

©Copyright 2013

Sayan Banerjee

On Particle Interaction Models

Sayan Banerjee

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2013

Reading Committee:

Krzysztof Burdzy, Chair

Zhen-Qing Chen

Soumik Pal

Peter Hoff

Program Authorized to Offer Degree:
UW Mathematics

University of Washington

Abstract

On Particle Interaction Models

Sayan Banerjee

Chair of the Supervisory Committee:
Professor Krzysztof Burdzy
UW Mathematics

This dissertation deals with three problems in Stochastic Analysis which broadly involve interactions, either between particles (Chapters 1 and 2), or between particles and the boundary of a C^2 domain (Chapter 3).

- In Chapter 1, we introduce a new model called the **Brownian Conga Line**. It is a random curve evolving in time, generated when a particle performing a two dimensional Gaussian random walk leads a long chain of particles connected to each other by cohesive forces. We approximate the discrete Conga line in some sense by a smooth random curve and subsequently study the properties of this smooth curve.
- In Chapter 2 (joint work with Chris Hoffman), we investigate a **Random Mass Splitting Model** and the closely related random walk in a random environment (RWRE) whose heat kernel at time t turns out to be the mass splitting distribution at t . We prove a quenched invariance principle (QIP) and consequently a quenched central limit theorem for this RWRE using techniques from Rassoul-Agha and Seppäläinen [12] which in turn was based on the work of Kipnis and Varadhan [7] and others.
- In Chapter 3, we deal with a particle performing a Brownian motion inside a bounded C^2 domain with reflection and diffusion at the boundary. We call this model **Brownian Motion with Boundary diffusion** following [1], and study its properties.

TABLE OF CONTENTS

	Page
List of Figures	ii
Chapter 1: The Brownian Conga Line	1
1 Introduction	1
2 The discrete Conga line	6
3 The continuous Conga Line	13
4 The two dimensional Conga line	33
5 Loops and singularities in particle paths	48
6 Freezing in the tail	54
7 Simulations	58
Bibliography	59
Chapter 2: Random Mass Splitting and a quenched invariance principle	60
1 Description of the model	61
2 Quenched invariance principle in general space-time random environments	66
3 Proof of Theorem 9	70
Bibliography	85
Chapter 3: Brownian Motion with Boundary Diffusion	87
1 Stationary Distribution	88
2 A quantity conserved under conformal maps	104
3 A closer look at the trace process	106
Bibliography	111

LIST OF FIGURES

Figure Number	Page
1.1 The Conga line for $n = 1$ (black), $n = 2$ (red), $n = 3$ (green) and $n = 4$ (blue) when $\alpha = 0.5$. (courtesy: Mary Solbrig)	2
1.2 10000 steps (red) and 10000-100 steps (blue) of the discrete two dimensional Conga line. (courtesy: Krzysztof Burdzy)	58
1.3 2000 steps of the discrete one dimensional Conga line with $\alpha = 0.5$. (courtesy: Shirshendu Ganguly)	58
1.4 Near the tip of the Conga line. Showing the first twenty particles. (courtesy: Shirshendu Ganguly)	58

ACKNOWLEDGMENTS

I wish to thank my adviser Krzysztof Burdzy for providing me advice and guidance in both academic and non-academic matters throughout, and Chris Hoffman for collaborating with me and providing immense help in the second project (Chapter 2). I also thank Ma and Baba for standing by my side and believing in me and what I wanted even at times when I did not. I thank my friends for always being there for me and patiently tolerating my anxiety when it often came. It is only when I have been alone that I have realised how important it is to talk to people, especially when they care for you. Thank you Abhirup, Rito, Swarnali, Debashree, Gourab, Pramita, Boga, Yannick, Matt, Chloe, Bharath, Shirshendu, Jair, Tvrtko, Soumik Da, GChat, Dev, Suvankar Da, Arnab for being willing ears. Apologies to people I missed, and I am sure there are lots of them. Thank you Mary for being my refuge and my accomplice in the thick and thin of life.

And last but not the least by any means, I would really like to thank the wonderful coffee shops in Seattle where I did most of my research, sometimes cherished new ideas and far more times fought through the obstacles and failures and loopholes of the mind while gazing at the multitude of coffee lovers who came and talked and celebrated life.

On the financial side, I gratefully acknowledge support from NSF grant number DMS-1206276 and Chris Hoffman's grants from the NSF and the NSA (grants number DMS-1308645 and H98230-13-1-0827).

DEDICATION

to Dadua and Dadabhai

Chapter 1

THE BROWNIAN CONGA LINE

1 Introduction

The Conga Line is a Cuban carnival march that has become popular in many cultures over time. It consists of a queue of people, each one holding onto the person in front of him. The person at the front of the line can move as he will, and the person holding onto him from behind follows him. The third person in the queue follows the second, and so on. Often people keep on joining the line over time by attaching themselves to the last person in the line. As the Conga Line grows in time, it displays interesting motion patterns where the randomness in the motion of the first person propagates down the line, diminishing in influence as it moves further down. In this article, we devise a mathematical formulation of the Conga Line and study its properties.

The formulation is as follows.

Let Z_k , $k \geq 1$, be *i.i.d* standard 2-dimensional normal random variables. Fix some $\alpha \in (0, 1)$. Let $X_1(0) = 0$ and $X_1(n) = \sum_{i=1}^n Z_i$ for $n \geq 1$. This denotes the leading particle, or the *tip* of the Conga line.

Now, we define processes X_k inductively as follows. Suppose that $\{X_k(n), n \geq 0\}$ have already been defined for $1 \leq k \leq j$. Then we let $X_{j+1}(0) = 0$ and

$$X_{j+1}(n) = (1 - \alpha)X_{j+1}(n - 1) + \alpha X_j(n - 1) \quad (1.1)$$

for $n \geq 1$. Here, the process X_k denotes the motion over time of the particle at distance k from the leading particle. The relation (1.1) describes the manner in which a particle X_{j+1} follows the preceding particle X_j . It is easy to check from (1.1) that $X_j(n) = 0$ for all $j > n$. These represent the particles at rest at the origin at time n . Note that the j -th particle X_j joins the Conga Line at time j . See Figure 1.1 for the construction of the Conga line for $n = 1, 2, 3, 4$.

The Conga line at time n is defined as the collection of random variables $\{X_k(n), k \leq n\}$.

One can also think of this model as a discrete version of a long string or molecule whose tip is moving randomly under the effect of an erratic force and the rest of it performs a constrained motion governed by the tip together with the cohesive forces. Burdzy and Pal [8] performed some simulations (see Figure 1.2) which led them to make the following observations:

1. For a fixed large n , the locations of the particles $\{X_k(n), k \geq 1\}$ sufficiently away from the tip look like a ‘smooth’ curve, and the smoothness increases as we move away from the tip.
2. For k significantly larger than 1, there is very little variability in the location of the particles over short periods of time.
3. The small loops in the curve tend to die out over time. Just before death, they look ‘elongated’ and their death site forms a cusp.
4. The particles near the origin seem to *freeze* showing very little movement over time.

All the above observations need precise mathematical formulations. Once the rigorous foundations are established, we can ask the correct questions and try to answer them. This, broadly, is the goal of the article.

We give a brief outline of the content of each section.

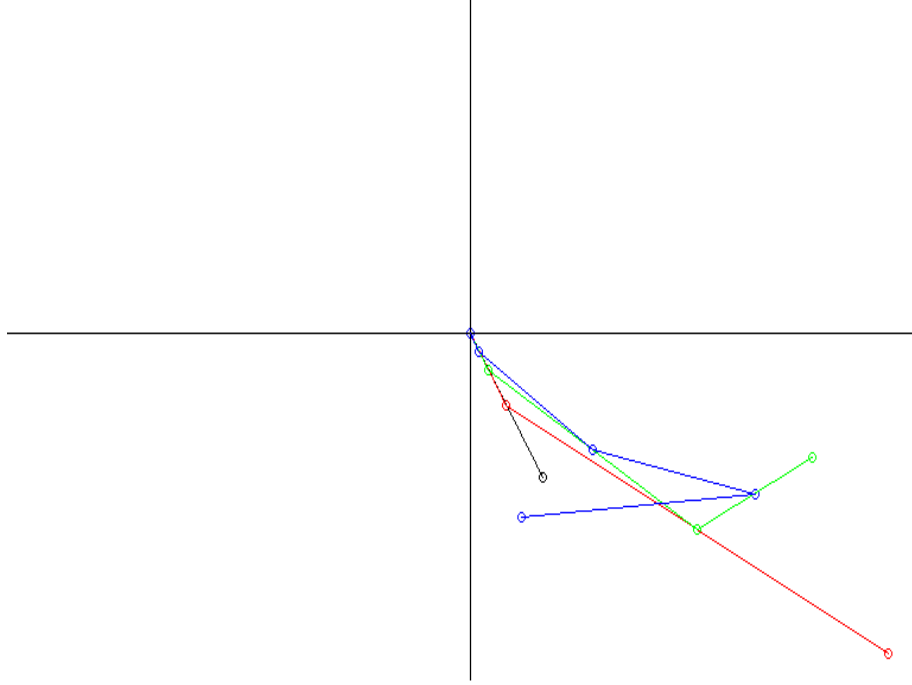


Figure 1.1: The Conga line for $n = 1$ (black), $n = 2$ (red), $n = 3$ (green) and $n = 4$ (blue) when $\alpha = 0.5$. (courtesy: Mary Solbrig)

In Section 2, we try to make mathematical sense of the statement ‘the process looks like a smooth curve’. This is the toughest challenge as the Conga line, unlike most known stochastic processes which can be approximated by continuous models, does not seem to have an interesting scaling limit. This is because if we look at the Conga Line for any fixed n , the distance between the particles decays exponentially as we move away from the tip, i.e., increase k . But it is precisely why this is a novel model, which exhibits *particles moving in different scales ‘in the same picture’*. The particles near the tip are wider spaced and their paths mostly resemble a Gaussian random walk, but those for large k are more closely packed and the Conga Line looks very smooth in this region (see Figure 1.2). To circumvent this problem, we describe a *coupling* between our discrete process $\{X_k(n), k \leq n\}$ and a smooth random process $\{u(x, t) : (x, t) \in \mathbb{R}^{+2}\}$ such that, when observed sufficiently away from the tip, more precisely for $k \geq n^\epsilon$ for any fixed $\epsilon > 0$ and large n , the points $X_{k+1}(n)$ are uniformly close to the points $u(k, n)$. Thus, u serves as a *smooth approximation* to the

discrete process X in a suitable sense. The x variable of u represents distance from the tip and the t variable represents time. Future references to the Conga line refer to this smooth version u . We close the section by presenting another smooth process \bar{u} that also serves as an approximation in the same sense, and is more intuitive when considering the motion of individual particles, i.e. trajectories of the form $\{X_k(n) : n \geq k\}$ for fixed k . It is also used in Section 6 to study the phenomenon of freezing of the Conga line near the origin.

In Section 3, we study the properties of the continuous, one dimensional Conga line u . First, we investigate the phenomenon of the particles at different distances from the tip moving in ‘different scales’ suggested by their different order of variances. The particles near the tip wiggle wildly indicated by their variance being $O(t)$, while those far away from the tip show very little movement, indicated by exponentially decaying variances. Furthermore, there exists a *cutoff* near $x = \alpha t$, where the *variance shows a sharp transition from ‘very large to very small’*. We identify this and study the fine changes in variance around this point.

Next, using the scaling properties of Brownian motion, we show that for fixed t , the Conga line can be scaled so that the space variable x runs in $[0, 1]$. We call this scaled version u_t and study its analytical properties. Upper bounds on the growth rate of the derivatives show that u_t is real analytic. We also make a detailed study of the covariance structure of the derivatives. This turns out to be a major tool in studying the subsequent properties like critical points, length, loops, etc.

With the basic framework of the Conga line established, we set out to investigate its finer properties. We investigate the distribution of critical points of u_t , i.e., points at which the derivative vanishes. The number of critical points in an interval serves as a measure of how wiggly the Conga line looks on that interval. The critical points are distributed as a point process on the real line and we show using an expectation meta-theorem for smooth Gaussian fields (see [1] p. 263) that its first intensity at x (for a large time t) is approximately of the form $\sqrt{t}x^{-1/2}$. This shows that, though the typical number of these points in a given interval is $O(\sqrt{t})$ for large t , the proportion of critical points around x decreases as $x^{-\frac{1}{2}}$ as we go farther away from the tip. We also show subsequently using second moment estimates that the critical points are reasonably *well-spaced* and they do

not tend to crowd around any point. Furthermore, we show that the *first intensity is a good estimate of the point process itself* as for a given interval I sufficiently away from the ends $x = 0$ and $x = 1$, the ratio $\frac{N_t(I)}{\mathbb{E}N_t(I)}$ goes to one in probability as t grows large.

In Section 4 we study properties of the scaled two dimensional Conga line, like length and number of loops. We also investigate a strange phenomenon. Although the mechanism of subsequent particles following the preceding ones and ‘cutting corners’ results in progressively smoothing out the roughness of the Gaussian random walk of the tip, we see that as t increases, the *scaled Conga line looks more and more like Brownian motion* in that the sup-norm distance between them on $[0, 1]$ is roughly of order $t^{-1/4}$ (with a log correction term). This can be explained by the fact that the noticeable smoothing of the paths of the unscaled Conga line takes place in a window of width \sqrt{t} around each point, which translates to a window of width $t^{-1/2}$ as we scale space and time by t . Thus in the scaled version, the *smoothing window becomes smaller with time*, resulting in this phenomenon. Thus, the scaled Conga line u_t for large t serves as a smooth approximation to Brownian motion which smooths out microscopic irregularities but retains its macroscopic characteristics.

In Section 5 we study the evolution of loops in the sequence of paths that the particles at successively larger distances from the tip trace out. We study this evolution under a metric similar to the Skorohod metric. It turns out that with probability one, every singularity, i.e. a point where the speed of the curve becomes zero, in a particle path is a *cuspl singularity* (looks like the graph of $y = x^{2/3}$ in a local co-ordinate frame). Furthermore, there is a *bijection between dying loops and cuspl singularities* in the sense that small loops die (i.e. the end points of the loop merge and loop shrinks to a point) creating cuspl singularities, and conversely, if such a singularity appears in the path of some particle, we can find a loop in the path of the immediately preceding particles, and it dies creating the singularity.

Finally, in Section 6, we investigate the phenomenon of *freezing near the origin*. We work with the smooth approximation \bar{u} , and show that for an appropriate choice of a sequence x_t of distances from the tip such that the particles at these distances remain sufficiently close to the origin, $\bar{u}(x_t, t)$ converges almost surely and in L^2 , and find the limiting function.

Notation: Before we proceed, we clarify the notation that we will be using here:

- (i) If g is a random function and V is a random variable with distribution function F and independent of g , then

$$\mathbb{E}_V g(V) = \int g(v) dF(v)$$

denotes the expectation with respect to V for a fixed realisation of g .

- (ii) Φ denotes the normal distribution function and $\bar{\Phi} = 1 - \Phi$.
- (iii) For any function f of several variables, $\partial_x^k f$ denotes the partial derivative of f with respect to the variable x taken k times.
- (iv) For functions $f, g : [0, \infty \rightarrow \mathbb{R}^+$, $f(t) \sim g(t)$ means that there f and g have the same growth rate in t , i.e., there exists a constant C such that

$$\frac{f(t)}{g(t)} \vee \frac{g(t)}{f(t)} \leq C$$

for all sufficiently large t .

- (v) For a family of real-valued functions $\{f_t : t \in (0, \infty)\}$ defined on a compact set $I \subseteq \mathbb{R}^k$ and a function $a : (0, \infty) \rightarrow [0, \infty)$, we say

$$f_t = O^\infty(a(t)) \text{ on } I$$

if

$$\sup_{t \in (0, \infty)} \frac{\sup_{x \in I} |f_t(x)|}{a(t)} \leq C$$

for some constant $C < \infty$. Sometimes, (by abuse of notation) we will write

$$f_t(x) = O^\infty(a(t)) \text{ for } x \in I$$

to denote the same.

2 The discrete Conga line

We set out by finding a neater expression for $X_k(n)$ in terms of $X_1(n)$.

Let T_1, T_2, \dots be i.i.d $\text{Geom}(\alpha)$ and let

$$\Theta_j = \sum_{i=1}^j T_i.$$

Then $\Theta_j \sim NB(j, 1 - \alpha)$, where $NB(a, b)$ represents the Negative Binomial distribution with parameters a and b . It is easy to see from the recursion relation (1.1) that one can write

$$X_k(n) = \mathbb{E}_{T_1} X_{k-1}(n - 1 - T_1).$$

By induction, we get

$$X_k(n) = \mathbb{E}_S X_1(n - k + 1 - \Theta_{k-1}) = \sum_{m=0}^{n-k+1} \binom{m+k-2}{m} (1-\alpha)^m \alpha^{k-1} X_1(n - k + 1 - m). \quad (1.2)$$

2.1 Approximation by a smooth process

Here we show that for any fixed $\epsilon > 0$, the discrete one dimensional Conga line can be approximated uniformly in k , for $n^\epsilon \leq k \leq n$, for large n , by a smooth process that arises as a smoothing kernel acting on Brownian motion.

Let $B_l \sim \text{Bin}(l, \alpha)$. From (1.2) and the fact that

$$P(\Theta_{k-1} \leq l - k + 1) = P(B_l \geq k - 1),$$

we get

$$\begin{aligned} X_k(n) &= EX_1(n - k + 1 - \Theta_{k-1}) = \sum_{j=0}^{n-k+1} P(\Theta_{k-1} = j) \sum_{l=0}^{n-k+1-j} Z_l \\ &= \sum_{l=0}^{n-k+1} Z_l \sum_{j=0}^{n-k+1-l} P(\Theta_{k-1} = j) = \sum_{l=k-1}^n P(\Theta_{k-1} \leq l - k + 1) Z_{n-l} \\ &= \sum_{l=k-1}^n P(B_l \geq k - 1) Z_{n-l} \end{aligned}$$

The next step is the key to the approximation. We obtain a *coupling* between a Brownian motion and our process X . Let (Ω, \mathcal{F}, P) be a probability space supporting a Brownian motion W . Then

$$X_k(n) = \sum_{l=k-1}^n P(B_l \geq k-1)(W(n-l) - W(n-l-1))$$

gives the desired coupling on this space. Note that we can write

$$X_{k+1}(n) = \int_0^n g(k, z) dW_z^n$$

where

$$g(k, z) = P(B_{[z]} \geq k),$$

$[\cdot]$ being the greatest integer function, and $W_z^t = W(t) - W(t-z)$, $0 \leq z \leq t$, is the time reversed Brownian motion from time t .

Let $\sigma = \sqrt{\alpha(1-\alpha)}$. Consider the “space-time” process

$$\begin{aligned} u(x, t) &= \int_0^t \bar{\Phi} \left(\frac{x - \alpha z}{\sigma \sqrt{z}} \right) dW_z^t \\ &= \int_0^t W(t-z) (\sqrt{2\pi})^{-1} \left(\frac{x + \alpha z}{2\sigma z^{3/2}} \right) \exp \left(-\frac{(x - \alpha z)^2}{2\sigma^2 z} \right) dz \end{aligned}$$

(We obtain the second expression from the first by an application of the *Stochastic Fubini Theorem*, see [7]).

We prove in what follows that for large n , and for $n^\epsilon \leq k \leq n$ for any fixed $\epsilon > 0$, the points $X_{k+1}(n)$ are “uniformly close” to the points $u(k, n)$ (u evaluated at integer points) for the given range of k . Our strategy is to first consider a discretized version of the process $u(x, t)$, given by

$$\hat{u}(k, n) = \sum_{l=0}^n \bar{\Phi} \left(\frac{k - \alpha l}{\sigma \sqrt{l}} \right) (W(n-l) - W(n-l-1))$$

In Lemma 1, we give a bound on the L^2 distance between $X_{k+1}(n)$ and $\hat{u}(k, n)$ for large n when $n^\epsilon \leq k \leq n$. In Lemma 2, a similar bound is achieved for the L^2 distance between

$\hat{u}(k, n)$ and $u(k, n)$. In Theorem 1, we prove using a Borel Cantelli argument that for large n the two processes X and u (evaluated at integer points) come uniformly close on $n^\epsilon \leq k \leq n$.

In the following, C_1, C_2, \dots represent absolute constants, C_ϵ, C'_ϵ denote constants that depend only on ϵ , C_p denotes a constant depending only on p and $D_{\epsilon,p}, D'_{\epsilon,p}$ denote constants depending upon both ϵ and p .

Lemma 1. Fix $\epsilon > 0$. For $n^\epsilon \leq k \leq n$,

$$\sum_{l=0}^n \left[P(B_l \geq k) - \bar{\Phi} \left(\frac{k - \alpha l}{\sigma \sqrt{l}} \right) \right]^2 \leq C_\epsilon \frac{\sqrt{\log k}}{\sqrt{k}} \quad (1.3)$$

where $\sigma = \sqrt{\alpha(1 - \alpha)}$. Consequently,

$$\mathbb{E}(X_{k+1}(n) - \hat{u}(k, n))^2 \leq C_\epsilon \frac{\sqrt{\log k}}{\sqrt{k}}$$

uniformly on $n^\epsilon \leq k \leq n$.

Proof: Choose $C > 0$ such that $\epsilon(C - 1) \geq 2$ and $C \geq \frac{3}{2}$. Take $L_k = [\alpha^{-1} \sqrt{Ck \log k}]$, where $[\cdot]$ represents the greatest integer function. Then, we can write

$$\begin{aligned} \sum_{l=0}^n \left[P(B_l \geq k) - \bar{\Phi} \left(\frac{k - \alpha l}{\sigma \sqrt{l}} \right) \right]^2 &\leq \sum_{l=0}^{[\frac{k}{\alpha}] - L_k} \left[P(B_l \geq k) - \bar{\Phi} \left(\frac{k - \alpha l}{\sigma \sqrt{l}} \right) \right]^2 \\ &\quad + \sum_{l=[\frac{k}{\alpha}] - L_k}^{[\frac{k}{\alpha}] + L_k} \left[P(B_l \geq k) - \bar{\Phi} \left(\frac{k - \alpha l}{\sigma \sqrt{l}} \right) \right]^2 \\ &\quad + \sum_{l=[\frac{k}{\alpha}] + L_k}^n \left[P(B_l \geq k) - \bar{\Phi} \left(\frac{k - \alpha l}{\sigma \sqrt{l}} \right) \right]^2 \\ &= S_1^{(k)} + S_2^{(k)} + S_3^{(k)}. \end{aligned}$$

Here, $S_1^{(k)}$ and $S_3^{(k)}$ correspond to the *tails* of the distribution functions, and we shall show that they are negligible compared to $S_2^{(k)}$. To this end, note that

$$S_1^{(k)} \leq 2k\alpha^{-1} \left(P^2(B_{[\frac{k}{\alpha}] - L_k} \geq k) + \bar{\Phi}^2 \left(\frac{\alpha L_k}{\sigma \sqrt{\frac{k}{\alpha}}} \right) \right).$$

Now, by Bernstein's inequality,

$$P(B_{[\frac{k}{\alpha}] - L_k} \geq k) \leq \exp\left(-\frac{\alpha^2 L_k^2 / 2}{\alpha([\frac{k}{\alpha}] - L_k) + \alpha L_k / 3}\right) \leq \exp\left(-\frac{\alpha^2 L_k^2}{2k}\right).$$

We also have

$$\bar{\Phi}\left(\frac{\alpha L_k}{\sigma \sqrt{\frac{k}{\alpha}}}\right) \leq \frac{\sigma \sqrt{k}}{\alpha^{3/2} L_k} \exp\left(-\frac{\alpha^3 L_k^2}{2\sigma^2 k}\right).$$

Therefore, for large k ,

$$S_1^{(k)} \leq \frac{4k}{\alpha} \exp\left(-\frac{\alpha^2 L_k^2}{k}\right) \leq \frac{C_1}{k^{C-1}} \leq \frac{C_1}{\sqrt{k}}.$$

Similarly, for $S_3^{(k)}$, we get

$$\begin{aligned} S_3^{(k)} &\leq 2(n - [k\alpha^{-1}]) \left(P^2(B_{[\frac{k}{\alpha}] + L_k} < k) + \Phi^2\left(\frac{-\alpha L_k}{\sigma \sqrt{\frac{2k}{\alpha}}}\right) \right) \\ &\leq 4(n - [k\alpha^{-1}]) \exp\left(-\frac{\alpha^2 L_k^2}{2k}\right) \leq C_3(n - [k\alpha^{-1}]) k^{-C/2} \\ &\leq \frac{C_3(n - [k\alpha^{-1}])}{\sqrt{k} n^{\epsilon(C-1)/2}} \leq \frac{C_3}{\sqrt{k}}. \end{aligned}$$

Now, for $S_2^{(k)}$, we use the *Berry Esseen Theorem* (see [3]).

$$S_2^{(k)} \leq \sum_{l=[\frac{k}{\alpha}] - L_k}^{[\frac{k}{\alpha}] + L_k} \frac{1}{l} \leq C_4 \frac{\sqrt{C \log k}}{\sqrt{k}}.$$

The above yield the lemma. □

Lemma 2. $\mathbb{E}(\hat{u}(k, n) - u(k, n))^2 \leq C_\epsilon \frac{\sqrt{\log k}}{\sqrt{k}}$ uniformly on $n^\epsilon \leq k \leq n$.

Proof: Write $\hat{u}(k, n) = \int_0^n \hat{f}(k, z) dW_z^n$ and $u(k, n) = \int_0^n f(k, z) dW_z^n$ where

$$\hat{f}(k, z) = \sum_{j=0}^n \bar{\Phi}\left(\frac{k - \alpha j}{\sigma \sqrt{j}}\right) \mathbb{I}(j \leq z < j + 1) \text{ and } f(k, z) = \bar{\Phi}\left(\frac{k - \alpha z}{\sigma \sqrt{z}}\right).$$

Then, we can decompose $\mathbb{E}(\hat{u}(k, n) - u(k, n))^2$ as in the proof of Lemma 1 as follows:

$$\begin{aligned}
\mathbb{E}(\hat{u}(k, n) - u(k, n))^2 &= \int_0^n \left[\sum_{j=0}^n \left(\bar{\Phi} \left(\frac{k - \alpha z}{\sigma \sqrt{z}} \right) - \bar{\Phi} \left(\frac{k - \alpha j}{\sigma \sqrt{j}} \right) \right) \mathbb{I}(j \leq z < j + 1) \right]^2 dz \\
&\leq \int_0^{\lfloor \frac{k}{\alpha} \rfloor - L_k} \left[\sum_{j=0}^n \left(\bar{\Phi} \left(\frac{k - \alpha z}{\sigma \sqrt{z}} \right) - \bar{\Phi} \left(\frac{k - \alpha j}{\sigma \sqrt{j}} \right) \right) \mathbb{I}(j \leq z < j + 1) \right]^2 dz \\
&\quad + \int_{\lfloor \frac{k}{\alpha} \rfloor - L_k}^{\lfloor \frac{k}{\alpha} \rfloor + L_k} \left[\sum_{j=0}^n \left(\bar{\Phi} \left(\frac{k - \alpha z}{\sigma \sqrt{z}} \right) - \bar{\Phi} \left(\frac{k - \alpha j}{\sigma \sqrt{j}} \right) \right) \mathbb{I}(j \leq z < j + 1) \right]^2 dz \\
&\quad + \int_{\lfloor \frac{k}{\alpha} \rfloor + L_k}^n \left[\sum_{j=0}^n \left(\bar{\Phi} \left(\frac{k - \alpha z}{\sigma \sqrt{z}} \right) - \bar{\Phi} \left(\frac{k - \alpha j}{\sigma \sqrt{j}} \right) \right) \mathbb{I}(j \leq z < j + 1) \right]^2 dz \\
&= I_1^{(k)} + I_2^{(k)} + I_3^{(k)}.
\end{aligned}$$

Now,

$$\frac{\partial f}{\partial z}(k, z) = (\sqrt{2\pi})^{-1} \left(\frac{k + \alpha z}{2\sigma z^{3/2}} \right) \exp \left(-\frac{(k - \alpha z)^2}{2\sigma^2 z} \right). \quad (1.4)$$

We can follow the same argument as in the proof of Lemma 1 and verify that $I_1^{(k)}$ and $I_3^{(k)}$ are bounded above by $C_5(\sqrt{k})^{-1}$. To handle the second term, note that by (1.4), we have

$$\left| \frac{\partial f}{\partial z}(k, z) \right| \leq C_6(\sqrt{k})^{-1}.$$

on $\lfloor \frac{k}{\alpha} \rfloor - L_k \leq z \leq \lfloor \frac{k}{\alpha} \rfloor + L_k$. So,

$$I_2^{(k)} \leq \int_{\lfloor \frac{k}{\alpha} \rfloor - L_k}^{\lfloor \frac{k}{\alpha} \rfloor + L_k} \left(C_6(\sqrt{k})^{-1} \right)^2 dz = 2C_6 \frac{\sqrt{\log k}}{\sqrt{k}}.$$

This proves the lemma. \square

So, by the preceding lemmas, we have proved that

$$\mathbb{E}(X_{k+1}(n) - u(k, n))^2 \leq C'_\epsilon \frac{\sqrt{\log k}}{\sqrt{k}}$$

uniformly on $n^\epsilon \leq k \leq n$.

Now, $X_{k+1}(n) - u(k, n) = \int_0^n (g(k, z) - f(k, z)) dW_z^n$. Now, by the fact that this is a centred Gaussian random variable,

$$\begin{aligned} \mathbb{E}|X_{k+1}(n) - u(k, n)|^{2p} &\leq C_p \left[\int_0^n (g(k, z) - f(k, z))^2 dz \right]^p \\ &\leq C_p C_\epsilon'^p \left(\frac{\log k}{k} \right)^{p/2}. \end{aligned} \quad (1.5)$$

We use this to obtain the following theorem.

Theorem 1. *For any $\delta > 0$, $\epsilon > 0$ and $\eta > 0$, define the following events:*

$$A_{k,n} = \left\{ |X_{k+1}(n) - u(k, n)| > \delta k^{-(\frac{1}{4}-\eta)} \right\}$$

and

$$B_n^\epsilon = \bigcup_{k=n^\epsilon}^n A_{k,n}$$

Then $P(\limsup_n B_n) = 0$.

Proof: By Chebychev type Inequality (for $2p$ -th moment) and (1.5), we get for any $p \geq 1$,

$$\begin{aligned} P(A_{k,n}) &\leq (\delta k^{-(\frac{1}{4}-\eta)})^{-2p} \mathbb{E}(X_{k+1}(n) - u(k, n))^{2p} \\ &\leq \frac{C_p C_\epsilon'^p}{\delta^{2p}} k^{p/2-2\eta p} \left(\frac{\log k}{k} \right)^{p/2} \leq D_{\epsilon,p} k^{-(2\eta p-1)}. \end{aligned}$$

Hence,

$$P(B_n^\epsilon) \leq D_{\epsilon,p} \sum_{k=n^\epsilon}^n k^{-(2\eta p-1)} \leq \frac{D'_{\epsilon,p}}{2\eta p-2} n^{-\epsilon(2\eta p-2)}.$$

Now, choose p large enough such that $P(B_n^\epsilon) \leq Cn^{-2}$ for some constant C . The result now follows by the Borel Cantelli lemma. \square

Note: The above theorem suggests that the distance between the points $X_k(n)$ and $u(k, n)$ decreases as we increase k . This is exactly what is suggested by .

2.2 A related process

Here we give another smooth approximation \bar{u} to X given by

$$\bar{u}(x, t) = E_Z W \left(t - \frac{x}{\alpha} - Z_{\rho^2 x} \right), \quad (1.6)$$

that proves to be more useful and intuitive while investigating the paths of individual particles and study the phenomenon of freezing near the origin. Note that \bar{u} has the following properties:

- Increasing x (i.e. going away from the tip) results in the paths $\bar{u}(x, \cdot)$ being progressively smoother, indicated by the increasing variance of the Z variable.
- For fixed x , varying t represents the path of the particle at distance x from the tip. As successive particles ‘cut corners’, the paths become smoother.

