Probabilistic and geometric methods in last passage percolation

by

Milind Hegde

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Mathematics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Alan Hammond, Co-chair
Professor Shirshendu Ganguly, Co-chair
Professor Fraydoun Rezakhanlou
Professor Alistair Sinclair

Spring 2021

Probabilistic and geometric methods in last passage percolation

Abstract

Probabilistic and geometric methods in last passage percolation

by

Milind Hegde

Doctor of Philosophy in Mathematics

University of California, Berkeley

Professor Alan Hammond, Co-chair

Professor Shirshendu Ganguly, Co-chair

Last passage percolation (LPP) refers to a broad class of models thought to lie within the Kardar-Parisi-Zhang universality class of one-dimensional stochastic growth models. In LPP models, there is a planar random noise environment through which directed paths travel; paths are assigned a weight based on their journey through the environment, usually by, in some sense, integrating the noise over the path. For given points y and x, the weight from y to x is defined by maximizing the weight over all paths from y to x. A path which achieves the maximum weight is called a *geodesic*.

A few last passage percolation models are exactly solvable, i.e., possess what is called integrable structure. This gives rise to valuable information, such as explicit probabilistic resampling properties, distributional convergence information, or one point tail bounds of the weight profile as the starting point y is fixed and the ending point x varies. However, much of the behaviour currently proved within only the special models with integrable structure is expected to hold in a much larger class of models; further, integrable techniques alone seem unable to obtain certain types of information, such as process-level properties of natural limiting objects. This thesis explores probabilistic and geometric approaches which assume limited integrable input to infer information unavailable via exactly solvable techniques or to develop methods which will prove be to robust and eventually applicable to a broader class of non-integrable models.

In the first part, we study the parabolic Airy_2 process $\mathcal{P}_1$ and the parabolic Airy line ensemble $\mathcal{P}$ via a probabilistic resampling structure enjoyed by the latter. The parabolic Airy line ensemble is an infinite $\mathbb{N}$-indexed collection of random continuous non-intersecting curves, whose top (and lowest indexed) curve is $\mathcal{P}_1$, which is a limiting weight profile in LPP when the starting point is fixed (which can be thought of as a particular form of initial data) and the ending point is allowed to vary along a line. The ensemble $\mathcal{P}$ possess a probabilistic resampling property known as the Brownian Gibbs property, which gives an explicit description of a certain conditional

distribution of $\mathcal{P}$ in terms of non-intersecting Brownian bridges. This property gives a qualitative comparison—absolute continuity—of $\mathcal{P}_1$ to Brownian motion, as proved in [CH14]. Using a framework introduced in [Ham19a], we prove here a strong quantitative form of comparison, showing that an event which has probability ε under the law of Brownian motion has probability at most $\varepsilon^{1-o(1)}$ under the law of an increment of $\mathcal{P}_1$, with an explicit form of $\exp(O(1)(\log \varepsilon^{-1})^{5/6})$ for the $\varepsilon^{-o(1)}$ error factor. Up to the error factor, this is expected to be sharp. One consequence is that the Radon-Nikodym derivative of the increment of $\mathcal{P}_1$ with respect to Brownian motion lies in all L^p spaces with $p \in [1, \infty)$, i.e., has finite polynomial moments of all orders. The bounds also hold for lower curves of $\mathcal{P}$ and for weight profiles in an LPP model known as Brownian LPP.

In the second part, we work in an LPP model on the lattice $\mathbb{Z}^2$ and make use of a geometric approach, i.e., we study properties of geodesics and other weight maximizing structures. The random environment is given by i.i.d. non-negative random variables associated to the vertices of $\mathbb{Z}^2$, and the weight of an up-right path is the sum of the random variables it passes through. Now, in exactly solvable models such as when the vertex weight distribution is geometric or exponential, the GUE Tracy-Widom distribution is known to be a scaling limit of the last passage value X_r from $(1, 1)$ to (r, r). The GUE Tracy-Widom distribution is well-known to have upper and lower tail exponents of $\frac{3}{2}$ and 3, which is also known for the prelimiting X_r in the mentioned exactly solvable models. Here we work in a more general setup and adopt some natural assumptions of curvature and weak one-point upper and lower tail estimates on the weight profile—with no assumptions on the vertex weight distribution and, hence, a non-integrable setting—which we bootstrap up to obtain the optimal tail exponents for both tails, in terms of both upper and lower bounds. We also obtain sharp upper and lower bounds for the lower tail of the maximum weight over all paths which are *constrained* to lie in a strip of given width, a result which was not previously known in even the integrable models and does not seem easily accessible via integrable techniques.

Finally, in the third part, we combine the geometric and probabilistic approaches to obtain an estimate on a certain probability of the KPZ fixed point. The KPZ fixed point $(t, x) \mapsto \mathfrak{h}_t(x)$ is a space-time process constructed in [MQR17] which should be thought of as a limiting object analogous to $\mathcal{P}_1$ under general initial data. We show that the event, for fixed $t > 0$, that $\mathfrak{h}_t$ has ε-*twin peaks*, i.e., there exists a point at a given distance away from the maximizer of $\mathfrak{h}_t$ at which $\mathfrak{h}_t$ comes within ε of its maximum, has probability at least of order ε. This served as an important input in [CHHM21] to prove that the Hausdorff dimension of the set of exceptional times t where $\mathfrak{h}_t$ has multiple maximizers is almost surely $\frac{2}{3}$, on the positive-probability event that the set is non-empty. Geometric and probabilistic arguments are combined by making use of a remarkable identity of last passage values between an original and a certain transformed environment proven in [DOV18], which allows us to consider geodesics through an environment which itself has the Brownian Gibbs property.

To family, friends, mentors, and the simple things which suffice.

Contents

List of Figures

Acknowledgments

And when we got there, I wasn't sure
Had the world spun a little more than
I had expected when we set out?
Had I got wiser overnight somehow?

–Mumm-ra, *What Would Steve Do?*[1]

I have somehow finally come to the task of writing the acknowledgments of my thesis. It is a strange feeling, partly because during much of graduate school, when bored, I would often look up the theses of others and read their acknowledgments, to get a shallow window into their experience—and to wonder about what I might write in my own.

Graduate school has been quite a journey, and almost completely because of the people I spent time with and learned from along the way. I will try to express my feelings to everyone who has affected me, but I apologize if I somehow forget someone.

First, of course, is my adviser, Alan. As I imagine it is for every graduate student, it's hard to concisely describe the effect my adviser has had on me. Alan was the one who introduced me to research in probability, and gave me problems and projects that I could think about and which were within my abilities to solve. His style of thinking about probability seems, to me, creatively unique and I have attempted to learn from it. And I consider myself lucky to know a senior mathematician who so obviously has a variety of interests outside of mathematics, to which he also puts in time and energy, and maintains an enviable balance of work and life—it has helped me to also think about my own priorities and give myself the time to do the same.

In many ways, I have had a second adviser—Shirshendu—recently also officially. I remember how I was initially quite intimidated by his abilities, and put off trying to talk to him for at least a year! This was pure foolishness on my part. The amount I have gained from interacting with him is immense—in terms of mathematical ideas, collaborations, and also advice about so much. The slow blending of mentorship into aspects of friendship is also something I value very much, and hope to continue.

I found my first semester in Berkeley to be a somewhat turbulent time. I wasn't sure who might be my adviser, wasn't getting very excited about courses or finding much of a coherent probability student community, and was questioning some decisions. But much of this changed in my second semester, and in large part because of Jim Pitman's course on combinatorial stochastic processes. More than the subject matter, Jim's incredible energy and infectious enthusiasm about probability rekindled my own, and his efforts to organize weekly post-seminar pizza and drinks laid the groundwork for the probability students to continue on our own in later semesters. And,

[1] Like many pop songs, this one appears to be about romance; also like many pop songs, larger themes can sometimes be discerned if you're only half-listening.

not to discount the subject matter, he has patiently answered all my questions and reference requests on Brownian motion, and taught me perhaps my favorite theorem of probability: the decomposition of Brownian motion into a Poisson point process of Brownian excursions.

I thank Fraydoun Rezakhanlou, for his many thorough courses that I learned a lot from, as well his weekly seminar. I deeply enjoyed teaching for Craig Evans, and his lectures that I attended have been a model of clarity and flow that I would like to emulate. Alistair Sinclair's courses on randomness and computation and mixing times of Markov chains exposed me to a new perspective on probability that I hope to further explore.

I also have to thank my teachers and mentors from before graduate school. Foremost among them is Manjunath, who first introduced me to and inspired my interest in probability—I remember coming to the first class of his graduate probability course with the intent of understanding the difference between the strong and weak laws of large numbers—and became a larger role model to me as a teacher and person. I think I can safely say we are now also friends, and I have very much enjoyed our correspondence over the last few years on a wider variety of subjects. I also thank all the faculty at IISc for teaching me much of the math I know.

Going further back, I should also mention my 11^{th} grade physics teacher, BGK, who handed me her copy of Griffith's *Electrodynamics* when I asked too many questions; K.D. Joshi, whose book *Educative JEE Mathematics* revealed mathematics to me as a subject of depth (and whose response a couple years ago to an email I had sent him in high school made me very happy); Mrs. Fohner, who spurred my interest in science back in 7^{th} grade; and Ms. Lawlor, my 6^{th} grade teacher who I am certain influenced me in deep ways, but they are too deep to pick apart now.

The math department graduate advisers—Vicky, Isabel, Jon—have been such a boon and made things so much easier. I thank Vicky in particular for her help when, like in every student's nightmare, I forgot to show up for an online (and pre-COVID!) final exam in my very first year. Since then I have never credited a course I didn't absolutely have to.

Apart from the above, I must also thank my collaborators—Ivan, Konstantin, Riddhi—for working with me, sharing their knowledge, and reading and giving feedback on my own mathematics. Thanks also to Ivan for helping and giving advice during the months I was applying for postdoctoral positions, to Allan Sly for support during that time, and to Riddhi for giving me valuable advice in my first semester! And finally, to Erik for giving me very detailed advice on going about applications, and equally helpful feedback on my research statement.

Of course, my time at Berkeley has been so lovely mainly because of my friends. Nick and I met and became housemates rather randomly, but I've loved our many conversations on so many topics on that sofa we very wisely picked out early on. I will miss cooking, and joking, and mocking TV shows, and more, with Koulik. With Anamika, long phone calls—is it "Hello Mr. Milind" or "Hello Mr. Hegde"?—and walks, and the relentless teasing. With Jimmy, the discussions about India, and with Siddhesh and Anushka, the hikes!

Satyaki's singularity is hard to communicate to those who haven't met him—all I can say is that he is a force—but he made me more clearly articulate so many thoughts over the years, and the world should make more space for people like him. It's funny to think about how my first conver-

sation with Dan—about the proto-Indo-European language, if I remember correctly—became a template for so many later ones, but I think the real basis for our friendship might be a strongly shared sense of the absurd, which I believe is rather uncommon to share with someone. Ella, Jake, and Adam have been great companions, probabilistic and otherwise, especially in La Val's. I've been inspired a lot by Ella's moral convictions, and have loved walks with Jake, a habit we developed well before COVID. Ahmed, with whom I could feel the spark nearly instantly—I would have loved if we had somehow been at the same graduate school. And Sri, another friend met randomly, but with whom a spark was also immediate and sustained.

There are also my many friends from college. With Sahana I've had many, many long conversations fleshing out the things we were going through, and also a friend with whom I could stay in New York—I'm looking forward to returning the favour! With Harsha I've exchanged so many meandering emails—can't believe that was purely post-undergrad!—and so many lovely memories for such geographic distance; the visit to Europe and our conversations over the last year are perhaps some of my favorites from all of grad school. With Pritish, Rohit, Anamay, Naren, Sabareesh, it's amazing how conversations seem to be exactly the same as they were at IISc, even if they have at times been slightly more sporadic. I look forward to the day when I'm able to convince Sabareesh of my point of view on any topic. More reflections on how we are changing will surely come with Pranav and Aditya.

And finally, I thank my family: Appa, Amma, Sahana. It's probably because of Appa's urging that I don't become a software engineer that I ended up on the path I did, not to mention my parents' continuous nurturing of my love of books. Amma indulged my five-year-old self's craze of multiplying large numbers by two, if I remember correctly. You were all asking me about progress on my thesis several years before I started thinking about my work in those terms, but I can finally say: here it is!

Part I

Introduction

Chapter 1

The KPZ universality class

The Kardar-Parisi-Zhang (KPZ) universality class refers to a broad family of one-dimensional stochastic growth models which are believed to showcase certain *universal* behaviour. The universality class was introduced by the physicists it is named after in a paper [KPZ86] from 1986, in which this universality was conjectured and a qualitative description of the defining characteristics of models which should lie in the class was proposed.

In this chapter we will describe a number of models thought to or known to lie in the KPZ universality class to illustrate the connections and state of knowledge, though this thesis will study only a few models; a good survey article is [Cor12]. Our results will be discussed in Chapter 2.

1.1 The Eden model and first passage percolation

Let us first consider a simple model known as the *Eden model*, which was introduced by the physical chemist and electrical engineer Murray Eden to model bacterial growth [Ede61]. The setting is $\mathbb{Z}^2$, and we consider its faces (one can also consider an equivalent model on the vertices

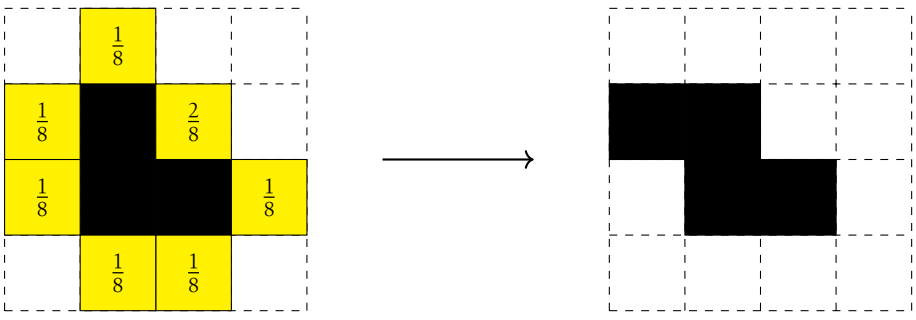


Figure 1.1: One step of the Eden model. The currently occupied faces are black, the adjacent faces are yellow, and the non-adjacent unoccupied faces are white. The fraction in each yellow boxes is the probability of it being picked to be newly occupied.

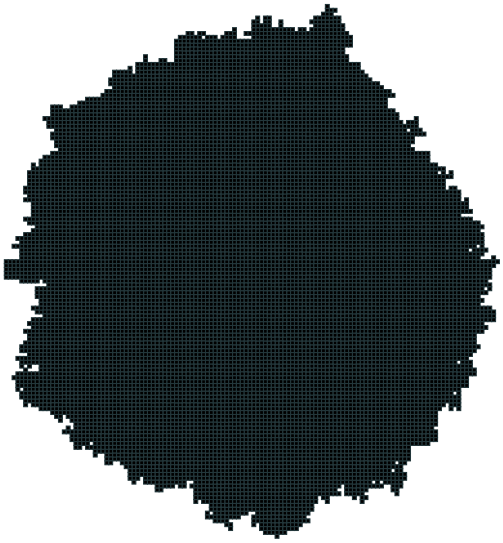


Figure 1.2: A simulation of the Eden model after about 12,000 discrete time steps.

of $\mathbb{Z}^2$ by going to the dual graph). Initially, the face at the origin is occupied—representing a single bacterial cell. At each discrete step of time, a face adjacent to one of the currently occupied faces is picked with probability proportional to the number of occupied faces it is adjacent to, and becomes occupied—representing the event that one of the boundary bacterial cells divides into two. See Figure 1.1. One can also consider a continuous time version of the model by assigning to each boundary edge separating an occupied and an unoccupied face an independent rate one exponential clock (i.e., random variable), and declaring the unoccupied face corresponding to the first clock that rings to be newly occupied.

As might be expected, of interest is the long-time behaviour of the set of occupied faces. In Figure 1.2 is a simulation after about 12,000 steps. To first order, the beginnings of a limiting shape can perhaps be discerned from the figure. Indeed, the existence of a deterministic limit shape has been proved in all dimensions by Cox-Durrett [CD81].

With the existence of a limit shape, two directions naturally arise to pursue enquiry:

(i) the properties of the limit shape, and
(ii) the nature of the fluctuations in the prelimit around the limit shape.

It is fairly immediate from the definition of the model that the limit shape must respect the symmetries of the lattice, i.e., be unchanged under reflection about the axes and rotation by multiples of 90°. The model also possesses an additive structure known as *sub-additivity* which implies that the limit shape must be convex. (However, in spite of these properties and unlike what may be suggested by Figure 1.2, it can be seen from larger simulations that the limit shape is almost certainly *not* a circle.)

To state this subadditivity, let us consider a broader class of models known as *first passage percolation* (FPP), of which the variant of the Eden model with exponential waiting times mentioned

above is a special case. To each edge e of $\mathbb{Z}^2$ (or, more generally, $\mathbb{Z}^d$) is associated an i.i.d. non-negative random variable ξ_e. To a non-self-intersecting path γ in $\mathbb{Z}^2$ is associated a *weight* (also often called a waiting time) $w(\gamma)$ which is given by $\sum_{e\in\gamma}\xi_e$, i.e., the sum of random values associated to the edges on γ. Then for given vertices $x, y \in \mathbb{Z}^2$, the weight from x to y is

$$\tau_{x,y} = \min_{\gamma: x\to y} w(\gamma),$$

with the minimum over all paths connecting x to y. We note, to contrast with later, that the paths are allowed to *backtrack*, i.e., need not be functions of any particular coordinate.

Unlike the biological origins of the Eden model, FPP was introduced by Hammersley and Welsh in [HW65] to model the physical phenomenon of a liquid flowing from one side of a porous medium to the other. This can also be seen as a notion of random growth by thinking of the liquid as growing into the medium.

The mathematical connection of FPP to the (continuous time version of the) Eden model is the following: In the Eden model, a given site x could have become occupied only by a path of occupied vertices connecting the origin to it. Each such path takes a given time to reach x, which is the sum of the waiting times (i.e., weights, which are exponentially distributed) of each edge on it. Thus x will be occupied at a time which is the *minimum* waiting time over all the paths that reach x from 0. The set of occupied vertices at a given time t, like what is shown in Figure 1.2, is exactly the set $\{x \in \mathbb{Z}^2 : \tau_{0,x} \le t\}$. So, the Eden model is FPP in two dimensions in the special case that ξ_e are distributed as i.i.d. exponential random variables.

The variational formula for $\tau_{x,y}$ yields the aforementioned subadditive relation easily: we have

$$\tau_{x,y} \le \tau_{x,z} + \tau_{z,y},$$

for any $z \in \mathbb{Z}^2$, i.e., the triangle inequality. The reason is that any two paths from x to z and z to y respectively can be concatenated to give a path from x to y, and the weight of the concatenated path is the sum of the weights of the individual paths.

This subadditive relation is immensely useful because of Kingman's subadditive ergodic theorem (which, in fact, was introduced in its original form precisely to study FPP [HW65]). Though we do not state a precise version of this theorem here, it essentially yields the existence of the limit shape mentioned above by giving the almost sure existence of the following closely related limit for x a vector in $\mathbb{Z}^2$:

$$\lim_{n\to\infty} \frac{\tau_{0,[nx]}}{n} = \mu(x),$$

where $\mu(x)$ is a deterministic constant usually known as the *time constant*. In terms of the time constant, the normalized limiting shape of the set of sites reached by a given time (as time goes to infinity) is given by $\{x : \mu(x) \le 1\}$.

We can now reformulate the questions listed on page 3 in a slightly more precise way. We view the first as asking about the properties of the function μ, and the second as asking about the fluctuations of $\tau_{0,[nx]} - n\mu(x)$. The subadditivity relation implies that μ is a convex function,

and so an initial form of the first question is to establish *strict* convexity of μ; an initial form of the second is to identify the *scale* (eg. n^α for some $\alpha > 0$) of the fluctuations of $\tau_{0,[nx]} - n\mu(x)$.

Unfortunately, we must confess that we have been leading the reader thus far to an anticlimactic conclusion. *Essentially nothing is known about the answers to either of these questions.* This is the state of knowledge for the vast majority of models believed to lie within the KPZ universality class. The models which are exceptions are those which have what is known as *integrable* structure.

1.2 Integrability within KPZ

While we have mentioned the KPZ class, we have not yet provided even a qualitative description of what features should identify a model as belonging to the class. We start this section by rectifying this omission.

Models in the class typically have an associated *height function* or *profile*. In FPP, this could be the profile of weights of all vertices at a given distance from the origin, or the boundary of the set of vertices reached by a given time. Three features of the evolution of the profile are believed to be hallmarks of the models in the KPZ class:

(i) *a smoothening tendency*: depressions in the profile are quickly filled,
(ii) *slope-dependent growth*: faster growth where the profile's local gradient is larger, and
(iii) *space-time white noise roughening*.

Let us illustrate this in the context of the Eden model/FPP. The first two points are related and say, for example, that regions of the boundary of the set of occupied sites which form local depressions will experience faster growth, thereby pushing the overall profile towards smoothness. In terms of the dynamics, this is manifested through the fact that such macroscopically rougher portions of the boundary will have a greater number of unoccupied sites incident—a sort of isoperimetric phenomenon—and so there is a greater probability of an incident site in that region being picked to be occupied than in macroscopically smoother boundary areas. The third point on white noise roughening simply refers to the i.i.d. randomness of the environment creating fluctuations and local roughness.

While the above three qualities are present in every KPZ model, they are extremely explicit in the *KPZ equation*, a stochastic PDE thought to describe the scaling limit of the height function of many KPZ models in certain parameter regimes, and the first of our models which displays integrable structure. We turn to it next.

The KPZ equation is a stochastic PDE stated formally as

$$\partial_t \mathcal{H}(t,x) = \frac{1}{2}\partial_x^2 \mathcal{H}(t,x) + \frac{1}{2}\big(\partial_x \mathcal{H}(t,x)\big)^2 + \xi(t,x), \tag{1.1}$$

where $\xi(t,x)$ is space-time white noise, i.e., a Gaussian process defined on $\mathbb{R}^2$ with covariance given by

$$\mathbb{E}[\xi(s,y)\xi(t,x)] = \delta_{t=s}\delta_{x=y}; \tag{1.2}$$

it is rigorously defined as a random generalized function.

The solution $\mathcal{H}$ of (1.1) is called the *height* function, and it can be thought of as the height of a one-dimensional interface at position x and time t as the interface evolves.

As mentioned, the three features of the KPZ class are explicit on the righthand side of (1.1): the first term, a Laplacian, represents smoothening; the second, the square of the gradient, represents (non-linear) slope-dependent growth; while the third is white noise in the flesh.

Note that we said that the KPZ equation is stated *formally* as (1.1). This is because it is quite non-trivial to make sense of what (1.1) actually means. Essentially, the difficulty comes from the space-time white noise $\xi(t,x)$. Its presence suggests that $\mathcal{H}$ should be locally rough, perhaps Brownian in some sense, in which case the derivative $\partial_x \mathcal{H}(t,x)$ will only exist in the sense of distributions, i.e., generalized functions. While generalized functions can always be differentiated, there is in general no way to directly multiply them, and so the second term of (1.1) is not well-defined.

Given this issue, there has been much recent work to make sense of (1.1). While not getting into the details, we mention here some of the main advances and papers: Hairer's theory of regularity structures [Hai13], paracontrolled distributions [GIP12, GP17], and energy solutions [GJ14]. These are all quite sophisticated approaches, and the first two are aimed at not only making sense of the KPZ equation, but also other non-linear and singular stochastic PDEs. For our purposes, since we are concerned with only (1.1), however, there is a simpler, physically relevant notion of solution.

This physically relevant solution is known as the *Cole-Hopf* solution. It is given by *defining* $\mathcal{H}(t,x)$ to be $\log \mathcal{Z}(t,x)$, where $\mathcal{Z}(t,x)$ solves the multiplicative stochastic heat equation (SHE)

$$\partial_t \mathcal{Z}(t,x) = \frac{1}{2}\partial_x^2 \mathcal{Z}(t,x) + \xi(t,x)\mathcal{Z}(t,x), \tag{1.3}$$

with initial condition given by $\mathcal{Z}(0,x) = \exp(\mathcal{H}(0,x))$. One can use a formal chain rule computation to check that, if $\mathcal{Z}(t,x)$ solves (1.3), then $\mathcal{H}(t,x)$ should solve something like (1.1); also, it is known that $\mathcal{Z}$ remains positive for all positive time if it has non-negative initial data [Mue91], and so $\mathcal{H}$ is well-defined. Note further that the stochastic heat equation is linear in $\mathcal{Z}$ and so doesn't require making sense of things like multiplying generalized functions; this means that it is much more straightforward to give a meaning to the solution of (1.3).

A great deal is now known about the KPZ equation under a general class of initial data, but initial progress was restricted to what is called the *narrow wedge* solution. The narrow wedge case is when $\mathcal{Z}$ is the fundamental solution to the SHE, i.e., $\mathcal{Z}(0,x) = \delta_{x=0}$, a Dirac mass at zero; in this case, $\mathcal{H}(0,x)$ can be thought of as being 0 at $x = 0$ and $-\infty$ elsewhere, so that growth is initially from the single point $x = 0$, analogous to the single starting site of the Eden model.

The reason the narrow-wedge condition is special is because it has *integrable* structure, also often referred to as *exactly solvable* structure. (This will also be the case for the analogues of narrow wedge initial conditions in other models which this thesis is concerned with that we will introduce in Chapter 2.) What does this mean?

Essentially, it means that there are *exact* descriptions or formulas for the distribution of the relevant random quantity. Since the random variables of interest are usually quite complicated and non-linear functions of the noise, this is by no means typical. In the case of the KPZ equation with narrow wedge initial condition, one such random quantity is $\mathcal{Z}(t,0)$; the exact description available for $\mathcal{Z}(t,0)$ is its distribution function and its moments (or Laplace transform) [ACQ11, BG16].

These formulas are given in terms of classical objects like the Airy function (and the associated Airy point process, which we will say more about later) and can be expressed as contour integrals; in other models, combinatorial bijections and algebraic identities may come into play, which can then be related to similar analytic descriptions using devices like classical orthogonal polynomial theory. These contour integral formulas, with work, can be made amenable to analysis such that sharp asymptotic information can be extracted. The great deal of analytic control available over these objects is what makes integrable models mathematically tractable.

The continuum directed random polymer

The KPZ equation at finite time t can be thought of as a "positive temperature" model. What this means can be made more clear by way of a polymer model, known as the *continuum directed random polymer* (CDRP), which is closely associated to the KPZ equation. This discussion will be rather heuristic and not concern itself with the technical subtleties of the construction of the object we are trying to introduce; the interested reader is referred to [AKQ14] for the full details.

Random polymer models have a rich history in probability theory (see, for example, the book [Com17]), and they typically have three ingredients: a (discrete or continuous) space on which everything is defined, a random environment given by some kind of discrete or continuous space-time white noise, and a base measure on paths in the space. The model consists of specifying a way that paths picked according to the base measure are reweighted by their journey through the random environment to give a new measure. The random path picked according to this reweighted distribution is called a *polymer*, and the distribution a *polymer path measure*.

For the CDRP, the space we are working on is $[0,1]\times\mathbb{R}$. This should be thought of as an infinite strip of height one, with the height variable indexing time. The random environment ξ is space-time white noise, i.e., the random generalized function defined earlier by (1.2); we will denote the law of ξ by Q. Finally, the base measure on paths is given by the law of Brownian bridge (with specified endpoints $(0,0)$ and (t,x) with $t\in(0,1)$), a probability measure on the space of continuous functions on $[0,t]$ which we will denote $\boldsymbol{P}_{t,x}$; the fact that paths are continuous *functions* is the source of the adjective "directed" in CDRP: paths have a unique position at each time (i.e., height) and so cannot backtrack, unlike in FPP.

For a fixed endpoint (x,t), the joint distribution $\mathbb{P}_{t,x}$ of random environment ξ and polymer Γ is defined by a reweighting of the Brownian bridge (with endpoints $(0,0)$ and (x,t)) measure, and

is formally given by

$$\mathbb{P}_{t,x}(\mathrm{d}\xi, \mathrm{d}\Gamma) = \frac{1}{\mathcal{Z}_\beta} \exp\left\{\beta \int_0^1 \xi(s, \Gamma_s)\, \mathrm{d}s\right\} \boldsymbol{P}_{t,x}(\mathrm{d}\Gamma)\, Q(\mathrm{d}\xi); \tag{1.4}$$

here $\mathcal{Z}_\beta = \mathcal{Z}_\beta(t,x)$ is a normalization constant, known as the *partition function*, such that the righthand side, conditionally on ξ, is a (ξ-dependent, so random) measure with total mass one, i.e.,

$$\mathcal{Z}_\beta(t,x) = \int \exp\left\{\beta \int_0^1 \xi(s, \Gamma_s)\, \mathrm{d}s\right\} \boldsymbol{P}_{t,x}(\mathrm{d}\Gamma).$$

In words, (1.4) says that paths are given an exponential reweighting based on their integral through the random environment that they traverse, modified by a multiplicative factor β, which can be thought of as an interaction strength parameter. This parameter is typically called an *inverse temperature*, for this is how its effect behaves: when β is low, there is not much of a reweighting effect in (1.4), and the base pure randomness of the Brownian bridge measure $\boldsymbol{P}_{t,x}$ is dominant, as might be expected on physical grounds if the temperature is high and entropy reigns supreme; on the other hand, if β is high, paths which are well-suited to the environment (i.e., pass through regions where ξ is large) are exponentially more preferred, which can be interpreted physically as a low temperature phenomenon, when low energy ground states are favored.

We say (1.4) is a *formal* description because it is difficult to make sense of the integral of white noise along a Brownian bridge path, and even more so to interpret the exponential of that integral; again, the issue is that we are dealing with non-linear functions of the generalized function that is white noise. Further, it is shown in [AKQ14] that the polymer path measure of the CDRP is *singular* with respect to Brownian bridge, and so a Radon-Nikodym derivative, or measure reweighting, description as in (1.4) cannot hold. Nevertheless, the above is the intuition for the model.

Now we may briefly point out the connection of the CDRP to the KPZ equation. In statistical physical models, there is an important quantity known as the *free energy* $\mathcal{F}$ which is defined as the logarithm of the partition function:

$$\mathcal{F}_\beta(t,x) = \log \mathcal{Z}_\beta(t,x).$$

If we take $\beta = 1$, it turns out that the partition function for the CDRP, i.e., $\mathcal{Z} = \mathcal{Z}_{\beta=1}$ solves the multiplicative stochastic heat equation (1.3). Thus, recalling the definition of the Cole-Hopf solution, it follows that the free-energy of the CDRP solves the KPZ equation. (For general β, $\mathcal{Z}_\beta$ solves the SHE with coefficients depending on β; by linear transformations, we can reduce back to the standard SHE given in (1.3), and so the general case of any fixed β is essentially captured by the $\beta = 1$ case. The same is not true if we wish to consider $\mathcal{Z}_\beta$ as a function of β, and we will say something about that topic in Section 1.4.)

1.3 Last passage percolation

Having introduced the CDRP with general β parameter, it is natural to consider the extreme cases. The $\beta = 0$ case is of no interest, as there is no interaction between the path and the environment. The other extreme, $\beta = \infty$, is, however, extremely rich (though in different but related settings), and we turn to it now; this is the *zero temperature* case.

Let us look at (1.4) again. Remember that all the manipulations being performed here are purely formal. With that in mind, we see that, taking β large in (1.4), the path Γ which maximizes the weight through the environment, i.e., maximizes $\int_0^1 \xi(s, \Gamma_s)\, \mathrm{d}s$, has outsize contribution of mass to the polymer path measure defined by the path-marginal of $\mathbb{P}$; in fact, in the limit of $\beta \to \infty$, the marginal path measure should (intuitively) concentrate on that single maximizing path. Observe that, assuming such a maximizing path is unique, there is no longer any randomness in the selection of path given the environment. We have come upon a model which falls in the class of models known as *last passage percolation* (LPP), which is the family of models that this thesis will be studying.

(Unfortunately, while the CDRP can be rigorously constructed in the case of finite β, the same has not been done in the case of infinite β. Nevertheless, we will continue, for the time being, to use the same intuitive picture of maximizing weights of continuous paths through a continuum space-time white noise, before later introducing discrete analogues that are well-defined. The troubled reader may wish to keep instead in the background a very fine mesh of i.i.d. random variables and appropriately discretized paths instead of white noise in the continuum and arbitrary continuous functions as paths.)

Here is a summary of the heuristic continuum LPP model we have arrived at. We work on an infinite strip of height one, with the height indexing time, and an associated white noise. We

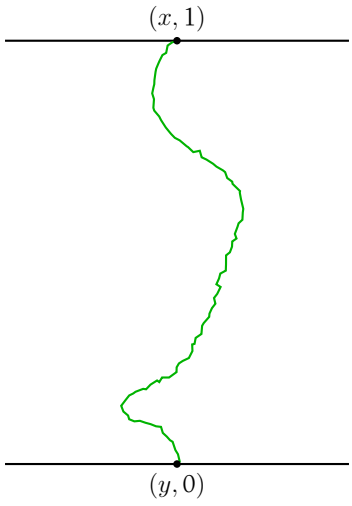


Figure 1.3: A depiction of a geodesic in the continuum LPP model.

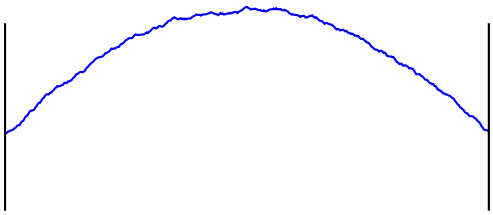


Figure 1.4: A depiction of the parabolic Airy_2 process.

consider directed paths which are continuous functions $\gamma : [0, 1] \to \mathbb{R}$. Each path is assigned a random weight given by its integral through the white noise environment. We fix two endpoints on the boundaries of the strip, $(y, 0)$ and $(x, 1)$, and maximize the weight over all paths with those two endpoints. For reasons of convention, the maximizing path is usually no longer called a polymer in the zero temperature case, but instead a *geodesic* (in spite of the fact that the object is defined by a maximization rather than a minimization as in the usual definition of geodesics in metric spaces). See Figure 1.3.

We emphasize again that this is *not* a rigorously defined model, in part due to the difficulty of defining the maximization over all continuous paths of a weight through the singular space-time white noise; any naive definition will lead to a result of ∞. We will introduce discrete analogues which are well-defined shortly, but the continuum version will be useful to explain some of the limiting objects that we are also interested in.

The parabolic Airy_2 process

The first of these limiting objects is known as the *parabolic Airy_2 process*, denoted by $\mathcal{P}_1 : \mathbb{R} \to \mathbb{R}$ (the reason for the subscript one will be explained later). In terms of the continuum LPP model, $\mathcal{P}_1(x)$ is the weight of the geodesic with starting point $(0, 0)$ (i.e., $y = 0$ in our earlier discussion) and endpoint $(x, 1)$; for this reason, we will think of it as a *weight profile*.

The parabolic Airy_2 process was originally defined by Prähofer and Spohn [PS02], who showed that it is the scaling limit of an analogous prelimiting weight profile in a certain discrete LPP model. It has a number of interesting properties. Firstly, it is called parabolic because it has a global downward parabolic curvature; see Figure 1.4. On adding back in the parabola x^2, the resulting process is stationary, i.e., $x \mapsto \mathcal{P}_1(x) + x^2$ is stationary; this process is known as the *Airy_2 process* $\mathcal{A}_1$.

The one-point distribution of $\mathcal{A}_1$ is the GUE Tracy-Widom distribution. This distribution was discovered nearly a decade before the Airy_2 process, in the seemingly very different context of random matrix theory [TW94]; we will see in Chapter 2 that there are in fact some rather remarkable connections between random matrices and certain LPP models underlying this coincidence. More precisely, the GUE Tracy-Widom distribution is the scaling limit of the fluctuations of the largest eigenvalue of the Gaussian Unitary Ensemble (GUE) (for the unfamiliar but interested reader, this will also be defined in Chapter 2).

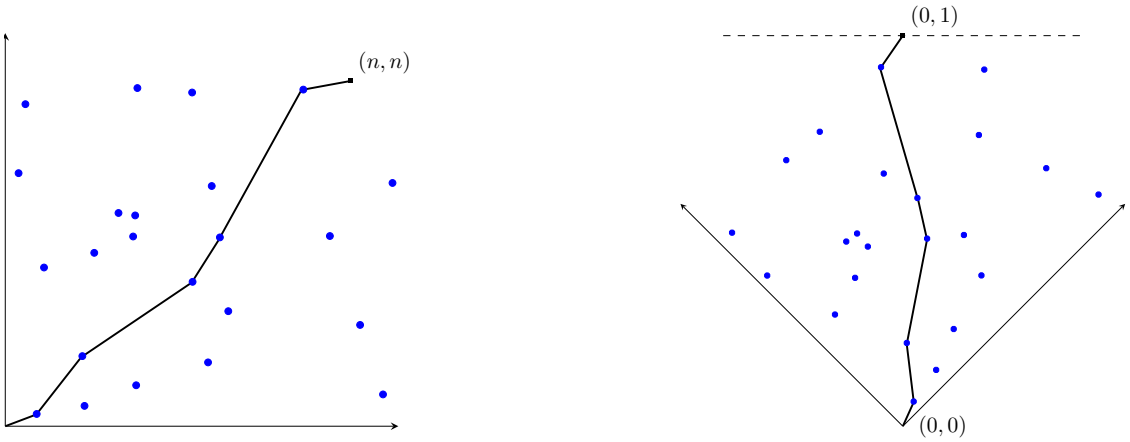


Figure 1.5: Poissonian LPP, and the rotation and vertical scaling to fit into the heuristic continuum LPP model. The blue dots are the Poisson points, and in black is an up-right path from $(0, 0)$ to (n, n); its weight is 5 as it passes through 5 blue points. The dashed horizontal line in the right panel is the line on which the endpoint of the geodesic varies to give a weight profile which converges to $\mathcal{P}_1$ under an appropriate scaling.

This information and more about the parabolic Airy_2 process was obtained via integrable connections which we will indicate soon. In fact, the discovery of the GUE Tracy-Widom distribution in the context of LPP preceded the discovery of the parabolic Airy_2 process—though both occurred in the exactly solvable model of *Poissonian LPP*.

Our first discrete model: Poissonian LPP

Poissonian LPP is defined as follows. We work in $\mathbb{R}^2$. The random environment is given by a rate one Poisson point process on $\mathbb{R}^2$, and directed paths are piecewise linear paths which move in an up-right manner, i.e., have north-east direction. The weight of a path is the number of Poissonian points it passes through. The LPP weight between say $(0, 0)$ and (x, y) with $x, y > 0$ is the maximum weight over all up-right paths between the endpoints, i.e., the maximum number of Poisson points which can be picked up by up-right paths; see Figure 1.5.

This is related to the heuristic continuum LPP model from above by a rotation and scaling. More precisely, consider the region between the lines $x+y = 0$ and $x+y = 2n$. If the model is rotated by 45° so that the $x + y = 0$ line is horizontal, and scaled in the resulting vertical direction by $2n$, so that the original $x + y = 2n$ line is at height 1, then one is in the analogue of the continuum LPP setup (modulo an important horizontal scaling which we have not yet touched upon); note that up-right paths in the original setup are piecewise linear *functions of the height* after rotation. However, for the remaining discussion we will stick with the original, unrotated, coordinate system.

Incidentally, suppose one fixes the square $[0,1]^2$ and conditions on there being n Poisson points inside it. Then, if one orders the points by their x-coordinates and considers the permutation obtained by looking at the ranks of their y-coordinates, it is easy to see that it is distributed as a *uniform* random permutation σ_n on $\{1,\dots,n\}$ (because under the conditioning, the set of Poisson points has the distribution of n i.i.d. samples from the uniform distribution on $[0,1]^2$). Further, the LPP weight from $(0,0)$ to $(1,1)$ is exactly the length of the longest increasing subsequence of σ_n, i.e., the largest ℓ such that there are $i_1 < i_2 < \ldots < i_\ell$ with $\sigma(i_1) < \ldots < \sigma(i_\ell)$.

This is the first hint of integrable structure in this model. For there is a classical bijection on permutations known as the Robinson-Schensted correspondence which maps a permutation on $\{1,\dots,n\}$ to a pair of standard Young tableaux of the same shape (see Figure 1.6); further, the length of the first row of these tableaux is precisely the length of the longest increasing subsequence of the original permutation. (An important generalization by Knuth, known as the Robinson-Schensted-Knuth, or RSK, correspondence, will be briefly returned to later.)

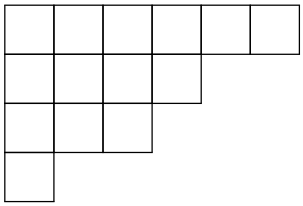


Figure 1.6: A Young diagram; the defining property is that the length of the rows is decreasing (in general, non-increasing) as we go down. A (standard) Young tableau is one where the diagram is filled with positive integers which are increasing in each row and column, but for the purposes of our exposition here it suffices to consider only the shape of the Young tableau, i.e, the Young diagram.

This is discussed with great clarity and much greater detail in the book [Rom15]. But for us, we observe that a statistic of interest in this model, namely the LPP value, is encoded in an explicit fashion as the length of the top row of a concrete algebraic object picked according to a certain measure (corresponding to the Poisson randomness of the points). This is what makes Poissonian LPP an exactly solvable model.

It was by making use of this algebraic connection that Baik, Deift, and Johansson proved in a breakthrough paper [BDJ99] that the LPP value from $(0,0)$ to (n,n), suitably centred and scaled, converges in law to the GUE Tracy-Widom distribution.

Theorem 1.1 (Baik-Deift-Johansson). *Let X_n be the LPP value from $(0,0)$ to (n,n) and F_{TW} be the GUE Tracy-Widom distribution. As $n \to \infty$,*

$$\frac{X_n - 2n}{n^{1/3}} \to F_{\mathrm{TW}},$$

where the convergence is in distribution.

The KPZ scaling exponents

Observe that, in Theorem 1.1, the scaling of X_n, a quantity growing linearly in n to first order, is $n^{1/3}$, instead of the more classical $n^{1/2}$ of the central limit theorem. In other words, the fluctuations of X_n are on the scale of $n^{1/3}$. This $\frac{1}{3}$ is the first of the characteristic KPZ scaling exponents (which we have taken an anomalously long time to come to in this introduction in comparison to most overviews of KPZ). It is often called the longitudinal fluctuation exponent and denoted χ; we will instead refer to it as the *weight fluctuation* exponent.

The second characteristic KPZ scaling exponent is the *transversal fluctuation* exponent, usually denoted ξ. To see where it arises, recall that we have seen that the GUE Tracy-Widom distribution is the one-point distribution of the stationary Airy_2 process, and that it is also the scaling limit of the LPP value to (n, n) in Poissonian LPP. We have not yet explored the obvious avenue: what is the statistic of Poissonian LPP that converges to the Airy_2 process? (In fact, the more natural limiting object will be the parabolic Airy_2 process $\mathcal{P}_1$.) It is in describing the correct spatial scalings for this functional convergence that we will need ξ.

The statistic of Poissonian LPP that will converge to $\mathcal{P}_1$ is the weight profile with fixed starting point $(0, 0)$ and ending point varying on the line $x + y = 2n$; in the right panel of Figure 1.5, this is the horizontal line through $(0, 1)$ (after vertical scaling). The convergence will hold after a spatial, horizontal scaling by n^{ξ}—in other words, this is the scaling at which correlations are non-trivial. In $1 + 1$ dimensional KPZ models, the universal value of the transversal fluctuation exponent predicted is

$$\xi = \frac{2}{3}.$$

(Another way of thinking of this transversal fluctuation exponent is that the width of the smallest diagonal strip that contains the geodesic from $(0, 0)$ to (n, n) will typically be of order $n^{2/3}$: see Figure 1.7.)

Returning to the interpretation of the transversal fluctuation exponent as describing the scale of non-trivial correlations, consider the map

$$x \mapsto L_n(x) := n^{-1/3} \left(X_{n+xn^{2/3}, n-xn^{2/3}} - 2n \right),$$

where $X_{x,y}$ is the LPP value from $(0, 0)$ to (x, y) and the map is defined for $|x| \leq n^{1/3}$. Then, one meaning of the above statement that $n^{2/3}$ is the right spatial scale for non-trivial correlations is that

$$\limsup_{n\to\infty} \operatorname{Corr}\big(L_n(x), L_n(y)\big) \in (0, 1)$$

for all fixed $x, y \in \mathbb{R}$, and the same with lim inf in place of lim sup. (The Fortuin–Kasteleyn–Ginibre, or FKG, correlation inequality from percolation theory tells us that the correlation must be non-negative.)

With L_n defined, we can state the convergence theorem to $\mathcal{P}_1$, proved in [PS02].

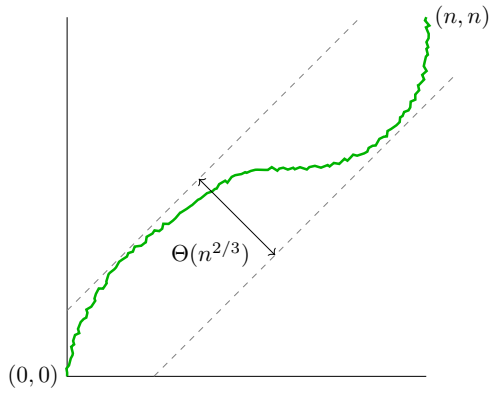


Figure 1.7: In green is the geodesic from $(0,0)$ to (n,n). One interpretation of the statement that the transversal fluctuation exponent $\xi = \frac{2}{3}$ is that the width of the minimum width diagonal strip which contains the geodesic, here depicted in dashed light gray lines, is typically of order $n^{2/3}$.

Theorem 1.2 (Prähofer-Spohn)**.** *The following holds in the sense of convergence of finite dimensional distributions:*

$$\lim_{n\to\infty} L_n(x) = \mathcal{P}_1.$$

This was later improved to a functional convergence, i.e., distributional convergence as processes, by Johansson [Joh03], though in a discrete variant of the Poissonian model.

A heuristic for $\xi = \frac{2}{3}$

We pause here to give a quick heuristic for why $\xi = \frac{2}{3}$ is the transversal fluctuation exponent. Suppose we know that, to first order, the weight *loss* to $(n+x, n-x)$ compared to (n,n) is $O(x^2/n)$, as it is in all known integrable LPP models and should be anticipated by the parabolic curvature of $\mathcal{P}_1$. (However, this statement has also not yet been shown in any non-integrable model.) Because the underlying randomness of the environment is i.i.d., it is natural to believe that the weight profile should exhibit Brownian features, at least on a local scale before the parabolic curvature becomes powerful. Indeed, a main result of this thesis will be proving a precise and quantitative version of this statement. Then, when moving a distance of x from (n,n), we have two opposing forces: a weight loss of order x^2/n, and a Brownian weight fluctuation of $x^{1/2}$. We see that the scale of x where these are of the same order is $x \approx n^{2/3}$.

The KPZ relation

There is a famous relation, called the KPZ relation, that is expected to hold between χ and ξ for stochastic growth models in *all* dimensions. This is that

$$\chi = 2\xi - 1. \tag{1.5}$$

This is of course true for the integrable models in $1+1$ dimensions that we have been discussing. In arbitrary dimensions, it has been shown to hold for FPP under certain quite strong assumptions [Cha13] (see also [AD14]) which have till date not been verified for any particular distribution of the edge weights. There is, however, no consensus on even the level of conjecture about what the individual values of χ and ξ should be in any dimension higher than two, and even suggestions that, for high enough dimension, $\chi = 0$. See [ADH17, Section 3.1] for some brief discussion and references.

The other rows of the Young tableaux

We've seen that the length of the *top* row of the pair of Young tableaux associated to the Poisson point environment encodes information which is very valuable for our purposes. What of the remaining rows?

In the context of permutations, and as a question about the Robinson-Schensted correspondence, this was answered by Greene [Gre74]: given a permutation σ, the *sum* of the lengths of the top k rows of the associated Young diagram is exactly the size of the longest subsequence of σ which is the union of k increasing subsequences of σ.

In Poissonian LPP, this has a nice equivalent meaning. Consider a collection of k *disjoint* paths in the environment. We ascribe such a collection a weight, namely the sum of the weights of the constituent curves. Then the maximum weight over all such collections of k disjoint paths is the sum of the lengths of the top k rows of the associated Young diagram.

(Note that the best collection of k disjoint paths need not include the highest weight single path, i.e., the geodesic. In particular, the collection is *not* formed by taking the geodesic, finding and including in the collection the highest weight path disjoint from the geodesic, and iterating.)

The LPP description also gives a nice interpretation to the fact that the Young diagram row lengths are non-increasing: the amount of weight *gained* by giving oneself $k+1$ disjoint paths instead of k can be no more than the same gained by going from $k-1$ paths to k. This is true by definition when comparing the second and first rows (as the first row length is the weight of the best path). While plausible as a sort of law of diminishing returns, it is, however, not immediate for larger k. A proof which works with the LPP paths directly (in a related LPP model) and does not make use of the RS (or RSK) correspondence is given in [BGHH20], a paper whose contents are not entirely included in this thesis but on which I was a coauthor; it will be discussed briefly in Section 2.6.

The Airy point process and line ensemble

Now, as we saw earlier, Baik, Deift, and Johansson proved that the length of the top row of the random Young tableaux picked according to the measure induced by the Poisson randomness of points converges, in law, to the GUE Tracy-Widom distribution (after a certain centring and scaling), and that the weight profile of Poissonian LPP converges to $\mathcal{P}_1$. It is natural to wonder what happens to the lengths of the top k rows jointly.

The GUE Tracy-Widom distribution is the scaling limit of the top eigenvalue of the Gaussian Unitary Ensemble. Might the scaling limit of the top k row length be the scaling limit of the top k eigenvalues of GUE?

The scaling limit of the latter is a random point process known as the Airy point process. Its top point is thus distributed according to the GUE Tracy-Widom distribution. The Airy point process is a *determinantal* point process, with kernel $K^{\mathrm{Ai}}(x,y)$ given by

$$K^{\mathrm{Ai}}(x,y) = \begin{cases} \dfrac{\mathrm{Ai}(x)\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\mathrm{Ai}(y)}{x-y} & x \neq y \\ \mathrm{Ai}'(x)^2 - x\mathrm{Ai}(x)^2 & x = y, \end{cases}$$

where $\mathrm{Ai} : \mathbb{R} \to \mathbb{R}$ is the classical Airy function.

In essence, this means that the probability (rather, joint density) of finding points of the point process at each of the locations $x_1, \ldots, x_k$ is given by $\det(K^{\mathrm{Ai}}(x_i, x_j))_{i,j=1}^k$ for each $k \in \mathbb{N}$. A good introduction to determinantal point processes is [HKPV09]. This determinantal structure, and the appearance of the classical Airy function, is another indication of the exactly solvable nature of the objects we are encountering.

Given this description of the GUE Tracy-Widom distribution and the analogy between row lengths of random Young diagrams and eigenvalues of the GUE, [BDJ99] conjectured that the lengths of the top k rows too jointly converge to the top k points of the Airy point process. This was proved shortly afterwards, independently, by Borodin, Okounkov and Olshanski [BOO00]; Johansson [Joh01]; and Okounkov [Oko00].

We now recall that we originally introduced the GUE Tracy-Widom distribution not on its own, but as the one point distribution of the stationary process known as the Airy_2 process $\mathcal{A}_1$. To complete the picture we've been sketching, we next introduce the *Airy line ensemble* $\mathcal{A}$. This is an infinite ($\mathbb{N}$-indexed) collection of continuous, non-intersecting curves $\mathcal{A}_1 > \mathcal{A}_2 > \ldots$ which is stationary and whose one point distribution is the Airy point process.

In the heuristic LPP model defined previously in the continuum, the weight profile from fixed starting point as the ending point varied was given by the parabolic Airy_2 process $\mathcal{P}_1$, which was $x \mapsto \mathcal{A}_1(x) - x^2$. So, if we subtract a parabola from the Airy line ensemble $\mathcal{A}$, we obtain the parabolic Airy line ensemble $\mathcal{P}$, whose top indexed curve is $\mathcal{P}_1$; see Figure 1.8. This also explains the reason for the subscript of 1 for $\mathcal{P}_1$.

The parabolic Airy line ensemble was described in terms of its finite dimensional distributions in [PS02] by the name of the multi-line Airy process; the finite dimensional distributions can be described via a determinantal point process, with kernel known as the extended Airy kernel. That there exists a law on the space of infinite collections of continuous curves which has these finite dimensional distributions, and such that samples from that law are non-intersecting, is not trivial. The construction of this law, i.e., of $\mathcal{P}$, was undertaken in the paper [CH14].

Taking the interpretation of the Young diagram row lengths in Poissonian LPP to the limit, we can extend the interpretation of $\mathcal{P}_1(x)$ as the weight of the geodesic from $(0,0)$ to $(x,1)$ to the

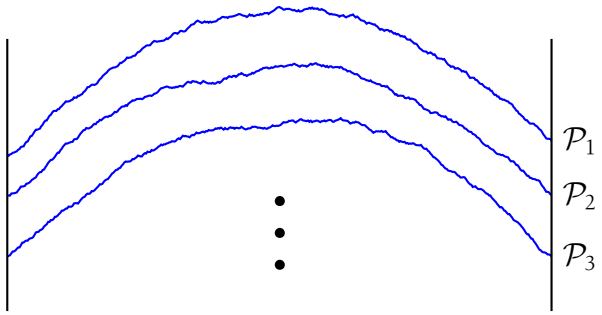


Figure 1.8: A depiction of the parabolic Airy line ensemble.

other curves. In particular, we see that the sum of the top k values of $\mathcal{P}(x)$ for a given x should be thought of as the weight of the best collection of k disjoint paths from $(0,0)$ to $(x,1)$ in the heuristic continuum LPP model. (In fact, this has recently been proven [DZ21] for a rigorously defined version of the continuum LPP model—which is different from the heuristic model in that the underlying noise environment is a specific distribution which is *not* i.i.d.)

Part II of this thesis will be devoted to a study of the parabolic Airy line ensemble.

1.4 The KPZ equation and the KPZ fixed point

Before closing out this chapter, we pause and return to elucidate a connection we skipped over earlier: namely, what is the precise relation between the positive temperature object, the KPZ equation, and the various zero temperature objects and models we have encountered, such as last passage percolation, the parabolic Airy$_2$ process, and the parabolic Airy line ensemble?

There is an obvious similarity between the continuum directed random polymer and the continuum LPP model which was stressed before. This suggests that some statistic of the continuum directed random polymer with fixed starting point (say 0) should, in some sense, converge to the weight profile in the continuum LPP model with the same fixed point, i.e., $\mathcal{P}_1$.

In fact, the statistic is the free energy of the CDRP, which, recall, is described by the Cole-Hopf solution to the KPZ equation (with narrow wedge initial condition, since the polymers are fixed to start from 0). That the free energy converges to $\mathcal{P}_1$ was first conjectured in [ACQ11, Conjecture 1.5], later extended in [CH16, Conjecture 2.17], and recently proven independently by Quastel-Sarkar [QS20] and Virág [Vir20].

Theorem 1.3 (Quastel-Sarkar, Virág)**.** *Let $\mathcal{H} : [0,\infty)\times\mathbb{R} \to \mathbb{R}$ be the solution to the KPZ equation (1.1) with narrow-wedge initial condition. Then,*

$$2^{1/3}\alpha^{-1/3}\left(\mathcal{H}(\alpha, 2^{1/3}\alpha^{2/3}x) + \frac{\alpha}{24}\right) \to \mathcal{P}_1(x) \tag{1.6}$$

as $\alpha \to \infty$, where the convergence is in distribution as continuous processes on $\mathbb{R}$ in the topology of uniform convergence on compact sets.

(We have switched from the temporal argument being represented by t to α for a reason which will become clear ahead.)

In fact, the spatial scaling factor of $2^{1/3}$ was not included in the aforementioned conjectures, and it is not completely clear from the formulations of the theorems in [QS20, Vir20] whether it has been taken into account. However, a scaling factor of this sort is necessary for this convergence to hold, as we explain now.

The solution to the KPZ equation (1.1) can be expected to be in some sense locally Brownian, with a diffusion rate of *one*; heuristically, this is due to the presence of the space-time white noise of rate one. Further, it is known (eg. [CH16, Proposition 1.17]) that

$$\alpha^{-1/3}\left(\mathcal{H}(\alpha,\alpha^{2/3}x)+\frac{\alpha}{24}\right)+\frac{x^2}{2}$$

is stationary in x. These two facts should be contrasted with their analogues for $\mathcal{P}_1(x)$: first, $\mathcal{P}_1$ is locally Brownian (again in a sense which we have not yet specified) of rate *two*, and second, $\mathcal{P}_1$ needs an addition of x^2, not $x^2/2$, to be stationary. Thus, without the spatial scaling factor, there is no reason that the KPZ equation should change its local diffusion rate from 1 to 2 and its asymptotic decay rate from $-x^2/2$ to $-x^2$ in the limit.

