^{-1}\int_0^T (\nabla_\epsilon^+ \mathcal{Z}_S^\epsilon \cdot \nabla_\epsilon^- \mathcal{Z}_S^\epsilon\ ,\ \varphi^2)_\epsilon dS$$

the proof will be completed if we can show that $\mathbb{E}[R_T^\epsilon(\varphi)^2] \to 0$ as $\epsilon \to 0$. Changing back to microscopic variables, we note that

$$R_T^\epsilon(\varphi) = \epsilon^2 \int_0^{\epsilon^{-2}T} \sum_{x\in\Lambda} \nabla^+ Z_s(x)\nabla^- Z_s(x)\varphi(\epsilon x)^2 ds.$$

Using the identity $\big(\int_0^t f(s)ds\big)^2 = 2\int_0^t\int_0^s f(s)f(s')ds'ds$, we find that

$$R_T^\epsilon(\varphi)^2 = 2\epsilon^4 \int_0^{\epsilon^{-2}T}\int_0^s \sum_{x,y\in\Lambda} \varphi(\epsilon x)^2\varphi(\epsilon y)^2 \nabla^+ Z_{s'}(x)\nabla^- Z_{s'}(x)\nabla^+ Z_s(y)\nabla^- Z_s(y)\ ds'ds.$$

Recall the filtration $\mathcal{F}_s = \sigma(\{\eta_s(x) : x \in \Lambda,\ s \le t\})$, and define

$$U^\epsilon(y, s', s) := \mathbb{E}\big[\nabla^+ Z_s(y)\nabla^- Z_s(y)\big|\mathcal{F}_{s'}\big].$$

Then we find that

$$\mathbb{E}[R_T^\epsilon(\varphi)^2] = 2\epsilon^4 \int_0^{\epsilon^{-2}T} \int_0^s \sum_{x,y\in\Lambda} \varphi(\epsilon x)^2 \varphi(\epsilon y)^2 \mathbb{E}\big[\nabla^+ Z_s(x) \nabla^- Z_{s'}(x) U^\epsilon(y,s',s)\big] \; ds'ds.$$

Now since $Z_t(x) = e^{2\sqrt{\epsilon}h_t(x)-\nu t}$, and since $|e^q - e^p| \leq |q-p|e^q$ for $p<q$, it follows that we have the “brute-force" bound: $|\nabla^\pm Z_t(x)| \leq C\epsilon^{1/2} Z_t(x)$, hence the preceding expression yields

$$\mathbb{E}[R_T^\epsilon(\varphi)^2] \leq C\epsilon^5 \int_0^{\epsilon^{-2}T} \int_0^s \sum_{x,y\in\Lambda} \varphi(\epsilon x)^2 \varphi(\epsilon y)^2 \mathbb{E}\big[Z_{s'}(x)^2 U^\epsilon(y,s',s)\big] \; ds'ds. \tag{2.37}$$

Now, we make the following claim, the proof of which we will postpone until a bit later:

$$\sup_{x\in\Lambda} e^{-2a\epsilon x} \mathbb{E}|U^\epsilon(x,s,t)| \leq C\epsilon^{1/8}(t-s)^{-1/2}, \qquad \epsilon^{-3/2} \leq s \leq t \leq \epsilon^{-2}T. \qquad \textbf{(Claim 1)}$$

The $1/8$ appearing in the power is not sharp, it is just a convenient bound which suffices for our purposes. Before proving (**Claim 1**), let us see why it implies the result. First we are going to split up the integral in (2.37) as three terms:

$$\int_0^{\epsilon^{-3/2}} \int_0^s (-) ds'ds \; + \; \int_{\epsilon^{-3/2}}^{\epsilon^{-2}T} \int_0^{\epsilon^{-3/2}} (-) ds'ds \; + \; \int_{\epsilon^{-3/2}}^{\epsilon^{-2}T} \int_{\epsilon^{-3/2}}^s (-) ds'ds. \tag{2.38}$$

To deal with the first two terms, we use the brute-force bound $|\nabla^+ Z_s(y) \nabla^- Z_s(y)| \leq \epsilon Z_s(y)^2$, which in turn implies that the expectation appearing in (2.37) is bounded by $\epsilon \cdot \mathbb{E}[Z_{s'}(x)^2 Z_s(y)^2]$. By Cauchy-Schwarz, this is in turn bounded by $\epsilon \cdot \|Z_{s'}(x)\|_4^2 \|Z_s(y)\|_4^2$, which by (2.22) is further bounded by $\epsilon \cdot Ce^{2a(x+y)}$. Noting (for instance by Riemann sum approximation) that

$$\epsilon^2 \sum_{x,y\in\Lambda} \varphi(\epsilon x)^2 \varphi(\epsilon y)^2 e^{Ca\epsilon(x+y)} \leq C, \tag{2.39}$$

it follows that the first integral in (2.38) will be bounded by $C\epsilon$ and the second one will be bounded by $C\epsilon^{1/2}$, which is sufficient, since these quantities approach zero as $\epsilon \to 0$.

In order to deal with the third integral in (2.38), we define events $A_{K,s',x} := \{|Z_{s'}(x)| \leq K\}$, and we are going to split up the expectation in (2.37) according to whether or not $A_{K,x,s'}$ occurs or not:

$$\mathbb{E}\big[Z_{s'}(x)^2 \cdot 1_{A_{K,x,s'}} U^\epsilon(y,s',s)\big] \quad + \quad \mathbb{E}\big[Z_{s'}(x)^2 \cdot 1_{A^c_{K,x,s'}} \nabla^+ Z_s(y) \nabla^- Z_s(y)\big] \tag{2.40}$$

where A^c denotes the complement of A. The first term of (2.40) can be bounded by $K^2\mathbb{E}[U^\epsilon(y,s',s)]$, which by (**Claim 1**) is bounded by $CK^2\epsilon^{1/8}(t-s)^{-1/2}e^{a\epsilon y}$. Now for the second term. Since $1_{A^c_{K,x,s'}} \leq K^{-2}Z_{s'}(x,y)$, and since $\nabla^+ Z_s(y)\nabla^- Z_s(y) \leq \epsilon Z_s(y)^2$, the second term of (2.40) is bounded by $K^{-2}\epsilon\mathbb{E}[Z_{s'}(x)^4 Z_s(y)^2]$, which (by the Cauchy-Schwarz inequality) is in turn bounded by $K^{-2}\epsilon\|Z_{s'}(x)\|_8^4\|Z_s(y)\|_4^2$. By (2.22), this is further bounded by $K^{-2}\epsilon \cdot Ce^{4a\epsilon x+2a\epsilon y}$. Then applying (2.39), it will follow that the third term in (2.38) will be bounded by $C(\epsilon^{1/8}K^2 + K^{-2})$. Letting $\epsilon \to 0$ and then $K \to \infty$, it finally follows that the RHS of (2.37) approaches zero as $\epsilon \to 0$.

Thus all that is left to do is to prove (**Claim 1**). This is done separately below. □

Proof of (**Claim 1**)*:* To prove the "key estimate" (as [16] calls it) we are going to write:

$$Z_t(x) = L_t(x) + D_t^t(x)$$

where L stands for the solution to the **L**inear equation:

$$L_t(x) := \mathbf{p}_t^R * Z_0(x)$$

and D stands for the **D**uhamel contribution from the random noise:

$$D_s^t(x) := \int_0^s \sum_y \mathbf{p}_{t-u}^R(x,y)dM_u(y).$$

We note that

$$\begin{aligned}\nabla^+ Z_t(x)\nabla^- Z_t(x) =& \nabla^+ L_t(x)\nabla^- L_t(x) + \nabla^+ D_t^t(x)\nabla^- D_t^t(x)\\ &+ \nabla^+ D_t^t(x)\nabla^- L_t(x) + \nabla^+ L_t(x)\nabla^- D_t^t(x).\end{aligned}$$

Noting that $D_s^t(x)$ is a martingale in s, it follows that $\nabla^\pm D_s^t(x)$ is a martingale in s, and hence $\nabla^+ D_s^t(x)\nabla^- D_s^t(x) - \langle \nabla^+ D^t(x), \nabla^- D^t(x)\rangle_s$ is a martingale. Using the condition on the martingales that $\langle M(x), M(y)\rangle_t = 0$ if $x \neq y$, it follows that

$$\langle \nabla^+ D^t(x), \nabla^- D^t(x)\rangle_s = \int_0^s \sum_y K_{t-u}(x, y; A) d\langle M(y)\rangle_u,$$

where

$$K_t(x, y; A) := \nabla^+ \mathbf{p}_t^R(x, y; A)\nabla^- \mathbf{p}_t^R(x, y; A).$$

The reason we specify A here but not elsewhere will be made clear below. Summarizing the last paragraph, we find that if $s \le r \le t$ then

$$\mathbb{E}\big[\nabla^+ D_r^t(x)\nabla^- D_r^t(x) \bigm| \mathcal{F}_s\big] = \nabla^+ D_s^t(x)\nabla^- D_s^t(x) + \mathbb{E}\bigg[\int_s^r K_{t-u}(x, y; A) d\langle M(y)\rangle_u \biggm| \mathcal{F}_s\bigg].$$

Hence

$$\begin{aligned}U^\epsilon(x, s, t) =& \nabla^+ L_t(x)\nabla^- L_t(x) + \nabla^+ D_s^t(x)\nabla^- D_s^t(x) + \mathbb{E}\bigg[\int_s^t K_{t-u}(x, y; A) d\langle M(y)\rangle_u \biggm| \mathcal{F}_s\bigg]\\ &+ \nabla^+ D_s^t(x)\nabla^- L_t(x) + \nabla^+ L_t(x)\nabla^- D_s^t(x).\end{aligned}$$

Let us call the terms on the RHS as $I_1, ..., I_5$, respectively. In order to bound the expectations of I_1, I_2, I_4, and I_5, it suffices (by the Cauchy-Schwarz inequality) to bound both $\mathbb{E}[(\nabla^\pm L_t(x))^2]$ and $\mathbb{E}[(\nabla^\pm D_s^t(x))^2]$ by $C\epsilon^{1/8}(t-s)^{-1/2}e^{2a\epsilon x}$, where C does not depend on x, ϵ, or $s, t \in [\epsilon^{-3/2}, \epsilon^{-2}T]$. The bound for I_3 is more involved, and will be done later.

Let us start be getting the desired bound on $\mathbb{E}[(\nabla^{\pm} L_t(x))^2]$. Using Assumption 2.5.2 and then Corollary 2.3.3 or 2.3.17, we have that

$$\begin{aligned}\|\nabla^{\pm} L_t(x)\|_2 &\le \sum_y |\nabla^{\pm}\mathbf{p}_t^R(x,y)| \cdot \|Z_0(y)\|_2 \\ &\le C\sum_y |\nabla^{\pm}\mathbf{p}_t^R(x,y)| e^{a\epsilon y} \\ &\le Ct^{-1/2}e^{a\epsilon x}.\end{aligned}$$

Squaring both sides, we get

$$\mathbb{E}[(\nabla^{\pm} L_t(x))^2] \le Ct^{-1}e^{2a\epsilon x}.$$

Using the assumption that $t \ge \epsilon^{-3/2}$, we have that $t^{-1/2} \le \epsilon^{3/4}$, so that $t^{-1} \le \epsilon^{3/4}t^{-1/2} \le \epsilon^{3/4}(t-s)^{-1/2}$. Noting that $\epsilon^{3/4} \le \epsilon^{1/8}$ when $\epsilon \le 1$, the desired bound for $\mathbb{E}[(\nabla^{\pm} L_t(x))^2]$ follows.

Next, we are going to bound $\mathbb{E}[(\nabla^{\pm} D_s^t(x))^2]$. First note that $\|\frac{d}{du}\langle M(y)\rangle_u\|_p \le C\epsilon e^{2a\epsilon y}$, where C does not depend on y. This follows by (2.1), the "brute-force" bound $|\nabla^+ Z_u(x)\nabla^- Z_u(x)| \le C\epsilon Z_u(x)^2$, and (2.22). Note that

$$\begin{aligned}\mathbb{E}[(\nabla^{\pm} D_s^t(x))^2] &= \mathbb{E}[\langle \nabla^{\pm} D^t(x)\rangle_s] \\ &= \mathbb{E}\Big[\int_0^s \sum_y (\nabla^{\pm}\mathbf{p}_{t-u}^R(x,y))^2 d\langle M(y)\rangle_u\Big] \\ &\le \int_0^s \sum_y (\nabla^{\pm}\mathbf{p}_{t-u}^R(x,y))^2 \cdot \mathbb{E}\Big|\frac{d}{du}\langle M(y)\rangle_u\Big| du \\ &\le C\epsilon \int_0^s (t-u)^{-1}\sum_y \nabla^{\pm}\mathbf{p}_{t-u}^R(x,y)e^{2a\epsilon y} du \\ &\le C\epsilon e^{2a\epsilon x}\int_0^s (t-u)^{-3/2} du \\ &\le C\epsilon e^{2a\epsilon x}(t-s)^{-1/2}.\end{aligned}$$

We used Proposition 2.3.2 or 2.3.16 in the fourth line, and Corollary 2.3.3 or 2.3.17 in the fifth line. In the last line we just performed the integral and omitted the subtracted term. The desired claim now follows by noting that $\epsilon \leq \epsilon^{1/8}$ for $\epsilon \leq 1$. This completes the proof of desired bounds for I_1, I_2, I_4, and I_5.

Now we just need to bound I_3, which was defined as $\mathbb{E}\big[\int_s^t \sum_y K_{t-u}(x,y;A)d\langle M(y)\rangle_u \big| \mathcal{F}_s\big]$. Using Equation (2.1), we can expand $d\langle M(y)\rangle_u$ into $(\epsilon + o(\epsilon))Z_u(y)^2 du + \nabla^+ Z_u(y)\nabla^- Z_u(y)du$. Consequently, we find that

$$|I_3| \leq C\epsilon \bigg| \int_s^t \sum_y K_{t-u}(x,y;A)\mathbb{E}[Z_u(y)^2|\mathcal{F}_s]du \bigg| + \int_s^t \sum_y |K_{t-u}(x,y;A)||U^\epsilon(y,s,u)|du. \tag{2.41}$$

Now we make a claim which we hold off until later:

$$\mathbb{E}\bigg[\epsilon\bigg|\int_s^t \sum_y K_{t-u}(x,y;A)\mathbb{E}[Z_u(y)^2|\mathcal{F}_s]du\bigg|\bigg] \leq C\epsilon^{1/8}e^{2a\epsilon x}(t-s)^{-1/2}. \qquad \textbf{(Claim 2)}$$

Before proving this, let us first see why it implies (**Claim 1**). Using the bounds for I_1, I_2, I_4, and I_5 together with (2.41) and (**Claim 2**) shows that

$$\mathbb{E}|U^\epsilon(x,s,t)| \leq C\epsilon^{1/8}e^{2a\epsilon x}(t-s)^{-1/2} + \int_s^t \sum_y |K_{t-u}(x,y;A)|\mathbb{E}|U^\epsilon(y,s,u)|du. \tag{2.42}$$

Now we can iterate this, and we get that

$$\mathbb{E}|U^\epsilon(x,s,t)| \leq C\epsilon^{1/8}e^{2a\epsilon x}(t-s)^{-1/2}$$

$$+C\epsilon^{1/8}\sum_{n=1}^\infty \int_{\Delta_n(s,t)} (u_n-s)^{-1/2} \sum_{y_1,...,y_n} |K_{t-u_1}(x,y_1;A)| \prod_{i=1}^{n-1} |K_{u_i-u_{i+1}}(y_i,y_{i+1};A)| e^{2a\epsilon y_n} du_n \cdots du_1$$

where $\Delta_n(s,t) := \{(u_1,...,u_n) : s \leq u_n \leq \cdots \leq u_1 \leq t\}$. Now note by the triangle inequality that $e^{2a\epsilon y_n} \leq e^{2a\epsilon x}e^{2a\epsilon|y_1-x|}\prod_{i=1}^n e^{2a\epsilon|y_{i+1}-y_i|}$. Now using Propositions 2.3.11 and 2.3.12 (or

2.3.28), one may check that the integral over $\Delta_n(s,t)$ is at most $C\epsilon \cdot c_*^n \cdot e^{2a\epsilon x}$, where $c_* < 1$ is a constant independent of x and ϵ. Using the fact that $\epsilon \leq C(t-s)^{-1/2}$, the desired claim now follows easily. □

Proof of (**Claim 2**)*:* This will be the final part of the proof. We will prove that

$$\mathbb{E}\left[\epsilon\left|\sum_{y\in\Lambda}\int_s^t K_{t-u}(x,y;A)\mathbb{E}[Z_u(y)^2|\mathcal{F}_s]du\right|\right] \leq C\epsilon^{1/8}e^{2a\epsilon x}(t-s)^{-1/2}$$

uniformly in $x \in \mathbb{Z}_{\geq 0}$ and $s,t \in [0,\epsilon^{-2}\tau]$, where

$$K_t(x,y;A) = \nabla^+\mathbf{p}_t^R(x,y;A)\nabla^-\mathbf{p}_t^R(x,y;A)$$

and $\mathbf{p}_t^R(\cdot,\cdot;A)$ is the Robin heat kernel with parameter μ_A, and $\mathcal{F}_t = \sigma(\eta_s(x) : s \leq t, x \in \Lambda)$.

In the special case of ASEP-H, the same result holds with $\epsilon^{1/8}$ improved to $\epsilon^{1/2-\delta}$. This can be seen by dissecting the proof below, and it is mainly because of the difference between the methods used to prove Lemmas 2.3.9 and 2.3.26.

In order to prove the above inequality, we write

$$\begin{aligned}
&\epsilon\sum_{y\in\Lambda}\int_s^t |K_{t-u}(x,y;A)|\mathbb{E}[Z_u(y)^2|\mathcal{F}_s]du \\
&\qquad\qquad = \epsilon\sum_{y\in\Lambda}\int_s^t |K_{t-u}(x,y;A)|\mathbb{E}[Z_u(y)^2 - Z_t(x)^2|\mathcal{F}_s]du \\
&\qquad\qquad\quad + \epsilon\mathbb{E}[Z_t(x)^2|\mathcal{F}_s]\sum_{y\in\Lambda}\int_s^t |K_{t-u}(x,y;A) - K_{t-u}(x,y;0)|du \\
&\qquad\qquad\quad + \epsilon\mathbb{E}[Z_t(x)^2|\mathcal{F}_s]\sum_{y\in\Lambda}\int_s^t |K_{t-u}(x,y;0)|du.
\end{aligned}$$

Let us name the three terms on the RHS as J_1, J_2, J_3, in that specific order. To bound J_1, we write $Z_u(y)^2 - Z_t(x)^2 = (Z_u(y) - Z_t(x))(Z_u(y) + Z_t(x))$, so by Cauchy-Schwarz we have that $\|Z_u(y)^2 - Z_t(x)^2\|_1 \leq \|Z_u(y) + Z_t(x)\|_2 \|Z_u(y) - Z_t(x)\|_2$. Using (2.22), we have that $\|Z_u(y) + Z_t(x)\|_2 \leq Ce^{a\epsilon x}$. Using (2.23) and (2.24), we have that for $\alpha < 1/2$

$$\begin{aligned}\|Z_u(y) - Z_t(x)\|_2 &\leq \|Z_u(y) - Z_u(x)\|_2 + \|Z_u(x) - Z_t(x)\|_2 \\ &\leq C(\epsilon|x-y|)^\alpha e^{a\epsilon(x+y)} + C\epsilon^\alpha(1 \vee |t-u|^{\alpha/2})e^{2a\epsilon x}.\end{aligned}$$

Let us take $\alpha = 1/8$ for this proof. Note from Proposition 2.3.2 (or 2.3.16) that $|K_t(x,y;A)| \leq (1 \wedge t^{-1})|\nabla^+ \mathbf{p}_t^R(x,y)|$. Now we make a few observations: firstly, note that $(1 \wedge t^{-1})(1 \vee t^{\alpha/2}) = (1 \wedge t^{-(2-\alpha)/2})$. Secondly, by the identity $r^\alpha \leq e^r$, we find that $(\epsilon|x-y|)^\alpha \leq \epsilon^\alpha(1 \vee t^{\alpha/2})e^{-(1\wedge t^{-1/2})|x-y|}$. Combining all of these observations, one finds that

$$\begin{aligned}\mathbb{E}[|J_1|] &\leq \epsilon \cdot C\epsilon^\alpha \int_s^t (1 \wedge (t-u)^{-(2-\alpha)/2}) \sum_{y \in \Lambda} |\nabla^+ \mathbf{p}_{t-u}^R(x,y)| \big[e^{a\epsilon(x+y)}e^{-(1\wedge(t-u)^{-1/2})|x-y|} + e^{2a\epsilon x}\big] \\ &\leq C\epsilon \cdot \epsilon^{1/8} \int_0^t (1 \wedge (t-u)^{-3/2})e^{2a\epsilon x}du \\ &\leq C(t-s)^{-1/2} \cdot \epsilon^{1/8}e^{2a\epsilon x} \int_0^\infty (1 \wedge u^{-3/2})du.\end{aligned}$$

In the second line we substituted $\alpha = 1/8$, and we used Corollary 2.3.3 (or 2.3.17) to bound the sum over y. In the next line we noted that $\epsilon \leq C(t-s)^{-1/2}$ since $t - s \leq \epsilon^{-2}\tau$. In the final line we made the substitution $u \to t - u$. This proves that J_1 satisfies the desired bound.

To show that J_2 satisfies a bound of the desired type, we apply Equation (2.22) together with Lemma 2.3.9 or 2.3.26 to say that

$$\begin{aligned}\mathbb{E}[|J_2|] &\leq \epsilon\|Z_t(x)\|_2^2 \sum_{y \geq 0} \int_0^{\epsilon^{-2}\tau} |K_t(x,y;A) - K_t(x,y;0)|dt \\ &\leq \epsilon \cdot Ce^{2a\epsilon x} \cdot \epsilon^{1/8} \\ &\leq C(t-s)^{-1/2}e^{2a\epsilon x}\epsilon^{1/8}\end{aligned}$$

where we used the fact that $t-s \leq \epsilon^{-2}\tau$.

As for J_3, we can split it into two further terms

$$J_3 = \epsilon \mathbb{E}[Z_t(x)^2|\mathcal{F}_s]\bigg(\sum_{y\geq 0}\int_0^\infty K_u(x,y;0)du - \sum_{y\geq 0}\int_{t-s}^\infty K_u(x,y;0)du\bigg).$$

As a consequence of Proposition 2.3.10 or 2.3.27, the first term satisfies

$$\epsilon \mathbb{E}[Z_t(x)^2|\mathcal{F}_s]\sum_{y\geq 0}\int_0^\infty K_\tau(x,y;0)d\tau = O(\epsilon^2)\mathbb{E}[Z_t(x)^2|\mathcal{F}_s].$$

After taking expectations, this will be bounded by $C\epsilon^2\|Z_t(x)\|_2^2$, which by (2.22) (and the fact that $t-s\leq \epsilon^{-2}\tau$) is bounded by $C\epsilon(t-s)^{-1/2}e^{2a\epsilon x}$. To bound the second term on the RHS, note by Proposition 2.3.2 (or 2.3.16) and Corollary 2.3.3 (or 2.3.17) that $\sum_y |K_u(x,y)| \leq 1\wedge u^{-3/2}$. After integrating from $t-s$ to ∞, this will be bounded by $C(t-s)^{-1/2}$, which completes the proof. $\square$

2.6 Extension of Results to Narrow-Wedge Initial Data

As usual, we adopt the notation $\Lambda = \{0,...,N\}$ or $\Lambda = \mathbb{Z}_{\geq 0}$, and respectively $I=[0,1]$ or $I=[0,\infty)$, depending on the model under consideration.

In this section we will consider the weakly asymmetric limit of the height functions for open ASEP started from an initial configuration which has zero particles. There are several reasons why such initial data poses a problem. The first problem is as follows:

- The associated sequence of initial data for the rescaled Gärtner-transformed height functions is $\mathcal{Z}_0^\epsilon(X) = e^{-\epsilon^{-1/2}X}$. Note that this converges weakly (in the PDE sense) to 0 as $\epsilon\to 0$, therefore one may expect that $\mathcal{Z}^\epsilon(T,X)$ will just converge almost surely to zero as $\epsilon \to 0$, which is indeed the case. So the limit is trivial.

Hence it is clear that some sort of normalization is necessary in order to obtain a nontrivial limiting object. The solution to this problem is to introduce a logarithmically diverging correction to the height function, as in [5, 40]. More specifically, we want our Gärtner-transformed initial data $\mathscr{Z}_0^\epsilon(X) = e^{-\epsilon^{-1/2}X}$ to converge to something nontrivial as $\epsilon \to 0$, and the only sensible way to do this is to multiply by a factor of $\epsilon^{-1/2}$ since that will make $\mathscr{Z}_0^\epsilon$ converge weakly (in the PDE sense) to δ_0 as $\epsilon \to 0$. Hence we will *redefine* $Z_t(x)$ and $\mathscr{Z}^\epsilon(T,X)$ as the following quantities:

$$Z_t(x) = \frac{\varrho}{\epsilon^{1/2}} e^{\lambda h_t(x) + \nu t}$$

$$\mathscr{Z}^\epsilon(T,X) = Z_{\epsilon^{-2}T}(\epsilon^{-1}X)$$

where $\varrho := \big(\int_I e^{-Z} dZ \big)^{-1}$. So $\varrho = 1$ if $I = [0,\infty)$, and $\varrho = 1/(1-e^{-1})$ if $I = [0,1]$. The choice for this constant will be made clear later (see Lemma 2.6.6 below).

We hope that it is clear that this is **not** the same $Z_t(x)$ and $\mathscr{Z}^\epsilon(T,X)$ appearing in Sections 2 and 4, due to the normalizing factor which is $\epsilon^{-1/2}\varrho$.

Although this normalization scheme should presumably give us a nontrivial object in the limit, the proof involves some subtleties. The main issue is that

- The associated sequence of initial data for the (redefined) Gärtner-transformed height functions $\mathscr{Z}_0^\epsilon(T,X) = \frac{\varrho}{\epsilon^{1/2}} e^{-\epsilon^{-1/2}X}$ is no longer "near-equlilbrium," as defined in Assumption 2.5.2. Therefore, the results of Section 5 no longer apply in our case.

The solution to this issue involves quite a few subtleties, and therefore we will devote this section to proving convergence of $\mathscr{Z}^\epsilon$ to the SHE in this case.

Recall that $\|X\|_p = \mathbb{E}[|X|^p]^{1/p}$.

Lemma 2.6.1. *For all $T \geq 0$, there exists $C = C(T)$ such that for every $\epsilon > 0$ sufficiently small*

$$\sup_{t \leq T} \|Z_t(x)\|_p \leq C Z_0(x).$$

Proof. Since we are starting from an empty configuration of particles and since the exponential jump rate satisfies $p \sim \frac{1}{2} + O(\sqrt{\epsilon}) \leq 1$, we know that the position of the largest occupied site at time t is stochastically dominated by a Poisson random variable $N(t)$, with mean t. Therefore $h_t(x)$ is stochastically dominated by $N(t) + h_0(x)$. Therefore for all $x \in \Lambda$ and $t \leq T$, we have that

$$Z_t(x) = \epsilon^{-1/2} e^{\epsilon^{1/2} h_t(x) + \nu t} \leq \epsilon^{-1/2} e^{\epsilon^{1/2} N(t)} e^{\epsilon^{1/2} h_0(x) + \nu t} = e^{\epsilon^{1/2} N(t)} Z_0(x)$$

and consequently

$$\|Z_t(x)\|_p \leq \|e^{\epsilon^{1/2} N(t)}\|_p Z_0(x) = e^{\epsilon^{1/2}(e^{pt} - 1)} Z_0(x)$$

from which one may deduce the claim. □

Proposition 2.6.2. *Fix a terminal time $\tau \geq 0$. Let $\alpha \in [0, 1/2)$. We have the following bounds, uniformly over all (small enough) $\epsilon > 0$, $x, y \in \Lambda$, and $s, t \in [1, \epsilon^{-2}\tau]$ with $s < t$:*

$$\|Z_t(x)\|_p \leq C(\epsilon^2 t)^{-1/2}, \tag{2.43}$$

$$\|Z_t(x) - Z_t(y)\|_p \leq C(\epsilon|x - y|)^{\alpha}(\epsilon^2 t)^{-(1+\alpha)/2}, \tag{2.44}$$

$$\|Z_t(x) - Z_s(x)\|_p \leq C\epsilon^{\alpha}(1 \vee |t - s|)^{\alpha/2}(\epsilon^2 s)^{-(1+\alpha)/2}. \tag{2.45}$$

Proof. The proof here will be loosely based on the one given in [, Proposition 1.8], however we found a small mistake there (the same one mentioned in the proof of Proposition 2.5.4) so several new ideas will also be used. These will involve the long-time estimates, Propositions 2.3.6 and 2.3.22.

We have from Lemma 2.2.7 that

$$Z_t(x) = \mathbf{p}_t^R * Z_0(x) + \int_0^t \sum_{y\geq 0} \mathbf{p}_{t-s}^R(x,y) dM_s(y)$$

Using Lemma 2.5.3 and (2.21), and we obtain for $t \geq 1$ that

$$\Big\| \int_0^t \sum_{y\geq 0} \mathbf{p}_s^R(x,y) dM_s(y) \Big\|_p^2 \leq C\epsilon \int_0^t (t-s)^{-1/2} \sum_{y\geq 0} \mathbf{p}_{t-s+1}^R(x,y) \|Z_s(y)\|_p^2 ds.$$

Since $(x+y)^2 \leq 2x^2 + 2y^2$, we have

$$\|Z_t(x)\|_p^2 \leq 2\big|\mathbf{p}_t * Z_0(x)\big|^2 + C\epsilon \int_0^t (t-s)^{-1/2} \sum_{y\geq 0} \mathbf{p}_{t-s+1}^R(x,y) \|Z_s(y)\|_p^2 ds.$$

Using $\mathbf{p}_t^R(x,y) \leq Ct^{-1/2}$ together with $\sum_{y\geq 0} e^{-\epsilon^{1/2} y} = (1-e^{-\epsilon^{1/2}})^{-1} \sim \frac{1}{2}\epsilon^{-1/2}$ implies that

$$\mathbf{p}_t^R * Z_0(x) = \varrho\epsilon^{-1/2} \sum_y \mathbf{p}_t^R(x,y) e^{-\epsilon^{1/2} y} \leq C\epsilon^{-1} t^{-1/2} = C(\epsilon^2 t)^{-1/2} \tag{2.46}$$

which in turn implies that $|\mathbf{p}_t^R * Z_0(x)|^2 \leq C(\epsilon^2 t)^{-1/2} (\mathbf{p}_t * Z_0)(x)$.

Summarizing the above computations, we have proved that for all $t \in [1, \epsilon^{-2}T]$, we have

$$\|Z_t(x)\|_p^2 \leq C(\epsilon^2 t)^{-1/2} (\mathbf{p}_t^R * Z_0)(x) + C\epsilon \int_0^t (t-s)^{-1/2} \sum_{y\geq 0} \mathbf{p}_{t-s+1}^R(x,y) \|Z_s(y)\|_p^2 ds. \tag{2.47}$$

Now we would like to iterate this inequality, but the problem is that we have only proved (2.47) for $t \geq 1$, however we need to prove it for all $t \geq 0$ in order to apply the iteration. By Lemma 2.6.1 and Proposition 2.3.13 (or 2.3.29), for $t \leq 1$ we have

$$\|Z_t(x)\|_p \leq CZ_0(x) \leq Ce^t \mathbf{p}_t(x,x) Z_0(x) \leq Ce^1 \sum_y \mathbf{p}_t(x,y) Z_0(y) = C(\mathbf{p}_t * Z_0)(x). \tag{2.48}$$

Now squaring the LHS and RHS of (2.48), and then applying (2.46), we obtain that $\|Z_t(x)\|_p^2 \leq C(\epsilon^2 t)^{-1/2}(\mathbf{p}_t^R * Z_0)(x)$ for $t \leq 1$. This proves that (2.47) still holds when $t \leq 1$, thus justifying our ability to iterate it.

Iterating (2.47) and applying the semigroup property of $\mathbf{p}_t^R$ proves that

$$\begin{aligned}\|Z_t(x)\|_p^2 &\leq C(\epsilon^2 t)^{-1/2}(\mathbf{p}_t^R * Z_0)(x) \\ &\quad + \sum_{k=0}^{\infty} C^k \epsilon^k \bigg(\int_{\Delta_k(t)} t_0^{-1/2} \prod_{j=1}^{k} (t_j - t_{j-1})^{-1/2} (t-t_k)^{-1/2} dt_0 \cdots dt_k \bigg) (\mathbf{p}_{t+k}^R * Z_0)(x)\end{aligned} \tag{2.49}$$

where $\Delta_k(t) = \{(t_0, ..., t_k) \in \mathbb{R}^{k+1} \mid t_0 < \cdots < t_k < t\}$. A recursion will reveal that the integral within the parentheses is bounded $Ct^{k/2}/(k/2)!$. Recall the long-time estimates (Proposition 2.3.6 or 2.3.22) which show that

$$\mathbf{p}_t^R(x,y) \leq C(t^{-1/2} + \epsilon)e^{K\epsilon^2 t}$$

where C, K are constants **not** depending on the terminal time τ. This in turn implies that

$$\mathbf{p}_t^R * Z_0(x) \leq C(\epsilon^{-1} t^{-1/2} + 1)e^{K\epsilon^2 t} \tag{2.50}$$

Using (2.49) and the remark underneath, we then see that

$$\begin{aligned}\|Z_t(x)\|_p^2 &\leq C(\epsilon^2 t)^{-1/2}(\mathbf{p}_t^R * Z_0)(x) + C\sum_{k=0}^{\infty} \frac{C^k \epsilon^k t^{k/2}}{(k/2)!}(\epsilon^{-1}t^{-1/2}+1)e^{K\epsilon^2(t+k)} \\ &\leq C(\epsilon^2 t)^{-1/2}(\mathbf{p}_t^R * Z_0)(x) + C\big[(\epsilon^2 t)^{-1/2} + 1\big]e^{K'\epsilon^2 t} \\ &\leq C(\epsilon^2 t)^{-1/2}\big[(\mathbf{p}_t^R * Z_0)(x) + 1\big].\end{aligned} \tag{2.51}$$

In the final inequality we used the fact that $\epsilon^2 t \leq \tau$, and $1 \leq \tau^{1/2}(\epsilon^2 t)^{-1/2}$. The proof of the first bound (2.43) is completed by noting from (2.46) that $(\mathbf{p}_t^R * Z_0)(x) \leq C(\epsilon^2 t)^{-1/2}$, so that (2.51) is bounded by $C(\epsilon^2 t)^{-1}$.

To prove the second bound (2.44), note that

$$\|Z_t(x) - Z_t(y)\|_p \le |\mathbf{p}_t^R * Z_0(x) - \mathbf{p}_t^R * Z_0(y)| + \left\| \int_0^t \sum_z \big(\mathbf{p}_t^R(x,z) - \mathbf{p}_t^R(y,z)\big) dM_s(z) \right\|_p .$$

Let us call the terms on the RHS as J_1, J_2, respectively. We will show that each J_i satisfies a bound of the desired type. For J_1, note by Proposition 2.3.2 (or 2.3.16) that

$$J_1 \le \epsilon^{-1/2} \sum_z \big|\mathbf{p}_t^R(x,z) - \mathbf{p}_t^R(y,z)\big| e^{-\epsilon^{1/2} z} \le C\epsilon^{-1/2} t^{-(1+\alpha)/2} |x-y|^\alpha \sum_z e^{-\epsilon^{1/2} z}$$

$$\le C\epsilon^{-1} t^{-(1+\alpha)/2} |x-y|^\alpha = C(\epsilon^2 t)^{-(1+\alpha)/2} (\epsilon|x-y|)^\alpha .$$

For J_2, we are going to apply Lemma 2.5.3 with $F(s,z) = \mathbf{p}_{t-s}^R(x,z) - \mathbf{p}_{t-s}^R(y,z)$. Note that if $|t-s-u| \le 1$, then by the triangle inequality, Proposition 2.3.2 (or 2.3.16), and Proposition 2.3.13 (or 2.3.29), we have

$$\begin{aligned}
(\mathbf{p}_u^R(x,z) - \mathbf{p}_u^R(y,z))^2 &\le |\mathbf{p}_u^R(x,z) - \mathbf{p}_u^R(y,z)|(\mathbf{p}_u^R(x,z) + \mathbf{p}_u^R(y,z)) \\
&\le C(1 \wedge u^{-(1+\alpha)/2})|x-y|^\alpha \cdot e^1(\mathbf{p}_{t-s+1}^R(x,z) + \mathbf{p}_{t-s+1}^R(y,z)) \\
&\le C(t-s)^{-(1+\alpha)/2} |x-y|^\alpha \cdot (\mathbf{p}_{t-s+1}^R(x,z) + \mathbf{p}_{t-s+1}^R(y,z)).
\end{aligned}$$

Hence $\sup_{|s-u|\le 1} F_u(s,z)^2$ is bounded by the last expression. Thus by Lemma 2.5.3 and Equation (2.51) above, we find that for $t \ge 1$ and $\alpha \in [0,1)$ we have:

$$\begin{aligned}
J_2^2 &\le C\epsilon|x-y|^\alpha \int_0^t (t-s)^{-(1+\alpha)/2} \sum_z \big(\mathbf{p}_{t-s+1}^R(x,z) + \mathbf{p}_{t-s+1}^R(y,z)\big) \|Z_s(z)\|_p^2 ds \\
&\le C|x-y|^\alpha \int_0^t (t-s)^{-(1+\alpha)/2} s^{-1/2} \sum_z \big(\mathbf{p}_{t-s+1}^R(x,z) + \mathbf{p}_{t-s+1}^R(y,z)\big) \big[(\mathbf{p}_s^R * Z_0)(z) + 1\big] ds \\
&= C|x-y|^\alpha t^{-\alpha/2} \big((\mathbf{p}_{t+1}^R * Z_0)(x) + (\mathbf{p}_{t+1}^R * Z_0)(y) + 1 + 1\big) \\
&\le C(\epsilon|x-y|)^\alpha (\epsilon^2 t)^{-(1+\alpha)/2}.
\end{aligned}$$

Let us justify each of the inequalities above. In the second line we used (2.51) to bound $\|Z_s(z)\|_p^2$ by $C(\epsilon^2 s)^{-1/2}\big[(\mathbf{p}_s^R * Z_0)(z)+1\big]$. In the third line, we used the semigroup property to rewrite the sum over z, moreover we applied Corollaries 2.3.3 and 2.3.17 with $a_i = 0$ in order to bound the $+1$ term next to $\mathbf{p}_s * Z_0$, and then we used the fact that $\int_0^t (t-s)^{-(1+\alpha)/2} s^{-1/2} ds = Ct^{-\alpha/2}$ which can be proved by making the substitution $s = tu$. In the final line, we used (2.46) to bound $\mathbf{p}_{t+1}^R * Z_0$ by $C(\epsilon^2 t)^{-1/2}$. Taking square roots in the above expression, we find that for $\alpha < 1$,

$$J_2 \le C(\epsilon|x-y|)^{\alpha/2}(\epsilon^2 t)^{-(1+\alpha)/4}$$

and now we note that $(\epsilon^2 t)^{-(1+\alpha)/4} \le T^{(1+\alpha)/4}(\epsilon^2 t)^{-(1+\alpha)/2}$ which completes the proof of the second bound (2.44).

Now for the final bound (2.45). We again use the semigroup property and Proposition 2.2.7 to write

$$Z_t(x) = (\mathbf{p}_{t-s}^R * Z_s)(x) + \int_s^t \sum_y \mathbf{p}_{t-u}^R(x,y) dM_s(y).$$

Therefore

$$\|Z_t(x) - Z_s(x)\|_p \le \big\|(\mathbf{p}_{t-s}^R * Z_s)(x) - Z_s(x)\big\|_p + \bigg\| \int_s^t \sum_y \mathbf{p}_{t-u}^R(x,y) dM_s(y) \bigg\|_p.$$

Let us call the terms on the RHS as I_1, I_2 respectively. Note that

$$\begin{aligned} I_1 &= \bigg\| \sum_y \mathbf{p}_{t-s}^R(x,y) Z_s(y) - Z_s(x) \bigg\|_p \\ &\le \bigg\| \sum_y \mathbf{p}_{t-s}^R(x,y)\big(Z_s(y) - Z_s(x)\big) \bigg\|_p + \bigg| \sum_y \mathbf{p}_{t-s}^R(x,y) - 1 \bigg| \cdot \|Z_s(x)\|_p \\ &\le \sum_y \mathbf{p}_{t-s}^R(x,y)(\epsilon^2 s)^{-(1+\alpha)/2}(\epsilon|y-x|)^\alpha + C\epsilon(t-s)^{1/2} \cdot (\epsilon^2 s)^{-1/2}. \end{aligned} \tag{2.52}$$

In the final inequality we applied Proposition 2.3.7 (or 2.3.23) together with (2.43) and (2.44). Just as in (2.30) with $a = 0$, we have that

$$\sum_y \mathbf{p}^R_{t-s}(x,y)(\epsilon|x-y|)^\alpha \le C\epsilon^\alpha(1 \vee |t-s|^{\alpha/2}).$$

Also, since $s, t \le \epsilon^{-2}T$, it follows that $\epsilon(t-s)^{1/2} \le T^{(1-\alpha)/2}\epsilon^\alpha(t-s)^{\alpha/2}$ and similarly, $(\epsilon^2 s)^{-1/2} \le T^{\alpha/2}(\epsilon^2 s)^{-(1+\alpha)/2}$ for $\alpha < 1$, hence by absorbing those powers of T into the constant C, we get

$$\epsilon(t-s)^{1/2}(\epsilon^2 s)^{-1/2} \le C\epsilon^\alpha(t-s)^{\alpha/2}(\epsilon^2 s)^{-(1+\alpha)/2}.$$

Together with (2.52), the preceding two expressions give the desired bound on I_1. In fact the bound holds for all $\alpha < 1$, not just $\alpha < 1/2$.

Now we consider I_2. By the first bound (2.43), note that $\|Z_u(y)\|_p \le C(\epsilon^2 u)^{-1/2} \le C(\epsilon^2 s)^{-1/2}$ whenever $u \in [s,t]$. Furthermore, if $r \le \epsilon^{-2}T$, then we know by Corollary 2.3.3 (with $a_1 = a_2 = 0$) that $\sum_y \mathbf{p}^R_r(x,y)$ is bounded by a constant independent of x and r. Using these two facts,

$$\begin{aligned}
I_2^2 &\le C\epsilon \int_s^t (t-u)^{-1/2} \sum_y \mathbf{p}^R_{t-u+1}(x,y)\|Z_u(y)\|_p^2 du \\
&\le C\epsilon(\epsilon^2 s)^{-1} \int_s^t (t-u)^{-1/2} \sum_y \mathbf{p}^R_{t-u+1}(x,y) du \\
&= C\epsilon(\epsilon^2 s)^{-1} \int_s^t (t-u)^{-1/2} du \\
&= C(\epsilon^2 s)^{-1} \cdot \epsilon(t-s)^{1/2}.
\end{aligned}$$

So taking square roots, we see that for $\alpha < 1$ we have

$$I_2 \le C(\epsilon^2 s)^{-1/2}\epsilon^{1/2}(t-s)^{1/4} \le C(\epsilon^2 s)^{-\alpha/2}\epsilon^{\alpha/2}(t-s)^{\alpha/4}$$

where we again used the fact that $s, t \le \epsilon^{-2}T$ in the final line. This is equivalent to (2.45), thus completing the proof of the estimates. $\square$

Corollary 2.6.3. *For any $0 < \delta \leq \tau$, the laws of the rescaled processes $\{\mathcal{Z}^\epsilon\}_{\epsilon>0}$ are tight on the Skorokhod space $D([\delta,\tau], C(I))$. Moreover, any limit point lies in $C([\delta,\tau], C(I))$.*

For $T \in [\delta,\tau]$, let $\mathcal{L}(T) : C([\delta,\tau], C(I)) \to C(I)$ denote the evaluation map at T. Let $\mathbb{P}$ be a limit point of the $\{\mathcal{Z}^\epsilon\}$. Then the process $(\mathcal{L}(T+\delta))_{T\in[0,\tau-\delta]}$ has the same distribution under $\mathbb{P}$ as the solution of the SHE started from initial data distributed as that of $\mathcal{L}(\delta)$ under $\mathbb{P}$.

Proof. To show tightness on $D([\delta,\tau], C(I))$ for $0 < \delta \leq \tau$, we rewrite the estimates of Proposition 2.6.2 in terms of the rescaled macroscopic processes $\mathcal{Z}^\epsilon$: for $\alpha < 1/2$, $S, T \in [\delta,\tau]$, and $X, Y \in I$ we have

$$\|\mathcal{Z}^\epsilon(T,X)\|_p \leq CT^{-1/2}$$

$$\|\mathcal{Z}^\epsilon(T,X) - \mathcal{Z}^\epsilon(T,Y)\|_p \leq CT^{-(1+\alpha)/2}|X-Y|^\alpha$$

$$\|\mathcal{Z}^\epsilon(T,X) - \mathcal{Z}^\epsilon(S,X)\|_p \leq CS^{-(1+\alpha)/2}(|T-S|^{\alpha/2} \vee \epsilon^\alpha)$$

With the assumption that $S, T \geq \delta$, it follows that $T^{-1/2} \leq \delta^{-1/2}$, $T^{-(1+\alpha)/2}, S^{-(1+\alpha)/2} \leq \delta^{-(1+\alpha)/2}$. Therefore. The $\{\mathcal{Z}^\epsilon\}_{\epsilon\in(0,1]}$ are uniformly bounded, uniformly spatially Hölder, and uniformly temporally Hölder (except for jumps of order ϵ^α) with large probability. Now we apply the version of Arzela-Ascoli for Skorokhod spaces (together with Prohorov's theorem) to obtain tightness, see [17, Chapter 3].

The fact that any limit point lies in $C([\delta,\tau], C(I))$ is a straightforward consequence of Kolmogorov's continuity criterion.

To prove the final statement, note that the above bounds imply that the sequence of initial data $\{\mathcal{Z}^\epsilon(\delta,\cdot)\}$ is "near-equilibrium" as defined in Assumption 2.5.2. Thus the results of Section 5 apply, and the proof is finished. □

Definition 2.6.4. *Let $D((0,\infty), C(I))$ denote the set of all functions from $(0,\infty) \to C(I)$ which* *are continuous on the right and have left limits. For $0 < \delta \leq \tau$, let $\mathcal{L}_{\delta,\tau} : D((0,\infty), C(I)) \to D([\delta,\tau], C(I))$ be given by $\varphi \mapsto \varphi|_{[\delta,\tau]}$. Henceforth, we will equip $D((0,\infty), C(I))$ with the smallest topology for which each of the maps $\mathcal{L}_{\delta,\tau}$ is continuous. We will also equip $C((0,\infty), C(I))$ with the analogous topology.*

As a notational convention, $f_*\mu$ will denote the pushforward of a measure μ under a map f.

A few remarks about the topology on $D((0,\infty), C(I))$:

- Note that this topology is metrizable via $\sum_n 2^{-n}(1 \wedge \rho_n)$, where ρ_n is a metric inducing the topology of $D([n^{-1}, n], C(I))$.
- Note $\phi_n \stackrel{n\to\infty}{\longrightarrow} \phi$ in $D((0,\infty), C(I))$ if and only if $\phi_n|_{[\delta,\tau]} \stackrel{n\to\infty}{\longrightarrow} \phi|_{[\delta,\tau]}$ in $D([\delta,\tau], C(I))$, for every $0 < \delta \leq \tau$. Similarly, for a sequence of probability measures $\mathbb{P}_n$ on $D((0,\infty), C(I))$, we have that $\mathbb{P}_n \to \mathbb{P}$ weakly iff $(\mathcal{L}_{\delta,\tau})_*\mathbb{P}_n \to (\mathcal{L}_{\delta,\tau})_*\mathbb{P}$ weakly for all $0 < \delta \leq \tau$.
- This topology is an analogue of the topology of uniform convergence on compact sets, but for the Skorokhod Space.

Lemma 2.6.5. *Let $\mathbb{Q}^\epsilon$ denote the law of $\mathcal{Z}^\epsilon$ on $D((0,\infty), C(I))$. Then there exists a measure $\mathbb{Q}$ on $C((0,\infty), C(I))$ which is a limit point of the $\mathbb{Q}^\epsilon$ on $D((0,\infty), C(I))$.*

Proof. The basic idea is to use the Kolmogorov Extension Theorem in conjunction with the previous results. For $0 < \delta \leq \tau$, we will let $\mathbb{P}^\epsilon_{\delta,\tau}$ denote the law of $\mathcal{Z}^\epsilon$ on $D([\delta,\tau], C(I))$. For $0 < \delta' \leq \delta \leq \tau \leq \tau'$, we define $\mathcal{L}^{\delta',\tau'}_{\delta,\tau} : D([\delta',\tau'], C(I)) \to D([\delta,\tau], C(I))$ be the map $\varphi \mapsto \varphi|_{[\delta,\tau]}$.

We will use an inductive construction. Let $\mathbb{P}_1$ be a limit point of the $\{\mathbb{P}^\epsilon_{1,1}\}_{\epsilon\in(0,1]}$ on $D(\{1\}, C(I))$. Then by Corollary 2.6.3 we can find a sequence $\epsilon_j \downarrow 0$ such that $\mathbb{P}^{\epsilon_j}_{1,1} \to \mathbb{P}_1$ weakly. Suppose (for the inductive hypothesis) that for each $k \leq n-1$ we have constructed a probability measure $\mathbb{P}_k$ on $C([k^{-1}, k], C(I))$ and a sequence $(\epsilon^k_j)_{j=1}^\infty$ with the following two properties:

1. For each $1 \leq k \leq n-1$, the sequence $\mathbb{P}^{\epsilon^k_j}_{k^{-1},k}$ converges weakly to $\mathbb{P}_k$ as $j \to \infty$.

2. For each $2 \leq k \leq n-1$, $(\epsilon_j^k)_{j=1}^\infty$ is a subsequence of $(\epsilon_j^{k-1})_{j=1}^\infty$

By Corollary 2.6.3, the sequence $\{\mathbb{P}_{n^{-1},n}^{\epsilon_j^{n-1}}\}_{j\geq 1}$ is tight, therefore we can find a probability measure $\mathbb{P}_n$ on $C([n^{-1},n],C(I))$ and subsequence $(\epsilon_j^n)_{j=1}^\infty$ of $(\epsilon_j^{n-1})_j$ such that $\mathbb{P}_{n^{-1},n}^{\epsilon_j^n} \to \mathbb{P}_n$ weakly as $j\to\infty$. Therefore the inductive hypothesis holds true for $k=n$.

In the end, we obtain a sequence $\mathbb{P}_k$ of probability measures on $C([k^{-1},k],C(I))$ which is consistent in the sense that $(\mathcal{L}_{k^{-1},k}^{(k+1)^{-1},k+1})_*\mathbb{P}_{k+1} = \mathbb{P}_k$, for every k. By the Kolmogorov Extension Theorem, there exists a unique probability measure $\mathbb{Q}$ on $C((0,\infty),C(I))$ such that $(\mathcal{L}_{k^{-1},k})_*\mathbb{Q} = \mathbb{P}_k$ for all k.

To show that the measure $\mathbb{Q}$ is actually a limit point of the $\mathbb{Q}^\epsilon$, consider the sequences (ϵ_j^n) from before. Then the "diagonal" sequence $\mathbb{Q}^{\epsilon_n^n}$ converges weakly to $\mathbb{Q}$ as $n\to\infty$. Indeed ϵ_n^n is an eventual subsequence of ϵ_n^k for every k, and therefore $(\mathcal{L}_{k^{-1},k})_*\mathbb{Q}^{\epsilon_n^n} = \mathbb{P}_{k^{-1},k}^{\epsilon_n^n}$ converges weakly (as $n\to\infty$) to $(\mathcal{L}_{k^{-1},k})_*\mathbb{Q} = \mathbb{P}_k$ for every k. $\square$

Lemma 2.6.6. *Let $\mathbb{Q}$ be a limit point of the $\{\mathcal{Z}^\epsilon\}$ on $C((0,\infty),C(I))$, as constructed in the previous lemma, and let $\mathcal{L}_T : C((0,\infty),C(I)) \to C(I)$ denote the evaluation map at T. For any $p \geq 1$, there exists a constant $C = C(p)$ such that for $X \in I$ and $T \leq 1$,*

$$\|\mathcal{L}_T(X)\|_p^2 \leq CT^{-1/2}P_T(X,0)$$

$$\|\mathcal{L}_T(X) - P_T(X,0)\|_p^2 \leq CP_T(X,0)$$

where as usual, $\|F\|_p = \big(\int |F|^p d\mathbb{Q}\big)^{1/p}$.

Proof. Recall Equation (2.49) and the remark underneath, which say that

$$\|Z_t(x)\|_p^2 \leq C(\epsilon^2 t)^{-1/2}(\mathbf{p}_t^R * Z_0)(x) + C\sum_{k=0}^\infty \frac{C^k\epsilon^k t^{k/2}}{(k/2)!}(\mathbf{p}_{t+k}^R * Z_0)(x).$$

Rewriting the preceding expression in terms of the macroscopic variables, we have

$$\|\mathcal{Z}_T^\epsilon(X)\|_p^2 \le CT^{-1/2}(P_T^\epsilon *_\epsilon \mathcal{Z}_0^\epsilon)(X) + \sum_{k=0}^{\infty} \frac{C^k T^{k/2}}{(k/2)!}(P_{T+\epsilon^2 k}^\epsilon *_\epsilon \mathcal{Z}_0^\epsilon)(X) \tag{2.53}$$

where $(f *_\epsilon g)(X) := \epsilon \sum_{y\in\epsilon\Lambda} f(X,Y)g(Y)$, for any functions $f : (\epsilon\Lambda)^2 \to \mathbb{R}$ and $g : \epsilon\Lambda \to \mathbb{R}$.

We define the quantity $F(T, X, \epsilon)$ to be the RHS of (2.53). We are going to prove that

$$\limsup_{\epsilon\to 0} F(T, X, \epsilon) \le CT^{-1/2} P_T(X, 0)$$

where the P_T on the RHS is the continuum Robin heat kernel. As a first step, we claim that for any $k > 0$,

$$\lim_{\epsilon\to 0}(P_{T+\epsilon^2 k}^\epsilon *_\epsilon \mathcal{Z}_0^\epsilon)(X) = P_T(X, 0) \tag{2.54}$$

Indeed, we have that

$$\begin{aligned}
\varrho\epsilon^{1/2} \sum_{Y\in\epsilon\Lambda} P_{T+\epsilon^2 k}^\epsilon(X,Y)e^{-\epsilon^{-1/2}Y} &= \varrho\epsilon^{-1/2} \int_I P_{T+\epsilon^2 k}^\epsilon(X, \epsilon\lfloor\epsilon^{-1}Y\rfloor)e^{-\epsilon^{1/2}\lfloor\epsilon^{-1}Y\rfloor}dY \\
&= \varrho \int_I P_{T+\epsilon^2 k}^\epsilon(X, \epsilon\lfloor\epsilon^{-1/2}Z\rfloor)e^{-\epsilon^{1/2}\lfloor\epsilon^{-1/2}Z\rfloor}dZ \\
&\xrightarrow{\epsilon\to 0} \varrho \int_I P_T(X,0)e^{-Z}dZ \\
&= P_T(X,0).
\end{aligned} \tag{2.55}$$

In the first equality we used that $\epsilon\sum_{Y\in\epsilon\Lambda} f(Y) = \int_I f(\epsilon\lfloor\epsilon^{-1}Y\rfloor)dY$, for any function f. In the second equality we made the substitution $Z = \epsilon^{-1/2}Y$. In the third line we used uniform convergence (see Theorem 2.3.30) of $P_{T+\epsilon^2 k}^\epsilon(X,\cdot)$ to $P_T(X,\cdot)$ together with the fact that $\epsilon\lfloor\epsilon^{-1/2}Z\rfloor \to 0$ and $\epsilon^{1/2}\lfloor\epsilon^{-1/2}Z\rfloor \to Z$. In the final line we used the fact that $\varrho = \big(\int_I e^{-Z}dZ\big)^{-1}$, as defined earlier in this section. This proves (2.54).

In order to finish showing that $\lim_{\epsilon\to 0} F(T,X,\epsilon) \le CT^{-1/2}P_T(0,X)$, we will take the limit as $\epsilon \to 0$ on the RHS of Equation (2.53), and then pass the limit through the infinite sum and apply (2.54). However, we need to justify interchanging the $\lim_{\epsilon\to 0}$ with the infinite sum $\sum_{k=0}^{\infty}$. To justify this interchange, we recall Equation (2.50) which says (after passing to macroscopic variables) that

$$\sup_{\epsilon\in(0,1]} (P_T^\epsilon *_\epsilon \mathcal{Z}_0^\epsilon)(X) \le C(T^{-1/2}+1)e^{KT}$$

where C, K do **not** depend on the terminal time τ. Note that $T^{-1/2}+1 \le CT^{-1/2}e^{KT}$, hence the RHS of the last expression may be further bounded by $C'T^{-1/2}e^{2KT}$. Using this bound shows that

$$\begin{aligned}\sum_{k=0}^{\infty} \frac{C^k T^{k/2}}{(k/2)!} \sup_{\epsilon\in(0,1]} (P_{T+\epsilon^2 k}^\epsilon *_\epsilon \mathcal{Z}_0^\epsilon)(X) &\le CT^{-1/2}e^{2KT} \sum_{k=0}^{\infty} \frac{C^k T^{k/2}}{(k/2)!} e^{2Kk} \\ &\le CT^{-1/2}e^{K'T}. \end{aligned} \tag{2.56}$$

By (2.56) and the dominated convergence theorem, we can interchange the limit as $\epsilon \to 0$ with the infinite sum in (2.53), as discussed before. This completes the proof that

$$\limsup_{\epsilon\to 0} F(T,X,\epsilon) \le CT^{-1/2}P_T(X,0). \tag{2.57}$$

In order to prove the second bound, we note from Lemma 2.5.3 and (2.21) that for $t \ge 1$,

$$\big\| Z_t(x) - \mathbf{p}_t^R * Z_0(x) \big\|_p^2 \le C\epsilon \int_0^t (t-s)^{-1/2} \sum_y \mathbf{p}_{t-s+1}^R(x,y) \|Z_s(y)\|_p^2 ds.$$

In terms of macroscopic variables, this says that for $T \ge \epsilon^2$

$$\begin{aligned}\big\| \mathcal{Z}_T^\epsilon(X) - (P_T^\epsilon *_\epsilon \mathcal{Z}_0^\epsilon)(X) \big\|_p^2 &\le C \int_0^T (T-S)^{-1/2} \big(P_{T-S+\epsilon^2}^\epsilon *_\epsilon \|\mathcal{Z}_S\|_p^2 \big) dS \\ &\le C \int_0^T (T-S)^{-1/2} \big(P_{T-S+\epsilon^2}^\epsilon *_\epsilon F(S,\cdot,\epsilon) \big)(X) dS \end{aligned} \tag{2.58}$$

where we used (2.53) in the last line. Now using (2.56), it is easily shown that

$$S \;\mapsto\; (T-S)^{-1/2} \cdot \sup_{\epsilon \in (0,1]} \big(P^{\epsilon}_{T-S+\epsilon^2} *_{\epsilon} F(S,\cdot,\epsilon)\big)(X)$$

is a function which is integrable over $[0,T]$. Moreover, by (2.57) we have that

$$\limsup_{\epsilon \to 0} \big(P^{\epsilon}_{T-S+\epsilon^2} *_{\epsilon} F(S,\cdot,\epsilon)\big)(X) \le S^{-1/2} C\big(P_{T-S} * P_S(\cdot,0)\big)(X) = CS^{-1/2}P_T(X,0).$$

To finish the proof, let $\epsilon \to 0$ in (2.58) and apply the dominated convergence theorem to interchange the $\lim$ and the $\int_0^T$, and finally note that $\int_0^T (T-S)^{-1/2}S^{-1/2}dS$ is a constant not depending on T. $\square$

Lemma 2.6.7. *Let $\mathbb{Q}$ and $\mathcal{L}_T$ be as in the previous lemma. Then we can enlarge the probability space $\big(C((0,\infty),C(I)),\ \mathcal{B}orel,\ \mathbb{Q}\big)$ so as to admit a space-time white noise W such that $\mathbb{Q}$-almost surely, for every $0<S<T$:*

$$\mathcal{L}_T = P_{T-S} * \mathcal{L}_S + \int_S^T P_{T-U} * (\mathcal{L}_U dW_U).$$

Proof. Let us fix $S \ge 0$ for now. Define

$$Y_T^S(\varphi) = (\mathcal{L}_{T+S},\varphi) - (\mathcal{L}_S,\varphi) - \frac{1}{2}\int_S^{S+T} (\mathcal{L}_U,\varphi'')dU.$$

By Corollary 2.6.3, we know that $(\mathcal{L}_{S+T})_{T\ge 0}$ has the same distribution (under $\mathbb{Q}$) as the solution of the SHE started from $\mathcal{L}(S)_*\mathbb{Q}$. Therefore (see (2.31)), the law of $(\mathcal{L}_{S+T})_{T\ge 0}$ solves the martingale problem for the SHE, as defined in 2.5.5. Just as in the proof of Proposition 2.5.6, this shows that Y^S is an orthogonal martingale measure. So we let $\overline{W}$ be an independent white noise on $[0,\infty)\times I$ (*not* depending on S), then define the following space-time process for $T \ge 0$:

$$W_T^S(\varphi) := \int_0^T \int_I \mathcal{L}_{S+U}(X)^{-1}\varphi(X)1_{\{L_{S+U}(X)\neq 0\}}Y^S(dX\ dU) + \int_S^{S+T}\int_I 1_{\{\mathcal{L}_U(X)=0\}}\varphi(X)\overline{W}(dX\ dU).$$

Just as in the proof of proposition 2.5.6, W^S is a cylindrial Wiener process such that $\mathbb{Q}$-almost surely, for any $T \geq 0$ we have that

$$\mathcal{L}_{S+T} = P_T * \mathcal{L}_S + \int_0^T P_{T-U} * (\mathcal{L}_U dW_U^S). \tag{2.59}$$

This construction actually shows that the white noises W^S are consistent, in the sense that for any $S_1 < S_2$ we have that $W_T^{S_2} = W^{S_1}_{S_2-S_1+T} - W^{S_1}_{S_2-S_1}$ (because the Y^S satisfy the same relation).

Finally, if $T > 0$, then we define

$$W_T := W^1_{(T-1)\vee 0} + \sum_{k=1}^{\infty} W^{2^{-k}}_{2^{-k}\wedge(T-2^{-k})\vee 0}.$$

Using the property that that $W_T^{S_2} = W^{S_1}_{S_2-S_1+T} - W^{S_1}_{S_2-S_1}$ for $S_1 < S_2$, it follows that the terms appearing in the infinite sum are independent. Hence the convergence of the infinite series may be checked by the martingale convergence theorem. The noise W may be equivalently described as the a.s. and L^2 limit of the W^S as $S \to 0$. One may then directly check that W is a cylindrical Wiener process with the property that $W_T^S = W_{S+T} - W_S$. This, together with (2.59), shows that the desired relation holds. $\square$

Theorem 2.6.8 (Main Result of this Section)**.** *The rescaled processes $\mathcal{Z}^\epsilon$ converge weakly in $D((0,\infty), C(I))$ to the solution of the SHE with δ_0 initial data (as constructed in Proposition 2.4.3).*

Proof. Let $\mathbb{Q}$ be *any* limit point of the laws of the $\{\mathcal{Z}^\epsilon\}$ on $C((0,\infty), C(I))$. By Lemma 2.6.5, we know that at least one such $\mathbb{Q}$ exists.

As usual, we let $\mathcal{L}_T : C((0,\infty), C(I)) \to C(I)$ denote the canonical T-coordinate, and we will let $\|F\|_p := \big(\int |F|^p d\mathbb{Q}\big)^{1/p}$. By Lemma 2.6.7, after possibly extending the probability space, there exists a white noise W such that $\mathbb{Q}$-a.s., for any $0 < S < T$ we have that

$$\mathcal{L}_T = P_{T-S} * \mathcal{L}_S + \int_S^T P_{T-U} * (\mathcal{L}_U dW_U).$$

Therefore, if we can prove that $\int_0^T P_{T-U} * (\mathcal{L}_U dW_U)$ exists and is in $L^2(\mathbb{Q})$ for any $T > 0$, then we would have that

$$\begin{aligned} &\left\| \mathcal{L}_T(X) - P_T(X,0) - \int_0^T P_{T-U} * (\mathcal{L}_U dW_U)\big|_X \right\|_2 \\ &\leq \big\| P_{T-S} * \mathcal{L}_S(X) - P_T(X,0) \big\|_2 + \left\| \int_0^S P_{T-U} * (\mathcal{L}_U dW_U)\big|_X \right\|_2. \end{aligned} \tag{2.60}$$

Notice that the top expression does not depend on S, whereas the bottom one does. So if we can prove that (for any fixed T) both of the terms in (2.60) tend to 0 as $S \to 0$, then it will immediately follow that $\mathbb{Q}$ is the distribution of the solution of the SHE started from δ_0. Thus the remainder of the proof will be dedicated to this purpose.

For the remainder of this proof, we will fix some time $T > 0$, and we will always consider $S < 1 \wedge T$.

Let us start with the first term in (2.60). By the semigroup property and Minkowski's inequality, we see that

$$\begin{aligned} \big\| P_{T-S} * \mathcal{L}_S(X) - P_T(X,0) \big\|_p &= \left\| \int_I P_{T-S}(X,Y) \big[\mathcal{L}_S(Y) - P_S(Y,0) \big] dY \right\|_p \\ &\leq \int_I P_{T-S}(X,Y) \big\| \mathcal{L}_S(Y) - P_S(Y,0) \big\|_p dY \\ &\leq C \int_I P_{T-S}(X,Y) \cdot P_S(Y,0)^{1/2} dY \end{aligned}$$

where we used Lemma 2.6.6 in the final line. By applying Proposition 2.3.31 with $b = 0$, we see that $P_{T-S}(X,Y) \leq C(T-S)^{-1/2}$. Similarly, applying Proposition 2.3.31 with $b = 2$, we see that

$P_S(Y,0) \le CS^{-1/2}e^{-2Y/\sqrt{S}}$. Therefore, continuing the above chain of inequalities, we find that

$$\begin{aligned}\int_I P_{T-S}(X,Y)\cdot P_S(Y,0)^{1/2}dY &\le C(T-S)^{-1/2}S^{-1/4}\int_I e^{-Y/\sqrt{S}}dY \\ &\le C(T-S)^{-1/2}S^{-1/4}\int_0^\infty e^{-Z}(S^{1/2}dZ) \\ &= C(T-S)^{-1/2}S^{1/4}\end{aligned}$$

where we made the substitution $Z = S^{-1/2}Y$ in the second line. Since T was assumed to be fixed, and since $(T-S)^{-1/2} \le \sqrt{2}T^{-1/2}$ whenever $S < T/2$, this computation shows that

$$\limsup_{S\to 0} \big\|P_{T-S} * \mathcal{L}_S(X) - P_T(X,0)\big\|_p \lesssim \lim_{S\to 0} S^{1/4} = 0.$$

This shows that the first term appearing on the RHS of (2.60) approaches 0 as $S \to 0$.

Next, we are going to consider the second term appearing on the RHS of (2.60). As stated above, we first need to prove existence of the stochastic integral $\int_0^T P_{T-U} * (\mathcal{L}_U dW_U)$. This amounts to showing that

$$\int_0^T \int_I P_{T-U}(X,Y)^2 \mathbb{E}[\mathcal{L}_U(Y)^2]dY\,dU < \infty$$

where the expectation is with respect to $\mathbb{Q}$. Using Lemma (2.6.6), it holds that $\mathbb{E}[\mathcal{L}_U(Y)^2] \le CU^{-1/2}P_U(Y,0)$. Using the first bound in Proposition 2.3.31 with $b = 0$, we also have that $P_{T-U}(X,Y)^2 \le C(T-U)^{-1/2}P_{T-U}(X,Y)$. Therefore,

$$\begin{aligned}&\int_0^T \int_I P_{T-U}(X,Y)^2\mathbb{E}[\mathcal{L}_U(Y)^2]dY\,dU \\ &\qquad \le C\int_0^T (T-U)^{-1/2}U^{-1/2}\bigg[\int_I P_{T-U}(X,Y)P_U(Y,0)dY\bigg]dU \\ &\qquad \le C\int_0^T (T-U)^{-1/2}U^{-1/2}dU\cdot P_T(X,0) \\ &\qquad = CP_T(X,0) < \infty.\end{aligned}$$

In the second line, we used the semigroup property of the P_T, and in the third line we used the fact that $\int_0^T (T-U)^{-1/2} U^{-1/2} dU$ is a constant not depending on T, which can be proved by making the substitution $V = U/T$. This proves that the stochastic integral $\int_0^T P_{T-U} * (\mathcal{L}_U dW_U)$ exists. Now we just need to show that the second term in (2.60) tends to 0 as $S \to 0$. Applying the Itô isometry and then using the same exact arguments as above, we see that

$$\begin{aligned}\left\| \int_0^S P_{T-U} * (\mathcal{L}_U dW_U)\big|_X \right\|_2^2 &= \int_0^S \int_I P_{T-U}(X,Y)^2 \mathbb{E}[\mathcal{L}_U(Y)^2] dY dU \\ &\leq C \int_0^S (T-U)^{-1/2} U^{-1/2} dU \cdot P_T(X,0).\end{aligned}$$

Clearly $P_T(X,0)$ does not depend on S, and as long as $U < S < T/2$ we have $(T-U)^{-1/2} \leq (T-S)^{-1/2} \leq (T/2)^{-1/2}$, which no longer depends on U or S. Therefore the last integral is $O(S^{1/2})$ as $S \to 0$, which completes the proof. $\square$

Chapter 3: Positive random walks and an identity for half-space SPDE's

The purpose of this chapter is to investigate the continuum limit of a distributional identity for half-space directed polymers given in [14]. The limiting identity turns out to relate the multiplicative-noise half-space stochastic heat equation with Dirichlet boundary condition to the same equation with Robin boundary condition. We view this identity as a precursor for proving Gaussian fluctuation behavior of the supercritical half-space KPZ equation at the origin.

3.1 Introduction and context

We will focus on three related objects: uniform measures on collections of nearest-neighbor *non-negative* paths (e.g., Brownian meander), directed polymers weighted by such measures, and multiplicative-noise stochastic partial differential equations (SPDE) in a half-space.

3.1.1 Half-space stochastic heat equations

We begin our discussion with SPDE's. The multiplicative-noise stochastic heat equation has been a frequent subject of research within stochastic analysis and mathematical physics in recent years. This equation arises naturally in the context of directed polymers and interacting particle systems, as a *weak scaling limit* [40]. In spatial dimension one, the multiplicative-noise stochastic heat equation is also related to the so-called KPZ equation via the Hopf-Cole transform, and may be solved by the classical Itô-Walsh construction [150] or by more modern techniques such as regularity structures [88]. In the present article, we consider the stochastic heat equation with multiplicative noise on a *half-line*:

$$\partial_T Z(T,X) = \frac{1}{2}\partial_X^2 Z(T,X) + Z(T,X)\cdot\xi(T,X), \qquad X \geq 0, T \geq 0, \tag{SHE}$$

where ξ is a Gaussian space-time white noise on $\mathbb{R}_+ \times \mathbb{R}_+$. Naturally one needs to impose boundary conditions at $X = 0$ in order to make sense of this equation. In the present work we consider two types of boundary conditions, Robin and Dirichlet. First let us write the Robin boundary condition of parameter $A \in \mathbb{R}$:

$$\partial_X Z(T, 0) = AZ(T, 0). \tag{3.1}$$

This type of homogeneous boundary condition has been considered in [45, 130, 74, 15] in the context of interacting particle systems, and a robust solution theory has been developed in [70] using techniques of [83]. This boundary condition transforms into a Neumann boundary condition for the half-space KPZ equation upon taking the logarithm. Next, we consider the Dirichlet boundary condition for the half-space SHE:

$$Z(T, 0) = 0. \tag{3.2}$$

This type of boundary condition was considered, for instance, in [82], in the context of directed polymers near an absorbing wall. Again, one can make sense of the equation using classical techniques of [150] or more modern ones such as [83]. Our main result compares these two types of boundary conditions; specifically it allows us to interchange information about the initial data with that of the boundary condition imposed on the SHE:

Theorem 3.1.1. *Fix $A \in \mathbb{R}$. Let $Z^{(A)}_{Rob}(T, X)$ denote the solution of* (SHE) *with Robin boundary parameter A as in* (3.1) *and delta initial data $Z^{(A)}_{Rob}(0, X) = \delta_0(X)$. Let $Z^{(A)}_{Dir}(T, X)$ be the solution to* (SHE) *with Dirichlet boundary condition* (3.2) *and initial data $Z^{(A)}_{Dir}(0, X) = e^{B_X - (A + \frac{1}{2})X}$, where B is a standard Brownian motion independent of ξ. Then for each $T \geq 0$ we have the following equality of distributions:*

$$Z^{(A)}_{Rob}(T, 0) \stackrel{d}{=} \lim_{X \to 0} \frac{Z^{(A)}_{Dir}(T, X)}{X}. \tag{3.3}$$

This will be a consequence of Theorem 3.2.4, which in turn will use the work of [14, 151] and Theorem 3.2.2 (our main technical result) as inputs. Let us now discuss the motivation for this result, the contexts in which it has arisen, and the methods used to prove it.

To give some motivation towards (3.3), we now explain it using the exact solvability framework developed in [14], which is a crucial input to the proof of (3.3). Both the left and right sides of (3.3) have interpretations in terms of partition functions of a certain family of probabilistic models known as *directed polymers* (see Section 1.2). Specifically, the left side of (3.3) can be related to a polymer that is modeled on a Brownian motion which gets reweighted according to its local time at zero, whereas the right side can be related to a polymer that is modeled on a Brownian motion conditioned to remain positive. In [14], the authors use certain nontrivial symmetries of Macdonald polynomials in order to obtain information about the large-scale behavior of *discrete* versions of these polymer models and others (which is similar in theme to, and builds on, older works of [24, 43, 126, 136, 11]). One particular result in that paper (Proposition 8.1) is a highly non-obvious identity in distribution for directed polymers with log-gamma weights, that effectively allows one to switch some of the bulk weights of the random environment with those on the boundary without changing the distribution of the associated partition function. Our main goal was to take the SPDE limit of that identity, which effectively gives Theorem 3.1.1 under the appropriate scaling. Hence our result can be viewed as a special case of more general algebraic principles that may be used to extract certain nontrivial symmetries in certain half-space models.

The right side of (3.3) equals $(\partial_X Z^{(A)}_{Dir})(T,0)$. It is not clear why this derivative should even exist in the first place, since the spatial regularity of Z_{Dir} is much worse than C^1. One of our main technical results, given in Section 4, is that the limit in (3.3) is indeed well-defined (Corollary 3.4.3). In fact we will prove something stronger: the limit in the right side of (3.3) simultaneously exists for all $T \geq 0$ almost surely, and is Hölder 1/4$-$ as a function of T almost surely.

In order to convince the reader that (3.3) is at least plausible, let us verify formally that the expectations are the same on both sides of the equation. Let $P^{(A)}_{Rob}(T;X,Y)$ denote the Robin boundary heat kernel and let $P_{Dir}(T;X,Y)$ denote the Dirichlet boundary one, where by *heat kernel* we mean the fundamental solution of the heat equation with the associated boundary condition started from the delta measure at point X. Letting $P(T;X) = \frac{1}{\sqrt{2\pi T}} e^{-X^2/2T}$, one may verify directly that these kernels are given by the following explicit formulas for $T, X, Y \geq 0$:

$$
\begin{aligned}
P^{(A)}_{Rob}(T;X,Y) &= P(T;X+Y) + P(T;X-Y) - 2A\int_0^\infty P(T;X+Y+Z)e^{-AZ}dZ, \\
P_{Dir}(T;X,Y) &= \lim_{A\to\infty} P^{(A)}_{Rob}(T;X,Y) = P(T;X-Y) - P(T;X+Y).
\end{aligned}
$$

By the Duhamel principle (see Definition 2.1) it holds that $\mathbb{E}[Z^{(A)}_{Rob}(T,X)] = P^{(A)}_{Rob}(T;0,X)$ and $\mathbb{E}[Z^{(A)}_{Dir}(T,X)] = \mathbb{E}[\int_0^\infty P_{Dir}(T;X,Y)e^{B_Y-(A+1/2)Y}dY] = \int_0^\infty P_{Dir}(T;X,Y)e^{-AY}dY$. One then formally interchanges an expectation and a derivative to obtain

$$
\begin{aligned}
\mathbb{E}[\partial_X\big|_{X=0} Z^{(A)}_{Dir}(T,X)] &= \int_0^\infty \partial_X\big|_{X=0} P_{Dir}(T;X,Y)e^{-AY}dY \\
&= 2\int_0^\infty \partial_Y P_{Dir}(T;0,Y)e^{-AY}dY = P^{(A)}_{Rob}(T;0,X) = \mathbb{E}[Z^{(A)}_{Rob}(T,0)],
\end{aligned}
$$

where we integrate by parts in the third equality. This shows at a purely formal level that the expectations on either side of (3.3) are the same.

Theorem 3.1.1 suggests a duality between the initial data of a solution to the half-space SHE and the boundary conditions one imposes on it. It may be interesting to see if more general versions of this hold. For example, could it be possible that the identity holds as a *process* in T and not just in the one-point sense? Using this type of idea, one may potentially obtain useful information about objects of interest, such as the Neumann-boundary Kardar-Parisi-Zhang (KPZ) equation that was considered in [45]. It was conjectured in Chapter 2 that one has the almost-sure convergence

$$
\lim_{T\to\infty} \frac{1}{T}\log Z^{(A)}_{Rob}(T,0) = \begin{cases} -\frac{1}{24}, & A \geq -1/2 \\ (A+1/2)^2 - \frac{1}{24}, & A \leq -1/2, \end{cases}
$$

which would give the exact law of large numbers for Neumann-boundary KPZ. Unfortunately Theorem 3.1.1 alone is not enough to obtain this result. Nevertheless, it is plausible and even hopeful that a clever use of (3.3) (perhaps combined with some new ideas and techniques) could lead to quantitative results that are close to the above expression. Indeed, despite the fact that on the Robin side of (3.3) there is no *visible* phase transition at $A = -1/2$, the appearance of the term $A + 1/2$ on the *Dirichlet* side already indicates the presence of a nontrivial change in large-scale behavior at $A = -1/2$. The introduction to Section 2 includes a further discussion of this. More than just computing the above limit, we are also interested in computing the limiting distribution of the fluctuations around the mean value. These should be of order $T^{1/2}$ and Gaussian in the $A < -1/2$ case, and they should be of order $T^{1/3}$ and random-matrix theoretic otherwise (with separate cases when $A = -1/2$ and $A > -1/2$). See for instance [130, 15, 10, 23].

The main technical difficulties in the present work are of an analytic nature: translating the discrete identity in [14] to that of (3.3) required us to prove a general convergence result for directed polymers, stated below as Theorem 3.1.2. As we will now see, this involves the analysis of an interesting object in its own right: the Brownian meander.

3.1.2 Directed polymers weighted by positive random walks

This brings us to the method of proof of Theorem 3.1.1. As suggested above, it will be proved using an approximation via *directed polymers* with very specific weights, where a discrete version of this identity holds.

Directed polymers are natural probabilistic objects that were first introduced in [94, 96]. They generalize directed first- and last-passage percolation and have deep connections to statistical mechanics and stochastic analysis. Specifically, we consider an environment $\{\omega_{i,j}\}_{(i,j)\in\mathbb{Z}_{\geq 0}\times\mathbb{Z}}$ consisting of i.i.d., mean-zero, finite-variance random variables. The standard deviation of the weights is referred to as the *inverse temperature*. One may define a partition function $Z^\omega(n,x)$ as a sum over all *directed* nearest-neighbor simple random walk paths $(i, \gamma_i)_{0 \le i \le n}$ of length n starting from $(0, x)$, of the product of all weights $e^{\omega_{i,\gamma_i}}$ along the path. Similarly, there is also a natural way to define *random* Markovian transition densities associated to this environment ω, wherein a nearest-neighbor path γ has probability proportional to the product of weights $e^{\omega_{i,\gamma_i}}$ along it. As is standard practice in statistical mechanics, one may then ask questions about the existence of infinite-volume limits of these path measures and their typical fluctuation scale, as well as the typical scale and shape of the fluctuations of the partition function itself [37].

Many seminal results in these directions have been proved, perhaps most notably that there is a phase transition which becomes apparent in high dimensions. Specifically, in spatial dimensions greater than two, there is a strictly positive critical value of the inverse temperature below which weak disorder holds, meaning that the fluctuations of a typical polymer path look like Brownian motion and one may construct infinite-length path measures [39, 37]. Such polymers are said to exhibit *weak disorder.* In contrast, lower-dimensional polymers at any nonzero inverse temperature are known to be characterized by *strong disorder*, meaning that the path fluctuations are quite different and there is no sensible notion of an infinite volume Gibbs measure [37]. The results of [3, 2] examined the partition function in a regime that lies *between* strong and weak disorder. Specifically, in spatial dimension one, they scaled the inverse temperature of the model like $n^{-1/4}$ and simultaneously applied diffusive scaling to the partition function, and there they observed that the fluctuations are governed by (SHE) and that the path measures themselves have a continuum analogue. Recent work of [34, 32] has investigated the intermediate-disorder behavior in two spatial dimensions, where the scaling $n^{-1/4}$ is replaced by $(\log n)^{-1/2}$. In a different direction, [151] extended the work of [3] to the case of half-space polymers with Robin boundary condition.

We will be interested in the analogous half-space question of intermediate-disorder fluctuations of the directed polymer partition function associated to *uniform non-negative path measures*. Specifically, let

- $\mathbf{P}^n_x$ denote the uniform probability measure on the collection of all paths $(\gamma_i)_{0\le i\le n}$ such that $\gamma_0 = x$, $|\gamma_{i+1}-\gamma_i| = 1$ for $i<n$, and $\gamma_i \ge 0$ for all $i \le n$.

- $\omega_{i,j}$ be i.i.d. mean-zero, variance-one random variables that are uniformly bounded from below by a deterministic constant.

- f_n be a sequence of functions bounded uniformly by a function growing at-worst exponentially fast near infinity such that $f_n(n^{1/2}\,\cdot)$ converges (as $n\to\infty$) to some function $f(\cdot)$ in the Hölder space $C^{\alpha}_{loc}(\mathbb{R}_+)$, for all $\alpha\in(0,1/2)$.

Letting $\mathbf{E}^n_x$ denote the expectation with respect to $\mathbf{P}^n_x$, and setting S to be the canonical process associated to $\mathbf{P}^n_x$, one defines a directed-polymer partition function as follows:

$$Z^{\omega}_k(n,x) := \mathbf{E}^n_x\bigg[f_k(S_n)\prod_{i=0}^{n}(1+k^{-1/4}\omega_{i,S_i})\bigg].$$

Note that the expectation is taken only with respect to the random walk, *conditional* on the environment $\omega_{i,j}$, which is always assumed to be independent of the walk. We consider the rescaled partition function

$$\mathscr{Z}_n(T,X) := Z^{\omega}_n(nT, n^{1/2}X), \tag{3.4}$$

where the quantity on the right side is defined by linear interpolation between points of the lattice $L := \{(x,n)\in\mathbb{Z}^2_{\ge 0} : n-x\in 2\mathbb{Z}\}$.

In a manner analogous to [3] we show that $\mathscr{Z}_n$ converges in law to a random continuous space-time field. The natural candidate for such a limit would be a continuum analogue of $Z^{\omega}_k(n,x)$, where the expectation $\mathbf{E}^n_x$ over positive *discrete* random walks is replaced by that of *continuous* ones. Indeed the limiting space-time field can be described as follows: it has the formal Feynman-Kac interpretation that takes as its input the so-called *Brownian meander* [66, 65] on a finite time interval, and exponentially weighs it by its integral against a space-time white noise field. More precisely, if $\mathscr{P}^T_t(X,Y)$ denotes the inhomogeneous Markov transition density at time t of Brownian motion started from X and conditioned to stay positive until time $T \geq t$, then this limiting space-time field $\mathscr{Z}$ necessarily solves the multiplicative-noise SPDE on the half-space that is given in Duhamel form by

$$\mathscr{Z}(T,X) = \int_{\mathbb{R}^+} \mathscr{P}^T_T(X,Y) f(Y) dY + \int_0^T \int_{\mathbb{R}^+} \mathscr{P}^T_{T-S}(X,Y) \mathscr{Z}(S,Y) \xi(dY dS), \tag{3.5}$$

where ξ is a space-time white noise and f is the limiting function from the third bullet point above. An important step towards proving Theorem 3.1.1 will be to show that a solution of (3.5) exists and makes sense even when $X = 0$, and then to show that it can in turn be related to the derivative of the solution of the Dirichlet-boundary SHE at the origin. This will all be done in Section 4; more specifically we will show that the solution of (3.5) equals

$$\mathscr{Z}(T,X) = \frac{Z_{Dir}(T,X)}{2\Phi(X/\sqrt{T}) - 1}, \qquad\qquad T, X > 0, \tag{3.6}$$

where Z_{Dir} solves (SHE) with Dirichlet boundary condition (3.2) with the same initial data as $\mathscr{Z}$, and Φ is the cdf of a standard normal variable so that $\mathscr{Z}(T,0) = (2\pi T)^{1/2} \lim_{X \to 0} \frac{Z_{Dir}(T,X)}{X}$. We then have the following result.