Theorem 2. *The result of Theorem 1 holds with u replaced by \bar{u} .*

Proof: Let $\rho = \sigma/\alpha^{3/2}$. Consider another continuous process u^* given by

$$u^*(x, t) = \int_0^t \bar{\Phi} \left(\frac{\frac{x}{\alpha} - s}{\rho\sqrt{x}} \right) dW_s^t.$$

Rewrite u^* as follows:

$$\begin{aligned} u^*(x, t) &= \int_0^t \int_{-\infty}^s \frac{1}{\sqrt{2\pi x \rho}} \phi \left(\frac{\frac{x}{\alpha} - w}{\rho\sqrt{x}} \right) dw dW_s^t \\ &= \int_{-\infty}^t \int_{\max\{w, 0\}}^t dW_s^t \frac{1}{\sqrt{2\pi x \rho}} \phi \left(\frac{\frac{x}{\alpha} - w}{\rho\sqrt{x}} \right) dw \\ &= \int_0^t \frac{1}{\sqrt{2\pi x \rho}} \phi \left(\frac{\frac{x}{\alpha} - w}{\rho\sqrt{x}} \right) W(t-w) dw + W(t) \bar{\Phi} \left(\frac{\sqrt{x}}{\alpha \rho} \right) \\ &= \int_{-\infty}^t \frac{1}{\sqrt{2\pi x \rho}} \phi \left(\frac{\frac{x}{\alpha} - w}{\rho\sqrt{x}} \right) W(t-w) dw \\ &\quad - \int_{-\infty}^0 \frac{1}{\sqrt{2\pi x \rho}} \phi \left(\frac{\frac{x}{\alpha} - w}{\rho\sqrt{x}} \right) W(t-w) dw + W(t) \bar{\Phi} \left(\frac{\sqrt{x}}{\alpha \rho} \right) \\ &= \bar{u}(x, t) + e_1(x, t) + e_2(x, t). \end{aligned}$$

Clearly,

$$\mathbb{E}(e_2^2(k, n)) \leq \frac{\alpha^2 \rho^2 n}{k} \exp \left\{ -\frac{k}{\alpha^2 \rho^2} \right\}.$$

So, on $n^\epsilon \leq k \leq n$,

$$\mathbb{E}(e_2^2(k, n)) \leq \alpha^2 \rho^2 n \exp \left\{ -\frac{n^\epsilon}{\alpha^2 \rho^2} \right\}.$$

Furthermore, it is easy to verify that

$$\mathbb{E}(e_1^2(x, t)) = t \bar{\Phi}^2 \left(\frac{\sqrt{x}}{\alpha \rho} \right) + \int_0^\infty \bar{\Phi}^2 \left(\frac{s + \frac{x}{\alpha}}{\rho \sqrt{x}} \right) ds.$$

So, on $n^\epsilon \leq k \leq n$,

$$\begin{aligned} \mathbb{E}(e_1^2(k, n)) &\leq n \bar{\Phi}^2 \left(\frac{n^{\epsilon/2}}{\rho \alpha} \right) + \int_0^\infty \bar{\Phi}^2 \left(\frac{s}{\rho \sqrt{n}} + \frac{n^{\epsilon/2}}{\rho \alpha} \right) ds \\ &\leq \frac{\alpha^2 \rho^2}{n^\epsilon} \exp \left\{ -\frac{n^\epsilon}{\alpha^2 \rho^2} \right\} + \rho \sqrt{n} \int_{\frac{n^{\epsilon/2}}{\rho \alpha}}^\infty \frac{1}{y^2} e^{-y^2} dy \\ &\leq \frac{\alpha^2 \rho^2}{n^\epsilon} \exp \left\{ -\frac{n^\epsilon}{\alpha^2 \rho^2} \right\} + \frac{\rho \sqrt{n}}{2 \left(\frac{n^{\epsilon/2}}{\rho \alpha} \right)^3} \exp \left\{ -\frac{n^\epsilon}{\rho^2 \alpha^2} \right\}. \end{aligned}$$

Now, by calculations similar to those in the proof of Lemma 2,

$$\begin{aligned} \mathbb{E}(u(k, n) - u^*(k, n))^2 &= \int_0^n \left[\bar{\Phi} \left(\frac{k - \alpha s}{\sigma \sqrt{s}} \right) - \bar{\Phi} \left(\frac{k - \alpha s}{\alpha \rho \sqrt{k}} \right) \right]^2 ds \\ &\leq C \frac{(\log k)^{\frac{5}{2}}}{\sqrt{k}}. \end{aligned}$$

Now, the same proof as Theorem 1 yields the result. \square

3 The continuous Conga Line

Here we investigate properties of the continuous one dimensional Conga line u obtained in the previous section as an approximation to the discrete Conga line X .

3.1 Particles moving in different scales

It is not hard to observe by estimating

$$\text{Var}(u(x, t)) = \int_0^t \bar{\Phi}^2 \left(\frac{x - \alpha y}{\sigma \sqrt{y}} \right) dy$$

that particles at distances ct from the leading particle have variance $O(t)$ if $c < \alpha$ and $o(1)$ (in fact, the variance decays exponentially with t) if $c > \alpha$. Also in a window of width $c\sqrt{t}$ about αt , the variance is $O(\sqrt{t})$. In particular, this indicates that there is a small window somewhere between 0 and αt where the variance changes from being 'very small to very large', i.e., there is a *cut-off* below which the variance goes to zero with t , and above which the variance grows to infinity.

Theorem 3.

(i) For $\lambda > 0$ and $1/2 \leq \beta < 1$, $\text{Var}(u(\alpha t - \lambda t^\beta, t)) \sim t^\beta$.

(ii) $\text{Var}(u(\alpha t + \sigma\sqrt{\lambda t \log t}, t)) \sim \frac{t^{1/2-\lambda}}{(\log t)^{3/2}}$.

Proof: (i) Take any $c > \lambda/\alpha$. Then, decomposing the variance,

$$\begin{aligned} \text{Var}(u(\alpha t - \lambda t^\beta, t)) &= \int_0^t \bar{\Phi}^2\left(\frac{\alpha t - \lambda t^\beta - \alpha y}{\sigma\sqrt{y}}\right) dy = \int_0^{t-ct^\beta} \bar{\Phi}^2\left(\frac{\alpha t - \lambda t^\beta - \alpha y}{\sigma\sqrt{y}}\right) dy \\ &\quad + \int_{t-ct^\beta}^t \bar{\Phi}^2\left(\frac{\alpha t - \lambda t^\beta - \alpha y}{\sigma\sqrt{y}}\right) dy. \end{aligned}$$

The first integral satisfies

$$\begin{aligned} \int_0^{t-ct^\beta} \bar{\Phi}^2\left(\frac{\alpha t - \lambda t^\beta - \alpha y}{\sigma\sqrt{y}}\right) dy &\leq Ct^{1/2} \int_0^{t-ct^\beta} \bar{\Phi}^2\left(\frac{\alpha t - \lambda t^\beta - \alpha y}{\sigma\sqrt{y}}\right) \left(\frac{\alpha t - \lambda t^\beta + \alpha y}{2\sigma y\sqrt{y}}\right) dy \\ &\leq Ct^{1/2} \int_{\frac{\alpha c - \lambda}{\sigma} t^{\beta-\frac{1}{2}}}^\infty \bar{\Phi}^2(z) dz \\ &\leq Ct^{1/2} \exp\left(-\frac{(\alpha c - \lambda)^2 t^{2\beta-1}}{\sigma^2}\right), \end{aligned}$$

where the second step above follows from a change of variables.

It is easy to check that

$$\int_{t-ct^\beta}^t \bar{\Phi}^2\left(\frac{\alpha t - \lambda t^\beta - \alpha y}{\sigma\sqrt{y}}\right) dy \sim t^\beta$$

proving part (i).

(ii) We decompose the variance as

$$\begin{aligned} \text{Var}(u(\alpha t + \sigma\sqrt{\lambda t \log t}, t)) &= \int_0^{t/2} \bar{\Phi}^2\left(\frac{\alpha t + \sigma\sqrt{\lambda t \log t} - \alpha y}{\sigma\sqrt{y}}\right) dy \\ &\quad + \int_{t/2}^t \bar{\Phi}^2\left(\frac{\alpha t + \sigma\sqrt{\lambda t \log t} - \alpha y}{\sigma\sqrt{y}}\right) dy. \end{aligned}$$

The first integral decays like e^{-Ct} for some constant C . For the second integral, we make a change of variables similar to (i) and standard estimates for the normal c.d.f. to get

$$\begin{aligned} \int_{t/2}^t \bar{\Phi}^2\left(\frac{\alpha t + \sigma\sqrt{\lambda t \log t} - \alpha y}{\sigma\sqrt{y}}\right) dy &\sim t^{1/2} \int_{\sqrt{\lambda \log t}}^{\infty} \bar{\Phi}^2(z) dz \sim t^{1/2} \int_{\sqrt{\lambda \log t}}^{\infty} z^{-2} \exp(-z^2) dz \\ &\sim \frac{t^{1/2}}{(\log t)^{3/2}} \int_{\sqrt{\lambda \log t}}^{\infty} 2z \exp(-z^2) dz \sim \frac{t^{1/2-\lambda}}{(\log t)^{3/2}}, \end{aligned}$$

proving (ii). □

Part (ii) of the above theorem has the following interesting consequence, demonstrating a *cut-off phenomenon* for the variance of $u(x, t)$ in the vicinity of $x = \alpha t + \sigma\sqrt{\frac{1}{2}t \log t}$.

Corollary 3.1. *As $t \rightarrow \infty$*

(i) $\text{Var}(u(\alpha t + \sigma\sqrt{\lambda t \log t}, t)) \rightarrow 0$ if $\lambda \geq 1/2$.

(ii) $\text{Var}(u(\alpha t + \sigma\sqrt{\lambda t \log t}, t)) \rightarrow \infty$ if $\lambda < 1/2$.

(iii) For $0 < \delta < \infty$, $\text{Var}(u(\alpha t + \sigma\sqrt{t((1/2) \log t - (3/2) \log \log t - \log \delta)}, t)) \sim \delta$.

So, the variance exhibits a sharp jump around $\alpha t + \sigma\sqrt{t((1/2) \log t - (3/2) \log \log t)}$.

The proof follows easily from part (ii) of Theorem 3.

3.2 Analyticity of the scaled Conga Line

For a fixed time t , the Conga line satisfies the following equality in distribution:

$$\left\{ \int_0^t \bar{\Phi}\left(\frac{tx - \alpha s}{\sigma\sqrt{s}}\right) dW_s^t : 0 \leq x \leq 1 \right\} \stackrel{d}{=} \left\{ \sqrt{t} \int_0^1 \bar{\Phi}\left(\frac{x - \alpha s}{\sigma_t \sqrt{s}}\right) dW_s^1 : 0 \leq x \leq 1 \right\}, \quad (1.7)$$

where $\sigma_t = \frac{\sigma}{\sqrt{t}}$. This can be checked from the fact that both are Gaussian processes with covariance function K where

$$K(x_1, x_2) = \int_0^t \bar{\Phi}\left(\frac{tx_1 - \alpha s}{\sigma\sqrt{s}}\right) \bar{\Phi}\left(\frac{tx_2 - \alpha s}{\sigma\sqrt{s}}\right) ds = t \int_0^1 \bar{\Phi}\left(\frac{x_1 - \alpha s}{\sigma_t\sqrt{s}}\right) \bar{\Phi}\left(\frac{x_2 - \alpha s}{\sigma_t\sqrt{s}}\right) ds.$$

So, to study the Conga line for fixed t , we study the scaled process

$$\left\{ u_t(x) = \int_0^1 \bar{\Phi}\left(\frac{x - \alpha s}{\sigma_t\sqrt{s}}\right) dW_s^1 : 0 \leq x \leq 1 \right\}.$$

Now, we take a look at the derivatives of this process. It is easy to check that we can differentiate under the integral. Thus,

$$\begin{aligned} \partial_x u_t(x) &= - \int_0^1 \frac{1}{\sigma_t\sqrt{s}} \phi\left(\frac{x - \alpha s}{\sigma_t\sqrt{s}}\right) dW_s^1, \\ \partial_x^2 u_t(x) &= - \int_0^1 \frac{1}{\sigma_t^2 s} \phi'\left(\frac{x - \alpha s}{\sigma_t\sqrt{s}}\right) dW_s^1 \\ &= \int_0^1 \frac{1}{\sigma_t^2 s} \left(\frac{x - \alpha s}{\sigma_t\sqrt{s}}\right) \phi\left(\frac{x - \alpha s}{\sigma_t\sqrt{s}}\right) dW_s^1. \end{aligned}$$

In general, the $(n + 1)$ th derivative takes the following form:

$$\partial_x^{n+1} u_t(x) = \int_0^1 (\sigma_t\sqrt{s})^{-(n+1)} (-1)^{n+1} \text{He}_n\left(\frac{x - \alpha s}{\sigma_t\sqrt{s}}\right) \phi\left(\frac{x - \alpha s}{\sigma_t\sqrt{s}}\right) dW_s^1,$$

where He_n is the n -th *Hermite polynomial (probabilist version)* given by

$$\text{H}_n(x) = (-1)^n e^{x^2/2} \frac{d}{dx} e^{-x^2/2}.$$

In the following lemma, we give an upper bound on the growth rate of the derivatives. Using this, we will prove that, for fixed t , u_t is *real analytic* on the interval $(0, 1)$, and the radius of convergence around x_0 is comparable to $|x_0|$. This is natural as the Conga line gets smoother as we move away from the tip. We start off with the following lemma.

Lemma 3. For $0 < \epsilon < \frac{x}{\alpha}$,

$$\begin{aligned} |\partial_x^{n+1} u_t(x)| &\leq (2\pi)^{1/4} \|W\| \left\{ \left(\frac{\sqrt{2}}{x - \alpha\epsilon}\right)^{n+1} + \frac{1}{x} \left(\frac{\sqrt{2}}{x - \alpha\epsilon}\right)^n \right\} (n+1)! \\ &\quad + \|W\| \left\{ \left(\frac{1}{\sigma_t\sqrt{\epsilon}}\right)^{n+1} + \frac{n+1}{x} \left(\frac{1}{\sigma_t\sqrt{\epsilon}}\right)^n \right\} \sqrt{(n+1)!}, \end{aligned}$$

where $\|W\| = \sup_{0 \leq s \leq 1} |W_s|$.

Proof: In the proof, we consider C as a generic positive constant whose value might change in between steps. Let $K_t^n(x, s) = (\sigma_t \sqrt{s})^{-(n+1)} (-1)^{n+1} \text{He}_n \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \phi \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right)$. Then

$$\begin{aligned} \partial_x^{n+1} u_t(x) &= \int_0^1 (\sigma_t \sqrt{s})^{-(n+1)} (-1)^{n+1} \text{He}_n \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \phi \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) dW_s^1 \\ &= \int_0^1 W(1-s) \partial_s K_t^n(x, s) ds. \end{aligned}$$

So, $|\partial_x^{n+1} u_t(x)| \leq \|W\| \int_0^1 |\partial_s K_t^n(x, s)| ds$. So, we have to estimate the integral $\int_0^1 |\partial_s K_t^n(x, s)| ds$.

Now,

$$\begin{aligned} \partial_s K_t^n(x, s) &= (-1)^n \frac{n+1}{2\sigma_t^{n+1} s^{(n+3)/2}} \text{He}_n \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \phi \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \\ &\quad + (-1)^{n+1} \frac{x + \alpha s}{2\sigma_t s^{3/2}} (\sigma_t \sqrt{s})^{-(n+1)} \text{He}_{n+1} \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \phi \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right). \end{aligned}$$

So,

$$\begin{aligned} \int_0^1 |\partial_s K_t^n(x, s)| ds &\leq \frac{n+1}{x} \int_0^1 (\sigma_t \sqrt{s})^{-n} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \left| \text{He}_n \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \phi \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \right| ds \\ &\quad + \int_0^1 (\sigma_t \sqrt{s})^{-(n+1)} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \left| \text{He}_{n+1} \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \phi \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \right| ds. \end{aligned}$$

From the above, it is clear that estimating the second integral suffices.

$$\int_0^1 (\sigma_t \sqrt{s})^{-(n+1)} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \left| \text{He}_{n+1} \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \phi \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \right| ds \leq I_t^x + J_t^x,$$

where

$$\begin{aligned}
I_t^x &= \int_0^\epsilon (\sigma_t \sqrt{s})^{-(n+1)} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \left| \text{He}_{n+1} \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \phi \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \right| ds \\
&= \int_0^\epsilon (x - \alpha s)^{-(n+1)} \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right)^{n+1} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \left| \text{He}_{n+1} \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \phi \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \right| ds \\
&\leq (x - \alpha \epsilon)^{-(n+1)} \int_0^\epsilon \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right)^{n+1} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \left| \text{He}_{n+1} \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \phi \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \right| ds \\
&= (x - \alpha \epsilon)^{-(n+1)} \int_{\frac{x - \alpha \epsilon}{\sigma_t \sqrt{\epsilon}}}^\infty s^{n+1} |\text{He}_{n+1}(s)| \phi(s) ds \\
&\leq (x - \alpha \epsilon)^{-(n+1)} \left(\int_{\frac{x - \alpha \epsilon}{\sigma_t \sqrt{\epsilon}}}^\infty s^{2n+2} \phi(s) ds \right)^{1/2} \left(\int_{\frac{x - \alpha \epsilon}{\sigma_t \sqrt{\epsilon}}}^\infty \text{He}_{n+1}^2(s) \phi(s) ds \right)^{1/2} \\
&\leq (x - \alpha \epsilon)^{-(n+1)} (2\pi)^{1/4} \sqrt{(n+1)!} \left(\int_0^\infty s^{2n+2} \phi(s) ds \right)^{1/2} \\
&\leq (2\pi)^{1/4} \left(\frac{\sqrt{2}}{x - \alpha \epsilon} \right)^{n+1} (n+1)!.
\end{aligned}$$

Here we use the facts that

$$\int_{-\infty}^\infty \text{He}_n^2(s) \phi(s) ds = n!$$

and

$$\int_0^\infty s^{2n+2} \phi(s) ds = \frac{(2n+2)!}{2^{n+1}(n+1)!} < 2^{n+1}(n+1)!.$$

Similarly,

$$\begin{aligned}
J_t^x &= \int_\epsilon^1 (\sigma_t \sqrt{s})^{-(n+1)} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \left| \text{He}_{n+1} \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \phi \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \right| ds \\
&\leq (\sigma_t \sqrt{\epsilon})^{-(n+1)} \int_{-\infty}^\infty |\text{He}_{n+1}(s)| \phi(s) ds \\
&\leq (\sigma_t \sqrt{\epsilon})^{-(n+1)} \left(\int_{-\infty}^\infty \text{He}_{n+1}^2(s) \phi(s) ds \right)^{1/2} \left(\int_{-\infty}^\infty \phi(s) ds \right)^{1/2} \\
&= \sqrt{(n+1)!} (\sigma_t \sqrt{\epsilon})^{-(n+1)}.
\end{aligned}$$

The lemma follows from the above. \square

From this lemma, it is not too hard to see that u_t is real analytic on $(0, 1)$. Let x_0 be any point in this interval. For $0 < \delta < x_0$, define

$$\Delta_t^n(x_0, \delta) = \left(\sup_{x \in (x_0 - \delta, x_0 + \delta)} |\partial_x^{n+1} u_t(x)| \right) \frac{\delta^{n+1}}{(n+1)!}.$$

The n -th order Taylor polynomial based at x_0 is given by $T_t^n(x) = \sum_{i=0}^n \frac{\partial_x^i u_t(x_0)}{i!} (x - x_0)^i$.

By Taylor's Inequality,

$$\sup_{x \in (x_0 - \delta, x_0 + \delta)} |u_t(x) - T_t^n(x)| \leq \Delta_t^n(x_0, \delta).$$

From the above lemma, we know that, for $\epsilon < \frac{x_0}{\alpha}$,

$$\begin{aligned} \Delta_t^n(x_0, \delta) &\leq (2\pi)^{1/4} \|W\| \left\{ \left(\frac{\sqrt{2}\delta}{x_0 - \delta - \alpha\epsilon} \right)^{n+1} + \frac{\delta}{x_0 - \delta} \left(\frac{\sqrt{2}\delta}{x_0 - \delta - \alpha\epsilon} \right)^n \right\} \\ &\quad + \|W\| \left\{ \left(\frac{\delta}{\sigma_t \sqrt{\epsilon}} \right)^{n+1} + \frac{(n+1)\delta}{x_0 - \delta} \left(\frac{1}{\sigma_t \sqrt{\epsilon}} \right)^n \right\} \frac{1}{\sqrt{(n+1)!}}. \end{aligned}$$

The above error will go to zero only when $\frac{\sqrt{2}\delta}{x_0 - \delta - \alpha\epsilon} < 1$, i.e., $\delta < \frac{x_0 - \alpha\epsilon}{1 + \sqrt{2}}$. We can make ϵ arbitrarily small to get the following:

Corollary 3.2. *The scaled Conga line u_t is real analytic on $(0, 1)$. The power series expansion of u_t around $x_0 \in (0, 1)$ converges in $\left(\frac{\sqrt{2}x_0}{1 + \sqrt{2}}, \min\{\sqrt{2}x_0, 1\} \right)$.*

We are going to use this property of the Conga line multiple times in this article.

3.3 Covariance structure of the derivatives

In the following sections, we will analyse the finer properties of the Conga line like distribution of critical points, length and shape and number of loops. For all of these, fine estimates on the covariance structure of the derivatives are of utmost importance. This section is devoted to finding these estimates.

The next lemma is about the covariance between the first derivatives at two points. In what follows, we write $L_t^x(M) = \alpha^{-1} \sqrt{-M \frac{x}{t} \log \frac{x}{t}}$.

Lemma 4. *For $\delta \leq x, y \leq \alpha$, $\text{Cov}(u_t'(x), u_t'(y)) \geq 0$ and satisfies*

$$\begin{aligned} \text{Cov}(u_t'(x), u_t'(y)) &= \exp \left\{ -\frac{2\alpha t}{\sigma^2} \left(\sqrt{\frac{x^2 + y^2}{2}} - \frac{x + y}{2} \right) \right\} \left(\frac{\sqrt{t}}{2\sqrt{\pi\alpha}\sigma \left(\frac{x^2 + y^2}{2} \right)^{1/4}} \right) \\ &\quad \times \left(1 + O^\infty \left(\sqrt{\frac{\log t}{t}} \right) \right). \end{aligned}$$

Consequently, the correlation function $\rho_t(x, y) = \text{Corr}(u'_t(x), u'_t(y))$ is always non-negative and has the following decay rate

$$\begin{aligned}\rho_t(x, y) &\geq \exp\{-C_1 t(x-y)^2\} \left(\frac{(xy)^{1/4}}{\left(\frac{x^2+y^2}{2}\right)^{1/4}} \right) \left(1 + O^\infty\left(\sqrt{\frac{\log t}{t}}\right) \right), \\ \rho_t(x, y) &\leq \exp\{-C_2 t(x-y)^2\} \left(\frac{(xy)^{1/4}}{\left(\frac{x^2+y^2}{2}\right)^{1/4}} \right) \left(1 + O^\infty\left(\sqrt{\frac{\log t}{t}}\right) \right),\end{aligned}$$

where constants C_1, C_2 depend only on δ and α .

Proof: To prove this lemma, note that, by completing squares in the exponent, we get

$$\begin{aligned}\text{Cov}(u'_t(x), u'_t(y)) &= \int_0^1 \frac{1}{\sigma_t^2 s} \phi\left(\frac{x-\alpha s}{\sigma_t \sqrt{s}}\right) \phi\left(\frac{y-\alpha s}{\sigma_t \sqrt{s}}\right) ds \\ &= \exp\left\{-\frac{2\alpha}{\sigma_t^2} \left(\sqrt{\frac{x^2+y^2}{2}} - \frac{x+y}{2}\right)\right\} \int_0^1 \frac{1}{\sigma_t^2 s} \phi^2\left(\frac{\sqrt{\frac{x^2+y^2}{2}} - \alpha s}{\sigma_t \sqrt{s}}\right) ds.\end{aligned}$$

Now we want to estimate the integral $\int_0^1 \frac{1}{\sigma_t^2 s} \phi^2\left(\frac{x-\alpha s}{\sigma_t \sqrt{s}}\right) ds$ where $\delta \leq x \leq \alpha$. By choosing M large enough, we can ensure that

$$\int_0^1 \frac{1}{\sigma_t^2 s} \phi^2\left(\frac{x-\alpha s}{\sigma_t \sqrt{s}}\right) ds = \int_{\frac{x}{\alpha} - L_t^x(M)}^{\frac{x}{\alpha} + L_t^x(M)} \frac{1}{\sigma_t^2 s} \phi^2\left(\frac{x-\alpha s}{\sigma_t \sqrt{s}}\right) ds + O^\infty\left(\frac{1}{t}\right).$$

Notice that

$$\begin{aligned}\int_{\frac{x}{\alpha} - L_t^x(M)}^{\frac{x}{\alpha} + L_t^x(M)} \frac{1}{\sigma_t^2 s} \phi^2\left(\frac{x-\alpha s}{\sigma_t \sqrt{s}}\right) ds &= \frac{1}{2\pi} \int_{\frac{x}{\alpha} - L_t^x(M)}^{\frac{x}{\alpha} + L_t^x(M)} \frac{x+\alpha s}{2\sigma_t s^{3/2}} \frac{2\sqrt{s}}{\sigma_t(x+\alpha s)} \exp\left\{-\left(\frac{x-\alpha s}{\sigma_t \sqrt{s}}\right)^2\right\} ds \\ &= \frac{1}{2\pi\sigma_t\sqrt{\alpha x}} \int_{\frac{x}{\alpha} - L_t^x(M)}^{\frac{x}{\alpha} + L_t^x(M)} \frac{x+\alpha s}{2\sigma_t s^{3/2}} \exp\left\{-\left(\frac{x-\alpha s}{\sigma_t \sqrt{s}}\right)^2\right\} ds \\ &\quad + O^\infty(\sqrt{\log t}) \\ &= \frac{1}{2\pi\sigma_t\sqrt{\alpha x}} \int_{-\infty}^{\infty} e^{-s^2} ds + O^\infty(\sqrt{\log t}) \\ &= \frac{\sqrt{t}}{2\sigma\sqrt{\pi\alpha x}} + O^\infty(\sqrt{\log t}).\end{aligned}$$

Substituting $\sqrt{\frac{x^2+y^2}{2}}$ in place of x proves the lemma. \square

Lemma 5. For $\delta \leq x \leq \alpha$,

$$\text{Var}(u_t''(x)) = \frac{\sqrt{\alpha} t^{3/2}}{4\sqrt{\pi}\sigma^3 x^{3/2}} \left(1 + O^\infty \left(\sqrt{\frac{\log t}{t}} \right) \right).$$

Proof: Follows along the same lines as the proof of Lemma 4. \square

Lemma 6. For $\delta \leq x \leq \alpha$,

$$\text{Cov}(u_t'(x), u_t''(x)) = O^\infty(\sqrt{t \log t}).$$

Proof:

$$\begin{aligned} \text{Cov}(u_t'(x), u_t''(x)) &= \int_0^1 \frac{1}{(\sigma_t \sqrt{s})^3} \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \phi^2 \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) ds \\ &= \frac{1}{\pi} \int_{\frac{x}{\alpha} - L_t^x(M)}^{\frac{x}{\alpha} + L_t^x(M)} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \frac{1}{\sigma_t^2 (x + \alpha s)} \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \exp \left\{ - \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right)^2 \right\} ds \\ &\quad + O^\infty(t^{-1}) \\ &= \frac{1}{\pi} \int_{\frac{x}{\alpha} - L_t^x(M)}^{\frac{x}{\alpha} + L_t^x(M)} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \frac{1}{2\sigma_t^2 x} (1 + f_t(x, s)) \\ &\quad \times \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \exp \left\{ - \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right)^2 \right\} ds \\ &\quad + O^\infty(t^{-1}), \end{aligned}$$

where $f_t(x, s) = \frac{2x}{x + \alpha s} - 1$.

Using the fact that $\int_{-\infty}^{\infty} s \exp(-s^2) ds = 0$, we get

$$\frac{1}{\pi} \int_{\frac{x}{\alpha} - L_t^x(M)}^{\frac{x}{\alpha} + L_t^x(M)} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \frac{1}{2\sigma_t^2 x} \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \exp \left\{ - \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right)^2 \right\} ds = O^\infty(t^{-1}) \quad (1.8)$$

choosing sufficiently large M .

Also note that

$$f_t(x, s) = O^\infty \left(\sqrt{\frac{\log t}{t}} \right)$$

for $\delta \leq x \leq \alpha$, $\frac{x}{\alpha} - L_t^x(M) \leq s \leq \frac{x}{\alpha} + L_t^x(M)$, which yields

$$\frac{1}{\pi} \int_{\frac{x}{\alpha} - L_t^x(M)}^{\frac{x}{\alpha} + L_t^x(M)} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \frac{f_t(x, s)}{2\sigma_t^2 x} \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \exp \left\{ - \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right)^2 \right\} ds = O^\infty(\sqrt{t \log t}). \quad (1.9)$$

(1.8) and (1.9) prove the lemma. \square

Corollary 3.3.

$$\text{Corr}(u_t'(x), u_t''(x)) = O^\infty \left(\sqrt{\frac{\log t}{t}} \right).$$

Proof: This follows from Lemmas 4, 5 and 6. \square

Let $\Sigma_t(x, y)$ be the covariance matrix of $(u_t'(x), u_t'(y))$. We need the following lemma to estimate the determinant of the matrix.

Lemma 7. *There exist constants C^*, C_1, C_2 such that, for $\delta \leq x, y \leq \alpha$ with $|x - y| \leq \frac{C^*}{\sqrt{t}}$,*

$$C_1 t^2 (y - x)^2 \leq \det \Sigma_t(x, y) \leq C_2 t^2 (y - x)^2$$

Proof: We fix $x \in [\delta, \alpha]$ and consider the function $\Psi_{t,x}(y) = \det \Sigma_t(x, y)$. Consider the function $g_t(y) = \text{Var } u_t'(y) = \int_0^1 \frac{1}{\sigma_t^2 s} \phi^2 \left(\frac{y - \alpha s}{\sigma_t \sqrt{s}} \right) ds$. Let H_n denote the n -th *Hermite polynomial (physicist version)* given by

$$H_n(x) = (-1)^n e^{x^2} \frac{d}{dx} e^{-x^2} = 2^{n/2} \text{He}_n(\sqrt{2}x).$$

Then we can write the n -th derivative of g_t as

$$g_t^{(n)}(x) = (-1)^n \int_0^1 \frac{1}{\sigma_t^2 s} \frac{1}{(\sigma_t \sqrt{s})^n} H_n \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) \phi^2 \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) ds.$$

Using the fact that $\int_{-\infty}^{\infty} H_n(s) \exp\{-s^2\} ds = 0$ and the same technique as the proof of Lemma 6, one can show that, for $n \geq 1$,

$$g_t^{(n)}(\cdot) = O^\infty(t^{n/2} \sqrt{\log t}) \quad (1.10)$$

Let

$$E_{t,x}(y) = g_t(x)g_t(y) - g_t^2(\eta)$$

and

$$F_{t,x}(y) = \left[1 - \exp \left\{ -\frac{4\alpha t}{\sigma^2} \left(\sqrt{\frac{x^2 + y^2}{2}} - \frac{x + y}{2} \right) \right\} \right] g_t^2(\eta).$$

Then, writing down $\text{Cov}(u'_t(x), u'_t(y))$ as in the proof of Lemma 4, we get

$$\Psi_{t,x}(y) = F_{t,x}(y) + E_{t,x}(y). \quad (1.11)$$

It is easy to check that

$$E_{t,x}(x) = E'_{t,x}(x) = 0. \quad (1.12)$$

The double derivative of $E_{t,x}$ takes the form

$$E''_{t,x}(y) = g_t(x)g_t''(y) - 2(g_t'(\eta))^2(\partial_y\eta)^2 - 2(g_t(\eta))(g_t''(\eta))(\partial_y\eta)^2 - 2(g_t(\eta))(g_t'(\eta))(\partial_y^2\eta).$$

Using (1.10) we deduce

$$E_{t,x}(\cdot) = O^\infty(t^{3/2}\sqrt{\log t}),$$

which, along with (1.12) yields

$$|E_{t,x}(y)| \leq Ct^{3/2}\sqrt{\log t} (y - x)^2 \quad (1.13)$$

for some constant $C < \infty$.

Now, to estimate $F_{t,x}$, note that in the region $\delta \leq x, y \leq \alpha$,

$$\exp \left\{ -\frac{\alpha t(y - x)^2}{2\delta\sigma^2} \right\} \leq \exp \left\{ -\frac{4\alpha t}{\sigma^2} \left(\sqrt{\frac{x^2 + y^2}{2}} - \frac{x + y}{2} \right) \right\} \leq \exp \left\{ -\frac{t(y - x)^2}{2\sigma^2} \right\}. \quad (1.14)$$

Using (1.14) along with the fact that $e^{-C}x \leq 1 - e^{-x} \leq x$ on $0 \leq x \leq C$, and Lemma 4, we get

$$C^*t^2(y - x)^2 \leq F_{t,x}(y) \leq Ct^2(y - x)^2, \quad (1.15)$$

where C, C^* are positive, finite constants.

(1.13) and (1.15) together prove the lemma. \square

3.4 Analyzing the distribution of critical points

Let $N_t(I)$ denote the number of critical points of u_t in an interval $I \in [\delta, \alpha]$. Then N_t defines a *simple point process* on $[\delta, \alpha]$. Our first goal is to find out the *first intensity* of this process. For this, we use the *Expectation meta-theorem for smooth Gaussian fields* (see [1] p. 263), which implies the following:

$$\mathbb{E}(N_t(I)) = \int_I \mathbb{E}(|u_t''(y)| |u_t'(y) = 0) p_t^y(0) dy, \quad (1.16)$$

where p_t^y is the density of $u_t'(y)$. Before we go further, we remark that the meta-theorem from [1] mentioned above is a very general theorem which holds in a much wider set-up under a set of assumptions. In our case, it is easy to check that all the assumptions hold. Now, we utilize (1.16) and the developments in subsection 3.3 to derive a nice expression for the first intensity.

Lemma 8. *The first intensity ρ_t for N_t satisfies*

$$\rho_t(x) = \frac{\sqrt{\alpha t}}{\pi \sigma \sqrt{2x}} \left(1 + O^\infty \left(\sqrt{\frac{\log t}{t}} \right) \right) \quad (1.17)$$

for $\delta \leq x \leq \alpha$.

Proof: By standard formulae for normal conditional densities and the lemmas proved in Subsection 3.3, we manipulate (1.16) as follows:

$$\begin{aligned} \mathbb{E}(N_t(I)) &= \sqrt{\frac{2}{\pi}} \int_I \left[\frac{\text{Var}(u_t'(y)) \cdot \text{Var}(u_t''(y)) - \text{Cov}^2(u_t'(y), u_t''(y))}{\text{Var}(u_t'(y))} \right]^{\frac{1}{2}} \frac{1}{\sqrt{2\pi \text{Var}(u_t'(y))}} dy \\ &= \sqrt{\frac{2}{\pi}} \int_I \frac{\left(1 + O^\infty \left(\frac{\log t}{t} \right) \right)}{\sqrt{2\pi}} \left[\frac{\text{Var}(u_t''(y))}{\text{Var}(u_t'(y))} \right]^{\frac{1}{2}} dy \\ &= \int_I \frac{\left(1 + O^\infty \left(\sqrt{\frac{\log t}{t}} \right) \right)}{\pi} \left[\frac{\alpha t}{2\sigma^2 y} \right]^{\frac{1}{2}} dy \\ &= \int_I \left(1 + O^\infty \left(\sqrt{\frac{\log t}{t}} \right) \right) \frac{\sqrt{\alpha t}}{\sqrt{2\pi\sigma}} \frac{1}{\sqrt{y}} dy. \end{aligned}$$

Here, we use Corollary 3.3 for to get the second equality and the estimates proved in Lemma 4 and Lemma 5 to get the third equality. This proves the lemma. \square

Thus $\hat{\rho}_t(x) = \frac{\sqrt{\alpha t}}{\pi\sigma\sqrt{2x}}$ gives us the approximate first intensity for N_t . From this, we see that the expected number of critical points in a small interval $[x, x+h]$ is approximately $\frac{\sqrt{\alpha t}h}{\pi\sigma\sqrt{2x}}$.

Now that we know the first intensity reasonably accurately, we can ask finer questions about the distribution of critical points, such as

(i) What can we say about the spacings of the critical points? Are there points in $[\delta, \alpha]$ around which there is a large concentration of critical points, or are they more or less well-spaced?