Including the scaling factor of $2^{1/3}$ in the spatial argument resolves both these issues. Indeed, let $\mathcal{H}^{\mathrm{sc},\alpha}$ be the scaled narrow-wedge solution to the KPZ equation given by

$$\mathcal{H}^{\mathrm{sc},\alpha}(t,x)=r^{-1/2}\alpha^{-1/3}\left(\mathcal{H}(\alpha t,2\alpha^{2/3}rx)+\frac{\alpha t}{24}\right),$$

where r is a scaling parameter to be set. For any value of r, this rescaled object has a locally Brownian rate of two. We need to find r which is also such that the rescaled object decays like $-x^2$. In other words, since $\mathcal{H}^{\mathrm{sc},\alpha}(1,x)$ decays like $r^{-1/2}\times 4r^2x^2/2$, we need

$$2r^{3/2}=1 \;\Rightarrow\; r=2^{-2/3}.$$

The lefthand side of (1.6) is exactly $\mathcal{H}^{\mathrm{sc},\alpha}(1,x)$ with this value of r, and we now define $\mathcal{H}^{\mathrm{sc},\alpha}$ to always be associated with $r=2^{-2/3}$.

The KPZ fixed point

Having explained the scaling factors, let us take a look again at (1.6). It can be interpreted in two ways. The first is that we are simply taking time to infinity. The second is that we are fixing time to be 1 (which, recall, is also the height of the strip in both the CDRP and the heuristic continuum LPP model) and performing a certain rescaling of the noise environment and path measure; in other words, we are fixing time to be 1 but taking the inverse temperature parameter β to ∞. Indeed, the second perspective is more useful, as it suggests we also consider the following: what may be the limit, as $\alpha\to\infty$, of $\mathcal{H}^{\mathrm{sc},\alpha}(t,x)$, as a space-time process?

A second natural question arises by considering (1.6) again. That convergence was stated for the initial condition of the KPZ equation being narrow-wedge. What happens for other initial conditions?

For the CDRP, the interpretation of a general initial condition $x \mapsto \mathcal{H}(0,x)$ is the following. Instead of forcing the polymer to begin at $(0,0)$, it may begin at any point $(x,0)$; depending on its starting location, a path is given a reward determined by $\mathcal{H}(0,x)$, which thus influences the polymer path measure.

We can now give a unified answer to both the just introduced questions in terms of our heuristic continuum LPP model. The first is simpler: intuitively, instead of restricting our infinite strip to have height 1, we should allow any height $t \in (0,\infty)$ to describe the limit of large α of $\mathcal{H}^{\mathrm{sc},\alpha}(t,x)$. To answer the second, we need to define a notion of initial condition for the LPP model.

Let $\mathfrak{h}_0 : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ be a given deterministic function. The augmented weight of a path starting at $(y,0)$ in the continuum LPP model will be $\mathfrak{h}_0(y)$ plus its usual weight determined by its journey through the noise environment. (We allow $\mathfrak{h}_0$ to take the value $-\infty$ so that certain starting points can be excluded, as in, for example, the narrow-wedge case.) Then, the LPP problem is to maximize the augmented weight overall from the *line* at height 0 to the specified endpoint. So paths are essentially given a (possibly negative) reward described by $\mathfrak{h}_0$ based on their starting point.

The endpoint need not be at height 1, but can be (x,t) for any $x \in \mathbb{R}$ and $t \in (0,\infty)$. Let us call the corresponding LPP value to be $\mathfrak{h}_t(x)$, which depends on the choice of $\mathfrak{h}_0$. The space-time process $(x,t) \mapsto \mathfrak{h}_t(x)$ is known as the *KPZ fixed point*. It is expected to be a universal scaling limit, parametrized by initial conditions, for all models in the KPZ universality class. For narrow wedge initial condition, $\mathfrak{h}_1$ is the same as $\mathcal{P}_1$ (though the subscripts of one have a different meaning in each).

The KPZ fixed point was first constructed as the scaling limit of a particular model in the KPZ class, known as TASEP (which is equivalent to a particular LPP model, exponential LPP, that we will introduce in the next chapter), by Matetski-Quastel-Remenik [MQR17].

With this, we may state the full convergence theorem of the KPZ equation proved by Quastel-Sarkar and Virág.

Theorem 1.4 (Quastel-Sarkar, Virág)**.** *Let $\mathcal{H} : (0,\infty)\times\mathbb{R} \to \mathbb{R}$ be the solution to the KPZ equation (1.1) with initial condition $\mathcal{H}(0,\cdot) : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$. Then,*

$$\mathcal{H}^{\mathrm{sc},\alpha} \to \mathfrak{h}$$

as $\alpha \to \infty$, where $\mathfrak{h}_0 = \lim_{\alpha\to\infty} \mathcal{H}^{\mathrm{sc},\alpha}(0,\cdot)$, and the asserted convergence is in the sense of finite dimensional distributions.

A stronger space-time process convergence is also expected but has not yet been achieved; for fixed time, a process level convergence of the spatial marginal has been proved in the above result.

We can give a slightly more precise description of $\mathfrak{h}_t$. Suppose the LPP value from $(y, 0)$ to (x, t) in the continuum LPP model is given by $\mathcal{L}(y, 0; x, t)$. From the discussion of the LPP interpretation of $\mathfrak{h}_t$ above, we see that the following variational formula should hold:

$$\mathfrak{h}_t(x) = \sup_{y \in \mathbb{R}} \Big\{ \mathfrak{h}_0(y) + \mathcal{L}(y, 0; x, t) \Big\}.$$

In fact, this is true in a rigorous sense, as has been shown quite recently. Originally, however, the construction in [MQR17] of $\mathfrak{h}$ did not yield a variational formula, because it was at the time an open problem to rigorously construct $\mathcal{L}$ as a scaling limit of any LPP model. The construction of $\mathcal{L}$—now called the *directed landscape*—was achieved in the breakthrough paper [DOV18], aided by [DV18], in the model of Brownian LPP, which we will introduce in the next chapter. Finally, the resulting variational formula and the earlier definition of $\mathfrak{h}$ from [MQR17] were shown to be equivalent in [NQR20].

Conclusion of Chapter 1

We now come to the end of this chapter introducing the KPZ universality class. We focused mainly on integrable models and hinted at integrable techniques. But while the perspective of this thesis is not an integrable one, along the way, we have also encountered the main objects that this thesis will study: the parabolic Airy line ensemble, the KPZ fixed point, the GUE Tracy-Widom distribution, and discrete LPP models. In the next chapter we will discuss the probabilistic and geometric viewpoint on the study of KPZ and explain our main results.

Chapter 2

Probabilistic and geometric viewpoints

In the previous chapter we introduced a number of objects, most of which were mathematically tractable to analyze because of integrable connections. In this chapter, we will adopt instead a probabilistic and geometric perspective, one which has informed much of the work in this thesis, as well as introduce our main results.

This alternative viewpoint will consist of two approaches. The first is to initially use integrable information to identify convenient probabilistic features of models, and then pursue all subsequent analysis using only these probabilistic methods. The second is to investigate the implications of assumptions in certain models, like last passage percolation, which can currently be verified in only the integrable cases.

In both approaches, the underlying theme or aim is to develop techniques which can achieve results beyond the reach of integrable methods. For example, techniques which are robust may be applicable in models defined only by assumptions which create a more general, non-integrable framework, and may eventually be applicable to a wider class of models, if at a later date we have enough of an understanding to resolve some of the fundamental difficulties which plague all non-integrable KPZ models (such as the lack of proof of the weight fluctuation exponent being $\frac{1}{3}$, or of the parabolic curvature of the weight profile).

This thesis broadly consists of three pillars. The first is a study of the parabolic Airy_2 process using the parabolic Airy line ensemble and a valuable probabilistic resampling property it possesses. The second is an investigation of aspects of LPP values using geometric properties of geodesics. Finally, the third combines aspects of the first two to study the KPZ fixed point. These three pillars form Parts II, III, and IV. While the terms "probabilistic" and "geometric" are not unrelated in our context, in this discussion we will use them to respectively refer to the first and second pillars.

The probabilistic and geometric viewpoint on KPZ largely came into its own in the last decade. Complementing our first two pillars, we focus on two streams: the first was initiated by Corwin and Hammond in [CH14], where they introduced the *Brownian Gibbs* property to construct and study the parabolic Airy line ensemble; the second by Basu, Sidoravicius, and Sly in [BSS14], where they wielded an understanding of geodesic geometry in last passage percolation to settle

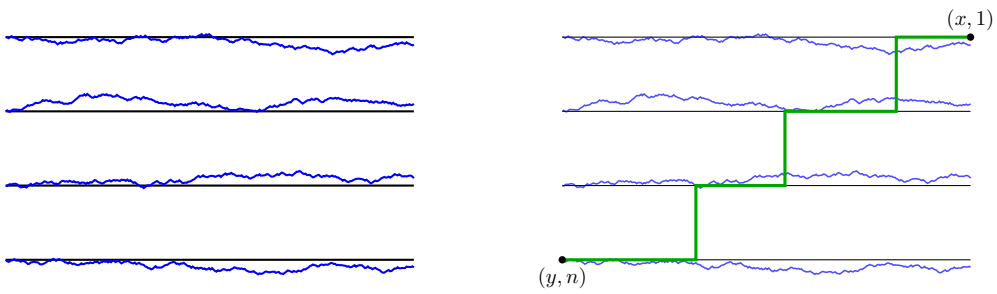


Figure 2.1: Left: A depiction of the environment given by independent Brownian motions B. The functions B_i corresponding to each line are graphed in red on the corresponding black line for visual clarity; the function values themselves need not be ordered. Right: An upright path γ from (y, n) to $(x, 1)$ is depicted in green; note that in a formal sense the depicted vertical portions are not part of the path. The path's weight is the sum of the increments of f_i along the portion of the i^{th} line γ spends on it.

the so-called *slow bond problem*.

We start by discussing the first stream in Section 2.1. We will discuss the results of this thesis that fall in this stream in Section 2.2, and then turn to background on the geometric stream and results in that stream in Sections 2.3 and 2.4. Results that form the third pillar of the thesis, which combine the geometric and probabilistic viewpoints, will be discussed in Section 2.5. Finally, we will briefly discuss some other work undertaken by me during graduate school that is not included in this thesis in Section 2.6.

2.1 Background on the probabilistic approach

Brownian LPP

The Brownian Gibbs property is an explicit and important resampling property enjoyed by the parabolic Airy line ensemble. To motivate it, we introduce an integrable LPP model known as *Brownian LPP* which is the setting for one of our main results and where we can define a prelimiting version of the parabolic Airy line ensemble .

Brownian LPP is a semi-discrete LPP model, which here means that the space on which the noise and directed paths are defined is discrete in one dimension and continuous in the other. More precisely, let $n \in \mathbb{N}$ and $[\![1, n]\!]$ denote the integer interval $\{1, \dots, n\}$. Then the space is $\mathbb{R} \times [\![1, n]\!]$. This should be imagined as a collection of n horizontal lines, arranged vertically. The noise is a collection $B = (B_1, \dots, B_n)$ of i.i.d. two-sided rate one Brownian motions, i.e., one Brownian motion for each element of $[\![1, n]\!]$ and each defined on $\mathbb{R}$. These may be pictured as being associated to the horizontal lines. While these lines are depicted as vertically ordered, the values of the Brownian motions themselves do not have any such ordering, of course. We

label the horizontal lines in increasing index with height, i.e., the lowest is labeled n, and the highest 1. See Figure 2.1.

Directed paths are paths which are up-right, i.e., remain on a given horizontal line for some interval before moving up to the next line above. See again Figure 2.1. An up-right path is parametrized by its *jump times* $\{t_i\}_{i=1}^{n-1}$ at which it jumps from the $(i+1)^{\text{st}}$ line to the i^{th} line. Define $\Pi_{y,x}^n$ to be the set of up-right paths from (y,n) to $(x,1)$. The *weight* of an up-right path $\gamma \in \Pi_{y,x}^n$ in B is denoted $B[\gamma]$ and defined by

$$B[\gamma] = \sum_{i=1}^{n-1} \Big(B_i(t_{i-1}) - B_i(t_i) \Big),$$

where $\{t_i\}_{i=1}^{n-1}$ are the jump times of γ, with $t_n = y$ and $t_0 = x$. This expression is thus the sum of increments of B along the portions of γ on each line. As usual, the last passage value in B from (y,n) to $(x,1)$ is defined by maximizing over all up-right paths with those endpoints:

$$B[(y,n) \to (x,1)] = \sup_{\gamma \in \Pi_{y,x}^n} B[\gamma].$$

If the set $\Pi_{y,x}^n$ is empty, i.e., if $y > x$, we define the passage value to be $-\infty$.

(Alternatively, the noise can be thought of as a collection of independent one-dimensional white noises, which allows the model to fit in more directly to the description of LPP models given in Chapter 1, where paths were assigned a weight by integrating against the noise, since increments of Brownian motion are integrals of white noise. That is, with the white noise environment, the weight of the path is exactly its integral against the white noise. But the Brownian description is more standard.)

Recall that, in Poissonian LPP, while the LPP value was equal to the length of the top row of the associated Young diagram, the lengths of the lower rows were related to weights of maximal weight collections of disjoint curves. A similar equivalence will hold for Brownian LPP. To describe it, let $\Pi_{y,x}^{n,k}$ be the collection of sets of k disjoint curves from (y,n) to $(x,1)$; we do not impose the disjointness conditions at the common starting and ending point. Though the staring and ending points are common, it is possible to have paths which are disjoint except for those points because, for any j, a path can jump from line n to line $n-j$ (or from line j to line 1) at the very first instant y (or final instant x).

Now we define, for $k \in [\![1,n]\!]$,

$$B[(y,n)^k \to (x,1)^k] = \sup_{\boldsymbol{\gamma} \in \Pi_{y,x}^{n,k}} \sum_{i=1}^{k} B[\gamma_i],$$

to be the LPP value for k disjoint paths with the given endpoints, where $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_k)$.

Recall that in Poissonian LPP, it was the *sum* of the lengths of the rows of the Young diagram which corresponded to the weights of the maximal weight disjoint collections. Further, it was

the row lengths which had the scaling limit of the Airy point process. Here, the statistic corresponding to the Young diagram row lengths is given by

$$P_{n,k}(x) := B[(0,n)^k \to (x,1)^k] - B[(0,n)^{k-1} \to (x,1)^{k-1}], \tag{2.1}$$

for $k \geq 2$, with $P_{n,1}(x) = B[(0,n) \to (x,1)]$. Thus, just as with the row lengths of the Young diagram, we have

$$B[(0,n)^k \to (x,1)^k] = \sum_{i=1}^{k} P_{n,k}(x).$$

Also as in the Poissonian case, we have that $P_{n,1} \geq P_{n,2} \geq \ldots \geq P_{n,n}$.

In order to get the parabolic Airy line ensemble $\mathcal{P}$ in the limit, we will need to introduce the scalings corresponding to the weight and transversal fluctuations. We will return to this shortly, after discussing the exactly solvable nature of Brownian LPP.

GUE and Dyson's Brownian motion

Brownian LPP is an integrable model. Its integrability is due to its direct equivalence with the well-known random matrix model, the Gaussian Unitary Ensemble (GUE).

The size-n GUE is defined as a $n \times n$ random Hermitian matrix A, where the diagonal entries a_{ii} are i.i.d. normal random variables with mean zero and variance one, and the above diagonal entries a_{ij} with $i > j$ are i.i.d. *complex* normal random variables with mean zero and variance one, i.e., $X + \sqrt{-1}Y$ with X, Y i.i.d. real normal random variables with mean zero and variance $\frac{1}{2}$; the below diagonal entries are defined by $a_{ji} = \overline{a}_{ij}$ to give the Hermitian property. The "U" in "GUE" arises because the resulting distribution over matrices is invariant under conjugation by unitary matrices.

The remarkable fact is that there is a distributional equality between Brownian LPP and GUE: if λ_1 is the largest eigenvalue of a size-n GUE, then

$$P_{n,1}(1) = B[(0,n) \to (1,1)] \stackrel{d}{=} \lambda_1. \tag{2.2}$$

This was first proved in [Bar01, GTW01]. But even more amazingly, the identity extends to hold between much richer objects, and this is of great importance. To state this, we need to introduce Dyson's Brownian motion.

There are two descriptions of Dyson's Brownian motion. The first is as a dynamic on the eigenvalues of the GUE. We simply replace the real normal random variables (including the real and imaginary components of the complex normal entries) which were involved in the entries of the GUE by independent Brownian motions of diffusion rate equal to the variance of the entry (i.e., diffusion rate $\frac{1}{2}$ if the variance was $\frac{1}{2}$, and 1 if it was 1). This gives a *process* of random matrices. At any given time t, the marginal distribution of this process at that time is exactly the same as the GUE except scaled by $\sqrt{t}$ (since a variance t normal random variable has the same distribution as $\sqrt{t}$ times a variance one normal random variable).

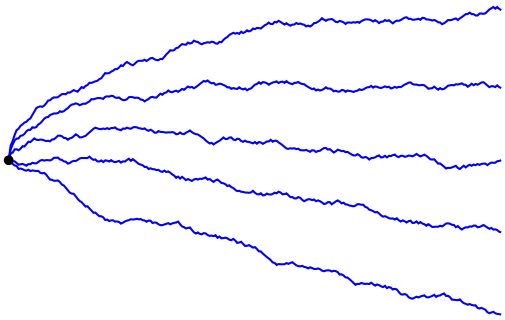


Figure 2.2: A depiction of Dyson's Brownian motion. A vertical slice at any fixed time gives n random points which have the same distribution as the eigenvalues of size-n GUE, except scaled by a factor of the square of the time.

The random matrix process induces a process on its eigenvalues: we get $\lambda_1(t) > \lambda_2(t) > \ldots > \lambda_n(t)$ by following the paths of the ordered eigenvalues as time proceeds. This eigenvalue process is exactly Dyson's Brownian motion. This was the original definition of Dyson's Brownian motion given by Dyson [Dys62]; it also has a description in terms of a system of stochastic differential equations. A depiction is given in Figure 2.2. Now, the righthand side of (2.2) is $\lambda_1(1)$, the top line of n-level Dyson's Brownian motion at time one.

The second description of Dyson's Brownian motion came later [Gra99], but is better suited to serve as motivation for the upcoming Brownian Gibbs property. Notice that the eigenvalues remain ordered throughout time, and are strictly ordered at any given time, since this is true of the eigenvalues of GUE; while in principle two of them could coincide at some random time, it should be plausible that in fact they don't. Quite surprisingly, the second description of Dyson's Brownian motion is simply that it is a collection of n independent Brownian motions, all started at zero, which are conditioned to *never intersect*.

Of course, this is not a trivial conditioning, for two reasons: (i) independent Brownian motions with common starting point will almost surely intersect infinitely often in any time-interval containing zero, and (ii) independent Brownian motions with any starting points will almost surely intersect infinitely often over the course of their entire lifetimes on $[T, \infty)$ for any $T \geq 0$. So the conditioning we desire is singular in two distinct ways, at zero and at infinity.

We won't get into the details of how to overcome this singular conditioning, but, essentially, this is resolved by making use of something known as the Doob h-transform, which allows us to condition a Markov process X on a singular event if we can find a function h which is harmonic with respect to the generator of X and is such that the event that $h(X_t) = 0$ for some t is the same as the singular event we wish to condition on. For Dyson's Brownian motion, X is n-dimensional Brownian motion killed at the time it exits the Weyl chamber $\{x \in \mathbb{R}^n : x_1 \leq x_2 \leq \ldots \leq x_n\}$, and $h : \mathbb{R}^n \to \mathbb{R}$ is the Vandermonde determinant,

$$h(x) = \prod_{i<j} (x_i - x_j).$$

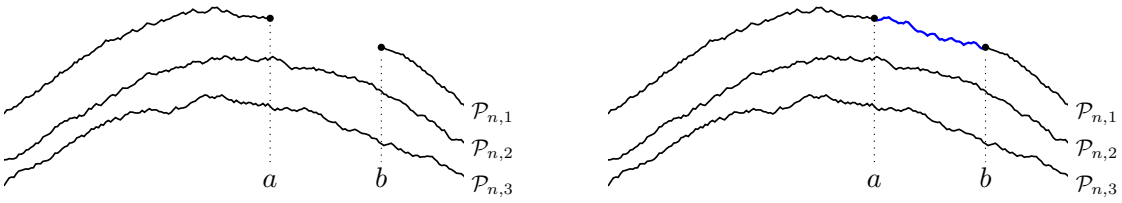


Figure 2.3: The Brownian Gibbs property: the conditional distribution of the "erased" portion of the top curve on $[a, b]$, given everything else (as depicted in the left panel), is that of a Brownian bridge (depicted in blue in the right panel) between the endpoint values marked by black circles (which have been conditioned upon), conditioned to not intersect the lower curve.

(Note that X is not simply n-dimensional Brownian motion, essentially because the event that that process will eventually exit the Weyl chamber has probability 1.) See, for example, [War07, Section 3] for more details.

With this brief description of the precise definition, we will proceed with an intuitive understanding of the meaning of non-intersecting Brownian motions; in any case, we will only be concerned with the behaviour of Dyson's Brownian motion on compact intervals away from the origin, so, heuristically, neither of the singular conditionings arise.

Having introduced Dyson's Brownian motion, we may now state the remarkable generalization of the identity (2.2) proven in [OY02], where $P_n = (P_{n,1}, \dots, P_{n,n})$ and DBM_n is n-level Dyson's Brownian motion:

$$P_n \text{ and } \mathrm{DBM}_n \text{ are equal in law as processes on } [\![1, n]\!] \times [0, \infty).$$

The Brownian Gibbs property

The Brownian Gibbs property gives an explicit description of the conditional law of P_n given a selection of data. Let $[a, b] \subseteq (0, \infty)$ be a given interval avoiding zero and $k \in [\![1, n]\!]$. Let $\mathcal{F}$ be the σ-algebra generated by the data of $P_{n,k+1}, \dots, P_{n,n}$ on $[0, \infty)$ as well as $P_{n,1}, \dots, P_{n,k}$ outside of $[a, b]$. In other words, we condition on everything apart from the top k curves on $[a, b]$.

The Brownian Gibbs property says that the conditional law of $(P_{n,1}, \dots, P_{n,k})$ on $[a, b]$ given $\mathcal{F}$ is that of k independent Brownian bridges $B_1^{\mathrm{br}}, \dots, B_k^{\mathrm{br}}$ of rate one, with B_i^{br} between the coordinates $(a, P_{n,i}(a))$ and $(b, P_{n,i}(b))$, conditioned on intersecting neither each other nor the lower curve $P_{n,k+1}$. See Figure 2.3.

That this is true is easy to see from the description of Dyson's Brownian motion as Brownian motions conditioned to not intersect. Indeed, the conditional distribution of a single Brownian motion on $[a, b]$, given its path outside of $[a, b]$, is a Brownian bridge between the appropriate endpoints (as can be checked by a covariance calculation); considering k of them together and imposing the non-intersection conditioning from the definition of Dyson's Brownian motion then (intuitively) gives the Brownian Gibbs description. (Of course, that Dyson's Brownian motion has

the Brownian Gibbs property needs a precise proof, and this was in fact rigorously proved only recently, essentially in [DM20]—see Lemma 11.12 ahead for details.)

What does the fact that P_n has the Brownian Gibbs property have to do with the parabolic Airy line ensemble $\mathcal{P}$? Similar to how the weight profile in Poissonian LPP converged, under appropriate centring and scaling, to the top line of $\mathcal{P}$, in Brownian LPP we have that P_n converges under scaling to the whole of $\mathcal{P}$. More precisely, let $\mathcal{P}_n : [\![1, n]\!] \times [-\frac{1}{2}n^{1/3}, \infty) \to \mathbb{R}$ be defined by

$$\mathcal{P}_{n,j}(x) = n^{-1/3} \left(P_{n,j}(n + 2n^{2/3}x) - 2n - 2n^{2/3}x \right). \tag{2.3}$$

While not performing the calculation here, we remark that $2n + 2n^{2/3}x$ is the mean of $P_{n,j}(n + 2n^{2/3}x)$ after excluding terms of lower order than the fluctuation order, i.e., $n^{1/3}$; this can be seen from the fact that $P_{n,1}(1) = \lambda_1$ is $2\sqrt{n}$ to first order (from knowledge of GUE), Brownian square-root scaling, and a Taylor expansion of the square-root. The factor of 2 in front of $n^{2/3}x$ is included in order to have $\mathcal{P}$ as the limit of $\mathcal{P}_n$ since, as mentioned before, by convention $\mathcal{P}$ is locally of diffusion rate two, while P_n is locally of diffusion rate one.

Now, the Brownian Gibbs property will also hold for any affine transformation of P_n, such as $\mathcal{P}_n$, though possibly of a different rate. This is because affine transformations of Brownian motions are still Brownian motions, just with a possibly different drift and diffusion rate, and the drift does not play a role in the conditional distribution in terms of Brownian bridges on an interval. In particular, $\mathcal{P}_n$ has the Brownian Gibbs property of rate two.

Finally, we have the promised convergence.

Theorem 2.1 (Corwin-Hammond). *As $n \to \infty$,*

$$\mathcal{P}_n \to \mathcal{P}$$

as processes in the topology of uniform convergence of compact sets of $\mathbb{N} \times \mathbb{R}$. Further, the Brownian Gibbs property (of rate two) is preserved in the limit.

In fact, [CH14] proved the convergence when the prelimiting object was non-intersecting Brownian bridges, rather than Brownian motions as in Dyson's Brownian motion. However, this immediately implies the same convergence for Dyson's Brownian motion by the use of a transformation which takes Brownian bridges to Brownian motions; see [DV18, equation (2)].

As an immediate consequence of Theorem 2.1's assertion that $\mathcal{P}$ has the Brownian Gibbs property, it follows that $\mathcal{P}_j(\cdot) - \mathcal{P}_j(a)$ is absolutely continuous to rate two Brownian motion on any compact interval $[a, b]$. This is a qualitative comparison statement, and a main result of Part II will be a strong quantitative formulation, as we will discuss shortly in Section 2.2.

However, even the qualitative comparison was sufficient to resolve a conjecture made by Johansson about a decade earlier [Joh03, Conjecture 1.5]. In terms of the heuristic continuum LPP model from Chapter 1, the conjecture was that the geodesic whose starting point is fixed at $(0, 0)$ will have an almost surely unique ending point if the ending point is allowed to vary along the line $(x, 1)$; in other words, the *point-to-line* geodesic has an almost surely unique ending

point. Equivalently, since $\mathcal{P}_1$ is the weight profile as the ending point varies on exactly the same line, the conjecture asserts that $\mathcal{P}_1$ has an almost surely unique maximizer on $\mathbb{R}$.

Since Brownian motion has a unique maximizer on each compact interval, combining the absolute continuity of $\mathcal{P}_1$ to Brownian motion on compact intervals with the parabolic rate of decay of $\mathcal{P}_1$ yielded Johansson's conjecture in [CH14]. Shortly afterwards, alternate proofs of the conjecture not making use of the Brownian comparison were provided in [FQR13, Pim14].

Part IV of this thesis will be devoted to proving an estimate needed to investigate the set of random exceptional times when Johansson's conjecture fails to be true, as we will discuss in Section 2.5.

2.2 Results in the probabilistic stream: Brownian regularity

In this section we will motivate and state our main result of Part II and some of its applications, which were originally proved in [CHH19] with Alan Hammond and Jacob Calvert. To do so, we return to the parabolic Airy_2 process $\mathcal{P}_1$ and the larger system of which it is a part, the parabolic Airy line ensemble $\mathcal{P}$. We saw in Section 2.1 that $\mathcal{P}_1$ enjoys a qualitative comparison to Brownian motion by way of the Brownian Gibbs property: increments of $\mathcal{P}_1$ on compact intervals are absolutely continuous to rate two Brownian motion.

While for some situations, such as for the resolution to Johansson's conjecture, this level of information may be sufficient, it is easy to imagine that there are other situations where a more quantitative comparison may be necessary. For an example of a more quantitative comparison, while absolute continuity yields the existence of the Radon-Nikodym derivative of the increment of $\mathcal{P}_1$ with respect to Brownian motion, we may want to know whether it lies in an L^p space for a given $p > 1$, or all $p \geq 1$.

It is an easy consequence of the Hölder inequality that, if the Radon-Nikodym derivative lies in an L^p space for a given $p \geq 1$, then an event which has a Brownian probability of $\varepsilon \in (0,1)$ can have probability under the law of an increment of $\mathcal{P}_1$ at most $\varepsilon^{1-1/p}$. Thus, if the Radon-Nikodym derivative lay in every L^p space for $p \geq 1$, then an event of Brownian probability of ε would have probability at most $\varepsilon^{1-o(1)}$ under the law of an Airy increment. The first main result of this thesis establishes such a bound, with the error factor $\varepsilon^{-o(1)}$ taking the more explicit form $\exp(O(1)(\log \varepsilon^{-1})^{5/6})$. The result also applies to the prelimiting $\mathcal{P}_n$ under a mild lower bound condition on the value of ε.

Main Theorem 2.2. *Let $k \in \mathbb{N}$, $d \geq 1$, $\mathcal{C}$ be the space of continuous functions on $[-d,d]$ which vanish at $-d$, and A a Borel measurable subset of $\mathcal{C}$. Let $\varepsilon = \mathcal{B}(A)$, where $\mathcal{B}$ is the law of rate two Brownian motion on $[-d,d]$. Then there exists an absolute finite constant G such that*

$$\mathbb{P}\Big(\mathcal{P}_k(\cdot) - \mathcal{P}_k(-d) \in A\Big) \leq \varepsilon \cdot \exp\big(Gd(\log \varepsilon^{-1})^{5/6}\big).$$

Further, there exists a constant c depending on only k such that, if $\varepsilon \geq \exp(-cn^{1/12})$, then the same bound holds with $\mathcal{P}_{n,k}$ in place of $\mathcal{P}_k$.

We will state a slightly more general version of this theorem as Theorem 4.8 in Chapter 4 that does not require the interval under consideration, here $[-d, d]$, to be centred at zero.

Theorem 2.2 is the culmination across several papers of a probabilistic study of narrow wedge KPZ structure via the probabilistic Brownian Gibbs technique. The road begins with the absolute continuity comparison made by [CH14]. An important intermediate step was achieved in [Ham19a], in which a comparison formally very similar to that of Theorem 2.2 was made, but with the compared processes being affinely shifted so that their interval endpoint values vanish, and the comparison being with respect to Brownian bridge. The relation of Theorem 2.2 to this counterpart result in [Ham19a] is important both for its formal similarity and its striking differences, and for the technique of proof. We will discuss this more in Chapter 3.

The form of Brownian comparison made by this result and by its upcoming prelimiting counterpart Theorem 4.8 is strong enough to open up an exciting array of applications concerning KPZ universality and last passage percolation models, and has, in fact, already been an important tool in a number of subsequent studies. We will briefly mention these works of other authors in which Theorem 2.2 has played an important role shortly. First, we give an easy application of Theorem 2.2 to illustrate its utility and power; further applications are discussed and proven in the original paper [CHH19].

Movement of Airy$_2$ in an interval

An immediate application of Theorem 2.2 which illustrates its utility is the following corollary, which gives a tail bound on the amount the Airy process or its parabolic version moves in a unit-order interval.

Corollary 2.3. *Let $d \geq 1$ and $k \in \mathbb{N}$. Then there exist $C < \infty$ and $C' < \infty$ such that, for $x > 0$,*

$$\mathbb{P}\left(\sup_{s\in[-d,d]} |\mathcal{P}_k(s) - \mathcal{P}_k(-d)| \geq x\right) \leq e^{-x^2/8d + Cd^{1/6}x^{5/3}}$$

and

$$\mathbb{P}\left(\sup_{s\in[-d,d]} |\mathcal{A}_k(s) - \mathcal{A}_k(-d)| \geq x\right) \leq e^{-x^2/8d + C'dx^{5/3}}.$$

Recall that $\mathcal{P}_k$ and $\mathcal{A}_k$ are both, in an idealised sense, rate two processes, and that we are considering the tail probability of an increment over an interval of length $2d$. For Brownian motions of rates two, these probabilities can be understood as being roughly $\exp(-x^2/8d)$, and this accounts for the dominant terms in the exponents in the bounds in Corollary 2.3; the remaining terms in the exponents of the form $d^{1/6}x^{5/3}$ or $dx^{5/3}$ are sub-dominant corrections arising from Theorem 2.2.

Estimates on $\mathcal{P}_{n,k}$ have previously appeared in the literature with a weaker tail bound exponent of $3/2$, instead of 2 as obtained here, such as in [DV18, Proposition 1.6] and [Ham19a, Theorem

2.14]; additionally, we obtain an explicit coefficient for the $-x^2$ term in the exponent, as well as a quantified sub-dominant correction.

Proof of Corollary 2.3. Let B be a rate two Brownian motion on $[-d, d]$ started at zero. Writing $N(0, 4d)$ for a normal random variable with mean zero and variance $4d$, we see that, by the reflection principle for Brownian motion,

$$\mathbb{P}\left(\sup_{s\in[-d,d]} |B(s)| \geq x\right) \leq 2 \cdot \mathbb{P}\left(\sup_{s\in[-d,d]} B(s) \geq x\right) = 4 \cdot \mathbb{P}\Big(N(0, 4d) \geq x\Big) \leq 4 \cdot e^{-x^2/8d}.$$

The last inequality is due to the Chernoff bound. Now we apply Theorem 2.2, and raise the value of G obtained from Theorem 2.2 to absorb the multiplicative constant of 4, to get the first bound in Corollary 2.3. The second follows from the first by noting that $\sup_{s\in[-d,d]} |\mathcal{P}_k(s) - \mathcal{P}_k(-d)|$ differs from $\sup_{s\in[-d,d]} |\mathcal{A}_k(s) - \mathcal{A}_k(-d)|$ by at most d^2, and by bounding $xd/4 + Cd^{1/6}x^{5/3}$ by $C'dx^{5/3}$, where C' is defined by say $C' = C + 1$. □

The proof of Corollary 2.3 illustrates that the usefulness of Theorem 2.2 lies in allowing us to use all of the many powerful probabilistic tools and symmetries available for Brownian motion and the normal distribution in the problem of estimating the probabilities of very naturally arising events for the parabolic Airy_2 process.

Use of Theorem 2.2 in others' work

As mentioned, Theorem 2.2 has proved to be an important tool in a number of KPZ studies subsequent to [CHH19]. We discuss them briefly here.

First, Theorem 2.2 is used by [BGZ19] in an argument establishing the temporal covariance exponent in exponential LPP (which we will introduce in the next section) with flat initial conditions. In terms of Brownian LPP, the analogous question would be to identify the order of growth in r and n of $\text{Cov}(\mathcal{P}^{\text{flat}}_{n,1}, \mathcal{P}^{\text{flat}}_{r,1})$, where $\mathcal{P}^{\text{flat}}_{m,1}$ is the scaled weight profile as defined in (2.3), except the starting point may be anywhere on the line indexed by n and ending point at $(0, n - m + 1)$. In other words, we are looking at the correlation of geodesic weight to height r and n, when the corresponding geodesics may begin anywhere on the same bottom line; this initial condition is the meaning of "flat". The corresponding quantity in exponential LPP is shown in [BGZ19] to be of order r/n as $n \to \infty$ if r grows linearly with n. Interestingly, this is a different rate than the narrow-wedge case, where it was independently found to be $(r/n)^{1/3}$in [BG18] and [FO19].

Next, Theorem 2.2 was used by [GH20a, GH20b]. [GH20b], with the companion paper [GH20a] studies stability and chaos of geodesic overlap in a model of dynamical Brownian LPP, where the Brownian environment is given an evolution over time. This evolution in time is not the same as that of the KPZ fixed point, as additional environment (i.e., additional lines in Brownian LPP) is not being revealed; instead, the entire environment (on a fixed number of lines) is changing with time. The question they answer is, what is the time scale at which the geodesic in the original

environment and in the perturbed environment are essentially statistically independent, with the measurement in terms of the statistic of overlap of the geodesics? They show that the critical time scale is of order $n^{-1/3}$. Theorem 2.2 is used to obtain strong control over the probability of that there is a path far from the geodesic whose weight is quite close to the geodesic weight; the third pillar of this thesis mentioned earlier will study a similar event under general initial conditions.

A third work which makes use of Theorem 2.2 is [DSV20]. This work studies limiting geodesics in Brownian LPP, and proves that they possess a finite three-halves variation, analogous to how Brownian motion possesses finite quadratic variation. They also show that the weight function along the geodesic possesses cubic variation. Their analysis uses the form of Theorem 2.2 that the Radon-Nikodym derivative of $\mathcal{P}_1$ with respect to Brownian motion lies in all L^p spaces.

The final work that has thus far made use of Theorem 2.2, also in its Radon-Nikodym form, is the earlier mentioned [DZ21]. One of the results of this work is that the Airy line ensemble captures weights of maximal weight disjoint collections of paths in a continuum LPP setting (where the environment is given by the directed landscape), analogous to the role of the lower rows of the Young diagram in Poissonian LPP.

2.3 Background on the geometric approach

Now we turn to the other stream of probabilistic and geometric investigations of KPZ. Unlike the previous section, this does not rely on precise probabilistic structures like the Brownian Gibbs property. Instead, it studies LPP problems by focusing on geometric properties of geodesics, often with the central idea of considering the geodesic at *different scales*.

This line of approach was initiated by Basu, Sidoravicius, and Sly in [BSS14]. While that paper works in both Poissonian LPP and an LPP model on the lattice $\mathbb{Z}^2$, much of the subsequent work, including that pursued in Part III of this thesis, has focused on the lattice LPP model, and so we introduce it now.

This time, the space is $\mathbb{Z}^2$. The random environment is given by an i.i.d. collection of non-negative random variables $\{\xi_v\}_{v\in\mathbb{Z}^2}$, one random variable for each vertex of $\mathbb{Z}^2$. The directed paths are up-right paths which move from one vertex to the one immediately above or immediately to the right at each step. The weight of a path γ is the sum $\sum_{v\in\gamma}\xi_v$. We will label the LPP value from $(1,1)$ to $(r+z, r-z)$ by X_r^z. (The notation does not match that for the similar object $\mathcal{P}_n$ from Brownian LPP because we will be concerned with values of z up to $|r|$, i.e., much larger than the transversal fluctuation scale $r^{2/3}$.)

The exactly solvable cases

While universality suggests that similar behaviour should hold for a wide class of distributions of the vertex weights, two distributions are distinguished for rendering the model exactly solvable. These are the geometric distributions (for any parameter $p \in (0,1)$) and the exponential distributions (which we may, without loss of generality, take to have rate one). Here, by the geometric

distribution of parameter q, we mean the law of the random variable G such that

$$\mathbb{P}(G \geq k) = (1-q)q^k \text{ for } k \in \mathbb{N}.$$

In these cases, Johansson showed convergence of X_r^z, suitable centred and scaled, to the GUE Tracy-Widom distribution [Joh00]:

Theorem 2.4 (Johansson). *Let the vertex weight distribution be geometric of parameter $q \in (0,1)$ or exponential of rate one. Let $z = \alpha r$ for some $\alpha \in (-1,1)$, and let F_{TW} be the GUE Tracy-Widom distribution. There exist constants $\mu = \mu(\alpha, q)$ and $\sigma = \sigma(\alpha, q)$ such that the following weak convergence holds:*

$$\lim_{r\to\infty} \frac{X_r^z - \mu r}{\sigma r^{1/3}} \xrightarrow{d} F_{\mathrm{TW}}.$$

Johansson's result is explicit about what the constants μ and σ are, but we have not included them as they are not very illuminating for our purposes. But we remark that the limit shape given by μ as a function of α has non-trivial curvature throughout the domain $\alpha \in (0,1)$; i.e., without scaling, in the whole windows of $O(r)$ and not just in the window of size $O(r^{2/3})$ that converges to the parabolic Airy_2 process (this convergence is also known for these vertex weight distributions).

We pause here to recall that an important feature of the GUE Tracy-Widom distribution, namely the "non-Gaussian" behavior of its upper and lower tails. In particular, it is known, for example from [Sep98b, page 224] or [RRV11, Theorem 1.3], that as $\theta \to \infty$,

$$\begin{aligned} F_{\mathrm{TW}}\big([\theta,\infty)\big) &= \exp\left(-\frac{4}{3}\theta^{3/2}\left(1+o(1)\right)\right) \qquad \text{and} \\ F_{\mathrm{TW}}\big((-\infty,-\theta]\big) &= \exp\left(-\frac{1}{12}\theta^{3}\left(1+o(1)\right)\right). \end{aligned} \tag{2.4}$$

(In fact, these tail exponents of $3/2$ and 3 are more universal in KPZ than just the GUE Tracy-Widom distribution. While the latter is the limiting distribution for narrow-wedge initial data. the same tail exponents are expected for a much wider class of initial data. For example, the results of [CG18] assert that the suitably scaled solution to the KPZ equation has upper bounds on the one-point upper and lower tails with the same tail exponents, up to a certain depth into the tail, under a wide class of *general* initial data. Similar upper bounds are known for the KPZ fixed point under general initial data are known as well [MQR17, Proposition 4.7]. The same exponents are also known from [RRV11] for the entire class of Tracy-Widom(β) distributions, where the GUE case corresponds to $\beta = 2$.)

The estimates (2.4) are obtained by precise asymptotic analysis of structures which are revealed by the exactly solvable nature of objects like the GUE Tracy-Widom distribution. Part III will be devoted to obtaining a geometric understanding and derivation of the upper and lower tail exponents of $3/2$ and 3, as we will explain shortly in Section 2.4.

Now, importantly for investigations in the prelimit, Johansson was also able to obtain a remarkable identity between X_r^z and certain determinantal point processes. This is simpler in the exponential case and we restrict to that. Here, X_r^z has the same distribution as the top eigenvalue of a random matrix ensemble known as the Laguerre Unitary Ensemble (LUE), also known as complex Wishart matrices.

Using this identification, Ledoux and Rider [LR10], as well as Basu, Ganguly, Krishnapur, and myself [BGHK19], were able to study the LUE to obtain tail estimates for X_r^z with the same tail exponents as of the GUE Tracy-Widom from (2.4). We also mention that an upper bound on the upper tail was previously obtained by Seppaläinen [Sep98a] using large deviation and superadditivity techniques. We state only the $z = 0$ case, where we set $X_r = X_r^0$; for exponential LPP, $\mu = 4$ in this case (in the notation of Theorem 2.4).

Theorem 2.5 (Basu-Ganguly-H.-Krishnapur, Ledoux-Rider, Seppaläinen)**.** *There exist positive finite constants c_1, c_2, c_3, θ_0, and r_0 such that, for $r > r_0$ and $\theta_0 < \theta < r^{2/3}$,*

$$\mathbb{P}\left(X_r > 4r + \theta r^{1/3}\right) \leq \exp\left(-c_1\theta^{3/2}\right) \qquad \text{and}$$
$$\exp\left(-c_2\theta^3\right) \leq \mathbb{P}\left(X_r < 4r - \theta r^{1/3}\right) \leq \exp\left(-c_3\theta^3\right).$$

We remark that one of our main results will obtain the missing lower bound on the upper tail; see Theorem 2.6.

The assumptions in the geometric approach

As mentioned earlier, the geometric approach typically adopts certain assumptions (which are only known to be true in the exactly solvable cases currently) and utilizes geometric arguments involving the geodesic, usually on many different scales, to obtain robust conclusions. Having explained what is known in the exactly solvable cases, we can now introduce and appreciate the forms of the assumptions adopted. Qualitatively, there are two main assumptions:

1. **Curvature of the limit shape:** The limit shape $\alpha \mapsto \mu(\alpha)$ has curvature. This can sometimes be for all $\alpha \in (-1, 1)$ (i.e., at all points away from the coordinate axes), all α in a neighbourhood of zero, or just at zero itself.
2. **One point upper and lower tail bounds on the $n^{1/3}$ scale**: Typically upper bounds on both tails as in Theorem 2.5, though sometimes lower bounds may also be required. The tail exponents often—but not always—need to be the optimal ones. Because there is no general proof that fluctuations occur on the scale of $n^{1/3}$, this must be assumed.

There have been many papers which have combined precise assumptions which fit this qualitative description with the study of geodesic geometry at different scales. As mentioned, the first was [BSS14], where the following question was studied. Suppose we are in exponential LPP, but the n vertex weights along the diagonal $x = y$ within the square $[\![1, n]\!]^2$ are distributed as independent $\mathrm{Exp}(1-\varepsilon)$ random variables for some $\varepsilon > 0$ (and so each has mean $(1-\varepsilon)^{-1}$, i.e.,

greater than 1). Then does the LPP value to (n, n) change to first order? That is, is it the case that

$$\lim_{n\to\infty} \frac{\mathbb{E}[X_n]}{n} > 4$$

for all $\varepsilon > 0$? Note that the perturbation is a microscopic one in that only $O(n)$ many vertices of the environment have a different distribution assigned, out of the $O(n^2)$ many vertices in total, but the question concerns whether this leads to a change which is in some sense macroscopically apparent.

This question was raised by physicists but had no clear consensus on its answer within the physics community. It was answered in the affirmative in [BSS14], by a strategy which relied on a sophisticated analysis that studied the geodesic at different scales, showing a degree of independence between then, and the effect of the perturbation on paths which were near geodesics, i.e., had weight very close to the maximum.

2.4 Results in the geometric stream: Bootstrapping

The main aim of this part of this thesis is to demonstrate a geometric source for the tail exponents of $3/2$ and 3 for the Tracy-Widom in the context of LPP. In other words, we will prove results like Theorem 2.5, but under assumptions in the vein of the two listed in the previous section—which is to say, without exactly solvable connections such as those available in the geometric and exponential cases. The results in this section were originally proven in [GH20c] with Shirshendu Ganguly.

First, we can observe a simple geometric source for why the upper tail exponent, $\frac{3}{2}$, is smaller than the lower tail exponent, 3: for the LPP weight to be large, it is enough to have a *single* path with large weight, while, for the LPP weight to be small, *all* paths need to have small weight. Of course, this does not indicate at all why the tail exponents take their precise values, and that is the main point of interest of the main result.

Here we will describe a simplified version of the assumptions that are adopted and give an informal version of the main theorem; the assumptions will be precisely stated, along with a precise version of the main theorem, in Chapter 7.

There are essentially three assumptions. We set $\mu = \lim_{r\to\infty} \mathbb{E}[X_r]/r$, i.e., the limiting linear coefficient in the $x = y$ direction.

1. **Sharp concavity of limit shape and non-random fluctuations in a neighbourhood of zero:** The limit shape is, to second order, parabolic, with a lower order correction, and $\mathbb{E}[X_r^z]$ is within $O(r^{1/3})$ of the limit shape, for $z \in [-\rho r, \rho r]$ with ρ some positive constant.
2. **Upper bounds on tail probabilities, uniformly in directions:** There is an $\alpha > 0$ such that, if $z \in [-(1-\varepsilon)r, (1-\varepsilon)r]$, then we have tail bounds with tail exponent α:
   a) $\mathbb{P}\big(X_r^z - \mathbb{E}[X_r^z] > \theta r^{1/3}\big) \le \exp(-c\theta^\alpha)$
   b) $\mathbb{P}\big(X_r^z - \mathbb{E}[X_r^z] < -\theta r^{1/3}\big) \le \exp(-c\theta^\alpha)$.

Note that $\alpha > 0$ is arbitrary. The main result will bootstrap this tail exponent to the optimal $3/2$ and 3 for the upper and lower tails.

3. **Lower bounds on tail probabilities, in the diagonal direction:** There are non-trivial fluctuations of scale $r^{1/3}$ in the diagonal direction, i.e., there is a $\delta > 0$ and $C > 0$ such that
   a) $\mathbb{P}\big(X_r - \mu r > Cr^{1/3}\big) \geq \delta$
   b) $\mathbb{P}\big(X_r - \mu r < -Cr^{1/3}\big) \geq \delta$.

We discuss the assumptions (in their precise form) in more detail in Chapter 7.

While the assumptions are currently only known to hold for exponential and geometric LPP, we emphasize again that their form precludes anything like the previous integrable techniques used to prove the tail bound in those models. We have to rely completely on an understanding of the geometry of weight maximizing paths and the interaction across scales.

Without further ado, here is an informal version of the main result.

Main Theorem 2.6 (Informal). *Under the assumptions, there exist constants c_1, c_2, c_3, and c_4 such that, for all large enough r and appropriate ranges of θ,*

$$\exp\big(-c_1\theta^{3/2}\big) \leq \mathbb{P}\left(X_r > \mu r + \theta r^{1/3}\right) \leq \exp\big(-c_2\theta^{3/2}(\log\theta)^{-1/2}\big)$$
$$\exp\big(-c_3\theta^3\big) \leq \mathbb{P}\left(X_r < \mu r - \theta r^{1/3}\right) \leq \exp\big(-c_4\theta^3\big).$$

This is given in its precise form as Theorems 7.4–7.7. The point of these results is the obtaining of the correct upper and lower tail exponents even though we work in a non-integrable setting and rely on only geometric arguments.

Extremely briefly, the techniques which prove Theorem 2.6 rely on looking at the geodesic at many different scales and utilizing a tension arising from the fact that fluctuations must have exponent $\frac{1}{3}$ on all of them; indeed, the results can be thought of as connecting the tail exponents to the weight fluctuation exponent, with the formula $\frac{3}{2} = 1/(1-\frac{1}{3})$ capturing the relation. (We note in passing that the Gaussian tail exponent of 2 and the classical scaling exponent of $\frac{1}{2}$ also satisfy this relation.) For the lower tail bounds, similar ideas are combined with an analysis of the weight and geometry of a large number of high weight *disjoint* curves. These aspects are discussed in the form of a much more detailed overview in Section 7.3.

The techniques also get sharp bounds for the lower tail of *constrained* LPP values, a result which was not previously known even in the case of integrable models. By constrained LPP values we mean the following. Let U be a rectangle with one axis parallel to the line $x = y$, and having the other set of parallel sides with midpoints $(1,1)$ and (r,r). We define X_r^U to be the maximum weight over all paths from $(1,1)$ to (r,r) which are constrained to remain within U. Suboptimal tail bounds for X_r^U have played important roles in previous geometric investigations of LPP, such as [BSS14, BGHH20].

Main Theorem 2.7 (Informal). *Under the assumptions, there exist positive constants c_1 and c_2 such that, if U has width (measured in the direction of the line $x = -y$) $\ell r^{2/3}$, then, for appropriate ranges*

of θ,

$$\exp\left(-c_1 \min(\ell\theta^{5/2}, \theta^3)\right) \leq \mathbb{P}\left(X_r^U - \mu r < -\theta r^{1/3}\right) \leq \exp\left(-c_2 \min(\ell\theta^{5/2}, \theta^3)\right)$$

This is stated in its precise form as Theorem 7.13.

2.5 Results combining the perspectives: Fractal geometry

So far in this chapter we have seen two broad approaches: the first made use of probabilistic properties of weight profiles, such as the parabolic Airy line ensemble $\mathcal{P}$, of natural LPP objects, while the second studied geodesic geometry in discrete LPP models. And in fact, there will be no interactions between the probabilistic properties of weight profiles and the geometry of geodesics in the respective proofs.

In this section we introduce a result whose proof will require reasoning about both properties of geodesic as well as probabilistic properties of weight profiles.

The result, while of interest in its own right, served as a technical input for a project [CHHM21] regarding *fractal* properties of the KPZ fixed point that I undertook with Ivan Corwin, Alan Hammond, and Konstantin Matetski. To motivate the result included in this thesis, we will first spend a little time explaining the larger goal.

A return to Johansson's conjecture

Recall Johansson's conjecture [Joh03], discussed above, that the parabolic Airy$_2$ process $\mathcal{P}_1$ almost surely has a unique maximizer, which was proved in [CH14], as well as in [FQR13, Pim14]. Now, $\mathcal{P}_1$ is an instance of the KPZ fixed point, introduced at the end of Chapter 1, with narrow-wedge initial condition and at time one.

The resolution to Johansson's conjecture also implies that at any fixed time $t > 0$, $\mathfrak{h}_t$ almost surely has a unique maximizer when started from narrow-wedge initial condition. But it does *not* say that there do not exist exceptional times t where $\mathfrak{h}_t$ has multiple maximizers; for, since there are uncountably many times in any given interval, a fixed time bound and a union bound tell us nothing.

This can be thought of via an analogous situation involving a Brownian motion B. The probability that $B_t = 0$ for any fixed t is zero. Yet, for any given interval there is positive probability that B will hit zero in that interval; and, if B starts at zero and the interval is $(0, x)$ for some x, then the interval will almost surely contain a zero.

Continuing this example, note that the fixed time probability being zero implies, via Fubini's theorem, that the Lebesgue measure of the set of times where B equals zero is almost surely zero:

$$\mathbb{E}\Big[\mathrm{Leb}\big(\{t \in [0,T] : B(t) = 0\}\big)\Big] = \mathbb{E}\left[\int_0^T \mathbb{1}_{B(t)=0}\,\mathrm{d}t\right] = \int_0^T \mathbb{P}\big(B(t) = 0\big)\,\mathrm{d}t = 0.$$

So if we are interested in understanding the size of the set of exceptional times when B is zero, the Lebesgue measure is too crude a tool. We instead need a finer notion of sparsity: typically, this is one of *fractal dimension*. While there are a number of different formulations of fractal dimension, the most common one (and the one adopted in [CHHM21]) is that of *Hausdorff dimension*.

While we won't give the precise definition here, the unfamiliar reader can heuristically think of the Hausdorff dimension of a set X as the number $\alpha > 0$ such that one needs order $\varepsilon^{-\alpha}$ many sets of diameter ε to cover X. It is easy to see that this notion recovers that the dimension of a line is 1, of a square is 2, and so on. A good reference is [Mat99].

Now, it is a well-known fact that the zero set of Brownian motion on an interval, conditionally on it being non-empty, almost surely has Hausdorff dimension $\frac{1}{2}$. [CHHM21] identifies the Hausdorff dimension of the set of exceptional times where $\mathfrak{h}$ has multiple spatial maximizers, for a broad class of initial data (along the way also showing that the KPZ fixed point with these initial data almost surely has a unique maximizer at any fixed time, thus extending the validity of Johansson's conjecture to beyond the narrow-wedge case).

For $T > 0$ and $A \geq 0$, let

$$\mathcal{T}_{T,A} = \Big\{ t \in [0,T] : \mathfrak{h}_t \text{ has multiple maximizers at distance greater than } A \text{ apart} \Big\}.$$

The following is the main result of [CHHM21].

Theorem 2.8 (Corwin-Hammond-H.-Matetski)**.** *Fix $T > 0$ and $A \geq 0$. Then, for a broad class of initial data $\mathfrak{h}_0$, $\mathcal{T}_T \neq \emptyset$ with positive probability, and, conditionally on this event, $\mathcal{T}_T$ almost surely has Hausdorff dimension $\frac{2}{3}$.*

We will define the precise class of initial conditions in Chapter 11.

In terms of geodesics, the exceptional times in $\mathcal{T}_T$ are exactly those that exhibit *polymer instability*: the instants where the endpoint of the geodesic with the given initial condition and unconstrained endpoint performs a macroscopic, discontinuous, jump. See Figure 2.4.

A common technique to prove bounds on Hausdorff dimension is to construct a measure supported on the set in question. It can be useful to first consider a proxy set which is "fattened" so as to have positive Lebesgue measure. In our case, the proxy set will be defined as the set of times when an event we call *ε-twin peaks* occurs. This event will be denoted $\mathsf{TP}_A^\varepsilon$ and is defined as a subset of continuous functions by

$$\mathsf{TP}_A^\varepsilon = \Big\{ f : \exists x_1, x_2 \text{ s.t. } |x_1 - x_2| > A \text{ and } f(x_i) > \mathrm{Max}(f) - \varepsilon \text{ for } i = 1, 2 \Big\};$$

in words, there are two peaks (with one being the maximum) of the weight profile at least distance A apart, such that they are within ε of each other.