Theorem 3.1.2. *The sequence of processes $\mathscr{Z}_n$ defined in* (3.4) *converge in law to the solution of* (3.5) *as $n \to \infty$. The convergence occurs in the sense of finite-dimensional distributions. If we assume that the $\omega_{i,j}$ have $p > 8$ moments, then distributional convergence holds when the space $C(\mathbb{R}_+ \times \mathbb{R}_+)$ is equipped with the topology of uniform convergence on compact sets.*

This theorem will be proved in Section 5.2 in greater generality (where the distribution of the weights ω may vary with n), see Proposition 3.5.9 and Theorem 3.5.11. It is actually a simplified version of Theorem 3.2.2 which is the true input to proving Theorem 3.1.1. The main difficulty towards this result will be in obtaining the necessary estimates for the inhomogeneous transition densities (and their discrete analogues) appearing in (3.5).

Thus the proof of Theorem 3.1.2 will lead to some new technical results related to the uniform measures $\mathbf{P}_x^n$ and their continuum analogues. These will be collected in appendices at the end of the paper. To illustrate a few such results, we will prove a coupling result for such random walks in the nearest-neighbor case, and then we will use that coupling to show the following concentration property: there exist constants $c, C > 0$ (independent of $n, x \geq 0$) such that for all $u > 0$ and all $k \leq n$ one has that

$$\mathbf{P}_x^n\big(\sup_{0 \leq i \leq k} |S_i - x| > u \big) \leq Ce^{-cu^2/k}.$$

We remind the reader that S_i is the *conditioned* walk. The study of such random walks started with the invariance principle of [95], further generalized in [19]. Later, the study expanded considerably, with local limit theorems [29] and expansions to heavy-tailed increments [30]. We will see that some of the estimates we derive are similar in spirit to some of those works, but the intricate details are somewhat different. We will give proofs of many of these technical results because the highly specific estimates needed to prove Theorem 3.1.2 were not found in those references (since our random walk does not necessarily start at zero).

It should be noted that we work with a simplified version of the partition function as opposed to much of the previous literature: [3, 38] and related works. There the partition function $Z_k^\omega(n, x)$ is defined with weights $e^{k^{-1/4}\omega_{i,S_i}}$ instead of the quantity $1 + k^{-1/4}\omega_{i,S_i}$ that we have used in (3.4) above. The reason for this is that the latter object is mathematically simpler because it is already renormalized (has expectation exactly 1 rather than approximately 1), and hence leads to simpler proofs and less stringent moment restrictions. However, it should be noted that the exponential version is more natural from the physical point of view, and entire results such as [60] have been devoted to finding the correct renormalization and phase transition behavior for that version as a function of the moment assumptions.

Outline: In Section 2, we prove Theorem 3.1.1 as Theorem 3.2.4, which uses [14] and [151] as important inputs. In Section 3, we will introduce and state some estimates about the transition densities associated to positive random walks, though the proofs are postponed to the appendices. In Section 4, we will develop the existence and uniqueness theory of the limiting SPDE (3.5) from Theorem 3.1.2, and as a corollary we prove that $\partial_X Z_{Dir}(T,0)$ exists. In Section 5, we prove Theorem 3.1.2 by using the estimates developed in the appendices. In the appendices we derive some elementary but powerful bounds related to the measures $\mathbf{P}^n_x$, which are crucial for the proofs in the main body.

Acknowledgements: Foremost, the author is tremendously grateful to two anonymous referees who gave extremely diligent reports, recommendations, and errata relating to the original article. We thank Ivan Corwin for suggesting that interesting identities could potentially be obtained from results in [14], and also for reading portions of the preliminary draft. We also thank Yier Lin for pointing out a mistake in some of the estimates. The author was partially supported by the Fernholz Foundation's "Summer Minerva Fellows" program, as well as summer support from Ivan Corwin's NSF grant DMS:1811143.

3.2 Main results

In this section, we show how to prove Theorem 3.1.1. We denote non-negative reals as $\mathbb{R}_+$ and non-negative integers as $\mathbb{Z}_{\geq 0}$.

We will use the notion of mild solutions for SPDEs throughout this article. Thus for completeness, we begin by giving the formal definition of such a solution, although it is peripheral to the main goals of the section.

Definition 3.2.1 (Mild Solution)**.** *Recall the Dirichlet-boundary heat kernel*

$$P_t^{Dir}(X,Y) := \frac{1}{\sqrt{2\pi t}}\big(e^{-(X-Y)^2/2t} - e^{-(X+Y)^2/2t}\big). \tag{3.7}$$

Let ξ be a space-time white noise defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let μ be an indepen-

dent random Borel measure on $\mathbb{R}_+$. A continuous space-time process $Z_{Dir} = (Z_{Dir}(T, X))_{T,X\geq 0}$ is a mild solution of the Dirichlet-boundary SHE with initial data μ if $\mathbb{P}$-almost surely, for all $X, T \geq 0$ one has that

$$Z_{Dir}(T, X) = \int_{\mathbb{R}_+} P_T^{Dir}(X, Y)\mu(dY) + \int_0^T \int_{\mathbb{R}_+} P_{T-S}^{Dir}(X, Y) Z_{Dir}(S, Y)\xi(dS, dY),$$

where the integral against ξ is meant to be interpreted in the Itô-Walsh sense [150].

The fact that this object exists will be established as a special case of the results in Section 4. The definition of the Robin boundary version $Z_{Rob}^{(A)}$ of (SHE) is very similar, but one replaces the Dirichlet heat kernel with the Robin boundary one throughout. We refer the reader to Section 2.4 above for more details, including the existence/uniqueness of this Robin boundary version.

The proof of Theorem 3.1.1 will be obtained by approximating both $Z_{Dir}^{(A)}$ and $Z_{Rob}^{(A)}$ by the partition function of a directed polymer with log-gamma weights. For these weights we use a known identity that allows us to switch the boundary weights with those on the initial data without changing the distribution of the partition function along the boundary [14] (Proposition 8.1). The approximation argument will strongly emulate the arguments given in [151, 3] although there are new challenges that make the convergence result rather difficult and technical. These additional difficulties are a byproduct of the inhomogeneous Markov transition densities for random walks conditioned to stay above zero.

Let us explicitly state the Dirichlet-boundary approximation result now. For each $n \in \mathbb{N}$, let $\omega^n = \{\omega_{i,j}^n\}_{i\geq j\geq 0}$ denote a random environment indexed by the principal octant of $\mathbb{Z}^2$ with the following properties:

- The "bulk-environment" random variables $\{\omega_{i,j}^n\}_{i\geq j\geq 1}$ are i.i.d., and the "lower-boundary" random variables $\{\omega_{i,0}^n\}_{i\geq 0}$ are also i.i.d. These two collections are independent.
- For $j > 0$ (the bulk variables) $\omega_{i,j}^n$ have finite second moment. Furthermore one has $\mathbb{E}[\omega_{i,j}^n] =$

0 and $\mathbb{E}[(\omega_{i,j}^n)^2] = 1 + o(1)$ as $n \to \infty$.

- For $j = 0$ (at the lower boundary) $\log(1 + n^{-1/4}\omega_{i,j}^n)$ has finite second moment; moreover there exist $\mu, \sigma \in \mathbb{R}$ such that $\mathbb{E}[\omega_{i,j}^n] = \mu n^{-1/4} + o(n^{-1/4})$, and $var(\omega_{i,j}^n) = \sigma^2 + o(1)$ as $n \to \infty$. We also assume $\omega_{i,0}$ have $2 + \epsilon$ moments for some $\epsilon > 0$.

An upright path in $\mathbb{Z}^2$ is a function $\gamma = (\gamma_1, \gamma_2) : \{0, ..., n\} \to \mathbb{Z}^2$ such that both γ_1 and γ_2 are weakly increasing, and $\gamma_1(i) + \gamma_2(i) - i$ is constant in i. For $p \geq q \geq 0$ define the random partition function

$$Z_n(p, q) := \sum_{\gamma:(0,0)\to(p,q)} 2^{-\#\{i \leq p+q \,:\, \gamma_2(i) \neq 0\}} \prod_{i=0}^{p+q} (1 + n^{-1/4}\omega^n_{\gamma_1(i),\gamma_2(i)}),$$

where the sum is taken over all upright paths γ from $(0, 0)$ to (p, q) that stay in the octant $\{(i, j) : i \geq j \geq 0\}$. As a convention, we also set $Z_n(p, q) = Z_n(p, 0)$ for $q \leq 0$. Let Φ denote the cdf of a standard normal distribution. We define the rescaled processes

$$\mathscr{Z}_n(T, X) := \frac{1}{2\Phi\big(\frac{X+n^{-1/2}}{\sqrt{T}}\big) - 1} \cdot Z_n(nT + n^{1/2}X, nT - n^{1/2}X), \qquad T, X \geq 0$$

where we interpolate linearly between integer values of Z_n.

The following result is the primary technical contribution of this work.

Theorem 3.2.2. *In the above notations and assumptions, the sequence of processes $\mathscr{Z}_n$ converges in distribution (in the sense of finite-dimensional marginals, as $n \to \infty$) to the unique space-time process satisfying* (3.5) *(equivalently given by* (3.6)*) with initial data $\mathscr{Z}(0, X) = Z_{Dir}(0, X) = e^{\sigma B_X + (\mu - \frac{1}{2}\sigma^2)X}$, where B is a standard Brownian motion independent of the space-time white noise ξ. If we assume that all weights $\omega_{i,j}^n$ have more than eight moments bounded independently of n, then distributional convergence holds when the space $C(\mathbb{R}_+ \times \mathbb{R}_+)$ is equipped with the topology of uniform convergence on compact sets.*

We will see that Theorem 3.2.2 is essentially equivalent to a more complicated version of Theorem 3.1.2, where the distribution of the weights ω depends on n and the domain of the polymer paths has been changed from a quadrant to an octant of $\mathbb{Z}^2$, which makes the geometry more challenging to work with. Accordingly, the proof of this theorem will proceed in two steps: first by reducing the claim of the theorem to that of Theorem 3.1.2 with a specific initial data (which will be achieved in Section 5.1), and then proving Theorem 3.1.2 which is simpler thanks to known methods and is done in Section 5.2.

Remark 3.2.3. *There are really two different regimes in which one should interpret Theorem 3.2.2. One regime is $X > 0$, where the result merely says that $Z_n(nT + n^{1/2}X, nT)$ converges to $Z_{Dir}(T, X)$. The more interesting regime is $X = 0$, in which case the theorem says that $(\pi nT/2)^{1/2} Z_n(nT, nT)$ converges in law to $\lim_{X\to 0} \frac{Z_{Dir}(T,X)}{2\Phi(X/\sqrt{T})-1}$, i.e.,*

$$n^{1/2} Z_n(nT, nT) \xrightarrow{d} \lim_{X\to 0} \frac{Z_{Dir}(T,X)}{X}.$$

An advantage of our approach is that the proof will simultaneously cover both regimes. In fact, we will see that convergence even takes place in a parabolic Hölder space of the appropriate regularity provided that the weights have more than eight moments.

We now combine this result with the Robin boundary result of [151] and the log-gamma identities of [14] in order to obtain the following result, which clearly implies Theorem 3.1.1. In what follows, we denote by $\Gamma^{-1}(\theta, c)$ the inverse-gamma distribution of shape parameter θ and scale parameter c, i.e., the law of the random variable cX, where X has pdf given by

$$f(x) = \frac{x^{-\theta-1}}{\Gamma(\theta)} e^{-1/x}, \qquad x > 0.$$

We will also write $\mathbb{E}[\Gamma^{-1}(\theta, c)] = \frac{c}{\theta-1}$ and $var(\Gamma^{-1}(\theta, c)) = \frac{c^2}{(\theta-1)^2(\theta-2)}$ to denote respectively the expectation and variance of such a random variable.

For $n \in \mathbb{N}$, let $\zeta_n^1 = \{\zeta_n^1(i,j)\}_{i\geq j\geq 0}$ and $\zeta_n^2 = \{\zeta_n^2(i,j)\}_{i\geq j\geq 0}$ be fields of independent random variables with the following distributions

$$\zeta_n^1(i,j) \sim \begin{cases} \Gamma^{-1}(2\sqrt{n},\ \frac{1}{2}\mathbb{E}[\Gamma^{-1}(2\sqrt{n},1)]^{-1}), & i \neq j \\ \\ \Gamma^{-1}(\sqrt{n}+A+\frac{1}{2},\ \frac{1}{2}\mathbb{E}[\Gamma^{-1}(2\sqrt{n},1)]^{-1}), & i = j \end{cases}$$

$$\zeta_n^2(i,j) \sim \begin{cases} \Gamma^{-1}(2\sqrt{n},\ \frac{1}{2}\mathbb{E}[\Gamma^{-1}(2\sqrt{n},1)]^{-1}), & j \neq 0 \\ \\ \Gamma^{-1}(\sqrt{n}+A+\frac{1}{2},\ \frac{1}{2}\mathbb{E}[\Gamma^{-1}(2\sqrt{n},1)]^{-1}), & j = 0. \end{cases}$$

Let Z_n^1 and Z_n^2 denote the associated partition functions, i.e.,

$$Z_n^\alpha := \sum_{\gamma:(0,0)\to(\lfloor nT\rfloor,\lfloor nT\rfloor)} \prod_{i=0}^{2\lfloor nT\rfloor} \zeta_n^\alpha(\gamma_1(i),\gamma_2(i)), \qquad \text{for } \alpha\in\{1,2\}. \tag{3.8}$$

Here the sum is taken over all upright paths γ from $(0,0)$ to $(\lfloor nT\rfloor,\lfloor nT\rfloor)$ that stay in the octant $\{(i,j) : i \geq j \geq 0\}$.

Theorem 3.2.4 (Joint with [14, 151])**.** *With Z_n^1 and Z_n^2 defined in* (3.8)*, the following are true:*

1. $\sqrt{n}Z_n^1$ *converges in distribution as $n\to\infty$ to the left-hand side of* (3.3).
2. $\sqrt{n}Z_n^2$ *converges in distribution as $n\to\infty$ to the right-hand side of* (3.3).
3. *For every n, one has $Z_n^1 \overset{d}{=} Z_n^2$.*

Proof. Item *(1)* is proved as Theorem 5.1(B) of [151] using techniques from [3]. Item *(3)* is proved in Proposition 8.1 of [14] by developing the theory of half-space Macdonald processes. Thus we only need to prove Item *(2)*, and this will be done using Theorem 3.2.2, in the special case where $X = 0$. As in Theorem 4.5 of [3], we define a family of independent weights $\omega^n = \{\omega_{i,j}^n\}_{i\geq j\geq 0}$ according to the rule:

$$2\zeta_n^2(i,j) = 1 + (4n)^{-1/4}\omega_{i,j}^n, \qquad j > 0,$$

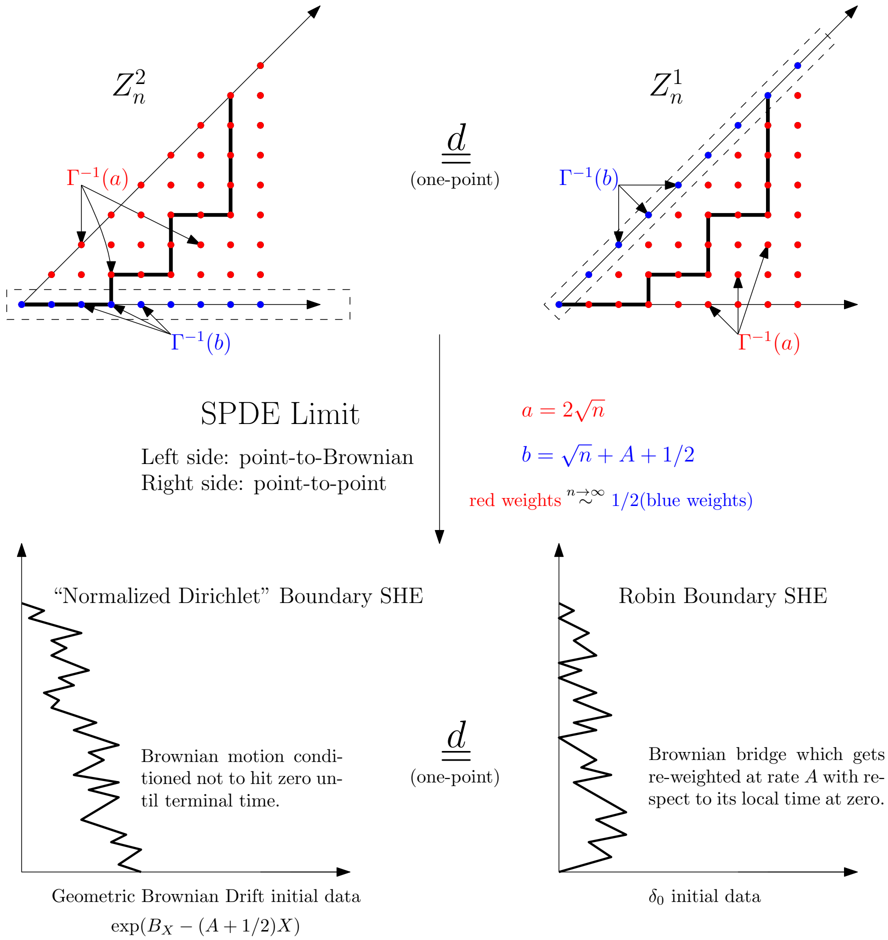

Figure 3.1: A graphical description of Theorem 3.2.4. The weight of a given path is the product of the weights along it, and the partition function Z_n^α for $\alpha \in \{1, 2\}$ is given by summing the weights of all upright paths from $(0, 0)$ to $(\lfloor nT \rfloor, \lfloor nT \rfloor)$ that stay in the octant. We have represented the SPDE limits by their respective (purely formal) Feynman-Kac representations.

$$\zeta_n^2(i, 0) = 1 + n^{-1/4}\omega_{i,0}^n.$$

There are now three things to verify, corresponding to the three bullet points preceding Theorem 3.2.2. Using the fact that

$$\mathbb{E}[\Gamma^{-1}(\theta, 1)] = \frac{1}{\theta - 1}, \qquad var(\Gamma^{-1}(\theta, 1)) = \frac{1}{(\theta - 1)^2(\theta - 2)},$$

one gets the desired asymptotics on $\mathbb{E}[\omega_{i,j}]$ and on $\mathbb{E}[(\omega_{i,j})^2]$, with $\mu = -A$ and $\sigma^2 = 1$. This proves the corollary (and thus also Theorem 3.1.1). $\square$

Once again we would like to emphasize the tremendous importance of [14] as the primary input to proving the preceding theorem, and thus the main result (3.3). It may be interesting to explore more robust methods that might give a direct proof of (3.3) using purely stochastic analytic methods instead of exact solvability, but we have tried and this seems out of reach for us at the moment. With Theorem 3.2.4 in place, we will now shift the goals of the paper to the analytical and technical aspects focusing on the methods used to prove Theorem 3.2.2.

Since the sum defining the partition function in the preceding results is over all upright paths that stay in the principal octant of $\mathbb{Z}^2$, it is natural to relate those quantities to reflecting random walk measures. However, if one does asymptotics in Corollary 3.2.4, she may verify that $\zeta_n^2(j,j) \to 1/2$ in probability as $n \to \infty$. What this means is that instead of pure reflection, our random walk path loses mass by a factor of $1/2$ each time it hits zero. Hence, it is clear that the analysis in proving Theorem 3.2.2 will involve taking a close look at these random walk measures, as well as directed polymers weighted by such measures, as suggested in the introduction.

More precisely, fix some $x \in \mathbb{Z}_{\geq 0}$, and define a sample space of non-negative random walk trajectories by

$$\Omega_x^n := \{(s_0, ..., s_n) \in \mathbb{Z}^n : |s_{i+1} - s_i| = 1, s_i \geq 0, s_0 = x\}.$$

Define a sub-probability measure μ_x^n and a probability measure $\mathbf{P}_x^n$ on Ω_x^n by

$$\mu_x^n(S) := 2^{-n}, \qquad \mathbf{P}_x^n(S) := \frac{1}{\#\Omega_x^n} = \frac{\mu_x^n(S)}{\mu_x^n(\Omega_x^n)}, \qquad \text{for all } S \in \Omega_x^n.$$

As an intermediate step in proving Theorem 3.2.2, we obtain the following result.

Theorem 3.2.5. *With the above notation, the following are true.*

1. *(Markov Property) Fix $n, x \geq 0$. Let $S = (S_k)_{k=0}^n$ denote the coordinate process associated to $\mathbf{P}_x^n$, i.e., S is a Ω_x^n-valued random variable with law $\mathbf{P}_x^n$. Then $(S_k)_{k=0}^n$ is a time-inhomogeneous Markov process, in fact conditionally on $(S_k)_{k=0}^K$ with $K < n$, the*

process $(S_{k+K})_{k=0}^{n-K}$ *is distributed according to* $\mathbf{P}_{S_K}^{n-K}$. *One has explicit transition densities for* $0 \le i_1 < ... < i_k \le n$*:*

$$\mathbf{P}_x^n(S_{i_1} = s_1, ..., S_{i_k} = s_k) = \mathfrak{p}_{i_1}^n(x, s_1)\mathfrak{p}_{i_2-i_1}^{n-i_1}(s_1, s_2) \cdots \mathfrak{p}_{i_k-i_{k-1}}^{n-i_{k-1}}(s_{k-1}, s_k),$$

where $\mathfrak{p}_i^n$ *is given in Definition 3.3.2 below.*

2. *(Mass) For every* $x \in \mathbb{Z}_{\ge 0}$, *the total mass of* μ_x^n *is asymptotically* $(x+1)\sqrt{\frac{2}{\pi n}}$*:*

$$\lim_{n\to\infty} n^{1/2}\mu_x^n(\Omega_x^n) = (x+1)\sqrt{2/\pi}.$$

3. *(Concentration) There exist* $C, c > 0$ *such that for every* $x \ge 0$, *every* $0 \le m \le k \le n$, *and every* $u > 0$ *one has that*

$$\mathbf{P}_x^n\big(\sup_{m\le i\le k} |S_i - S_m| > u\big) \le Ce^{-cu^2/(k-m)}.$$

4. *(Convergence of Transition Densities) Let* $\mathfrak{p}_n^N$ *be as in Item (1). One has the convergence*

$$(n/2)^{1/2}\mathfrak{p}_{2\lfloor tn\rfloor}^{2\lfloor Tn\rfloor}(2\lfloor n^{1/2}X/\sqrt{2}\rfloor, 2\lfloor n^{1/2}Y/\sqrt{2}\rfloor) \stackrel{n\to\infty}{\longrightarrow} \mathscr{P}_t^T(X,Y),$$

where $\mathscr{P}_t^T$ *is the transition probability for a certain (inhomogeneous) Markov process defined in Definition 3.3.4 below. Moreover, for fixed* (t, T, X) *the convergence in the* Y*-variable occurs in* $L^p(\mathbb{R}_+, e^{aY}dY)$ *for every* $p \in [1, \infty)$.

The first part of the theorem is elementary and the last part is a more local version of the results of [95, 19]. The third part is new as far as we know, and the second part will simply follow from the local central limit theorem. All proofs may be found in the appendices, except for *(1)* which is proved in Section 3.

Remark 3.2.6. *One can actually formulate an invariance principle for this family of measures.*

This was done in greater generality in [95, 19]. Fix $X, T \geq 0$. For each $x, N \geq 0$, let $(S_n^{x,N})_{n=0}^N$ be distributed according to $\mathbf{P}_x^N$. Then the processes $(N^{-1/2} S_{Nt}^{N^{1/2}X, NT})_{t\in[0,T]}$ converge in law (with respect to the uniform topology on $C[0,T]$, as $N \to \infty$) to a time-inhomogeneous Markov process B on $[0,T]$ whose transition densities $\mathscr{P}_t^T(X,Y)$ are given by the limit in Item (4). This limiting process B may be interpreted as a standard Brownian motion conditioned to stay positive until time T; see Proposition 3.3.5. This invariance principle will be immediate from the results of Appendix A, but it will not be needed for the results above.

Let us now discuss the basic idea of the proof of Theorem 3.2.2 in the special case when $(T, X) = (1, 0)$ because this is enough to give the main idea. Denote by $\mathbf{E}_{KRW}$ the expectation with respect to a reflected random walk of length $2n$ that is started from 0 and killed at the origin with probability $1/2$, i.e., the one whose transition density is equal to $p_n^{(1/2)}$ which is defined in Section 3 below. By rotating the picture appropriately, one rewrites the partition function appearing in Theorem 3.2.2 as a discrete Feynman-Kac formula for this killed walk:

$$\begin{aligned} Z_n &= \sum_{\gamma:(0,0)\to(n,n)} 2^{-\#\{i\leq 2n\,:\,\gamma_2(i)>0\}} \prod_{i=0}^{2n} (1 + n^{-1/4}\omega^n_{\gamma_1(i),\gamma_2(i)}) \\ &= \mathbf{E}_{KRW}\Bigg[z_0^n(S_{T_n}) \prod_{i=0}^{T_n-1} (1 + n^{-1/4}\hat{\omega}^n_{i,S_i}) \cdot 1_{\{\text{survival}\}} \Bigg], \end{aligned} \tag{3.9}$$

where

- $\hat{\omega}^n_{i,j}$ is defined to be $\omega^n_{(n-\frac{i-j}{2}),(n-\frac{i+j}{2})}$ for all i, j.
- The expectation $\mathbf{E}_{KRW}$ is taken *only* with respect to the random walk S, i.e., **conditional** on the $\omega^n_{i,j}$ (which are always assumed to be independent of S).
- T_n is the first time that (i, S_i) hits the diagonal line $\{(2n-j, j) : 0 \leq j \leq 2n\}$.
- $z_0^n(x) := \prod_{i=0}^x (1 + n^{-1/4}\omega^n_{i,0})$ can be thought of as a sort of “initial data" for the above discrete Feynman-Kac representation.

- $\{\text{survival}\}$ is the event that the random walk survives up to time $2n$ (or equivalently, up to time T_n).

Now, using Theorem 3.2.5(2) with $x = 0$, one finds that $\mathbf{P}_{KRW}(\text{survival}) \approx \sqrt{2/\pi n}$. Moreover, we can make the approximation $T_n \approx 2n$ for reasons justified later, see Proposition 3.5.8. This essentially reduces the octant geometry to that of a quadrant, thus reducing the theorem statement to that of Theorem 3.1.2, which is simpler as we see below. Combining this with the above gives

$$\begin{aligned}\sqrt{\frac{\pi n}{2}} Z_n &\approx \mathbf{E}_{KRW}\left[z_0^n(S_{2n}) \prod_{i=0}^{2n} (1 + n^{-1/4} \hat{\omega}_{i,S_i}^n) \,\middle|\, \text{survival} \right] \\ &= \mathbf{E}_{KRW}\left[z_0^n(S_{2n}) \sum_{k=0}^{2n} n^{-k/4} \sum_{1 \le i_1 < \ldots < i_k \le 2n} \prod_{j=1}^{k} \hat{\omega}_{i_j, S_{i_j}}^n \,\middle|\, \text{survival} \right]. \end{aligned} \tag{3.10}$$

In the notation of Theorem 3.2.5, the killed random walk conditioned to survive has law $\mathbf{P}_x^n$ and the associated Markov process has transition densities $\mathfrak{p}_n^N$. Using theorem 3.2.5*(1)*, the expectation in the preceding expression may be expanded as

$$\sum_{k=0}^{2n} n^{-k/4} \sum_{0 \le i_1 < \ldots < i_k \le 2n} \sum_{(x_1, \ldots, x_{k+1}) \in \mathbb{Z}_{\ge 0}^{k+1}} z_0^n(x_{k+1}) \prod_{j=1}^{k+1} \mathfrak{p}_{i_j - i_{j-1}}^{2n - i_{j-1}}(x_{j-1}, x_j) \prod_{j=1}^{k} \hat{\omega}_{i_j, x_j}^n, \tag{3.11}$$

with $x_0 := 0$, $i_0 := 0$, and $i_{k+1} := 2n$. Recall that $\log(1+u) \approx u - \frac{1}{2}u^2$, so by writing

$$\begin{aligned} z_0^n(x) = e^{\sum_0^x \log(1 + n^{-1/4} \omega_{i,0}^n)} &\approx e^{\sum_0^x \left(n^{-1/4} \omega_{i,0}^n - \frac{1}{2} n^{-1/2} (\omega_{i,0}^n)^2 \right)} \\ &= e^{n^{-1/4} \sum_0^x (\omega_{i,0}^n - n^{-1/4}\mu) + n^{-1/2}\mu x - \frac{1}{2n^{1/2}} \sum_0^x (\omega_{i,0}^n)^2}, \end{aligned} \tag{3.12}$$

one may convince herself (using Donsker's principle and the law of large numbers together with the third bullet point preceding Theorem 3.2.2) that as $n \to \infty$,

$$\left(z_0^n(n^{1/2} X) \right)_{X \ge 0} \xrightarrow{d} \left(e^{\sigma B_X + (\mu - \frac{1}{2}\sigma^2) X} \right)_{X \ge 0} \tag{3.13}$$

for a Brownian motion B. Then taking the limit of (3.11) as $n \to \infty$ by using Theorem 3.2.5*(4)* (with some uniformity estimates), one obtains the Wiener-Itô chaos series

$$\sum_{k=0}^{\infty}\int_{0\le t_1<\ldots<t_k\le 1}\int_{\mathbb{R}_+^{k+1}} e^{\sigma B_{x_{k+1}}+(\mu-\frac{1}{2}\sigma^2)x_{k+1}}\prod_{j=1}^{k+1}\mathscr{P}_{t_j-t_{j-1}}^{1-t_{j-1}}(x_{j-1},x_j)dx_{k+1}\xi(dx_k,dt_k)\cdots\xi(dx_1,dt_1),$$

with the convention $x_0 = 0$, $t_0 = 0$, $t_{k+1} = 1$, and where the $\mathscr{P}_t^T$ are the conditional heat kernels from the limit in Theorem 3.2.5*(4)*, and ξ is a space-time white noise. But (as we will see in Proposition 3.4.2 below) this chaos series is precisely equal to

$$\lim_{X\to 0}\frac{Z_{Dir}(1,X)}{2\Phi(X)-1} = \sqrt{\pi/2}\lim_{X\to 0}\frac{Z_{Dir}(1,X)}{X},$$

where the initial data is $e^{\sigma B_X+(\mu-\frac{1}{2}\sigma^2)X}$, and Φ is the cdf of a standard normal, which implies that $\Phi(0) = 1/2$ and $\Phi'(0) = \sqrt{\pi/2}$ giving the equality above. This will complete the argument for Theorem 2.2. Note that no part of the argument relies on the finer details of the weights $\omega_{i,j}^n$ beyond their mean and variance.

3.3 Uniform measures on collections of positive paths

In this section we will introduce the inhomogeneous heat kernels $\mathfrak{p}_n^N$ associated to random walks conditioned to stay positive. We begin with an elementary discussion of the properties of these measures, and later we state technical estimates about these measures that will be necessary in subsequent sections, though their proofs are postponed to the appendices.

Definition 3.3.1. *For $n \in \mathbb{Z}_{\ge 0}$ and $x \in \mathbb{Z}$, let $p_n(x)$ denote the standard heat kernel on $\mathbb{Z}$ (i.e., the transition function for a discrete-time simple symmetric random walk started from zero). Then we define*

$$p_n^{(1/2)}(x,y) = p_n(x-y) - p_n(x+y+2), \qquad n,x,y\ge 0.$$

The kernels $p_n^{(1/2)}$ have the following probabilistic interpretation. Consider a simple symmetric random walk $(S_n)_{n\ge 0}$ with $S_0 = 0$ on the integer lattice $\mathbb{Z}$. Impose the condition that this random walk gets killed, i.e., enters an auxiliary death state, at the first instance that it hits the value -1.

Equivalently one can consider a random walk reflected at 0 that dies independently with probability $1/2$ each time it attempts to move from site 0 to site 1. Then $p_n^{(1/2)}(x,y)$ is the probability of the following event: the walk started from x is at position y at time n.

Definition 3.3.2. *We define the following quantity for integers* $0 \le n \le N$

$$\mathfrak{p}_n^N(x,y) := p_n^{(1/2)}(x,y)\frac{\psi(y;N-n)}{\psi(x;N)}, \qquad \textit{where} \qquad \psi(x;n) := \sum_{y\ge 0} p_n^{(1/2)}(x,y).$$

The probabilistic relevance of these kernels $\mathfrak{p}_n^N$ will be demonstrated shortly in Proposition 3.3.3. As in Theorem 3.2.5, let

$$\Omega_x^N := \{(s_0,...,s_N) \in \mathbb{Z}_{\ge 0}^{N+1} : |s_{i+1}-s_i| = 1, s_0 = x\}.$$

Then denote by $\mathbf{P}_x^N$ the uniform probability measure on Ω_x^N, and let S denote the coordinate process associated to this measure (e.g., S can be the identity map on Ω_x^N). In plainer terms, S is a simple symmetric random walk of length N conditioned to stay non-negative throughout its course.

Proposition 3.3.3. *S is an inhomogeneous Markov process on the discrete time interval* $\{0,...,N\}$. *In fact, for* $0 \le i_1 < ... < i_n \le N$ *one has*

$$\begin{aligned}\mathbf{P}_x^N(S_{i_1}=s_1,...,S_{i_n}=s_n) &= \mathfrak{p}_{i_1}^N(x,s_1)\mathfrak{p}_{i_2-i_1}^{N-i_1}(s_1,s_2)\cdots\mathfrak{p}_{i_n-i_{n-1}}^{N-i_{n-1}}(s_{n-1},s_n)\\ &= p_{i_1}^{(1/2)}(x,s_1)p_{i_2-i_1}^{(1/2)}(s_1,s_2)\cdots p_{i_n-i_{n-1}}^{(1/2)}(s_{n-1},s_n)\frac{\psi(s_n,N-i_n)}{\psi(x,N)}.\end{aligned}$$

In particular, for $M<N$ the conditional law of $(S_{M+k})_{k=0}^{N-M}$ given $(S_k)_{k=0}^M$ is distributed according to $\mathbf{P}_{S_M}^{N-M}$.

This proves Theorem 3.2.5*(1)* and shows that the $\mathfrak{p}_n^N(x,\cdot)$ are probability measures.

Proof. Write $S_{[0,M]}$ for the restriction of S to $\{0,1,...,M\}$, and write $S^{[M,N]}$ for the restriction of S to $\{M,...,N\}$ shifted by M places (so $S^{[M,N]}$ is defined on $\{0,...,N-M\}$). For nearest-neighbor paths s_1 and s_2 of lengths M and $N-M$, respectively, such that $s_1(M) = s_2(0)$ one computes that

$$\mathbf{P}_x^N(S^{[M,N]} = s_2 | S_{[0,M]} = s_1) = \frac{\mathbf{P}_x^N(S = s_1 * s_2)}{\mathbf{P}_x^N(S_{[0,M]} = s_1)} = \frac{\frac{1}{\#\Omega_x^N}}{\frac{\#\{\pi \in \Omega_x^N : \pi|_{[0,M]} = s_1\}}{\#\Omega_x^N}}$$

$$= \frac{1}{\#\{\pi \in \Omega_x^N : \pi|_{[0,M]} = s_1\}} = \frac{1}{\#\Omega_{s_1(M)}^{N-M}} = \mathbf{P}_{s_1(M)}^{N-M}(S = s_2),$$

where $s_1 * s_2$ denotes the concatenation of paths. This immediately implies that given $(S_k)_{k=0}^M$ the law of $(S_{M+k})_{k=0}^{N-M}$ is distributed according to $\mathbf{P}_{S_M}^{N-M}$. This also implies that $(S_k)_{k=0}^M$ and $(S_{M+k})_{k=0}^{N-M}$ are conditionally independent given S_M. Therefore, in order to prove the given formula for transition densities, it suffices to prove the claim for $n = 1$; then the claim for general n follows from the conditional independence and induction (recall that n is the number of indices $0 \le i_1 < \dots < i_n \le N$ appearing in the transition formula).

To prove the formula for $n = 1$ it suffices by conditional independence to assume that $i_n = N$. Note that $\mathbf{P}_x^N$ is the probability measure associated to the killed random walk conditioned to survive, so that

$$\mathbf{P}_x^N(S_N = s) = \frac{p_N^{(1/2)}(x,s)}{\sum_{y \ge 0} p_N^{(1/2)}(x,y)} = p_N^{(1/2)}(x,s)\frac{1}{\psi(x,N)},$$

which proves the claim. $\square$

Next we introduce the continuum analogues of the previously introduced measures. We will generally use capital letters to distinguish macroscopic variables from lowercase microscopic ones.

Definition 3.3.4. *Let* $P_t(X) := e^{-X^2/2t}/\sqrt{2\pi t}$ *denote the standard heat kernel on the whole line* $\mathbb{R}$. *Recall the Dirichlet boundary heat kernel*

$$P_t^{Dir}(X,Y) := P_t(X-Y) - P_t(X+Y).$$

We then define the inhomogeneous kernel for $0 \le t \le T$ *and* $X, Y > 0$:

$$\mathscr{P}_t^T(X,Y) := \begin{cases} P_t^{Dir}(X,Y)\frac{2\Phi(Y/\sqrt{T-t})-1}{2\Phi(X/\sqrt{T})-1} & t<T \\ P_T^{Dir}(X,Y)\frac{1}{2\Phi(X/\sqrt{T})-1} & t=T, \end{cases}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x e^{-u^2/2}du$ *is the cdf of a standard normal. For* $X = 0$*, one analogously defines the quantity for* $Y > 0$ *and* $T \ge t \ge 0$*:*

$$\mathscr{P}_t^T(0,Y) = \begin{cases} Y(T/t^3)^{1/2}e^{-Y^2/2t}\big(2\Phi(Y/\sqrt{T-t})-1\big) & t<T \\ (Y/T)e^{-Y^2/2T} & t=T, \end{cases}$$

which is the limit of the previously defined $\mathscr{P}_t^T(X,Y)$ *as* $X \to 0$*.*

We now discuss the relevance of these kernels as Markov transition densities. Specifically, for $X > 0$ define $\mathbf{W}_X^T$ to be the probability measure on $C([0,T],\mathbb{R}_+)$ obtained by conditioning Brownian motion on $[0,T]$ started from X to stay strictly positive until time T.[1] We define B to be the canonical process associated to $\mathbf{W}_X^T$. One can also define $\mathbf{W}_0^T$ as the weak limit of the $\mathbf{W}_X^T$ as $X \to 0$. The fact that this limiting measure exists is not difficult but not entirely trivial either (see the appendices). It is called the *Brownian meander* and has been studied extensively in [66, 65, 49, 95] and subsequent papers on the subject.

Proposition 3.3.5. *Fix some* $T, X > 0$ *and let* $\mathbf{W}_X^T$ *be as defined above, and let* B *denote the associated canonical process. Consider the kernels* $\mathscr{P}_t^T$ *defined before. Then for* $0 \le t_1 < ... < t_n \le T$ *and* $Y_1, ..., Y_n > 0$*,*

$$\begin{aligned} &\mathbf{W}_X^T(B_{t_1} \in dY_1,\ ...,\ B_{t_n} \in dY_n) \\ &\qquad = \mathscr{P}_{t_1}^T(X,Y_1)\mathscr{P}_{t_2-t_1}^{T-t_1}(Y_1,Y_2)\cdots\mathscr{P}_{t_n-t_{n-1}}^{T-t_{n-1}}(Y_{n-1},Y_n)\, dY_1\cdots dY_n \end{aligned}$$

[1] This is not the same as a 3D Bessel process, which is BM conditioned to stay positive for *all* time and is time-homogeneous.

In particular, if $T < S$ then the conditional law of $(B_{t+S})_{t\in[0,T-S]}$ given $(B_t)_{t\in[0,S]}$ is equal to $\mathbf{W}_{B_S}^{T-S}$. The same statements hold true for $X = 0$.

Before moving on to the proof, we remark that when $X \neq 0$ and $t_n \neq T$, the above formula for transition densities reduces to

$$P_{t_1}^{Dir}(X, Y_1) P_{t_2-t_1}^{Dir}(Y_1, Y_2) \cdots P_{t_n-t_{n-1}}^{Dir}(Y_{n-1}, Y_n) \frac{2\Phi(Y_n/\sqrt{T-t_n}) - 1}{2\Phi(X/\sqrt{T}) - 1}\, dY_1 \cdots dY_n.$$

When $t_n = T$ the numerator $2\Phi(Y_n/\sqrt{T-t_n}) - 1$ should be interpreted as 1. When $X = 0$ this expression becomes $0/0$, and one needs to take the limit, which gives the formula stated in the above proposition.

Proof. Assuming $X > 0$ the proof is analogous to that of Proposition 3.3.3. Basically one first shows that if $S < T$ then the conditional law of $(B_{t+S})_{t\in[0,T-S]}$ given $(B_t)_{t\in[0,S]}$ is equal to $\mathbf{W}_{B_S}^{T-S}$, and furthermore that $(B_{t+S})_{t\in[0,T-S]}$ and $(B_t)_{t\in[0,S]}$ are conditionally independent given B_S. This may be proven by a single computation using the basic properties of standard Brownian motion.

As in the proof of Proposition 3.3.3, this then reduces the claim to proving the formula for $n = 1$ and $t_n = T$. In turn, this follows by noticing that $\mathbf{W}_X^T$ is the same as Brownian motion killed at zero but conditioned to survive. Hence one finds that

$$\mathbf{W}_X^T(B_T \in dY) = \frac{P_T^{Dir}(X, Y)dY}{\int_0^\infty P_T^{Dir}(X, Z)dZ} = \frac{P_T^{Dir}(X, Y)dY}{2\Phi(X/\sqrt{T}) - 1},$$

which proves the claim. □

This concludes the introductory material on the subject, and we now state several technical estimates on these inhomogeneous heat kernels that are used heavily in the sequel. The proofs may be found in Appendix B.

Proposition 3.3.6. *Fix $\tau \geq 0$. Then for $n \geq 0$, define*

$$\mathscr{P}_n(t,T;X,Y) := (n/2)^{1/2}\mathfrak{p}_{2\lfloor tn\rfloor}^{2\lfloor Tn\rfloor}(2\lfloor n^{1/2}X/\sqrt{2}\rfloor, 2\lfloor n^{1/2}Y/\sqrt{2}\rfloor).$$

Then for each fixed $X,T,t \geq 0$, as $n \to \infty$ the map $Y \mapsto \mathscr{P}_n(t,T;X,Y)$ converges pointwise and in $L^p(\mathbb{R}_+, e^{aY}dY)$ to $\mathscr{P}_t^T(X,Y)$ for all $p \geq 1$ and $a \geq 0$. Furthermore for all $X,T \geq 0$, the map $(t,Y) \mapsto \mathscr{P}_n(t,T;X,Y)$ converges pointwise and in $L^p(dt \otimes e^{aY}dY)$ to $\mathscr{P}_t^T(X,Y)$ for all $p \in [1,3)$ and $a \geq 0$ (as $n \to \infty$).

We refer the reader to Proposition 3.7.7 of the appendix for the proof. We remark that the annoying factors of 2 appearing in the definition of $\mathscr{P}_n$ are only necessary due to the periodicity of the simple random walk.

Proposition 3.3.7. *Let $a, \tau > 0$ and let $\mathscr{P}_t^T$ be the kernels from Definition 3.3.4. Then there exists a constant $C = C(\tau, a)$ such that for all $X, Y \geq 0$, all $\theta \in [0, 1/2]$, and all $s \leq t \leq T \leq \tau$ one has the following*

$$\int_{\mathbb{R}_+} \mathscr{P}_t^T(X,Z)e^{aZ}dZ \leq Ce^{aX}, \tag{3.14}$$

$$\int_{\mathbb{R}_+} \mathscr{P}_t^T(X,Z)^2e^{aZ}dZ \leq Ct^{-1/2}e^{aX}, \tag{3.15}$$

$$\int_{\mathbb{R}_+} \big(\mathscr{P}_t^T(X,Z) - \mathscr{P}_t^T(Y,Z)\big)^2 e^{aZ}dZ \leq Ct^{-\frac{1}{2}-\theta}e^{a(X+Y)}|X-Y|^{2\theta}, \tag{3.16}$$

$$\int_{\mathbb{R}_+} \big(\mathscr{P}_s^{T-t+s}(X,Z) - \mathscr{P}_t^T(X,Z)\big)^2 e^{aZ}dZ \leq Cs^{-\frac{1}{2}-\theta}e^{2aX}|t-s|^{\theta} \tag{3.17}$$

The proof may be found as the very last thing in Appendix B. We remark that these bounds will be the key behind the proofs of Section 4 below.

3.4 Existence of the derivative in Dirichlet SHE

Note that in order to prove the identity (3.3), one first needs to prove that the mild solution of Z_{Dir} exists and that the limit on the right-hand side of (3.3) also exists. In this section we actually do something much stronger. We will prove that the mild solution of Z_{Dir} and the aforementioned limits not only exist, but in fact one almost surely has the *simultaneous* existence of $\lim_{X\to 0}\frac{Z_{Dir}(T,X)}{X}$ for *all* $T\geq 0$, for a fixed initial data. Furthermore this limit is Hölder-continuous as a function of T.

All of this will essentially be proved in a single step by showing that for $X,T\geq 0$ the chaos series

$$\sum_{k=0}^{\infty}\int_{0\leq t_1<\ldots<t_k\leq T}\int_{\mathbb{R}_+^{k+1}} f(X_{k+1})\prod_{j=1}^{k+1}\mathscr{P}_{t_j-t_{j-1}}^{T-t_{j-1}}(X_{j-1},X_j)dX_{k+1}\xi_T(dX_k,dt_k)\cdots\xi_T(dX_1,dt_1),$$

converges uniformly over compact subsets of $(T,X)\in\mathbb{R}_+\times\mathbb{R}_+$, where $t_0:=0, X_0:=X$, $\xi_T(X,t):=\xi(X,T-t)$ for a space-time white noise ξ, and f is some random initial data with at-worst exponential growth at infinity. Then we will show almost trivially that when $X,T>0$ this chaos series equals $Z_{Dir}(T,X)/\big(2\Phi(X/\sqrt{T})-1\big)$, where Φ is the cdf of a standard normal and Z_{Dir} satisfies the conditions of Definition 3.2.1. This would simultaneously prove existence of Z_{Dir} and also the desired limit. This is because we know the above chaos series extends continuously to $X=0$, which means $\lim_{X\to 0}\frac{Z_{Dir}(T,X)}{2\Phi(X/\sqrt{T})-1}$ exists, which is equivalent to showing that $\lim_{X\to 0}\frac{Z_{Dir}(T,X)}{X}$ exists (for all $T\geq 0$, a.s.).

In order to prove the uniform convergence of this chaos series, we are going to use the inhomogeneous heat kernel estimates stated at the end of Section 3. The proofs may be skipped without any effect on the readability of Section 5, although some ideas are similar to ones used there.

With this motivation, we now move on to the main results of this section. Given some possibly random initial data $f:\mathbb{R}_+\to\mathbb{R}_+$, recall from (3.5) the following Duhamel-form SPDE:

$$\mathscr{Z}(T,X)=\int_{\mathbb{R}^+}\mathscr{P}_T^T(X,Y)f(Y)dY+\int_0^T\int_{\mathbb{R}^+}\mathscr{P}_{T-S}^T(X,Y)\mathscr{Z}(S,Y)\xi(dYdS), \tag{3.18}$$

where ξ is a space-time white noise and so the above should be interpreted as an Itô integral. Since $\mathscr{Z}$ appears on both sides of this relation, it is not clear that a solution would even exist. Thus we have the following result, which will be proved by rigorously expanding (3.18) into the chaos series mentioned above.

Theorem 3.4.1. *Fix $a, \tau > 0$ and suppose that we have some random function-valued initial data f satisfying*

$$\sup_{X \geq 0} e^{-aX} \mathbb{E}[f(X)^2] < \infty.$$

Then, a unique solution to the SPDE (3.18) *with initial data f exists in the class of space-time functions $\mathscr{Z}(T, X)$ that satisfy*

$$\sup_{\substack{X \geq 0 \\ T \in [0,\tau]}} e^{-aX} \mathbb{E}[\mathscr{Z}(T, X)^2] < \infty.$$

Furthermore, the solution $\mathscr{Z}$ may be constructed in such a way that its law is supported on the space of functions that are Hölder-continuous of exponent $1/2 - \epsilon$ in the X variable and $1/4 - \epsilon$ in the T variable on any compact subset of $(T, X) \in (0, \infty) \times [0, \infty)$ for any $\epsilon > 0$.

Proof. This is adapted from the proofs given in Section 2.4 above. Informally, one argues as follows: define the following sequence of iterates:

$$u_0(T, X) = \int_{\mathbb{R}_+} \mathscr{P}_T^T(X, Y) f(Y) dY,$$

$$u_{n+1}(T, X) = \int_0^T \int_{\mathbb{R}^+} \mathscr{P}_{T-S}^T(X, Y) u_n(S, Y) \xi(dY dS).$$

In other words, u_n is just the n^{th} term of a chaos series given by the expansion of (3.18). Thus it is clear that the desired solution to (3.18) should be given by $\sum_{n \geq 0} u_n$. Hence, in order to formalize these ideas, we will show that the series $\sum_{n \geq 0} u_n$ converges in the appropriate Banach space of random space-time functions.

To this end, let us define a Banach space $\mathcal{B}$ of $C(\mathbb{R}_+)$-valued processes $u = (u(T,\cdot))_{T\in[0,\tau]}$ that are adapted to the natural filtration of ξ and with norm given by

$$\|u\|_{\mathcal{B}}^2 := \sup_{\substack{X\geq 0\\ T\in[0,\tau]}} e^{-aX}\mathbb{E}[u(T,X)^2].$$

Then define a sequence of functions $F_n : [0,\tau] \to \mathbb{R}$ for $n \geq 0$ by

$$F_n(T) := \sup_{\substack{X\geq 0\\ S\in[0,T]}} e^{-aX}\mathbb{E}[u_n(S,X)^2],$$

where u_n are the iterates defined above. By Itô isometry, it is clear that

$$\begin{aligned}\mathbb{E}[u_{n+1}(T,X)^2] &= \int_0^T \int_{\mathbb{R}_+} \mathscr{P}_{T-S}^T(X,Y)^2\mathbb{E}[u_n(S,Y)^2]dY\,dS\\ &\leq \int_0^T \left[\int_{\mathbb{R}_+} \mathscr{P}_{T-S}^T(X,Y)^2 e^{aY}dY\right]F_n(S)dS. \end{aligned}\tag{3.19}$$

Now by (3.15) we have that

$$\int_{\mathbb{R}_+} \mathscr{P}_{T-S}^T(X,Y)^2 e^{aY}dY \leq C(T-S)^{-1/2}e^{aX}, \qquad \forall T\in[0,\tau], X\geq 0, \tag{3.20}$$

where C may depend on a and τ. Furthermore one notes that the F_n are increasing functions of T, and therefore $T \mapsto \int_0^T (T-S)^{-1/2}F_n(S)dS$ is also increasing (which may be verified by making the substitution $S = TU$). Combining this fact with (3.19) and (3.20), one obtains

$$F_{n+1}(T) \leq C\int_0^T (T-S)^{-1/2}F_n(S)dS, \tag{3.21}$$

where C does not depend on n. Now, we claim that $F_0(T) \leq C$ (with $C = C(a,\tau)$). Indeed, by Jensen's inequality and Fubini's theorem, one has

$$\mathbb{E}[u_0(T,X)^2] = \mathbb{E}\left[\left(\int_{\mathbb{R}_+} \mathscr{P}_T^T(X,Y)f(Y)dY\right)^2\right] \leq \int_{\mathbb{R}_+} \mathscr{P}_T^T(X,Y)\mathbb{E}[f(Y)^2]dY \leq Ce^{aX},$$

where in the last inequality we used (3.14) together with the assumption that $\mathbb{E}[f(X)^2] \leq Ce^{aX}$. This proves that $F_0 \leq C$, which means that one may iterate (3.21) to obtain

$$F_n(T) \lesssim C^n T^{n/2}/(n/2)!, \tag{3.22}$$

which implies that $\sum_{n\geq 0} \|u_n\|_{\mathcal{B}} < \infty$. This completes the proof of existence.

The proof of uniqueness is essentially the same. Indeed, if $\mathscr{Z}$ and $\mathscr{Z}'$ were two solutions in $\mathcal{B}$ that are started from the same initial data f, then an application of Itô's isometry reveals that

$$\mathbb{E}\big[(\mathscr{Z}(T,X) - \mathscr{Z}'(T,X))^2\big] = \int_0^T \int_{\mathbb{R}_+} \mathscr{P}_{T-S}^T(X,Y)^2 \mathbb{E}[(\mathscr{Z}(S,Y) - \mathscr{Z}'(S,Y))^2] dSdY.$$

Then one iterates as above and may obtain that the left-hand side is bounded above (uniformly in T, X) by $C^n T^{n/2}/(n/2)!$, and by letting $n \to \infty$ this tends to zero.

Now we address the Hölder regularity. Let u_n be the iterates defined above. We know that u_0 is a smooth function of $(T,X) \in (0,\infty) \times [0,\infty)$ because it is the solution to the deterministic (i.e., noiseless) version of SPDE (3.18) which is just an inhomogeneous heat equation (e.g., one may simply differentiate u_0 under the integral sign). Thus, it suffices to prove that the function $\mathscr{Z}_0 := \mathscr{Z} - u_0 = \sum_{n\geq 1} u_n$ has the required Hölder regularity, so this is what we will do.

Henceforth fix an exponent $\gamma \in (0, 1/2)$. For the spatial regularity, one computes that

$$\mathbb{E}[(u_{n+1}(T,X) - u_{n+1}(T,Y))^2] = \int_0^T \int_{\mathbb{R}_+} \big(\mathscr{P}_{T-S}^T(X,Z) - \mathscr{P}_{T-S}^T(Y,Z)\big)^2 \mathbb{E}[u_n(S,Z)^2] dZ dS$$

$$
\begin{aligned}
&\le \int_0^T \left[\int_{\mathbb{R}_+} \big(\mathscr{P}^T_{T-S}(X,Z) - \mathscr{P}^T_{T-S}(Y,Z)\big)^2 e^{aZ} dZ \right] F_n(S) dS \\
&\le C \int_0^T (T-S)^{\gamma-1} |X-Y|^{1-2\gamma} e^{a(X+Y)} F_n(S) dS \\
&\overset{(3.22)}{\le} C e^{a(X+Y)} |X-Y|^{1-2\gamma} \int_0^T (T-S)^{\gamma-1} \frac{C^n S^{n/2}}{(n/2)!} dS \\
&\le C^{n+1} e^{a(X+Y)} |X-Y|^{1-2\gamma} \frac{T^{(n+2\gamma)/2}}{(n/2)!} \int_0^1 (1-a)^{\gamma-1} a^{n/2} da\ .
\end{aligned}
$$

In the third line we used (3.16) with $\theta = \frac{1}{2} - \gamma$, and in the final line we made a substitution $S = Ta$. Note that the final integral is bounded independently of n, so it may be absorbed into the constant (which will then depend on γ). Using hypercontractivity of the Ornstein-Uhlenbeck semigroup associated to the Gaussian noise ξ, we can bound the p^{th} moments of elements of the homogeneous Wiener chaoses in terms of their second moments. Specifically, if $p \ge 2$ then Equation 7.2 of [84] says that:

$$
\begin{aligned}
\mathbb{E}\Big[\big|u_{n+1}(T,X) - u_{n+1}(T,Y)\big|^p\Big]^{1/p} &\le (p-1)^{(n+1)/2} \mathbb{E}\Big[\big(u_{n+1}(T,X) - u_{n+1}(T,Y)\big)^2\Big]^{1/2} \\
&\le C^{(n+1)/2} p^{(n+1)/2} e^{a(X+Y)/2} \frac{T^{(n+1)/4}}{\sqrt{(n/2)!}} |X-Y|^{\frac{1}{2}-\gamma}.
\end{aligned}
$$

Using Minkowski's inequality and summing over all n, we then obtain

$$
\mathbb{E}\Big[\big|\mathscr{Z}_0(T,X) - \mathscr{Z}_0(T,Y)\big|^p\Big]^{1/p} \le \sum_{n\ge 1} \mathbb{E}\Big[\big|u_n(T,X) - u_n(T,Y)\big|^p\Big]^{1/p} \le D(p,T) e^{a(X+Y)/2} |X-Y|^{\frac{1}{2}-\gamma}.
$$

Here $D(p,T) := \sum_n \frac{(CpT^{1/2})^{(n+1)/2}}{\sqrt{(n/2)!}}$, which is independent of X, Y and increasing as a function of T. This is enough by Kolmogorov's criterion to ensure that $\mathscr{Z}_0$ is Hölder continuous of exponent $1/2 - \gamma - \epsilon$ (on compact sets) in the spatial variable.

For the temporal regularity, one computes

$$\begin{aligned}
&\mathbb{E}[(u_{n+1}(T,X)-u_{n+1}(S,X))^2]\\
=&\mathbb{E}\Big[\Big(\int_0^T\int_{\mathbb{R}_+}\mathscr{P}^T_{T-U}(X,Z)u_n(U,Z)\xi(dZdU)-\int_0^S\int_{\mathbb{R}_+}\mathscr{P}^S_{S-U}(X,Z)u_n(U,Z)\xi(dZdU)\Big)^2\Big]
\end{aligned}$$

$$\begin{aligned}
=\int_0^S\int_{\mathbb{R}_+}\big(\mathscr{P}^T_{T-U}(X,Z)-\mathscr{P}^S_{S-U}(X,Z)\big)^2\mathbb{E}[u_n(U,Z)^2]dZdU\\
+\int_S^T\int_{\mathbb{R}_+}\mathscr{P}^T_{T-U}(X,Z)^2\mathbb{E}[u_n(U,Z)^2]dZdU.
\end{aligned}$$

Let us call the integrals in the final expression I_1 and I_2 respectively. As before, one has the bound $\mathbb{E}[u_n(U,Z)^2]\le e^{aZ}F_n(U)\le e^{aZ}\frac{C^nU^{n/2}}{(n/2)!}$. Then one uses (3.17) with $\theta=\frac{1}{2}-\gamma$ to bound the inner integral of I_1 by

$$\int_{\mathbb{R}_+}\big(\mathscr{P}^T_{T-U}(X,Z)-\mathscr{P}^S_{S-U}(X,Z)\big)^2e^{aZ}dZ\le Ce^{2aX}(S-U)^{\gamma-1}|T-S|^{\frac{1}{2}-\gamma},$$

and one also uses (3.15) to bound the inner integral of I_2 as

$$\int_{\mathbb{R}_+}\mathscr{P}^T_{T-U}(X,Z)^2e^{aZ}dZ\le C(T-U)^{-1/2}e^{aX}.$$

Then one finally performs the integral over U on the respective domains, and one can obtain (as in the spatial case) that $I_1+I_2\le C^{n+1}e^{2aX}T^{(n+1)/2}|T-S|^{\frac{1}{2}-\gamma}/(n/2)!$. Then one uses hypercontractivity and sums over n (exactly as in the spatial case), to get that

$$\mathbb{E}\big[\big|\mathscr{Z}_0(T,X)-\mathscr{Z}_0(S,X)\big|^p\big]^{1/p}\le D(p,T)e^{2aX}|T-S|^{\frac{1}{4}-\frac{\gamma}{2}}.$$

Here $D(p,T)$ is an increasing function of T (the same one as before), so it can be bounded from above on any compact set of (T,X). This is enough to give Hölder regularity of $\frac{1}{4}-\frac{\gamma}{2}-\epsilon$ in time, by Kolmogorov's criterion. $\square$

Next, we discuss the relationship of the $\mathscr{Z}$ that we have constructed in Theorem 3.4.1 with the Dirichlet-boundary SHE.

Proposition 3.4.2. *Any solution of the SPDE* (3.18) *must a.s. satisfy the following relation for all* $T, X > 0$

$$\mathscr{Z}(T,X)\big(2\Phi(X/\sqrt{T})-1\big) = Z_{Dir}(T,X)$$

where Z_{Dir} solves the Dirichlet-boundary SHE as in Definition 3.2.1 with the same initial data f.

Proof. One notes the following relation for $X > 0$, which is immediate from Definition 3.3.4:

$$\mathscr{P}_t^T(X,Y)\big(2\Phi(X/\sqrt{T})-1\big) = \begin{cases} P_t^{Dir}(X,Y)\big(2\Phi(Y/\sqrt{T-t})-1\big), & t<T \\ P_T^{Dir}(X,Y), & t=T. \end{cases} \tag{3.23}$$

So suppose $\mathscr{Z}$ solves (3.18), and define

$$A(T,X) := \mathscr{Z}(T,X)\big(2\Phi(X/\sqrt{T})-1\big).$$

By multiplying both sides of (3.18) by $2\Phi(X/\sqrt{T})-1$ and applying (3.23), one has the relation

$$\begin{aligned} A(T,X) &= \int_{\mathbb{R}_+} P_T^{Dir}(X,Y)f(Y)dY + \int_0^T\int_{\mathbb{R}_+} P_{T-S}^{Dir}(X,Y)\Big[\mathscr{Z}(S,Y)\big(2\Phi(Y/\sqrt{S})-1\big)\Big]\xi(dY,dS) \\ &= \int_{\mathbb{R}_+} P_T^{Dir}(X,Y)f(Y)dY + \int_0^T\int_{\mathbb{R}_+} P_{T-S}^{Dir}(X,Y)A(S,Y)\xi(dY,dS), \end{aligned}$$

so that A is indeed a mild solution to the Dirichlet-boundary SHE. $\square$

One thing we have not addressed is the *uniqueness* of solutions to the Dirichlet-boundary SHE in some large enough class of random space-time functions. This can be obtained from Theorem 3.4.1 with minimal work, and with the same conditions on the initial data f, one can in fact obtain existence/uniqueness in the space of ξ-adapted space-time functions A satisfying the bound

$\sup_{T \le \tau,\, X \ge 0} \mathbb{E}[A(T,X)^2] < \infty$.

Corollary 3.4.3. *Consider any solution Z_{Dir} of the Dirichlet-boundary SHE, started from any initial data f satisfying the assumptions of Theorem 3.4.1. Then almost surely, for every $T > 0$ the limit of $\frac{Z_{Dir}(T,X)}{X}$ exists as $X \to 0$.*

Proof. Consider the solution $\mathscr{Z}$ to (3.18) started from initial data f. By the preceding proposition, we can couple this with the solution to the Dirichlet-boundary SHE in such a way so that

$$\mathscr{Z}(T,X) = \frac{Z_{Dir}(T,X)}{2\Phi(X/\sqrt{T}) - 1}$$

for all $X > 0$ and $T \ge 0$. But we know that $\mathscr{Z}$ extends continuously to $X = 0$ by Theorem 3.4.1, hence we know that

$$\lim_{X \to 0} \frac{Z_{Dir}(T,X)}{2\Phi(X/\sqrt{T}) - 1}$$

exists, and since $2\Phi(X/\sqrt{T}) - 1$ has nonzero derivative at $X = 0$, the claim follows. □

3.5 Convergence of the partition function to SHE

In this section we use a discrete chaos expansion together with the methods of [3, 31] and the heat kernel estimates of the previous sections in order to prove Theorem 3.2.2. The first step (Section 5.1) is to simplify the geometry of the region where our directed polymer lives, and then (in Section 5.2) we will prove the convergence result in the simpler domain.

As a notational convention, we will usually write C for constants, and we will not generally specify when irrelevant terms are being absorbed into the constants. We will also write $C(a), C(a,p)$, or $C(a,p,K)$ whenever we want to specify exactly which parameters the constant depends on. This will not always be specified, though. This applies throughout the paper. Please be warned that we will freely use many different bounds from the appendices in the following proofs, so the reader may wish to skim those estimates first.

3.5.1 Reduction from the octant model to the quadrant model

In this section, we reduce the technicality of working with the partition function in an octant to working with it in a quadrant, which simplifies many computations. The dichotomy here is that the quadrant has a simple geometry that makes polymer-convergence results of the desired type quite straightforward; on the other hand, the octant has the advantage that one has nice identities such as those of Corollary 3.2.4(3) which fail for a quadrant. Hence, one viewpoint is simpler for technical computations while the other is well-adapted for exact solvability. The results of this section are specific to the case of our positive random walk measures; however, the general outline and arguments that will be given may be easily modified for other random walk measures, such as the reflecting walk, as long as the analogous heat kernel bounds hold. Thus this section may potentially prove useful to other works of a similar flavor.

In what follows, we fix a sequence $\omega^n = \{\omega^n_{i,j}\}_{i,j\geq 0}$ of i.i.d. random environments with $n \in \mathbb{N}$. As always, we denote by $\mathbb{E}$ (resp. $\mathbb{P}$) the expectation (resp. probability) with respect to the environment $\omega^n_{i,j}$ and we denote by $\mathbf{E}^n_x$ (resp. $\mathbf{P}^n_x$) the expectation (resp. probability) with respect to the positive random walk measures of Section 3. Furthermore T_n will denote the first time that this random walk (i, S_i), started from $(0, x)$ with $x \geq 0$, hits the diagonal line $\{(j, 2n-j) : j \geq 0\}$.

First we need an estimate on the variance of the discrete chaos series terms.

Lemma 3.5.1. *Let $\mathfrak{p}^N_n(x,y)$ be the positive random walk transition probabilities given in Definition 3.3.2. Then there exist constants $B, C, K > 0$ such that for all $x, n, k \geq 0$ and $a \geq 0$,*

$$\sum_{\substack{0\leq i_1<\ldots<i_k\leq n\\ (x_1,\ldots,x_k)\in\mathbb{Z}^k_{\geq 0}}} \mathfrak{p}^n_{i_1}(x,x_1)^2\mathfrak{p}^{n-i_1}_{i_2-i_1}(x_1,x_2)^2\cdots\mathfrak{p}^{n-i_{k-1}}_{i_k-i_{k-1}}(x_{k-1},x_k)^2e^{ax_k} \leq Be^{ax+Ka^2n}C^kn^{k/2}/(k/2)!,$$

where $(k/2)!$ is a shorthand for $\Gamma(1+k/2)$.

Proof. We first state a bound, which is Proposition 3.7.3 in the appendix: there exist constants

$C, K > 0$ such that for all $x \geq 0$, all $N \geq n \geq 0$, all $a \geq 0$, and all $p \geq 1$ one has that

$$\sum_{y \geq 0} \mathfrak{p}_n^N(x, y)^p e^{ay} \leq C^p (n+1)^{-(p-1)/2} e^{ax + Ka^2 n}.$$

Applying this k times, one sees that

$$\sum_{(x_1, \ldots, x_k) \in \mathbb{Z}_{\geq 0}^k} \mathfrak{p}_{i_1}^n(x, x_1)^2 \mathfrak{p}_{i_2 - i_1}^{n-i_1}(x_1, x_2)^2 \cdots \mathfrak{p}_{i_k - i_{k-1}}^{n - i_{k-1}}(x_{k-1}, x_k)^2 e^{ax_k}$$

$$\leq C^k e^{ax + Ka^2 n} (i_1 + 1)^{-1/2} (i_2 - i_1 + 1)^{-1/2} \cdots (i_k - i_{k-1} + 1)^{-1/2}.$$

Thus the desired sum is bounded above by

$$e^{ax + Ka^2 n} \sum_{1 \leq i_1 < \ldots < i_k \leq n+1} i_1^{-1/2} (i_2 - i_1)^{-1/2} \cdots (i_k - i_{k-1})^{-1/2}.$$

Now one recognizes that

$$n^{-k/2} \sum_{1 \leq i_1 < \ldots < i_k \leq n+1} i_1^{-1/2} (i_2 - i_1)^{-1/2} \cdots (i_k - i_{k-1})^{-1/2}$$

$$= \frac{1}{n^k} \sum_{1 \leq i_1 < \ldots < i_k \leq n+1} \big(\frac{i_1}{n}\big)^{-1/2} \big(\frac{i_2}{n} - \frac{i_1}{n}\big)^{-1/2} \cdots \big(\frac{i_k}{n} - \frac{i_{k-1}}{n}\big)^{-1/2}, \tag{3.24}$$

which as a Riemann sum approximation is bounded above by (say) twice

$$\int_{0 \leq t_1 < \ldots < t_k \leq 1} t_1^{-1/2} (t_2 - t_1)^{-1/2} \cdots (t_k - t_{k-1})^{-1/2} dt_1 \cdots dt_k \leq B/(k/2)!,$$

where $B > 0$. Hence the lemma is proved. □

Now we use the variance bound in conjunction with Doob's martingale inequality to get a bound on the expected supremum in the partition function.

Lemma 3.5.2. *Take a sequence $\omega^n = \{\omega_{i,j}^n\}$ of random environments with variance uniformly* *bounded above by* 1*. Furthermore let* $\{z_0^n(x)\}_{x\geq 0}$ *be some sequence of non-negative stochastic processes, independent of the* ω^n*, with the property that* $\mathbb{E}[z_0^n(x)^2]\leq Ke^{an^{-1/2}x}$ *for some constants* K,a *that are independent of* n *and* x*. Then there exists a constant* C *such that for all* $n,x\geq 0$ *one has that*

$$\mathbb{E}\bigg[\sup_{0\leq k\leq n}\mathbf{E}_x^n\bigg[z_0^n(S_n)\prod_{i=0}^k(1+n^{-1/4}\omega_{i,S_i}^n)\bigg]^2\bigg]\leq Ce^{an^{-1/2}x}.$$

Proof. First we fix some $n\in\mathbb{N}$ and we note that the process

$$M_k^n:=\mathbf{E}_x^n\bigg[z_0^n(S_n)\prod_{i=1}^k(1+n^{-1/4}\omega_{i,S_i}^n)\bigg]$$

is a $\mathbb{P}$-martingale in the k variable with respect to the filtration $(\mathcal{F}_k^n)_{k\geq 0}$, where $\mathcal{F}_k^n$ is generated by z_0^n and $\{\omega_{i,j}^n\}_{0\leq j\leq i\leq k}$. Therefore by Doob's martingale inequality $\mathbb{E}[\sup_{0\leq k\leq n}(M_k^n)^2]\leq 4\mathbb{E}[(M_n^n)^2]$. This reduces our work to proving the claim without the supremum inside the expectation (and replacing k by n in the upper limit of the product). To do this, we set $x_0:=x$ and we write

$$\begin{aligned}
&\mathbb{E}\bigg[\mathbf{E}_x^n\bigg[z_0^n(S_n)\prod_{i=1}^n(1+n^{-1/4}\omega_{i,S_i}^n)\bigg]^2\bigg]\\
&=\mathbb{E}\bigg[\bigg(\sum_{\substack{0\leq k\leq n\\ 0\leq i_1<\ldots<i_k\leq n\\ (x_1,\ldots,x_{k+1})\in\mathbb{Z}_{\geq 0}^{k+1}}} n^{-k/4}z_0^n(x_{k+1})\prod_{j=1}^k\mathfrak{p}_{i_j-i_{j-1}}^{n-i_{j-1}}(x_{j-1},x_j)\omega_{i_jx_j}\cdot\mathfrak{p}_{n-i_k}^{n-i_k}(x_k,x_{k+1})\bigg)^2\bigg]\\
&=\sum_{\substack{0\leq k\leq n\\ 0\leq i_1<\ldots<i_k\leq n\\ (x_1,\ldots,x_k)\in\mathbb{Z}_{\geq 0}^{k}}} n^{-k/2}\prod_{j=1}^k\mathfrak{p}_{i_j-i_{j-1}}^{n-i_{j-1}}(x_{j-1},x_j)^2\mathbb{E}\bigg[\bigg(\sum_{x_{k+1}\in\mathbb{Z}_{\geq 0}}z_0^n(x_{k+1})\mathfrak{p}_{n-i_k}^{n-i_k}(x_k,x_{k+1})\bigg)^2\bigg],
\end{aligned}$$

where $x_0:=x$ by convention. By Jensen we then have that

$$\bigg(\sum_{x_{k+1}\geq 0}z_0^n(x_{k+1})\mathfrak{p}_{n-i_k}^{n-i_k}(x_k,x_{k+1})\bigg)^2\leq\sum_{x_{k+1}\geq 0}z_0^n(x_{k+1})^2\mathfrak{p}_{n-i_k}^{n-i_k}(x_k,x_{k+1}).$$

We know by assumption that $\mathbb{E}[z_0^n(x_{k+1})^2]\leq e^{an^{-1/2}x_{k+1}}$. Thus we find that the expectation of the last expression is bounded above by $Ce^{an^{-1/2}x_k}$ because of the inequality $\sum_{y\geq 0}\mathfrak{p}_n^N(x,y)e^{ay} \leq Ce^{ax+Ka^2n}$, which holds by Proposition 3.7.1 in the appendix. Thus by Lemma 3.5.1 we have

$$\begin{aligned}
\mathbb{E}\Bigg[\mathbf{E}_x^n\Big[z_0^n(S_n)\prod_{i=1}^n(1+n^{-1/4}\omega_{i,S_i}^n)\Big]^2\Bigg] &\leq \sum_{\substack{0\leq k\leq n\\ 0\leq i_1<\ldots<i_k\leq n\\ (x_1,\ldots,x_k)\in\mathbb{Z}_{\geq 0}^k}} n^{-k/2}\prod_{j=1}^k \mathfrak{p}_{i_j-i_{j-1}}^{n-i_{j-1}}(x_{j-1},x_j)^2\cdot Ce^{an^{-1/2}x_k} \\
&\leq \sum_{k=0}^n n^{-k/2}BC^{k+1}e^{an^{-1/2}x}n^{k/2}/(k/2)! \\
&\leq Be^{an^{-1/2}x}\sum_{k=0}^\infty C^{k+1}/(k/2)!.
\end{aligned}$$

This completes the proof. $\square$

We now introduce a class of Banach spaces that will be useful for describing convergence of initial data:

Definition 3.5.3. *Let $\alpha,\delta\in(0,1)$. A function $f:\mathbb{R}\to\mathbb{R}$ is said to be in the exponentially δ-weighted α-Hölder space $\mathscr{C}_{e(\delta)}^\alpha(\mathbb{R})$ if*

$$\sup_{x\in\mathbb{R}}\frac{|f(x)|}{e^{\delta|x|}}+\sup_{\substack{x,y\in\mathbb{R}\\ |x-y|\leq 1}}\frac{|f(x)-f(y)|}{e^{\delta|x|}|x-y|^\alpha}<\infty.$$

We turn $\mathscr{C}_\delta^\alpha$ into a Banach space by defining the norm of f to be the above quantity.

A straightforward consequence of Arzela-Ascoli is that $\mathscr{C}_{e(\delta)}^\alpha$ embeds compactly into $\mathscr{C}_{e(\delta')}^{\alpha'}$ for $\alpha'<\alpha$ and $\delta<\delta'$. The key estimate of this section is as follows:

Theorem 3.5.4 (Key Estimate)**.** *Fix $\alpha\in(0,1)$. Suppose that $(z_0^n(x))_{x\in\mathbb{Z}_{\geq 0}}$ is a family of deterministic non-negative functions such that the linearly interpolated and rescaled family $z_0^n(n^{1/2}x)$ are bounded with respect to the norm of $\mathscr{C}_{e(\delta)}^\gamma$ for some $\gamma,\delta\in(0,1)$. Define the "error" random variable*

$$\mathcal{E}(x,n):=\sup_{k\in[n-n^\alpha,n]}\Bigg|\mathbf{E}_x^n\Big[z_0^n(S_n)\prod_{i=1}^n\big(1+n^{-1/4}\omega_{i,S_i}^n\big)-z_0^n(S_k)\prod_{i=1}^k\big(1+n^{-1/4}\omega_{i,S_i}^n\big)\Big]\Bigg|.$$

Then $\sup_{x\geq 0} e^{-3an^{-1/2}x}\mathbb{E}[\mathcal{E}(x,n)] \to 0$ *as* $n\to\infty$.

Proof. By the triangle inequality, we have $\mathcal{E}(x,n) \leq \mathcal{E}_1(x,n)+\mathcal{E}_2(x,n)$, where

$$\mathcal{E}_1(x,n) := \sup_{k\in[n-n^\alpha,n]} \left| \mathbf{E}_x^n\left[z_0^n(S_n)\bigg(\prod_{i=1}^n \big(1+n^{-1/4}\omega_{i,S_i}^n\big) - \prod_{i=1}^k \big(1+n^{-1/4}\omega_{i,S_i}^n\big)\bigg)\right]\right|,$$

$$\mathcal{E}_2(x,n) := \sup_{k\in[n-n^\alpha,n]} \left| \mathbf{E}_x^n\left[\big(z_0^n(S_n) - z_0^n(S_k)\big) \prod_{i=1}^k \big(1+n^{-1/4}\omega_{i,S_i}^n\big)\right]\right|.$$

We separately show that both of these satisfy the desired bound. Henceforth when we write n^α we actually mean $\lceil n^\alpha\rceil$.

First we consider $\mathcal{E}_1$. First we establish a martingale inequality. If $(M_k)_{k\geq 0}$ is a martingale defined on any probability space, then note that for $r\leq n$ one has that

$$\sup_{r\leq k\leq n} |M_n - M_k| \leq |M_n - M_r| + \sup_{r\leq k\leq n}|M_k - M_r|,$$

and by Doob's inequality one has that $\big\|\sup_{r\leq k\leq n}|M_k-M_r|\big\|_p \leq \frac{p}{p-1}\|M_n-M_r\|_p$, therefore one has that

$$\big\|\sup_{r\leq k\leq n}|M_n-M_k|\big\|_p \leq \|M_n-M_r\|_p + \frac{p}{p-1}\|M_n-M_r\|_p = \frac{2p-1}{p-1}\|M_n-M_r\|_p. \tag{3.25}$$

Now let us fix some $n\in\mathbb{N}$. Let us define a martingale

$$M_k^n := \mathbf{E}_x^n\left[z_0^n(S_n)\prod_{i=1}^k (1+n^{-1/4}\omega_{i,S_i}^n)\right].$$

This is a $\mathbb{P}$-martingale in the k variable, for fixed $n\in\mathbb{N}$. Consequently, using (3.25) with $p=2$ gives us

$$\mathbb{E}\big[\sup_{k\in[n-n^\alpha,n]}(M_n-M_k)^2\big] \leq 9\mathbb{E}[(M_n^n - M_{n-n^\alpha}^n)^2]. \tag{3.26}$$

Computing the right-hand side, one gets

$$\mathbb{E}[(M_n^n - M_{n-n^\alpha}^n)^2] = \mathbb{E}\Bigg[\mathbf{E}_x^n\Bigg[z_0^n(S_n)\Bigg(\prod_{i=1}^{n}\big(1+n^{-1/4}\omega_{i,S_i}^n\big) - \prod_{i=1}^{n-n^\alpha}\big(1+n^{-1/4}\omega_{i,S_i}^n\big)\Bigg)\Bigg]^2\Bigg]$$
$$= \sum_{\substack{0\le k\le n-n^\alpha \\ 0\le i_1<\ldots<i_k\le n-n^\alpha \\ (x_1,\ldots,x_k)\in\mathbb{Z}_{\ge 0}^k}} n^{-k/2}\cdot \mathfrak{p}_{i_1}^n(x,x_1)^2\prod_{j=1}^{k-1}\mathfrak{p}_{i_j-i_{j-1}}^{n-i_{j-1}}(x_j,x_{j+1})^2\mathfrak{F}_n(i_k,x_k), \tag{3.27}$$

where $\mathfrak{F}_n(i_k,x_k)$ is given by

$$\sum_{\substack{1\le \ell\le n^\alpha \\ 0\le j_1<\ldots<j_\ell\le n^\alpha \\ (u_1,\ldots,u_\ell)\in\mathbb{Z}_{\ge 0}^\ell}} n^{-\ell/2}\mathfrak{p}_{n-n^\alpha+j_1-i_k}^n(x_k,u_1)^2$$
$$\cdot\prod_{v=1}^{\ell-1}\mathfrak{p}_{j_v-j_{v-1}}^{n^\alpha-j_{v-1}}(u_j,u_{j+1})^2\Bigg(\sum_{u_{\ell+1}\ge 0} z_0^n(u_{\ell+1})\mathfrak{p}_{n^\alpha-j_\ell}^{n^\alpha-j_\ell}(u_\ell,u_{\ell+1})\Bigg)^2.$$

Note that the latter sum starts at $\ell=1$ rather than $\ell=0$ which is crucial. These expressions come from writing

$$\prod_{i=1}^{n}\big(1+n^{-1/4}\omega_{i,S_i}^n\big) - \prod_{i=1}^{n-n^\alpha}\big(1+n^{-1/4}\omega_{i,S_i}^n\big)$$
$$= \prod_{k=1}^{n-n^\alpha}(1+n^{-1/4}\omega_{k,S_k}^n)\Bigg(\prod_{\ell=1}^{n^\alpha}(1+n^{-1/4}\omega_{\ell+n-n^\alpha,S_{\ell+n-n^\alpha}}^n)\ -\ 1\Bigg),$$

and then expanding both products and taking expectations. The subtraction of 1 from the second product is what causes the sum defining $\mathfrak{F}_n$ to start at $\ell=1$ rather than $\ell=0$. By Jensen and the fact that $z_0^n(x)\le Ce^{an^{-1/2}x}$ (with say $a=2\delta$ where δ is the same as in the theorem statement) we then have that

$$\Bigg(\sum_{u_{\ell+1}\ge 0} z_0^n(u_{\ell+1})\mathfrak{p}_{n^\alpha-j_\ell}^{n^\alpha-j_\ell}(u_\ell,u_{\ell+1})\Bigg)^2 \le \sum_{u_{\ell+1}\ge 0} z_0^n(u_{\ell+1})^2\mathfrak{p}_{n^\alpha-j_\ell}^{n^\alpha-j_\ell}(u_\ell,u_{\ell+1}) \le Ce^{an^{-1/2}u_\ell},$$

where we used Proposition 3.7.1 in the last bound. Then by repeatedly applying Proposition 3.7.3, note that $\mathfrak{F}_n(i_k,x_k)$ is bounded above by

$$\sum_{\ell=1}^{n^\alpha+1} n^{-\ell/2}C^\ell e^{an^{-1/2}x_k}\sum_{1\le j_1<\ldots<j_\ell\le n^\alpha+1}(n-n^\alpha+j_1-i_k)^{-1/2}(j_2-j_1)^{-1/2}\cdots(j_\ell-j_{\ell-1})^{-1/2}.$$

Consequently the entirety of (3.27) is bounded above, after again applying Proposition 3.7.3 several more times, by

$$\sum_{\substack{0\le k\le n-n^\alpha\\ 1\le i_1<\ldots<i_k\le n-n^\alpha+1\\ 1\le \ell\le n^\alpha\\ 1\le j_1<\ldots<j_\ell\le n^\alpha+1}} n^{-(k+\ell)/2}C^{k+\ell}e^{an^{-1/2}x}i_1^{-1/2}$$
$$\cdot\prod_{r=1}^{k-1}(i_r-i_{r-1})^{-1/2}(n-n^\alpha+j_1-i_k)^{-1/2}\prod_{v=1}^{\ell-1}(j_v-j_{v-1})^{-1/2}.$$

We rewrite that as $e^{an^{-1/2}x}$ multiplied by

$$\sum_{\substack{0\le k\le n-n^\alpha\\ 1\le i_1<\ldots<i_k\le n-n^\alpha+1\\ 1\le \ell\le n^\alpha\\ 1\le j_1<\ldots<j_\ell\le n^\alpha+1}} n^{-\ell(1-\alpha)/2}n^{-k}n^{-\ell\alpha}C^{k+\ell}(\frac{i_1}{n})^{-1/2}$$
$$\cdot\prod_{r=1}^{k-1}(\frac{i_r}{n}-\frac{i_{r-1}}{n})^{-1/2}(\frac{n-n^\alpha+j_1-i_k}{n^\alpha})^{-1/2}\prod_{v=1}^{\ell-1}(\frac{j_v}{n^\alpha}-\frac{j_{v-1}}{n^\alpha})^{-1/2}.$$

Except for the factor $n^{-\ell(1-\alpha)/2}$ we recognize a Riemann sum approximation for

$$\sum_{\substack{k\ge0\\ \ell\ge1}}C^{k+\ell}\int_{0\le t_1<\ldots<t_k\le1}\int_{t_k\le s_1<\ldots<s_\ell\le1}t_1^{-1/2}\cdots(t_k-t_{k-1})^{-1/2}(s_1-t_k)^{-1/2}\cdots(s_\ell-s_{\ell-1})^{-1/2}d\mathbf{t}d\mathbf{s}.$$

This series may be bounded by

$$\sum_{\substack{k\ge0\\ \ell\ge1}}C^{k+\ell}/\big((k+\ell)/2\big)!$$

which converges absolutely to a constant independently of n. Since $\ell\ge1$ in all expressions above, the left over factor $n^{-\ell(1-\alpha)/2}$ is at worst $n^{-(1-\alpha)/2}$. Summarizing the bounds, we showed that $\mathbb{E}[\mathcal{E}_1(n,x)]$ is bounded above by at worst $Ce^{an^{-1/2}x}n^{-(1-\alpha)/2}$ which implies the desired result on $\mathcal{E}_1$.