(ii) Given an interval $I \in [\delta, \alpha]$, how good is $\mathbb{E}N_t(I)$ as an estimate of $N_t(I)$?

The next lemma answers (i) by estimating the second intensity of N_t . First we present a formula for the second intensity of N_t taken from [1].

$$\mathbb{E}(N_t(I)^2 - N_t(I)) = \int_{I \times I} \mathbb{E}(|u_t''(y)||u_t''(z)| | u_t'(y) = 0, u_t'(z) = 0) p_t^{y,z}(0) dy dz, \quad (1.18)$$

where $p_t^{y,z}$ is the joint density of $(u_t'(y), u_t'(z))$.

In the following, C^* represents a positive constant.

Lemma 9. For $t > \frac{4(1 + \sqrt{2})^2}{2\delta^2}$, and $\delta \leq y, z \leq \alpha$ with $|y - z| \leq \frac{C^*}{\sqrt{t}}$,

$$\mathbb{E}(|u_t''(y)||u_t''(z)| | u_t'(y) = 0, u_t'(z) = 0) \leq C(y - z)^2 t^{5/2}$$

for some constant $C > 0$.

Proof: The hypothesis of the lemma tells us that y and z lie in the region of analyticity of each other, i.e. we can write

$$u_t'(y) = \sum_{n=0}^{\infty} u_t^{(n+1)}(z) \frac{(y-z)^n}{n!},$$

and the same holds with y and z interchanged. If we know that $u'_t(y) = 0$ and $u'_t(z) = 0$, the above equation becomes

$$0 = \sum_{n=1}^{\infty} u_t^{(n+1)}(z) \frac{(y-z)^n}{n!}.$$

From this, we can solve for $u_t''(z)$ to get

$$u_t''(z) = - \sum_{n=2}^{\infty} u_t^{(n+1)}(z) \frac{(y-z)^{n-1}}{n!},$$

and the same holds with y and z interchanged. Thus, the conditional expectation in (1.18) becomes

$$\begin{aligned} & \mathbb{E} \left[\left| - \sum_{n=1}^{\infty} u_t^{(n+2)}(z) \frac{(y-z)^n}{(n+1)!} \right| \left| - \sum_{n=1}^{\infty} u_t^{(n+2)}(y) \frac{(y-z)^n}{(n+1)!} \right| \left| u'_t(y) = 0, u'_t(z) = 0 \right] \\ &= (y-z)^2 \mathbb{E} \left[\left| - \sum_{n=0}^{\infty} u_t^{(n+3)}(z) \frac{(y-z)^n}{(n+2)!} \right| \left| - \sum_{n=0}^{\infty} u_t^{(n+3)}(y) \frac{(y-z)^n}{(n+2)!} \right| \left| u'_t(y) = 0, u'_t(z) = 0 \right]. \end{aligned}$$

Now, by the Cauchy-Schwarz inequality and the fact that the conditional variance is bounded above by the total variance, we have

$$\begin{aligned} & \mathbb{E} \left(|u_t^{(m)}(y)| |u_t^{(n)}(z)| \mid u'_t(y) = 0, u'_t(z) = 0 \right) \\ & \leq \sqrt{\mathbb{E} \left(u_t^{(m)}(y)^2 \mid u'_t(y) = 0, u'_t(z) = 0 \right)} \sqrt{\mathbb{E} \left(u_t^{(n)}(z)^2 \mid u'_t(y) = 0, u'_t(z) = 0 \right)} \\ & \leq \sqrt{\mathbb{E} \left(u_t^{(m)}(y) \right)^2} \sqrt{\mathbb{E} \left(u_t^{(n)}(z) \right)^2}. \end{aligned}$$

We know that

$$u_t^{(m+3)}(y) = \int_0^1 (\sigma_t \sqrt{s})^{-(m+3)} (-1)^{m+3} \text{He}_{m+2} \left(\frac{y - \alpha s}{\sigma_t \sqrt{s}} \right) \phi \left(\frac{y - \alpha s}{\sigma_t \sqrt{s}} \right) dW_s^1.$$

So, using the same techniques as in the proof of Lemma 3, for $0 < \epsilon < \frac{\delta}{2\alpha}$, we estimate the

variance as

$$\begin{aligned} \mathbb{E} \left(u_t^{(m+3)}(y) \right)^2 &= \int_0^1 (\sigma_t \sqrt{s})^{-2(m+3)} \text{He}_{m+2}^2 \left(\frac{y - \alpha s}{\sigma_t \sqrt{s}} \right) \phi^2 \left(\frac{y - \alpha s}{\sigma_t \sqrt{s}} \right) ds \\ &\leq \frac{C_1 t}{\sigma^2 (y - \alpha \epsilon)^{2m+4}} \int_0^\infty s^{2m+3} \text{He}_{m+2}^2(s) \phi^2(s) ds \\ &\quad + \frac{C_2}{(\sigma_t \sqrt{\epsilon})^{2m+5}} \int_0^\infty \text{He}_{m+2}^2(s) \phi^2(s) ds. \end{aligned}$$

To estimate the first integral, we note that the function

$$g_m(s) = s^{2m+3} \exp\{-s^2/2\}$$

is maximised at $s = \sqrt{2m+3}$. So, $g_m(s) \leq (2m+3)^{(2m+3)/2} \exp\{-(2m+3)/2\}$.

By Stirling's Formula,

$$\Gamma(n) = \sqrt{\frac{2\pi}{n}} \left(\frac{n}{e}\right)^n (1 + O(n^{-1})),$$

where $\Gamma(\cdot)$ is the Gamma function. Using this, we get $g_m(s) \leq C\sqrt{m}2^m(m+1)!$. So,

$$\int_0^\infty s^{2m+3} \text{He}_{m+2}^2(s) \phi^2(s) ds \leq C\sqrt{m}2^m(m+1)! \int_{-\infty}^\infty \text{He}_{m+2}^2(s) \phi(s) ds \leq C2^m \{(m+2)!\}^2.$$

The second integral is easier to estimate. Finally, we get

$$\mathbb{E} \left(u_t^{(m+3)}(y) \right)^2 \leq \frac{C_1 t}{\sigma^2 (y - \alpha \epsilon)^{2m+4}} 2^m \{(m+2)!\}^2 + \frac{C_2}{(\sigma_t \sqrt{\epsilon})^{2m+5}} \{(m+2)!\}.$$

Therefore,

$$\begin{aligned} \frac{\sqrt{\mathbb{E} \left(u_t^{(m+3)}(y) \right)^2}}{(m+2)!} |y-z|^m &\leq C \left[t \left(\frac{\sqrt{2}|y-z|}{y-\alpha\epsilon} \right)^{2m} \left(\frac{1}{y-\alpha\epsilon} \right)^4 \right. \\ &\quad \left. + \frac{1}{(m+2)!} \left(\frac{\sqrt{t}|y-z|}{\sigma\sqrt{\epsilon}} \right)^{2m} \left(\frac{\sqrt{t}}{\sigma\sqrt{\epsilon}} \right)^5 \right]^{1/2} \\ &\leq Ct^{5/4} a_m(t, y, z), \end{aligned}$$

where, by the assumptions of the lemma, $\sum_{m=0}^\infty a_m(t, y, z) \leq C$, where C is a constant that

does not depend on t, y, z . Thus, we have

$$\begin{aligned} \mathbb{E}(|u_t''(y)||u_t''(z)| \mid u_t'(y) = 0, u_t'(z) = 0) &\leq (y-z)^2 \sum_{m,n=0}^{\infty} \frac{\sqrt{\mathbb{E}(u_t^{(m+3)}(y))^2}}{(m+2)!} |y-z|^m \\ &\quad \times \frac{\sqrt{\mathbb{E}(u_t^{(n+3)}(y))^2}}{(n+2)!} |y-z|^n \\ &\leq C(y-z)^2 t^{5/2} \left(\sum_{m=0}^{\infty} a_m(t, y, z) \right)^2, \end{aligned}$$

which proves the lemma. \square

We know that $p_t^{y,z}(0) = \frac{1}{2\pi\sqrt{\det \Sigma_t(y, z)}}$. Using Lemmas 7 and 9, we get

Lemma 10. *For $t > \frac{4(1+\sqrt{2})^2}{2\delta^2}$, and $h \leq C^*t^{-1/2}$,*

$$\mathbb{E}(N_t^2([x, x+h]) - N_t([x, x+h])) \leq C(\delta)h^3t^{3/2}.$$

In particular, we get

$$\mathbb{E}(N_t([x, x+h])\mathbb{I}(N_t([x, x+h]) \geq 2)) \leq C(\delta)h^3t^{3/2}.$$

Using this lemma, we can deduce that if we divide $[\delta, \alpha]$ into subintervals of sufficiently small length, the number of critical points in any of these should not exceed one. The following corollary makes this precise.

Corollary 3.4. *Let $\{a_t\}$ be any sequence such that $a_t = o(t^{-1/4})$. Divide the interval $[\delta, \alpha]$ into subintervals $I_1, \dots, I_{\lfloor \sqrt{t}/a_t \rfloor + 1}$ of length at most $\frac{a_t}{\sqrt{t}}$. Then*

$$P\left(\max_{1 \leq j \leq \lfloor \sqrt{t}/a_t \rfloor + 1} N_t(I_j) \geq 2\right) \leq C(\delta)a_t^2t^{1/2} \rightarrow 0$$

as $t \rightarrow \infty$.

This follows easily from Lemma 10 using the union bound.

Now, we answer (ii).

Note that for a Poisson point process, the first intensity determines the whole process. The Conga line lacks the Markov property. We can think of it as a process that ‘*gains smoothness at the cost of Markov property*’. But Lemma 4 tells us that there is *exponential decorrelation*, i.e. pieces of the Conga line that are reasonably far apart are almost independent. Thus, we expect that the first intensity of N_t should give us a lot of information about the process N_t itself. We conclude this section on critical points by giving basis to this intuition by showing the following:

Lemma 11. *Let $I \subseteq [\delta, \alpha]$ be an interval. Then*

$$\frac{N_t(I)}{\mathbb{E}N_t(I)} \xrightarrow{P} 1$$

as $t \rightarrow \infty$.

Proof: It suffices to prove the result for $I = [\delta, \alpha]$.

Consider a collection of intervals

$$\mathcal{C} = \{I_j : 1 \leq j \leq C[\sqrt{t}/r]\}$$

where each interval is of length $\frac{1}{\sqrt{t}}$ in $[\delta, \alpha]$, and $d(I_j, I_k) \geq \frac{r}{\sqrt{t}}$ for a sufficiently large r (which can be a function of t), whose optimal choice will be made later, and $d(A, B)$ represents the usual distance between sets A and B . Using the long range independence of the Conga line (see Lemma 4), we will prove that $\text{Var}\left(N_t\left(\bigcup_{j=1}^{C[\sqrt{t}/r]} I_j\right)\right)$ is very small compared to $\mathbb{E}\left(N_t\left(\bigcup_{j=1}^{C[\sqrt{t}/r]} I_j\right)\right)^2$. The proof is completed by covering $[\delta, \alpha]$ with $[r]$ translates $\mathcal{C}_1, \dots, \mathcal{C}_{[r]}$ of such collections and an application of Chebychev Inequality.

Note that all the constants used in this proof depend on δ .

We begin by computing $\mathbb{E}(N_t(I_1)N_t(I_2))$ using an analogue of the Expectation meta-theorem (which can also be derived from the second intensity formula (1.18)) as follows:

$$\begin{aligned} \mathbb{E}(N_t(I_1)N_t(I_2)) &= \int_{I_1 \times I_2} \mathbb{E}(|u_t''(y)u_t''(z)| \mid u_t'(y) = 0, u_t'(z) = 0) \\ &\quad \times \frac{1}{2\pi\sqrt{\det \Sigma_t(y, z)}} dydz \end{aligned} \quad (1.19)$$

where $\Sigma_t(y, z)$ is the covariance matrix for $(u'_t(y), u'_t(z))$. We know that if

$$(u''_t(y), u''_t(z), u'_t(y), u'_t(z)) \sim N(0, \Sigma),$$

then

$$(u''_t(y), u''_t(z) \mid u'_t(y) = 0, u'_t(z) = 0) \sim N(0, \Sigma^*),$$

where $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$ and $\Sigma^* = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. Now

$$\begin{aligned} \text{Cov}(|u''_t(y)|, |u''_t(z)| \mid u'_t(y) = 0, u'_t(z) = 0) &= \frac{2}{\pi} \sqrt{\sigma_{11}^* \sigma_{22}^*} \left(\rho_{12}^* \arcsin \rho_{12}^* + \sqrt{1 - \rho_{12}^{*2}} - 1 \right) \\ &\leq \frac{2}{\pi} \sqrt{\sigma_{11}^* \sigma_{22}^*} (\rho_{12}^* \arcsin \rho_{12}^*) \\ &\leq \sigma_{12}^* = \sigma_{12} - (\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})_{12}. \end{aligned} \quad (1.20)$$

Take $(y, z) \in I_j \times I_k$, where $I_j, I_k \in \mathcal{C}$ with $j \neq k$.

The proof of Lemma 4 shows that for $\eta = \sqrt{\frac{y^2+z^2}{2}}$,

$$\text{Cov}(u'_t(y), u'_t(z)) \leq \exp\{-Ct(y-z)^2\} \text{Var}(u'_t(\eta))$$

and

$$\text{Var}(u'_t(\eta)) \leq C_1 \sqrt{t}.$$

So, as $|y-z| \geq \frac{r}{\sqrt{t}}$,

$$\begin{aligned} \text{Cov}(u'_t(y), u'_t(z)) &\leq \exp\{-Cr^2\} \text{Var}(u'_t(\eta)) \\ &\leq C_1 \sqrt{t} \exp\{-C_2 r^2\}. \end{aligned} \quad (1.21)$$

Calculations similar to those in the proof of Lemma 4 show

$$\begin{aligned} \text{Cov}(u'_t(y), u''_t(z)) &= - \int_0^1 \frac{1}{(\sigma_t \sqrt{s})^3} \left(\frac{z - \alpha s}{\sigma_t \sqrt{s}} \right) \phi \left(\frac{y - \alpha s}{\sigma_t \sqrt{s}} \right) \phi \left(\frac{z - \alpha s}{\sigma_t \sqrt{s}} \right) ds \\ &= - \exp \left\{ -\frac{2\alpha}{\sigma_t^2} \left(\sqrt{\frac{y^2+z^2}{2}} - \frac{y+z}{2} \right) \right\} \\ &\quad \times \int_0^1 \frac{1}{(\sigma_t \sqrt{s})^3} \left(\frac{z - \alpha s}{\sigma_t \sqrt{s}} \right) \phi^2 \left(\frac{\sqrt{\frac{y^2+z^2}{2}} - \alpha s}{\sigma_t \sqrt{s}} \right) ds. \end{aligned}$$

and thus

$$\begin{aligned}
|\text{Cov}(u'_t(y), u''_t(z))| &\leq C_1 \exp\{-C_2 r^2\} \left(\int_0^1 |z - \eta| \frac{1}{(\sigma_t \sqrt{s})^4} \phi^2 \left(\frac{\eta - \alpha s}{\sigma_t \sqrt{s}} \right) ds \right. \\
&\quad \left. + \int_0^1 \frac{|\eta - \alpha s|}{(\sigma_t \sqrt{s})^4} \phi^2 \left(\frac{\eta - \alpha s}{\sigma_t \sqrt{s}} \right) ds \right) \\
&\leq C_1 \exp\{-C_2 r^2\} t^{3/2}.
\end{aligned} \tag{1.22}$$

Also, from Lemma 6,

$$|\text{Cov}(u'_t(y), u''_t(y))| \leq C \sqrt{t \log t}. \tag{1.23}$$

Similar calculations also show

$$\begin{aligned}
|\text{Cov}(u''_t(y), u''_t(z))| &\leq C_1 \exp\{-C_2 r^2\} \int_0^1 \frac{1}{(\sigma_t \sqrt{s})^4} \left| \frac{y - \alpha s}{\sigma_t \sqrt{s}} \right| \left| \frac{z - \alpha s}{\sigma_t \sqrt{s}} \right| \phi^2 \left(\frac{\sqrt{\frac{y^2 + z^2}{2}} - \alpha s}{\sigma_t \sqrt{s}} \right) ds \\
&\leq C_1 \exp\{-C_2 r^2\} t^{5/2},
\end{aligned} \tag{1.24}$$

writing $\frac{y - \alpha s}{\sigma_t \sqrt{s}}$ as $\frac{y - \eta}{\sigma_t \sqrt{s}} + \frac{\eta - \alpha s}{\sigma_t \sqrt{s}}$ and similarly for $\frac{z - \alpha s}{\sigma_t \sqrt{s}}$. Furthermore, we see that

$$\begin{aligned}
\det \Sigma_{22} = \det \Sigma_t(y, z) &= \text{Var}(u'_t(y)) \text{Var}(u'_t(z)) - \text{Cov}^2(u'_t(y), u'_t(z)) \\
&\geq \text{Var}(u'_t(y)) \text{Var}(u'_t(z)) - \exp\{-Cr^2\} \text{Var}^2(u'_t(\eta)) \\
&\geq \frac{1}{2} \text{Var}(u'_t(y)) \text{Var}(u'_t(z))
\end{aligned} \tag{1.25}$$

for sufficiently large r . Using equations (1.21), ..., (1.25) to estimate the right side of equation (1.20), we see that there is a $K > 0$ for which

$$\text{Cov}(|u''_t(y)|, |u''_t(z)| \mid u'_t(y) = 0, u'_t(z) = 0) \leq C_1 \exp\{-C_2 r^2\} t^K.$$

Plugging this into the expression (1.19), we get

$$\mathbb{E}(N_t(I_1)N_t(I_2)) \leq C \int_{I_1 \times I_2} \frac{C_1 \exp\{-C_2 r^2\} t^K + \frac{2}{\pi} \sqrt{\text{Var}(u''_t(y))} \sqrt{\text{Var}(u''_t(z))}}{2\pi \sqrt{\text{Var}(u'_t(y))} \text{Var}(u'_t(z))} dy dz. \tag{1.26}$$

We know from Lemma 5 that

$$\text{Var}(u''_t(y)) \leq Ct^{3/2}.$$

Thus

$$\frac{2}{\pi} \sqrt{\text{Var}(u_t''(y))} \sqrt{\text{Var}(u_t''(z))} = O(t^{3/2}).$$

If we choose $r = \sqrt{M \log t}$ for a large enough M , then

$$C_1 \exp\{-C_2 r^2\} t^K \ll \frac{2}{\pi} \sqrt{\text{Var}(u_t''(y))} \sqrt{\text{Var}(u_t''(z))}.$$

Consequently, from (1.26),

$$\begin{aligned} \mathbb{E}(N_t(I_1)N_t(I_2)) &\leq C \int_{I_1 \times I_2} \frac{\sqrt{\text{Var}(u_t''(y))} \sqrt{\text{Var}(u_t''(z))}}{\pi^2 \sqrt{\text{Var}(u_t'(y))} \sqrt{\text{Var}(u_t'(z))}} dy dz \\ &= C \left(\int_{I_1} \frac{\sqrt{\text{Var}(u_t''(y))}}{\pi \sqrt{\text{Var}(u_t'(y))}} dy \right) \left(\int_{I_2} \frac{\sqrt{\text{Var}(u_t''(z))}}{\pi \sqrt{\text{Var}(u_t'(z))}} dz \right) \\ &= \left(1 + O\left(\sqrt{\frac{\log t}{t}}\right) \right) \mathbb{E}(N_t(I_1)) \mathbb{E}(N_t(I_2)), \end{aligned}$$

where the last step above follows from (1.16) using Corollary 3.3 (see the proof of Lemma 8).

Thus,

$$\text{Cov}(N_t(I_1), N_t(I_2)) = O\left(\sqrt{\frac{\log t}{t}}\right) \quad (1.27)$$

for this choice of r .

Now, we have all we need to compute the variance of $N_t\left(\bigcup_{j=1}^{C[\sqrt{t}/r]} I_j\right)$.

$$\begin{aligned} \text{Var} N_t \left(\bigcup_{j=1}^{C[\sqrt{t}/r]} I_j \right) &= \sum_{j=1}^{C[\sqrt{t}/r]} \text{Var} N_t(I_j) + 2 \sum_{i < j} \text{Cov}(N_t(I_i), N_t(I_j)) \\ &\leq C_1 \frac{\sqrt{t}}{r} + C_2 (\sqrt{t \log t}) r^{-2}, \end{aligned}$$

where we used Lemma 10 crucially in putting the constant bound on $\text{Var} N_t(I_j)$.

With our choice of $r = \sqrt{M \log t}$, the above becomes

$$\text{Var} N_t \left(\bigcup_{j=1}^{C[\sqrt{t}/r]} I_j \right) \leq C \sqrt{\frac{t}{\log t}}. \quad (1.28)$$

Finally, we have, for small $\epsilon > 0$,

$$\begin{aligned}
P\left(\left|\frac{N_t([\delta, \alpha])}{\mathbb{E}N_t([\delta, \alpha])} - 1\right| \geq \epsilon\right) &= P(|N_t([\delta, \alpha]) - \mathbb{E}N_t([\delta, \alpha])| \geq \epsilon \mathbb{E}N_t([\delta, \alpha])) \\
&\leq \sum_{l=1}^{\lfloor \sqrt{M \log t} \rfloor} P\left(\left|N_t\left(\bigcup_{j \in \mathcal{C}_l} I_j\right) - \mathbb{E}N_t\left(\bigcup_{j \in \mathcal{C}_l} I_j\right)\right| \geq \epsilon \mathbb{E}N_t\left(\bigcup_{j \in \mathcal{C}_l} I_j\right)\right) \\
&\leq \sum_{l=1}^{\lfloor \sqrt{M \log t} \rfloor} \frac{\text{Var } N_t\left(\bigcup_{j \in \mathcal{C}_l} I_j\right)}{\epsilon^2 \left(\mathbb{E}N_t\left(\bigcup_{j \in \mathcal{C}_l} I_j\right)\right)^2} \\
&\leq \frac{C}{\epsilon^2} \sqrt{\log t} \frac{\sqrt{\frac{t}{\log t}}}{\left(\sqrt{\frac{t}{\log t}}\right)^2} = \frac{C \log t}{\epsilon^2 \sqrt{t}},
\end{aligned}$$

which goes to zero as $t \rightarrow \infty$. □

4 The two dimensional Conga line

Here we study properties of the two dimensional continuous Conga line which is the approximation of our original discrete model in the plane.

4.1 Analyzing length

The length of the Conga line in the interval $[\delta, \alpha]$ is given by $l_t = \int_{\delta}^{\alpha} |u'_t(x)| dx$. In this section, we give estimates for the expected length and its concentration about the mean.

Lemma 12. $\mathbb{E}(l_t) \sim t^{1/4}$.

Proof: From Lemma 4, we see that

$$\begin{aligned}
\mathbb{E}(l_t) = \int_{\delta}^{\alpha} \mathbb{E}|u'_t(x)| dx &= \int_{\delta}^{\alpha} \left(\frac{t^{1/4}}{(\pi^3 \alpha \sigma^2)^{1/4} x^{1/4}}\right) dx + O\left(\frac{\sqrt{\log t}}{t^{1/4}}\right) \\
&= \frac{4t^{1/4}}{3(\pi^3 \alpha \sigma^2)^{1/4}} \left(\alpha^{3/4} - \delta^{3/4}\right) + O\left(\frac{\sqrt{\log t}}{t^{1/4}}\right).
\end{aligned}$$

□

But this gives us only a rough estimate of the behaviour of length for large t . To get a better idea of how the length behaves for large time t , we need higher moments and, if

possible, some form of concentration about the mean. The next lemma gives us an estimate of the variance of l_t .

Lemma 13.

$$\text{Var}(l_t) = O(1).$$

Proof: From Lemma 4, we know that, for $\delta \leq x, y \leq \alpha$,

$$0 \leq \text{Cov}(u'_t(x), u'_t(y)) \leq C_1 \exp\{-C_2 t(x-y)^2\} \left(\frac{\sqrt{t}}{\left(\frac{x^2+y^2}{2}\right)^{1/4}} \right) \left(1 + O^\infty \left(\sqrt{\frac{\log t}{t}} \right) \right).$$

Also, note that

$$\begin{aligned} \text{Cov}(|u'_t(x)|, |u'_t(y)|) &= \frac{2}{\pi} \sqrt{\text{Var}(u'_t(x))} \sqrt{\text{Var}(u'_t(y))} \left[\rho_t(x, y) \arcsin(\rho_t(x, y)) + \sqrt{1 - \rho_t^2(x, y)} - 1 \right] \\ &\leq \text{Cov}(u'_t(x), u'_t(y)). \end{aligned}$$

Thus,

$$\begin{aligned} \text{Var}(l_t) &= \int_\delta^\alpha \int_\delta^\alpha \text{Cov}(|u'_t(x)|, |u'_t(y)|) dx dy \\ &\leq \int_\delta^\alpha \int_\delta^\alpha C \exp\{-C_2 t(x-y)^2\} \left(\frac{\sqrt{t}}{\left(\frac{x^2+y^2}{2}\right)^{1/4}} \right) dx dy \\ &= O(1). \end{aligned}$$

□

Thus, although the expected length grows like $t^{1/4}$, the variance is bounded. This already tells us that the actual length cannot deviate much from the expected length.

In what we do next, we get *Gaussian concentration* of length about the mean in a window of scale $O(\sqrt{\log t})$.

We know that for most useful concentration results, we need some ‘independence’ in our

model. Our strategy here is to construct a new process \hat{u}_t which is very ‘close’ to the original process u_t and is nicer to analyze as $\hat{u}_t(x)$ and $\hat{u}_t(x')$ are *independent* whenever x and x' are sufficiently far apart. As this yields a useful tool which is going to be used in later sections, we give a detailed construction.

Construction of \hat{u}_t

By Lemma 4, we see that the correlation between points x and y in the Conga line with $|x - y| = \frac{\lambda}{\sqrt{t}}$ decays like $e^{-C\lambda^2}$ as λ increases. We make use of this fact.

Divide the interval $\left[\delta - \sqrt{\frac{M \log t}{t}}, \alpha + \sqrt{\frac{M \log t}{t}} \right]$ into subintervals

$$I_k = [y_k, y_{k+1}], \quad -1 \leq k \leq \left[(\alpha - \delta) \frac{\sqrt{t}}{\sqrt{M \log t}} \right] + 1$$

of length at most $\frac{\sqrt{M \log t}}{\sqrt{t}}$. Define the process \hat{u}_t as

$$\hat{u}_t(x) = \int_{y_k/\alpha}^{y_{k+3}/\alpha} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \phi\left(\frac{x - \alpha s}{\sigma_t \sqrt{s}}\right) [W(1 - s) - W(1 - y_k/\alpha)] ds \quad (1.29)$$

if $x \in [y_{k+1}, y_{k+2}]$.

In the following, all constants C, C_1, C_2, \dots depend only on δ and α .

Lemma 14. \hat{u}_t satisfies the following properties:

(i) \hat{u}_t is smooth everywhere except possibly at the points y_k .

(ii) $\{\hat{u}_t(x) : x \in I_k\}$ is independent of $\{\hat{u}_t(x) : x \in I_{k+3}\}$ for all k .

(iii) For $x \in I_{k+1}$,

$$\left| \hat{u}_t(x) - u_t(x) + W\left(1 - \frac{y_k}{\alpha}\right) \right| \leq Ct^{-M/2} \|W\|$$

and

$$\left| u'_t(x) - \hat{u}'_t(x) \right| \leq \frac{C}{\sqrt{M}} \left(\frac{x}{t}\right)^{\frac{M}{4} - \frac{1}{2}} \|W\|.$$

Proof: Properties (i) and (ii) follow from the definition of \hat{u}_t .

To prove property (iii) notice that, for $x \in I_{k+1}$,

$$\begin{aligned} u_t(x) - \hat{u}_t(x) &= \int_{[y_k/\alpha, y_{k+3}/\alpha]^c \cap [0,1]} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \phi\left(\frac{x - \alpha s}{\sigma_t \sqrt{s}}\right) W(1-s) ds \\ &\quad + W(1 - y_k/\alpha) - W(1 - y_k/\alpha) \int_{[y_k/\alpha, y_{k+3}/\alpha]^c \cap [0,\infty]} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \phi\left(\frac{x - \alpha s}{\sigma_t \sqrt{s}}\right) ds. \end{aligned}$$

From (??) and a similar equation for the derivatives of \hat{u}_t , we obtain

$$\begin{aligned} u'_t(x) - \hat{u}'_t(x) &= \int_{[y_k/\alpha, y_{k+3}/\alpha]^c \cap [0,1]} \partial_s K_t^0(x, s) [W(1-s) - W(1 - y_k/\alpha)] ds \\ &\quad - W(1 - y_k/\alpha) \int_1^\infty \partial_s K_t^0(x, s) ds, \end{aligned}$$

where K_t^n is defined as in (??). If $x \in I_{k+1}$, then $\frac{x}{\alpha} - L_t^x(M) \geq \frac{y_k}{\alpha}$ and $\frac{x}{\alpha} + L_t^x(M) \leq \frac{y_{k+3}}{\alpha}$. So, the above differences yield part (iii). \square

Consequently, if \hat{l}_t is the length of the curve \hat{u}_t restricted to $[\delta, \alpha]$, then

$$P(|l_t - \hat{l}_t| \geq r) \leq C_1 \exp\left\{-C_2 r^2 t^{\frac{M}{2}-1}\right\} \quad (1.30)$$

and

$$E(\hat{l}_t) = \frac{4t^{1/4}}{3(\pi^3 \alpha \sigma^2)^{1/4}} \left(\alpha^{3/4} - \delta^{3/4}\right) + O\left(\frac{\sqrt{\log t}}{t^{1/4}}\right) + O\left(\frac{1}{t^{\frac{M}{4}-\frac{1}{2}}}\right). \quad (1.31)$$

So, to find the concentration of the length around the mean at time t , we look at the length of the curve \hat{u}_t . Let W_k be the Brownian motion defined on $I = \left[0, \frac{3}{\alpha} \sqrt{\frac{M \log t}{t}}\right]$ by

$$W_k(s) = W\left(1 - \frac{y_k}{\alpha} - s\right) - W\left(1 - \frac{y_k}{\alpha}\right).$$

For each k , the Brownian motions W_k and W_{k+3} so defined are clearly independent.

As length is an additive functional, we can find the length on subintervals I_k and add them together. Heuristically, we can see that this gives us concentration as the length of the curve on every third interval is independent of each other, and as these are summed up,

the errors get averaged out.

Now, we give the rigorous arguments. In the following, we fix the probability space $(\Omega, \mathcal{B}(\Omega), \mathcal{P})$, where $\Omega = C \left[0, \frac{3}{\alpha} \sqrt{\frac{M \log t}{t}} \right]$ denote the set of continuous complex valued functions on I equipped with the sup-norm metric d , and \mathcal{P} is the Wiener measure.

We need some concepts from *Concentration of Measure* Theory. See [5] for an excellent survey of techniques in this area. We give a very brief outline of the concepts we need.

Transportation Cost Inequalities and Concentration: Let (χ, d) be a complete separable metric space equipped with the Borel sigma algebra $\mathcal{B}(\chi)$. Consider the p -th Wasserstein distance between two probability measures P and Q on this space, defined as

$$\mathcal{W}_p(P, Q) = \inf_{\pi} [\mathbb{E}d(X, X')^p]^{1/p},$$

where the infimum is over all couplings π of a pair of random elements (X, X') with the marginal of X being P and that of X' being Q .

Now, fix a probability measure P . Suppose there is a constant $C > 0$ such that for all probability measures $Q \ll P$, we have

$$\mathcal{W}_p(P, Q) \leq \sqrt{2CH(Q | P)},$$

where H refers to the relative entropy $H(Q | P) = \mathbb{E}^Q \log(dQ/dP)$. Then we say that P satisfies the L^p *Transportation Cost Inequality*. In short, we write $P \in \mathcal{T}_p(C)$.

Now, we present one of the key results which connects Transportation Cost Inequalities and Concentration of Measures.

Lemma 15. *Suppose P is a probability measure on $(\chi, \mathcal{B}(\chi))$. Suppose further that each P is in $\mathcal{T}_1(C)$. Then, for any 1-Lipschitz map $F : \chi \rightarrow \mathbb{R}$ and any $r > 0$,*

$$P \left(\left| F - \int F dP \right| > r \right) \leq \exp \left\{ -\frac{r^2}{2C} \right\}.$$

It is easy to see that $\mathcal{T}_2(C)$ implies $\mathcal{T}_1(C)$. But the main advantage in dealing with $\mathcal{T}_2(C)$ comes from its *tensorization property* described in the following lemma.

Lemma 16. Suppose $P_i, i = 1, 2, \dots, n$ are probability measures on $(\chi, \mathcal{B}(\chi))$. Suppose further that each P_i is in $\mathcal{T}_2(C)$. On χ^n , define the distance between $x^n = (x_1, x_2, \dots)$ and $y^n = (y_1, y_2, \dots)$ by

$$d^n(x^n, y^n) = \sqrt{\sum_{i=1}^n d^2(x_i, y_i)}.$$

Then $\bigotimes_{i=1}^n P_i \in \mathcal{T}_2(C)$ on (χ^n, d^n) .

The following lemma, which follows from the developments in [6], is of key importance to us.

Lemma 17. The Wiener measure on $C[0, T]$ satisfies the transportation inequality $\mathcal{T}_2(T)$ with respect to the sup-norm metric.

These tools are all we need to establish a concentration result for l_t .

Let us define the function $T_t^k : \Omega \rightarrow \mathbb{R}$ as follows:

$$T_t^k(f) = \int_{y_{k+1}}^{y_{k+2}} \left| \int_0^{\frac{y_{k+3} - y_k}{\alpha}} \partial_s K_t^0 \left(x, s + \frac{y_k}{\alpha} \right) f(s) ds \right| dx.$$

Notice that $\hat{l}_t = \sum_{k=0}^{\frac{\sqrt{t}}{\sqrt{M \log t}}} T_t^k(W_k)$. Suppose we prove that T_t^k is Lipschitz with respect to d with Lipschitz constant C_t . Then, with $N = \left\lceil \frac{\alpha - \delta}{3} \sqrt{\frac{t}{M \log t}} \right\rceil$, the functions $\{T_t^{(i)} : \Omega^N \rightarrow \mathbb{R} : i = 0, 1, 2\}$ defined by

$$T_t^{(i)}(f) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} T_t^{3k+i}(f_{k+1})$$

where $f = (f_1, \dots, f_N)$, is also Lipschitz with respect to d^N with the same constant C_t .