Then our proxy set is defined by

$$\mathcal{T}_{T,A}^\varepsilon = \Big\{ t \in [0,T] : \mathfrak{h}_t \in \mathsf{TP}_A^\varepsilon \Big\}.$$

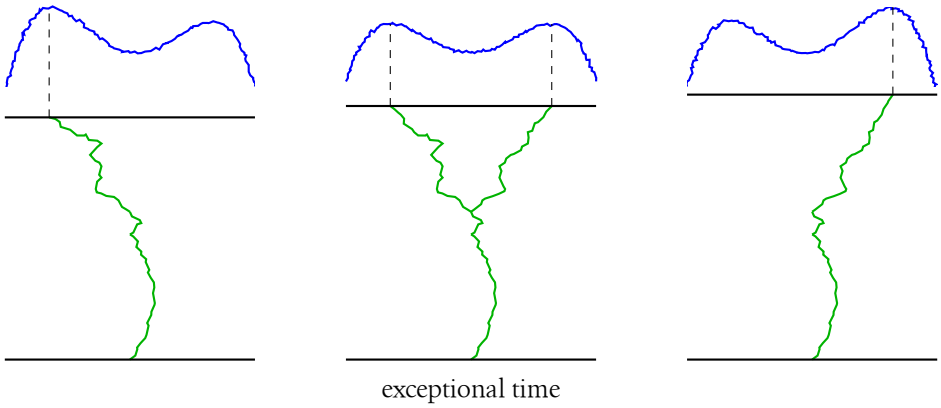


Figure 2.4: In blue is the KPZ fixed point at various times (which advances as we go from the left panel to the right one, as illustrated by the increasing height of the upper black line), while in green is the geodesic in the heuristic continuum LPP model. Notice that the endpoint of the geodesic is the maximizer of the KPZ fixed point at each time. In the left panel, the geodesic is unique and has endpoint on the left, while a competing peak develops on the right. In the centre panel, the KPZ fixed point has *two* maximizers, as the right peak has caught up to the original left peak, and thus the corresponding time is an exceptional one where Johansson's conjecture on uniqueness of maximizer is broken. Correspondingly, there are two geodesics; we believe, but it is not yet proven, that it is typical (in a precise sense, in spite of working on an exceptional event) for this to occur without the two geodesics being completely disjoint, i.e., as depicted. In the final panel we see that the right peak has overtaken the left peak and is now the maximizer; the geodesic endpoint is thus now on the right. In particular, as we advance across time, and move across the exceptional instant of the middle panel, the geodesic endpoint performs a macroscopic jump.

To construct a measure μ on $\mathcal{T}_{T,A}$, we instead consider the uniform measure μ^ε on $\mathcal{T}^\varepsilon_{T,A}$ and take a limit as $\varepsilon \to 0$. Of course, we need to scale μ^ε to ensure that the limiting measure is non-trivial. The existence of a scaling such that the limit is non-trivial is the main result of this part of this thesis.

Main Theorem 2.9 (Informal)**.** *Let the initial state $\mathfrak{h}_0$ lie in the broad class from Theorem 2.8, and suppose that $A > 0$ and $t > 0$. There exists $\eta > 0$ such that, for all $\varepsilon \in (0,1)$,*

$$\mathbb{P}\left(\mathfrak{h}_t \in \mathsf{TP}^\varepsilon_A\right) \geq \eta\varepsilon. \tag{2.5}$$

An implication of this result is that $\lim_{\varepsilon\to 0} \varepsilon^{-1}\mu^\varepsilon$ has total mass at least η, and so is non-trivial. The precise version of Theorem 2.9 is stated as Theorem 11.2.

Remarks on the proof

An upper bound of order ε on the same probability is also proved in [CHHM21]. The fact that the probability should be of order ε is easy to see if one expects $\mathfrak{h}_t$ to be Brownian: Brownian motion has probability of order ε of coming within ε of its maximum on a given interval.

But note that while we have a very refined understanding of the Brownian nature of $\mathcal{P}_1$, that is only the narrow-wedge case, while Theorem 2.9 applied to a wide class of initial data. For general initial data, an infinite ensemble analogous to the parabolic Airy line ensemble has not been constructed, and if it were constructed, would not have a simple and explicit resampling property like the Brownian Gibbs property.

It is for this reason that new techniques are needed to derive Theorem 2.9. The techniques developed combine a study of geodesic geometry and the Brownian Gibbs property of the narrow-wedge case to obtain an understanding for general initial data. While a more detailed explanation is given in Chapter 11, we will try to communicate here how it is that both types of reasoning can be brought to bear upon the problem.

The key is a remarkable generalization of the RSK correspondence proved in [DOV18], and used to great effect there to construct the parabolic Airy sheet (in terms of the heuristic continuum LPP model from Chapter 1, this is a limiting object which encodes the joint distribution of LPP values from any $(y, 0)$ to $(x, 1)$, i.e., with starting point also unconstrained, unlike $\mathcal{P}_1$).

Recall that given a Brownian environment $B = (B_1, \ldots, B_n)$, we denoted LPP values through the environment by $B[(y, 1) \to (x, n)]$, and defined $P_n = (P_{n,1}, \ldots, P_{n,n})$ in (2.1) in terms of last passage values starting at $(0, 1)$ through B. In particular, we had

$$P_{n,1}(x) := B[(0, n) \to (x, 1)]$$

and that $P_{n,1} > P_{n,2} > \ldots > P_{n,n}$. Because the curves of P_n are ordered and all start at the common value of zero at zero, it is easy to see that, *if we consider P_n to be a new environment*, then

$$P_n[(0, n) \to (x, 1)] = P_{n,1}(x) = B[(0, n) \to (x, 1)].$$

In other words, LPP values are preserved when starting from $(0, n)$ when we swap the B environment for the P_n environment. The remarkable fact proved in [DOV18, Proposition 4.1] is that they are preserved no matter the starting point, so long as it is on line n.

Proposition 2.10 (Dauvergne-Ortmann-Virág)**.** *Let $y \leq x$. Then,*

$$B[(y, n) \to (x, 1)] = P_n[(y, n) \to (x, 1)].$$

Note that this is not a statement of equality in distribution, but of actual values. In fact, this is a particular case of an equality which holds deterministically between any environment defined by a continuous $f = (f_1, \ldots, f_n)$ (analogous to B) and the environment analogous to P_n that is defined from f in the same way that P_n is defined from B.

Now recall the variational formula for the KPZ fixed point, given by

$$\mathfrak{h}_t(x) = \sup_{y\in\mathbb{R}}\Big\{\mathfrak{h}_0(y) + \mathcal{L}(y,0;x,t)\Big\},$$

where $\mathcal{L}$ is the directed landscape and $\mathcal{L}(y,0;x,t)$, in terms of the heuristic continuum LPP model of Chapter 1, is the LPP value from $(y,0)$ to (x,t). Taking $t=1$, this formula suggests a prelimiting analogue $\mathfrak{h}^{(n)}$ of $\mathfrak{h}_1$ that can be defined for Brownian LPP, taking into account the correct scalings:

$$\mathfrak{h}^{(n)}(x) = \sup_{y\geq -\frac{1}{2}n^{1/3}}\Big\{\mathfrak{h}_0(y) + n^{-1/3}\left(B[(2yn^{2/3},n)\to(n+2xn^{2/3})] - 2n - 2n^{2/3}(x-y)\right)\Big\}.$$

The important point of this formula is that the Brownian LPP term has starting point on line n and ending point on line 1, and so can be replaced by the same LPP value in the environment given by P_n via Proposition 2.10. Thus $\mathfrak{h}^{(n)}$ can be defined by an LPP problem through P_n.

Now, we know P_n is the same as Dyson's Brownian motion, and so enjoys the Brownian Gibbs property. Thus we have come upon an LPP problem in an environment that enjoys a convenient probabilistic resampling property. The arguments to prove Theorem 2.9 will use both these aspects (and then use the convergence of $\mathfrak{h}^{(n)}$ to $\mathfrak{h}_1$, and invariance properties of $\mathfrak{h}$ to handle other t), thus combining the probabilistic and geometric streams.

2.6 Other graduate work

In this final section of the introductory portion of this thesis, I will briefly describe some of the others projects relating to KPZ that I have worked on during my time in graduate school.

A weight difference profile and Brownian local time

The first project [GH21] I mention was done with Shirshendu Ganguly, and also concerns fractal geometry of a certain limiting LPP process. More precisely, let $\mathcal{S}:\mathbb{R}^2\to\mathbb{R}$ be the previously mentioned *parabolic Airy sheet*: in terms of the heuristic continuum LPP model of Chapter 1, $\mathcal{S}(y,x)$ is the LPP weight from $(y,0)$ to $(x,1)$. The process under consideration is a *weight difference profile* $\mathcal{D}:\mathbb{R}\to\mathbb{R}$, given by $x\mapsto\mathcal{S}(y_b,x)-\mathcal{S}(y_a,x)$ for some fixed $y_b>y_a$.

This process was first studied in [BGH19]. There, it was shown via a planarity argument that $\mathcal{D}$ is non-decreasing and that, almost surely, its set of non-constant points has Hausdorff dimension $\frac{1}{2}$. Given that this is also the Hausdorff dimension of the zero set of Brownian motion, Manjunath Krishnapur raised the question of whether there is some precise connection between $\mathcal{D}$ and Brownian local time.

In [GH21] we show that there is indeed a connection, on both global and local scales. On a global scale, we prove a form of absolute continuity of $\mathcal{D}$ to Brownian local time (of rate four)

by proving the existence of certain random intervals, or *patches*, such that the restriction of $\mathcal{D}$ to each patch is absolutely continuous to Brownian local time. We call this a *Brownian local time patchwork quilt*, in analogy with similar terminology for an absolute continuity comparison of the KPZ fixed point (at fixed time) with Brownian motion that was previously introduced and proved by Alan Hammond [Ham19d]. On a local scale, we show that, under Brownian scaling, the local limit of $\mathcal{D}$, taken at various points of increase, is explicitly Brownian local time, i.e.,

$$\lim_{\varepsilon \to 0} \varepsilon^{-1/2} \left(\mathcal{D}(\tau + \varepsilon t) - \mathcal{D}(\tau) \right)$$

is Brownian local time of rate four, where the convergence is as processes in the topology of uniform convergence on compact sets, and τ is, for example, the first point of increase of $\mathcal{D}$ after a given point $\lambda \in \mathbb{R}$, or a point of increase chosen uniformly from an interval $[a, b]$ by independently sampling from the probability measure defined by regarding $\mathcal{D}$ as a distribution function (after suitable centering and normalizing) on this interval.

Geodesic watermelons in lattice LPP

This project [BGHH20] was undertaken with Riddhipratim Basu, Shirshendu Ganguly, and Alan Hammond. The object of study was the *geodesic watermelon*, i.e., for a given $n \in \mathbb{N}$ and $k \in [\![1, n]\!]$, the maximum weight collection of k disjoint paths from $(1, 1)$ to (n, n) (more precisely, from and to a set of k points adjacent to $(1, 1)$ and (n, n) respectively). The name "watermelon" arises from the fact that the maximizing collection of disjoint curves look vaguely like the stripes of the watermelon fruit, and the nomenclature originated in the physics literature. As we have seen in earlier sections, in the exactly solvable cases, the weight of the geodesic watermelon has important connections to determinantal objects like eigenvalues of the LUE and, in the limit, the Airy point process.

This work, following the theme of geometric investigations, adopted some general assumptions regarding curvature of the limit shape and estimates on one point tail probabilities (of only the weight of the geodesic and *not* the geodesic watermelon) to study both the weight and geometry of geodesic watermelons. We showed that the the weight of the k-geodesic watermelon has fluctuations of order $k^{5/3}n^{1/3}$ below the leading order linear term of order kn—matching predictions from asymptotics of the Airy point process. We also show that the width of the geodesic watermelon is of order $k^{1/3}n^{2/3}$.

One interesting point is that starting with just information about the $k = 1$ case of the weight of the geodesic—that it fluctuates on scale $n^{1/3}$—we are able to prove the correct order of fluctuation for the weight of the k-geodesic watermelon. Another interesting feature is an important deterministic property of geodesic watermelons that is proved and centrally used, namely that the curves of the k- and $(k + 1)$-geodesic watermelons *interlace*; see Figure 2.5.

Finally, as one half of the argument proving the weight fluctuation statement, the paper constructs k disjoint paths with collective weight of the appropriate order. This was used later in the bootstrapping argument of [GH20c], which forms Part III of this thesis, and, as a result, we also include the construction in Chapter 10 of this thesis.

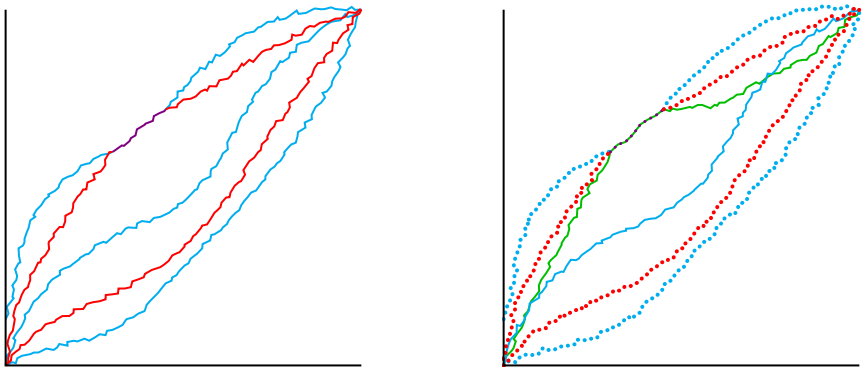


Figure 2.5: The interlacing of geodesic watermelons. In blue is the 3-geodesic watermelon and in red the 2-geodesic watermelon. In green is the geodesic itself, which crosses over the blue curves but interlaces with the red, illustrating that interlacing does not give much information about the geometry of geodesic watermelons beyond consecutive values of k.

A law of iterated logarithm in exponential LPP

The final project we mention was in collaboration with Riddhipratim Basu, Shirshendu Ganguly, and Manjunath Krishnapur. We proved an upper law of iterated logarithm for the LPP value from $(1, 1)$ to (n, n) in exponential LPP. More precisely, we showed that there is a deterministic constant $c > 0$ such that, almost surely,

$$\liminf_{n\to\infty} \frac{X_n - 4n}{n^{1/3}(\log\log n)^{1/3}} = -c.$$

(Recall that, to first order, X_n grows like $4n$ in exponential LPP.)

This complements results of Ledoux [Led18], who proved that the same lim inf is not $-\infty$, and that the lim sup of the same quantity, with the power of $\log\log n$ being $\frac{2}{3}$ instead of $\frac{1}{3}$, is also a finite positive deterministic constant.

As part of our argument in [BGHK19], we proved an optimal (in terms of tail exponent) lower bound for the lower tail of the top eigenvalue of the Laguerre β-ensemble—which, in the $\beta = 2$ case, translates to the lower bound on the lower tail of X_n in exponential LPP that was stated in Theorem 2.5.

Part II

Brownian regularity

Chapter 3

Proof context

This part of the thesis, i.e., Chapters 3–6, forms the first pillar of this thesis, concerning the probabilistic resampling viewpoint within KPZ. It consists of material from [CHH19] in work done with Jacob Calvert and Alan Hammond.

It will be convenient for the Airy-type processes to have, in some sense, a local diffusion rate of *one* instead of their usual two. To do this, we will henceforth, in this part, work with a version denoted $\mathcal{L}$ of the parabolic Airy line ensemble, for which we will use the additional adjective "standardized", defined by

$$\mathcal{L}(i,x) = 2^{-1/2}\mathcal{P}(i,x) = 2^{-1/2}(\mathcal{A}(i,x) - x^2),$$

where $\mathcal{P}$ is the parabolic Airy line ensemble, $\mathcal{A}$ is the Airy line ensemble, $\mathcal{P}(i,x) = \mathcal{P}_i(x)$ and similarly for $\mathcal{A}$.

3.1 Pertinent recent work

We start by giving a brief account of the general background of the Airy_2 process. The interested reader is referred to the survey [QR14] for a more detailed review, though from the slightly different viewpoint of integrable probability.

The locally Brownian nature of the Airy_2 process has been previously established in a number of different formulations. One relatively weak version is to consider *local limits* of the Airy_2 process; i.e., to study the Gaussianity of $\varepsilon^{-1/2}\left(\mathcal{A}(x+\varepsilon) - \mathcal{A}(x)\right)$ for a given $x \in \mathbb{R}$ as $\varepsilon \searrow 0$. The appearance of Brownian motion in this limit was proven in [Häg08, CP15, QR13]. The final of these three articles, [QR13], also establishes Hölder $\frac{1}{2}-$ continuity of the Airy_2 (as well as Airy_1) process, which is extended to limiting weight profiles arising from a very general class of initial conditions in [MQR17, Theorem 4.13]. A stronger notion of the locally Brownian nature of the Airy_2 process is absolute continuity of $\mathcal{A}$ with respect to Brownian motion on a unit order compact interval. This was first proved in [CH14].

Another line of work has established various Brownian features in the *pre-limiting* weight profiles. For instance, [BG18] establishes local Brownian fluctuations (in the sense of sub-Gaussian tails) in the weight profile of point-to-point exponential LPP, while [Ham19c] establishes a (sharp) version of the Holder $\frac{1}{2}-$ continuity mentioned above for the pre-limiting weight profiles in Brownian LPP (which also applies with quite general initial conditions).

However, none of these results addresses the question of bounding probabilities involving the Airy_2 process in terms of Brownian probabilities, or, equivalently, providing growth bounds on the Radon-Nikodym derivative with respect to some Brownian process.

A result in this direction was proved in [Ham19a]. There the comparison was between a modification of $\mathcal{L}$, denoted $\mathcal{L}^{[-d,d]}$, that is defined by affinely shifting $\mathcal{L}$ to be zero at both endpoints of $[-d,d]$, and Brownian bridge, instead of between a vertically shifted version of $\mathcal{L}$ and Brownian motion as in Theorem 2.2. The form of the result, however, is otherwise much the same:

Theorem 3.1 (Theorem 1.10 of [Ham19a]). *Let $d \geq 1$, $k \in \mathbb{N}$, and let $\mathcal{C}_{0,0}$ be the space of continuous functions which vanish at both endpoints of $[-d,d]$. Let A be a Borel measurable subset of $\mathcal{C}_{0,0}$, and let $\varepsilon = \mathcal{B}^{[-d,d]}(A)$, where $\mathcal{B}^{[-d,d]}$ is the law of standard Brownian bridge on $[-d,d]$ (i.e., with vanishing endpoints). There exists $\varepsilon_0 = \varepsilon_0(d) > 0$ and an absolute finite constant G such that, if $\varepsilon \in (0, \varepsilon_0)$, then*

$$\mathbb{P}\left(\mathcal{L}_k^{[-d,d]}(\,\cdot\,) \in A\right) \leq \varepsilon \cdot \exp\left(Gd^2(\log \varepsilon^{-1})^{5/6}\right).$$

This also follows immediately from Theorem 2.2 and the fact that performing the affine shift described on Brownian motion results in Brownian bridge.

Theorem 3.1 and our new Theorem 2.2 are formally very similar, the latter obtained merely by substituting Brownian motion for Brownian bridge. However, it is found in many contexts that Theorem 3.1 is unable to provide the kind of information that is desired. This is because, though the process $\mathcal{L}(\,\cdot\,) - \mathcal{L}(-d)$ can be obtained from the bridge $\mathcal{L}^{[-d,d]}(\,\cdot\,)$ and the endpoint $\mathcal{L}(d)$, the desired information gets away from us due to potentially pathological correlations between these two random objects. Controlling this correlation is especially required to understand the slope or maximum of $\mathcal{L}$ on an interval; the slope or maximum are often of relevance in LPP problems, as can be seen in the applications discussed in Section 2.2.

The proof of Theorem 2.2 is significantly more involved and subtle than the proof of Theorem 3.1 in [Ham19a] because of the need to handle these correlations. We make some more comments contrasting the proofs in Section 3.2.

Theorem 3.1 was a crucial tool in the four-part study of Brownian LPP undertaken in [Ham19a, Ham19c, Ham19b, Ham19d]. In the final paper [Ham19d], a form of Brownian regularity was proved for pre-limiting weight profiles for general initial conditions, to which we return shortly. But we first turn to discussing the Brownian Gibbs property, a crucial idea in the proofs of Theorem 3.1 as well as our own main result.

The Brownian Gibbs property

A central player in our approach is the *Brownian Gibbs property*, and here we discuss previous work in this line of study. The Brownian Gibbs property was first employed in [CH14], to study the *Airy line ensemble*. Recall from Chapter 1 that the Airy line ensemble is an $\mathbb{N}$-indexed collection of continuous, non-intersecting curves, whose uppermost curve is the Airy_2 process. The Brownian Gibbs property is an explicit spatial Markov property enjoyed by the Airy line ensemble after a parabolic shift and multiplication by a factor $2^{-1/2}$ (to allow comparison to Brownian objects of diffusion rate *one*), resulting in the *standardized parabolic Airy line ensemble*. In short, the Brownian Gibbs property says that the conditional distribution of any set of k consecutive curves on an interval $[a, b]$, conditionally on all the other curves on all of $\mathbb{R}$ and the k curves themselves on $(a, b)^c$, is given by k independent rate one Brownian bridges between appropriate endpoints and conditioned to intersect neither each other nor the preceding and succeeding curves.

The Brownian Gibbs property and various softenings of it have proved to be a versatile tool in probabilistic investigations of KPZ. Beyond the already mentioned [CH14], there have been numerous works on line ensembles enjoying this or an analogous property, which we briefly discuss.

The Brownian Gibbs property itself was a central theme in the previously mentioned four-part study [Ham19a, Ham19c, Ham19b, Ham19d] of Brownian LPP. While [CH14] established that the Brownian Gibbs property is enjoyed by the Airy line ensemble, and hence by the limiting weight profiles in a number of LPP models, Brownian LPP is special in that its weight profile satisfies the Brownian Gibbs property even in the *pre-limit*. This is a crucial integrable input first observed by [OY02] (who related the energy profiles in Brownian LPP to Dyson Brownian motion), and is the reason why Brownian LPP is the setting of the mentioned four-part study, as well as why our main results will apply to it. Apart from this four-part study, we mention some other works in this vein. The work [CS14] establishes the ergodicity of the Airy line ensemble using the Brownian Gibbs property. The fractal nature of a certain limiting weight difference profile in Brownian LPP is investigated in [BGH19], using inputs from the four-part study mentioned earlier. The Brownian Gibbs property is used in [CIW19a, CIW19b] to analyse tightness of families of non-intersecting Brownian bridges above a hard wall, subject to a tilting of measure in terms of the area the curves capture below them; they also establish that an area-tilted form of the Brownian Gibbs property is enjoyed by the limiting ensemble.

A softened version of Brownian Gibbs, in which intersection is not prohibited but suffers an energetic penalty, was used in an investigation of the *scaled* solution to the KPZ equation with narrow-wedge initial condition [CH16], establishing for that process absolute continuity with respect to Brownian motion on compact intervals. This form of Brownian Gibbs was also used in the recent [CG18] to obtain bounds on the one-point upper and lower tails for the solution to the KPZ equation from quite general initial data, and in [CGH19] to establish the rate of decay of correlations with time of the narrow wedge solution at the origin. A discrete Gibbsian property was used in [CD18] to study the transversal fluctuation exponent and tightness of the

appropriately scaled height function in the asymmetric simple exclusion process and stochastic six vertex model, started with step initial conditions. A sequence of discrete line ensembles associated to the inverse gamma directed polymer, which obeys a softened discrete version of the Brownian Gibbs property, was shown to be tight in [Wu19].

Finally, we mention the valuable contribution [DOV18], aided by [DV18], which establishes the existence of the space-time Airy sheet using Brownian LPP and the Brownian Gibbs property.

3.2 Method of proof

In this final section of the chapter we compare, on a high level, the method of proof of the main theorem with that of Theorem 3.1 as given in [Ham19a]. A more general form of the main result will be stated in Chapter 4 as Theorem 4.8. Chapter 5 is devoted to describing the general framework for the proof of our main result, with Section 5.2 addressing the conceptual framework specific to the main result Theorem 4.8.

At the highest level, the method of proof of Theorem 3.1 in [Ham19a] relies on embedding the parabolic Airy_2 curve as the uppermost curve in the parabolic Airy line ensemble and employing the Brownian Gibbs property. In [Ham19a], a significant amount of additional technical apparatus, known as the jump ensemble, was developed to further this technique, which culminated in the proof of Theorem 3.1.

The proof of our main theorem is based squarely on the Brownian Gibbs property via the jump ensemble as well, but the details of the proof differ quite substantially from that of Theorem 3.1 because of the difficulties that arise from possibly pathological correlations between the bridge $\mathcal{L}^{[-d,d]}$ and the endpoints $\mathcal{L}(-d)$ and $\mathcal{L}(d)$.

A flavour of this difficulty can be seen even in a purely Brownian toy example quite easily, and this example will be fairly representative because of the Brownian Gibbs property. Suppose we are trying to bound the probability that a Brownian process lies in a particular measurable subset of continuous functions. We are contrasting the situation when the Brownian process is Brownian motion with when it is Brownian bridge; we note that applying the affine shifting procedure described before Theorem 3.1, which defines $\mathcal{L}^{[-d,d]}$ from $\mathcal{L}$, to Brownian motion results in Brownian bridge. Let B be a standard rate one Brownian motion on $[-d, d]$ started at zero, and let $B^{[-d,d]}$ be the Brownian bridge on $[-d, d]$ resulting from the affine shifting procedure.

A standard fact is that $B^{[-d,d]}$ is *independent of the original endpoint value* $B(d)$ of the Brownian motion. Thus, when evaluating the probability that $B^{[-d,d]}$ lies in some subset of continuous functions, one simply has to integrate over $B(d)$; the conditional probability given $B(d)$ is the same for all of them.

In contrast, consider the probability that B lies in a subset A of continuous functions. If we here try to decompose the process by conditioning on its endpoint value $B(d)$, the conditional probability of A depends on $B(d)$. More importantly, the nature of the dependence is not the same for all A, and so there is no clear way to decouple the conditional probability of A from the endpoint values in an event-agnostic way.

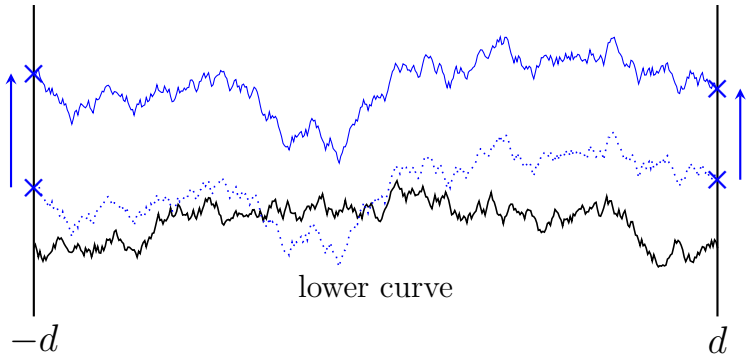


Figure 3.1: The Brownian Gibbs property and an independence property of Brownian bridges essentially reduces the proof of Theorem 3.1 given in [Ham19a] to understanding the probability of non-intersection with the lower curve (i.e., the second curve of the parabolic Airy line ensemble) conditionally on the endpoint values at $-d$ and d. This probability has an important monotonicity property which substantially simplifies the proof of Theorem 3.1: if we raise the endpoint values, the top curve is more likely to be fully above the lower curve. This can be seen by the stochastic domination depicted here, as the top curve with lower endpoint values (blue and dotted) intersects the lower curve (thick and in black), while on raising the endpoint values, non-intersection is achieved. This monotonicity will *not* be available in the proof of the main result here. This is because we do not have access to the independence property of Brownian bridges, which is what allowed the decoupling of the probability of the event under consideration from the probability of non-intersection.

The Brownian Gibbs property in some sense relates the statement to be proved, here regarding the process $\mathcal{L}$ in the form of Theorem 2.2, to considerations similar to this toy example. Recall that, in a loose sense, the Brownian Gibbs property says that the conditional distribution of $\mathcal{L}$ on an interval is that of a Brownian bridge with appropriate endpoints conditioned on being above a lower curve over the whole interval; the lower curve is the second curve of the standardized parabolic Airy line ensemble. We are considering the probability that $\mathcal{L}^{[-d,d]}$ belongs to an event A. On a heuristic level, applying the Brownian Gibbs property and the independence from endpoints enjoyed by Brownian bridge, bounding the conditional probability of A given non-intersection and the endpoint values $\mathcal{L}(-d)$ and $\mathcal{L}(d)$ reduces to bounding the probability of non-intersection given $\mathcal{L}(-d)$ and $\mathcal{L}(d)$; the probability of A under Brownian bridge factors out.

A simplifying feature of the conditional probability of non-intersection given $\mathcal{L}(-d)$ and $\mathcal{L}(d)$ is that it enjoys an intuitive monotonicity in the endpoint values: when they are higher, avoiding the lower curve is more probable (see Figure 3.1). Using this monotonicity, it is sufficient for the proof of Theorem 3.1 to bound the non-intersection probability by obtaining a bound on the endpoint value density in only the case when the endpoints are very low. This is a crucial technical result in [Ham19a], stated as Lemma 5.17. (This description is not completely accurate as in the proof of Theorem 3.1 the technical apparatus of the jump ensemble allows the non-

intersection condition to be not with the entire lower curve but only a certain subset of it. We ignore this point here.)

However, for the process $\mathcal{L}(\,\cdot\,) - \mathcal{L}(-d)$, analogous to the Brownian motion discussion, the probability of an event A and the probability of non-intersection cannot be decoupled given the endpoint values, and the probability of the combined event does not enjoy a monotonicity property in the endpoint values. (Of course, for certain events this monotonicity property would be true, but it does not hold in an event-agnostic manner.) Thus, while in the proof of Theorem 3.1 it was sufficient to have an endpoint value density bound in only the case where the endpoint values are very low, for the main result here we will need corresponding density bounds for the remaining ranges of endpoint values as well. The case of low endpoint values is handled by using the same statement of [Ham19a], Lemma 5.17 there, here stated as Proposition 6.1, but the other ranges of endpoint values give rise to additional cases of greater technical difficulty.

A fuller discussion of the ideas and approach of the proof is provided in Chapter 5, with Section 5.1 discussing the jump ensemble and Section 5.2 discussing the framework specific to the proof of the main result.

3.3 Organization of Part II

In Chapter 4, we introduce the Brownian Gibbs property and the more general objects to which our results apply, and then state the main result in its general form as Theorem 4.8. Chapter 5 sets up the framework in which our proof operates: in Section 5.1 we introduce the jump ensemble, and in Section 5.2 we provide a conceptual framework for the proof of the principal result. Finally, the main theorem is proved in Chapter 6 across four sections, each covering a different case.

Chapter 4

Notation and setup

In this chapter we introduce some notation we will be using throughout Part II of the thesis; give the definitions of the main objects of study; and then state the main result, Theorem 4.8.

4.1 Notation, Brownian Gibbs, and regular ensembles

General notation

We take the set of natural numbers $\mathbb{N}$ to be $\{1, 2, \ldots\}$. For $k \in \mathbb{N}$, we use an overbar to denote a k-vector, i.e., $\overline{x} \in \mathbb{R}^k$. We denote the integer interval $\{i, i+1, \ldots, j\}$ by $[\![i, j]\!]$. For a function $f : [\![k]\!] \times \mathbb{R} \to \mathbb{R}$, we write $\overline{f}(x)$ for $(f(1, x), \ldots, f(k, x))$. A k-vector $\overline{x} = (x_1, \ldots, x_k) \in \mathbb{R}^k$ is called a k-decreasing list if $x_1 > x_2 > \ldots > x_k$. For a set $I \subseteq \mathbb{R}$, let $I^k_> \subseteq I^k$ be the set of k-decreasing lists of elements of I, and $I^k_\geq$ be the analogous set of k-non-increasing lists.

For a real valued function f whose domain of definition contains an interval $[a, b]$, we define $f^{[a,b]} : [a, b] \to \mathbb{R}$ to be the affinely shifted bridge version of f that is zero at both endpoints, i.e., for $x \in [a, b]$,

$$f^{[a,b]}(x) := f(x) - \frac{x-a}{b-a} \cdot f(b) - \frac{b-x}{b-a} \cdot f(a).$$

For an interval $[a, b] \subseteq \mathbb{R}$, we denote the space of continuous functions with domain $[a, b]$ which vanish at a by $\mathcal{C}_{0,*}([a, b], \mathbb{R})$, and the space of continuous functions which may take any value at the endpoints by $\mathcal{C}_{*,*}([a, b], \mathbb{R})$. The asterisk should be thought of as a wildcard indicating that any value may be taken.

Line ensembles and the Brownian Gibbs property

Definition 4.1 (Line ensembles)**.** Let Σ be an (possibly infinite) interval of $\mathbb{Z}$, and let Λ be a (possibly unbounded) interval of $\mathbb{R}$. Let $\mathcal{X}$ be the set of continuous functions $f : \Sigma \times \Lambda \to \mathbb{R}$

endowed with the topology of uniform convergence on compact subsets of $\Sigma \times \Lambda$, and let $\mathcal{C}$ denote the Borel σ-algebra of $\mathcal{X}$.

A Σ-*indexed line ensemble* $\mathcal{L}$ is a random variable defined on a probability space $(\Omega, \mathcal{B}, \mathbb{P})$, taking values in $\mathcal{X}$ such that $\mathcal{L}$ is a $(\mathcal{B}, \mathcal{C})$-measurable function. We regard $\mathcal{L}$ as a Σ-indexed collection of random continuous curves (despite the usage of the word "line"), each of which maps Λ into $\mathbb{R}$. We will slightly abuse notation and write $\mathcal{L} : \Sigma \times \Lambda \to \mathbb{R}$, even though it is not $\mathcal{L}$ which is such a function, but rather $\mathcal{L}(\omega)$ for each $\omega \in \Omega$. A line ensemble is *ordered* if, for all $i, j \in \Sigma$ with $i < j$, it holds that $\mathcal{L}(i, x) > \mathcal{L}(j, x)$ for all $x \in \Lambda$. Statements such as this are understood as being asserted almost surely with respect to $\mathbb{P}$.

Definition 4.2 (Normal, Brownian bridge, and Brownian motion laws)**.** We will use $N(m, \sigma^2)$ to denote the normal distribution with mean m and variance σ^2, and sometimes, with abuse of notation, a random variable with this distribution.

Let $k \in \mathbb{N}$, $a, b \in \mathbb{R}$ with $a < b$, and $\overline{x}, \overline{y} \in \mathbb{R}^k_>$. We write $\mathcal{B}^{[a,b]}_{k;\overline{x},\overline{y}}$ for the law of k independent Brownian bridges $(B_1, \ldots, B_k)$ of diffusion parameter one, with $B_i : [a, b] \to \mathbb{R}$ and $B_i(a) = x_i$ and $B_i(b) = y_i$, for $i = 1, \ldots, k$.

We will also need the law of standard Brownian motion started at 0 on the interval $[a, b]$, which we will denote by $\mathcal{B}^{[a,b]}_{0,*}$; i.e., $\mathcal{B}^{[a,b]}_{0,*}$ is the law of a rate one Brownian motion B with $B(a) = 0$.

Now let $f : [a, b] \to \mathbb{R} \cup \{-\infty\}$ be a measurable function such that $x_k > f(a)$ and $y_k > f(b)$. Define the non-intersection event on a set $A \subseteq [a, b]$ with lower boundary curve f by

$$\mathsf{NonInt}^A_f = \Big\{ \text{for all } x \in A\,,\ B(i, x) > B(i{+}1, x) \text{ for each } 1 \le i \le k{-}1, \text{ and } B(k, x) > f(x) \Big\}.$$

When $A = [a, b]$, we omit its mention in the notation, i.e., we write NonInt_f.

With this definition, we can move to defining the Brownian Gibbs property.

Definition 4.3 (Brownian Gibbs property)**.** Let $n \in \mathbb{N}$, $I \subseteq \mathbb{R}$ be an interval, $k \in [\![n]\!]$, and $a, b \in I$ with $a < b$. Let $D_{k;a,b} = [\![k]\!] \times (a, b)$ and $D^c_{k;a,b} = ([\![n]\!] \times I) \setminus D_{k;a,b}$. Let $\mathcal{L} : [\![n]\!] \times I \to \mathbb{R}$ be an ordered line ensemble. We say that $\mathcal{L}$ has the Brownian Gibbs property if the following holds for all such choices of k, a, and b:

$$\text{Law}\Big(\mathcal{L}|_{D_{k;a,b}} \text{ conditionally on } \mathcal{L}|_{D^c_{k;a,b}} \Big) = \mathcal{B}^{[a,b]}_{k;\overline{x},\overline{y}}\,(\,\cdot \mid \mathsf{NonInt}_f)\,,$$

where $\overline{x} = \overline{\mathcal{L}}(a)$, $\overline{y} = \overline{\mathcal{L}}(b)$, and $f(\,\cdot\,) = \mathcal{L}(k+1, \cdot\,)$ on $[a, b]$.

In words, the conditional distribution of the top k curves of $\mathcal{L}$ on $[a, b]$, given the form on $\mathcal{L}$ on $D^c_{k;a,b}$, is the law of k independent Brownian bridges, the i^{th} from $\mathcal{L}(i, a)$ to $\mathcal{L}(i, b)$, which are conditioned to intersect neither each other nor the lower curve $\mathcal{L}(k+1, \cdot\,)$ on $[a, b]$.

In the next definition we define *regular ensembles*, which are the general objects to which our main result will apply. The definition is the same as [Ham19a, Definition 2.4], with the parameter

$\overline{\varphi}$ in that definition taking the value $(1/3, 1/9, \infty)$; the value of ∞ for the third parameter is a formal device to indicate that the range of s in point (2) below is $[1, \infty)$ instead of $[1, n^{\varphi_3}]$ for a finite value of φ_3.

Definition 4.4 (Regular Brownian Gibbs ensemble)**.** Consider a Brownian Gibbs ensemble that has the form

$$\mathcal{L} : [\![n]\!] \times \big[-z_n, \infty\big) \to \mathbb{R}\,,$$

and which is defined on a probability space under the law $\mathbb{P}$. The number $n = n(\mathcal{L})$ of ensemble curves and the absolute value z_n of the finite endpoint may take any values in $\mathbb{N}$ and $[0, \infty)$. (In fact, we may also take $z_n = \infty$, in which case we would take the domain of $\mathcal{L}$ to be $[\![n]\!] \times \mathbb{R}$.) Let C and c be two positive constants. The ensemble $\mathcal{L}$ is said to be (c, C)-regular if the following conditions are satisfied.

1. **Endpoint escape.** $z_n \geq cn^{1/3}$.
2. **One-point lower tail.** If $z \geq -z_n$ satisfies $|z| \leq cn^{1/9}$, then

$$\mathbb{P}\Big(\mathcal{L}\big(1, z\big) + 2^{-1/2}z^2 \leq -s\Big) \leq C \exp\big\{ -cs^{3/2}\big\}$$

for all $s \in \big[1, \infty\big)$.

3. **One-point upper tail.** If $z \geq -z_n$ satisfies $|z| \leq cn^{1/9}$, then

$$\mathbb{P}\Big(\mathcal{L}\big(1, z\big) + 2^{-1/2}z^2 \geq s\Big) \leq C \exp\big\{ -cs^{3/2}\big\}$$

for all $s \in [1, \infty)$.

We reserve the symbols c and C for this usage in the remainder of this part of the thesis.

The symbol n will be reserved in the rest of this part of the thesis for the number of curves in the regular ensemble under consideration, which we will denote by $\mathcal{L}_n$.

Though the definition of regular ensembles only includes one-point tail information for the top curve, this actually extends to the lower curves as well [Ham19a, Proposition 2.7]. Though we do not state this result, we will have need of two associated sequences of constants for the statement of our main results. For a (c, C)-regular ensemble, define $C_1 = 140C$, $c_1 = 2^{-5/2}c \wedge 1/8$; and, for each $k \geq 2$,

$$C_k = \max\Big\{10 \cdot 20^{k-1}5^{k/2}\Big(\tfrac{10}{3-2^{3/2}}\Big)^{k(k-1)/2}C\,,\, e^{c/2}\Big\}$$

and

$$c_k = \big((3 - 2^{3/2})^{3/2}2^{-1}5^{-3/2}\big)^{k-1}c_1. \tag{4.1}$$

These symbols will retain these meanings throughout this part of the thesis.

One example of a regular Brownian Gibbs line ensemble is the *standardized parabolic Airy line ensemble*, given by

$$\mathcal{L}(i, x) = 2^{-1/2}(\mathcal{A}(i, x) - x^2),$$

for $(i,x) \in \mathbb{N} \times \mathbb{R}$, where $\mathcal{A} : \mathbb{N} \times \mathbb{R} \to \mathbb{R}$ is the Airy line ensemble. (We include a factor of $2^{-1/2}$ to allow comparisons to be made with rate *one* Brownian objects.) The Airy line ensemble was constructed as an ensemble of continuous non-intersecting curves in [CH14, Theorem 3.1], and tightness estimates furnished by each of [DV18], [DNV19], and [Ham19a] lead to simplified constructions. It is defined as follows.

Definition 4.5 (Airy line ensemble)**.** The Airy line ensemble $\mathcal{A} : \mathbb{N} \times \mathbb{R} \to \mathbb{R}$ is a collection of random continuous curves $\mathcal{A}(j, \cdot\,)$ for $j \in \mathbb{N}$. For any finite set $I \subset \mathbb{R}$, define the random object $\mathcal{A}[I]$ to be the point process on $I \times \mathbb{R}$ given by $\big\{\big(s, \mathcal{A}(j,s)\big) \mid j \in \mathbb{N},\, s \in I\big\}$. The law of $\mathcal{A}$ is defined as the unique distribution supported on such collections of continuous curves such that, for each finite $I = \{t_1, \cdots, t_m\}$, $\mathcal{A}[I]$ is a determinantal point process whose kernel is the extended Airy$_2$ kernel K_2^{ext}, specified by

$$K_2^{\text{ext}}\big(s_1, x_1; s_2, x_2\big) = \begin{cases} \int_0^\infty e^{-\lambda(s_1 - s_2)} \mathrm{Ai}\big(x_1 + \lambda\big) \mathrm{Ai}\big(x_2 + \lambda\big) \mathrm{d}\lambda & \text{if } s_1 \geq s_2\,, \\ -\int_{-\infty}^0 e^{-\lambda(s_1 - s_2)} \mathrm{Ai}\big(x_1 + \lambda\big) \mathrm{Ai}\big(x_2 + \lambda\big) \mathrm{d}\lambda & \text{if } s_1 < s_2\,, \end{cases}$$

where $\mathrm{Ai} : \mathbb{R} \to \mathbb{R}$ is the Airy function. The Airy line ensemble's curves are ordered, with $\mathcal{A}(1, \cdot\,)$ uppermost.

4.2 An important example of regular ensembles: Brownian LPP weight profiles

Here we introduce the Brownian last passage percolation model, which will generate an important example of regular ensembles via the RSK correspondence, and weight profiles from general initial conditions. These definitions are not logically required for the proof of our main theorem, but do motivate our decision to prove the result in the more general context of regular ensembles. Many of these objects were introduced in Section 2.1, but we recall aspects of their definitions here.

First is the weight profile of Brownian LPP, $\mathcal{P}_{n,1}$, from (2.1). As we saw in Chapter 2, the function $y \mapsto 2^{-1/2}\mathcal{P}_{n,1}(y)$ (which we call the *weight profile*) is a tight sequence of random functions which converges to $\mathcal{L}_1$, the standardized parabolic Airy$_2$ process, which is the top curve in the standardized parabolic Airy line ensemble mentioned above. These inferences follow from the relation between $\mathcal{P}_n$ and Dyson Brownian motion proved in [OY02] and the fact that the scaling limit of Dyson Brownian motion is the Airy$_2$ process [AVM05] in the sense of finite-dimensional distributions, upgraded to the space of continuous functions by [CH14]. The equality in distribution with Dyson Brownian motion for the top line alone was proved earlier in [GTW01] and also [Bar01].

Again recalling from Chapter 2, we may regard this function $y \mapsto \mathcal{P}_{n,1}(y)$ as the top line in an ensemble of n continuous curves which we denote $\mathcal{P}_n : [\![n]\!] \times [-\frac{1}{2}n^{1/3}, \infty) \to \mathbb{R}$. We also recall Theorem 2.1 that $\mathcal{P}_n$ converges to the parabolic Airy line ensemble under the standard notion

of weak convergence given the locally uniform topology on curves. This is proved by the same references mentioned in the previous paragraph for the top line of the ensemble.

Our reason for considering this ensemble of curves is that it enjoys the Brownian Gibbs property and is in fact regular.

Proposition 4.6. *There exist choices of the positive constants c and C such that each of the scaled Brownian LPP line ensembles $2^{-1/2}\mathcal{P}_n : [\![n]\!] \times \big[-\frac{1}{2}n^{1/3}, \infty\big) \to \mathbb{R}$, $n \in \mathbb{N}$, is $\big(c, C\big)$-regular.*

In fact, this is almost [Ham19a, Proposition 2.5], which proves that $2^{-1/2}\mathcal{P}_n$ is (c, C) regular in a slightly weaker sense, namely with point (2) in Definition 4.4 holding for $s \in [1, n^{1/3}]$ only. Below we point out how to modify that proof to obtain the complete claimed range of s.

Proof. [Ham19a, Proposition 2.5] proves that $2^{-1/2}\mathcal{P}_n$ satisfies points (1) and (3) of Definition 4.4, and so we only need to prove point (2) for all of $s \in [1, \infty)$. To do this, we simply replace the use of [Ham19a, Lemma A.1(1)] in the proof of [Ham19a, Proposition 2.5] with [DV18, Theorem 3.1]; this latter theorem is an improved moderate deviation bound for the k^{th} line of Dyson Brownian motion (equivalently, the k^{th} eigenvalue of the Gaussian Unitary Ensemble), which we need for only $k = 1$. □

We have stated this slightly improved regularity of $2^{-1/2}\mathcal{P}_n$ in comparison to the statement of [Ham19a, Proposition 2.5] for completeness. If Proposition 4.6 were used in place of [Ham19a, Proposition 2.5] in the arguments of [Ham19a], minor improvements to certain statements quoted from [Ham19a] that we use later, in Section 5.1, could be made. However, in view of the minor and technical nature of these improvements, we do not formally claim, state, or use them in our arguments, and therefore we will not carry through these improved effects of Proposition 4.6 further.

Basic parabolic symmetry of regular ensembles.

Here we record a straightforward proposition that allows us to translate the interval of consideration and still retain a regular ensemble (with an extra linear term).

Let $Q : \mathbb{R} \to \mathbb{R}$ denote the parabola $Q(x) = 2^{-1/2}x^2$, and let $l : \mathbb{R}^2 \to \mathbb{R}$ be given by $l(x, y) = -2^{-1/2}y^2 - 2^{1/2}y(x - y)$. Note that $x \mapsto l(x, y)$ is the tangent line of the parabola $x \mapsto -Q(x)$ at the point $(y, -Q(y))$. Note also that, for any $x, y \in \mathbb{R}$,

$$Q(x) = -l(x, y) + Q(x - y). \tag{4.2}$$

For $z_n \geq 0$, consider a regular ensemble $\mathcal{L}_n : [\![n]\!] \times [-z_n, \infty) \to \mathbb{R}$. For any $y_n > -z_n$, define $\mathcal{L}^{\text{shift}}_{n,y_n} : [1, n] \times [-z_n - y_n, \infty) \to \mathbb{R}$ to be the shifted ensemble given by

$$\mathcal{L}^{\text{shift}}_{n,y_n}(i, x) = \mathcal{L}_n\left(i, x + y_n\right) - l\left(x + y_n, y_n\right)$$

By (4.2), $\mathcal{L}^{\text{shift}}_{n,y_n} = \mathcal{L}_n(i, x + y_n) + Q(x + y_n) - Q(x)$.

Lemma 4.7 (Lemma 2.26 of [Ham19a]). *Let $c, C > 0$ and $n \in \mathbb{N}$. Suppose that $\mathcal{L}_n : [\![n]\!] \times [-z_n, \infty) \to \mathbb{R}$ is a (c, C)-regular ensemble. Whenever $y_n \in \mathbb{R}$ satisfies $|y_n| \leq c/2 \cdot n^{1/9}$, the ensemble $\mathcal{L}^{\text{shift}}_{n,y_n}$ is $(c/2, C)$-regular.*

This lemma will allow our main result to apply to an interval $[K - d, K + d]$ not necessarily centred at the origin.

4.3 Main result

For $k \in \mathbb{N}$, let D_k be a sequence of constants depending only on k, given by

$$D_k = \max\left\{k^{1/3}c_k^{-1/3}\left(2^{-9/2} - 2^{-5}\right)^{-1/3}, 36\left(k^2 - 1\right), 2\right\} \tag{4.3}$$

for $k \geq 2$, and set $D_1 = D_2$; here c_k is as given in (4.1). This will be the value of D_k for the rest of Part II of the thesis.

Our main result will concern an interval $[K - d, K + d]$ for $K \in \mathbb{R}$ and $d \geq 1$. For such K, define the linear function $\ell_{K,d} : \mathbb{R} \to \mathbb{R}$ by

$$\ell_{K,d}(x) = 2^{1/2}K(x - K + d).$$

Our main result is a generalization of Theorem 2.2 that applies even when the interval under consideration is not centred at zero.

Theorem 4.8. *Suppose that $\mathcal{L}_n$ is an n-curve (c, C) regular ensemble for some $(c, C) \in (0, \infty)^2$. Let $d \geq 1$ denote a parameter. Let $K \in \mathbb{R}$ satisfy $[K - d, K + d] \subset c/2 \cdot [-n^{1/9}, n^{1/9}]$, and let $k \in \mathbb{N}$.*

Suppose that $n \geq k \vee (c/3)^{-18} \vee 6^{36}$. For any Borel measurable $A \subset \mathcal{C}_{0,*}\big([K - d, K + d]\big)$, write $\varepsilon = \mathcal{B}^{[K-d,K+d]}_{0,*}(A)$. Suppose that ε satisfies the (k, d)-dependent upper bound $\varepsilon < e^{-1} \wedge (17)^{-1/k}C_k^{-1/k}D_k^{-1} \wedge \exp(-(24)^6 d^6/D_k^3)$; as well as the n-dependent lower bound

$$\varepsilon \geq \exp\left\{-\left(c/2 \wedge 2^{1/2}\right)D_k^{-1}n^{1/12}\right\}. \tag{4.4}$$

Then there exists $G < \infty$ such that

$$\mathbb{P}\left(\mathcal{L}_n\big(k, \cdot\big) - \mathcal{L}_n\big(k, K - d\big) + \ell_{K,d}(\,\cdot\,) \in A\right) \leq \varepsilon \cdot G \cdot \exp\left\{4932 \cdot d \cdot D_k^{5/2}\big(\log \varepsilon^{-1}\big)^{5/6}\right\}.$$

Specifically, this probability is $\varepsilon \cdot \exp\left\{d(\log \varepsilon^{-1})^{5/6}O_k(1)\right\}$, where $O_k(1)$ denotes a k-dependent term that is independent of ε and d.

Remark 4.9. The upper bound on ε is only a technical one and is of no real consequence. The rapid decay in n of the lower bound (4.4) means that no difficulty is created in applications, since, roughly put, events whose probabilities have decay that is superpolynomial in n are in practice irrelevant. In the case that $n = \infty$, such as for the standardized parabolic Airy line ensemble, this lower bound becomes the vacuous $\varepsilon > 0$.

Remark 4.10. The linear term $\ell_{K,d}$ introduced in the event in the general result is necessary. It arises from the parabolic curvature of regular ensembles, which cannot be ignored when the interval $[K-d, K+d]$ is far from the origin. In fact, we will prove the theorem for $K = 0$, and then use the parabolic invariance introduced in Section 4.2 to get the general statement, as we have $\mathcal{L}^{\text{shift}}_{n,K}(k,\cdot) - \mathcal{L}^{\text{shift}}_{n,K}(k,-d) = \mathcal{L}_n\big(k,\cdot+K\big) - \mathcal{L}_n\big(k,K-d\big) + \ell_{K,d}(\cdot+K)$ on $[-d,d]$.

Next, we present a result that permits us to dispense with Theorem 4.8's inconsequential but practically irksome Brownian probability hypothesis of a d- and k-dependent upper bound on ε.

Corollary 4.11. *Under the circumstances of Theorem 4.8, suppose on the parameter ε we assume only the lower bound $\varepsilon \geq \exp(-gn^{1/12})$ holds, where $g = g(k) = (c/2 \wedge 2^{1/2})D_k^{-1}$. Then there exists $H = H(k)$ such that*

$$\mathbb{P}\Big(\mathcal{L}_n(k,\cdot) - \mathcal{L}_n(k,K-d) + \ell_{K,d}(\cdot) \in A\Big) \;\leq\; \varepsilon \cdot \exp\Big\{Hd^6\Big\} \exp\Big\{Hd\big(\log\varepsilon^{-1}\big)^{5/6}\Big\}.$$

Proof. Let $h = e^{-1} \wedge (17)^{-1/k} C_k^{-1/k} D_k^{-1} \wedge \exp(-(24)^6 d^6/D_k^3)$ be the expression upper bounding ε in Theorem 4.8.

When ε satisfies the upper bound $\varepsilon \leq h$ of Theorem 4.8 in addition to the assumed lower bound $\varepsilon \geq \exp(-gn^{1/12})$, the claim follows from the same theorem by taking $H = H(k) \geq 4932 D_k^{5/2}$ large enough that $\exp(Hd^6) \geq G$, where G is as in that theorem; this can be done such that H does not depend on d since $d \geq 1$.

When we instead have $\varepsilon \geq h$, note that $\varepsilon \cdot h^{-1} \geq 1$, and so

$$\mathbb{P}\Big(\mathcal{L}_n(k,\cdot) - \mathcal{L}_n(k,K-d) + \ell_{K,d}(\cdot) \in A\Big) \leq 1 \leq \varepsilon \cdot h^{-1}.$$

Now let $H = H(k)$ be such that $h^{-1} \leq \exp(Hd^6)$, which is possible by the definition of h and since $d \geq 1$. This yields our claim since $\exp(Hd(\log\varepsilon^{-1})^{5/6}) \geq 1$. □

Chapter 5

Proof framework

In this chapter we introduce the two frameworks required for our proof: the first is the *jump ensemble*, a general technique introduced in [Ham19a] which allows one to analyse regular Brownian Gibbs line ensembles using a more explicitly Brownian proxy; while the second is specific to our proof of Theorem 4.8 and is a conceptual framework of *costs*. We will also reduce the proof of Theorem 4.8 to a statement, Theorem 5.11, about the jump ensemble, and Chapter 6 will be devoted to providing a major part of the proof of this statement using the introduced framework of costs.

5.1 The jump ensemble

We start with a working description of the technical framework in which our proof approach operates, known as the jump ensemble. The jump ensemble should be thought of as a sort of "half-way house" between Brownian motion and the line ensemble $\mathcal{L}_n$ that we wish to study. Roughly speaking, what we mean by this is that the jump ensemble conditioned on a certain manageable event has the same distribution as $\mathcal{L}_n$; but since the jump ensemble can be described in terms of Brownian objects, we can estimate probabilities involving the jump ensemble using knowledge about Brownian motion.

The construction we describe is the same as that in [Ham19a, Chapter 4]. The reader is referred to that article for a fuller discussion; here we restrict ourselves to providing a complete, though perhaps sometimes not fully motivated, description of the jump ensemble that allows the reader to understand the proofs. The notation used in this section is largely the same as in [Ham19a] for the convenience of the reader. We stress that some of the proofs underlying the correctness and usefulness of the jump ensemble as given in [Ham19a] are technically involved, and so we choose to not reproduce them here, instead focusing only on illustrating the ideas and statements of the jump ensemble.

We use only three statements from [Ham19a], reproduced here as Lemma 5.3, Proposition 5.9, and Lemma 5.10. We call these three statements *the side interval test*; the *jump ensemble candidate*

proficiency; and the *high probability of the favourable event*. The reason for the use of these names will become clearer over the next few subsections.

Motivation and main themes

Before turning to the details of the jump ensemble, let us bring to focus some of the main themes. Recall that we aim to study the k^{th} curve of $\mathcal{L}_n$. To do this, we initially consider the top k curves together. The basic tool we have at our disposal in studying regular line ensembles is the Brownian Gibbs property. To recall it, let $\mathcal{F}_{\text{BB}}$ be the σ-algebra generated by the following collection of random variables (where the BB subscript stands for "Brownian bridge"):

- all the lower curves $\mathcal{L}_n : [\![k+1, n]\!] \times [-z_n, \infty) \to \mathbb{R}$;
- and the top k curves $\mathcal{L}_n : [\![k]\!] \times ([-z_n, \ell] \cup [r, \infty)) \to \mathbb{R}$ outside (ℓ, r).

(Though this σ-algebra's definition clearly depends on k, we suppress this dependence in the notation $\mathcal{F}_{\text{BB}}$.)

The statement of the Brownian Gibbs property is then that, conditionally on $\mathcal{F}_{\text{BB}}$, the top k curves of $\mathcal{L}_n$ on $[\ell, r]$ have the same distribution as a collection of k independent Brownian bridges, the i^{th} from $\mathcal{L}_n(i, \ell)$ to $\mathcal{L}_n(i, r)$, with the curves in the collection conditioned on intersecting neither $\mathcal{L}_n(k+1, \cdot)$ nor each other on all of $[\ell, r]$.