Now we consider $\mathcal{E}_2(x,n)$. Since z_0^n is bounded in $\mathscr{C}^{\gamma}_{e(\delta)}$ we have the following bound with C independent of x,y,n:

$$|z_0^n(n^{1/2}x)-z_0^n(n^{1/2}y)|\leq C|x-y|^{\gamma}e^{\delta(x+y)}.$$

Using positivity of $B_k^n:=\prod_1^k(1+n^{-1/2}\omega^n_{i,S_i})$ we then find that

$$\begin{aligned}\mathcal{E}_2(x,n)&\leq \sup_{k\in[n-n^{\alpha},n]}\mathbf{E}_x^n\Big[\big|z_0^n(S_n)-z_0^n(S_k)\big|B_k^n\Big]\\&\leq Cn^{-\gamma/2}\sup_{k\in[n-n^{\alpha},n]}\mathbf{E}_x^n\Big[|S_n-S_k|^{\gamma}e^{\delta n^{-1/2}(S_n+S_k)}B_k^n\Big]\\&=Cn^{-\gamma/2}\sup_{k\in[n-n^{\alpha},n]}\mathbf{E}_x^n\Big[\mathbf{E}^{n-k}_{S_k}[|\tilde{S}_{n-k}-\tilde{S}_0|^{\gamma}e^{\delta n^{-1/2}\tilde{S}_{n-k}}]e^{\delta n^{-1/2}S_k}B_k^n\Big],\end{aligned}$$

where the final equality follows from the Markov property of the positive random walk S. Now we recognize that

$$\mathbf{E}_y^N[|S_N-S_0|^{\gamma}e^{\delta S_N}]\leq \mathbf{E}_y^N[|S_N-S_0|^{2\gamma}]^{1/2}\mathbf{E}_y^N[e^{2\delta S_N}]^{1/2}\leq CN^{\gamma/2}e^{\delta y+K\delta^2N},$$

where C,K are independent of y,N, by Propositions 3.6.10 and 3.7.1. Consequently we find for $k\in[n-n^{\alpha},n]$ that

$$\mathbf{E}^{n-k}_{S_k}[|\tilde{S}_{n-k}-\tilde{S}_0|^{\gamma}e^{\delta n^{-1/2}\tilde{S}_{n-k}}]\leq C(n-k)^{\gamma/2}e^{\delta n^{-1/2}S_k}\leq Cn^{\alpha\gamma/2}e^{\delta n^{-1/2}S_k}.$$

Combining our bounds, we find that

$$\mathcal{E}_2(x,n)\leq Cn^{-(1-\alpha)\gamma/2}\sup_{k\in[n-n^{\alpha},n]}\mathbf{E}_x^n[e^{2\delta n^{-1/2}S_k}B_k^n]. \tag{3.28}$$

Now for any $\lambda>0$, $(e^{\lambda S_k})_k$ is a $\mathbf{P}_x^n$-submartingale because (S_k) is a submartingale (Lemma 3.6.3) and since $x \mapsto e^{\lambda x}$ is increasing and convex for any λ. Thus letting $\mathcal{G}_k$ denote the filtration generated by the first k steps of the n-step positive random walk S, we find

$$\mathbf{E}_x^n[e^{\lambda S_k} B_k^n] \leq \mathbf{E}_x^n[\mathbf{E}_x^n[e^{\lambda S_n}|\mathcal{G}_k] B_k^n] = \mathbf{E}_x^n[e^{\lambda S_n} B_k^n],$$

for all $k \leq n, \lambda > 0$ because B_k^n is $\mathcal{G}_k$-measurable. Setting $\lambda = 2\delta n^{-1/2}$, this means that

$$\mathbb{E}\left[\sup_{k\leq n} \mathbf{E}_x^n[e^{2\delta n^{-1/2} S_k} B_k^n]^2\right] = \mathbb{E}[\sup_{k\leq n} \mathbf{E}_x^n[e^{2\delta n^{-1/2} S_n} B_k^n]^2] \leq C e^{2\delta n^{-1/2} x}, \tag{3.29}$$

where we used Lemma 3.5.2 in the last bound, with $z_0^n(x) := e^{2\delta n^{-1/2} x}$. Combining (3.28) and (3.29) gives the required result. $\square$

Next we give some Kolmogorov-type moment conditions that ensure tightness of the sequence z_0^n of initial data in $\mathscr{C}^{\alpha}_{e(\delta)}$.

Proposition 3.5.5. *Suppose that $\{z^n\}_{n\geq 1}$ is a family of random functions on $\mathbb{R}$ that satisfies the following moment conditions for some constants a, p, β, C independent of n, x, y.*

- $\mathbb{E}[|z^n(x) - z^n(y)|^p] \leq C|x-y|^{p\beta/2} e^{a(|x|+|y|)}$.
- *there exist positive integrable random variables $D(n)$ such that $\sup_n \mathbb{E}[D(n)] < \infty$ and $z^n(x) \leq D(n) e^{a|x|}$.*

Then assuming $p > 1/\beta$, there exist $\delta > a$ and $\alpha < \beta - p^{-1}$ such that (z^n) is tight with respect to the topology of $\mathscr{C}^{\alpha}_{e(\delta)}$.

Before the proof, we remark that when we apply this result, the functions will be defined on $\mathbb{R}_+$ as opposed to all of $\mathbb{R}$ and thus the absolute values on x, y are unnecessary. Furthermore, the z^n appearing in the proposition statement will actually be rescaled and linearly interpolated functions $z_0^n(n^{1/2}x)$.

Proof. Recall from earlier that $\mathscr{C}^{\alpha}_{e(\delta)}$ embeds *compactly* into $\mathscr{C}^{\alpha'}_{e(\delta')}$ whenever $\delta' > \delta$ and $\alpha' < \alpha$. Therefore to prove the lemma, it suffices to show that if the two inequalities in the lemma statement hold uniformly over a family $\mathcal{F}$ of real-valued functions, then there exist α, δ such that

$$\lim_{a\to\infty} \sup_{z\in\mathcal{F}} \mathbb{P}(\|z\|_{\mathscr{C}^\alpha_{e(\delta)}} > a) = 0.$$

We actually show something stronger, namely that under the given assumptions, there exists $C > 0$ such that for all $a > 0$

$$\sup_{n\in\mathbb{N}} \mathbb{P}(\|z^n\|_{\mathscr{C}^\alpha_{e(\delta)}} > a) \le Ca^{-1}. \tag{3.30}$$

To prove this, for a function z we write $\|z\|_{\mathscr{C}^\alpha_{e(\delta)}} = \|z\|_\delta + [z]_{\alpha,\delta}$ where $\|z\|_\delta := \sup_{x\in\mathbb{R}} \frac{|z(x)|}{e^{\delta|x|}}$ and $[z]_{\alpha,\delta} := \sup_{x\in\mathbb{R}} e^{-\delta|x|} \sup_{|y-x|\le 1} \frac{|z(x)-z(y)|}{|x-y|^\alpha}$.

To prove (3.30), the following fact will be useful to us: For any $\gamma \in (0,1)$, the γ-Hölder seminorm $[f]_\gamma$ of a function $f : [0,1] \to \mathbb{R}$ is equivalent (as a seminorm) to the quantity given by $\sup_{v\in\mathbb{N},1\le k\le 2^v} 2^{\gamma v}|f(k2^{-v}) - f((k-1)2^{-v})|$. This is proved as an intermediate step in the standard proof of the classical Kolmogorov-Chentsov criterion.

The exact choices of α, δ will be specified later, but for now let them denote generic constants. Now to prove (3.30) let us write for a function z,

$$\begin{aligned}
\|z\|_\delta &\le \sup_{v\in\mathbb{Z}} e^{-\delta|v|}\big(|z(v)| + \sup_{x\in[v,v+1]} |z(x) - z(v)|\big) \\
&\le \sup_{v\in\mathbb{Z}} e^{-\delta|v|}\left(|z(v)| + \sup_{x\in[v,v+1]} \frac{|z(x) - z(v)|}{|x-v|^\alpha}\right) \\
&\lesssim \sup_{v\in\mathbb{Z}} e^{-\delta|v|}\left(|z(v)| + \sup_{r\in\mathbb{N},1\le k\le 2^r} 2^{\alpha r}|z(v + k2^{-r}) - z(v + (k-1)2^{-r})|\right),
\end{aligned}$$

where $\lesssim$ denotes the absorption of some universal constant which can depend on α, δ but not on the function z. Likewise let us note that

$$[z]_{\alpha,\delta} \lesssim \sup_{v\in\mathbb{Z}} e^{-\delta|v|} \sup_{r\in\mathbb{N},1\le k\le 2^r} 2^{\alpha r}|z(v + k2^{-r}) - z(v + (k-1)2^{-r})|.$$

Consequently we find that

$$\|z\|_{\mathscr{C}_\delta^\alpha} \lesssim A(z,\delta) + B(z,\alpha,\delta),$$

where

$$\begin{aligned}
A(z,\delta) &:= \sup_{v\in\mathbb{Z}} e^{-\delta|v|}|z(v)|, \\
B(z,\alpha,\delta) &:= \sup_{v\in\mathbb{Z}} e^{-\delta|v|} \sup_{r\in\mathbb{N}, 1\le k\le 2^r} 2^{\alpha r}|z(v+k2^{-r}) - z(v+(k-1)2^{-r})|.
\end{aligned}$$

Now, with z^n uniformly satisfying the bounds given in the lemma statement, let us bound these terms $A(z^n,\delta)$ and $B(z^n,\alpha,\delta)$ individually to obtain (3.30). We will do this by using the hypotheses in the lemma. Note that by a brutal union bound and Markov's inequality followed by the hypothesis $z^n \le D(n)e^{an^{-1/2}x}$, we have

$$\begin{aligned}
\mathbb{P}(A(z^n,\delta) > a) &\le \sum_{v\in\mathbb{Z}} \mathbb{P}(|h^\epsilon(v)| > e^{\delta|v|}a) \\
&\le \sum_{v\in\mathbb{Z}} a^{-1}e^{-\delta|v|}\mathbb{E}[z^n(v)] \\
&\le \sup_j \mathbb{E}[D(j)] \cdot a^{-1} \sum_{v\in\mathbb{Z}} e^{-(\delta-a)|v|},
\end{aligned}$$

The series converges to a finite value independent of n as long as δ is chosen larger than a. Next we control B, which will also just use a brutal union bound and Markov's inequality:

$$\begin{aligned}\mathbb{P}(B(z^n,\alpha,\delta)>a) &\le \sum_{\substack{v\in\mathbb{Z}\\ r\in\mathbb{N}\\ 1\le k\le 2^r}} \mathbb{P}(2^{\alpha r}|z^n(v+k2^{-r})-z^n(v+(k-1)2^{-r})|>e^{\delta|v|}a)\\ &\le \sum_{\substack{v\in\mathbb{Z}\\ r\in\mathbb{N}\\ 1\le k\le 2^r}} a^{-p}2^{\alpha pr}e^{-\delta|v|p}\mathbb{E}\big|z^n(n+k2^{-r})-z^n(v+(k-1)2^{-r})\big|^p\\ &\le a^{-p}\sum_{\substack{v\in\mathbb{Z}\\ r\in\mathbb{N}\\ 1\le k\le 2^r}} 2^{(\alpha-\beta)pr}e^{(2a-\delta)p|v|}\\ &= a^{-p}\sum_{\substack{v\in\mathbb{Z}\\ r\in\mathbb{N}}} 2^{\big[1+(\alpha-\beta)p\big]r}e^{(2a-\delta)p|v|}\end{aligned}$$

The double series converges to a finite value independent of n so long as δ,α are chosen so as to satisfy $\delta>2a$ and $1+(\alpha-\beta)p<0$. This is permissible so long as $p>\beta^{-1}$. $\square$

Lemma 3.5.6. *Let $(X_n)_{n\ge0}$ be a non-negative L^1 supermartingale. Then*

$$\mathbb{P}\big(\sup_n X_n>a\big)\le\frac{\mathbb{E}[X_0]}{a}.$$

Proof. We apply Doob-Meyer decomposition to write $X=M-A$, where M is a martingale with $M_0=X_0$, and A_0 is a non-decreasing process with $A_0=0$. Then M is a positive martingale and $X\le M$. Doob's first martingale inequality then gives

$$\mathbb{P}\big(\sup_{n\le N}X_n>a\big)\le\mathbb{P}\big(\sup_{n\le N}M_n>a\big)\le\frac{\mathbb{E}[M_N]}{a}=\frac{\mathbb{E}[M_0]}{a}.$$

Since $M_0=X_0$, letting $N\to\infty$ gives the claim because the right side does not depend on N and the left side approaches $\mathbb{P}\big(\sup_n X_n>a\big)$ by monotone convergence. $\square$

Proposition 3.5.7. *For each $n\in\mathbb{N}$, let $\{\omega^n_{i,0}\}_{i\ge1}$ be a family of i.i.d. random variables such that $\omega^n_{i,0}$ has finite p^{th} moment, with $p>2$. Also assume that $1+n^{-1/4}\omega^n_{i,0}>0$ a.s. and that $\sup_n\mathbb{E}[|\omega^n_{1,0}|^p]<\infty$. Furthermore assume that $\mathbb{E}[\omega^n_{i,0}]=\mu n^{-1/4}+o(n^{-1/4})$ and var$(\omega^n_{i,0})=$*

$\sigma^2 + o(1)$ as $n \to \infty$. Define $z_0^n(x) := \prod_{i=1}^x (1 + n^{-1/4}\omega_{i,0}^n)$. Then z_0^n satisfies the first two conditions of Proposition 3.5.5:

- *$\mathbb{E}[|z_0^n(x) - z_0^n(y)|^p] \le Cn^{-p/4}|x-y|^{p/2}e^{an^{-1/2}(x+y)}$ for some constants C, a independent of n, x, y.*
- *with the same a, there exist square-integrable random variables $D(n)$ such that $\sup_n \mathbb{E}[D(n)^2] < \infty$ and $z_0^n(x) \le D(n)e^{an^{-1/2}x}$ for all n, x almost surely.*

Proof. Before proving either bullet point, we prove a preliminary bound. By Taylor expanding u^p near $u = 1$ we see $(1 + n^{-1/4}\omega_{i,0}^n)^p = 1 + pn^{-1/4}\omega_{i,0}^n + \frac{1}{2}(p^2 - p)n^{-1/2}(\omega_{i,0}^n)^2 + o(n^{-1/2})$, which has expectation roughly $1 + n^{-1/2}(p\mu + \frac{p^2-p}{2}\sigma^2) + o(n^{-1/2})$. For some $a = a(p)$ this is bounded above by $1 + an^{-1/2}$, and so we see that

$$\mathbb{E}[z_0^n(x)^p] = \prod_{i=1}^x \mathbb{E}[(1 + n^{-1/4}\omega_{i,0}^n)^p] \le (1 + n^{-1/2}a)^x \le e^{an^{-1/2}x}, \tag{3.31}$$

since $1 + v \le e^v$. With this preliminary bound in mind, we proceed to the proof of the first bullet point. It suffices to prove the claim when $y = 0$ (i.e., $z_0^n(y) = 1$), by independence of the multiplicative increments of z_0^n. Let us begin by writing

$$\mathbb{E}[|z_0^n(x) - 1|^p] \le 2^p\bigg(\mathbb{E}\bigg[\bigg|z_0^n(x) - \frac{z_0^n(x)}{\mathbb{E}[z_0^n(x)]}\bigg|^p\bigg] + \mathbb{E}\bigg[\bigg|\frac{z_0^n(x)}{\mathbb{E}[z_0^n(x)]} - 1\bigg|^p\bigg]\bigg).$$

Let us denote these expectations on the right side as E_1 and E_2, respectively. We bound each of these separately. For E_1, one notes by using (3.31) that

$$\begin{aligned} E_1 = \mathbb{E}[z_0^n(x)^p]\bigg|1 - \frac{1}{\mathbb{E}[z_0^n(x)]}\bigg|^p &\le e^{an^{-1/2}x}\big|1 - e^{-an^{1/2}x}\big|^p \\ &\le e^{an^{-1/2}x}\big(an^{-1/2}x\big)^p = a^p e^{an^{-1/2}x}n^{-p/2}x^p, \end{aligned}$$

where we used $\mathbb{E}[z_0^n(x)] \le \mathbb{E}[z_0^n(x)^p]^{1/p} \le e^{an^{-1/2}x}$ (by (3.31)) in the first inequality, and we used $1 - e^{-v} \le v$ in the second one. Finally, note that $u^pe^u \le Cu^{p/2}e^{2u}$ for some $C > 0$ independent of u, and applying this with $u = an^{-1/2}x$ already gives the desired bound on E_1.

Now we bound E_2. This is the difficult part, and one needs to somehow exploit cancellations that occur at the quadratic scale (e.g., via a Burkholder-type inequality). To do this, first note that the process $M_x^n := \frac{z_0^n(x)}{\mathbb{E}[z_0^n(x)]}$ is a martingale in the x-variable (for fixed n). Define $\zeta_i^n := \frac{1+n^{-1/4}\omega_{i,0}^n}{\mathbb{E}[1+n^{-1/4}\omega_{i,0}]}$. Then Burkholder-Davis-Gundy says

$$E_2 \le C\mathbb{E}\bigg[\bigg(\sum_{i=1}^x (M_i^n - M_{i-1}^n)^2\bigg)^{p/2}\bigg] = C\mathbb{E}\bigg[\bigg(\sum_{i=1}^x (\zeta_1^n)^2 \cdots (\zeta_{i-1}^n)^2(\zeta_i^n - 1)^2\bigg)^{p/2}\bigg] \tag{3.32}$$

Now, using the given conditions, $|\zeta_i^n - 1|$ is easily seen to be bounded above by $C(n^{-1/4}|\omega_{i,0}^n| + n^{-1/2})$, so the square is bounded by $C(n^{-1/2}(\omega_{i,0}^n)^2 + n^{-1})$. Writing $\|A\|_p := \mathbb{E}[|A|^p]^{1/p}$, we notice by the triangle inequality and independence of ζ_i^n that

$$\bigg\|\sum_{i=1}^x (\zeta_1^n)^2 \cdots (\zeta_{i-1}^n)^2(\zeta_i^n - 1)^2\bigg\|_{p/2} \le Cn^{-1/2}\sum_{i=1}^x \|(\zeta_1^n)^2\|_{p/2}\cdots\|(\zeta_{i-1}^n)^2\|_{p/2}\|(\omega_{i,0}^n)^2 + n^{-1/2}\|_{p/2}.$$

Now, it holds that $\|(\zeta_1^n)^2\|_{p/2} \le e^{2an^{-1/2}/p}$, by (3.31) (with $x = 1$). Hence each term of the sum can be bounded above by $e^{2an^{-1/2}x/p}$. The contribution of the $n^{-1/2}$ term next to $(\omega_{i,0}^n)^2$ is then seen to be negligible, so we disregard it. Hence the the entire sum may be bounded by $Cn^{-1/2}xe^{2an^{-1/2}x/p}$, which, combined with (3.32) and the fact that $\|(\omega_{i,0}^n)^2\|_{p/2}$ is bounded independently of n by assumption, completes the proof.

Now we prove the second bullet point. Note that $\frac{z_0^n(x)^p}{\mathbb{E}[z_0^n(x)^p]}$ is a positive martingale in the x-variable. Let $D(n) := \sup_{x\ge 0} z_0^n(x)/\mathbb{E}[z_0^n(x)^p]^{1/p}$. Then it is clear from Lemma 3.5.6 that $\mathbb{P}(D(n)^p > a) \le a^{-1}$, so that $\mathbb{P}(D(n) > a) \le a^{-p}$. If $p > 2$, then this easily implies that $\sup_n \mathbb{E}[D(n)^2] < \infty$. But (3.31) tells us that $\mathbb{E}[z_0^n(x)^p]^{1/p} \le Ce^{an^{-1/2}x}$, so we are done. $\square$

Next, we finally prove the octant-quadrant reduction theorem, i.e., that we can replace T_n with $2n$

as discussed in the proof sketch at the end of Section 2. Let us reformulate the main notational conventions here:

- S is a simple symmetric random walk of length n started from x and conditioned to stay positive throughout its course (i.e., the canonical process associated to the measures $\mathbf{P}_x^n$). We assume $n-x$ is even.
- $\hat{\omega}_{i,j}^n$ is defined to be $\omega^n_{(n-\frac{i-j}{2}),(n-\frac{i+j}{2})}$ for all i,j of the same parity, where $\omega_{i,j}^n$ is a family of random environments satisfying the conditions of the three bullet points before Theorem 3.2.2, but now the bulk random variables are indexed by all pairs (i,j) with $|i|\geq j$.
- T_n is the first time that $n-i=S_i$.
- $z_0^n(x) := \prod_{i=0}^x (1+n^{-1/4}\bar{\omega}_{i,0}^n)$, where the $\bar{\omega}_{i,0}$ have $p>2$ moments.

We remark that all conditions of Theorem 3.5.4 are *almost* satisfied by this environment. The only caveat is that the sequence of initial data is not deterministic, however by Propositions 3.5.5 and 3.5.7 and Skorohod's Lemma (and the fact that z_0^n are independent of the bulk weights) we may choose a probability space on which $z_0^n \to z_0$ almost surely with respect to the topology of $\mathscr{C}^\alpha_{e(\delta)}$ for some choice of $\alpha,\delta\in(0,1)$. Here $z_0(x)$ is a geometric Brownian motion with the appropriate diffusion and drift coefficients. Note that a.s. convergence is stronger than a.s. boundedness in that norm which is the condition required in Theorem 3.5.4. Thus there is no loss of generality in assuming that the initial data are in fact deterministic.

Proposition 3.5.8 (Octant-Quadrant Reduction)**.** *In the notation of the bullet points immediately above, we define the following random variable for* $n,x\geq 0$*:*

$$\mathscr{E}(x,n) := \mathbf{E}_x^n\bigg[z_0^n(S_n)\prod_{i=0}^{n-1}(1+n^{-1/4}\hat{\omega}_{i,S_i}^n)\bigg] - \mathbf{E}_x^n\bigg[z_0^n(S_{T_n})\prod_{i=0}^{T_n-1}(1+n^{-1/4}\hat{\omega}_{i,S_i}^n)\bigg].$$

Let x_n *be a sequence of non-negative integers such that* $x_n\leq Cn^{1/2}$ *for some* $C>0$*. Then* $\mathscr{E}(x_n,n)\to 0$ *in probability.*

Proof. First we will show that $\sum_n \mathbf{P}^n_{x_n}(T_n \le n - n^{2/3}) < \infty$. By Borel-Cantelli, this would imply that all $\mathbf{P}^n_{x_n}$ may be coupled to the same probability space in such a way that one almost surely has $T_n > n - n^{2/3}$ for large enough n. Then the result follows immediately from Theorem 3.5.4 by taking $\alpha = 2/3$ in the definition of $\mathcal{E}(x,n)$. Note that the choice of exponent $2/3$ is arbitrary and could be replaced by any $\alpha > 1/2$.

To prove that $\sum_n \mathbf{P}^n_{x_n}(T_n \le n - n^{2/3}) < \infty$, one first notes that the event $\{T_n \le n - n^{2/3}\}$ can only happen if $\sup_{i \le n} S_i \ge n^{2/3}$. But by Theorem 3.6.8, we know that there are universal constants $C, c, c' > 0$ so that

$$\mathbf{P}^n_{x_n}\big(\sup_{i \le n} S_i \ge n^{2/3}\big) \le Ce^{-c(n^{2/3}-x_n)^2/n} \le Ce^{-c(n^{2/3}-Cn^{1/2})^2/n} \le Ce^{-c'n^{1/3}}.$$

The right side is summable as a function of n, completing the proof. □

Note that by equations (3.9) and (3.10) and the surrounding discussion (but replacing n above by $2n$), the above proposition reduces the proof of Theorem 3.2.2 to that of Theorem 3.1.2 but with varying weights, so this is what we focus on now.

3.5.2 Convergence for the quadrant model

In this section we finally complete the main goals of the paper. Unless otherwise stated, we always implicitly assume the following:

- All families $\{\omega^n_{i,j}\}$ of i.i.d. weights satisfy the assumptions that were stated in the bullet points before Theorem 3.2.2.

With the reduction (Proposition 3.5.8) finished, we define a partition function in the quadrant that is modified to take parity into account. Specifically, given (n,x) in the lattice $L := \{(n,x) \in \mathbb{Z}^2_{\ge 0} :$

$n - x \equiv 0 \pmod 2\}$ we define

$$Z_k(n,x) := \mathbf{E}_x^n\Big[z_0^k(S_n) \prod_{i=1}^n (1 + k^{-1/4}\omega_{n-i,S_i}^k)\Big]$$
$$= \sum_{r=0}^n k^{-r/4} \sum_{\substack{1\le i_1<\ldots<i_r\le n \\ (x_1,\ldots,x_{r+1})\in\mathbb{Z}_{\ge 0}^{r+1}}} \prod_{j=1}^r \mathfrak{p}_{i_j-i_{j-1}}^{n-i_{j-1}}(x_{j-1},x_j)\omega_{n-i_j,x_j}^k \cdot \big(z_0^k(x_{r+1})\mathfrak{p}_{n-i_r}^{n-i_r}(x_r,x_{r+1})\big), \tag{3.33}$$

with $i_0 := 0, x_0 := x$, and $z_0^k(x) = \prod_{i=0}^x (1+k^{-1/4}\omega_{i,0}^k)$ (in fact z_0^k can be any sequence of functions converging weakly and also satisfying the two bullet points of Proposition 3.5.7). Consider the following family of diffusively rescaled processes

$$\mathscr{Z}_n(T,X) := Z_n(nT, n^{1/2}X), \qquad T, X \ge 0, \tag{3.34}$$

where we interpolate linearly between points of the lattice L. We will now show that $\mathscr{Z}_n$ converge in law as $n \to \infty$ with respect to the topology of uniform convergence on compact subsets of $\mathbb{R}_+ \times \mathbb{R}_+$ to the solution of (3.18). The first step for doing this is proving tightness in the appropriate Hölder space. This part is not necessary if one is only interested in following the minimal logical flow for the proof of Theorem 3.1.1, and thus some of the proofs are not included. As always we denote $\|X\|_p := \mathbb{E}[|X|^p]^{1/p}$.

Proposition 3.5.9 (Tightness)**.** *Let $\mathscr{Z}_n$ be defined as in* (3.34)*, and assume that (for each k), the i.i.d. weights $\{\omega_{i,j}^k\}_{i,j}$ have $p > 8$ moments, bounded independently of k. Then for every $a \ge 0$, $\theta \in [0,1)$, and compact set $K \subset [0,\infty)^2$ there exists $C = C(a,p,\theta,K) > 0$ such that one has the following estimates uniformly over all pairs of space-time points $(T,X),(S,Y) \in K$:*

$$\|\mathscr{Z}_n(T,X)\|_p \le C, \tag{3.35}$$

$$\|\mathscr{Z}_n(T,X) - \mathscr{Z}_n(T,Y)\|_p \le C|X-Y|^{\theta/2}, \tag{3.36}$$

$$\|\mathscr{Z}_n(T,X) - \mathscr{Z}_n(S,X)\|_p \le C|T-S|^{\theta/4}. \tag{3.37}$$

In particular, the laws of the $\mathscr{Z}_n$ are tight with respect to the topology of uniform convergence on compact subsets of $C(\mathbb{R}_+ \times \mathbb{R}_+)$.

The restriction $p > 8$ is only necessary to obtain tightness in the Hölder space. Using more elegant arguments, this may be extended to $p \geq 6$ (see Appendix B of [3]). The one-point convergence result will only require two moments though.

Proof. Note that the functions Z_k defined in (3.33) satisfy the following Duhamel-form relation

$$Z_k(n,x) = \sum_{y \geq 0} \mathfrak{p}_n^n(x,y) z_0^k(y) + k^{-1/4} \sum_{i=0}^{n-1} \sum_{y \geq 0} \mathfrak{p}_{n-i}^n(x,y) Z_k(i,y) \omega_{i,y}^k. \tag{3.38}$$

Define the martingale $M_r(x,n,k) := k^{-1/4} \sum_{i=0}^{r-1} \sum_{y \geq 0} \mathfrak{p}_{n-i}^n(x,y) Z_k(i,y) \omega_{i,y}^k$. This is a martingale in the r-variable, with respect to the filtration $\mathcal{F}_r^k := \sigma(\{\omega_{i,j}^k\}_{1 \leq i \leq r; j \geq 0})$. This is because $Z_k(i,y)$ is $\mathcal{F}_r^k$-measurable, and $\mathcal{F}_r^k$ is independent of the mean-zero random variables $\omega_{r,y}^k$ with $y \geq 0$. Applying Burkholder-Davis-Gundy and then Minkowski's inequality to $M_r(x,n,k)$ shows that

$$\begin{aligned} \|M_r(x,n,k)\|_p^2 &\leq C \left\| k^{-1/2} \sum_{i=0}^{r-1} \left[\sum_{y \geq 0} \mathfrak{p}_{n-i}^n(x,y) Z_k(i,y) \omega_{i,y}^k \right]^2 \right\|_{p/2} \\ &\leq C k^{-1/2} \sum_{i=0}^{r-1} \left\| \sum_{y \geq 0} \mathfrak{p}_{n-i}^n(x,y) Z_k(i,y) \omega_{i,y}^k \right\|_p^2. \end{aligned} \tag{3.39}$$

Next, we notice that since the $\omega_{i,y}^k$ are independent of $Z_k(i,y)$, another application of Burkholder-Davis-Gundy (or in this case, its more elementary version for independent sums, the Marcinkiewicz-Zygmund inequality) shows that

$$\left\| \sum_{y \geq 0} \mathfrak{p}_{n-i}^n(x,y) Z_k(i,y) \omega_{i,y}^k \right\|_p^2 \leq C \sum_{y \geq 0} \mathfrak{p}_{n-i}^n(x,y)^2 \|Z_k(i,y)\|_p^2 \|\omega_{i,y}^k\|_p^2. \tag{3.40}$$

Since $p \leq p_0$ and the p_0^{th} moments of $\omega_{i,y}^k$ are bounded independently of k, i, y it follows that

$\|\omega_{i,y}^k\|_p^2$ may be absorbed into the constant. Combining (3.38),(3.39),(3.40), one finds that

$$\|Z_k(n,x)\|_p^2 \le C\bigg(\sum_{y\ge 0}\mathfrak{p}_n^n(x,y)\|z_0^k(y)\|_p\bigg)^2 + Ck^{-1/2}\sum_{i=0}^{n-1}\sum_{y\ge 0}\mathfrak{p}_i^n(x,y)^2\|Z_k(n-i,y)\|_p^2. \quad (3.41)$$

Now, we note that $\|z_0^k(y)\|_p \le e^{ak^{-1/2}y}$ by (3.31). Hence, $\sum_y \mathfrak{p}_n^n(x,y)\|z_0^k(y)\|_p$ may be bounded above by $Ce^{ak^{-1/2}x+Ka^2k^{-1}n}$ by Proposition 3.7.1. After this, we set $x_0 := x$ and $i_0 := 0$ and we iterate (3.41). Then we get

$$\begin{aligned}\|Z_k(n,x)\|_p^2 &\le C\sum_{r=0}^n k^{-r/2}\sum_{\substack{0\le i_1<\ldots<i_r<n\\(x_1,\ldots,x_r)\in\mathbb{Z}_{\ge 0}^r}}\prod_{j=1}^r \mathfrak{p}_{n-i_{j-1}}^{n-i_j}(x_{i-1},x_i)^2\cdot e^{ak^{-1/2}x_r+Ka^2n/k}\\ &\overset{\text{Lemma 3.5.1}}{\le} Ce^{ak^{-1/2}x+Ka^2n/k}\sum_{r=0}^n C^k k^{-r/2}n^{r/2}/(r/2)!\\ &\le Ce^{ak^{-1/2}x+Bn/k}, \end{aligned} \quad (3.42)$$

where we use $\sum_r C^k k^{-r/2}n^{r/2}/(r/2)! \le e^{C^2n/k}$ and then rename $B := Ka^2+C^2$. Now replace x by $n^{1/2}X$, n by nT, and k by n. This will give $\|\mathscr{Z}_n(T,X)\|_p^2 \le Ce^{aX+BT}$. But e^{aX+BT} can be bounded from above on any compact set, proving (3.35).

Now we will prove (3.36). By applying Burkholder-Davis-Gundy (twice) in the same way which was used in proving (3.41), one sees that

$$\begin{aligned}\|Z_k(n,x)-Z_k(n,y)\|_p^2 \le C\bigg\|\sum_{w\ge 0}\big(\mathfrak{p}_n^n(x,w)-\mathfrak{p}_n^n(y,w)\big)z_0^k(w)\bigg\|_p^2&\\ + Ck^{-1/2}\sum_{i=0}^{n-1}\sum_{w\ge 0}\big(\mathfrak{p}_{n-i}^n(x,w)-\mathfrak{p}_{n-i}^n(y,w)\big)^2\|Z_k(i,w)\|_p^2.&\end{aligned} \quad (3.43)$$

We will bound the first term using the coupling lemma. Specifically, let P (and its expectation operator E) denote a coupling of $\mathbf{E}_x^n$ and $\mathbf{E}_y^n$ as in Proposition 3.6.4, and let (S^x, S^y) be the associated coordinate process. Recall from Proposition 3.5.7 that $\mathbb{E}[(z_0^k(x)-z_0^k(y))^4] \le$

$Ck^{-1}|x-y|^2e^{ak^{-1/2}(x+y)}$ for some constants C,a independent of n,x,y. Then by independence of z_0^k and S, one may apply Minkowski and Jensen to commute the respective expectations and obtain

$$\begin{aligned}\Big\|\sum_{w\geq 0}\big(\mathfrak{p}_n^n(x,w)-\mathfrak{p}_n^n(y,w)\big)z_0^k(w)\Big\|_p^2 &= \big\|\mathbf{E}_x^n[z_0^k(S_n)]-\mathbf{E}_y^n[z_0^k(S_n)]\big\|_p^2 \\ &= \big\|E[z_0^k(S_n^x)-z_0^k(S_n^y)]\big\|_p^2 \leq E\big[\|z_0^k(S_n^x)-z_0^k(S_n^y)\|_p^2\big] \\ \leq CE\big[k^{-1/2}|S_n^x-S_n^y|e^{ak^{-1/2}(S_n^x+S_n^y)}\big] &\leq Ck^{-1/2}|x-y|\mathbf{E}_x^n[e^{2ak^{-1/2}S_n}]^{1/2}\mathbf{E}_y^n[e^{2ak^{-1/2}S_n}]^{1/2} \\ &\overset{\text{Prop. 3.7.1}}{\leq} Ck^{-1/2}|x-y|e^{ak^{-1/2}(x+y)},\end{aligned}$$

where we noted that $e^c+e^d\leq 2e^{c+d}$. Next, we geometrically interpolate (i.e., $c\wedge d\leq c^\theta d^{1-\theta}$ for $\theta\in[0,1]$) between the bound of Proposition 3.7.3 and that of (3.56) (with $p=2$ for both). This will yield the following for all $\alpha\geq 0$:

$$\sum_{z\geq 0}\big(\mathfrak{p}_n^N(x,z)-\mathfrak{p}_n^N(y,z)\big)^2e^{\alpha z}\leq Ce^{\alpha(x+y)+K\alpha^2 n}(n^{-\frac12-\frac12\theta}+\alpha^\theta n^{-\frac12})|x-y|^\theta \tag{3.44}$$

Using these bounds and using equation (3.43) in macroscopic coordinates, we will obtain:

$$\begin{aligned}&\|Z_k(n,x)-Z_k(n,y)\|_p^2 \\ &\leq Ck^{-1/2}|x-y|e^{ak^{-1/2}(x+y)}+Ck^{-1/2}\sum_{i=0}^{n-1}\sum_{w\geq 0}\big(\mathfrak{p}_{n-i}^n(x,w)-\mathfrak{p}_{n-i}^n(y,w)\big)^2\|Z_k(i,w)\|_p^2 \\ &\overset{(3.42)}{\leq} Ck^{-1/2}|x-y|e^{ak^{-1/2}(x+y)}+Ck^{-1/2}\sum_{i=1}^{n}\sum_{w\geq 0}\big(\mathfrak{p}_{n-i}^n(x,w)-\mathfrak{p}_{n-i}^n(y,w)\big)^2Ce^{ak^{-1/2}w+Bi/k} \\ &\overset{(3.44)}{\leq} Ck^{-1/2}|x-y|e^{ak^{-1/2}(x+y)} \\ &\qquad +Ck^{-1/2}\sum_{i=1}^n e^{ak^{-1/2}(x+y)}\big[(n-i)^{-\frac12-\frac12\theta}+a^\theta k^{-\theta/2}(n-i)^{-1/2}\big]|x-y|^\theta e^{Bn/k} \\ &\leq Ck^{-1/2}|x-y|e^{ak^{-1/2}(x+y)}+C(n^{\frac12-\frac12\theta}+k^{-\theta/2}n^{1/2})k^{-1/2}e^{ak^{-1/2}(x+y)}|x-y|^\theta e^{Bn/k}.\end{aligned} \tag{3.45}$$

In the last line, we used the bound $\sum_i(n-i)^{-\frac12-\frac12\theta}\leq Cn^{\frac12-\frac12\theta}$ for $\theta<1$. Now we convert to macroscopic coordinates ($k \to n; n \to nT; x \to n^{1/2}X; y \to n^{1/2}Y$), to get

$$\|\mathscr{Z}_n(T,X) - \mathscr{Z}_n(T,Y)\|_p^2 \le Ce^{2a(X+Y)}\big(|X-Y| + (T^{\frac{1}{2}-\frac{1}{2}\theta} + T^{1/2})|X-Y|^\theta\big)e^{BT}.$$

On any compact set $|X-Y|$ may be bounded by $C|X-Y|^\theta$ (since $\theta < 1$).Similarly, we can also absorb $(1 + T^{\frac{1}{2}-\frac{1}{2}\theta} + T^{1/2})e^{2a(X+Y)+BT}$ into the constant, proving (3.36).

Now we will prove (3.37). Let $m \le n$. For this, one writes

$$Z_k(n,x) = \sum_{y\ge 0} \mathfrak{p}_{n-m}^n(x,y)Z_k(m,y) + k^{-1/4}\sum_{i=1}^{n-m}\sum_{y\ge 0}\mathfrak{p}_{n-m-i}^n(x,y)Z_k(i+m,y)\omega_{(i+m)y}^k.$$

Again imitating the proof of (3.41) and using the fact that $\mathfrak{p}_{n-m}^n(x,\cdot)$ is a probability measure (then applying Jensen), one sees

$$\begin{aligned}\|Z_k(n,x) - Z_k(m,x)\|_p^2 \le C\sum_{y\ge 0}&\mathfrak{p}_{n-m}^n(x,y)\|Z_k(m,y) - Z_k(m,x)\|_p^2 \\ &+ Ck^{-1/2}\sum_{i=1}^{n-m}\sum_{y\ge 0}\mathfrak{p}_{n-m-i}^n(x,y)^2\|Z_k(i+m,y)\|_p^2.\end{aligned}$$

Let us call the sums on the right side $S_1(m,n,k,x), S_2(m,n,k,x)$, respectively. We bound these separately. We first compute that

$$\begin{aligned}&\sum_{y\ge 0}\mathfrak{p}_n^N(x,y)|x-y|^\theta e^{a(x+y)} = \mathbf{E}_x^N[|S_n - x|^\theta e^{a(S_n+x)}] \\ &\le \mathbf{E}_x^N[|S_n - x|^{2\theta}]^{1/2}\mathbf{E}_x^N[e^{2a(S_n+x)}]^{1/2} \le Cn^{\theta/2}e^{2ax+Ka^2n},\end{aligned}$$

where the last inequality follows from Propositions 3.6.10 and 3.7.1. Using this and (3.45) we see that

$$\begin{aligned}
S_1 &\le C \sum_{y\ge 0} \mathfrak{p}^n_{n-m}(x,y)\cdot \big[k^{-1/2}|x-y|e^{ak^{-1/2}(x+y)} \\
&\qquad\qquad + C(m^{\frac{1}{2}-\frac{1}{2}\theta} + k^{-\theta/2}m^{1/2})k^{-1/2}e^{ak^{-1/2}(x+y)}|x-y|^{\theta}e^{Bm/k}\big] \\
&\le Ck^{-1/2}(n-m)^{1/2}e^{2ak^{-1/2}x+Ka^2n/k} \\
&\qquad\qquad + C(m^{\frac{1}{2}-\frac{1}{2}\theta} + k^{-\theta/2}m^{1/2})k^{-1/2}e^{Bm/k}(n-m)^{\theta/2}e^{2ak^{-1/2}x+Ka^2n/k}.
\end{aligned}$$

Next, to bound S_2, we are going to use (3.42) with Proposition (3.7.1) and we obtain

$$\begin{aligned}
S_2 &\le Ck^{-1/4}\sum_{i=1}^{n-m}\sum_{y\ge 0}\mathfrak{p}^n_{n-m-i}(x,y)^2 e^{ak^{-1/2}y+B(i+m)/k} \\
&\le Ck^{-1/2}\sum_{i=1}^{n-m}(n-m-i)^{-1/2}e^{ak^{-1/2}x+Bn/k} \\
&\le Ck^{-1/2}(n-m)^{1/2}e^{ak^{-1/2}x+Bn/k}.
\end{aligned}$$

Combining the bounds for S_1, S_2 and then converting to macroscopic coordinates $(n\to nT; m\to nS; k\to n; x\to n^{-1/2}X)$ will yield the following bound:

$$\|\mathscr{Z}_n(T,X)-\mathscr{Z}_n(S,X)\|_p^2 \le C(|T-S|^{1/2} + (S^{\frac{1}{2}-\frac{1}{2}\theta}+S^{1/2})|T-S|^{\theta/2})e^{2aX+B'T},$$

where B' is a large constant depending on a^2 and B. Since $|T-S|^{1/2}\le C|T-S|^{\theta/2}$ on compact sets and since $(1+S^{\frac{1}{2}-\frac{1}{2}\theta}+S^{1/2})e^{2aX+B'T}$ may be bounded from above on compact sets, this finishes the proof of (3.37). □

Now that we have proved tightness, we only need to obtain convergence of finite-dimensional marginals of $\mathscr{Z}_n$ to those of SPDE (3.18). Thanks to the Cramer-Wold device (and linearity of integration with respect to space-time white noise) this will not be more difficult than just proving convergence of *one-point* marginals. This can be done by using the convergence result in Proposition 3.3.6 together with the machinery developed in the papers [3, 31].

Specifically, we will use Theorem 2.3 of [31], which in turn was inspired by the results of Section 4 in [3]. We state this result in a version that is adapted to our own context. Throughout, we will fix $T>0$ and we will denote $\Delta_k(T) := \{(t_1,...,t_k) : 0<t_1<...<t_k<T, t_i\in\mathbb{R}\}$. Also denote by $\Delta_k^n(T) := \{(\frac{t_1}{n},...,\frac{t_k}{n}) : 0<t_1<...<t_k<Tn, t_i\in\mathbb{Z}\}$, and let $(\mathbb{R}^d)_n := (n^{-1/2}\mathbb{Z})^d$. Then define

$$\mathcal{L}_k^n := \Delta_k^n(T)\times(\mathbb{R}^k)_n,$$

and equip $\mathcal{L}_k^n$ with the σ-finite measure that assigns mass $n^{-3/2} = n^{-1}\cdot n^{-1/2}$ to each distinct space-time point $(\frac{t}{n},\frac{x}{\sqrt{n}})$. We denote by $L^2(\mathcal{L}_k^n)$ the L^2-space associated to this measure.

Theorem 3.5.10 (Theorem 2.3 of [31]). *For each $n\in\mathbb{N}$, let $\{\omega_{i,j}^n\}_{i,j\geq 0}$ be a family of random weights with mean zero and $var(\omega_{i,j}^n) = \sigma^2 + o(1)$ (as $n\to\infty$). Let $\{F_k^n\}_{n,k\in\mathbb{N}}$ be a family of functions defined on $\mathcal{L}_k^n$. Suppose that $F_k : \Delta_k(T)\times\mathbb{R}^k\to\mathbb{R}$ is a family of continuous functions such that $\|F_k^n - F_k\|_{L^2(\mathcal{L}_k^n)}\to 0$ as $n\to\infty$, for every $k\in\mathbb{N}$. Furthermore assume that*

$$\sup_n \sum_{k\geq 0} \|F_k^n\|_{L^2(\mathcal{L}_k^n)}^2 < \infty.$$

Then define random variables

$$X_n := \sum_{k\geq 0} n^{-3k/4} \sum_{(\vec{t},\vec{x})\in\mathcal{L}_k^n} F_k^n(\vec{t},\vec{x})\omega_{(nt_1),(n^{1/2}x_1)}\cdots\omega_{(nt_k),(n^{1/2}x_k)}.$$

Then X_n converges in distribution as $n\to\infty$ to the random variable

$$\sum_{k=0}^{\infty}\sigma^k\int_{\Delta_k(T)}\int_{\mathbb{R}_+^k} F_k(t_1,...,t_k;x_1,...,x_k)\xi(dx_1dt_1)\cdots\xi(dx_kdt_k),$$

where ξ is a space-time white noise on $\mathbb{R}_+\times\mathbb{R}$.

We refer the reader to Section 4 of [3] for an explanation of the scaling exponent $n^{-3k/4}$. With this result in place, we are now ready to prove the main result of this section, which is a generalization of Theorem 3.1.2 to the case where the weights ω vary with n.

Theorem 3.5.11. *Let $\mathscr{Z}_n$ be as defined in* (3.34). *Then the finite-dimensional marginals of $\mathscr{Z}_n$ converge to those of SPDE* (3.18). *More precisely, if $F \subset \mathbb{R}_+ \times \mathbb{R}_+$ is finite, then $(Z_n(T,X))_{(T,X)\in F}$ converges in law to $(\mathscr{Z}(T,X))_{(T,X)\in F}$ where $\mathscr{Z}$ solves* (3.18) *with initial data given by $\mathscr{Z}(0,X) = e^{\sigma B_X + (\mu - \sigma^2/2)X}$ for a standard Brownian motion B.*

Proof. Using the discussion at the end of Section 2 (more specifically, equations (3.12) and (3.13)), we know that $z_0^n(n^{1/2}X)$ converges in law to a geometric Brownian motion with drift, specifically $e^{\sigma B_X + (\mu-\sigma^2/2)X}$. We exploit Skorohod's lemma to couple all of the z_0^n to the same probability space in such a way so that this convergence occurs a.s. uniformly on compact sets.

Fix $x, t > 0$. In our case, we set

$$F_k^n(t_1, ..., t_k; x_1, ..., x_k) := \sum_{x_{k+1} \in n^{-1/2}\mathbb{Z}_{\geq 0}} z_0^n(n^{1/2}x_{k+1}) \prod_{j=1}^{k+1} \mathscr{P}_n(t_j - t_{j-1}, T - t_{j-1}; x_{j-1}, x_j),$$

$$F_k(t_1, ..., t_k; x_1, ..., x_k) := \int_{\mathbb{R}_+} e^{B_{x_{k+1}} - (A+1/2)x_{k+1}} \prod_{j=1}^{k+1} \mathscr{P}_{t_j - t_{j-1}}^{T-t_{j-1}}(x_{j-1}, x_j)\, dx_{k+1},$$

where $\mathscr{P}_t^T$ was given in Definition 3.3.4, $\mathscr{P}_n$ was defined in Proposition 3.3.6 and where $(x_0, t_0) := (x, t)$. The condition that

$$\sup_n \sum_{k \geq 0} \|F_k^n\|^2_{L^2(\mathcal{L}_k^n)} < \infty,$$

follows quite simply from Lemma 3.5.1. Also the condition that $\|F_k^n - F_k\|_{L^2(\mathcal{L}_k^n)} \to 0$ as $n \to \infty$, follows by inducting on the last statement in Proposition 3.3.6.

By Theorem 3.5.10, we conclude that the one-point marginals of $\mathscr{Z}_n$ converge to those of the solution of (3.18). The proof for multi-point marginals is similar, but one defines a new family $\tilde{F}_k^n$ by taking linear combinations of the F_k^n that are defined above, then one applies the Cramer-Wold device to make the conclusion. $\square$

Note that (via Proposition 3.5.8) this result also implies Theorem 3.2.2, thus completing the main goals of the paper. One thing that we have not yet explained the normalization $\big(2\Phi\big(\frac{X+n^{-1/2}}{\sqrt{T}}\big)-1\big)^{-1}$ appearing in Theorem 3.2.2. This is an easy consequence of the fact that (by the local central limit theorem and (3.47)), the asymptotic mass of the measures $\mu^{nT}_{n^{1/2}X}$ appearing in Theorem 3.2.5 is equal to $2\Phi\big(\frac{X+n^{-1/2}}{\sqrt{T}}\big)-1+o(n^{-1/2})$.

3.6 Appendix 1: Preliminary estimates and concentration of measure

The purpose of this appendix is to gather estimates for the simple symmetric random walk conditioned to stay positive. The results and proofs are classical in spirit, and the literature on such measures is extensive [95, 19, 29, 30, 66] etc. However, we will only give a brief exposition of those selected estimates that apply to our nearest-neighbor weights, many of which we could not find in the above references, and might be applicable to other models.

We recall the uniform positive random walk measures $\mathbf{P}^n_x$ and the three associated quantities ($\mathfrak{p}^N_n, p^{(1/2)}_n$, and ψ) that were defined in Section 3. The main goal of this appendix will be to prove the following concentration inequality for the measures $\mathbf{P}^n_x$:

$$\mathbf{P}^n_x\Big(\sup_{1\le j\le k}|S_j-x|>u\Big)\le Ce^{-cu^2/k},$$

where C, c are independent of n, x, k with $k\le n$. This will in turn allow us to prove various L^p moment bounds that are used in Section 5. The methods used in proving these results will be coupling arguments and martingale techniques, some of which might be useful in and of themselves. More specifically, the main key will be to notice that for fixed $n\in\mathbb{N}$, the process

$$M^n_k := \frac{S_k+1}{\psi(S_k, n-k)}, \qquad 0\le k\le n,$$

is a $\mathbf{P}^n_x$-martingale with respect to the k-variable. Moreover we will use the fact that (S_k) is itself a submartingale. First we state a few preliminary lemmas.

Lemma 3.6.1. *Let $\psi(x,N)$ be as in Definition 3.3.2. Then there exists a constant $C>0$ such that for all $x,N\geq 0$ one has*

$$\frac{x+1}{x+1+C\sqrt{N}}\leq \psi(x,N)\leq 1\wedge\left(\frac{C(x+1)}{\sqrt{N}}\right).$$

Furthermore for each $x\geq 0$ one has that

$$\lim_{N\to\infty}\sqrt{N}\psi(x,N)=(x+1)\sqrt{2/\pi}.$$

Note that this already proves Theorem 3.2.5*(2)*. Furthermore note that the upper and lower bounds on ψ are strong enough to give an upper and lower envelope on ψ, i.e.,

$$C^{-1}\frac{x+1}{x+1+\sqrt{N}}\leq \psi(x,N)\leq C\frac{x+1}{x+1+\sqrt{N}}. \tag{3.46}$$

This is because $1\wedge w\leq \frac{2w}{1+w}$. We now proceed to the proof.

Proof. First we prove the upper bound. Let p_N denote the standard heat kernel on the whole line $\mathbb{Z}$. Using Definitions 3.3.1 and 3.3.2 and the fact that p_N is symmetric and sums to 1, it holds that

$$\psi(x,N)=\sum_{y\geq 0}\big(p_N(x-y)-p_N(x+y+2)\big)=p_N(x+1)+p_N(0)+2\sum_{1\leq u\leq x}p_N(u). \tag{3.47}$$

Now, we use the simple uniform bound $p_N\leq CN^{-1/2}$ to see that the right side of the last expression is bounded above by $2C(x+1)N^{-1/2}$. On the other hand, it is obvious that $\psi(x,N)\leq 1$ for all x,N. So, we have obtained the desired upper bound.

Next, we prove the lower bound. We consider two different cases: $x\leq 2\sqrt{N}$ and $x>2\sqrt{N}$.

First we consider the case $x>2\sqrt{N}$. One may apply Hoeffding's inequality for the simple random walk to deduce that

$$\psi(x,N) = p_N(x+1) + p_N(0) + 2\sum_{1\le u\le x} p_N(u) \ge \sum_{-x\le u\le x} p_N(u) \ge 1 - 2e^{-(x+1)^2/2N}.$$

Now set $q := \frac{(x+1)^2}{2N}$. Then $q \ge 2$, so $q+2 \le e^q$, and thus $\frac{1}{1-2e^{-q}} \le 1 + \frac{2}{q}$. This means that $\psi(x,N)^{-1} \le 1 + \frac{N}{2(x+1)^2}$. But since $x+1 \ge \sqrt{N}$, it follows that $\frac{N}{(x+1)^2} \le \frac{\sqrt{N}}{x+1}$. Hence we obtain $\psi(x,N) \ge \frac{x+1}{x+1+0.5\sqrt{N}}$, whenever $x > 2\sqrt{N}$.

Now we consider the case $x \le 2\sqrt{N}$. The local central limit theorem tells us that $p_N(u) \ge \frac{c}{\sqrt{N}} e^{-2u^2/N} \ge \frac{c}{\sqrt{N}} e^{-8}$, for some $c > 0$ and all u, N with $u \le 2\sqrt{N}$. Hence

$$\psi(x,N) = p_N(x+1) + p_N(0) + 2\sum_{1\le u\le x} p_N(u) \ge \sum_{0\le u\le x} p_N(u) \ge \frac{ce^{-8}}{\sqrt{N}}(x+1).$$

Now one simply notes that $\frac{ce^{-8}}{\sqrt{N}} \ge \frac{1}{x+1+c^{-1}e^8\sqrt{N}}$. This proves the lower bound.

Finally, we prove the last statement about the limit. For this, let us write

$$\psi(x,N) = p_N(x+1) + p_N(0) + 2\sum_{1\le u\le x} p_N(u)$$

The local limit theorem tells us that for each u, the quantity $\sqrt{N}p_N(u)$ oscillates back and forth between $\sqrt{2/\pi}$ and zero (depending on the parity of N) as N becomes large. This already implies that $N^{1/2}$ times the right side converges to $\big(1+x\big)\sqrt{2/\pi}$. $\square$

Lemma 3.6.2. *Let $(a_x)_{x\ge 0}$ be a sequence of non-negative numbers such that*

$$a_x \le a_{x+1}, \qquad a_{x+2} - a_{x+1} \le a_{x+1} - a_x, \qquad \textit{for all } x \ge 0.$$

Then for all $x \geq 0$ and $k \geq 0$, one has that

$$\frac{a_x}{a_{x+k}} \leq \frac{a_{x+1}}{a_{x+k+1}}.$$

Proof. It suffices to prove the claim when $k = 1$, because then one has that

$$\frac{a_x}{a_{x+k}} = \prod_{j=0}^{k-1} \frac{a_{x+j}}{a_{x+j+1}} \leq \prod_{j=0}^{k-1} \frac{a_{x+j+1}}{a_{x+j+2}} = \frac{a_{x+1}}{a_{x+k+1}}.$$

To prove the claim for $k = 1$, one uses the mean value theorem to extract $u \in [a_x, a_{x+1}]$ and $v \in [a_{x+1}, a_{x+2}]$ such that

$$\log a_{x+1} - \log a_x = \frac{1}{u}(a_{x+1} - a_x), \qquad \log a_{x+2} - \log a_{x+1} = \frac{1}{v}(a_{x+2} - a_{x+1}).$$

Then clearly $\frac{1}{v} \leq \frac{1}{u}$, and by hypothesis, it is also true that $a_{x+2} - a_{x+1} \leq a_{x+1} - a_x$. So we conclude that $\log a_{x+2} - \log a_{x+1} \leq \log a_{x+1} - \log a_x$. □

Lemma 3.6.3 (Monotonicity)**.** *Fix $n \in \mathbb{N}$. Then $\psi(x,n)$ is an increasing function of x. Thus, $\mathfrak{p}_1^n(x, x+1) \geq 1/2 \geq \mathfrak{p}_1^n(x, x-1)$ for all $x, n \geq 0$. Furthermore, $\mathfrak{p}_1^n(x, x+1)$ is a decreasing function of x, and $\mathfrak{p}_1^n(x, x-1)$ is an increasing function of x.*

Proof. As in the proof of Lemma 3.6.1, we write

$$\psi(x,n) = p_n(0) + p_n(x+1) + 2\sum_{y=1}^{x} p_n(y).$$

Consequently, it holds that

$$\psi(x+1,n) - \psi(x,n) = p_n(x+1) + p_n(x+2), \tag{3.48}$$

and the right side is clearly non-negative, which proves the first statement. For the second statement, we just note that

$$\mathfrak{p}_1^n(x, x+1) = \frac{\psi(x+1, n-1)}{2\psi(x,n)} \geq \frac{\psi(x-1, n-1)}{2\psi(x,n)} = \mathfrak{p}_1^n(x, x-1).$$

To prove the final statement, we note that p_n is a non-increasing function of $|x|$, and thus the right side of (3.48) is also a non-increasing function of x. Thus we may apply Lemma 3.6.2 with $k = 2$ and $a_x = \psi(x, n)$, to conclude that $\frac{\psi(x,n)}{\psi(x+2,n)}$ is an increasing function of x. Now we write

$$\frac{1}{\mathfrak{p}_1^n(x, x+1)} = 2\frac{\psi(x,n)}{\psi(x+1, n-1)} = 1 + \frac{\psi(x-1, n-1)}{\psi(x+1, n-1)},$$

where we use the relation $\psi(x, n) = \frac{1}{2}\psi(x+1, n-1) + \frac{1}{2}\psi(x-1, n-1)$. By the discussion of the previous paragraph, the right side is an increasing function of x, and so $\mathfrak{p}_1^n(x, x+1)$ is a decreasing function of x. Finally, this implies that $\mathfrak{p}_1^n(x, x-1) = 1 - \mathfrak{p}_1^n(x, x+1)$ is an increasing function of x. □

Proposition 3.6.4 (Coupling lemma for Positive Walks)**.** *Fix $n \in \mathbb{N}$ and $x \geq 0$. There exists a coupling $\mathbf{Q}^n_{x,x+1}$ of the measures $\mathbf{P}^n_x$ and $\mathbf{P}^n_{x+1}$ that is supported on pairs (γ, γ') of paths such that $|\gamma_i' - \gamma_i| \leq 1$ for all $i \leq n$. More generally, for fixed $n \in \mathbb{N}$, the measures $\{\mathbf{P}^n_x\}_{x \geq 0}$ may all be coupled together in such a way that the coordinate processes associated to neighboring values of x are never more than distance 1 apart.*

Proof. Let $\{U_i\}_{i=1}^n$ be a sequence of i.i.d. Uniform$[0,1]$ random variables. We make an inductive construction as follows. Let $S_0 = x$ and $S_0' = x+1$.

Suppose that $S_0, ..., S_k$ and $S_0', ..., S_k'$ have been constructed in such a way that $|S_i - S_i'| = 1$

for all k. If $S'_k = S_k + 1$, we define

$$(S_{k+1}, S'_{k+1}) := \begin{cases} (S_k - 1, S'_k - 1), & U_{k+1} > \mathfrak{p}_1^{n-k}(S_k, S_k + 1) \\ (S_k + 1, S'_k - 1), & \mathfrak{p}_1^{n-k}(S_k, S_k + 1) > U_{k+1} > \mathfrak{p}_1^{n-k}(S'_k, S'_k + 1) \\ (S_k + 1, S'_k + 1), & U_{k+1} < \mathfrak{p}_1^{n-k}(S'_k, S'_k + 1) \end{cases} .$$

We know by lemma 3.6.3 that one of these cases must hold. Similarly, if $S'_k = S_k - 1$, then we define (S_{k+1}, S'_{k+1}) in a symmetric fashion. This completes the inductive step.

A close look at this construction reveals that for $x_1, ..., x_n \geq 0$ one has

$$P(S_1 = x_1, S_2 = x_2, ..., S_n = x_n) = \mathfrak{p}_1^n(x, x_1) \prod_{j=1}^{n-1} \mathfrak{p}_1^{n-j}(x_j, x_{j+1}),$$
$$P(S'_1 = x_1, S'_2 = x_2, ..., S'_n = x_n) = \mathfrak{p}_1^n(x+1, x_1) \prod_{j=1}^{n-1} \mathfrak{p}_1^{n-j}(x_j, x_{j+1}).$$

By Proposition 3.3.3, S is distributed as $\mathbf{P}_x^n$ and S' is distributed as $\mathbf{P}_{x+1}^n$.

The proof of the more general statement is very similar. One simply uses a uniform coupling together with the Lemma 3.6.3, and the argument is a straightforward generalization of the one given above for two values of x. $\square$

Proposition 3.6.5 (Martingales for Positive Walks)**.** *Fix $x, n, k \geq 0$ with $k \leq n$. Let S be distributed according to $\mathbf{P}_x^n$. For $i \leq k$ define a function $f^{(k,n)}(x, i) := \mathbf{E}_x^{n-i}[S_{k-i}]$. Then the process*

$$M_i = M_i^{(k,n)} := f^{(k,n)}(S_i, i), \qquad 0 \leq i \leq k$$

is a martingale with respect to the natural filtration of S. Furthermore it has bounded increments

$$|M_{i+1} - M_i| \leq 2, \qquad 0 \leq i \leq k-1.$$

In the special case when $k = n$, one has the explicit form $f^{(n,n)}(x,i) = -1 + \frac{x+1}{\psi(x,n-i)}$

Proof. We suppress the superscript (k,n) on f from now on. Letting $\mathcal{F}_k$ denote the natural filtration of S, it is a consequence of the Markov property that $f(S_i,i) = \mathbf{E}_x^n[S_k|\mathcal{F}_i]$, which shows that M is a martingale in the i-variable for fixed x,n,k.

To prove that it has bounded increments, first note that

$$f(x,i) - f(x+1,i) = \mathbf{E}_x^{n-i}[S_{k-i}] - \mathbf{E}_{x+1}^{n-i}[S_{k-i}].$$

By the coupling lemma (Proposition 3.6.4), this is bounded in absolute value by 1. Consequently, one finds that

$$\begin{aligned}|f(x\pm 1,k+1) - f(x,k)| &= \Big|f(x\pm 1,k+1) - \sum_{y\in\{x-1,x+1\}} \mathfrak{p}_1^{n-k}(x,y) f(y,k+1)\Big| \\ &\le \sum_{y\in\{x-1,x+1\}} \mathfrak{p}_1^{n-k}(x,y)\big|f(x\pm 1,k+1) - f(y,k+1)\big| = \mathfrak{p}_1^{n-k}(x,x\mp 1)\cdot 2 \le 2,\end{aligned}$$

which gives the desired result.

For the final statement, if $k = n$ one may compute $\mathbf{E}_x^{n-i}[S_{n-i}] = \sum_{y\ge 0} y\mathfrak{p}_{n-i}^{n-i}(x,y) = \psi(x,n-i)^{-1}\sum_{y\ge 0} y p_{n-i}^{(1/2)}(x,y)$. Now the claim follows from the fact that $y \mapsto y+1$ is a unipotent eigenfunction of the semigroup $p^{(1/2)}$, i.e., $\sum_{y\ge 0}(y+1)p_n^{(1/2)}(x,y) = x+1$ for every $n, x \ge 0$. □

Lemma 3.6.6. *Let $b \ge 0$. There exists a constant $C = C(b) > 0$ such that for all $n \ge 0$ and all $x,y,z \ge 0$ one has*

$$p_n^{(1/2)}(x,y) \le C\left[\frac{1}{\sqrt{n+1}} \wedge \frac{x+1}{n+1}\right] e^{-bn^{-1/2}|x-y|}.$$

$$\big|p_n^{(1/2)}(x,y) - p_n^{(1/2)}(x,z)\big| \le C\left[\frac{1}{n+1} \wedge \frac{x+1}{(n+1)^{3/2}}\right]|z-y|e^{-bn^{-1/2}\big(|x-y|\wedge|x-z|\big)}.$$

Proof. The proof given here is inspired by the methods of Appendix A of [57].

Let us start with the first bound. Note that it suffices to prove the bound for $y \geq x$. Indeed, if $x \geq y$, then by symmetry one has that $p_n^{(1/2)}(x,y) = p_n^{(1/2)}(y,x) \leq C\frac{y+1}{n+1}e^{-b|x-y|n^{-1/2}} \leq C\frac{x+1}{n+1}e^{-b|x-y|n^{-1/2}}$.

Let $p_n(x)$ denote the standard discrete heat kernel on $\mathbb{Z}$. One first notes that for $z \in \mathbb{C}$ one has that $\sum_{x\in\mathbb{Z}} p_n(x)z^x = 2^{-n}(z+z^{-1})^n$, i.e., $\mathbf{E}[z^{U_n}] = \mathbf{E}[z^{U_1}]^n$ for a simple random walk U on $\mathbb{Z}$. Letting C denote the unit circle of $\mathbb{C}$ oriented counterclockwise, Cauchy's integral formula says

$$p_n(x) = \frac{1}{2\pi i}\oint_C z^{-x-1}2^{-n}(z+z^{-1})^n dz.$$

Since the integrand is analytic away from the origin, one may deform the contour without changing the value. Specifically, we will expand the radius of the circle to $e^{bn^{-1/2}}$. Parametrizing this as $z = e^{bn^{-1/2}}e^{it}$, one finds that

$$p_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-xbn^{-1/2}}e^{-itx}\Big[\cosh(b/\sqrt{n})\cos t + i\sinh(b/\sqrt{n})\sin t\Big]^n dt. \tag{3.49}$$

Now, by Taylor expanding $\sinh$ and $\cosh$, the key observation is that

$$\begin{aligned} f_n(t) :&= \Big|\cosh(b/\sqrt{n})\cos t + i\sinh(b/\sqrt{n})\sin t\Big|^n \\ &= \Big[\cosh^2(b/\sqrt{n})\cos^2 t + \sinh^2(b/\sqrt{n})\sin^2 t\Big]^{n/2} \\ &= \Big[\big(1+b^2/(2n)+O(n^{-2})\big)\cos^2 t + \big(b^2/n + O(n^{-2})\big)\sin^2 t\Big]^{n/2} \\ &\leq \Big[\cos^2 t + \frac{b^2}{n} + O(n^{-2})\Big]^{n/2}, \end{aligned}$$

where the $O(n^{-2})$ terms denote quantities which are uniformly bounded above by $e^b b^4 n^{-2}$. Since n^{-2} decays faster than n^{-1}, this means that the last expression is bounded above by $\big[\cos^2 t +$

$Cn^{-1}\big]^{n/2}$, where $C = C(b)$. Now, there exist constants $c, D > 0$ such that for $t \in [-\pi/2, \pi/2]$ and $n \geq 1$ we have

$$\Big[\cos^2 t + Cn^{-1}\Big]^{n/2} \leq De^{-nct^2}. \tag{3.50}$$

Indeed, on $[\pi/2, \pi/2]$ it holds that $\cos^2 t \leq e^{-2ct^2}$ for some (small enough) $c > 0$ so that the left side of (3.50) is bounded above by $e^{-nct^2}(1 + \frac{Ce^{2ct^2}}{n})^{n/2}$. Since $1 + u \leq e^u$, note that $(1 + \frac{Ce^{2ct^2}}{n})^{n/2} \leq e^{\frac{C}{2}e^{2ct^2}} \leq e^{\frac{C}{2}e^{c\pi^2/2}} =: D$ on $[-\pi/2, \pi/2]$, giving (3.50).

Because the left side of (3.50) is an upper bound for $f_n(t)$, it easily follows from (3.50) that $\int_{-\pi/2}^{\pi/2} f_n(t)dt \leq Cn^{-1/2}$ and $\int_{-\pi/2}^{\pi/2} |t| f_n(t) \leq Cn^{-1}$ (where additional terms have been absorbed into C). With this in mind, we compute via (3.49) that

$$p_n(y-x) - p_n(y+x+2) \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \big|e^{-(y-x)(bn^{-1/2}+it)} - e^{-(y+x+2)(bn^{-1/2}+it)}\big| f_n(t)dt. \tag{3.51}$$

Thanks to the absolute value and the trigonometric nature of the integrand, we may replace the integral over $[-\pi, \pi]$ with twice the integral over $[-\pi/2, \pi/2]$. Furthermore, using $|e^{is} - 1| \leq |s|$ one computes that

$$\begin{aligned}
\big|e^{-(y-x)(bn^{-1/2}+it)} - e^{-(y+x+2)(bn^{-1/2}+it)}\big| &= \big|e^{-(y-x)bn^{-1/2}} e^{-(y-x)it}\big(1 - e^{-2(x+1)(bn^{-1/2}+it)}\big)\big| \\
&= e^{-(y-x)bn^{-1/2}}\big|1 - e^{-2(x+1)(bn^{-1/2}+it)}\big| \\
&\leq e^{-(y-x)bn^{-1/2}}\big|1 - e^{-2(x+1)bn^{-1/2}}\big| + e^{-(y+x+2)bn^{-1/2}}\big|1 - e^{-2(x+1)it}\big| \\
&\leq 2be^{-(y-x)bn^{-1/2}}\frac{x+1}{n^{1/2}} + 2e^{-(y+x+2)bn^{-1/2}}(x+1)|t|.
\end{aligned} \tag{3.52}$$

Now we note that $e^{-(y+x+2)bn^{-1/2}} \leq e^{-(y-x)bn^{-1/2}}$ since $x, y \geq 0$. Combining (3.51) and (3.52), together with the fact that $\int_{-\pi/2}^{\pi/2} f_n(t)dt \leq Cn^{-1/2}$ and $\int_{-\pi/2}^{\pi/2} |t| f_n(t)dt \leq Cn^{-1}$, proves that for $y \geq x$ one has $p_n^{(1/2)}(x, y) \leq C\frac{x+1}{n+1}e^{-bn^{-1/2}(y-x)}$, as desired. In order to obtain the other bound

$p_n^{(1/2)}(x,y) \le \frac{1}{\sqrt{n+1}} e^{-bn^{-1/2}|x-y|}$, one replaces (3.52) with the easier bound

$$\left|e^{-(y-x)(bn^{-1/2}+it)} - e^{-(y+x+2)(bn^{-1/2}+it)}\right| \le 2e^{-bn^{1/2}|x-y|},$$

where we used the triangle inequality and the fact that $|e^{is}| \le 1$. Then one uses (3.51) and the fact that $\int f_n \le Cn^{-1/2}$. This completes the proof of the first bound.

Now we prove the second bound stated in the proposition. For this, one uses the same arguments, but one needs to replace (3.52) with the appropriate bound. Specifically, we need to consider

$$e^{-(y-x)(bn^{-1/2}+it)} - e^{-(y+x+2)(bn^{-1/2}+it)} - \left(e^{-(z-x)(bn^{-1/2}+it)} - e^{-(z+x+2)(bn^{-1/2}+it)}\right).$$

We write this as

$$g(y-x) - g(y+x+2) - g(z-x) + g(z+x+2) = \int_{y-x}^{y+x+2} \int_{v}^{v+z-y} g''(u) du dv,$$

where $g(u) = e^{-u(bn^{-1/2}+it)}$. So one computes $g''(u) = (bn^{-1/2}+it)^2 g(u)$. Now, for u in the relevant range, it is clear that $|g(u)| = e^{-bn^{-1/2}u} \le e^{-bn^{-1/2}(|z-x|\wedge|y-x|)}$. Hence

$$|g''(u)| = |bn^{-1/2}+it|^2 |g(u)| \le \left(2b^2n^{-1} + 2t^2\right) e^{-bn^{-1/2}(|z-x|\wedge|y-x|)},$$

where we used $|p+q|^2 \le 2|p|^2 + 2|q|^2$. Combining the previous two expressions, we find that

$$\left|g(y-x) - g(y+x+2) - g(z-x) + g(z+x+2)\right|$$

$$\le 2(x+1)|z-y|\left(2b^2n^{-1} + 2t^2\right) e^{-bn^{-1/2}(|z-x|\wedge|y-x|)}.$$

Now multiplying by $f_n(t)$ and integrating over $[-\pi,\pi]$, we finally obtain

$$|p_n^{(1/2)}(x,y) - p_n^{(1/2)}(x,z)| \le C(b)(x+1)|z-y|e^{-bn^{-1/2}(|z-x|\wedge|y-x|)}\int_{-\pi}^{\pi}(n^{-1}+t^2)f_n(t)dt$$

$$\le C(x+1)|y-z|(n+1)^{-3/2},$$

where we use the bound (3.50) for $f_n(t)$ in the last inequality. This already proves one part of the second bound, namely $|p_n^{(1/2)}(x,y) - p_n^{(1/2)}(x,z)| \le \frac{C|y-z|(x+1)}{(n+1)^{3/2}}e^{-bn^{-1/2}(|z-x|\wedge|y-x|)}$. For the other bound $p_n^{(1/2)}(x,y) \le \frac{C|y-z|}{n+1}e^{-bn^{-1/2}(|z-x|\wedge|y-x|)}$, we simply note that

$$|g(y-x) - g(y+x+2) - g(z-x) + g(z+x+2)|$$

$$\le |g(y-x) - g(z-x)| + |g(y+x+2) - g(z+x+2)|$$
$$\le \int_{y-x}^{z-x}|g'(u)|du + \int_{y+x+2}^{z+x+2}|g'(u)|du,$$

and then we apply similar arguments as before, noting $|g'(u)| \le |bn^{-1/2}+it||g(u)|$. $\square$

Lemma 3.6.7. *There exists a constant $C>0$ such that for all $x\ge 0$ and all $n\ge k\ge 1$ one has that*

$$\mathbf{E}_x^n[S_k] \le x + Ck^{1/2}.$$

Proof. We consider two cases, $k>n/2$ and $k\le n/2$.

Case 1. $k>n/2$. First, we claim that $\mathbf{E}_x^n[S_k] \le \mathbf{E}_x^n[S_n]$. In fact, it is even true that S forms a $\mathbf{P}_x^n$-submartingale and thus $\mathbf{E}_x^n[S_k]$ is an increasing function of k for every n. This follows immediately from Lemma 3.6.3 after noticing that $\mathbf{E}_x^n[S_{k+1}|\mathcal{F}_k] = S_k + (2\mathfrak{p}_1^{n-k}(S_k,S_k+1)-1) \ge S_k$. Now, from the preceding proposition, we know that $M_k := \frac{S_k+1}{\psi(S_k,n-k)}$ forms a martingale. Thus, we see that

$$\mathbf{E}_0^n[S_n+1] = \mathbf{E}_x^n[M_0] = \frac{x+1}{\psi(x,n)} \le x+1+Cn^{1/2},$$

where we applied the lower bound of Lemma 3.6.1 in the final bound. Since $k > n/2$, we see that $n^{1/2} \le 2^{1/2}k^{1/2}$, which gives the desired bound in this case.

Case 2. $k \le n/2$. First we use the coupling lemma (Proposition 3.6.4) to see that $\mathbf{E}_x^n[S_k] \le 1 + \mathbf{E}_{x-1}^n[S_k]$. Iterating this x times shows that

$$\mathbf{E}_x^n[S_k] \le x + \mathbf{E}_0^n[S_k].$$

Thus we only need to show that $\mathbf{E}_0^n[S_k] \le Ck^{1/2}$. To prove this, write $\mathbf{E}_0^n[S_k] = \sum_{y\ge 0} \mathfrak{p}_k^n(0,y)y$. Now we write $\mathfrak{p}_k^n(0,y) = p_k^{(1/2)}(0,y)\frac{\psi(y,n-k)}{\psi(0,n)}$. By Lemma 3.6.1 we know $\frac{1}{\psi(0,n)} \le C\sqrt{n}$. Furthermore we also know from the same lemma that $\psi(y,n-k)$ is bounded above by $1\wedge(Cy(n-k)^{-1/2})$, which is in turn bounded above by $1 \wedge (Cyn^{-1/2})$ since $k \le n/2$. Moreover, we also know from Lemma 3.6.6 that $p_k^{(1/2)}(0,y) \le \frac{C}{k+1}e^{-y/\sqrt{k}}$. Thus, we find that

$$\mathbf{E}_0^n[S_k] \le \frac{C}{k+1}\bigg[\sum_{0\le y\le\sqrt{n}} e^{-y/\sqrt{k}}n^{1/2}(n^{-1/2}y^2) + \sum_{y\ge\sqrt{n}} e^{-y/\sqrt{k}}(n^{1/2}y)\bigg]. \tag{3.53}$$

Let us refer to the two sums inside the square brackets on the right side as J_1 and J_2, respectively.

First we bound J_1. Now, we use the bound $\sum_{r\ge 0} r^2\alpha^r \le \frac{2}{(1-\alpha)^3}$ (valid for $\alpha < 1$) and we see that

$$J_1 \le C\sum_{y\ge 0} y^2 e^{-y/\sqrt{k}} \le \frac{C}{(1-e^{-1/\sqrt{k}})^3} \le Ck^{3/2}.$$

In the last bound, we used the elementary bound $(1-e^{-q})^{-1} \le 1 + q^{-1}$ (which in turn implies $(1-e^{-q})^{-3} \le 2^3(1+q^{-3})$) with $q = k^{-1/2}$.

Next, we bound J_2. Using the bound $\sum_{r\ge s} r\alpha^r \le C\big[\frac{\alpha^s}{(1-\alpha)^2} + \frac{s\alpha^s}{1-\alpha}\big]$, we see that

$$J_2 = n^{1/2}\sum_{y\ge\sqrt{n}} e^{-y/\sqrt{k}}y \le n^{1/2}\bigg[\frac{e^{-\sqrt{n/k}}}{(1-e^{-1/\sqrt{k}})^2} + \frac{n^{1/2}e^{-\sqrt{n/k}}}{1-e^{-1/\sqrt{k}}}\bigg].$$

Now

$$n^{1/2}e^{-\sqrt{n/k}} = k^{1/2}(n/k)^{1/2}e^{-\sqrt{n/k}} \le k^{1/2} \sup_{u>0} ue^{-u} = Ck^{1/2}.$$

Similarly, one finds that $ne^{-\sqrt{n/k}} \le Ck$. We also note that $(1-e^{-q})^{-1} \le 1+q^{-1}$, and thus $(1-e^{-q})^{-2} \le 2+2q^{-2}$. Taking $q = k^{-1/2}$ and then combining the last few expressions, one finally gets $J_2 \le Ck^{3/2}$.

Combining the bounds of J_1 and J_2 with (3.53), we obtain the desired bound. □

Finally we have our concentration theorem, the main result of this appendix.

Theorem 3.6.8 (Concentration)**.** *As before, let $S = (S_k)_{0\le k\le n}$ denote the canonical process associated to $\mathbf{P}^n_x$. Then there exist $C, c > 0$ such that for every $x \ge 0$, every $0 \le k \le n$, and every $u > 0$ one has that*

$$\mathbf{P}^n_x\big(\sup_{0\le i\le k} |S_i - x| > u\big) \le Ce^{-cu^2/k}.$$

In other words, on time scales of length k, the path measure $\mathbf{P}^n_x$ concentrates on spatial scales of order $\sqrt{k}$ around x. The idea of the proof is to exploit the martingales from Proposition 3.6.5 and apply well-known concentration inequalities for bounded-increment martingales. The Gaussian decay constant c will be obtained as $1/32$, which is not sharp (presumably $c = 1/2$ should be possible, but we do not have a proof).

Proof. Throughout this proof, x, n, and k will be **fixed**. Let us write

$$\mathbf{P}^n_x\big(\sup_{0\le i\le k} |S_i - x| > u\big) = \mathbf{P}^n_x\big(\sup_{0\le i\le k} S_i > x+u\big) + \mathbf{P}^n_x\big(\inf_{0\le i\le k} S_i < x-u\big).$$

Let us refer to the terms on the right side as p_1, p_2 respectively.

First we bound p_2. Recall from Lemma 3.6.3 that $\mathfrak{p}^n_1(x, x+1) \ge 1/2 \ge \mathfrak{p}^n_1(x, x-1)$ for all $n, x \ge 0$. This trivially shows that S is a submartingale, which directly gives the claim for p_2 by Azuma's inequality [9] for submartingales, with $c = 1/2$.

Now we will bound p_1, which is more difficult. Letting $M = (M_i^{(n,k)})_{i=0}^k$ denote the martingale from Proposition 3.6.5, it is clear that $S_k = M_k$. Furthermore $M_0 = f(x,0) = \mathbf{E}_x^n[S_k] \leq Ck^{1/2} + x$ by Proposition 3.6.7. Since the increments of M are bounded above by 2, we may apply Azuma's inequality again to see that

$$\begin{aligned}\mathbf{P}_x^n(S_k > x+u) = \mathbf{P}_x^n\big(M_k > x+u\big) &\leq \mathbf{P}_x^n\big(M_k - M_0 > u - Ck^{1/2}\big) \\ &\leq e^{-(u-Ck^{1/2})^2/8k} \leq Ce^{-u^2/16k}.\end{aligned}$$

In the last inequality, we used the fact that $(u-Ck^{1/2})^2 \geq \frac{1}{2}u^2 - C^2k$. This, in turn, is because $(a+b)^2 \leq 2(a^2+b^2)$. Combining the bounds on p_1 and p_2 shows that

$$\mathbf{P}_x^n(|S_k - x| > u) \leq Ce^{-u^2/16k}. \tag{3.54}$$

Since (S_i) is a submartingale (Lemma 3.6.3) and since $x \mapsto e^{\lambda x}$ is increasing and convex it follows that the process $(e^{\lambda S_i})_{i=0}^n$ is a $\mathbf{P}_x^n$-submartingale as well. Thus, we may apply Doob's martingale inequality to see that

$$p_1 \leq Ce^{-\lambda(x+u)}\mathbf{E}_x^n[e^{\lambda S_k}] = Ce^{-\lambda(x+u)}\bigg(1 + \int_0^\infty \lambda e^{\lambda y}\mathbf{P}_x^n(S_k > y)du\bigg).$$

Now we split the integral as $\int_0^x$ plus $\int_x^\infty$. We use the crude bound $\mathbf{P}_x^n(S_k > y) \leq 1$ for the integral over $[0,x]$, and we use the bound (3.54) for the other. This gives

$$p_1 \leq Ce^{-\lambda u} + Ce^{-\lambda u}\int_x^\infty \lambda e^{\lambda(y-x)-(y-x)^2/16k}dy \leq C(e^{-\lambda u} + \lambda k^{1/2}e^{4\lambda^2 k - \lambda u}).$$

Setting $\lambda = \frac{u}{8k}$ gives a bound of $C(e^{-u^2/8k} + uk^{-1/2}e^{-u^2/16k})$. Now one simply notes that $r \leq Ce^{r^2/32}$, so that $uk^{-1/2} \leq Ce^{u^2/32k}$. This gives the desired bound on p_1, where the constant appearing in the theorem statement is $c := 1/32$. $\square$

We now give a slightly generalized version of the concentration theorem.

Corollary 3.6.9. *In the same setting as the previous theorem, there exist $C, c > 0$ such that for every $x \geq 0$, every $0 \leq m \leq k \leq n$, and every $u > 0$ one has that*

$$\mathbf{P}_x^n\big(\sup_{m \leq i \leq k} |S_i - S_m| > u\big) \leq Ce^{-cu^2/(k-m)}.$$

Here, C, c are the same as in the previous theorem.

Proof. Define

$$g(k, n, x, u) := \mathbf{P}_x^n\big(\sup_{0 \leq i \leq k} |S_i - x| > u\big).$$

By the Markov property (conditioning on the first m steps), we have that

$$\mathbf{P}_x^n\big(\sup_{m \leq i \leq k} |S_i - S_m| > u\big) = \mathbf{E}_x^n\Big[g(k-m, n-m, S_m, u)\Big].$$

But Theorem 3.6.8 tells us that $g(k, n, x, u) \leq Ce^{-cu^2/k}$ independently of x, n. $\square$

Corollary 3.6.10. *Let $p > 0$. There exists a constant $C = C_p > 0$ such that for every $x \geq 0$ and every $0 \leq k \leq m \leq n$, one has*

$$\mathbf{E}_x^n\big[|S_k - S_m|^p\big] \leq C|k-m|^{p/2}.$$

Proof. Let us write

$$\mathbf{E}_x^n\big[|S_k - S_m|^p\big] = \int_0^\infty pu^{p-1}\mathbf{P}_x^n(|S_k - S_m| > u)du.$$

By Corollary 3.6.9, this is bounded above by

$$C\int_0^\infty pu^{p-1}e^{-cu^2/(k-m)}du = Cp(k-m)^{p/2}\int_0^\infty v^{p-1}e^{-cv^2}dv = C_p(k-m)^{p/2},$$

where we made a substitution $y = (k-m)^{-1/2}u$ in the first equality. $\square$

By Arzela-Ascoli, the preceding corollary clearly implies tightness of the diffusively rescaled process mentioned in Remark 3.2.6. Indeed we can use this to easily recover classical results such as [95, 27] in this nearest-neighbor case, for instance by showing that any subsequential limit has the same finite-dimensional marginal distributions as $\mathbf{W}_X^T$ which in turn can be shown e.g. by Proposition B.6 below.