Lemma 18. For each k , T_t^k is Lipschitz on (Ω, d) with Lipschitz constant $C \sqrt{M \log t}$ where C is a constant depending only on δ and α .

Proof: For f_1, f_2 in Ω ,

$$\begin{aligned} |T_t^k(f_1) - T_t^k(f_2)| &\leq \int_{y_{k+1}}^{y_{k+2}} \int_0^{\frac{y_{k+3} - y_k}{\alpha}} \left| \partial_s K_t^0 \left(x, s + \frac{y_k}{\alpha} \right) \right| |f_1(s) - f_2(s)| ds dx \\ &\leq \|f_1 - f_2\| \int_{y_{k+1}}^{y_{k+2}} \int_0^{\frac{y_{k+3} - y_k}{\alpha}} \left| \partial_s K_t^0 \left(x, s + \frac{y_k}{\alpha} \right) \right| ds dx. \end{aligned}$$

By the estimates obtained in the proof of Lemma 3,

$$\int_{y_{k+1}}^{y_{k+2}} \int_0^{\frac{y_{k+3}-y_k}{\alpha}} \left| \partial_s K_t^0 \left(x, s + \frac{y_k}{\alpha} \right) \right| ds dx \leq C\sqrt{t} \int_{y_{k+1}}^{y_{k+2}} dx \leq C\sqrt{M \log t}.$$

This proves the lemma. \square

Now, for any $f \in \Omega^{3N}$, define the following functions in Ω^N : $f^{(1)} = (f_1, f_4, \dots, f_{3N-2})$, $f^{(2)} = (f_2, f_5, \dots, f_{3N-1})$ and $f^{(3)} = (f_3, f_6, \dots, f_{3N})$. Notice that

$$\hat{l}_t = \sqrt{N} \left(T_t^{(1)}(\tilde{W}^{(1)}) + T_t^{(2)}(\tilde{W}^{(2)}) + T_t^{(3)}(\tilde{W}^{(3)}) \right)$$

where $\tilde{W} = (W_0, W_1, \dots, W_{3N-1}) \in \Omega^{3N}$. Using this fact and Lemmas 16 and 15, we get for any $r > 0$,

$$P(|\hat{l}_t - E\hat{l}_t| \geq r\sqrt{M \log t}) \leq C_1 \exp \{-C_2 r^2\}$$

This, along with (1.30) and (1.31) gives us our main conclusion:

Theorem 4. *We have*

$$P(|l_t - El_t| \geq r\sqrt{M \log t}) \leq C_1 \exp \{-C_2 r^2\},$$

where C_1, C_2 are constants depending only on δ and α .

4.2 How close is the scaled Conga Line to Brownian motion?

Though the unscaled Conga line seen far away from the tip ‘smooths out’ Brownian motion more and more with increasing t , we see that in the simulations of the scaled Conga line, *making t larger actually makes the curve rougher* and resemble Brownian motion more and more. Closer analysis reveals that this in fact results from the *scaling*. Again, before we supply the rigorous arguments, we give a heuristic reasoning. Looking at equation (1.7), we see that although the scaling takes the Brownian motion W^t on $[0, t]$ to a Brownian motion W^1 on $[0, 1]$, the width of the window on which the smoothing takes place in the unscaled Conga line, which is comparable to \sqrt{t} , is taken to $O(t^{-1/2})$ in the scaled version, which shrinks with time t .

In the following, we consider the sequence of random curves $u_t(\cdot)$ indexed by t , and $L_t^x = \alpha^{-1} \sqrt{-M\sigma_t^2 x \log \sigma_t^2 x}$.

Theorem 5. *There exists a deterministic constant C_1 such that, almost surely, there is $T = T(\omega) > 0$ for which*

$$\left| u_t(x) - W \left(1 - \frac{x}{\alpha} \right) \right| \leq C_1 \sqrt{- \left(\frac{x}{t} \right)^{1/2} \log \left(\frac{x}{t} \right)} \quad (1.32)$$

for all $x \in (0, \alpha]$ satisfying $x > \alpha L_t^x$ for all $t \geq T$. In particular, for any fixed $\beta < 1$, the above holds almost surely for $x \in [t^{-\beta}, \alpha]$ for all $t \geq T$.

Thus, the scaled Conga line is close to Brownian motion for large t although the unscaled one is not, as can be seen from the right side of equation (1.32). This subsection is devoted to proving the above theorem.

For any continuous function $f : [0, 1] \rightarrow \mathbb{C}$, define

$$P_t f(x) = \int_0^1 \frac{x + \alpha s}{2\sigma_t s^{3/2}} \phi \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) f(1 - s) ds.$$

Note that the Conga line is given by $u_t(x) = P_t W(x)$. $P_t f$ can be thought of as a *smoothing kernel* acting on the function $x \mapsto f(1 - x/\alpha)$. The following lemma shows that if f is Lipschitz, then for large t , $P_t f(x)$ is close to $f(1 - x/\alpha)$.

Lemma 19. *If f is Lipschitz with constant C , then for large enough t and for $x \in (0, \alpha]$ satisfying $x > \alpha L_t^x$,*

$$\left| P_t f(x) - f \left(1 - \frac{x}{\alpha} \right) \right| \leq C \sigma_t \sqrt{x}. \quad (1.33)$$

Note that

$$\begin{aligned} \left| P_t f(x) - f \left(1 - \frac{x}{\alpha} \right) \right| &\leq C \alpha^{-1} \int_0^1 \frac{x + \alpha s}{2\sigma_t s^{3/2}} \phi \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) |x - \alpha s| ds \\ &= I_t^x + J_t^x + S_t^x \end{aligned}$$

where

$$\begin{aligned} I_t^x &= C \alpha^{-1} \int_0^{(x/\alpha) - L_t^x} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \phi \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) |x - \alpha s| ds \leq C \alpha^{-1} \int_0^{(x/\alpha) - L_t^x} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \phi \left(\frac{x - \alpha s}{\sigma_t \sqrt{s}} \right) ds \\ &\leq C \alpha^{-1} \overline{\Phi}(\sqrt{-M \log \sigma_t^2 x}) \leq C \alpha^{-1} (\sigma_t \sqrt{x})^M \end{aligned}$$

and

$$\begin{aligned}
J_t^x &= C\alpha^{-1} \int_{(x/\alpha)-L_t^x}^{\min((x/\alpha)+L_t^x, 1)} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \phi\left(\frac{x - \alpha s}{\sigma_t \sqrt{s}}\right) |x - \alpha s| ds \\
&\leq C\alpha^{-1} \sigma_t \sqrt{\frac{2x}{\alpha}} \int_{(x/\alpha)-L_t^x}^{\min((x/\alpha)+L_t^x, 1)} \frac{x + \alpha s}{2\sigma_t s^{3/2}} \phi\left(\frac{x - \alpha s}{\sigma_t \sqrt{s}}\right) \left|\frac{x - \alpha s}{\sigma_t \sqrt{s}}\right| ds \\
&\leq C\alpha^{-1} \sigma_t \sqrt{\frac{2x}{\alpha}} \int_{-\infty}^{\infty} |s| \phi(s) ds.
\end{aligned}$$

Similarly as I_t^x , S_t^x is small compared to J_t^x . \square

Now, Brownian motion is not Lipschitz, but it can be uniformly approximated on $[0, 1]$ by piecewise linear random functions whose Lipschitz constants can be controlled using *Levy's Construction of Brownian motion* which we now briefly describe following [2]. Define the n -th level dyadic partition $\mathcal{D}_n = \{\frac{k}{2^n} : 0 \leq k \leq 2^n\}$ and let $\mathcal{D} = \cup_{n=0}^{\infty} \mathcal{D}_n$. Let $\{Z_n : n \in \mathbb{N}\}$ be *i.i.d* standard normal random variables. Define the random piecewise linear functions F_n as follows:

$$F_n(x) = \begin{cases} 2^{-\frac{n+1}{2}} Z_x & x \in \mathcal{D}_n \setminus \mathcal{D}_{n-1} \\ 0 & x = 0 \\ \text{linear} & \text{in between} \end{cases}$$

With this, Levy's construction says that a Brownian motion W can be constructed via

$$W(x) = \sum_{n=0}^{\infty} F_n(x).$$

for $x \in [0, 1]$.

Let $W_N(x) = \sum_{n=0}^N F_n(x)$. This function serves as the piecewise linear (hence Lipschitz) approximation to W . From Lemma 19, for any N ,

$$\left| P_t W(x) - W\left(1 - \frac{x}{\alpha}\right) \right| \leq C \sum_{n=0}^N \sigma \sqrt{\frac{x}{t}} \|F_n'\|_{\infty} + 2 \sum_{n=N}^{\infty} \|F_n\|_{\infty}.$$

Fix $c > \sqrt{2 \log 2}$. Let $N^* = \inf\{n : |Z_d| \leq c\sqrt{n} \forall d \in \mathcal{D} \setminus \mathcal{D}_n\}$.

$$\sum_{n=0}^{\infty} P(\text{there exists } d \in \mathcal{D}_n \text{ with } |Z_d| \leq c\sqrt{n}) \leq \sum_{n=0}^{\infty} (2^n + 1) \exp\left(\frac{-c^2 n}{2}\right) < \infty.$$

So, by Borel-Cantelli Lemma, $P(N^* < \infty) = 1$.

Now, for $n > N^*$, $\|F_n\|_\infty \leq c\sqrt{n}2^{-n/2}$ and $\|F'_n\|_\infty \leq \frac{2\|F_n\|_\infty}{2^{-n}} \leq 2c\sqrt{n}2^{n/2}$. So, for $l > N^*$, we get

$$\begin{aligned} \left| P_t W(x) - W\left(1 - \frac{x}{\alpha}\right) \right| &\leq C \sum_{n=0}^{N^*} \sigma \sqrt{\frac{x}{t}} \|F'_n\|_\infty + 2C \sum_{n=N^*}^l \sigma \sqrt{\frac{x}{t}} c\sqrt{n}2^{n/2} \\ &\quad + 2 \sum_{n=l}^{\infty} c\sqrt{n}2^{-n/2}. \end{aligned} \quad (1.34)$$

Now, take t large enough that, for every $x \in (0, \alpha]$, the first term is less than $\sqrt{-\left(\frac{x}{t}\right)^{1/2} \log\left(\frac{x}{t}\right)}$ and $\sqrt{\frac{x}{t}} \in (2^{-l}, 2^{-l+1}]$ for some $l > N^*$. Plugging this l into equation (1.34) and using the fact that the last sum above is dominated by its leading term, we get the result of the theorem.

4.3 Analyzing number of loops

A **loop** L in a continuous curve $f : \mathbb{R} \rightarrow \mathbb{C}$ is defined as a restriction of the form $f|_{[a,b]}$ where $f(a) = f(b)$ and f is injective on $[a, b)$. Note that L divides the plane into a bounded component and an unbounded component. Define the size of the loop

$$s(L) = \max\{R > 0 : \exists x \in \text{the bounded component } \mathcal{B} \text{ of } L \text{ such that } B(x, R) \subseteq \mathcal{B}\}.$$

It can be shown (the quick way is to look at the expectation meta-theorem from [1]) that if f is a continuously differentiable Gaussian process, then with probability one, it has no *singularities* (points where the first derivative of both $\operatorname{Re} f$ and $\operatorname{Im} f$ vanish). Using this fact, it is easy to see that if I is a compact interval on which f is not injective, then $f|_I$ has at least one loop L of positive size.

As the number of loops is bounded above by the number of critical points of $\operatorname{Re} f$ (equivalently $\operatorname{Im} f$), we see that by Lemma 11, for a large fixed t , the number of loops in the Conga line is bounded above by $C\sqrt{t}$ with very high probability. This section is dedicated to achieving a lower bound. The simulation (Figure 1.2) shows a number of loops, most of

them being small. In the following, we obtain a lower bound for the number of *small loops*, which differs from the upper bound by a logarithmic factor. For this, our main ingredient is the *Support Theorem for Brownian Motion* which says the following:

Theorem 6. *If $f : [0, 1] \rightarrow \mathbb{C}$ is continuous and W is a complex Brownian motion on $[0, 1]$, then for any $\epsilon > 0$,*

$$P(\|W - f\| < \epsilon) > 0,$$

where $\|g\| = \sup_{x \in [0, 1]} |g(x)|$.

We will not prove the theorem, but this can be proved either by approximating f by piecewise linear functions and using Levy's construction of Brownian motion, or by an application of the *Girsanov Theorem* (see [7]).

We also need to exploit the exponentially decaying correlation between $u_t(x)$ and $u_t(x')$ as $|x - x'|$ increases (see Lemma 4) by bringing into play the approximation of u_t by the process \hat{u}_t introduced in Subsection 4.1.

Now, we state the main theorem of this section.

Theorem 7. *Choose $R > 6C_1$, where C_1 is the constant in Theorem 5. Let N_t^l be the number of loops of size less than or equal to $2R \left(\frac{\log t}{t}\right)^{1/4}$ in the (scaled) Conga line u_t in $[\delta, \alpha]$ at time t . Then there exist constants C and C' such that*

$$P\left(C\sqrt{\frac{t}{\log t}} \leq N_t^l \leq C'\sqrt{t}\right) \rightarrow 1$$

as $t \rightarrow \infty$.

Proof: The upper bound follows from Lemma 11.

Proving the lower bound is more involved.

Our strategy is to choose a function f which has a loop and run the Brownian motion

W in a narrow tube around f , which, by Theorem 6, we can do with positive probability. Now by Theorem 5, we know that for large t , u_t is ‘close’ to the Brownian motion W with very high probability, and thus the curve u_t is forced to run in a narrow sausage around f thereby inducing a loop.

Such a function is $f(x) = C((4x - 2)^3 - (4x - 2), 1 - (4x - 2)^2)$ for $x \in [0, 1]$, where C is a suitably chosen constant to make the size of the loop in f to be R . Let us denote the continuous functions restricted to the ϵ -sausage around $f|_{[a,b]}$ as

$$S(f; \epsilon, [a, b]) = \{g \in C[a, b] : \|f - g\| < \epsilon\}.$$

Fix α' such that $\delta < \alpha' < \alpha$. For $x \in [\delta, \alpha']$, define

$$f_x^{(t)}(s) = \left(\frac{M \log t}{t}\right)^{1/4} f\left(\frac{1}{\alpha} \sqrt{\frac{t}{M \log t}}(s - x)\right), \quad x \leq s \leq x + \alpha \sqrt{\frac{M \log t}{t}}.$$

Then for any continuously differentiable complex-valued Gaussian process g , defined on a subset of $[0, 1]$ containing $\left[x, x + \alpha \sqrt{\frac{M \log t}{t}}\right]$, and any complex number c , $g \in S\left(c + f_x^{(t)}; \frac{R}{2} \left(\frac{M \log t}{t}\right)^{1/4}, \left[x, x + \alpha \sqrt{\frac{M \log t}{t}}\right]\right)$ implies that g has a self-intersection on $\left[x, x + \alpha \sqrt{\frac{M \log t}{t}}\right]$ and thus, due to absence of singularities with probability one, g has a loop of positive size on this interval.

We break up the proof into parts:

- (i) In Lemma 20, we prove that the probability of $u_t|_{\left[x, x + \alpha \sqrt{\frac{M \log t}{t}}\right]}$ having a loop of size comparable to $\left(\frac{\log t}{t}\right)^{1/4}$ is bounded below uniformly for all $x \in [\delta, \alpha']$ by a fixed positive constant p independent of x and t .
- (ii) We use part (iii) of Lemma 14 and Lemma 20 to deduce that the probability of \hat{u}_t having a loop of size comparable to $\left(\frac{\log t}{t}\right)^{1/4}$ on each interval I_{k+1} is bounded below by $p/2$.

(iii) We use the independence of $\hat{u}|_{I_k}$ and $\hat{u}|_{I_{k+3}}$ for every k to deduce in Lemma 21 that the total number of such loops in \hat{u}_t is bounded below by $\frac{p}{4} \left[(\alpha' - \delta) \sqrt{\frac{t}{M \log t}} \right]$ with very high probability.

(iv) We finally use part (iii) of Lemma 14 again to translate the result of Lemma 21 to the original process u_t in Lemma 22.

Lemma 20. *There is a constant $p > 0$ independent of x and t such that*

$$P \left(u_t \Big|_{\left[x, x + \alpha \sqrt{\frac{M \log t}{t}} \right]} \in S \left(W \left(1 - \frac{x}{\alpha} \right) + f_x^{(t)}; \frac{R}{2} \left(\frac{M \log t}{t} \right)^{1/4}, \left[x, x + \alpha \sqrt{\frac{M \log t}{t}} \right] \right) \right) \geq p > 0$$

for all $x \in [\delta, \alpha']$, for all sufficiently large t .

Proof: Choose and fix any $x \in [\delta, \alpha']$. By Theorem 5 and by the translation and scaling invariance of Brownian motion, we get for $R > 6C_1$ (here C_1 is the constant in Theorem 5) and large t ,

$$\begin{aligned} & P \left(u_t \Big|_{\left[x, x + \alpha \sqrt{\frac{M \log t}{t}} \right]} \in S \left(W \left(1 - \frac{x}{\alpha} \right) + f_x^{(t)}; \frac{R}{2} \left(\frac{M \log t}{t} \right)^{1/4}, \left[x, x + \alpha \sqrt{\frac{M \log t}{t}} \right] \right) \right) \\ & \geq P \left(\sup_{s \in \left[x, x + \alpha \sqrt{\frac{M \log t}{t}} \right]} \left| u_t(s) - W \left(1 - \frac{s}{\alpha} \right) \right| \leq C_1 \left(\frac{M \log t}{t} \right)^{1/4} \text{ and} \right. \\ & \quad \left. \sup_{s \in \left[x, x + \alpha \sqrt{\frac{M \log t}{t}} \right]} \left| W \left(1 - \frac{s}{\alpha} \right) - W \left(1 - \frac{x}{\alpha} \right) - f_x^{(t)}(s) \right| \leq \frac{R}{3} \left(\frac{M \log t}{t} \right)^{1/4} \right) \\ & \geq P \left(\sup_{s \in \left[x, x + \alpha \sqrt{\frac{M \log t}{t}} \right]} \left| W \left(\frac{s-x}{\alpha} \right) - f_x^{(t)}(s) \right| \leq \frac{R}{3} \left(\frac{M \log t}{t} \right)^{1/4} \right) \\ & \quad - P \left(\sup_{s \in \left[x, x + \alpha \sqrt{\frac{M \log t}{t}} \right]} \left| u_t(s) - W \left(1 - \frac{s}{\alpha} \right) \right| > C_1 \left(\frac{M \log t}{t} \right)^{1/4} \right) \\ & = P \left(\sup_{s \in \left[0, \alpha \sqrt{\frac{M \log t}{t}} \right]} \left| W_0^{(t)}(s) - f_0^{(t)}(s) \right| \leq \frac{R}{3} \left(\frac{M \log t}{t} \right)^{1/4} \right) \end{aligned}$$

$$\begin{aligned}
& - P \left(\sup_{s \in \left[x, x + \alpha \sqrt{\frac{M \log t}{t}} \right]} \left| u_t(s) - W \left(1 - \frac{s}{\alpha} \right) \right| > C_1 \left(\frac{M \log t}{t} \right)^{1/4} \right) \\
\geq & P \left(\sup_{s \in [0,1]} |W(s) - f(s)| \leq \frac{R}{3} \right) \\
& - P \left(\sup_{s \in [\delta, \alpha]} \left| u_t(s) - W \left(1 - \frac{s}{\alpha} \right) \right| > C_1 \left(\frac{M \log t}{t} \right)^{1/4} \right) \\
\geq & p > 0.
\end{aligned}$$

Here we used Theorem 5 and Theorem 6 for the last step. By virtue of the second last step above, we can choose p independent of x and t , and the above lower bound works uniformly for all $x \in [\delta, \alpha']$. \square

Recall that by part (iii) of Lemma 14, we know that for $x \in I_{k+1}$,

$$\left| \hat{u}_t(x) - u_t(x) + W \left(1 - \frac{y_k}{\alpha} \right) \right| \leq C(\delta) t^{-M/2} \|W\|.$$

Define the event

$$A_k = \left\{ \hat{u}_t |_{I_{k+1}} \in S \left(W(1 - y_{k+1}/\alpha) - W(1 - y_k/\alpha) + f_{y_{k+1}}^{(t)}; \frac{R}{2} \left(\frac{M \log t}{t} \right)^{1/4}, I_{k+1} \right) \right\}.$$

If A_k holds, then \hat{u}_t has a loop in I_{k+1} . Write

$$S_t = \sum_{k=1}^{\lfloor (\alpha' - \delta) \sqrt{\frac{t}{M \log t}} \rfloor} \mathbb{I}_{A_k}.$$

Then the following holds.

Lemma 21.

$$P \left(S_t < \frac{p}{4} \left\lfloor (\alpha' - \delta) \sqrt{\frac{t}{M \log t}} \right\rfloor \right) \leq 3 \exp \left\{ -\frac{p^2}{8} \left\lfloor (\alpha' - \delta) \sqrt{\frac{t}{M \log t}} \right\rfloor \right\}.$$

Proof: By Lemma 20, it is easy to see that

$$\begin{aligned}
P(A_k) & \geq P \left\{ u_t |_{I_{k+1}} \in S \left(W(1 - y_{k+1}/\alpha) + f_{y_{k+1}}^{(t)}; \left(\frac{R}{2} - \epsilon \right) \left(\frac{M \log t}{t} \right)^{1/4}, I_{k+1} \right) \right. \\
& \quad \left. \text{and } \|W\| \leq \frac{\epsilon t^{M/2}}{C(\delta)} \left(\frac{M \log t}{t} \right)^{1/4} \right\} \\
& \geq \frac{p}{2} > 0
\end{aligned}$$

for large enough t and small enough ϵ . Thus we see that $ES_t \geq \frac{p}{2} \left[(\alpha' - \delta) \sqrt{\frac{t}{M \log t}} \right]$.

Now, as \hat{u}_t is independent on every third interval, so A_k is independent of A_{k+3} for every k . The result now follows by Bernstein's Inequality. \square

The above implies that with very high probability $S_t \geq \frac{p}{4} \left[(\alpha' - \delta) \sqrt{\frac{t}{M \log t}} \right]$.

Define the event

$$B_k = \{u_t \text{ has a loop on } I_{k+1}\}$$

and the corresponding sum

$$\tilde{N}_t^l = \sum_{k=1}^{\left\lfloor (\alpha' - \delta) \sqrt{\frac{t}{M \log t}} \right\rfloor} \mathbb{1}_{B_k}.$$

Our final lemma is the following.

Lemma 22.

$$P \left(N_t^l \geq \frac{p}{4} \left[(\alpha' - \delta) \sqrt{\frac{t}{M \log t}} \right] \right) \rightarrow 1$$

as $t \rightarrow \infty$.

Proof: Note that $N_t^l \geq \tilde{N}_t^l$.

By part (iii) of Lemma 14, we note that for small enough $\epsilon > 0$, the events A_k and $\left\{ \|W\| \leq \frac{\epsilon t^{M/2}}{C(\delta)} \left(\frac{M \log t}{t} \right)^{1/4} \right\}$ imply that B_k holds. We see that, for large t ,

$$\begin{aligned} P \left(\tilde{N}_t^l \geq \frac{p}{4} \left[(\alpha' - \delta) \sqrt{\frac{t}{M \log t}} \right] \right) &\geq P \left(S_t \geq \frac{p}{4} \left[(\alpha' - \delta) \sqrt{\frac{t}{M \log t}} \right] \right. \\ &\quad \left. \text{and } \|W\| \leq \frac{\epsilon t^{M/2}}{C(\delta)} \left(\frac{M \log t}{t} \right)^{1/4} \right) \\ &\geq 1 - P \left(S_t < \frac{p}{4} \left[(\alpha' - \delta) \sqrt{\frac{t}{M \log t}} \right] \right) \\ &\quad - P \left(\|W\| > \frac{\epsilon t^{M/2}}{C(\delta)} \left(\frac{M \log t}{t} \right)^{1/4} \right), \end{aligned}$$

which goes to one as $t \rightarrow \infty$ by Lemma 21. \square

The proof of the lower bound in Theorem 7 follows from the above lemmas. \square

5 Loops and singularities in particle paths

We start off this section by describing an interesting phenomenon that one notices in simulations of the paths of the individual particles in the discrete Conga line. The leading particle ($k = 1$) performs an erratic Gaussian random walk. But as k increases, the successive particles are seen to *cut corners* in the paths of the preceding particles making them smoother. This can be heuristically explained by the fact that a particle following another one in front directs itself along the shortest path between itself and the preceding particle (see equation (1.1)), and hence cuts corners. This phenomenon is captured by the process \bar{u} described in Subsection 2.2. So, we use the approximation of the discrete Conga line X by the smooth process \bar{u} in this section.

Recall that a **singularity** of a curve $\gamma : \mathbb{R} \rightarrow \mathbb{C}$ is a point t_0 at which its speed vanishes, i.e. $|\gamma'(t_0)| = 0$. A singularity t_0 of an *analytic* curve γ is called a **cusp singularity** if there exists a translation and rotation of co-ordinates taking $\gamma(t_0)$ to the origin, under which, γ has the representation $\gamma^* = (\gamma_1^*, \gamma_2^*)$ with the following power series expansions:

$$\gamma_1^*(t) = \sum_{i=2}^{\infty} a_i t^i$$

$$\gamma_2^*(t) = \sum_{i=3}^{\infty} b_i t^i$$

with $a_2 \neq 0$, $b_3 \neq 0$, for t in a neighborhood $[t_0 - \delta, t_0 + \delta]$ around t_0 for some $\delta > 0$.

Intuitively, this means that the graph of γ locally around t_0 looks like $y = x^{2/3}$ under a rigid motion of co-ordinates taking $\gamma(t_0)$ to the origin.

Making a change of variables $p = t - \frac{x}{\alpha}$ and $\tau = \rho^2 x$, we can rewrite \bar{u} as

$$f(p, \tau) = \bar{u}(x, t) = E_Z W(p - Z_\tau)$$

We restrict our attention to $p > 0, \tau > 0$. Fixing τ and varying p in the above expression for f yields the path of the particle at distance $x = \tau/\rho^2$ from the tip. Note that this is precisely the solution to the *heat equation* with the initial function being the Brownian motion W , the space variable represented by p and time by τ .

Another interesting observation, which was described briefly in the Introduction, is the *evolution of loops* as we look at the paths of successive particles. If a particle in the (two dimensional) Conga line goes through a loop, the particle following it, which cuts corners and tries to “catch it”, will go through a smaller loop. This is suggested by the simulations, where small loops are seen to die, and just before death, they look somewhat ‘elongated’, and the death site looks like a *cuspl singularity*. Other loops are seen to break after some time, that is, their end points come apart. Figure 1.2, though not depicting the particle paths, gives an idea of the loops in various stages of evolution. In this section, we investigate evolving loops in the paths of successive particles, especially the relationship between dying loops and formation of singularities.

Before we can start off, we give some definitions that will be useful in describing the evolution of loops.

We define a metric space (\mathcal{M}, d) , with a metric similar to the Skorohod metric on RCLL paths, on which we want to study loop evolutions:

$$\mathcal{M} = \{f : [a, b] \rightarrow \mathbb{C}; -\infty < a < b < \infty\}$$

If $f : [a_1, b_1] \rightarrow \mathbb{C}, g : [a_2, b_2] \rightarrow \mathbb{C} \in \mathcal{M}$, define

$$d(f, g) = \inf \{ \|\lambda_1 - \lambda_2\| + \|f \circ \lambda_1 - g \circ \lambda_2\| \mid \lambda_i : [0, 1] \rightarrow [a_i, b_i] \text{ is a homeomorphism} \}$$

where $\|\cdot\|$ denotes the sup-norm metric. It can be easily checked that (\mathcal{M}, d) is a metric space.

Define the **evolution of a loop** L as a continuous function $L : [0, T) \rightarrow \mathcal{M}$ such that $L_0 = L$ and L_t is a loop for every $0 \leq t < T$. If $f : \mathbb{R} \rightarrow \mathbb{C}$ is a space-time process, and $L_t = f(\cdot, t) \big|_{[a_t, b_t]}$ is a loop evolution, we say that L is a **loop evolution of f (starting from L)**. Say that a loop L of f **vanishes after time T** if T is the maximal time such that there exists a loop evolution $L : [0, T) \rightarrow \mathcal{M}$ of f starting from L .

Note: Although f has *no singularities for a fixed time τ* , it can be easily verified by an application of the expectation meta-theorem of [1] that the expected number of singularities of $f(\cdot, \tau)$ for (p, τ) lying in a compact set $K = [a, b] \times [c, d]$ is positive, and thus *singularities do occur with positive probability if we allow both space and time to vary*.

It is easy to see that if a loop dies at a site (p_0, τ_0) , then p_0 is a singularity for the curve $f(\cdot, \tau_0)$. We prove in Lemma 23 that with probability one, *any singularity looks is a cusp singularity*. In Lemma 24 we prove that for any (cusp) singularity in the Conga line at some time τ_0 , *there exists a loop in some small time interval $(\tau_0 - \delta, \tau_0)$* .

Lemma 23. *With probability one, any singularity p_0 of the (analytic) curve $f(\cdot, \tau_0)$ is a cusp singularity.*

Proof: Write $f = (f^1, f^2)$. It suffices to prove the lemma for (p_0, τ_0) lying in a rectangle $K = [a, b] \times [c, d]$. Our first step is showing the following:

$$P(\exists(p_0, \tau_0) \in K \text{ such that } \partial_p f(p_0, \tau_0) = 0 \text{ and the vectors } (\partial_p^2 f^1(p_0, \tau_0), \partial_p^3 f^1(p_0, \tau_0)) \text{ and } (\partial_p^2 f^2(p_0, \tau_0), \partial_p^3 f^2(p_0, \tau_0)) \text{ are linearly dependent}) = 0 \tag{1.35}$$

To show this, define A_n to be the event which holds when all the following are satisfied:

- (i) There exists $(p_0, \tau_0) \in K$ for which $\partial_p f(p_0, \tau_0) = 0$ and

$$(\partial_p^2 f^2(p_0, \tau_0), \partial_p^3 f^2(p_0, \tau_0)) = \lambda (\partial_p^2 f^1(p_0, \tau_0), \partial_p^3 f^1(p_0, \tau_0))$$

for some $\lambda \in [-n, n]$.

- (ii) The Lipschitz constants of the functions $\{\partial_p^i f^j(p, \tau) : (p, \tau) \in K; i = 1, 2, 3; j = 1, 2\}$ are less than or equal to n .

We will show that $P(A_n) = 0$ which will yield (1.35).

Partition the rectangle into a grid of sub-rectangles of side length $\leq \epsilon$, where ϵ is small. Call the set of grid points \hat{K} .

Now, suppose A_n holds. Let (p_0, τ_0) lie in a sub-rectangle R and let $(p_i, \tau_j) \in \hat{K}$ be a grid point adjacent to R . Note that as the Lipschitz constants of the above functions and λ are bounded by n , the following event holds:

$$A_n^{ij} = \left\{ |\partial_p f(p_i, \tau_j)| \leq \sqrt{2}n\epsilon \text{ and} \right.$$

$$\left. |(\partial_p^2 f^2(p_i, \tau_j), \partial_p^3 f^2(p_i, \tau_j)) - \lambda(\partial_p^2 f^1(p_i, \tau_j), \partial_p^3 f^1(p_i, \tau_j))| \leq 4n^2\epsilon \right.$$

for some $\lambda \in [-n, n]$.

Thus we have

$$A_n \subseteq \bigcup_{i,j} A_n^{ij}.$$

We show that there is a constant C depending on n such that $P(A_n^{ij}) \leq C\epsilon$.

To save us notation, call

$$X = (X_1, X_2, X_3) = (\partial_p f^1(p_i, \tau_j), \partial_p^2 f^1(p_i, \tau_j), \partial_p^3 f^1(p_i, \tau_j))$$

and similarly Y for f^2 . X and Y are independent and each follows a centred trivariate normal distribution. Let us call the density function of X p_{ij} and the distribution of Y as

Q_{ij} . Then, as X and Y have uniformly bounded densities,

$$\begin{aligned} P(A_n^{ij}) &\leq \int_{x \in [-\sqrt{2n\epsilon}, \sqrt{2n\epsilon}] \times \mathbb{R}^2} p_{ij}(x) \\ &\quad \times Q_{ij} \left(|Y_1| \leq \sqrt{2n\epsilon}; (Y_2, Y_3) \in \text{the } 4n^2\epsilon \text{ neighbourhood of the linear span of } (x_2, x_3) \right) \\ &\leq \int_{x \in [-\sqrt{2n\epsilon}, \sqrt{2n\epsilon}] \times \mathbb{R}^2} p_{ij}(x) C'_{ij} \epsilon^2 dx \leq C_{ij} \epsilon^3, \end{aligned}$$

where C'_{ij}, C_{ij} depend on n . Note that the determinants of the covariance matrices of X and Y are continuous and do not vanish at any point on the compact set K . Thus we can bound C_{ij} by C (which depends on n) uniformly over i, j, ϵ . Using these facts, we get

$$P(A_n) \leq \sum_{i,j} P(A_n^{ij}) \leq C\epsilon.$$

As ϵ is arbitrary, we get $P(A_n) = 0$.