Candidate ensembles

We interpret this description as a *resampling property*, which is to say that, given the data in $\mathcal{F}_{\text{BB}}$, the top k curves of $\mathcal{L}_n$ on $[\ell, r]$ are obtained by rejection sampling collections of k independent Brownian bridges with the given endpoints until they fully avoid $\mathcal{L}_n(k+1, \cdot)$ and each other on $[\ell, r]$. We call the curves' avoidance of each other on $[\ell, r]$ *internal* non-intersection.

This resampling interpretation suggests a slightly different viewpoint on the Brownian Gibbs property. Let us call the collection of k independent Brownian bridges with the given endpoints a *candidate* ensemble; it has forgotten all information about the lower curve, as its definition only involves $\mathcal{L}_n$ on $[\![k]\!] \times \{\ell, r\}$. Our desire is for the candidate ensemble to gain the correct $\mathcal{F}_{\text{BB}}$-conditional distribution of the top k curves of $\mathcal{L}_n$ on $[\ell, r]$. In order for this to happen, the candidate ensemble must *reconstruct* the effect of the forgotten data as well as satisfy the other constraints that $\mathcal{L}_n$ does.

The basic relation between the $(k+1)^{\text{st}}$ curve of $\mathcal{L}_n$ and the top k-curves of $\mathcal{L}_n$ is that the top k curves must not intersect the $(k+1)^{\text{st}}$; beyond this, the additional constraint that the top k curves of $\mathcal{L}_n$ satisfy which the Brownian bridge candidate does not necessarily is of internal non-intersection. The mentioned reconstruction is done by passing a *test of non-intersection*, both with the lower curve and internally. The reinterpretation of the Brownian Gibbs property is that the candidate ensemble, on passing the non-intersection test, gains the target line ensemble's

$\mathcal{F}_{\text{BB}}$-conditional distribution. In terms of rejection sampling, the rejection sampling probability is exactly the probability of the candidate ensemble passing this test.

Here, the role of $\mathcal{F}_{\text{BB}}$ is to specify the data which the candidate ensemble, on passing the non-intersection test, must conform to. In particular, the distribution attained by the candidate ensemble on passing the non-intersection test is the $\mathcal{F}_{\text{BB}}$-conditional distribution of $\mathcal{L}_n$. The data in $\mathcal{F}_{\text{BB}}$ should be thought of as the data conditioned on, and so available to the candidate ensemble, some of which it then forgets. The data not in $\mathcal{F}_{\text{BB}}$ is, of course, not available to the candidate ensemble at all.

This idea that the candidate ensemble forgets some amount of data available to it is an important one, and one which we will develop further over the next few pages. In particular, we will consider the effects of retaining and forgetting different quantities of data; as we shall see shortly, the example here, of forgetting the entire bottom curve, is too extreme and will not be useful for our purposes.

So the broad theme may be described as follows in a two-step process. First, we condition on a certain selection of data, here represented by $\mathcal{F}_{\text{BB}}$; and second, we consider candidate ensembles which retain some subset of this data and forget the rest. The candidate ensemble recovers the correct conditional distribution, specified by the data first conditioned on, in spite of the forgotten data, by resampling till the appropriate constraint is met, which is that of non-intersection.

Remark 5.1. The language of "retaining data" we are using in this discussion is slightly at odds with the usage in [Ham19a]. There, retained data refers to the data contained in a σ-algebra such as $\mathcal{F}_{\text{BB}}$, with respect to which the conditional distribution of $\mathcal{L}_n$ is considered. Here, by retained data we mean the data contained in this σ-algebra which is *further* retained by the candidate ensemble, in the sense that the further retained data is involved in the specification of the candidate ensemble. Thus, by retaining different quantities of data in this sense, we can generate various candidate ensembles which, on passing the respective non-intersection tests, will each have the distribution of $\mathcal{L}_n$ conditionally on the same σ-algebra.

Features of a useful candidate

So we see that we must consider other candidate ensembles, and the jump ensemble will be one such. What are the features of a useful candidate?

The final aim is to estimate probabilities for the k^{th} curve of the line ensemble $\mathcal{L}_n$. So the features we need of a candidate ensemble to reach this aim is that we must be able to

(i) estimate probabilities of interest for the candidate ensemble; and
(ii) translate them to estimates on probabilities for the line ensemble.

To successfully estimate probabilities for the candidate ensemble, it must be amenable to the tools at our disposal, which in practice means it must be sufficiently Brownian (this is also imposed by our intention to use the Brownian Gibbs property); while to successfully translate estimates

to the line ensemble, it should be intuitively clear that we need the probability of passing the non-intersection test to be not too low.

The high jump difficulty

So let us consider how the Brownian bridge candidate fares in meeting the aims (i) and (ii). Since the candidate is an ensemble of independent Brownian bridges, the point (i) from the previous paragraph is clearly easily met. But on (ii) unfortunately, because of the weak control that we have over $\mathcal{L}_n(k+1,\cdot)$ and the intricacies of that random function, it is difficult to obtain sufficiently strong control on the probability of passing the non-intersection test with the lower curve. (Roughly speaking, the Brownian bridge candidate was the one used in [CH14] to analyse the absolute continuity of Brownian Gibbs ensembles with respect to Brownian motion, and a large part of that paper was spent obtaining control over exactly this non-intersection test passing probability.)

In these terms, we do not have good control over the test passing probability of pure Brownian bridge, and so this candidate is not directly useful. This points to the need to look for better-suited candidate processes. To understand how a better candidate process should be designed, let us consider what made the Brownian bridge candidate have a low test passing probability.

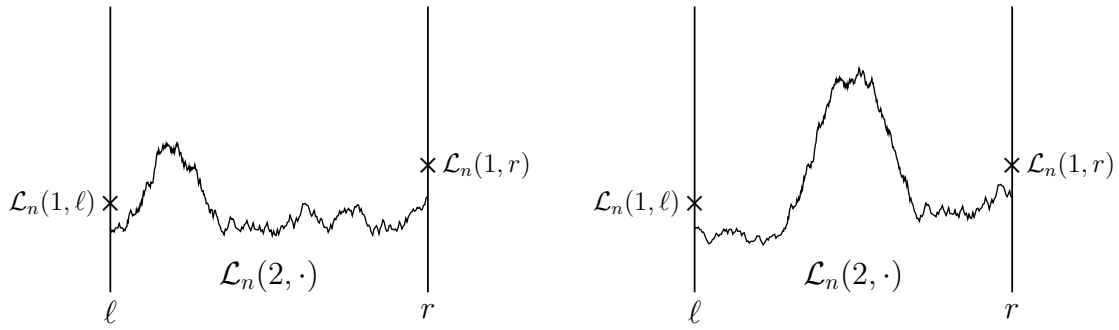


Figure 5.1: Two illustrations in the $k=1$ case of instances of data from the lower curve which are difficult for the Brownian bridge candidate to handle. The black crosses indicate the values of $\mathcal{L}_n(1,\cdot)$ at the endpoints, which are the points between which the Brownian bridge candidate must move. In the left panel there is a moderate sized peak very close to the left side of $[\ell, r]$, which causes difficulty because of the immediacy of the jump required by the candidate. In the right panel there is a large peak which causes difficulty because of its height.

In essence, the Brownian bridge candidate ensemble forgot too much of the data in $\mathcal{F}_{\text{BB}}$, and is thus too far in nature from $\mathcal{L}_n$, to have a high probability of lower curve non-intersection. In particular, it forgot *all* data of the profile of $\mathcal{L}_n(k+1,\cdot)$ that it might have used to increase its probability of avoidance. We shall consider two instances of the lower curve data which is difficult for the Brownian bridge to avoid in order to illustrate two features that our replacement candidate will need. These instances are depicted in Figure 5.1.

For the first instance, suppose that the lower curve has a peak inside $[\ell, r]$ which is close to one side of the interval, say ℓ, as illustrated in the first panel of Figure 5.1. Then the Brownian bridge candidate, to succeed in the non-intersection test, must execute a jump immediately. The difficulty is that since the space to make the jump is limited, a more extreme jump is needed even if the peak is not very large. The low probability of Brownian bridges making such a jump in turn makes the non-intersection test passing probability of the Brownian bridge candidate ensemble low. This discussion suggests the first feature that will aid a successful candidate: we can provide it extra space to make a *run-up* before any required jump.

Turning to the second panel of Figure 5.1, the second instance of difficult lower curve data is when it exhibits an extremely large peak somewhere inside $[\ell, r]$ (which is not necessarily close to either side). It may seem that giving space for a run-up would address this difficulty as well, as a Brownian bridge is clearly more likely to make a bigger jump over a larger interval. However, giving a run-up is in fact not sufficient to handle this sort of data while maintaining a not-too-low non-intersection probability; quantitative reasoning for this conclusion is explored more fully in the beginning of [Ham19a, Chapter 4] and also briefly, in the context of the jump ensemble, in Remark 5.7 ahead. (The discussion in [Ham19a] concerns a setup incorporating a run-up which we will introduce shortly in Subsection 5.1.)

Heuristically, the reason for the difficulty of this data is that the Brownian bridge, having forgotten the entirety of the lower curve, does not know when the jump is required. And so, as alluded to earlier, the second feature of assistance a successful candidate should make use of is to retain more information about the lower curve. More formally, by "using retaining data", we mean that the candidate ensemble will be conditioned to avoid intersection with a curve formed from the retained data (apart from retaining data to specify the values of the candidate ensemble at the endpoints). This will become clearer as our discussion progresses.

(One might wonder about the likelihood of encountering the sort of lower curve data we have been discussing, and whether we cannot exclude such difficult data from $\mathcal{F}_{\mathrm{BB}}$ in the analysis. In our final argument we will indeed restrict ourselves to data in a σ-algebra analogous to $\mathcal{F}_{\mathrm{BB}}$ which is favourable and extremely likely. However, even under such a restriction to more favourable data, our control on the lower curve is not strong enough to exclude data such as what has been discussed.)

Next we move to discussing in more detail the two changes we have mentioned: retaining more data and giving a run-up.

A coarsened lower curve profile

We first discuss the second feature we mentioned, namely retaining a selection of data from the lower curve profile. It should be clear that we should not retain all the data, as this would result in the candidate essentially being the same as the top k curves of $\mathcal{L}_n$ itself; it is difficult to estimate the probabilities of such an ensemble. So we must make a careful selection which balances between retaining no information, as in the Brownian bridge ensemble, and retaining

full information, as in the pure line ensemble; further, the retained information must provide a rough view of the overall geometry of the lower curve.

In fact, we will have the candidate ensemble retain a *coarsened* version of the lower curve. More precisely, the candidate will be conditioned on avoiding this coarsened version. This coarsened non-avoidance can be thought of as a preliminary test to the full non-intersection test; the candidate, on passing the preliminary test, will naturally have a more suitable overall geometry to pass the final test, and thus will have a higher probability of doing so. The exact form of this coarsening, which we will describe in Section 5.1, is at the heart of the jump ensemble method.

Making space

Now let us turn to see how we can provide the first kind of assistance, namely to provide the candidate ensemble with space to make a run-up to more successfully jump over the lower curve. The only way to make space is to step back from the interval $[\ell, r]$. In fact, we will work in an interval $[-2T, 2T]$ which contains $[\ell, r]$, with the parameter T's value to be assigned later. Let us label as *side intervals* the intervals $[-2T, \ell]$ and $[r, 2T]$, and as the *middle interval* the interval $[\ell, r]$.

Working in $[-2T, 2T]$ means that the values of the candidate are not pre-determined at ℓ and r, as in Figure 5.1, but at $\pm 2T$. Of course, simply working on a bigger interval does not gain us anything immediately, since, in our current setup of conditioning on $\mathcal{F}_{\text{BB}}$, the non-intersection must now be done on the larger interval.

To deal with this, we change the setup by changing the data we condition on. Instead of conditioning on $\mathcal{F}_{\text{BB}}$, we consider the σ-algebra $\mathcal{F}$ generated by the following collection of random variables:

- all the lower curves $\mathcal{L}_n : [\![k+1, n]\!] \times [-z_n, \infty) \to \mathbb{R}$;
- the top k curves $\mathcal{L}_n : [\![k]\!] \times ([-z_n, -2T] \cup [2T, \infty)) \to \mathbb{R}$ outside $(-2T, 2T)$;
- and the $2k$ standard bridges $\mathcal{L}_n^{[-2T,\ell]}(i, \cdot\,)$ and $\mathcal{L}_n^{[r,2T]}(i, \cdot\,)$ for $i = 1, \dots, k$.

(Recall here the notation $f^{[a,b]}$ introduced in Section 4.1 for the affinely shifted bridge version of a function f, though mildly abused here to refer to the bridge version of the i^{th} curve of the ensemble and not the i^{th} curve of an undefined bridge version of the ensemble.) We again suppress the k dependence of the σ-algebra in the notation $\mathcal{F}$.

In words, we retain data of the entirety of all the lower curves; the top k curves outside $(-2T, 2T)$; and, on $[-2T, \ell]$ and $[r, 2T]$, the standard bridge paths of the top k curves on these intervals, which we will call the *side bridges*. Nothing is retained on $[\ell, r]$, and, in particular, the values of the candidate ensemble at ℓ and r are not determined.

Remark 5.2. The side bridges may appear to be complicated objects to condition upon; in fact, they are easy to handle because of the Brownian Gibbs property and an independence property possessed by the corresponding side bridge decomposition of Brownian bridges. See Lemma 5.4 ahead.

The σ-algebra $\mathcal{F}$ and the selection of data included in it is of great importance for the jump ensemble method, and will be used throughout the arguments of Theorem 4.8. As such, the conditional law $\mathbb{P}(\,\cdot \mid \mathcal{F})$ will be used extensively, and so we use the notation

$$\mathbb{P}_{\mathcal{F}}(\,\cdot\,) := \mathbb{P}(\,\cdot \mid \mathcal{F})$$

to denote it. In this notation, our aim is to understand the law of the top k curves of $\mathcal{L}_n$ on $[-2T, 2T]$ under $\mathbb{P}_{\mathcal{F}}$.

Why does conditioning on $\mathcal{F}$ help? Our reasoning was that lower curve avoidance on $[\ell, r]$ without a run-up is difficult, and so we need to give a run-up. This was done by expanding the interval to $[-2T, 2T]$. However, we then need to enforce lower curve avoidance on all of $[-2T, 2T]$, which is more difficult. But by including the side bridges of the top k curves of $\mathcal{L}_n$, we can use that data to help make the non-intersection easier on the side intervals. This is because the geometries of the top k side bridges of $\mathcal{L}_n$ are *already* well suited for lower curve avoidance with $\mathcal{L}_n(k+1, \cdot\,)$, and the candidate ensemble can piggyback on this success. Thus, we get almost the best of both worlds: the lower curve avoidance is made easier in the middle interval of $[\ell, r]$ due to the space for a run-up, while at the same time the lower curve avoidance on the side intervals is manageable using the data of the top k side bridges of $\mathcal{L}_n$.

How do we make use of this data? We will combine the candidate ensemble on $[\ell, r]$ with the data from $\mathcal{F}$ to create a new ensemble on $[-2T, 2T]$. The form of this combination is dictated by the Brownian Gibbs property and the linear operation involved in the definition of $f^{[a,b]}$.

Let the candidate ensemble be denoted by $X : [\![k]\!] \times [\ell, r] \to \mathbb{R}$; the new ensemble created using X and data from $\mathcal{F}$ will be called the *resampled* ensemble $\mathcal{L}^{\mathrm{re},X} : [\![k]\!] \times [-2T, 2T] \to \mathbb{R}$. Intuitively, the values of the candidate ensemble at ℓ and r are used to affinely shift the side bridges; the affinely shifted bridges define the resampled ensemble on $[-2T, \ell] \cup [r, 2T]$, while the candidate ensemble determines the resampled ensemble on $[\ell, r]$. This is illustrated in Figure 5.2, and the formal definition of $\mathcal{L}^{\mathrm{re},X}$ is given by the following, for $i = 1, \dots, k$:

$$\mathcal{L}^{\mathrm{re},X}(i,x) = \begin{cases} \mathcal{L}_n^{[-2T,\ell]}(i,x) + \frac{x+2T}{\ell+2T} \cdot X(i,\ell) + \frac{\ell-x}{\ell+2T} \cdot \mathcal{L}_n(i,-2T) & x \in [-2T, \ell] \\ \mathcal{L}_n^{[r,2T]}(i,x) + \frac{2T-x}{2T-r} \cdot X(i,r) + \frac{x-r}{2T-r} \cdot \mathcal{L}_n(i,2T) & x \in [r, 2T] \\ X(i,x) & x \in [\ell, r]. \end{cases} \tag{5.1}$$

(Implicit in the above discussion is the promise that the resampled ensemble $\mathcal{L}^{\mathrm{re},X}$ will be able to pass the non-intersection test on the side intervals sufficiently well due to the use of data in $\mathcal{F}$. We discuss and make good on this promise in Section 5.1.)

As with the earlier discussion of the Brownian bridge ensemble and $\mathcal{F}_{\mathrm{BB}}$, the Brownian Gibbs property says that for certain candidate ensembles X, the distribution of $\mathcal{L}^{\mathrm{re},X}$, conditioned on passing the non-intersection tests on $[-2T, 2T]$, will be the $\mathcal{F}$-conditional law of $\mathcal{L}_n$ on $[\![k]\!] \times [-2T, 2T]$. The candidate ensembles X for which this is true are Brownian bridge ensembles conditioned on avoiding the lower curve on some subset of $[\ell, r]$; the jump ensemble,

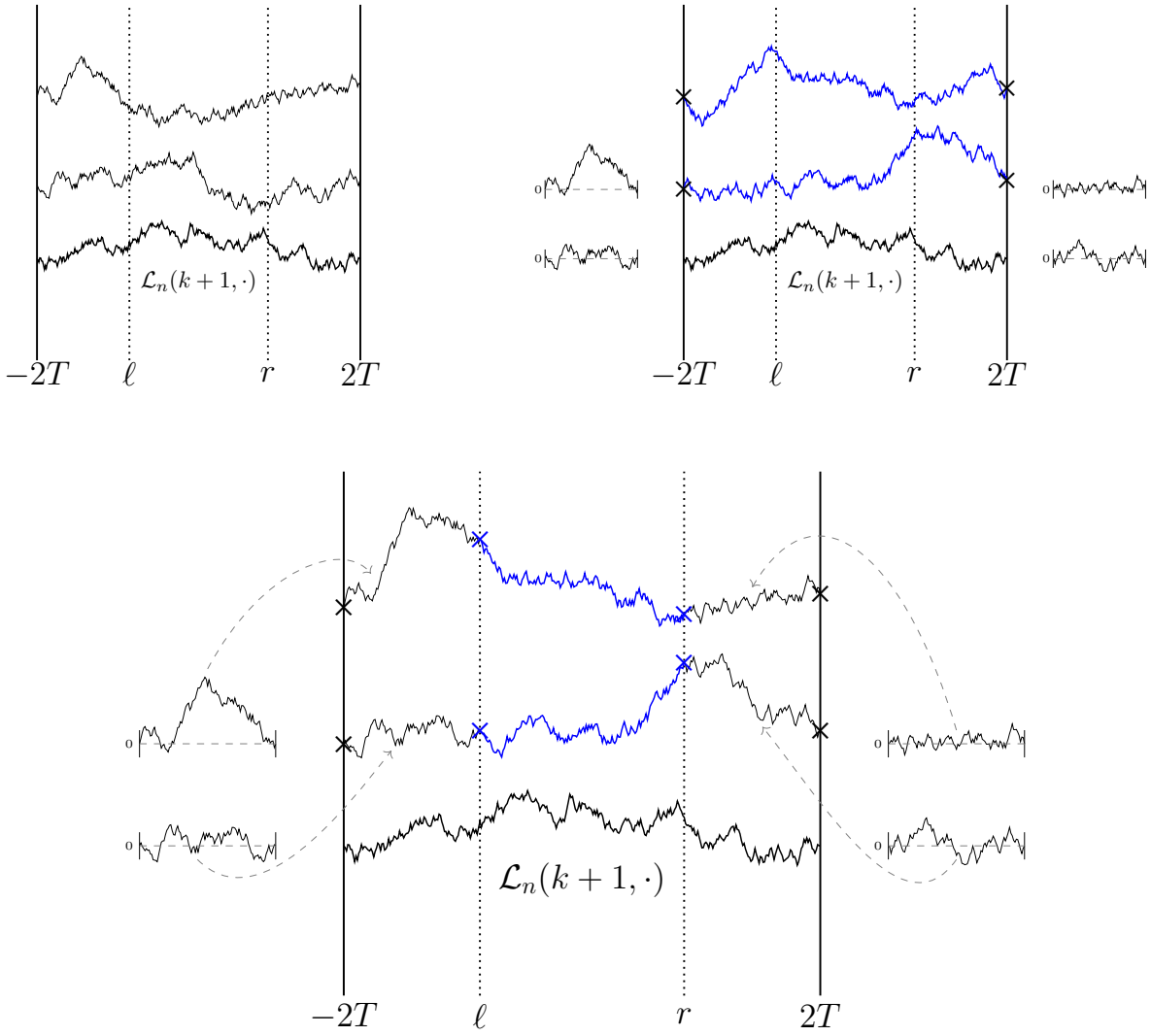


Figure 5.2: Constructing $\mathcal{L}^{\mathrm{re},X}$ from the candidate process X (top two curves in blue on $[-2T, 2T]$ in the second figure, and on $[\ell, r]$ in the third) when $k = 2$. In the first figure we have the original line ensemble $\mathcal{L}_n$. In the second figure, the black elements are the data available in $\mathcal{F}$, namely the entirety of the $(k+1)^{\mathrm{st}}$ curve, the positions at $\pm 2T$ of the first k curves (denoted by crosses), and the bridges obtained from the side intervals by affine shift. The blue curves comprise the candidate process X (though technically X is restricted to $[\ell, r]$). In the final figure we complete the reconstruction by pasting the side interval bridges according to the positions dictated by X (blue crosses at ℓ and r) on $[\ell, r]$. Note that in this figure, $\mathcal{L}^{\mathrm{re},X}$ passes both the side interval tests and the middle interval test.

which will be conditioned on avoiding a coarsened version of the lower curve, will fit this description. Section 5.1 is devoted to setting up a precise version of this statement, which is recorded in Lemma 5.5.

In summary, we are looking to define a candidate process which has estimable probabilities by virtue of being in some sense Brownian, and which has a not-too-low probability of passing the non-intersection test. To accomplish this, we saw in this subsection and the previous that the candidate process will make use of a coarsened version of the lower curve profile in its definition; and will use extra space for a run-up, for which we work with a more sophisticated selection of data captured by $\mathcal{F}$. This data will be combined with the candidate ensemble to give the resampled ensemble.

In the next subsection we expand on the idea that including the data of the side bridges in $\mathcal{F}$ makes it easy for the candidate ensemble to pass the non-intersection test on the side intervals.

The side intervals test

We formulate the non-intersection test on the side intervals as *the side intervals test*. The side intervals test has two parts: that $\mathcal{L}^{\mathrm{re},X}(k,\cdot)$ does not intersect $\mathcal{L}_n(k+1,\cdot)$; and that $\mathcal{L}^{\mathrm{re},X}(i,\cdot)$ does not intersect $\mathcal{L}^{\mathrm{re},X}(i+1,\cdot)$ for $i = 1,\dots,k-1$—both of these on $[-2T,\ell] \cup [r,2T]$. A look at the first two cases in the definition (5.1) of $\mathcal{L}^{\mathrm{re},X}$ suggests that whether this test is passed is simply a question of whether $X(i,\ell)$ and $X(i,r)$ are high enough in value, as the remaining quantities are $\mathcal{F}$-measurable and thus not affected by the candidate X. This intuition is roughly correct, and a precise version is the content of the next lemma, which we refer to as *the side intervals test criterion*.

Lemma 5.3 (Side intervals test criterion, Lemma 3.8 of [Ham19a])**.** *There exist $\mathcal{F}$-measurable random vectors $\overline{\mathrm{Corner}}^{\ell,\mathcal{F}}, \overline{\mathrm{Corner}}^{r,\mathcal{F}} \in \mathbb{R}^k_{\geq}$ such that $\mathcal{L}^{\mathrm{re},X}$ passes the side intervals tests if and only if $\bar{X}(x) - \overline{\mathrm{Corner}}^{x,\mathcal{F}} \in (0,\infty)^k_{>}$ for $x = \ell, r$.*

The proof is given in [Ham19a], but we include here Figure 5.3 which captures the essential argument.

The conclusion we draw from Lemma 5.3 is that analysing the passing of the side interval tests by the candidate X is very simple in practice: we merely need to consider the event that $\bar{X}(\ell) - \overline{\mathrm{Corner}}^{\ell,\mathcal{F}} \in (0,\infty)^k_{>}$ and $\bar{X}(r) - \overline{\mathrm{Corner}}^{r,\mathcal{F}} \in (0,\infty)^k_{>}$.

Next we give a precise description of candidate ensembles X which are such that $\mathcal{L}^{\mathrm{re},X}$, conditioned on the non-intersection tests, has the $\mathcal{F}$-conditional distribution of the top k curves of $\mathcal{L}_n$.

Applying Brownian Gibbs to $\mathcal{L}^{\mathrm{re},X}$

The resampled ensemble $\mathcal{L}^{\mathrm{re},X}$ passing the non-intersection tests on $[-2T,2T]$ requires

- $\mathcal{L}^{\mathrm{re},X}(i,x) > \mathcal{L}^{\mathrm{re},X}(i+1,x)$ for $i = 1,\dots,k-1$ and $x \in [-2T,2T]$; and
- $\mathcal{L}^{\mathrm{re},X}(k,x) > \mathcal{L}_n(k+1,x)$ for $x \in [-2T,2T]$.

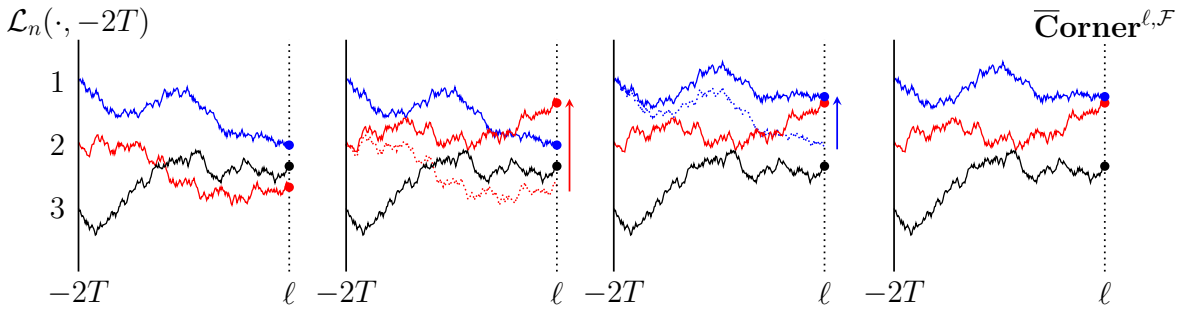


Figure 5.3: Building $\overline{\mathrm{C}}\mathrm{orner}^{\ell,\mathcal{F}}$ for $k = 3$. $\overline{\mathrm{C}}\mathrm{orner}^{\ell,\mathcal{F}}$ is a vector, the $(i-1)^{\text{st}}$ entry of which is the unique smallest value which $\mathcal{L}_n(i-1,\ell)$ can adopt via affine translation before crossing $\mathcal{L}_n(i,\cdot)$ on $[-2T,\ell]$. The line $\mathcal{L}(2,\cdot)$ (red) intersects $\mathcal{L}_n(3,\cdot)$ (black) on $[-2T,\ell]$ (first panel), so $\mathcal{L}_n(2,\cdot)$ (dotted red) is affinely translated until just touching, but not crossing $\mathcal{L}_n(3,\cdot)$ (second panel); the translated value of $\mathcal{L}_n(2,\ell)$ is $\mathrm{Corner}_2^{\ell,\mathcal{F}}$. This process is then repeated for the line $\mathcal{L}_n(1,\cdot)$ (blue). Then $\mathcal{L}_n(1,\cdot)$ (dotted blue) intersects the translated curve $\mathcal{L}_n(2,\cdot)$ (red), so $\mathcal{L}_n(1,\cdot)$ is affinely translated until just touching, but not crossing $\mathcal{L}_n(2,\cdot)$ (third panel); the translated value of $\mathcal{L}_n(1,\ell)$ is $\mathrm{Corner}_1^{\ell,\mathcal{F}}$. The result of these translations is shown in the fourth panel. This procedure depends on the collection of bridges on $[-2T,\ell]$, justifying the dependence of $\overline{\mathrm{C}}\mathrm{orner}^{\ell,\mathcal{F}}$ on $\mathcal{F}$.

We denote by $\mathsf{Pass}(X)$ the indicator for the event described by these two bullet points. (In [Ham19a], an analogous indicator obtained by restricting these two bullet points to $x \in [\ell, r]$ is denoted $T_3(X)$, where 3 represents the test of non-intersection on the middle interval being the third in a sequence of tests.)

Now we may describe a class of candidate ensembles which, conditioned on $\{\mathsf{Pass}(X) = 1\}$, have the desired $\mathcal{F}$-conditional distribution. Let $A \subseteq [\ell, r]$ be an $\mathcal{F}$-measurable random closed set. Define the candidate ensemble $X : [\![k]\!] \times [\ell, r] \to \mathbb{R}$ as a collection of k independent Brownian bridges, the i^{th} one from $(-2T, \mathcal{L}_n(i,-2T))$ to $(2T, \mathcal{L}_n(i,2T))$, conditioned on $\mathsf{NonInt}^A_{\mathcal{L}_n(k+1,\cdot)}$. (Recall that NonInt^A_f is the event that the bottom curve $X(k,x)$ is larger than $f(x)$ for all $x \in A$.) Also define $X' : [\![k]\!] \times [\ell, r] \to \mathbb{R}$ in the same way, with the additional conditioning that $\bar{X}'(x) - \overline{\mathrm{C}}\mathrm{orner}^{x,\mathcal{F}} \in (0,\infty)^k_>$ for $x \in \{\ell, r\}$; we introduce this variant candidate ensemble as it is the form that the jump ensemble will take.

Both X and X' have the desired $\mathcal{F}$-conditional distribution on passing the non-intersection tests. To prove this, we first need a fact about decompositions of Brownian bridges. This property of Brownian bridges also explains why there is no difficulty in conditioning on the potentially complicated objects, the side bridges of $\mathcal{L}_n$, as mentioned in Remark 5.2. The proof of this fact is a straightforward checking of covariances and is omitted.

Lemma 5.4. *Let $T > 0$ and $x_1, \ldots, x_m \in [-2T, 2T]$ with $x_1 < \ldots < x_m$, for some $m \in \mathbb{N}$. Let $x_0 = -2T$ and $x_{m+1} = 2T$. Let B be a Brownian bridge (with arbitrary fixed starting and ending point values) on $[-2T, 2T]$. Then, conditionally on $\big(B(x_1), \ldots, B(x_m)\big)$, the distribution of*

$\big(B^{[x_i,x_{i+1}]}\big)_{i=0}^{m}$ *is that of* $m+1$ *independent Brownian bridges, with the* i^{th} *one of duration* $x_i - x_{i-1}$.

Lemma 5.5. *For* X *as defined above, conditionally on* $\mathcal{F}$*, the following two laws on* $\mathcal{C}_{*,*}([-2T,2T],\mathbb{R})^k$ *are equal:*

$$\mathbb{P}_{\mathcal{F}}\big(\mathcal{L}^{\mathrm{re},X} \in \cdot \mid \mathsf{Pass}(X) = 1\big) \quad \textit{and} \quad \mathbb{P}_{\mathcal{F}}\big(\mathcal{L}_n \in \cdot\big).$$

The same holds with X' *in place of* X.

Proof. Let $B : [\![k]\!] \times [\ell, r] \to \mathbb{R}$ be the restriction to $[\ell, r]$ of a collection of k independent Brownian bridges on $[-2T, 2T]$, with the i^{th} having starting and ending points $(-2T, \mathcal{L}_n(i, -2T))$ and $(2T, \mathcal{L}_n(i, 2T))$. Lemma 5.4, combined with the Brownian Gibbs property possessed by $\mathcal{L}_n$, implies that the $\mathcal{F}$-conditional distribution of $\mathcal{L}_n$ restricted to the top k curves on $[\ell, r]$ is that of $\mathcal{L}^{\mathrm{re},B}$ conditioned on the event $\{\mathsf{Pass}(B) = 1\}$. It is immediate that this latter distribution is the same as that of $\mathcal{L}^{\mathrm{re},X}$ conditioned on $\{\mathsf{Pass}(X) = 1\}$, as the law of X is just that of B with an additional conditioning that is consistent with $\{\mathsf{Pass}(B) = 1\}$. That is, $\{\mathsf{Pass}(B) = 1\}$ is a subset of $\mathsf{NonInt}^A_{\mathcal{L}_n(k+1,\cdot)}$, and the distribution of B conditioned on $\mathsf{NonInt}^A_{\mathcal{L}_n(k+1,\cdot)}$ is that of X. Since the law of X' is that of X conditioned on passing the side interval tests, and since conditioning on $\{\mathsf{Pass}(X) = 1\}$ is a stronger one than conditioning on X passing the side intervals test, i.e., the former event is contained in the latter, the argument of the previous paragraph holds for the candidate ensemble X' as well. □

Having completed the general set up and groundwork of candidate ensembles, we may now turn to describing the jump ensemble itself.

Parameters of the jump ensemble

We start with two parameters, $k \in \mathbb{N}$ and $\varepsilon > 0$. The first is simply the number of curves of $\mathcal{L}_n$ that we are studying, which will also be the number of curves in the jump ensemble. The second is to be understood as the Brownian probability of the event that we wish to analyse under the law of $\mathcal{L}_n$, but is formally simply a positive parameter. The logic of the jump ensemble is to set the parameters according to the event we wish to study.

Though in the discussion in the preceding subsections we were working with a deterministic interval $[\ell, r]$, for the jump ensemble we will in fact need to work on a particular random subinterval $[\mathfrak{l}, \mathfrak{r}]$ of $[-2T, 2T]$ which will be defined shortly. All the arguments and statements of the previous subsections of Section 5.1 will hold true with $\mathfrak{l}$ and $\mathfrak{r}$ in place of ℓ and r, as can be easily checked, since $\mathfrak{l}$ and $\mathfrak{r}$ will be defined in terms of only the lower curve $\mathcal{L}_n(k+1, \cdot)$; this data is present in $\mathcal{F}$, and so, conditional on $\mathcal{F}$, $\mathfrak{l}$ and $\mathfrak{r}$ can safely be thought of as being deterministic.

In Section 5.1 we introduced a parameter T. For the jump ensemble, the value of T is determined by both parameters k and ε, and is given by

$$T := D_k(\log \varepsilon^{-1})^{1/3}, \tag{5.2}$$

where D_k is given by (4.3). The value of T given in (5.2) will be its fixed value for the remainder of Part II.

Remark 5.6. Though not needed for our arguments, here is a heuristic idea of why this is the form of T we select; a fuller discussion is available in the beginning of [Ham19a, Chapter 4]. We see that a larger value of T gives more space for a run-up, which helps the candidate pass the non-intersection test on the middle interval. However, as T gets larger we must also grapple with the globally parabolic curvature of $\mathcal{L}_n$, which means that the starting and ending points of the candidate ensemble will fall quadratically. Simply put, there is more space for the run-up before the jump, but the required jump is higher as the starting point is lower. This selection of T—in particular the 1/3 exponent of $\log \varepsilon^{-1}$—balances these opposing forces and in some sense maximises the non-intersection test probability of the to-be-defined jump ensemble.

Remark 5.7. This same reasoning of balancing these opposing forces of curvature and run-up advantage is what shows that a run-up is not sufficient to handle the second instance of lower curve data discussed in Subsection 5.1, as explained in the beginning of [Ham19a, Chapter 4]. Indeed, what that discussion essentially shows is that even with the well-tuned choice of T made above, there exists data in $\mathcal{F}$ with sufficiently high probability for which the Brownian bridge candidate ensemble (which has forgotten all data about the lower curve) is unable to pass the non-intersection test with sufficiently high probability.

We now record a certain upper and lower bound that the parameter ε is required to meet for technical reasons; these constraints also previously appeared in the statement of Theorem 4.8.

$$\begin{aligned} &\varepsilon < e^{-1} \wedge (17)^{-1/k} C_k^{-1/k} D_k^{-1} \wedge \exp\Big\{ -(24)^6 d^6 / D_k^3 \Big\} \quad \text{and} \\ &\varepsilon > \exp\Big\{ -\big(c/2 \wedge 2^{1/2}\big) D_k^{-1} n^{1/12} \Big\}. \end{aligned} \tag{5.3}$$

As we noted in Remark 4.9, both these bounds do not cause any difficulties in practice. And, in the case that $n = \infty$, the lower bound becomes simply $\varepsilon > 0$.

With these definitions, we may start making precise the notion introduced earlier of a coarsened version of the underlying curve.

Coarsening the lower curve

Let $\mathfrak{c}_+ : [-T, T] \to \mathbb{R}$ be the least concave majorant of $\mathcal{L}_n(k+1, \cdot) : [-T, T] \to \mathbb{R}$, and define a random interval $[\mathfrak{l}, \mathfrak{r}]$ by

$$\begin{aligned} \mathfrak{l} &= \inf\big\{ x \in [-T, T] \mid \mathfrak{c}_+'(x) \le 4T \big\} \quad \text{and} \\ \mathfrak{r} &= \sup\big\{ x \in [-T, T] \mid \mathfrak{c}_+'(x) \ge -4T \big\}. \end{aligned}$$

We can think of $\mathfrak{c}_+$ as a first coarsening of the lower curve $\mathcal{L}_n(k+1, \cdot)$. As indicated earlier, the interval $[\mathfrak{l}, \mathfrak{r}]$ will play the role of $[\ell, r]$ in Section 5.1. Note that though random, $\mathfrak{l}$ and $\mathfrak{r}$ are

functions of the curve $\mathcal{L}_n(k+1,\cdot)$ and not of the k curves we are attempting to study. The purpose of defining this random interval is that on it we are guaranteed some control over the coarsened lower curve profile (which we will be further coarsening before using in the definition of the jump ensemble); we will not use any data about $\mathcal{L}_n(k+1,\cdot)$ outside of $[\mathfrak{l},\mathfrak{r}]$ in defining the jump ensemble, though it is available.

By the concavity of $\mathfrak{c}_+$, it follows that $\mathfrak{l} \le \mathfrak{r}$, and in fact on a high probability favourable event we will discuss in Section 5.1, $[\mathfrak{l},\mathfrak{r}]$ will be an interval of length at least T. We use the σ-algebra $\mathcal{F}$ defined in Section 5.1 on page 62, except with $\mathfrak{l}$ and $\mathfrak{r}$ in the place of ℓ and r. Though $\mathfrak{l}$ and $\mathfrak{r}$ are random, the definition of $\mathcal{F}$ is adequate as $\mathfrak{l}$ and $\mathfrak{r}$ are determined by the lower curve data, which is already present in its entirety in $\mathcal{F}$.

We came to the conclusion in the heuristic discussion in Section 5.1 that the candidate process X would need to use information about a coarsened version of the lower curve $\mathcal{L}_n(k+1,\cdot)$ in its definition, in order to have a high enough probability of passing the middle interval non-intersection test. So far we have defined a preliminary coarsening, the least concave majorant $\mathfrak{c}_+$, which was used to define the interval $[\mathfrak{l},\mathfrak{r}]$. Now we define a further (and final) coarsening, which we will then use in the next subsection to finally define the jump ensemble J.

To precisely describe the final coarsening of $\mathcal{L}_n(k+1,\cdot)$ on $[\mathfrak{l},\mathfrak{r}]$, we first define a subset of extreme points of $\mathfrak{c}_+$. Let $\mathrm{xExt}(\mathfrak{c}_+)$ be the x-coordinates of the set of extreme (or corner) points of the convex set $\{(x,y) : \mathfrak{l} \le x \le \mathfrak{r}, y \le \mathfrak{c}_+(x)\}$. Note that necessarily $\mathfrak{l},\mathfrak{r} \in \mathrm{xExt}(\mathfrak{c}_+)$. Then define the *pole set* P to be a subset of $\mathrm{xExt}(\mathfrak{c}_+)$ such that

- $\mathfrak{l},\mathfrak{r} \in P$,
- $p_1,p_2 \in P, p_1 \ne p_2 \implies |p_1-p_2| \ge d_{\mathrm{ip}}$, and
- if $x \in \mathrm{xExt}(\mathfrak{c}_+)$, then some element $p \in P$ satisfies $|p-x| \le d_{\mathrm{ip}}$.

Here $d_{\mathrm{ip}} \in [1, \mathfrak{r}-\mathfrak{l}]$ is a parameter called the *inter-pole distance*; typically it is set to a unit order quantity independent of ε and k, and usually it is comparable to the interval of interest under study. For example, in the proof of our results, we will set it to be a multiple of d. The parameter d_{ip} defines the minimum separation between consecutive elements of P. These elements of the pole set P will be called *poles*. These above three properties do not necessarily define P uniquely, and to address this we take P to be the subset satisfying these conditions of maximal cardinality, and then maximal in lexicographic order.

Remark 5.8. By the definition of d_{ip}, it is clear that the size of the pole set, $|P|$, is at most $2T/d_{\mathrm{ip}}$. So for constant order values of d_{ip}, $|P|$ is potentially rather large. However, in arguments the only poles which must be considered are essentially the ones within or adjacent to the interval under study. In the proof of our main result our arguments will be focused on the single pole contained in the interval $[-2d, 2d]$ (if it is present), and will in one instance make use of the preceding and succeeding poles. We guarantee ourselves this control on the number of poles in $[-d,d]$ by making an appropriate choice of d_{ip}.

The coarsened profile of $\mathcal{L}_n(k+1,\cdot)$ that will be used in defining the jump ensemble J is exactly

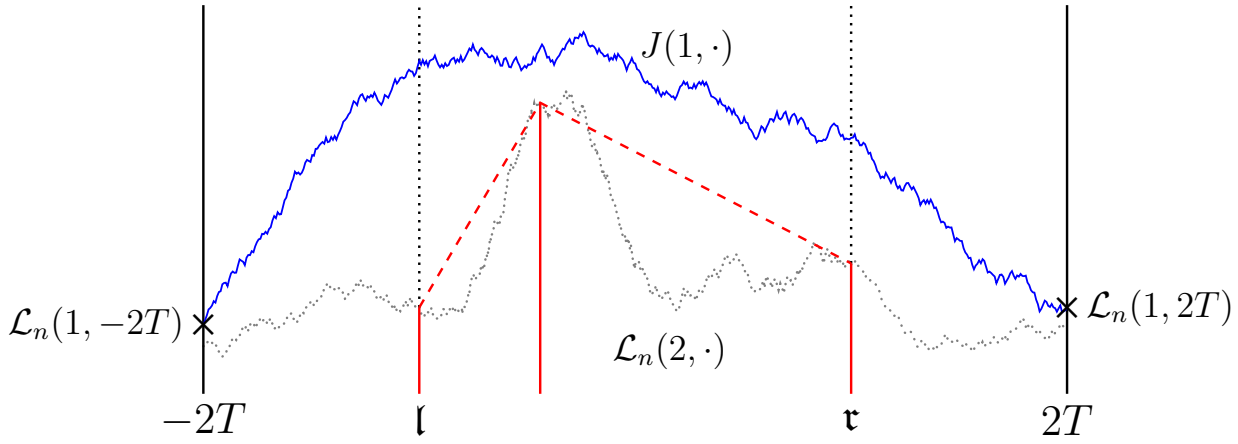


Figure 5.4: The jump ensemble candidate J (in blue) when $k = 1$. The lines in red are the *poles*, i.e., the elements of the pole set P, which J is conditioned to jump over (recall that necessarily $\mathfrak{l}, \mathfrak{r} \in P$). The dashed red piecewise linear function is the Tent map. The lower curve is drawn dotted and in light gray to indicate that J does not have access to this data, only the heights of the solid poles. Making this jump forces the candidate to avoid a coarsened version of the lower curve profile, which makes it more likely that it also avoids the full lower curve, as in this figure. Moving back to $[-2T, 2T]$ from $[\mathfrak{l}, \mathfrak{r}]$ gives the candidate space to make the jump. Though we have shown the blue curve on the entirety of $[-2T, 2T]$, J is only the restriction to $[\mathfrak{l}, \mathfrak{r}]$, and so does *not* need to avoid the lower curve outside $[\mathfrak{l}, \mathfrak{r}]$ (though it does in the illustrated instance). The avoidance on the full interval $[-2T, 2T]$ is a requirement imposed only on $\mathcal{L}^{\mathrm{re},J}$, which is made by combining J with the data in $\mathcal{F}$ as in Figure 5.2.

the set $\{(p, \mathcal{L}_n(k+1, p) \mid p \in P)\}$, which J will be conditioned to *jump* over. We next make precise what we mean by J jumping over the poles, and also give the definition of J.

Defining the jump ensemble

Conditional on $\mathcal{F}$, let $B : [\![k]\!] \times [-2T, 2T] \to \mathbb{R}$, with $\{B(i,\cdot)\}_{i=1}^k$ a collection of k independent Brownian bridges on $[-2T, 2T]$ and $B(i,\cdot)$ having endpoints $(-2T, \mathcal{L}_n(i, -2T))$ and $(2T, \mathcal{L}_n(i, 2T))$ for $i = 1, \ldots, k$. Note that the required information about the endpoint values of $\mathcal{L}_n$ is present in $\mathcal{F}$. The jump ensemble $J : [\![k]\!] \times [\mathfrak{l}, \mathfrak{r}] \to \mathbb{R}$ is the restriction to $[\mathfrak{l}, \mathfrak{r}]$ of B conditioned on

(i) $\overline{B}(x) - \overline{\mathrm{Corner}}^{x,\mathcal{F}} \in (0,\infty)^k_>$ for $x \in \{\mathfrak{l}, \mathfrak{r}\}$; and
(ii) $B(i,p) \geq \mathcal{L}_n(k+1, p)$ for all $p \in P$ and $i = 1, \ldots, k$.

As we saw in Lemma 5.3, the conditioning present in point (i) ensures that J passes the side interval tests. We will refer to the event in point (ii), namely $J(p) \geq \mathcal{L}_n(k+1, p)$ for $p \in P$, as *jumping over the pole* p.

As we noted when we stated its definition, the pole set P represents the coarsened version of the lower curve $\mathcal{L}_n(k+1,\cdot)$ that J has access to. By conditioning J to avoid this coarsened version of $\mathcal{L}_n(k+1,\cdot)$, we increase the probability of J successfully avoiding all of $\mathcal{L}_n(k+1,\cdot)$ on $[\mathfrak{l},\mathfrak{r}]$, compared (heuristically) to a candidate process with no information about the underlying curve. Indeed we will see ahead in Proposition 5.9 that this increased probability is high enough to be useful for our intended application.

It will be necessary in our arguments to consider how much J deviates from the shape defined by the poles. To do this, define the $\mathcal{F}$-measurable random piecewise affine function $\text{Tent} : [\mathfrak{l},\mathfrak{r}] \to \mathbb{R}$ which linearly interpolates between the points $(p, \mathcal{L}_n(k+1,p))$ for $p \in P$. Note that from the definition of $\mathfrak{c}_+$ we have that Tent is concave, and from the definition of $\mathfrak{l}$ and $\mathfrak{r}$ we have that the slope of every linear segment of Tent lies in $[-4T, 4T]$, which for future reference we will express (with abuse of notation, as Tent is only piecewise linear and not linear) as

$$\text{slope}(\text{Tent}) \in [-4T, 4T]. \tag{5.4}$$

See Figure 5.4 for an illustration of the jump ensemble and the Tent map.

We also note here that Lemma 5.5 implies that $\mathcal{L}^{\text{re},J}$, conditionally on $\mathcal{F}$ and $\{\mathsf{Pass}(J) = 1\}$, has the $\mathcal{F}$-conditional distribution of the top k curves of $\mathcal{L}_n$ on $[-2T, 2T]$. Since $\mathcal{L}^{\text{re},J} = J$ on $[\![k]\!] \times [\mathfrak{l},\mathfrak{r}]$, this implies that the distribution of J, conditionally on $\mathcal{F}$ and $\{\mathsf{Pass}(J) = 1\}$, is that of the top k curves of $\mathcal{L}_n$ on $[\mathfrak{l},\mathfrak{r}]$.

The probability that J passes the non-intersection test

Now we shall address whether the jump ensemble is in fact able to pass the non-intersection test on the whole interval $[-2T, 2T]$ with sufficiently high probability, the task for which we specifically defined the coarsened version of $\mathcal{L}_n(k+1,\cdot)$ that J is conditioned to jump over. Recall from Section 5.1 that we have an easy criterion for J passing the non-intersection test on the side intervals of $[-2T, \mathfrak{l}]$ and $[\mathfrak{r}, 2T]$, which we called the side-intervals test. This criterion is that $(J(i,x) - \text{Corner}_i^{x,\mathcal{F}})_{i=1}^k \in (0,\infty)^k_>$ for $x \in \{\mathfrak{l},\mathfrak{r}\}$, which J is in fact conditioned to satisfy above in point (i) of its definition. Thus all that remains is for J to pass the non-intersection test on the middle interval $[\mathfrak{l},\mathfrak{r}]$, i.e., for J to satisfy

$$J(1,x) > J(2,x) > \ldots > J(k,x) > \mathcal{L}_n(k+1,x) \quad \forall x \in [\mathfrak{l},\mathfrak{r}].$$

In other words, the indicator of this last event is the same as the indicator $\mathsf{Pass}(J)$. It is important for our approach that the event $\{\mathsf{Pass}(J) = 1\}$ being conditioned on does not have too low a probability. Unlike the side intervals test, there is no simple criterion for the middle interval test. In fact, an analysis was undertaken in [Ham19a] to obtain an appropriately strong lower bound on this probability, which holds on a high probability $\mathcal{F}$-measurable favourable event Fav (that we will define shortly). As that argument does not serve out expository purpose, we do not present it here; instead we reproduce the statement from [Ham19a] in the next Proposition 5.9. This is the statement we previously referenced as *jump ensemble candidate proficiency*.

Proposition 5.9 (Jump ensemble candidate proficiency, Proposition 4.2 of [Ham19a]). *We have that*

$$\mathbb{P}_{\mathcal{F}}\Big(\mathsf{Pass}(J) = 1\Big) \geq \exp\left\{-3973k^{7/2}d_{\mathrm{ip}}^2 D_k^2 \left(\log \varepsilon^{-1}\right)^{2/3}\right\} \cdot \mathbb{1}_{\mathrm{Fav}}.$$

We will now define Fav, before returning to discuss the important role of Proposition 5.9 in our approach.

The definition of Fav & the role of Proposition 5.9

The favourable event Fav is defined as the intersection

$$\mathsf{Fav} = \mathsf{F}_1 \cap \mathsf{F}_2 \cap \mathsf{F}_3,$$

where

$$\begin{aligned}
\mathsf{F}_1 &= \left\{\mathcal{L}_n(i,x) \in T^2[-2\sqrt{2}-1, -2\sqrt{2}+1] \text{ for } (i,x) \in [\![k]\!] \times \{-2T, 2T\}\right\} \\
\mathsf{F}_2 &= \left\{-T^2 \leq \mathcal{L}_n(k+1,x) \leq T^2 \text{ for } x \in [-T,T]\right\}, \\
\mathsf{F}_3 &= \bigcap_{i \in [\![k]\!]} \left\{\mathrm{Corner}_i^{\mathfrak{l},\mathcal{F}} \in \left[-T^2, T^2\right]\right\} \cap \left\{\mathrm{Corner}_i^{\mathfrak{r},\mathcal{F}} \in \left[-T^2, T^2\right]\right\}.
\end{aligned}$$

Note that Fav is an $\mathcal{F}$-measurable event. As its name suggests, this event fixes good data in $\mathcal{F}$ on which we have strong enough control to make our arguments. The reader should view this data as being fixed in the arguments involving the jump ensemble, as we will be working only on this event; the bound on $\mathbb{P}(\mathsf{Fav}^c)$ just ahead allows us to take this liberty.

The form of the favourable event respects the parabolic curvature possessed by $\mathcal{L}_n$. In particular, since we are working on the interval $[-2T, 2T]$, we expect that at the endpoints the location of $\mathcal{L}_n$ will be $O(-T^2)$, which dictates the form of the three subevents F_1, F_2, and F_3 above.

It is a simple calculation based on the definition of $\mathfrak{l}$ and $\mathfrak{r}$ that, on F_2,

$$\mathfrak{l} \leq -T/2 \quad \text{and} \quad \mathfrak{r} \geq T/2.$$

We need the knowledge that the favourable event occurs with sufficiently high probability; this is provided to us from [Ham19a]:

Lemma 5.10 (High probability of favourable event, Lemma 4.1 of [Ham19a]).

$$\mathbb{P}\left(\mathsf{Fav}^c\right) \leq \varepsilon^{2^{-5}c_k D_k^3}.$$

In fact for our purposes it will be sufficient to note that $2^{-5}c_k D_k^3 \geq 1$ for all k, and so the upper bound above is further bounded by ε.

Now we may discuss the central role of Proposition 5.9 in our argument. We will use it to reduce the problem of understanding the probability of an event under the $\mathcal{F}$-conditional law of $\mathcal{L}_n$ to

understanding the same under the law of J. For concreteness, let us illustrate this by attempting to bound the probability that the vertically shifted k^{th} curve $\mathcal{L}_n(k,\cdot) - \mathcal{L}_n(k,-d)$ lies in a measurable subset $A \subseteq \mathcal{C}_{0,*}([-d,d],\mathbb{R})$ of continuous functions vanishing at $-d$, where $d > 0$. Recall again that the $\mathcal{F}$-conditional distribution of $\mathcal{L}_n$ on $[\mathfrak{l},\mathfrak{r}]$ is the same as the distribution of $\mathcal{L}^{\text{re},J}$ on $[\mathfrak{l},\mathfrak{r}]$ conditioned on the event that $\mathsf{Pass}(J) = 1$, and also that $\mathcal{L}^{\text{re},J}(i,\cdot) = J(i,\cdot)$ on $[\mathfrak{l},\mathfrak{r}]$. We assume that ε is small enough that $[-d,d] \subseteq [-T/2,T/2] \subseteq [\mathfrak{l},\mathfrak{r}]$, the last inclusion on the event Fav. (This assumption on ε is implied by the condition that $\varepsilon < \exp(-(24)^6 d^6/D_k^3)$ that is imposed in Theorem 4.8.) We also have to set the last parameter of the jump ensemble, the inter-pole distance d_{ip}, which we set as

$$d_{\text{ip}} = 5d, \tag{5.5}$$

which will be its value in our application of the jump ensemble. Then we see that

$$\begin{aligned}
&\mathbb{P}\Big(\mathcal{L}_n(k,\cdot) - \mathcal{L}_n(k,-d) \in A\Big) \\
&\qquad = \mathbb{E}\left[\mathbb{P}_{\mathcal{F}}(\mathcal{L}_n(k,\cdot) - \mathcal{L}_n(k,-d) \in A)\cdot \mathbb{1}_{\mathsf{Fav}}\right] + \mathbb{P}(\mathsf{Fav}^c) \\
&\qquad = \mathbb{E}\left[\mathbb{P}_{\mathcal{F}}\big(J(k,\cdot) - J(k,-d) \in A \mid \mathsf{Pass}(J) = 1\big)\cdot \mathbb{1}_{\mathsf{Fav}}\right] + \mathbb{P}(\mathsf{Fav}^c) \\
&\qquad \le \mathbb{E}\left[\frac{\mathbb{P}_{\mathcal{F}}(J(k,\cdot) - J(k,-d) \in A)}{\mathbb{P}_{\mathcal{F}}\left(\mathsf{Pass}(J) = 1\right)}\cdot \mathbb{1}_{\mathsf{Fav}}\right] + \mathbb{P}(\mathsf{Fav}^c) \\
&\qquad \le \mathbb{E}\left[\mathbb{P}_{\mathcal{F}}\big(J(k,\cdot) - J(k,-d) \in A\big)\cdot \mathbb{1}_{\mathsf{Fav}}\right]\cdot \exp\left\{O_k(1)\left(\log \varepsilon^{-1}\right)^{2/3}\right\} \\
&\qquad\qquad + \mathbb{P}(\mathsf{Fav}^c),
\end{aligned} \tag{5.6}$$

using Proposition 5.9 in the last inequality.