3.7 Appendix 2: Heat kernel estimates for conditioned walks

We now prove various estimates for the heat kernels $\mathfrak{p}_n^N$ defined in Section 3. Not much motivation will be given here, but the content of Sections 4 and 5 has illustrated the applicability of these estimates. The methods used in proving these bounds will be elementary bounds together with the results of Appendix A (specifically Propositions 3.6.6 and 3.6.4, and Theorem 3.6.8).

Proposition 3.7.1. *There exist constants $C, K > 0$ such that for all $x \geq 0$, all $N \geq n \geq 0$, and all $a \geq 0$ one has that*

$$\sum_{y \geq 0} \mathfrak{p}_n^N(x,y)e^{ay} \leq Ce^{ax+Ka^2n}.$$

Proof. In the notation of Appendix A, let us write

$$\sum_{y\geq 0} \mathfrak{p}_n^N(x,y)e^{ay} = \mathbf{E}_x^N[e^{aS_n}] = 1 + \int_0^\infty ae^{au}\mathbf{P}_x^N(S_n > u)du.$$

Now we split the integral as $\int_0^x + \int_x^\infty$. For the integral over $[0,x]$ we use the crude bound $\mathbf{P}_x^N(S_n > u) \leq 1$. For the integral over $[x,\infty)$, we use the result of Theorem 3.6.8. This will give

$$\mathbf{E}_x^N[e^{aS_n}] \leq e^{ax} + C\int_x^\infty ae^{au}e^{-c(u-x)^2/n}du \leq e^{ax} + C\cdot an^{1/2}e^{ax+\frac{a^2n}{4c}}.$$

Since $an^{1/2} \leq e^{a^2n}$, this gives the result with $K := 1 + \frac{1}{4c}$. $\square$

We remark that $c = 1/32$ from the proof of Theorem 3.6.8, so we can obtain $K = 9$ in the preceding proposition. Conjecturally, the optimal value of K should be $1/2$, as is the case for the simple random walk (as seen from $\cosh(a) \leq e^{\frac{1}{2}a^2}$).

Lemma 3.7.2. *Fix $b > 0$. There exists $C = C(b) > 0$ such that for all $x \geq 0$ and all $N \geq n \geq 0$ one has that*

$$\mathfrak{p}_n^N(x,y) \leq \frac{C}{\sqrt{n+1}} e^{-b|x-y|/\sqrt{n}}.$$

We remark that this bound is fairly strong, and many of our estimates could have been derived from this result rather than from the concentration theorem (but only in a weaker form because the decay is merely exponential rather than Gaussian).

Proof. We consider four different cases.

Case 1. $x \geq \sqrt{N}$. Then, one has $\frac{\psi(y,N-n)}{\psi(x,N)} \leq \frac{1}{\psi(x,N)} \leq C$ by Lemma 3.6.1. Thus it holds that $\mathfrak{p}_n^N(x,y) \leq C p_n^{(1/2)}(x-y) \leq C(n+1)^{-1/2} e^{-b|x-y|/\sqrt{n}}$. The final inequality comes from the first bound of Lemma 3.6.6.

Case 2. $n < N/2$ and $y \leq x$. Then one has

$$\begin{aligned}
\mathfrak{p}_n^N(x,y) &\leq C p_n^{(1/2)}(x,y) \left[\frac{x+1+\sqrt{N}}{x+1}\right]\left[\frac{y+1}{y+1+\sqrt{N-n}}\right] \\
&\leq C(n+1)^{-1/2} e^{-b|x-y|/\sqrt{n}} \left[\frac{x+1+\sqrt{N}}{x+1}\right]\left[\frac{x+1}{x+1+\sqrt{N-n}}\right] \\
&\leq C(n+1)^{-1/2} e^{-b|x-y|/\sqrt{n}} \left[\frac{N}{N-n}\right]^{1/2}.
\end{aligned}$$

We used (3.46) in the first line and we used Lemma 3.6.6 and that $y \mapsto \frac{y+1}{y+1+\sqrt{N-n}}$ is monotone increasing in the second line. Then we canceled the $x+1$ and used the fact that $x \mapsto \frac{x+1+\sqrt{N}}{x+1+\sqrt{N-n}}$ is monotone decreasing in the last line. Since $n < N/2$ it follows that $\left[\frac{N}{N-n}\right]^{1/2} \leq 2^{1/2}$ so that term may be absorbed into C.

Case 3. $n < N/2$ and $y \geq x$. Then

$$\begin{aligned}
\mathfrak{p}_n^N(x,y) &\leq C p_n^{(1/2)}(x,y) \left[\frac{x+1+\sqrt{N}}{x+1}\right]\left[\frac{y+1}{y+1+\sqrt{N-n}}\right] \\
&\leq C p_n^{(1/2)}(x,y) \left[\frac{x+1+\sqrt{N}}{x+1}\right]\left[\frac{y+1}{x+1+\sqrt{N-n}}\right] \\
&\leq C\left[\frac{N}{N-n}\right]^{1/2} p_n^{(1/2)}(x,y)\frac{y+1}{x+1} \\
&= C\left[\frac{N}{N-n}\right]^{1/2} p_n^{(1/2)}(x,y)\left[\frac{y-x}{x+1} \ + \ 1\right] \\
&\leq C\left[\frac{y-x}{n+1} + C(n+1)^{-1/2}\right]e^{-b|x-y|/\sqrt{n}}. \qquad (3.55)
\end{aligned}$$

Here we noted $y \geq x$ in the second line, and we used the fact that $x \mapsto \frac{x+1+\sqrt{N}}{x+1+\sqrt{N-n}}$ is monotone decreasing in the third line. In the final line, we used $\left[\frac{N}{N-n}\right]^{1/2} \leq 2^{1/2}$ (since $n < N/2$) and also the first bound of Lemma 3.6.6. Now, we know that the bound (3.55) is true for all b, in particular it is true with b replaced by $b+1$, after perhaps making the constant bigger. Thus we see that

$$\frac{|x-y|}{n+1}e^{-(b+1)|x-y|/\sqrt{n}} \leq \frac{1}{\sqrt{n+1}}e^{-b|x-y|/\sqrt{n}}\left[\frac{|x-y|}{\sqrt{n}}e^{-|x-y|/\sqrt{n}}\right].$$

$$\leq \frac{1}{\sqrt{n+1}}e^{-b|x-y|/\sqrt{n}} \sup_{u>0} ue^{-u} = \frac{C}{\sqrt{n+1}}e^{-b|x-y|/\sqrt{n}}.$$

Case 4. $x \leq \sqrt{N}$ and $n \geq N/2$. Since $x \leq \sqrt{N} \leq \sqrt{2n}$, we can apply Lemmas 3.6.1 and 3.6.6 to see that

$$\mathfrak{p}_n^N(x,y) \leq C p_n^{(1/2)}(x,y)\frac{x+1+\sqrt{N}}{x+1}$$

$$\leq C\frac{x+1}{n+1}e^{-b|x-y|/\sqrt{n}} \cdot \frac{2\sqrt{2n}+1}{x+1} \leq C(n+1)^{-1/2}e^{-b|x-y|/\sqrt{n}}.$$

This completes the proof of all cases. $\square$

Proposition 3.7.3. *There exist constants $C, K > 0$ such that for all $x \geq 0$, all $N \geq n \geq 0$, all*

$a \geq 0$, *and all* $p \geq 1$ *one has that*

$$\sum_{y\geq 0} \mathfrak{p}_n^N(x,y)^p e^{ay} \leq C^p (n+1)^{-(p-1)/2} e^{ax+Ka^2 n}.$$

Proof. Using Lemma 3.7.2 with $b = 0$, one finds that

$$\mathfrak{p}_n^N(x,y)^p = \mathfrak{p}_n^N(x,y)^{p-1}\mathfrak{p}_n^N(x,y) \leq \frac{C^{p-1}}{(n+1)^{(p-1)/2}}\mathfrak{p}_n^N(x,y).$$

Then the claim follows immediately from Proposition 3.7.1. $\square$

We now bound space-time differences of the heat kernels $\mathfrak{p}_n^N$.

Lemma 3.7.4. *There exists a constant* $C > 0$ *such that for all* $x, y, z \geq 0$ *one has that*

$$\left|\mathfrak{p}_n^N(x,y) - \mathfrak{p}_n^N(x,z)\right| \leq \frac{C}{n+1}\left[\frac{N+1}{N-n+1}\right]^{1/2} |y-z|.$$

Proof. Without loss of generality, assume $y \geq z$. It suffices to prove the bound in the case $y = z+1$. In the general case, one simply adds the bound $y - z$ times. Let us write

$$\left|\mathfrak{p}_n^N(x,z+1) - \mathfrak{p}_n^N(x,z)\right| = \left|\frac{p_n^{(1/2)}(x,z+1)\psi(z+1,N-n) - p_n^{(1/2)}(x,z)\psi(z,N-n)}{\psi(x,N)}\right|$$

$$\leq |p_n^{(1/2)}(x,z+1) - p_n^{(1/2)}(x,z)|\frac{\psi(z+1,N-n)}{\psi(x,N)} + p_n^{(1/2)}(x,z)\frac{|\psi(z+1,N-n)-\psi(z,N-n)|}{\psi(x,N)}.$$

Let us call the two terms of the last expression I_1, I_2 respectively. From here, one considers two cases ($x \leq \sqrt{N}$ and $x \geq \sqrt{N}$) and bound I_1, I_2 separately each time. The arguments are similar to the ones above, so the proof is not included. $\square$

Proposition 3.7.5. *Fix* $p \geq 1$. *There exists a constant* $C = C(p) > 0$ *such that for all* $x, y \geq 0$, *all*

$N \geq n \geq m \geq 0$*, and all* $a \geq 0$ *one has that*

$$\sum_{z\geq 0} \left|\mathfrak{p}_n^N(x,z) - \mathfrak{p}_n^N(y,z)\right|^{2p} e^{az} \leq Ce^{a(x+y)+Ka^2 n}\left(n^{\frac{1}{2}-\frac{3}{2}p} + a^p n^{\frac{1}{2}-p}\right)|x-y|^p, \tag{3.56}$$

$$\sum_{z\geq 0} \left|\mathfrak{p}_m^{N-n+m}(x,z) - \mathfrak{p}_n^N(x,z)\right|^{2p} e^{az} \leq Ce^{2ax+Ka^2 n}\left(m^{\frac{1}{2}-\frac{3}{2}p} + a^p m^{\frac{1}{2}-p}\right)|n-m|^{p/2}. \tag{3.57}$$

In the spatial bound (3.56)*, the constant* C *grows at worst exponentially in* p*.*

We remark that in the special case that $p = 1$ and $a \leq Cn^{-1/2}$, one has that $n^{\frac{1}{2}-\frac{3}{2}p} + a^p n^{\frac{1}{2}-p} \leq Cn^{-1}$ and similarly for m. This is the case in which this bound will most often be applied.

Proof. We first start out by proving an auxiliary bound:

$$\sum_{z\geq 0} \left(\mathfrak{p}_n^N(x,z) - \mathfrak{p}_n^N(y,z)\right)^2 e^{az} \leq Ce^{a(x+y)+Ka^2 n}\left(n^{-1} + an^{-1/2}\right)\left[\frac{N+1}{N-n+1}\right]^{1/2}|x-y|. \tag{3.58}$$

Let us prove this. The coupling lemma (3.6.4) and the preceding lemma will be key here. First, by the coupling lemma, we know that $\mathbf{P}_x^N$ and $\mathbf{P}_y^N$ may be coupled in such a way so that the respective coordinate processes (call them $(S_n^x)_{n=0}^N$ and $(S_n^y)_{n=0}^N$) are never a distance more than $|y-x|$ apart (i.e., $\sup_{n\leq N}|S_n^x - S_n^y| \leq |x-y|$ a.s.). Let E denote the expectation with respect to the coupled measure. Now, by writing $\left(\mathfrak{p}_n^N(x,z) - \mathfrak{p}_n^N(y,z)\right)^2 = \mathfrak{p}_n^N(x,z)\left(\mathfrak{p}_n^N(x,z) - \mathfrak{p}_n^N(y,z)\right) - \mathfrak{p}_n^N(y,z)\left(\mathfrak{p}_n^N(x,z) - \mathfrak{p}_n^N(y,z)\right)$ we may write

$$\begin{aligned}
&\sum_{z\geq 0}\left(\mathfrak{p}_n^N(x,z) - \mathfrak{p}_n^N(y,z)\right)^2 e^{az} \\
&\qquad = \mathbf{E}_x^N[(\mathfrak{p}_n^N(x,S_n) - \mathfrak{p}_n^N(y,S_n))e^{aS_n}] - \mathbf{E}_y^N[(\mathfrak{p}_n^N(x,S_n) - \mathfrak{p}_n^N(y,S_n))e^{aS_n}] \\
&\qquad = E[(\mathfrak{p}_n^N(x,S_n^x) - \mathfrak{p}_n^N(y,S_n^x))e^{aS_n^x}] - E[(\mathfrak{p}_n^N(x,S_n^y) - \mathfrak{p}_n^N(y,S_n^y))e^{aS_n^y}] \\
&\qquad = E[(\mathfrak{p}_n^N(x,S_n^x) - \mathfrak{p}_n^N(x,S_n^y))e^{aS_n^x}] + E[\mathfrak{p}_n^N(x,S_n^y)(e^{aS_n^x} - e^{aS_n^y})] \\
&\qquad\qquad + E[(\mathfrak{p}_n^N(y,S_n^y) - \mathfrak{p}_n^N(y,S_n^x))e^{aS_n^y}] + E[\mathfrak{p}_n^N(y,S_n^x)(e^{aS_n^y} - e^{aS_n^x})].
\end{aligned}$$

Let us refer to the terms in the last expression as J_1, J_2, J_3, J_4, respectively. Since J_1 and J_3 occupy symmetric roles, it suffices to bound J_1 and then the analogous bound for J_3 automatically follows. The same thing happens for J_2 and J_4. With this understanding, we will only prove the desired bound for J_1 and J_2.

Let us start by bounding J_1. By Lemma 3.7.4, we see that

$$\begin{aligned}|\mathfrak{p}_n^N(x,S_n^x)-\mathfrak{p}_n^N(x,S_n^y)| &\leq \frac{C}{n+1}\bigg[\frac{N+1}{N-n+1}\bigg]^{1/2}|S_n^x-S_n^y| \\ &\leq \frac{C}{n+1}\bigg[\frac{N+1}{N-n+1}\bigg]^{1/2}|x-y|.\end{aligned}$$

Here we applied the coupling in the second inequality. Applying the definition of J_1 and then Proposition 3.7.1, we therefore obtain that

$$J_1 \leq \frac{C}{n+1}\bigg[\frac{N+1}{N-n+1}\bigg]^{1/2}|x-y|E[e^{aS_n^x}] \leq \frac{C}{n+1}\bigg[\frac{N+1}{N-n+1}\bigg]^{1/2}|x-y|e^{ax+Ka^2n}.$$

This already gives the desired bound on J_1. As discussed, the analogous bound on J_3 is obtained in an identical fashion, but one will get e^{ay} instead of e^{ax}. The final bound on J_1+J_3 is then obtained by noting that $e^{ax}+e^{ay} \leq 2e^{a(x+y)}$.

Now we bound J_2. First note that $|e^u-e^v| \leq |u-v|e^{u\vee v}$ for all $u,v\in\mathbb{R}$. Thus $|e^{aS_n^y}-e^{aS_n^x}| \leq a|S_n^y-S_n^x|e^{a(S_n^y\vee S_n^x)} \leq a|y-x|e^{a(S_n^y+S_n^x)}$. By Cauchy-Schwarz, we in turn bound $E[e^{a(S_n^y+S_n^x)}] \leq \mathbf{E}_x^N[e^{2aS_n}]^{1/2}\mathbf{E}_y^N[e^{2aS_n}]^{1/2} \leq Ce^{a(x+y)+Ka^2n}$, by Proposition 3.7.3. Now, we also know from Lemma 3.7.2 that $\mathfrak{p}_n^N(x,S_n^y) \leq C(n+1)^{-1/2}$. Using these facts, we find that

$$J_2 \leq Ca|y-x|E[\mathfrak{p}_n^N(x,S_n^y)e^{a(S_n^y+S_n^x)}] \leq Can^{-1/2}|x-y|e^{a(y+x)+Ka^2n}.$$

Already this proves the required bound on J_2. The analogous bound on J_4 follows immediately. This completes the proof of (3.58).

Now let us prove the spatial estimate (3.56). For $m \le n$, we use the semigroup property to write $\mathfrak{p}_n^N(x,z) = \sum_{y\ge 0} \mathfrak{p}_m^N(x,y)\mathfrak{p}_{n-m}^{N-m}(y,z)$ and then using Jensen's inequality, we find that

$$\begin{aligned}
\left|\mathfrak{p}_n^N(x,z) - \mathfrak{p}_n^N(y,z)\right|^{2p} &= \left|\sum_{w\ge 0}\big(\mathfrak{p}_m^N(x,w) - \mathfrak{p}_m^N(y,w)\big)\mathfrak{p}_{n-m}^{N-m}(w,z)\right|^{2p} \\
&\le \left(\sum_{w\ge 0}\big(\mathfrak{p}_m^N(x,w) - \mathfrak{p}_m^N(y,w)\big)^2\mathfrak{p}_{n-m}^{N-m}(w,z)\right)^p
\end{aligned}$$

Denoting by I the left-hand side of (3.56), we then find by Minkowski's inequality that

$$\begin{aligned}
I^{1/p} &\le \left(\sum_{z\ge 0}\left[\sum_{w\ge 0}\big(\mathfrak{p}_m^N(x,w) - \mathfrak{p}_m^N(y,w)\big)^2\mathfrak{p}_{n-m}^{N-m}(w,z)e^{az/p}\right]^p\right)^{1/p} \\
&\overset{\text{Minkowski}}{\le} \sum_{w\ge 0}\left[\sum_{z\ge 0}\big(\mathfrak{p}_m^N(x,w) - \mathfrak{p}_m^N(y,w)\big)^{2p}\mathfrak{p}_{n-m}^{N-m}(w,z)^p e^{az}\right]^{1/p} \\
&= \sum_{w\ge 0}\big(\mathfrak{p}_m^N(x,w) - \mathfrak{p}_m^N(y,w)\big)^2\left[\sum_{z\ge 0}\mathfrak{p}_{n-m}^{N-m}(w,z)^p e^{az}\right]^{1/p} \\
&\overset{Prop.3.7.3}{\le} C\sum_{w\ge 0}\big(\mathfrak{p}_m^N(x,w) - \mathfrak{p}_m^N(y,w)\big)^2(n-m)^{\frac{1-p}{2p}}e^{(aw+Ka^2(n-m))/p} \\
&\overset{(3.58)}{\le} C(n-m)^{\frac{1-p}{2p}}\big(m^{-1} + am^{-1/2}\big)\left[\frac{N+1}{N-m+1}\right]^{1/2}|x-y|e^{\left(a(x+y)+Ka^2n\right)/p}.
\end{aligned}$$

Setting $m := n/2$ then gives (3.56), because $\left[\frac{N+1}{N-\frac{1}{2}n+1}\right]^{1/2} \le \left[\frac{N+1}{\frac{1}{2}N+1}\right]^{1/2} \le 2^{1/2}$. Note that the constant C does not depend on p, which also proves the final sentence given in the theorem statement after noting that $\big(n^{-1} + an^{-1/2}\big)^p \le 2^p(n^{-p} + a^p n^{-p/2})$.

We move on to the temporal estimate (3.57). The main idea is to use Jensen's inequality together with the spatial estimate. Specifically, we start off by writing

$$\begin{aligned}\big|\mathfrak{p}_m^{N-n+m}(x,z)-\mathfrak{p}_n^N(x,z)\big|^{2p} &= \Big|\mathfrak{p}_n^{N-n+m}(x,z)-\sum_{y\geq 0}\mathfrak{p}_{n-m}^N(x,y)\mathfrak{p}_m^{N-n+m}(y,z)\Big|^{2p}\\ &= \Big|\sum_{y\geq 0}\mathfrak{p}_{n-m}^N(x,y)\big(\mathfrak{p}_m^{N-n+m}(x,z)-\mathfrak{p}_m^{N-n+m}(y,z)\big)\Big|^{2p}\\ &\overset{\text{Jensen}}{\leq} \sum_{y\geq 0}\mathfrak{p}_{n-m}^N(x,y)\big|\mathfrak{p}_m^{N-n+m}(x,z)-\mathfrak{p}_m^{N-n+M}(y,z)\big|^{2p}.\end{aligned}$$

Next, we multiply by e^{az}, then sum over z, and interchange the sum over z with the sum over y. Letting J denote the left-hand side of (3.57), this gives

$$\begin{aligned}J &\leq \sum_{y\geq 0}\mathfrak{p}_{n-m}^N(x,y)\sum_{z\geq 0}\big|\mathfrak{p}_m^{N-n+m}(x,z)-\mathfrak{p}_m^{N-n+m}(y,z)\big|^{2p}e^{az}\\ &\leq C^p\sum_{y\geq 0}\mathfrak{p}_{n-m}^N(x,y)e^{a(x+y)+Ka^2m}\big(m^{\frac12-\frac32p}+a^pm^{\frac12-p}\big)|y-x|^p\\ &= C^pe^{ax+Ka^2m}\big(m^{\frac12-\frac32p}+a^pm^{\frac12-p}\big)\mathbf{E}_x^N[|S_{n-m}-x|^pe^{aS_{n-m}}].\end{aligned}$$

All that is left to do is to show that one has $\mathbf{E}_x^N[|S_{n-m}-x|^pe^{aS_{n-m}}]\leq Ce^{ax+Ka^2(n-m)}|n-m|^{p/2}$. This is an easy consequence of the concentration theorem. Indeed, for any $k\leq N$ one may write

$$\mathbf{E}_x^N[|S_k-x|^pe^{aS_k}]\leq \mathbf{E}_x^N[|S_k-x|^{2p}]^{1/2}\mathbf{E}_x^N[e^{2aS_k}]^{1/2},$$

and then the claim follows immediately from Propositions 3.7.1 and Corollary 3.6.10. □

Corollary 3.7.6 (Spatial/Temporal Estimates)**.** *There exists $C>0$ such that for all $x,y,z\geq 0$ and all $N\geq n\geq m\geq 0$ one has that*

$$\big|\mathfrak{p}_n^N(x,z)-\mathfrak{p}_n^N(y,z)\big|\leq C(n+1)^{-3/4}|x-y|^{1/2}, \tag{3.59}$$

$$\big|\mathfrak{p}_m^{N-n+m}(x,z)-\mathfrak{p}_n^N(x,z)\big|\leq C(m+1)^{-3/4}|n-m|^{1/4}. \tag{3.60}$$

These pointwise bounds are quite useful, in the sense that that the exponents (despite not being sharp) are ones which actually give meaningful information. However, we will not actually need this estimate, but it could potentially be useful if one wanted to develop the results of Section 4 with Dirac initial data (for instance).

Proof. Let $I(2p)$ denote the left-hand side of (3.56) with $a = 0$. Then $I(2p)^{\frac{1}{2p}}$ is bounded above by $C(n+1)^{\frac{1}{4p}-\frac{3}{4}}|x-y|^{1/2}$, where the constant C is independent of p (by the final sentence in the statement of Proposition 3.7.5). Letting $p \to \infty$ already proves the first bound (since ℓ^p norms converge to the ℓ^∞ norm).

For the second bound, we cannot do the same thing, since the constant in (3.57) could (in principle) have worse-than-exponential dependence on p. However, we can use the semigroup property to write

$$\big|\mathfrak{p}_m^{N-n+m}(x,z) - \mathfrak{p}_n^N(x,z)\big| \le \sum_{y\ge 0} \mathfrak{p}_{n-m}^N(x,y)\big|\mathfrak{p}_m^{N-n+m}(x,z) - \mathfrak{p}_m^{N-n+m}(y,z)\big|,$$

and then one may use the spatial bound (3.59) with Corollary 3.6.10 to obtain the result. □

Next we prove a strong convergence result for the discrete kernels $\mathfrak{p}_n^N$ to the continuous ones $\mathscr{P}_t^T$ from Definition 3.3.4, from which we can easily obtain estimates for the continuous kernels as well. In the case of Brownian meander at terminal time ($X = 0$ and $t = T$), the following result is weaker than the local convergence result of [29], but we actually need it for all (t, T) so we give an original proof.

Proposition 3.7.7. *Fix $\tau \ge 0$. Then for $n \ge 0$, define*

$$\mathscr{P}_n(t,T;X,Y) := (n/2)^{1/2}\mathfrak{p}_{2\lfloor tn\rfloor}^{2\lfloor Tn\rfloor}(2\lfloor n^{1/2}X/\sqrt{2}\rfloor, 2\lfloor n^{1/2}Y/\sqrt{2}\rfloor).$$

Then for each fixed $X, T, t \ge 0$, the map $Y \mapsto \mathscr{P}_n(t,T;X,Y)$ converges pointwise and in $L^p(\mathbb{R}_+, e^{aY}dY)$ to $\mathscr{P}_t^T(X,Y)$ for all $p \ge 1$ and $a \ge 0$ (as $n \to \infty$).

Furthermore for all $X, T \geq 0$, the map $(t, Y) \mapsto \mathscr{P}_n(t, T; X, Y)$ converges pointwise and in $L^p(dt \otimes e^{aY} dY)$ to $\mathscr{P}_t^T(X, Y)$ for all $p \in [1, 3)$ and $a \geq 0$ (as $n \to \infty$).

From now on, we will abbreviate quantities such as $\mathfrak{p}_{2\lfloor tn \rfloor}^{2\lfloor Tn \rfloor}(2\lfloor n^{1/2}X/\sqrt{2}\rfloor, 2\lfloor n^{1/2}Y/\sqrt{2}\rfloor)$ by just writing $\mathfrak{p}_{2nt}^{2nT}((2n)^{1/2}X, (2n)^{1/2}Y)$ instead. We hope that this abuse of notation will not cause any confusion, but in reality one should keep in mind that all quantities are only defined with *even* integers. The reason for this is the periodicity of the simple random walk: $\mathfrak{p}_n^N(x, y)$ vanishes if n and $x - y$ have different parity. If it were not for this parity consideration, we could take a limit of the simpler quantity $n^{1/2}\mathfrak{p}_{\lfloor nt \rfloor}^{\lfloor nT \rfloor}(\lfloor n^{1/2}X\rfloor, \lfloor n^{1/2}Y\rfloor)$.

Proof. First, let us prove pointwise convergence. Letting p_n denote the standard heat kernel on all $\mathbb{Z}$, we recall that

$$p_n^{(1/2)}(x, y) = p_n(x - y) - p_n(x + y + 2).$$

$$\psi(x, n) = p_n(0) + p_n(x + 1) + 2\sum_{1 \leq y \leq x} p_n(y) = \sum_{-x \leq y \leq x+1} p_n(y).$$

Let F_n denote the cdf associated to p_n, so that $\psi(x, n) = F_n(x + 1) - F_n(-x) = F_n(x) + F_n(x + 1) - 1$. By uniformity of convergence of cdf's in the central limit theorem we know that $F_n(n^{1/2}x)$ converges uniformly (on $\mathbb{R}$) to $\Phi(x)$, where Φ is the cdf of a standard normal. From this it is clear that $\psi(n, n^{1/2}x) = F_n(n^{1/2}x) + F_n(n^{1/2}x + 1) - 1$ converges uniformly to $2\Phi(x) - 1$ (because Φ has no atoms). In turn, one deduces that $\psi(2nT, (2n)^{1/2}X) = \psi(2nT, (2nT)^{1/2}X/\sqrt{T})$ converges to $2\Phi(X/\sqrt{T}) - 1$. From here, completing the proof of pointwise convergence is easy using the local central limit theorem (though notice that $X = 0$ requires a separate proof) as done in earlier proofs.

Now we will fix t, T, X, and we will address convergence in $L^p(\mathbb{R}_+, e^{aY} dY)$. The main idea is simply to use dominated convergence in conjunction with Lemma 3.7.2. Specifically, that lemma (applied with $b = 2at^{1/2}/p$) tells us that

$$\mathscr{P}_n(t, T; X, Y) \leq Ct^{-1/2}e^{-2a|X-Y|/p}. \tag{3.61}$$

Here C is a constant independent of Y (but it will depend on t, a, p). Letting $p \geq 1$, it is then clear from (3.61) that for *fixed* X, T, t, the sequence of maps

$$Y \mapsto \mathscr{P}_n(t, T; X, Y)^p e^{aY}$$

is dominated (uniformly in n) by a function that is integrable on $\mathbb{R}_+$. This is enough to guarantee by dominated convergence that

$$\int_{\mathbb{R}_+} |\mathscr{P}_n(t, T; X, Y) - \mathscr{P}_t^T(X, Y)|^p e^{aY} dY \to 0.$$

Similarly, one uses (3.61) in conjunction with the dominated convergence theorem to obtain convergence in $L^p(\mathbb{R}_+ \times \mathbb{R}_+, dt \otimes e^{aY} dY)$ of $(Y, t) \mapsto \mathscr{P}_n(t, T; X, Y)$. This argument only works for $p \in [1, 3)$, since the singularity of $\int_{\mathbb{R}_+} t^{-p/2} e^{-pt^{-1/2}|X-Y|} dY \sim t^{-(p-1)/2}$ fails to be absolutely integrable near $t = 0$, if $p \geq 3$. $\square$

Proposition 3.7.8. *Let $a, \tau > 0$ and let $\mathscr{P}_t^T$ be the kernels from Definition 3.3.4. Then there exists a constant $C = C(\tau, a)$ such that for all $X, Y \geq 0$, all $\theta \in [0, 1/2]$, and all $s \leq t \leq T \leq \tau$ one has the following*

$$\int_{\mathbb{R}_+} \mathscr{P}_t^T(X, Z) e^{aZ} dZ \leq C e^{aX}, \tag{3.62}$$

$$\int_{\mathbb{R}_+} \mathscr{P}_t^T(X, Z)^2 e^{aZ} dZ \leq C t^{-1/2} e^{aX}, \tag{3.63}$$

$$\int_{\mathbb{R}_+} \big(\mathscr{P}_t^T(X, Z) - \mathscr{P}_t^T(Y, Z)\big)^2 e^{aZ} dZ \leq C t^{-\frac{1}{2}-\theta} e^{a(X+Y)} |X - Y|^{2\theta}, \tag{3.64}$$

$$\int_{\mathbb{R}_+} \big(\mathscr{P}_s^{T-t+s}(X, Z) - \mathscr{P}_t^T(X, Z)\big)^2 e^{aZ} dZ \leq C s^{-\frac{1}{2}-\theta} e^{2aX} |t - s|^{\theta} \tag{3.65}$$

Proof. The claims follow from the L^1 and L^2 convergence in Proposition 3.7.7. More specifically, (3.62) follows from Proposition 3.7.1 and convergence in $L^1(\mathbb{R}_+, e^{aY} dY)$. Next, (3.63) follows from Proposition 3.7.3 and convergence in $L^2(e^{aY} dY)$. Expressions (3.64) and (3.65) with $\theta = 1/2$ follow immediately from Proposition 3.7.5 and convergence in $L^2(e^{aY} dY)$. The appearance of the terms Ka^2n in the exponent will be absorbed into the constant because a effectively becomes replaced by $n^{-1/2}a$. The $\theta = 0$ cases of (3.64) and (3.65) follow immediately from (3.63) and the fact that $e^{aX} + e^{aY} \leq 2e^{a(X+Y)}$. The proofs for general θ then follow easily by geometric interpolation (i.e., $\min\{a, b\} \leq a^\theta b^{1-\theta}$ for all $a, b \geq 0$ and all $\theta \in [0, 1]$). $\square$

Chapter 4: Convergence of ASEP to KPZ with basic coupling of the dynamics

We prove an extension of a seminal result of Bertini and Giacomin. Namely we consider weakly asymmetric exclusion processes with several distinct initial data simultaneously, then run according to the basic coupling, and we show joint convergence to the solution of the KPZ equation with the same driving noise in the limiting equation. Along the way, we analyze fine properties of nontrivially coupled solutions-in-law of KPZ-type equations.

4.1 Introduction and context

Interacting particle systems on the integer lattice have been a popular area of research in recent years. Of particular interest is the asymmetric simple exclusion process (ASEP), which was introduced by Spitzer [140] and subsequently generalized and explored in many works. ASEP is a Feller process on $\{0,1\}^{\mathbb{Z}}$ in which one starts with an initial configuration on $\mathbb{Z}$ consisting of some particles (1's) and some empty sites (0's), and the evolution of the dynamics can be described by having the particles independently perform asymmetric nearest-neighbor (continuous-time) random walks on $\mathbb{Z}$, but with jumps suppressed whenever one particle tries to jump onto another one. This hard-core repulsion effect between the particles makes the system physically interesting but also mathematically difficult to analyze.

In a seminal paper, Bertini and Giacomin [16] showed that under a certain scaling regime and specific tuning of the jump parameters, the fluctuations of ASEP are described by a nonlinear stochastic partial differential equation called the Kardar-Parisi-Zhang (KPZ) equation:

$$\partial_t h(t,x) = \partial_x^2 h(t,x) + (\partial_x h(t,x))^2 + \xi(t,x), \qquad x \in \mathbb{R}, t \geq 0,$$

where ξ is Gaussian space-time white noise, specified by the formal space-time covariance function $\mathbb{E}[\xi(t,x)\xi(s,y)] = \delta(t-s)\delta(x-y)$. More specifically, Bertini and Giacomin considered ASEP where the right jump rate for each particle equals $1+\epsilon^{1/2}$ and the left jump rate equals 1. They consider initial data which is "near stationarity" in a certain precise way. They then define a discrete height function $h_t^\epsilon(x)$ ($t \geq 0, x \in \mathbb{Z}$) as follows: $h_t^\epsilon(0)$ is the number of particles up to time t that have passed from 0 to 1, minus the number of particles that have passed from 1 to 0. Then $h_t^\epsilon(x)$ equals $h_t^\epsilon(0)$, plus the number of particles at time t which are between 0 and x (inclusive), minus the number of vacant sites between 0 and x (understood to be linearly interpolated when x is not an integer). They then prove that $\epsilon^{1/2}h^\epsilon(\epsilon^{-2}t, \epsilon^{-1}x) - \epsilon^{-1}t - t/24$ converges as $\epsilon \to 0$, to the Hopf-Cole solution of the KPZ equation (see Theorem 4.3.1 for a precise version).

The result was striking because it was one of the first examples of a particle system in a regime that exhibited non-Gaussian fluctuation behavior, and it was one of the works that paved the way to the field of KPZ universality for random growth models, see the survey [40] as well as subsequent recent work on particle systems that built on, generalized, or was inspired by the work of Bertini and Giacomin, e.g. [5, 12, 73, 57, 48, 47, 45, 153] just to name a few.

The main goal of the present work is to prove that in the fluctuation regime of [16], if one starts with *two or more* different initial data, and then one runs the particle system according to the *same* dynamics, then convergence to KPZ holds *jointly* with the same realization of the noise appearing in the limiting equation. When we refer to the "same dynamics," we are referring to the so-called *basic coupling*, a natural and important object that appears in many contexts when dealing with exclusion systems, e.g. in providing a full description of the ergodic theory of exclusion processes, see [111, 112, 72]. This basic coupling is described as follows: for each pair of sites $(x, y) \in \mathbb{Z}^2$ if a particle from both systems is present at x, and if a particle from one system jumps from x to y, then a particle from the other system also jumps from x to y at the same time assuming the target site is not blocked in the other configuration. This coupling can be constructed by the so-called "graphical construction" of ASEP, which randomly assigns directed arrows to each bond in $\mathbb{Z}$ according to independent Poisson point processes, see e.g. [93, 137]. A slightly more general definition and construction of the coupling is given in Subsection 3.1.

With this setup, let us now state our main result. We will say that an ϵ-indexed sequence of macroscopically scaled pairs $(\epsilon^{1/2}h^{1,\epsilon}(0, \epsilon^{-1}x), \epsilon^{1/2}h^{2,\epsilon}(0, \epsilon^{-1}x))$ of initial data for the height functions is *admissible* if it converges (jointly) in law to some limiting pair of height functions, and if it is tight in the sense that the L^p moments of its absolute value and of its spatial differences can be bounded via a Kolmogorov-Chentsov criterion with a sublinear growth rate at infinity. The precise assumption is given in Section 3, see Theorem 4.3.8.

Theorem 4.1.1. *Consider the weakly asymmetric scaling of ASEP from [16]. Let $(h_0^{1,\epsilon}, h_0^{2,\epsilon})$ be an admissible sequence of initial data. Evolve the corresponding height functions $h^{1,\epsilon}(t, x), h^{2,\epsilon}(t, x)$ according to the basic coupling described above. Then $(h^{1,\epsilon}, h^{2,\epsilon})$ converge jointly as $\epsilon \to 0$ to the solution of the KPZ equation driven by the same noise.*

This result will be stated more precisely and proved as Theorems 4.3.7 and 4.3.8 below. The main difficulty lies in the fact that for interacting particle systems such as ASEP, some of the jumps are suppressed due to the fact that particles are not allowed to jump onto other particles. To prove the result, one may convince themselves that it is somehow necessary to keep track of the noise as well as the height profile in the limit, not just the latter. At first glance, one might try to show that $(h^{1,\epsilon}, h^{2,\epsilon}, \xi^\epsilon)$ converge jointly as $\epsilon \to 0$ to (h^1, h^2, ξ), where $h^{i,\epsilon}$ are the rescaled, renormalized, and basically coupled height profile as described earlier, where ξ^ϵ keeps track of the Poisson clocks which excite the particles to jump, and where h^1, h^2 both solve the KPZ equation with the same noise ξ. Unfortunately, this approach is bound to fail because approximately half of the Poisson clocks go unused by the system due to suppressed jumps. In reality, the “correct" discretization of the noise consists of only those Poisson clocks which are used by the system. But this depends intricately on the initial data of the system. In other words, there is no natural choice of ξ^ϵ above: there is always a $\xi^{1,\epsilon}$ associated with $h^{1,\epsilon}$ and likewise there is $\xi^{2,\epsilon}$ for $h^{2,\epsilon}$. And the primary technical task is to relate the $\xi^{i,\epsilon}$ for $i = 1, 2$, in particular to prove that these converge to the same noise in the limit. So one runs into a vicious cycle which creates a difficulty in the arguments.

In terms of applications of our theorem, one can recover a few results about how joint solutions of KPZ behave when run according to the same noise ξ. Here is just one example: consider the stochastic Burgers equation

$$\partial_t u = \partial_x^2 u + \partial_x(u^2) + \partial_x \xi,$$

which is formally related to the KPZ equation by $u = \partial_x h$, and indeed one can define the solution this way interpreted in terms of distributions. Consider two solutions u^1, u^2 of stochastic Burgers driven by the same realization of ξ, started from two initial data u_0^1, u_0^2 respectively. Suppose that the initial data are ordered, i.e., $u_0^1 \leq u_0^2$ deterministically in the sense that $u_0^2 - u_0^1$ is a positive Borel measure. Then Theorem 4.1.1 implies that $u^1(t, \bullet) \leq u^2(t, \bullet)$ almost surely for all t, again interpreted in the sense that the difference is a positive measure. In other words, the KPZ dynamics preserve the property that the difference of height functions is nondecreasing. This is because the ordering is preserved at the level of the particle systems, see (A) below. This result can very likely be proved using other methods as well, for instance proving the result first for smooth noises ξ (see for instance Section 3 of [64]) and then using an approximation of space-time white noise by spatial mollifications and using the fact that the desired result is stable under limits and that the associated solutions converge after height renormalization (see e.g. [132]). One advantage in our discretization via ASEP is that the result is already obvious at the level of the particle system without using PDE techniques.

The input to proving our main theorem will require two steps. First we will prove a result (Theorem 4.2.3 below) about nontrivially coupled KPZ's, which says that two solutions-in-law of the KPZ equation with the property that their difference has zero quadratic variation in the x variable must in fact be driven by the same noise. This result may be of independent interest, and it will be the main tool to identify joint limit points of the coupled height functions. The other tool we will use is the monotonicity and attractivity properties of ASEP and related systems. It should be noted that our methods are easily generalizable to other types of basically coupled systems that satisfy these properties as well, such as joint convergence of the *symmetric* simple exclusion process to the Edwards-Wilkinson fixed point as well as higher-spin processes for which KPZ fluctuations are known, such as ASEP(q, J) [47]. We discuss the latter model in Subsection 3.5.

Outline: In Section 2, we prove a result about coupled solutions-in-law of the KPZ equation. In Section 3 we prove Theorem 4.1.1. Subsection 3.1 introduces the basic coupling model and the notations, Subsection 3.2 describes the result of Bertini-Giacomin in some detail, Subsection 3.3 contains the proof of our main result in the case of deterministic initial data (Theorem 4.3.7) and then Subsection 3.4 contains the main result for randomized initial data, Theorem 4.3.8. Subsection 3.5 then includes a discussion of how to generalize our argument to more complex models.

Acknowledgements: We thank Ivan Corwin for suggesting the problem. The author was partially supported by the Fernholz Foundation's Summer Minerva Fellows program, as well as summer support from Ivan Corwin's NSF grant DMS:1811143.

4.2 A result about nontrivially coupled KPZ's

To prove the main result, we use a continuum apparatus which allows us to efficiently identify joint limit points of the coupled particle system. To formulate our result we consider a slightly more general version of the KPZ equation with a parameter $\lambda \in \mathbb{R}$:

$$\partial_t h(t,x) = \partial_x^2 h(t,x) + \lambda(\partial_x h(t,x))^2 + \xi(t,x), \qquad x \in \mathbb{R}, t \geq 0. \tag{KPZ}$$

We use the notion of the so-called Hopf-Cole solution, which uses the fact that if h solves (KPZ) then $Z := e^{\lambda h}$ solves the multiplicative noise equation given by $\partial_t Z = \partial_x^2 Z + \lambda Z\xi$ which actually turns out to be well-posed using classical methods from [150]. To make this rigorous, one formulates all of this using the Duhamel principle:

Definition 4.2.1 (Hopf-Cole solution)**.** *Let $P(T,X) = \frac{1}{\sqrt{2\pi T}}e^{-X^2/2T}$, and let ξ denote a space-time white noise on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let Z_0 denote some (random) Borel measure on $\mathbb{R}$. We say that a continuous space-time process $h = (h(T,X))_{T>0, X\in\mathbb{R}}$ is a solution of* (KPZ) *if $\mathbb{P}$-almost surely, for every $T > 0$ and $X \in \mathbb{R}$, the process $Z(T,X) := e^{\lambda h(T,X)}$ satisfies the identity*

$$Z(T,X) = \int_{\mathbb{R}} P(T, X-Y) Z_0(dY) + \lambda \int_0^T \int_{\mathbb{R}} P(T-S, X-Y) Z(S,Y)\xi(dS, dY),$$

where the integral against the white noise is meant to be interpreted in the Itô-Walsh sense [150].

Next we will define the class of initial data for which our apparatus will be applicable. This class of functions will also be used extensively in later sections of the paper.

Definition 4.2.2. *Let $\alpha, \delta \in (0,1)$. A function $f : \mathbb{R} \to \mathbb{R}$ is said to be in the δ-weighted α-Hölder space $\mathscr{C}_\delta^\alpha(\mathbb{R})$ if*

$$\sup_{x\in\mathbb{R}} \frac{|f(x)|}{(1+|x|)^\delta} + \sup_{\substack{x,y\in\mathbb{R} \\ |x-y|\le 1}} \frac{|f(x)-f(y)|}{(1+|x|)^\delta |x-y|^\alpha} < \infty.$$

We turn $\mathscr{C}_\delta^\alpha$ into a Banach space by defining the norm of f to be the above quantity.

We are going to prove a result which roughly says that if we have two space-time processes defined on the same probability space, each solving (KPZ) in law, not necessarily driven by the same noise but their difference satisfies some specific nontrivial deterministic condition, then the two noises must in fact be the same.

Theorem 4.2.3. *Suppose we have two standard space-time white noises ξ^1, ξ^2 coupled onto the same probability space. Suppose furthermore that they satisfy the following conditions:*

1. $\mathbb{E}[(\xi^1, f)(\xi^2, g)] = 0$ *for all $f, g \in L^2(\mathbb{R}_+ \times \mathbb{R})$ which have disjoint supports.*

2. *For every $t > 0$ the spatial process $h^2(t,\cdot) - h^1(t,\cdot)$ has a.s. finite p-variation for some $p < 2$, where h^i is a solution of $\partial_t h^i = \partial_x^2 h^i + \lambda_i(\partial_x h^i)^2 + \xi^i$, for $i = 1, 2$. Here $\lambda_1, \lambda_2 \in \mathbb{R}$ and furthermore we assume that the initial data $h^i(0,\cdot) \in \mathscr{C}^\alpha_\delta$ for some $\alpha, \delta \in (0,1)$.*

Then $\xi^1 = \xi^2$.

We remark that the two noises are **not** assumed to be *jointly* Gaussian. This will be important while applying the theorem later.

Proof. Define a bilinear form I on $L^2(\mathbb{R}_+ \times \mathbb{R})$ by $I(\phi, \psi) := \mathbb{E}[(\xi^1, \phi)(\xi^2, \psi)]$. By Cauchy-Schwarz

$$|I(\phi, \psi)| \le \mathbb{E}[(\xi^1, \phi)^2]^{1/2}\mathbb{E}[(\xi^2, \psi)^2]^{1/2} = \|\phi\|_2 \|\psi\|_2.$$

Thus I is bounded, so by Riesz representation theorem there exists some bounded operator $A : L^2(\mathbb{R}_+ \times \mathbb{R}) \to L^2(\mathbb{R}_+ \times \mathbb{R})$ such that $I(\phi, \psi) = \langle \phi, A\psi \rangle_{L^2(\mathbb{R}_+\times\mathbb{R})}$ and $\|A\| \le 1$.

Note that $\langle \phi, A\psi \rangle_{L^2(\mathbb{R}_+\times\mathbb{R})} = 0$ whenever ϕ and ψ have disjoint supports on $\mathbb{R}_+ \times \mathbb{R}$. The reader may show that any operator on an L^2 space (associated with a sigma finite measure) which satisfies this property is necessarily a multiplication operator. Thus there exists some $\mathbf{v} \in L^\infty(\mathbb{R}_+ \times \mathbb{R})$ such that $A\phi = \mathbf{v} \cdot \phi$ for all $\phi \in L^2(\mathbb{R}_+ \times \mathbb{R})$. Note that $\|\mathbf{v}\|_{L^\infty(\mathbb{R}_+\times\mathbb{R})} = \|A\| \le 1$.

We have shown that if $\phi, \psi \in L^2(\mathbb{R}_+ \times \mathbb{R})$ then

$$\mathbb{E}[(\xi^1, \phi)(\xi^2, \psi)] = \int_{\mathbb{R}_+\times\mathbb{R}} \phi(t,x)\psi(t,x)\mathbf{v}(t,x)dt\, dx, \tag{4.1}$$

where $|\mathbf{v}(t,x)| \le 1$ a.e. Note that ξ^1, ξ^2 have *not* been shown or assumed to be jointly Gaussian. Our goal is now to show that $\mathbf{v} = 1$ a.e. on $\mathbb{R}_+ \times \mathbb{R}$.

For $i = 1, 2$ we define $X^i(t,x)$ for $t \ge 0$ and $x \in \mathbb{R}$ as the solution of the linear SPDE

$$\partial_t X^i = \partial_x^2 X^i + \xi^i,$$

with $X^i(0,x) = 0$. Letting h^i be as in the theorem statement, we can write $h^i(t,x) = X^i(t,x) + v^i(t,x)$, where $h_0^i(x) = h^i(0,x)$ and v^i is a remainder term which is locally Holder continuous of exponent strictly greater than $1/2$ in the spatial variable. For the KPZ equation on the *circle* $\mathbb{T}$, the existence of such a remainder term v^i was first proved as Theorem 1.10 in [86] using a preliminary version of the theory of regularity structures. We believe that the result on the *full line* $\mathbb{R}$ (which is what we need) can also be proved using regularity structures, however it has not been done in the literature thus far (in the introduction of [88], there is a discussion of the difficulties involved with making direct sense of the full-line KPZ equation). However, the full line result can instead be deduced from Definitions 3.2, 3.3, and Theorem 3.19 in [132] which uses the theory of paracontrolled products [79] to make direct sense of the full-line KPZ equation. The fact the notion of solution used there coincides with the Hopf-Cole solution also follows Theorem 3.19 there. However, that theorem assumes that the initial data lie in $\mathscr{C}_\delta^\alpha$ (see Assumptions 3.7 and Remark 3.8 in [132]) which is the only reason we have assumed such a restriction on the class of initial data in this theorem and in later parts of this paper. This assumption can likely be relaxed, but it does not seem to have been done in the literature thus far.

Now let $Y := X^2 - X^1$. Then

$$Y(t,x) = \big[h^2(t,x) - h^1(t,x)\big] + \big[v^1(t,x) - v^2(t,x)\big].$$

By assumption, for each fixed $t > 0$, each of the two terms in the square brackets have a.s. finite p-variation in the x variable, for some $p < 2$ (since the v^i are spatially Holder continuous of exponent strictly greater than $1/2$). Thus, Y has a.s. finite p-variation in the x variable.

Define a sequence of random variables

$$Q_N(t) := \sum_{k=1}^{2^N} \big(Y(t, 2^{-N}k) - Y(t, 2^{-N}(k+1))\big)^2.$$

Since Y is of finite p-variation in the x variable with $p < 2$, and since Q_N is approximating the *quadratic* variation, it follows that $Q_N(t) \to 0$ almost surely as $N \to \infty$. We claim that $\mathbb{E}[Q_N(t)] \to 0$ as well. To prove this, it suffices to show that $\sup_N \mathbb{E}[Q_N(t)^q] < \infty$ for some $q > 1$, as that implies uniform integrability. To show this uniform L^q bound, note that $(a-b)^2 \leq 2a^2+2b^2$ for all a, b, and recall that $Y = X^1 - X^2$. Therefore

$$Q_N(t) \leq 2 \sum_{k=1}^{2^N} \sum_{i=1,2} \big(X^i(t, 2^{-N}k) - X^i(t, 2^{-N}(k+1))\big)^2,$$

so that

$$\begin{aligned}
\mathbb{E}[Q_N(t)^q] &\leq 2 \cdot 2^{(N+1)(q-1)} \sum_{k=1}^{2^N} \sum_{i=1,2} \mathbb{E}\big[\big|X^i(t, 2^{-N}k) - X^i(t, 2^{-N}(k+1))\big|^{2q}\big] \\
&= 4 \cdot 2^{(N+1)(q-1)} \sum_{k=1}^{2^N} \mathbb{E}\big[\big|X^1(t, 2^{-N}k) - X^1(t, 2^{-N}(k+1))\big|^{2q}\big] \\
&= 2^{Nq+q+1} \mathbb{E}\big[\big|X^1(t, 2^{-N}) - X^1(t, 0))\big|^{2q}\big] \\
&\leq C_q \cdot 2^{Nq} \mathbb{E}\big[\big(X^1(t, 2^{-N}) - X^1(t, 0))\big)^2\big]^q
\end{aligned}$$

Here the first inequality is obtained by using the Hölder (or Jensen) inequality on the double sum from the previous expression, which allows us to bring the q^{th} power inside the sum at a cost of an extra factor $2^{(N+1)(q-1)}$. The equality in the second line holds because X^1 and X^2 have the same distribution as space-time fields, so the sum over $i = 1, 2$ simply doubles the expectation of the $i = 1$ case. The equality in the third line holds because X^1 is stationary in x (recall that it was started from zero initial data) and thus the terms in the sum do not depend on k. In the last inequality C_q is a constant depending on q but not N, and it holds because $X^1(t, 2^{-N}) - X^1(t, 0)$ has a centered normal distribution, and thus satisfies the standard “reverse Jensen" bounds. With all of this in place, we just need to show that $\mathbb{E}[\big(X^1(t, 2^{-N}) - X^1(t, 0))\big)^2] \leq C2^{-N}$. But this is standard, see for instance Section 2.3 of [85] for a precise computation which shows that $\mathbb{E}[(X^1(t, x) - X^1(t, y))^2] \leq C|x - y|$ where C is independent of t, x, y. Thus we have shown that $\mathbb{E}[Q_N(t)] \to 0$ as $N \to \infty$.

Recall that the goal is to show that $\mathbf{v} = 1$ a.e. To do this, we will now compute $\mathbb{E}[Q_N(t)]$ in a different manner using $\mathbf{v}$. We can write Y in mild form as $Y_t = p * (\xi^1 - \xi^2)$, where $*$ denotes space-time convolution and p is the standard heat kernel as always. Thus, by using (4.1) we see that

$$\mathbb{E}[Q_N(t)] = 2 \int_0^t \int_{\mathbb{R}} \left[\sum_{k=1}^{2^N} \big(p_{t-s}(x_k - z) - p_{t-s}(x_{k+1} - z)\big)^2 \right] \big(1 - \mathbf{v}(s,z)\big) dz\, ds, \tag{4.2}$$

where $x_k := k \cdot 2^{-N}$. Now we will show that the limit of this quantity is strictly positive for some $t > 0$ unless $1 - \mathbf{v}$ vanishes a.e. on $\mathbb{R}_+ \times \mathbb{R}$. The only major difficulty is that $\mathbf{v}$ has L^∞ regularity at best, and the part of the integrand in the square brackets is converging weakly as $N \to \infty$ to a measure which is singular with respect to 2D Lebesgue measure, so taking a limit of the above integral is somewhat tricky and will involve using the Lebesgue differentiation theorem from measure theory. Define

$$\alpha := \min_{\substack{t \in [1,2] \\ x \in [1,3]}} \big(p_t(x-1) - p_t(x+1)\big) > 0.$$

By using the relation $p_t(x) = \epsilon^{-1} p_{\epsilon^{-2}t}(\epsilon^{-1}x)$, valid for all $\epsilon, t > 0$ and $x \in \mathbb{R}$, we see that

$$\min_{\substack{t \in [\epsilon^2, 2\epsilon^2] \\ x \in [\epsilon, 3\epsilon]}} \big(p_t(x-\epsilon) - p_t(x+\epsilon)\big) = \epsilon^{-1}\alpha, \qquad \text{for all } \epsilon > 0.$$

Thus $\big(p_t(x-\epsilon) - p_t(x+\epsilon)\big)^2 \geq \alpha^2 \epsilon^{-2} \big(1_{[\epsilon,3\epsilon]} + 1_{[-3\epsilon,-\epsilon]}\big)(x)$, for all $t \in [\epsilon^2, 2\epsilon^2]$. Taking $\epsilon = 2^{-N-1}$, we see that

$$\sum_{k=1}^{2^N} \big(p_{t-s}(x_k - z) - p_{t-s}(x_{k+1} - z)\big)^2 \geq 4^{N+1} \alpha^2, \tag{4.3}$$

whenever $s \in [t - 2 \cdot 4^{-N-1}, t - 4^{-N-1}]$ and $z \in [0,1]$. Here $x_k = 2^{-N}k$ as always.

For $t \geq 0$ define $u(t) := \int_0^1 \int_0^t (1 - \mathbf{v}(s,z))dsdz$. By combining (4.2) and (4.3), we see that

$$\mathbb{E}[Q_N(t)] \geq 4^{N+1}\alpha^2\big(u(t-4^{-N-1}) - u(t - 2\cdot 4^{-N-1})\big).$$

By the Lebesgue differentiation theorem for nicely shrinking sets (see Theorem 3.21 in [68]), there exists a measure zero zet $S \subset [0,\infty)$ such that for $t \notin S$, the right side of the last expression converges as $N \to \infty$ to $\alpha^2 u'(t) = \alpha^2 \int_0^1 (1 - \mathbf{v}(t,z))dz$. But we know that $\mathbb{E}[Q_N(t)] \to 0$ for every t, so we have shown that $u'(t) = 0$ for all $t \notin S$. Thus $u(t) = u(0) = 0$. Since $\mathbf{v} \leq 1$, this implies that $\mathbf{v}(s,z) = 1$ for a.e. $(s,z) \in [0,\infty) \times [0,1]$. Of course, there is nothing special about the interval $[0,1]$ here. By changing the definition of $Q_N(t)$ so that the sum ranges over all k from $\lfloor 2^N a \rfloor$ to $\lfloor 2^N b \rfloor$, we can obtain the same result on $[0,\infty)\times[a,b]$ for any real numbers $a < b$.

We conclude that $\mathbf{v} = 1$ a.e. Thus by (4.1) we see that $\mathbb{E}[(\xi^1 - \xi^2, \phi)^2] = 0$ for all $\phi \in L^2(\mathbb{R}_+ \times \mathbb{R})$, and thus $\xi^1 = \xi^2$. $\square$

We now define a *cylindrical Wiener process* to be a family of Brownian motions $W_T(f)$, indexed by $f \in L^2(\mathbb{R})$, defined on the same probability space, and satisfying

$$\mathbb{E}[W_T(f)W_S(g)] = (S \wedge T)\langle f, g\rangle_{L^2(\mathbb{R})}.$$

for all $f, g \in L^2(\mathbb{R})$. Any space-time white noise ξ uniquely defines a cylindrical Wiener process W such that $W_T(f) = (\xi, f \otimes 1_{[0,T]})_{L^2(\mathbb{R}_+\times\mathbb{R})}$, and vice versa, so the two may be viewed as equivalent objects [150, 85].

Corollary 4.2.4. *Suppose we have two standard cylindrical Wiener processes W^1, W^2 coupled onto the same probability space. Suppose furthermore that they satisfy the following conditions:*

1. $\mathbb{E}[W_T^1(f)W_T^2(g)] = 0$ *for all $T \geq 0$ and all $f, g \in L^2(\mathbb{R})$ which have disjoint supports.*
2. *For $f, g \in L^2(\mathbb{R})$, the processes $(W_T^1(f))_{T\geq 0}$ and $(W_T^2(g))_{T\geq 0}$ are both martingales with respect to their **joint** filtration.*

3. *For every $t > 0$ the spatial process $h^2(t,\cdot) - h^1(t,\cdot)$ has a.s. finite p-variation for some $p < 2$, where h^i is a solution of $\partial_t h^i = \partial_x^2 h^i + F^i(\partial_x h^i) + dW^i$. Here F^1, F^2 are admissible nonlinearities as mentioned above.*

Then $W^1 = W^2$.

Proof. Define ξ^1, ξ^2 to be the random elements of $\mathcal{S}'(\mathbb{R}_+ \times \mathbb{R})$ such that

$$(\xi^i, \phi) := \int_0^\infty \langle \phi(T,\cdot), dW_T^i \rangle_{L^2(\mathbb{R})}; \text{ for all } \phi \in \mathcal{S}(\mathbb{R}_+ \times \mathbb{R}).$$

Note that $\mathbb{E}[(\xi^i,\phi)^2] = \|\phi\|^2_{L^2(\mathbb{R}_+\times\mathbb{R})}$ so the ξ^i are space-time white noises and we can stochastically extend the definition of (ξ^i, ϕ) to all $\phi \in L^2(\mathbb{R}_+ \times \mathbb{R})$.

Let $f, g \in L^2(\mathbb{R})$. Since $W^1(f)$ and $W^2(g)$ are martingales in their *joint* filtration we see that

$$\mathbb{E}\big[\big(W_T^1(f) - W_S^1(f)\big)\big(W_T^2(g) - W_S^2(g)\big)\big] = \mathbb{E}[W_T^1(f)W_T^2(g)] - \mathbb{E}[W_S^1(f)W_S^2(g)],$$

which equals zero whenever f, g have disjoint supports. From this it follows (using approximation by elementary integrands) that $\mathbb{E}[(\xi^1,\phi)(\xi^2,\psi)] = 0$ whenever ϕ and ψ have disjoint supports on $\mathbb{R}_+ \times \mathbb{R}$. Thus the conditions of Theorem 4.2.3 are satisfied, so $\xi^1 = \xi^2$, i.e., $W^1 = W^2$. $\square$

4.3 Proof of the main theorem

We will now derive some consequences of Theorem 4.2.3 in the context of interacting particle systems. In particular we will prove Theorem 4.1.1 in the case of [16] and [47]. Although our results are for WASEP, they can be extended quite easily to some other systems, so we describe in some generality a class of particle systems that we use.

4.3.1 The basic coupling, height functions, and notation

Although we consider ASEP for most of the paper, we would like to describe some extensions to more complicated models in later subsections. Thus we give a slightly more general description of the types of processes that are covered by our result.

In order to describe our result in full generality, fix $J \in \mathbb{N}$ and consider a function $b : \{-1,1\} \times \{0,...,J\}^2 \to [0,1]$. We consider Feller processes on the state space $S := \{0,...,J\}^{\mathbb{Z}}$ which are described by the following dynamics. Each ordered pair $(x, x+1)$ and $(x, x-1)$ has a Poisson clock of rate 1. Every time the clock associated to (x,y) rings, one particle jumps from x to y with probability $b(y-x, \eta(x), \eta(y))$ and stays there with probability $1-b(y-x,\eta(x),\eta(y))$. However, the jump is suppressed if there is no particle at x, or if there are already J particles at y (equivalently we can just impose that $b(i,0,\cdot) = 0 = b(i,\cdot,J)$ for all $i = -1,1$). The pre-generator of such a process acts on local functions f by the formula

$$Lf(\eta) = \sum_{x,y\in\mathbb{Z}:|x-y|=1} b(y-x,\eta(x),\eta(y))\big(f(\eta+e_y-e_x)-f(\eta)\big), \tag{4.4}$$

where $e_x(z) = 1_{\{x=z\}}$, and $f : S \to \mathbb{R}$ is some local function. This process is called a *nearest-neighbor generalized-misanthrope process* if b is increasing in the $\eta(x)$ variable and decreasing in the $\eta(y)$ variable. Examples include ASEP and more generally ASEP(q,j) as considered in [47]. See Subsection 3.5 for more on the latter.

For nearest-neighbor generalized misanthrope processes there is a natural way to run the dynamics associated to several initial data coupled together. This is usually called the **basic coupling.** Specifically for $x \in \mathbb{Z}$ we associate to each directed bond $(x,x+1)$ and $(x,x+1)$ Poisson clocks of rate one, as well as iid uniform random variables $\{U_i(x,x+1)\}_{i\geq1}$ and $\{U_i(x+1,x)\}_{i\geq1}$ which are independent of the Poisson clocks on that bond. Whenever the i^{th} Poisson clock associated to $(x,x+1)$ rings, a particle jumps from x to $x+1$ only when $b(1,\eta(x),\eta(x+1)) < U_i(x,x+1)$, and similarly for $(x, x-1)$ with $b(-1, \eta(x), \eta(x-1))$ and $U_i(x, x-1)$. In this way, we can define a Markov process on the product $S \times S$ of the individual state spaces which describes the evolution of two particle systems coupled so that each marginal onto S is a Feller process with generator L given above, and moreover (by the monotonicity properties of b) the two individual particle systems stay dominated for all time if they start dominated (see (A) below). When $J = 1$ there is a straightforward way to describe the coupling without any uniform variables, instead using Poisson clocks of different rates on each bond. For the seminal work on coupled processes, see e.g. [111, 93]. Our description of the basic coupling is in the spirit of [93], while [111] instead chooses to explicitly write the generator for the entire coupled system on the product space.

If $(\eta_t(x))_{t\geq 0}$ is a generalized misanthrope process on the state space $\{0, ..., J\}$ then we define the *height function*

$$h_t(x) := \begin{cases} h_t(0) + \sum_{k=0}^{x}(2\eta_t(k) - J), & x \geq 0, \\ h_t(0) + \sum_{k=0}^{-x}(2\eta_t(-k) - J), & x < 0, \end{cases},$$

where $h_t(0)$ equals twice the current through the origin up to time t, i.e., twice the number of particles which have moved from the site $x = 0$ to the site $x = 1$ minus twice the number of particles which have moved from the site $x = 1$ to the site $x = 0$ up to time t.

The height functions associated to nearest-neighbor misanthrope processes have two useful properties. The dynamics preserve their ordering as well as the ordering of their spatial derivative:

$$h_t^1(x) \leq h_t^2(x) \text{ for all } t \geq 0, x \geq 0 \text{ if } h_0^1(x) \leq h_0^2(x) \text{ for all } x \geq 0, \tag{M}$$

$$\eta_t^1(x) \leq \eta_t^2(x) \text{ for all } t \geq 0, x \geq 0 \text{ if } \eta_0^1(x) \leq \eta_0^2(x) \text{ for all } x \geq 0. \tag{A}$$

Property (M) is usually called *monotonicity* of the particle system, whereas property (A) is usually called *attractivity* of the system. Both properties are easily proved by considering the action of a single jump excitation in the joint system. In terms of the SPDE limits, (M) says that the limiting height functions h^1 and h^2 are coupled so that $h^1 \leq h^2$ if $h^1(0,\cdot) \leq h^2(0,\cdot)$, and (A) says that $h^2(t,\cdot) - h^1(t,\cdot)$ is a nondecreasing function for every $t > 0$ if it is nondecreasing for $t = 0$.

Let us now establish some notation. A function $h : \mathbb{Z} \to \mathbb{Z}$ is called **viable** if there is a particle system associated to it, in other words if $h_t(x+1) - h_t(x) \in \{-J, -J+2, ..., J-2, J\}$ for all x. Likewise a function from $\mathbb{R} \to \mathbb{R}$ will be called viable if its restriction to $\mathbb{Z}$ is viable and if its value at non-integers is linearly interpolated from the two nearest integer values. An obvious but important property used below is that the class of admissible height functions is closed under the operations $\max$ and $\min$.

Given some collection $h^1, ..., h^n : \mathbb{R}_+ \times \mathbb{Z} \to \mathbb{R}$ of time-evolving height profiles, we will often define "rescaled and renormalized" versions of them which converge in law to the solution of the KPZ equation. In all of these cases what we will mean is that there exist some constants a_ϵ, b_ϵ such that

$$h^{i,\epsilon}(t,x) := a_\epsilon h^i(\epsilon^{-2}t, \epsilon^{-1}x) + b_\epsilon t \tag{4.5}$$

converges in law to the solution of (KPZ).

Whenever we have an evolving height function $h(t,x)$ in our model, we will denote by h^ϵ its rescaled and renormalized version converging to KPZ. Thus h^ϵ is a random function from $\mathbb{R}_+ \times \epsilon\mathbb{Z} \to \mathbb{R}$ that depends on ϵ in three different ways: through the initial data which is generally changing with ϵ, through the parameters of the model which are being weakly scaled as $p = \frac{1}{2} + \frac{1}{2}\sqrt{\epsilon}$ and $q = \frac{1}{2} - \frac{1}{2}\sqrt{\epsilon}$ (this will be explained below), and through the renormalization constants and diffusive scaling as in (4.5). Often we will have several height functions $h^1, h^2, ..., h^n$ which are coupled via the same dynamics, we will denote their rescaled versions as $h^{1,\epsilon}, h^{2,\epsilon}, ..., h^{n,\epsilon}$. We will use the capital letters $(H^1, H^2, ..., H^n)$ to denote the *joint* continuum limits of the rescaled fields $(h^{1,\epsilon}, h^{2,\epsilon}, ..., h^{n,\epsilon})$. Thus the H^i are random continuous functions from $\mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ which are defined on the same probability space as each other. We will always use the subscript 0 to denote the initial data both in the prelimit and the limit, i.e., $H_0^i = H^i(0,\cdot)$, $h_0^{i,\epsilon} = h^{i,\epsilon}(0,\cdot)$, and so on.

Often we will have some initial data $h_0^{1,\epsilon}, h_0^{2,\epsilon}, ..., h_0^{k,\epsilon}$ and from these we will build more initial data $h_0^{k+1,\epsilon}, ..., h_0^{k+n,\epsilon}$. We will always denote by $h^{i,\epsilon}$ (i.e., without the zero subscript) to denote the evolution of the *coupled* the process started from $h_0^{i,\epsilon}$. In other words, the dynamics of the newly constructed $h^{i,\epsilon}$ are always implicitly assumed to be driven by the same realization of the Poissonian clocks (and uniform variables, if $J > 1$) as those of the original $h^{i,\epsilon}$.

Whenever we refer to "convergence" of $(h_0^{i,\epsilon})_{i=1}^k$ to $(H_0^i)_{i=1}^k$, we mean convergence in $C(\mathbb{R})^k$, where $C(\mathbb{R})$ is the space of continuous functions on $\mathbb{R}$ equipped with the the topology of uniform convergence on compacts, which is completely metrizable via the same metric

$$d(f,g) := \sum_{n\geq 1} 2^{-n} \max\big\{1, \sup_{x\in[-n,n]} |f(x)-g(x)|\big\}.$$

Sometimes we use the stronger topology of $\mathscr{C}_\delta^\alpha(\mathbb{R})$ from Definition 4.2.2 and we specify whenever we do this. When we refer to the convergence of the entire height profile $h^{i,\epsilon}$ to H^i, we mean in the Skorohod space $D([0,T], C(\mathbb{R})^k)$ for every $T > 0$.

4.3.2 The convergence result of Bertini-Giacomin

Throughout Subsections 3.2, 3.2, and 3.4 we consider ASEP, which corresponds in (4.4) to the choices $J = 1$, $b(1,1,0) = p$, and $b(-1,1,0) = q$ where $p, q \geq 0$. In our ϵ-dependent model below, p will be scaled as $\frac{1}{2} + \frac{1}{2}\epsilon^{1/2}$ while q will be scales as $\frac{1}{2} - \frac{1}{2}\epsilon^{1/2}$.

The main result of [16] can be formulated as follows. We would like to emphasize once again that the height functions considered in the theorem below depend on ϵ in three different ways: through the initial data which is generally changing with ϵ, through the parameters of the model which are being weakly scaled as $p = \frac{1}{2} + \frac{1}{2}\epsilon^{1/2}$ and $q = \frac{1}{2} - \frac{1}{2}\epsilon^{1/2}$, and through the renormalization constants and diffusive scaling as in (4.5).

Theorem 4.3.1 (Theorem 2.3 of [16])**.** *Let h_0^ϵ be a deterministic sequence of initial data such that $\epsilon^{1/2}h_0^\epsilon(\epsilon^{-1}x)$ converges in $\mathscr{C}_\delta^\alpha$ to some H_0, where $0 < \alpha < 1/2$ and $0 < \delta < 1$. Let h^ϵ denote the rescaled and renormalized height function as in* (4.5)*, with $a_\epsilon = \epsilon^{1/2}$ and $b_\epsilon = \frac{1}{2}\epsilon^{-1} + \frac{1}{24}$. Then h^ϵ converges in law to the Hopf-Cole solution of* (KPZ)*. The initial data of the limiting object is given by the limit in $\mathscr{C}_\delta^\alpha$ of $\epsilon^{1/2}h_0^\epsilon(\epsilon^{-1}x)$. The convergence is obtained with respect to the topology of the Skorohod space $D([0,T], C(\mathbb{R}))$, for all $T > 0$.*

Let us remark that convergence in $\mathscr{C}_\delta^\alpha$ is slightly different than the actual assumption on the initial data given in [16]. Specifically, in Definition 2.2 of [16], the authors considered *possibly random* initial data which are "near stationarity" in the sense that if $Z_0^\epsilon := \exp(h_0^\epsilon)$ then one has the moment bounds $\|Z_0^\epsilon(x)\|_p \leq Ce^{ax}$ and $\|Z_0^\epsilon(x) - Z_0^\epsilon(y)\|_p \leq C|x-y|^{1/2}e^{a(|x|+|y|)}$, uniformly in x, y, ϵ. Here p is some exponent larger than 10 and $\|A\|_p := \mathbb{E}[|A|^p]^{1/p}$. The substance of their proof is unchanged when the exponent $1/2$ in the second bound is changed to arbitrary $\alpha \in (1/p, 1/2)$. For technical reasons we will find it convenient to work with deterministic initial data which converge in $\mathscr{C}_\delta^\alpha$, which clearly satisfy these bounds. In fact even functions of *linear* growth would satisfy these bounds, so our assumption of sublinear growth and deterministic data is actually substantially *more* restrictive. We will randomize the assumptions on our initial data in Subsection 3.4.

4.3.3 Main result: joint convergence for ASEP

Our goal is to extend Theorem 4.3.1 so that one may consider the limiting height field started from any finite collection of (sequences of) initial data $(h_0^{i,\epsilon})_{i=1}^k$ whose dynamics are jointly run according to the basic coupling. The goal is to obtain convergence in $D([0,T], C(\mathbb{R})^k)$. We are going to do this in a manner which is essentially orthogonal to proof of the original convergence result of [16], by exploiting Theorem 4.2.3 and (M) and (A).

Lemma 4.3.2. *Suppose that we have two deterministic sequences of initial data $h_0^{1,\epsilon}$ and $h_0^{2,\epsilon}$ which both converge in $\mathscr{C}_\delta^\alpha$ to the same initial data. For any joint limit point (H^1, H^2) of the basically coupled space-time processes, we have $H^1(t,x) = H^2(t,x)$ for all t, x a.s.*

Proof. One readily checks that if two height functions are viable, then so are their maximum and minimum. We thus define $h_0^{3,\epsilon} := \max\{h_0^{1,\epsilon}, h_0^{2,\epsilon}\}$ and $h_0^{4,\epsilon} := \min\{h_0^{1,\epsilon}, h_0^{2,\epsilon}\}$. It is clear that $h_0^{3,\epsilon}$ and $h_0^{4,\epsilon}$ both converge in $\mathscr{C}_\delta^\alpha$ to the same initial data as $h_0^{1,\epsilon}$ and $h_0^{2,\epsilon}$. By (M) is also clear that $h^{4,\epsilon}(t,x) \le h^{i,\epsilon}(t,x) \le h^{3,\epsilon}(t,x)$ for $i = 1, 2$ and all t, x, ϵ.

Letting (H^1, H^2, H^3, H^4) denote a joint limit point of all four processes, we see that it must satisfy $H^4 \le H^i \le H^3$ for $i = 1, 2$. It is also true that $H_0^1 = H_0^2 = H_0^3 = H_0^4$ because $h_0^{1,\epsilon}$ and $h_0^{2,\epsilon}$ converge in $\mathscr{C}_\delta^\alpha$ to the same function and hence so do their max and min. The KPZ equation satisfies uniqueness in law, thus two solutions started from the same initial data have the same expectation, i.e., $\mathbb{E}[H^4(t,x)] = \mathbb{E}[H^3(t,x)]$.

Since $H^4(t,x) \le H^3(t,x)$ and $\mathbb{E}[H^4(t,x)] = \mathbb{E}[H^3(t,x)]$, we conclude that $H^4 = H^3$ a.s. Since H^1, H^2 are nested in between H^3 and H^4, we conclude that $H^1 = H^2$. $\square$

The next lemma will be the key behind all subsequent results. It proves the main result in the very special case that the two initial data are ordered as in (A), and it will be proved using the results of Section 2.

Lemma 4.3.3. *If $h_0^{1,\epsilon}$ and $h_0^{2,\epsilon}$ which are both deterministic, their difference is nondecreasing for every ϵ, and they converge weakly to initial data H_0^1 and H_0^2, then $h^{1,\epsilon}$ and $h^{2,\epsilon}$ converge jointly to the solution of the KPZ equation driven by the same noise.*

Proof. Note by (A) that the dynamic of the particle system preserves the condition that the difference of height functions is nondecreasing. Thus if $h_0^{2,\epsilon} - h_0^{1,\epsilon}$ is nondecreasing, then we know that $h^{2,\epsilon}(t,\cdot) - h^{1,\epsilon}(t,\cdot)$ is a.s. nondecreasing for every t. In particular, if (H^1, H^2) is a joint limit point of $(h^{1,\epsilon}, h^{2,\epsilon})$, then $H^2(t,\cdot) - H^1(t,\cdot)$ is nondecreasing (and in particular, of finite variation) for every t. Thus condition *(3)* of Corollary 4.2.4 is satisfied.

Now we just need to make sure that the conditions *(1)* and *(2)* of Corollary 4.2.4 is satisfied. For this we need to look into the precise details of how exactly Bertini and Giacomin proved their result. They first noted that of one defines

$$Z_t^{i,\epsilon}(x) := \exp\big(\epsilon^{1/2} h^{i,\epsilon}(t,x) - (\frac{1}{2}\epsilon^{-1} - \frac{1}{24})t\big),$$

then the $Z^{i,\epsilon}$ solve a discrete parabolic martingale-driven SPDE:

$$dZ_t(x) = (1+2\epsilon^{1/2})^{1/2} \Delta Z^{i,\epsilon}(x)dt + dM_t^{i,\epsilon}(x), \tag{4.6}$$

where $\Delta f(x) := \frac{1}{2}(f(x+1)+f(x-1)-2f(x))$, and $M^{i,\epsilon}(x)$ are jump martingales with the property that

$$\langle M^{i,\epsilon}(x), M^{j,\epsilon}(y)\rangle_t = 0 \qquad \text{if} \qquad x \neq y \tag{4.7}$$

for all $i,j = 1,2$. See equation (3.13) in [16].

Bertini and Giacomin then proceed to show that, for smooth functions $\phi \in C_c^\infty(\mathbb{R})$, if one defines the martingale $M_t^{i,\epsilon}(\phi) := \epsilon \sum_{x\in\mathbb{Z}} \phi(\epsilon x) M_t^{i,\epsilon}(x)$, then any limit point (joint over all $\phi \in C_c^\infty$ and all $i = 1,2$) of $M_t^{i,\epsilon}(\phi)$ as $\epsilon \to 0$ is a continuous martingale $M_t^i(\phi)$. In the language of [150], the collection of martingales $M_t^i(\phi)$, as ϕ ranges over all smooth functions, form an *orthogonal martingale measure*, in the sense that $\langle M_t^i(\phi), M_t^j(\psi)\rangle = 0$ whenever ϕ, ψ have disjoint supports and $i,j = 1,2$ (this is clear because the corresponding statement is true even in the prelimit, by the property that $\langle M^{i,\epsilon}(x), M^{j,\epsilon}(y)\rangle_t = 0$ if $x \neq y$ and $i,j = 1,2$).

Now consider any joint limit point (H^1, H^2) of $(h^{1,\epsilon}, h^{2,\epsilon})$. Let $Z^1 := e^{H^1}$ and $Z^2 := e^{H^2}$. Bertini and Giacomin show using (4.6) that the Z^i must satisfy the relation

$$(Z^i_t,\phi)_{L^2(\mathbb{R})} - \int_0^t (Z^i_s,\phi'')_{L^2(\mathbb{R})}ds = M^i_t(\phi).$$

Using the language of [100] and [150], Bertini and Giacomin then use this to show that for each $i=1,2$ one can construct the driving noise W^i of Z^i as an Ito-Walsh stochastic integral against M^i. By the properties of Ito-Walsh stochastic integrals, it then automatically follows that $\langle W^i(\phi),W^j(\psi)\rangle = 0$ for $i,j=1,2$ and ϕ,ψ of disjoint supports. Indeed, thus is is because the corresponding property is true for M^i and because W^i is a stochastic integral against M^i. Thus the conditions of Corollary 4.2.4 are satisfied and so $H^1=H^2$. $\square$

Remark 4.3.4. *Note that the results of the previous two lemmas generalize fairly straightforwardly to the case where we have $k>2$ distinct initial data converging in $\mathscr{C}^\alpha_\delta(\mathbb{R})^k$. Indeed, if $(H^1,...,H^k)$ is a joint limit point of the height functions and if any subpair (H^i,H^j) is driven by the same noise, then they are all driven by the same noise. Here we are implicitly using the fact that the driving noise can be deterministically recovered from any solution-in-law of the KPZ equation, which is a nontrivial fact that can be deduced by combining the orthomartingale theory of [150] with the Hopf-Cole transform and a positivity result of [122]. Alternatively this fact can also be deduced more directly from the pathwise theories developed by [86, 132]. Alternatively, even without using any of those aforementioned results, one can recognize that our proof strategy in both lemmas was done in such a way that the proofs generalize directly to several initial data. Indeed, in the proof of Lemma 4.3.2 one can consider the max and min of k distinct initial profiles, and these are still viable and convergent to the same limit in $\mathscr{C}^\alpha_\delta$. In the proof of Lemma 4.3.3, it is clear that one can keep track of both the noises as well as the height functions in the limit. Likewise, subsequent results such as Proposition 4.3.6, Theorem 4.3.7, and Theorem 4.3.8 also generalize to more than two initial profiles, either by using the nontrivial fact mentioned earlier or by working through the logic in the proofs directly.*

Now that we have proved the main result in the special case when the two initial data are ordered deterministically at the level of the particle system, the next step will be to prove the claim for two initial data which are smooth or at least differentiable in some strong enough sense. Then we can dominate both of the initial data by some third initial data whose derivative is larger than both of the individual initial data, and then apply Lemma 4.3.3 to conclude that the noises for all three are the same. This will be done in Proposition 4.3.6 below, but first we need a lemma.

We henceforth define $\mathcal{V}_\epsilon$ to be all functions of the form $\epsilon^{1/2} f(\epsilon^{-1}x)$ where f is a viable height function as defined in Subsection 3.1. We also let $\mathscr{C}^1_\delta$ to be the set of all continuously differentiable functions on $\mathbb{R}$ such that $\sup_{x\in\mathbb{R}}(1+|x|)^{-\delta}(|f'(x)|+\int_0^x |f'(u)|du) < \infty$. It is clear that $\mathscr{C}^1_\delta$ is a Banach space if we define its norm to be $|f(0)|$ plus that quantity[1] and that it embeds compactly into $\mathscr{C}^\alpha_{\delta'}$ whenever $\alpha < 1$ and $\delta' > \delta$. More generally we will often use the fact that if $0 < \alpha_1 < \alpha_2 \leq 1$ and $1 > \delta_1 > \delta_2 > 0$ then $\mathscr{C}^{\alpha_2}_{\delta_2}$ embeds compactly into $\mathscr{C}^{\alpha_1}_{\delta_1}$. This follows from Arzela-Ascoli together with the interpolation properties of Hölder seminorms, see e.g. Lemma 24.14 of [62] for the elementary proof, or [118] for a more general theory on Hölder spaces and their embeddings via Littlewood-Paley theory (Section 2 of [132] also has a nice discussion of the latter). We now give an approximation algorithm $\mathcal{A}^\epsilon$ for smooth functions by rescaled viable functions, and moreover the algorithm preserves the property that the difference of two functions is nondecreasing.

Lemma 4.3.5. *Fix $\alpha \in (0, 1/4)$ and $\delta < \delta' \in (0, 1)$. Then there exists a family of maps $\mathcal{A}^\epsilon : \mathscr{C}^1_\delta \to \mathscr{C}^\alpha_\delta \cap \mathcal{V}_\epsilon$ with the following properties:*

- *For all $f \in \mathscr{C}^1_\delta$, we have that $\mathcal{A}^\epsilon(f) \to f$ in $\mathscr{C}^\alpha_{\delta'}$ as $\epsilon \to 0$.*
- *$\mathcal{A}^\epsilon(g) - \mathcal{A}^\epsilon(f)$ is nondecreasing whenever $g - f$ is nondecreasing.*

Proof. We will construct $\mathcal{A}^\epsilon(f)$ on $\epsilon\mathbb{Z}$. The values in between are understood to be linearly interpolated.

[1]This is **not** a standard definition of $\mathscr{C}^1_\delta$, we have only defined it in this way for convenience of the arguments given later. Strictly speaking, we should really call this space BV_δ or something similar due to the defining condition that the variation is bounded by $C|x|^\delta$.

Note that $(1+|x|)^{-\delta'}(|f(x)|+|f'(x)|)$ can be viewed as a continuous function on the *closed* interval $[-\infty,\infty]$, which vanishes at the endpoints $-\infty$ and ∞. Suppose $(1+|x|)^{-\delta'}(|f(x)|+|f'(x)|) < 2\epsilon^{1/2}$ on $[-\infty,-M]\cup[M,\infty]$ where implicitly $M=M_\epsilon\geq 1$. We define $\mathcal{A}^\epsilon(f)$ on $(-\infty,-M]$ to just oscillate between the two values in $\epsilon^{1/2}\mathbb{Z}$ which are closest to $f(-M)$.