Now if p_0 is a singularity occurring at time τ_0 , i.e. $\exists (p_0, \tau_0) \in K$ for which $\partial_p f(p_0, \tau_0) = 0$, we can apply a rigid motion of co-ordinates such that $f(p_0, \tau_0)$ is the new origin and the rotation angle θ is chosen to satisfy the equation

$$A_\theta \begin{bmatrix} X_2 & X_3 \\ Y_2 & Y_3 \end{bmatrix} = \begin{bmatrix} a_2 & a_3 \\ 0 & b_3 \end{bmatrix}$$

where A_θ is the rotation matrix corresponding to θ . By (1.35), we see that a_2 and b_3 are non-zero. Then, the result follows by taking this new co-ordinate frame. \square

Lemma 24. *If (p_0, τ_0) is a (cusp) singularity, then $\exists \delta > 0$ such that $f(\cdot, \tau)$ has a loop on some interval containing p_0 for all $\tau \in [\tau_0 - \delta, \tau_0)$.*

Proof: First we show that f is jointly analytic in (p, τ) . For this, note that

$$\partial_p^n f(p, \tau) = \int_0^\infty (-1)^n \tau^{-(n+1)/2} \text{He}_n \left(\frac{y-p}{\sqrt{\tau}} \right) \phi \left(\frac{y-p}{\sqrt{\tau}} \right) W(y) dy.$$

So, by using the fact that $\lim_{y \rightarrow \infty} \frac{W(y)}{y} = 0$ almost surely, we get that with probability one,

$$|\partial_p^n f(p, \tau)| \leq \frac{C}{\tau^{n/2}} \sqrt{n!} (p^2 + \tau) \quad (1.36)$$

for some random constant C . From this, we know that $f(\cdot, \tau_0)$ has an analytic representation in the space variable p as

$$f(p, \tau_0) = \sum_{i=0}^{\infty} a_i(p_0, \tau_0)(p - p_0)^i. \quad (1.37)$$

Now, we prove joint analyticity. Note that if we can prove

$$\sum_{i=0}^{\infty} \mathbb{E}_Z |a_i(p_0, \tau_0)| (|p - p_0| + |Z_\delta|)^i < \infty \quad (1.38)$$

for $p \in \mathbb{R}$ and $0 \leq \delta < \epsilon$ for some $\epsilon > 0$, then we can write

$$f(p, \tau) = \sum_{i=0}^{\infty} \mathbb{E}_Z a_i(p_0, \tau_0) ((p - p_0) - Z_{\tau - \tau_0})^i \quad (1.39)$$

for $p \in \mathbb{R}$ and $\tau \in [\tau_0, \tau_0 + \epsilon)$ and it follows that f is jointly analytic on $\mathbb{R} \times [\tau_0, \tau_0 + \epsilon)$. Joint analyticity on $\mathbb{R} \times \mathbb{R}^+$ is immediate as a result.

From the bound (1.36) on $\partial_p^n f(p, \tau)$, we see that (1.38) holds when

$$\sum_{i=0}^{\infty} \frac{C}{\tau_0^{i/2} \sqrt{i!}} (p_0^2 + \tau_0) 2^i (|p - p_0|^i + 2^{i/2} \sqrt{i!} \delta^{i/2}) < \infty$$

which is satisfied for $\delta < \frac{\tau_0}{8}$ and all $p \in \mathbb{R}$.

Now, if f has a cusp singularity at (p_0, τ_0) , then by the rigid motion of co-ordinates used in Lemma 23 and joint analyticity around (p_0, τ_0) , we can write $f = (f_1, f_2)$ in the new co-ordinate frame, where

$$\begin{aligned} f^1(p, \tau_0 - s) &= a_2(p^2 - s) + E^1(p, s) \\ f^2(p, \tau_0 - s) &= b_3(p^3 - 3ps) + E^2(p, s) \end{aligned}$$

for small enough s . It can be checked that

$$g_s(p) = (a_2(p^2 - s), b_3(p^3 - 3ps))$$

has a loop in $[-\sqrt{3s}, \sqrt{3s}]$. So, if we choose and fix $M > \sqrt{3}$, then we see that there is $\delta > 0$ such that for all $s < \delta$, and $p \in [-M\sqrt{s}, M\sqrt{s}]$,

$$|E^1(p, s)| \leq Cs^{3/2}$$

and

$$|E^2(p, s)| \leq Cs^2$$

for some random constant C . This, along with the fact that g_s has a loop on $[-M\sqrt{s}, M\sqrt{s}]$ forces f to have a loop on this interval. \square

From Lemma 23 and Lemma 24, it is clear that *there is a bijection between dying loops and singularities in the particle path evolution.*

Shape of a dying loop

To understand the limiting shape of the dying loop, we rescale the space co-ordinate p as $\sqrt{s}P$, and rescale f^1 by s , f^2 by $s^{3/2}$. It can be easily checked that as $s \rightarrow 0$, this new scaled function

$$\hat{f}_s(P) = \left(s^{-1} f^1(\sqrt{s}P), s^{-3/2} f^2(\sqrt{s}P) \right)$$

converges uniformly on $[-M, M]$ to the function

$$\hat{g}(P) = (a_2(P^2 - 1), b_3(P^3 - 3P))$$

which contains a loop $\hat{g}|_{[-\sqrt{3}, \sqrt{3}]}$. It is also not too hard to check that any loop in \hat{f}_s on $[-M, M]$ converges to the loop of \hat{g} in the topology on the space (\mathcal{M}, d) described above.

This gives us the limiting shape of a dying loop. The difference in the scaling exponent explains why *the loops look elongated before death.*

6 Freezing in the tail

This section addresses Observation 4 of Burdzy and Pal [8]. The *tail of the Conga line* refers to the particles at distance $x > \alpha t$ from the leading particle.

So far, we have studied the behaviour for $x \leq \alpha t$. This part is more dynamic in time and the particles in this region have variances going to infinity with time, indicating appreciable motion. Also, a particle at any *fixed distance from the tip* eventually steps into this regime.

The tail on the other hand seems to *freeze in time* and the angle at which the Conga line comes out of the origin shows very little change with time after a while (see Figure 1.2). Further, the very small variance of any particle in the tail region for large t indicates that we indeed need to *rescale* the tail to study its properties. In other words, the tail behaves in a very different manner compared to particles near the tip.

To study the phenomenon of ‘Freezing in the tail’, we use the continuous version \bar{u} described in Subsection 2.2. For any fixed $\eta > 0$, we choose a sequence of distances from the tip

$$x_t = \alpha(1 + \eta)t$$

and study

$$\bar{v}_t(\eta) = \bar{u}(x_t, t).$$

The distances x_t from the tip increasing with time t ensures that these particles remain in the tail region for all times. We rescale \bar{v}_t as

$$v_t(\eta) = \sqrt{2\pi\rho^2\alpha(1+\eta)t} \exp\left\{\frac{\eta^2 t}{2\rho^2\alpha(1+\eta)t}\right\} \bar{v}_t(\eta). \quad (1.40)$$

Also define

$$v(\eta) = \int_0^\infty \exp\left\{-\frac{s\eta}{\rho^2\alpha(1+\eta)}\right\} W(s) ds, \quad (1.41)$$

where W is the driving Brownian motion in expression (1.6).

Theorem 8. *For any fixed $\eta > 0$,*

$$v_t(\eta) \rightarrow v(\eta)$$

almost surely and in L^2 as $t \rightarrow \infty$.

Proof: In the following, C_1, C_2, \dots represent finite, positive constants.

It follows from (1.6) that

$$v_t(\eta) = \int_0^{(1-\eta)t} W(s) \exp\left\{-\frac{s^2}{2\rho^2 t \alpha(1+\eta)} - \frac{s\eta}{\rho^2\alpha(1+\eta)}\right\} ds.$$

Almost sure convergence follows from the fact that

$$\int_0^\infty |W(s)| \exp \left\{ -\frac{s\eta}{\rho^2\alpha(1+\eta)} \right\} ds < \infty$$

and the Dominated Convergence Theorem.

To prove L^2 convergence, note that

$$v(\eta) - v_t(\eta) = \int_0^{(1-\eta)t} g_t(s)W(s)ds + \int_{(1-\eta)t}^\infty \exp \left\{ -\frac{s\eta}{\rho^2\alpha(1+\eta)} \right\} W(s)ds, \quad (1.42)$$

where

$$g_t(s) = \exp \left\{ -\frac{s\eta}{\rho^2\alpha(1+\eta)} \right\} \left(1 - \exp \left\{ -\frac{s^2}{2\rho^2t\alpha(1+\eta)} \right\} \right).$$

It is easy to see that

$$g_t(s) \leq \frac{C_1}{\rho^2\alpha(1+\eta)t} s^2 \exp \left\{ -\frac{s\eta}{\rho^2\alpha(1+\eta)} \right\} I(s \leq \sqrt{t}) + C_2 \exp \left\{ -\frac{s\eta}{\rho^2\alpha(1+\eta)} \right\} I(s > \sqrt{t}). \quad (1.43)$$

Now, notice that the first term in (1.42) can be written as

$$E_1(t) = \int_0^{(1-\eta)t} \left(\int_a^{(1-\eta)t} g_t(s)ds \right) dW(a).$$

From (1.43), we get

$$\int_a^{(1-\eta)t} g_t(s)ds \leq \frac{C_3}{t} I(a < \sqrt{t}) + C_4 \exp \left\{ -\frac{\eta(\sqrt{t} \vee a)}{\rho^2\alpha(1+\eta)} \right\}.$$

Thus

$$\mathbb{E}(E_1(t))^2 = \int_0^{(1-\eta)t} \left(\int_a^{(1-\eta)t} g_t(s)ds \right)^2 da \leq \frac{C_5}{t^{3/2}}. \quad (1.44)$$

Similarly, for the second term in (1.42), say $E_2(t)$, we get

$$\mathbb{E}(E_2(t))^2 \leq C_6 t \exp \left\{ -\frac{2\eta(1-\eta)t}{\rho^2\alpha(1+\eta)} \right\}. \quad (1.45)$$

L^2 convergence follows from (1.44) and (1.45). \square

Acknowledgements

I thank my adviser Krzysztof Burdzy for guiding me through the project and giving interesting ideas to work on. I also thank Shirshendu Ganguly and Soumik Pal for helpful discussions. I thank Krzysztof Burdzy, Shirshendu Ganguly and Mary Solbrig for the figures, and Bharathwaj Palvannan and Yannick Van Huele for helping me with computer graphics issues.

7 Simulations

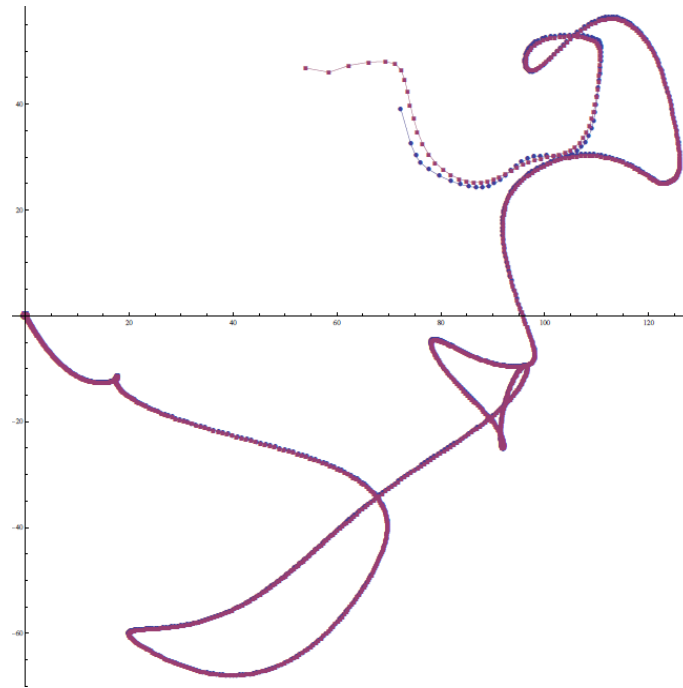


Figure 1.2: 10000 steps (red) and 10000-100 steps (blue) of the discrete two dimensional Conga line. (courtesy: Krzysztof Burdzy)

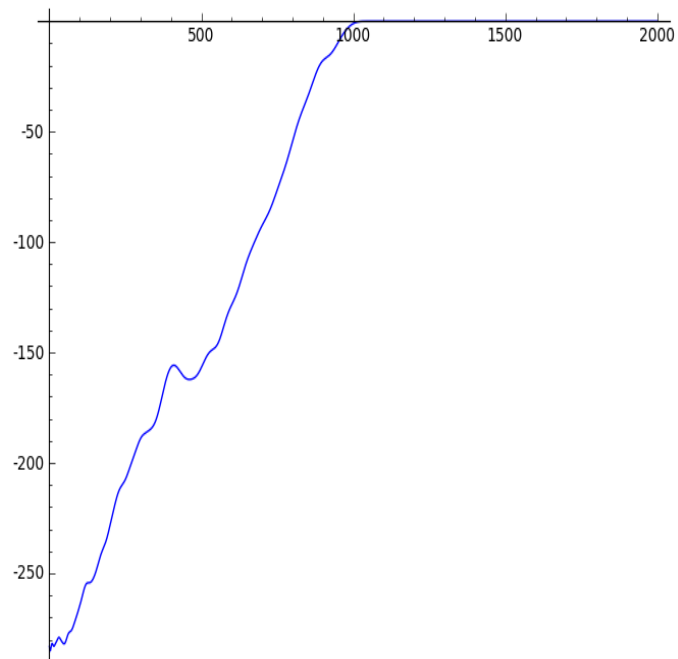


Figure 1.3: 2000 steps of the discrete one dimensional Conga line with $\alpha = 0.5$. (courtesy: Shirshendu Ganguly)

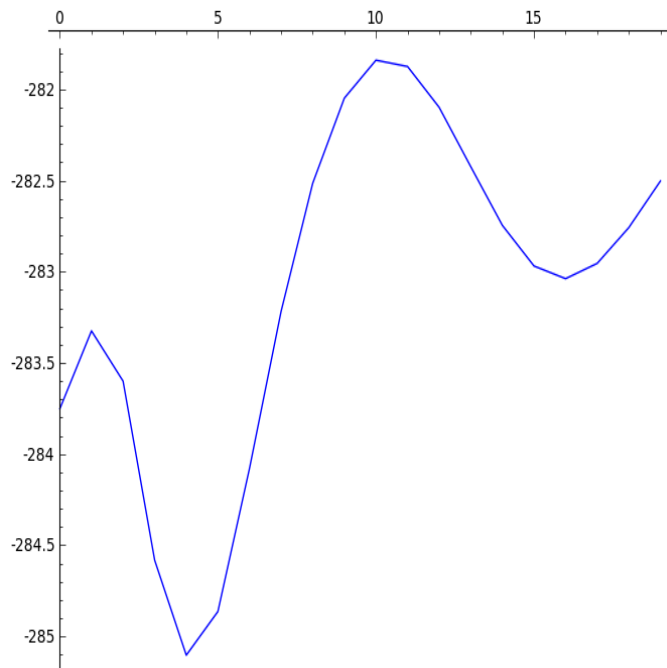


Figure 1.4: Near the tip of the Conga line. Showing the first twenty particles. (courtesy: Shirshendu Ganguly)

BIBLIOGRAPHY

- [1] Adler, Robert J., and Jonathan E. Taylor. Random fields and geometry. Vol. 115. Springer, 2007.
- [2] Mörters, Peter, and Yuval Peres. Brownian motion. Vol. 30. Cambridge University Press, 2010.
- [3] Feller, William. An introduction to probability theory and its applications. Vol. 2. John Wiley Sons, 2008.
- [4] Chou, Kai-Seng, and Xi-Ping Zhu. The curve shortening problem. CRC Press, 2010.
- [5] Ledoux, Michel. The concentration of measure phenomenon. Vol. 89. AMS Bookstore, 2005.
- [6] Pal, Soumik. "Concentration for multidimensional diffusions and their boundary local times." *Probability Theory and Related Fields* 154.1-2 (2012): 225-254.
- [7] Revuz, Daniel, and Marc Yor. Continuous martingales and Brownian motion. Vol. 293. Springer, 1999.
- [8] Burdzy, K. and Pal, S. Private communication.

Chapter 2

RANDOM MASS SPLITTING AND A QUENCHED INVARIANCE PRINCIPLE

(Joint work with Christopher Hoffman)

Imagine a one dimensional city (motivation from Abbott's *Flatland*) on the Y -axis with houses at the points $(0, k)$. Suppose the city is stricken with an epidemic and things are getting worse by the day. As the death toll rises, each house $(0, k)$ has only a Poisson(1) number of survivors $v(0, k)$ when the long awaited medical breakthrough suddenly happens. The surviving residents (assume at least one survives) of house $(0, 0)$ are scientists who were working on a cure for a while and they finally have an antidote! But the quantity that is produced is limited (say has mass 1 unit). So, they get out of their house carrying an equal proportion of the medicine $(1/v(0, 0))$, and start performing simple random walks, with time along the positive X -axis, to share it with the survivors. Meanwhile, the remaining survivors decide that staying inside the house in unhygienic conditions is dangerous, and they get out of their respective houses and start performing simple random walks at the same time as the scientists. Whenever a group of people meet on the way, say at (t, k) , and at least one person in the group has some medicine, it gets equally divided among all the people in the group. This process continues. The question we ask in this article is: What is the distribution of this mass for large times t ?

We call the above model the *Random Mass Splitting* model. In addition to the distribution of the medicine at time t we will study another property of the model. We use the movement of the villagers over time to define a random environment. Then we study the movement of a random walker in this random environment (RWRE). This defines a random

walk in a random environment (which will be precisely defined in the next section). We are interested in this random walk in a random environment because it turns out that the heat kernel of the walk conditioned on the environment ω at time t is precisely the distribution of the medicine at time t when the villagers are moving according to ω . Random walk in a random environment models have been studied by many authors. This can be a very difficult field as even the simplest properties such as transience and recurrence are difficult to establish [6]. But a theory of random walk in random environments has been developed. We will harness that theory to find the asymptotic mass distribution.

In many examples, such as simple random walk on supercritical percolation clusters, we get a quenched invariance principle (i.e. an almost sure convergence of the RWRE to Brownian motion) [2] [13]. However in other examples we find behaviors that are very different from the usual diffusive behavior of simple random walk on \mathbb{Z}^d [14] [1]. Many of the proofs of invariance principles for random walks in random environments have their origins in the seminal work of Kipnis and Varadhan [7]. This paper laid down the foundation for quenched invariance principles for reversible Markov chains. Maxwell and Woodroffe [8] and Derriennic and Lin [4] subsequently extended their approach to the non-reversible set-up. Rassoul-Agha and Seppäläinen [12] developed further on these techniques to give a set of conditions under which a QIP holds for random walks in space-time random environments. They verified these conditions to prove a QIP for *i.i.d. environments* (see [12]). In this article, we follow the approach of [12]. Verifying their conditions for a QIP presents significantly more complications in our situation than it does with their i.i.d. environments. That these complications can be overcome demonstrates the robustness of the approach in [12]

1 Description of the model

We now present a more rigorous mathematical description of our model. Consider i.i.d. Poisson(1) random variables $\{v(0, k) : k \in \mathbb{Z}\}$ representing particles at points $(0, k)$ of the Y -axis at time $t = 0$. Each particle starts performing a simple random walk $\{S^{(k,i)}(\cdot) : 1 \leq i \leq v(0, k), k \in \mathbb{Z}\}$, with $S^{(k,i)}(0) = k$ and time represented along the positive X -axis. This

gives a collection of random variables $\{e^+(t, y), e^-(t, y), v(t, y)\}$ where

$$e^+(t, y) = \#\{(k, i) : S^{(k, i)}(t) = y, S^{(k, i)}(t+1) = y+1\}$$

$$e^-(t, y) = \#\{(k, i) : S^{(k, i)}(t) = y, S^{(k, i)}(t+1) = y-1\}$$

$$v(t, y) = \#\{(k, i) : S^{(k, i)}(t) = y\} = e^+(t, y) + e^-(t, y)$$

Note that

$$v(t, y) = e^+(t-1, y-1) + e^-(t-1, y+1).$$

For technical convenience, we take the $S^{(k, i)}$ to be two-sided random walks and thus the above random variables are defined for all $(t, y) \in \mathbb{Z}^2$.

Let $p(t, y)$ denote the (random amount of) mass at (t, y) . Then, for each t ,

$$\sum_{y=-t}^t p(t, y) = 1$$

Thus, $p(0, 0) = 1$. At time 1, we have

$$p(1, 1) = \frac{e^+(0, 0)}{v(0, 0)} \quad \text{and} \quad p(1, -1) = \frac{e^-(0, 0)}{v(0, 0)}$$

Now, suppose $e^+(0, 0) > 0$. After the mass gets split equally among the particles at $(1, 1)$, each one has mass $\frac{e^+(0, 0)}{v(0, 0)v(1, 1)}$. A similar statement holds if $e^-(0, 0) > 0$. Thus at time 2, the mass distribution is

$$p(2, 2) = \frac{e^+(0, 0)e^+(1, 1)}{v(0, 0)v(1, 1)}$$

$$p(2, 0) = \frac{e^+(0, 0)e^-(1, 1)}{v(0, 0)v(1, 1)} + \frac{e^-(0, 0)e^+(1, -1)}{v(0, 0)v(1, -1)} \quad \text{and}$$

$$p(2, -2) = \frac{e^-(0, 0)e^-(1, -1)}{v(0, 0)v(1, -1)}.$$

Here we make a convention of taking any term in the above expression whose denominator (and hence numerator) vanishes to be equal to zero.

To write down the explicit distribution of mass at time t , we define the *environment* to be the collection of ‘random edge crossings’

$$\omega = \{(e^+(t, y), e^-(t, y)) : (t, y) \in \mathbb{Z}^2\} \in ((\mathbb{Z}_+ \cup \{0\})^2)^{\mathbb{Z}^2}$$

We call the space of environments $\Omega = ((\mathbb{Z}_+ \cup \{0\})^2)^{\mathbb{Z}^2} = \{(a(t, y), b(t, y)) \in (\mathbb{Z}_+ \cup \{0\})^2 : (t, y) \in \mathbb{Z}^2\}$. With these definitions we can inductively define the mass distribution with the equation

$$p(t+1, y; \omega) = p(t, y+1; \omega) \frac{e^-(t, y+1)}{v(t, y+1)} + p(t, y-1; \omega) \frac{e^+(t, y-1)}{v(t, y-1)} \quad (2.1)$$

The space Ω is equipped with the canonical product sigma field \mathcal{G} and the natural shift transformation $T_{(t', y')}$ such that

$$(a(t, y), b(t, y)) (T_{(t', y')} \omega) = (a(t+t', y+y'), b(t+t', y+y')) (\omega)$$

The random edge crossings induce a measure \mathbb{P} on (Ω, \mathcal{G}) and it is easy to see that \mathbb{P} is invariant under the shift $T_{(t', y')}$. It can be further proved that $(\Omega, \mathcal{G}, \{T_{(t', y')}\}_{(t', y') \in \mathbb{Z}^2}, \mathbb{P})$ is ergodic. We leave this verification to the reader. We shall denote the corresponding expectation by \mathbb{E} .

Now fix an ω . This generates a mass distribution at (t, y) which we write as $p(t, y; \omega)$. Let $\Gamma(0, y)$ denote the set of all paths $\gamma : \mathbb{Z} \rightarrow \mathbb{Z}$ with $\gamma(0) = 0$ and $\gamma(t) = y$. Then the mass at (t, y) is given by

$$p(t, y; \omega) = \sum_{\gamma \in \Gamma(0, y)} \prod_{i=0}^{t-1} \frac{e^{(\gamma(i+1) - \gamma(i))} (i, \gamma(i))}{v(i, \gamma(i))} 1_{(v(i, \gamma(i)) > 0)}$$

where a product in the above sum is interpreted as zero if one of the terms in the product vanishes.

Note that the mass remains 1 at $(0, 0)$ at all times when $v(0, 0) = 0$. To avoid this, we make a change of measure to $\tilde{\mathbb{P}}$ where

$$\tilde{\mathbb{P}}(A) = \mathbb{E}(v(0, 0) 1_A)$$

for $A \in \mathcal{G}$. $\tilde{\mathbb{P}}$ is a probability measure as $\mathbb{E}(v(0,0)) = 1$. This change of measure ensures $\tilde{\mathbb{P}}(v(0,0) > 0) = 1$. So from now on we only consider ω with $v(0,0) > 0$.

Given any ω and any (t, y) we can define a random walk in a random environment which is started at y at time t and has transition probabilities generated by ω . We write $X_n(t, y)$ for the random variable that represents the location of the random walk after n steps (at time $t + n$). Its distribution is determined by $\mathbb{P}_{(t,y)}^\omega$ which is defined as follows.

$$\mathbb{P}_{(t,y)}^\omega(X_0 = (t, y)) = 1.$$

The X_n are a time inhomogeneous Markov chain with transition probabilities

$$\mathbb{P}_{(t,y)}^\omega(X_{n+1} = (t' + 1, y' + 1) | X_n = (t', y')) = \frac{e^+(t', y')}{v(t', y')} \quad \text{and} \quad (2.2)$$

$$\mathbb{P}_{(t,y)}^\omega(X_{n+1} = (t' + 1, y' - 1) | X_n = (t', y')) = \frac{e^-(t', y')}{v(t', y')} \quad (2.3)$$

Now we present the key observation that connects the random walk in a random environment with the random mass splitting model.

Lemma 25. *For $t \in \mathbb{N}$ and $y \in \mathbb{Z}$*

$$p(t, y; \omega) = \mathbb{P}_{(0,0)}^\omega(X_t = (t, y)).$$

Proof. This follows by induction on t . Fix an environment ω . Then we have

$$p(0, 0; \omega) = \mathbb{P}_{(0,0)}^\omega(X_0 = (0, 0)) = 1$$

and

$$p(0, y; \omega) = \mathbb{P}_{(0,0)}^\omega(X_0 = (0, y)) = 0$$

for all $y \neq 0$. Now suppose the lemma is true for some t and all $y \in \mathbb{Z}$. The transition probabilities for X_{t+1} given in (2.2) and (2.3) and the mass splitting rule in (2.1) show that for an arbitrary y

$$\begin{aligned} \mathbb{P}_{(0,0)}^\omega(X_{t+1} = (t+1, y)) &= \mathbb{P}_{(0,0)}^\omega(X_t = (t, y+1)) \frac{e^-(t, y+1)}{v(t, y+1)} \\ &\quad + \mathbb{P}_{(0,0)}^\omega(X_t = (t, y-1)) \frac{e^+(t, y-1)}{v(t, y-1)} \\ &= p(t, y+1; \omega) \frac{e^-(t, y+1)}{v(t, y+1)} + p(t, y-1; \omega) \frac{e^+(t, y-1)}{v(t, y-1)} \\ &= p(t+1, y; \omega). \end{aligned}$$

Thus the induction hypothesis holds for $t + 1$ and all $y \in \mathbb{Z}$. This completes the proof. \square

This gives us a Markov chain $\hat{X} = (X_n)_{n \geq 0}$ for a fixed environment ω . \mathbb{P}^ω is called the *quenched law*. The *joint annealed law* is given by

$$\tilde{\mathbb{P}}_{(t,y)}(d\hat{X}, d\omega) = \mathbb{P}_{(t,y)}^\omega(d\hat{X})\tilde{\mathbb{P}}(d\omega)$$

The law $\tilde{\mathbb{P}}_{(t,y)}(d\hat{X}, \Omega)$ is called the *marginal annealed law* or just *annealed law*. Let

$$\frac{X_{[nt]} - ([nt], 0)}{\sqrt{n}} = \left(0, S^{(n)}(t)\right) \quad (2.4)$$

The major features that make our model different from the models considered in literature so far are the following:

- (i) Our model is not symmetric, even for fixed ω , as the first co-ordinate (time) always takes a deterministic step of 1 with each transition of the Markov chain \hat{X} . This makes it different from the random conductance model, see [3]. This environment thus falls in the category of *random walk in random space-time environments* due to the deterministic increase in the time at each step.
- (ii) Our model is far from elliptic as there are points in space-time where the walker's next move is entirely determined by the environment. These points (t, y) occurs where all of the walkers in ω moved in the same direction ($e^+(t, y) > e^-(t, y) = 0$ or vice versa).
- (iii) For *space-time environments*, the only models considered so far, as far as we know, are the *i.i.d.* product environment considered by [12]. Our model is far from *i.i.d.* in that there are a lot of dependence between successive times.

These differences from the models in the literature present some technical challenges, although we adapt the same rough outline as in [12] and to prove a quenched invariance principle. Now we present the main theorem of this article:

Theorem 9. *For almost all ω with respect to $\tilde{\mathbb{P}}$, the process $S^{(n)}$ converges weakly to a standard Brownian motion B in the usual Skorohod topology under $\mathbb{P}_{(0,0)}^\omega$.*

From this we deduce the answer to the question that was posed at the beginning of this paper.

Corollary 9.1. *Almost surely the mass distribution (scaled by \sqrt{t}) converges weakly to the standard normal distribution.*

Proof. Fix an environment ω . As we have shown in Lemma 25 $p(t, y; \omega) = \mathbb{P}^\omega(X_t = (t, y))$ for all $t \in \mathbb{N}$ and all $y \in \mathbb{Z}$ so the scaled mass distribution at time t is given by $S^{(n)}(t)$. Then the Theorem follows directly from Theorem 9. \square

Before we set off proving the theorem, we recall some of the machinery involved in proving QIP in general space-time random environments, mostly taken from [12].

2 Quenched invariance principle in general space-time random environments

Let

$$D(\omega) = \mathbb{E}_{(0,0)}^\omega(X_1) = \left(1, \frac{e^+(0,0) - e^-(0,0)}{v(0,0)} \right)$$

For a bounded measurable function h on Ω define the operator

$$\Pi h(\omega) = \frac{e^+(0,0)h(T_{(1,1)}\omega) + e^-(0,0)h(T_{(1,-1)}\omega)}{v(0,0)}$$

The operator $h \mapsto \Pi h - h$ defines the generator of the Markov process of the environment as ‘seen from the particle’ with transition probability

$$\pi(\omega, A) = \mathbb{P}_{(0,0)}^\omega(T_{X_1}\omega \in A)$$

This Markov process, first conceived by [10] plays a key role in the arguments.

Now suppose that this Markov process has a stationary ergodic measure \mathbb{P}^∞ . Then we can extend the operator Π to a contraction on $L^p(\Omega, \mathbb{P}^\infty)$ for any $p \in [1, \infty]$. Denote the joint law of $(\omega, T_{X_1}\omega)$ by

$$\mu_2^\infty(d\omega_0, d\omega_1) = \pi(\omega_0, d\omega_1)\mathbb{P}^\infty(d\omega_0)$$

For a fixed ω , define the filtration $(\mathcal{F}_n^\omega)_{n \geq 0}$ where $\mathcal{F}_n^\omega = \sigma(X_0, X_1, \dots, X_n)$. The main idea involved in proving QIP involves creating a martingale M with respect to $(\mathbb{P}_{(0,0)}^\omega, \mathcal{F}_n^\omega)$ that is ‘close’ to the process X in some sense, and then applying martingale central limit theorems to M and translate the results to X . The main challenge involved is in the last step: to control the error involved in estimating X by M . To prove convergence of X to a Brownian motion in the Skorohod topology, we need to show that the error is uniformly small along a ‘typical’ path of X . The way to do this, which we elaborate now, has its roots in the seminal work of Kipnis and Varadhan, and subsequently extended to the non-reversible set-up by Maxwell and Woodroffe [8] and Derriennic and Lin [4].

Towards this end, note that the process

$$\bar{X}_n = X_n - \sum_{k=0}^{n-1} D(T_{X_k} \omega)$$

is a $(\mathbb{P}_{(0,0)}^\omega, \mathcal{F}_n^\omega)$ martingale. So, if we can prove that $\sum_{k=0}^{n-1} D(T_{X_k} \omega) - (n, 0)$ can be expressed as a $(\mathbb{P}_{(0,0)}^\omega, \mathcal{F}_n^\omega)$ martingale plus some error that can be controlled, a quenched CLT can be established for $(X_n)_{n \geq 0}$. Now, a standard way to have this decomposition is to investigate whether there exists a solution h to *Poisson’s Equation*

$$h = \Pi h + D - (1, 0)$$

If such a solution h did exist, we could write down

$$\sum_{k=0}^{n-1} D(T_{X_k} \omega) - (n, 0) = \sum_{k=0}^{n-1} [h(T_{X_{k+1}} \omega) - \Pi h(T_{X_k} \omega)] + [h(T_{X_0} \omega) - \Pi h(T_{X_{n-1}} \omega)]$$

Note that the first term in the above expression is a $(\mathbb{P}_{(0,0)}^\omega, \mathcal{F}_n^\omega)$ martingale and the second term is L^2 bounded with respect to $\mathbb{P}^\infty \otimes \mathbb{P}_{(0,0)}^\omega$. This would immediately give us the result. But unfortunately, a solution to the Poisson equation might not exist. To see this, note that if it did, then

$$\mathbb{E}^\infty \left| \mathbb{E}_{(0,0)}^\omega (X_n - (n, 0)) \right|^2 = \mathbb{E}^\infty \left| \mathbb{E}_{(0,0)}^\omega h(T_{X_0} \omega) - \mathbb{E}_{(0,0)}^\omega \Pi h(T_{X_{n-1}} \omega) \right|^2 = O(1)$$

but the above does not hold for the i.i.d. space-time product environment where $\mathbb{E}^\infty \left| \mathbb{E}_{(0,0)}^\omega (X_n - (n, 0)) \right|^2$ grows like \sqrt{n} . Rassoul-Agha and Seppäläinen [12] follow an extension of this idea, first in-

roduced by [7], and later extended by Toth [15] and subsequently Maxwell and Woodroffe [8], to deal with the case where a solution to Poisson equation does not exist.

Let h_ϵ be a solution of

$$(1 + \epsilon)h = \Pi h + D - (1, 0)$$

Call $g = D - (1, 0)$. Then, the solution of the above can be written as

$$h_\epsilon = \sum_{k=1}^{\infty} \frac{\Pi^{k-1}g}{(1 + \epsilon)^k} = \epsilon \sum_{n=1}^{\infty} \frac{\sum_{k=0}^{n-1} \Pi^{k-1}g}{(1 + \epsilon)^{n+1}} \quad (2.5)$$

which is in $L^2(\Omega, \mathbb{P}^\infty)$. Define

$$H_\epsilon(\omega_0, \omega_1) = h_\epsilon(\omega_1) - \Pi h_\epsilon(\omega_0).$$

Then, as our random walk has bounded range, Theorem 2 of [12] says the following:

Theorem 10. *Let \mathbb{P}^∞ be any stationary ergodic probability measure for the Markov process on the environment Ω generated by $\Pi - I$. Also assume that there is $\alpha < 1$ such that*

$$\mathbb{E}^\infty \left| \mathbb{E}_{(0,0)}^\omega (X_n - (n, 0)) \right|^2 = O(n^\alpha) \quad (2.6)$$

Then $H = \lim_{\epsilon \rightarrow 0} H_\epsilon$ exists in $L^2(\mu_2^\infty)$ and for \mathbb{P}^∞ a.e. ω , $S^{(n)}$ converges weakly to a Brownian motion with variance

$$\sigma^2 = \mathbb{E}^\infty \mathbb{E}_{(0,0)}^\omega (X_1 - D(\omega) + H(\omega, T_{X_1}\omega))^2 \quad (2.7)$$

Proof. We outline the main steps of the proof from [12] (and some from [8]) for completeness.