Theorem 4.8 asserts a bound of the form $\varepsilon \cdot \exp(O_k(1)(\log \varepsilon^{-1})^{5/6})$ on the left-hand side of the first line of the above display, where ε is the probability of A under the law of Brownian motion. So in order to prove Theorem 4.8, the main step to be effected is to bound the first term after the last inequality by a quantity of the same form. The notational equivalence we have just made between the parameter ε of the jump ensemble and the Brownian motion probability of the event of interest is one we will adopt formally: in the remainder of the proof of Theorem 4.8, the parameter ε of the jump ensemble will have the value

$$\varepsilon := \mathcal{B}_{0,*}^{[-d,d]}(A). \tag{5.7}$$

(As we will record formally soon, the parabolic invariance Lemma 4.7 allows us to reduce Theorem 4.8 to the case where $K = 0$.) With this choice of ε, the importance of Proposition 5.9 in achieving the goal mentioned in the last paragraph is now clear, in particular that the exponent of the $\log \varepsilon^{-1}$ in the exponent of the statement of Proposition 5.9 is $2/3 < 5/6$. Looking back at (5.6), to actually achieve this goal we need two bounds: that $\mathbb{P}_{\mathcal{F}}\big(J(k,\cdot) - J(k,-d) \in A\big)\cdot \mathbb{1}_{\mathsf{Fav}}$ and $\mathbb{P}(\mathsf{Fav}^c)$ are both bounded by $\varepsilon \exp(O_k(1)(\log \varepsilon^{-1})^{5/6})$.

The second bound is implied by Lemma 5.10. Finally, in the following Theorem 5.11, we have the first bound, proving which will be the work of the next chapter. After stating this theorem and

an immediate corollary, we will end this section by giving a brief summary of the jump ensemble. Recall that $\mathcal{B}_{0,*}^{[-d,d]}$ is the law on $\mathcal{C}_{0,*}([-d,d],\mathbb{R})$ of a Brownian motion started at coordinates $(-d,0)$.

Theorem 5.11. *Let $d \geq 1$ and $A \subseteq \mathcal{C}_{0,*}([-d,d],\mathbb{R})$. There exist $\varepsilon_0 = \varepsilon_0(d,k) = \exp(-(24)^6 d^6/D_k^3)$ and absolute constant $G < \infty$ such that, if $\mathcal{B}_{0,*}^{[-d,d]}(A) = \varepsilon < \varepsilon_0$,*

$$\mathbb{P}_{\mathcal{F}}\Big(J(k,\cdot) - J(k,-d) \in A\Big) \cdot \mathbb{1}_{\mathsf{Fav}} \leq \varepsilon \cdot G d^{\frac{1}{2}} \cdot D_k^4 (\log \varepsilon^{-1})^{\frac{4}{3}} \cdot \exp\left(792 \cdot d \cdot D_k^{5/2} \cdot (\log \varepsilon^{-1})^{5/6}\right).$$

Given this theorem we may prove the main Theorem 4.8:

Proof of Theorem 4.8. By applying the parabolic invariance Lemma 4.7 with $y_n = -K$, proving Theorem 4.8 reduces to the case $K = 0$. The condition that $[K-d, K+d] \subset c/2 \cdot [-n^{1/9}, n^{1/9}]$ is exactly the one required to apply Lemma 4.7.

We have that

$$\begin{aligned}
&\mathbb{P}\big(\mathcal{L}_n(k,\cdot) - \mathcal{L}_n(k,-d) \in A\big) \\
&\qquad = \mathbb{E}\Big[\mathbb{P}_{\mathcal{F}}\big(J(k,\cdot) - J(k,-d) \in A \;\big|\; \mathsf{Pass}(J) = 1\big) \cdot \mathbb{1}_{\mathsf{Fav}}\Big] + \mathbb{P}(\mathsf{Fav}^c) \\
&\qquad \leq \mathbb{E}\left[\frac{\mathbb{P}_{\mathcal{F}}\big(J(k,\cdot) - J(k,-d) \in A\big)}{\mathbb{P}_{\mathcal{F}}\left(\mathsf{Pass}(J) = 1\right)} \cdot \mathbb{1}_{\mathsf{Fav}}\right] + \mathbb{P}(\mathsf{Fav}^c).
\end{aligned}$$

By Lemma 5.10, $\mathbb{P}(\mathsf{Fav}^c) < \varepsilon$ for the choice of D_k we have assumed. This bound requires (4.4), which we have also assumed. By Proposition 5.9 and Theorem 5.11, we find that the last expression is bounded up to a constant factor by

$$\begin{aligned}
&\varepsilon \cdot d^{\frac{1}{2}} \cdot D_k^4 (\log \varepsilon^{-1})^{\frac{4}{3}} \cdot \exp\left(792 \cdot d \cdot D_k^{5/2} \cdot (\log \varepsilon^{-1})^{5/6} + 3973 k^{7/2} d_{\mathrm{ip}}^2 D_k^2 \left(\log \varepsilon^{-1}\right)^{2/3}\right) + \varepsilon \\
&\qquad\qquad \leq 2\varepsilon \cdot d^{\frac{1}{2}} \cdot D_k^4 (\log \varepsilon^{-1})^{\frac{4}{3}} \cdot \exp\left(4931 \cdot d \cdot k^{7/2} \cdot D_k^{5/2} \cdot (\log \varepsilon^{-1})^{5/6}\right);
\end{aligned}$$

we have used that $d_{\mathrm{ip}} = 5d$ from (5.5), $d \geq 1$, $D_k^{1/2} (\log \varepsilon^{-1})^{1/6} \geq 24d$ from the assumed upper bound on ε, and $3973 \times 5^2/24 + 792 \leq 4931$. Since $x^{11/5} \leq e^x$ for all $x \geq 1$, we may absorb the factor $d^{\frac{3}{2}} D_k^{\frac{11}{2}} (\log \varepsilon^{-1})^{\frac{11}{6}}$ by increasing the coefficient of the exponent by 1. This proves Theorem 4.8. □

Hence the remaining task is to prove Theorem 5.11. This is accomplished in Chapter 6. Before proceeding to this, we give in the next subsection an important statement about stochastic domination relations between the jump ensemble and certain Brownian bridges, and finally a concluding subsection giving a brief summary of the jump ensemble which may act as a quick reference for the reader.

A stochastic domination property of the jump ensemble

In our proof of Theorem 5.11, at various points we will need to stochastically dominate $J(k,\cdot)$ by or have $J(k,\cdot)$ stochastically dominate certain Brownian bridges. For this we will make use of the following statement, which will be proven essentially using Lemma 5.4 to reduce stochastic domination of processes to obvious stochastic dominations of point-values.

Lemma 5.12. *Fix $n \in \mathbb{N}$ and $k \le n$. Let $\phi_{\text{start}} \ge \mathcal{L}_n(k,-2T)$ and $\phi_{\text{end}} \ge \mathcal{L}_n(k,2T)$, and let $x \in [\mathfrak{l},\mathfrak{r}]$. Then, on* Fav *and conditionally on $\mathcal{F}$,*

(i) The law of $J(k,\cdot)$ (as a law on $[\mathfrak{l},\mathfrak{r}]$) stochastically dominates the law of a Brownian bridge from $(\mathfrak{l},-T^2)$ to $(\mathfrak{r},-T^2)$.

(ii) Conditionally on $J(k,x)$, the law of $J(k,\cdot)$ restricted to $[\mathfrak{l},x]$ stochastically dominates the law of a Brownian bridge from $(\mathfrak{l},-T^2)$ to $(x,J(k,x))$. Under the same conditioning, the law of $J(k,\cdot)$ restricted to $[x,\mathfrak{r}]$ stochastically dominates the law of a Brownian bridge from $(x,J(k,x))$ to $(\mathfrak{r},-T^2)$.

(iii) The law of $J(k,\cdot)$ is stochastically dominated by the law of the restriction to $[\mathfrak{l},\mathfrak{r}]$ of a Brownian bridge from $(-2T,\phi_{\text{start}})$ to $(2T,\phi_{\text{end}})$ which is conditioned to be above all poles, above $(\mathfrak{l},\text{Corner}_k^{\mathfrak{l},\mathcal{F}})$, and above $(\mathfrak{r},\text{Corner}_k^{\mathfrak{r},\mathcal{F}})$.

(iv) Conditionally on $J(k,x)$, the law of $J(k,\cdot)$, restricted to $[x,\mathfrak{r}]$, is stochastically dominated by the law of the restriction to $[x,\mathfrak{r}]$ of a Brownian bridge from $(x,J(k,x))$ to $(2T,\phi_{\text{end}})$ which is conditioned to be above all poles in $[x,\mathfrak{r}]$ and above $(\mathfrak{r},\text{Corner}_k^{\mathfrak{r},\mathcal{F}})$.

Proof. Lemma 5.12(i) follows from (ii) by taking $x=\mathfrak{r}$ and averaging, and noting that, on Fav, $J(k,\mathfrak{r}) \ge \text{Corner}_k^{\mathfrak{r},\mathcal{F}} \ge -T^2$. So we prove (ii); in fact, we prove the first part of (ii) as the second part is analogous.

Let B be a Brownian bridge from $(\mathfrak{l},J(k,\mathfrak{l}))$ to $(x,J(k,x))$. Letting $m=|P|$ and $x_1,\dots,x_m$ be the elements of P, we may apply Lemma 5.4 to decompose B at the elements of P. Since, conditionally on $J(k,\mathfrak{l})$ and $J(k,x)$, the distribution of $J(k,\cdot)$ is that of B conditioned on the values of B at elements of P being sufficiently high, the decomposition provided by Lemma 5.4 yields that the law of $J(k,\cdot)$, conditionally on its values at $\mathfrak{l}$ and x, stochastically dominates that of B. Since, on Fav, $J(k,\mathfrak{l}) \ge \text{Corner}_k^{\mathfrak{l},\mathcal{F}} \ge -T^2$, it is clear that B stochastically dominates the Brownian bridge described in Lemma 5.12(i), yielding the claim.

Now we turn to (iii). Given two intervals I_x for $x \in \{\mathfrak{l},\mathfrak{r}\}$, consider the restriction to $[\mathfrak{l},\mathfrak{r}]$ of a Brownian bridge B with starting point $(-2T,\mathcal{L}_n(k,-2T))$ and ending point $(2T,\mathcal{L}_n(k,2T))$, conditioned on $B(x) \ge \mathcal{L}_n(k+1,x)$ for all $x \in P$, and on $B(x) \in I_x$ for $x \in \{\mathfrak{l},\mathfrak{r}\}$. Call this law $\mathcal{B}(I_\mathfrak{l},I_\mathfrak{r})$.

Conditionally on $J(k-1,\cdot)$ as a process on $[\mathfrak{l},\mathfrak{r}]$, the law of $J(k,\cdot)$ is $\mathcal{B}(I_\mathfrak{l},I_\mathfrak{r})$ with I_x a *finite* interval determined by $J(k-1,\cdot)$, in such a way that $\inf I_x = \text{Corner}_k^{x,\mathcal{F}}$, for $x \in \{\mathfrak{l},\mathfrak{r}\}$. Using the same decomposition from Lemma 5.4, it is clear that $\mathcal{B}(I_\mathfrak{l},I_\mathfrak{r})$ is stochastically dominated by $\mathcal{B}(I'_\mathfrak{l},I'_\mathfrak{r})$, where $I'_x = [\text{Corner}_k^{x,\mathcal{F}},\infty)$ for $x \in \{\mathfrak{l},\mathfrak{r}\}$.

Averaging over $J(k-1,\cdot)$, we find that the law of $J(k,\cdot)$ is stochastically dominated by $\mathcal{B}(I'_{\mathfrak{l}}, I'_{\mathfrak{r}})$. On Fav, the law of the process described in Lemma 5.12(iii) is that of the Brownian bridge B in the definition of $\mathcal{B}(I'_{\mathfrak{l}}, I'_{\mathfrak{r}})$ with the endpoints shifted vertically upwards, which clearly maintains the described stochastic domination.

The proof of (iv) is along the same lines as (iii). □

Summary of the jump ensemble

The definition of the jump ensemble was rather involved, and here we provide a quick summary of the main aspects of its definition which the reader should keep in mind in order to understand the arguments leading to the proof of Theorem 5.11.

The jump ensemble has three parameters, $\varepsilon > 0$, $k \in \mathbb{N}$, and $d_{\mathrm{ip}} > 0$; and is defined on an interval $[-2T, 2T]$, where $T = D_k(\log \varepsilon^{-1})^{1/3}$, with D_k given by (4.3). The interval $[-2T, 2T]$ contains a random subinterval $[\mathfrak{l}, \mathfrak{r}]$. The jump ensemble J is the restriction to $[\mathfrak{l}, \mathfrak{r}]$ of a collection of k independent Brownian bridges, the i^{th} from $(-2T, \mathcal{L}_n(i, -2T))$ to $(2T, \mathcal{L}_n(i, 2T))$, conditioned on $\overline{J}(x) - \overline{\mathrm{C}}\mathrm{orner}^{\mathfrak{l},\mathcal{F}} \in (0,\infty)^k_>$ for $x \in \{\mathfrak{l}, \mathfrak{r}\}$ and on $J(p) > \mathcal{L}_n(k+1, p)$ for all p in the pole set P. The elements of the pole set P have a minimum separation of d_{ip}, and are a subset of $[\mathfrak{l}, \mathfrak{r}]$. This interval is defined in terms of the underlying curve $\mathcal{L}_n(k+1, \cdot)$, and the relevant consequence of its definition is that the slope (of each linear segment) of the Tent map (which linearly interpolates the points $(p, \mathcal{L}_n(k+1, p))$) lies in $[-4T, 4T]$.

In our application, ε is set according to the Brownian motion probability of the event under consideration, and $d_{\mathrm{ip}} = 5d$.

We gain control over several of the random objects present in the above paragraph on a high probability favourable event Fav. On this event, we have that $[-T/2, T/2] \subseteq [\mathfrak{l}, \mathfrak{r}]$; that $\overline{\mathrm{C}}\mathrm{orner}^{\mathfrak{l},\mathcal{F}}$, $\overline{\mathrm{C}}\mathrm{orner}^{\mathfrak{r},\mathcal{F}} \in [-T^2, T^2]^k$; that $\mathcal{L}_n(k+1, x) \in [-T^2, T^2]$ for $x \in [-T, T]$; and that $\mathcal{L}_n(i, x) \in T^2[-2\sqrt{2}-1, -2\sqrt{2}+1]$ for $(i, x) \in [\![k]\!] \times \{-2T, 2T\}$.

In the next section we discuss the conceptual framework underlying the proof of Theorem 5.11.

5.2 A conceptual framework in terms of costs

Notation. We will use the notation $A \lesssim B$ to indicate that there exists a constant $G < \infty$ which is independent of ε, k, and d such that $A \leq GB$. The value of G may however vary from line to line.

We now begin discussing the ideas underlying the approach of the proof of Theorem 5.11. We fix the curve index k that we are studying, and for the proof of Theorem 5.11 in this section and in Chapter 6 adopt the abuse of notation

$$J(\,\cdot\,) = J(k, \cdot\,),$$

which we will refer to as the *jump curve*; none of our arguments will involve the other curves of the full jump ensemble.

The overarching conceptual framework is one of *costs* that a certain joint density of the jump curve J must satisfy. This section will set up the quantities the proof will be working with, introduce the costs these quantities must interact with, and conclude by showing that the proof of Theorem 5.11 can be reduced to showing that these costs are met.

We would like to study the probability that the jump curve J lies in some $A \subseteq \mathcal{C}_{0,*}([-d, d], \mathbb{R})$, where the probability of A under the law of Brownian motion is ε. As mentioned at the end of the last section, the parameters of the jump ensemble are ε and k, with $d_{\text{ip}} = 5d$. Our aim is to get a bound of ε on the jump curve probability of A, up to a constant multiplicative factor which is subpolynomial in ε^{-1}.

Recall that the jump ensemble involves the notion of a pole set. We are provided control over the minimum separation of consecutive poles, and so by our choice of $d_{\text{ip}} = 5d$ we ensure that there is at most one pole in the larger interval $[-2d, 2d]$. Thus there are two possibilities: either the interval $[-2d, 2d]$ contains a single pole, or no poles.

In the case where there are no poles in $[-2d, 2d]$, there can of course be poles which are arbitrarily close to $[-2d, 2d]$, and we will see in our analysis that our bounds become too weak if a pole is too close to a point under consideration. This is why we consider the presence of poles in $[-2d, 2d]$ even though the interval of interest is $[-d, d]$: when there is no pole in $[-2d, 2d]$ we can focus the analysis at the points $\pm d$, which are then ensured a distance of at least d from the nearest pole. When there *is* a pole, we will adopt a trick that will be described later to allow us to step back from it, again giving us an order d distance from it.

In the scenario where there is no pole in $[-2d, 2d]$, the jump curve is essentially just a Brownian bridge on $[-2d, 2d]$, and the argument for the probability comparison we claim is much more straightforward. So for the purposes of the exposition in this section we discuss the more difficult case where there is a pole in $[-2d, 2d]$.

By the Brownian Gibbs property, we have a direct way to write the probability of A for the jump curve, given its endpoint values $J(-2d)$ and $J(2d)$, in terms of the Brownian bridge probability of A between those endpoint values. In some sense, our task is to show that this Brownian bridge probability becomes a Brownian motion probability when we take expectations over $J(-2d)$ and $J(2d)$.

Translating between Brownian motion and Brownian bridge

It is instructive to look at how Brownian motion probabilities translate to Brownian bridge probabilities. We remind the reader of a useful property of Brownian motion: if B is a Brownian motion, then the distribution of B on an interval $[x_1, x_2]$, conditionally on $B(x_1)$ and $B(x_2)$, is that of Brownian bridge from $(x_1, B(x_1))$ to $(x_2, B(x_2))$.

Now suppose we are again working on $[-2d, 2d]$, and suppose B is a rate one Brownian motion started from x-coordinate $-4d$ according to some probability distribution μ. Right now it is

not clear why we need the device of the measure μ, but we will retain it to give ourselves some freedom which, by proper choice at a later point, will make the requirements more lenient. At present it can be thought of as, and indeed will later be defined to be, the uniform measure on a very large interval centred at 0.

On a heuristic level, the probability that $B(-4d) \in [y, y + \mathrm{d}y]$ and $B(4d) \in [z, z + \mathrm{d}z]$ is given by $(16\pi d)^{-\frac{1}{2}} \exp(-(y-z)^2/16d)\,\mathrm{d}\mu(y)\,\mathrm{d}z$. By the Markov property, $B(\,\cdot\,) - B(-d)$ is standard Brownian motion started at coordinates $(-d, 0)$ when restricted to $[-d, \infty)$. Now, recall that $\mathcal{B}_{0,*}^{[-d,d]}$ is the law of standard Brownian motion on $[-d, d]$, started at coordinates $(-d, 0)$, and that $\mathcal{B}_{y,z}^{[-4d,4d]}$ is the law of rate one Brownian bridge from $(-4d, y)$ to $(4d, z)$. Then for $A \subseteq \mathcal{C}_{0,*}([-d, d], \mathbb{R})$, we have the calculation

$$\begin{aligned}
\mathcal{B}_{0,*}^{[-d,d]}\big(A\big) &= \mathbb{P}\big(B(\,\cdot\,) - B(-d) \in A\big) \\
&= \mathbb{E}\Big[\mathbb{P}\big(B(\,\cdot\,) - B(-d) \in A \mid B(-4d), B(4d)\big)\Big] \\
&= \frac{1}{\sqrt{16\pi d}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{B}_{y,z}^{[-4d,4d]}\big(\tilde{A}\big) \cdot e^{-(y-z)^2/16d}\,\mathrm{d}\mu(y)\,\mathrm{d}z, \qquad (5.8)
\end{aligned}$$

where $\tilde{A}$ is the set of functions f in $\mathcal{C}_{*,*}([-4d, 4d], \mathbb{R})$ such that $f(\,\cdot\,) - f(-d)$, regarded as a function with domain $[-d, d]$, lies in A. As we said before, $(16\pi d)^{-\frac{1}{2}} \exp(-(y-z)^2/16d)$ is exactly the conditional density of $B(4d)$ at z given that $B(-4d)$ is y. A very similar calculation holds for the jump curve, which points us to what we should try to prove. Let $f_J(y, z)$ be the joint density of

$$Y^* := J(-4d) \quad \text{and} \quad Z^* := J(4d)$$

at the point $(y, z) \in \mathbb{R}^2$; the $*$ is an adornment that will be removed in the final definition of Y and Z that will be used in our actual arguments.

Suppose now, for simplicity, that the pole p in $[-2d, 2d]$ is at zero and has height zero (i.e., $p = 0$ and $\mathrm{Tent}(p) = 0$). Then, on $\mathsf{Fav} \cap \{P \cap [-2d, 2d] \neq \emptyset\}$,

$$\begin{aligned}
\mathbb{P}_{\mathcal{F}}\big(J(\,\cdot\,) - J(-d) \in A\big) &= \mathbb{E}_{\mathcal{F}}\Big[\mathbb{P}_{\mathcal{F}}\Big(J(\,\cdot\,) - J(-d) \in A \,\Big|\, J(-4d), J(4d)\Big)\Big] \\
&= \mathbb{E}_{\mathcal{F}}\Big[\mathcal{B}_{J(-4d),J(4d)}^{[-4d,4d]}\big(\tilde{A} \mid J(0) \geq 0\big)\Big] \\
&\leq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\mathcal{B}_{y,z}^{[-4d,4d]}\big(\tilde{A}\big)}{\mathcal{B}_{y,z}^{[-4d,4d]}\big(J(0) \geq 0\big)} \cdot f_J(y, z)\,\mathrm{d}y\,\mathrm{d}z. \qquad (5.9)
\end{aligned}$$

In essence, our aim is to run this calculation forward, and the previous one backwards, in order to get from the jump curve probability of A to the standard Brownian motion probability of A. Then by direct comparison of the integrands, in order for the last line of the second calculation to be roughly equal to the last line of the first calculation we would need

$$f_J(y, z) \text{ “} = \text{ ”} \mathcal{B}_{y,z}^{[-4d,4d]}\big(J(0) \geq 0\big) \cdot d^{-\frac{1}{2}} \exp\left(-\frac{(y-z)^2}{16d}\right) \cdot \frac{\mathrm{d}\mu}{dy}(y)$$

$$= \mathbb{P}\left(N\left(\frac{y+z}{2}, 2d\right) \geq 0\right) \cdot d^{-\frac{1}{2}} \exp\left(-\frac{(y-z)^2}{16d}\right) \cdot \frac{\mathrm{d}\mu}{dy}(y).$$

Of course, we will have to provide ourselves some error margins in order to succeed. This is hiding in the "=" symbol above, which means that the left side is bounded above by the right side, possibly multiplied by a constant of the form $\exp(G(\log \varepsilon^{-1})^{5/6})$, which we will often refer to as the *leeway factor*; the notation of "=", however, we use only in this instance. Since $T = D_k(\log \varepsilon^{-1})^{1/3}$, we will also often refer to constants of the form $\exp(GT^{5/2})$ as leeway factors.

So equivalently, what we require is that

$$f_J(y,z) \cdot \left[\mathbb{P}\left(N\left(\frac{y+z}{2}, 2d\right) \geq 0\right)\right]^{-1} \cdot d^{\frac{1}{2}} \exp\left(\frac{(y-z)^2}{16d}\right) \leq \exp\left(G(\log \varepsilon^{-1})^{5/6}\right). \quad (5.10)$$

Notice that we have temporarily ignored the issue of choosing μ and the slightly stronger demand (than the above inequality) that will arise from its Radon-Nikodym derivative with respect to Lebesgue measure. We will return to this in a few paragraphs, but only promise here that the cost will be polynomial in $\log \varepsilon^{-1}$ instead of exponential, and so will not substantively affect the analysis.

Remark 5.13. This correspondence between the random variables, Y and Z (temporarily with the $*$ adornment added), and the arguments of their joint density, y and z, is one that will be maintained throughout the proof of Theorem 5.11. Sometimes we will also refer to the marginal densities of Y or Z. By an abuse of notation (as Y and Z do not necessarily have the same distribution), we will refer to both densities as f_J, distinguishing whether we mean that of Y or Z based on whether the argument is y or z. This will not cause any confusion as at no point will we refer to a marginal density at a specific value. Similarly, $f_J(z \mid y)$ will be the conditional density of Z given $Y = y$, evaluated at the point z, and so on.

The vault and slope costs

With this simplification, let us focus on the two terms multiplying $f_J(y,z)$ on the left-hand side of (5.10). They arise from two conceptually distinct sources. The first of these two terms comes from the potential difficulty a Brownian bridge faces in order to jump or *vault* over the pole, a task which J is conditioned to accomplish. Accordingly we refer to this factor as the *vault* cost, and it will be denoted by V^* (the $*$ being again an adornment that will be removed in the final corrected version of V which we will use). The second of the two terms, on the other hand, is the potential difficulty faced by a Brownian motion to attain the slope specified by y and z, which may be equivalently be thought of as an increment, across an interval of length $4d$. We refer to this as the *slope* cost, and denote it by S^*. Thus,

$$\begin{aligned} V^* &:= \left[\mathbb{P}\left(N\left(\frac{y+z}{2}, 2d\right) \geq 0\right)\right]^{-1} \\ S^* &:= d^{\frac{1}{2}} \cdot \exp\left(\frac{(y-z)^2}{16d}\right). \end{aligned} \quad (5.11)$$

Note that V^* behaves differently depending on the values of y and z; for example, if $y+z$ is positive, the probability is bounded below by a constant, and so V^* is bounded above by a constant and is easily managed by the margin of error we have provided. On the other hand, if $y+z$ is negative, V^* can be seen to be roughly $d^{\frac{1}{2}} \exp((y+z)^2/16d)$, thus posing a much more serious demand. More precisely,

$$V^* \lesssim \begin{cases} 1 & y+z>0 \\ d^{\frac{1}{2}} \cdot \exp\left(\frac{(y+z)^2}{16d}\right) & y+z \leq 0, \end{cases} \tag{5.12}$$

Our analysis will later break into cases based on this fact.

Thus, (5.10) says that roughly what we need to show, for some $G < \infty$, is that

$$f_J(y,z) \cdot V^* \cdot S^* \leq \exp\left(G(\log \varepsilon^{-1})^{5/6}\right).$$

The above heuristic description was idealised to highlight the main features of the approach, but is in essence correct. We now discuss which aspects of the description change in the actual approach. There were three simplifying assumptions: that the pole position p equals zero; that the height of the pole is zero; and the postponement of the choice of μ. In addressing each of these simplifications, we will come to the final quantities V and S, and we will see that V and S respectively equal V^* and S^* up to the leeway factor.

Addressing the simplifications

Pole position $p = 0$.

The first simplifying assumption we made in the heuristic description was that the pole in $[-2d, 2d]$ lies at 0. In general, of course, we have very limited control over the pole position as it is determined by the $(k+1)^{\text{st}}$ curve of the original line ensemble. This point will be addressed essentially by changing our frame of reference horizontally.

Call the pole position $p \in [-2d, 2d]$ and suppose $p \neq 0$. The coefficient of $1/2$ for each of y and z in V^* was due to the pole position of 0 being equidistant from $-4d$ and $4d$. One option would be to maintain the random variables Y^* and Z^* to be the values at $\pm 4d$, in which case the coefficients would not be $1/2$ but λ and $1-\lambda$ for some $\lambda \in [0,1]$. While the subsequent analysis could possibly be adapted for this case, this would introduce an undesirable level of complication to the formulas. So to maintain the symmetric coefficient of $1/2$, we instead have Y^* and Z^* be the values at $p-4d$ and $p+4d$, i.e., $J(p-4d)$ and $J(p+4d)$ respectively. (We will introduce a further modification shortly which will be the final definition of Y and Z, without the $*$ adornment.)

While on some level this is merely a trick, it is one which leads to very useful simplifications. At a technical level, though p is a random variable, it is an $\mathcal{F}$-measurable one, and so this trick is sound: the relevant point is that, given the $\mathcal{F}$-data, there is no obstacle to writing the law of

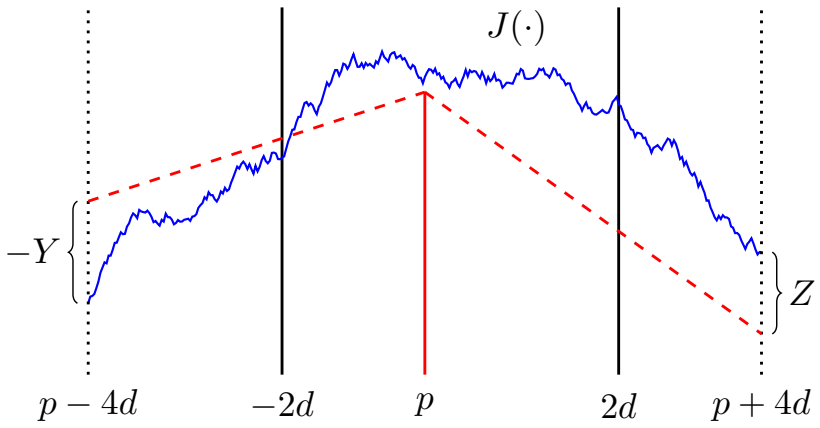


Figure 5.5: Illustrating the definition of Y and Z when $P \cap [-2d, 2d] \neq \emptyset$. The Tent map is depicted as the piecewise linear function drawn in dashed red. The single pole p in $[-2d, 2d]$ is depicted as a vertical red line. The jump curve is blue. In this figure Y is negative as J is below Tent at $p - 4d$, and thus the distance indicated is $-Y$. The quantity Z is positive as J is above Tent at $p + 4d$.

the jump curve on $[-2d, 2d]$ as the marginal of Brownian bridge on $[p - 4d, p + 4d]$, between correctly distributed endpoints, which is conditioned to jump over the pole at p.

Pole height Tent$(p) = 0$.

Now we address the height of the pole. Here also we essentially employ a change in the frame of reference. At this point in our discussion, we have Y^* and Z^* being the values of $J(p - 4d)$ and $J(p + 4d)$ themselves, i.e., the deviation from $0 = \text{Tent}(p)$. When $\text{Tent}(p) \neq 0$, the obvious choice is to let Y and Z respectively represent the deviations of $J(p - 4d)$ and $J(p + 4d)$ from $\text{Tent}(p)$. But to make certain derivations slightly simpler in the sequel, we instead let the final definitions of Y and Z be the respective deviations of J from Tent at $p - 4d$ and $p + 4d$, i.e.,

$$\begin{aligned} Y &:= J(p - 4d) - \text{Tent}(p - 4d), \\ Z &:= J(p + 4d) - \text{Tent}(p + 4d). \end{aligned} \tag{5.13}$$

See Figure 5.5. This is the final definition of Y and Z which will be maintained for the rest of the argument, at least on the event that $P \cap [-2d, 2d] \neq \emptyset$. (On the other event, where there is no pole, they will have conceptually analogous but different definitions.)

The choice of μ.

We still have to make a suitable choice for μ. We start by providing some intuition as to the role of μ.

Since our approach to move from jump curve probabilities (5.9) to Brownian motion probabilities (5.8) is somewhat crude, in the sense that we are trying to directly compare the integrands to conclude that the integrals are comparable, in order to succeed we cannot allow the integrand in (5.8) to be zero when the integrand in (5.9) is not small. This suggests that the support of μ should be the support of the law of Y, i.e., the support of $f_J(y)$ (recall from Remark 5.13 that this is the marginal density of Y). However, it is reasonable to assume that $f_J(y) > 0$ on the entire real line, and we run into a problem when we try to find a full support measure μ that also satisfies the constraints imposed by the costs V and S. This will be shown in the discussion in the following paragraphs. Heuristically, the solution will be to let μ have support only where $f_J(y)$ is not too small in some sense that we will specify.

From the above expressions for V^* and S^*, we see that, in the worst case, the total cost to be paid is $\exp\left(\frac{y^2}{8d} + \frac{z^2}{8d}\right)$. This tells us that we cannot afford for $\frac{\mathrm{d}\mu}{\mathrm{d}y}$ to be too small anywhere, as the additional cost corresponding to the choice of μ will be the reciprocal of this derivative. The logic behind this inference is as follows: in the worst case, we have (without loss of generality) $y < 0$, as well as the condition $J(0) \geq 0$, where for the sake of simplified discussion we have again assumed $p = 0$ and $\mathrm{Tent}(p) = 0$. A Brownian bridge forced to make a large jump (from y to 0) in unit order time will, because of the difficulty, make the required jump with very little extra margin. So heuristically we may think of the $J(0) \geq 0$ condition as being $J(0) = 0$. But the density of a Brownian motion which is 0 at 0 having value y at $-4d$ and z at $4d$ is exactly $(8\pi d)^{-1} \exp\left(-y^2/8d - z^2/8d\right)$. Thus we expect that this is the best density bound we can hope for J as well, which means any extra requirement imposed by $\frac{\mathrm{d}\mu}{\mathrm{d}y}$ must be absorbable in the leeway factor of $\exp\left(G(\log \varepsilon^{-1})^{5/6}\right)$.

In other words, we require

$$\frac{\mathrm{d}\mu}{\mathrm{d}y}(y) \geq \exp\left(-G(\log \varepsilon^{-1})^{5/6}\right). \tag{5.14}$$

This bound cannot hold for all y for the density of a probability measure. But a closer look at (5.9) shows a modification we may make to that calculation: we can look for the density bound on $f_J(y, z)$ that we have been discussing for y and z in a *good* region, and find a separate argument for why the contribution of the integral from the bad regions is $O(\varepsilon)$. More precisely, call the good region $\mathcal{G}_R^{(1)} \subseteq \mathbb{R}^2$ (R is a parameter we will set later, and the 1 in the superscript refers to the fact that this is the first case, when a pole is present), and let $\tilde{y} = y + \mathrm{Tent}(p - 4d)$, $\tilde{z} = z + \mathrm{Tent}(p + 4d)$; this is so that $Y = y$ implies $J(p - 4d) = \tilde{y}$, and similarly $Z = z$ implies $J(p + 4d) = \tilde{z}$. Then, on $\mathsf{Fav} \cap \{P \cap [-2d, 2d] \neq \emptyset\}$,

$$\begin{aligned}
&\mathbb{P}_{\mathcal{F}}\big(J(\,\cdot\,) - J(-d) \in A\big) \qquad (5.15)\\
&\quad = \mathbb{E}_{\mathcal{F}}\Big[\mathbb{P}_{\mathcal{F}}\big(J(\,\cdot\,) - J(-d) \in A \mid J(p - 4d), J(p + 4d)\big)\Big]\\
&\quad \leq \mathbb{E}_{\mathcal{F}}\Big[\mathbb{P}_{\mathcal{F}}\big(J(\,\cdot\,) - J(-d) \in A \mid J(p - 4d), J(p + 4d)\big) \cdot \mathbb{1}_{(Y,Z)\in\mathcal{G}_R^{(1)}}\Big] + \mathbb{P}_{\mathcal{F}}\big((Y, Z) \notin \mathcal{G}_R^{(1)}\big)\\
&\quad = \mathbb{E}_{\mathcal{F}}\Big[\mathcal{B}^{[p-4d,p+4d]}_{J(p-4d),J(p+4d)}\big(\widetilde{A} \mid J(p) \geq \mathrm{Tent}(p)\big) \cdot \mathbb{1}_{(Y,Z)\in\mathcal{G}_R^{(1)}}\Big] + \mathbb{P}_{\mathcal{F}}\big((Y, Z) \notin \mathcal{G}_R^{(1)}\big)
\end{aligned}$$

$$\leq \iint_{\mathcal{G}_R^{(1)}} \frac{\mathcal{B}_{\tilde{y},\tilde{z}}^{[p-4d,p+4d]}\big(\widetilde{A}\big)}{\mathcal{B}_{\tilde{y},\tilde{z}}^{[p-4d,p+4d]}\big(J(p) \geq \mathrm{Tent}(p)\big)} \cdot f_J(y,z)\,\mathrm{d}y\,\mathrm{d}z + \mathbb{P}_{\mathcal{F}}\big((Y,Z) \notin \mathcal{G}_R^{(1)}\big),$$

where, recall, $\tilde{A}$ is the set of functions f in $\mathcal{C}_{*,*}([-4d,4d],\mathbb{R})$ such that $f(\,\cdot\,) - f(-d)$, regarded as a function with domain $[-d,d]$, lies in A. Thus we see that, if we can show the second term in the final displayed line is $O(\varepsilon)$, then we only require the density bounds associated with V and S for $(y,z) \in \mathcal{G}_R^{(1)}$, and the same holds true for the lower bound on $\frac{\mathrm{d}\mu}{\mathrm{d}y}(y)$. It will turn out that including the condition $(y,z) \in [-RT^{3/2}, RT^2]^2$ in the definition of $\mathcal{G}_R^{(1)}$ allows us to obtain the $O(\varepsilon)$ bound on the second term. (We will remark later in Section 5.2 on why we require the lower bound to be $-RT^{3/2}$ and not $-RT^2$.) So we only need to meet the condition (5.14) for $y,z \in [-RT^{3/2}, RT^2]$. An easy choice which meets this requirement is the uniform measure on $[-RT^2, RT^2]$, and so we let μ be this measure. Thus, we set

$$\frac{\mathrm{d}\mu}{\mathrm{d}y}(y) = (2RT^2)^{-1},$$

for $y \in [-RT^2, RT^2]$. We will treat separately the corresponding cost, which is equal to $2RT^2$, and not include it in either V or S.

The good regions

We now give the definition of $\mathcal{G}_R^{(1)}$ as a subset of $\mathbb{R}^2$.

$$\mathcal{G}_R^{(1)} = \left\{ (y,z) \in \mathbb{R}^2 \;:\; \begin{array}{c} y \in (-RT^{3/2}, RT^2), \\ z \in (-RT^{3/2}, RT^2), \\ |y-z| < 2RT^{3/2} \end{array} \right\}. \tag{5.16}$$

We have included an extra condition that $|y-z| \leq 2RT^{3/2}$. To have the $O(\varepsilon)$ upper bound on the probability of $(Y,Z) \in \mathcal{G}_R^{(1)}$ means that the increment of J across an interval of $8d$ must be bounded by $O(T^{3/2}) = O((\log \varepsilon^{-1})^{1/2})$ with probability at least $1-\varepsilon$. This is plausible, on the basis that the probability that a Brownian motion has an increment of size greater than $O((\log \varepsilon^{-1})^{1/2})$ over a unit order interval is polynomial in ε. In fact, we have the following bound on the event that $(Y,Z) \in \mathcal{G}_R^{(1)}$.

Lemma 5.14. *We have for $R \geq 3$ and $d \geq 1$*

$$\mathbb{P}_{\mathcal{F}}\Big((Y,Z) \notin \mathcal{G}_R^{(1)}\Big) \cdot \mathbb{1}_{\mathsf{Fav},P\cap[-2d,2d]\neq\mathcal{O}} \lesssim \left(\varepsilon^{R^2D_k^3/8d} + \varepsilon^{(R-3)^2D_k^3/2}\right) \cdot \exp\left(13R^2D_k^{5/2}\left(\log\varepsilon^{-1}\right)^{5/6}\right).$$

To prove this, we will actually make use of a similar but weaker statement about a good region $\mathcal{G}_R^{(2)} \subseteq \mathbb{R}^2$, the difference being that the bound on the probability that $(Y,Z) \in \mathcal{G}_R^{(2)}$ will hold regardless of the presence or absence of a pole in $[-2d,2d]$. We define

$$\mathcal{G}_R^{(2)} = \left\{ (y,z) \in \mathbb{R}^2 \;:\; \begin{array}{c} y \in (-RT^2, RT^2), \\ z \in (-RT^2, RT^2), \\ |y-z| < 2RT^{3/2} \end{array} \right\}. \tag{5.17}$$

Our earlier definitions of Y and Z were on the $\mathcal{F}$-measurable event that $P \cap [-2d, 2d] \neq \emptyset$ and do not make sense otherwise, as they are deviations of J from Tent at points defined relative to p. So on the event that $P \cap [-2d, 2d] = \emptyset$, we define

$$\begin{aligned} Y &:= J(-d) - \mathrm{Tent}(-d) \\ Z &:= J(d) - \mathrm{Tent}(d). \end{aligned} \tag{5.18}$$

In the following lemma, Y and Z are defined in this case-specific manner.

Lemma 5.15. *We have that for $R \geq 3$*

$$\mathbb{P}_{\mathcal{F}}\left((Y,Z) \notin \mathcal{G}_R^{(2)}\right) \cdot \mathbb{1}_{\mathsf{Fav}} \lesssim \left(\varepsilon^{R^2 D_k^3/4d} + \varepsilon^{(R-3)^2 D_k^3/2}\right) \exp\left(13 R^2 D_k^{5/2} \left(\log \varepsilon^{-1}\right)^{5/6}\right).$$

While we defer the proofs of these lemmas to Section 6.1, let us say a few words about its approach. The statements about Y and Z being bounded above in absolute value by RT^2 are proved by stochastically dominating J by an appropriate Brownian bridge on Fav, and similarly with the roles of J and the Brownian bridge reversed for the lower bound. The slightly trickier issue is bounding the increment across the interval of length $8d$. We postpone discussing this point.

Based on the form of the statements of Lemmas 5.14 and 5.15, we set R (somewhat arbitrarily) as

$$R = 6\sqrt{d}, \tag{5.19}$$

so that, for small enough ε, and since $d \geq 1$ and $D_k \geq 2$ (see (4.3)),

$$\mathbb{P}_{\mathcal{F}}\left((Y,Z) \notin \mathcal{G}_R^{(1)}\right) \cdot \mathbb{1}_{\mathsf{Fav}, P \cap [-2d,2d] \neq \emptyset} \leq \varepsilon \quad \text{and} \quad \mathbb{P}_{\mathcal{F}}\left((Y,Z) \notin \mathcal{G}_R^{(2)}\right) \cdot \mathbb{1}_{\mathsf{Fav}} \leq \varepsilon.$$

The final costs

Having addressed the three simplifications in the heuristic derivation, let us see how the costs have changed. Recall that the random variables Y and Z are respectively the quantities $J(p-4d) - \mathrm{Tent}(p-4d)$ and $J(p+4d) - \mathrm{Tent}(p+4d)$. As immediately preceding (5.15), let $\tilde{y} = y + \mathrm{Tent}(p-4d)$ and $\tilde{z} = z + \mathrm{Tent}(p+4d)$, so that $Y = y$ implies $J(p-4d) = \tilde{y}$, and similarly $Z = z$ implies $J(p+4d) = \tilde{z}$. The final definitions of V and S are

$$\begin{aligned} V &:= \mathcal{B}_{\tilde{y},\tilde{z}}^{[p-4d,p+4d]}\left(J(p) \geq \mathrm{Tent}(p)\right)^{-1} \\ &= \mathbb{P}\left(N\left(\frac{y+z}{2}, 2d\right) \geq \mathrm{Tent}(p) - \frac{\mathrm{Tent}(p-4d) + \mathrm{Tent}(p+4d)}{2}\right)^{-1} \\ S &:= d^{\frac{1}{2}} \cdot \exp\left(\frac{1}{16d}(\tilde{y} - \tilde{z})^2\right) \\ &= d^{\frac{1}{2}} \cdot \exp\left(\frac{1}{16d}\left(y - z + \mathrm{Tent}(p-4d) - \mathrm{Tent}(p+4d)\right)^2\right). \end{aligned}$$

For $(y,z) \in \mathcal{G}_R^{(1)}$, we have that S and V differ from S^* and V^* by factors which can be absorbed within the leeway factor, which we show now. We will have, on expanding the second square in the exponent of S, some extra terms apart from the $(y-z)^2/16d$ present in S^*. The extra factor is

$$\exp\left(\frac{1}{16d}(\mathrm{Tent}(p-4d)-\mathrm{Tent}(p+4d))^2+\frac{1}{8d}(y-z)(\mathrm{Tent}(p-4d)-\mathrm{Tent}(p+4d))\right). \tag{5.20}$$

In this expression, we see immediately that the first term in the exponent does not cause a problem: since Tent has slope bounded in absolute value by $4T$ (recall (5.4)), it follows that $|\mathrm{Tent}(p-4d)-\mathrm{Tent}(p+4d)| \le 32Td$, so that

$$\frac{1}{16d}(\mathrm{Tent}(p-4d)-\mathrm{Tent}(p+4d))^2 \le 64T^2d,$$

which is well below $O((\log\varepsilon^{-1})^{5/6})$. We need the same bound to hold for the second term in the exponent in (5.20) as well. Again using that Tent has absolute value of slope bounded by $4T$, and that $|y-z| \le 2RT^{3/2}$ on $\mathcal{G}_R^{(1)}$, what we find is

$$(y-z)(\mathrm{Tent}(p-4d)-\mathrm{Tent}(p+4d))) \le 2RT^{3/2}\cdot 32Td = 64RT^{5/2}d.$$

This is why we included the bound on $|y-z|$ of order $T^{3/2}$ in the definition of $\mathcal{G}_R^{(1)}$. In other words, we have that

$$S \lesssim S^*\cdot\exp\left(64T^2d+8RT^{5/2}\right) \le d^{\frac{1}{2}}\cdot\exp\left(\frac{1}{16d}(y-z)^2+10RT^{5/2}\right), \tag{5.21}$$

the last inequality since $d \le \sqrt{T}/24$, $R \ge 6$ from (5.19), and $64/24 \le 3$.

Remark 5.16. We here used that $d \le \sqrt{T}/24$, which is equivalent to the assumption $\varepsilon \le \exp\left(-(24)^6d^6/D_k^3\right)$ made in Theorems 4.8 and 5.11; we had previously made use of this inequality in the proof of Theorem 4.8. We will be making use of this inequality many times in the sequel as well, as here to bound expressions of the form T^2d by $T^{5/2}/24$, and also to reduce coefficients when we have a margin to convert a small power of d to T.

Let us now finally analyse V in comparison with V^*, which will show us why we included in $\mathcal{G}_R^{(1)}$ that $y,z \ge -RT^{3/2}$.

We first focus on the right-hand quantity in the probability expression of V, namely

$$\mathrm{Tent}(p)-\frac{\mathrm{Tent}(p-4d)+\mathrm{Tent}(p+4d)}{2}.$$

By the fact that the slope of Tent is bounded in absolute value by $4T$, this quantity is bounded above by

$$\frac{1}{2}\times 4d\times 4T+\frac{1}{2}\times 4d\times 4T = 16Td.$$

Thus V is bounded as

$$V \leq \mathbb{P}\left(N\left(\frac{y+z}{2}, 2d\right) \geq 16Td\right)^{-1}.$$

As in the analysis of V^* in (5.12), we see that this quantity behaves differently depending on the value of $y+z$:

$$V \lesssim \begin{cases} 1 & y+z > 32Td \\ d^{\frac{1}{2}} \cdot \exp\left(\frac{1}{16d}(y+z-32Td)^2\right) & y+z \leq 32Td. \end{cases} \tag{5.22}$$

The bound in the worse case, on expanding the exponent and including only the factors which differ from those of V^*, is

$$\exp\left(64T^2d - 4(y+z)T\right).$$

Again the first term in the exponent, being $(\log \varepsilon^{-1})^{2/3}$, does not cause a problem, and so we turn to the second term. Firstly we see that we have the trivial upper bound of 0 on the term $-4(y+z)T$ when $y+z \geq 0$. When $y+z < 0$, we use that $y, z \geq -RT^{3/2}$ from the definition of $\mathcal{G}_R^{(1)}$ to see that

$$-(y+z) \cdot T \leq 2RT^{5/2}.$$

We note that any weaker lower bound on y, z would not have been sufficient to obtain an upper bound of order $T^{5/2}$ on the above quantity. From the previous few paragraphs we conclude that, when $y+z \leq 32Td$,

$$V \lesssim d^{\frac{1}{2}} \cdot \exp\left(\frac{1}{16d}(y+z)^2 + 11RT^{5/2}d\right), \tag{5.23}$$

since $d \leq \sqrt{T}/24$ and $R \geq 1$ imply $64T^2d + 8RT^{5/2} \leq (64/24)T^{5/2} + 8RT^{5/2} \leq 11RT^{5/2}$.

So overall what we have observed is that for $(y, z) \in \mathcal{G}_R^{(1)}$, it is true that

$$V \approx V^*$$
$$S \approx S^*,$$

in the sense that the left sides are bounded by the right sides up to multiplication by the factor $\exp\left(G(\log \varepsilon^{-1})^{5/6}\right)$. Thus, while we will work with V and S, the reader is advised to keep in mind the more convenient expressions from (5.21) and (5.23) after ignoring the leeway factor.

The bound to be proved

With the quantities V and S defined, the following is what we will prove for (y, z) in the good region $\mathcal{G}_R^{(1)}$ and when $P \cap [-2d, 2d] \neq \emptyset$:

$$f_J(y, z) \cdot V \cdot S \cdot \left(\frac{\mathrm{d}\mu}{dy}(y)\right)^{-1} \leq G_1 \cdot \exp\left(G_2(\log \varepsilon^{-1})^{5/6}\right), \tag{5.24}$$

for some finite constants $G_1 = G_1(\varepsilon, k, d)$ and $G_2 = G_2(k, d)$, and for $\varepsilon < \varepsilon_0$ for some $\varepsilon_0 > 0$. Since the Radon-Nikodym derivative term is $2RT^2$, it is therefore sufficient to prove

$$f_J(y,z) \cdot V \cdot S \le G_1' \cdot \exp\left(G_2'(\log \varepsilon^{-1})^{5/6}\right), \tag{5.25}$$

for some $G_1' = G_1'(\varepsilon, k, d) < \infty$ and $G_2' = G_2'(k, d) > 0$; then (5.25) implies (5.24) with $G_1 = G_1' \cdot 2RT^2$ and $G_2 = G_2'$. We end this section by showing that if we have (5.24), then we will almost have our main Theorem 5.11. More precisely, we have the following lemma, which, along with a proposition about the no-pole case, will allow us to prove Theorem 5.11 modulo proving these input statements.

Lemma 5.17. *Suppose for all $(y,z) \in \mathcal{G}_R^{(1)}$ we have (5.24). Then with G_1 and G_2 as in (5.24) we have*

$$\mathbb{P}_{\mathcal{F}}\big(J(\cdot) - J(-d) \in A, (Y,Z) \in \mathcal{G}_R^{(1)}\big) \cdot \mathbb{1}_{\mathsf{Fav}, P \cap [-2d, 2d] \neq \emptyset} \lesssim G_1 \cdot \varepsilon \cdot \exp\left(G_2(\log \varepsilon^{-1})^{5/6}\right).$$

Proof. Let B be a Brownian motion begun at $p - 4d$ according to the distribution μ. We have on $\mathsf{Fav} \cap \{P \cap [-2d, 2d] \neq \emptyset\}$,

$$\mathbb{P}_{\mathcal{F}}\Big(J(\cdot) - J(-d) \in A,\ (Y,Z) \in \mathcal{G}_R^{(1)}\Big) \le \iint_{\mathcal{G}_R^{(1)}} \frac{\mathcal{B}_{\tilde{y},\tilde{z}}^{[p-4d,p+4d]}\big(\tilde{A}\big)}{\mathcal{B}_{\tilde{y},\tilde{z}}^{[p-4d,p+4d]}\big(J(p) \ge \mathrm{Tent}(p)\big)} \cdot f_J(y,z)\,\mathrm{d}y\,\mathrm{d}z$$

Using (5.24), this integral is bounded by

$$\begin{aligned} &G_1 d^{-\frac{1}{2}} \exp\left(G_2(\log \varepsilon^{-1})^{5/6}\right) \iint_{\mathcal{G}_R^{(1)}} \mathcal{B}_{\tilde{y},\tilde{z}}^{[p-4d,p+4d]}\big(\tilde{A}\big) e^{-\frac{1}{16d}(\tilde{y}-\tilde{z})^2}\,\mathrm{d}\mu(y)\,\mathrm{d}z \\ &\qquad \lesssim \mathbb{P}\big(B(\cdot) - B(-d) \in A\big) \cdot G_1 \cdot \exp\left(G_2(\log \varepsilon^{-1})^{5/6}\right) \\ &\qquad = \mathcal{B}_{0,*}^{[-d,d]}(A) \cdot G_1 \cdot \exp\left(G_2(\log \varepsilon^{-1})^{5/6}\right) = G_1 \cdot \varepsilon \cdot \exp\left(G_2(\log \varepsilon^{-1})^{5/6}\right). \end{aligned}$$

The second-to-last equality follows from the Markov property of Brownian motion. □

In fact, we will establish (5.24) with $G_2 = 792 \cdot d \cdot D_k^{5/2}$ and $G_1 = G'RT^4$, where G' is an absolute constant independent of d, k, ε. This is the content of Lemmas 6.3, 6.6, and 6.14 ahead, as $792d \ge \max(22R^2, 41R)$ from R's value set in (5.19), and the observation above that to go from (5.25) to (5.24) we must multiply G_1' by $2RT^2$. We will also establish the following proposition in the no-pole case:

Proposition 5.18. *There exists a positive constant $\varepsilon_0 = \varepsilon_0(d) > 0$ such that if $\varepsilon < \varepsilon_0$, then, on $\mathsf{Fav} \cap \{P \cap [-2d, 2d] = \emptyset\}$,*

$$\mathbb{P}_{\mathcal{F}}\Big(J(\cdot) - J(-d) \in A,\ (Y,Z) \in \mathcal{G}_R^{(2)}\Big) \lesssim \varepsilon \cdot \exp\left(756 \cdot d \cdot D_k^{5/2} \cdot (\log \varepsilon^{-1})^{5/6}\right).$$

Admitting these statements for now, namely Lemmas 5.14, 5.15, Proposition 5.18, and that we have (5.24) with $G_1 = G'RT^4$ and $G_2 = 792 \cdot d \cdot D_k^{5/2}$, we may complete the proof of Theorem 5.11.

Proof of Theorem 5.11. The quantity $\mathbb{P}\left(J(\cdot) - J(-d) \in A\right) \cdot \mathbb{1}_{\mathsf{Fav}}$ satisfies the following upper bound:

$$\begin{aligned}
&\mathbb{P}\left(J(\cdot) - J(-d) \in A\right) \cdot \mathbb{1}_{\mathsf{Fav}} \\
&\quad \le \left[\mathbb{P}_{\mathcal{F}}\left(J(\cdot) - J(-d) \in A,\, (Y,Z) \in \mathcal{G}_R^{(1)}\right) + \mathbb{P}_{\mathcal{F}}\left((Y,Z) \notin \mathcal{G}_R^{(1)}\right)\right] \mathbb{1}_{\mathsf{Fav}, P \cap [-2d,2d] \neq \emptyset} \\
&\qquad + \left[\mathbb{P}_{\mathcal{F}}\left(J(\cdot) - J(-d) \in A,\, (Y,Z) \in \mathcal{G}_R^{(2)}\right) + \mathbb{P}_{\mathcal{F}}\left((Y,Z) \notin \mathcal{G}_R^{(2)}\right)\right] \mathbb{1}_{\mathsf{Fav}, P \cap [-2d,2d] = \emptyset}.
\end{aligned} \tag{5.26}$$

Focus on the first term after the inequality of (5.26). By Lemma 5.14 and our choice of $R = 6\sqrt{d}$ from (5.19),

$$\mathbb{P}_{\mathcal{F}}\left((Y,Z) \notin \mathcal{G}_R^{(1)}\right) \cdot \mathbb{1}_{\mathsf{Fav}, P \cap [-2d,2d] \neq \emptyset} \lesssim \varepsilon \cdot \exp\left(468 \cdot d \cdot D_k^{5/2} (\log \varepsilon^{-1})^{5/6}\right).$$

From our assumption that we have (5.24) with $G_2 = 792 \cdot d \cdot D_k^{5/2}$ and $G_1 = G' R T^4$ for all $(y,z) \in \mathcal{G}_R^{(1)}$, we get from Lemma 5.17

$$\begin{aligned}
&\mathbb{P}_{\mathcal{F}}\left(J(\cdot) - J(-d) \in A,\, (Y,Z) \in \mathcal{G}_R^{(1)}\right) \cdot \mathbb{1}_{\mathsf{Fav}, P \cap [-2d,2d] \neq \emptyset} \\
&\qquad \lesssim \varepsilon \cdot d^{\frac{1}{2}} \cdot D_k^4 (\log \varepsilon^{-1})^{\frac{4}{3}} \cdot \exp\left(792 \cdot d \cdot D_k^{5/2} (\log \varepsilon^{-1})^{5/6}\right)
\end{aligned} \tag{5.27}$$

Now we turn to the second line of (5.26). From Lemma 5.15 and our choice of $R = 6\sqrt{d}$ in (5.19),

$$\mathbb{P}_{\mathcal{F}}\left((Y,Z) \notin \mathcal{G}_R^{(2)}\right) \cdot \mathbb{1}_{\mathsf{Fav}, P \cap [-2d,2d] = \emptyset} \lesssim \varepsilon \cdot \exp\left(468 \cdot d \cdot D_k^{5/2} (\log \varepsilon^{-1})^{5/6}\right).$$

Finally from Proposition 5.18 we have

$$\mathbb{P}_{\mathcal{F}}\left(J(\cdot) - J(-d) \in A,\, (Y,Z) \in \mathcal{G}_R^{(2)}\right) \lesssim \varepsilon \cdot \exp\left(756 \cdot d \cdot D_k^{5/2} \cdot (\log \varepsilon^{-1})^{5/6}\right).$$

Substituting these bounds into (5.26) gives

$$\mathbb{P}_{\mathcal{F}}\left(J(\cdot) - J(-d) \in A\right) \lesssim \varepsilon \cdot d^{\frac{1}{2}} \cdot D_k^4 (\log \varepsilon^{-1})^{\frac{4}{3}} \cdot \exp\left(792 \cdot d \cdot D_k^{5/2} \cdot (\log \varepsilon^{-1})^{5/6}\right),$$

completing the proof. □

At this point we pause to review our progress. We have defined good regions $\mathcal{G}_R^{(1)}$ and $\mathcal{G}_R^{(2)}$ to which we can restrict our analysis. We have also defined the costs that need to be met on $\mathcal{G}_R^{(1)}$ in order to prove Theorem 5.11, and indeed we have given the proof of Theorem 5.11 modulo these costs being met and a few additional statements being proven. Concretely, our task is now to establish (5.25) with the claimed values of G_1' and G_2'; to prove Lemmas 5.14 and 5.15; and to prove Proposition 5.18.