Next we define $\mathcal{A}^\epsilon(f)$ on the interval $[-M,M]$. Break $[-M,M]$ into $\lfloor\epsilon^{-1/4}\rfloor$ equally sized intervals of length $2M/\lfloor\epsilon^{-1/4}\rfloor$. On each of those intervals I, let $x_I^\epsilon := \min I\cap\epsilon\mathbb{Z}$. For $x\in\epsilon\mathbb{Z}$ such that $x_I\leq x\leq \max\{x_I+|f'(x_I)|\epsilon^{3/4}, \max I\cap\epsilon\mathbb{Z}\}$ we define f inductively by the formula $\mathcal{A}^\epsilon f(x+\epsilon)-\mathcal{A}^\epsilon f(x) := \text{sign}(f'(x_I))\epsilon^{1/2}$. For $x\in\epsilon\mathbb{Z}$ such that $x_I+|f'(x_I)|\epsilon^{3/4}\leq x\leq \max I\cap\epsilon\mathbb{Z}$ we simply define $\mathcal{A}^\epsilon f(x+\epsilon)-\mathcal{A}^\epsilon f(x) := \epsilon^{1/2}(-1)^{x/\epsilon}$.

Finally, define $\mathcal{A}^\epsilon(f)$ on $[M,\infty)$ by the formula $\mathcal{A}^\epsilon(f)(x+\epsilon)-\mathcal{A}^\epsilon(f)(x) := \epsilon^{1/2}(-1)^{x/\epsilon}$.

From our construction it is clear that $\mathcal{A}^\epsilon(g)-\mathcal{A}^\epsilon(f)$ is nondecreasing whenever $g-f$ is nondecreasing. This is because the latter is equivalent to $g'\geq f'$.

Note that for all $x\in[-M,M]$, the quantity $\mathcal{A}^\epsilon(f)(x)$ is always within $O(\epsilon^{1/4})$ of $f(-M)+\epsilon^{1/4}\sum_{I:\sup I<x} f'(x_I)$, which is a Riemann sum approximation to $f(-M)+\int_{-M}^x f'(t)dt$. Consequently $\mathcal{A}^\epsilon(f)$ converges pointwise to f as $\epsilon\to 0$.

Next we prove that there exists $C>0$ independent of x,y,ϵ (but in general dependent on f) such that $|\mathcal{A}^\epsilon f(x)-\mathcal{A}^\epsilon f(y)|\leq C|x|^\delta|x-y|^{1/4}$ whenever $|y-x|\leq 1$ and $x\in\mathbb{R}$. This is enough to prove relative precompactness of $\{\mathcal{A}^\epsilon(f)\}_{\epsilon\in(0,1]}$ inside of $\mathscr{C}^\alpha_{\delta'}$ (because $\alpha<1/2$ and $\delta'>\delta$), which would finish the proof. To prove this inequality, we first consider the case where $x,y\in[-M,M]$ and $1\geq|x-y|>\epsilon^{3/4}$. In this case, note that $\mathcal{A}^\epsilon f(x)-\mathcal{A}^\epsilon f(y)$ is always within $\|f'\|_{L^\infty([y,x])}\epsilon^{1/4}$ of $\epsilon^{1/4}\sum_{I:y<\sup I<x} f'(x_I)$. Now the number of intervals I in the approximation

scheme such that such that $y < \sup I < x$ is bounded above by $\epsilon^{-1/4}/2M \leq \epsilon^{-1/4}$. Consequently we find that $|\mathcal{A}^\epsilon f(x) - \mathcal{A}^\epsilon f(y)| \leq 2\epsilon^{1/4}\|f'\|_{L^\infty([y,x])}$. Now since $f \in \mathscr{C}^1_\delta$ and $|y-x| \leq 1$ we find that $\|f'\|_{L^\infty[y,x]} \leq \|f\|_{\mathscr{C}^1_\delta}|x|^\delta$, proving the claim in this case since $\epsilon^{1/4}$ can be bounded above by $|x-y|^{1/3}$ (recall we assumed $|x-y| > \epsilon^{3/4}$). Next we consider the case where $\epsilon < |x-y| \leq \epsilon^{3/4}$. Then we can use the naive bound

$$\begin{aligned}|\mathcal{A}^\epsilon f(x) - \mathcal{A}^\epsilon f(y)| &\leq \epsilon^{1/2} + \sum_{u\in\epsilon\mathbb{Z}\cap[x,y]} |\mathcal{A}^\epsilon f(u+\epsilon) - \mathcal{A}^\epsilon f(u)| \\ &\leq \epsilon^{1/2} + \epsilon^{-1/2}|x-y| \\ &\leq \epsilon^{1/2} + \epsilon^{1/4} \leq 2|x-y|^{1/4}.\end{aligned}$$

Finally, we consider the case where $|x-y| < \epsilon$. In this case it is clear that since the global Lipchitz constant of $\mathcal{A}^\epsilon(f)$ never exceeds $\epsilon^{-1/2}$ that one has

$$|\mathcal{A}^\epsilon(f)(x) - \mathcal{A}^\epsilon(f)(y)| \leq \epsilon^{-1/2}|x-y| \leq \epsilon^{-1/2}\epsilon^{1/2}|x-y|^{1/2} = |x-y|^{1/2}.$$

$\square$

Proposition 4.3.6. *Suppose that we have two viable deterministic sequences of initial data such that their re-scaled versions $h_0^{1,\epsilon}$ and $h_0^{2,\epsilon}$ converge in $\mathscr{C}^\alpha_\delta$ to functions H_0^1 and H_0^2 respectively, where $0 < \alpha < 1/2$ and $0 < \delta < 1$. Assume that $H_0^1, H_0^2 \in \mathscr{C}^1_\delta$. Then for any joint limit point (H^1, H^2) of $(h^{1,\epsilon}, h^{2,\epsilon})$, H^1 and H^2 are solutions of the KPZ equation driven by the same noise.*

Proof. Choose some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which one may define, for each $\epsilon \in (0,1]$, a system of i.i.d. Poisson clocks of rate $p = \frac{1}{2} + \frac{1}{2}\sqrt{\epsilon}$ and rate $q = \frac{1}{2} - \frac{1}{2}\sqrt{\epsilon}$ associated to each bond $\{x, x+1\}$ with $x \in \mathbb{Z}$. For different values of ϵ these can be coupled in an arbitrary manner; ultimately it is irrelevant.

Define $H_0^3(x) := \int_0^x \max\{\partial_x H_0^1(u), \partial_x H_0^2(u)\}du$, so that $H_0^3 - H_0^1$ and $H_0^3 - H_0^2$ are both nondecreasing functions. Note that H_0^3 also lies in $\mathscr{C}^1_\delta$.

Next, use the algorithm in Lemma 4.3.5 to construct viable height functions $h_0^{i,\epsilon}$ for $3 \le i \le 5$ in such a way that

- $\epsilon^{1/2} h_0^{3,\epsilon}(\epsilon^{-1}x)$ converges in $\mathscr{C}_{\delta'}^{\alpha'}$ to $H^3(x)$ for some $\delta' > \delta$ and $\alpha' < \alpha$.
- $\epsilon^{1/2} h_0^{4,\epsilon}(\epsilon^{-1}x)$ converges in $\mathscr{C}_{\delta'}^{\alpha'}$ to $H^1(x)$ for some $\delta' > \delta$ and $\alpha' < \alpha$.
- $\epsilon^{1/2} h_0^{5,\epsilon}(\epsilon^{-1}x)$ converges in $\mathscr{C}_{\delta'}^{\alpha'}$ to $H^2(x)$ for some $\delta' > \delta$ and $\alpha' < \alpha$.
- $h_0^{3,\epsilon} - h_0^{i,\epsilon}$ are nondecreasing in x for $i = 4, 5$ and all ϵ.

On the same probability space, let $h^{1,\epsilon}$ and $h^{2,\epsilon}$ be the (time-evolving) height profiles started from initial data $h_0^{1,\epsilon}$ and $h_0^{2,\epsilon}$ (respectively) and whose dynamics are governed by the Poisson clocks described above. Then let $h^{3,\epsilon}, h^{4,\epsilon}$ and $h^{5,\epsilon}$ be the height functions associated with initial data $h_0^{3,\epsilon}, h_0^{4,\epsilon}, h_0^{5,\epsilon}$, respectively.

Let $(H^1, H^2, H^3, H^4, H^5)$ be a joint limit point of $(h^{1,\epsilon}, h^{2,\epsilon}, h^{3,\epsilon}, h^{4,\epsilon}, h^{5,\epsilon})$, which is **not** necessarily defined on the same probability space as above. Then let $(\xi^1, \xi^2, \xi^3, \xi^4, \xi^5)$ denote the respective driving noises. By Lemma 4.3.2 we know that $H^4 = H^1$ and $H^5 = H^2$, therefore $\xi^4 = \xi^1$ and $\xi^5 = \xi^2$ (for instance by Theorem 4.2.3). But since $h_0^{3,\epsilon} - h_0^{i,\epsilon}$ are nondecreasing in x for $i \in \{4, 5\}$, Lemma 4.3.3 implies that $\xi^4 = \xi^3$ and $\xi^5 = \xi^3$. So all noises are equal, and in particular $\xi^1 = \xi^2$. $\square$

We are now ready to state and prove the main result for two arbitrary initial data which converge in $\mathscr{C}_\delta^\alpha$. The idea will be to nest the two initial data between smooth initial data satisfying the hypotheses of Proposition 4.3.6, and then take advantage of the monotonicity (M). Recall our notation that the subscript “0" in H_0^i denotes the (deterministic) initial data of the height profile, while H^i without the subscript denotes the entire space-time profile viewed as a random variable in some Skorohod space $D([0, \infty), C(\mathbb{R}))$.

Theorem 4.3.7. *Let $\eta_t^{1,\epsilon}$ and $\eta_t^{2,\epsilon}$ be sequences (indexed by $\epsilon \in (0,1]$) of exclusion processes with generator* (4.4) *on $\{0,1\}^{\mathbb{Z}}$ with $b(-1,1,0) = \frac{1}{2} + \frac{1}{2}\sqrt{\epsilon}$ and $b(1,1,0) = \frac{1}{2} - \frac{1}{2}\sqrt{\epsilon}$. Assume that the dynamics are run via the basic coupling as described above in Subsection 3.1. Let $h^{1,\epsilon}, h^{2,\epsilon}$ denote the rescaled height functions as in* (4.5). *Suppose that the deterministic sequences of initial data $h_0^{1,\epsilon}, h_0^{2,\epsilon}$ converge in $\mathscr{C}_\delta^\alpha$ as $\epsilon \to 0$ to H_0^1, H_0^2 respectively, where $0 < \alpha < 1/2$ and $0 < \delta < 1$. Then one has joint convergence in law as $\epsilon \to 0$ of the entire time-evloving height profile $(h^{1,\epsilon}, h^{2,\epsilon})$ to (H^1, H^2) where H^1, H^2 both solve the KPZ equation with the same noise and with initial data H_0^1, H_0^2, resp. The convergence holds with respect to the topology of $D([0,T], C(\mathbb{R})^2)$.*

Proof. Choose some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which one may define, for each $\epsilon \in (0,1]$, a system of i.i.d. Poisson clocks of rate $p = \frac{1}{2} + \frac{1}{2}\sqrt{\epsilon}$ and rate $q = \frac{1}{2} - \frac{1}{2}\sqrt{\epsilon}$ associated to each bond $\{x, x+1\}$ with $x \in \mathbb{Z}$. For different values of ϵ these can be coupled in an arbitrary manner; ultimately it is irrelevant.

Choose arbitrary sequences of smooth approximating height functions $a_0^N, b_0^N, c_0^N, d_0^N$, for $N \in \mathbb{N}$, satisfying the following five properties:

- $a_0^N, b_0^N, c_0^N, d_0^N \in \mathscr{C}_{\delta'}^1$ for some $\delta' > \delta$.
- $b_0^N \le H_0^1 \le a_0^N$.
- $d_0^N \le H_0^2 \le c_0^N$.
- a_0^N, b_0^N converge in $\mathscr{C}_{\delta'}^{\alpha'}$ to H_0^1 for some $\alpha' < \alpha$ and δ' as above.
- c_0^N, d_0^N converge in $\mathscr{C}_{\delta'}^{\alpha'}$ to H_0^2 with α', δ' as above.

The existence of such sequences is straightforward. Indeed, one can even choose (α', δ') arbitrarily from $(0, \alpha) \times (\delta, 1)$ and then take a_0^N and c_0^N to coincide with $|x|^{(\delta+\delta')/2}$ in some neighborhood of $\pm\infty$ and similarly one can choose b_0^N and d_0^N to coincide with $-|x|^{(\delta+\delta')/2}$ in some neighborhood in $\pm\infty$ (this neighborhood will obviously depend on N though).

Now use Lemma 4.3.5 to define viable height functions $a_0^{N,\epsilon}, b_0^{N,\epsilon}, c_0^{N,\epsilon}, d_0^{N,\epsilon}$ which are jointly admissible and converge under the appropriate scaling to $a_0^N, b_0^N, c_0^N, d_0^N$, respectively. Denote by $(h_t^{1,\epsilon}, h_t^{2,\epsilon}, a_t^{N,\epsilon}, b_t^{N,\epsilon}, c_t^{N,\epsilon}, d_t^{N,\epsilon})_{t\geq 0}$ the (time-evolving) height profiles associated with initial data $(h_0^{1,\epsilon}, h_0^{2,\epsilon}, a_0^{N,\epsilon}, b_0^{N,\epsilon}, c_0^{N,\epsilon}, d_0^{N,\epsilon})$, respectively. The dynamics for each of these objects are run according to the Poisson clocks described above.

Let $(H^1, H^2, a^N, b^N, c^N, d^N)$ denote a joint limit point of all of these objects (as $\epsilon \to 0$), which is not necessarily defined on the same probability space. By Proposition 4.3.6, all of a^N, b^N, c^N, d^N solve the KPZ equation with the same realization of the noise ξ (we are using Remark 4.3.4 here).

Recall by construction, a_0^N and b_0^N converge in $\mathscr{C}_{\delta'}^{\alpha'}$ to H_0^1 as $N \to \infty$, and moreover $b_0^N \leq H_0^1 \leq a_0^N$. Note that for a fixed realization of ξ, the solution of the KPZ equation is continuous as a function of the initial data, viewed as a function from $\mathscr{C}_{\delta'}^{\alpha'}(\mathbb{R}) \to C(\mathbb{R}_+ \times \mathbb{R})$. This can be proved directly from Definition 4.2.1 by exploiting Mueller's positivity result [122], and it can also be proved more directly by using more modern techniques such as [86, 132]. Thus as $N \to \infty$, the function $b^N - a^N$ converges uniformly to 0 on compact sets of $\mathbb{R}_+ \times \mathbb{R}$, and moreover by (M) it is true that $b^N \leq H^1 \leq a^N$ for all N. Thus a^N, b^N both converge uniformly to H^1 on compact subsets of $\mathbb{R}_+ \times \mathbb{R}$, and on the other hand they also converge to the solution of the KPZ equation driven by the common noise of the a^i, b^i and initial data H_0^1. Thus, we conclude that H^1 is driven by the same noise as the a^i, b^i.

A completely analogous argument will show that H^2 is driven by the same noise as the c^i, d^i, completing the proof. $\square$

One can ask why, in the above proof, one could not have defined simpler approximations $H^i * \phi_\delta$ and then just used the fact the the KPZ equation is continuous as a function of the initial data on $\mathscr{C}_{\delta'}^{\alpha'}(\mathbb{R})$, and used Proposition 4.3.6 without relying on (M). The problem with this idea is that it would be circular: we do not know beforehand that the joint limit points all solve the KPZ equation with the same noise: therefore we do not know that they are continuous as a function of the initial data. Hence some kind of monotonicity property must be leveraged.

4.3.4 Random initial conditions near stationarity

In the above theorem we assumed that $h_0^{1,\epsilon}$ and $h_0^{2,\epsilon}$ were deterministic and converged with respect to the topology of $\mathscr{C}_\delta^\alpha(\mathbb{R})$ for some $0 < \alpha < 1/2$ and some $0 < \delta < 1$. In this subsection we relax these conditions slightly to allow for *random* sequences of pairs of initial data that may only converge in distribution and satisfy some p^{th} moment bounds that are generally easy to check in practice. The prototypical examples to keep in mind for this subsection are the height function pairs generated by iid Bernoulli configurations. These two product Bernoulli configurations may be independent or correlated by some parameter; it does not matter so long as the finite-dimensional marginals for the pair of height functions converge *jointly* in law.

We denote by $\|X\|_p := \mathbb{E}[|X|^p]^{1/p}$ for a random variable X defined on some probability space.

Theorem 4.3.8. *The conclusion of Theorem 4.3.7 still holds for random initial data* $(h_0^{1,\epsilon}, h_0^{2,\epsilon})$ *so long as this pair converges jointly in the sense of finite dimensional distributions to* (H_0^1, H_0^2) *and there exist* $\alpha \in (0, 1/2]$,$\delta \in (0, 1)$, $C > 0$, *and* $p > \max\{\alpha^{-1}, (1-\delta)^{-1}\}$ *such that for all* $\epsilon \in (0, 1]$ *the pair satisfies the moment bounds*

$$\|h_0^{i,\epsilon}(x)\|_p \leq C(1 + |x|)^\delta \text{ for all } x \in \mathbb{R}$$

$$\|h_0^{i,\epsilon}(x) - h_0^{i,\epsilon}(y)\|_p \leq C(1 + |x|)^\delta |x - y|^\alpha \text{ whenever } |x - y| \leq 1.$$

The proof is immediately obtained by combining the results of Lemmas 4.3.9 and 4.3.10 given just below. Note that the rescaled height functions associated to iid Bernoulli configurations satisfy these bounds with $\alpha = \delta = 1/2$ (in fact, one does not even need the extra factor of $(1 + |x|)^\delta$ in the second bound).

Lemma 4.3.9. *The conclusion of Theorem 4.3.7 still holds for random initial data* $(h_0^{1,\epsilon}, h_0^{2,\epsilon})$ *so long as this pair converges in law to* (H_0^1, H_0^2) *with respect to the topology of* $\mathscr{C}_\delta^\alpha \times \mathscr{C}_\delta^\alpha$ *for some* $\alpha \in (0, 1/2)$ *and* $\delta \in (0, 1)$.

Proof. Recall $\mathcal{V}^\epsilon$ which is the set of all functions of the form $\epsilon^{1/2} h(\epsilon^{-1}x)$ where h is a viable height function. Fix $T > 0$ and let $\mathcal{M}$ denote the set of all probability measures on $D([0, T], C(\mathbb{R})^2)$. Define $Q^\epsilon : (V^\epsilon \cap \mathscr{C}_\delta^\alpha)^2 \to \mathcal{M}$ by sending a rescaled pair of viable height functions (h_0^1, h_0^2) to the law of the (entire time evolution of the) basically coupled ASEP height process started from (h_0^1, h_0^2), with right jump parameter $\frac{1}{2} + \frac{1}{2}\epsilon^{1/2}$ and left jump parameter $\frac{1}{2} - \frac{1}{2}\epsilon^{1/2}$.

Likewise, define $Q : \mathscr{C}_\delta^\alpha(\mathbb{R})^2 \to \mathcal{M}$ by sending (h_0^1, h_0^2) the solution of the KPZ equation driven by the same realization of ξ started from h_0^1, h_0^2 respectively. Theorem 4.3.7 says precisely that $Q^\epsilon(h_0^{1,\epsilon}, h_0^{2,\epsilon}) \to Q(h_0^1, h_0^2)$ whenever $(h_0^{1,\epsilon}, h_0^{2,\epsilon}) \to (h_0^1, h_0^2)$ in $(\mathscr{C}_\delta^\alpha)^2$.

Now suppose that the hypothesis of the lemma holds, i.e., $(h_0^{1,\epsilon}, h_0^{2,\epsilon})$ converges in law to (H_0^1, H_0^2) with respect to the topology of $\mathscr{C}_\delta^\alpha(\mathbb{R})^2$. By Skorohod's representation theorem[2] we may find a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $(h_0^{1,\epsilon}, h_0^{2,\epsilon}) \to (h_0^1, h_0^2)$ in $(\mathscr{C}_\delta^\alpha)^2$ almost surely. Then by the discussion above, $Q^\epsilon(h_0^{1\epsilon}, h_0^{2,\epsilon}) \to Q(h_0^1, h_0^2)$ in $\mathcal{M}$ almost surely. This is enough to give the required result. Indeed it shows that $\mathbb{E}[f\big(Q^\epsilon(h_0^{1\epsilon}, h_0^{2,\epsilon})\big)] \to \mathbb{E}[f\big(Q(h_0^1, h_0^2)\big)]$ for all bounded continuous $f : \mathcal{M} \to \mathbb{R}$. To finish the proof one simply takes f of the form $f(\nu) := \int_{D([0,T],C(\mathbb{R}))} g(h)\nu(dh)$ where g is a bounded real-valued continuous function on $D([0, T], C(\mathbb{R}))$. Then one may disintegrate the law of $h^{i,\epsilon}$ by decoupling the initial data and the dynamics to obtain the desired result. $\square$

Lemma 4.3.10. *Suppose that* $\{h^\epsilon\}_{\epsilon \in (0,1]}$ *is a family of* $C(\mathbb{R})$*-valued random variables such that there exist* $\alpha \in (0, 1/2)$*,*$\delta \in (0, 1), C > 0$*, and* $p > \max\{\alpha^{-1}, (1-\delta)^{-1}\}$ *which satisfy the*

[2] One technical remark here is that the spaces $\mathscr{C}_\delta^\alpha$ are not separable and thus Skorohod's representation theorem may not hold, strictly speaking. In practice this is not an issue, because for $\delta < \delta'$ and $\alpha' < \alpha$ it is actually true that $\mathscr{C}_\delta^\alpha$ embeds *compactly* into $\mathscr{C}_{\delta'}^{\alpha'}$, as we already mentioned earlier. Any compact metric space is separable, thus in our argument above, one should instead use almost sure convergence with respect to the weaker topology of $\mathscr{C}_{\delta'}^{\alpha'}$ for some $\alpha' < \alpha$ and $\delta' > \delta$. This does not cause any issues for the proof.

following moment bounds uniformly over all $\epsilon \in (0,1]$:

$$\|h^\epsilon(x)\|_p \le C(1+|x|)^\delta \text{ for all } x \in \mathbb{R}$$

$$\|h^\epsilon(x) - h^\epsilon(y)\|_p \le C(1+|x|)^\delta |x-y|^\alpha \text{ whenever } |x-y| \le 1.$$

Then there exist $\alpha' \in (0,\alpha)$ and $\delta' \in (\delta, 1)$ such that $\{h^\epsilon\}_{\epsilon\in(0,1]}$ is tight with respect to the topology of $\mathscr{C}^{\alpha'}_{\delta'}$.

Proof. Recall from earlier that $\mathscr{C}^\alpha_\delta$ embeds *compactly* into $\mathscr{C}^{\alpha'}_{\delta'}$ whenever $\delta' > \delta$ and $\alpha' < \alpha$. Therefore to prove the lemma, it suffices to show that if the two inequalities in the lemma statement hold, then there exist α', δ' such that

$$\lim_{a\to\infty} \sup_{\epsilon\in(0,1]} \mathbb{P}(\|h^\epsilon\|_{\mathscr{C}^{\alpha'}_{\delta'}} > a) = 0.$$

We actually show something stronger, namely that under the given assumptions, there exists $C' > 0$ such that for all $a > 0$

$$\sup_{\epsilon\in(0,1]} \mathbb{P}(\|h^\epsilon\|_{\mathscr{C}^{\alpha'}_{\delta'}} > a) \le C' a^{-p}, \tag{4.8}$$

where p is the same exponent given in the lemma statement. To prove this we write $\|h^\epsilon\|_{\mathscr{C}^{\alpha'}_{\delta'}} = \|h^\epsilon\|_{\delta'} + [h^\epsilon]_{\alpha',\delta'}$ where

$$\|h\|_{\delta'} := \sup_{x\in\mathbb{R}} \frac{|h(x)|}{(1+|x|)^{\delta'}}, \quad \text{and} \quad [h]_{\alpha',\delta'} := \sup_{x\in\mathbb{R}} (1+|x|)^{-\delta'} \sup_{|y-x|\le 1} \frac{|h(x)-h(y)|}{|x-y|^{\alpha'}}.$$

To prove (4.8), the following fact will be useful to us: For any $\gamma \in (0,1)$, the γ-Hölder seminorm $[f]_\gamma$ of a function $f : [0,1] \to \mathbb{R}$ is equivalent (as a seminorm) to the quantity given by $\sup_{n\in\mathbb{N}, 1\le k\le 2^n} 2^{\gamma n} |f(k2^{-n}) - f((k-1)2^{-n})|$. This is proved as an intermediate step in the standard proof of the classical Kolmogorov-Chentsov criterion.

The exact choices of α', δ' will be specified later, but for now let them denote generic constants.

Now to prove (4.8) let us write for a function h,

$$\begin{aligned}
\|h\|_{\delta'} &\leq \sup_{n\in\mathbb{Z}}(1+|n|)^{-\delta'}\big(|h(n)| + \sup_{x\in[n,n+1]}|h(x)-h(n)|\big) \\
&\leq \sup_{n\in\mathbb{Z}}(1+|n|)^{-\delta'}\bigg(|h(n)| + \sup_{x\in[n,n+1]}\frac{|h(x)-h(n)|}{|x-n|^{\alpha'}}\bigg) \\
&\lesssim \sup_{n\in\mathbb{Z}}(1+|n|)^{-\delta'}\bigg(|h(n)| + \sup_{r\in\mathbb{N},1\leq k\leq 2^r} 2^{\alpha' r}|h(n+k2^{-r})-h(n+(k-1)2^{-r})|\bigg),
\end{aligned}$$

where $\lesssim$ denotes the absorption of some universal constant which can depend on α', δ' but not on the function h. Likewise let us note that

$$[h]_{\alpha',\delta'} \lesssim \sup_{n\in\mathbb{Z}}(1+|n|)^{-\delta} \sup_{r\in\mathbb{N},1\leq k\leq 2^r} 2^{\alpha' r}|h(n+k2^{-r})-h(n+(k-1)2^{-r})|.$$

Consequently we find that

$$\|h\|_{\mathscr{C}^{\alpha'}_{\delta'}} \lesssim A(h,\delta') + B(h,\alpha',\delta'),$$

where

$$\begin{aligned}
A(h,\delta') &:= \sup_{n\in\mathbb{Z}}(1+|n|)^{-\delta'}|h(n)|, \\
B(h,\alpha',\delta') &:= \sup_{n\in\mathbb{Z}}(1+|n|)^{-\delta'} \sup_{r\in\mathbb{N},1\leq k\leq 2^r} 2^{\alpha' r}|h(n+k2^{-r})-h(n+(k-1)2^{-r})|.
\end{aligned}$$

Now, with h^ϵ as given in the lemma statement, let us bound these terms $A(h^\epsilon,\delta')$ and $B(h^\epsilon,\alpha',\delta')$ individually to obtain (4.8). We will do this by using the hypotheses in the lemma. Note that by a brutal union bound and Markov's inequality followed by the hypothesis $\|h^\epsilon(x)\|_p \leq (1+|x|)^\delta$, we have

$$
\begin{aligned}
\mathbb{P}(A(h^\epsilon,\delta') > a) &\leq \sum_{n\in\mathbb{Z}} \mathbb{P}(|h^\epsilon(n)| > (1+|n|)^{\delta'} a) \\
&\leq \sum_{n\in\mathbb{Z}} a^{-p}(1+|n|)^{-\delta' p} E[|h^\epsilon(n)|^p] \\
&\leq a^{-p}\sum_{n\in\mathbb{Z}} (1+|n|)^{(\delta-\delta')p},
\end{aligned}
$$

The series converges as long as δ' is chosen so that $(\delta-\delta')p < -1$, for instance $\delta' := \frac{1}{2}(1+\delta+\frac{1}{p})$ which is less than 1 by the hypothesis that $p > (1-\delta)^{-1}$. Next we control B, which will also just use a brutal union bound and Markov's inequality:

$$
\begin{aligned}
\mathbb{P}(B(h^\epsilon,\alpha',\delta') > a) &\leq \sum_{\substack{n\in\mathbb{Z}\\ r\in\mathbb{N}\\ 1\leq k\leq 2^r}} \mathbb{P}(2^{\alpha' r}|h^\epsilon(n+k2^{-r}) - h^\epsilon(n+(k-1)2^{-r})| > (1+|n|)^{\delta'} a) \\
&\leq \sum_{\substack{n\in\mathbb{Z}\\ r\in\mathbb{N}\\ 1\leq k\leq 2^r}} a^{-p} 2^{\alpha' p r}(1+|n|)^{-\delta' p}\mathbb{E}\big|h^\epsilon(n+k2^{-r}) - h^\epsilon(n+(k-1)2^{-r})\big|^p \\
&\leq a^{-p} \sum_{\substack{n\in\mathbb{Z}\\ r\in\mathbb{N}\\ 1\leq k\leq 2^r}} 2^{(\alpha'-\alpha)pr}(1+|n|)^{(\delta-\delta')p} \\
&= a^{-p}\sum_{\substack{n\in\mathbb{Z}\\ r\in\mathbb{N}}} 2^{\big[1+(\alpha'-\alpha)p\big] r}(1+|n|)^{(\delta-\delta')p}
\end{aligned}
$$

The series converges so long as $(\delta-\delta')p < -1$ and $1+(\alpha'-\alpha)p < 0$. We already chose δ' earlier so as to satisfy the condition $(\delta-\delta')p < -1$. Now α' can be chosen for instance $\frac{1}{2}(\alpha-\frac{1}{p})$ which is positive since $p > \alpha^{-1}$. □

4.3.5 More general models and further problems

One may ask the question of how robust the above method of proof is. The answer is that it is generalizable to more complex systems than ASEP, but it is not all-encompassing. More precisely, the method is applicable to any particle system where

- both (M) and (A) hold.

- one has a discrete martingale equation as in (4.6).

- the discrete martingales from (4.6) satisfy (4.7).

Then one can essentially copy and paste the proof above (with minor modifications) to prove joint convergence in those systems as well.

For instance, by taking $\lambda_i = 0$ in Theorem 4.2.3, our method will also work to show joint convergence of the nearest-neighbor *symmetric* simple exclusion process to the Edwards-Wilkinson fixed point. Actually this is even simpler, as one need not perform a nonlinear transform to obtain a discrete SPDE as in (4.6). The height function itself will satisfy an equation similar to (4.6) with the martingales satisfying (4.7). The proofs of all other propositions and lemmas work in precisely the same way as done above.

Less trivial examples of systems satisfying all three of the points above are higher-spin misanthrope processes. One concrete example of such a particle system is the ASEP(q, J) model from [47]. This comes from the generator (4.4) on $\{0, ..., J\}^{\mathbb{Z}}$ by taking

$$b(1, a, b) := \frac{1}{2[J]_q} q^{a-b-(J+1)}[a]_q[J-b]_q, \qquad b(-1, a, b) := \frac{1}{2[J]_q} q^{a-b-(J+1)}[J-a]_q[b]_q,$$

where $[a]_q := \frac{q^q - q^{-a}}{q - q^{-1}}$ for $q \in (0, 1)$. ASEP(q, J) satisfies (A) as well as (M) thanks to the nearest-neighbor interaction. The main result of [47] then proves convergence of the associated (diffusively scaled and renormalized) height function to the KPZ equation by scaling the model parameter as $q = e^{-\epsilon}$. Note that this recovers the results of [16] by setting $J = 1$. Proposition 2.1 in [47] says precisely that (4.6) and (4.7) are satisfied with Z_t defined in expression (1.8) there and a_ϵ, b_ϵ defined accordingly in (4.5). We then have the following result:

Theorem 4.3.11. *Theorems 4.3.7 and 4.3.8 still hold if we replace ASEP by ASEP(q, J), scaling q*

as $e^{-\epsilon}$ in the model parameters above.

Proof. The proof of Lemma 4.3.2 holds essentially verbatim as given. For the proof of Lemma 4.3.3, we need to replace (4.6) with the appropriate modification and then verify that (4.7) still holds. See equation (1.8) of [47] for the appropriate modification of the discrete equation (4.6), and see Proposition 2.1 of [47] for the proof that (4.7) still holds. The proof of Lemma 4.3.5 still holds verbatim, since height functions which are viable for ASEP are still viable for ASEP(q, J) (after perhaps multiplying by 2 in the case that J is even). In the proofs of Proposition 4.3.6 and Theorem 4.3.7, the argument requires a slight modification: on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ one should just take the Poisson clocks to be of rate one, and to account for the jump rate differences one should instead add i.i.d. uniform variables to each bond, which are independent of the Poisson clocks. The reason for this is discussed in Subsection 3.1: if $J > 1$ then the construction of the basic coupling is slightly more complicated than for single-spin systems. The proof of Theorem 4.3.8 is unchanged. $\square$

Examples of interesting systems that do **not** satisfy property (M) are the non-simple exclusion processes studied for instance in [57, 153]. These processes have a generator similar to (4.4), the only difference is that non-neighboring sites may interact with one another, so b can be a function from $\mathbb{Z} \times \{0, ..., J\}^2 \to [0, 1]$ and the sum in (4.4) would be over all pairs $(x, y) \in \mathbb{Z}^2$. Particles may jump over other particles in these systems, which locally allows height functions to overtake one another. These systems still satisfy (A), and thus our proof still works as long as the two sequences of near-stationary initial data are coupled so that one always dominates the other (Lemma 4.3.3), however for arbitrary sequences one probably needs to use a different method without appealing to a black box like Theorem 4.2.3. For instance one can hope to directly study the quadratic variations appearing in (4.6).

Then there are also open boundary systems such as those considered in [45]. These do seem to satisfy (M) but the missing part of the argument is the boundary analogue of Theorem 4.2.3. Furthermore, there are also discrete-time vertex models and their degenerations, such as those studied in [48, 71, 41, 113], for which KPZ fluctuations are known. We do not know if these systems fall within the scope of our work, since the rules of their evolution are more complex and do not exactly fit the framework of the misanthrope-type exclusion processes we have described in Subsection 3.1. In particular it is unclear what exactly the basic coupling even means for these models. Some of the aforementioned systems may be explored in future work.

Here is another direction in which one can hope to generalize Theorem 4.3.7. Rather than making the model more complicated, one can instead hope to strengthen the topology in which convergence occurs. Specifically one can hope to prove *uniform* convergence of the entire stochastic flow of ASEP to that of the KPZ equation, on compact subsets of the space of continuous functions. More precisely, fix a compact set $K \subset \mathscr{C}_\delta^\alpha(\mathbb{R})$ and let $K_\epsilon \subset \mathcal{V}^\epsilon \cap \mathscr{C}_\delta^\alpha$ be a sequence of compact sets that converge to K in the sense of Hausdorff distance, as $\epsilon \to 0$ (where $\mathcal{V}^\epsilon$ was defined just before Lemma 4.3.5). Consider the random maps $\Phi_t^\epsilon : K_\epsilon \to C(\mathbb{R})$ which (for a fixed realization of the Poisson clocks) sends a rescaled initial height function h to the height profile at time t of the ASEP profile started from h and whose dynamics are run according to those Poisson clocks. Consider also the continuum version $\Phi : K \to C(\mathbb{R})$ which (for a fixed realization of ξ) sends a function h to the time t solution of the KPZ equation started from h and driven by ξ. Let $G(\Phi_t^\epsilon)$ denote the set of all $(h, \Phi_t^\epsilon(h))$ such that $h \in K_\epsilon$, and likewise let $G(\Phi_t)$ denote the set of all $(h, \Phi_t(h))$ such that $h \in K$. Also let d_H denote the Hausdorff distance on compact subsets of $C(\mathbb{R}) \times C(\mathbb{R})$ (one could also hope to use the stronger topology of $\mathscr{C}_\delta^\alpha(\mathbb{R}) \times \mathscr{C}_\delta^\alpha(\mathbb{R})$). Then one can hope to prove convergence of the entire flow $(\Phi_t^\epsilon)_{t\in[0,T]}$ to $(\Phi_t)_{t\in[0,T]}$ where the convergence is meant to be interpreted, for instance, in the sense that (via Skorohod's representation theorem) there exists a coupling of all Φ_t^ϵ, Φ_t onto some probability space and some $\alpha \in (0, 1/2)$ such that

$$\limsup_{\delta\to 0}\limsup_{\epsilon\to 0}\left[\sup_{t\in[0,T]} d_H(G(\Phi_t^\epsilon), G(\Phi_t)) + \sup_{|s-t|\le\delta} \frac{d_H(G(\Phi_t^\epsilon), G(\Phi_s^\epsilon))}{|t-s|^\alpha \vee \epsilon^\alpha}\right] = 0,$$

where the extra factor ϵ^α in the denominator is to account for jumps. Theorem 4.3.7 and Remark

4.3.4 show (in some sense) that convergence of these flows holds in the sense of finite-dimensional distributions, but extending the convergence to this uniform Hölder sense might be more interesting. The goal would be to prove this for arbitrary compact sets K and arbitrary approximating sequences K_ϵ. We do not have strong enough spatial or temporal estimates required to do this, except for the trivial case where K, K_ϵ are all finite sets with cardinality bounded in ϵ, in which case the methods of [16] combined with our methods used to prove Theorem 4.3.7 are enough.

Chapter 5: A dynamical systems perspective on Strassen's Law

This chapter is based on joint work with Yier Lin. A result of Arcones [6] implies that if a measure-preserving linear operator S on an abstract Wiener space (X, H, μ) is mixing, then the set of limit points of the random sequence $((2\log n)^{-1/2} S^n x)_{n\in\mathbb{N}}$ equals the unit ball of H for a.e. $x \in X$. We extend this result to the case of a continuous parameter n and higher Gaussian chaoses, and we also prove a contraction-type principle for Strassen laws. We then use these extensions to recover or prove Strassen-type laws for a broad collection of processes derived from a Gaussian measure, culminating with a general machinery to derive "nonlinear" Strassen laws for singular SPDEs such as Φ^4_2 and KPZ.

5.1 Introduction

The law of the iterated logarithm for Brownian motion states that if B is a standard Brownian motion then $\limsup_{t\to 0}(2t\log\log(1/t))^{-1/2}B_t = 1$. Strassen in a seminal work [141] generalized this statement to show the functional form of this statement, namely that if we let $B^\epsilon(t) = \epsilon^{-1/2}B(\epsilon t)$ then the set of limit points as $\epsilon \to 0$ in $C[0,1]$ of the sequence $\{(2\log\log(1/\epsilon))^{-1/2}B^\epsilon\}_\epsilon$ is almost surely equal to the unit ball of its Cameron-Martin space. Since Strassen's original work there has been a tremendous effort resulting in a large literature expanding the scope of the theorem into many different settings including invariance principles and Banach space-valued processes [104, 103, 76, 4], Gaussian processes and higher chaoses [127, 77, 75, 6], iterated processes [28, 6, 50, 123], stronger topologies [13], sharper envelopes [134], as well as more complicated stochastic processes driven by multiparameter fields and fractional processes [131, 55, 114, 128, 67], etc. For surveys of classical topics and results on the law of the iterated logarithm, see [109, Chapter 8] or [102, 18].

The goal of the present paper is to extend Strassen's law in yet another general direction, related to many of the aforementioned extensions, eventually culminating with a compact limit set theorem for the small-noise regime of subcritical singular SPDEs such as KPZ and Φ^4_d, which have been of popular interest in the probability literature recently, see e.g. [40, 33, 46], etc. One difference from the aforementioned results is that we take a dynamical-systems and semigroup-theory perspective on Strassen's law, rather than considering an arbitrary sequence of random variables in a Banach space. We are uncertain if this perspective is novel, or if it merely provides a convenient notational framework to summarize proof methods that may already be well-known to experts in the area. Nonetheless the dynamics perspective does allow us to concisely recover and generalize some of the more classical results cited above, in particular our main result Theorem 5.1.3 below will recover some of the main results from [104, 131, 13, 75, 28, 75, 50, 6, 123, 55], see Section 5.4.

First we establish some notation. If H is any real Hilbert space and $S : H \to H$ is any bounded operator, we denote by S^* its adjoint operator with respect to the inner product of H. Throughout this work we will use the notion of abstract Wiener spaces introduced by Leonard Gross [78]. These are formal triples (X, H, μ) where X is a Banach space, μ is a centered Gaussian measure on X, and $H \hookrightarrow X$ is the embedded Cameron-Martin space (see e.g. [85]).

Let (X, H, μ) be an abstract Wiener space. Given any bounded linear map $S : H \to H$ satisfying $SS^* = I$, there exists a μ-a.e.-defined Borel-measurable linear extension $\hat{S}$ on X which is unique up to a.e. equivalence (see Proposition 3.46 and Theorem 3.47 of [85]). Moreover the condition $SS^* = I$ guarantees that this extension is measure-preserving. In the sequel we will not distinguish between $\hat{S}$ and S and simply write $\hat{S} = S$.

The following result is most likely attributable to Arcones [6, Theorem 2.1], and the main focus of the present paper will be to extend it to continuous-parameter settings and higher Gaussian chaoses, ultimately culminating in a nonlinear version of Strassen's Law for some singular SPDEs.

Proposition 5.1.1. *Let (X, H, μ) be an abstract Wiener space. Let $S : H \to H$ be a bounded* *operator and write* $S_N := S^N$*. Suppose* S *satisfies the following two properties:*

1. $SS^* = I$.

2. $\langle S_N x, y\rangle_H \to 0$ *as* $N \to \infty$ *for all* $x, y \in H$.

Then for μ*-almost every* $x \in X$*, the set of limit points of* $\{\frac{S_N x}{\sqrt{2\log N}} : N \in \mathbb{N}\}$ *is equal to the unit ball of* H.

Just to be clear, we are talking about limits points with respect to the topology of X and $S_N x$ is meant to be understood in terms of the unique measurable linear extensions mentioned above if $x \in X$. We will see later (Lemma 5.2.3) that condition *(2)* can be stated as $\bigcap_{n\in\mathbb{N}} \mathrm{Im}(S_N^*) = \{0\}$, where $\mathrm{Im}(S)$ denotes the image of S. This is easier to check in certain instances.

Now we will formulate a continuous-time version of Proposition 5.1.1. Recall that a strongly continuous semigroup on a Banach space X is a semigroup $(S_t)_{t\geq 0}$ of bounded operators from $X \to X$ such that $\|S_t x - x\|_X \to 0$ as $t \to 0$ for all $x \in X$. Moreover, if X is a Banach space and if $\gamma : [0,\infty) \to X$ is any function, then we call $x \in X$ a *cluster point at infinity* of γ if there exists a sequence $t_n \uparrow \infty$ such that $\|\gamma(t_n) - x\|_X \to 0$. Similarly if $\gamma : (0,1] \to X$ then we will call $x \in X$ a *cluster point at zero* of γ if there exists a sequence $t_n \downarrow 0$ such that $\|\gamma(t_n) - x\|_X \to 0$.

Proposition 5.1.2. *Let* (X, H, μ) *be an abstract Wiener space. Suppose that* $(S_t)_{t\geq 0}$ *is a family of bounded operators from* $H \to H$ *satisfying the following four properties:*

1. $S_{t+u} = S_t S_u$ *for all* $t, u \geq 0$.

2. $S_t S_t^* = I$.

3. $\langle S_t x, y\rangle_H \to 0$ *as* $t \to \infty$ *for all* $x, y \in H$.

4. $(S_t)_{t\geq 0}$ *extends to a strongly continuous semigroup on the larger space* X.

Then for μ*-almost every* $x \in X$*, the set of cluster points at infinity of* $\{\frac{S_t x}{\sqrt{2\log t}} : t \geq e\}$ *is equal to the unit ball of* H.

Now let us relate this to the usual law of the iterated logarithm. Recall that a strongly continuous *multiplicative* semigroup is a family of bounded operators $(R_\epsilon)_{\epsilon\in(0,1]}$ satisfying $R_\epsilon R_\delta = R_{\epsilon\delta}$ and moreover $\|R_\epsilon x - x\|_X \to 0$ as $\epsilon \uparrow 1$ for all $x \in X$. Then by setting $S_t := R_{e^{-t}}$, Proposition 5.1.2 can be restated in terms of multiplicative semigroups as follows.

Let (X, H, μ) be an abstract Wiener space. Suppose that $(R_\epsilon)_{\epsilon\in(0,1]}$ is a family of bounded operators from $H \to H$ satisfying the following four properties:

1. $R_\epsilon R_\delta = R_{\epsilon\delta}$ for all $\epsilon, \delta \in (0, 1]$.
2. $R_\epsilon R_\epsilon^* = I$.
3. $\langle R_\epsilon x, y\rangle_H \to 0$ as $\epsilon \to 0$ for all $x, y \in H$.
4. $(R_\epsilon)_{\epsilon\in(0,1]}$ extends to a strongly continuous semigroup on the larger space X.

Then for μ-almost every $x \in X$, the set of cluster points at zero of $\{\frac{R_\epsilon x}{\sqrt{2\log\log(1/\epsilon)}} : \epsilon \in (0, e^{-e}]\}$ is equal to the unit ball of H.

The quintessential example of such a setup is the classical Wiener space $X = C[0, 1], H = H_0^1[0, 1]$, and μ is the law of a standard Brownian motion. Then one sets $R_\epsilon f(x) = \epsilon^{-1} f(\epsilon^2 x)$ and the above statement recovers the classical version of Strassen's law. We will give a self-contained proof of Propositions 5.1.1 and 5.1.2 in Subsection 5.2.2 below, as well as a partial converse which says that Strassen's law necessarily implies ergodicity of S.

Next let us discuss our main result of the paper, the generalization to higher chaos and the contraction principle. If (X, H, μ) is an abstract Wiener space, then the n^{th} homogeneous Wiener chaos, denoted by $\mathcal{H}^k(X, \mu)$ is defined to be the closure in $L^2(X, \mu)$ of the linear span of $H_k \circ g$ as g varies through all elements of the continuous dual space X^*, where H_k denotes the k^{th} Hermite polynomial (normalized so that $H_k(Z)$ has unit variance for a standard normal Z). Letting Y be another separable Banach space, a Borel-measurable map $T : X \to Y$ is called a chaos of order k over X if $f \circ T \in \mathcal{H}^k(X, \mu)$ for all $f \in Y^*$. For such a chaos we may define (following [108, 91]) its "homogeneous form" $T_{hom} : H \to Y_i$ by the Bochner integral formula

$$T_{hom}(h) := \int_X T(x+h)\mu(dx) = \frac{1}{k!}\int_X T(x)\langle x, h\rangle^k \mu(dx),$$

which may be shown to converge as shown in the appendix. Our main result in this work is the following, which will be restated as Corollary 5.2.21 below.

Theorem 5.1.3. *Let (X, H, μ) be an abstract Wiener space, let $(R_\epsilon)_{\epsilon\in(0,1]}$ be a family of Borel-measurable a.e. linear maps from $X \to X$ which are measure-preserving and satisfy $\langle R_\epsilon x, y\rangle_H \to 0$ as $\epsilon \to 0$ for all $x, y \in H$. Let $T^i : X \to Y_i$ be a chaos of degree k_i over X for $1 \le i \le m$. Suppose that there exist strongly continuous semigroups $(Q^i_\epsilon)_{\epsilon\in(0,1]}$ of operators from $Y_i \to Y_i$ for $1 \le i \le m$ with the property that*

$$T^i \circ R_\epsilon = Q^i_\epsilon \circ T^i, \qquad \mu\text{-a.e.} \qquad \text{for all } \epsilon \in (0,1] \text{ and } 1 \le i \le m. \tag{5.1}$$

Let Z be a Banach space, and let $\mathcal{M} \subset Y_1 \times \cdots Y_m$ be a closed subset, such that for all $\delta > 0$

$$\mu(\{x \in X : (\delta^{k_1}T^1(x), ..., \delta^{k_m}T^m(x)) \in \mathcal{M}\}) = 1.$$

Let $\Phi : \mathcal{M} \to Z$ be continuous (possibly nonlinear). Then the compact set

$$K := \{(T^1_{hom}(h), ..., T^m_{hom}(h)) : h \in B(H)\}$$

is necessarily contained in $\mathcal{M}$, and moreover the set of cluster points at zero of the random set

$$\big\{\Phi\big((2\log\log(1/\epsilon))^{-k_1/2}T^1(R_\epsilon x),\ \dots,(2\log\log(1/\epsilon))^{-k_m/2}T^m(R_\epsilon x)\big) : \epsilon \in \mathbb{Q} \cap (0,1]\big\}$$

is almost surely equal to $\Phi(K)$.

Note that one recovers the previous result by setting $m = 1, k_1 = 1, Y_1 = X = Z = \mathcal{M}$, and making Φ the identity map. It may seem unclear what the application of such a generalization is, but a number of examples recovering previously known results are given in Section 5.4 below. One particular example of interest (which sparked our interest in this problem) is the following.

Theorem 5.1.4. *For $\epsilon \in (0, 1/3]$ let $C_\epsilon := (\log\log(1/\epsilon))^{1/2}$, and let h^ϵ denote the Hopf-Cole solution to the KPZ equation*

$$\partial_t h^\epsilon = \partial_x^2 h^\epsilon + C_\epsilon(\partial_x h^\epsilon)^2 + \xi,$$

with initial data $h(0, x) = 0$. Then for any $s, y \geq 0$ the set of limit points as $\epsilon \downarrow 0$ in $\mathcal{C}_{s,y} := C([0, s] \times [-y, y])$ of the sequence of functions $C_\epsilon^{-1}\epsilon^{-1/2}h^\epsilon(\epsilon^2 t, \epsilon x)$ is a.s. equal to the compact set K_{Zero} given by the closure in $\mathcal{C}_{s,y}$ of the set of smooth functions h satisfying

$$h(0, x) = 0, \qquad \left\|\partial_t h - \partial_x^2 h - (\partial_x h)^2\right\|_{L^2([0,s]\times[-y,y])} \leq 1.$$

If we instead let $h^\epsilon(0, x)$ be a two sided Brownian motion (fixed for different values of ϵ) then the same result holds but with compact limit set K_{Br} given by the closure of smooth functions $h \in \mathcal{C}_{s,y}$ satisfying

$$h(0, 0) = 0, \qquad \|\partial_x h(0, \cdot)\|_{L^2[-y,y]} \leq 1, \qquad \left\|\partial_t h - \partial_x^2 h - (\partial_x h)^2\right\|_{L^2([0,s]\times[-y,y])} \leq 1.$$

If we likewise define k^ϵ to be the solution of

$$\partial_t k^\epsilon = \partial_x^2 k^\epsilon + C_\epsilon^{-1}(\partial_x k^\epsilon)^2 + \xi,$$

then the same compact limit set results hold for $C_\epsilon^{-1}\epsilon^{1/2}k^\epsilon(\epsilon^{-2}t, \epsilon^{-1}x)$. Moreover, the same results hold in stronger topologies given by parabolic Holder seminorms up to but excluding exponent 1/2 (see Definition 5.4.7).

In the theorem statement, we are interpreting the derivatives in a weak sense, so the L^2 norms are defined (by convention) to be infinite if the weak derivatives $\partial_t h, \partial_x h, \partial_x^2 h$ do not exist.

Note that in the first result stated in the theorem (for the family h^ϵ), the nonlinearity must be scaled along with the parameter ϵ, so that it blows up in the $\epsilon \to 0$ limit. If we did not do this then the limiting compact set would simply agree with that of the linearized equation $\partial_t h_{Linear} = \partial_x^2 h_{Linear} + \xi$, namely

$$K_{Linear} := \{h \in \mathcal{C}_{s,y} : h(0,x) = 0, \left\|\partial_t h - \partial_x^2 h\right\|_{L^2([0,s]\times[-y,y])} \leq 1\}.$$

Indeed this can be proved by decomposing $h_{KPZ} = h_{Linear} + v$ where h_{KPZ} is the Hopf-Cole solution to KPZ with initial data zero and v is a remainder term which has better regularity than h_{Linear} (see e.g. Theorem 3.19 of [132]). Then under the scaling necessary to obtain Strassen's law, it is easy to check that the remainder term converges a.s. to zero in the topology of $C([0,s] \times [-y,y])$ and the set of limit points for the part corresponding to h_{Linear} can be shown to be K_{Linear} by applying Proposition 5.1.2 above (see Example 5.4.6 below).

Likewise in the second result stated above (for the family k^ϵ), the nonlinearity must be scaled along with the parameter ϵ so that it vanishes in the $\epsilon \to 0$ limit. If we did not do this then the asymptotics would be wrong entirely and instead one would need to apply a scaling that respects the tail behavior of the so-called KPZ fixed point [117], namely $(\log\log(1/\epsilon))^{2/3}$, see [53].

The fact that the nonlinearity of the KPZ equation must be scaled along with the parabolic scaling of space-time to obtain a nontrivial limit set in the Strassen law can be seen as a manifestation of the so-called *weak KPZ universality,* which roughly states that the KPZ equation is only scale-invariant up to a one-parameter family of equations which interpolates between two *bona fide* fixed points [40]. The manner in which we prove the above theorem uses the theory of regularity structures [86] and is robust enough to prove similar theorems for other rough equations such as Φ^4_2 and Φ^4_3 (modulo the difficulty of developing the solution theory for the latter equation on the full space

$\mathbb{R}^3$ using regularity structures). We also prove a similar result for mollified smooth versions of the noise, specifically we show that even though mollifying is not a measure-preserving operation on the noise, it is enough to be “asymptotically rapidly measure-preserving," see Subsection 5.4.3.

In principle other theories such as paracontrolled products [79] should also be able to prove such SPDE results, but the main problem is the nonlocality of the Bony paraproduct decomposition makes it difficult to check the scale invariance condition (i.e. the commutation relation) in Theorem 5.1.3, which is inherently a local phenomenon.

5.2 Proofs of main theorems

5.2.1 Ergodicity properties of measure preserving linear operators

In this subsection we review the fact that the ergodic properties of measure-preserving linear operators on an abstract Wiener space are necessarily determined by their action on the Cameron-Martin space.

Definition 5.2.1. *Let (X, H, μ) be an abstract Wiener space, and let $E \subset X$ be a Borel-measurable linear subspace with $\mu(E) = 1$. We say that $S : E \to X$ is an a.e. defined measure-preserving linear map if $\mu(S^{-1}(F)) = \mu(F)$ for all Borel subsets $F \subset X$.*

Note that any a.e. defined measure preserving linear map may be measurably extended to all of X by defining it to be zero outside of E (though this extension is no longer linear), and in this case one still has $\mu(S^{-1}(F)) = \mu(F)$ for all Borel sets F.

Lemma 5.2.2. *Consider the set of a.e.-defined measure-preserving linear maps $S : E \to X$ where E is some Borel linear subspace of measure 1. The set of all such linear transformations (modulo a.e. equivalence) is in bijection with the set of bounded linear maps from $H \to H$ satisfying $SS^* = I$. Moreover the bijection is given by simply restricting S to H.*

Proof. We need to show that any a.e. defined measure-preserving linear transformation from $X \to X$ necessarily maps $H \to H$ boundedly and satisfies $SS^* = I$ on H. To prove this, let $S : E \to$

X be a measure preserving Borel-measurable linear map, where $E \subset X$ is a Borel measurable linear subspace satisfying $\mu(E) = 1$. It is known (see Proposition 3.42 in [85]) that H equals the intersection of all Borel-measurable linear subspaces of X of measure 1. If F is any Borel-measurable linear subspace of measure 1, then so is $E \cap S^{-1}(F)$, and thus $x \in H$ implies $x \in S^{-1}(F)$ so that $Sx \in F$. Since F is arbitrary, we have shown that $Sx \in H$ for all $x \in H$. Thus S is a *globally defined* Borel measurable linear map from $H \to H$, from which it follows that S is automatically bounded (see remark 3.38 in [85]). In order for S to be measure-reserving, it must clearly satisfy $SS^* = I$ (e.g. by computing the covariance structure of $S_*\mu$).

Conversely, given any bounded linear map $S : H \to H$ satisfying $SS^* = I$, there exists a μ-a.e.-defined Borel-measurable linear extension $\hat{S}$ on X which is unique up to a.e. equivalence (this follows from e.g. Proposition 3.46 and Theorem 3.47 of [85]). Moreover the condition $SS^* = I$ guarantees that this extension is measure-preserving (in the sequel we simply write $\hat{S} = S$ without specifying that it actually denotes the unique extension). □

Next we have a lemma about the structure of such measure-preserving transformations, namely that they can be orthogonally decomposed into a unitary part and a part converging strongly to zero.

Lemma 5.2.3. *Let H be a Hilbert space and consider any linear operator $S : H \to H$ satisfying $SS^* = I$. Then we can orthogonally decompose $H = A \oplus B$ where A and B are invariant under S and S^*. Moreover $S : A \to A$ is unitary and $\|S_n x\|_H \to 0$ for all $x \in B$. Explicitly one can write $A = \bigcap_n$ Im(S_n^*) and $B = \overline{\bigcup_{n \in \mathbb{N}} \ker(S_n)}$, which is the closure of $\bigcup_{n \in \mathbb{N}} \ker(S_n)$ in H.*

Proof. Since $SS^* = I$, we have $S_n S_n^* = I$. It is clear that S and S^* both leave $\bigcap_{n \in \mathbb{N}} \mathrm{Im}(S_n^*)$ invariant and S^*S is the identity there. For each n, one has $H = \mathrm{Im}(S_n^*) \oplus \ker(S_n)$ via the decomposition $x = S_n^* S_n x + (x - S_n^* S_n x)$. Hence, $S_n^* S_n$ is merely the projection map from H onto $\mathrm{Im}(S_n^*)$ and indeed the two given subspaces A, B are orthogonal. Moreover it is clear that $\|S_n x\| \to 0$ on the closure of $\bigcup_{n \in \mathbb{N}} \ker(S_n)$, since $\|S_n x\|$ is eventually zero for all x in the dense subspace $\bigcup \ker(S_n)$. □

Given Lemma 5.2.2, it is natural then to ask what conditions on S, when viewed as a map from $H \to H$, ensure that the measure-preserving extension S is ergodic, weakly mixing and strongly mixing. Our first proposition addresses this.

Proposition 5.2.4. *Let (X, H, μ) be an abstract Wiener space. Consider any linear operator $S : H \to H$ satisfying $SS^* = I$. Then we have the following equivalences.*

1. *$\bigcap_{n\geq 1} \sigma(S_n)$ is a 0-1 sigma algebra if and only if $\|S_n x\|_H \to 0$ as $n \to \infty$ for all $x \in H$.*

2. *S is mixing if and only if $\langle S_n x, y\rangle_H \to 0$ as $n \to \infty$ for all $x, y \in H$.*

3. *S is ergodic if and only if any of the following equivalent conditions hold:*

   (a) $\frac{1}{n}\sum_{j=1}^{n} \langle S_j x, y\rangle_H^k \to 0$ as $n \to \infty$ for all $x, y \in H$ and $k \in \mathbb{N}$.

   (b) $\frac{1}{n}\sum_{j=1}^{n} |\langle S_j x, y\rangle_H| \to 0$ as $n \to \infty$ for all $x, y \in H$.

   (c) S is weakly mixing.

   (d) S does not admit an invariant subspace of dimension two, on which it acts by rotation.

   (e) The spectral measure μ_x of S is atomless for every x in the unitary part of S (the latter was defined in Lemma 5.2.3).

In various formulations, these statements have been well-known for several decades. For example, Item *(3e)* is a reformulation of the well-known Maruyama theorem on ergodicity of shifts of Gaussian fields [116]. The equivalence of ergodicity and weak mixing is a special case of e.g. [135]. Regarding *(2)* and *(3)*, Ustunel and Zakai have proved stronger results than these, even considering ergodic *random* (nonlinear) rotations of the Wiener space, see e.g. [142, 143, 144]. We are not aware of Item *(1)* in the literature, but it is basically a version of Blumenthal's 0-1 law for Gaussian processes (and indeed implies it for Brownian motion). We give a proof of the proposition only for the sake of completeness.

In the proof we will use the standard fact that for each $h \in H$ there exists an a.e. defined linear extension of the map from $H \to \mathbb{R}$ given by $v \mapsto \langle v, h\rangle_H$. By an abuse of notation we denote this linear extension as $\langle \cdot, h\rangle$ as well, and the map from $H \to L^2(X,\mu)$ given by $h \mapsto \langle \cdot, h\rangle$ is a linear isometry (in particular the law of each $\langle \cdot, h\rangle$ is a Gaussian of variance $\|h\|_H^2$ with respect to μ), see e.g. Section 3.2 in [85].

Proof of Item (1). Assume first that $\|S_n x\|_H \to 0$ for all $x \in H$. Then one can decompose H into an orthogonal direct sum: $H = \bigoplus_{n\geq 0} H_n$ where $H_n := \ker(S_{n+1}) \cap \mathrm{Im}(S_n^*)$, via the formula

$$x = \sum_{k=0}^{\infty} S_k^* S_k x - S_{k+1}^* S_{k+1} x.$$

The series converges to x in H because the N^{th} partial sum equals $x - S_{N+1}^* S_{N+1} x$ and we know that $\|S_{N+1}^* S_{N+1} x\|_H \leq \|S_{N+1} x\|_H \to 0$. Moreover one easily checks that $S_k^* S_k x - S_{k+1}^* S_{k+1} x \in \ker(S_{k+1}) \cap \mathrm{Im}(S_k^*)$, and that $\ker(S_i) \cap \mathrm{Im}(S_{i-1}^*)$ is orthogonal to $\ker(S_j) \cap \mathrm{Im}(S_{j-1}^*)$ for $i < j$.

Let ξ be a B-valued random variable with law μ and define ξ_n to be the projection of ξ onto H_n. Note that the ξ_n are independent X-valued random variables (but not necessarily iid). Note also that $S_n \xi$ is measurable with respect to $\{\xi_j\}_{j\geq n}$. Consequently $\bigcap_{n\geq 1} \sigma(S_n\xi) \subseteq \bigcap_{n\geq 1} \sigma(\{\xi_j : j \geq n\})$, which by Kolmogorov's 0-1 law is a 0-1 sigma algebra.

Conversely, suppose that $\|S_n x\| \not\to 0$ for some $x \in H$. Then by Lemma 5.2.3, the closed subspace $A := \bigcap_{n\geq 0} \mathrm{Im}(S_n^*)$ is nonzero. Letting ξ denote a random variable in X with law μ, let ξ_A denote the projection onto A applied to ξ. Since $S|_A$ is unitary (again by Lemma 5.2.3) it is clear that $\bigcap_n \sigma(S_n)$ contains at least $\sigma(\xi_A)$, which is nontrivial since $A \neq \{0\}$ so that at least one nonzero Gaussian random variable is measurable with respect to it. $\square$

Proof of Item (2). Suppose that $\langle S_n x, y\rangle_H \to 0$ for all $x, y \in H$. We wish to show that as $n \to \infty$,

$$\int_X f(S_n x) g(x) \mu(dx) \to 0$$

for all bounded measurable functions $f, g : X \to \mathbb{R}$ such that $\int_X f d\mu = \int_X g d\mu = 0$. By an application of Cauchy-Schwarz and the measure-preserving property of S_n, it suffices to prove this in a dense subspace of $L^2(X,\mu)$.

By using the Wick forumla (aka Isserlis' theorem), one can easily show that the claim is at least true whenever f and g are both of the form $x \mapsto p(\langle x,e_1\rangle,...,\langle x,e_k\rangle)$ for some $k\in\mathbb{N}$, some polynomial $p:\mathbb{R}^k\to\mathbb{R}$ and some orthonormal set of vectors $e_1,...,e_k$ in H. Then by a density argument and the fact that the $\langle x,e_i\rangle$ has Gaussian tails, one can extend this from k-variable polynomials p to all continuous functions $f,g:\mathbb{R}^k\to\mathbb{R}$ with at-worst polynomial growth at infinity. One can then extend to all bounded Borel-measurable functions $f,g:\mathbb{R}^k\to\mathbb{R}$ by density of continuous functions in $L^2(\mathbb{R}^k,\gamma_k)$ where γ_k is the standard Gaussian measure on $\mathbb{R}^k$.

Thus, to finish the argument, it suffices to show that the set of all functions of the form $x\mapsto f(\langle x,e_1\rangle,...,\langle x,e_k\rangle)$, where $f:\mathbb{R}^k\to\mathbb{R}$ is bounded and measurable, is dense in $L^2(X,\mu)$. To show this, choose an orthonormal basis $\{e_j\}$ for H and Let $\mathcal{F}_n$ denote the sigma algebra generated by $\langle x,e_1\rangle,...,\langle x,e_n\rangle$. Let $f\in L^\infty(X,\mu)$ and let $f_n:=\mathbb{E}[f|\mathcal{F}_n]$. Then f_n is bounded and measurable and of the form $x\mapsto h(\langle x,e_1\rangle,...,\langle x,e_n\rangle)$. Furthermore by martingale convergence $\|f_n-f\|_{L^2}\to 0$, completing the proof.

The converse direction is striaghtforward: if S is mixing, then apply the mixing definition to $f(x):=\langle x,a\rangle$ and $g(x):=\langle x,b\rangle$ to conclude that $\langle S_na,b\rangle\to 0$. □

Proof of Item (3). We are going to show that ergodicity implies (a) which implies (b) which implies (c). Clearly (c) implies ergodicity. Then we show that (d) is equivalent to (e) for the unitary part of the operator from the decomposition in Lemma 5.2.3. Then we will show that for unitary operators, (e) holds if and only if (b) holds.

So assume that S is ergodic as a map from $X\to X$. This is equivalent to the statement that $\int_X\big(\frac{1}{n}\sum_{j=1}^n f(S_jx)\big)g(x)\mu(dx)\to 0$ for all $f,g\in L^2(X,\mu)$ such that $\int_X fd\mu=\int_X gd\mu=0$. Now fix $k\in\mathbb{N}$ and $a,b\in H$. Letting H_k denote the k^{th} Hermite polynomial, we set $f(x):=\frac{1}{\sqrt{k!}}H_k(\langle x,a\rangle)$ and $g(x):=\frac{1}{\sqrt{k!}}H_k(\langle x,b\rangle)$. Then $\int_X f(S_jx)g(x)d\mu=\langle a,S_jb\rangle_H^k$ (see e.g. [125]), so by ergodicity we obtain (a).

Now assume (a) holds. Let $a, b \in H$ with $\|a\| = \|b\| = 1$. Fix $\epsilon > 0$ and let $p : [-1, 1] \to \mathbb{R}$ be a polynomial such that $\sup_{x\in[-1,1]} \big|p(x) - |x|\big| < \epsilon$. Since $|\langle a, S_j b\rangle| \le 1$ for all $j \in \mathbb{N}$, it follows that $\frac{1}{n}\sum_{j=1}^n \big|p(\langle a, S_j b\rangle) - |\langle a, S_j b\rangle|\big| < \epsilon$. Moreover, since (a) holds we know that $\frac{1}{n}\sum_{j=1}^n p(\langle a, S_j b\rangle) \to 0$. Consequently we find that $\limsup_n \frac{1}{n}\sum_{j=1}^n |\langle a, S_j b\rangle_H| \le \epsilon$. Since ϵ is arbitrary, it follows that (b) holds.

Now assume that (b) holds. To show (c), we want that $\frac{1}{n}\sum_{j=1}^n \big|\int_X f(S_j x)g(x)\mu(dx)\big| \to 0$ for all $f, g \in L^2(X, \mu)$ that have mean zero. First note that, by essentially the same series of density arguments given in the proof of Item *(2)*, it suffices to prove this whenever f, g are of the form $x \mapsto p(\langle x, e_1\rangle, ..., \langle x, e_k\rangle)$ for some $k \in \mathbb{N}$, some polynomial $p : \mathbb{R}^k \to \mathbb{R}$ and some orthonormal set of vectors $e_1, ..., e_k$ in H. In turn, by Wick formula it suffices to show that $\frac{1}{n}\sum_{j=1}^n |\langle a, S_j b\rangle_H|^k \to 0$ for all $a, b \in H$ and all $k \in \mathbb{N}$. Note that we have $|\langle a, S_j b\rangle| \le \|a\|\|b\|$, so $|\langle a, S_j b\rangle_H|^k \le \|a\|^{k-1}\|b\|^{k-1}|\langle a, S_j b\rangle_H|$. Summing over j and applying (b), we conclude $\frac{1}{n}\sum_{j=1}^n |\langle a, S_j b\rangle_H|^k \to 0$.

It is straightforward to show that (using the spectral theorem of the unitary operator) atoms of μ_x correspond precisely to complex eigenvalues of S, i.e., two-dimensional subspaces on which S acts by rotation. Thus (d) implies (e) and vice versa (by focusing only on the unitary part of the operator).

Finally we explain why (e) is equivalent to (b). The spectral measure μ_x of S is supported on $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$ and defined via its Fourier transform: $\hat{\mu}_x(k) := \langle S_k x, x\rangle$ where $k \in \mathbb{Z}$ and $S_k := S^*_{-k}$ if $k < 0$. Thus, to show that (b) and (e) are equivalent, we just need to show that a finite measure μ on $\mathbb{T}$ is atomless if and only if $\frac{1}{n}\sum_{k=-n}^n |\hat{\mu}(k)| \to 0$ as $n \to \infty$. This is a direct consequence of Wiener's Lemma.

□

5.2.2 Strassen's law for mixing linear operators

In this subsection we prove Propositions 5.1.1 and 5.1.2. First we need three preliminary lemmas and then we will formulate the result as Theorem 5.2.10. Some of the results given in this subsection are due to Arcones [6] and others, but we state the proofs because some of these will be important in generalizing later.

Lemma 5.2.5. *Let $(X_i)_{i\geq 1}$ be a stationary sequence of real-valued jointly Gaussian random variables. Suppose that $var(X_0) = 1$ and $cov(X_0, X_n) \to 0$ as $n \to \infty$. Then*

$$\limsup_{n\to\infty} \frac{X_n}{\sqrt{2\log n}} = 1, \qquad a.s.$$

This lemma is classical and is a special case of Lemma 2.1 of [6], which in turn is an improvement of results from [106, 124]. Nonetheless we give a proof just to illustrate the fact that this is the only place where the mixing condition is used, and is also the reason why we are not able to extend the result to the ergodic case (i.e., we cannot prove or find a counterexample to the above fact under the weaker assumption $\frac{1}{n}\sum_1^n |cov(X_0, X_n)| \to 0$).

Proof. Define $c_n := cov(X_0, X_n)$ and fix $\epsilon > 0$. Choose $r \in \mathbb{N}$ such that $c_n < \epsilon$ for $n \geq r$. Let $Y_n := \epsilon\xi + \sqrt{1-\epsilon}Z_n$ where ξ, Z_n are iid standard Gaussians.

Note that if $i < j$ then $cov(Y_i, Y_j) = \epsilon > c_{r(j-i)} = cov(X_{ri}, X_{rj})$. Therefore by Slepian's Lemma [139] we find that

$$\mathbb{E}[\max\{Y_0, Y_1, Y_2 ..., Y_n\}] \leq \mathbb{E}[\max\{X_0, X_r, X_{2r}, ..., X_{nr}\}]$$

for any $n \in \mathbb{N}$. Since the Z_n are i.i.d. standard normals it is clear that

$$\lim_{n\to\infty} (2\log n)^{-1/2}\mathbb{E}[\max_{1\leq i\leq n} Z_i] = 1$$

and therefore

$$\lim_{n\to\infty}(2\log n)^{-1/2}\mathbb{E}[\max_{1\leq i\leq n} Y_i] = \sqrt{1-\epsilon}.$$

Consequently we find that

$$\limsup_{n\to\infty}(2\log n)^{-1/2}\mathbb{E}[\max_{0\leq i\leq n} X_{ri}] \geq \sqrt{1-\epsilon}.$$

Because $\lim_{n\to\infty}\frac{\log n)^{1/2}}{(\log(rn))^{1/2}} = 1$ for any $r \in \mathbb{N}$, we get

$$\limsup_{n\to\infty}(2\log n)^{-1/2}\mathbb{E}[\max_{0\leq i\leq n} X_i] \geq \sqrt{1-\epsilon}.$$

Since ϵ is arbitrary this means that

$$\limsup_{n\to\infty}(2\log n)^{-1/2}\mathbb{E}[\max_{0\leq i\leq n} X_i] \geq 1.$$

Letting $G_n := (2\log n)^{-1/2}\max_{1\leq i\leq n} X_i$ we claim that there exists a random variable $W \geq 0$ with $\mathbb{E}[W] < \infty$ such that $\sup_n G_n \leq W$. Indeed, it is clear that

$$G_n \leq \frac{\max_{1\leq i\leq n}|X_i|}{\max\{1,\log i\}^{1/2}} \leq \frac{\sup_{k\in\mathbb{N}} X_k}{\max\{1,\log k\}^{1/2}} =: W.$$

To see that W is integrable, note that if $a > 2$, then by a union bound we have

$$\mathbb{P}(W > a) \leq \sum_{k\in\mathbb{N}}\mathbb{P}(X_k > a(\log k)^{1/2}) \leq C\sum_{k\geq 1} k^{-a^2/2} \leq Ce^{-ca^2}$$

where $C, c > 0$ are appropriately chosen constants, and we have used the fact that each X_k is a standard Gaussian. Thus by the “reverse Fatou Lemma" we find that

$$\mathbb{E}[\limsup_{n\to\infty}(2\log n)^{-1/2}\max_{0\leq i\leq n} X_i] \geq \limsup_{n\to\infty}(2\log n)^{-1/2}\mathbb{E}[\max_{0\leq i\leq n} X_i] \geq 1.$$

Note that, by replacing $\log n$ (and $\log k$) with $\log(n+N)$ (and $\log(k+N)$) throughout the proof so far, we can show that

$$\mathbb{E}[\limsup_{n\to\infty}(2\log(n+N))^{-1/2}\max_{0\le i\le n} X_i] \ge 1,$$

for any $N \in \mathbb{N}$. Since

$$(2\log(n+N))^{-1/2}\max_{0\le i\le n} X_i \le \max_{0\le i\le n}(2\log(i+N))^{-1/2}X_i,$$

the previous expression implies that

$$\mathbb{E}[\sup_{k\in\mathbb{N}}(2\log(k+N))^{-1/2}X_k] \ge 1$$

for all $N \in \mathbb{N}$. By stationarity of $(X_i)_i$, we can replace X_k by X_{k+N} here to obtain

$$\mathbb{E}[\sup_{k>N}(2\log k)^{-1/2}X_k] \ge 1$$

for all $N \in \mathbb{N}$. Since $\sup_{k>N}(2\log k)^{-1/2}X_k \le W$ for all $N \ge 2$, and since $\mathbb{E}[W] < \infty$, the dominated convergence theorem shows that

$$\mathbb{E}[\limsup_{k\to\infty}(2\log k)^{-1/2}X_k] = \lim_{N\to\infty}\mathbb{E}[\sup_{k>N}(2\log k)^{-1/2}X_k] \ge 1. \tag{5.2}$$

On the other hand, since each X_k is standard normal, it is clear that for any $\delta > 0$,

$$\sum_{k\in\mathbb{N}}\mathbb{P}(X_k > \sqrt{(2+\delta)\log k}) \le \sum_{k\in\mathbb{N}} Ck^{-1-\delta/2} < \infty,$$

which implies via Borel Cantelli lemma that

$$\limsup_{k\to\infty}(2\log k)^{-1/2}X_k \le 1 \tag{5.3}$$

almost surely. Now combining (5.2) and (5.3) completes the proof. □

Next we deduce an easy corollary which generalizes the above lemma to $\mathbb{R}^N$. We denote the unit sphere of $\mathbb{R}^N$ to be the set of all points $(a_1, ..., a_N)$ with $\sum_1^N a_i^2 = 1$.

Corollary 5.2.6. *Let $(\vec{X}_n)_{n\geq 1}$ be a stationary sequence of jointly Gaussian random variables in $\mathbb{R}^N$, say $\vec{X}_n = (X_n^1, X_n^2, ..., X_n^N)$. Suppose that $cov(X_0^i, X_0^j) = \delta_{ij}$ and that $cov(X_0^i, X_n^j) \to 0$ as $n \to \infty$ for all $1 \leq i, j \leq N$. Then the unit sphere of $\mathbb{R}^N$ is contained in the set of limit points of the random sequence $((2\log n)^{-1/2}\vec{X}_n)_{n\geq 1}$.*

Proof. By Lemma 5.2.5, if we restrict our attention to only the first coordinate, then the set of limit points of the random sequence $((2\log n)^{-1/2}\vec{X}_n)_{n\geq 1}$ must contain a point on the set $A := \{(a_1, ..., a_N) : a_1 = 1\}$. On the other hand, by the same Borel-Cantelli argument used in proving (5.3) (with X_k replaced by $\|\vec{X}_k\|$ where $\|\cdot\|$ denotes the Euclidean norm), it follows that the set of limit points must be contained in the set $B := \{(a_1, ..., a_N) : \sum_{i=1}^N a_i^2 \leq 1\}$.

Since $A \cap B = \{(1, 0, ..., 0)\}$ it follows that $(1, 0, ..., 0)$ is a.s. a limit point of the random sequence $((2\log n)^{-1/2}\vec{X}_n)_{n\geq 1}$. The fact that any point on the unit sphere is a limit point then follows from rotational invariance of the conditions that $cov(X_0^i, X_0^j) = \delta_{ij}$ and $cov(X_0^i, X_n^j) \to 0$ (i.e., these conditions remain true if we replace $(\vec{X}_n)_n$ by $(U(\vec{X}_n))_n$ for some orthogonal $N \times N$ matrix U). $\square$

Lemma 5.2.7. *Let (X, H, μ) be an abstract Wiener space of infinite dimension. Let $S(H) := \{h \in H : \|h\|_H = 1\}$ and let $B(H) := \{h \in H : \|h\|_H \leq 1\}$. Then the closure in X of $S(H)$ is $B(H)$.*

Proof. $B(H)$ is a compact (hence closed) subset of X which contains $S(H)$, so the set of limit points of $S(H)$ must be contained in $B(H)$. Choose an orthonormal basis $\{e_n\}_n$ for H. Since $e_n \in B(H)$, since $B(H)$ is compact in X, and since $e_n \to 0$ weakly in H, it follows that $\|e_n\|_X \to 0$ as $n \to \infty$. Now fix $h \in H$ with $\|h\|_H \leq 1$. Choose $c_n \in \mathbb{R}$ such that $\|h + c_n e_n\|_H = 1$. It is clear that $|c_n| \leq 2$ (otherwise $1 = \|c_n e_n + h\|_H \geq |c_n|\|e_n\| - \|h\| > 2 - 1 = 1$), and therefore $\|c_n e_n\|_X \to 0$. Thus $h + c_n e_n$ is a sequence in $S(H)$ converging to $h \in B(H)$ with respect to the topology of X. $\square$

Lemma 5.2.8. *Let (X, H, μ) be an abstract Wiener space. Suppose that $S : H \to H$ is some operator satisfying $SS^* = I$. Let $S_N := S^N$. If $\|S_N x\|_H \to 0$, then for any $a \notin K$ there exists $\epsilon > 0$ such that*

$$\mu\Big(\Big\{x \in X : \frac{S_k x}{\sqrt{\log k}} \in B(a, \epsilon) \ \ i.o.\Big\}\Big) = 0.$$

Proof. By definition of the Cameron Martin space, for every $x \in X$, we have

$$\sup\{\ell(x) : \ell \in X^*, \int \ell^2 d\mu = 1\} = \|x\|_H$$

(we set $\|x\|_H = \infty$ when $x \notin H$). For any $x \in X \backslash K$, there exists $\ell \in X^*$ such that $\ell(x) > \sqrt{2}$ and $\int_X \ell^2 d\mu = 1$. For each $n \in \mathbb{N}$, under μ, $\ell(S_n x)$ is a standard Gaussian random variables. Using Borel-Cantelli lemma and Gaussian tail bound, together with $\ell(x) > \sqrt{2}$, there exist small enough $\epsilon_x > 0$ such that

$$\mathbb{P}\Big(\{\frac{\ell(S_n \xi)}{\sqrt{\log n}}\}_{n=1}^{\infty} \in B(\ell(x), \epsilon_x) \quad i.o.\Big) = 0$$

By making ϵ_x smaller, we have

$$\mathbb{P}\Big(\{\frac{S_n \xi}{\sqrt{\log n}}\}_{n=1}^{\infty} \in B(x, \epsilon_x) \quad i.o.\Big) = 0.$$

□

Lemma 5.2.9. *[108, Equation (4.4)] Let μ be a centered Gaussian measure on a separable Banach space, then for all $a > 0$*

$$\mu(\{x \in X : \|x\|_X > a + \int_X \|u\|_X \mu(du)\}) \le e^{-a^2/(2\sigma^2)},$$

where $\sigma := \sup_{\|f\|_{X^*} \le 1} \int_X f(x)^2 \mu(dx) = \sup_{\|h\|_H \le 1} \|h\|_X$.

Finally we are ready to prove the main result of the section, a restatement of Proposition 5.1.1 in addition to a partial converse.

Theorem 5.2.10. *Let (X, H, μ) be an abstract Wiener space. Let $E \subset X$ be a Borel measurable linear subspace of measure 1, and suppose $S : E \to X$ is linear and measure-preserving. If S is* *mixing, then the set of limit points of the random sequence* $((2\log n)^{-1/2}S^n x)_{n\in\mathbb{N}}$ *equals the unit ball of* H *for a.e.* $x \in X$*. Conversely, if* S *is any measure-preserving linear operator such that the set of limit points of the random sequence* $((2\log n)^{-1/2}S^n x)_{n\in\mathbb{N}}$ *equals the unit ball of* H *for a.e.* $x \in X$*, then* S *must be ergodic.*

Proof. Fix $h \in H$ with $\|h\|_H = 1$, and let $\epsilon > 0$. We wish to show that $\|(2\log n)^{-1/2}S_n x - h\|_H < \epsilon$ infinitely often. To this end, let $\{e_i\}_i$ be an orthonormal basis of H with $e_1 = h$. Let P_k denote the orthogonal projection onto the subspace M_k which is spanned by $\{e_i\}_{i=1}^k$.

Choose some $k \in \mathbb{N}$ so that $\int_X \|x - P_k x\|_X^2 \mu(dx) < (\epsilon/5)^2$. Then for all n we have that

$$\mathbb{P}(\|S_n\xi - P_k(S_n\xi)\|_X > a) = \mathbb{P}(\|\xi - P_k(\xi)\|_X > a) \le e^{-Ma^2}\mathbb{E}[e^{M\|\xi - P_k(\xi)\|_X^2}] \tag{5.4}$$

for any $a > 0$ and for any $M > 0$ such that the expectation on the right side is finite. By Lemma 5.2.9, it is true that $\int_X e^{\alpha\|x\|^2}\nu(dx)$ is finite whenever $\alpha < (2\int_X \|x\|^2\nu(dx))^{-1}$, for any centered Gaussian measure ν on X. Consequently in (5.4) we can take M to be $(5/\epsilon)^2$ and we can take a to be $\epsilon\sqrt{2\log n}/4$ and we find

$$\sum_{n\in\mathbb{N}} \mathbb{P}\bigg(\frac{\|S_n\xi - P_k(S_n\xi)\|_X}{\sqrt{2\log n}} > \epsilon/4\bigg) < \infty,$$

where ξ is distributed according to μ. Consequently by Borel-Cantelli lemma, we find that

$$(2\log n)^{-1/2}\|S_n\xi - P_k(S_n\xi)\|_X < \epsilon/4$$

for all but finitely many n, a.s.. Thus we just need to show that $\|(2\log n)^{-1/2}P_k(S_n x) - h\|_X < \epsilon/4$ infinitely often. But this is just a finite dimensional statement which immediately follows from Corollary 5.2.6. Indeed, by assumption h lies on the unit sphere of M_k, and the joint covariances tend to zero precisely because of the condition that S is mixing (via Item *(2)* in Proposition 5.2.4).

In the notation of Lemma 5.2.7, we have shown that any point on $S(H)$ is almost surely a limit point of $(2\log n)^{-1/2}S_nx$. By that same lemma, any point of $B(H)$ is also a limit point. It is clear from Lemma 5.2.8 that no point outside of $B(H)$ can be a limit point of $(2\log n)^{-1/2}S_nx$. That lemma is still valid in this case, since it only relies on the measure-preserving property of S_n and nothing else. This completes the proof of the first statement.

Next we prove that S must be ergodic if the set of limit points of $((2\log n)^{-1/2}S^nx)_{n\in\mathbb{N}}$ equals the unit ball of H for a.e. $x \in X$. Indeed, if S is not ergodic, then by Item *(3d)* in Proposition 5.2.4, there is a two-dimensional invariant subspace M on which S acts by rotation. We claim that no nonzero point which lies in M can be visited infinitely often, since $(2\log n)^{-1/2}S_nx$ converges to zero a.s. for any $x \in M$. Indeed, let P be the projection onto M and let $Q = I - P$ denote the projection onto $M^\perp$. Let ξ be sampled from μ, so that $Q\xi$ has Cameron-Martin space $M^\perp$. Since M and $M^\perp$ are invariant under S, it follows that $PS_n = S_nP$ and $QS_n = S_nQ$. By Lemma 5.2.8 it is clear that the set of limit points of $(2\log n)^{-1/2}QS_n\xi = (2\log n)^{-1/2}S_n(Q\xi)$ must be contained in the unit ball of the Cameron-Martin space of $Q\xi$, namely $M^\perp$ (this lemma only relies on the measure-preserving property of S_n and nothing else). Furthermore, since $(2\log n)^{-1/2}S_nx$ converges to zero a.s. for any $x \in M$, it follows that $(2\log n)^{-1/2}PS_n\xi = (2\log n)^{-1/2}S_n(P\xi)$ converges to zero a.s. Consequently the set of limit points of $(2\log n)^{-1/2}S_n\xi = (2\log n)^{-1/2}PS_n\xi + (2\log n)^{-1/2}QS_n\xi$ must also be contained in $M^\perp$, and hence contain no nonzero points of M, proving the claim. $\square$

This already proves Proposition 5.1.1, and to prove Proposition 5.1.2 one just needs the following lemma.