Denote

$$\begin{aligned} M_n^\epsilon &= \sum_{k=0}^{n-1} H_\epsilon(T_{X_k}\omega, T_{X_{k+1}}\omega) \\ S_n^\epsilon &= \sum_{k=0}^{n-1} h_\epsilon(T_{X_k}\omega) \\ R_n^\epsilon &= h_\epsilon(\omega) - \Pi h_\epsilon(T_{X_n}\omega) \end{aligned}$$

Then,

$$X_n - (n, 0) = \bar{X}_n + M_n^\epsilon + \epsilon S_n^\epsilon + R_n^\epsilon$$

From the second equality in 2.5, by using 2.6 and comparing with $\int_0^\infty x^\alpha(1+\epsilon)^{-x}dx$, we get

$$\mathbb{E}^\infty |h_\epsilon|^2 = O(\epsilon^{-2\alpha})$$

This, along with the fact that $\|H_\epsilon - H_\delta\|_{L^2(\mu_2^\infty)} \leq (\epsilon + \delta) \left(\mathbb{E}^\infty |h_\epsilon|^2 + \mathbb{E}^\infty |h_\delta|^2 \right)$ implies the existence of the limit $H = \lim_{\epsilon \rightarrow 0} H_\epsilon$ in $L^2(\mu_2^\infty)$. It furthermore implies that for fixed n , $\lim_{\epsilon \rightarrow 0} S_n^\epsilon = 0$ in L^2 , under the appropriate measure. Thus, we conclude

$$M_n = \sum_{k=0}^{n-1} H(T_{X_k}\omega, T_{X_{k+1}}\omega)$$

is \mathbb{P}^∞ almost surely a $(\mathbb{P}_{(0,0)}^\omega, \mathcal{F}_n^\omega)$ square integrable martingale. Furthermore, by choosing k_n to be the unique k satisfying $2^{k-1} \leq n < 2^k$ and $\epsilon_n = 2^{k_n}$, we conclude that the error given by

$$R_n = X_n - (n, 0) - \bar{X}_n - M_n = M_n^\epsilon - M_n + \epsilon S_n^\epsilon + R_n^\epsilon$$

satisfies

$$\mathbb{E}^\infty \left| \mathbb{E}_{(0,0)}^\omega(R_n) \right|^2 = O(n^{2\alpha}) \quad (2.8)$$

Let

$$n^{-1/2} (\bar{X}_{[nt]} + M_{[nt]}) = (0, M_n^*(t))$$

Then, by martingale invariance principles, M_n^* converges weakly to a Brownian motion with the given variance under $\mathbb{P}_{(0,0)}^\omega$ for \mathbb{P}^∞ almost sure ω . Here, the assumption on the ergodicity of T_{X_1} is needed for the existence of the appropriate limits required for the Martingale Functional CLT to hold (see Theorem 2.11 of [3]).

As we have

$$\sup_{0 \leq t \leq 1} |S^{(n)}(t) - M_n^*(t)| \leq n^{-1/2} \max_{k \leq n} |R_k| \quad (2.9)$$

therefore, the QIP for $S^{(n)}$ will follow if we can show that

$$n^{-1/2} \max_{k \leq n} |R_k| \rightarrow 0$$

in $\mathbb{P}_{(0,0)}^\omega$ -probability as $n \rightarrow \infty$ \mathbb{P}^∞ almost every ω . This follows from the theory of ‘fractional coboundaries’ developed in Derriennic and Lin [4] and the estimate (2.8). \square

So, to prove Theorem 9 with the aid of Theorem 10, we need to prove the existence of a stationary distribution \mathbb{P}^∞ for the Markov process of the environment as seen by the particle (generated by $\Pi - I$), the moment bound (2.6) and the ergodicity of the Markov shift T_{X_1} with respect to \mathbb{P}^∞ . We prove these in the following section.

3 Proof of Theorem 9

PROOF OF THEOREM 9. We break this proof up into three parts corresponding to the three conditions in Theorem 10. In Section 3.1, we show the existence of a stationary measure \mathbb{P}^∞ for the Markov process of the environment with Lemma 26. In Section 3.2, we prove Lemma 27 which gives a moment condition that verifies (2.6). In Section 3.3, we prove the ergodicity of the Markov shift process $\{T_{X_n}\}_{n \geq 1}$ with respect to the stationary measure \mathbb{P}^∞ . This is in Lemma 37. These three lemmas verify the hypothesis of Theorem 10 thus we prove the desired almost sure invariance principle. That the convergence is to a standard Brownian motion can be seen by the calculation of the variance of $\sigma^2 = 1$ in the note in Section 3.1. \square

3.1 Stationarity

In Lemma 26, we show that the measure $\tilde{\mathbb{P}}$ is in fact a stationary measure for the Markov process of the environment.

Lemma 26. *The measure $\tilde{\mathbb{P}}$ is stationary for the Markov process on Ω as seen from the particle, that is, the Markov process generated by the random shifts $(T_{X_n})_{n \geq 0}$.*

Proof. Let $\tilde{\mathbb{E}}$ denote the corresponding expectation. Take ψ measurable with respect to \mathcal{G} .

It will suffice to show that $\tilde{\mathbb{E}}(\Pi\psi) = \tilde{\mathbb{E}}(\psi)$. Note that

$$\begin{aligned}
\tilde{\mathbb{E}}(\Pi\psi) &= \mathbb{E}(v(0,0)\Pi\psi) \\
&= \mathbb{E}(e^+(0,0)\psi(T_{(1,1)}\omega)) + \mathbb{E}(e^-(0,0)\psi(T_{(1,-1)}\omega)) \\
&= \mathbb{E}(e^+(-1,-1)\psi(\omega)) + \mathbb{E}(e^-(-1,1)\psi(\omega)) \quad (\text{by translation invariance of } \mathbb{P}) \\
&= \mathbb{E}(v(0,0)\psi(\omega)) \quad (\text{as } e^+(-1,-1) + e^-(-1,1) = v(0,0)) \\
&= \tilde{\mathbb{E}}(\psi)
\end{aligned}$$

This proves the lemma. \square

Note: We claimed in Theorem 9 that the limit is a standard Brownian motion. But Theorem 10 presents the diffusion coefficient of the Brownian motion in a more abstract form. Here we show that the expression in (2.7) indeed reduces to one in our case.

$$\begin{aligned}
\sigma^2 &= \tilde{\mathbb{E}}_{(0,0)} |X_1 - D(\omega) + H(\omega, T_{X_1}\omega)|^2 \\
&= \lim_{\epsilon \rightarrow 0} \tilde{\mathbb{E}}_{(0,0)} |X_1 - (1,0) - D(\omega) + h_\epsilon(T_{X_1}\omega) - \Pi h_\epsilon(\omega)|^2 \\
&= \lim_{\epsilon \rightarrow 0} \tilde{\mathbb{E}}_{(0,0)} |X_1 - (1,0) - \epsilon h_\epsilon(\omega)|^2 \\
&= \tilde{\mathbb{E}}_{(0,0)} |X_1 - (1,0)|^2 \quad (\text{as } \tilde{\mathbb{E}}|h_\epsilon|^2 = O(\epsilon^{-2\alpha})) \\
&= \mathbb{E}v(0,0) \left(\frac{e^+(0,0) + e^-(0,0)}{v(0,0)} \right) = 1
\end{aligned}$$

where the limit in the second equality above follows from the L^2 convergence of H_ϵ to H described in Theorem 10.

3.2 Moment condition

In this section we verify the moment condition by proving the following lemma.

Lemma 27. $\tilde{\mathbb{E}} \left| \mathbb{E}_{(0,0)}^\omega(X_n) - (n, 0) \right|^2 = O(\sqrt{n}(\log n)^3)$.

SKETCH OF PROOF OF LEMMA 27. Our strategy is as follows. Fix an environment ω . Let X and \tilde{X} be two independent copies of the random walk starting from $(0,0)$ and running in the *same environment* ω . Also define Y_i by $(0, Y_i) = X_i - \tilde{X}_i$ and let $\tilde{\mathbb{P}}^*$ be the law of the Markov process

$$(X_i, \tilde{X}_i, T_{X_i}\omega, T_{\tilde{X}_i}\omega).$$

Finally let $\tilde{\mathbb{E}}^*$ be its expectation.

Step 1 Our first task is to establish

$$\tilde{\mathbb{E}} \left| \mathbb{E}_{(0,0)}^\omega(X_n) - (n, 0) \right|^2 \leq \tilde{\mathbb{E}}^* \left(\#\{i : Y_i = 0\} \right) \quad (2.10)$$

Step 2 Next we show that Y_i behaves like (two times) a lazy random walk when $Y_i \neq 0$. But when $Y_i = 0$ the distribution of Y_{i+1} has a complicated dependence structure.

Step 3 We use the sequence $\{Y_i\}_{i=0}^n$ to divide the interval $[0, n]$ into excursions (maximal connected intervals where $Y_i \neq 0$) and holding times (maximal connected intervals where $Y_i = 0$). By the previous step we know the distribution of excursion lengths. Much of the work of this proof is in estimating the distribution of the lengths of the holding times.

Step 4 We bound the right hand side of (2.10) by breaking it up into two parts. One is when the longest holding time is at most $k \log^3(n)$ and the other is when the longest holding time is longer than $k \log^3(n)$. Using the fact that excursions have the same distribution as a lazy random walk we are able to show that for a typical sequence Y_i there are about \sqrt{n} holds (and also about \sqrt{n} excursions). So the first part contributes approximately $\sqrt{n} \log^3(n)$ to the expectation. We show that the probability of a long holding time is going to zero sufficiently quickly so the second part is contributing very little to the expectation (provided that k is sufficiently large). Adding up these two bounds completes the proof.

□

Lemma 28. *Fix an environment ω . Let X and \tilde{X} being two independent copies of the random walk starting from $(0, 0)$ and running in the environment ω . Also define Y_i by $(0, Y_i) = X_i - \tilde{X}_i$ and let $\tilde{\mathbb{P}}^*$ be the law of the Markov process*

$$\left(X_i, \tilde{X}_i, T_{X_i} \omega, T_{\tilde{X}_i} \omega \right).$$

Finally let $\tilde{\mathbb{E}}^*$ be its expectation. Then

$$\tilde{\mathbb{E}} \left| \mathbb{E}_{(0,0)}^\omega(X_n) - (n, 0) \right|^2 \leq \sum_{i=0}^{n-1} \tilde{\mathbb{P}}^*(Y_i = 0) = \tilde{\mathbb{E}}^* \left(\#\{i : Y_i = 0\} \right) \quad (2.11)$$

Proof. We can write down

$$\begin{aligned} \mathbb{E}_{(0,0)}^\omega(X_n) &= \sum_{i=0}^{n-1} [\mathbb{E}_{(0,0)}^\omega(X_{i+1}) - \mathbb{E}_{(0,0)}^\omega(X_i)] \\ &= \left[\mathbb{E}_{(0,0)}^\omega \left((X_i + (1, 1)) \frac{e^+(X_i)}{v(X_i)} + (X_i + (1, -1)) \frac{e^-(X_i)}{v(X_i)} \right) - \mathbb{E}_{(0,0)}^\omega(X_i) \right] \\ &= \sum_{i=0}^{n-1} \mathbb{E}_{(0,0)}^\omega D(T_{X_i}\omega) \end{aligned}$$

Recall that $g = D - (1, 0)$. Therefore,

$$\begin{aligned} \tilde{\mathbb{E}} \left| \mathbb{E}_{(0,0)}^\omega(X_n) - (n, 0) \right|^2 &= \tilde{\mathbb{E}} \left| \sum_{i=0}^{n-1} \mathbb{E}_{(0,0)}^\omega g(T_{X_i}\omega) \right|^2 \\ &= \tilde{\mathbb{E}} \left| \sum_{i=0}^{n-1} \sum_{y \in \mathbb{Z}} \mathbb{P}_{(0,0)}^\omega(X_i = (i, y)) g(T_{(i,y)}\omega) \right|^2 \\ &= \tilde{\mathbb{E}} \left[\sum_{i,j=0}^{n-1} \sum_{y,z \in \mathbb{Z}} \mathbb{P}_{(0,0)}^\omega(X_i = (i, y)) \mathbb{P}_{(0,0)}^\omega(X_j = (j, z)) g(T_{(i,y)}\omega) \cdot g(T_{(j,z)}\omega) \right] \end{aligned}$$

where in the above sum, we take a term to be zero if $\mathbb{P}_{(0,0)}^\omega(X_i = (i, y)) = 0$ or $\mathbb{P}_{(0,0)}^\omega(X_j = (j, z)) = 0$. Note that consequently, we only have terms with $v(i, y) \geq 1$ and $v(j, z) \geq 1$.

If $i > j$, we can condition on

$$\mathcal{F}_i = \sigma\{e^+(t, y), e^-(t, y) : t \leq i, y \in \mathbb{Z}\} \quad (2.12)$$

to get

$$\begin{aligned} &\tilde{\mathbb{E}} \left(\mathbb{P}_{(0,0)}^\omega(X_i = (i, y)) \mathbb{P}_{(0,0)}^\omega(X_j = (j, z)) g(T_{(i,y)}\omega) \cdot g(T_{(j,z)}\omega) \middle| \mathcal{F}_i \right) \\ &= \mathbb{P}_{(0,0)}^\omega(X_i = (i, y)) \mathbb{P}_{(0,0)}^\omega(X_i = (j, z)) g(T_{(j,z)}\omega) \cdot \tilde{\mathbb{E}}(g(T_{(i,y)}\omega) \middle| \mathcal{F}_i) = 0 \end{aligned}$$

using the fact that

$$\begin{aligned} e^+(i, y)|\mathcal{F}_i &\sim \text{Bin}(v(i, y), 1/2) \\ e^-(i, y)|\mathcal{F}_i &\sim \text{Bin}(v(i, y), 1/2) \end{aligned} \quad (2.13)$$

Also, the terms in the above sum corresponding to $i = j$ and $y \neq z$ vanish as $(e^+(i, y), e^-(i, y))$ and $(e^+(i, z), e^-(i, z))$ are conditionally independent given \mathcal{F}_i implying

$$\tilde{\mathbb{E}}(g(T_{(i,y)}\omega) \cdot g(T_{(i,z)}\omega)|\mathcal{F}_i) = \tilde{\mathbb{E}}(g(T_{(i,y)}\omega)|\mathcal{F}_i) \cdot \tilde{\mathbb{E}}(g(T_{(i,z)}\omega)|\mathcal{F}_i) = 0$$

Thus the only terms in the above sum which do not vanish are those with $i = j$ and $y = z$.

For these terms, we see that

$$\begin{aligned} \tilde{\mathbb{E}}(|g(T_{(i,y)}\omega)|^2|\mathcal{F}_i) &= \tilde{\mathbb{E}}\left(\frac{(e^+(i, y) - e^-(i, y))^2}{v^2(i, y)}|\mathcal{F}_i\right) \\ &= 4\tilde{\mathbb{E}}\left(\frac{(e^+(i, y) - \frac{1}{2}v(i, y))^2}{v^2(i, y)}|\mathcal{F}_i\right) = \frac{1}{v(i, y)} \end{aligned}$$

by (2.13). Thus we get

$$\begin{aligned} \tilde{\mathbb{E}}\left|\mathbb{E}_{(0,0)}^\omega(X_n) - (n, 0)\right|^2 &= \tilde{\mathbb{E}}\left(\sum_{i=0}^{n-1} \sum_{y \in \mathbb{Z}} \left[\mathbb{P}_{(0,0)}^\omega(X_i = (i, y))\right]^2 \frac{1}{v(i, y)}\right) \\ &\leq \tilde{\mathbb{E}}\left(\sum_{i=0}^{n-1} \sum_{y \in \mathbb{Z}} \left[\mathbb{P}_{(0,0)}^\omega(X_i = (i, y))\right]^2\right) \\ &= \sum_{i=0}^{n-1} \tilde{\mathbb{P}}^*(Y_i = 0) \end{aligned} \quad (2.14)$$

□

Also, notice that although the two walks X and \tilde{X} are conditionally independent given ω , the annealed joint law of (X, \tilde{X}) is far from being the product of the annealed marginal laws of X and \tilde{X} , as the subsequent calculations will indicate. This is what creates a major technical obstacle.

Now we start to understand the annealed law of the process Y_i . We will show that it behaves like a lazy random walk provided that $Y_i \neq 0$, but at 0 it behaves very differently. Our first step in making that precise is the following.

Lemma 29. *Let Y_i be defined as in Lemma 28. If $r \neq 0$ then*

$$\tilde{\mathbb{P}}^*(Y_{n+1} = s \mid Y_n = r, Y_{n-1} = r_{n-1}, \dots, Y_1 = r_1) = \begin{cases} \frac{1}{4} & \text{if } |s - r| = 2 \\ \frac{1}{2} & \text{if } s = r \end{cases}$$

Proof. Note that for r, s with $r \neq 0$ and $s - r = 0, +2$ or -2 , we get

$$\begin{aligned} & \tilde{\mathbb{P}}^*(Y_{n+1} = s, Y_n = r, Y_{n-1} = r_{n-1}, \dots, Y_1 = r_1) \\ &= \tilde{\mathbb{E}} \sum_{y_i - z_i = r_i, y - z = r} \mathbb{P}_{(0,0)}^\omega(X_1 = (1, y_1), \dots, X_n = (n, y)) \mathbb{P}_{(0,0)}^\omega(X_1 = (1, z_1), \dots, X_n = (n, z)) \\ & \quad \times \tilde{\mathbb{E}} \left[\sum_{y' - z' = s} \mathbb{P}_{(n,y)}^\omega(X_1 = (1, y')) \mathbb{P}_{(n,z)}^\omega(X_1 = (1, z')) \mid \mathcal{F}_n \right] \end{aligned}$$

where the filtration $\{\mathcal{F}_n\}$ is defined in (2.12). By (2.13),

$$\tilde{\mathbb{E}} \left[\sum_{y' - z' = s} \mathbb{P}_{(n,y)}^\omega(X_1 = (1, y')) \mathbb{P}_{(n,z)}^\omega(X_1 = (1, z')) \mid \mathcal{F}_n \right] = \begin{cases} 1/4 & \text{when } |s - r| = 2 \\ 1/2 & \text{when } s = r \end{cases}$$

Thus we get

$$\begin{aligned} & \tilde{\mathbb{P}}^*(Y_{n+1} = s, Y_n = r, Y_{n-1} = r_{n-1}, \dots, Y_1 = r_1) \\ &= \begin{cases} \frac{1}{4} \tilde{\mathbb{P}}^*(Y_n = r, Y_{n-1} = r_{n-1}, \dots, Y_1 = r_1) & \text{if } |s - r| = 2 \\ \frac{1}{2} \tilde{\mathbb{P}}^*(Y_n = r, Y_{n-1} = r_{n-1}, \dots, Y_1 = r_1) & \text{if } s = r \end{cases} \end{aligned}$$

Therefore,

$$\begin{aligned} & \tilde{\mathbb{P}}^*(Y_{n+1} = s \mid Y_n = r, Y_{n-1} = r_{n-1}, \dots, Y_1 = r_1) \\ &= \tilde{\mathbb{P}}^*(Y_{n+1} = s \mid Y_n = r) = \begin{cases} \frac{1}{4} & \text{if } |s - r| = 2 \\ \frac{1}{2} & \text{if } s = r \end{cases} \end{aligned}$$

□

Thus we see that away from zero, the process Y behaves as a homogeneous Markov process with the given transition probabilities. But at zero, things are not so nice. If $Y_n = 0$ and $v(X_n) = 1$ then $Y_{n+1} = 0$ as well. But we will now show that if $Y_n = 0$ and $v(X_n) > 1$ then $Y_{n+1} \neq 0$ with probability at least $1/4$.

Lemma 30. *Let Y_i be defined as in Lemma 28. Then*

$$\begin{aligned} & \tilde{\mathbb{P}}^*(Y_{n+1} = 0, Y_n = 0, Y_{n-1} = r_{n-1}, \dots, Y_1 = r_1) \\ &= \tilde{\mathbb{E}} \sum_{y_i - z_i = r_i, y \in \mathbb{Z}} \mathbb{P}_{(0,0)}^\omega(X_1 = (1, y_1), \dots, X_n = (n, y)) \mathbb{P}_{(0,0)}^\omega(X_1 = (1, z_1), \dots, X_n = (n, y)) \\ & \quad \times \frac{1}{2} \left(1 + \frac{1}{v(n, y)} \right) \end{aligned}$$

Proof. A simple calculation proves the lemma.

$$\begin{aligned} & \tilde{\mathbb{P}}^*(Y_{n+1} = 0, Y_n = 0, Y_{n-1} = r_{n-1}, \dots, Y_1 = r_1) \\ &= \tilde{\mathbb{E}} \sum_{y_i - z_i = r_i, y \in \mathbb{Z}} \mathbb{P}_{(0,0)}^\omega(X_1 = (1, y_1), \dots, X_n = (n, y)) \mathbb{P}_{(0,0)}^\omega(X_1 = (1, z_1), \dots, X_n = (n, y)) \\ & \quad \times \tilde{\mathbb{E}} \left[\frac{(e^+(n, y))^2 + (e^-(n, y))^2}{v^2(n, y)} \middle| \mathcal{F}_n \right] \\ &= \tilde{\mathbb{E}} \sum_{y_i - z_i = r_i, y \in \mathbb{Z}} \mathbb{P}_{(0,0)}^\omega(X_1 = (1, y_1), \dots, X_n = (n, y)) \mathbb{P}_{(0,0)}^\omega(X_1 = (1, z_1), \dots, X_n = (n, y)) \\ & \quad \times \frac{1}{2} \left(1 + \frac{1}{v(n, y)} \right) \end{aligned}$$

□

Note that the presence of the term $\left(1 + \frac{1}{v(n, y)} \right)$ in the summand makes this process depend on its entire past while making transitions from zero, thus destroying its Markov property. This is a major difference from its analogue in the i.i.d. case studied by [12], where the process Y is a homogeneous Markov process perturbed at zero.

To deal with this problem, we use $\{Y_i\}$ to decompose $[0, n]$ into *excursions* away from zero and *holding times* at zero. Let

$$\{e_j : 1 \leq j \leq a(n)\}$$

be the excursions of Y away from zero on $[0, n]$. Let the time interval spanned by e_j be $[\alpha_j, \beta_j]$, i.e., $e_j(0) = |Y_{\alpha_j}| = 2$ and $\beta_j = \inf\{k > \alpha_j : Y_k = 0\}$. Let the first holding time at 0 be $\gamma_0 = \alpha_1 - 1$ and let $\gamma_j = \alpha_{j+1} - \beta_j - 1$ be the holding time at zero between e_j and e_{j+1} . With these defined, we can write down

$$\sum_{i=0}^{n-1} 1(Y_i = 0) \leq \sum_{i=0}^{a(n)} \gamma_i \leq (a(n) + 1) \left(\sup_{j \leq n} \gamma_j \right) \quad (2.15)$$

Let the time duration of the j -th excursion be denoted by $T_j = \beta_j - \alpha_j$.

Denote by R the homogeneous random walk starting from zero with transition probabilities to $2, 0, -2$ being $1/4, 1/2, 1/4$ respectively, and let

$$R_n^* = \sup_{k \leq n} R_k$$

Denote the law of R by \mathbb{P}_R and the corresponding expectation by \mathbb{E}_R . Let

$$A_n = \sup\{k \geq 0 : T_1 + \dots + T_k \leq n\}$$

It is easy to see that $T_1 + \dots + T_k$ has the same distribution as the hitting time of level $2k$ by R . Thus $\{A_n : n \geq 1\}$ and $\{R_n^*/2 : n \geq 1\}$ have the same distribution. Note that $a(n) \leq A_n + 1$. Therefore,

$$\begin{aligned} \tilde{\mathbb{E}}^* a(n) &\leq \tilde{\mathbb{E}}^*(A_n + 1) \\ &= \mathbb{E}_R(R_n^*/2 + 1) \\ &\leq C\sqrt{n} \end{aligned} \quad (2.16)$$

for some constant $C < \infty$.

The excursions away from zero have the same law as those of a lazy random walk R so their distribution is well understood. Our main challenge is to provide an upper bound on

the supremum of the holding times γ_j . For a lazy random walk R , the holding times are i.i.d. geometric random variables and we can derive an upper bound on the probability that a holding time is at least $\log n$. The γ_j are far from i.i.d. and they depend on the entire past till that time. In spite of this dependence we will bound the probability that a holding time γ_j is bigger than $(\log n)^k$ for some positive number k . Our bound will be uniform in j and independent of all previous holding times. Using these bounds we will be able to bound the probability that the supremum of the holding times is large.

To do this we note that if $Y_i = 0$ and $v(X_i) \geq 2$ then there is at least a probability of $1/4$ that $Y_{i+1} \neq 0$, independent of anything in the process or the environment up to time i . We now define a set of stopping times to indicate when these times occur. These stopping times will give us our bound on the holding times.

We now define some stopping times for the Markov process $(X_i, \tilde{X}_i, T_{X_i}, T_{\tilde{X}_i})$. For each $j \geq 1$, define:

$$I_1^{(j)} = \beta_j \wedge n$$

$$I_{i+1}^{(j)} = \begin{cases} \inf\{k > I_i^{(j)} : v(X_k) \geq 2\} \wedge n & \text{if } X_{I_i^{(j)}+1} = \tilde{X}_{I_i^{(j)}+1}, \\ I_i^{(j)} & \text{otherwise.} \end{cases}$$

Thus we get a bi-indexed family $\{I_i^{(j)} : 1 \leq i < \infty\}$ which is well defined with probability one. Note that for each j , the sequence indexed by i eventually becomes constant.

For $j \geq 1$ with $\beta_j < n$, these stopping times indexed by i represent the times $t \in [\beta_j, \alpha_{j+1} - 1] \cap [0, n]$ (the portion of the holding time at zero between e_j and e_{j+1} lying in $[0, n]$) where there are at least two particles at $X_t (= \tilde{X}_t)$ and the sequence becomes constant if either time n or time $\alpha_{j+1} - 1$ is reached.

Call the corresponding stopped sigma fields $\{\mathcal{F}_{I_i^{(j)}}^* : 1 \leq i < \infty\}$. Let $k_j = \sup\{i \geq 1 : I_1^{(j)} < I_2^{(j)} < \dots < I_i^{(j)}\}$ with $k_j = 1$ for a constant sequence. With this notation, $\gamma_j = I_{k_j}^{(j)} - I_1^{(j)}$.

Control over the holding times γ_j is obtained through the following two lemmas.

Lemma 31.

$$\tilde{\mathbb{P}}^*(k_j \geq k) \leq \left(\frac{3}{4}\right)^{k-1} \quad (2.17)$$

Proof. Note that, by the strong Markov property applied at these stopping times,

$$\begin{aligned} \tilde{\mathbb{P}}^*(k_j \geq k) &= \tilde{\mathbb{E}}^* \mathbf{1}(I_1^{(j)} < I_2^{(j)} < \dots < I_k^{(j)}) \\ &\leq \tilde{\mathbb{E}}^* \mathbf{1}(I_1^{(j)} < I_2^{(j)} < \dots < I_{k-1}^{(j)} < n) \tilde{\mathbb{E}}^* \left(\mathbf{1} \left(X_{I_{k-1}^{(j)}+1}^{(j)} = \tilde{X}_{I_{k-1}^{(j)}+1} \right) \middle| \mathcal{F}_{I_{k-1}^{(j)}}^* \right) \\ &= \tilde{\mathbb{E}}^* \mathbf{1}(I_1^{(j)} < I_2^{(j)} < \dots < I_{k-1}^{(j)} < n) \tilde{\mathbb{E}}^* \left[\frac{e^+ \left(X_{I_{k-1}^{(j)}}^{(j)} \right)^2 + e^- \left(X_{I_{k-1}^{(j)}}^{(j)} \right)^2}{v \left(X_{I_{k-1}^{(j)}}^{(j)} \right)^2} \middle| T_{X_{I_{k-1}^{(j)}}} \right] \quad (\text{by (2.13)}) \\ &= \tilde{\mathbb{E}}^* \mathbf{1}(I_1^{(j)} < I_2^{(j)} < \dots < I_{k-1}^{(j)} < n) \cdot \frac{1}{2} \left(1 + \frac{1}{v \left(X_{I_{k-1}^{(j)}}^{(j)} \right)} \right) \\ &\leq \frac{3}{4} \tilde{\mathbb{E}}^* \mathbf{1}(I_1^{(j)} < I_2^{(j)} < \dots < I_{k-1}^{(j)}) \leq \left(\frac{3}{4}\right)^{k-1} \end{aligned}$$

where the last step follows by induction. \square

Lemma 32. For sufficiently large $M > 0$,

$$\tilde{\mathbb{P}}^* \left(\bigcup_{0 \leq i, j \leq n} \{I_{i+1}^{(j)} - I_i^{(j)} \geq M^2 \log^2 n\} \right) \leq \frac{2}{n^{M-1}} \quad (2.18)$$

Proof. For each $u \leq n$, let

$$\tau_u = \inf\{k \geq u : v(X_k) \geq 2\}$$

Note that by the stationarity of T_{X_k} under $\tilde{\mathbb{P}}^*$,

$$\tilde{\mathbb{P}}^*(\tau_u \geq M^2 \log^2 n) = \tilde{\mathbb{P}}^*(\tau_0 \geq M^2 \log^2 n)$$

Under $\tilde{\mathbb{P}}^*$, $v(0, k) \sim \text{Poi}(1)$ for $k \neq 0$ and $v(0, 0)$ has a *size-biased* $\text{Poi}(1)$ distribution. Let

$$\begin{aligned} N_t^+ &= \#\{(k, i) : k \geq 1, S^{(k, i)}(t) < 0\} \\ N_t^- &= \#\{(k, i) : k \leq -1, S^{(k, i)}(t) > 0\} \end{aligned}$$

Then, it is easy to check that $N_t^+, N_t^- \sim \text{Poi}(\mu_t)$ where

$$\mu_t = \sum_{k=1}^{\infty} \mathbb{P}(S_t \geq k) = \mathbb{E}S_t^+$$

where S is a simple random walk starting from zero. Note that if both N_t^+ and N_t^- are non-zero, then X_k starting from $(0, 0)$ must intersect at least one random walk not starting from the origin before time t , and at the time of intersection, say k , $v(X_k) \geq 2$. Thus,

$$\begin{aligned} \tilde{\mathbb{P}}^*(\tau_0 \geq t) &\leq \tilde{\mathbb{P}}^*(N_t^+ = 0 \text{ or } N_t^- = 0) \\ &\leq 2e^{-\mu t} \\ &\leq 2e^{-C\sqrt{t}} \end{aligned}$$

as $\mathbb{E}S_t^+ = (1/2)\mathbb{E}|S_t| \geq C\sqrt{t}$. Thus,

$$\tilde{\mathbb{P}}^*(\tau_0 \geq M^2 \log^2 n) \leq 2e^{-M \log n} = \frac{2}{n^M}$$

Consequently,

$$\begin{aligned} \tilde{\mathbb{P}}^* \left(\bigcup_{0 \leq i, j \leq n} \{I_{i+1}^{(j)} - I_i^{(j)} \geq M^2 \log^2 n\} \right) &\leq \tilde{\mathbb{P}}^* (\exists u \leq n \text{ with } \tau_u \geq M^2 \log^2 n) \\ &\leq \frac{2}{n^{M-1}} \end{aligned}$$

by a simple union bound. □

Lemmas 31 and 32 yield the following corollary:

Corollary 10.1. *There exists a constant C that does not depend on n such that for any $M > 0$, we can choose M' (depending on M) satisfying*

$$\tilde{\mathbb{P}}^*(\sup_{j \leq n} \gamma_j \geq M' \log^3 n) \leq Cn^{-M} \tag{2.19}$$

for all n .

Proof. For sufficiently large M and any $j, k \geq 1$, we have:

$$\begin{aligned}
\tilde{\mathbb{P}}^*(\gamma_j \geq M^2 k \log^2 n) &= \tilde{\mathbb{P}}^*(\gamma_j \geq M^2 k \log^2 n, I_k^{(j)} - I_1^{(j)} < M^2 k \log^2 n) \\
&+ \tilde{\mathbb{P}}^*(\gamma_j \geq M^2 k \log^2 n, I_k^{(j)} - I_1^{(j)} \geq M^2 k \log^2 n) \\
&\leq \tilde{\mathbb{P}}^*(\gamma_j > I_k^{(j)} - I_1^{(j)}) + \tilde{\mathbb{P}}^*\left(\sum_{i=0}^{k-1} (I_{i+1}^{(j)} - I_i^{(j)}) \geq M^2 k \log^2 n\right) \\
&\leq \tilde{\mathbb{P}}^*(k_j \geq k) + \tilde{\mathbb{P}}^*\left(\bigcup_{0 \leq i, j \leq n} \{I_{i+1}^{(j)} - I_i^{(j)} \geq M^2 \log^2 n\}\right) \\
&\leq \left(\frac{3}{4}\right)^{k-1} + \frac{2}{n^{M-1}}
\end{aligned}$$

where the last step follows from Lemmas 31 and 32.

The assertion then follows by taking $k = M \log n$ and the union bound. \square

Now, to prove Lemma 27, notice that by (2.15),

$$\begin{aligned}
\tilde{\mathbb{E}}^*\left(\sum_{i=0}^{n-1} 1(Y_i = 0)\right) &\leq \tilde{\mathbb{E}}^*\left((a(n) + 1) \left(\sup_{j \leq n} \gamma_j\right)\right) \\
&= \tilde{\mathbb{E}}^*\left((a(n) + 1) \left(\sup_{j \leq n} \gamma_j\right) 1(\sup_{j \leq n} \gamma_j < M' \log^3 n)\right) \\
&+ \tilde{\mathbb{E}}^*\left((a(n) + 1) \left(\sup_{j \leq n} \gamma_j\right) 1(\sup_{j \leq n} \gamma_j \geq M' \log^3 n)\right) \\
&\leq (M' \log^3 n) \tilde{\mathbb{E}}^*(a(n) + 1) + n^2 \tilde{\mathbb{P}}^*(\sup_{j \leq n} \gamma_j \geq M' \log^3 n) \\
&\leq C_1 \sqrt{n} \log^3 n + C n^{-(M-2)} \quad (\text{by (2.16)}) \\
&\leq C_2 \sqrt{n} \log^3 n
\end{aligned}$$

choosing $M > 2$.