Establishing (5.25) will break into separate cases that will each require different arguments which will all be handled in the next Chapter 6. An easy and a moderate case, respectively when $y, z < 0$ or $y + z > 0$, will also supply us the bounds we need to prove Lemmas 5.14 and 5.15. Chapter 6 will also address the case where a pole is not present in $[-2d, 2d]$, i.e., Proposition 5.18.

Chapter 6

Proving the density bounds

This chapter proves the statements of Chapter 5 that are needed in the proof of Theorem 5.11 as given in Section 5.2; in particular, we prove here Lemmas 5.14 and 5.15, Proposition 5.18, and that equation (5.24) holds with the claimed constants. In essence, these all follow from the last item, i.e., (5.24), which is a bound on the density $f_J(y, z)$.

The proof of this density bound is broken up into four sections. The first three are when a pole is present in $[-2d, 2d]$ and are distinguished by the values of $Y = J(p-4d) - \mathrm{Tent}(p-4d)$ and $Z = J(p+4d) - \mathrm{Tent}(p+4d)$, roughly corresponding to an easy case of being below the pole on both sides (Section 6.1); a moderate case of being above the pole on both sides (Section 6.2); and a difficult case of being above and below the pole on either side (Section 6.3). Lemmas 5.14 and 5.15 are proved at the end of the moderate case, Section 6.2. The last section, Section 6.4, addresses when there is no pole in $[-2d, 2d]$, i.e., Proposition 5.18.

6.1 The easy case: Below the pole on both sides

Having set up the problem and identified what bounds (5.25) we require on $f_J(y, z)$, we now prove such a bound in a simple case. This case is when $y < 0$ and $z < 0$, which implies, from (5.21) and (5.23), that

$$V \cdot S \lesssim \exp\left(\frac{y^2}{8d} + \frac{z^2}{8d} + 21RT^{5/2}\right).$$

So, it is sufficient to prove

$$f_J(y, z) \lesssim d^{-1} \cdot \exp\left(-\frac{y^2}{8d} - \frac{z^2}{8d}\right). \tag{6.1}$$

In this case we are aided by the presence of the pole. Essentially, the desired density of J at (y, z) is bounded by the density of a particular pair of independent Brownian bridges at (y, z). More precisely, on the event $\mathsf{Fav} \cap \{P \cap [-2d, 2d] \neq \emptyset\}$, let p^- and p^+ be the elements of P immediately

preceding and succeeding p; and define

$$\sigma^2_{-4d} = 4d \cdot \frac{(p - p^- - 4d)}{p - p^-}$$
$$\sigma^2_{4d} = 4d \cdot \frac{(p^+ - p - 4d)}{p^+ - p}.$$

The first quantity is the variance of a Brownian bridge defined on the interval $[p^-, p]$ at the point $p - 4d$, while the second is the same for a Brownian bridge defined on $[p, p^+]$ at the point $p + 4d$. The following proposition is exactly the case which arose in [Ham19a] in the analysis of the Brownian bridge regularity of regular ensembles, where it was Proposition 5.17. The proof is fairly straightforward and we will shortly reproduce it here for completeness and because it aids our exposition.

Proposition 6.1. *We have that for $y, z < 0$,*

1. *Joint bound*

$$f_J(y, z) \cdot \mathbb{1}_{\mathsf{Fav}, P \cap [-2d, 2d] \neq \emptyset} \lesssim d^{-1} \cdot \exp\left(-\frac{1}{2\sigma^2_{-4d}} y^2 - \frac{1}{2\sigma^2_{4d}} z^2\right) \cdot \mathbb{1}_{\mathsf{Fav}, P \cap [-2d, 2d] \neq \emptyset}.$$

2. *Marginal bounds*

$$f_J(y) \cdot \mathbb{1}_{\mathsf{Fav}, P \cap [-2d, 2d] \neq \emptyset} \lesssim d^{-\frac{1}{2}} \cdot \exp\left(-\frac{1}{2\sigma^2_{-4d}} y^2\right) \cdot \mathbb{1}_{\mathsf{Fav}, P \cap [-2d, 2d] \neq \emptyset}$$
$$f_J(z) \cdot \mathbb{1}_{\mathsf{Fav}, P \cap [-2d, 2d] \neq \emptyset} \lesssim d^{-\frac{1}{2}} \cdot \exp\left(-\frac{1}{2\sigma^2_{4d}} z^2\right) \cdot \mathbb{1}_{\mathsf{Fav}, P \cap [-2d, 2d] \neq \emptyset}.$$

We note that we have the following simple bounds on the variances:

Lemma 6.2. *On $\mathsf{Fav} \cap \{P \cap [-2d, 2d] \neq \emptyset\}$, we have $\sigma^2_{4d}, \sigma^2_{-4d} \in [\frac{4}{5}d, 4d]$.*

Proof. The upper bound is obvious from the defining expressions. For the lower bound, we have

$$\sigma^2_{-4d} = 4d\left(1 - \frac{4d}{p - p^-}\right) \geq 4d\left(1 - \frac{4d}{5d}\right) = \frac{4}{5}d,$$

since $p - p^- \geq d_{\mathrm{ip}} = 5d$. A similar argument proves the corresponding bound for σ^2_{4d}. □

With these variance bounds and the density bounds of Proposition 6.1, the sufficient bound (5.25) is immediate:

Lemma 6.3. *When $y < 0$ and $z < 0$, we have (5.25) with $G_2' = 21RD_k^{5/2}$ and with G_1' independent of ε, k, and d.*

Proof. From (5.21) and (5.23), the total cost is bounded above as

$$V \cdot S \lesssim d \cdot \exp\left(\frac{y^2}{8d} + \frac{z^2}{8d} + 21RT^{5/2}\right)$$

Proposition 6.1 combined with Lemma 6.2 says that $f_J(y,z) \lesssim d^{-1}\exp\left(-\frac{y^2}{8d} - \frac{z^2}{8d}\right)$, so we are done. □

Proof of Proposition 6.1. The second statement follows from the first by integrating out one of the variables, so we prove only the first statement on the joint density bound. We are in the situation where $P \cap [-2d, 2d] \neq \emptyset$, and p is the unique element in this intersection. Let p^- and p^+ be the adjacent elements of P. Let $\mathcal{F}[p^-, p, p^+]$ denote the σ-algebra generated by $\mathcal{F}$ and the random variables $J(x)$ for $x \in \{p^-, p, p^+\}$. (These random variables provide extra information only when $P \cap [-2d, 2d] \neq \emptyset$.) The density $f_J(y,z)$ has a counterpart $f_J^{\mathcal{F}[p^-,p,p^+]}$ under the augmented σ-algebra, and it is enough to show that

$$f_J^{\mathcal{F}[p^-,p,p^+]}(y,z) \cdot \mathbb{1}_{P\cap[-2d,2d]\neq\emptyset} \lesssim \sigma_{-4d}^{-1} \cdot \sigma_{4d}^{-1} \cdot \exp\left(-\frac{y^2}{2\sigma_{-4d}^2} - \frac{z^2}{2\sigma_{4d}^2}\right),$$

since then Proposition 6.1(1) will arise by averaging.

Under the law $\mathbb{P}_{\mathcal{F}[p^-,p,p^+]}$, the processes $J(\,\cdot\,)$ on $[p^-, p]$ and $[p, p^+]$ are conditionally independent. Since the data in $\mathcal{F}[p^-, p, p^+]$ causes $\{P \cap [-2d, 2d] \neq \emptyset\}$ to occur, it is thus enough to argue that

- the conditional density of Y at $s \leq 0$ is at most a constant multiple of $\sigma_{-4d}^{-1}\exp\left(-s^2/2\sigma_{-4d}^2\right)$;
- and the conditional density of Z at $t \leq 0$ is at most $\sigma_{4d}^{-1}\exp\left(-t^2/2\sigma_{4d}^2\right)$.

These statements are straightforward to verify. Indeed, the conditional law under $\mathcal{F}[p^-, p, p^+]$ of $J(p-4d)$ is normal with mean $\left(1 - \frac{4d}{p-p^-}\right)J(p) + \frac{4d}{p-p^-}J(p^-)$ and variance σ_{-4d}^2. Note that $J(p^-) \geq \mathrm{Tent}(p^-)$ and $J(p) \geq \mathrm{Tent}(p)$ since $p^-, p \in P$, and that Tent is affine on the interval between consecutive pole set elements p^- and p; thus, we see that this mean is at least $\mathrm{Tent}(p - 4d)$. The first bullet point statement follows from the form of the normal density since we have shown that $\mathbb{E}_{\mathcal{F}[p^-,p,p^+]}[Y] \geq 0$, and we are concerned with the density only on $(-\infty, 0]$. The second bullet point is proved in the same fashion. This proves Proposition 6.1(1). □

6.2 The moderate case: Above the pole on both sides

In this section, we address the case of bounding $f_J(y,z)$ when $y + z > 0$. Here is the main proposition to be proved.

Proposition 6.4 (Density bound on increment). *Let $\mathfrak{l}/2 < x_1 < x_2 < \mathfrak{r}/2$ be $\mathcal{F}$-measurable, $\sigma^2 = x_2 - x_1 \geq d$, R be as in (5.19), and suppose that $\sigma \leq T^2 \cdot \sqrt{d/2}$. Suppose also that*

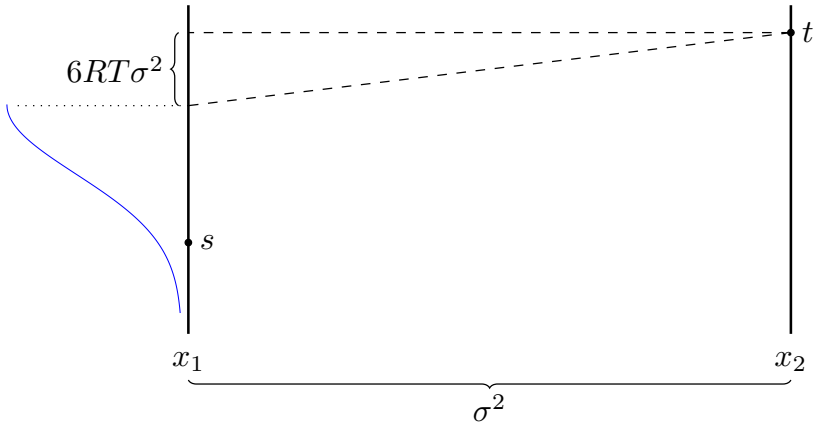


Figure 6.1: Illustrating Proposition 6.4 when $s < t$. The requirement that $|s - t| \geq 6RT\sigma^2$ comes from an error in an estimate of a certain slope; the blue curve on the left shows the Gaussian-like density centred at $t - 6RT\sigma^2$ that dominates the joint density $f_J^{x_1,x_2}(s,t)$. Note that the proposition does not require or use the presence of a pole in $[x_1, x_2]$.

$[x_1 - d, x_1 + d] \cap P = \emptyset$ *and* $[x_2 - d, x_2 + d] \cap P = \emptyset$. *Let* $f_J^{x_1,x_2}(s,t)$ *be the joint density of* $(J(x_1) - \mathrm{Tent}(x_1), J(x_2) - \mathrm{Tent}(x_2))$ *at* (s,t). *If* $s, t \in [-RT^2, RT^2]$ *and* $|s - t| > 6RT\sigma^2$, *then*

$$f_J^{x_1,x_2}(s,t) \cdot \mathbb{1}_{\mathsf{Fav}} \lesssim \sigma^{-1} \cdot d^{-\frac{1}{2}} T^2 \cdot \exp\left(-\frac{1}{2\sigma^2} \big(|s-t| - 6RT\sigma^2 \big)^2 + \frac{4R^2T^2}{\sigma^2} + 36R^2\sigma^2 \right).$$

We also have that $f_J^{x_1,x_2}(s,t) \lesssim d^{-1}$ *for all* $s, t \in \mathbb{R}$.

Remark 6.5. The condition that $|s - t| > 6RT\sigma^2$ arises in the proof of Proposition 6.4 from the error of an estimate on certain slopes; see Figure 6.1. For our purposes it does not cause any difficulty, as when $|y - z| < 6RT\sigma^2$ the costs will be always absorbable in the leeway factor. For example, this is seen in the proof of Lemma 6.6 below.

It may not be immediately clear what is the relation of this proposition to the case where $y+z > 0$. In fact, this proposition has been carefully stated to apply to a more general situation than just the case of this section. For example, unlike Proposition 6.1, this proposition does not require a pole to be present, and we will make use of it in the no-pole case addressed in Section 6.4 as well. We will also use Proposition 6.4 to prove Lemmas 5.14 and 5.15 near the end of the section, as it provides a density bound in terms of the *increment* $|y - z|$, and so can be easily used to bound the increment of J across an interval, a requirement which was briefly discussed in Section 5.2.

Before proving Proposition 6.4, we apply it to show that it yields the sufficient bound (5.25) in the case that $y + z > 0$. To see that Proposition 6.4 is sufficient for this purpose, note that by (5.22), when $y + z > 0$, V can be absorbed in the leeway factor, and so we essentially only need to consider S. This cost, being $d^{\frac{1}{2}} \exp((y+z)^2/16d)$ up to the leeway factor, is in essence met by the density bound provided by Proposition 6.4 by taking $x_1 = p - 4d$ and $x_2 = p + 4d$; this

is consistent with the ideas that Proposition 6.4 controls the increment and that the slope cost S is a cost associated with large increments.

Lemma 6.6. *When $y+z>0$ and $(y,z)\in\mathcal{G}_R^{(1)}$, we have (5.25) with $G_2'=22R^2D_k^{5/2}$ and $G_1'=G'T^2$, where G' is a constant independent of ε,k, and d.*

Proof. In this case, we see that

$$V\lesssim\begin{cases}1 & y+z>32Td\\ d^{\frac{1}{2}}\cdot\exp\left(64T^2d\right) & 0<y+z<32Td,\end{cases}$$

i.e., in this case V is bounded by the leeway factor $d^{\frac{1}{2}}\cdot\exp\left(G_2T^{5/6}\right)$ with $G_2=8$, since $d\leq\sqrt{T}/24$ by the assumption $\varepsilon\leq\exp(-(24)^6d^6/D_k^3)$ (note that $8<22R^2$ since $R\geq1$ from (5.19)). So we merely need to handle S, which from (5.21) is bounded above as

$$S\lesssim d^{\frac{1}{2}}\cdot\exp\left(\frac{1}{16d}(y-z)^2+10RT^{5/2}\right).$$

Note that if we set $x_1=p-4d$ and $x_2=p+4d$ in Proposition 6.4, then $\sigma^2=8d$. So with these parameters, we see from the second part of Proposition 6.4 that, when $|y-z|<48RTd$,

$$\begin{aligned}f_J(y,z)\cdot V\cdot S&\lesssim\exp\left(\frac{1}{16d}|y-z|^2+64T^2d+10RT^{5/2}\right)\\&\leq\exp\left((48\times3R^2+64)T^2d+10RT^{5/2}\right)\\&\leq\exp\left(19R^2T^{5/2}\right),\end{aligned}$$

the last inequality since $R\geq1$ and again using that $d\leq\sqrt{T}/24$; we have also used that $64/24\leq3$.

When $|y-z|>48RTd$, we again make use of Proposition 6.4 with $x_2=p+4d$ and $x_1=p-4d$. The assumptions of Proposition 6.4 are satisfied since $d_{\text{ip}}=5d$. So we obtain the bound

$$f_J(y,z)\cdot\mathbb{1}_{\mathsf{Fav}}\lesssim d^{-1}T^2\exp\left(-\frac{1}{16d}(|y-z|-48RTd)^2+\frac{4R^2T^2}{8d}+36R^2\times8d\right)$$

when $|y-z|>48RTd$. So, for $|y-z|>48RTd$,

$$\begin{aligned}&f_J(y,z)\cdot V\cdot S\\&\quad\lesssim T^2\cdot\exp\left(6RT|y-z|-48\times3R^2T^2d+\frac{1}{2}R^2T^2+36R^2\times8d+64T^2d+10RT^{5/2}\right)\\&\quad\leq T^2\cdot\exp\left(12R^2T^{5/2}+10RT^{5/2}\right)=T^2\cdot\exp\left(22R^2T^{5/2}\right),\end{aligned}$$

the last inequality since we have $|y-z|\leq2RT^{3/2}$ on $\mathcal{G}_R^{(1)}$ and using that $1\leq d\leq\sqrt{T}/24$ and $R\geq1$ from (5.19) to see that $36R^2\times8d\leq12R^2T^{1/2}$, allowing us to drop the middle four terms. This verifies (5.25) with the claimed values. □

We may now turn to discussing the proof strategy of Proposition 6.4. In the proof of Proposition 6.1, we were greatly aided by the presence of the pole and the *difficulty* that a Brownian bridge faces in making large jumps while remaining negative on either side. An interesting feature of that argument is that no extra reasoning was needed to obtain a density bound from a comparison with a Brownian object. Typically, such comparisons easily yield bounds on tail probabilities, but these do not immediately imply a pointwise density bound.

In the proof of Proposition 6.4, as well as in the case addressed in Section 6.3, both of these features will be missing. Firstly, when J (still thought of as essentially a Brownian bridge) is allowed to be positive on one side, the pole *assists* its attainment of the values on either side, and so it is not clear why the Brownian density should bound that of the jump curve. This is especially true for the case analysed in the next Section 6.3, where essentially the same bound as that proved in Section 6.1 must hold. Secondly, we will not be able to access density bounds directly, but will need to make further technical arguments to move from tail probability bounds to density bounds. To accomplish the latter, we will make use of a technique of *local randomization*, in both the case of this section as well as that of Section 6.3.

A short description of what we mean by local randomization is the following: Suppose that we wish to obtain a bound on the density of $J(x)$ for some x. We first obtain bounds on the tail probabilities of J at certain points, say $x-\eta$ and $x+\eta$ for some $\eta > 0$, with no pole contained in $[x-\eta, x+\eta]$. To convert this to a density bound at x, we use that J, conditionally on its values at $x-\eta$ and $x+\eta$, is a Brownian bridge on $[x-\eta, x+\eta]$. Then the distribution of J at x can be written as a convolution of the distributions at $x-\eta$ and $x+\eta$ with a normal random variable, which, when combined with the tail bounds, can be used to give a density bound. Heuristically, the tail bound is being propagated and smoothed by the Brownian bridge to a density bound.

At the level of this description, no importance is given to the exact value of η and we have not explained what we mean by "local" in local randomization. To aid our discussion, let us say that a random variable X has a *pseudo-variance at most* σ^2 if we have a tail bound of the form $\mathbb{P}(X < t) \leq \exp(-t^2/2\sigma^2)$. As the discussion of V and S in Section 5.2 showed, we require the pseudo-variance we obtain in sub-Gaussian density bounds to be essentially optimal. And indeed, if we knew that the distributions of $J(x-\eta)$ and $J(x+\eta)$ were actually Gaussian, then the density bound arising from the convolution mentioned would be precisely the correct one. However, when all we have is a sub-Gaussian tail bound and not an actual Gaussian distribution, there is some extra gain in the pseudo-variance we obtain for the final density bound. This is captured in the following lemma, whose proof will be given at the end of the section.

Lemma 6.7. *Let X be a random variable such that $\mathbb{P}(X < x) \leq A\exp(-\frac{1}{2\sigma_2^2}(x-x_0)^2)$ for $x < x_0$, and let N be a normal random variable with mean 0 and variance σ_1^2 which is independent of X. Then the density f of $X+N$ satisfies*

$$f(x) \leq \frac{A+1}{\sqrt{2\pi}\sigma_1} \cdot \exp\left(-\frac{(x-x_0)^2}{2(\sigma_1+\sigma_2)^2}\right) \tag{6.2}$$

for $x < x_0$, and is bounded by $1/\sqrt{2\pi}\sigma_1$ for all $x \in \mathbb{R}$.

The point we were expressing is seen in this formula by the fact that the pseudo-variance guaranteed by this bound is $(\sigma_1 + \sigma_2)^2$, which is greater than $\sigma_1^2 + \sigma_2^2$ as it would be had we known, in the notation of the lemma, that X has Gaussian distribution with variance σ_2^2.

For η a positive constant, this gain in pseudo-variance gives a density bound that is too weak for our purposes; indeed, the bound is weaker than that claimed in Proposition 6.4. The solution is, roughly, to take $\eta \to 0$. In the language of the lemma, if $\sigma_2(\eta) \to \sigma$ and $\sigma_1(\eta) \to 0$ as $\eta \to 0$, then

$$\lim_{\eta \to 0} (\sigma_1 + \sigma_2)^2 = \sigma^2 = \lim_{\eta \to 0} (\sigma_1^2 + \sigma_2^2)\,;$$

i.e., there is no gain in pseudo-variance in the limit. However, taking $\eta \to 0$ leads to a blow up in the constant in front of the exponential in (6.2), and so we actually take η to be a small ε-dependent quantity, small enough that the gain in pseudo-variance is manageable. This is the argument of Proposition 6.4, which we turn to next. We will then give the pending proofs of Lemmas 5.14 and 5.15, and finish the section by proving the technical tool Lemma 6.7.

The proof of Proposition 6.4 will actually obtain the claimed bound on the conditional density $f_J^{x_1,x_2}(s \mid t)$, so we also need that the marginal density $f_J^{x_2}(t)$ is bounded.

Lemma 6.8 (Marginal density is bounded)**.** *For any $\mathcal{F}$-measurable $x \in [-2T, 2T]$ such that $[x - d, x + d] \cap P = \emptyset$, let f_J^x be the density of $J(x) - \mathrm{Tent}(x)$ conditionally on $\mathcal{F}$. Then we have on* Fav *that $f_J^x(s) \le \pi^{-\frac{1}{2}} d^{-\frac{1}{2}}$ for all $s \in \mathbb{R}$.*

Proof. By assumption there is no pole in $[x - d, x + d]$. So $J(x)$, conditionally on $J(x - d)$ and $J(x + d)$, is given by

$$J(x) = \frac{1}{2} J(x - d) + \frac{1}{2} J(x + d) + N\left(0, \frac{1}{2} d\right).$$

Thus a formula for f_J^x is

$$f_J^x(s) = \frac{1}{\sqrt{\pi d}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{d}(t - s)^2\right) \mathrm{d}\nu(t),$$

where ν is the law of $\frac{1}{2} J(x - d) + \frac{1}{2} J(x + d)$. From this formula the claim follows. □

We next cite a standard bound on normal probabilities before turning to the proof of Proposition 6.4.

Lemma 6.9 (Normal bounds)**.** *Let $\sigma^2 > 0$. If $t > \sigma$ for the first inequality and $t > 0$ for the second,*

$$\frac{\sigma}{2\sqrt{2\pi} t} \exp\left(-\frac{t^2}{2\sigma^2}\right) \le \mathbb{P}\Big(N(0, \sigma^2) > t\Big) \le \exp\left(-\frac{t^2}{2\sigma^2}\right).$$

Proof. Replacing t by σt, it suffices to take $\sigma = 1$. The standard lower bound

$$\mathbb{P}(N(0,1) > t) \geq (2\pi)^{-1/2} \frac{t}{t^2+1} \exp\left(-t^2/2\right)$$

for $t \geq 0$ may be found in [Wil91, Section 14.8]. Note that $\frac{t}{t^2+1} \geq (2t)^{-1}$ for $t \geq 1$. The upper bound is simply the Chernoff bound. □

The proof of Proposition 6.4 has two steps, as described in the earlier discussion. The first is a tail bound on a quantity close to $J(x_1)$, conditionally on $J(x_2)$ (and with the roles of x_1 and x_2 reversed); the second is to convert this tail bound into a density bound using Lemma 6.7. The first step is isolated in the next lemma, while the second step is performed in the immediately following proof of Proposition 6.4.

Lemma 6.10. *Let $\mathfrak{l}/2 < x_1 < x_2 < \mathfrak{r}/2$ be $\mathcal{F}$-measurable, $\sigma^2 = x_2 - x_1 \geq d$, and R be as in (5.19). Suppose also that $[x_1 - d, x_1 + d] \cap P = [x_2 - d, x_2 + d] \cap P = \emptyset$. Then on* Fav, *for $r < t + \mathrm{Tent}(x_2) - 4\sigma^2(R+2)T$ and any $\eta < d/2$,*

$$\mathbb{P}_{\mathcal{F}}\Big(\tfrac{1}{2}J(x_1+2\eta) + \tfrac{1}{2}J(x_1-2\eta) < r \;\Big|\; J(x_2) + \mathrm{Tent}(x_2) = t\Big) \leq \exp\left(-\frac{1}{2\sigma_\eta^2}\big(r - t - \mathrm{Tent}(x_2) + 4\sigma^2(R+2)T\big)^2\right),$$

where $\sigma_\eta^2 = \sigma^2 \cdot \frac{x_1 - \mathfrak{l}}{x_2 - \mathfrak{l}} - \eta$.

Similarly, for $r < t + \mathrm{Tent}(x_1) - 4\sigma^2(R+2)T$ and on Fav,

$$\mathbb{P}_{\mathcal{F}}\Big(\tfrac{1}{2}J(x_2+2\eta) + \tfrac{1}{2}J(x_2-2\eta) < r \;\Big|\; J(x_1) + \mathrm{Tent}(x_1) = t\Big) \leq \exp\left(-\frac{1}{2\tilde{\sigma}_\eta^2}\big(r - t - \mathrm{Tent}(x_1) + 4\sigma^2(R+2)T\big)^2\right),$$

where $\tilde{\sigma}_\eta^2 = \sigma^2 \cdot \frac{\mathfrak{r} - x_2}{\mathfrak{r} - x_1} - \eta$.

Proof. We will only prove the first bound as the second bound is analogous, by repeating the below argument with the roles of x_1 and x_2 switched, and $\mathfrak{r}$ in place of $\mathfrak{l}$.

By assumption, there is no pole in $[x_1 - 2\eta, x_1 + 2\eta]$ for all $\eta \leq d/2$. For every such η, the distribution of $J(x_1)$ given $J(x_2)$, $J(x_1 - 2\eta)$, and $J(x_1 + 2\eta)$ depends on only $J(x_1 - 2\eta)$ and $J(x_1 + 2\eta)$, and is given by

$$J(x_1) = \frac{1}{2}J(x_1 - 2\eta) + \frac{1}{2}J(x_1 + 2\eta) + N(0, \eta). \tag{6.3}$$

Given $J(x_2) = t + \mathrm{Tent}(x_2)$ and on Fav, Lemma 5.12(ii) implies that J, restricted to $[\mathfrak{l}, x_2]$, stochastically dominates the Brownian bridge with endpoints $(\mathfrak{l}, -T^2)$ and $(x_2, t + \mathrm{Tent}(x_2))$. We call this Brownian bridge B. Then the slope of the line connecting these two points is

$$m := \frac{t + \mathrm{Tent}(x_2) + T^2}{x_2 - \mathfrak{l}},$$

and so $\mathbb{E}[B(x_1+r)] = t + \mathrm{Tent}(x_2) - (x_2 - x_1 - r)m$ for any r such that $x_1 + r \in [\mathfrak{l}, x_2]$. Now, conditionally on $J(x_2)$, we have a coupling such that

$$\frac{1}{2}J(x_1+2\eta) + \frac{1}{2}J(x_1-2\eta) \geq \frac{1}{2}B(x_1+2\eta) + \frac{1}{2}B(x_1-2\eta). \tag{6.4}$$

Since the covariance of B is given for $r_1 \leq r_2$ by

$$\mathrm{Cov}(B(r_1), B(r_2)) = \frac{(r_1 - \mathfrak{l})(x_2 - r_2)}{x_2 - \mathfrak{l}},$$

it follows after some algebraic simplification that the variance of the right hand side of (6.4) is σ_η^2. The mean of $\frac{1}{2}B(x_1+2\eta) + \frac{1}{2}B(x_1-2\eta)$ is

$$t + \mathrm{Tent}(x_2) - \tfrac{1}{2}(x_2 - x_1 - 2\eta)m - \tfrac{1}{2}(x_2 - x_1 + 2\eta)m = t + \mathrm{Tent}(x_2) - \sigma^2 m.$$

Thus we have that, for $r < t + \mathrm{Tent}(x_2) - \sigma^2 m$, on $\mathsf{Fav} \cap \{[x_1 - d, x_1 + d] \cap P = \emptyset\}$,

$$\begin{aligned}
&\mathbb{P}_{\mathcal{F}}\Big(\tfrac{1}{2}J(x_1+2\eta) + \tfrac{1}{2}J(x_1-2\eta) < r \;\Big|\; J(x_2) + \mathrm{Tent}(x_2) = t\Big) \\
&\qquad \leq \mathbb{P}_{\mathcal{F}}\Big(N\big(t + \mathrm{Tent}(x_2) - \sigma^2 m, \sigma_\eta^2\big) < r\Big) \\
&\qquad \leq \exp\left(-\frac{1}{2\sigma_\eta^2}\big(r - t - \mathrm{Tent}(x_2) + \sigma^2 m\big)^2\right),
\end{aligned} \tag{6.5}$$

the last inequality obtained for $r < t + \mathrm{Tent}(x_2) - \sigma^2 m$ via the upper bound from Lemma 6.9.

Now returning to the definition of m, on Fav,

$$m = \frac{t + \mathrm{Tent}(x_2) + T^2}{x_2 - \mathfrak{l}} \leq \frac{(R+2)T^2}{T/4} = 4\,(R+2)\,T,$$

since we have assumed that $t \leq RT^2$; that $x_2 \geq \mathfrak{l}/2$; and since, on Fav, $\mathfrak{l} \leq -T/2$ and $\mathrm{Tent}(x_2) \leq T^2$.

Using this bound on m in (6.5) completes the proof of Lemma 6.10. □

Proof of Proposition 6.4. We prove only the case of $s < t$; the other case is analogous, making use of the second inequality of Lemma 6.10 instead of the first as we do in the case of $s < t$.

We first note that

$$\sigma_\eta^2 = \frac{(x_2 - x_1)(x_1 - \mathfrak{l})}{x_2 - \mathfrak{l}} - \eta \leq \sigma^2 - \eta. \tag{6.6}$$

We will apply Lemma 6.7 to Lemma 6.10 using (6.3). The parameters of Lemma 6.7 are set as follows (the formal notational conflict between σ_1 or σ_2 and σ_η should not cause confusion):

$X = \frac{1}{2}J(x_1-2\eta)+\frac{1}{2}J(x_1+2\eta)-\mathrm{Tent}(x_1)$, $\sigma_1^2 = \eta$, $\sigma_2^2 = \sigma_\eta^2$, $x_0 = t+\mathrm{Tent}(x_2)-\mathrm{Tent}(x_1)-\sigma^2 m$, and A specified by the constant represented by $\lesssim$ in the first inequality of Lemma 6.10.

This yields, on the event $\mathsf{Fav}\cap\{[x_1-d, x_1+d]\cap P = \emptyset\}$, for each $\eta < d/2$, the following bound on the conditional density of $J(x_1) - \mathrm{Tent}(x_1)$, conditionally on $J(x_2) - \mathrm{Tent}(x_2) = t$:

$$\begin{aligned} f_J^{x_1,x_2}(s \mid t) &\lesssim \eta^{-\frac{1}{2}} \cdot \exp\left(-\frac{1}{2\left(\sigma_\eta + \eta^{1/2}\right)^2}\left(s - t + \mathrm{Tent}(x_1) - \mathrm{Tent}(x_2) + 4\sigma^2(R+2)T\right)^2\right) \\ &\leq \eta^{-\frac{1}{2}} \cdot \exp\left(-\frac{1}{2\left(\sigma^2 + 2\eta^{1/2}\sigma\right)}\left(s - t + \mathrm{Tent}(x_1) - \mathrm{Tent}(x_2) + 4\sigma^2(R+2)T\right)^2\right) \end{aligned}$$

for $s < t + \mathrm{Tent}(x_2) - \mathrm{Tent}(x_1) - 4\sigma^2(R+2)T$. We used (6.6) when expanding the square in the denominator of the exponent in the last inequality. Note that

$$\left|\mathrm{Tent}(x_1) - \mathrm{Tent}(x_2)\right| \leq 4T(x_2 - x_1) = 4T\sigma^2.$$

Using the previous equation, we obtain, on $\mathsf{Fav}\cap\{[x_1-d, x_1+d]\cap P = \emptyset\}$, that

$$f_J^{x_1,x_2}(s \mid t) \lesssim \eta^{-\frac{1}{2}} \cdot \exp\left(-\frac{1}{2\left(\sigma^2 + 2\eta^{1/2}\sigma\right)}\left(s - t + MT\sigma^2\right)^2\right),$$

for $s \leq t - MT\sigma^2$, where $M = 4 + 4(R+2) = 4(R+3)$. Now using the inequality $(1+x)^{-1} \geq 1 - x$ for $x = 2\eta^{1/2}\sigma^{-1}$, we find

$$f_J^{x_1,x_2}(s \mid t) \lesssim \eta^{-\frac{1}{2}} \cdot \exp\left(-\frac{1}{2\sigma^2}\left(s - t + MT\sigma^2\right)^2 + \frac{\eta^{1/2}}{\sigma^3}\left(s - t + MT\sigma^2\right)^2\right). \tag{6.7}$$

Let us focus on bounding the second term in the exponent. We expand the square and drop the cross-term, since $s - t \leq 0$, to get that the second term is bounded above by

$$\frac{\eta^{1/2}}{\sigma^3}\left((s-t)^2 + M^2T^2\sigma^4\right) \leq \frac{\eta^{1/2}}{\sigma^3}\left(4R^2T^4 + M^2T^2\sigma^4\right),$$

the last inequality since $s - t \in [-2RT^2, 0]$. We now use this bound in (6.7) and set $\eta^{1/2} = T^{-2}\sigma$ (which satisfies $\eta < d/2$ by assumption), to obtain

$$f_J^{x_1,x_2}(s \mid t) \lesssim \frac{T^2}{\sigma}\exp\left(-\frac{1}{2\sigma^2}(s - t + MT\sigma^2)^2 + \frac{4R^2T^2}{\sigma^2} + M^2\sigma^2\right).$$

The argument is complete by noting that $f_J^{x_2}(t) \lesssim d^{-\frac{1}{2}}$ by Lemma 6.8 and that $M \leq 6R$, since $M = 4R + 12 \leq 4R + 2 \times 6\sqrt{d} = 6R$ from (5.19) and $d \geq 1$.

The final statement in Proposition 6.4 of a constant bound on $f_J^{x_1,x_2}(s,t)$ for all values of s and t follows immediately from (6.3) and the latter assertion of Lemma 6.7 with the parameters $\sigma_1^2 = \eta = d/4$, and again using that the marginal density satisfies $f_J^{x_2}(t) \lesssim d^{-\frac{1}{2}}$ from Lemma 6.8. □

We will now move towards the proofs of Lemmas 5.14 and 5.15, which use Propositions 6.1 and 6.4. Then we will conclude this section and this part of the argument by proving Lemma 6.7.

For the proofs of Lemmas 5.14 and 5.15 we will need two further statements, the first bounding certain Gaussian integrals, and the other a standard tail bound on the supremum of a Brownian bridge. These are the next two lemmas. We will make frequent use of Lemma 6.11 in the next section as well.

Lemma 6.11. *For $a > 0$,*

$$\int_0^\infty \exp\left(-ax^2 + bx\right)\,\mathrm{d}x \lesssim \begin{cases} a^{-\frac{1}{2}} & b \le 0 \\ a^{-\frac{1}{2}} \exp\left(\frac{b^2}{4a}\right) & b \in \mathbb{R}. \end{cases}$$

Proof. Completing the square, we find

$$\begin{aligned} \int_0^\infty \exp\left(-ax^2 + bx\right)\,\mathrm{d}x &= \int_0^\infty \exp\left(-a\left(x^2 - \frac{bx}{a} + \frac{b^2}{4a^2}\right) + \frac{b^2}{4a}\right)\,\mathrm{d}x \\ &= \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2}{4a}\right) \mathbb{P}\left(N\left(\frac{b}{2a}, \frac{1}{2a}\right) > 0\right) \\ &= \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2}{4a}\right) \mathbb{P}\left(N(0,1) > -\frac{1}{\sqrt{2a}} b\right). \end{aligned}$$

For all $b \in \mathbb{R}$ this probability factor is bounded by a constant, which yields Case 2 of the statement. If $b < 0$ we may use the Chernoff bound for normal random variables to obtain Case 1:

$$\sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2}{4a}\right) \mathbb{P}\left(N(0,1) > -\frac{1}{\sqrt{2a}} b\right) \le \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2}{4a} - \frac{1}{2} \cdot \frac{b^2}{2a}\right) = \sqrt{\frac{\pi}{a}}. \qquad \square$$

Lemma 6.12. *Let B be a Brownian bridge of length T from $(0,0)$ to $(T,0)$. Then we have*

$$\mathbb{P}\left(\sup_{[0,T]} B(x) \ge r\right) = \mathbb{P}\left(\inf_{[0,T]} B(x) \le -r\right) = e^{-2r^2/T}.$$

Proof. The equality of the two quantities follows from Brownian symmetry. By Brownian scaling the statement reduces to when $T = 1$, which is given by equation (3.40) in [KS98, Chapter 4]. $\square$

Proof of Lemma 5.15. By Lemma 5.12(iii) we have that J is stochastically dominated by the restriction to $[\mathfrak{l}, \mathfrak{r}]$ of a Brownian bridge from $(-2T, 2T^2)$ to $(2T, 2T^2)$ conditioned to jump over all the poles as well as $\mathrm{Corner}_k^{\mathfrak{l},\mathcal{F}}$ and $\mathrm{Corner}_k^{\mathfrak{r},\mathcal{F}}$. This event being conditioned on has a constant probability since, on Fav, the value of Tent at the poles, $\mathrm{Corner}_k^{\mathfrak{l},\mathcal{F}}$, and $\mathrm{Corner}_k^{\mathfrak{r},\mathcal{F}}$ are all below T^2 and $T \ge 1$. We also have on Fav that $-T^2 \le \mathrm{Tent} \le T^2$. Thus, for both the pole and no-pole cases of the definition of (Y, Z) from (5.13) and (5.18), on Fav,

$$\mathbb{P}_{\mathcal{F}}\Big(\max\{Y, Z\} > RT^2\Big) \lesssim \mathcal{B}^{[-2T,2T]}_{2T^2,2T^2}\left(\sup_{t\in[-2T,2T]} B(t) > (R-1)T^2\right)$$

$$= \exp\left(-\frac{1}{2}(R-3)^2T^3\right) = \varepsilon^{(R-3)^2D_k^3/2},$$

using Lemma 6.12 in the second inequality. Similarly for the lower side, we have by Lemma 5.12(i) that, on Fav, J stochastically dominates a Brownian bridge from $(\mathfrak{l}, -T^2)$ to $(\mathfrak{r}, -T^2)$. Thus again using Lemma 6.12 and that $-T^2 \le \mathrm{Tent} \le T^2$, and for both cases of the definition of (Y, Z),

$$\begin{aligned}\mathbb{P}_{\mathcal{F}}\Big(\min\{Y,Z\} < -RT^2\Big)\cdot \mathbb{1}_{\mathsf{Fav}} &\le \mathcal{B}^{[\mathfrak{l},\mathfrak{r}]}_{-T^2,-T^2}\left(\inf_{t\in[\mathfrak{l},\mathfrak{r}]} B(t) < -(R-1)T^2\right)\\ &= \exp\left(-\frac{2}{(\mathfrak{r}-\mathfrak{l})}(R-2)^2T^4\right) \le \varepsilon^{(R-2)^2D_k^3},\end{aligned}$$

since $|\mathfrak{r}|, |\mathfrak{l}| \le T$. We note that for our range of R, $\varepsilon^{(R-2)^2D_k^3} \le \varepsilon^{(R-3)^2D_k^3/2}$.

We are left with bounding

$$\mathbb{P}_{\mathcal{F}}\left(|Y-Z| > 2RT^{3/2}, -RT^2 < Y, Z < RT^2\right)$$

on Fav. Since the definition of Y and Z depends on whether $P \cap [-2d, 2d] = \emptyset$ or not (see (5.13) and (5.18)), the bound we can obtain on the above probability depends on the same as well. However, the bound in the case where a pole is present (which is the one claimed in the statement of Lemma 5.15) actually holds for both cases; this is because the distance between the points where Y and Z measure the deviation of J from Tent is $8d$ when a pole is present, larger than the $2d$ it is when the pole is absent in $[-2d, 2d]$. So we will present the case where $P \cap [-2d, 2d] \neq \emptyset$, but exactly the same argument works in the other case as well, where it yields a slightly stronger bound corresponding to $2d$ in place of $8d$. Let us define

$$\mathcal{G} = \left\{(y,z) : -RT^2 < y, z < RT^2, |y-z| > 2RT^{3/2}\right\},$$

so that

$$\mathbb{P}_{\mathcal{F}}\left(|Y-Z| > 2RT^{3/2}, -RT^2 < Y, Z < RT^2\right) = \iint_{\mathcal{G}} f_J(y,z)\,\mathrm{d}y\,\mathrm{d}z \cdot \mathbb{1}_{\mathsf{Fav}}.$$

To bound this integral, we will use Proposition 6.4 to bound the density and make the change of variables $(u, v) = (y - z, y)$. Note that the range of y and z satisfies the hypotheses of Proposition 6.4 with $x_1 = p - 4d, x_2 = p + 4d$, so that $\sigma^2 = 8d$. These parameter choices satisfy the hypotheses of Proposition 6.4 since $d_{\mathrm{ip}} = 5d$; in particular, $|y - z| \ge 48RTd$ for all $(y, z) \in \mathcal{G}$ as $2RT^{3/2} \ge 48RTd$ since $d \le \sqrt{T}/24$. Note that $36\sigma^2 = 36 \times 8d = 288d$. So, on Fav,

$$\begin{aligned}&\iint_{\mathcal{G}} f_J(y,z)\,\mathrm{d}y\,\mathrm{d}z\\ &\lesssim d^{-1}T^2 \iint_{\mathcal{G}} \exp\left(-\frac{1}{16d}(|y-z| - 48RTd)^2 + \frac{R^2T^2}{2d} + 288R^2d\right)\mathrm{d}y\,\mathrm{d}z\end{aligned}$$

$$
\begin{aligned}
&\leq 2d^{-1}T^2 \int_{-RT^2}^{RT^2} \int_{2RT^{3/2}}^{\infty} \exp\left(-\frac{1}{16d}(u-48RTd)^2 + \frac{R^2T^2}{2d} + 288R^2d\right) \mathrm{d}u\,\mathrm{d}v \\
&= 2d^{-1}T^2 \int_{-RT^2}^{RT^2} \int_{0}^{\infty} \exp\left(-\frac{1}{16d}(u+2RT^{3/2}-48RTd)^2 + \frac{R^2T^2}{2d} + 288R^2d\right) \mathrm{d}u\,\mathrm{d}v \\
&\lesssim RT^4d^{-1} \exp\left(\frac{R^2T^2}{2d} + 288R^2d\right) \\
&\times \int_0^{\infty} \exp\left\{-\frac{1}{16d}\left(u^2 + 2u(2RT^{3/2}-48RTd) + 4R^2T^3 - 192R^2T^{5/2}d + 48^2R^2T^2d^2\right)\right\} \mathrm{d}u \\
&\lesssim RT^4d^{-\frac{1}{2}} \exp\left(-\frac{1}{4d}R^2T^3 + 12R^2T^{5/2} - 48\times 3R^2T^2d + \frac{R^2T^2}{2d} + 288R^2d\right) \\
&\leq RT^4d^{-\frac{1}{2}} \exp\left(12R^2T^{5/2}\right) \varepsilon^{R^2D_k^3/4d}.
\end{aligned}
$$

We have used Case 1 of Lemma 6.11 with $a = 1/(16d)$ and $b = -2(2RT^{3/2} - 48RTd)$ for the integral in the second-to-last line, since $d \leq \sqrt{T}/24$ implies that $2RT^{3/2} - 48RTd \geq 0$, and thus that $b \leq 0$. In the last line, since $1 \leq d \leq \sqrt{T}/24$, we see that $288R^2d \leq 12R^2T^{1/2}$, and thus the sum of the last three terms in the exponent of the penultimate line is negative and may be dropped. Finally, since $x \leq \exp(x^{5/8})$ for $x \geq 1$, and since $1 \leq d \leq \sqrt{T}/24$ and $R \geq 1$ from (5.19), we have that $RT^4d^{-\frac{1}{2}} \leq \exp(R^2T^{5/2})$. This completes the proof of Lemma 5.15. □

Proof of Lemma 5.14. Using Lemma 5.15, it is enough to show that, on Fav and when $P \cap [-2d, 2d] \neq \emptyset$, $\mathbb{P}\left(Y < -T^{3/2}\right) + \mathbb{P}_{\mathcal{F}}\left(Z < -T^{3/2}\right)$ is bounded by the right-hand side in the Lemma 5.14's statement.

Since we are considering the situation where Y and Z are negative, we may use Proposition 6.1. So from Proposition 6.1(2) and Lemma 6.2 we have that, on $\mathsf{Fav} \cap \{P \cap [-2d, 2d] \neq \emptyset\}$,

$$
\mathbb{P}_{\mathcal{F}}\big(Y < -RT^{3/2}\big) \lesssim d^{-\frac{1}{2}} \int_{RT^{3/2}}^{\infty} \exp\left(-\frac{y^2}{8d}\right) \mathrm{d}y \lesssim \exp\left(-\frac{1}{8d}R^2T^3\right) = \varepsilon^{R^2D_k^3/8d},
$$

where we have performed the change of variables $y \mapsto y + RT^{3/2}$ and applied Case 1 of Lemma 6.11 with $a = 1/(8d)$ and $b = -RT^{3/2}/(4d)$ in the second inequality. Similarly, we have

$$
\mathbb{P}_{\mathcal{F}}\big(J(p+4d) - \mathrm{Tent}(p+4d) < -RT^{3/2}\big) \lesssim \varepsilon^{R^2D_k^3/8d}. \qquad \square
$$

We conclude the section by providing the proof of the technical tool Lemma 6.7.

Proof of Lemma 6.7. Let ν be the law of X. For $x < x_0$ and $0 < \delta < 1$, let $\tilde{x} = \delta(x_0 - x)$. Since X and N are independent, we have

$$
f(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \int_{-\infty}^{\infty} e^{-(x-y)^2/2\sigma_1^2} \mathrm{d}\nu(y) = \frac{1}{\sqrt{2\pi}\sigma_1} \left[\int_{[x-\tilde{x},x+\tilde{x}]} + \int_{[x-\tilde{x},x+\tilde{x}]^c} e^{-(x-y)^2/2\sigma_1^2} \mathrm{d}\nu(y)\right].
$$

From the first equality we see that the density is bounded by $(\sqrt{2\pi}\sigma_1)^{-1}$ for all x. For the stronger bound for small enough x, note using the hypothesis on ν that the first integral in the right-hand side is bounded by

$$\int_{[x-\tilde{x},x+\tilde{x}]} \mathrm{d}\nu(y) \le \nu\big((-\infty, x+\tilde{x})\big) \le A \cdot \exp\left(-\frac{(x+\tilde{x}-x_0)^2}{2\sigma_2^2}\right) = A \cdot \exp\left(-\frac{(1-\delta)^2(x-x_0)^2}{2\sigma_2^2}\right),$$

where we have used that $x+\tilde{x}$ is less than x_0; this is due to $\delta < 1$ and $x < x_0$.

The second integral is bounded by $\exp\left(-\frac{\tilde{x}^2}{2\sigma_1^2}\right) = \exp\left(-\frac{\delta^2(x-x_0)^2}{2\sigma_1^2}\right)$. These inequalities hold for all $0 < \delta < 1$, and so if we set $\delta = \sigma_1/(\sigma_1+\sigma_2)$, we obtain our result. □

6.3 The difficult case: Above and below the pole on either side

At this stage we have proved the required bound on $f_J(y,z)$ in the two cases where $y, z < 0$ or $y + z > 0$. This leaves the case where $y < 0$, $z > 0$, and $y + z < 0$ (the case where $y > 0$ and $z < 0$ is clearly symmetric). Perhaps surprisingly, this turns out to be the most difficult case. However, we now give a heuristic reason why we should expect the density $f_J(y,z)$ to be largest in this case, as a proxy for why this case is most difficult.

Recall that $f_J(y,z)$ is the density of (Y,Z), which, from (5.13), are respectively the deviations of J from Tent at $p-4d$ and $p+4d$. So, the size of the density $f_J(y,z)$ essentially represents a comparison of the probability that J takes the values $y + \mathrm{Tent}(p-4d)$ and $z + \mathrm{Tent}(p+4d)$ respectively at $p-4d$ and $p+4d$ to the probability of the same for Brownian motion. A larger value of this density is obtained if J finds it easier to adopt the specified values than a Brownian motion does. This is precisely what happens when $y < 0$, $z > 0$, and $y + z < 0$, as J has the pole at p which pushes it up and helps it attain the value of z at $p+4d$; a Brownian motion has no such assistance. Thus the density should be highest for this case. Considering the situation in the two cases we have already analysed in Sections 6.1 and 6.2 should convince the reader that in Section 6.1 the pole actually makes J's task more difficult than B's, which has no pole, while in Section 6.2, the pole has essentially no effect.

In this section, since $y + z < 0$, the vault cost V cannot be ignored. Thus we need a stronger bound on $f_J(y,z)$ than was required in Section 6.2; in fact, we need a bound of the same basic form as that proved in Section 6.1. This is why the previous paragraph's conclusion that the density is highest in this case indicates that the required argument will be more delicate.

Let $f_J(z \mid y)$ be the conditional density of Z at z given $Y = y$. Our aim will be the following proposition.

Proposition 6.13. *Let R be as in (5.19). If $40Td \le z \le 2RT^{3/2}$ and $y + z < 0$,*

$$f_J(z \mid y) \cdot \mathbb{1}_{\mathsf{Fav},P\cap[-2d,2d]\neq\mathcal{O}} \lesssim d^{-\frac{1}{2}} \cdot \exp\left(-\frac{z^2}{8d} + 20RT^{5/2}\right).$$

Further, $f_J(z \mid y) \lesssim d^{-\frac{1}{2}}$ *for all* $y, z \in \mathbb{R}$ *with* $y + z < 0$.

First we show that Proposition 6.13 implies the sufficient bound (5.25) when $y < 0, z > 0$, and $y + z < 0$ (and symmetrically when $y > 0$ and $z < 0$).

Lemma 6.14. *When* $yz < 0$, $y + z < 0$, *and* $(y, z) \in \mathcal{G}_R^{(1)}$, *we have* (5.25) *with* $G_2' = 41RD_k^{5/2}$ *and* $G_1' = G'$, *where* G' *is a constant independent of* ε, k, *and* d.

Proof. We give the proof for when $y < 0$ and $z > 0$. From (5.21) and (5.23), we see that in this case

$$V \cdot S \lesssim d \cdot \exp\left(\frac{y^2}{8d} + \frac{z^2}{8d} + 21RT^{5/2}\right). \tag{6.8}$$

So it suffices to prove, for some $G < \infty$,

$$f_J(y, z) \lesssim d^{-1} \cdot \exp\left(-\frac{y^2}{8d} - \frac{z^2}{8d} + GT^{5/2}\right).$$

There are two cases to consider. If $z \leq 40Td$, then, from (6.8), $V \cdot S$ is bounded by

$$d \cdot \exp\left(\frac{y^2}{8d} + \frac{40^2T^2d^2}{8d} + 21RT^{5/2}\right) \leq d \cdot \exp\left(\frac{y^2}{8d} + 30RT^{5/2}\right)$$

since $d \leq \sqrt{T}/24$ and $40^2/(8 \times 24) \leq 9$. Noting that $30 < 41$, it suffices to prove

$$f_J(y, z) \lesssim d^{-1} \cdot \exp\left(-\frac{y^2}{8d}\right).$$

This is provided by Proposition 6.1 and Lemma 6.2 since $y < 0$, and by Proposition 6.13's latter statement that $f_J(z \mid y) \lesssim d^{-\frac{1}{2}}$.

Now suppose $z > 40Td$. Note that we also have $z \leq 2RT^{3/2}$, since $y < 0$ and $(y, z) \in \mathcal{G}_R^{(1)}$ implies that $|y - z| \leq 2RT^{3/2}$. So from Proposition 6.1, Lemma 6.2, and Proposition 6.13, we obtain

$$f_J(y, z) \lesssim d^{-1} \cdot \exp\left(-\frac{y^2}{8d} - \frac{z^2}{8d}\right) \cdot \exp\left(20RT^{5/2}\right).$$

This completes the proof, after taking into account the extra factor of $\exp\left(21RT^{5/2}\right)$ which arises from the expressions for V and S as in (6.8). □

We next turn to discussing the proof ideas of Proposition 6.13. The claim of the proposition may be a surprising one at first glance, for, in a slight abuse of the language of pseudo-variance used in Section 6.2, Proposition 6.13 says that $J(p+4d) - \mathrm{Tent}(p+4d)$, conditionally on $J(p-4d) - \mathrm{Tent}(p-4d)$ being negative, has pseudo-variance at most $4d$; in contrast, a Brownian motion, conditionally on its value at $p - 4d$, would have a much higher variance of $8d$ at the position

$p + 4d$. So we must crucially use both that $J(p - 4d) - \mathrm{Tent}(p - 4d) < 0$ (i.e., $y < 0$) and that J must satisfy $J(p) \geq \mathrm{Tent}(p)$.

Heuristically, because J is jumping over $\mathrm{Tent}(p)$ from a negative value at $p - 4d$, it will make the jump with a very low margin. Thus the variance at p is not $4d$ as it would be for a Brownian motion, but essentially 0. This explains how we can get a pseudo-variance of at most $4d$ at $p+4d$ for J. This intuition is captured in Lemma 6.16, which says that we may safely restrict our analysis to the case where J jumps over the pole at p by at most 1. To prove Lemma 6.16, we need a technical lemma about the monotonicity of conditional probabilities of Gaussians. The result is identical to [Ham19a, Lemma 2.21], but we include its short proof here for completeness.

Lemma 6.15. *Fix $r > 0$, $m \in \mathbb{R}$, and $\sigma^2 > 0$, and let X be distributed as $N(m, \sigma^2)$. Then the quantity $\mathbb{P}(X \geq r + s \mid X \geq s)$ is a strictly decreasing function of $s \in \mathbb{R}$.*

Proof. Note that

$$\log \mathbb{P}(X \geq s + r \mid X \geq s) = \log \int_{s+r}^{\infty} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} \mathrm{d}x - \log \int_{s}^{\infty} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} \mathrm{d}x$$

has derivative in s given by

$$\frac{\exp\left\{-\frac{(s-m)^2}{2\sigma^2}\right\} \int_{s+r}^{\infty} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} \mathrm{d}x - \exp\left\{-\frac{(s+r-m)^2}{2\sigma^2}\right\} \int_{s}^{\infty} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} \mathrm{d}x}{\int_{s+r}^{\infty} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} \mathrm{d}x \cdot \int_{s}^{\infty} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} \mathrm{d}x}.$$

The denominator is clearly positive. Performing the change of variable $x \mapsto x + r$ in the first integral of the numerator and manipulating the exponents show that the numerator equals

$$\int_{s}^{\infty} \exp\left\{-\frac{(x-m)^2+(s-m)^2+r^2}{2\sigma^2}\right\} \left(\exp\left\{-\frac{(x-m)r}{\sigma^2}\right\} - \exp\left\{-\frac{(s-m)r}{\sigma^2}\right\}\right) \mathrm{d}x\,.$$

The proof is complete by noting that this integrand is strictly negative for all $x > s$. □

In order to state Lemma 6.16, we define the random variable U to be the deviation of the jump ensemble from the Tent map at the pole p. For later use, we also take this opportunity to define W_η to be the same at $p + 4d + \eta$ for $\eta < d$. So, we define

$$\begin{aligned} U &:= J(p) - \mathrm{Tent}(p) \\ W_\eta &:= J(p + 4d + \eta) - \mathrm{Tent}(p + 4d + \eta). \end{aligned} \tag{6.9}$$

The parameter η, as in the previous section, will be set to a specific small value in a local randomization argument later. Recall also from (5.13) that Y and Z are respectively the deviation of J from Tent at $p - 4d$ and $p + 4d$. See Figure 6.2. We now turn to our assertion that J typically makes a narrow jump over p.