Lemma 5.2.11. *Let (X, H, μ) be an abstract Wiener space. Suppose that $S_t : H \to H$ is some family of operators satisfying $S_tS_t^* = I$. Moreover assume that (S_t) extends to a strongly continuous semigroup on X. Then there exists a deterministic function $C : [0,1] \to \mathbb{R}_+$ such that*

$C(\rho) \to 0$ *as* $\rho \downarrow 0$ *and such that*

$$\mu\left(\left\{x \in X : \limsup_{k\to\infty} \frac{\sup_{t\in[k\rho,(k+1)\rho]} \|S_t x - S_{k\rho}x\|_X}{\sqrt{\log k}} \leq C(\rho)\right\}\right) = 1$$

for all $\rho \in [0, 1]$. *In particular* $(S_t x)_{t\geq 0}$ *has the same set of cluster points as* $(S_N x)_{N\in\mathbb{N}}$ *for* x *in a set of full* μ*-measure.*

Proof. Let ξ denote a random variable in X with law μ, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then $(S_t\xi)_{t\in[0,1]}$ is continuous in t, thus it can be viewed as a Gaussian random variable taking values in the Banach space $C([0, 1], X)$ of continuous paths in X, endowed with norm $\|F\|_{C([0,1],X)} := \sup_{t\in[0,1]} \|F(t)\|_X$. By Fernique's theorem, $\mathbb{E}[\sup_{t\in[0,1]} \|S_t\xi - \xi\|^p_X] < \infty$ for all $p \geq 1$. By continuity, we also know that $\sup_{t\in[0,\rho]} \|S_t\xi - \xi\|_X \to 0$ as $\rho \downarrow 0$. Letting

$$C(\rho) := \mathbb{E}[\sup_{t\in[0,\rho]} \|S_t\xi - \xi\|^2_X],$$

we thus have by uniform integrability that $\lim_{\rho\downarrow 0} C(\rho) = 0$. Letting

$$A(k, \rho) := \sup_{t\in[k\rho,(k+1)\rho]} \|S_t x - S_{k\rho}x\|_X,$$

the stationarity of the process $(S_t\xi)_{t\geq 0}$ implies that $A(k, \rho) \overset{d}{=} A(0, \rho)$ for all $k \in \mathbb{N}$. We thus find that

$$\mathbb{P}(A(k, \rho) > \mathbb{E}[A(0, \rho)] + u) = \mathbb{P}(A(0, \rho) > \mathbb{E}[A(0, \rho)] + u) \leq e^{-u^2/(2C(\rho))},$$

where we used Lemma 5.2.9 in the last inequality. But $\mathbb{E}[A(0, \rho)] \leq \mathbb{E}[A(0, \rho)^2]^{1/2} = C(\rho)^{1/2}$, and therefore

$$\mathbb{P}(A(k, \rho) > u) \leq e^{-(u-C(\rho)^{1/2})^2/(2C(\rho))} \leq e^{1/2} \cdot e^{-u^2/(4C(\rho))}$$

where we used $(a - b)^2 \geq \frac{1}{2}a^2 - b^2$ in the last inequality. Thus

$$\mathbb{P}(A(k, \rho) > 4C(\rho)\sqrt{\log k}) \lesssim k^{-4},$$

so by the Borel-Cantelli lemma we find that

$$\limsup_{k\to\infty} \frac{A(k,\rho)}{\sqrt{\log k}} \leq 4C(\rho), \quad a.s.$$

As we already observed eariler $C(\rho) \to 0$ as $\rho \downarrow 0$, completing the proof. $\square$

To summarize the main results of this section, let (X, H, μ) be an abstract Wiener space, let $S : X \to X$ be measure-preserving and a.e. linear, and consider the following statements:

1. $\bigcap_n \sigma(S_n)$ is a trivial sigma algebra.

2. S is mixing.

3. The set of limit points of the random sequence $((2\log n)^{-1/2} S^n x)_{n\in\mathbb{N}}$ equals the unit ball of H for a.e. $x \in X$.

4. S is ergodic/weakly mixing.

What we have shown is that (1) implies (2) (which is actually true for any dynamical system, by e.g. the reverse martingale convergence theorem), that (2) implies (3), and that (3) implies (4). From Proposition 5.2.4, it is clear that (2) does not imply (1) in general, and that (4) does not imply (2) in general.

We do not know if (3) implies (2) or if (4) implies (3), although obviously both cannot be true in general. To show that (4) implies (3), one would need to prove Lemma 5.2.5 with the condition $cov(X_0, X_n) \to 0$ replaced by the condition $\frac{1}{n}\sum_{j=1}^n |cov(X_0, X_j)| \to 0$. We do not know how to prove this, nor are we certain that it is even true. We have reason to suspect that “(4) implies (3)" may actually be false, and a counterexample might be given by an operator whose spectral measure is atomless but highly singular with respect to Lebesgue measure. For instance, in the case that the spectral measure looks like the usual two-thirds Cantor measure, we have some reason to suspect that the set of limit points of the random sequence $((2\log n)^{-1/2} S^n x)_{n\in\mathbb{N}}$ equals the ball in H of radius $\sqrt{\log_3 2}$ for a.e. $x \in X$, rather than the unit ball. However, we do not have a proof of this.

5.2.3 Higher Gaussian Chaos

The main goal of this subsection is to prove a version of Strassen's law for higher Gaussian chaoses, which reduces to the previous version for first-order chaos.

If (X, H, μ) is an abstract Wiener space, then the k^{th} homogeneous Wiener chaos, denoted by $\mathcal{H}^k(X, \mu)$ is defined to be the closure in $L^2(X, \mu)$ of the linear span of $H_k \circ g$ as g varies through all elements of the continuous dual space X^*, where H_k denotes the k^{th} Hermite polynomial

$$H_k(x) := (-1)^k e^{\frac{x^2}{2}} \frac{d^k}{dx^k} e^{-\frac{x^2}{2}}.$$

Equivalently it can be described as the closure in $L^2(X, \mu)$ of the linear span of $H_k(\langle \cdot, v \rangle)$ as v ranges through all elements of H. One always has the orthonormal decomposition $L^2(X, \mu) = \bigoplus_{k \geq 0} \mathcal{H}^k(X, \mu)$, see e.g. [125, Section 1.1]. Sometimes the k^{th} chaos is also described slightly differently as linear combinations of products of Hermite polynomials, for instance in [91], but our formulation is equivalent by the umbral identity for Hermite polynomials, see Corollary 2.3 in [84].

Definition 5.2.12. *Let (X, H, μ) be an abstract Wiener space, and let Y be another separable Banach space. A Borel-measurable map $T : X \to Y$ is called homogeneous of order k if $f \circ T \in \mathcal{H}^k(X, \mu)$ for all $f \in Y^*$.*

Before formulating the main result of this section, we collect a few important results about homogeneous variables, whose proofs may be found in Appendix 1 below.

Proposition 5.2.13. *Let (X, H, μ) be an abstract Wiener space and let $T : X \to Y$ be homogeneous of order k. Let $\|T\|^2_{L^2(X,\mu;Y)} := \int_X \|T(x)\|^2_Y \mu(dx)$. Then $\|T\|_{L^2(X,\mu;Y)} < \infty$ and moreover*

$$\mu(\{x \in X : \|T(x)\|_Y > u\}) < C \exp\left[-\alpha\big(u/\|T\|_{L^2(X,\mu;Y)}\big)^{2/k}\right], \tag{5.5}$$

where $C, \alpha > 0$ depend on k but are independent of the choice of X, H, μ, Y, T, and $u > 0$.

This is proved as Corollary 5.5.14 of the appendix.

Let (X, H, μ) be an abstract Wiener space. Choose an orthonormal basis $\{e_i\}_i$ for H, and let

$$P_N x := \sum_{j=1}^{N} \langle x, e_j \rangle e_j, \qquad\qquad Q_N x := \sum_{j=N+1}^{\infty} \langle x, e_j \rangle e_j = x - P_N x.$$

Note that $P_N x, Q_N x$ are independent. Therefore if (x, y) is sampled from $\mu^{\otimes 2}$, then $P_N x + Q_N y$ is distributed as μ. If $T : X \to Y$ is homogeneous of order k, we thus define a sequence of "finite-rank Cameron-Martin projections for T" by the formula

$$T_N(x) := \int_X T(P_N x + Q_N y)\mu(dy). \tag{5.6}$$

This is a well-defined Bochner integral for μ a.e. $x \in X$. Indeed, since

$$\int_X \int_X \|T(P_N x + Q_N y)\|_Y \mu(dy)\mu(dx) = \int_X \|T(u)\|_Y \mu(du) < \infty,$$

it follows that $\int_X \|T(P_N x + Q_N y)\|_Y \mu(dy)$ is finite for a.e. x.

Proposition 5.2.14. *Let (X, H, μ) be an abstract Wiener space, and let $T : X \to Y$ be homogeneous of order k. If T_N is defined as in* (5.6)*, then T_N is also homogeneous of order k, and moreover $\|T_N - T\|_Y \to 0$ a.e. and in every $L^p(X, \mu)$ as $N \to \infty$. In fact, one has the following super-polynomial convergence bound:*

$$\mu(\{x \in X : \|T_N(x) - T(x)\|_Y > u\}) \leq C \exp\big[-\alpha\big(u/\|T_N - T\|^2_{L^2(X,\mu;Y)}\big)^{2/k}\big] \tag{5.7}$$

where C, α are independent of X, H, μ, Y, T, u, N and the choice of basis $\{e_i\}_i$, but may depend on the homogeneity k.

This is proved as Proposition 5.5.15 in the appendix.

If (X, H, μ) is an abstract Wiener space and $T : X \to Y$ is homogeneous of order k, then we define the associated map $T_{hom} : H \to Y$ by

$$T_{hom}(h) := \int_X T(x+h)\mu(dx) = \frac{1}{k!}\int_X T(x)\langle x, h\rangle^k \mu(dx).$$

See (5.45) and (5.46) of the appendix for the proof of the equality.

Lemma 5.2.15. *Let (X, H, μ) be an abstract Wiener space, and let $T : X \to Y$ be homogeneous of order k. Choose an orthonormal basis $\{e_i\}$ of H and let T_N be the finite-rank approximation given in* (5.6)*. Then we have the uniform convergence*

$$\lim_{N\to\infty} \sup_{\|h\|_H \le 1} \left\| (T_N)_{hom}(h) - T_{hom}(h) \right\|_Y = 0.$$

Letting $B(H)$ denote the unit ball of H, it follows that T_{hom} is continuous from $B(H) \to Y$, where $B(H)$ is given the topology of X (not of H).

Corollary 5.2.16. *Let (X, H, μ) be an abstract Wiener space, and let $T^i : X \to Y_i$ be homogeneous of order k_i for $1 \le i \le m$, where $m \in \mathbb{N}$. Then the set*

$$\{(T^1_{hom}(h), ..., T^m_{hom}(h)) : h \in B(H)\}$$

is a compact subset of $Y_1 \times \cdots Y_m$.

These results are proved as Lemma 5.5.21 and Corollary 5.5.22 in the Appendix. With all of these preliminaries in place, we are ready to formulate the main theorem of this subsection, which is a generalization of the main theorem to homogeneous variables of order k.

Theorem 5.2.17. *Let (X, H, μ) be an abstract Wiener space, and let $T^i : X \to Y_i$ be homogeneous of order k_i for $1 \le i \le m$, where $m \in \mathbb{N}$. Let $S : X \to X$ be a.e. linear, measure-preserving, and mixing. Then almost surely, the set of limit points of the random set*

$$\big\{ \big((2\log n)^{-k_1/2} T^1(S_n x),\ \ldots,\ (2\log n)^{-k_m/2} T^m(S_n x) \big) : n \in \mathbb{N} \big\}$$

is equal to the compact set

$$K := \{(T^1_{hom}(h), ..., T^m_{hom}(h)) : h \in B(H)\}.$$

Note that Proposition 5.1.1 can be viewed as a special case where $k_1 = 1, m = 1, Y = X$, and T^1 is the identity on X.

Proof. Let $S(H) := \{h \in H : \|h\|_H = 1\}$. Note by Lemma 5.2.7 that $S(H)$ is dense in $B(H)$ with respect to the topology of X. Therefore, the set

$$D := \{(T^1_{hom}(h), ..., T^m_{hom}(h)) : h \in S(H)\}$$

is dense in K. Thus is suffices to show that any point in D is a limit point of the given sequence. So fix a point $(T^1_{hom}(h), ..., T^m_{hom}(h)) \in D$, where $\|h\|_H = 1$, and let $\epsilon > 0$. We wish to show that

$$\sum_{i=1}^m \left\|(2\log n)^{-k_i/2} T^i(S_n x) - T^i_{hom}(h)\right\|_{Y_i} < \epsilon \tag{5.8}$$

infinitely often. Fix an orthonormal basis $\{e_i\}$ for H, with $e_1 = h$, and let T^i_N be the associated finite rank Cameron Martin projections as in (5.6). We claim that it is enough to prove (5.8) with each $T^i(S_n x)$ replaced by $T^i_N(S_n x)$ and $T^i_{hom}(h)$ replaced by $(T^i_N)_{hom}(h)$ (and also replacing ϵ by $\epsilon/2$), for some large enough N.

Indeed, by Proposition 5.2.14 we can choose N so large that $\int_X \|T^i_N - T^i\|^2_{Y_i} d\mu < \epsilon/(2m)$. Then by (5.7) and the fact that S_n is measure preserving, it is clear that

$$\sum_{n \geq 2} \mu(\{x : (2\log n)^{-k_i/2} \|T^i_N(S_n x) - T^i(S_n x)\|_Y > \epsilon/m\}) < \infty, \tag{5.9}$$

so by Borel Cantelli lemma, $(2\log n)^{-k_i/2}\|T^i_N(S_n x) - T^i(S_n x)\|_Y < \epsilon/m$ for all but finitely many $n \in \mathbb{N}$ almost surely. Furthermore, by Lemma 5.2.15 we can (by making N larger) ensure that $\|(T^i_N)_{hom}(h) - T^i_{hom}(h)\|_Y < \epsilon/m$.

Thus we just need to show that

$$\sum_{i=1}^{m} \left\|(2\log n)^{-k_i/2} T_N^i(S_n x) - (T_N^i)_{hom}(h)\right\|_{Y_i} < \epsilon/2 \tag{5.10}$$

infinitely often. Note that each T^i is homogeneous of order k and measurable with respect to a *finite* collection $\{\langle\cdot, e_i\rangle\}_{i=1}^N$ of i.i.d. standard Gaussians, therefore one can verify that each T_N^i can be written as

$$T_N^i(x) = \sum_{j=1}^{M_i} y_j^i H_{k_i}(\langle v_j^i, x\rangle), \tag{5.11}$$

for some *finite* collection of vectors $\{v_j^i\} \subset span(\{e_i\}_{i=1}^N)$, $\{y_j^i\} \subset Y$, and $M_i \in \mathbb{N}$. Here H_k is the k^{th} Hermite polynomial. Using Cameron Martin theorem, one can see that

$$(T_N^i)_{hom}(h) = \sum_{j=1}^{M_i} y_j^i \langle v_j^i, h\rangle^{k_i}. \tag{5.12}$$

Note that

$$\left|\langle v_j^i, (2\log n)^{-1/2} S_n x\rangle^{k_i} - (2\log n)^{-k_i/2} H_{k_i}(\langle v_j^i, S_n x\rangle)\right| \to 0 \qquad a.s.,$$

by Borel-Cantelli lemma and the fact that $x \mapsto \langle v_j^i, S_n x\rangle$ are standard Gaussians under μ. Thus it suffices to show that

$$\sum_{\substack{1\le i\le m \\ 1\le j\le M_i}} \|y_j^i\|_Y \left|\langle v_j^i, (2\log n)^{-1/2} S_n x\rangle^{k_i} - \langle v_j^i, h\rangle^{k_i}\right| < \epsilon/2$$

happens infinitely often. Letting $P_N : H \to span\{e_i\}_{i=1}^N$ denote the orthogonal projection, it is clear that $v_j^i = P_N(v_j^i)$, thus by exploiting self-adjointness of P_N, the previous expression is equivalent to showing that

$$\sum_{\substack{1\le i\le m \\ 1\le j\le M_i}} \|y_j^i\|_Y \left|\langle v_j^i, (2\log n)^{-1/2} P_N(S_n x)\rangle^{k_i} - \langle v_j^i, h\rangle^{k_i}\right| < \epsilon/2$$

infinitely often. Since $h = e_1$, we know from Corollary 5.2.6 that $\|(\log n)^{-1/2}P_N(S_nx)-h\|_H < \delta$ infinitely often, for arbitrary $\delta > 0$. By choosing δ small enough and noting that $\langle v^i_j, \cdot\rangle$ is a *continuous* function on $span\{e_i\}_{i=1}^N$ (because there are only finitely many variables), the claim (5.10) immediately follows, and thus (5.8) is proved.

Now the only thing left to show is that the set of limit points of the given sequence cannot contain points outside of the set K. This will be done in the following lemma. □

Lemma 5.2.18. *In the setting of Theorem 5.2.17, let*

$$a_n(x) := \big((2\log n)^{-k_1/2}T^1(S_nx),\ \dots, (2\log n)^{-k_m/2}T^m(S_nx)\big).$$

For a.e $x \in X$ one has that $dist(a_n(x), K) < \epsilon$ for all but finitely many n.

Proof. Note that if our abstract Wiener space (X, H, μ) is finite dimensional, then the statement is straightforward, since T and T_{hom} are of the form (5.11) and (5.12) respectively, and since all of the relevant quantities are continuous functions.

Now we move to the infinite-dimensional case. Suppose for contradiction the claim was false. Let U denote a neighborhood of size ϵ around K. Then since S is mixing (hence ergodic) and since the event "$dist(U^c, a_n(x)) < \epsilon/5$ infinitely often" is shift invariant, it follows that it actually occurs with probability 1. By the same argument used in deriving (5.9), we can choose N so large that $\sum_{n=1}^m(2\log n)^{-k_i/2}\|T^i(S_nx) - T^i_N(S_nx)\|_{Y_i} < \epsilon/5$ for all but finitely many n almost surely, and moreover by Lemma 5.2.15 we can ensure (by making N possibly larger) that $\sum_{n=1}^m \sup_{\|h\|\le 1}\|T^i_{hom}(h) - (T^i_N)_{hom}(h)\|_{Y_i} < \epsilon/5$. By the latter bound and the definition of U it is clear that

$$dist\bigg(\ U^c\ ,\ \big((T^1_{hom})_N(h), ..., (T^m_N)_{hom}(h)\big)\ \bigg) > 4\epsilon/5$$

for any h such that $\|h\|_H \le 1$. On the other hand the former bound and our shift-invariant event of full probability guarantees that $dist(U^c, a_n^N(x)) < 2\epsilon/5$ infinitely often (for a.e. x), where

$$a_n^N(x) := \big((2\log n)^{-k_1/2} T_N^1(S_n x), ..., (2\log n)^{-k_m/2} T_N^m(S_n x)\big).$$

The preceding two sentences imply (by finite dimensionality) that the sequence $(a_n^N(x))_{n\geq 1}$ contains a limit point outside of the set

$$\{((T_N^1)_{hom}(h), ..., (T_N^m)_{hom}(h)) : \|h\|_H \leq 1\},$$

since the distance of $a_n^N(x)$ to that set must be greater than $2\epsilon/5$ infinitely often. This contradicts the finite dimensional version of the statement that the set of limit points must be contained in K, which is impossible as noted earlier. $\square$

Next we formulate a continuous-time version of the above results. If Y is a Banach space, we denote by $C([0,1], Y)$ the space of continuous maps from $[0,1] \to Y$, equipped with the Banach space norm $\|\gamma\|_{C([0,1],Y)} := \sup_{t\in[0,1]} \|\gamma(t)\|_Y$. For $t \in [0,1]$ we define $\pi_t : C([0,1], Y) \to Y$ by sending $\gamma \mapsto \gamma(t)$.

Theorem 5.2.19. *Let (X, H, μ) be an abstract Wiener space, let $(S_t)_{t\geq 0}$ be a family of Borel-measurable a.e. linear maps from $X \to X$ which are measure-preserving and mixing, and let $T^i : X \to Y_i$ be homogeneous of degree k_i for $1 \leq i \leq m$. Suppose that there exist strongly continuous semigroups $(G_t^i)_{t\geq 0}$ of operators from $Y_i \to Y_i$ for $1 \leq i \leq m$ with the property that*

$$T \circ S_t = G_t^i \circ T, \qquad \mu\text{-a.e.} \qquad \text{for all } t \geq 0. \tag{5.13}$$

Then almost surely, the set of cluster points at infinity of the random set

$$\big\{\big((2\log t)^{-k_1/2} T^1(S_t x),\ \ldots, (2\log t)^{-k_m/2} T^m(S_t x)\big) : t \in \mathbb{Q} \cap [0,\infty)\big\}$$

is equal to the compact set

$$K := \{(T^1_{hom}(h), ..., T^m_{hom}(h)) : h \in B(H)\}.$$

Note that we impose no semigroup condition on (S_t) itself. This is because we do not need to, though in practice (S_t) will usually be a strongly continuous semigroup on X. The reason we are intersecting with $\mathbb{Q}$ is purely technical: we need to ensure that the given event is actually Borel-measurable. Also note that this theorem implies Proposition 5.1.2: set $k = 1$, $m = 1$, $Y = X$, let T^1 be the identity on X, and let $G^1_t = S_t$.

Proof. First we claim that that $(G^i_t(T^i(\xi)))_{t\in[0,1]}$ is homogeneous variable of order k_i taking values in the space $C([0,1], Y_i)$. To prove this, note that if Y, Z are Banach spaces, if $T : X \to Y$ is homogeneous of order k, and if $A : Y \to Z$ is a bounded linear map, then $A \circ T$ is also homogeneous of order k. We simply apply this to the case where $Y = Y_i$, $Z = C([0,1], Y_i)$ and $A : Y_i \to C([0,1], Y_i)$ sends a point y to $(G^i_t y)_{t\in[0,1]}$. This linear map is bounded by the uniform boundedness principle.

Next, note that the set of cluster points must contain K by Theorem 5.2.17. Thus we just need to show it contains no other points. Note that by (5.13), we can instead equivalently consider the random set

$$\big\{\big((2\log t)^{-k_1/2} G^1_t(T^1(x)),\ \ldots, (2\log t)^{-k_m/2} G^m_t(T^m(x))\big) : t \in [0,\infty)\big\}.$$

The argument that this set contains no cluster points outside of K is very similar to that of Lemma 5.2.11. More precisely, we show that there exists deterministic functions $C_i : [0,1] \to \mathbb{R}_+$ such that $C_i(\rho) \to 0$ as $\rho \downarrow 0$ and such that

$$\mu\Bigg(\Bigg\{x \in X : \limsup_{n\to\infty} \frac{\sup_{t\in[n\rho,(n+1)\rho]} \|G^i_t(T^i(x)) - G^i_{n\rho}(T^i(x))\|_{Y_i}}{(\log n)^{k_i/2}} \le C_i(\rho)\Bigg\}\Bigg) = 1$$

for all $\rho \in [0,1]$ and $1 \le i \le m$. In particular $(G^i_t(T^i(x)))_{t\ge 0}$ has the same set of cluster points as $(G_N(T^i(x)))_{N\in\mathbb{N}}$ for x in a set of full μ-measure.

To prove the above claim, one uses precisely the same arguments as we did in the proof of Lemma 5.2.11. Namely one defines $C_i(\rho) := C\mathbb{E}[\sup_{t\in[0,\rho]} \|G^i_t(T^i(\xi)) - T^i(\xi)\|^2_{Y_i}]$, where ξ is sampled from μ and $C > 0$ is to be determined later. Then one uses the fact that $(G^i_t(T^i(\xi)))_{t\in[0,1]}$ is homogeneous variable of order k_i taking values in the space $C([0,1], Y_i)$, by the discussion above. Finally one uses the associated tail bounds for such homogeneous variables as given in Proposition 5.2.13, and concludes using Borel-Cantelli. □

One can also formulate the preceding theorem in multiplicative form, as we did in the introduction.

5.2.4 Pushforward or "contraction principle" for Strassen's Law

Next we derive a corollary that will be used in deriving Strassen's Law for singular semilinear SPDEs later. The following can be viewed as a sort of "contraction principle" for Strassen's law.

Corollary 5.2.20. *Let (X, H, μ) be an abstract Wiener space, and let $T^i : X \to Y_i$ be homogeneous of order k_i for $1 \le i \le m$, where $m \in \mathbb{N}$. Let Z be a Banach space, and let $\mathcal{M} \subset Y_1 \times \cdots Y_m$ be a closed subset, such that for all $\delta > 0$ one has*

$$\mu(\{x \in X : (\delta^{k_1} T^1(x), ..., \delta^{k_m} T^m(x)) \in \mathcal{M}\}) = 1. \tag{5.14}$$

Let $\Phi : \mathcal{M} \to Z$ be continuous (possibly nonlinear). Then the compact set

$$K := \{(T^1_{hom}(h), ..., T^m_{hom}(h)) : h \in B(H)\}$$

is necessarily contained in $\mathcal{M}$, and moreover the set of cluster points at infinity of the random set

$$\big\{\Phi\big((2\log n)^{-k_1/2} T^1(S_n x),\ \ldots, (2\log n)^{-k_m/2} T^m(S_n x)\big) : n \in \mathbb{N}\big\}$$

is almost surely equal to $\Phi(K)$.

Proof. Note that if $(a_1, ..., a_m) \in K$ then by Theorem 5.2.17 there exists a subset $E \subset X$ of full measure such that for each $x \in E$, then $\big((2\log n)^{-k_1/2}T^1(S_n x),\ \ldots, (2\log n)^{-k_m/2}T^m(S_n x)\big)$ converges to $(a_1, ..., a_m)$ along some (x-dependent) subsequence. Since S_n are measure-preserving it holds by (5.14) that

$$\big((2\log n)^{-k_1/2}T^1(S_n x),\ \ldots, (2\log n)^{-k_m/2}T^m(S_n x)\big) \in \mathcal{M}$$

for a.e. $x \in X$. Consequently $(a_1, ..., a_m)$ is a limit point of $\mathcal{M}$ and thus belongs to $\mathcal{M}$ (since M is closed). This implies that K is contained in $\mathcal{M}$.

Now we prove that the limit set is necessarily $\Phi(K)$. The fact that any point $z \in \Phi(K)$ is a limit point is due to the fact that Φ is continuous and any point of K is a limit point of the sequence

$$\big((2\log n)^{-k_1/2}T^1(S_n x),\ \ldots, (2\log n)^{-k_m/2}T^m(S_n x)\big)$$

as $n \to \infty$ (see Theorem 5.2.17).

Now we need to prove that points outside of $\Phi(K)$ are not limit points. Suppose $z \notin \Phi(K)$. The latter set is closed so we may choose $\epsilon > 0$ such that $\|z - b\|_Z > \epsilon$ for all $b \in \Phi(K)$. Choose $\delta > 0$ so that $dist(\Phi(a), \Phi(K)) < \epsilon$ whenever $dist(a, K) < \delta$ (this δ exists by compactness of K). (need to specify what is $dist$) We choose points $a_1, ..., a_N$ so that $B(a_i, \epsilon)$ form an open cover of $\Phi(K)$, then consider the open cover $U_i := \Phi^{-1}(B(a_i, \epsilon))$ of K, then let U denote the union of the U_i, and take $\delta := \min_{x \in K} dist(x, U^c) > 0$). Letting

$$a_n(x) := \big((2\log n)^{-k_1/2}T^1(S_n x),\ \ldots, (2\log n)^{-k_m/2}T^m(S_n x)\big),$$

by Lemma 5.2.18 we know that for a.e $x \in X$ that $dist(a_n(x), K) < \delta$ for all but finitely many n

and therefore z is not a limit point of $\Phi(a_n(x))$. $\square$

The next corollary deals with the continuous setting.

Corollary 5.2.21. *Let (X, H, μ) be an abstract Wiener space, let $(S_t)_{t\geq 0}$ be a family of Borel-measurable a.e. linear maps from $X \to X$ which are measure-preserving and mixing, and let $T^i : X \to Y_i$ be homogeneous of degree k_i for $1 \leq i \leq m$. Suppose that there exist strongly continuous semigroups $(G^i_t)_{t\geq 0}$ of operators from $Y_i \to Y_i$ for $1 \leq i \leq m$ with the property that*

$$T^i \circ S_t = G^i_t \circ T^i, \qquad \mu\text{-a.e.} \qquad \text{for all } t \geq 0. \tag{5.15}$$

Let Z be a Banach space, and let $\mathcal{M} \subset Y_1 \times \cdots Y_m$ be a closed subset, such that for all $\delta > 0$

$$\mu(\{x \in X : (\delta^{k_1} T^1(x), ..., \delta^{k_m} T^m(x)) \in \mathcal{M}\}) = 1.$$

Let $\Phi : \mathcal{M} \to Z$ be continuous (possibly nonlinear). Then the compact set

$$K := \{(T^1_{hom}(h), ..., T^m_{hom}(h)) : h \in B(H)\}$$

is necessarily contained in $\mathcal{M}$, and moreover the set of cluster points at infinity of the random set

$$\big\{\Phi\big((2\log t)^{-k_1/2} T^1(S_t x),\ \ldots, (2\log t)^{-k_m/2} T^m(S_t x)\big) : t \in \mathbb{Q} \cap [0, \infty)\big\}$$

is almost surely equal to $\Phi(K)$.

Note that the previous two corollaries are the most general version of the Strassen's Law that we have stated so far (e.g. set $Z = Y_1 \times \cdots Y_m = \mathcal{M}$ and let Φ be the identity). Note also that Corollary 5.2.21 will be used in multiplicative form later (this is precisely Theorem 5.1.3 from the introduction), not the additive form stated above.

5.3 Technical lemmas for Gaussian measures on spaces of distributions

Recall that the goal of this work is to give a broad collection of examples of Strassen's Law for stochastic processes, cluminating with SPDEs. Thus we devote a section to some basic lemmas which will be useful in this regard.

Our first lemma is a result which roughly says that if X is a Banach space of distributions that is "nested" in between $\mathcal{S}(\mathbb{R}^d)$ and $\mathcal{S}'(\mathbb{R}^d)$ in such a way that X densely contains $\mathcal{S}(\mathbb{R}^d)$, then its dual space also admits a natural "nesting" between $\mathcal{S}(\mathbb{R}^d)$ and $\mathcal{S}'(\mathbb{R}^d)$. The dual space may no longer densely contain $\mathcal{S}(\mathbb{R}^d)$, but in the weak* topology it still does. We will use $(\cdot,\cdot)$ to denote the natural pairing between $\mathcal{S}(\mathbb{R}^d)$ and $\mathcal{S}'(\mathbb{R}^d)$.

Let $\mathcal{S}(\mathbb{R}^d)$ be the set of Schwartz functions on $\mathbb{R}^d$ and let $\mathcal{S}'(\mathbb{R}^d)$ be the set of tempered distribution.

Lemma 5.3.1. *Suppose X is a Banach space of tempered distributions on $\mathbb{R}^d$ such that one has continuous inclusion maps $\mathcal{S}(\mathbb{R}^d) \hookrightarrow X \hookrightarrow \mathcal{S}'(\mathbb{R}^d)$. Suppose that the image of the first embedding is dense with respect to the norm topology on X. Then there exists a Banach space X^{du} of distributions on $\mathbb{R}^d$ such that one has continuous inclusions $\mathcal{S}(\mathbb{R}^d) \hookrightarrow X^{du} \hookrightarrow \mathcal{S}'(\mathbb{R}^d)$, and furthermore there exists a "canonical" bilinear pairing $B : X \times X^{du} \to \mathbb{R}$ with the following properties:*

- $|B(x,f)| \le \|x\|_X \|f\|_{X^{du}}$.
- *The map from $X^{du} \to X^*$ given by $f \mapsto B(\cdot, f)$ is an isomorphism and a linear isometry.*
- $B(\phi, f) = (f, \phi)$ *and* $B(x,\phi) = (x,\phi)$ *for all $x \in X$, all $f \in X^{du}$, and all $\phi \in \mathcal{S}(\mathbb{R}^d)$.*

*The image of the first embedding $\mathcal{S}(\mathbb{R}^d) \hookrightarrow X^{du}$ may **not** be dense in the norm topology of X^{du}, but it is always dense with respect to the weak* topology on X^{du}.*

Before proving the lemma, let's give a few examples. If $X = L^p(\mathbb{R}^d)$ with $1 \le p < \infty$, then $X^{du} = L^q(\mathbb{R}^d)$ with $\frac{1}{p} + \frac{1}{q} = 1$. If $X = C_0(\mathbb{R}^d)$ then X^{du} consists of finite signed Borel measures on $\mathbb{R}^d$ equipped with total variation norm. Note in this case that the norm-closure of $\mathcal{S}(\mathbb{R}^d)$ in

X^{du} is $L^1(\mathbb{R}^d)$, which is a proper closed subspace of X^{du}, but still dense with respect to weak (i.e. Prohorov) convergence of measures. Further examples are given by Sobolev spaces $X = H^s(\mathbb{R}^d)$ and $X^{du} = H^{-s}(\mathbb{R}^d)$ with $s \in \mathbb{R}$. A rich class of examples is given by more general Sobolev spaces and Besov spaces (of exponent less than infinity) including those defined with weight functions. Note that the assumptions in the lemma automatically imply separability of X (because $\mathcal{S}(\mathbb{R}^d)$ is second countable and completely metrizable), though X^{du} may not be separable in its norm topology (e.g. take X to be $L^1(\mathbb{R})$ or $C_0(\mathbb{R})$).

Proof. Note that we have a map $\mathcal{G}$ from $X^* \to \mathcal{S}'(\mathbb{R}^d)$ given by restriction to the Schwartz class, i.e., $f \overset{\mathcal{G}}{\mapsto} f|_{\mathcal{S}(\mathbb{R}^d)}$. The map clearly defines a continuous linear operator from $X^* \to \mathcal{S}'(\mathbb{R}^d)$. We define X^{du} to be the image of X^* under $\mathcal{G}$. We also note that $\mathcal{G}$ is injective since $\mathcal{S}(\mathbb{R}^d)$ is norm-dense in X. We thus define the norm on X^{du} by $\|\mathcal{G}f\|_{X^{du}} = \|f\|_{X^*}$. Then clearly X^{du} is a Banach space that is isometric to X^* (via $\mathcal{G}$) and the inclusion $X^{du} \hookrightarrow \mathcal{S}'(\mathbb{R}^d)$ is clearly continuous.

We now claim that $\mathcal{S}(\mathbb{R}^d)$ is contained in X^{du}. To prove this we need to check that if $\phi \in \mathcal{S}(\mathbb{R}^d)$ then the map a_ϕ from $X \to \mathbb{R}$ given by $x \mapsto (x, \phi)$ is continuous. This is clear because if $x_n \to x$ in X, then $x_n \to x$ in $\mathcal{S}'(\mathbb{R}^d)$ so that $(x_n, \phi) \to (x, \phi)$. Next, we note that $\phi = \mathcal{G}a_\phi \in X^{du}$, proving the claim. By the closed graph theorem, it follows that the inclusion map $\mathcal{S}(\mathbb{R}^d) \hookrightarrow X^{du}$ is automatically continuous.

Now we construct the bilinear map B. For this, we simply define $B(x, \mathcal{G}f) := f(x)$ whenever $x \in X$ and $f \in X^*$. Clearly $|B(x, \mathcal{G}f)| \leq \|f\|_{X^*}\|x\|_X = \|\mathcal{G}f\|_{X^{du}}\|x\|_X$, so that B is bounded. Note that if $\phi \in \mathcal{S}(\mathbb{R}^d)$ then $B(\phi, \mathcal{G}f) = f(\phi) = (f|_{\mathcal{S}(\mathbb{R}^d)}, \phi) = (\mathcal{G}f, \phi)$, as desired. Also $B(x, \phi) = B(x, \mathcal{G}a_\phi) = a_\phi(x) = (x, \phi)$, completing the proof. The map $f \mapsto B(\cdot, f)$ is an isometry from $X^{du} \to X^*$ because it is inverse to the isometry $\mathcal{G}$.

Finally we need to show that $\mathcal{S}(\mathbb{R}^d)$ is weak* dense in X^{du}. This follows immediately from the fact that $\mathcal{S}(\mathbb{R}^d)$ is a total set in X^{du}, i.e., it separates points of X (which is clear because it separates points of the larger space $\mathcal{S}'$). For the equivalence of totality and weak* density, see [63, p. 439, Exercise 41]. $\square$

A corollary of this lemma is that if we want to check that a given Hilbert space H is the Cameron Martin space of a Gaussian measure μ on some separable Banach space X of distributions as above, then it suffices to check the action of the covariance function only on smooth test functions in the Schwartz class, as opposed to the entire dual space of X.

Lemma 5.3.2. *Let H be a Hilbert space and let X be a Banach space such that one has continuous inclusions $\mathcal{S}(\mathbb{R}^d) \hookrightarrow H \hookrightarrow \mathcal{S}'(\mathbb{R}^d)$ and $\mathcal{S}(\mathbb{R}^d) \hookrightarrow X \hookrightarrow \mathcal{S}'(\mathbb{R}^d)$. Suppose that $\mathcal{S}(\mathbb{R}^d)$ is dense in H and in X. Let μ be a Gaussian measure on X. If*

$$\int_X (x,\phi)(x,\psi)\mu(dx) = \langle \phi, \psi \rangle_H \tag{5.16}$$

for all $\phi, \psi \in \mathcal{S}(\mathbb{R}^d)$, then μ has Cameron Martin space H^{du}.

For instance, a simple intuitive corollary of this lemma is that regardless of the choice of Banach space X taken to contain a simple Brownian motion, it always has Cameron martin space H_0^1.

Proof. The Cameron Martin norm may be defined for $h \in X$ by

$$\|h\|_{CM} = \sup\{f(h) : f \in X^*, \int_X f^2 d\mu \leq 1\}.$$

By weak* density of $\mathcal{S}(\mathbb{R}^d)$ in X^{du} the supremum on the right is the same as

$$\sup\{(h,\phi) : \phi \in \mathcal{S}(\mathbb{R}^d), \int_X (x,\phi)^2 \mu(dx) \leq 1\},$$

which by our assumption (5.16) is the same as

$$\sup\{(h,\phi) : \phi \in \mathcal{S}(\mathbb{R}^d), \|\phi\|_H^2 \leq 1\},$$

which by definition (and density of $\mathcal{S}$ in H) equals the operator norm of the linear functional on H given by $(\cdot, h)$, which by construction equals the conjugate norm $\|h\|_{H^{du}}$ (to be understood as $+\infty$ if $h \notin H^{du}$).

So far this argument shows that the Cameron-Martin space of μ equals $X \cap H^{du}$ (with the norm of H^{du}), which is therefore closed in H^{du}. To finish the proof we need to show that H^{du} does not contain any vectors outside of X. Assume for contradiction that such a vector does exist; then $X \cap H^{du}$ would have a nonzero orthogonal complement with respect to the inner product of H^{du}. Take some nonzero bounded linear functional $u : H^{du} \to \mathbb{R}$ which vanishes on the closed subspace $X \cap H^{du}$. Since H is reflexive, every linear functional on H^{du} is represented as $B(f, \cdot)$ for some $f \in H$ where B is the bilinear form constructed in the previous lemma. Thus write $u = B(f, \cdot)$ for some $f \in H$. Since $X \cap H^{du}$ contains $\mathcal{S}$ we see that $(f, \phi) = B(f, \phi) = 0$ for all $\phi \in \mathcal{S}$. This means that $f = 0$ so that $u = 0$, a contradiction. $\square$

Remark 5.3.3. *Often in cases of interest the condition that H densely contains $\mathcal{S}(\mathbb{R}^d)$ fails, for examples in linear SPDEs where we want to impose a initial or boundary condition (see example 5.4.6 below). To resolve this one may define for an open set $U \subset \mathbb{R}^n$ the class $\mathcal{S}(U)$ of Schwartz functions on $\mathbb{R}^d$ vanishing on the complement of U. This is a Frechet space under the same family of seminorms (in fact it is a closed subspace of $\mathcal{S}(\mathbb{R}^d)$), and one can define its continuous dual $\mathcal{S}'(U)$. Moreover $\mathcal{S}(U)$ embeds into its dual using the L^2 pairing as in the case of the full space. Then all of the above lemmas still hold with $\mathbb{R}^d$ replaced by U. Usually we will take U to be $\mathbb{R}_+ \times \mathbb{R}^d$ in examples below, e.g. Example 5.4.6.*

Next we identify conditions for a semigroup S_t of operators on a Banach space X embedded in $\mathcal{S}'$ to be strongly continuous.

Lemma 5.3.4. *Suppose X is a Banach space of distributions on $\mathbb{R}^d$ such that one has continuous inclusion maps $\mathcal{S}(\mathbb{R}^d) \hookrightarrow X \hookrightarrow \mathcal{S}'(\mathbb{R}^d)$ and such that $\mathcal{S}(\mathbb{R}^d)$ is dense in X. Let $S_t : X \to X$ be a semigroup of bounded operators for $t \geq 0$ such that*

- *each S_t maps $\mathcal{S}(\mathbb{R}^d)$ to itself.*

- $S_t\phi \overset{\mathcal{S}(\mathbb{R}^d)}{\to} \phi$ *as* $t \to 0$ *for all* $\phi \in \mathcal{S}(\mathbb{R}^d)$.

- *such that* $\|S_t\|_{X\to X} \le Ce^{at}$ *for some* $C, a > 0$.

Then S_t is a strongly continuous semigroup on X.

The proof follows immediately from Exercise 4.2 in [85].

5.4 Examples and applications to SPDEs

With the above lemmas in place, we now move onto the core examples.

Example 5.4.1 (Brownian motion)**.** *Consider the Gaussian Banach space (X, H, μ) where $X = C[0,1]$, $H = H_0^1([0,1])$ and μ is the Wiener measure. For $w \in C[0,1]$, define $R_\epsilon : C[0,1] \to C[0,1]$ by $w \mapsto \epsilon^{-1}w(\epsilon^2\bullet)$. It is standard to check that $\{R_\epsilon\}_{\epsilon\in(0,1]}$ satisfies the four conditions stated after Proposition 5.1.2. Thus with probability 1, the set of cluster points as $\epsilon \to 0$ of $\{(\epsilon\log\log(1/\epsilon))^{-\frac{1}{2}}w(\epsilon\bullet)\}_{\epsilon\in(0,1]}$ is equal to $\{x \in H_0^1[0,1] : \frac{1}{2}\int_0^1 \dot{x}(s)^2ds \le 1\}$.*

But we can say much more than this. We claim that the set of limit points of $\epsilon^{-1}w(\epsilon^2\bullet)$ equals the unit ball of $H_0^1[0,1]$ not only with respect to the uniform norm, but also with respect to all of the Hölder seminorms up to (but not including) $1/2$. To prove this, we would like to set X to be the Hölder space of exponent $\alpha < 1/2$. But this space is not separable, so Proposition 5.1.2 is not applicable. However, the Hölder space of exponent α embeds into the fractional Sobolev space $X := W^{\alpha-\gamma,p}$ for any $\gamma > 0$ and $p > 1$, and this space is indeed separable and satisfies the conditions stated after Proposition 5.1.2. Moreover, by the Sobolev embedding $W^{\alpha-\gamma,p}$ embeds into the Hölder space of exponent $\alpha - p^{-1} - 2\gamma$ for any $\gamma > 0$, so by making p very large, γ very small, and α close to $1/2$, we can approach Hölder spaces of exponent arbitrarily close to $1/2$. As an alternative to Sobolev embedding, one may also realize Brownian motion as a Gaussian measure on the closure of smooth functions with respect to a Holder seminorm, which will be a separable space. This recovers the main result of [13].

Example 5.4.2. *Note that the space* $X = \{f \in C[0,\infty) : \lim_{t\to\infty} \frac{|f(t)|}{t} = 0\}$ *with norm given by* $\sup_{t>0} \frac{|f(t)|}{1+t}$ *is a separable Banach space which supports Brownian motion* B *(for all time, not just the unit interval) almost surely, and moreover the Cameron martin space is given by* $H_0^1[0,\infty) = \{f : \int_0^\infty f'(s)^2 < \infty\}$. *Furthermore one checks that the family of operators* $R_\epsilon f(x) = \epsilon^{1/2} f(\epsilon^{-1}x)$ *is mixing with respect to this norm (note the large-time as opposed to small-time regime). Consequently one obtains the Strassen's Law for* $\epsilon^{1/2}B(\epsilon^{-1}t)/\sqrt{2\log\log(1/\epsilon)}$ *with respect to the norm of* X. *Note here that one could just as well take multidimensional Brownian motion or even Brownian motion* $B_Y(t)$ *taking values in some Banach space* Y *(defining the norm of* X *to be* $\sup_{t>0} \frac{\|f(t)\|_Y}{1+t}$*), in which case the Cameron-Martin space is given by* $H_0^1([0,\infty),K) := \{f \in C([0,1],Y) : \int_0^\infty \|f'(s)\|_K^2 ds < \infty\}$ *where* $K \subset Y$ *denotes the embedded Cameron-Martin space of the random variable* $B_Y(1)$. *This easily recovers the main result of [104] without any additional difficulty, see that paper for more details.*

Example 5.4.3 (Fractional Brownian motion (FBM))**.** *Consider the FBM* $(B^H(t), t \in [0,1])$ *with Hurst parameter* $H \in (0,1)$. *By [55], the Cameron Martin space of fractional Brownian motion is given by*

$$\mathcal{H}_H = \Big\{h : \exists\, \ell \in L^2([0,1],\mathbb{R}) \text{ s.t. } h(t) = \int_0^t K_H(t,s)\ell(s)ds, t \in [0,1]\Big\},$$

where K_H *may be viewed as the inverse kernel for the fractional Laplacian* $(-\Delta)^{H/2+1/4}$. *We set* $\|h\|_{\mathcal{H}_H} := \|\ell\|_{L^2}$ *(*ℓ *is uniquely determined by* h*). The expression of the kernel* K_H *is given by [55, Eq. (6)]. Note that* $K_H(\epsilon t, \epsilon s) = \epsilon^{H-\frac{1}{2}} K_H(t,s)$. *In addition, the map* $\ell \to \int_0^{\cdot} K_H(\cdot,s)\ell(s)ds$ *is an isometry between* $L^2([0,1],\mathbb{R})$ *and* $\mathcal{H}_H$. *To prove the Strassen's law for the FBM note for any path* $w \in \mathcal{H}_H$, *let* $w_\epsilon(\cdot) = \epsilon^{\frac{H}{2}} w(\epsilon^{-1}\cdot)1_{[0,\epsilon]}(\cdot)$ *and* $\ell_\epsilon(\cdot) = \epsilon^{\frac{1}{2}}\ell(\epsilon^{-1}\cdot)1_{[0,\epsilon]}(\cdot)$. *Then* $\|w_\epsilon\|_{\mathcal{H}_H} = \|\ell\|_{L^2}$ *and therefore,*

$$\lim_{\epsilon\to 0} \|w_\epsilon\|_{\mathcal{H}_H} = \lim_{\epsilon\to 0} \|\ell_\epsilon\|_{L^2} = 0.$$

The Strassen's law follows from (the multiplicative form of) Proposition 5.1.2.

Example 5.4.4. *Here is a simple example in the second chaos. Let* $X = C([0,1],\mathbb{R}^2)$*, take* $\mu =$

2d BM, and $H = \{(f_1, f_2) \in X : \int_0^1 (f_1'(t)^2 + f_2'(t)^2)dt < \infty\}$*. then let* $Y = C[0,1]$ *and consider the (discontinuous) map* $\psi : X \to Y$ *given by*

$$\psi(B_1, B_2) = \int_0^\bullet B_1(t)dB_2(t).$$

For $f \in Y$ *and* $(f_1, f_2) \in X$

$$Q_\epsilon f(t) := \epsilon^{-1} f(\epsilon t),$$

$$R_\epsilon(f_1, f_2)(t) := (\epsilon^{-1/2} f_1(\epsilon t), \epsilon^{-1/2} f_2(\epsilon t)).$$

Then it is clear that $\psi \circ R_\epsilon = Q_\epsilon \circ \psi$ *a.s., and thus one can obtain the LIL (as* $\epsilon \to 0$*) for the family of processes*

$$\left\{ \left((2\epsilon \log\log(1/\epsilon))^{-1} \int_0^{\epsilon t} B_1(s)dB_2(s) \right)_{t\in[0,1]} \right\}_{0<\epsilon\le 1/3}.$$

The compact limit set is easily checked to be $\{\int_0^\cdot f_1(s)f_2'(s)ds : \int_0^1 f_1'(s)^2 + f_2'(s)^2 \le 1\}$*. Note that one may strengthen the topology of* Y *to the closure of smooth functions with respect to the Holder norm of any exponent less than 1/2.*

Example 5.4.5 (Compositions of processes)**.** *Burdzy in [28] proved that if* B, W *are two independent two-sided Brownian motions, then*

$$\limsup_{t\to 0} \frac{B(W(t))}{t^{1/4}\log(\log(1/t))^{3/4}} = 2^{5/4}3^{-3/4}.$$

In [50], the authors extend this to prove a functional version of this theorem, namely that that the set of limit points of the family of random functions $Z^\epsilon : [-1,1] \to \mathbb{R}$ *defined by*

$$Z^\epsilon(t) := \epsilon^{-1/4}(\log\log(1/\epsilon))^{-3/4} B(W(\epsilon t)),$$

as $\epsilon \to 0$ *equals the compact set of functions given by*

$$Q := \{f \circ g : f, g \in C([-1,1]), f(0) = g(0) = 0, \int_{-1}^1 f'(t)^2 + g'(t)^2 \le 1\}.$$

Note that the composition is well-defined because the integral condition implies that g maps $[-1,1]$ to itself. We can easily recover this result using our framework as follows. Define $B^\epsilon(t) := (\epsilon \log\log(1/\epsilon))^{-1/2}B(\epsilon t)$ and define $W^\epsilon(t) = (\epsilon \log\log(1/\epsilon))^{-1/2}W(\epsilon t)$. Note that Z^ϵ has the same distribution as $B^\epsilon \circ W^\epsilon$. This is true not only for fixed ϵ, but also as a process in ϵ, so it suffices to study the limit points of $B^\epsilon \circ W^\epsilon$ as $\epsilon \to 0$. Note by Lemmas 5.2.11 and 5.2.18 that with probability 1, W^ϵ has image contained in $[-2,2]$ for small ϵ almost surely. Note that the map F from $C[-2,2] \times C([-1,1],[-2,2]) \to C[-1,1]$ defined by $(f,g) \mapsto f \circ g$ is continuous, thus by Theorem 5.2.21, the set of limit points of $B^\epsilon \circ W^\epsilon$ is given by the image of the set $K := \{(f,g) : \int_{-2}^2 f'(t)^2dt + \int_{-1}^1 g'(t)^2dt \leq 1\}$ under the continuous map F, which equals the set of $f \circ g$ such that $\int_{-2}^2 f'(t)^2dt + \int_{-1}^1 g'(t)^2dt \leq 1$. But this condition automatically implies that $|g(x)| \leq 1$ so that this set necessarily coincides with the set K defined above. Note that, as in the first example, we can straightforwardly obtain this result in stronger topologies of Holder-continuous functions.

We can also prove similar theorems for composition processes derived from higher chaoses. For instance in [123], the author proves that if W, B_1, B_2 are independent two-sided standard Brownian motions, and if one defines $Z(t) := A(W(t))$ where $A(t) := \frac{1}{2}\int_0^t B_2(s)dB_1(s) - B_1(s)dB_2(s)$, then

$$\limsup_{t\to 0} \frac{A(W(t))}{t^{1/2}(\log\log(1/t))^{3/2}} = (2/3)^{-3/2}/\pi.$$

We can obtain a functional version of this using essentially the same method as above. Define $Z^\epsilon(t) := \epsilon^{-1/2}(\log\log(1/\epsilon))^{-3/2}A(W(\epsilon t))$, and note that Z^ϵ is equal in law to $A^\epsilon \circ W^\epsilon$, where $A^\epsilon(t) := (\epsilon\log\log(1/\epsilon))^{-1}A(\epsilon t)$ and $W^\epsilon(t) = (\epsilon \log\log(1/\epsilon))^{-1/2}W(\epsilon t)$. Hence, mimicking the proof from above but using the higher chaos results in lieu of the first chaos version (see Example 5.4.4), we see that $A^\epsilon \circ W^\epsilon$ has the set of limit points given by $f \circ g$ where $f, g \in C[-1,1]$ and f is of the form $\frac{1}{2}\int_0^\bullet h_2h_1' - h_1h_2'$ where $\int_{-1}^1 g'(t)^2 + h_1'(t)^2 + h_2'(t)^2dt \leq 1$, and $g(0) = h_i(0) = 0$ for $i = 1, 2$.

Example 5.4.6. *Consider the stochastic heat equation*

$$\partial_t h(t,x) = \partial_x^2 h(t,x) + \xi(t,x), \quad h(0,x) = 0, \quad t \geq 0, x \in \mathbb{R}^d,$$

where ξ is a Gaussian space-time white noise.

We define the Frechet space $\mathcal{S}_0'(\mathbb{R}_+ \times \mathbb{R}^d)$ to be the set of all $f \in \mathcal{S}'(\mathbb{R} \times \mathbb{R}^d)$ such that $f(\phi) = 0$ for all $\phi \in \mathcal{S}(\mathbb{R} \times \mathbb{R}^d)$ such that ϕ is supported on $(-\infty, 0) \times \mathbb{R}^d$.

Suppose that Y is some separable Banach space which embeds continuously into $\mathcal{S}_0'(\mathbb{R}_+ \times \mathbb{R}^d)$, such that the law of h can be realized as a Gaussian measure on Y. When $d = 1$, an explicit example of such a space Y is the weighted space

$$Y := \Big\{ h \in C(\mathbb{R}_+ \times \mathbb{R}) : h(0,x) = 0, \ \limsup_{t+|x| \to \infty} \frac{h(t,x)}{t + |x|} = 0 \Big\} \tag{5.17}$$

and with norm given by $\|h\|_Y := \sup_{t,x} \frac{h(t,x)}{1+t+|x|}$. For $d > 1$ the equation cannot be realized as a continuous function, thus one has to use a space of generalized functions for Y. An example would be the closure of Schwartz functions in a weighted parabolic Besov-Holder space of negative exponent $\alpha := \frac{1}{2} - \frac{d}{2} - \kappa$ where $\kappa > 0$ is arbitrary. More specifically, one takes the closure of $\mathcal{S}(\mathbb{R}^{d+1})$ under the norm

$$\|\phi\|_Y := \sup_{\|\psi\|_{C^r} \leq 1} \ \sup_{\lambda \in (0,1], (t,x) \in \mathbb{R}_+ \times \mathbb{R}} \frac{\lambda^{-\alpha} |(\phi, \psi_{t,x}^\lambda)|}{w(t,x)}, \tag{5.18}$$

where $r = -\lfloor \alpha \rfloor$, $\psi_{t,x}^\lambda(s,y) = \lambda^{-d-2} \phi(\lambda^{-2}(t-s), \lambda^{-1}(x-y))$. Here w is some specific weight function, see for instance [33] for more information on these spaces and see Section 2 in particular for the proof that the solution may indeed be realized in these spaces.

We claim that the solution of the equation has Cameron Martin space given by

$$H := \{h \in \mathcal{S}_0'(\mathbb{R} \times \mathbb{R}^d) : \partial_t h - \partial_x^2 h \in L^2(\mathbb{R}_+ \times \mathbb{R}^d)\},$$

with norm given by $\|h\|_H := \|\partial_t h - \partial_x^2 h\|_{L^2(\mathbb{R}_+\times\mathbb{R}^d)}$. *To prove this, note that if* h *solves the equation, then* $h = K^s * \xi$ *where* $K^s(t,x) = \frac{1}{\sqrt{2\pi|t|}} e^{-x^2/2|t|} 1_{\{t\le 0\}}$ *and thus* $\mathbb{E}[(h,\phi)(h,\psi)] = \langle K^s * \phi, K^s * \psi\rangle_{L^2(\mathbb{R}_+\times\mathbb{R})}$ *where* $*$ *denotes convolution in both space and time. Since* K^s *is the kernel for the inverse operator of* $(\partial_t + \partial_x^2)$ *which is adjoint to* $\partial_t - \partial_x^2$, *it follows immediately that the by Lemma 5.3.2 that* H *is indeed the Cameron-Martin space.*

If we define $R_\epsilon h(x) = \epsilon^{\frac{d}{2}-1} h(\epsilon^2 t, \epsilon x)$, *(to be interpreted in the sense of distributions, i.e., integration against a test function). Then one easily verifies that* R_ϵ *sends* Y *boundedly to itself and satisfies all of the conditions of Lemma 5.3.4 for either choice* (5.17) *or* (5.18) *of* Y. *Consequently by Proposition 5.1.2 (or the discussion afterwards) one finds that the set* $\{(\log\log(1/\epsilon))^{-1/2} R_\epsilon h\}_{\epsilon\in(0,1]}$ *is precompact and that its set of limit points in* Y *as* $\epsilon \to 0$ *equals the unit ball of* H *as defined above.*

5.4.1 Strassen Law for Φ_2^4

In this subsection we finally show how the above results can be used to prove a "nonlinear" Strassen law for singular stochastic PDEs. The two specific examples we focus on are Φ_2^4 and KPZ, though our method is more general and can be used in much broader contexts. In the specific case of KPZ our result can be a manifestation of weak universality, as discussed in the introduction.

Definition 5.4.7 (weighted Holder space)**.** *Let* $\mathfrak{s} = (k_1, 1, ..., 1)$ *be a scaling of* $\mathbb{R}^d$, *that is a formal* $(d+1)$*-tuple. The Euclidean scaling is defined with* $k_1 = 1$ *and the parabolic scaling is defined with* $k_1 = 2$. *Let* $w : \mathbb{R}^{d+1} \to \mathbb{R}$ *be a weight function, that is, there exists* $C > 0$ *such that uniformly over* $x, y \in \mathbb{R}^d$

$$C^{-1} < w(x)/w(y) < C.$$

Define the weighted parabolic Holder space $C^{\alpha}_{\mathfrak{s},w}(\mathbb{R}^{d+1})$ to be the the closure of $\mathcal{S}(\mathbb{R}^{d+1})$ under the norm

$$\|\phi\| := \sup_{\|\psi\|_{C^r}\le 1} \sup_{\lambda\in(0,1]} \sup_{(t,x)\in\mathbb{R}\times\mathbb{R}^d} \frac{\lambda^{-\alpha}|(\phi,\psi^{\lambda}_{t,x})|}{w(t,x)}, \tag{5.19}$$

where $r=-\lfloor\alpha\rfloor$, $\psi^{\lambda}_{t,x}(s,y)=\lambda^{-d-k_1}\phi(\lambda^{-k_1}(t-s),\lambda^{-1}(x-y))$.

Following Hairer [83], we usually omit the subscript $\mathfrak{s}$ whenever we use the Euclidean scaling and we include it whenever we use the parabolic scaling.

Now let us explain what exactly is a solution for Φ^4_2, and how to obtain Strassen's Law for it. Based on seminal work of Da Prato and Debuscche [51], it was proved in [121] that if we fix a smooth even compactly supported function $\psi:\mathbb{R}^3\to\mathbb{R}$ which integrates to 1, and if we define $\xi^\delta=\xi*\psi_\delta$ where $\psi_\delta(t,x)=\delta^{-3}\psi(\delta^{-2}t,\delta^{-1}x)$ then for any choice of $h(0,x)\in C^{-\nu}_w(\mathbb{R}^2)$ and $\alpha>0$ the classical solutions of the equation

$$\partial_t h_\delta = \Delta h_\delta - (h_\delta^3 - \alpha^2 C_\delta h_\delta) + \alpha\xi^\delta$$

converge in $C([0,T],C^{-\nu}_w(\mathbb{R}^2))$ to a limiting object independent of the choice of mollifier, for a suitable choice of constant C_δ depending on the mollifier. We call this object the solution of the Φ^4_2 equation

$$\partial_t h = \Delta h - h^{:3:} + \alpha\xi.$$

Theorem 5.4.8. *For any $T>0$ the random set h^ϵ of solutions to*

$$\partial_t h^\epsilon(t,x) = \Delta h^\epsilon(t,x) - h^\epsilon(t,x)^{:3:} + (\log\log(1/\epsilon))^{-1/2}\epsilon^2\xi(\epsilon^2 t,\epsilon x)$$

with $h^\epsilon(0,x)=0$ is almost surely precompact in $C([0,T],C^{-\nu}_w(\mathbb{R}^2))$ with limit set equal to the closure in that space of the set of space-time Schwartz functions f such that

$$f(0,x)=0, \qquad \|\partial_t f - \Delta f + f^3\|_{L^2([0,T]\times\mathbb{R}^2)} \le 1.$$

To prove this, we consider the first, second, and third Wick powers of the 2+1 dimensional additive SHE. This is important because one can show that there exists a locally Lipchitz continuous and globally-defined solution map $\mathbf{S} : C\big([0,T], C_w^{-\nu}(\mathbb{R}^2)\big)^3 \to C_w^{-\nu}(\mathbb{R}^2)$ such that $\mathbf{S}$ sends this Wick-power triple to the renormalized solution of the ϕ^4_2 equation

$$\partial_t \phi = \Delta\phi - \phi^{:3:} + \xi,$$

where ξ is the same noise from which the wick powers of the linearized equation are derived, and $C_w^{-\nu}(\mathbb{R}^2)$ denotes an appropriately weighted Besov-Hölder space of negative exponent on $\mathbb{R}^2$. For the definition $\mathbf{S}$, see for instance Theorem 3.2 of [120] or equations (1.6) and (1.7) of the companion paper on which that theorem is based [121].

We use the letters z, a, b as a shorthand to denote space-time points $(t, x_1, x_2) \in \mathbb{R}_+ \times \mathbb{R}^2$, and we write $K(z) := \frac{1}{2\pi t} e^{-(x_1^2+x_2^2)/2t} 1_{\{t\geq 0\}}$. Given a space-time white noise ξ, these Wick Powers are defined by the following formulas:

$$Z(z) := \int_{\mathbb{R}_+\times\mathbb{R}^2} K(z-a)\xi(da),$$

$$Z^{:2:}(z) := \int_{(\mathbb{R}_+\times\mathbb{R}^2)^2} K(z-a)K(z-b)\xi(da)\xi(db),$$

$$Z^{:3:}(z) := \int_{(\mathbb{R}_+\times\mathbb{R}^2)^3} K(z-a)K(z-b)K(z-c)\xi(da)\xi(db)\xi(dc).$$

These objects are at-best distributions in $C\big([0,T], C_w^{-\nu}(\mathbb{R}^2)\big)$, they are not actually definable as pointwise functions and therefore the above formulas are somewhat nonsensical: one actually needs to interpret the above by integrating against a Schwartz test function in space-time. For instance in the case of $Z^{:2:}$, if ϕ is a smooth compactly supported test function then the quantity

$$(Z^{:2:}, \phi)_{L^2} := \int_{(\mathbb{R}_+\times\mathbb{R}^2)^2} \bigg[\int_{\mathbb{R}_+\times\mathbb{R}^2} \phi(z)K(z-a)K(z-b)dz\bigg]\xi(da)\xi(db),$$

will be a well-defined stochastic integral (i.e., the integrand in square brackets is in $L^2\big((\mathbb{R}_+ \times \mathbb{R}^2)^2\big)$) and therefore it defines a *bona fide* random variable on the probability space of ξ. Moreover it lies in the second homogeneous chaos for all ϕ.

With this setup in place, define a measure-preserving SCS $(R_\epsilon)_{\epsilon\in(0,1]}$ to act on ξ by the formula

$$R_\epsilon\xi(t,x_1,x_2) := \epsilon^2\xi(\epsilon^2 t, \epsilon x_1, \epsilon x_2),$$

which as usual needs to be interpreted by integrating against a test function. Letting $\psi_1(\xi) := Z$, $\psi_2(\xi) := Z^{:2:}$ and $\psi_3(\xi) := Z^{:3:}$ one can formally verify (and then make rigorous via test functions) that $Q^i_\epsilon \circ \psi_i = \psi_i \circ R_\epsilon$ a.s. for $1 \le i \le 3$, where

$$Q^i_\epsilon f(t,x_1,x_2) = f(\epsilon^2 t, \epsilon x_1, \epsilon x_2).$$

This uses the fact that for all $t > 0$ and $x_1, x_2 \in \mathbb{R}$ the (2+1)-dimensional heat kernel K satisfies $K(c^2t, cx_1, cx_2) = c^{-2}K(t, x_1, x_2)$ for any $c > 0$. Once we choose the appropriate Banach spaces for these chaoses to reside, it is fairly clear that for any $\delta > 0$ the random distribution $\mathbf{S}(\delta\psi_1(R_\epsilon\xi), \delta^2\psi_2(R_\epsilon\xi), \delta^3\psi_3(R_\epsilon\xi)$ will solve the equation

$$\partial_t h(t,x) = \Delta h(t,x) - h(t,x)^{:3:} + \delta\epsilon^2\xi(\epsilon^2 t, \epsilon x),$$

which is precisely the object we want to study (with $\delta = \log\log(1/\epsilon)^{-1/2}$). To make all of this precise, on needs to choose the appropriate separable Banach spaces on which $\xi, Z, Z^{:2:}, Z^{:3:}$ will lie such that the operator semigroups defined above are strongly continuous on these spaces. For this we can choose $X := C^{-5/2-\kappa}(\mathbb{R}_+ \times \mathbb{R})$ and $Y_k = C([0,T], C_w^{-\nu}(\mathbb{R}^2))$ and it is shown in [121] that these spaces do indeed contain these objects. The proof is done by using Lemma 9 in Section 5 of that paper. The fact that the operator semigroups are strongly continuous on these spaces follows from Lemma 5.3.4 (note that all of these spaces are *defined* in that reference as the closure of Schwartz functions under the appropriate norm, so Schwartz functions are trivially dense there).

It remains to be seen why the compact limit set is the one described in our theorem above. This boils down to the fact that $(Q^i)_{hom}(h) = (K * h)^i$ (where $*$ denotes space-time convolution and K is the heat kernel) for $i = 1, 2, 3$, and $h \in L^2(\mathbb{R}_+ \times \mathbb{R})$, and also the fact that if $f : \mathbb{R}_+ \times \mathbb{R}$ is any Schwartz function and if $\mathbf{S}$ denotes the solution map above, then $\mathbf{S}(K * f, (K * f)^2, (K * f)^3)$ is the classical solution of the equation $\partial_t v = \partial_x^2 v - v^3 + f$, with zero initial data, which is clear from the way that the solution map $\mathbf{S}$ is defined in Equations (1.6) and (1.7) of [121]. Thus the proof of Theorem 5.4.8 follows immediately from Corollary 5.2.21 (in multiplicative form) applied to the chaoses ψ_i composed with the solution map $\mathbf{S}$.

5.4.2 Strassen's Law for KPZ

Here we do the computation for some of the higher-order objects appearing in the solution map for the KPZ equation, using methods from [88]. Let ξ be a standard space-time white noise on $\mathbb{R}_+ \times \mathbb{R}$ and define the following distribution-valued chaoses for $z \in \mathbb{R}_+ \times \mathbb{R}$:

$$\begin{aligned}
\Pi_z^\xi \Xi(\psi) &:= \int_{\mathbb{R}_+ \times \mathbb{R}} \psi(w)\xi(dw), \\
\Pi_z^\xi \big[\Xi\mathcal{I}[\Xi]\big](\psi) &:= \int_{(\mathbb{R}_+ \times \mathbb{R})^2} \psi(w)(K(w-a) - K(z-a))\xi(da)\xi(dw), \\
\Pi_z^\xi \big[\Xi\mathcal{I}[\Xi\mathcal{I}[\Xi]]\big](\psi) &:= \int_{(\mathbb{R}_+ \times \mathbb{R})^3} \psi(w)(K(w-a) - K(z-a)) \\
&\qquad \cdot (K(a-b) - K(z-b))\xi(db)\xi(da)\xi(dw), \\
\Pi_z^\xi \big[\Xi\mathcal{I}[\Xi\mathcal{I}[\Xi\mathcal{I}[\Xi]]]\big](\psi) &:= \int_{(\mathbb{R}_+ \times \mathbb{R})^4} \psi(w)(K(w-a) - K(z-a))(K(a-b) - K(z-b)) \\
&\qquad \cdot \big(K(b-c) - K(z-c) - (t_z - t_c)\partial_t K(z-c) - (x_z - x_c)\partial_x K(z-c)\big) \\
&\qquad \xi(dc)\xi(db)\xi(da)\xi(dw),
\end{aligned}$$

$$\Pi_z^\xi \big[\Xi\mathcal{I}[X_1\Xi]\big](\psi) := \int_{(\mathbb{R}_+ \times \mathbb{R})^2} \psi(w)K(w-a)(x_z - x_a)\xi(da)\xi(dw)$$

where $z = (t_z, x_z), a = (t_a, x_a), b = (t_b, x_b)$, and $c = (t_c, x_c) \in \mathbb{R}_+ \times \mathbb{R}$. Now for $\tau \in \mathcal{W} := \{\Xi, \Xi\mathcal{I}[\Xi], \Xi\mathcal{I}[\Xi\mathcal{I}[\Xi]], \Xi\mathcal{I}[\Xi\mathcal{I}[\Xi\mathcal{I}[\Xi]]], \Xi\mathcal{I}[X_1\Xi]\}$, we define the order $|\tau|$ to be $-3/2-\kappa, -1-2\kappa, -1/2-3\kappa, -4\kappa, -2\kappa$ respectively, where $\kappa \in (0, 1/8)$ is arbitrary but fixed.

Given a Schwartz function $\Pi : (\mathbb{R} \times \mathbb{R})^2 \to \mathbb{R}$ and $-\alpha < 0$ one may define a norm

$$\|\Pi\| := \sup_{z \in \mathbb{R}^d} w(z)^{-1} \|\Pi(z, \cdot)\|_{C^{-\alpha}_{\mathfrak{s},w}},$$

where $C^{-\alpha}_{\mathfrak{s},w}$ is a weighted Holder space of exponent α and weight w defined in the previous subsection. One may then define E^{α}_{w} to be the closure of $\mathcal{S}\big((\mathbb{R} \times \mathbb{R})^2\big)$ under this norm.

[88] prove that the above symbols $\Pi\tau$ defined above are supported in the spaces $E^{|\tau|}_{w_\tau}$ for the appropriate choice of w_τ. [88] actually prove the following: there exists a certain nonlinear subspace $\mathcal{X}$ (called the space of "admissible models") contained in $\bigoplus_{\tau \in \mathcal{W}} E^{|\tau|}_{w_\tau}$ with the property that $(\Pi\tau)_{\tau \in \mathcal{W}}$ are supported on $\mathcal{X}$ and there exists a deterministic continuous map $\Phi : \mathcal{X} \to C^{\alpha}_{w}$ such that Φ sends $(\Pi\tau)_{\tau \in \mathcal{W}}$ to the Hopf-Cole solution of the KPZ equation (actually the SHE, but log is a continuous map on positive functions, so the result follows) driven by the same realization of the noise ξ that appears in the integrals defining $\Pi\tau$ above. In fact, [88] prove something stronger, namely if we replace ξ by $\delta\xi$ in the above chaoses, then Φ sends $(\Pi\tau)_{\tau \in \mathcal{W}}$ to the Hopf-Cole solution of KPZ driven by $\delta\xi$, for **any** $\delta > 0$. The remarkable thing is that the solution map Φ itself does not depend on δ. This is elaborated in Appendix 2.

Here is another important property possessed by Φ. For $\tau \in \mathcal{W}$, let $\psi_\tau(z; \xi) := \Pi^{\xi}_{z}\tau$. as defined above. Then it is clear that $(\psi_\tau)_{hom}(z; f)$ is given by the expression for $\Pi^{\xi}_{z}\tau$ but with $\xi(dw), \xi(da), \xi(db), \xi(dc)$ replaced by $f(w)dw, f(a)da, f(b)db, f(c)dc$, respectively (so these are now classical integrals, not stochastic Itô integrals as above). Here f can be an arbitrary smooth function. Then using the method of [91], it is easy to show that the map Φ described above above send $(\psi_\tau)_{hom}(\cdot; f)$ to the classical solution of the deterministic equation $\partial_t h = \partial_x^2 h + \frac{1}{2}(\partial_x h)^2 + f$, with initial data zero. This is also explained in Appendix 2, see the discussion of canonical models.

With this in mind, let us now prove Strassen's Law for KPZ (Theorem 5.1.4) by defining the appropriate semigroups on the spaces E_τ. The semigroup acting on the noise is $R_\epsilon f(t,x) = \epsilon^{3/2} f(\epsilon^2 t, \epsilon x)$. Then by using the fact that each of the functions $K(t,x), t\cdot\partial_t K(t,x), x\cdot\partial_x K(t,x)$ are invariant under the scaling $cf(c^2t, cx)$ one easily checks that the follow semigroups Q^τ on E_w^τ will satisfy the commutation relation $Q_\epsilon^\tau \circ \psi_\tau = \psi_\tau \circ R_\epsilon$:

$$Q_\epsilon^{\Xi} f(t,x,s,y) = \epsilon^{3/2} f(\epsilon^2 t, \epsilon x, \epsilon^2 s, \epsilon y),$$

$$Q_\epsilon^{\Xi\mathcal{I}[\Xi]} f(t,x,s,y) = \epsilon f(\epsilon^2 t, \epsilon x, \epsilon^2 s, \epsilon y),$$

$$Q_\epsilon^{\Xi\mathcal{I}[\Xi\mathcal{I}[\Xi]]} f(t,x,s,y) = \epsilon^{1/2} f(\epsilon^2 t, \epsilon x, \epsilon^2 s, \epsilon y),$$

$$Q_\epsilon^{\Xi\mathcal{I}[\Xi\mathcal{I}[\Xi\mathcal{I}[\Xi]]]} f(t,x,s,y) = f(\epsilon^2 t, \epsilon x, \epsilon^2 s, \epsilon y),$$

$$Q_\epsilon^{\Xi\mathcal{I}[X_1\Xi]} f(t,x,s,y) = f(\epsilon^2 t, \epsilon x, \epsilon^2 s, \epsilon y)$$

where the (t,x) stands for the z variable and the (s,y) stands for the w variable in the integrals above (note that the f appearing here will be distributional in the latter variable and thus, as usual, the rescaling of coordinates needs to be interpreted in an integrated sense). In other words one simply has $Q_\epsilon^\tau f(t,x,s,y) = \epsilon^{-|\tau|_0} f(\epsilon^2 t, \epsilon x, \epsilon^2 s, \epsilon y)$ where $|\tau|_0$ denotes the order of τ without the κ term. Furthermore they are strongly continuous by Lemma 5.3.4.

5.4.3 "Noise-smoothing" and non-measure-preserving systems

Recall that we defined solutions to the Φ_2^4 equation in terms of the limit of a mollification procedure. Thus one may want to consider proving Strassen's law for a family of processes that is being diffusively scaled while being simultaneously approximated by mollified noise, similar to what was done in [91].

Specifically consider $\delta = \delta(\epsilon)$ and look at the family of processes h^ϵ defined similarly to the Φ^4_2 example above but with smooth noises at scale $\delta(\epsilon)$, i.e., letting $L_2 := \log\log$ we want to consider

$$\partial_t h^\epsilon = \Delta h^\epsilon + (C_{\delta(\epsilon)/\epsilon} L_2(1/\epsilon)^{-1} h^\epsilon - (h^\epsilon)^3) + (L_2(1/\epsilon))^{-1/2} \epsilon^{3/2} \xi_{\delta(\epsilon)}(\epsilon^2 t, \epsilon x),$$

and the goal is to try to find the set of limit points of the sequence of functions given by

$$(L_2(1/\epsilon))^{-1/2} h^\epsilon(\epsilon^2 t, \epsilon x).$$

The main result of this section will be that the limit points in $C([0,T], C_w^{-\nu}(\mathbb{R}^2))$ coincide with those found previously if $\delta(\epsilon) = \epsilon^{1+u}$ for some $u > 0$, they depend on the mollifier if $\delta(\epsilon) = \epsilon$, and it is a trivial set consisting of a one-dimensional family of functions if $\delta(\epsilon) = \epsilon^{1-u}$.

To prove this consider a multiplicative SCS $(R_\epsilon)_{\epsilon \in (0,1]}$ on a Banach space X. Writing $S_t := R_{e^{-t}}$ note that $\|S_{t+s}\| \leq \|S_t\| \|S_s\|$ and thus by Fekete's subadditive lemma we know that the following quantity exists as a real number

$$\kappa(R) := \lim_{t \to \infty} t^{-1} \log \|S_t\| = \inf_{t > 0} t^{-1} \log \|S_t\|.$$

In other words $\|R_\epsilon\|_{X \to X} \leq \epsilon^{-\kappa(R)}$ and $\kappa(R)$ is the optimal such exponent. Suppose we have a Gaussian measure μ on X such that R_ϵ is measure-preserving and satisfies the mixing condition on H.

Now suppose we are given a family of bounded linear operators A_ϵ on X (where $x \mapsto A_\epsilon(x)$ is continuous for all x but A_ϵ is not necessarily a semigroup). Suppose that there exists a Banach space Y and a constant $C > 0$ such that

1. Y embeds continuously into X.

2. $\mu(Y) = 1$.

3. $\|A_\epsilon y - y\|_X \le C\epsilon^{\kappa(R)}$ for all y with $\|y\|_Y \le 1$.

Then we have the bounds $\|R_\epsilon A_\epsilon x - R_\epsilon x\|_X \le \|R_\epsilon\| \|A_\epsilon x - x\| \le \epsilon^{-\kappa(R)} \|A_\epsilon x - x\|$. Now for all x in a set of full measure (namely $x \in Y$) we can bound $\|A_\epsilon x - x\| \le C\|x\|_Y \epsilon^{\kappa(R)}$. Consequently we find that

$$\|R_\epsilon A_\epsilon x - R_\epsilon x\|_X \le C\|x\|_Y,$$

and thus as long as $x \in Y$ (which is almost every x) we have proved that the set of limit points of $(\log\log(1/\epsilon))^{-1/2} R_\epsilon A_\epsilon x$ coincides with those of $(\log\log(1/\epsilon))^{-1/2} R_\epsilon x$, namely $B(H)$.

Intuitively what this implies is the seemingly obvious fact that the Strassen's law still holds as long as we mollify fast enough relative to the rescaling operation (R_ϵ). Consider the simple example of two-sided Brownian motion for instance, say Wiener measure on $X = C[-1,1]$. Let $A_\epsilon f(x) = f * \phi_\epsilon$ where ϕ is a smooth even mollifier and $\phi_\epsilon(x) = \epsilon^{-1}\phi(\epsilon^{-1}x)$ (by convention we let $A_\epsilon f$ be constant on each of $[-1, -1+\epsilon]$ and $[1-\epsilon, 1]$). Let $R_\epsilon f(x) = \epsilon^{-u/2} f(\epsilon^u x)$ where $u > 0$, so that $\kappa(R) = -u/2$. If $u < 1$ then we may define Y to be the closure of smooth functions with respect to the Holder norm of exponent $u/2$ and we know that this space supports Wiener measure almost surely. Moreover an easy computation shows that $\|y * \phi^\epsilon - y\|_{C[-1,1]} \le C\epsilon^{u/2}$ as long as $\|y\|_Y \le 1$ (where $C = \int_{\mathbb{R}} \phi(v)|v|^{u/2} dv$). Consequently by the discussion in the previous paragraph, one obtains the same Strassen law as Brownian motion for the family $R_\epsilon A_\epsilon$ on $C[-1,1]$. On the other hand if $u = 1$ then the limit points will be the mollifier-dependent compact set $\{\phi * f : f(0) = 0, \|f'\|_{L^2[-1,1]} \le 1\}$, since one easily verifies that $R_\epsilon A_\epsilon f = \phi * R_\epsilon f$ in this case. Thus we find that $u < 1$ is necessary to obtain a nontrivial limit set that coincides with the non-mollified case. For $u > 1$ the smoothing dominates the rescaling, so one obtains a trivial limit set consisting only of constant functions of absolute value bounded above by 1, as may be checked by hand.

We thus formulate the following abstract result:

Theorem 5.4.9. *Let (X,H,μ) be an abstract Wiener space, let $(R_\epsilon)_{\epsilon\in(0,1]}$ be a family of Borel-measurable a.e. linear maps from $X\to X$ which are measure-preserving and mixing, and let $T^i:X\to Y_i$ be homogeneous of degree k_i for $1\le i\le m$. Suppose that there exist strongly continuous semigroups $(Q^i_\epsilon)_{t\ge 0}$ operators from $Y_i\to Y_i$ for $1\le i\le m$ with the property that*

$$T^i\circ R_\epsilon = Q^i_\epsilon\circ T^i, \qquad \mu\text{-}a.e. \qquad \textit{for all } \epsilon\in(0,1].$$

For $\epsilon\in(0,1]$ suppose that $J^i_\epsilon: X\to Y_i$ is a family of chaoses of order k_i such that there exist measurable functions $C^i:X\to\mathbb{R}_+$ such that one has the bound

$$\|J^i_\epsilon(x)-T^i(x)\|_{Y_i}\le C^i(x)\epsilon^{\kappa(Q^i)}.$$

Let Z be a Banach space, and let $\mathcal{M}\subset Y_1\times\cdots Y_m$ be a closed subset, such that the semigroup $Q^1_\epsilon\oplus\cdots\oplus Q^m_\epsilon$ sends $\mathcal{M}$ to itself, and moreover for all $\delta,\epsilon>0$

$$\mu(\{x\in X:(\delta^{k_1}T^1(x),...,\delta^{k_m}T^m(x))\in\mathcal{M}\})=1,$$

$$\mu(\{x\in X:(\delta^{k_1}J^1_\epsilon(x),...,\delta^{k_m}J^m_\epsilon(x))\in\mathcal{M}\})=1.$$

Let $\Phi:\mathcal{M}\to Z$ be uniformly continuous on bounded sets. Then the compact set

$$K:=\{(T^1_{hom}(h),...,T^m_{hom}(h)):h\in B(H)\}$$

is necessarily contained in $\mathcal{M}$, and moreover the set of cluster points at infinity of the random set

$$\big\{\Phi\big((2\log\log(1/\epsilon))^{-k_1/2}Q^1_\epsilon J^1_\epsilon(x),\ \ldots,(2\log\log(1/\epsilon))^{-k_m/2}Q^m_\epsilon J^m_\epsilon(x)\big):\epsilon\in(0,1]\}$$

is almost surely equal to $\Phi(K)$.

The proof is fairly straightforward despite the long statement; one makes the bound

$$\|Q^i_\epsilon J^i_\epsilon(x) - Q^i_\epsilon T^i(x)\| \le \|Q^i_\epsilon\|_{Y^i\to Y^i}\|J^i_\epsilon(x) - T^i(x)\| \le \|Q^i_\epsilon\|_{Y^i\to Y^i}C^i(x)\epsilon^{\kappa(Q^i)} \le C^i(x)$$

where we used by definition of $\kappa(Q^i)$ that $\|Q^i_\epsilon\|\epsilon^{\kappa(Q^i)} \le 1$. Hence the set of limit points of the desired sequence must coincide with that of Theorem 5.2.21. Note that one needs the additional assumption of *uniform* continuity of Φ on bounded subsets of $Y_1 \times \cdots \times Y_m$ to argue this last part. Finally we note that we recover the result of the discussion above (for first order chaos) when we set $m = 1$ with $X = Y_1 = Z = \mathcal{M}$ and $J^1_\epsilon = A_\epsilon$ and $Q^i_\epsilon = R_\epsilon$ and $\Phi = I$ and $C^1(x) = C\|x\|_Y$. Note also that we do *not* make the requirement that $Q^i_\epsilon J^i_\epsilon = J^i_\epsilon R_\epsilon$ and in general this will be false. Finally we remark that when $J^i_\epsilon = T^i$ the above theorem recovers Theorem 5.2.21, at least in the case that Φ is uniformly continuous and $\mathcal{M}$ is invariant under the semigroup (which is usually true in practice).