This, together with (2.14), gives Lemma 27. \square

3.3 Ergodicity of the Markov Shift

In this section we prove the ergodicity of the shift from the point of view of a tagged particle.

Lemma 33. *The vertical shift $(\Omega, T_{(0,1)}, \mathbb{P})$ is strong mixing and thus totally ergodic. In particular $(\Omega, T_{(0,2)}, \mathbb{P})$ is ergodic.*

Proof. For any cylinder set $R \subset \Omega$ which is defined by the values in a finite rectangle we have that for any sufficiently large M the sets $T_{(0,M)}(R)$ and R are independent. As these cylinder sets generate the σ -algebra, the shift $T_{(0,1)}$ is mixing and totally ergodic. Thus $T_{(0,2)}$ is ergodic. A (very slightly) more involved analysis shows that $T_{(0,1)}$ is weak Bernoulli and isomorphic to the shift on an infinite entropy i.i.d. measure. \square

Now we want to define a new transformation built from $(\Omega, T_{(0,2)}, \mathbb{P})$. This transformation is constructed from $(\Omega, T_{(0,2)}, \mathbb{P})$ by inducing on the subset $\{v(0,0) > 0\}$ and then building a “Kakutani skyscraper” of height $v(0,0)$. The techniques of inducing and building skyscrapers are standard in ergodic theory and have been well studied in the theory of Kakutani equivalence [9] [5] [11]. We start by describing the state space and the measure.

Let $\hat{\Omega} \subset \Omega \times \mathbb{N}$ consist of points of the form (ω, y) with $y \leq v(0,0)$. Let $\hat{\mathbb{P}}$ be a measure on $\hat{\Omega}$ defined as follows. Fix any $A \subset \Omega$ and $i \in \mathbb{N}$. Then

$$\hat{\mathbb{P}}(A, i) = \mathbb{P}(A \cap \{i \leq v(0,0)\}).$$

Let $\pi : \hat{\Omega} \rightarrow \Omega$ be the projection map, $\pi(x, y) = x$. For any $A \subset \Omega$, let $\hat{A} \subset \hat{\Omega}$ be defined by $\hat{A} = \pi^{-1}(A)$. Also, recall that $\tilde{\mathbb{P}}$ is an invariant measure for our process $\{T_{X_n} : n \in \mathbb{N}\}$. These definitions give us the following lemma.

Lemma 34. *For any $A \subset \Omega$*

$$\tilde{\mathbb{P}}(A) = \hat{\mathbb{P}}(\hat{A})$$

as both are the size biased version of the measure \mathbb{P} .

Now we define the new transformation. For $\omega \in \Omega$ let

$$n(\omega) = \inf\{m > 0 : 2 \mid m, v(0, m) > 0\} \quad \text{and} \quad \tilde{\omega} = T_{(0, n(\omega))}(\omega).$$

Next we define $\hat{T} : \hat{\Omega} \rightarrow \hat{\Omega}$ by

$$\hat{T}(\omega, y) = \begin{cases} (\omega, y + 1) & \text{if } y < v(0, 0) \\ (\tilde{\omega}, 1) & \text{if } y = v(0, 0) \end{cases}$$

Lemma 35. *The transformation $(\hat{\Omega}, \hat{T}, \hat{\mathbb{P}})$ is measure preserving and ergodic.*

Proof. The first half of the lemma is standard in ergodic theory (see Chapter 2.3 of [11]) so we only check the second half. Suppose A is invariant under \hat{T} . Set

$$A' = \left[\bigcup_{m=-\infty}^{\infty} T_{(0,2)}^m(\pi(A)) \right] \cap \{v(0,0) = 0\}.$$

If $\omega \in \pi(A) \cup A'$ then either $v(0,0) > 0$ and there exists y such that $(\omega, y) \in A$ or $v(0,0) = 0$ and there exists ω' and y' and m' such that $(\omega', y') \in A$ and $T_{(0,2)}^{m'}(\omega') = \omega$. It is easy to check that in either case we have $T_{(0,2)}\omega \in \pi(A) \cup A'$ and $\pi(A) \cup A'$ is invariant.

Also note that as every point in $\pi^{-1}(\pi(A))$ is in the \hat{T} -orbit of a point in A we have

$$A \subset \pi^{-1}(\pi(A)) \subset \bigcup_{m=-\infty}^{\infty} \hat{T}^m(A) = A \quad (2.20)$$

where the last containment is by the invariance of A .

Now assume that $0 < \hat{\mathbb{P}}(A)$. Then $0 < \mathbb{P}(\pi(A))$. So by the ergodicity of $T_{(0,2)}$ and the invariance of $\pi(A) \cup A'$ we have that

$$\mathbb{P}(\pi(A) \cup A') = 1.$$

Then

$$\mathbb{P}\{v(0,0) > 0\} = \mathbb{P}(\pi(A))$$

and by (2.20)

$$1 = \hat{\mathbb{P}}\{\pi^{-1}(\pi(A))\} = \hat{\mathbb{P}}(A)$$

so $\hat{\mathbb{P}}(A) = 1$ and \hat{T} is ergodic. □

Now we show that ergodicity of $(\hat{\Omega}, \hat{T}, \hat{\mathbb{P}})$ implies the ergodicity of T_{X_1} .

Lemma 36. *If an event A is invariant for T_{X_1} , then \hat{A} is invariant for \hat{T} .*

Proof. Let A be an invariant set for T_{X_1} with $\tilde{\mathbb{P}}(A) > 0$. Then for every n , $\mathbb{P}^\omega(T_{X_n} \in A) = 1$ for $\tilde{\mathbb{P}}$ a.e. $\omega \in A$ and $\mathbb{P}^\omega(T_{X_n} \in A) = 0$ for $\tilde{\mathbb{P}}$ a.e. $\omega \notin A$. Note that for $\tilde{\mathbb{P}}$ a.e. $\omega \in A$ with $v(0,0) > 0$, there exist paths $\{\gamma_1(k) \in \mathbb{Z} : k \leq N\}$ and $\{\gamma_2(k) \in \mathbb{Z} : k \leq N\}$, with γ_1, γ_2, N depending on ω , such that $\gamma_1(0) = 0$, $\gamma_2(0) = n(\omega)$, $\gamma_1(N) = \gamma_2(N)$ and

$$e^{(\gamma_i(j+1) - \gamma_i(j))} (j, \gamma_i(j)) > 0$$

for $i = 1, 2, j \leq N - 1$. Let X and \hat{X} denote two random walks on the same environment ω starting from $(0, 0)$ and $(0, n(\omega))$ respectively. This gives, for a given such ω , a natural coupling λ between the laws of $(T_{X_n}(\omega))$ and $(T_{\hat{X}_n}(\tilde{\omega}))$ such that

$$\lambda \left(T_{X_N}(\omega) = T_{\hat{X}_N}(\tilde{\omega}) \right) > 0$$

Thus, a.s. $\omega \in A$, $\mathbb{P}^{\tilde{\omega}}(T_{X_N} \in A) > 0$ implying $\tilde{\omega} \in A$. Now by the definition of \hat{A} and \hat{T} we have that \hat{A} is a.s. invariant under \hat{T} . \square

Lemma 37. $\tilde{\mathbb{P}}$ is ergodic for the Markov shift process $(T_{X_n})_{n \geq 0}$.

Proof. Let $A \subset \Omega$ be an invariant set for $(T_{X_n})_{n \geq 0}$. Then by Lemma 36 then \hat{A} is invariant for \hat{T} . As $(\hat{\Omega}, \hat{T}, \hat{\mathbb{P}})$ is ergodic we have that $\hat{\mathbb{P}}(\hat{A}) = 0$ or 1 . By Lemma 34, $\tilde{\mathbb{P}}(A)$ is also equal to 0 or 1 . \square

Acknowledgements: We wish to thank Krzysztof Burdzy for suggesting a similar problem which led to this article, and providing guidance throughout this project. S.B. gratefully acknowledges support from NSF grant number DMS-1206276. C.H. gratefully acknowledges support from the NSF and the NSA (grants number DMS-1308645 and H98230-13-1-0827).

BIBLIOGRAPHY

- [1] N. Berger, M. Biskup, C. E. Hoffman, and G. Kozma. Anomalous heat-kernel decay for random walk among bounded random conductances. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(2):374–392, 2008.
- [2] Noam Berger and Marek Biskup. Quenched invariance principle for simple random walk on percolation clusters. *Probab. Theory Related Fields*, 137(1-2):83–120, 2007.
- [3] Marek Biskup. Recent progress on the random conductance model. *Probab. Surv.*, 8:294–373, 2011.
- [4] Yves Derriennic and Michael Lin. Fractional Poisson equations and ergodic theorems for fractional coboundaries. *Israel J. Math.*, 123:93–130, 2001.
- [5] Shizuo Kakutani. Induced measure preserving transformations. *Proc. Imp. Acad. Tokyo*, 19:635–641, 1943.
- [6] Steven A. Kalikow. Generalized random walk in a random environment. *Ann. Probab.*, 9(5):753–768, 1981.
- [7] C. Kipnis and S. R. S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.*, 104(1):1–19, 1986.
- [8] Michael Maxwell and Michael Woodroffe. Central limit theorems for additive functionals of Markov chains. *Ann. Probab.*, 28(2):713–724, 2000.
- [9] Donald S. Ornstein, Daniel J. Rudolph, and Benjamin Weiss. Equivalence of measure preserving transformations. *Mem. Amer. Math. Soc.*, 37(262):xii+116, 1982.
- [10] George C. Papanicolaou and S. R. S. Varadhan. Diffusions with random coefficients. In *Statistics and probability: essays in honor of C. R. Rao*, pages 547–552. North-Holland, Amsterdam, 1982.
- [11] Karl Petersen. *Ergodic theory*, volume 2 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1989. Corrected reprint of the 1983 original.

- [12] Firas Rassoul-Agha and Timo Seppäläinen. An almost sure invariance principle for random walks in a space-time random environment. *Probab. Theory Related Fields*, 133(3):299–314, 2005.
- [13] Vladas Sidoravicius and Alain-Sol Sznitman. Quenched invariance principles for walks on clusters of percolation or among random conductances. *Probab. Theory Related Fields*, 129(2):219–244, 2004.
- [14] Ya. G. Sinai. The limit behavior of a one-dimensional random walk in a random environment. *Teor. Veroyatnost. i Primenen.*, 27(2):247–258, 1982.
- [15] Bálint Tóth. Persistent random walks in random environment. *Probab. Theory Relat. Fields*, 71(4):615–625, 1986.

Chapter 3

BROWNIAN MOTION WITH BOUNDARY DIFFUSION

There has been a prolific amount of work in the past fifty years regarding reflected Brownian motion on general domains. The most general form of reflected Brownian motion is Obliquely Reflected Brownian motion (ORBM), which involves reflection at a boundary point x along the vector $\mathbf{v}_\theta(x) = \mathbf{n}(x) + \tan \theta(x)\mathbf{t}(x)$ and has the following Skorohod representation on a C^2 domain:

$$X_t = X_0 + B_t + \int_0^t \mathbf{v}_\theta(X_s) dL_s$$

Brownian motion with Boundary diffusion (BMBD) is a similar process but with diffusion on the boundary. Intuitively, it can be thought about as a deterministic reflection at the boundary accompanied by a Gaussian noise. It has the Skorohod representation on a C^2 domain D :

$$\begin{aligned} X_t &= X_0 + B_t + \int_0^t \alpha(X_s)\mathbf{t}(X_s)dB_{L_s^*}^* + \int_0^t \beta(X_s)\mathbf{t}(X_s)dL_s^X \\ &+ \frac{1}{2} \int_0^t (1 + \alpha^2(X_s)\kappa(X_s))\mathbf{n}(X_s)dL_s^X \end{aligned} \quad (3.1)$$

where B^* is a one dimensional Brownian motion independent of B , L_t^X is the boundary local time of X on D and $\kappa : \partial D \rightarrow \mathbb{R}$ is the curvature of ∂D . Here $\alpha : \partial D \rightarrow \mathbb{R}$ is called the (boundary) *diffusion* and $\beta : \partial D \rightarrow \mathbb{R}$ is called the (boundary) *drift*. From here onwards, we denote this process by $BMBD(\alpha, \beta)$.

Unlike ORBM, BMBD is trickier to handle as it is not conformally invariant in the sense that the diffusion and drift terms change, but it still remains a BMBD with these changed diffusion and drift functions. Ryan Card [1], in his thesis, showed that if $f : D \rightarrow \tilde{D}$ is a conformal map between C^2 domains, then the process defined by $Y_t = X(c^{-1}(t))$, for

$c(t) = \int_0^t |f'(X_s)|^2 ds$, is a BMBD with diffusion function:

$$\tilde{\alpha}(f(x)) = \alpha(x)\sqrt{|f'(x)|} \quad (3.2)$$

and drift function:

$$\tilde{\beta}(f(x)) = \beta(x) + \frac{1}{2}\tilde{\alpha}^2(f(x))\frac{\tilde{\mathbf{t}} \cdot \Delta_{\partial D} f(x)}{|f'(x)|^2} \quad (3.3)$$

for $x \in \partial D$, where $\Delta_{\partial D} f = (\partial_{\mathbf{tt}} \text{Re}f, \partial_{\mathbf{tt}} \text{Im}f)$ represents the Laplacian of f along the boundary of D and $\tilde{\mathbf{t}}$ represents the unit tangent at $f(x) \in \partial \tilde{D}$.

1 Stationary Distribution

As BMBD on a bounded domain is a compactly supported measure, there exists a stationary distribution on D . Furthermore, it can be shown that it is unique. In this section, we investigate the stationary distribution of BMBD and see that the stationary density is the solution of a specific boundary value problem.

Theorem 11. *The (unique) stationary distribution of BMBD(α, β) on a bounded C^2 domain D with $\beta \in C^1(\overline{D})$ and $\alpha \in C^2(\overline{D})$ has a density in $C^2(\overline{D})$ iff there exists a non-negative function $h \in C^2(\overline{D})$ with $\|h\|_1 = 1$ which solves the boundary value problem*

$$\begin{aligned} \Delta h &= 0 \quad \text{on } D \\ \partial_{\mathbf{n}} h + \partial_{\mathbf{tt}}(\alpha^2 h) - 2\partial_{\mathbf{t}}(\beta h) &= 0 \quad \text{on } \partial D \end{aligned} \quad (3.4)$$

In this case, the solution is unique and gives the stationary density.

Proof: Let f be any C^2 function on \overline{D} . Then by the Ito Formula,

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t \nabla f(X_s) \cdot dB_s + \int_0^t \alpha(X_s) \nabla f(X_s) \cdot \mathbf{t}(X_s) dB_{L_s^X}^* \\ &+ \int_0^t \beta(X_s) \nabla f(X_s) \cdot \mathbf{t}(X_s) dL_s^X + \frac{1}{2} \int_0^t (1 + \alpha^2(X_s) \kappa(X_s)) \nabla f(X_s) \cdot \mathbf{n}(X_s) dL_s^X \\ &+ \frac{1}{2} \int_0^t \Delta f(X_s) ds + \frac{1}{2} \int_0^t \alpha^2(X_s) \Delta_{\partial D} f(X_s) dL_s^X \end{aligned}$$

Now, choosing f in $C_c^2(D)$, and taking expectation under the stationary distribution, we see that the boundary terms and martingale terms disappear and we get

$$\int_D h \Delta f = 0$$

By Green's second identity, this yields us that $\int_D f \Delta h = 0$ for all such f and consequently, $\Delta h = 0$ on D .

Now, for any $f \in C^2(\bar{D})$, we compute the expectation under the stationary measure Q .

$$\begin{aligned} E^Q f(X_t) &= E^Q f(X_0) + \frac{t}{2} \int_D \Delta f(x) h(x) dx \\ &+ E^Q \left[\frac{1}{2} \int_0^t (1 + \alpha^2(X_s) \kappa(X_s)) \partial_{\mathbf{n}} f(X_s) dL_s^X + \int_0^t \beta(X_s) \nabla f(X_s) \cdot \mathbf{t}(X_s) dL_s^X \right. \\ &\left. + \frac{1}{2} \int_0^t \alpha^2(X_s) \Delta_{\partial D} f(X_s) dL_s^X \right] \end{aligned} \quad (3.5)$$

By stationarity, $E^Q f(X_t) = E^Q f(X_0)$. By Green's second identity and the fact that h is harmonic in D ,

$$\int_D \Delta f(x) h(x) dx = \int_{\partial D} f(x) \partial_{\mathbf{n}} h(x) dx - \int_{\partial D} h(x) \partial_{\mathbf{n}} f(x) dx$$

Recall that the Revuz measure of L^X is a measure ν that satisfies

$$E^Q \int_0^1 f(X_s) dL_s^X = \int_{\partial D} f(x) \nu(dx)$$

Following the recent work of Burdzy, Chen, Marshall and Ramanan [2], we know that the Revuz measure of L^X is $h(x) dx$ restricted to ∂D . We give an outline of the proof. Let g is continuous on ∂D . Its harmonic extension to \bar{D} (also denoted by g) is continuous on \bar{D} .

Now, for any $\epsilon \in (0, 1)$,

$$E^Q \int_0^1 1(1 - \epsilon < |X_s| < 1) g(X_s) ds = E^Q \int_{\partial D} 1(1 - \epsilon < |z| < 1) g(z) h(z) dz$$

and therefore,

$$E^Q \int_0^1 \frac{1}{\epsilon} 1(1 - \epsilon < |X_s| < 1) g(X_s) ds = E^Q \int_{\partial D} \frac{1}{\epsilon} 1(1 - \epsilon < |z| < 1) g(z) h(z) dz$$

Now, by the continuity and boundedness of g and h , the limit on the right hand side exists and equals $\int_{\partial D} g(x)h(x)dx$. Now, by the definition of local time, $\int_0^1 \frac{1}{\epsilon} 1(1 - \epsilon < |X_s| < 1)g(X_s)ds$ converges to $\int_0^1 g(X_s)dL_s^X$. Now, one can refer to the the aforementioned paper for an argument of the fact that the family

$$\left\{ \int_0^1 \frac{1}{\epsilon} 1(1 - \epsilon < |X_s| < 1)g(X_s)ds, \epsilon \in \left(0, \frac{1}{2}\right) \right\}$$

is uniformly integrable. The result then follows by standard arguments. This result yields us the following:

$$\begin{aligned} E^Q \left[\frac{1}{2} \int_0^1 (1 + \alpha^2(X_s)\kappa(X_s))\partial_{\mathbf{n}}f(X_s)dL_s^X + \int_0^1 \beta(X_s)\nabla f(X_s)\cdot\mathbf{t}(X_s)dL_s^X + \frac{1}{2} \int_0^1 \alpha^2(X_s)\Delta_{\partial D}f(X_s)dL_s^X \right] \\ = \frac{1}{2} \int_{\partial D} (1 + \alpha^2(x)\kappa(x))\partial_{\mathbf{n}}f(x)h(x)dx + \int_{\partial D} \beta(x)\partial_{\mathbf{t}}f(x)h(x)dx \\ + \frac{1}{2} \int_{\partial D} \alpha^2(x)\Delta_{\partial D}f(x)h(x)dx \end{aligned}$$

Now, we want to compute the quantity $\Delta_{\partial D}f(x)$. note that we can locally parametrize the boundary in a small neighbourhood of a boundary point x by a curve $\gamma : (-\epsilon, \epsilon) \rightarrow \partial D$ such that $\gamma(0) = x$ such that $\mathbf{t}(\gamma(s)) = \gamma'(s)$ (parametrized by arclength). Then it is easy to check that $\gamma''(s) = \kappa(\gamma(s))\mathbf{n}(\gamma(s))$. Thus,

$$\Delta_{\partial D}f(\gamma(s)) = \left(\frac{d}{ds} \nabla f(\gamma(s)) \right) \cdot \gamma'(s) = \partial_{\mathbf{tt}}f - \nabla f(\gamma(s)) \cdot \gamma''(s) = \partial_{\mathbf{tt}}f(\gamma(s)) - \kappa(\gamma(s))\partial_{\mathbf{n}}f(\gamma(s))$$

Putting $s = 0$, we get

$$\Delta_{\partial D}f(x) = \partial_{\mathbf{tt}}f(x) - \kappa(x)\partial_{\mathbf{n}}f(x) \tag{3.6}$$

With all these in hand, we rewrite 3.5 as follows:

$$\begin{aligned} 0 &= \int_{\partial D} f(x)\partial_{\mathbf{n}}h(x)dx - \int_{\partial D} h(x)\partial_{\mathbf{n}}f(x)dx + \int_{\partial D} (1 + \alpha^2(x)\kappa(x))\partial_{\mathbf{n}}f(x)h(x)dx \\ &+ 2 \int_{\partial D} \beta(x)\partial_{\mathbf{t}}f(x)h(x)dx + \int_{\partial D} \alpha^2(x) (\partial_{\mathbf{tt}}f(x) - \kappa(x)\partial_{\mathbf{n}}f(x)) h(x)dx \\ &= \int_{\partial D} f(x)\partial_{\mathbf{n}}h(x)dx + 2 \int_{\partial D} \beta(x)\partial_{\mathbf{t}}f(x)h(x)dx + \int_{\partial D} \alpha^2(x)\partial_{\mathbf{tt}}f(x)h(x)dx \end{aligned}$$

Now by applying integration by parts on the third term, we get the following:

$$\int_{\partial D} f(x)\partial_{\mathbf{n}}h(x)dx - 2 \int_{\partial D} f(x)\partial_{\mathbf{t}}(\beta h)(x)dx + \int_{\partial D} f(x)\partial_{\mathbf{tt}}(\alpha^2 h)(x)dx = 0$$

for all f in $C^2(\overline{D})$ which in turn yields the result.

The uniqueness of the solution follows from the uniqueness of the stationary distribution. \square

The above theorem yields the following facts:

Corollary 11.1. (i) *If $\beta = 0$, then the stationary distribution is uniform if and only if α is a constant.*

(ii) *For any diffusion function α , then the stationary distribution is uniform if and only if $\beta = c + \alpha \partial_t \alpha$ for some constant c .*

Proof: (i) We plug in $\beta = 0$ and $h = \frac{1}{\text{Area}(D)}$ in (3.4), we see that α satisfies $\partial_{tt}\alpha^2 = 0$. So, $\partial_t \alpha^2 = c$ for some constant c . As $\partial_t \alpha^2$ is conservative (its integral on closed loops is zero), so $c = 0$. Thus α is a constant.

(ii) Plugging in $h = \frac{1}{\text{Area}(D)}$ in (3.4), we get

$$\partial_{tt}\alpha^2 - 2\partial_t\beta = 0$$

This gives the result.

Example: The stationary distribution of $BMBD(\sin^2 \theta, \sin^2 \theta \sin 2\theta)$ on the unit disk $\overline{\mathbb{D}}$ is uniform.

In what follows, we apply a sequence of conformal maps to map the unit disk \mathbb{D} onto different domains. This is helpful because the stationary distribution remains invariant under conformal maps in the sense that if $f : D \rightarrow \tilde{D}$ is a conformal map and h and h' be the densities corresponding to the stationary distributions of X_t and $Y_{c^{-1}(t)}$, then $h(z) = \tilde{h}(f(z))$ for all $z \in D$. This can be easily seen from the fact that if we take a small ball around z in D , its image in \tilde{D} is dilated approximately by a factor of $|f'(z)|^2$ (the Jacobian). Also the time change happens by the same factor. So they cancel out each other in the expected average time spent in the respective (infinitesimal) sets. Before we proceed, we would like

to point out a special case of the formulas (3.2) and (3.3) which we would use time and again.

Lemma 38. (i) *Assume α is never 0. If $f : \mathbb{H} \rightarrow D$ is a conformal map that satisfies $f'(x) = \frac{1}{\alpha^2(x)}$ for $x \in \partial\mathbb{H}$, and X is a BMBD(α, β) on \mathbb{H} , then the process $Y_t = f(X(c^{-1}(t)))$ is a BMBD($1, \tilde{\beta}$) on D , where $\tilde{\beta}$ satisfies (3.3).*

(ii) *If $f : \mathbb{H} \rightarrow \mathbb{D}$ is the conformal map $f(z) = e^{iz}$, and X is a BMBD($\alpha \circ f, \beta \circ f$), then the process $Y_t = f(X(c^{-1}(t)))$ is a BMBD(α, β) on $\overline{\mathbb{D}}$.*

BMBD on the disk \mathbb{D} : An inversion formula

We consider $BMBD(\alpha, 0)$ on the disk \mathbb{D} , where $\alpha : \partial\mathbb{D} \rightarrow \mathbb{R}^+$ is a C^2 function. In polar co-ordinates (r, θ) , the boundary condition in (3.4) takes the following form:

$$\frac{\partial h}{\partial r} = \frac{\partial^2}{\partial \theta^2}(\alpha^2 h) \quad \text{on } \partial\mathbb{D}$$

We know that if we can solve (3.4) for a non-negative function $h \in C^2(\overline{\mathbb{D}})$ with $\|h\|_1 = 1$, we have a stationary density. Now we ask about the inverse problem, that is, given such a function h , can we solve for α . It turns out that for any such positive function h , we can solve for function α , but not uniquely. What we need in addition for the uniqueness is a real number σ_0 which has the following interpretation:

$$\sigma_0 = \frac{1}{2\pi} \int_0^{2\pi} \alpha^2(\theta) h(\theta) d\theta$$

(Here and in what follows, by abuse of notation, we denote the boundary values $h(1, \theta)$ by $h(\theta)$). We make this precise in the following theorem:

Theorem 12. *For every positive function $h \in C^2(\overline{\mathbb{D}})$ with $\|h\|_1 = 1$ satisfying (3.4) and any real number*

$$\sigma_0 > \sup_{0 \leq \theta < 2\pi} \int_0^1 \frac{1}{s} (h(s, \theta) - h(0)) ds$$

there exists a unique choice of $\alpha \in C^2(\partial\mathbb{D})$ for which $BMBD(\alpha, 0)$ has a stationary distribution whose density is given by h , and it is given by

$$\alpha(\theta) = \left(\frac{\sigma_0 - \int_0^1 \frac{1}{s} (h(s, \theta) - h(0)) ds}{h(\theta)} \right)^{\frac{1}{2}} \quad (3.7)$$

For this α and h ,

$$\sigma_0 = \frac{1}{2\pi} \int_0^{2\pi} \alpha^2(\theta) h(\theta) d\theta \quad (3.8)$$

Note: The constraint in the range of values of σ_0 is only to assure that the quantity under the square root in (3.7) is positive. Also note that the fact $\|h\|_1 = 1$ forces $h(0) = \frac{1}{\pi}$.

Proof. Let \tilde{h} be a harmonic conjugate of h . As $h \in C^2(\overline{\mathbb{D}})$, the Cauchy-Riemann equations remain valid at the boundary. Using them, we get

$$\frac{\partial \tilde{h}}{\partial \theta} = \frac{\partial^2}{\partial \theta^2}(\alpha^2 h) \quad \text{on } \partial \mathbb{D}$$

Thus, there is a choice of \tilde{h} (uniquely given by $\tilde{h}(0) = 0$) such that

$$\tilde{h} = \frac{\partial}{\partial \theta}(\alpha^2 h) \quad \text{on } \partial \mathbb{D} \tag{3.9}$$

holds. Let

$$P(r, \theta) = \frac{1}{2\pi} \operatorname{Re} \left(\frac{1 + e^{i\theta}}{1 - e^{i\theta}} \right) = \frac{1}{2\pi} \frac{1 - r^2}{1 - 2r \cos \theta + r^2}$$

represent the Poisson kernel on the disk \mathbb{D} and let

$$Q(r, \theta) = \frac{1}{2\pi} \operatorname{Im} \left(\frac{1 + e^{i\theta}}{1 - e^{i\theta}} \right)$$

Taking harmonic extensions of both sides of (3.9) onto the disk \mathbb{D} , we get

$$\tilde{h}(r, \theta) = \int_0^{2\pi} P(r, \theta - \eta) \frac{\partial}{\partial \eta}(\alpha^2 h)(\eta) d\eta$$

Thus

$$\begin{aligned} h(r, \theta) &= h(0) - \int_0^{2\pi} Q(r, \theta - \eta) \frac{\partial}{\partial \eta}(\alpha^2 h)(\eta) d\eta \\ &= h(0) + \int_0^{2\pi} \frac{\partial}{\partial \eta} Q(r, \theta - \eta) (\alpha^2 h)(\eta) d\eta \\ &= h(0) - \int_0^{2\pi} \frac{\partial}{\partial \theta} Q(r, \theta - \eta) (\alpha^2 h)(\eta) d\eta \\ &= h(0) - \int_0^{2\pi} r \frac{\partial}{\partial r} P(r, \theta - \eta) (\alpha^2 h)(\eta) d\eta \\ &= h(0) - r \frac{\partial}{\partial r} \int_0^{2\pi} P(r, \theta - \eta) (\alpha^2 h)(\eta) d\eta \\ &= h(0) - r \frac{\partial \Phi}{\partial r} \end{aligned}$$

where Φ is the harmonic extension of $\alpha^2 h$ onto the disk \mathbb{D} . Thus, we get

$$\frac{\partial \Phi}{\partial r} = -\frac{1}{r}(h(r, \theta) - h(0))$$

Integrating both sides, we get

$$\Phi(r, \theta) - \Phi(0) = - \int_0^r \frac{1}{s} (h(s, \theta) - h(0)) ds$$

As $\alpha^2 h$ is continuous on $\overline{\mathbb{D}}$, we can take a limit as $r \rightarrow 1$ to get

$$\alpha^2(\theta)h(\theta) = \Phi(0) - \int_0^1 \frac{1}{s}(h(s, \theta) - h(0))ds$$

which proves the result upon noting that, by harmonicity of Φ ,

$$\Phi(0) = \frac{1}{2\pi} \int_0^{2\pi} \alpha^2(\theta)h(\theta)d\theta = \sigma_0$$

□

Another proof: Let us consider the Fourier series

$$h(\theta) = \sum_{n=0}^{\infty} a_n e^{in\theta}$$

and

$$\alpha^2(\theta)h(\theta) = \sum_{n=0}^{\infty} b_n e^{in\theta}$$

Then h can be represented as

$$h(r, \theta) = \sum_{n=0}^{\infty} a_n r^{|n|} e^{in\theta}$$

and (3.4) yields

$$a_n + |n|b_n = 0$$

for every $n \neq 0$, which yields us the result upon noting that $\sigma_0 = b_0 = \frac{1}{2\pi} \int_0^{2\pi} \alpha^2(\theta)h(\theta)d\theta$ by definition of Fourier coefficients.

Example: If $h(r, \theta) = \frac{4 + r \cos \theta}{4\pi}$, then we can check that h satisfies all the conditions of the previous theorem. For this h and any $\sigma_0 > 1/(4\pi)$, the inversion formula yields

$$\alpha(\theta) = \sqrt{\frac{\sigma_0 - \cos \theta}{4 + \cos \theta}}$$

Relation to Obliquely Reflected Brownian motion

Consider $BMBD(1, \beta)$ on the half plane given by the Skorohod equations:

$$\hat{X}_t = \hat{X}_0 + \hat{B}_t^{(1)} + \hat{B}_{\hat{L}_t}^* + \int_0^t \beta(\hat{X}_s) d\hat{L}_s \quad (3.10)$$

$$\hat{Y}_t = \hat{Y}_0 + \hat{B}_t^{(2)} + \frac{1}{2} \hat{L}_t. \quad (3.11)$$

Note that the pair (\hat{Y}, \hat{L}) are independent of β and they are determined by the Brownian motion $\hat{B}^{(2)}$ by the Skorohod map. Let $G : \mathbb{H} \rightarrow \mathbb{D}$ be the conformal map $G(z) = e^{iz}$. Let $Z_t = G(\hat{X}(c^{-1}(t)) + i\hat{Y}(c^{-1}(t)))$ where $c(t) = \int_0^t |G'(\hat{X}(s) + i\hat{Y}(s))|^2 ds = \int_0^t e^{-2\hat{Y}_s} ds$. Then we know by Lemma 38 that Z_t is $BMBD(1, \beta)$ on $\overline{\mathbb{D}}$ with local time given by $L(t) = \hat{L}(c(t))$. This representation yields us the following interesting facts:

(i) When β is a constant, then we get $Z_t = \tilde{Z}_t e^{i\hat{B}_{L_t}^*}$, where \tilde{Z} is an $ORBM$ with constant reflection angle $\theta = \arctan \beta$. Note that this clarifies our intuition that the boundary diffusion *smooths things out* near the boundary.

A special choice of filtrations

Note that in the Skorohod representation of $BMBD$ in the half plane, the y coordinate is just a reflected Brownian motion and the x coordinate depends on \hat{Y}_t only through its local time, so a valid construction of the process is to *run the entire process \hat{Y} and construct the process \hat{X} conditioned on the process \hat{Y}* . The filtration which makes this possible is the following:

$$\mathcal{F}_0 = \sigma(\hat{B}^{(2)}) ; \mathcal{F}_t = \sigma(\hat{B}^{(2)}; \hat{B}_s^{(1)}, 0 \leq s \leq t; \hat{B}_s^*, 0 \leq s \leq \hat{L}_t) \quad (3.12)$$

for $t > 0$. We will be working with this filtration in what follows.

Investigating Polarity of Points

We know that for RBM in the half plane, and a large class of $ORBM$ s, the points are polar sets. There is a vast literature on $ORBM$'s, see [3], that deals with the relationship between the angle of reflection θ and polarity of points. [3] devise an integral test towards this end. We show here that for $BMBD$, under very general conditions, any given point is visited with positive probability by the process. This is a stark difference from $ORBM$.