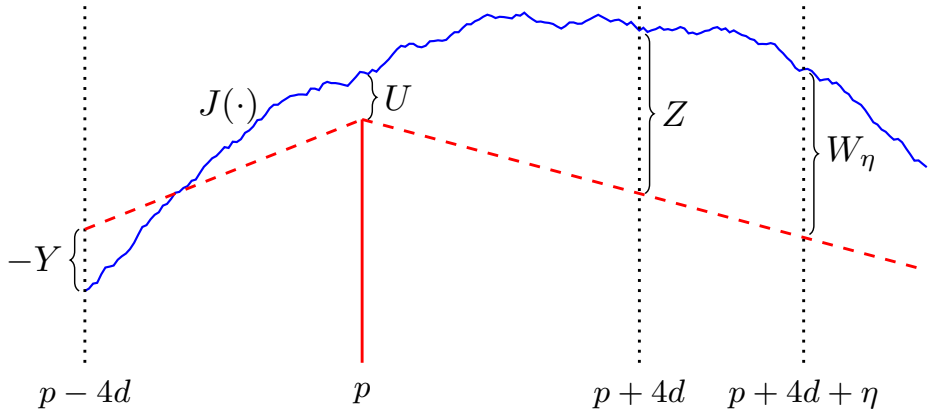


Figure 6.2: Illustrating the definitions of U, W, Y and Z in the subcase being addressed in this section; Y is negative, and so the length being shown is $-Y$. The blue curve is J, while the red dashed function is Tent. The red vertical line emphasises the height of the pole at p.

Lemma 6.16 (Narrow jump over p)**.** *Let $f_J(z,u \mid y)$ be the joint conditional density of Z and U given Y. For $y, z \in \mathbb{R}$ such that $y+z<0$,*

$$f_J(z \mid y) \lesssim d^{\frac{1}{2}} \int_0^1 f_J(z,u \mid y)\,\mathrm{d}u.$$

Proof. We have

$$\int_0^1 f_J(z,u \mid y)\,\mathrm{d}u = f_J(z \mid y) \int_0^1 f_J(u \mid y,z)\,\mathrm{d}u. \tag{6.10}$$

Let X be distributed as $N(0,2d)$, which is the distribution of a Brownian bridge of duration $8d$ from 0 to 0 at its midpoint, and let $\Delta = \frac{1}{2}(\mathrm{Tent}(p-4d)+\mathrm{Tent}(p+4d)) - \mathrm{Tent}(p)$. Note that $\Delta \le 0$ by the concavity of the Tent map. Then the second factor on the right hand side is

$$\begin{aligned}\mathbb{P}_{\mathcal{F}}\big(U \in [0,1] \,\big|\, Y=y, Z=z\big) &= \mathbb{P}\left(X + \frac{y+z}{2} + \Delta \in [0,1] \,\middle|\, X + \frac{y+z}{2} + \Delta \ge 0\right) \\ &= 1 - \mathbb{P}\left(X + \frac{y+z}{2} + \Delta \ge 1 \,\middle|\, X + \frac{y+z}{2} + \Delta \ge 0\right).\end{aligned}$$

Now by Lemma 6.15 with $r=1$ and $s = -(y+z)/2 - \Delta$, we have that

$$\mathbb{P}\left(X \ge 1 - \frac{y+z}{2} - \Delta \,\middle|\, X \ge -\frac{y+z}{2} - \Delta\right)$$

is an increasing function of $y+z$, i.e., $\mathbb{P}_{\mathcal{F}}(U \in [0,1] \mid Y=y, Z=z)$ is a decreasing function of $y+z$. So for $y+z<0$, we obtain that

$$\mathbb{P}_{\mathcal{F}}\big(U \in [0,1] \mid Y=y, Z=z\big) \ge \mathbb{P}\big(X \in [0,1] \mid X \ge 0\big) = \mathbb{P}\big(|X| \le 1\big);$$

the second quantity is the value of the first at $y + z = -2\Delta \geq 0$. An easy bound using the standard normal density gives that $\mathbb{P}(|X| \leq 1)^{-1} \lesssim d^{\frac{1}{2}}$ for $d \geq 1$; tracing back the relations and using this last bound in (6.10) yields

$$f_J(z \mid y) \lesssim d^{\frac{1}{2}} \int_0^1 f_J(z, u \mid y)\, \mathrm{d}u,$$

completing the proof of Lemma 6.16. □

In the previous case in Section 6.1 we made use of the tool Lemma 6.7 to convert a tail bound to a density bound, and we will make use of essentially the same tool in this section; it is simply restated in a form involving the upper tail.

Lemma 6.17. *Let X be a random variable such that $\mathbb{P}(X > x) \leq A \exp(-\frac{1}{2\sigma_2^2}(x - x_0)^2)$ for $x > x_0$, and let N be a normal random variable with mean 0 and variance σ_1^2 which is independent of X. Then the density f of $X + N$ satisfies*

$$f(x) \leq \frac{A+1}{\sqrt{2\pi}\sigma_1} \cdot \exp\left(-\frac{(x - x_0)^2}{2(\sigma_1 + \sigma_2)^2}\right)$$

for $x > x_0$, and is bounded by $1/\sqrt{2\pi}\sigma_1$ for all $x \in \mathbb{R}$.

Proof. This follows by applying Lemma 6.7 to $\tilde{X} := -X$, since mean zero normal distributions are symmetric, and where X is as in Lemma 6.17. □

The broad idea of the proof of Proposition 6.13 is to write the distribution of Z, conditionally on $U = J(p) - \mathrm{Tent}(p)$ and $W_\eta = J(p + 4d + \eta) - \mathrm{Tent}(p + 4d + \eta)$, in terms of a Brownian bridge using the definition of J, and then use this Brownian structure to obtain a bound on the conditional density of Z given Y via Lemma 6.17 and Lemma 6.16. This is the same strategy of local randomization that was used in Section 6.2 to prove Proposition 6.4.

So, as before, we will need tail probabilities on the distribution of W_η given U and Y. By the Markov property of Brownian bridges, W_η is conditionally independent of Y given U, and so in fact we need a tail probability for W_η given only U. Such a tail bound is the content of Lemma 6.18, whose argument is essentially again stochastically dominating J by a Brownian bridge.

Lemma 6.18. *We have for $0 < \eta \leq d$, $u \geq 0$, $w > u + 9T(4d + \eta)$ and on the event $\mathsf{Fav} \cap \{P \cap [-2d, 2d] \neq \emptyset\}$,*

$$\mathbb{P}_{\mathcal{F}}\Big(W_\eta > w \;\Big|\; U = u\Big) \lesssim \exp\left(-\frac{(w - u - 9T(4d + \eta))^2}{2(4d + \eta)} + 7T\right).$$

To prove Lemma 6.18 we will need a lower bound on the probability of a Brownian bridge jumping over poles. This is recorded in the next lemma, whose straightforward proof is deferred to the end of the section to permit the flow of the overall argument.

Lemma 6.19. *Let $N \geq 1$, $x_0 \in [\mathfrak{l}, \mathfrak{r}]$, and $x_N = \mathfrak{r}$. Let $x_1 < \ldots < x_{N-1}$ be $\mathcal{F}$-measurable points in (x_0, x_N), and B be a Brownian bridge with law $\mathcal{B}_{0,0}^{[x_0,2T]}$. Then there exists a $G < \infty$ such that*

$$\mathbb{P}_{\mathcal{F}}\Big(B(x_i) > 0, i = 1, \ldots, N-1\Big) > G^{-1}N^{-1/2}\exp(-3N).$$

Proof of Lemma 6.18. Define the coordinate ϕ_{end} by

$$\phi_{\text{end}} := u + \text{Tent}(p) + 5T(2T - p),$$

which is the y-coordinate at $x = 2T$ of a line with slope $5T$ started at $(p, u + \text{Tent}(p))$. Since on Fav we have $|\text{slope}(\text{Tent})| \leq 4T$, it follows that $\phi_{\text{end}} > \mathcal{L}(2T)$. Lemma 5.12(iv) tells us that the Brownian bridge B from $(p, u+\text{Tent}(p))$ to $(2T, \phi_{\text{end}})$ conditioned to jump over all poles in $[p, 2T]$ and to be above $\text{Corner}_k^{\mathfrak{r},\mathcal{F}}$ stochastically dominates J on $[p, 2T]$. Let A be the conditioning event just mentioned. Then we have, on $\mathsf{Fav} \cap \{P \cap [-2d, 2d] \neq \emptyset\}$,

$$\mathbb{P}_{\mathcal{F}}\Big(W_\eta > w \;\Big|\; U = u\Big) \leq \mathcal{B}^{[p,2T]}_{u+\text{Tent}(p),\phi_{\text{end}}}\Big(B(p + 4d + \eta) > w + \text{Tent}(p + 4d + \eta) \;\Big|\; A\Big).$$

We need to lower bound $\mathbb{P}_{\mathcal{F}}(A)$ on Fav. Note that on Fav we have $\text{Corner}_k^{\mathfrak{r},\mathcal{F}} \leq T^2$, while $B(\mathfrak{r})$ has mean bounded below by

$$J(p) + 5T \cdot (\mathfrak{r} - p) \geq -T^2 + 5T \cdot (T/2 - d) = \frac{3}{2}T^2 - 5Td;$$

here we used $J(p) \geq \text{Tent}(p) = \mathcal{L}_n(k+1, p) \geq -T^2$ on Fav. These bounds, along with the concavity of Tent and Lemma 6.19 (with $N = |P| \leq 2T$, as $d_{\text{ip}} \geq 1$ implies this bound), implies that $\mathbb{P}_{\mathcal{F}}(A) \cdot \mathbb{1}_{\mathsf{Fav}}$ is bounded below, up to an absolute constant, by $T^{-\frac{1}{2}}\exp(-6T) \geq \exp(-7T)$ as $T \geq 1$. So, on $\mathsf{Fav} \cap \{P \cap [-2d, 2d] \neq \emptyset\}$,

$$\mathbb{P}_{\mathcal{F}}\Big(W_\eta > w \;\Big|\; U = u\Big) \lesssim \exp(7T) \cdot \mathcal{B}^{[p,2T]}_{u+\text{Tent}(p),\phi_{\text{end}}}\Big(B(p + 4d + \eta) > w + \text{Tent}(p + 4d + \eta)\Big).$$

Let $\rho = \frac{4d+\eta}{2T-p}$. Then $\mathbb{E}_{\mathcal{F}}[B(p+4d+\eta)] = u + \text{Tent}(p) + 5T(4d+\eta)$ and $\text{Var}_{\mathcal{F}}(B(p+4d+\eta)) = (1-\rho)(4d+\eta)$. So, for $w > u + 9T(4d+\eta)$ and on Fav,

$$\begin{aligned}
&\mathcal{B}^{[p,2T]}_{u+\text{Tent}(p),\phi_{\text{end}}}\Big(B(p + 4d + \eta) > w + \text{Tent}(p + 4d + \eta)\Big) \\
&\qquad = \mathbb{P}_{\mathcal{F}}\Big(N(u + \text{Tent}(p) + 5T(4d+\eta), (1-\rho)(4d+\eta)) > w + \text{Tent}(p + 4d + \eta)\Big) \\
&\qquad \leq \mathbb{P}_{\mathcal{F}}\Big(N(0, (1-\rho)(4d+\eta)) > w - u - 9T(4d+\eta)\Big) \\
&\qquad \leq \exp\left(-\frac{(w - u - 9T(4d+\eta))^2}{2(4d+\eta)}\right)
\end{aligned}$$

The third line uses that $\text{Tent}(p+4d+\eta) - \text{Tent}(p) \geq -4T(4d+\eta)$ from (5.4); the final inequality uses the upper bound of Lemma 6.9 with $\sigma^2 = (1-\rho)(4d+\eta)$ and $t = w - u - 9T(4d+\eta)$. □

With this tail bound we may turn to the proof of Proposition 6.13.

Proof of Proposition 6.13. Let λ be defined as

$$\lambda = \frac{\eta}{4d+\eta}$$

so that, conditionally on Y, U, and W_η, the distribution of Z is

$$Z = \lambda U + (1-\lambda)W_\eta + N(0, 4\lambda d). \tag{6.11}$$

This relation holds because, by the definition of J, the distribution of J on $[p, p+4d+\eta]$, conditionally on $J(p)$ and $J(p+4d+\eta)$, is a Brownian bridge with the prescribed endpoint values, and because Tent is affine on $[p, p+4d+\eta]$ for $\eta < d$ as $d_{\mathrm{ip}} = 5d$. Note that the variance of this Brownian bridge at $p+4d$ is $\frac{4d\eta}{4d+\eta} = 4\lambda d$.

Let $\nu_p(\cdot \mid y)$ be the conditional law of U given that $Y = y$. From the narrow jump Lemma 6.16 we see

$$f_J(z \mid y) \lesssim d^{\frac{1}{2}} \cdot \int_0^1 f_J(z, u \mid y)\, \mathrm{d}u = d^{\frac{1}{2}} \cdot \int_0^1 f_J(z \mid u, y)\, \mathrm{d}\nu_p(u \mid y), \tag{6.12}$$

where $f_J(z \mid u, y)$ is the conditional density of Z given U and Y. So our task is to bound $f_J(z \mid u, y)$ when $u \in [0,1]$. By the Markov property, this object does not depend on y. We will obtain this bound by converting the conditional tail bound of Lemma 6.18 to a conditional density bound using Lemma 6.17.

We have from Lemma 6.18, by taking $w = \frac{t-\lambda u}{1-\lambda}$ and simplifying the resulting expression, that, for $t > u + 36Td$,

$$\mathbb{P}_{\mathcal{F}}\Big(\lambda U + (1-\lambda)W_\eta > t \mid U = u, Y = y\Big) \lesssim \exp\left(-\frac{1}{8d(1-\lambda)}\left(t - u - 36Td\right)^2 + 7T\right). \tag{6.13}$$

While simplifying we used that $(4d+\eta)(1-\lambda) = 4d$. Now using (6.11) and (6.13), we apply Lemma 6.17. The parameters of this application are $\sigma_1^2 = 4\lambda d$, $\sigma_2^2 = 4(1-\lambda)d$, $x_0 = u + 36Td$, and A equal to the constant specified by $\lesssim$ in (6.13) multiplied by $\exp(7T)$. Thus we obtain, for every $\eta < d$ and $z > u + 36Td$,

$$f_J(z \mid u, y) \lesssim \eta^{-\frac{1}{2}} \cdot \exp\left(-\frac{1}{2\sigma^2}\left(z - u - 36Td\right)^2 + 7T\right),$$

where $\sigma^2 = 4d(\lambda^{1/2} + (1-\lambda)^{1/2})^2 \le 4d(1 + 2\lambda^{1/2}) \le 4d(1+\eta^{1/2})$ since $0 \le \lambda \le \eta/4$. Setting $\eta^{1/2} = 8T^{-2}d$ (which satisfies $\eta < d$ as $d \le \sqrt{T}/24$ and $T \ge 1$) and using that $(1+x)^{-1} \ge 1 - x$ for $x = \eta^{1/2}$ yields

$$f_J(z \mid u, y) \lesssim d^{-1}T^2 \cdot \exp\left(-\frac{1}{8d}\left(z - u - 36Td\right)^2 + T^{-2}\left(z - u - 36Td\right)^2 + 7T\right).$$

Using this in (6.12) and bounding the integral by a trivial bound on the integrand then shows

$$f_J(z \mid y) \lesssim d^{-\frac{1}{2}} T^2 \cdot \exp\left(-\frac{1}{8d}(z - 1 - 36Td)^2 + T^{-2}z^2 + 7T\right)$$

for $z > 36Td + 1$. Since $4Td > 1$, and since $T^2 \leq e^T$ as $T \geq 1$, it also holds that

$$f_J(z \mid y) \lesssim d^{-\frac{1}{2}} \cdot \exp\left(-\frac{1}{8d}(z - 40Td)^2 + T^{-2}z^2 + 8T\right)$$

for $z > 40Td$. Expanding the expression in the exponent gives the expression

$$-\frac{z^2}{8d} + \frac{z(40Td)}{4d} - \frac{40^2T^2d^2}{8d} + z^2T^{-2} + 8T \leq -\frac{z^2}{8d} + 20RT^{5/2} - 200T^2d + 4R^2T + 8T,$$

using that $z \leq 2RT^{3/2}$. Since $R = 6\sqrt{d}$ from (5.19) and since $T, d \geq 1$, the last three terms in the previous display are collectively negative and may be dropped. This proves the first part of Proposition 6.13.

Now we turn to the latter claim of Proposition 6.13. The tail bound (6.13), combined with the latter part of Lemma 6.17, says that $f_J(z \mid y, u) \lesssim d^{-\frac{1}{2}}$. Then (6.12) gives the latter claim of Proposition 6.13. This completes the proof of Proposition 6.13. □

Finally, we provide the last piece of the proof of Proposition 6.13 by proving Lemma 6.19.

Proof of Lemma 6.19. Let A denote the event that a Brownian motion B' with law $\mathcal{B}_{0,*}^{[x_0,2T]}$ is negative at $2T$; write $\widetilde{B}$ for the process B' conditioned on A. Since Brownian motion on $[x_0, 2T]$ conditioned on its value at $2T$ is a Brownian bridge on $[x_0, 2T]$ with appropriate endpoints, Lemma 5.4 implies that the Brownian bridge B in the statement of Lemma 6.19 stochastically dominates $\widetilde{B}$. Thus,

$$\mathbb{P}_{\mathcal{F}}\Big(\bigcap_{i=1}^{N-1}\{B(x_i) > 0\}\Big) \geq \mathbb{P}\Big(\bigcap_{i=1}^{N-1}\Big\{\widetilde{B}(x_i) > 0\Big\} \mid A\Big) = \frac{\mathbb{P}\Big(\bigcap_{i=1}^{N}\Big\{\widetilde{B}(x_i) > 0\Big\}, \widetilde{B}(2T) < 0\Big)}{\mathbb{P}\Big(\widetilde{B}(2T) < 0\Big)}.$$

The denominator is equal to $\frac{1}{2}$. We may lower bound the numerator by

$$\mathbb{P}\left(\bigcap_{i=1}^{N}\Big\{\widetilde{B}(x_i) - \widetilde{B}(x_{i-1}) \in [0, \sqrt{x_i - x_{i-1}}]\Big\}, \widetilde{B}(2T) - \widetilde{B}(x_N) < -\sum_{i=1}^{N}\sqrt{x_i - x_{i-1}}\right)$$
$$= p^N \cdot \mathbb{P}\left(N(0, 2T - x_N) < -\sum_{i=1}^{N}\sqrt{x_i - x_{i-1}}\right), \quad (6.14)$$

using the independence of Brownian motion increments, where $p = \mathbb{P}\big(N(0,1) \in [0,1]\big) > (2\pi)^{-1/2}e^{-1/2} > e^{-2}$; the first inequality is by lower bounding the standard normal density on $[0,1]$ by $(2\pi)^{-1/2}e^{-1/2}$, and the last inequality is by numerical calculation. Now,

$$\mathbb{P}\left(N(0,2T-x_N) < -\sum_{i=1}^{N}\sqrt{x_i - x_{i-1}}\right) \geq \mathbb{P}\left(N(0,1) > \frac{\sqrt{N}\sqrt{x}_N}{\sqrt{2T-x_N}}\right)$$
$$\geq \mathbb{P}\left(N(0,1) > \sqrt{N}\right) \gtrsim \frac{e^{-N/2}}{\sqrt{N}}.$$

We have used the Cauchy-Schwarz inequality in the first inequality; that $x_N = \mathfrak{r} \leq T$ in the second; and Lemma 6.9 in the last. Combining with (6.14) and that $\mathbb{P}\big(\tilde{B}(2T) < 0\big) = \frac{1}{2}$ completes the proof. □

6.4 When no pole is present

The aim of this section is to perform the final task in the proof of Theorem 5.11, namely to provide the proof of Proposition 5.18. The analysis follows similar lines to the case where a pole was present, but here we are aided by the fact that there is no vault cost V. Because there is no pole, we cannot make the choice of stepping back to $p-4d$ and $p+4d$ in decomposing the jump curve probability. Since we know only that there is no pole in $[-2d, 2d]$, and we need to give ourselves some distance from the nearest pole, we make the choice to not step back and instead decompose at $-d$ and d. More precisely we recall, as mentioned before Lemma 5.15, that Y and Z are defined on the event $P \cap [-2d, 2d] = \emptyset$ as

$$Y := J(-d) - \mathrm{Tent}(-d)$$
$$Z := J(d) - \mathrm{Tent}(d).$$

Correspondingly, on $P \cap [-2d, 2d] = \emptyset$, $f_J(y,z)$ is the joint density of (Y,Z) under this definition.

As there is no vault cost, we only need to consider the slope cost S. In this context, since the interval $[-d,d]$ has length $2d$, S is

$$S = d^{\frac{1}{2}} \cdot \exp\left(-\frac{1}{4d}(\tilde{y}-\tilde{z})^2\right);$$

here, $\tilde{y} = y + \mathrm{Tent}(-d)$ and $\tilde{z} = z + \mathrm{Tent}(d)$, so that $Y = y$ implies $J(-d) = \tilde{y}$ and similarly for $Z = z$ and $J(d) = \tilde{z}$; thus the tilde plays the same role as it did in Section 5.2 of going from the value of Y or Z to the value of J at the corresponding points.

The bound we will obtain from Proposition 6.4 is in terms of y and z, and here we see the exact form of the leeway factor needed to write S in terms of y and z, as was done in (5.21) in the pole case; this bound will hold for $(y,z) \in \mathcal{G}_R^{(2)}$, where we recall the definition of $\mathcal{G}_R^{(2)}$ from

(5.17). Using from (5.4) that the slope of the Tent map is bounded in absolute value by $4T$, so that $|\mathrm{Tent}(d) - \mathrm{Tent}(-d)| \le 8Td$, we see that, for $(y,z) \in \mathcal{G}_R^{(2)}$,

$$\begin{aligned} S &= d^{\frac{1}{2}} \cdot \exp\left(\frac{1}{4d}(\tilde{y}-\tilde{z})^2\right) \\ &\le d^{\frac{1}{2}} \cdot \exp\left(\frac{1}{4d}(y-z)^2 + 16T^2 d + \frac{1}{2d}(y-z)(\mathrm{Tent}(d) - \mathrm{Tent}(-d))\right) \\ &\le d^{\frac{1}{2}} \cdot \exp\left(-\frac{1}{4d}(y-z)^2 + T^{5/2} + 8RT^{5/2}\right) \\ &\le d^{\frac{1}{2}} \cdot \exp\left(-\frac{1}{4d}(y-z)^2 + 9RT^{5/2}\right); \end{aligned} \tag{6.15}$$

the first inequality using the mentioned bound on $|\mathrm{Tent}(d) - \mathrm{Tent}(-d)|$; the second since $(y,z) \in \mathcal{G}_R^{(2)}$ implies $|y-z| \le 2RT^{3/2}$, and since $d \le \sqrt{T}/24$ and $16/24 \le 1$; and third since $R \ge 1$ from (5.19).

Apart from this, we will also need to make a suitable choice for the distribution μ, discussed in Section 5.2, which is here the distribution of the Brownian motion we start at $-d$ to which we compare J. We select μ to be uniform on $[-RT^2, RT^2]$ as this interval contains all values y may take when $(y,z) \in \mathcal{G}_R^{(2)}$. As before, the cost for this choice of μ is only polynomial in T, and so does not affect the bound we need to prove on $f_J(y,z)$.

So overall our aim is to get a bound on the joint density of the form

$$f_J(y,z) \le \exp\left(-\frac{1}{4d}(y-z)^2\right) \cdot \exp\left(GT^{5/2}\right)$$

for some $G < \infty$. This of course is essentially immediate from Proposition 6.4 when $|y-z|$ is sufficiently large; recall that Proposition 6.4 was stated carefully to not assume the presence of a pole in the interval of consideration.

Proof of Proposition 5.18. On $\mathsf{Fav} \cap \{P \cap [-2d, 2d] = \emptyset\}$,

$$\begin{aligned} &\mathbb{P}_{\mathcal{F}}\Big(J(\,\cdot\,) - J(-d) \in A, (Y,Z) \in \mathcal{G}_R^{(2)}\Big) \\ &\qquad = \mathbb{E}_{\mathcal{F}}\Big[\mathbb{P}_{\mathcal{F}}\Big(J(\,\cdot\,) - J(-d) \in A \;\Big|\; J(d), J(-d)\Big)\mathbb{1}_{(Y,Z)\in\mathcal{G}_R^{(2)}}\Big] \\ &\qquad = \mathbb{E}_{\mathcal{F}}\left[\mathcal{B}^{[-d,d]}_{J(-d),J(d)}(\tilde{A})\,\mathbb{1}_{(Y,Z)\in\mathcal{G}_R^{(2)}}\right], \end{aligned}$$

where $\tilde{A}$ is the set of functions f in $\mathcal{C}_{*,*}([-d,d],\mathbb{R})$ such that $f(\,\cdot\,) - f(-d) \in A$. Recall the notation $\tilde{y} = y + \mathrm{Tent}(-d)$ and $\tilde{z} = z + \mathrm{Tent}(d)$. We may write the last displayed expression as an integral on the event $\mathsf{Fav} \cap \{P \cap [-2d, 2d] = \emptyset\}$:

$$\mathbb{E}_{\mathcal{F}}\left[\mathcal{B}^{[-d,d]}_{J(-d),J(d)}(\tilde{A})\,\mathbb{1}_{(Y,Z)\in\mathcal{G}_R^{(2)}}\right] = \iint_{\mathcal{G}_R^{(2)}} \mathcal{B}^{[-d,d]}_{\tilde{y},\tilde{z}}(\tilde{A})\, f_J(y,z)\,\mathrm{d}y\,\mathrm{d}z$$

$$= d^{-\frac{1}{2}} \iint_{\mathcal{G}_R^{(2)}} \mathcal{B}^{[-d,d]}_{\tilde{y},\tilde{z}}\big(\tilde{A}\big) e^{-\frac{1}{4d}(\tilde{y}-\tilde{z})^2} \cdot \big(f_J(y,z)\cdot S\big)\,\mathrm{d}y\,\mathrm{d}z. \quad (6.16)$$

From Proposition 6.4 with parameters $x_1 = -d$ and $x_2 = d$, we know that on Fav and when $|y-z| > 12RTd$ and $(y,z) \in \mathcal{G}_R^{(2)}$,

$$\begin{aligned} f_J(y,z) &\lesssim d^{-1}T^2 \cdot \exp\left(-\frac{1}{4d}(|y-z| - 12RTd)^2 + \frac{4R^2T^2}{2d} + 72R^2d\right) \\ &= d^{-1}T^2 \cdot \exp\left(-\frac{1}{4d}\left((y-z)^2 - 24RTd|y-z| + 144R^2T^2d^2\right) + \frac{4R^2T^2}{2d} + 72R^2d\right) \\ &\le d^{-1} \cdot \exp\left(-\frac{1}{4d}(y-z)^2 + 6RT|y-z| - 36R^2T^2d + \frac{4R^2T^2}{2d} + 72R^2d + T^2\right) \\ &\le d^{-1} \cdot \exp\left(-\frac{1}{4d}(y-z)^2 + 12R^2T^{5/2}\right), \end{aligned}$$

the penultimate inequality since $T^2 \le \exp(T^2)$ as $T \ge 1$; and the last inequality since on $\mathcal{G}_R^{(2)}$ we have $|y-z| < 2RT^{3/2}$ and since $1 \le d \le \sqrt{T}/24$, which implies that the sum of the last four terms in the exponent in the penultimate line is negative and may be dropped.

For when $|y-z| < 12RTd$, we also use Proposition 6.4 to say that $f_J(y,z) \lesssim d^{-1}$; and so we see that

$$\begin{aligned} f_J(y,z) \lesssim d^{-1} \cdot \exp\left(-\frac{(y-z)^2}{4d} + \frac{(y-z)^2}{4d}\right) &\le d^{-1} \cdot \exp\left(-\frac{1}{4d}(y-z)^2 + \frac{144R^2T^2d^2}{4d}\right) \\ &\le d^{-1} \cdot \exp\left(-\frac{1}{4d}(y-z)^2 + 2R^2T^{5/2}\right), \end{aligned}$$

the last inequality since $d \le \sqrt{T}/24$ and $144/(4\times 24) \le 2$.

So we see from (6.15) and the above two bounds on $f_J(y,z)$ depending on the size of $|y-z|$, that

$$f_J(y,z)\cdot S \le \exp\left(21R^2T^{5/2}\right),$$

using that $R \le R^2$ and $d \ge 1$.

Thus, we obtain that (6.16) is bounded up to an absolute multiplicative constant by

$$\exp\left(21R^2T^{5/2}\right) \cdot \frac{1}{\sqrt{4\pi d}} \iint_{\mathbb{R}^2} \mathcal{B}^{[-d,d]}_{\tilde{y},\tilde{z}}\Big(\tilde{A}\Big) e^{-\frac{1}{4d}(\tilde{y}-\tilde{z})^2}\,\mathrm{d}\mu(y)\,\mathrm{d}z. \quad (6.17)$$

Let B be a Brownian motion started with distribution μ at $-d$. Focusing on the integral,

$$\frac{1}{\sqrt{4\pi d}} \iint_{\mathbb{R}^2} \mathcal{B}^{[-d,d]}_{\tilde{y},\tilde{z}}\Big(\tilde{A}\Big) e^{-\frac{1}{4d}(\tilde{y}-\tilde{z})^2}\,\mathrm{d}\mu(y)\,\mathrm{d}z = \mathbb{P}\Big(B(\,\cdot\,) - B(-d) \in A\Big) = \mathcal{B}^{[-d,d]}_{0,*}\big(A\big) = \varepsilon,$$

using the Markov property of Brownian motion for the penultimate equality.

Combining this with the ignored factor in (6.17) and using that $R^2 = 36d$ from (5.19) gives that (6.16) is bounded above, up to an absolute multiplicative constant, by

$$\varepsilon \cdot \exp\left(756 \cdot D_k^{5/2} \cdot d \cdot (\log \varepsilon^{-1})^{5/6}\right),$$

since $21 \times 36 = 756$. □

Part III

Bootstrapping

Chapter 7

The basic idea of bootstrapping

This part of the thesis, namely Chapters 7–10, forms the second pillar, concerning the geometric perspective on LPP. Chapters 7–9 consist of material from [GH20c] in work done with Shirshendu Ganguly, while Chapter 10 contains material from [BGHH20] done with Riddhipratim Basu, Alan Hammond, and Shirshendu Ganguly.

7.1 Introduction, the model, and assumptions

We need to state the precise version of Theorem 2.6.

Recall that in LPP on $\mathbb{Z}^2$ one assigns i.i.d. non-negative weights $\{\xi_v : v \in \mathbb{Z}^2\}$ to the vertices of $\mathbb{Z}^2$ and studies the weight and geometry of weight-maximising directed paths. The weight of a given up-right nearest neighbor path γ is $\ell(\gamma) := \sum_{v\in\gamma} \xi_v$. For given vertices $u = (u_1, u_2), v = (v_1, v_2) \in \mathbb{Z}^2$ with $u_i \leq v_i$ for $i = 1$ and 2 (i.e., the natural partial order), the last passage value $X_{u,v}$ is defined by $X_{u,v} = \max_{\gamma:u\to v} \ell(\gamma)$, where the maximization is over the set of up-right paths from u to v; maximizing paths are called *geodesics*. For $r \in \mathbb{N}$, we adopt the shorthand $X_r := X_{(1,1),(r,r)}$.

A few special distributions of the vertex weights $\{\xi_v : v \in \mathbb{Z}^2\}$ render the model integrable, i.e., admitting exact connections to algebraic objects such as random matrices and Young diagrams, as we saw in Chapters 1 and 2. This allows exact computations which lead to the appearance of the Airy_2 process and hence the GUE Tracy-Widom distribution. For concreteness, we recall next the special case of exponentially distributed (with rate one) vertex weights, previously stated as Theorem 2.4. In this case, Johansson proved the following [Joh00] via the development of the aforementioned connections to representation theory.

Theorem 7.1 (Theorem 1.6 of [Joh00]). *Let $\{\xi_v : v \in \mathbb{Z}^2\}$ be i.i.d. exponential rate one random variables. As $r \to \infty$ it holds that*

$$\frac{X_r - 4r}{2^{4/3}r^{1/3}} \xrightarrow{d} F_{\text{TW}},$$

where F_{TW} is the GUE Tracy-Widom distribution, and $\xrightarrow{d}$ denotes convergence in distribution.

This one-point convergence was later upgraded to convergence to the Airy_2 process in the sense of finite dimensional distributions by considering a suitable observable (see for example [BF08]).

We recall the important feature of the GUE Tracy-Widom distribution mentioned in Chapter 2, the "non-Gaussian" behavior of its upper and lower tails. In particular, it is known, for example from [Sep98b, page 224] or [RRV11, Theorem 1.3], that as $\theta \to \infty$,

$$\begin{aligned} F_{\text{TW}}\big([\theta,\infty)\big) &= \exp\left(-\frac{4}{3}\theta^{3/2}\left(1+o(1)\right)\right) \qquad \text{and} \\ F_{\text{TW}}\big((-\infty,-\theta]\big) &= \exp\left(-\frac{1}{12}\theta^{3}\left(1+o(1)\right)\right). \end{aligned} \tag{7.1}$$

Given the distributional convergence asserted by Theorem 7.1, it is natural to ask whether tail bounds similar to (7.1) are satisfied by X_r at the finite r level. Indeed, again in the case of exponential weights, estimates along these lines have been attained which achieve the correct upper and lower tail exponents of $3/2$ and 3. The first result in this direction was proved by Seppäläinen, who obtained an upper bound for the upper tail (with the correct leading exponent coefficient $4/3$) in [Sep98a, page 622] via a coupling with the totally asymmetric simple exclusion process and an evaluation and expansion of the large deviation rate function. The large deviation bound yields a finite r estimate using superadditivity properties of the upper tail probabilities (see (7.6) ahead for a discussion). But this strategy does not give a lower bound or bounds for the lower tail, and these bounds were proven using connections to random matrix theory. In more detail, Johansson proved in [Joh00, Remark 1.5] via representation theoretic techniques that X_r is equal in distribution to the top eigenvalue of the Laguerre Unitary Ensemble, and upper bounds on the upper and lower tails on this eigenvalue were proved in [LR10, Theorem 2]. [LR10] remarks, but does not prove, that a lower bound on the upper tail should be achievable by methods in the paper, but not a lower bound on the lower tail; the latter was proved very recently in [BGHK19, Theorem 2]. This discussion may be summarized as the following theorem, which was previously stated as Theorem 2.5.

Theorem 7.2 ([Sep98a, Joh00, LR10, BGHK19])**.** *Let $\{\xi_v : v \in \mathbb{Z}^2\}$ be i.i.d. exponential random variables. There exist positive finite constants c_1, c_2, c_3, θ_0, and r_0 such that, for $r > r_0$ and $\theta_0 < \theta < r^{2/3}$,*

$$\begin{aligned} \mathbb{P}\left(X_r > 4r + \theta r^{1/3}\right) &\le \exp\left(-c_1\theta^{3/2}\right) \qquad \textit{and} \\ \exp\left(-c_2\theta^{3}\right) \le \mathbb{P}\left(X_r < 4r - \theta r^{1/3}\right) &\le \exp\left(-c_3\theta^{3}\right). \end{aligned}$$

Remark 7.3. In fact, the missing lower bound on the upper tail is a straightforward consequence of one of our results (Theorem 7.5) along with the distributional convergence in Theorem 7.1 and an application of the Portmanteau theorem.

That the above bounds hold only for $\theta \le r^{2/3}$ is an important fact because one should not expect universality beyond this threshold. The lower tail is trivially zero for $\theta > 4r^{2/3}$ since the

vertex weights are non-negative; for the upper tail, beyond this level, we enter the large deviation regime, where the tail behavior is dictated by the individual vertex distribution. Thus in the case of exponential LPP, the upper tail decays exponentially in $\theta r^{1/3}$ for $\theta > r^{2/3}$.

Similar bounds as Theorem 7.2 are available in only a handful of other LPP models; these are when the vertex weights are geometric [Joh00, BDM+01], and the related models of Poissonian LPP [Sep98b, LM01, LMR02] and Brownian LPP [OY02, LR10]. While [Sep98b] relies on coupling Poissonian LPP to Hammersley's process (a continuous version of the exclusion process), the remaining arguments use powerful identities with random matrix theory and connections to representation theory, combined with precise analysis of the resulting formulas.

However, the conjectured universality of KPZ behavior suggests that similar bounds should hold under rather minimal assumptions on the vertex weight distribution, i.e., even when special connections to random matrix theory and representation theory are unavailable. Thus it is an important goal to develop more robust methods of investigation that may apply to a wider class of models, an objective that has driven a significant amount of work in this field, with the eventual aim to go beyond integrability.

Nonetheless, despite various attempts, so far only a few results are known to be true in a universal sense. These include the existence of a limiting geodesic weight profile (i.e., the expected geodesic weight as the endpoint varies) and its concavity under mild moment assumptions on the vertex weights [Mar06]. This is a relatively straightforward consequence of super-additivity properties exhibited by the geodesic weights, as we elaborate on later. This and certain general concentration estimates based on martingale methods were first developed in Kesten's seminal work on first passage percolation (FPP) [Kes86]; FPP is a notoriously difficult to analyze and canonical non-integrable model in the KPZ class where the setting is the same as that of LPP, but one instead minimizes the weight among all paths between two points, without any orientation constraint. Similar arguments extend to the case of general LPP models. Note that while the precise limiting profile is expected to depend on the model, properties such as concavity as well as local fluctuation behavior are predicted to be universal.

Following Kesten's work, there has been significant progress in FPP in providing rigorous proofs assuming certain natural conditions, such as strong curvature properties of limit shapes and the existence of critical exponents dictating fluctuations. Thus an important broad goal is to extract the minimal set of key properties of such models that govern other more refined behavior. The recent work of the myself with Shirshendu Ganguly, Riddhipratim Basu, and Alan Hammond in [BGHH20], as well as the content of this part of this thesis, are guided by the same philosophy. We will revisit this discussion in more detail after the statements of our main results.

To initiate the geometric perspective of this part of the thesis, we point out the disparity in the upper and lower tail exponents in Theorem 7.2. This is not surprising since, while the upper tail event enforces the existence of a *single* path of high weight, the lower tail event is global and forces *all* paths to have low weight.

However, the precise exponents of $3/2$ and 3 might appear mysterious, and it is natural to seek a geometric explanation for them. This is the goal of this work. More precisely, we establish

bounds with optimal exponents in the nature of Theorem 7.2, starting from certain much weaker tail bounds as well as local strong concavity assumptions on the limit shape (Theorems 7.4–7.7). In particular, we do not make use of any algebraic techniques in our arguments; indeed, the nature of our assumptions do not allow such techniques to be applicable. Instead, our methods are strongly informed by an understanding of the geometry of geodesics and other weight maximising path ensembles in last passage percolation.

We also mention that, while our main result is known in integrable models such as exponential LPP in view of Theorem 7.2, our techniques also obtain sharp tail exponents for a related LPP problem, namely the lower tail of the maximum weight over all paths *constrained* to lie inside a strip of given width (Theorem 7.13 ahead); the precise exponent depends on this width. Estimates of these probabilities have played important roles in previous geometric investigations [BSS14, BGHH20], but sharp forms had not been proven even in integrable models, and do not seem amenable to exactly solvable analysis. The form of the exponent we prove in Theorem 7.13 is also suggestive of the anticipated answer to the question of typical transverse fluctuations of the geodesic when conditioned on having low weight in the moderate deviation regime; we elaborate on this slightly following Theorem 7.13. The large deviation version of the same question was investigated in [BGS19], and the related upper tail large deviation version in FPP in [BGS17].

We next set up precisely the framework of last passage percolation on $\mathbb{Z}^2$, describe our assumptions, and state our main results.

Model, notation, and assumptions

We denote the set $\{1, 2, \ldots\}$ by $\mathbb{N}$, and, for $i, j \in \mathbb{Z}$, we denote the integer interval $\{i, i+1, \ldots, j\}$ by $[\![i, j]\!]$.

We start with a random field $\{\xi_v : v \in \mathbb{Z}^2\}$ of i.i.d. random variables following a distribution ν supported on $[0, \infty)$. We consider up-right nearest neighbor paths, which we will refer to as *directed* paths. For a directed path γ, the associated *weight* is denoted $\ell(\gamma)$ and is defined by

$$\ell(\gamma) := \sum_{u \in \gamma} \xi_u.$$

For $u, v \in \mathbb{Z}^2_+$, with $u \preceq v$ in the natural partial order mentioned earlier, we denote by $X_{u,v}$ the *last passage value* or *weight* from u to v, i.e.,

$$X_{u,v} := \max_{\gamma : u \to v} \ell(\gamma),$$

where the maximization is over all directed paths from u to v; for definiteness, when u and v are not ordered in this way and there is no directed path from u to v, we define $X_{u,v} = -\infty$. Now for ease of notation, for sets $A, B \subseteq \mathbb{Z}^2$, we also adopt the intuitive shorthand

$$X_{A,B} := \sup_{u \in A, v \in B} X_{u,v}.$$

For $v \in \mathbb{Z}^2_+$, X_v will denote $X_{(1,1),v}$, and for $r, z \in \mathbb{Z}$, we will denote $X_{(1,1),(r-z,r+z)}$ by X^z_r. We will also denote the case of $z = 0$ by X_r, as above. Notational confusion between X_v and X_r is avoided in practice in this usage as v will always be represented by a pair of coordinates, while r is a scalar. Recall that a path (which may not be unique) which achieves the last passage value is called a geodesic.

For an up-right path γ from $(1, 1)$ to $(r - z, r + z)$, we define the *transversal fluctuation* of γ by

$$\mathrm{TF}(\gamma) := \min \{w : \gamma \subseteq U_{r,w,z}\},$$

where $U_{r,w,z}$ is the strip of width w around the interpolating line, i.e., the set of vertices $v \in \mathbb{Z}^2$ such that $v + t \cdot (-1, 1)$ lies on the line $y = \frac{r+z}{r-z} \cdot x$ for some $t \in \mathbb{R}$ with $|t| \leq w/2$.

Assumptions

The general form of our assumptions is quite similar to the ones in the recent work [BGHH20] devoted to the study of geodesic watermelons, a path ensemble generalizing the geodesic, and also bears resemblence to ones adopted in recent studies [Ale20, Gan20] in FPP. We start by recalling that ν is the distribution of the vertex weights and has support contained in $[0, \infty)$. The limit shape is the map $[-r, r] \to \mathbb{R} : z \mapsto \lim_{r\to\infty} r^{-1}\mathbb{E}[X^z_r]$. It follows from standard super-additivity arguments that this limit exists (though possibly infinite if the upper tail of ν is too heavy) for each $z \in [-r, r]$ and that this map is *concave* [Mar06, Proposition 2.1]. Let

$$\mu = \lim_{r\to\infty} r^{-1}\mathbb{E}[X_r]$$

be this map evaluated at zero. Also note from Theorems 7.1 and 7.2 that the fluctuations of X_r around μ can be expected to be on scale $r^{1/3}$. Finally, we point out that the normalized limit shape map in the exactly solvable models of Exponential, Geometric, Brownian, and Poissonian LPP is, up to translation and scaling by constants,

$$\sqrt{r^2 - z^2} = r - \frac{z^2}{2r} - O\left(\frac{z^4}{r^3}\right); \tag{7.2}$$

this will be relevant in motivating the form of our second assumption. Note that the first term of the right hand side of (7.2) denotes the expected linear growth of the model, while the second encodes a form of strong concavity of the limit shape. Also, the non-random fluctuation, i.e., how much the mean of X^z_r falls below (7.2), is expected to be $\Theta(r^{1/3})$, which is known in the aforementioned exactly solvable models.

Given the setting, we state our assumptions; not all the assumptions are required for each of the main results, and we will specify which ones are in force in each case. We will elaborate more on the content of each assumption following their statements.

1. **Limit shape existence:** The vertex weight distribution ν is such that $\mu < \infty$.

2. **Strong concavity of limit shape and non-random fluctuations:** There exist positive finite constants ρ, G, H, g_1, and g_2 such that, for large enough r and $z \in [-\rho r, \rho r]$,

$$\mathbb{E}[X_r^z] \in \mu r - G\frac{z^2}{r} + \left[-H\frac{z^4}{r^3}, 0\right] + \left[-g_1 r^{1/3}, -g_2 r^{1/3}\right].$$

The first three terms on the right hand side encode the limit shape and its strong concavity as in (7.2), while the final interval captures the non-random fluctuation.

3. **Upper bound on moderate deviation probabilities, uniform in direction:** There exists $\alpha > 0$ such that the following hold. Fix any $\varepsilon > 0$, and let $|z| \in [0, (1-\varepsilon)r]$. Then, there exist positive finite constants c, θ_0, and r_0 (all depending on only ε) such that, for $r > r_0$ and $\theta > \theta_0$,
   a) $\mathbb{P}\left(X_r^z - \mathbb{E}[X_r^z] > \theta r^{1/3}\right) \leq \exp(-c\theta^\alpha)$,
   b) $\mathbb{P}\left(X_r^z - \mathbb{E}[X_r^z] < -\theta r^{1/3}\right) \leq \exp(-c\theta^\alpha)$.

4. **Lower bound on diagonal moderate deviation probabilities:** There exist constants $\delta > 0$, $C > 0$, r_0 such that, for $r > r_0$,
   a) $\mathbb{P}\left(X_r - \mu r > C r^{1/3}\right) \geq \delta$,
   b) $\mathbb{P}\left(X_r - \mu r < -C r^{1/3}\right) \geq \delta$.

These will be respectively referred to as Assumptions 1–4 in this part of the thesis. Assumption 1, which is known to be true under mild moment conditions on ν, is stated to avoid any pathologies and will be in force henceforth without us explicitly mentioning it further. Assumption 3 is the a priori tail assumption that our work seeks to improve on. We will refer to the tail bounds as *stretched exponential* though this term usually refers to $0 < \alpha \leq 1$, (which is the case of primary interest for us).

Assumption 1 in fact follows from Assumption 3a, for the latter implies that $\nu([\theta, \infty)) \leq \exp(-c\theta^\alpha)$, for a possibly smaller c and sufficiently large θ (see Remark 7.10).

Observe that Assumption 2 is a mild relaxation of the form of the weight profile in all known integrable models, as we do not impose a lower order term of order $-z^4/r^3$ in the upper bound. Our arguments would also work if we replaced the third term $[-Hz^4/r^3, 0]$ of Assumption 2 with $[-Hz^4/r^3,\ Hz^4/r^3]$, but we have not included this relaxation so as to not introduce further complexity.

The additional translation by $-\Theta(r^{1/3})$ in Assumption 2 for the non-random fluctuation will be a crucial ingredient (note that $\mathbb{E}[X_r] \leq \mu r$ by super-additivity). As the reader might already be aware, non-random fluctuations are an important object of study and this will be further evident from their role in the arguments here (in particular that they are the same scale as the random fluctuations) as well as in past work: see, for example, [BGHH20, BHS18]. For applications in FPP, see [Cha13] and [AD14]. A powerful general theory to control such objects for general sub-additive sequences, particularly in the context of FPP, was developed in [Ale97].

We end this discussion by pointing out that Assumption 4b follows from Assumptions 2 and 3b; see Lemma 9.7. This is essentially because by assumption $\mu r > \mathbb{E}[X_r] + \Theta(r^{1/3})$ and we have

assumed deviation bounds from the expectation. However, this style of argument does not work to derive Assumption 4a from Assumptions 2 and 3, and this task seems more difficult.

7.2 Main results

Our main contribution is to obtain the optimal upper and lower tail exponents for X_r in terms of upper and lower bounds, starting from a selection of the assumptions just stated. Notice that all the assumptions except the first involve the weight fluctuations occurring on scale $r^{1/3}$, and our results essentially connect this fluctuation exponent of $1/3$ to the two tail exponents. Here are the precise statements, which expand on the informal version stated earlier as Theorem 2.6.

Theorem 7.4 (Upper-tail upper bound). *Under Assumptions 2 and 3a, there exist constants c, $\zeta \in (0, \frac{2}{25}]$, r_0, and θ_0 (all depending on α) such that, for $\theta_0 \le \theta \le r^{\zeta}$ and $r > r_0$,*

$$\mathbb{P}\left(X_r - \mathbb{E}[X_r] \ge \theta r^{1/3}\right) \le \exp\left(-c\theta^{3/2}(\log\theta)^{-1/2}\right).$$

Further, $\zeta(\alpha) \to 0$ as $\alpha \to 0$, and $\zeta(\alpha) = \frac{2}{25}$ if $\alpha \ge 1$.

Theorem 7.5 (Upper-tail lower bound). *Under Assumptions 2 and 4a (the former only at $z = 0$), there exist constants $c > 0$, $\eta > 0$ and r_0 such that, for $r > r_0$ and $\theta_0 < \theta < \eta r^{2/3}$,*

$$\mathbb{P}\left(X_r - \mathbb{E}[X_r] \ge \theta r^{1/3}\right) \ge \exp\left(-c\theta^{3/2}\right).$$

Theorem 7.6 (Lower-tail upper bound). *Under Assumptions 2 and 3, there exist constants $c > 0$, r_0, and θ_0 (all depending on α) such that, for $\theta > \theta_0$ and $r > r_0$,*

$$\mathbb{P}\left(X_r - \mathbb{E}[X_r] \le -\theta r^{1/3}\right) \le \exp\left(-c\theta^{3}\right).$$

Theorem 7.7 (Lower-tail lower bound). *Under Assumptions 2, 3, and 4b, there exist constants $c > 0$, $\eta > 0$, θ_0, and r_0 (all depending on α) such that, for $r > r_0$ and $\theta_0 < \theta < \eta r^{2/3}$,*

$$\mathbb{P}\left(X_r - \mathbb{E}[X_r] \le -\theta r^{1/3}\right) \ge \exp\left(-c\theta^{3}\right).$$

The constants θ_0 and r_0 in the theorems should not be confused with the ones appearing in the assumptions.

The aforementioned Theorem 7.13 on upper and lower bounds for last passage values when paths are constrained to lie inside a strip of given width will be stated ahead after Section 7.3, which elucidates the main arguments of Theorems 7.4–7.7.

As the reader might anticipate, one might be able to relax some of the assumptions, e.g., the precise form of Assumption 2 should not be essential, and we expect our arguments to go through under reasonable relaxations. For example, a polynomial lower order term in (7.2), say of the form $|z|^{2+\delta}/r^{1+\delta}$ for some $\delta > 0$, or the related assumption of local uniform strong concavity of the

limit shape may be sufficient. However we have not pursued this as we have sought to achieve the cleanest presentation to highlight the key geometric insights underlying the arguments.

Next we make some remarks and observations on the results, focusing mainly on aspects of Theorem 7.4.

Remark 7.8 (Relation of tail exponents to fluctuation exponents)**.** We have assumed that weight fluctuations occur on the scale $n^{1/3}$, and this is because this is thought to be the scale of fluctuations for LPP in two dimensions for all vertex weight distributions. The basis for this is the following heuristic. Let χ be the scale of weight fluctuations, and ξ the scale of transversal fluctuations (i.e., the scale of the width of the smallest rectangle containing the geodesic), also called the wandering exponent. In any dimension, these exponents are expected to satisfy the KPZ relations $\chi = 2\xi - 1$; this has been proven in FPP in [Cha13] for a particular precise definition of the exponents. (These results and statements is contingent on the existence of these exponents, which is non-trivial and an important open problem.) In two dimensions, the weight profile is additionally expected to exhibit Brownian fluctuations, which suggests $\chi = \xi/2$. These combine to imply $\chi = 1/3$ and $\xi = 2/3$.

For LPP in higher dimensions, the KPZ relation is still expected to hold, but not the Brownian nature of the weight profile (as it is no longer a one-dimensional function). Thus it is natural to ask what the tail exponents would be in this case. The algebra of our arguments suggests that for general dimension the upper and lower tail exponents should respectively be $1/(1-\chi)$ and $2/(1-\chi) = 1/(1-\xi)$ (by the KPZ relation); see [GH20c, Section 1.7] for some more details.

Even in two dimensions, the exponent χ need not be $1/3$ if the noise field is not i.i.d. [BH19]. As seen in [BH19], the KPZ relation need not hold in this setting, and there is no reason to expect Brownian fluctuations for the weight profile. But it is interesting to ask what relation may exist between the fluctuation exponents and the tail exponents. In fact, the argument for Theorem 7.5 (lower bound on the upper tail), which we discuss in Section 7.3, should apply quite generally, i.e., as long as correlation inequalities hold, and suggests that at least in models enjoying positive association the upper tail exponent may be $1/(1-\chi)$.

In a more classical setting, it is a nice exercise to use that the fluctuations of random walk of size n are of order $n^{1/2}$ to conclude that the tail exponents in that case should be $1/(1-\frac{1}{2}) = 2$, again via the arguments for the upper tail ahead. For the lower bound, the argument does not make use of concentration of measure estimates, and thus provides a simple geometric indication of the source of the Gaussian distribution's tail exponent that we were not previously aware of. (The above prediction of a higher exponent for the lower tail does not apply since this is not a model of last passage percolation.)

Remark 7.9 (Suboptimal log factor in Theorem 7.4)**.** The reader would have noticed that the tail in Theorem 7.4 is not optimal, due to the appearance of $(\log\theta)^{-1/2}$. This arises due to the lack of sub-additivity of the sequence $\{X_r\}_{r\in\mathbb{N}}$ (which is super-additive instead), which necessitates considering a certain union bound; coping with the entropy from the union bound leads to the introduction of the factor of $(\log\theta)^{-1/2}$ in the exponent. We discuss this further in Section 7.3.

Remark 7.10 ($\zeta(\alpha) \to 0$ as $\alpha \to 0$)**.** The tail exponent claimed in Theorem 7.4 holds only for $\theta \leq r^{\zeta}$ for a positive $\zeta = \zeta(\alpha)$ with $\lim_{\alpha \to 0} \zeta(\alpha) = 0$, and as we will see now, this is indeed necessary. First note that Assumption 3 implies that the vertex weight distribution's upper tail decays with exponent at least α; to see this, observe that $\mathbb{P}(X_{(r-1,r)} - \mathbb{E}[X_{(r-1,r)}] > -0.5tr^{1/3}) > 1/2$ for all large enough t by Assumption 3b, and so

$$\begin{aligned}\frac{1}{2} \cdot \mathbb{P}\left(\xi_{(r,r)} \geq tr^{1/3}\right) &\leq \mathbb{P}\left(X_{(r-1,r)} - \mathbb{E}[X_{(r-1,r)}] > -0.5tr^{1/3}, \xi_{(r,r)} \geq tr^{1/3}\right) \\ &\leq \mathbb{P}\left(X_r - \mathbb{E}[X_r] \geq 0.25tr^{1/3}\right) \leq \exp(-ct^{\alpha}),\end{aligned}$$

using Assumption 3a in the last inequality, and bounding $\mathbb{E}[X_r] - \mathbb{E}[X_{(r-1,r)}]$ by $0.25tr^{1/3}$. This holds for all r and t large enough; taking $r = r_0$ large enough for the bound to hold and letting ξ be any random variable distributed according to ν shows that, for all large enough t,

$$\mathbb{P}\big(\xi \geq tr_0^{1/3}\big) \leq \exp(-ct^{\alpha}) \implies \mathbb{P}\big(\xi \geq t\big) \leq \exp(-\tilde{c}t^{\alpha}).$$

Conversely, assuming that $\mathbb{P}\big(\xi \geq t\big) \geq \exp(-\tilde{c}t^{\alpha})$, it follows that Assumption 3a cannot hold with any power $\beta > \alpha$ for the entire tail. Now recall, as mentioned after Theorem 7.2, that after a certain point the behavior of individual vertex weights is expected to govern the tail of point-to-point weights. So under the aforementioned assumption on ξ, an upper bound for $\zeta(\alpha)$ could be obtained by considering the value of ζ which solves

$$\exp(-c\theta^{3/2}) = \exp(-c(\theta r^{1/3})^{\alpha})$$

for $\theta = r^{\zeta}$, which is $\zeta = 2\alpha/(9 - 6\alpha)$. This goes to zero as $\alpha \to 0$, as in Theorem 7.4.