Now let us argue how this result allows us to recover the Strassen law for a version of Φ^4_2 driven by a smooth noise mollified in both time and space. One defines

$$\begin{aligned}
J^1_\epsilon(\xi)(z) &:= \int_{\mathbb{R}_+\times\mathbb{R}} K^\epsilon(z-z_1)\xi(dz_1),\\
J^2_\epsilon(\xi)(z) &:= \int_{(\mathbb{R}_+\times\mathbb{R})^2} K^\epsilon(z-z_1)K^\epsilon(z-z_2)\xi(dz_1)\xi(dz_2),\\
J^3_\epsilon(\xi)(z) &:= \int_{(\mathbb{R}_+\times\mathbb{R})^3} K^\epsilon(z-z_1)K^\epsilon(z-z_2)K^\epsilon(z-z_3)\xi(dz_1)\xi(dz_2)\xi(dz_3),
\end{aligned}$$

where $K^\epsilon = K * \phi_\epsilon$ for a smooth even compactly supported function $\phi : \mathbb{R}^3 \to \mathbb{R}$ and $\phi_\epsilon(t,x) = \epsilon^{-3}\phi(\epsilon^{-2}t, \epsilon^{-1}x)$, and K is the 2+1 dimensional heat kernel as usual. Letting $\mathbf{S}$ denote the solution map of the system and letting $R_\epsilon\xi(t,x) := \epsilon^{2u}\xi(\epsilon^{2u}t, \epsilon^u x)$ and $Q^i_\epsilon f(t,x) = f(\epsilon^{2u}t, \epsilon^u x)$, we claim for each $\epsilon > 0$ that $v := \mathbf{S}(\delta Q^1_\epsilon J^1_\epsilon \xi, \delta^2 Q^2_\epsilon J^2_\epsilon \xi, \delta^3 Q^3_\epsilon J^3_\epsilon \xi)$ is the solution of the equation

$$\partial_t v(t,x) = \Delta v(t,x) - (v(t,x)^3 - C_{\epsilon^{1-u}} \delta^2 v(t,x)) + \delta \epsilon^{2u} \xi_\epsilon(\epsilon^{2u} t, \epsilon^u x)$$

where $v(0,x) = 0$ and $\xi_\epsilon = \xi * \phi_\epsilon$ and C_ϵ is a logarithmic correction depending on ϕ. Indeed one may verify that $Q^i_\epsilon J^i_\epsilon \xi = J^i_{\epsilon^{1-u}}(R_\epsilon \xi)$. Furthermore by the results of Subsection 5.4.1 it is clear that $\mathcal{S}(\delta J^1_\gamma \eta, \delta^2 J^2_\gamma \eta, \delta^3 J^3_\gamma \eta)$ necessarily solves the equation

$$\partial_t u = \Delta u - (u^3 - C_\gamma \delta^2 u) + \delta \eta_\gamma$$

whenever η is sampled from the Gaussian measure μ. Setting $\gamma = \epsilon^{1-u}$ and $\eta = R_\epsilon \xi$ then proves the claim, since one may verify easily that $\eta_\gamma(t,x) = \epsilon^{2u} \xi_{\gamma\epsilon^u}(\epsilon^{2u} t, \epsilon^u x)$ for all $\gamma, \epsilon > 0$.

Now to finish our analysis of the mollified equation, one needs to show a bound of the type $\|J^i_\epsilon(\xi) - T^i(\xi)\|_{Y_i} \le C^i(x)\epsilon^{\kappa(Q^i)}$. The method of obtaining this is effectively summarized in [33, Section 2.3.1] and we do not reproduce the computation here. The main point is that $\kappa(Q^i) = \nu u$ whenever $Y_i = C([0,T], C^{-\nu}_w(\mathbb{R}^2))$ and we can obtain a bound as good as $\epsilon^{(1-\rho)\nu}$ for the norm of $J^i_\epsilon - T^i$ for any choice of $\rho > 0$. Thus we choose $\rho = 1 - u$, which works as long as $u < 1$.

Note that one may obtain very similar results for KPZ, however one cannot simply let $\mathcal{M}$ be the entire space $Y_1 \times ... \times Y_m$ in theorem 5.4.9. The local uniform continuity is guaranteed by [88, Theorem 5.3] which says that the solution map for the SHE is locally Lipchitz, and the Hopf-cole transform can also be shown to be locally Lipchitz wherever the solution is bounded away from zero.

5.5 Appendix 1: Exponential bounds on Gaussian chaoses in Banach spaces

In this appendix we collect some results about homogeneous variables of order k, which are "higher-order generalizations" of Gaussian variables in Banach spaces. Most of the material in this appendix is an adapted and slightly more general version of the material from Borell's original works on Gaussian chaoses [21, 20], as well as [7] and Section 2 of Chapter 3 of [109], and a few results from [91, Section 3]. The reason we included this material (despite being seemingly standard) is because it was difficult to find the exact results we needed in the literature; certain results seem to be sporadically given without proof or only partial proof in some of the aforementioned references, and we wanted to give a more explicit and comprehensive overview for convenience of usage.

If (X, H, μ) is an abstract Wiener space, then the n^{th} homogeneous Wiener chaos, denoted by $\mathcal{H}^k(X, \mu)$ is defined to be the closure in $L^2(X, \mu)$ of the linear span of $H_k \circ g$ as g varies through all elements of the continuous dual space X^*, where H_k denotes the k^{th} Hermite polynomial (normalized so that $H_k(Z)$ has unit variance for a standard normal Z). Equivalently it can be described as the closure in $L^2(X, \mu)$ of the linear span of $H_k(\langle \cdot, v\rangle)$ as v ranges through all elements of H. One always has the orthonormal decomposition $L^2(X, \mu) = \bigoplus_{k \geq 0} \mathcal{H}^k(X, \mu)$, see e.g. [125].

Definition 5.5.1. *Let (X, H, μ) be an abstract Wiener space, and let Y be another separable Banach space. A Borel-measurable map $T : X \to Y$ is called a chaos of order k if $f \circ T \in \mathcal{H}^k(X, \mu)$ for all $f \in Y^*$.*

The main goal of this appendix is to show that if T is homogeneous of order k for some k, then $\|T\|_Y$ has nice tail bounds (matching the k^{th} power of a Gaussian) and satisfies equivalence of moments. Our first step in this direction is to associate a natural multi-linear form to any homogeneous variable of order k.

Definition 5.5.2. *Let (X, H, μ) be an abstract Wiener space, and let X^k denote the k-fold product of X with itself. A subset $E \subset X^k$ is called a k-linear domain if the following hold:*

1. *E is a Borel set and $\mu^{\otimes k}(E) = 1$.*

2. *E is symmetric: if $(x_1, ..., x_k) \in E$ then $(x_{\sigma(1)}, ..., x_{\sigma(k)}) \in E$ for all $\sigma \in S_k$ where S_k is the symmetric group on k letters.*

3. *If* $(x_1, ..., x_{k-1}, a) \in E$ *and* $(x_1, ..., x_{k-1}, b) \in E$ *then* $(x_1, ..., x_{k-1}, a + rb) \in E$ *for all* $r \in \mathbb{R}$.

Let Y *be another Banach space. A function* $F : E \to Y$ *is called symmetric and multilinear if*

- F *is Borel measurable.*
- $F(x_1, ..., x_k) = F(x_{\sigma(1)}, ..., x_{\sigma(k)})$ *for all* $\sigma \in S_k$ *and all* $(x_1, ..., x_k) \in E$.
- $F(x_1, ..., x_{k-1}, a + rb) = F(x_1, ..., x_{k-1}, a) + rF(x_1, ..., x_{k-1}, b)$, *whenever* $r \in \mathbb{R}$ *and* $(x_1, ..., x_{k-1}, a) \in E$ *and* $(x_1, ..., x_{k-1}, b) \in E$.

Similarly $F : E \to \mathbb{R}$ *will be called symmetric and sub-multilinear if the equality in the third bullet point is replaced by* $\leq$ *and* r *is replaced by* $|r|$ *in the right-hand side.*

We make the trivial remark that the class of k-linear domains is closed under countable intersections, which will be implicitly used on occasion. The point of these domains is that there may be multilinear functions (even very simple ones) which are not *measurably* extendable to a full subspace of the form E^k with E a Borel subspace of X of full measure, see [154] for an example (it should be noted that a possibly non-measurable multilinear extension to the entire space X^k always exists, see [8], but this is not useful here).

Lemma 5.5.3. *Let* (X, H, μ) *be an abstract Wiener space, and let* $E \subset X^k$ *be a* k*-linear domain. Then* $H^k \subset E$ *(where* H^k *denotes the* k*-fold Cartesian product of the Cameron-Martin space).*

Proof. We proceed by induction on k. The $k = 0$ case is vacuously true.

Assume that any $(k-1)$-linear domain contains H^{k-1}. Assume $E \subset X^k$ is a k-linear domain. For $x \in X$ define $E_x := \{(a_1, ..., a_{k-1}) \in X^{k-1} : (x, a_1, ..., a_{k-1}) \in E\}$. Since $\mu^{\otimes k}(E) = 1$, an application of Fubini's theorem shows that $\mu(F) = 1$, where $F := \{x \in X : \mu^{\otimes(k-1)}(E_x) = 1\}$. It's clear that $E_x \cap E_y \subset E_{x+ry}$ for all $x, y \in X$ and $r \in \mathbb{R}$, so that F is a Borel measurable linear subspace of full measure and thus contains H. Thus $\mu^{\otimes(k-1)}(E_v) = 1$ for all $v \in H$, and thus E_v is a $(k-1)$-linear domain. Thus by the inductive hypothesis we have that $H^{k-1} \subset E_v$, which means that $(v, h_1, ..., h_{k-1}) \in E$ for all $h_1, ..., h_{k-1} \in H$. Thus $H^k \subset E$. $\square$

Lemma 5.5.4. *Let $E \subset X^k$ be a k-linear domain. Then there exists a Borel set $A \subset E$ of full measure such that for every $x = (x_1, ..., x_k) \in A$ and every $\ell \leq k$, the set $\{(a_1, ..., a_\ell) \in X^\ell : (x_1, ..., x_{k-\ell}, a_1, ..., a_\ell) \in E\}$ is an ℓ-linear domain. In particular, for every $(x_1, ..., x_k) \in A$, every $h = (h_1, ..., h_k) \in H^k$ and every $I \subset [k]$ one has that $(x_I, h_{I^c}) \in E$, where (x_I, h_{I^c}) denotes the vector whose i^{th} component equals x_i if $i \in I$ and equals h_i if $i \notin I$.*

Proof. For $\ell \leq k$, and $(x_1, ..., x_\ell) \in X^\ell$, let

$$B^{k-\ell}_{(x_1,...,x_\ell)} := \{(a_1, ..., a_{k-\ell}) \in X^{k-\ell} : (x_1, ..., x_\ell, a_1, ..., a_{k-\ell}) \in E\},$$
$$F_\ell := \{(x_1, ..., x_k) \in X^k : \mu^{\otimes(k-\ell)}(B^{k-\ell}_{(x_1,...,x_\ell)}) = 1\}.$$

Since $\mu^{\otimes k}(E) = 1$ it follows that $\mu^{\otimes k}(F_\ell) = 1$. Then let $A := \bigcap_{0 \leq \ell \leq k} F_\ell$. Clearly $\mu^{\otimes k}(A) = 1$. If $(x_1, ..., x_k) \in A$ then $B^{k-\ell}_{(x_1,...,x_\ell)}$ is a $(k-\ell)$-linear domain, and thus the required claim follows. Furthermore by Lemma 5.5.3 it follows that $B^{k-\ell}_{(x_1,...,x_\ell)}$ contains $H^{k-\ell}$. In other words if $(x_1, ..., x_k) \in A$ then $(x_1, ..., x_\ell, h_1, ..., h_{k-\ell}) \in E$ for all $h_1, ..., h_{k-\ell} \in H$. By permutation invariance of E it then follows that $(x_I, h_{I^c}) \in E$ for all $I \subset [k]$ with $|I| = \ell$. $\square$

Lemma 5.5.5. *Let (X, H, μ) be an abstract Wiener space, and let $\psi \in \mathcal{H}^k(X, \mu)$. There exists a multilinear domain $E \subset X^k$ and a symmetric multilinear function $\hat{\psi} : E \to \mathbb{R}$ such that the following relation holds for $\mu^{\otimes k}$-a.e. $(x_1, ..., x_k) \in X^k$:*

$$\hat{\psi}(x_1, ..., x_k) = \frac{k^{k/2}}{2^k k!} \sum_{\epsilon \in \{-1,1\}^k} \epsilon_1 \cdots \epsilon_k \cdot \psi\bigg(k^{-1/2} \sum_{i=1}^k \epsilon_i x_i\bigg). \tag{5.20}$$

Proof. The main point is that if $p : \mathbb{R} \to \mathbb{R}$ is any monic polynomial of degree k then

$$\frac{1}{2^k k!} \sum_{\epsilon \in \{-1,1\}^k} \epsilon_1 \cdots \epsilon_k \cdot p\bigg(\sum_{i=1}^k \epsilon_i x_i\bigg) = x_1 \cdots x_n. \tag{5.21}$$

This may be verified combinatorially.

Since $\psi \in \mathcal{H}^k(X,\mu)$, there exists a Borel subset $F \subset X$ of measure 1, a sequence $p_j \uparrow \infty$ of natural numbers, and a collection of vectors $\{f_{i,j}\}_{1 \le i \le p_j} \subset X^*$ and a collection of scalars $\{c_{i,j}\}_{1\le i\le p_j} \subset \mathbb{R}$, with the property that $\psi(x) = \lim_{j\to\infty} \sum_{i=1}^{p_j} c_{i,j} H_k(f_{i,j}(x))$ for all $x \in F$ (actually the limit is in $L^2(X,\mu)$, but we can always pass to a subsequence to obtain an a.e. limit of this form).

Let $G_\epsilon := \{(x_1,...,x_k) \in X^k : k^{-1/2}\sum_1^k \epsilon_i x_i \in F\}$ for $\epsilon \in \{-1,1\}^k$, and let $G := \bigcap_{\epsilon\in\{-1,1\}^k} G_\epsilon$. Then $\mu^{\otimes k}\big(G\big) = 1$ since $k^{-1/2}\sum_1^k \epsilon_i x_i$ is distributed as μ, for every ϵ.

Let $q(x_1,...,x_k)$ denote the right-hand side of (5.20). Then by (5.21), if $(x_1,...,x_k) \in G$,

$$q(x_1,...,x_k) = \lim_{j\to\infty} \sum_{i=1}^{p_j} c_{i,j} f_{i,j}(x_1)\cdots f_{i,j}(x_k). \tag{5.22}$$

We define E to be the set of all $(x_1,...,x_k) \in X^k$ such that the right hand side of (5.22) converges, and we define $\hat{\psi} : E \to \mathbb{R}$ to be the value of the limit in the right side of (5.22). Then E is a Borel set, since the set of values where a sequence of measurable functions converges is always measurable. Since E contains G, it follows that $\mu^{\otimes k}(E) = 1$. Furthermore it is clear that E satisfies conditions *(2)* and *(3)* of Definition 5.5.2. It is also clear that $\hat{\psi}$ is a symmetric and multilinear function from $E \to \mathbb{R}$. $\square$

Our next proposition gives an explicit formula for $\hat{\psi}$ on H^k.

Lemma 5.5.6. *Let (X,H,μ) be an abstract Wiener space. For $h_1,...,h_k \in H$ and $\psi \in \mathcal{H}^k(X,\mu)$*

one has that

$$\hat{\psi}(h_1, ..., h_k) = \frac{1}{k!} \int_X \psi(x) \langle x, h_1 \rangle \cdots \langle x, h_k \rangle \mu(dx) \tag{5.23}$$

$$= \int_{X^k} \hat{\psi}(x_1, ..., x_k) \langle x_1, h_1 \rangle \cdots \langle x_k, h_k \rangle \mu^{\otimes k}(dx_1, ..., dx_k). \tag{5.24}$$

In particular there exists a constant $C > 0$ which may depend on k but is independent of X, H, μ, ψ such that the following bounds hold:

$$\sup_{\|h_1\|_H, ..., \|h_k\|_H \le 1} \hat{\psi}(h_1, ..., h_k)^2 \le C \int_X \psi(x)^2 \mu(dx) \tag{5.25}$$

$$\sup_{\|h_1\|_H, ..., \|h_k\|_H \le 1} \hat{\psi}(h_1, ..., h_k)^2 \le C \int_{X^k} \hat{\psi}(x_1, ..., x_k)^2 \mu^{\otimes k}(dx_1, ..., dx_k). \tag{5.26}$$

Proof. First we prove (5.23). Given the construction of $\hat{\psi}$ in the proof of Lemma 5.5.5, it suffices to prove the claim when $\psi(x) = H_k(\langle x, v \rangle)$ for some fixed $v \in H$. Note by (5.21) that in this case $\hat{\psi}(h_1, ..., h_k) = \langle h_1, v \rangle \cdots \langle h_k, v \rangle$. Furthermore, since any symmetric multilinear map is uniquely determined by its value on “diagonal elements" of the form $(h, h, ..., h)$, it suffices to prove the claim when all h_i are equal. Summarizing this paragraph, it suffices to show that for all $v, h \in H$ one has

$$\langle h, v \rangle^k = \frac{1}{k!} \int_X H_k(\langle x, v \rangle) \langle x, h \rangle^k \mu(dx).$$

Since $H_k(\langle \cdot, v \rangle) \in \mathcal{H}^k(X, \mu)$, it suffices to prove the claim with $\langle x, h \rangle^k$ replaced by its projection onto $\mathcal{H}^k(X, \mu)$ in the right hand side. But that projection is just $H_k(\langle x, h \rangle)$. In other words we just need to prove that

$$\langle h, v \rangle^k = \frac{1}{k!} \int_X H_k(\langle x, v \rangle) H_k(\langle x, h \rangle) \mu(dx).$$

But this is true, see e.g. Corollary 2.3 in [84].

Next we prove (5.25). Note by Hölder's inequality and (5.23) that

$$|\hat{\psi}(h_1, ..., h_k))| \leq \frac{1}{k!} \int_X |\psi(x)\langle h_1, x\rangle \cdots \langle h_k, x\rangle| \mu(dx)$$
$$\leq \frac{1}{k!} \Big[\int_X \psi(x)^{2k} \mu(dx) \Big]^{1/(2k)} \prod_{j=1}^{k} \Big[\int_X \langle h_i, x\rangle^{2k} \mu(dx) \Big]^{1/(2k)}$$

Since $\langle h_i, \cdot\rangle$ is distributed under μ as a centered Gaussian of variance $\|h_i\|_H^2$, we have

$$[\int_X \langle h_i, x\rangle^{2k} \mu(dx)]^{1/(2k)} = C^{1/(2k)} \|h_i\|_H,$$

where C is the $(2k)^{th}$ moment of a standard (variance 1) Gaussian. Furthermore, since ψ lies in the k^{th} Wiener chaos, we have the hypercontractive bound

$$\Big[\int_X \psi(x)^{2k} \mu(dx) \Big]^{1/(2k)} \leq (2k-1)^{k/2} \Big[\int_X \psi(x)^2 \mu(dx) \Big]^{1/2},$$

see for instance equation (7.2) in [84].

The proof of (5.24) is completely analogous to the proof of (5.23) above: it suffices to prove it when $\psi = H_k(\langle \cdot, v\rangle)$ for some $v \in H$, in which case $\hat{\psi}(x_1, ..., x_k) = \langle x_1, v\rangle \cdots \langle x_k, v\rangle$. By splitting the integrals using Fubini's theorem, it in turn suffices to show that $\int_X \langle x_i, h_i\rangle \langle x_i, v\rangle \mu(dx_i) = \langle h_i, v\rangle$. But this is a trivial. The proof of (5.26) is then done in a completely analogous manner to the proof of (5.25), using the same hypercontractive bounds but exploiting (5.24) rather than (5.23). □

Equations (5.24) and (5.26) can be generalized as follows:

Lemma 5.5.7. *Let (X, H, μ) be an abstract Wiener space, and let $\psi \in \mathcal{H}^k(X, \mu)$. Let $\hat{\psi} : E \to \mathbb{R}$ denote the associated multilinear form, where E is a k-linear domain. Let $A \subset E$ denote the subset constructed in Lemma 5.5.4. Then for every $(x_1, ..., x_k) \in A$, every $h_1, ..., h_k \in H$ and* *every $I \subset [k]$ one has the following identity:*

$$\hat{\psi}(x_{I^c}, h_I) = \int_{X^k} \hat{\psi}(x_{I^c}, u_I) \prod_{i \in I} \langle u_i, h_i \rangle \mu^{\otimes k}(du). \tag{5.27}$$

In particular there exists $C > 0$ which may depend on k but is independent of X, H, μ, ψ such that

$$\sup_{\|h_i\|_H \leq 1, \forall i \in I} \hat{\psi}(x_{I^c}, h_I)^2 \leq C \int_{X^k} \hat{\psi}(x_{I^c}, u_I)^2 \mu^{\otimes k}(du). \tag{5.28}$$

For $I \subset [k]$, we have as usual used (x_{I^c}, u_I) to denote the vector whose i^{th} component equals x_i if $i \notin I$ and equals u_i if $i \in I$. Notice that integral over X^k integrates out only those components whose index lies in I. The proof of (5.27) is done in a completely analogous fashion to the proof of equations (5.24) and (5.23) above. One first proves the claim for $\hat{\psi}$ of the form $(x_1, ..., x_k) \mapsto \langle x_1, v \rangle \cdots \langle x_k, v \rangle$, where $v \in H$ such that $x \mapsto \langle x, v \rangle$ is in X^*. The claim then follows since $\hat{\psi}$ is a pointwise limit of finite linear combinations of such functions on its entire domain E (see the proof of Lemma 5.5.5). The proof of (5.28) is then done using hypercontractivity and (5.27) as before.

Definition 5.5.8. *Let (X, H, μ) be an abstract Wiener space, let Y be another separable Banach space, and let $T : X \to Y$ be homogeneous of order k. We define the decoupled chaos $Q_T : X^k \to Y$ as follows:*

$$Q_T(x_1, ..., x_k) := \frac{k^{k/2}}{2^k k!} \sum_{\epsilon \in \{-1,1\}^k} \epsilon_1 \cdots \epsilon_k \cdot T\bigg(k^{-1/2} \sum_{i=1}^k \epsilon_i x_i\bigg).$$

Notice that if $T : X \to Y$ is a homogeneous variable of order k on the abstract Wiener space (X, H, μ), then Q_T is also a homogeneous variable of order k on the abstract Wiener space $(X^k, H^k, \mu^{\otimes k})$. By Lemma 5.5.5 it is intuitively clear that there should exist a k-linear domain $E \subset X^k$ and a multilinear function $\hat{T} : E \to Y$ which coincides with Q_T almost everywhere. This will be proved later in Proposition 5.5.18, but first we need some preliminaries.

Proposition 5.5.9. *Let (X, H, μ) be an abstract Wiener space, let Y be another separable Banach space, and let $T : X \to Y$ be homogeneous of order k. There exists a multilinear domain $E \subset X^k$ and a symmetric sub-multilinear map $J_T : E \to \mathbb{R}$ with the property that the following relation holds for $\mu^{\otimes k}$-a.e. $(x_1, ..., x_k) \in X^k$:*

$$J_T(x_1, ..., x_k) = \|Q_T(x_1, ..., x_k)\|_Y. \tag{5.29}$$

Furthermore, there exists a universal constant C such that the following bounds hold:

$$\sup_{\|h_1\|_H,...,\|h_k\|_H \le 1} J_T(h_1, ..., h_k)^2 \le C \sup_{\|f\|_{Y^*} \le 1} \int_X f(T(x))^2 \mu(dx) < \infty. \tag{5.30}$$

$$\sup_{\|h_1\|_H,...,\|h_k\|_H \le 1} J_T(h_1, ..., h_k)^2 \le C \sup_{\|f\|_{Y^*} \le 1} \int_{X^k} f(Q_T(x_1, ..., x_k))^2 \mu^{\otimes k}(dx). \tag{5.31}$$

Here C may depend on the homogeneity k but is independent of the choice of (X, H, μ, Y, T).

Proof. Choose $\{f_n\}_{n \ge 1} \subset Y^*$ such that $\sup_n f_n(y) = \|y\|_Y$ for all $y \in Y$. Such a collection of functionals exists because Y is separable.

We an find by Lemma 5.5.5 a k-linear domain $E_n \subset X^k$ and a multilinear symmetric map $g_n : E_n \to \mathbb{R}$ such that $f_n(Q_T(x_1, ..., x_k)) = g_n(x_1, ..., x_k)$ whenever $(x_1, ..., x_k) \in F_n$ where F_n is Borel and $\mu^{\otimes k}(F_n) = 1$.

Define E to be the set of all $(x_1, ..., x_k)$ such that $\sup_n g_n(x_1, ..., x_k) < \infty$, and define $J_T : E \to \mathbb{R}$ to be that supremum. Since $\sup_n f_n(y) = \|y\|_Y$, it's clear that $\sup_n g_n(x_1, ..., x_k)$ equals the right side of (5.29) whenever $(x_1, ..., x_k) \in \bigcap_n F_n$. Consequently E contains $\bigcap_n F_n$ and thus $\mu^{\otimes k}(E) = 1$. Furthermore, it is obvious that E satisfies conditions *(2)* and *(3)* of Definition 5.5.2, so that E is a k-linear domain. It is also clear that J_T is sub-multilinear on E.

Next we prove (5.30). The first inequality is clear directly from (5.25) and the fact that $J_T =$

$\sup_n g_n$, since $g_n \overset{a.e}{=} f_n \circ Q_T \in \mathcal{H}^k(X, \mu)$ and since $\|f_n\|_{Y^*} = 1$. To prove that

$$\sup_{\|f\|_{Y^*} \le 1} \int_X f(T(x))^2 \mu(dx) < \infty,$$

just apply the closed graph theorem to the linear map from $Y^* \to L^2(X, \mu)$ given by $f \mapsto f \circ T$ to deduce that it must be bounded.

The proof of (5.31) is similar to the proof of (5.30), but one uses (5.26) rather than (5.25). □

Bound (5.31) generalizes as follows:

Lemma 5.5.10. *Let $J_T : E \to \mathbb{R}$ be as in the previous lemma, and let $A \subset E$ be as constructed in Lemma 5.5.4. Then there exists a Borel set $B \subset A$ of full measure such that for all $x \in B$ one has the following bound:*

$$\sup_{\|h_i\|_H \le 1, \forall i \in I} J_T(x_{I^c}, h_I)^2 \le C \sup_{\|f\|_{Y^*} \le 1} \int_{X^k} f(Q_T(x_{I^c}, u_I))^2 \mu^{\otimes k}(du) < \infty. \tag{5.32}$$

Proof. Let g_n be as in the proof of the previous lemma so that $J_T = \sup_n g_n$. Then for all $x \in A$ one has

$$\sup_{\|h_i\|_H \le 1, \forall i \in I} J_T(x_{I^c}, h_I)^2 \le C \sup_n \int_{X^k} g_n(x_{I^c}, u_I)^2 \mu^{\otimes k}(du). \tag{5.33}$$

The proof of this is completely analogous to the first inequality in (5.31), however one uses (5.28) rather than (5.26). Note that since $g_n = f_n \circ Q_T$ $\mu^{\otimes k}$-a.e., it follows by Fubini that

$$\int_{X^k} g_n(x_{I^c}, u_I)^2 \mu^{\otimes k}(du) = \int_{X^k} f_n(Q_T(x_{I^c}, u_I))^2 \mu^{\otimes k}(du), \qquad \text{for } \mu^{\otimes k}\text{-a.e. } x \in A.$$

We let B denote the set of all $x \in A$ such that the above equality holds for all $I \subset [k]$. Then clearly B has measure 1 and satisfies the required condition since $\|f_n\|_{Y^*} = 1$. The fact that the sup over $\|f\|_{Y^*} \le 1$ is finite follows once again from the closed graph theorem, as applied in the proof of the previous lemma. □

Next we prove that the sub-multilinear form associated to any homogeneous variable has finite moments of all orders and that all of its moments are equivalent. Later we will use this to show that the homogeneous variable itself has finite moments of all orders and all its moments are equivalent.

Proposition 5.5.11. *Let (X, H, μ) be an abstract Wiener space and let $T : X \to Y$ be homogeneous of order k, and let J_T denote the associated sub-multilinear form. Then*

$$\hat{m}_T(r) := \int_{X^k} J_T(x_1, ..., x_k)^r \mu^{\otimes k}(dx_1, ..., dx_k) < \infty$$

for all $r > 0$, and moreover

$$\mu^{\otimes k}(\{(x_1, ..., x_k) \in X^k : J_T(x_1, ..., x_k) > a\}) < Ce^{-\alpha\left(a^2/\hat{m}_T(2)\right)^{1/k}}, \tag{5.34}$$

where $C, \alpha > 0$ are universal constants which depend on k but are independent of the choice of X, H, μ, Y, T, and $a > 0$.

Proof. The proof given here uses the Gaussian isoperimetric inequality and is adapted from Chapter 3 of [109], which in turn is adapted from Borell's original work [21, 20].

Assume $E \subset X^k$ is a k-linear domain, let $A \subset E$ be a subset as constructed in Lemma 5.5.4, and let $B \subset A$ be a further subset as constructed in Lemma 5.5.10. Note by Lemma 5.5.10 that (5.32) holds for all $x \in B$ and all $I \subset [k]$. For $x \in B$,let $S_I(x)$ denote the left-hand side of (5.32). Define $S_{total} : B \to \mathbb{R}$ by

$$S_{total}(a) := \sum_{\substack{I\subset[k] \\ I\neq\emptyset}} S_I(a).$$

For $M > 0$ let

$$A_M := \{a \in B : S_{total}(a) \leq M \text{ and } J_T(a) \leq M\}. \tag{5.35}$$

Now choose $M_0 > 0$ such that $\mu^{\otimes k}(A_{M_0}) \geq 1/2$. For $t \geq 0$ let

$$A_{M_0}(t) := \{(a_1 + th_1, ..., a_k + th_k) : (a_1, ..., a_k) \in A_{M_0} \text{ and } \sum_1^k \|h_i\|_H^2 \leq 1\}.$$

By the Gaussian isoperimetric inequality, $\mu_*^{\otimes k}(A_{M_0}(t)) \geq 1 - \frac{1}{2}e^{-t^2/2}$, where $\mu_*^{\otimes k}$ denotes inner measure. On the other hand, if $(a_1 + th_1, ..., a_k + th_k) \in A_{M_0}(t)$ then by sub-multilinearity of J_T on E and the fact that $(a_I, h_{I^c}) \in E$ for all $a \in B$ and $h \in H^k$ (recall Lemma 5.5.4 and note $B \subset A$) we have that

$$\begin{aligned} J_T(a_1 + th_1, ..., a_k + th_k) &\leq \sum_{I \subset [k]} t^{|I|} J_T(a_{I^c}, h_I) \\ &\leq M_0 + \sum_{\substack{I \subset [k] \\ I \neq \emptyset}} t^{|I|} S_I(a) \leq M_0 + t^k S_{total}(a) \leq M_0 + t^k M_0. \end{aligned}$$

In particular, $\{(a_1, ..., a_k) \in B : J_T(a_1, ..., a_k) > M_0 + t^k M_0\}$ is contained in the complement of $A_{M_0}(t)$, and since B has full measure this means that

$$\mu^{\otimes k}(\{(a_1, ..., a_k) \in X^k : J_T(a_1, ..., a_k) > M_0 + t^k M_0\}) \leq \frac{1}{2}e^{-t^2/2}. \tag{5.36}$$

This already shows that $\hat{m}_T(r) < \infty$ for all $r \geq 0$. Thus it only remains to show (5.34). For this we will obtain precise bounds on M_0 and then use (5.36). First note by Chebyshev that if we define $M_1 := 2\hat{m}_T(2)^{1/2}$, then

$$\mu^{\otimes k}(\{(x_1, ..., x_k) : J_T(x_1, ..., x_k) > M_1\}) \leq M_1^{-2}\hat{m}_T(2) = 1/4. \tag{5.37}$$

Recall $J_T = \sup_n g_n$, so by (5.33) we have $S_I(a) \leq \left[\int_{X^k} J_T(a_{I^c}, u_I)\mu^{\otimes k}(du)\right]^{1/2}$, so that by

Chebyshev, if we define $M_2 := 6^k \hat{m}_T(2)^{1/2}$ then for every $\ell \le k$,

$$\mu^{\otimes k}(\{a \in X^k : S_I(a) > 2^{-k} M_2) \le 4^k M_2^{-2} \int_{X^k} S_I(a)^2 \mu^{\otimes k}(da)$$
$$\le 4^k M_2^{-2} \int_{X^k} \int_{X^k} J_T(a_{I^c}, u_I)^2 \mu^{\otimes k}(du) \mu^{\otimes k}(da) = 4^k M_2^{-2} \hat{m}_T(2) \le (4/36)^k = 9^{-k}.$$

Thus we can apply a union bound to obtain

$$\mu^{\otimes k}(\{a \in X^k : S_{total}(a) > M_2\}) \le \sum_{\substack{I \subset [k] \\ I \neq \emptyset}} \mu^{\otimes k}(\{a \in X^k : S_I(a) > 2^{-k} M_2\})$$
$$\le \sum_{\substack{I \subset [k] \\ I \neq \emptyset}} 9^{-k} \le 2^k 9^{-k} \le 1/4. \tag{5.38}$$

Combining equations (5.37) and (5.38) we see that $\mu^{\otimes k}(A^c_{\max\{M_1, M_2\}}) \le 1/2$ with A_M defined in (5.35). Thus we can take $M_0 := \max\{M_1, M_2\} = 6^k \hat{m}_T(2)^{1/2}$, and then by equation (5.36) we obtain

$$\mu^{\otimes k}(\{(x_1, ..., x_k) : J_T(x_1, ..., x_k) > a\}) \le \frac{1}{2} e^{-\frac{1}{2}(M_0^{-1} a - 1)^{2/k}}.$$

As long as $a > 2M_0$ we know that $M_0^{-1} a - 1 > (2M_0)^{-1} a$, hence if $a > 2M_0$ then the last probability is bounded above by $\frac{1}{2} e^{-2^{-1-2k^{-1}} a^{2/k} M_0^{-2/k}}$. $\square$

Since $J_T = \|Q_T\|_Y$ a.e. by construction, we have the following.

Corollary 5.5.12. *The results of Proposition 5.5.11 still apply if we replace $J_T(x_1, ..., x_k)$ by $\|Q_T(x_1, ..., x_k)\|_Y$ throughout the statement.*

Proposition 5.5.13. *Let (X, H, μ) be an abstract Wiener space and let $T : X \to Y$ be homogeneous of order k. For $r \ge 1$ let*

$$m_T(r) := \int_X \|T(x)\|_Y^r \mu(dx),$$

and let $\hat{m}_T(r)$ be as defined in Proposition 5.5.11. There exist constants $c, C > 0$ which may *depend on k but are independent of (X,H,μ,Y,T,r) such that*

$$c^r m_T(r) \le \hat{m}_T(r) \le C^r m(r). \tag{5.39}$$

In particular $m_T(r) < \infty$ for all $r \ge 1$.

Proof. First we show that $\hat{m}_T(r) \le C^r m(r)$. To do this note by Jensen that

$$\begin{aligned}\|Q_T(x_1,...,x_n)\|_Y^r &= \left\| \frac{1}{2^k} \sum_{\epsilon\in\{-1,1\}^k} \frac{k^{k/2}}{k!} \epsilon_1\cdots\epsilon_n T\bigg(k^{-1/2}\sum_1^k \epsilon_i x_i\bigg)\right\|_Y^r \\ &\le \frac{1}{2^k}\sum_{\epsilon\in\{-1,1\}^k} \frac{k^{kr/2}}{(k!)^r}\left\|T\bigg(k^{-1/2}\sum_1^k \epsilon_i x_i\bigg)\right\|_Y^r,\end{aligned}$$

for $\mu^{\otimes k}$-a.e. $(x_1,...,x_k) \in X^k$. Thus we integrate both sides of the above inequality over X^k and we use the fact that $k^{-1/2}\sum_1^k \epsilon_i x_i$ has law μ under $\mu^{\otimes k}$, and we see that

$$\hat{m}_T(r) \le 2^{-k}\sum_\epsilon \frac{k^{kr/2}}{(k!)^r} m(r) = C^r m(r),$$

where $C := k^{k/2}/k!$.

Now we need to prove that $m_T(r) \le c^{-r}\hat{m}_T(r)$ for small enough $c>0$. Henceforth if $F: X^\ell \to Y$ is any function, where $\ell \in \mathbb{N}$, then we will abbreviate $F(y_1,...,y_\ell)$ by $F((y_i)_{i=1}^{\ell-1}, y_\ell)$. We claim that the following relation holds for $\mu^{\otimes k}$-a.e. $(x_1,...,x_k)\in X^k$:

$$Q_T((2^{-i/2}k^{1/2}x_i)_{i=1}^{k-1}, 2^{-(k-1)/2}k^{1/2}x_k) = k^{k/2}2^{-(k+1)(k-1)/4}Q_T((x_i)_{1\le i\le k}). \tag{5.40}$$

To prove this, we choose $\{f_n\}_{n\ge1}\subset Y^*$ such that $\sup_n f_n(y) = \|y\|_Y$. Since $\{f_n\}_n$ separates points of Y, it suffices to show, for each n, that (5.40) holds a.e. after applying f_n to both sides. In turn it suffices to show that for any $f\in Y^*$, the composition of f with both sides can be modified on a set of measure zero to yield a multilinear function defined on some k-linear domain, since the right side is formally obtained from the left side by pulling constants out of each coordinate. For the right hand side, this is true, simply by Lemma 5.5.5 and the fact that $f \circ T \in \mathcal{H}^k(X, \mu)$. To prove this for the left side, note that f composed with the left side equals

$$\frac{k^{k/2}}{2^k k!} \sum_{\epsilon \in \{-1,1\}^k} \epsilon_1 \cdots \epsilon_k f \circ T\bigg(\sum_{i=1}^{k-1} 2^{-i/2} \epsilon_i x_i + 2^{-(k-1)/2} \epsilon_k x_k \bigg).$$

Note that the sum of squares of the coefficients of the x_i equals 1. In particular this means that $\sum_{i=1}^{k-1} 2^{-i/2} \epsilon_i x_i + 2^{-(k-1)/2} \epsilon_k x_k$ has the same distribution as μ. Thus the proof that the above function can be modified on a set of measure zero to yield a multilinear function can thus be done in a similar fashion to the proof of Lemma 5.5.5 (i.e., one defines $G_\epsilon := \{(x_1, ..., x_k) \in X^k : \sum_{i=1}^{k-1} 2^{-i/2} \epsilon_i x_i + 2^{-(k-1)/2} \epsilon_k x_k \in F\}$ and proceeds exactly in the same way as that proof, using equation (5.21)).

Next we will inductively define measurable maps $T_j : X^\ell \to Y$ for $1 \le \ell \le k$ as follows. Let $T_1(x) := T(x)$, and then define

$$T_{\ell+1}(x_1, ..., x_{\ell+1}) := T_\ell\big((x_i)_{i=1}^{\ell-1}, 2^{-1/2}(x_\ell - x_{\ell+1})\big) - T_\ell\big((x_i)_{i=1}^{\ell-1}, 2^{-1/2}(x_\ell + x_{\ell+1})\big).$$

It is clear that each T_ℓ is homogeneous of order k. By induction it is easy to show that T_ℓ is equivalently given by

$$T_\ell(x_1, ..., x_\ell) = \sum_{\epsilon \in \{-1,1\}^\ell} \epsilon_1 \cdots \epsilon_\ell T\bigg(\sum_{i=1}^{\ell-1} 2^{-i/2} \epsilon_i x_i + 2^{-(\ell-1)/2} \epsilon_\ell x_\ell \bigg).$$

In particular, $T_k(x_1, ..., x_k)$ equals the left side of (5.40) multiplied by $k^{-k/2} 2^k k!$. We also define another inductive family of functions $g_\ell : X^\ell \to Y$ for $1 \le \ell \le k$ by $g_k := 0$ and

$$g_\ell(x_1, ..., x_\ell) = \int_X \big[g_{\ell+1}\big(x_1, ..., x_{\ell-1}, 2^{-1/2}(x_\ell - x_{\ell+1}), 2^{-1/2}(x_\ell - x_{\ell+1})\big) \\ + T_\ell\big(x_1, ..., x_{\ell-1}, x_{\ell+1}\big)\big]\mu(dx_{\ell+1}).$$

Inductively one can see that g_ℓ does not actually depend on x_ℓ (the definition is somewhat misleading), thus $g_1 = 0$.

We now claim the following series of relations:

$$\begin{aligned}
&\int_{X^{\ell+1}} \|T_{\ell+1}(x_1, ..., x_{\ell+1}) - g_{\ell+1}(x_1, ..., x_{\ell+1})\|_Y^r \mu^{\otimes(\ell+1)}(dx) \\
&= \int_{X^{\ell+1}} \big\|T_{\ell+1}\big(x_1, ..., x_{\ell-1}, 2^{-1/2}(x_\ell - x_{\ell+1}), 2^{-1/2}(x_\ell + x_{\ell+1})\big) \\
&\qquad\qquad - g_{\ell+1}\big(x_1, ..., x_{\ell-1}, 2^{-1/2}(x_\ell - x_{\ell+1}), 2^{-1/2}(x_\ell + x_{\ell+1})\big)\big\|_Y^r \mu^{\otimes(\ell+1)}(dx) \\
&\ge \int_{X^\ell} \bigg\| \int_X T_{\ell+1}\big(x_1, ..., x_{\ell-1}, 2^{-1/2}(x_\ell - x_{\ell+1}), 2^{-1/2}(x_\ell + x_{\ell+1})\big) \\
&\qquad - g_{\ell+1}\big(x_1, ..., x_{\ell-1}, 2^{-1/2}(x_\ell - x_{\ell+1}), 2^{-1/2}(x_\ell + x_{\ell+1})\big)\mu(dx_{\ell+1})\bigg\|_Y^r \mu^{\otimes\ell}(dx_1, ..., dx_\ell) \\
&= \int_{X^\ell} \bigg\| \int_X \big[T_\ell(x_1, ..., x_{\ell-1}, x_\ell) - T_\ell(x_1, ..., x_{\ell-1}, x_{\ell+1}) \\
&\qquad - g_{\ell+1}\big(x_1, ..., x_{\ell-1}, 2^{-1/2}(x_\ell - x_{\ell+1}), 2^{-1/2}(x_\ell + x_{\ell+1})\big)\mu(dx_{\ell+1})\bigg\|_Y^r \mu^{\otimes\ell}(dx_1, ..., dx_\ell) \\
&= \int_{X^\ell} \big\|T_\ell(x_1, ..., x_\ell) - g_\ell(x_1, ..., x_\ell)\big\|_Y^r \mu^{\otimes\ell}(dx_1, ..., dx_\ell).
\end{aligned}$$

Here the integrals of T_ℓ over X are meant to be interpreted in the sense of Bochner. The first equality is from rotational invariance of the last two coordinates, the next inequality is by Jensen, the next equality is just the definition of $T_{\ell+1}$ in terms of T_ℓ, and the last equality is the definition of g_ℓ. Note that the inequality we have just derived is of the form $p(\ell+1) \ge p(\ell)$, so recalling that

$g_1 = 0$ and $g_k = 0$, we find that

$$\int_{X^k} \|T_k(x_1, ..., x_k)\|_Y^r \mu^{\otimes k}(dx_1, ..., dx_k) \geq \int_X \|T_1(x_1)\|_Y^r \mu(dx_1).$$

By (5.40) and our inductive scheme we know that $T_k(x_1, ..., x_k) = k! 2^{-(k-1)^2/4} Q_T(x_1, ..., x_k)$ $\mu^{\otimes k}$-a.e., which gives desired claim. $\square$

Corollary 5.5.14. *Let (X, H, μ) be an abstract Wiener space and let $T : X \to Y$ be homogeneous of order k. Let $m_T(r)$ be as defined in Proposition 5.5.13. Then there exists some $\alpha > 0$ such that*

$$\mu(\{x \in X : \|T(x)\|_Y > a\}) < Ce^{-\alpha\left(a^2/m_T(2)\right)^{1/k}}, \tag{5.41}$$

where $C, \alpha > 0$ depend on k but are independent of the choice of (X, H, μ, Y, T) and $a > 0$.

Proof. It suffices to show that $\int_X \exp\left(\lambda \cdot m_T(2)^{-1/k} \|T(x)\|_Y^{2/k}\right)\mu(dx) < C$ where λ, C depend on k but not on X, H, μ, Y, T. To do this, write

$$\exp\left(\lambda \cdot m_T(2)^{-1/k} \|T(x)\|_Y^{2/k}\right) = \sum_{r=0}^{\infty} \frac{\lambda^r m_T(2)^{-r/k} \|T(x)\|_Y^{2r/k}}{r!}.$$

Now notice that

$$\begin{aligned} m_T(r)^{2/r} &\overset{(5.39)}{\leq} C\hat{m}_T(r)^{2/r} \leq C\left[\int_0^\infty r x^{r-1} \mu^{\otimes k}(\{a \in X^k : \|Q_T(a)\|_Y > x\})dx\right]^{2/r} \\ &\overset{(5.34)}{\leq} C\hat{m}_T(2) f(r)^{2/r} \overset{(5.39)}{\leq} C m_T(2) f(r)^{2/r}, \end{aligned}$$

where C may grow with each inequality but only depends on k, and

$$f(r) := \int_0^\infty r x^{r-1} e^{-\alpha x^{2/k}} dx = \frac{1}{2} k r \alpha^{-k/2} \Gamma(kr/2).$$

Here α is the same constant appearing in (5.34). Combining the previous three expressions, we see that

$$\int_X \exp\left(\lambda \cdot m_T(2)^{-1/k}\|T(x)\|_Y^{2/k}\right)\mu(dx) = \sum_{r=0}^{\infty} \frac{\lambda^r m_T(2)^{-r/k} m_T(2r/k)}{r!}$$
$$\leq \sum_{r=0}^{\infty} \frac{\lambda^r C^r f(2r/k)}{r!}$$

Notice that $f(2r/k)/r! = \frac{1}{2}k\alpha^{-k/2}$ is a constant depending on k but nothing else, hence the above sum will indeed converge (independently of X, H, μ, Y, T) for $\lambda < 1/C$. □

Next we prove that any homogeneous variable of order k can be approximated in the strongest sense by a sequence of finite rank Cameron-Martin projections, in the same way as a Gaussian variable in any Banach space.

Let (X, H, μ) be an abstract Wiener space. Choose an orthonormal basis $\{e_n\}_n$ for H, and let

$$P_n x := \sum_{j=1}^{n} \langle x, e_j \rangle e_j, \qquad Q_n x := \sum_{j=n+1}^{\infty} \langle x, e_j \rangle e_j = x - P_n x.$$

Note that $P_n x, Q_n x$ are independent. Therefore if (x, y) is sampled from $\mu^{\otimes 2}$, then $P_n x + Q_n y$ is distributed as μ. If $T : X \to Y$ is homogeneous of order k, we thus define a sequence of "finite-rank Cameron-Martin projections for T" by the formula

$$T_n(x) := \int_X T(P_n x + Q_n y)\mu(dy). \tag{5.42}$$

This is a well-defined Bochner integral for a.e. $x \in X$. Indeed, since

$$\int_X \int_X \|T(P_n x + Q_n y)\|_Y \mu(dy)\mu(dx) = \int_X \|T(u)\|_Y \mu(du) < \infty,$$

it follows that $\int_X \|T(P_n x + Q_n y)\|_Y \mu(dy)$ is finite for a.e. x.

Proposition 5.5.15. *Let (X, H, μ) be an abstract Wiener space, and let $T : X \to Y$ be homo-*

geneous of order k. If T_n is defined as in (5.42)*, then T_n is also homogeneous of order k, and moreover $\|T_n - T\|_Y \to 0$ a.e. and in every $L^p(X,\mu)$ as $n \to \infty$. In fact, one has the following super-polynomial convergence bound:*

$$\mu(\{x \in X : \|T_n(x) - T(x)\|_Y > u\}) \leq C \exp\big[-\alpha\big(u/\|T_n - T\|_{L^2(X,\mu;Y)}\big)^{2/k}\big] \tag{5.43}$$

where C, α are independent of X, H, μ, Y, T, u, n and the choice of basis $\{e_n\}$.

Proof. The proof of a.e. convergence follows immediately from the Banach space-valued martingale convergence theorem [35]. Indeed, T_n can be viewed as the conditional expectation of T given $\{\langle x, e_i\rangle\}_{1\leq i\leq n}$. The proof of L^p convergence follows similarly, since $\sup_n \int_X \|T_n(x)\|^p \mu(dx) \leq \int_X \|T(x)\|^p \mu(dx) < \infty$ by Jensen's inequality. The fact that T_n is also homogeneous of order k follows from the fact that the random variable $(x,y) \mapsto T(P_n x + Q_n y)$ is homogeneous of order k on the abstract Wiener space $(X^2, H^2, \mu^{\otimes 2})$ and the fact that integrals of homogeneous variables are still homogeneous of the same order. Since $T_n - T$ is homogeneous of order k, (5.43) follows immediately from Corollary 5.5.14 and Proposition 5.5.15. $\square$

Definition 5.5.16. *Let (X, H, μ) be an abstract Wiener space, and let Y be another separable Banach space. For $p \geq 1$ we define the following Banach spaces:*

1. *$L^p(X,\mu;Y)$ is the space of all Borel measurable functions $\phi : X \to Y$ such that $\int_X \|\phi\|_Y^p d\mu < \infty$, up to a.e. equivalence.*

2. *$\mathcal{H}_0^k(X,\mu;Y)$ is the closed subspace of $L^2(X,\mu;Y)$ consisting of all ϕ such that $f \circ \phi \in \mathcal{H}^k(X,\mu)$ for all $f \in Y^*$.*

3. *$\mathcal{H}^k(X,\mu;Y)$ consists of the closure in $L^2(X,\mu;Y)$ of the linear span of $x \mapsto H_k(g(x))y$, as g varies through all elements of the continuous dual space X^* and y varies through Y. Here H_k is the k^{th} Hermite polynomial as usual.*

Note that $L^2(X,\mu;Y)$ is not a Hilbert space unless Y is a Hilbert space (which is seldom true in practical applications). Also note by Corollary 5.5.14 that all of the L^p norms on $\mathcal{H}_0^k(X,\mu;Y)$ are equivalent.

Lemma 5.5.17. *In the notation of Definition 5.5.16, we have* $\mathcal{H}_0^k(X,\mu;Y) = \mathcal{H}^k(X,\mu;Y)$.

Proof. It is clear that $\mathcal{H}^k \subset \mathcal{H}_0^k$. Conversely let $T \in \mathcal{H}_0^k(X,\mu;Y)$. Fix an orthonormal basis $\{e_n\}$ for H and let $T_n \in \mathcal{H}_0^k(X,\mu;Y)$ be given by (5.42). Then T_n are measurable with respect to a *finite* family $\{\langle x,e_i\rangle\}_{1\le i\le n}$ of iid standard Gaussians, therefore it is clear that T_n can be written as $x \mapsto \sum_{j=1}^{p_n} y_j^n H_k(\langle x, v_j^n\rangle)$, where $p_n \in \mathbb{N}$, $y_j^n \in Y$, and $v_j^n \in \text{span}\{e_i\}_{i=1}^n$. In particular $T_n \in \mathcal{H}^k(X,\mu;Y)$. By Proposition 5.5.15we know that $T_n \to T$ in $L^2(X,\mu;Y)$, and thus $T \in \mathcal{H}^k(X,\mu;Y)$. □

Next we prove that Q_T can be modified on a measure-zero set to yield a multilinear function.

Proposition 5.5.18. *Let (X,H,μ) be an abstract Wiener space, and let $T : X \to Y$ be homogeneous of order k. There exists a k-linear domain $E \subset X^k$ and a multilinear function $\hat{T} : X^k \to Y$ such that $\hat{T}(x_1,...,x_k) = Q_T(x_1,...,x_k)$ for $\mu^{\otimes k}$-a.e. $(x_1,...,x_k)$.*

Proof. By Lemma 5.5.17, we know that there exists a Borel set F with $\mu(F) = 1$ such that $\lim_{n\to\infty} \|T(x) - \sum_{j=1}^{p_n} y_j^n H_k(g_j^n(x))\|_Y = 0$, for every $x \in F$, where $p_n \in \mathbb{N}$, $y_j^n \in Y$, and $g_j^n \in X^*$ (actually the limit is in $L^2(X,\mu;Y)$, but we can always pass to a subsequence to extract an a.e. limit).

As we did in the proof of Lemma 5.5.5, we define $G_\epsilon := \{(x_1,...,x_k) \in X^k : k^{-1/2}\sum_1^k \epsilon_i x_i \in F\}$ for $\epsilon \in \{-1,1\}^k$, and let $G := \bigcap_{\epsilon\in\{-1,1\}^k} G_\epsilon$. Then $\mu^{\otimes k}(G) = 1$ since $k^{-1/2}\sum_1^k \epsilon_i x_i$ is distributed as μ, for every ϵ.

Recalling the definition of Q_T (see Definition 5.5.8) and equation (5.21), it is clear that whenever $(x_1,...,x_k) \in G$

$$Q_T(x_1,...,x_k) = \lim_{n\to\infty} \sum_{j=1}^{p_n} y_j^n g_j^n(x_1)\cdots g_j^n(x_k). \tag{5.44}$$

We define E to be the set of all $(x_1, ..., x_k) \in X^k$ such that the right hand side of (5.44) converges (with respect to the norm topology on Y), and we define $\hat{T} : E \to Y$ to be the value of the limit in the right side of (5.44). Then E is a Borel set, since the set of values where a sequence of measurable functions converges is always measurable. Since E contains G, it follows that $\mu^{\otimes k}(E) = 1$. Furthermore it is clear that E satisfies conditions *(2)* and *(3)* of Definition 5.5.2. It is also clear that $\hat{T}$ is a symmetric and multilinear function from $E \to Y$. $\square$

Next we characterize $\hat{T}$ on the subset $H^k \subset E$. The following is attributed to Section 3 of [91].

Lemma 5.5.19. *Let (X, H, μ) be an abstract Wiener space, and let $T : X \to Y$ be homogeneous of order k. For $h_1, ..., h_k \in H$, one has the Bochner integral representation:*

$$\hat{T}(h_1, ..., h_k) = \frac{1}{k!} \int_X T(x) \langle x, h_1 \rangle \cdots \langle x, h_k \rangle \mu(dx). \tag{5.45}$$

In particular for $h \in H$ we have

$$\hat{T}(h, ..., h) = \int_X T(x + h) \mu(dx). \tag{5.46}$$

Proof. To prove (5.45), it suffices to show that equality holds if after applying any $f \in Y^*$. But since $f \circ T \in \mathcal{H}^k(X, \mu)$, this follows easily from (5.23).

To prove (5.46), note by the Cameron-Martin formula that

$$\int_X T(x+h)\mu(dx) = \int_X T(x) e^{\langle x, h \rangle - \frac{1}{2}\|h\|_H^2} \mu(dx) = \int_X T(x) \sum_{r=1}^{\infty} \frac{1}{r!} H_r(\langle x, h \rangle) \mu(dx)$$
$$= \int_X T(x) \frac{1}{k!} H_k(\langle x, h \rangle) \mu(dx) = \int_X T(x) \frac{1}{k!} \langle x, h \rangle^k \mu(dx) = \hat{T}(h, ..., h).$$

In the second equality we expanded the exponential in terms of its Hermite polynomial generating series, in the third equality we used that T is homogeneous of order k so that only the k^{th} Hermite polynomial is not orthogonal to T, and in the fourth equality we again used the fact that T is homogeneous of order k so that T is orthogonal to any polynomial of degree less than k (in particular it is orthogonal to $H_k(\langle x,h\rangle) - \langle x,h\rangle^k$). In the final equality we used (5.45). $\square$

Definition 5.5.20. *Let (X,H,μ) be an abstract Wiener space, and let $T : X \to Y$ be homogeneous of order k. We define the canonical functional $T_{hom} : H \to Y$ by sending $h \in H$ to the expression in* (5.46)*.*

Lemma 5.5.21. *Let (X,H,μ) be an abstract Wiener space, and let $T : X \to Y$ be homogeneous of order k. Choose an orthonormal basis $\{e_n\}$ of H and let T_n be the finite-rank approximation given in* (5.42)*. Then we have the uniform convergence*

$$\lim_{n\to\infty} \sup_{\|h\|\le 1} \big\|(T_n)_{hom}(h) - T_{hom}(h)\big\|_Y = 0.$$

Letting $B(H)$ denote the unit ball of H, it follows that T_{hom} is continuous from $B(H) \to Y$, where $B(H)$ is given the topology of X (not of H).

Proof. The proof given here is taken directly from Section 3 of [91]. It is clear that each T_n is continuous from $(B(H), \|\cdot\|_X)$ to $(Y, \|\cdot\|_Y)$, since it can be identified with a polynomial function from $\mathbb{R}^n \to Y$ (i.e., a function of the form $(x_1,...,x_n) \mapsto \sum_{j=1}^N p_j(x_1,...,x_n)y_j$ where p_j are polynomials and $y_j \in Y$). To prove that $(T_n)_{hom} \to T_{hom}$ uniformly on $B(H)$, note that

$$\begin{aligned}
\|(T_n)_{hom}(h) - T_{hom}(h)\|_Y &\le \int_X \|T_n(x+h) - T(x+h)\|_Y \mu(dy) \\
&= \int_X \|T_n(x) - T(x)\|_Y e^{\langle x,h\rangle - \frac{1}{2}\|h\|_H^2} \mu(dx) \\
&\le \|T_n - T\|_{L^2(X,\mu;Y)} \bigg[\int_X e^{2\langle x,h\rangle - \|h\|^2} \mu(dx) \bigg]^{1/2} \\
&= \|T_n - T\|_{L^2(X,\mu;Y)} \cdot e^{\|h\|^2/2.}
\end{aligned}$$

Here we applied the definition of T_{hom} in the first inequality, then we used the Cameron-Martin formula in the next line, then we used Cauchy-Schwarz in the following line, then we computed the integral in the last line. Hence we can finally obtain that $\sup_{\|h\|\le 1} \|(T_n)_{hom}(h) - T_{hom}(h)\|_Y \le$

$e^{1/2}\|T_n - T\|_{L^2(X,\mu;Y)}$, which tends to 0 by Proposition 5.5.15. $\square$

Corollary 5.5.22. *Let (X, H, μ) be an abstract Wiener space, and let $T^i : X \to Y_i$ be homogeneous of order k_i for $1 \le i \le m$, where $m \in \mathbb{N}$. Then the set*

$$\{(h, T^1_{hom}(h), ..., T^m_{hom}(h)) : h \in B(H)\}$$

is a compact subset of $X \times Y_1 \times \cdots Y_m$.

Proof. It is well-known that $B(H)$ is a compact subset of X. By Lemma 5.5.21, the map $h \mapsto (h, T^1_{hom}(h), ..., T^m_{hom}(h))$ is continuous from $B(H)$ to $X \times Y_1 \times \cdots Y_m$, and the claim follows immediately. $\square$

We conclude this appendix with the remark that the bounds we have derived here are not sharp, but see the seminal work [107] for a derivation of sharp bounds, as well as subsequent papers on that subject such as [1].

5.6 Appendix 2: Regularity Structures and the solution map Φ

Let us explain the map above by giving a rapid introduction to the theory of regularity structures. We reproduce without motivation only the formalities which are important to us in the KPZ example, but gentler introductions with more background and intuition can be found in [33, 83, 87]. Whenever we write $\mathbb{R}^{d+1}$ below it should be thought of as $\mathbb{R} \times \mathbb{R}^d$ which stands for space-time.

Definition 5.6.1. *A **regularity structure** is defined to be a triple (T, A, G) where $T = \bigoplus_{\alpha \in A} T_\alpha$ is a graded vector space indexed by $A \subset \mathbb{R}$ which is a discrete set bounded from below, and G is a group of (invertible) linear transformations $T \to T$ satisfying $\Gamma\tau - \tau \in \bigoplus_{\beta<\alpha} T_\beta$ for all $\tau \in T_\alpha$. Furthermore each T_α is required to be finite-dimensional. A **model** for (T, A, G) on $\mathbb{R}^{d+1}$ consists of two maps $\Pi : \mathbb{R}^{d+1} \to \mathcal{L}(T, \mathcal{S}'(\mathbb{R}^d))$ and $\Gamma : \mathbb{R}^{d+1} \times \mathbb{R}^{d+1} \to G$ such that for any compact set*

$K \subset \mathbb{R}^{d+1}$ and $\gamma > 0$ there exists $C > 0$ such that for all $\beta < \alpha \leq \gamma$ and $x, y \in K$ one has the bounds

$$|\Pi_x(\phi_x^\lambda)| \leq C\lambda^\alpha \|\tau\|_\alpha$$

$$\|\Gamma_{xy}\tau - \tau\|_\beta \leq C|x-y|^{\alpha-\beta}\|\tau\|_\alpha$$

$$\Pi_x\Gamma_{xy} = \Pi_y, \qquad \Gamma_{xy}\Gamma_{yz} = \Gamma_{xz},$$

where we abbreviated $\|\tau\|_\alpha := \|\tau^\alpha\|_{T_\alpha}$ where τ^α is the component of τ in T_α, where $\Pi_x = \Pi(x)$, where $\Gamma_{xy} = \Gamma(x, y)$, and where $\phi_{(t,x)}^\lambda(s, a) := \lambda^{-d-2}\phi(\lambda^{-2}(t-s), \lambda^{-1}(x-a))$, with $t, s \in \mathbb{R}$, and $a, x \in \mathbb{R}^d$ and $\lambda > 0$.

Given a kernel $K : \mathbb{R}^{d+1} \to \mathbb{R}$ which can be decomposed as $\sum_{n\geq 0} K_n$ where K_0 is smooth and each K_n with $n \geq 1$ is smooth and compactly supported and satisfies the bounds $|\partial^k K_n(x)| \leq 2^{n(d+k-\rho)}$ for some $\rho > 0$ one may then define a model (Π, Γ) for (T, A, G) to be **admissible** with respect to the kernel K if the following properties hold:

1. (T, A, G) contains a copy of the polynomial regularity structure, that is $\{0, 1, 2, ...\} \subset A$ and each T_α with $\alpha = k$ contains a linear subspace spanned by formal indeterminates in $d+1$ variables, say X^k with $k \in \mathbb{N}^{d+1}$ in multi-index notation, such that each $\Gamma \in G$ acts on X^k by sending it to $(X-h)^k := \sum_{i=0}^k \binom{k}{i} h^{k-i}X^i$ for some $h \in \mathbb{R}$.

2. There exists a linear map $\mathcal{I} : T \to T$ which sends each T_α to $T_{\alpha+\rho}$ and sends the polynomial part of the regularity structure to zero. Moreover for each symbol $\tau \in T$ and $k \in \mathbb{N}$ one has the identities

$$\Pi_x X^k(z) = (z-x)^k,$$
$$\Pi_x \mathcal{I}\tau(z) = \int_{\mathbb{R}^{d+1}} K(z-a)\Pi_x\tau(da) - \sum_{|k|_s \leq |\tau|+\rho} \frac{(z-x)^k}{k!} \int_{\mathbb{R}^{d+1}} \partial^k K(x-a)\Pi_x\tau(da),$$

   where the integrals can be shown to be well-defined quantities by decomposing $K = \sum_n K_n$ and using the given bounds on K_n and their derivatives. Again k here is a multi-index in $\mathbb{N}^{d+1}$

and $a^k = \prod a_i^{k_i}$, with $k! = \prod k_i!$, and $|k|_s = 2k_0 + k_1 + ... + k_d$.

Definition 5.6.2. *Given a model* (Π, Γ) *for* (T, A, G) *one can define for* $\gamma > 0$ *Frechet spaces* $\mathcal{D}^\gamma(\Pi, \Gamma)$ *to be the set of all functions* $f : \mathbb{R}^{d+1} \to T_\gamma^-$ *(where the latter is shorthand for* $\bigoplus_{\alpha<\gamma} T_\alpha$*) such that one has for all compact sets* K *and* $\alpha < \gamma$

$$\sup_{x \in K} \|f(x)\|_\alpha + \sup_{x,y \in K} \frac{\|f(x) - \Gamma_{xy} f(y)\|_\alpha}{|x - y|^{\gamma - \alpha}} < \infty.$$

These spaces generalize Holder spaces and recover them in the case of the polynomial regularity structure.

Given two models (Π, Γ) and $(\bar{\Pi}, \bar{\Gamma})$ one can define for each compact set $K \subset \mathbb{R}^{d+1}$ a distance between them by the formula

$$\max_{\beta<\alpha<\gamma} \sup_{x,y \in K} \sup_{\lambda \in (0,1]} \sup_{\|\phi\|_{C^r} \leq 1} \sup_{\|\tau\|_\alpha = 1} \left[\frac{|(\Pi_x \tau - \bar{\Pi}_x \tau)(\phi_x^\lambda)|}{\lambda^\alpha} + \frac{\|\Gamma_{xy}\tau - \bar{\Gamma}_{xy}\tau\|_\beta}{|x - y|^{\alpha - \beta}} \right],$$

where $r := -(\lfloor \min A \rfloor \wedge 0)$. Given a model (Π, Γ) for (T, A, G) we say that (Π, Γ, f) is a γ-**relevant triple** if $f \in \mathcal{D}^\gamma(\Pi, \Gamma)$. One can likewise define for each compact $K \subset \mathbb{R}^{d+1}$ a pseudo-metric on the space of γ-relevant triples for (T, A, G) by the same formula above but with an additional term inside the supremum given by

$$\frac{\|f(x) - \bar{f}(x) - \Gamma_{xy} f(y) + \bar{\Gamma}_{xy} \bar{f}(y)\|_\alpha}{|x - y|^{\gamma - \alpha}}.$$

Instead of looking at compact sets, in many contexts one may also work spatially on a compact torus $\mathbb{R} \times \mathbb{T}^d$ to simplify things, or alternatively one may take the supremum over all of $\mathbb{R}^d$ after introducing a suitable weight function $w(x)$ in the denominators of each of the expressions above. One then has the following "**Reconstruction Theorem**" with the relevant "Schauder Estimate":

Theorem 5.6.3. *Given a model* (Π, Γ) *for* (T, A, G) *and* $\gamma > 0$ *there exists a unique continuous linear map* $\mathcal{R} = \mathcal{R}^{(\Pi,\Gamma)} : \mathcal{D}^\gamma(\Pi, \Gamma) \to \mathcal{S}'(\mathbb{R}^{d+1})$ *satisfying* $|\big(\Pi_x f(x) - \mathcal{R}f\big)(\phi_x^\lambda)| \lesssim \lambda^\gamma$ *uniformly* *on compact sets of x and $\lambda \in (0,1]$ and uniformly over $\|\phi\|_{C^r} \leq 1$ where $r := -(\lfloor \min A \rfloor \wedge 0)$.*

Furthermore the map $(\Pi, \Gamma, f) \mapsto \mathcal{R}^{(\Pi,\Gamma)} f$ is locally Lipchitz continuous from the space of γ-relevant triples to the Holder space $C^{-r-\epsilon}$ for any $\epsilon > 0$.

Assume furthermore that $\gamma > 0$ and that the model (Π, Γ) is admissible with respect to a kernel $K = \sum_n K_n$ as above. Assume each K_n with $n \geq 1$ integrates to zero against all polynomials of degree less or equal to $\gamma + \rho$. Then there exists a bounded linear operator $\mathcal{K}_\gamma = \mathcal{K}_\gamma^{(\Pi,\Gamma)} : D^\gamma(\Pi, \Gamma) \to \mathcal{D}^{\gamma+\rho}(\Pi, \Gamma)$ which satisfies $K * \mathcal{R} = \mathcal{R}\mathcal{K}_\gamma$.

Hairer proves that the heat kernel admits a decomposition of the above type, where each K_n with $n \geq 1$ does integrate to zero. From now on when we refer to an admissible model it will always be with respect to the heat kernel, for which we can optimally take $\rho = 2$. The latter part of the above theorem is what allows us to construct the continuous parts of solution maps to singular SPDEs on spaces of *admissible* models, which are inherently nonlinear spaces. First one must introduce products:

A **product** on (T, A, G) is simply a symmetric bilinear map $T \times T \to T$, usually denoted by a simple juxtaposition of elements, such that $\tau\bar{\tau} \in T_{\alpha+\beta}$ whenever $\tau \in T_\alpha, \bar{\tau} \in T_\beta$. From now onward assume (T, A, G) is “normal" meaning that it contains the polynomial structure and admits a linear map $\mathcal{I}$ as above, and has a unique element Ξ, which together with the product on T and the operator $\mathcal{I}$ and the polynomial elements of T generate the entire regularity structure.

Given a continuous function $\xi : \mathbb{R}^{d+1} \to \mathbb{R}$ and a “normal" regularity structure (T, A, G) as described above, an *admissible* model (Π, Γ) (admissible with respect to the heat kernel) will be said to be a **canonical model** lifting the continuous function ξ up to order γ if one has the following properties:

1. $\Pi_x \Xi(z) = \xi(z)$ for all x, z.

2. $\Pi_x \tau\bar{\tau}(z) = \Pi_x \tau(z) \Pi_x \bar{\tau}(z)$ for all x, z whenever $|\tau| + |\bar{\tau}| \le \gamma$.

Any “normal" regularity structure as described above and any continuous function ξ can be shown to admit a canonical model lifting it, denoted $(\Pi, \Gamma) =: \iota\xi$ by using a fairly straightforward recursive construction. The most important property of canonical models is that for all $f \in \mathcal{D}^\gamma(\iota\xi)$ the distribution $\mathcal{R}^{\iota\xi} f$ actually exists as a continuous function, moreover $\mathcal{R}^{\iota\xi} f(x) = \big(\Pi_x f(x)\big)(x)$, and most strikingly

$$\mathcal{R}^{\iota\xi}(fg) = \mathcal{R}^{\iota\xi} f \cdot \mathcal{R}^{\iota\xi} g,$$

for all $f, g \in \mathcal{D}^\gamma(\iota\xi)$ assuming $fg \in \mathcal{D}^\beta$ for some $\beta \in (0, \gamma]$ where the product on the left hand side is the abstract product on T and the product on the right hand side is pointwise product of functions. While it is the case that canonical models thusly constructed are automatically admissible, it is not the case that all admissible models are of this form, or even limits of such models with respect to the natural metric. This fact is crucial for the renormalization of singular SPDEs such as KPZ and Φ^4_3.

With all of this machinery in place, let us now show how to construct the continuous part Φ of the solution map for the multiplicative SHE mentioned above. The spatial dimension d is always set to 1 here. One constructs a regularity structure as follows. One first constructs a set of words W as follows: one imposes that formal symbols $\Xi, 1, X_0, X_1 \in W$ with homogeneity $-3/2 - \kappa, 0, 2, 1$ resp (here X_0 stands for the time variable which has parabolic degree two). Then one imposes $\tau\bar{\tau} \in W$ whenever $\tau, \bar{\tau} \in W$, and $\mathcal{I}[\tau] \in W$ whenever $\tau \in W$, subject to the additional constraints that

- $\Xi^2 = 0$,
- $\tau 1 = \tau$,
- $\tau\mathcal{I}[\Xi] = 0$ for $\tau \neq \Xi$,
- $\mathcal{I}[\mathcal{I}[\Xi]] = 0$,

- $X^j X^\ell = X^{j+\ell}$ for multi-indices k, ℓ, and
- $\mathcal{I}[X^k] = 0$ for all multi-indices k.

Modulo these constraints, the order of elements is defined so that $|\tau\bar{\tau}| = |\tau| + |\bar{\tau}|$ and $|\mathcal{I}[\tau]| = |\tau| + 2$. One then defines T to be the linear span of all of these formal symbols W, with a product on T defined suggestively by the given juxtaposition rules of the formal symbols (and extended bilinearly), and A is the set of all possible orders of symbols. Note that W can be decomposed as $U \cup F$ where U denotes all symbols of the form $\mathcal{I}[\Xi\tau]$ for some $\tau \in W$ and F denotes the set of all symbols of the form $\mathcal{I}[\tau]$ for some $\tau \in W$ together with polynomial elements. The structure group G in its entirety will not be defined here but the elements of G relevant to the problem will be defined just below.

For any map Π satisfying the requirements of an admissible model, there is a way to define a corresponding family of invertible transformations Γ_{xy} of T such that one has the relations $\Pi_x \Gamma_{xy} = \Pi_y$ and $\Gamma_{xy}\Gamma_{yz} = \Gamma_{xz}$. The first condition and the fact that $\Pi_x X^k(z) = (z-x)^k$ forces Γ_{xy} to act on elements $\mathcal{I}[\tau]$ by the formula

$$\Gamma_{xy}\mathcal{I}[\tau] = \mathcal{I}[\tau] + \sum_{|k|_s \le |\tau|+2} \frac{X^k}{k!} \int_{\mathbb{R}^{d+1}} \big(\partial^k K(y-a) - \partial^k K(x-a)\big)(\Pi_x \tau)(da)$$

$$+ \sum_{|k|_s \le |\tau|+2} \frac{(X+y-x)^k - X^k}{k!} \int_{\mathbb{R}^{d+1}} \partial^k K(y-a)(\Pi_x\tau)(da).$$

Furthermore it is clear that $\Gamma_{xy}\Xi = \Xi$. Finally one has $\Gamma_{xy}(\Xi X^k \mathcal{I}[\tau]) = \Xi X^k \Gamma_{xy}\mathcal{I}[\tau]$ and $\Gamma_{xy} X^k = (X+y-x)^k$, which finishes specifying how each Γ_{xy} acts on T. Note that Γ_{xy} leaves invariant both the span of each of U and F, meaning our regularity structure can be "orthogonally decomposed" as a direct sum of two separate structures related only by the abstract product on T. The part corresponding to U has elements of only positive order which is used in the arguments.

One then proves that for each admissible model (Π, Γ) the map $f \mapsto 1 + \mathcal{K}_\gamma(f\Xi)$ is a contraction from $\mathcal{D}^\gamma(\Pi,\Gamma;\mathcal{U})$ to itself, where $D^\gamma(\Pi,\Gamma;\mathcal{U})$ is the subspace consisting those elements $f\in\mathcal{D}^\gamma(\Pi,\Gamma)$ that take values in the linear span of U (note that one must take a suitable terminal time for the system, and introduce a suitable weight to make this argument work, and this is the subject of Sections 3 and 4 of [88]). Consequently for each model there is a unique fixed point $u^{(\Pi,\Gamma)}$ satisfying $u^{(\Pi,\Gamma)}=1+\mathcal{K}_\gamma(u^{(\Pi,\Gamma)}\Xi)$. Here we add "1" because the initial data of zero in the KPZ equation (for which we are trying to prove the LIL) corresponds to initial data 1 for the mSHE.

The local Lipchitz continuity of the fixed point map $(\Pi,\Gamma)\mapsto u^{(\Pi,\Gamma)}$ can be obtained using an extension of the aforementioned contraction estimate via the following general principle. Suppose that X,Y are complete metric spaces and $h:X\times Y\to X$ is a map satisfying

$$d_X(h(x,y),h(x',y'))\le \varrho d_X(x,x')+Cd_Y(y,y')$$

where $\varrho\in(0,1)$ and $C>0$. Then $h(\cdot,y)$ is a contraction on X for all $y\in Y$ and thus it has a unique fixed point $u(y)\in X$. To prove Lipchitz continuity of $u:Y\to X$ note one has the estimate $d_X(u(y),u(y'))=d_X(h(u(y),y),h(u(y'),y'))\le\varrho d_X(u(y),u(y'))+Cd_Y(y,y')$. By rearranging terms one then immediately obtains

$$d_X(u(y),u(y'))\le\frac{C}{1-\varrho}d_Y(y,y'),$$

as desired. The argument of [88] is of this type with Y being the set of admissible models (Π,Γ) and $X=\mathcal{D}^\gamma(\Pi,\Gamma;\mathcal{U})$. This is not a product space but the argument works in precisely the same way with $X\times Y$ replaced by the space of γ-relevant triples. Here the map $h:X\times Y\to X$ corresponds to the map sending $(\Pi,\Gamma,f)\mapsto 1+\mathcal{K}_\gamma(f\Xi)$.

Once we have local Lipchitz continuity of $(\Pi,\Gamma)\mapsto u^{(\Pi,\Gamma)}$ it automatically follows that the map $(\Pi,\Gamma)\mapsto(\Pi,\Gamma,u^{(\Pi,\Gamma)})$ is also locally Lipchitz continuous, so that by the reconstruction theorem the composition map $(\Pi,\Gamma)\mapsto(\Pi,\Gamma,u^{(\Pi,\Gamma)})\mapsto\mathcal{R}^{(\Pi,\Gamma)}u^{(\Pi,\Gamma)}$ is also locally Lipchitz.

This composition will be called Φ_0 below. This map Φ_0 is almost the map Φ that we want to address above except for the fact that Φ is defined not defined on admissible models (Π, Γ) but rather on objects of the form $(\Pi_z \tau)_{z\in\mathbb{R}^{d+1},\tau\in W^-}$ where W^- denotes all those symbols from W which are of negative order. We call such a collection an admissible pre-model (in the sense that it extends to an admissible model and uniquely determines the action of Π on all elements of negative order). It follows directly from Proposition 3.31 and Theorem 5.14 of [83] that there is a locally Lipchitz map $\mathcal{E}$ sending $\mathbf{\Pi} \mapsto (\Pi, \Gamma)$ from admissible pre-models to their unique extension to an admissible model. Then the map Φ can be factored as $\Phi_0\mathcal{E}$.

First we claim that if ξ is a smooth function, and if $\iota\xi$ denotes the canonical model associated to it, then $\Phi_0(\iota\xi) = \mathcal{R}^{\iota\xi}u^{\iota\xi}$ solves the PDE $\partial_t u = \partial_x^2 u + \xi u$. Indeed, letting $v := \Phi_0(\iota\xi)$ we find by the fundamental property of canonical models, namely $\mathcal{R}^{\iota\xi}(fg) = \mathcal{R}^{\iota\xi}f \cdot \mathcal{R}^{\iota\xi}g$, as well as the fact that u_0 is a fixed point of the map $f \mapsto 1 + \mathcal{K}_\gamma(f\Xi)$, that

$$v-1 = \mathcal{R}^{\iota\xi}(u^{\iota\xi}-1) = \mathcal{R}^{\iota\xi}(\mathcal{K}_\gamma(u^{\iota\xi}\Xi)) = K*(\mathcal{R}^{\iota\xi}(u^{\iota\xi}\Xi)) = K*\big((\mathcal{R}^{\iota\xi}u^{\iota\xi})\cdot(\mathcal{R}^{\iota\xi}\Xi)\big) = K*(v\cdot\xi),$$

which is precisely the equation $\partial_t v = \partial_x^2 v + \xi v$ with initial data identically 1, written in Duhamel form. Here we used the fact that ξ is smooth so that $\mathcal{R}^{\iota\xi}\Xi(x) = (\Pi_x\Xi)(x) = \xi(x)$ and likewise $\mathcal{R}^{\iota\xi}1 = 1$. From this it immediately follows that for $f \in L^2$ (approximating by smooth functions if necessary) if we replace $\xi(dw), \xi(da), \xi(db), \xi(dc)$ by $f(w)dw, f(a)da, f(b)db, f(c)dc$, then Φ maps the five objects above to the classical solution of the equation $\partial_t u = \partial_x^2 u + fu$. Indeed one easily checks that the **extension map** $\mathcal{E}$ sends those five (classical) integrals to the canonical model ιf.

Next we need to show that for all $\delta > 0$ if we replace ξ by $\delta\xi$ then Φ maps those five stochastic integrals to the Ito-Walsh solution of mSHE driven by $\delta\xi$.

Hairer and Pardoux actually showed the following: Letting $\mathscr{M}$ denote the space of admissible models for (T, A, G), there exists a finite dimensional subgroup $\mathfrak{R}$ of the space of isomorphisms $T \to T$, and a continuous "**renormalization map**" $\mathfrak{R} \times \mathscr{M} \to \mathscr{M}$ denoted by $(M, \Pi, \Gamma) \mapsto (\Pi^M, \Gamma^M)$ (see [89, Theorem 4.2]). Furthermore there are explicitly defined operators $L, L^{(1)} : T \to T$ such that for arbitrary $c, c^{(1)} \in \mathbb{R}$ if we define $M := \exp(-cL - c^{(1)}L^{(1)})$ then $M \in \mathfrak{R}$ (see [89, Proposition 4.3]) and furthermore the unique fixed point $\Phi_0(u^{(\Pi^M,\Gamma^M)})$, where $(\Pi, \Gamma) = \iota\xi$ for some continuous ξ, solves the equation $\partial_t v = \partial_x^2 v + v(\xi - c - c^{(1)})$ (see [89, Proposition 4.4]). Finally given some smooth even mollifier ρ, they show that there are explicit (fixed) choices of $c, c^{(1)}$ (depending on the mollifier) such that if we let $M^\epsilon := \exp(-c\epsilon^{-1}L - c^{(1)}L^{(1)})$, then $(\Pi_\epsilon^{M^\epsilon}, \Gamma_\epsilon^{M^\epsilon})$ obtained by applying the renormalization map to the canonical model $(\Pi_\epsilon, \Gamma_\epsilon) = \iota\xi^\epsilon$ (where $\xi^\epsilon = \xi * \rho_\epsilon$ with $\rho_\epsilon(t, x) = \epsilon^{-3}\rho(\epsilon^{-2}t, \epsilon^{-1}x)$), converges as $\epsilon \to 0$ to a limiting (admissible) model $(\Pi^\infty, \Gamma^\infty)$ (see [89, Theorem 4.5]) such that the admissible pre-model coincides with the five stochastic integrals we have defined above and such that $\Phi_0(u^{(\Pi^\infty,\Gamma^\infty)})$ coincides with the Ito-Walsh solution of mSHE (see [89, Corollary 6.5]).

The arguments from [89] described above are much more general and can be used to prove the following stronger fact: for $\delta > 0$ if one replaces ξ, ξ^ϵ by $\delta\xi, \delta\xi^\epsilon$ and if one replaces c by $\delta^2 c$ and $c^{(1)}$ by $\delta^4 c^{(1)}$ respectively, then one also obtains a limiting model $(\Pi^{\infty,\delta}, \Gamma^{\infty,\delta})$ as $\epsilon \to 0$ such that the admissible pre-model coincides with the five stochastic integrals we have defined above with each instance of ξ replaced by $\delta\xi$ and such that $\Phi_0(u^{(\Pi^{\infty,\delta},\Gamma^{\infty,\delta})})$ coincides with the Ito-Walsh solution of mSHE driven by $\delta\xi$. Indeed the correct choice of constants can be verified by setting $G(u) = \delta u$ and $H(u) = 0$ in their paper, which will give precisely these scalings of constants, see Proposition 4.4 and equations (1.4) and (1.7) as described in that proposition. This is enough to prove our claim. One caveat is that [89] works on the torus not on the full line, but this discrepancy is a completely superficial one after the relevant theorems have been proved for weighted spaces as in [88], as explained in Section 5 of [88].

Chapter 6: Limit shape of subpartition-maximizing partitions

This is joint work with Ivan Corwin. This is an expository note answering a question posed to us by Richard Stanley, in which we prove a limit shape theorem for partitions of n which maximize the number of subpartitions. The limit shape and the growth rate of the number of subpartitions are explicit. The key ideas are to use large deviations estimates for random walks, together with convex analysis and the Hardy-Ramanujan asymptotics. Our limit shape coincides with Vershik's limit shape for uniform random partitions.

6.1 Maximizing the number of subpartitions

Given a partition $\lambda = (\lambda_1 \geq ... \geq \lambda_k)$ of n, we can identify it with a 1-Lipschitz function which is a finite perturbation of $|x|$ by following the Russian convention for drawing it. Specifically, start with the English convention for the Young diagram for λ (λ_1 boxes on the top row, then λ_2 below it and so on, all justified to line up on the left) and rotate it by $135°$. Then we place this rotated picture immediately adjacent to the graph of the function $x \mapsto |x|$ so that each box has unit length. This defines a 1-Lipschitz function $g_\lambda(x)$ with the property that $g_\lambda(x) \geq |x|$ and $g_\lambda(x) = |x|$ for large x. We also define a rescaled version of g_λ as $f_\lambda(x) := n^{-1/2} g_\lambda(n^{1/2}x)$ so that each box has side length $n^{-1/2}$ and area n^{-1} when depicted beneath the graph of f_λ. In particular $\int_{\mathbb{R}}(f_\lambda(x) - |x|)dx = 1$.

A *subpartition* of a partition $\lambda = (\lambda_1 \geq ... \geq \lambda_k)$ is a partition $\mu = (\mu_1 \geq ... \geq \mu_\ell)$ such that $\ell \leq k$ and $\mu_i \leq \lambda_i$ for all $i \leq \ell$. Our main result is as follows.

Theorem 6.1.1 (Theorems 6.4.2 and 6.5.2). *For each n, let λ_n denote a partition of n which maximizes the number of subpartitions among all other partitions of n. Then the number of subpartitions of λ_n grows as $e^{\pi\sqrt{2n/3 - o(\sqrt{n})}}$ as $n \to \infty$. Moreover f_{λ_n} converges uniformly as $n \to \infty$*

to the function $f(x) = \frac{2\sqrt{3}}{\pi}\log\left(2\cosh(\frac{\pi}{2\sqrt{3}}x)\right)$.