First we deal with $BMBD(1, 0)$. Let $S(t) = t + L_t$. With respect to the filtration \mathcal{F}

described above and \mathcal{G} generated by $B^{(2)}$, we can write the co-ordinate processes as follows (here we suppress the hats for notational convenience):

$$X_t = X_0 + W_{t+L_t} \quad (3.13)$$

$$Y_t = Y_0 + B_t^{(2)} + \frac{1}{2}L_t \quad (3.14)$$

where $(W_t, \mathcal{F}_{S^{-1}(t)})$ is a Brownian motion and is independent of $B^{(2)}$ as $\mathcal{G}_\infty = \mathcal{F}_0$. So, if we consider the time change, $t \rightarrow S^{-1}(t)$, then $\tilde{X}_t = X_{S^{-1}(t)} = X_0 + W_t$ is a Brownian motion and $\tilde{Y}_t = Y_{S^{-1}(t)}$ is a sticky Brownian motion independent of \tilde{X} . Let $\tau = \inf\{t \geq 1 : \tilde{X}_t = 0\}$. Then τ is a stopping time with respect to the filtration \mathcal{F} and is independent of \tilde{Y} . Now, for a sticky Brownian motion, it is known that $P(\tilde{Y}_t = 0) > 0$ for every t (look at Theorem 18 for example). Thus

$$P((\tilde{X}_t, \tilde{Y}_t) = (0, 0) \text{ for some } t > 0) \geq P(\tilde{Y}_\tau = 0) = \int_1^\infty P(\tilde{Y}_t = 0)g(t)dt > 0$$

where g is the density corresponding to the distribution of τ . Another way to see this could be using the fact that g is strictly positive on $(1, \infty)$ and $\int_1^M P(\tilde{Y}_t = 0)g(t)dt > 0$ for any $M > 1$ from the definition of sticky Brownian motion.

Now we turn to $BMBD(1, \beta)$. Our strategy is to *kill the drift* by devising a version of the *Girsanov Theorem*. We recall that B_L^* is a (time-changed) Brownian motion under the described filtration \mathcal{F} . Define the random function $Z_t = \exp \left\{ - \int_0^t \beta(X_s)dB_{L_s}^* - \frac{1}{2} \int_0^t \beta^2(X_s)dL_s \right\}$. Then under very general conditions, Z is a *martingale* with respect to the filtration \mathcal{F} . Then we can define an absolutely continuous change of measure on \mathcal{F}_T for any $0 \leq T < \infty$ defined by

$$\bar{P}_T(A) = E(1_A Z_T) \quad A \in \mathcal{F}_T. \quad (3.15)$$

Our main tool is the following theorem whose proof is a slight modification of the proof found in Karatzas and Shreve:

Theorem 13. *Fix $0 \leq T < \infty$ and assume that Z is a martingale. If M is a continuous local martingale on $0 \leq T < \infty$ under P , then the process*

$$\bar{M}_t = M_t + \int_0^t \beta(X_s)d\langle M, B_L^* \rangle_s \quad 0 \leq T < \infty$$

is a continuous local martingale under the changed measure \bar{P}_T , and if N is another continuous local martingale on $0 \leq T < \infty$ under P and

$$\bar{N}_t = N_t + \int_0^t \beta(X_s) d\langle N, B_{L_s}^* \rangle_s \quad 0 \leq T < \infty$$

then $\langle \bar{M}, \bar{N} \rangle_t = \langle M, N \rangle_t$ $0 \leq T < \infty$ a.s. P and \bar{P}_T . (Here the cross variations are computed under the appropriate measures).

Applying the theorem, we see that under this changed measure \bar{P}_T , $\bar{B}_t = B_{L_t}^* + \int_0^t \beta(X_s) dL_s$ is a local martingale and $\langle \bar{B} \rangle_t = \langle B^* \rangle_{L_t} = L_t$. Also, as $\langle B^{(1)}, B_{L_t}^* \rangle_t = \langle B^{(1)}, \bar{B} \rangle_t = 0$ (the cross variations being computed under P and \bar{P}_T respectively) for all t , so by Knight's Theorem, the processes $W_t^{(1)} = \bar{B}_{\sigma_t}$ and $B_t^{(1)}$ are independent Brownian motions under \bar{P}_T (here σ denotes the inverse local time). By the given SDE for \bar{B} , we can write $\bar{B}_t = W_{L_t}^{(1)}$. Also note that as $\mathcal{G}_\infty = \mathcal{F}_0$, so each of these changed measures leaves the measure of sets in \mathcal{G}_∞ unchanged. So, $B^{(2)}$ remains a Brownian motion under this changed measure. Thus, under \bar{P}_T ,

$$X_t = X_0 + W_{L_t}^{(1)} + B_t^{(1)} \quad (3.16)$$

$$Y_t = Y_0 + B_t^{(2)} + \frac{1}{2}L_t. \quad (3.17)$$

Thus carrying on exactly the same argument, under \bar{P}_T , $\tilde{X}_t = X_{S^{-1}(t)}$ is a Brownian motion and $\tilde{Y}_t = Y_{S^{-1}(t)}$ is a sticky Brownian motion independent of \tilde{X} . As the change of measure holds upto every fixed T , we now work with the stopping time $\tau \wedge 2$ where τ is the first hitting time of 0 after time 1 of the process \tilde{X} . Note that as $S^{-1}(t) \leq t$ for every t , therefore $\mathcal{F}_{S^{-1}(2)} \subseteq \mathcal{F}_2$. Thus the event of hitting $(0, 0)$ of the process (\tilde{X}, \tilde{Y}) in $0 < t \leq 2$ is in \mathcal{F}_2 and we show that its probability under \bar{P}_2 is positive.

$$\bar{P}_2((\tilde{X}_t, \tilde{Y}_t) = (0, 0) \text{ for some } 0 < t \leq 2) \geq \bar{P}_2(\tilde{Y}_{\tau \wedge 2} = 0, \tau \wedge 2 < 2) = \int_1^2 \bar{P}_2(\tilde{Y}_t = 0)g(t)dt > 0,$$

where we use $\bar{P}_2(\tilde{Y}_t = 0) > 0$ for $t \in [1, 2]$ as the change of measure restricted to \mathcal{F}_2 is absolutely continuous. Again, by absolute continuity, we conclude that

$$P((\tilde{X}_t, \tilde{Y}_t) = (0, 0) \text{ for some } 0 < t \leq 2) > 0.$$

Thus, we have the following theorem:

Theorem 14. Consider a $BMBD(1, \beta)$ on the half plane, where β is such that the process Z defined by

$$Z_t = \exp \left\{ - \int_0^t \beta(X_s) dB_{L_s}^* - \frac{1}{2} \int_0^t \beta^2(X_s) dL_s \right\}$$

is a martingale with respect to the filtration \mathcal{F} given in (3.12). Then any given boundary point is visited with positive probability.

Intuitively, the above result can be interpreted as follows. The trace process for the ORBM is the Cauchy process with drift and this is a 1-stable process. When we add a boundary diffusion term, we add a Brownian motion (which is a 2-stable process) to the trace process. Thus on a small time scale, the properties of the Brownian motion part are more prominent, which is indicated by the positive hitting probability of boundary points. But on a larger time scale, the properties of the Cauchy process with drift are more prevalent, which is indicated by part (i) and (ii) of the following theorem, closely following [2]. Part (iii) yields a physical interpretation to the constant σ_0 appearing in theorem 12. Before we present the theorem, we briefly describe some quantities needed to state it.

Let Z be a $BMBD(\alpha, \beta)$ on the disk. Let \mathcal{E}_t be the collection of all excursions e_s with $s < t$. We say that e_s belongs to the family \mathcal{E}_t^L of excursions with *large winding number* if $|\arg Z_s - \arg Z_{s+\zeta(e_s)-}| > 2\pi$ and $s < t$. Here X_{u-} denotes the left-hand limit and $\zeta(e_t)$ denotes the lifetime of the excursion e_t . Define

$$\arg^* Z_t = \arg Z_t - \sum_{s: e_s \in \mathcal{E}_t^L} (\arg Z_{s+\zeta(e_s)-} - \arg Z_s). \quad (3.18)$$

Now, we are ready to state the theorem.

Theorem 15. Let Z be a $BMBD(\alpha, \beta)$ on the disk \mathbb{D} and let h denote its stationary density on $\overline{\mathbb{D}}$, assuming $h \in C^2(\overline{\mathbb{D}})$, and θ parametrize its boundary. Then,

(i) For any $z \in \overline{\mathbb{D}}$, the distribution of

$$\frac{\arg Z_t}{t} - \mu_0$$

converges the Cauchy distribution under P_z .

(ii) For any $z \in \overline{\mathbb{D}}$, P_z almost surely

$$\lim_{t \rightarrow \infty} \frac{\arg^* Z_t}{t} = \int_0^{2\pi} \beta(\theta) h(\theta) d\theta = \mu_0 \text{ (say)}. \quad (3.19)$$

(iii) For each t , denoting a partition of $[0, t]$ by $\pi = \{\pi(0), \dots, \pi(n)\}$ and its mesh by $\|\pi\|$, the following limit exists in probability

$$Q_t(\arg^* Z) = \lim_{\|\pi\| \rightarrow 0} \sum_{i=0}^n (\arg^* Z_{\sigma(\pi(i+1))} - \arg^* Z_{\sigma(\pi(i))})^2$$

where σ represents the inverse local time at the boundary. Furthermore, P_z almost surely

$$\lim_{t \rightarrow \infty} \frac{Q_t(\arg^* Z)}{t} = 4 + \int_0^{2\pi} \alpha^2(\theta) h(\theta) d\theta. \quad (3.20)$$

Proof: The proofs of (i) and (ii) are very similar to that Theorem 3.2 (ii) and (iii) in [2] with the presence of an extra diffusion term. We briefly outline them.

(i) Let $f(z) = e^{iz}$. Let $\hat{Z} = \hat{X} + i\hat{Y}$ be a $BMBD(\alpha \circ f, \beta \circ f)$ on the half plane \mathbb{H} . Then by Lemma 38 $Z_t = f(\hat{Z}(c^{-1}(t)))$ is a $BMBD(\alpha, \beta)$ on $\overline{\mathbb{D}}$ with local time given by $L_t = \hat{L}_{c^{-1}(t)}$ for all t . From this we see that the processes $A_t = \arg Z_{\sigma(t)}$ and $\hat{A}_t = \hat{X}_{\hat{\sigma}(t)}$ are indistinguishable processes. Here σ and $\hat{\sigma}$ denote the respective inverse local times.

Note that $\bar{Z}_t = \hat{Z}_t - \int_0^t \beta(f(\hat{X}_s)) d\hat{L}_s - \int_0^t \alpha(f(\hat{X}_s)) d\hat{B}_{\hat{L}_t}^*$ is a normally reflected Brownian motion. Thus the process

$$\bar{X}_{\hat{\sigma}(t)} = \hat{X}_{\hat{\sigma}(t)} - \int_0^{\hat{\sigma}(t)} \beta(f(\hat{X}_s)) d\hat{L}_s - \int_0^{\hat{\sigma}(t)} \alpha(f(\hat{X}_s)) d\hat{B}_{\hat{L}_t}^*$$

is a Cauchy process C_t . Thus

$$\begin{aligned} A_t &= C_t + \int_0^{\hat{\sigma}(t)} \beta(f(\hat{X}_s)) d\hat{L}_s + \int_0^{\hat{\sigma}(t)} \alpha(f(\hat{X}_s)) d\hat{B}_{\hat{L}_t}^* \\ &= C_t + \int_0^{\sigma(t)} \beta(Z_s) dL_s + \int_0^{\sigma(t)} \alpha(Z_s) d\hat{B}_{L_s}^* \end{aligned} \quad (3.21)$$

(as $\hat{\sigma}(t) = c^{-1}(\sigma(t))$).

Define random times $T_t = \sup\{u \leq t : Z_u \in \partial\mathbb{D}\}$ (with convention $T_t = 0$ if $Z_u \notin \partial\mathbb{D}$ for every $u \leq t$) and $S = \inf\{u \geq 0 : Z_u \in \partial\mathbb{D}\}$. Then we can write

$$\begin{aligned} \frac{1}{t} \arg Z_t - \mu_0 &= \frac{1}{t} C_{L_t} + \left(\frac{1}{t} \int_0^t \beta(Z_s) dL_s - \mu_0 \right) + \frac{1}{t} (\arg Z_t - \arg Z_{T_t}) + \frac{1}{t} \arg Z_S + \frac{1}{t} \int_0^t \alpha(Z_s) d\hat{B}_{L_s}^* \\ &= \frac{1}{t} C_t + \frac{1}{t} (C_t - C_{L_t}) + \left(\frac{1}{t} \int_0^t \beta(Z_s) dL_s - \mu_0 \right) + \frac{1}{t} (\arg Z_t - \arg Z_{T_t}) + \frac{1}{t} \arg Z_S \\ &\quad + \frac{1}{t} \int_0^t \alpha(Z_s) d\hat{B}_{L_s}^*. \end{aligned}$$

Now, the distribution of $(|Z|, L)$ is independent of (α, β) . This is because in the half plane representation, the distribution of \hat{Y} is independent of (α, β) and $|Z_t| = e^{-\hat{Y}_t}$. Thus, using the limit-quotient theorem for additive functionals, we get that for every z , P_z almost surely,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \beta(Z_s) dL_s = \mu_0 \quad \text{and} \quad \lim_{t \rightarrow \infty} \frac{L_t}{t} = 1. \quad (3.22)$$

Fix any small $\epsilon > 0$ and any $z \in \overline{\mathbb{D}}$ and let

$$p(t) = P_z(|\arg Z_t - \arg Z_{T_t}| > \epsilon t).$$

Let $T'_u = \inf\{t \geq u : Z_t \in \partial\mathbb{D}\}$. Then by the Markov property applied at time t and symmetry of Brownian motion

$$P_z(|\arg Z_{T'_t} - \arg Z_{T_t}| > \epsilon t) \geq p_1(t)/2.$$

Now, using the fact that for every fixed u , the Cauchy process is a.s. continuous at time u , and by scaling, we get a $\delta > 0$ such that

$$P \left(\sup_{(1-\delta)t \leq u, v \leq (1+\delta)t} |C_u - C_v| \geq \epsilon t / 2 \right) < \epsilon.$$

Also, by (3.22), for sufficiently large t ,

$$P_z(L_t \in ((1-\delta)t, (1+\delta)t)) \geq 1 - \epsilon.$$

As the last two integrals in (3.21) are continuous in t , the jumps of A and C occur simultaneously. Using this fact and the above, one can show that $p(t) < 4\epsilon$. Thus we have

$$P_z \left(\frac{1}{t} |\arg Z_t - \arg Z_{T_t}| > \epsilon \right) < 4\epsilon$$

and similarly

$$P_z \left(\frac{1}{t} |C_{L_t} - C_t| > \epsilon \right) < 2\epsilon$$

for sufficiently large t .

Trivially we have a.s.

$$\lim_{t \rightarrow \infty} \frac{1}{t} \arg Z_S = 0.$$

Next, to handle the diffusion term, note that we can get a Brownian motion W such that

$$\int_0^t \alpha(Z_s) d\hat{B}_{L_s}^* = W_{\int_0^t \alpha^2(Z_s) dL_s}.$$

As by the limit-quotient theorem again

$$\frac{1}{t} \int_0^t \alpha^2(Z_s) dL_s = \int_0^{2\pi} \alpha^2(\theta) h(\theta) d\theta.$$

Thus assuming α is bounded, we get by an argument similar to showing $\lim_{t \rightarrow \infty} \frac{W_t}{t} = 0$ a.s., we get, P_z a.s.

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \alpha(Z_s) d\hat{B}_{L_s}^* = 0.$$

As $\frac{C_t}{t}$ follows Cauchy distribution and the remaining terms go to zero in probability, we get (i).

(ii) If we remove the jumps of size greater than 2π from the Cauchy process C , we get a process C^* which is a zero mean martingale and a Levy process. Hence, by the law of large numbers, almost surely

$$\lim_{t \rightarrow \infty} \frac{C_t^*}{t} = 0 \tag{3.23}$$

Thus

$$\frac{1}{t} \arg^* Z_{\sigma(t)} = \frac{1}{t} C_t^* + \frac{1}{t} \int_0^{\sigma(t)} \beta(Z_s) dL_s + \frac{1}{t} \arg Z_S + \frac{1}{t} \int_0^{\sigma(t)} \alpha(Z_s) d\hat{B}_{L_s}^*.$$

As from (3.22), P_z almost surely,

$$\lim_{t \rightarrow \infty} \frac{\sigma(t)}{t} = 1$$

so, the second term above approaches μ_0 and the fourth term approaches 0 almost surely and by (3.23), the first term approaches 0 and trivially, the third term approaches 0. So, P_z almost surely,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \arg^* Z_{\sigma(t)} = \mu_0.$$

Now using excursion theory and the exit system formulas, the same computation of [2] show that almost surely, there exists a t_1 such that for all $t \geq t_1$ there is no excursion e_t such that $\zeta(e_t) > \sqrt{L_t}$. This, and (3.22) imply that almost surely

$$\lim_{t \rightarrow \infty} \frac{T_t}{t} = 1$$

and thus

$$\lim_{t \rightarrow \infty} \frac{1}{t} \arg^* Z_{T_t} = \mu_0.$$

Finally, to prove the result, again by the computation in the paper mentioned, if F_a denotes the family of excursions with the property that $|\arg e(0) - \arg e(\zeta -)| \leq 2\pi$ and $\sup_{t \in [0, \zeta(e))} |\arg e(0) - \arg e(t)| \geq a$, then for every $\alpha > 0$, with probability 1, there is a $s_\alpha < \infty$ such that there are no excursions $e_t \in F_{\alpha L_t}$ with $L_t > s_\alpha$. Now, fix an arbitrarily small $\alpha > 0$ and t_1 large enough so that $\frac{1}{t} \arg^* Z_{T_t} \leq \mu_0 + \alpha$ and $\frac{L_t}{t} \leq 2$ for all $t > t_1$. If $\frac{1}{u} \arg^* Z_u \geq \mu_0 + 5\alpha$ for some $u > t_1$ then $|\arg^* Z_u - \arg^* Z_{T_u}| \geq 4\alpha u \geq 2\alpha L_u$. This means that there is an excursion starting at T_u belongs to $F_{2\alpha L_u} = F_{2\alpha L_{T_u}}$. Since there are no such excursions beyond $t > (1/2)s_{2\alpha}$, it follows that $\lim_{u \rightarrow \infty} \frac{1}{u} \arg^* Z_u \leq \mu_0$ a.s. (varying α over all rationals > 0). A similar argument yields $\lim_{u \rightarrow \infty} \frac{1}{u} \arg^* Z_u \geq \mu_0$ a.s.

(iii) To prove (iii), note that

$$\arg^* Z_{\sigma(t)} = C_t^* + \int_0^{\sigma(t)} \alpha(Z_s) dB_{L_s}^* + \int_0^{\sigma(t)} \beta(Z_s) dL_s.$$

So, the given limit exists and

$$Q_t(\arg^* Z) = \langle C_t^* + \int_0^{\sigma(\cdot)} \alpha(Z_s) dB_{L_s}^* \rangle_t.$$

Now, we note that C^* is a *pure jump* Levy process which, by an application of the Ito formula for Levy processes, satisfies the following:

$$C_t^{*2} = 2 \int_0^t C_{s-}^* dC_s^* + \int_0^t \int_{\mathbb{R}} x^2 N^*(ds, dx)$$

where N^* is a Poisson random measure with intensity given by

$$dt \otimes \frac{1}{\pi x^2} 1(|x| \leq 2\pi) dx.$$

Thus,

$$\langle C^* \rangle_t = \int_0^t \int_{\mathbb{R}} x^2 N^*(ds, dx).$$

Also, by standard stochastic calculus,

$$\langle \int_0^{\sigma(\cdot)} \alpha(Z_s) dB_{L_s}^* \rangle_t = \int_0^{\sigma(t)} \alpha^2(Z_s) dL_s.$$

As the Cauchy process is a pure jump process and $\int_0^{\sigma(\cdot)} \alpha(Z_s) dB_{L_s}^*$ is a continuous martingale, we get

$$\langle C^*, \int_0^{\sigma(\cdot)} \alpha(Z_s) dB_{L_s}^* \rangle \equiv 0$$

Thus, we get

$$Q_t(\arg^* Z) = \int_0^t \int_{\mathbb{R}} x^2 N^*(ds, dx) + \int_0^{\sigma(t)} \alpha^2(Z_s) dL_s.$$

By the strong law applied to the first term and the limit-quotient theorem applied to the second, along with the fact that $\frac{\sigma(t)}{t} \rightarrow 1$ almost surely, we get the result. \square

Note: If β is constant, then the asymptotic rotation rate $\lim_{t \rightarrow \infty} \frac{\arg^* Z_t}{t}$ is same as that of *ORBM* with drift β , and thus is unchanged by the addition of the boundary diffusion term.

2 A quantity conserved under conformal maps

We know that BMBD is not conformally invariant in the sense that the drift and diffusion functions change, though the process still remains a BMBD. If $f : D \rightarrow \tilde{D}$ is a conformal map between C^2 domains and the ‘tilde’s represent the corresponding quantities of the

mapped (appropriately time-changed) process in the new domain, then the new drift and diffusion functions are given by (3.3) and (3.2) respectively. So, a natural question to ask can be: Is there a quantity Q that is conserved under conformal maps in the sense that $\tilde{Q}(f(x)) = Q(x)$? In the next theorem, we answer this question in the affirmative.

Theorem 16. *Let $f : D \rightarrow \tilde{D}$ is a conformal map between C^2 domains, which maps a $BMBD(\alpha, \beta)$ process X onto a $BMBD(\tilde{\alpha}, \tilde{\beta})$ process Y defined on \tilde{D} given by $Y_t = f(X(c^{-1}(t)))$ where $c(t) = \int_0^t |f'(X_s)|^2 ds$ and $\tilde{\alpha}, \tilde{\beta}$ are given by (3.2) and (3.3) respectively. Define*

$$Q(x) = \beta(x) - \alpha(x)\partial_t\alpha(x) \text{ for } x \in \partial D$$

and let \tilde{Q} represent the corresponding quantity on $\partial\tilde{D}$. Then

$$\tilde{Q}(f(x)) = Q(x)$$

Proof: Taking tangential derivatives with respect to ∂D at x on both sides of the equation

$$(\tilde{\alpha} \circ f)^2(x) = \alpha^2(x)|f'(x)|$$

we get

$$2(\tilde{\alpha} \circ f)(x)\partial_{\mathbf{t}}(\tilde{\alpha} \circ f)(x) = 2\alpha(x)\partial_{\mathbf{t}}\alpha(x)|f'(x)| + \alpha^2(x)\partial_{\mathbf{t}}|f'(x)|$$

Dividing both sides by $2|f'(x)|$, we get

$$\tilde{\alpha}(f(x))\partial_{\tilde{\mathbf{t}}}\tilde{\alpha}(f(x)) = \alpha(x)\partial_{\mathbf{t}}\alpha(x) + \alpha^2(x)\frac{1}{2|f'(x)|}\partial_{\mathbf{t}}|f'(x)|$$

where $\tilde{\mathbf{t}}$ represents the unit tangent to $\partial\tilde{D}$ at $f(x)$. Now, it can be checked (for example see calculations in [1]) that

$$\partial_{\mathbf{t}}|f'(x)| = \tilde{\mathbf{t}}.\Delta_{\partial D}f(x).$$

Plugging this in, and using (3.2), we get

$$\tilde{\alpha}(f(x))\partial_{\tilde{\mathbf{t}}}\tilde{\alpha}(f(x)) = \alpha(x)\partial_{\mathbf{t}}\alpha(x) + \tilde{\alpha}^2(f(x))\frac{1}{2|f'(x)|^2}\tilde{\mathbf{t}}.\Delta_{\partial D}f(x).$$

Using this and (3.3), we get

$$\tilde{\beta}(f(x)) - \tilde{\alpha}(f(x))\partial_{\mathbf{t}}\tilde{\alpha}(f(x)) = \beta(x) - \alpha(x)\partial_{\mathbf{t}}\alpha(x)$$

which proves the theorem. \square

3 A closer look at the trace process

We now focus on $BMBD(1,0)$ on the upper half plane \mathbb{H} . We see that the trace process $\Lambda_t = C_t + B_t^*$ is the sum of an independent Cauchy process C given by $C_t = B_{\sigma_t}^{(1)}$ and a Brownian motion B^* . This is an example of a *Levy Process* with Levy-Khintchine exponent given by

$$\Psi(\theta) = \frac{\theta^2}{2} + \int_{\mathbb{R}} (1 - e^{i\theta x} + i\theta x I(|x| < 1)) \Pi(dx) \quad (3.24)$$

where the jump measure $\Pi(dx) = \pi(x)dx$ is given by $\pi(x) = \frac{1}{\pi x^2}$ for $x \neq 0$. Note that the trace process can also be written down as B_{σ_t+t} where σ is the inverse local time of $B^{(2)}$ and B is a Brownian motion independent of $B^{(2)}$. The time process of this Brownian motion, $\gamma_t = \sigma_t + t$ is a *subordinator with drift 1*, and there are lots of facts known about this, most of which we take from [6]. Note that we can write down the exponent in the Laplace transform of this as

$$\Phi(\theta) = -\log E(e^{-\theta\gamma_1}) = \theta + \int_{(0,\infty)} (1 - e^{-\theta x})\pi(x)dx \quad (3.25)$$

where $\pi(x) = \frac{1}{2}x^{-3/2}$ on $(0, \infty)$.

It is known that the Cauchy process almost surely does not hit a given boundary point. But by what we saw for our process, it hits any boundary point almost surely. We try to have a closer look at this phenomenon. The main essence of this result is carried by the properties of the time variable γ_t . Before we can proceed, we define the *Potential measure* of a process X by

$$U(dx) = \int_0^\infty P(X_t \in dx)dt. \quad (3.26)$$

This measure is pivotal in questions of transience and recurrence because it gives an estimate of how long a process spends in different regions of space. Define $\tau_x^+ = \inf\{t > 0 : X_t > x\}$

to be the first passage time above level x . We now define the phenomenon of *creeping*. A process X is said to creep above a level x if $P(X_{\tau_x^+} = x) > 0$. The following theorem is taken from [6].

Theorem 17. (i) Let X be a (killed) subordinator with $\Pi(0, \infty) = \infty$. Then for all $x > 0$ we have

$$P(X_{\tau_x^+} = x, x - X_{\tau_x^+ -} > 0) = 0, \quad (3.27)$$

that is, with probability one, the process cannot jump to the level x on its first passage.

(ii) If X as above has drift 0, then for all $x > 0$,

$$P(X_{\tau_x^+} = x) = 0. \quad (3.28)$$

(iii) If X as above has drift $d > 0$, then U has a strictly positive and continuous density u on $(0, \infty)$ and

$$P(X_{\tau_x^+} = x) = du(x). \quad (3.29)$$

Let $p_x = P(X_{\tau_x^+} = x)$. Then if $p_x > 0$ for some $x > 0$,

$$\lim_{\epsilon \downarrow 0} p_\epsilon = 1 \quad (3.30)$$

and the mapping $x \mapsto p_x$ is strictly positive and continuous on $[0, \infty)$.

In our case, it is easy to check that

$$P(\gamma_t \in dx) = \frac{t}{\sqrt{2\pi}(x-t)^{3/2}} e^{-\frac{t^2}{2(x-t)}} dx, \quad x > t. \quad (3.31)$$

So, $u(x) = \int_0^x \frac{x-y}{\sqrt{2\pi}y^{3/2}} e^{-\frac{(x-y)^2}{2y}} dy$ which satisfies the condition (iii) of the above theorem.

So

$$P(\gamma_{\tau_x^+} = x) = \int_0^x \frac{x-y}{\sqrt{2\pi}y^{3/2}} e^{-\frac{(x-y)^2}{2y}} dy. \quad (3.32)$$

It can also be checked that the above quantity goes to one as $x \downarrow 0$. The above theorem then yields us the following useful corollary.

Corollary 17.1. *Consider the process $B(\gamma_t)$ where B is a standard Brownian motion starting from 0 and B and γ are independent. Then with probability one, the process visits 0 infinitely often in any arbitrarily small time interval.*

Proof: We will show that conditionally on the Brownian motion B , the probability that the process γ visits the zero set of B infinitely often in any small interval $[0, \epsilon)$ is one. (Note that we heavily make use of the fact that B and γ are independent).

We know that with probability one the Brownian motion path has infinitely many zeros in any arbitrarily small time interval starting at 0. Now, by part (iii) of the above theorem, there exist zeros $\{z_k\}$ such that $p_{z_k} \geq 1 - 2^{-k}$. Then by Borel-Cantelli lemma, γ visits all but finitely many of the z_k 's with probability one yielding the lemma. \square

As σ is an additive functional (i.e., $\sigma_t = \sigma_s + \tilde{\sigma}_{t-s} \circ \theta_s$ where θ is the shift operator), so an analogous property as that of the above corollary holds at any stopping time.

Note that this is another indication of the trace process *behaving on a small time scale similar to Brownian motion*.

Recall that after the time change $t \mapsto S^{-1}(t)$, we obtained new co-ordinate processes \tilde{X} and \tilde{Y} which are independent and the former is a Brownian motion and the latter is a Sticky Brownian motion. Now we study some properties of Sticky Brownian motion (SBM), taken from [4]. Formally, it is defined to be a continuous process adapted to the filtration $\tilde{\mathcal{F}}$ taking values in $[0, \infty)$ which satisfies the following SDE:

$$\tilde{Y}_t = y + \theta \int_0^t I_{\tilde{Y}_s=0} ds + \int_0^t I_{\tilde{Y}_s>0} dW_s, \quad (3.33)$$

where W is a standard Brownian motion adapted to the same filtration $\tilde{\mathcal{F}}$. It is an important process as it arises as the limit of storage processes and some random walks. Ikeda and Watanabe [7] showed that the above SDE admits a weak solution and enjoys a uniqueness in law property. But although the joint law of (\tilde{Y}, W) is unique, \tilde{Y} is not measurable with respect to W , and thus the above equation does not have a strong solution. Thus the

filtration $\tilde{\mathcal{F}}$ cannot be the augmented natural filtration of W and the process \tilde{Y} always contains some extra randomness. The following remarkable theorem is taken from [4]:

Theorem 18. *Let \tilde{Y} be a SBM starting from 0 and W be its driving Weiner process, also starting from 0. Letting $L_t^W = \sup_{s \leq t} (-W_s)$, the conditional law of X given W satisfies*

$$P(\tilde{Y}_t \leq x | \sigma(W)) = \exp(-2\theta(W_t + L_t^W - x)) \quad a.s. \quad (3.34)$$

for $x \in [0, W_t + L_t^W]$. In particular, $\tilde{Y}_t \in [0, W_t + L_t^W]$ a.s.

Now, we study the process $Z_t = B(\gamma(t))$ where $\gamma(t) = \sigma_t + t$. This process a particular case of a wide class of processes called *subordinate Brownian motions*. It has seen significant development in recent years, some of which we are going to enlist now. Its characteristic exponent is $\Psi(\theta) = \theta + \theta^2$ and its Levy measure has density is given by

$$\pi(x) = \int_0^\infty \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}} \frac{1}{2t^{3/2}} dt = \frac{1}{\sqrt{\pi x^2}} \quad (3.35)$$

for $x > 0$ and $\pi(x) = \pi(-x)$ for $x < 0$. This is expected as this is the Levy measure of a Cauchy process. Let $\bar{Z}_t = \sup_{s \leq t} Z_s$ be the running supremum process and $\bar{Z} - Z$ denote the process reflected at its running supremum. As the process is of unbounded variation, there exists a continuous version of the local time at zero of the process $\bar{Z} - Z$ and we denote it by L . And the inverse local time by L^{-1} . The inverse local time is a subordinator. The *ascending ladder height process* H is defined as $H_t = Z(L_t^{-1})$, and this is also a subordinator with Laplace exponent χ given by

$$\chi(\lambda) = \exp\left(\frac{1}{\pi} \int_0^\infty \frac{\log(\lambda^2 \theta^2 + \lambda \theta)}{1 + \theta^2} d\theta\right). \quad (3.36)$$

From [5], it follows that the ladder process H has drift 1 as $\lim_{\lambda \rightarrow \infty} \frac{\chi(\lambda)}{\lambda} = 1$. This yields another proof of the fact that the process creeps, though it does not yield the result of infinitely many returns to the starting point which we proved earlier. Also, it follows from the paper that the Levy measure of $(0, \infty)$ of the ladder height process is infinite.

Define the potential measure (or occupation measure) of the subordinator H as

$$V(A) = \int_0^\infty I(H_t \in A) dt \quad (3.37)$$

for $A \in \mathcal{B}[0, \infty)$, and $V(x) = V[0, x]$ for $x > 0$. V is invariant, and hence harmonic for Z in $(0, \infty)$. From this we get that if η is the time spent by Z (starting from a positive point) before it hits $(-\infty, 0]$ and T_r is the hitting time of $[r, \infty)$, then

$$V(x) \geq V(r)P_x(T_r < \eta) \quad (3.38)$$

for $0 < x < r < \infty$. [5] establishes the following sharp bounds on the Green's function of the process Z killed on exiting $(0, \infty)$:

$$G^{(0, \infty)}(x, y) \asymp \begin{cases} x(y^{-1/2} \wedge 1), & 0 < x < 1, \\ \log\left(\frac{1+x^{1/2}y^{1/2}}{1+y-x}\right), & 1 \leq x < y < 2x, \\ x^{1/2}y^{-1/2}, & 1 \leq x < 2x < y. \end{cases} \quad (3.39)$$

BIBLIOGRAPHY

- [1] Card, Ryan Kazuo. Brownian Motion with Boundary Diffusion. Proquest, Umi Dissertatio, 2011.
- [2] Burdzy, K., Chen, Z., Marshall D., Ramanan K. (2013) Obliquely Reflected Brownian Motion in Non-Smooth Planar Domains. preprint.
- [3] Burdzy, Krzysztof, and Donald Marshall. "Hitting a boundary point with reflected Brownian motion." Sminaire de Probabilits XXVI. Springer Berlin Heidelberg, 1992. 81-94.
- [4] Warren, Jonathan. "Branching processes, the Ray-Knight theorem, and sticky Brownian motion." Sminaire de Probabilits XXXI. Springer Berlin Heidelberg, 1997. 1-15.
- [5] Kim, Panki, Renming Song, and Zoran Vondracek. "On the potential theory of one-dimensional subordinate Brownian motions with continuous components." Potential analysis 33.2 (2010): 153-173.
- [6] Kyprianou, Andreas. Introductory Lectures on Fluctuations of Levy processes with Applications (Universitext). Springer, 2006.
- [7] Ikeda, Nobuyuki, and Shinzo Watanabe. "Stochastic differential equations and diffusion processes." (1989).