The limit shape here is known as *Vershik's curve* and was first described as the limit of uniformly sampled partitions of n in [146]. Our result can be shown by using *large-deviations estimates* for uniformly sampled partitions of n which were found in the follow-up paper [56]. In particular, to prove Theorem 6.1.1, first note by the Hardy-Ramanujan asymptotics that the number of subpartitions of any partition of n is bounded above (up to some constant factor) by $e^{\pi\sqrt{2n/3}}$. We let μ_n be a partition of n which is closest to Vershik's curve (after normalization by $\sqrt{n}$), among all other partitions of n. Fixing $\epsilon > 0$, it follows from Theorem 1 of [56] that for large enough n, "most" partitions of $\lceil(1-\epsilon)n\rceil$ are going to be subpartitions of μ_n, which means that the number of subpartitions of μ_n is bounded below by $\frac{1}{n}e^{\pi\sqrt{2(1-\epsilon)n/3}-o(\sqrt{n})}$. Since ϵ can be made arbitrarily small, this gives tight bounds on the exponential scale which can then be used (via elementary topological arguments) to show that the maximizing partitions λ_n are very close to μ_n on the $\sqrt{n}$ scale, so that the λ_n also converge to Vershik's curve.

The main purpose of this note is to exposit the power of large deviations theory in this particular context of partition/subpartition problems. Specifically we are going to give a proof of Theorem 6.1.1, which is essentially a more rigorous version of the sketch given in the preceding paragraph. However, our exposition is more self-contained and based entirely on foundational principles. Specifically we do not use [56] or any other previous work on integer partitions, but instead rely on the seminal result of Mogulskii [119] which gives a large deviations rate function for the full sample path of a random walk with iid increments, and is arguably a central result of large deviations theory.

We also have the following similar result for k-chains of subpartitions, i.e., simply ordered sets of k subpartitions. The ordering may be strict or unstrict; our results do not depend on this convention.

Theorem 6.1.2 (Section 6.6)**.** *Let $k \geq 1$, and let λ_n denote a partition of n which maximizes the number of k-chains of subpartitions, among all other partitions of n. Then the number of k-chains of subpartitions of λ_n grows as $e^{k\pi\sqrt{2n/3}-o(\sqrt{n})}$ as $n \to \infty$. Furthermore f_{λ_n} converges uniformly to the same limit shape as in Theorem 6.1.1.*

We close out this introduction by noting a few questions that may warrant further study. In some cases, there are related results though we do not attempt to make a survey of them.

One natural question is to consider fluctuations around limit curves, as done in [152, 147, 145, 98] for instance. For the problem we have considered, this is a bit difficult to phrase since for each n we expect only a few maximizing partitions. On the other hand, if we let $s(\lambda)$ denote the number of subpartitions of λ, then we may, for $\beta \geq 0$ define a measure on partitions of n with probability of λ proportional to $s(\lambda)^\beta$. When $\beta \to \infty$, this measure concentrates on those λ which maximize $s(\lambda)$, hence our problem. When $\beta = 0$, this measure reduces to the uniform measure on partitions considered by Vershik. While we expect (in particular, based on our arguments in this paper) that the limit shape does not depend on β, it would be interesting to probe the dependence of β on the fluctuations around that shape. It might also be interesting to obtain concentration and large deviations bounds for such measures, as established in [149, 56] for instance.

While there are many other types of measures on partitions, one of particular importance is the Plancherel measure. This involves defining the dimension of λ to be the number of standard Young Tableaux of that shape. In terms of subpartitions, this is the number of n-chains of subpartitions where we restrict that a subpartition cannot equal the partition. The Plancherel measure is then proportional to that dimension squared. For that measure, seminal and independent works of Logan-Shepp [115] and Vershik-Kerov [148] established a limit shape as $n \to \infty$ now known as the Logan-Shepp-Vershik-Kerov (LSVK) curve. This limit curve is not the same as Vershik's curve. Hence, a natural question is to find a way to interpolate the model so as to find limit shapes which likewise interpolate between these two curves.

Theorem 6.1.2 shows that taking k-chains for k fixed does not achieve this aim of crossing over between the Vershik and LSVK curves. However, we speculate that taking $k = k(n) = cn^{1/2}$ may result in such a crossover. In fact, this problem can be reduced to a rhombus tiling limit shape problem for which there are some methods which may be useful. Another natural question involves increasing the dimension and considering higher dimensional partitions. In three dimensions, these would correspond with plane partitions, which are also nicely interpreted as rhombus tilings.

Acknowledgements: The authors are thankful to Greg Martin and Richard Stanley who initiated a conversation on MathOverflow two years ago on this question, and in particular to Richard Stanley who posed this question to the first author of this work. I. Corwin was partially supported by a Packard Foundation Science and Engineering Fellowship as well as NSF grant DMS:1811143 and DMS:1664650. S. Parekh was partially supported by the Fernholz Foundation's "Summer Minerva Fellows" program, as well as summer support from I. Corwin's NSF grant DMS:1811143.

Outline: In Section 6.2 we will derive exponentially sharp upper bounds for the number of nearest-neighbor paths which stay below a given barrier. In Section 6.3 we introduce a certain functional which will describe the limit shape and the growth rate of the maximizing partitions; this functional appears naturally from the upper bounds of Section 6.2. In Section 6.4 we prove the limit shape theorem abstractly (without identifying the limit shape explicitly), by using nice convexity properties of the functional defined in Section 6.3. In Section 6.5 we use Lagrange multipliers and Hardy-Ramanujan asymptotics to derive the limit shape explicitly (thus completing the proof of Theorem 6.1.1). In Section 6.6 we prove Theorem 6.1.2.

6.2 Preliminary upper bounds

First we introduce some notation. Always I will denote a subinterval of $\mathbb{Z}$ or of $\mathbb{R}$. The specific type of interval will always be made clear from the context. For a (continuous) function $f : I \to \mathbb{R}$, we define the **lower convex envelope** of f to be the supremum of all convex functions which are less than or equal to f. Note that this is a convex function, which is also the supremum of a countable number of linear functions which are equal (and in fact tangent, if $I = [0, 1]$) to f at certain special points. We also define the ***decreasing*** **lower convex envelope** to be the sup of all *decreasing* convex functions less than or equal to f, which is a (weakly) decreasing convex function.

Our first lemma is elementary (albeit tedious to state precisely) and says that the lower convex envelope necessarily optimizes a certain type of convex functional over the set of functions less than a given one.

Lemma 6.2.1. *Let $\psi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be a convex function. Let I be the discrete interval $\{a, a+1, ..., b\} \subset \mathbb{Z}$. We let $C(I)$ denote the space of all functions from $I \to \mathbb{R}$. Define a functional $J : C(I) \to \mathbb{R}$ by the formula*

$$J(f) := \sum_{i \in I \setminus \{a\}} \psi\big(f(i) - f(i-1)\big),$$

Fix some $f \in C(I)$, and let $K_f := \{g \in C(I) : g \leq f, g(a) = f(a), g(b) = f(b)\}$. Then one has that $\inf_{g \in K_f} J(g) = J(h)$, where h is the lower convex envelope of f. Similarly, if $\bar{K}_f := \{g \in C(I) : g \leq f, g(a) = f(a)\}$, and if we also assume that ψ achieves its minimum at 0, then $\inf_{g \in \bar{K}_f} J(g) = J(\bar{h})$, where $\bar{h}$ is the decreasing lower convex envelope of f.

Proof. We will work with K_f rather than $\bar{K}_f$, briefly indicating the necessary modifications at the end of the proof. The argument is essentially a geometric one which proceeds in two steps.

Step 1. Firstly, we show that $J(f) \geq J(h)$ whenever $f(a) = h(a)$, $f(b) = h(b)$, and h is the lower convex envelope of f. Let $C := \{x \in I : f(x) = h(x)\}$. The complement of C is the union of some finite collection of disjoint intervals $\bigcup_n^N (a_n, b_n) \cap \mathbb{Z}$. On each interval $(a_i, b_i) \cap \mathbb{Z}$ it is clear from the definition of the lower convex envelope that h is just a linear function, i.e.,

$h(x) = \frac{x-a_n}{b_n-a_n} f(b_n) + \frac{b_n-x}{b_n-a_n} f(a_n)$ for $x \in [a_n, b_n]$. By Jensen's inequality, one sees that

$$\sum_{a_n+1}^{b_n} \psi(f(i) - f(i-1)) \geq (b_n - a_n)\psi\Big(\frac{f(b_n) - f(a_n)}{b_n - a_n}\Big) = \sum_{a_n+1}^{b_n} \psi(h(i) - h(i-1)).$$

This is already enough to prove Step 1, since f coincides with h outside of the $[a_n, b_n]$.

Step 2. Secondly, we show that $J(h) \geq J(k)$ whenever h, k are both convex functions with the property that $h(a) = k(a)$, $h(b) = k(b)$, and $h \leq k$. To do this, we inductively define a sequence $\{h_j\}_{j=a}^b$ of functions: $h_a = h$, and

$$h_{j+1}(x) = \max\{h_j(x), (x - j + 1)k(j) + (j - x)k(j-1)\}.$$

In more geometric terms, we are simply taking h_{j+1} to be the maximum of h_j with the "tangent line" to k at $\{j-1, j\}$. In particular each h_j is convex, and it follows from convexity of k that $h_b = k$. Thus the claim will be proved if we can show that $J(h_j) \geq J(h_{j+1})$ for all $j \in \{a, ..., b-1\}$. But this is clear, because $h_j(x)$ agrees with $h_{j+1}(x)$ except for x in some interval $[u, v]$ where it equals $\frac{x-u}{v-u} h_j(v) + \frac{v-x}{v-u} h_j(u)$. Hence the same argument from Step 1 (using Jensen's inequalty) applies to show $J(h_j) \geq J(h_{j+1})$. This completes the proof of step 2.

Step 1 and Step 2 easily imply the claim because if $g \leq f$ with $g(a) = f(a)$ and $g(b) = f(b)$, and if $h \leq k$ are their respective lower convex envelopes then we have that $J(g) \geq J(h) \geq J(k)$, where the first inequality is from Step 1 and the second is from Step 2.

Now suppose we replace K_f by $\bar{K}_f$. Let $c \in \{a, ..., b\}$ be the point at which f achieves its minimum value. Let h and $\bar{h}$ denote the lower convex envelope and decreasing lower convex envelope (respectively) of f. Note that $h = \bar{h}$ on $\{a, ..., c\}$, and $h(c) = f(c)$, and therefore if $g \leq f$ then the above argument gives $\sum_{a+1}^{c} \psi(g(i) - g(i-1)) \geq \sum_{a+1}^{c} \psi(\bar{h}(i) - \bar{h}(i-1))$. On the other hand, note that $\bar{h}(x) = f(c)$ for $x \in \{c, ..., b\}$, and thus by assuming that ψ achieves its minimum at 0, we get that $\sum_{c+1}^b \psi(\bar{h}(i)-\bar{h}(i-1)) = \sum_{c+1}^b \psi(0) \le \sum_{c+1}^b \psi(g(i)-g(i-1))$, as desired. $\square$

Lemma 6.2.2. *Let $f:\{0,...,n\}\to\mathbb{R}$ with $f(0)=0$. Let S denote a simple symmetric nearest-neighbor random walk on $\mathbb{Z}$. Also, let g denote the decreasing lower convex envelope of f. We also let Λ^* be the large deviation rate function associated with S, which means that Λ^* is the Legendre transform of $\lambda\mapsto \log\mathbb{E}[e^{\lambda S_1}]$. Then*

$$\mathbb{P}\big(S_i \le f(i), \forall i \le n\big) \le e^{-\sum_{i=1}^n \Lambda^*(g(i)-g(i-1))}. \tag{6.1}$$

Proof. The proof uses a standard method for obtaining LDP upper bounds [58]. Note that for real numbers $\lambda_1,...,\lambda_n$, and any Borel set $C\subset\mathbb{R}^n$,

$$\inf_{x\in C} e^{\sum_1^n \lambda_i(x_i-x_{i-1})}\mathbb{P}(S\in C) \le \mathbb{E}[e^{\sum_1^n \lambda_i(S_i-S_{i-1})}] = e^{\sum_{i=1}^n \Lambda(\lambda_i)},$$

where $\Lambda(\lambda)=\log\mathbb{E}[e^{\lambda S_1}]$ and we impose that $x_0:=0$ in the relevant sum. Rearranging this gives us

$$\mathbb{P}(S\in C)\le e^{-\inf_{x\in C}\sum_1^n \lambda_i(x_i-x_{i-1})-\Lambda(\lambda_i)}.$$

Now we optimize over all $\lambda_1,...,\lambda_n$. If we assume that C is compact and convex we can use the minimax theorem for concave-convex functions [138] to interchange the sup over λ with the inf over x, specifically

$$\begin{aligned}\mathbb{P}(S\in C) &\le e^{-\sup_{\lambda\in\mathbb{R}^n}\inf_{x\in C}\sum_1^n \lambda_i(x_i-x_{i-1})-\Lambda(\lambda_i)} \\ &= e^{-\inf_{x\in C}\sup_{\lambda\in\mathbb{R}^n}\sum_1^n \lambda_i(x_i-x_{i-1})-\Lambda(\lambda_i)} \\ &= e^{-\inf_{x\in C}\sum_1^n \sup_{\lambda\in\mathbb{R}}\big(\lambda(x_i-x_{i-1})-\Lambda(\lambda)\big)} \\ &= e^{-\inf_{x\in C}\sum_1^n \Lambda^*(x_i-x_{i-1})}.\end{aligned} \tag{6.2}$$

Now we let $C=\{x\in\mathbb{R}^n : -i\le x_i\le f(i), \forall i\}$, which is clearly compact and convex. Note that $S\in C$ is equivalent to the left-hand side of (6.1) (owing to the fact that S only takes ±1

sized jumps). Applying (6.2) and using Lemma 6.2.1 to show that $\inf_{x\in C}\sum_1^n \Lambda^*(x_i - x_{i-1}) = \sum_1^n \Lambda^*(g(i)-g(i-1))$, we arrive at (6.1). $\square$

Corollary 6.2.3. *Let $f:\{0,...,n\}\to\mathbb{R}$ with $f(0)=f(n)=0$, and let g denote the lower convex envelope of f (not the decreasing one). Then the number of nearest neighbor bridges which stay below f (i.e., functions $\gamma:\{0,...,n\}\to\mathbb{Z}$ such that $\gamma(0)=\gamma(n)=0$, and $|\gamma(i)-\gamma(i-1)|=1$, and $\gamma(i)\leq f(i)$ for all i) is bounded above by $2^n e^{-\sum_{i=1}^n \Lambda^*(g(i)-g(i-1))}$.*

Proof. Let us pick a point $k\in\{0,...,n\}$ at which g attains its minimum value. Note that $g(k) = f(k)$. Note by Lemma 6.2.2 the number of nearest neighbor paths of length k starting from 0 and lying below $f|_{\{0,...,k\}}$ is less than or equal to $2^k e^{-\sum_1^k \Lambda^*(g(i)-g(i-1))}$. Similarly the number of nearest neighbor paths of length $n-k$ starting from 0 and lying below $f|_{\{k+1,...,n\}}$ is less than or equal to $2^{n-k}e^{-\sum_{k+1}^n \Lambda^*(g(i)-g(i-1))}$. Note that the number of bridge paths of length n lying below f is less than the number of pairs of paths (γ,γ') where γ is of the former type and γ' is of the latter type. Thus the total number of such bridges is bounded above by product of the two individual upper bounds, which equals $2^n e^{-\sum_1^n \Lambda^*(g(i)-g(i-1))}$. $\square$

An important thing to keep in mind is that the bounds of Propositions 6.2.2 and 6.2.3 are actually sharp up to some subexponential decay factor (see Section 6.4). At an intuitive level, what this says is that if we condition a random walk to stay underneath a fixed barrier, then the path which minimizes the energy of the random walk is none other than the lower convex envelope of that barrier. Another thing to keep in mind is that the bounds of this section hold *uniformly* over all partitions, which makes them a little bit stronger than ordinary LDP upper bounds.

6.3 The functional describing the limit shape

For a partition λ, one recalls the definitions of f_λ and g_λ given at the beginning of Section 6.1. A 1-Lipschitz function will always refer to a real-valued function f with the property that $|f(x) - f(y)|\leq |x-y|$, or equivalently f is absolutely continuous and $|f'|\leq 1$.

Let us now estimate (or at least upper bound) the number of subpartitions of a given partition. Each subpartition of a given λ can be interpreted as a trajectory of a simple symmetric random walk *bridge* which stays below the graph of g_λ (or alternatively of f_λ after rescaling). By Corollary 6.2.3, the number of such bridges can be upper bounded quite easily. Specifically let h_λ denote the lower convex envelope of f_λ, and let k denote a large enough integer so that $g_\lambda(x) = |x|$ whenever $|x| \geq k$. Then by Corollary 6.2.3 we know that the number of subpartitions of λ (i.e., the number of unit-length random walk bridges which lie in between the graphs of $g_\lambda(x)$ and $|x|$) is upper bounded by

$$\begin{aligned} &2^{2k} e^{-\sum_{i=-k}^{k} \Lambda^*\left(n^{1/2}\left[h_\lambda(n^{-1/2}i) - h_\lambda(n^{-1/2}(i-1))\right]\right)} \\ &= e^{\sum_{-k}^{k}\left[\log 2 - \Lambda^*\left(n^{1/2}\left[h_\lambda(n^{-1/2}i) - h_\lambda(n^{-1/2}(i-1))\right]\right)\right]} = e^{\sqrt{2n}\int_{\mathbb{R}} \phi(h_\lambda'(x))dx}, \end{aligned} \tag{6.3}$$

where in the final equality we are using the piece-wise linearity of h_λ and defining $\phi(x) := \log 2 - \Lambda^*(x)$. This function ϕ will be very important in the ensuing analysis. In particular, note that $\phi(x)$ is a concave and even function defined on $[-1, 1]$ which achieves its maximum value of $\log 2$ at $x = 0$, and its minimum of 0 at $x = \pm 1$. The explicit expression for ϕ is given in Definition 6.4.1 below.

The functional $f \mapsto \int_{\mathbb{R}} \phi \circ f'$ appearing in (6.3) will describe the optimal rate of growth of the number of subpartitions, as we will show in the following section. Therefore the remainder of this section will be devoted to analyzing this functional. To start, we make the following important definition:

Definition 6.3.1. *We define $\mathcal{X}$ to be the space of all 1-Lipschitz functions $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) \geq |x|$ and furthermore $\int_{\mathbb{R}}(f(x) - |x|)dx \leq 1$. We equip $\mathcal{X}$ with the topology of uniform convergence on compact sets. Furthermore, we define the functional $F : \mathcal{X} \to \mathbb{R}_+$ by $F(h) := \int_{\mathbb{R}} \phi \circ h'$, where*

$$\phi(x) := \log 2 - \left(\frac{1}{2}\log(1 - x^2) + x \tanh^{-1} x\right),$$

and h' is the derivative of h.

A few remarks are in order about this definition. Firstly, note that $\mathcal{X}$ is a compact space. Indeed, this is a consequence of Arzela Ascoli: equicontinuity is obvious, and pointwise boundedness follows from the integral condition on elements of $\mathcal{X}$ combined with the 1-Lipschitz property (in fact any $f \in \mathcal{X}$ is bounded above by $x \mapsto \sqrt{x^2+2}$, since this curve is the locus of all points such that the rectangle which has one vertex at that point and another one at the origin, and is also adjacent to the graph of $|x|$, has area exactly 1).

Secondly, we remark that even though we equipped $\mathcal{X}$ with the topology of uniform convergence on compact sets, this convergence is actually equivalent to uniform convergence on all of $\mathbb{R}$. This once again follows from the fact that for all $f \in \mathcal{X}$ one has that $|x| \le f(x) \le \sqrt{x^2+2}$, and also because of the fact that $\sqrt{x^2+2} - |x| \to 0$ as $|x| \to \infty$. In particular it is true that $\mathcal{X}$ is a complete metric space with respect to the uniform metric

$$d(f,g) = \sup_{x\in\mathbb{R}} |f(x) - g(x)|.$$

The completeness is a consequence of Fatou's Lemma (to ensure that the value of the integral remains ≤ 1 after taking a limit). This metric will be used very briefly in the proof of a later theorem (6.4.2).

Thirdly, it is not immediately clear that the integral defining the functional $F(f)$ actually converges for every $f \in \mathcal{X}$, but this will be taken care of by the following proposition which also highlights the nicest and most important property of F, and will crucially be used later.

Proposition 6.3.2. *The integral defining the functional F converges for every $f \in \mathcal{X}$. Furthermore, F is upper semicontinuous on $\mathcal{X}$.*

Proof. We will prove that if f_n is a family of 1-Lipschitz functions such that $f_n \to f$ uniformly, then $\limsup_{n\to\infty} F(f_n) \le F(f) < \infty$. The key difficulty here is that F is defined on functions on the whole real line, which is not compact. The proof will therefore proceed in two steps: first we replace $\mathbb{R}$ with a large compact interval and prove the upper semicontinuity in this simpler case; second we prove a certain “tightness" property (6.5) for functions in $\mathcal{X}$ which will simultaneously also show that the integral defining $F(f)$ necessarily converges for all $f \in \mathcal{X}$.

The first step is to show that for each fixed (large) $A > 0$ one has that

$$\limsup_{n\to\infty} \int_{[-A,A]} \phi \circ f_n' \le \int_{[-A,A]} \phi \circ f'. \tag{6.4}$$

The proof of this is quite standard, and purely topological (e.g., does not rely on properties of the space $\mathcal{X}$). Nevertheless we include a proof of (6.4) for completeness.

For simplicity, let us replace the interval $[-A, A]$ by $[0, 1]$ (the same argument works in the former case with some extra scaling factors). Let $\mathcal{X}[0, 1]$ denote the space of 1-Lipschitz functions on $[0, 1]$ equipped with the uniform topology. We will show that the functional $G(f) := \int_0^1 \phi \circ f'$ is upper semicontinuous from $\mathcal{X}[0, 1] \to \mathbb{R}$. To prove this it suffices to write G as the infimum of some collection of continuous functionals. To do this, we consider partitions $\mathcal{P} = (0 \le t_1 \le ... \le t_n \le 1)$ of $[0, 1]$, and we define $G_{\mathcal{P}}(f) := \sum_1^n (t_i - t_{i-1})\phi\big(\frac{f(t_i)-f(t_{i-1})}{t_i-t_{i-1}}\big)$. It is then clear that each $G_{\mathcal{P}}$ is continuous from $\mathcal{X}[0, 1] \to \mathbb{R}$. We then claim that $G = \inf_{\mathcal{P}} G_{\mathcal{P}}$ (where the infimum is taken over all partitions of $[0, 1]$) which would prove upper semicontinuity. To prove this equality, first note by Jensen's inequality and concavity of ϕ that for all $a < b$ and all f one has $\int_a^b \phi \circ f' \le (b - a)\phi\big(\frac{f(b)-f(a)}{b-a}\big)$, which proves that $G \le \inf_{\mathcal{P}} G_{\mathcal{P}}$. To prove the other direction, we define the partition $\mathcal{P}_n$ to be the one consisting of dyadic intervals $[k2^{-n}, (k + 1)2^{-n})$ with $0 \le k \le 2^n - 1$. For a 1-Lipschitz function f let f_n denote the continuous function with $f_n(0) = 0$ and whose derivative $f_n'(x)$ takes the constant value $2^n\big(f((k + 1)2^{-n}) - f(k2^{-n})\big)$ for $x \in [k2^{-n}, (k + 1)2^{-n})$. Note that f_n' forms a bounded martingale with respect to the dyadic filtration on the probability space $[0, 1]$ (i.e., the filtration associated with the nested family of partitions $\{\mathcal{P}_n\}_n$). Consequently f_n' converges to f' a.e, and thus $\phi \circ f_n' \to \phi \circ f'$ a.e. Hence by the bounded convergence theorem we conclude that $G_{\mathcal{P}_n}(f) = \int_0^1 \phi \circ f_n' \to \int_0^1 \phi \circ f' = G(f)$. This shows that $G \geq \inf_{\mathcal{P}} G_{\mathcal{P}}$. This proves upper semicontinuity of G and in turn also proves (6.4).

Now given that (6.4) holds, we want to take $A \to \infty$, but this involves a nontrivial interchange of limits and this is where noncompactness of the real line gets in the way. So now we actually need to use special properties of the space $\mathcal{X}$.

We will show that for every $\epsilon > 0$, there exists some $A = A(\epsilon) > 0$ (large) so that for all $f \in \mathcal{X}$ one has that

$$\int_{\mathbb{R}\setminus[-A,A]} \phi \circ f' < \epsilon. \tag{6.5}$$

Note that together with (6.4), this is enough to complete the proof that $\limsup_n F(f_n) \leq F(f)$. The key here is, of course, that the bound of (6.5) is uniform over *all* functions $f \in \mathcal{X}$. Note that (6.5) also shows that $F(f) < \infty$ for all $f \in \mathcal{X}$.

To prove (6.5), we first note that if f is 1-Lipschitz, then $f(x) - x$ is necessarily (weakly) decreasing for $x \geq 0$, thus

$$1 - f(n) + f(n-1) = \big(f(n-1) - (n-1)\big) - \big(f(n) - n\big) \geq 0 \quad \text{for } n \geq 1. \tag{6.6}$$

The condition that $\int_{\mathbb{R}} (f(x) - |x|)dx \leq 1$ shows that $\sum_{n\geq 0} f(n) - n \leq 3$ (e.g., via an integral comparison test, since we know $f(n) - n$ is decreasing and $f(0) \leq \sqrt{2} < 2$). Then for all $N \geq 1$ we find that

$$\sum_{n=1}^{N} n(1 - f(n) + f(n-1)) = \sum_{n=1}^{N}\sum_{k=1}^{n}(1 - f(n) + f(n-1))$$
$$= \sum_{k=1}^{N}\sum_{n=k}^{N}(1 - f(n) + f(n-1)) = \left[\sum_{k=1}^{N} f(k-1) - (k-1)\right] - N(f(N) - N),$$

where in the last line we used (6.6) so that the inner sum telescopes. Since $N(f(N)-N) \geq 0$ we can upper bound the last expression by $\sum_{k\geq 0}(f(k)-k)$. Hence we can let $N \to \infty$ in the preceding expression and we see that

$$\sum_{n\geq 1} n(1-f(n)+f(n-1)) \leq \sum_{k\geq 0} f(k)-k \leq 3. \tag{6.7}$$

Appealing to the definition of $\phi(x)$ we see that in $(-1,1)$, $\phi'(x) = -\tanh^{-1} x$ which has logarithmic singluarities at ± 1. Thus, it follows that ϕ asymptotically looks like $x|\log x|$ near $x = \pm 1$, i.e., $\lim_{x\to\pm 1} \frac{\phi(x)}{|x\mp 1| \log |x\mp 1|}$ will be a finite nonzero value. Since $|\log x| \leq Cx^{-1/3}$ near $x = 0$, this implies that there exists some $C > 0$ such that $\phi(x) \leq C(1-|x|)^{2/3}$ for all $x \in [-1,1]$. In particular, for all $A \geq 0$ one has

$$\begin{aligned}
\sum_{n\geq A} \phi\big(f(n)-f(n-1)\big) &\leq C \sum_{n\geq A} \big(1-f(n)+f(n-1)\big)^{2/3} \\
&\leq C\Big(\sum_{n\geq A} n^{-2}\Big)^{1/3} \Big(\sum_{n} n(1-f(n)+f(n-1))\Big)^{2/3} \\
&\leq C \cdot A^{-1/3} \cdot 3^{2/3}.
\end{aligned}$$

For the second inequality, note that if a_n are nonnegative real numbers such that $\sum_n na_n < \infty$, then by Holder's inequality $\sum_{n\geq A} a_n^{2/3} \leq \big(\sum_{n\geq A} na_n\big)^{2/3}\big(\sum_{n\geq A} n^{-2}\big)^{1/3}$. The final inequality uses the bound derived in (6.7), as well as $\sum_{n\geq A} n^{-2} \leq A^{-1}$.

To finish our proof, observe that Jensen's inequality and the concavity of ϕ show that $\int_{[n-1,n]} \phi \circ f' \leq \phi(f(n)-f(n-1))$. This, together with the preceding arguments, then shows that

$$\int_A^\infty \phi \circ f' \leq \sum_{n=A}^{\infty} \phi\big(f(n)-f(n-1)\big) \lesssim A^{-1/3},$$

independently of f, which finally proves (6.5). $\square$

At this point it is important to remark that Proposition 6.3.2 is **not** just some technical and otherwise unimportant intermediate step. Really it is where the "meat" of the proof of the limit shape (Theorem 6.1.1) really lies. Specifically, the important thing here is the second half of the proof where we prove a type of "tightness" estimate (6.5) for functions in $\mathcal{X}$. In terms of partitions, what it really shows (in an equivalent formulation) is that the sequence of partitions maximizing the number of subpartitions, stays bounded on the $n^{1/2}$ scale, i.e., that the sequence f_{λ_n} from Theorem 6.1.1 does not lose any mass in the limit (meaning that any subsequential limit f of f_{λ_n} satisfies $\int(f(x)-|x|)dx = 1$). We remark that the bound $A^{-1/3}$ appearing at the end of the proof may actually be improved optimally to $\frac{\log A}{A}$, but this is slightly more difficult.

As a corollary of Proposition 6.3.2, we can combine it with compactness of the space $\mathcal{X}$ in order to obtain the following key result.

Corollary 6.3.3. *The functional F from Definition 6.3.1 admits a maximum $M(F)$ on the space $\mathcal{X}$. There is a unique function f at which the maximum is attained and this maximizer f is a convex and symmetric function (i.e. $f(x) = f(-x)$) and moreover $\int_{\mathbb{R}}(f(x)-|x|)dx = 1$.*

Proof. Any upper semicontinuous function on a compact space achieves its maximum.

The uniqueness of the maximizer is a concavity property. Specifically we note that ϕ is a strictly concave function, meaning $\phi((1-t)a+tb) > (1-t)\phi(a)+t\phi(b)$ whenever $t \in (0,1)$ and $a \neq b$. This then easily implies that $F((1-t)f+tg) > (1-t)F(f)+tF(g)$ for $t \in (0,1)$ and $f \neq g$. Clearly this rules out the existence of two distinct maxima.

Symmetry is another consequence of concavity. Specifically, if the maximizer f was not symmetric, then we can define its reflection $f_s(x) := f(-x)$. Clearly $F(f_s) = F(f)$ and thus if $f \neq f_s$ then as above we have that $F(\frac{1}{2}f + \frac{1}{2}f_s) > \frac{1}{2}F(f) + \frac{1}{2}F(f_s) = F(f)$, which is a contradiction.

Let f be the maximizer. To prove that $\int(f(x)-|x|)dx = 1$, suppose that this integral took some value $\alpha < 1$. Then we let $h(x) = \alpha^{-1/2}h(\alpha^{1/2}x)$. Clearly $\int(h(x) - |x|)dx = 1$, and h is 1-Lipschitz. Moreover a simple substitution reveals that $F(h) = \alpha^{-1/2}F(f) > F(f)$ which is a contradiction.

To prove convexity of f, suppose (for contradiction) that a, b are two points of $\mathbb{R}$ such that there is a linear function ℓ equal to f at both a and b, and such that $\ell < f$ on (a, b). We define h to be equal to f on $\mathbb{R}\backslash[a, b]$, and equal to ℓ on $[a, b]$. Then by Jensen's inequality one has that $\int_a^b \phi \circ f' < (b-a)\phi\big(\frac{f(b)-f(a)}{b-a}\big) = \int_a^b \phi \circ h'$, which means that $F(f) < F(h)$; a contradiction. This completes the proof. □

6.4 The limit shape Theorem

Note that in (6.3) we already proved that for any sequence λ_n of partitions of n, the number of subpartitions is bounded above by $e^{\sqrt{2n}M(F)}$ where $M(F)$ is the maximum value of the functional F from above. A natural question is whether there exists a sequence of partitions for which the number of subpartitions actually grows at this optimal rate. It turns out that the answer is yes (up to some subexponential factor which is irrelevant), which retrospectively justifies why we performed such an in-depth analysis of the functional F in the first place.

Proposition 6.4.1. *There exists a sequence of partitions μ_n of n such that the number of subpartitions of μ_n actually grows as $e^{\sqrt{2n}M(F)-o(\sqrt{n})}$ as $n \to \infty$.*

The key behind proving this proposition is Mogulskii's theorem [119], which is really the primary underlying idea behind this entire work. This result essentially says that the bound in (6.3) (and also in Propositions 6.2.2 and 6.2.3) is actually *sharp* (again, up to some subexponential factor which is not relevant to us). But before getting to the proof, let us first prove the following important corollary.

Theorem 6.4.2 (Limit shape theorem)**.** *Let λ_n and f_{λ_n} be as in Theorem 6.1.1. As $n \to \infty$, the sequence f_{λ_n} converges uniformly to the unique maximizer $f_{\max}$ of the functional F from Definition 6.3.1.*

Proof. Let $s(\lambda_n)$ denote the number of subpartitions of λ_n, and let h_{λ_n} denote the lower convex envelopes of f_{λ_n}. By Proposition 6.4.1 and equation (6.3) we have that

$$e^{\sqrt{2n}M(F)-o(\sqrt{n})} \leq s(\lambda_n) \leq e^{\sqrt{2n}F(h_{\lambda_n})} \leq e^{\sqrt{2n}M(F)}, \qquad \text{as } \ n \to \infty,$$

which means that $M(F) - o(1) \leq F(h_{\lambda_n}) \leq M(F)$ as $n \to \infty$.

Thus we see that $F(h_{\lambda_n}) \to M(F)$ as $n \to \infty$. This is already enough to imply that $h_{\lambda_n} \to f_{\max}$ uniformly on compact sets as $n \to \infty$. Indeed it is true that for every $\epsilon > 0$ there exists $\delta > 0$ such that (for all $f \in \mathcal{X}$) $F(f) > M(F) - \delta$ implies that $d(f, g) < \epsilon$ (here d denotes the metric on $\mathcal{X}$ which was specified following Definition 6.3.1). If this was not the case then we can choose an $\epsilon > 0$ such that $\sup_{d(f_{\max},g)\geq\epsilon} F(g) = M(F)$. But the space $\mathcal{A}$ of 1-Lipschitz functions g such that $d(f_{\max}, g) \geq \epsilon$ is again a compact subset of $\mathcal{X}$ (being a closed subset of $\mathcal{X}$). Furthermore F is still an upper semicontinuous function on $\mathcal{A}$, hence it achieves its maximum value which we already know is $M(F)$. Then there exists some $g_{\max} \in \mathcal{A}$ such that $F(g_{\max}) = M(F)$, which clearly contradicts uniqueness of the maximizer since $d(f_{\max}, g_{\max}) \geq \epsilon$ by construction.

So we have proved that the convex envelopes h_{λ_n} (though not necessarily the functions f_{λ_n} themselves) converge uniformly to $f_{\max}$. Note that since $f_\lambda \geq h_\lambda$ (by definition of the convex envelope) we have

$$\begin{aligned}
\int_{\mathbb{R}} |f_{\lambda_n} - h_{\lambda_n}| &= \int_{\mathbb{R}} (f_{\lambda_n}(x) - h_{\lambda_n}(x))dx \\
&= \int_{\mathbb{R}} \big((f_{\lambda_n}(x) - |x|) - (h_{\lambda_n}(x) - |x|)\big)dx \\
&= 1 - \int_{\mathbb{R}} (h_{\lambda_n}(x) - |x|)dx.
\end{aligned}$$

Now h_{λ_n} converges to $f_{\max}$ and by Corollary 6.3.3 we know that $\int (f_{\max}(x) - |x|)dx = 1$, therefore by applying the preceding calculation and then Fatou's Lemma, we see that

$$\limsup_n \int_{\mathbb{R}} |f_{\lambda_n} - h_{\lambda_n}| = 1 - \liminf_n \int_{\mathbb{R}} (h_{\lambda_n}(x) - |x|)dx \leq 1 - \int (f_{\max}(x) - |x|)dx = 0.$$

Therefore $\|f_{\lambda_n} - h_{\lambda_n}\|_{L^1(\mathbb{R})} \to 0$, and since all functions are 1-Lipschitz, this L^1 convergence also implies uniform convergence. $\square$

Although this abstractly proves convergence to some limit shape, we still do not know anything about what the limit shape looks like geometrically. For instance is it bounded, and if so, is it a triangular shape or something more complicated? This question will be addressed in the following section.

Let us now start to get to the proof of Proposition 6.4.1. As mentioned before, the key is the following result, which essentially gives matching lower bounds to the upper bounds which we gave in Section 6.2. A proof may be found in Theorem 5.1.2 of [58] or in the original paper [119].

Theorem 6.4.3 (Mogulskii 1992). *Let μ_n denote the law on $C_0[0,1]$ of $(\frac{1}{n}S_{nt})_{t\in[0,1]}$ where S is any i.i.d. random walk (whose increment distribution has exponential moments), and the values of S at non-integer points are understood to be linearly interpolated from the two nearest integer points. Then μ_n satisfies an LDP with rate n and good rate function*

$$I(f) = \int_0^1 \Lambda^* \circ f',$$

where Λ^* denotes the Legendre transform of $\lambda \mapsto \log \mathbb{E}[e^{\lambda S_1}]$, and the integral is meant to be understood as $+\infty$ if f is not absolutely continuous.

It should be noted that Mogulskii's result is a vast strengthening of Cramer's theorem (from just the endpoint of an iid sample path to its entire history), in the same way that Donsker's invariance principle for iid random walks is a strengthening of the classical central limit theorem. Finally we are ready to prove Proposition 6.4.1.

Proof of Proposition 6.4.1. Let $f_{\max}$ be the maximizer from Corollary 6.3.3. We choose a sequence μ_n of partitions of n such that f_{μ_n} converges uniformly to $f_{\max}$. This can be done as follows. First we construct an intermediate partition $\tilde{\mu}_n$ by putting boxes of side length $n^{-1/2}$ beneath the graph of of $f_{\max}$ until no more boxes can be put in such a way that the graph of $f_{\tilde{\mu}_n}$ remains below that of f. Since $f_{\tilde{\mu}_n} \leq f$, one notices that $\tilde{\mu}_n$ will not actually be a partition of n but rather of some number $k(n) \leq n$. However, it is true that $|f_{\tilde{\mu}_n} - f_{\max}| \leq Cn^{-1/2}$ for some constant independent of n (otherwise more boxes could be added to $\tilde{\mu}_n$ without eclipsing the graph of $f_{\max}$). Now we can define μ_n to be equal to $\tilde{\mu}_n$ but with the remaining $n - k(n)$ boxes added to the first column of $\tilde{\mu}_n$. This will not change the limiting function $f_{\max}$.

We define $f_\delta(x) := \max\{|x|, f_{\max}(x) - \delta\}$, and we define the *support* of f_δ to be the set of x where $f_\delta(x) > |x|$ (this is an interval centered at 0, by convexity and symmetry of $f_{\max}$). Note that for large enough values of n, the $\delta/2$ neighborhood of f_δ lies strictly below f_{μ_n} on the support of f_δ (this is because $f_{\mu_n} \to f_{\max}$ uniformly). We are now going to consider nearest-neighbor (random walk) paths of grid-size $n^{-1/2}$ which lie in between the graphs of $f_{\delta/2}$ and $f_{3\delta/2}$. Such a path will be called (δ, n)-admissible. Let $k = k(n, \delta)$ denote the positive integer such that $n^{-1/2}k = \text{argmin}_{y \in \frac{1}{\sqrt{n}}\mathbb{Z}}|y - a|$ where $a = a(\delta) := \inf\{x > 0 : f_\delta(x) = x\}$.

Note by Mogulskii's Theorem that the number of (δ, n)-admissible paths terminating on the vertical axis (i.e., nearest-neighbor functions $\gamma : n^{-1/2}\mathbb{Z}_{\leq 0} \to n^{-1/2}\mathbb{Z}$ where $\mathbb{Z}_{\leq 0}$ denotes non-positive integers) is greater or equal to $2^k e^{-\sqrt{2n}\int_{-n^{-1/2}k}^0 \Lambda^* \circ f_\delta' - o(\sqrt{n})} = e^{\sqrt{2n}\int_{-\infty}^0 \phi \circ f_\delta' - o(\sqrt{n})}$, as $n \to \infty$ (with δ fixed).

Now we notice that two independent such random walks which are started from $(-n^{-1/2}k, n^{-1/2}k)$ and conditioned to stay between $f_{\delta/2}$ and $f_{3\delta/2}$ have probability at least $\frac{1}{\delta\sqrt{n}}$ of terminating at the same point. Indeed, this is because there are at most $\delta\sqrt{n}$ possible points $\{x_i\}_{i=1}^{\delta\sqrt{n}}$ at which such a walk can terminate (because the grid size is $n^{-1/2}$), and if p_i is the probability of terminating at point x_i, then by Cauchy-Schwarz one finds that $1 = \sum_1^{\delta\sqrt{n}} p_i \leq (\sum_i p_i^2)^{1/2}(\delta\sqrt{n})^{1/2}$, and because the probability of two independent such walks terminating at the same point equals $\sum_i p_i^2$.

Now, a random walk *bridge* of grid size $n^{-1/2}$ which lies between $f_{\delta/2}$ and $f_{3\delta/2}$ (which defines a subpartition of μ_n for large enough n) can be viewed as the concatenation of a pair of these random walk paths started from $(-n^{-1/2}k, n^{-1/2}k)$ terminating at the same point on the vertical axis (note here that we are using the property that $f_\delta(x) = f_\delta(-x)$). By the observations of the preceding two paragraphs, the number of such pairs is bounded below by $\frac{2}{\delta\sqrt{n}}\left(e^{\sqrt{2n}\int_{-\infty}^{0}\phi\circ f_\delta' - o(\sqrt{n})}\right)^2$. The prefactor $\frac{2}{\delta\sqrt{n}}$ may be absorbed into the $o(\sqrt{n})$ term in the exponent, giving a lower bound of $e^{\sqrt{2n}F(f_\delta)-o(\sqrt{n})}$. The $o(\sqrt{n})$ term may depend on δ but this is not a problem.

Since this lower bound holds true for arbitrary $\delta > 0$, the claim now follows if we can show that $F(f_\delta) \to F(f_{\max})$ as $\delta \to 0$. To do this, note that $f_\delta' \to f_{\max}'$ pointwise (trivially by the definition of f_δ). Thus by Fatou's Lemma and maximality of $F(f_{\max})$ it is true that $F(f_{\max}) \leq \liminf_{\delta\to 0} F(f_\delta) \leq \limsup_\delta F(f_n) \leq \max_g F(g) = F(f_{\max})$, which completes the argument. $\square$

6.5 Characterizing the limit shape

So far, many of our methods could have been used for more general types of models than the simple symmetric random walk (replacing ϕ with a more general concave function). We now move onto trying to find the limit shape $f_{\max}$ exactly, which will involve working with specific details of the function ϕ, and thus most of the subsequent arguments and analysis will be specialized just to the case of the simple random walk. In particular we will show that $f_{\max}$ has, up to scaling and centering, the shape of the curve $x \mapsto \log\cosh x$. In particular it is not just the triangular function $x \mapsto \max\{1, |x|\}$, nor is it the Vershik-Kerov curve. It is, in fact, the Vershik curve which is the limit shape of *uniformly* sampled partitions of n [146].

Since $f_{\max}$ is an even convex function there exists a *maximal* interval $(-a_{max}, a_{max})$ (which we will henceforth refer to as the **support** of $f_{\max}$) on which $f(x) > |x|$. This interval is the interior of the largest closed interval containing the support (in the usual sense) of the second distributional derivative $f''_{\max}$ (which is a nonnegative Borel measure since $f_{\max}$ is convex). Note that it is possible that $a_{max} = +\infty$, and in a moment we will show that this is indeed the case.

Let ψ be a smooth function with support contained in $(-a_{max}, a_{max})$, such that $\int_{\mathbb{R}} \psi = 0$. Then we claim

$$\int_{\mathbb{R}} (\phi' \circ f'_{\max}) \cdot \psi' = 0. \tag{6.8}$$

Indeed, one easily checks that $\lim_{\epsilon \to 0} \epsilon^{-1} \big(F(f_{\max} + \epsilon\psi) - F(f_{\max}) \big) = \int_{\mathbb{R}} (\phi' \circ f'_{\max}) \cdot \psi'$. However, since $\int \psi = 0$ and since the support of ψ is contained in the support of $f_{\max}$, it follows that for ϵ in a small enough neighborhood of zero, the function $f_{\max} + \epsilon\psi$ is an element of $\mathcal{X}$, and thus $F(f_{\max} + \epsilon\psi) \leq F(f_{\max})$. Hence if $\lim_{\epsilon \to 0} \epsilon^{-1} \big(F(f_{\max} + \epsilon\psi) - F(f_{\max}) \big)$ exists then it must equal zero, proving (6.8).

Now if $h : [-a, a] \to \mathbb{R}$ is any measurable function such that $\int h \cdot \psi' = 0$ for every function $\psi \in C_c^\infty$ with $\int \psi = 0$, then this precisely means that the distributional derivative of h is orthogonal (with respect to the L^2 pairing) to all except the constant functions. In particular it means that h' is itself a constant function. Applying this principle to $h := \phi' \circ f'_{\max}$, we see that $\phi'(f'_{\max}(x)) = \beta x + C$ for some $\beta, C \in \mathbb{R}$. But ϕ and $f_{\max}$ are even functions, so $\phi' \circ f'_{\max}$ is and odd function, and thus $C = 0$. Now recall that $\phi = \log 2 - \Lambda^*$ where Λ^* is the Legenrde transform of $\Lambda(x) = \log\cosh x$. This implies that $\Lambda' \circ \phi'$ is the negative of the identity function on $[-1, 1]$. In particular $\phi'(f'_{\max}(x)) = \beta x$ implies that $-f'_{\max}(x) = \Lambda'(\beta x)$, which in turn implies that $f_{\max}(x) = -\frac{1}{\beta} \log\cosh(\beta x) + D$ for all x in the support of $f_{\max}$. Here D is some constant of integration. Of course, we know that $f_{\max}$ is convex, which implies $\beta \leq 0$. Thus, by renaming β to be $-\beta$ we have proved the following.

Proposition 6.5.1. *There exists some $\beta_{max} \geq 0$ and some $D_{max} > 0$ such that for every $x \in$*

$(-a_{max}, a_{max})$ *one has that* $f_{\max}(x) = \frac{1}{\beta_{max}} \log\cosh(\beta_{max} x) + D_{max}$.

In the possibility where $\beta_{max} = 0$, the statement of the above Proposition is of course nonsensical, but (because the condition $\int (f_{\max}(x) - |x|)dx = 1$ determines D_{max} uniquely from β_{max}) it is meant to be interpreted in the sense that $f_{\max}(x) = D_{max}$ on its support, meaning that the limit shape would be the triangular function $x \mapsto \max\{1, |x|\}$. We will rule out this possibility shortly.

Our next goal is to find out whether or not $a_{max} < +\infty$, i.e., whether the limit shape is something compact or not. The next theorem tells us that the answer is no.

Theorem 6.5.2. *In the notations of Proposition 6.5.1,* $a_{max} = +\infty$, $\beta_{max} = \frac{\pi}{2\sqrt{3}}$, $D_{max} = \frac{1}{\beta_{max}} \log 2$, *and* $F(f_{\max}) = \pi/\sqrt{3}$. *In particular,* $f_{\max}(x) = \frac{2\sqrt{3}}{\pi} \log\left(2\cosh(\frac{\pi}{2\sqrt{3}} x)\right)$.

Proof. The key will be to use the Hardy-Ramanujan asymptotics together with the identity

$$\int_0^\infty \log(1 + e^{-2x})dx = \int_0^\infty \sum_{n\geq 1} (-1)^{n+1} \frac{e^{-2nx}}{n} dx = \sum_{n \geq 1} \frac{(-1)^{n+1}}{2n^2} = \frac{\pi^2}{24}. \tag{6.9}$$

Here we Taylor expanded the logarithm and then used the identity $\sum_{n\geq 1} n^{-2} = \frac{\pi^2}{6}$ and its corollaries: $\sum_{n\ even} n^{-2} = \frac{\pi^2}{24}$ and $\sum_{n\ odd} n^{-2} = \frac{3\pi^2}{24}$.

We now recall the Hardy-Ramanujan asymptotics [92] for the partition numbers. Specifically, if $p(n)$ denotes the number of partitions of n, then $p(n) = e^{\pi\sqrt{2n/3} - o(\sqrt{n})}$ as $n \to \infty$. Notice that p is an increasing function of n, and every subpartition of a partition of n is a partition of some integer $i \leq n$. Thus the number of subpartitions of any given partition λ of n is upper bounded by $\sum_{i=0}^n p(i) \leq (n+1)p(n) = (n+1)e^{\pi\sqrt{2n/3} - o(\sqrt{n})}$. The prefactor $(n+1)$ may be absorbed into the $o(\sqrt{n})$ term in the exponent, and thus by Proposition 6.4.1 we conclude that $F(f_{\max}) \leq \pi/\sqrt{3}$.

Now, let $f(x) := \alpha^{-1/2} \log(2\cosh(\alpha^{1/2}x))$, where $\alpha := \int_{-\infty}^\infty (\log(2\cosh x) - |x|)dx = 2\int_0^\infty \log(1 + e^{-2x})dx = \frac{\pi^2}{12}$ by (6.9). Note that f is 1-Lipschitz (because it has derivative given by $\tanh(\alpha^{1/2}x)$ which is bounded in absolute value by 1), and also note (by substituting $u = \alpha^{1/2}x$) that $\int_{\mathbb{R}} (f(u) -$

$|u|)dx = 1$ so that $f \in \mathcal{X}$. Now we claim that $F(f) = \pi/\sqrt{3} = 2\alpha^{1/2}$, which would indeed prove that $f = f_{\max}$. To prove this, note that $F(f) = \alpha^{-1/2}\int_{\mathbb{R}} \phi(\tanh u)du$, so that proving that $F(f) = 2\alpha^{1/2}$ now amounts to showing that $\int_{\mathbb{R}} \phi(\tanh u)du = 2\alpha$. In other words, we want to show

$$\int_0^\infty \phi(\tanh u)du = 2\int_0^\infty \log(1+e^{-2u})du. \tag{6.10}$$

One readily checks that $\phi(\tanh u) = \log(e^u + e^{-u}) - u\tanh u$, from which proving (6.10) amounts to checking that $\int_0^\infty \big(\log(e^u+e^{-u}) - 2u + u\tanh u\big)du = 0$. But the integrand here has an explicit antiderivative given by $u\log(e^u+e^{-u}) - u^2$, which is readily seen to evaluate to zero at both $u = 0$ and as $u \to \infty$. This proves (6.10), which finally shows that $f = f_{\max}$. $\square$

A further direction of study is to try to gain more precise asymptotics on the exact number of subpartitions of the maximizing sequence. Specifically we would like to find precise asymptotics on the $o(\sqrt{n})$ term in the optimal growth rate $e^{\pi\sqrt{2n/3} - o(\sqrt{n})}$, and we believe this can be done using more precise large deviations estimates. A similarly difficult "local" asymptotic problem would be to find the rate at which the side lengths go to ∞ (note that Theorem 6.5.2 merely proves that they grow faster than $\sqrt{n}$).

6.6 Extension to k-chains of subpartitions

We now extend the limit shape theorem to the case of partitions which maximize the number of k-chains of subpartitions, which will prove theorem 6.1.2. Since the proof is not significantly more complicated, we briefly indicate the changes which need to be made at each stage of the argument.

First we address the necessary modifications in Section 6.2. In the notation of Corollary 6.2.3, consider k-chains $\gamma_k \leq ... \leq \gamma_2 \leq \gamma_1 \leq f$ of nearest-neighbor bridges which stay below f. Then (by viewing the chain as just a k-tuple of paths and disregarding the ordering) the same corollary says that the number of these k-chains is bounded above by

$$\left(2^n e^{-\sum_1^n \Lambda^*(g(i)-g(i-1))}\right)^k.$$

Then, in equation (6.3) at the beginning of Section 6.3, this bound will tell us that for a given partition λ of n, the number of k-chains of subpartitions of λ (i.e., k-chains of random walk bridges of grid size $n^{-1/2}$ which are nestled in between the graphs of $f_\lambda(x)$ and $|x|$) is upper-bounded by

$$e^{k\sqrt{2n}F(h'_\lambda)}, \tag{6.11}$$

where as usual h_λ is the lower convex envelope of f_λ, and F is the functional of Definition 6.3.1.

Hence, all that is left to do is to show that the upper bound (6.11) is actually sharp up to the exponential scale (after replacing $F(h'_\lambda)$ with $M(F) = \pi/\sqrt{3}$ there). The way to do this is by modifying the proof of Proposition 6.4.1 to lower bound the number of ensembles of k distinct paths staying below the graph of $f_{\max}$. In the notation of that proof, we consider ensembles (implicitly depending on n) of nearest neighbor bridges $(\gamma^i)_{i=1}^k$ from $n^{-1/2}\mathbb{Z} \to n^{-1/2}\mathbb{Z}$, with the property that $f_{i\delta} \leq \gamma_i \leq f_{(i+1)\delta}$ for each $1 \leq i \leq k$. Clearly each such ensemble defines a k-chain of subpartitions of μ_n. Moreover the number of such k-chains is merely the product over $i \in \{1, ..., k\}$, of the individual number of paths lying between the graphs of $f_{i\delta}$ and $f_{(i+1)\delta}$, and we already know a good individual lower bound from the proof of Proposition 6.4.1. Specifically, we can lower bound this number of k-chains by

$$\prod_{i=1}^k \left(e^{\sqrt{2n}F(f_{(i+\frac{1}{2})\delta})-o(\sqrt{n})}\right) = e^{\sqrt{2n}\sum_{i=1}^k F(f_{(i+\frac{1}{2})\delta})-o(\sqrt{n})}.$$

As we already showed in the proof of Proposition 6.4.1, $F(f_\eta) \to 0$ as $\eta \to 0$, which means (by making δ close to 0) that we can actually lower bound the maximal number of k-chains of subpartitions by $e^{k\sqrt{2n}M(F)-o(\sqrt{n})}$, as $n \to \infty$. This already proves Theorem 6.1.2. We remark here that the proof does *not* rely on whether or not the k-chains are strictly ordered or not, so the statement of Theorem 6.1.2 does not depend on this interpretation.

Unfortunately our proof makes it clear that we cannot easily generalize to the case of $k(n)$-chains, i.e., where k grows to $+\infty$ with n. As stated in the introduction, we actually expect that if $k(n)$ grows slowly enough (at a rate of $o(n^{1/2})$) then one has the same limit shape. One the other hand if $n^{1/2} = o(k(n))$, then we expect the limit to be the LSVK curve [115, 148]. We expect a nontrivial crossover when $k(n) \sim \alpha n^{1/2}$ (with limit depending on α), because this is precisely the minimal growth rate at which the typical ensemble of sub-paths no longer has a tendency to just concentrate near the boundary of the partition, but actually distributes itself throughout the bulk of the partition according to some density (as can be shown via a random matrix argument, or alternatively using variational principles for domino tilings). This may or may not be pursued in a future work, but we believe that a similar overall approach will work.

References

[1] R. Adamczak, R. Latała, and R. Meller, "Moments of Gaussian chaoses in Banach spaces," *Electron. J. Probab.*, vol. 26, Paper No. 11, 36, 2021.

[2] T. Alberts, K. Khanin, and J. Quastel, "The continuum directed random polymer," *J. Stat. Phys.*, vol. 154, no. 1-2, pp. 305–326, 2014.

[3] ——, "The intermediate disorder regime for directed polymers in dimension $1 + 1$," *Ann. Probab.*, vol. 42, no. 3, pp. 1212–1256, 2014.

[4] K. S. Alexander, "Characterization of the cluster set of the LIL sequence in Banach space," *Ann. Probab.*, vol. 17, no. 2, pp. 737–759, 1989.

[5] G. Amir, I. Corwin, and J. Quastel, "Probability distribution of the free energy of the continuum directed random polymer in $1 + 1$ dimensions," *Comm. Pure Appl. Math.*, vol. 64, no. 4, pp. 466–537, 2011.

[6] M. A. Arcones, "On the law of the iterated logarithm for Gaussian processes," *J. Theoret. Probab.*, vol. 8, no. 4, pp. 877–903, 1995.

[7] M. A. Arcones and E. Giné, "On decoupling, series expansions, and tail behavior of chaos processes," *J. Theoret. Probab.*, vol. 6, no. 1, pp. 101–122, 1993.

[8] L. M. Arutyunyan and I. S. Yaroslavtsev, "Measurable polynomials on infinite-dimensional spaces," *Dokl. Akad. Nauk*, vol. 449, no. 6, pp. 627–631, 2013.

[9] K. Azuma, "Weighted sums of certain dependent random variables," *Tohoku Math. J. (2)*, vol. 19, pp. 357–367, 1967.

[10] J. Baik, G. Barraquand, I. Corwin, and T. Suidan, "Pfaffian Schur processes and last passage percolation in a half-quadrant," *Ann. Probab.*, vol. 46, no. 6, pp. 3015–3089, 2018.

[11] J. Baik and E. M. Rains, "Algebraic aspects of increasing subsequences," *Duke Math. J.*, vol. 109, no. 1, pp. 1–65, 2001.

[12] M. Balázs, J. Quastel, and T. Seppäläinen, "Fluctuation exponent of the KPZ/stochastic Burgers equation," *J. Amer. Math. Soc.*, vol. 24, no. 3, pp. 683–708, 2011.

[13] P. Baldi, G. Ben Arous, and G. Kerkyacharian, "Large deviations and the Strassen theorem in Hölder norm," *Stochastic Process. Appl.*, vol. 42, no. 1, pp. 171–180, 1992.

[14] G. Barraquand, A. Borodin, and I. Corwin, “Half-space Macdonald processes,” *Forum Math. Pi*, vol. 8, e11, 150, 2020.

[15] G. Barraquand, A. Borodin, I. Corwin, and M. Wheeler, “Stochastic six-vertex model in a half-quadrant and half-line open asymmetric simple exclusion process,” *Duke Math. J.*, vol. 167, no. 13, pp. 2457–2529, 2018.

[16] L. Bertini and G. Giacomin, “Stochastic Burgers and KPZ equations from particle systems,” *Comm. Math. Phys.*, vol. 183, no. 3, pp. 571–607, 1997.

[17] P. Billingsley, *Convergence of probability measures*, Second, ser. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, 1999, pp. x+277, A Wiley-Interscience Publication, ISBN: 0-471-19745-9.

[18] N. H. Bingham, “Variants on the law of the iterated logarithm,” *Bull. London Math. Soc.*, vol. 18, no. 5, pp. 433–467, 1986.

[19] E. Bolthausen, “On a functional central limit theorem for random walks conditioned to stay positive,” *Ann. Probability*, vol. 4, no. 3, pp. 480–485, 1976.

[20] C. Borell, “On the Taylor series of a Wiener polynomial,” *Seminar Notes on multiple stochastic integration, polynomial chaos and their integration*, 1984.

[21] C. Borell, “On polynomial chaos and integrability,” *Probab. Math. Statist.*, vol. 3, no. 2, pp. 191–203, 1984.

[22] A. Borodin, “Stochastic higher spin six vertex model and Macdonald measures,” *J. Math. Phys.*, vol. 59, no. 2, pp. 023301, 17, 2018.

[23] A. Borodin, A. Bufetov, and I. Corwin, “Directed random polymers via nested contour integrals,” *Ann. Physics*, vol. 368, pp. 191–247, 2016.

[24] A. Borodin and I. Corwin, “Macdonald processes,” *Probab. Theory Related Fields*, vol. 158, no. 1-2, pp. 225–400, 2014.

[25] A. Borodin and V. Gorin, *Lectures on integrable probability*, 2015. arXiv: `1212.3351 [math.PR]`.

[26] ——, “Moments match between the KPZ equation and the Airy point process,” *SIGMA Symmetry Integrability Geom. Methods Appl.*, vol. 12, Paper No. 102, 7, 2016.

[27] A. Bryn-Jones and R. A. Doney, “A functional limit theorem for random walk conditioned to stay non-negative,” *J. London Math. Soc. (2)*, vol. 74, no. 1, pp. 244–258, 2006.

[28] K. Burdzy, "Some path properties of iterated Brownian motion," in *Seminar on Stochastic Processes, 1992 (Seattle, WA, 1992)*, ser. Progr. Probab. Vol. 33, Birkhäuser Boston, Boston, MA, 1993, pp. 67–87.

[29] F. Caravenna, "A local limit theorem for random walks conditioned to stay positive," *Probab. Theory Related Fields*, vol. 133, no. 4, pp. 508–530, 2005.

[30] F. Caravenna and L. Chaumont, "Invariance principles for random walks conditioned to stay positive," *Ann. Inst. Henri Poincaré Probab. Stat.*, vol. 44, no. 1, pp. 170–190, 2008.

[31] F. Caravenna, R. Sun, and N. Zygouras, "Polynomial chaos and scaling limits of disordered systems," *J. Eur. Math. Soc. (JEMS)*, vol. 19, no. 1, pp. 1–65, 2017.

[32] ——, "The two-dimensional KPZ equation in the entire subcritical regime," *Ann. Probab.*, vol. 48, no. 3, pp. 1086–1127, 2020.

[33] A. Chandra and H. Weber, "Stochastic PDEs, Regularity Structures, and Interacting Particle Systems," *arXiv e-prints*, arXiv:1508.03616, arXiv:1508.03616, Aug. 2015. arXiv: `1508.03616 [math.AP]`.

[34] S. Chatterjee and A. Dunlap, "Constructing a solution of the $(2+1)$-dimensional KPZ equation," *Ann. Probab.*, vol. 48, no. 2, pp. 1014–1055, 2020.

[35] S. D. Chatterji, "Martingale convergence and the Radon-Nikodym theorem in Banach spaces," *Math. Scand.*, vol. 22, pp. 21–41, 1968.

[36] H. Cohn, R. Kenyon, and J. Propp, "A variational principle for domino tilings," *J. Amer. Math. Soc.*, vol. 14, no. 2, pp. 297–346, 2001.

[37] F. Comets, *Directed polymers in random environments*, ser. Lecture Notes in Mathematics. Springer, Cham, 2017, vol. 2175, pp. xv+199, Lecture notes from the 46th Probability Summer School held in Saint-Flour, 2016, ISBN: 978-3-319-50486-5; 978-3-319-50487-2.

[38] F. Comets, T. Shiga, and N. Yoshida, "Directed polymers in a random environment: Path localization and strong disorder," *Bernoulli*, vol. 9, no. 4, pp. 705–723, 2003.

[39] F. Comets and N. Yoshida, "Directed Polymers in Random Environment are Diffusive at Weak Disorder," *arXiv Mathematics e-prints*, math/0411223, math/0411223, Nov. 2004. arXiv: `math/0411223 [math.PR]`.

[40] I. Corwin, "The Kardar-Parisi-Zhang equation and universality class," *Random Matrices Theory Appl.*, vol. 1, no. 1, pp. 1130001, 76, 2012.

[41] I. Corwin, P. Ghosal, H. Shen, and L.-C. Tsai, “Stochastic PDE limit of the six vertex model,” *Comm. Math. Phys.*, vol. 375, no. 3, pp. 1945–2038, 2020.

[42] I. Corwin and A. Knizel, “Stationary measure for the open KPZ equation,” *arXiv e-prints*, arXiv:2103.12253, arXiv:2103.12253, Mar. 2021. arXiv: `2103.12253 [math.PR]`.

[43] I. Corwin, N. O’Connell, T. Seppäläinen, and N. Zygouras, “Tropical combinatorics and Whittaker functions,” *Duke Math. J.*, vol. 163, no. 3, pp. 513–563, 2014.

[44] I. Corwin and S. Parekh, “Limit shape of subpartition-maximizing partitions,” *Journal of Statistical Physics*, vol. 180, no. 1–6, 597–611, 2020.

[45] I. Corwin and H. Shen, “Open ASEP in the weakly asymmetric regime,” *Comm. Pure Appl. Math.*, vol. 71, no. 10, pp. 2065–2128, 2018.

[46] ——, “Some recent progress in singular stochastic partial differential equations,” *Bull. Amer. Math. Soc. (N.S.)*, vol. 57, no. 3, pp. 409–454, 2020.

[47] I. Corwin, H. Shen, and L.-C. Tsai, “$\mathrm{ASEP}(q, j)$ converges to the KPZ equation,” *Ann. Inst. Henri Poincaré Probab. Stat.*, vol. 54, no. 2, pp. 995–1012, 2018.

[48] I. Corwin and L.-C. Tsai, “KPZ equation limit of higher-spin exclusion processes,” *Ann. Probab.*, vol. 45, no. 3, pp. 1771–1798, 2017.

[49] E. Csáki and S. G. Mohanty, “Excursion and meander in random walk,” *Canad. J. Statist.*, vol. 9, no. 1, pp. 57–70, 1981.

[50] E. Csáki, M. Csörgo, A. Földes, and P. Révész, “Global Strassen-type theorems for iterated Brownian motions,” *Stochastic Process. Appl.*, vol. 59, no. 2, pp. 321–341, 1995.

[51] G. Da Prato and A. Debussche, “Strong solutions to the stochastic quantization equations,” *Ann. Probab.*, vol. 31, no. 4, pp. 1900–1916, 2003.

[52] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Second, ser. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 2014, vol. 152, pp. xviii+493, ISBN: 978-1-107-05584-1.

[53] S. Das and P. Ghosal, “Law of Iterated Logarithms and Fractal Properties of the KPZ Equation,” *arXiv e-prints*, arXiv:2101.00730, arXiv:2101.00730, Jan. 2021. arXiv: `2101.00730 [math.PR]`.

[54] D. Dauvergne, J. Ortmann, and B. Virag, “The directed landscape,” *arXiv e-prints*, arXiv:1812.00309, arXiv:1812.00309, Dec. 2018. arXiv: `1812.00309 [math.PR]`.

[55] L. Decreusefond and A. S. Üstünel, “Stochastic analysis of the fractional Brownian motion,” *Potential Anal.*, vol. 10, no. 2, pp. 177–214, 1999.

[56] A. Dembo, A. Vershik, and O. Zeitouni, “Large deviations for integer partitions,” *Markov Process. Related Fields*, vol. 6, no. 2, pp. 147–179, 2000.

[57] A. Dembo and L.-C. Tsai, “Weakly asymmetric non-simple exclusion process and the Kardar-Parisi-Zhang equation,” *Comm. Math. Phys.*, vol. 341, no. 1, pp. 219–261, 2016.

[58] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, Second, ser. Applications of Mathematics (New York). Springer-Verlag, New York, 1998, vol. 38, pp. xvi+396, ISBN: 0-387-98406-2.

[59] B. Derrida, “Matrix ansatz large deviations of the density in exclusion processes,” in *International Congress of Mathematicians. Vol. III*, Eur. Math. Soc., Zürich, 2006, pp. 367–382.

[60] P. S. Dey and N. Zygouras, “High temperature limits for $(1+1)$-dimensional directed polymer with heavy-tailed disorder,” *Ann. Probab.*, vol. 44, no. 6, pp. 4006–4048, 2016.

[61] J. Diehl, M. Gubinelli, and N. Perkowski, “The Kardar-Parisi-Zhang equation as scaling limit of weakly asymmetric interacting Brownian motions,” *Comm. Math. Phys.*, vol. 354, no. 2, pp. 549–589, 2017.

[62] B. Driver, “Complied Analysis and PDE lecture notes,” *Unpublished*, Chapter 24,

[63] N. Dunford and J. T. Schwartz, *Linear operators. Part I*, ser. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1988, pp. xiv+858, General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication, ISBN: 0-471-60848-3.

[64] A. Dunlap, C. Graham, and L. Ryzhik, “Stationary solutions to the stochastic Burgers equation on the line,” *Comm. Math. Phys.*, vol. 382, no. 2, pp. 875–949, 2021.

[65] R. T. Durrett and D. L. Iglehart, “Functionals of Brownian meander and Brownian excursion,” *Ann. Probability*, vol. 5, no. 1, pp. 130–135, 1977.

[66] R. T. Durrett, D. L. Iglehart, and D. R. Miller, “Weak convergence to Brownian meander and Brownian excursion,” *Ann. Probability*, vol. 5, no. 1, pp. 117–129, 1977.

[67] P. Fatheddin, “The law of the iterated logarithm for a class of SPDEs,” *Stoch. Anal. Appl.*, vol. 39, no. 1, pp. 113–135, 2021.

[68] G. B. Folland, *Real analysis*, Second, ser. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, 1999, pp. xvi+386, Modern techniques and their applications, A Wiley-Interscience Publication, ISBN: 0-471-31716-0.

[69] J. Gärtner, "Convergence towards Burgers' equation and propagation of chaos for weakly asymmetric exclusion processes," *Stochastic Process. Appl.*, vol. 27, no. 2, pp. 233–260, 1988.

[70] M. Gerencsér and M. Hairer, "Singular SPDEs in domains with boundaries," *Probab. Theory Related Fields*, vol. 173, no. 3-4, pp. 697–758, 2019.

[71] P. Ghosal, "Hall-Littlewood-PushTASEP and its KPZ limit," *arXiv e-prints*, arXiv:1701.07308, arXiv:1701.07308, Jan. 2017. arXiv: `1701.07308 [math.PR]`.

[72] T. Gobron and E. Saada, "Couplings, attractiveness and hydrodynamics for conservative particle systems," *Ann. Inst. Henri Poincaré Probab. Stat.*, vol. 46, no. 4, pp. 1132–1177, 2010.

[73] P. Gonçalves and M. Jara, "Nonlinear fluctuations of weakly asymmetric interacting particle systems," *Arch. Ration. Mech. Anal.*, vol. 212, no. 2, pp. 597–644, 2014.

[74] P. Gonçalves, N. Perkowski, and M. Simon, "Derivation of the stochastic Burgers equation with Dirichlet boundary conditions from the WASEP," *Ann. H. Lebesgue*, vol. 3, pp. 87–167, 2020.

[75] V. Goodman and J. Kuelbs, "Gaussian chaos and functional laws of the iterated logarithm for Itô-Wiener integrals," *Ann. Inst. H. Poincaré Probab. Statist.*, vol. 29, no. 4, pp. 485–512, 1993.

[76] V. Goodman, J. Kuelbs, and J. Zinn, "Some results on the LIL in Banach space with applications to weighted empirical processes," *Ann. Probab.*, vol. 9, no. 5, pp. 713–752, 1981.

[77] V. Goodman and J. Kuelbs, "Rates of clustering for some Gaussian self-similar processes," *Probab. Theory Related Fields*, vol. 88, no. 1, pp. 47–75, 1991.

[78] L. Gross, "Abstract Wiener spaces," in *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66)*, vol. II: Contributions to Probability Theory, Part 1, Univ. California Press, Berkeley, Calif., 1967, pp. 31–42.

[79] M. Gubinelli, P. Imkeller, and N. Perkowski, "Paracontrolled distributions and singular PDEs," *Forum Math. Pi*, vol. 3, e6, 75, 2015.

[80] M. Gubinelli and N. Perkowski, "KPZ reloaded," *Comm. Math. Phys.*, vol. 349, no. 1, pp. 165–269, 2017.

[81] ——, "Energy solutions of KPZ are unique," *J. Amer. Math. Soc.*, vol. 31, no. 2, pp. 427–471, 2018.

[82] T. Gueudré and P. Le Doussal, "Directed polymer near a hard wall and KPZ equation in the half-space," *EPL (Europhysics Letters)*, vol. 100, no. 2, p. 26 006, Oct. 2012. arXiv: `1208.5669 [cond-mat.dis-nn]`.

[83] M. Hairer, "A theory of regularity structures," *Invent. Math.*, vol. 198, no. 2, pp. 269–504, 2014.

[84] ——, "Advanced stochastic analysis," *Unpublished lecture notes*, 2016.

[85] M. Hairer, "An introduction to stochastic pdes," *arXiv preprint arXiv:0907.4178*, 2009.

[86] ——, "Solving the KPZ equation," *Ann. of Math. (2)*, vol. 178, no. 2, pp. 559–664, 2013.

[87] ——, "Introduction to regularity structures," *Braz. J. Probab. Stat.*, vol. 29, no. 2, pp. 175–210, 2015.

[88] M. Hairer and C. Labbé, "Multiplicative stochastic heat equations on the whole space," *J. Eur. Math. Soc. (JEMS)*, vol. 20, no. 4, pp. 1005–1054, 2018.

[89] M. Hairer and E. Pardoux, "A Wong-Zakai theorem for stochastic PDEs," *J. Math. Soc. Japan*, vol. 67, no. 4, pp. 1551–1604, 2015.

[90] M. Hairer and J. Quastel, "A class of growth models rescaling to KPZ," *Forum Math. Pi*, vol. 6, e3, 112, 2018.

[91] M. Hairer and H. Weber, "Large deviations for white-noise driven, nonlinear stochastic PDEs in two and three dimensions," *Ann. Fac. Sci. Toulouse Math. (6)*, vol. 24, no. 1, pp. 55–92, 2015.

[92] G. H. Hardy and S. Ramanujan, "Asymptotic Formulaae in Combinatory Analysis," *Proc. London Math. Soc. (2)*, vol. 17, pp. 75–115, 1918.

[93] T. E. Harris, "Nearest-neighbor Markov interaction processes on multidimensional lattices," *Advances in Math.*, vol. 9, pp. 66–89, 1972.

[94] D. A. Huse and C. L. Henley, "Pinning and roughening of domain walls in Ising systems due to random impurities," *Phys. Rev. Lett.*, vol. 54, no. 25, pp. 2708–2711, Jun. 1985.

[95] D. L. Iglehart, "Functional central limit theorems for random walks conditioned to stay positive," *Ann. Probability*, vol. 2, pp. 608–619, 1974.

[96] J. Z. Imbrie and T. Spencer, “Diffusion of directed polymers in a random environment,” *J. Statist. Phys.*, vol. 52, no. 3-4, pp. 609–626, 1988.

[97] K. Itô and H. P. McKean Jr., “Brownian motions on a half line,” *Illinois J. Math.*, vol. 7, pp. 181–231, 1963.

[98] V. Ivanov and G. Olshanski, “Kerov’s central limit theorem for the Plancherel measure on Young diagrams,” in *Symmetric functions 2001: surveys of developments and perspectives*, ser. NATO Sci. Ser. II Math. Phys. Chem. Vol. 74, Kluwer Acad. Publ., Dordrecht, 2002, pp. 93–151.

[99] M. Kardar, G. Parisi, and Y. Zhang, “Dynamic scaling of growing interfaces,” *Phys. Rev. Lett.*, vol. 56, 1986.

[100] N. Konno and T. Shiga, “Stochastic partial differential equations for some measure-valued diffusions,” *Probab. Theory Related Fields*, vol. 79, no. 2, pp. 201–225, 1988.

[101] A. Krajenbrink and P. L. Doussal, “Replica Bethe Ansatz solution to the Kardar-Parisi-Zhang equation on the half-line,” *SciPost Phys.*, vol. 8, p. 35, 3 2020.

[102] J. Kuelbs, “The law of the iterated logarithm for Banach space valued random variables,” in *Probability in Banach spaces (Proc. First Internat. Conf., Oberwolfach, 1975)*, 1976, 131–142. Lecture Notes in Math., Vol. 526.

[103] ——, “Kolmogorov’s law of the iterated logarithm for Banach space valued random variables,” *Illinois J. Math.*, vol. 21, no. 4, pp. 784–800, 1977.

[104] J. Kuelbs and R. Lepage, “The law of the iterated logarithm for Brownian motion in a Banach space,” *Trans. Amer. Math. Soc.*, vol. 185, pp. 253–265, 1973.

[105] C. Labbé, “Weakly asymmetric bridges and the KPZ equation,” *Comm. Math. Phys.*, vol. 353, no. 3, pp. 1261–1298, 2017.

[106] T. L. Lai, “Reproducing kernel Hilbert spaces and the law of the iterated logarithm for Gaussian processes,” *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, vol. 29, pp. 7–19, 1974.

[107] R. Latała, “Estimates of moments and tails of Gaussian chaoses,” *Ann. Probab.*, vol. 34, no. 6, pp. 2315–2331, 2006.

[108] M. Ledoux, “Isoperimetry and gaussian analysis,” in *Lectures on probability theory and statistics*, Springer, 1996, pp. 165–294.

[109] M. Ledoux and M. Talagrand, *Probability in Banach spaces*, ser. Classics in Mathematics. Springer-Verlag, Berlin, 2011, pp. xii+480, Isoperimetry and processes, Reprint of the 1991 edition, ISBN: 978-3-642-20211-7.

[110] T. M. Liggett, "Ergodic theorems for the asymmetric simple exclusion process," *Trans. Amer. Math. Soc.*, vol. 213, pp. 237–261, 1975.

[111] ——, "Coupling the simple exclusion process," *Ann. Probability*, vol. 4, no. 3, pp. 339–356, 1976.

[112] ——, *Interacting particle systems*, ser. Classics in Mathematics. Springer-Verlag, Berlin, 2005, pp. xvi+496, Reprint of the 1985 original, ISBN: 3-540-22617-6.

[113] Y. Lin, "KPZ equation limit of stochastic higher spin six vertex model," *Math. Phys. Anal. Geom.*, vol. 23, no. 1, Paper No. 1, 118, 2020.

[114] Z. Lin, W. Wang, and Y.-K. Choi, "Strassen-type laws of iterated logarithm for a fractional Brownian sheet," *Stochastic Anal. Appl.*, vol. 22, no. 1, pp. 193–210, 2004.

[115] B. F. Logan and L. A. Shepp, "A variational problem for random Young tableaux," *Advances in Math.*, vol. 26, no. 2, pp. 206–222, 1977.

[116] G. Maruyama, "The harmonic analysis of stationary stochastic processes," *Mem. Fac. Sci. Kyūsyū Univ. A*, vol. 4, pp. 45–106, 1949.

[117] K. Matetski, J. Quastel, and D. Remenik, "The KPZ fixed point," *Acta Math.*, vol. 227, no. 1, pp. 115–203, 2021.

[118] Y. Meyer, *Wavelets and operators*, ser. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 1992, vol. 37, pp. xvi+224, Translated from the 1990 French original by D. H. Salinger, ISBN: 0-521-42000-8; 0-521-45869-2.

[119] A. A. Mogulskii, "Large deviations for processes with independent increments," *Ann. Probab.*, vol. 21, no. 1, pp. 202–215, 1993.

[120] J.-C. Mourrat and H. Weber, "Convergence of the two-dimensional dynamic Ising-Kac model to Φ^4_2," *Comm. Pure Appl. Math.*, vol. 70, no. 4, pp. 717–812, 2017.

[121] ——, "Global well-posedness of the dynamic Φ^4 model in the plane," *Ann. Probab.*, vol. 45, no. 4, pp. 2398–2476, 2017.

[122] C. Mueller, "On the support of solutions to the heat equation with noise," *Stochastics Stochastics Rep.*, vol. 37, no. 4, pp. 225–245, 1991.

[123] D. Neuenschwander, "Law of the iterated logarithm for Lévy's area process composed with Brownian motion," *Statist. Probab. Lett.*, vol. 40, no. 4, pp. 371–377, 1998.

[124] M. Nisio, "On the extreme values of Gaussian processes," *Osaka Math. J.*, vol. 4, pp. 313–326, 1967.

[125] D. Nualart, *The Malliavin calculus and related topics*, Second, ser. Probability and its Applications (New York). Springer-Verlag, Berlin, 2006, pp. xiv+382, ISBN: 978-3-540-28328-7; 3-540-28328-5.

[126] N. O'Connell, T. Seppäläinen, and N. Zygouras, "Geometric RSK correspondence, Whittaker functions and symmetrized random polymers," *Invent. Math.*, vol. 197, no. 2, pp. 361–416, 2014.

[127] H. Oodaira, "On Strassen's version of the law of the iterated logarithm for Gaussian processes," *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, vol. 21, pp. 289–299, 1972.

[128] M. A. Ouahra and M. Mellouk, "Strassen's law of the iterated logarithm for stochastic Volterra equations and applications," *Stochastics*, vol. 77, no. 2, pp. 191–203, 2005.

[129] V. G. Papanicolaou, "The probabilistic solution of the third boundary value problem for second order elliptic equations," *Probab. Theory Related Fields*, vol. 87, no. 1, pp. 27–77, 1990.

[130] S. Parekh, "The KPZ limit of ASEP with boundary," *Comm. Math. Phys.*, vol. 365, no. 2, pp. 569–649, 2019.

[131] W. J. Park, "The law of the iterated logarithm for Brownian sheets," *J. Appl. Probability*, vol. 12, no. 4, pp. 840–844, 1975.

[132] N. Perkowski and T. C. Rosati, "The KPZ equation on the real line," *Electron. J. Probab.*, vol. 24, Paper No. 117, 56, 2019.

[133] K. Razi Naqvi, K. J. Mork, and S. Waldenstrøm, "On describing the steady absorption of Brownian particles by a restricted random walk," *Chem. Phys. Lett.*, vol. 92, no. 2, pp. 156–159, 1982.

[134] P. Révész, "A generalization of Strassen's functional law of iterated logarithm," *Z. Wahrsch. Verw. Gebiete*, vol. 50, no. 3, pp. 257–264, 1979.

[135] J. Rosiński and T. Żak, "The equivalence of ergodicity and weak mixing for infinitely divisible processes," *J. Theoret. Probab.*, vol. 10, no. 1, pp. 73–86, 1997.

[136] T. Sasamoto and T. Imamura, "Fluctuations of the one-dimensional polynuclear growth model in half-space," *J. Statist. Phys.*, vol. 115, no. 3-4, pp. 749–803, 2004.

[137] T. Seppäläinen, “Entropy for translation-invariant random-cluster measures,” *Ann. Probab.*, vol. 26, no. 3, pp. 1139–1178, 1998.

[138] M. Sion, “On general minimax theorems,” *Pacific J. Math.*, vol. 8, pp. 171–176, 1958.

[139] D. Slepian, “The one-sided barrier problem for gaussian noise,” *Bell System Technical Journal*, vol. 41, no. 2, pp. 463–501, 1962.

[140] F. Spitzer, “Interaction of Markov processes,” *Advances in Math.*, vol. 5, 246–290 (1970), 1970.

[141] V. Strassen, “An invariance principle for the law of the iterated logarithm,” *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, vol. 3, 211–226 (1964), 1964.

[142] A. S. Üstünel and M. Zakai, “Random rotations of the Wiener path,” *Probab. Theory Related Fields*, vol. 103, no. 3, pp. 409–429, 1995.

[143] ——, “Ergodicité des rotations sur l’espace de Wiener,” *C. R. Acad. Sci. Paris Sér. I Math.*, vol. 330, no. 8, pp. 725–728, 2000.

[144] ——, “Some ergodic theorems for random rotations on wiener space,” *arXiv preprint arXiv 0110162*, 2001.

[145] A. Vershik and Y. Yakubovich, “The limit shape and fluctuations of random partitions of naturals with fixed number of summands,” *Mosc. Math. J.*, vol. 1, no. 3, pp. 457–468, 472, 2001.

[146] A. M. Vershik, “Statistical mechanics of combinatorial partitions, and their limit configurations,” *Funktsional. Anal. i Prilozhen.*, vol. 30, no. 2, pp. 19–39, 96, 1996.

[147] A. M. Vershik, G. A. Freĭman, and Y. V. Yakubovich, “A local limit theorem for random partitions of natural numbers,” *Teor. Veroyatnost. i Primenen.*, vol. 44, no. 3, pp. 506–525, 1999.

[148] A. M. Vershik and S. V. Kerov, “Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young tableaux,” *Dokl. Akad. Nauk SSSR*, vol. 233, no. 6, pp. 1024–1027, 1977.

[149] ——, “Asymptotic behavior of the maximum and generic dimensions of irreducible representations of the symmetric group,” *Funktsional. Anal. i Prilozhen.*, vol. 19, no. 1, pp. 25–36, 96, 1985.

[150] J. B. Walsh, “An introduction to stochastic partial differential equations,” in *École d’été de probabilités de Saint-Flour, XIV—1984*, ser. Lecture Notes in Math. Vol. 1180, Springer, Berlin, 1986, pp. 265–439.

[151] X. Wu, “Intermediate disorder regime for half-space directed polymers,” *J. Stat. Phys.*, vol. 181, no. 6, pp. 2372–2403, 2020.

[152] Y. V. Yakubovich, “The central limit theorem for normalized Young diagrams of partitions into different summands,” *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, vol. 256, no. Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 3, pp. 212–223, 267–268, 1999.

[153] K. Yang, “Kardar-Parisi-Zhang Equation from Long-Range Exclusion Processes,” *arXiv e-prints*, arXiv:2002.05176, arXiv:2002.05176, Feb. 2020. arXiv: `2002.05176 [math.PR]`.

[154] E. V. Yurova, “On continuous restrictions of measurable multilinear mappings,” *Mat. Zametki*, vol. 98, no. 6, pp. 930–936, 2015.