Remark 7.11 (Intermediate regimes for upper tail)**.** While Theorem 7.4 asserts the $3/2$ tail exponent up till r^{ζ}, its proof will also show the existence of a number of ranges of θ in the interval $[r^{\zeta}, \infty)$ in which the tail exponent transitions from $3/2$ to α. More precisely, there exists a finite n and numbers $\alpha = \beta_1 < \beta_2 < \ldots < \beta_n = 3/2$ and $\infty = \zeta_1 > \zeta_2 > \ldots > \zeta_n = \zeta$ such that, for $j \in [\![1, n-1]\!]$ and $\theta \in [r^{\zeta_{j+1}}, r^{\zeta_j}]$,

$$\mathbb{P}\left(X_r - \mathbb{E}[X_r] \geq \theta r^{1/3}\right) \leq \exp\left(-c\theta^{\beta_j}\right).$$

Recursive expressions are also derived for the β_j and ζ_j quantities; see Remark 9.2.

However, we believe that these intermediate regimes are an artifact of our proof, and that the true behavior is that the tail $\exp(-c\theta^{3/2})$ holds for θ till $r^{2\alpha/(9-6\alpha)}$, and $\exp(-c(\theta r^{1/3})^{\alpha})$ after (as in Remark 7.10). Note also that for $\alpha = 1$, $r^{2\alpha/(9-6\alpha)} = r^{2/3}$, matching Theorem 7.2.

Remark 7.12 (Extending to other values of z)**.** We have stated our results for the last passage value to (r, r), but some also extend to $(r - z, r + z)$ for certain ranges of z. For the upper tail the argument of Theorem 7.4 also applies for $|z| = O(r^{2/3})$, while Theorem 7.5 extends to all $|z| = O(r^{5/6})$; as mentioned after the assumptions, the source of the $5/6$ is that for z of this order, the upper and lower bounds of Assumption 2 differ by the weight fluctuation order,

i.e., $r^{1/3}$. Regarding the upper bound on the lower tail, the argument for Theorem 7.6 does not conceptually rely on $z = 0$, but formally uses a result from [BGHH20] which is not proven for $z \neq 0$. The latter result can be extended to larger z without much difficulty, but we have not pursued this here. Finally, the argument for Theorem 7.7 applies for $|z| \leq r^{5/6}$.

The set of assumptions we adopt bears similarities to the ones that have appeared in the past literature on FPP, some of which we discussed in Remark 7.8. The most prominent of these include the work of Newman and coauthors (see e.g. [NP95, ADH17]) which investigated the effect of limit shape curvature assumptions on the geometry of geodesics and the fluctuation exponents. More recently, the previously mentioned work [Cha13] of Chatterjee assumed a strong form of existence of the exponents governing geometric and weight fluctuations of the geodesics and verified the KPZ relation between them; see also [AD14]. Subsequently [DH14, Ale20, Gan20] studied geodesics and bi-geodesics under related assumptions.

Inspired by this, recently, results in the exactly solvable cases of LPP have been obtained, relying merely on inputs analogous to the ones stated in the assumptions. See, for example, the very recent work [BGHH20] which develops the theory of *geodesic watermelons* under a similar set of assumptions to deduce properties of all known integrable lattice models. Other examples include [BHS18, BSS14, FO18], which work in the specific case of LPP with exponential weights; and [Ham19b, Ham19c, Ham19d], in which geometric questions in the semi-discrete model of Brownian LPP are studied.

An intriguing and novel aspect of our arguments is the use of the concentration of measure phenomena for sums of independent stretched exponential random variables, which is in fact at the heart of our arguments. General concentration results have, of course, been widely investigated in recent times [BLM13], but they have not previously played a central role in studies of LPP. On the other hand, concentration of measure has played a more significant role in FPP. We mention here [DHS14] which proves exponential concentration of the passage time on a subdiffusive scale and the related line of work bounding the variance [Kes86, BKS03, BR08, DHS15]. Also related is [CD13] which proves a central limit theorem for certain constrained first passage times. We point the reader to [ADH17, Section 3] for a more in depth survey.

A common theme in concentration of measure is that sums of independent random variables have behavior which transitions, as we extend further into the tail, from being sub-Gaussian to being governed by the tail decay of the individual variables. When the variables have stretched exponential tails, a precise form of this is a bound that is a generalization of Bernstein's inequality for subexponential random variables. Though such results are not unexpected, the recent article [KC18] explicitly records many extensions of concentration results for sums of sub-Gaussian or subexponential random variables to the stretched exponential case with a high dimensional statistics motivation, in a form particularly convenient for our application.

We next move on to an outline of the key ideas driving our proofs.

7.3 The key ideas

Before turning to the ideas underlying our arguments, we deal with some matters of convention. We will use the words "width" and "height" to refer to measurements made along the antidiagonal and diagonal respectively. So, for example, the set of $(x, y) \in \mathbb{Z}^2$ such that $2 \leq x + y \leq 2r$ and $|x - y| \leq \ell r^{2/3}$ is a parallelogram of height r and width $\ell r^{2/3}$. This usage will continue throughout this part of the thesis.

In the overview we will at certain moments make use of a few refined tools, which have appeared previously in [BGHH20], and whose content is explained informally in this section; their precise statements are gathered in Section 7.6 ahead.

Now we turn to the mathematical discussion. The flavors of our arguments are different for the upper and lower bounds on the two tails. Super-additivity, in various guises, plays a recurring role in all except the upper bound on the lower tail. In all the arguments a parameter k appears which plays different roles, but is essentially always finally set to be a multiple of $\theta^{3/2}$, where θ measures the depth into the tail we are considering. The reader should keep in mind this value of k in the discussion. Also, we assume without loss of generality that $\alpha \leq 1$ in this section.

We briefly give a version of a common theme which underlies the different arguments, namely of looking at smaller scales, which further explains why we take $k = \Theta(\theta^{3/2})$. Consider a geodesic path from $(1, 1)$ to (r, r) which attains a weight of $\mu r + \theta r^{1/3}$ for large θ (the following also makes sense for $-\theta$). If we look at a given $1/k$-fraction of the geodesic, that fraction's weight should be close to $\mu r/k + \theta r^{1/3}/k$ if the geodesic gains weight roughly uniformly across its journey; but on the other hand, KPZ fluctuation dictates that the fraction's weight should typically be $\mu r/k + C(r/k)^{1/3}$. So we look for a scale at which the typical behavior is not in tension with the notion of the geodesic's weight being spread close to uniformly over much of its journey. This means finding k such that $\theta r^{1/3}/k$ and $C(r/k)^{1/3}$ are of the same order, which occurs if $k = \Theta(\theta^{3/2})$.

Now we come to the detailed descriptions.

Upper bound on upper tail.

We start by discussing a simplified argument for the upper tail of the upper bound to illustrate the idea of bootstrapping. The starting point is a concentration of measure phenomenon for stretched exponential random variables alluded to before. More precisely, sums of independent stretched exponential random variables have the same qualitative tail decay deep in the tail as that of a single one (see Proposition 8.1 ahead). Not so deep in the tail lies a regime of Gaussian decay, but we will never be in this regime in our arguments.

Let $X^{(i)}_{r/k}$ be the last passage value from $i(r/k, r/k) + (1, 0)$ to $(i + 1)(r/k, r/k)$. Suppose, for purposes of illustration, that we actually had that X_r are *sub*-additive rather than super-additive,

i.e., we had that $X_r \le \sum_{i=1}^{k} X_{r/k}^{(i)}$. Each $X_{r/k}^{(i)}$ fluctuates at scale $(r/k)^{1/3}$, and

$$\left| \sum_{i=1}^{k} \mathbb{E}[X_{r/k}^{(i)}] - \mathbb{E}[X_r] \right| \le k \cdot C(r/k)^{1/3} = Ck^{2/3} r^{1/3}, \tag{7.3}$$

using Assumption 2. So under this illustrative sub-additive assumption we would have

$$\begin{aligned} \mathbb{P}\left(X_r - \mathbb{E}[X_r] > \theta r^{1/3}\right) &\le \mathbb{P}\left(\sum_{i=1}^{k} (X_{r/k}^{(i)} - \mathbb{E}[X_{r/k}^{(i)}]) > \theta r^{1/3} - Ck^{2/3} r^{1/3} \right) \\ &\le \mathbb{P}\left(\sum_{i=1}^{k} (X_{r/k}^{(i)} - \mathbb{E}[X_{r/k}^{(i)}]) > \frac{1}{2}\theta k^{1/3} (r/k)^{1/3} \right), \end{aligned} \tag{7.4}$$

the last inequality for $k \le (2C)^{-3/2}\theta^{3/2}$, which dictates our choice of k. Now by Assumption 3a we know that

$$\begin{aligned} &\mathbb{P}\left(X_{r/k,i} - \mathbb{E}[X_{r/k,i}] > \theta (r/k)^{1/3}\right) \le \exp(-c\theta^{\alpha}) \\ \implies &\mathbb{P}\left(X_{r/k,i} - \mathbb{E}[X_{r/k,i}] > \theta k^{1/3} (r/k)^{1/3}\right) \le \exp(-c\theta^{\alpha} k^{\alpha/3}). \end{aligned}$$

Because sums of stretched exponentials have the same deep tail decay as a single one, (7.4) shows that the probability that $X_r - \mathbb{E}[X_r]$ is greater than $\theta r^{1/3}$ is essentially like that of $X_{r/k} - \mathbb{E}[X_{r/k}]$ being greater than $\theta k^{1/3}(r/k)^{1/3}$, which is at most $\exp(-c\theta^{\alpha} k^{\alpha/3})$. This gives an improved tail exponent of $3\alpha/2$ for the point-to-point's upper tail, compared to the input of α, since k can be at most $O(\theta^{3/2})$.

We can now repeat this argument, with the improved exponent as the input, and obtain an output exponent which is greater by a factor of $3/2$, and we can continue doing so as long as the input exponent is at most 1. If we perform the argument one last time with the input exponent as 1, we obtain the optimal exponent of $3/2$.

The reason we require the input exponent to be at most 1 is that, beyond this point, the concentration behavior changes: for $\alpha \le 1$ the deep tail behavior of a sum of independent stretched exponentials is governed by the event that a single variable is large, while for $\alpha > 1$ the behavior is governed by the event that the deviation is roughly equidistributed among all the variables. This is a result of the change of the function x^{α} from being concave to convex as α increases beyond 1. More precisely, suppose $\alpha \in (1, 3/2]$ is the point-to-point tail exponent and let us accept the equidistributed characterization of the deep tail (as is proved in [KC18]). Then the probability (7.4) would be at most the probability that each of the k variables $X_{r/k}^{(i)} - \mathbb{E}[X_{r/k}^{(i)}]$ is at least $(\theta k^{1/3}/k)(r/k)^{1/3} = \theta k^{-2/3}(r/k)^{1/3}$, which is in turn bounded by

$$\exp\left(-ck \cdot (\theta k^{-2/3})^{\alpha}\right) = \exp\left(-c\theta^{\alpha} k^{1-2\alpha/3}\right).$$

By taking $k = \eta\theta^{3/2}$, which, as mentioned earlier, is the largest possible value we can take, we see that this final expression is $\exp(-c\theta^{3/2})$. In other words, the exponent of $3/2$ is a natural fixed point for the bootstrapping procedure.

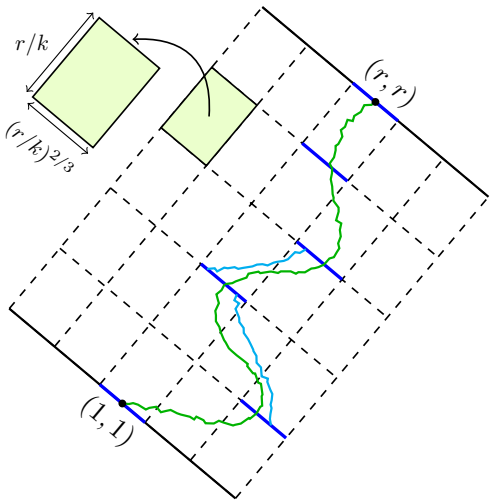


Figure 7.1: In green is depicted the heaviest path which passes through the selection of intervals in blue. The cyan curve between the second and third (similarly the third and fourth) blue intervals is the heaviest path with endpoints on those intervals. Because these consecutive cyan paths do not need to share endpoints, the weight of the green path is at most the sum of the interval-to-interval weights defined by the blue intervals, which provides the substitute sub-additive relation.

Now we turn to addressing the simplifications we made in the above discussion. Handling them correctly makes the argument significantly more complicated and technical, and reduces the tail from $\theta^{3/2}$ to $\theta^{3/2}(\log\theta)^{-1/2}$.

One simplification we skipped over is that the improvement in the tail bound after one iteration only holds for $\theta \leq r^{2/3}$ and not the entire tail (since k, the number of parts that the geodesic to (r, r) is divided into, can be at most r, and $k = \Theta(\theta^{3/2})$), which is a slight issue for the next round of the iteration. This is handled by a simple truncation.

But the main difficulty is that the X_r are super-additive, not sub-additive. To handle this, we consider a grid of height r and width $\text{poly}(\theta) \cdot r^{2/3}$. This width is set such that, with probability at most $\exp(-c\theta^{3/2})$, the geodesic exits the grid, using the bound recorded in Proposition 7.15 ahead on the transversal fluctuation; this allows us to restrict to the event that the geodesic stays within the grid. Intervals in the grid have width $(r/k)^{2/3}$ and are separated by a height of r/k.

The utility of the grid is that X_r can be bounded by a sum of interval-to-interval weights in terms of the intervals of the grid that the geodesic passes through; this bound can play the role of a sub-additive relation. See Figure 7.1. Then, just as we had a tail bound above for $X^{(i)}_{r/k}$ to bootstrap, a requisite step is to obtain an upper bound on the upper tail of the interval-to-interval weight, using only the point-to-point estimate available. We do this in Lemma 9.5 with the basic idea that the interval-to-interval weight being high will cause a point-to-point weight, from "backed up" points, to also be high; see Figure 9.2 for a more detailed illustration of the argument (such an argument of backing up has previously been implemented in [BSS14, BGHH20]).

With the interval-to-interval tail bound, we discretize the geodesic by considering which sequence of intervals it passes through, and bound the highest weight through a given sequence by the sum of interval-to-interval weights. This uses the bootstrapping idea and yields an improved tail estimate for the highest weight through a given sequence. Later we will take a union bound over all possible sequences of intervals; this union bound is what leads to the appearance of the suboptimal $(\log\theta)^{-1/2}$ in the bound as mentioned in Remark 7.9.

This strategy requires handling paths which are extremely "zig-zaggy"; to show that these paths are not competitive, we need upper bounds on upper tails of point-to-point weights, i.e. X_r^z, in a large number of directions indexed by z, though we are only ultimately proving a bound for paths ending at (r,r). (Recall that X_r^z is the weight to $(r-z,r+z)$ from $(1,1)$.) Further, in order to repeat the iterations of the bootstrap, the bounds in other directions must also be improving with each iteration. To achieve this, we in fact bound the deviations not from $\mathbb{E}[X_r^z]$ (which to second order is $\mu r - Gz^2/r$) in the j^{th} round of iteration, but from the bigger $\mu r - \lambda_j G z^2/r$, for a $\lambda_j \leq 1$ which decreases with the iteration number j. By adopting this relaxation we are able to obtain the improvement in the tail for all the required z with each iteration, which appears to be difficult if one insists that $\lambda_j = 1$ for all j.

A similar grid construction has been used previously, for example to obtain certain tail bounds in [BGHH20], to bound the number of disjoint geodesics in a parallelogram in [BHS18], and to study coalescence of geodesics in [BSS19].

Lower bound on upper tail

This is the easiest of the four arguments. Recall that we have C and δ from Assumption 4a such that $\mathbb{P}(X_{r/k} > \mu r/k + C(r/k)^{1/3}) \geq \delta$, and let $X_{r/k}^{(i)}$ be as in (7.3). By the super-additivity that the X_r genuinely enjoy, for any k it holds that $X_r \geq \sum_{i=1}^k X_{r/k}^{(i)}$. Choosing k to be an appropriate multiple of $\theta^{3/2}$, we obtain

$$\mathbb{P}\left(X_r > \mu r + \theta r^{1/3}\right) \geq \prod_{i=1}^{k} \mathbb{P}\left(X_{r/k}^{(i)} > \mu r/k + C(r/k)^{1/3}\right) \geq \delta^k = \exp(-c\theta^{3/2}).$$

Here we used the independence of $X_{r/k}^{(i)}$, but note that it would have sufficed for our purposes to have that they are positively associated, by the FKG inequality. Replacing μr by $\mathbb{E}[X_r]$ is a simple application of Assumption 2.

Upper bound on lower tail

The illustrative argument using sub-additivity given above for the upper bound on the upper tail is actually correct for the upper bound on the lower tail, as the super-additivity of X_r is in the favorable direction in this case. But, as we saw there, the approach can only bring the tail exponent up to $3/2$, and not 3. This is essentially because that argument focuses on the

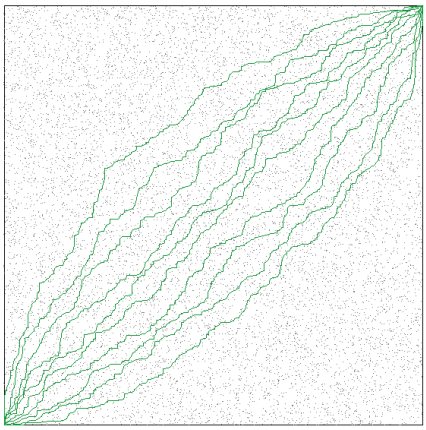


Figure 7.2: A simulation of the k-geodesic watermelon in the related model of Poissonian last passage percolation for $k = 10$.

weight of a single path, while the exponent of 3 for the lower tail is a result of *all* paths having low weight. Thus our strategy to prove the stronger bound is to construct $\theta^{3/2}$ disjoint paths moving through independent parts of the space, each suffering a weight loss of $\theta r^{1/3}$. By the discussion above and independence, the probability of each of them being small can be bounded by $\exp(-c\theta^{3/2} \cdot \theta^{3/2}) = \exp(-c\theta^3)$.

To do this formally, we rely on an important ingredient originally proved in [BGHH20], which studies the weight and geometry of maximal weight collections of k disjoint paths in $[\![1, r]\!]^2$, called *k-geodesic watermelons*. See Figure 7.2. It is shown in [BGHH20] that these k paths typically are each of weight $\mu r - Ck^{2/3}r^{1/3}$, and that they have a collective transversal fluctuation of order $k^{1/3}r^{2/3}$. In this thesis, we will prove the following quantitative bound on the weight X_r^k of the k-geodesic watermelon via a direct multi-scale construction of disjoint paths with correct order collective weight, which we formally state ahead as Theorem 7.16:

$$\mathbb{P}\left(X_r^k \leq \mu kr - Ck^{5/3}r^{1/3}\right) \leq \exp(-ck^2), \tag{7.5}$$

for all $k \leq \eta r$ for a small constant $\eta > 0$. We give the construction in Chapter 10.

For our purposes we observe that, for any $k \in \mathbb{N}$,

$$\mathbb{P}\left(X_r < \mu r - \theta r^{1/3}\right) \leq \mathbb{P}\left(X_r^k < \mu kr - \theta kr^{1/3}\right).$$

Taking $k = \eta\theta^{3/2}$ and noting that then $k\theta$ is of order $k^{5/3}$ and that $\theta < r^{2/3} \implies k < \eta r$ shows that

$$\mathbb{P}\left(X_r < \mu r - \theta r^{1/3}\right) \leq \exp(-ck^2) = \exp(-c\theta^3),$$

and it is a simple matter to replace μr by $\mathbb{E}[X_r]$ by possibly reducing the constant c.

Deferring a discussion of the details of the construction that the proof of (7.5) consists of, the construction requires three inputs. The first is the following:

1. Limit shape bounds, which we have by Assumption 2.

The next two inputs concern the maximum weight over all midpoint-to-midpoint paths constrained to lie in a given parallelogram $U = U_{r,\ell,z}$ of height r, width $\ell r^{2/3}$, and opposite side midpoints $(1, 1)$ and $(r - z, r + z)$. We will call such weights "constrained weights".

2. An exponential upper bound on the constrained weight's lower tail, which we will arrive at by bootstrapping. To elaborate, by using Assumption 3b and the previously mentioned Proposition 7.15 on the transversal fluctuation of the unconstrained geodesic, we can obtain an initial stretched exponential upper bound ((7.7) of Proposition 7.17 ahead) on the constrained weight's lower tail. Then, via a bootstrapping argument as above, we can upgrade this to a tail with exponent $3/2$ (see Proposition 8.3).
3. A lower bound on the mean of constrained weights using the above tail, provided by (7.8) of Proposition 7.17.

A more detailed overview of the construction is provided in Chapter 10.

Lower bound on lower tail

A detail about the construction described, which is captured in its formal statement Theorem 7.16, is that it fits inside a strip of width $4k^{1/3}r^{2/3}$ around the diagonal. To lower bound the lower tail probability, this suggests that we need to focus on paths which remain in the strip of this width (again we will be setting k to be a constant times $\theta^{3/2}$). Essentially this is because a consequence of the parabolic weight loss of Assumption 2 is that *any* path (not just the geodesic) which exits the strip of width $k^{1/3}r^{2/3}$ suffers a loss of $(k^{1/3}r^{2/3})^2/r = k^{2/3}r^{1/3}$, which is of order $\theta r^{1/3}$, with high probability. This is captured more precisely in Theorem 7.14 ahead.

Similar to the argument for the upper bound on the upper tail, we consider a grid where each cell has height r/k and width $(r/k)^{2/3}$, but with overall width $k^{1/3}r^{2/3}$. This gives k cells in each column and in each row, for a total of k^2 cells. See Figure 7.3.

Now consider the event that, for each interval in the grid, the maximum weight from that interval to the next row of intervals is less than $\mu r/k - C(r/k)^{1/3}$, and that the maximum weight of a path which exits the grid is at most $\mu r - Ck^{2/3}r^{1/3}$. This is an intersection of decreasing events, and on this event X_r is at most $\mu r - Ck^{2/3}r^{1/3}$: if it exits the grid it suffers a loss of $Ck^{2/3}r^{1/3}$ and if it stays in the grid it undergoes a loss of at least $C(r/k)^{1/3}$ for each of the k rows. Now if we know that there is a constant probability (say $\delta > 0$) lower bound on the event that a single interval-to-line weight is low, the FKG inequality (along with Theorem 7.14 to lower bound the probability of parabolic weight loss when exiting the grid) provides a lower bound of order δ^{k^2} on the described event's probability; setting k to be a multiple of $\theta^{3/2}$ will complete the proof.

To implement this we need a lower bound on the probability that the interval-to-line weight is small using the point-to-point lower bound of Assumption 4b. This is Lemma 9.9. The proof proceeds in two steps. First, a stepping back strategy as earlier gives a constant lower bound on the interval-to-interval weight's lower tail for intervals of size $\varepsilon r^{2/3}$, for some small $\varepsilon > 0$

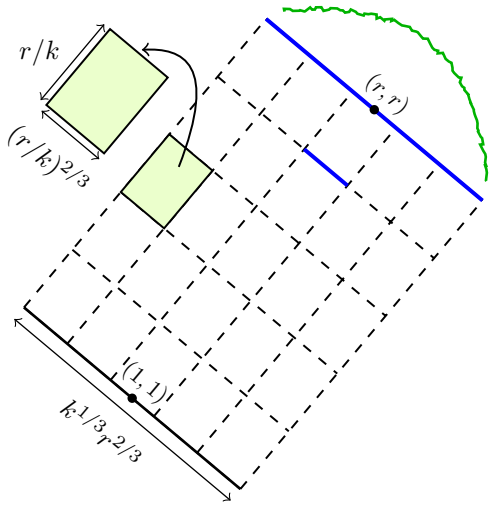


Figure 7.3: The grid of k^2 intervals for the lower bound of the lower tail. An interval and the following row of intervals are blue: consider the event that the heaviest path from the former to the latter is at most $\mu r/k - C(r/k)^{1/3}$. To prove that this has positive probability, we make use of parabolic curvature of the weight profile (shown in green) to argue that if the endpoint on the row is too extreme, it will typically suffer the loss we want; a separate backing up argument is employed for when the endpoint is near the center where the parabolic weight loss is not significant.

(as will be clear from the precise argument, the smallness of ε is crucial for this). By the FKG inequality, this is upgraded to a bound for intervals of length $r^{2/3}$; essentially, if each of the intervals are divided into ε^{-1} intervals of size $\varepsilon r^{2/3}$, and all ε^{-2} pairs of intervals have small weight (which is an intersection of decreasing events), then so must the original intervals. To get from this to an interval-to-line bound we again argue based on FKG. We divide the line into $r^{1/3}$ many intervals of size $r^{2/3}$ each. We can ensure that the weight is low whenever the destination interval is one of a constant number near (r, r) using the previous bound, and for the rest the parabolic curvature ensures that it is so likely to be low that the FKG inequality gives a positive lower bound independent of r, in spite of considering an intersection of $r^{1/3}$ many events; see Figure 7.3.

7.4 Tails for constrained weight

The ideas described in the previous section can be applied slightly more generally to yield the following theorem on the lower tail of the constrained weight X_r^U of the best path from $(1, 1)$ to (r, r) constrained to stay inside a parallelogram U.

Recall that $U = U_{r,\ell,z}$ denotes a parallelogram of height r, width $\ell r^{2/3}$, and opposite side midpoints $(1, 1)$ and $(r - z, r + z)$, defined to be the set of vertices $v = (v_x, v_y) \in \mathbb{Z}^2$ such that $v + t(-1, 1)$ lies on the line $y = \frac{r+z}{r-z} \cdot x$ for some $t \in \mathbb{R}$ with $|t| \leq \ell r^{2/3}/2$, and $2 \leq v_x + v_y \leq 2r$.

Let X_r^U be the maximum weight among all paths from $(1,1)$ to $(r-z, r+z)$ constrained to be inside U. The notation $U_{r,\ell,z}$ will be used for parallelograms throughout Part III of the thesis. Here we take $z = 0$.

Estimates on constrained weights have been crucial in several recent advances, see [BSS14]. The following theorem proves a sharp estimate on the tail as a function of the aspect ratio of the parallelogram, measured on the characteristic KPZ scale.

Theorem 7.13. *Under Assumptions 2, 3, and 4b, there exist finite positive constants c_1, c_2, η, C, θ_0, and r_0 (all independent of ℓ) such that, for $z = 0$, ranges of θ to be specified, $r > r_0$, and $C\theta^{-1} \le \ell \le 2r^{1/3}$,*

$$\exp\left(-c_1 \min(\ell\theta^{5/2}, \theta^3)\right) \le \mathbb{P}\left(X_r^U - \mu r \le -\theta r^{1/3}\right) \le \exp\left(-c_2 \min(\ell\theta^{5/2}, \theta^3)\right);$$

the second inequality holds for $\theta > \theta_0$ while the first holds for $\theta_0 < \theta < \eta r^{2/3}$. If ℓ is bounded below by a constant $\varepsilon > 0$ independent of θ, we can replace μr by $\mathbb{E}[X_r^U]$ for $r > \tilde{r}_0(\varepsilon)$ and with c_1 depending on ε.

We note that Theorem 7.6 and Theorem 7.7 are implied by Theorem 7.13 by taking $\ell = 2r^{1/3}$.

We also remark that the transition from $\ell\theta^{5/2}$ to θ^3 occurs when ℓ becomes of order $\theta^{1/2}$; this matches the belief (which comes from the parabolic curvature) that the geodesic, conditioned on its weight being less than $\mu r - \theta r^{1/3}$, will have typical transversal fluctuations of order $\theta^{1/2} r^{2/3}$.

The proof idea of Theorem 7.13 is a refinement of those of Theorems 7.6 and 7.7 described above, by picking the number of paths of average separation $k^{-2/3} = \theta^{-1}$ to be packed inside U, which turns out to be $\min(\ell k^{2/3}, k)$ (rather than k as before). We omit further outline to avoid repetition.

7.5 Related work

The main tools we use in our arguments are the super-additivity of the X_r (i.e., $X_{r+j} \ge X_r + X_{(r+1,r),(r+j,r+j)}$), geodesic watermelons, and concentration of measure results for sums of independent stretched exponential random variables. We have discussed aspects of the latter two that have appeared in various works, and here we briefly overview the first, i.e., super-additivity.

Not surprisingly, super-additivity of the weight has been an important tool in other investigations of non-integrable models; for example, the proof of the almost sure existence of a deterministic limit for X_r/r as $r \to \infty$ under a wide class of vertex distributions goes via Kingman's sub-additive theorem. Super-additivity was also crucial in [Led18], where a law of iterated logarithm for X_r was proved. More precisely, for exponential weights, $\limsup_{r\to\infty}(X_r - 4r)/r^{1/3}$ was shown to almost surely exist and be a finite, positive, deterministic constant. Super-additivity only aids in proving a result for the lim sup, and so the result on the lim inf in [Led18] is weaker. This was addressed in [BGHK19], where the lack of sub-additivity was handled by shifting perspective to also consider point-to-line passage times, which, as we have outlined, we will do in the arguments presented here as well. A usage of super-additivity was also made by Seppäläinen in [Sep98a]

and shortly after by Johansson in [Joh00], where, from a limiting large deviation theorem for the upper tail, it was pointed out that the same gives an explicit bound for finite r by a super-additivity argument. Briefly, and again in the context of Exponential LPP, the observation is that for every r and every $N \geq 1$,

$$\mathbb{P}\big(X_r > \theta\big)^N \leq \mathbb{P}\big(X_{Nr} > N\theta\big) \implies \mathbb{P}\big(X_r > \theta\big) \leq \lim_{N\to\infty} \big[\mathbb{P}\big(X_{Nr} > N\theta\big)\big]^{1/N}, \tag{7.6}$$

and the latter limit was shown to exist and explicitly identified in [Sep98a]. In a sense our arguments are dual to that of (7.6); while (7.6) uses super-additivity to go to *larger* r in order to obtain a bound, our arguments use super-additivity to reason about *smaller* r to obtain a bound.

Finally, we mention the recent work [EJS20] which proves a sharp upper bound (i.e., with the correct coefficient of $4/3$ as in (7.1)) on the right tail of X_r (centred by $\mu r = 4r$ and appropriately scaled) in exponential LPP via more probabilistic arguments, rather than precise analysis of integrable formulas. The technique utilizes calculations in an increment-stationary version of exponential LPP (where the vertex weight on the boundaries of $\mathbb{Z}^2_{\geq 0}$ differ in distribution from the rest) and a moment generating function identity specific to this model—features absent in the general setting under consideration here.

7.6 A few important tools

In this section we collect some refined tools for last passage percolation which we will use for our arguments as outlined in Section 7.3. There are four statements: the first asserts that it is typical for a path to suffer a weight loss which is quadratic in its transversal fluctuation, measured in the characteristic scalings of $r^{1/3}$ and $r^{2/3}$; the second is a related transversal fluctuation bound, but for paths with endpoint $(r-z, r+z)$ for $|z| \leq r^{5/6}$; the third is a high probability construction of a given number of disjoint paths which achieve a good collective weight; and the fourth provides bounds on the lower tail and mean of constrained weights.

We will import the proof ideas from [BGHH20] where similar statements have appeared. Our proofs are essentially the same but adapted suitably to work under the weaker tail exponent α assumed here; for this reason, we only explain the modifications that need to be made for the first, second and fourth tools in Appendix A. The proof of the third tool is provided in Chapter 10.

Parabolic weight loss for paths with large transversal fluctuation

The following is the precise statement of the first tool.

Theorem 7.14 (Refined transversal fluctuation loss). *Let $X_r^{s,t}$ be the maximum weight over all paths Γ from the line segment joining $(-tr^{2/3}, tr^{2/3})$ and $(tr^{2/3}, -tr^{2/3})$ to the line segment joining $(r - tr^{2/3}, r + tr^{2/3})$ and $(r + tr^{2/3}, r - tr^{2/3})$ such that $\mathrm{TF}(\Gamma) > (s+t)r^{2/3}$, with $t \leq s$. Under*

Assumptions 2 and 3a, there exist absolute constants r_0, s_0, $c > 0$ and $c_2 > 0$ such that, for $s > s_0$ and $r > r_0$,

$$\mathbb{P}\left(X_r^{s,t} > \mu r - c_2 s^2 r^{1/3}\right) < \exp\left(-cs^{2\alpha}\right).$$

The proof of this follows that of [BGHH20, Theorem 3.6]. We explain the necessary modifications in the appendix.

An important feature of Theorem 7.14 is that it bounds the probability of a *decreasing* event, which is useful as it allows the application of the FKG inequality.

Transversal fluctuation bound for $|z| \leq r^{5/6}$

The second tool is a result on the transversal fluctuation of geodesics to $(r - z, r + z)$ (note that Theorem 7.14 is related but only for $z = 0$), which is the following. We note in passing that the event of the geodesic having large transversal fluctuation is neither increasing nor decreasing.

Proposition 7.15 (Transversal fluctuations)**.** *For given z, let Γ_r^z be the geodesic from $(1, 1)$ to $(r - z, r + z)$ with maximum transversal fluctuation. Under Assumptions 2 and 3, there exist constants $c > 0$, s_0, and r_0 such that, for $r > r_0$, $s > s_0$, and $|z| \leq r^{5/6}$,*

$$\mathbb{P}\left(\mathrm{TF}(\Gamma_r^z) > sr^{2/3}\right) \leq \exp\left(-cs^{2\alpha}\right).$$

The proof of this is similar to that of [BSS14, Theorem 11.1] and appears in the appendix.

A high probability construction of disjoint paths with good collective weight

Here is the statement of our third tool, originally proved as [BGHH20, Theorem 3.1].

Theorem 7.16. *Under Assumptions 2 and 3, there exist $c, C_1 > 0, k_0 \in \mathbb{N}$ and $\eta > 0$ such that for all $k_0 \leq k \leq \eta r$ and $m \in [\![1, k]\!]$, with probability $1 - e^{-ckm}$, there exist m disjoint paths $\gamma_1, \ldots, \gamma_m$ in the square $[\![1, r]\!]^2$, with γ_i from $(1, m - i + 1)$ to $(r, r - i + 1)$ and $\max_i \mathrm{TF}(\gamma_i) \leq 2mk^{-2/3}r^{2/3}$, such that*

$$\sum_{i=1}^{m} \ell(\gamma_i) \geq \mu rm - C_1 mk^{2/3} r^{1/3}.$$

The proof of Theorem 7.16 will be given in Chapter 10 and will require as input our fourth tool on bounds for the lower tail and mean of constrained weights.

Bounds for constrained weights

To state our fourth and final tool, recall from Section 7.4 the notation for parallelograms $U_{r,\ell,z}$ of height r, width $\ell r^{2/3}$ and opposite midpoints $(1, 1)$ and $(r - z, r + z)$ as well as that for maximum weight of paths constrained inside U, X_r^U.

Proposition 7.17 (Lower tail & mean of constrained point-to-point, Proposition 3.7 of [BGHH20])**.** *Let positive constants L_1, L_2, and $K > 0$ be fixed. Let z and ℓ be such that $|z| \leq Kr^{2/3}$ and $L_1 \leq \ell \leq L_2$, and let $U = U_{r,\ell,z}$. There exist positive constants $r_0 = r_0(K, L_1, L_2)$ and $\theta_0 = \theta_0(K, L_1, L_2)$, and an absolute positive constant c, such that, for $r > r_0$ and $\theta > \theta_0$,*

$$\mathbb{P}\left(X_r^U \leq \mu r - \theta r^{1/3}\right) \leq \exp\left(-c\ell^{2\alpha/3}\theta^{2\alpha/3}\right). \tag{7.7}$$

As a consequence, there exists $C = C(K, L_1, L_2)$ such that, for $r > r_0$,

$$\mathbb{E}[X_r^U] \geq \mu r - Gz^2/r - Cr^{1/3}. \tag{7.8}$$

To be consistent with previous expressions we have included the parabolic term $-Gz^2/r$ in the previous, but note that for the ranges of z mentioned we can absorb it into the $Cr^{1/3}$ term.

7.7 Organization of Part III

In Chapter 8 we collect the concentration statements for stretched exponential random variables and prove an abstracted version of the bootstrap. In Chapter 9 we prove the main theorems. Finally, in Chapter 10, we construct a collection of disjoint high-weight paths and so prove Theorem 7.16.

Chapter 8

Concentration tools and the bootstrap

In this section we collect the concentration inequality for stretched exponential random variables from [KC18] and prove a slightly more flexible version which is more suitable for our applications. We then move to stating a general version of one iteration of the bootstrap, which will both illustrate the basic mechanism and be used later in Section 9.3.

To set the stage, let $\alpha \in (0, 1]$ and suppose Y_i are independent mean zero random variables which satisfy, for some $L, M < \infty$,

$$\inf\left\{\eta > 0 : \mathbb{E}\left[g_{\alpha,L}\left(\frac{|Y_i|}{\eta}\right)\right] \leq 1\right\} \leq M, \tag{8.1}$$

where $g_{\alpha,L}(x) = \exp\left(\min\{x^2, (x/L)^\alpha\}\right) - 1$. The above condition is equivalent to the finiteness of a certain Orlicz norm introduced in [KC18]; see Definition 2.3 and Proposition A.1 therein. The use of Orlicz norms to prove concentration inequalities is well known; see for example [Ver18, Wai19]. The reader not familiar with this notion can keep in mind mean zero random variables Y_i with the property that, for some $c > 0$ and C, and all $t \geq 0$,

$$\mathbb{P}(|Y_i| \geq t) \leq C \exp(-ct^\alpha), \tag{8.2}$$

which are known to satisfy (8.1).

Proposition 8.1. *Given the above setting, there exists $c = c(M, L) > 0$ such that for all $t \geq 0$ and all $k \in \mathbb{N}$,*

$$\mathbb{P}\left(\left|\sum_{i=1}^{k} Y_i\right| \geq t\right) \leq \begin{cases} 2\exp\left(-\dfrac{ct^2}{k}\right) & 0 \leq t \leq k^{1/(2-\alpha)} \\ 2\exp\left(-ct^\alpha\right) & t \geq k^{1/(2-\alpha)}. \end{cases}$$

These two regimes capture the transition from the Gaussian behavior in the immediate tail to stretched exponential behavior deep into the tail.

Proof of Proposition 8.1. [KC18, Theorem 3.1] and the discussion after Remark 2.1 therein imply that, for some constants C and $c > 0$ (depending on M and L), for all $t \geq 0$,

$$\mathbb{P}\left(\left|\sum_{i=1}^{k} Y_i\right| \geq C(\sqrt{kt} + t^{1/\alpha})\right) \leq 2\exp(-ct).$$

Evaluating the transition point where $\sqrt{kt} = t^{1/\alpha}$ yields the statement of Proposition 8.1 by modifying the value of c in the previous display. □

In our applications, we will only have an upper tail bound and hence not a direct verification of the hypothesis (8.1) which needs two sided bounds as in (8.2). It will also at times be convenient to center the variables not by their expectation but by some other constant for which a tail bound is available. These two aspects are handled in the next lemma.

Lemma 8.2. *Suppose $k \in \mathbb{N}$, $\{Y_i : i \in [\![1, k]\!]\}$ are independent, and there exist constants ν_i, $\alpha \in (0, 1]$, and $c > 0$ such that, for $t > t_0$, and $i \in [\![1, k]\!]$,*

$$\mathbb{P}\left(Y_i - \nu_i \geq t\right) \leq \exp(-ct^{\alpha}). \tag{8.3}$$

Then there exist constants $c_1 = c_1(c, \alpha, t_0)$ and $c' = c'(c, \alpha) > 0$ such that, for $t \geq 0$ and all $k \in \mathbb{N}$,

$$\mathbb{P}\left(\sum_{i=1}^{k}(Y_i - \nu_i) > t + kc_1\right) \leq \begin{cases} 2\exp\left(-\dfrac{c't^2}{k}\right) & 0 \leq t \leq k^{1/(2-\alpha)} \\ 2\exp\left(-c't^{\alpha}\right) & t \geq k^{1/(2-\alpha)}. \end{cases}$$

Proof. Let W_i be independent positive random variables whose distribution is defined by $\mathbb{P}(W_i > t) = \exp(-ct^{\alpha})$ for $t \geq 0$. Then the hypothesis on Y_i implies that $Y_i - \nu_i$ is stochastically dominated by $W_i + t_0$, and hence there is a coupling of the Y_i and W_i over all i simultaneously such that

$$Y_i - \nu_i \leq W_i + t_0,$$

by standard coupling arguments. It is a calculation that $\mathbb{E}[W_i] = \alpha c^{-1/\alpha}\Gamma(\alpha)$, where $\Gamma(z) = \int_0^{\infty} x^{z-1}e^{-x}\,\mathrm{d}x$ is the gamma function. Thus we get

$$P\left(\sum_{i=1}^{k}(Y_i - \nu_i) > t + kc_1\right) \leq \mathbb{P}\left(\sum_{i=1}^{k}(W_i - \mathbb{E}[W_i]) > t + k(c_1 - t_0 - \alpha c^{-1/\alpha}\Gamma(\alpha))\right).$$

Setting $c_1 = t_0 + \alpha c^{-1/\alpha}\Gamma(\alpha)$ and applying Proposition 8.1 completes the proof of Lemma 8.2, under the condition that $W_i - \mathbb{E}[W_i]$ satisfies (8.1) for some L and M depending only on α and c. We verify this next. [KC18, Proposition A.3] asserts that for any random variable Y satisfying, for all $t \geq 0$,

$$\mathbb{P}(|Y| \geq t) \leq 2\exp(-\tilde{c}t^{\alpha}), \tag{8.4}$$

there exist M and L, depending on α and $\tilde{c}$, such that (8.1) holds with Y in place of Y_i. Therefore it is sufficient to verify (8.4) for $Y = W_i - \mathbb{E}[W_i]$ for some $\tilde{c}$ depending on α and c. Since W_i is positive for each i, we have the bound

$$\mathbb{P}\Big(|W_i - \mathbb{E}[W_i]| \geq t\Big) \leq \begin{cases} 1 & 0 \leq t \leq \mathbb{E}[W_i] \\ \exp(-ct^\alpha) & t > \mathbb{E}[W_i] \end{cases},$$

which implies that (8.4) holds with $\tilde{c} = \min(c, \log 2) \cdot (\mathbb{E}[W_i])^{-\alpha}$, since $2\exp(-\tilde{c}t^\alpha) \geq 1$ for $0 \leq t \leq \mathbb{E}[W_i]$. Note that $\tilde{c}$ depends on only α and c. This completes the proof of Lemma 8.2. □

With the concentration tool Lemma 8.2 in hand, we next present the driving step of the bootstrapping argument. It is the formal statement and proof of one step of the iteration under a *sub*-additive assumption. As indicated in the outline of proof section, since X_r are *super*-additive, this will not be of use for the upper bound on the upper tail; but it will find application in the upper bound on the lower tail, where super-additivity is the favourable direction.

Proposition 8.3. *Suppose that for each $r, k \in \mathbb{N}$ with $k \leq r$, $\{Y_{r,i}^{(k)} : i \in [\![1, k]\!]\}$ is a collection of independent random variables. Suppose also that there exist $\alpha \in (0, 1]$, $c > 0$, r_0, and θ_0 such that, for $r \in \mathbb{N}$, $k \in \mathbb{N}$, $i \in [\![1, k]\!]$, and $\theta \in \mathbb{R}$ such that $r/k > r_0$ and $\theta > \theta_0$,*

$$\mathbb{P}\left(Y_{r,i}^{(k)} > \theta(r/k)^{1/3}\right) \leq \exp\left(-c\theta^\alpha\right). \tag{8.5}$$

Finally, let Y_r be a random variable such that $Y_r \leq \sum_{i=1}^k Y_{r,i}^{(k)}$ for any $k \in \mathbb{N}$ satisfying $r/k > r_0$. Then there exist $\widetilde{\theta}_0 = \widetilde{\theta}_0(c, \alpha, \theta_0, r_0)$ and $c' = c'(c, \alpha, \theta_0, r_0)$ such that, for $\widetilde{\theta}_0 < \theta < r^{2/3}$ and $r > r_0$,

$$\mathbb{P}\left(Y_r > \theta r^{1/3}\right) \leq \exp\left(-c'\theta^{3\alpha/2}\right).$$

Proposition 8.3 is written in a slightly more general way, without explicit reference to the LPP context it will be applied, to highlight the features of LPP that are relevant. In its application Y_r will be the weight of the heaviest path constrained to be in a certain parallelogram of height r, centred by μr, and $Y_{r,i}^{(k)}$ will be weights when constrained to be in disjoint subparallelograms of height r/k, centred by $\mu r/k$.

Finally, we mention a rounding convention we now adopt for the rest of this part of the thesis: the quantities k and r/k should always be integers and, when expressed as real numbers, will be rounded down without comment. The discrepancies of ± 1 which so arise will be absorbed into universal constants.

Proof of Proposition 8.3. By the bound $Y_r \leq \sum_{i=1}^k Y_{r,i}^{(k)}$, for every $r, k \in \mathbb{N}$ with $k \leq r$,

$$\mathbb{P}\left(Y_r > \theta r^{1/3}\right) \leq \mathbb{P}\left(\sum_{i=1}^k Y_{r,i}^{(k)} > \theta r^{1/3}\right). \tag{8.6}$$

We will choose $k = \eta\theta^{3/2}$ for some $\eta \in (0,1)$, a form which is guided by our desire to apply the concentration bound Lemma 8.2 with its input bound (8.3) provided by the hypothesis (8.5) of Proposition 8.3; we also need $k \geq 1$. The first consideration will determine an acceptable value for η via its development as the following two constraints:

1. Lemma 8.2 introduces a linear term kc_1, which, when multiplied by the scale $(r/k)^{1/3}$ of the $Y_{r,i}^{(k)}$ indicated by (8.5), is $c_1 k^{2/3} r^{1/3}$; we want this to be smaller than a constant, say $\frac{1}{2}$, times $\theta r^{1/3}$. Note that c_1 depends on α, c, and θ_0.
2. We require $r/k > r_0$ to apply the hypothesis of Proposition 8.3.

These two constraints, and that $\theta < r^{2/3}$ by hypothesis, force η to be smaller than r_0^{-1} and $2^{-3/2}c_1^{-3/2}$. We pick an η which satisfies these inequalities; thus η depends on c_1 and r_0. Set $\tilde{\theta}_0 = \eta^{-2/3}$; then $\theta \geq \tilde{\theta}_0$ implies $k \geq 1$. We will apply Lemma 8.2 with $Y_i = Y_{r,i}^{(k)}(r/k)^{-1/3}$, $\nu_i = \mu(r/k)^{2/3}$, and $t = \frac{1}{2}\theta k^{1/3}$. For $\theta \geq \tilde{\theta}_0$ and for a $\tilde{c}$ depending on only c and α,

$$\begin{aligned}
\mathbb{P}\left(Y_r > \theta r^{2/3}\right) &\leq \mathbb{P}\left(\sum_{i=1}^{k} Y_{r,i}^{(k)} > \frac{1}{2}\theta k^{1/3}\left(\frac{r}{k}\right)^{1/3} + kc_1\left(\frac{r}{k}\right)^{1/3}\right) \\
&\leq \begin{cases} 2\exp\left(-\tilde{c}\theta^2 k^{-1/3}\right) & \tilde{\theta}_0 k^{1/3} \leq \theta k^{1/3} \leq k^{1/(2-\alpha)} \\ 2\exp\left(-\tilde{c}\theta^{\alpha} k^{\alpha/3}\right) & \theta k^{1/3} \geq \max(\tilde{\theta}_0, k^{1/(2-\alpha)}) \end{cases} \qquad \text{(applying Lemma 8.2)} \\
&\leq 2\exp\left(-\tilde{c}\eta^{\alpha/3}\theta^{3\alpha/2}\right);
\end{aligned}$$

in the final line we have taken the second case of the preceding line. This is because $\alpha \leq 1$ implies $k^{1/(2-\alpha)} \leq k$, and the choice of k (and that $\eta < 1$) ensures that $\theta k^{1/3} \geq k$; so the second case holds, since we have already assumed $\theta > \tilde{\theta}_0$.

The proof of Proposition 8.3 is complete by absorbing the factor of 2 in the final display into the exponential, which we do by setting c' to $\tilde{c}\eta^{\alpha/3}/2$ and increasing $\tilde{\theta}$ (if needed), depending on c', so that $\exp(-c'(\tilde{\theta}_0)^{3\alpha/2}) \leq 1/2$. □

Chapter 9

The tail bound proofs

9.1 Upper tail bounds

In this section we prove Theorems 7.4 and 7.5, respectively the upper and lower bounds on the upper tail.

Upper bound on upper tail

As mentioned in Section 7.3, for the argument for the upper bound on the upper tail, we need a sub-additive relation, instead of the natural super-additive properties that point-to-point weights exhibit. To bypass this issue, we discretize the geodesic and bound the weights of the discretizations by interval-to-interval weights, which do have a sub-additive relation with the point-to-point weight; this allows us to appeal to a form of the basic bootstrapping argument outlined around (7.3); Then performing a union bound over all possible discretizations will complete the proof.

We next state a version of one iteration of the bootstrap for the upper bound on the upper tail. There are a number of parameters which we will provide more context for after the statement.

Proposition 9.1. *Let $\lambda_j = \frac{1}{2} + \frac{1}{2^j}$. Suppose there exist $\alpha \in (0,1]$, $\beta \in [\alpha, 1]$, $\zeta \in (0, \infty]$, $j \in \mathbb{N}$ and constants $c > 0$, θ_0, and r_0 such that, for $\theta > \theta_0$, $r > r_0$, and $|z| \le r^{5/6}$,*

$$\mathbb{P}\left(X_r^z \ge \mu r - \lambda_j \frac{Gz^2}{r} + \theta r^{1/3}\right) \le \begin{cases} \exp\left(-c\theta^\beta\right) & \theta_0 < \theta < r^\zeta \\ \exp\left(-c\theta^\alpha\right) & \theta \ge r^\zeta. \end{cases} \tag{9.1}$$

Let $\zeta' = \min\left(\frac{\alpha\zeta}{1+\alpha\zeta} \cdot \frac{3-\beta}{3\beta},\ \frac{2\alpha}{9+16\alpha}\right)$, with $\frac{\alpha\zeta}{1+\alpha\zeta}$ interpreted as 1 if $\zeta = \infty$. There exist $c' = c'(c, \alpha, \beta, j) > 0$, $\theta_0' = \theta_0'(\theta_0, c, \alpha, \beta, j)$, and $r_0' = r_0'(\alpha, j, r_0)$ such that, for $\theta > \theta_0'$, $r > r_0'$, and $|z| \le r^{5/6}$,

$$\mathbb{P}\left(X_r^z \ge \mu r - \lambda_{j+1} \frac{Gz^2}{r} + \theta r^{1/3}\right) \le \begin{cases} \exp\left(-c'\theta^{\frac{3\beta}{3-\beta}} (\log\theta)^{-\frac{\beta}{3-\beta}}\right) & \theta_0 < \theta < r^{\zeta'} \\ \exp(-c'\theta^\alpha) & \theta \ge r^{\zeta'}. \end{cases}$$

In particular, the input (9.1) *with parameters* $(\alpha, \beta, \zeta, j)$ *gives as output the same inequality with parameters* $(\alpha, \beta', \zeta', j+1)$*, where* $\beta' > \beta$ *may be taken to be* $\frac{3-\beta/2}{3-\beta} \cdot \beta$ *in order to absorb the logarithmic factor.*

We first explain in words the content of the above result and describe the role of the various quantifiers appearing in the statement.

The range of z Though Theorem 7.4 is stated only for $z = 0$, the discretization of the geodesic we adopt demands that we have the bootstrap improve the tail bound in a number of directions, defined by $|z| \le r^{5/6}$, in order to handle the potential "zig-zaggy" nature of the geodesic. Here we choose to consider $|z|$ till $r^{5/6}$ as till this level the lowest order term Hz^4/r^3 in Assumption 2 is at most of the order of fluctuations, namely $r^{1/3}$.

The role of λ_j One may expect to be able to obtain an improved tail for deviation from the expectation, which is $\mu r - Gz^2/r$ up to smaller order terms. However, for technical reasons, this proves to be difficult; we say a little more about this in the caption of Figure 9.2. Instead, Proposition 9.1 proves a bound for the deviation only from a point away from the expectation, reflected by the factor λ_j in front of the parabolic term, which decreases as j increases. Nonetheless, this weaker bound suffices for our application: the relaxation has no effect for the $z = 0$ direction asserted by Theorem 7.4 since the parabolic term is always zero in that case.

The role of ζ Notice that in the hypothesis (9.1) we allow two tail behaviors (with tail exponents α and β) for X_r^z in different regimes, with boundary at r^ζ. This is to allow the use of the conclusion of Proposition 9.1, which only improves the tail exponent for θ up to $r^{\zeta'}$, as input for subsequent applications of the same proposition. Theorem 7.4 will be obtained by applying Proposition 9.1 a finite number of times, with the output bound (with an increased exponent) of one application being the input for the next, till the exponent is raised from the initial value of $\beta = \alpha$ to a value greater than one for θ in the appropriate range of the tail. Then the same proposition will be applied one final time with $\beta = 1$; at this value of β,

$$\theta^{3\beta/(3-\beta)} (\log \theta)^{-\beta/(3-\beta)} = \theta^{3/2} (\log \theta)^{-1/2},$$

which will yield Theorem 7.4. The quantity

$$\zeta' = \min\left(\frac{\alpha\zeta}{1+\alpha\zeta} \cdot \frac{3-\beta}{3\beta}, \ \frac{2\alpha}{9+16\alpha} \right)$$

measures how far into the tail each improved exponent holds via our arguments. The above explicit expression we obtain is perhaps hard to parse and is not of great importance for our conclusions. Nonetheless, we point out two basic properties of ζ': (i) it is smaller than ζ, as may be seen by algebraic manipulations of the first of the two expressions being minimized in its definition (along with $\beta \ge \alpha$); and (ii) it decays to zero as $\alpha \to 0$ linearly.

We next prove Theorem 7.4 given Proposition 9.1, before turning to the proof of Proposition 9.1.

Proof of Theorem 7.4. First, if $\alpha \geq 1$, we apply Proposition 9.1 with $\alpha = \beta = 1$, $\zeta = \infty$, $j = 1$, and the hypothesis (9.1) provided by Assumption 3a. This yields Theorem 7.4 by taking $z = 0$.

If $\alpha \in (0, 1)$, we will apply Proposition 9.1 iteratively finitely many times. Let α_j, β_j, and ζ_j be values which we will specify shortly. We will select these values such that the hypothesis (9.1) of Proposition 9.1 holds with parameters $(\alpha_1, \beta_1, \zeta_1, 1)$ for all $|z| \leq r^{5/6}$, and, knowing that (9.1) holds with parameters $(\alpha_j, \beta_j, \zeta_j, j)$ for all $|z| \leq r^{5/6}$ and applying Proposition 9.1 will imply that (9.1) holds with parameters $(\alpha_{j+1}, \beta_{j+1}, \zeta_{j+1}, j + 1)$ for all $|z| \leq r^{5/6}$.

We set $\alpha_j = \alpha$ for all j, and adopt the initial settings $\beta_1 = \alpha$ and $\zeta_1 = \infty$; so again (9.1) is provided by Assumption 3a when $j = 1$. The subsequent values are read off of Proposition 9.1 as follows for $j \geq 2$:

$$\beta_j = \min\left(\frac{3 - \frac{1}{2}\beta_{j-1}}{3 - \beta_{j-1}} \cdot \beta_{j-1},\ 1\right) \quad \text{and} \quad \zeta_j = \min\left(\frac{\alpha\zeta_{j-1}}{1 + \alpha\zeta_{j-1}} \cdot \frac{3 - \beta_{j-1}}{3\beta_{j-1}},\ \frac{2\alpha}{9 + 16\alpha}\right), \tag{9.2}$$

where $\alpha\zeta_{j-1}/(1 + \alpha\zeta_{j-1})$ in the definition of ζ_j is interpreted as 1 when $\zeta_{j-1} = \infty$. We adopt the previous expression for β_j instead of the one given by Proposition 9.1 in order to absorb the log factor in the denominator of the exponent furnished by that proposition. Observe that $\beta_j > \beta_{j-1}$ whenever $\beta_{j-1} < 1$.

We define $n \in \mathbb{N}$ by

$$n := \min\left\{j : \beta_j = 1\right\}; \tag{9.3}$$

it can be checked that n is finite since, if $\beta_j < 1$,

$$\frac{\beta_j}{\beta_{j-1}} = \frac{3 - \frac{1}{2}\beta_{j-1}}{3 - \beta_{j-1}} = 1 + \frac{\beta_{j-1}}{2(3 - \beta_{j-1})} \geq 1 + \frac{\alpha}{2(3 - \alpha)},$$

as $\beta_{j-1} > \beta_{j-2} > \ldots > \beta_1 = \alpha$.

By the previous discussion, we know that (9.1) holds with parameters $(\alpha_n, \beta_n = 1, \zeta_n, n)$. Applying Proposition 9.1 with these parameters and taking $z = 0$ gives the statement of Theorem 7.4 with $\zeta = \zeta_{n+1} = \min(\frac{2}{3} \cdot \frac{\alpha\zeta_n}{1+\alpha\zeta_n}, \frac{2\alpha}{9+16\alpha})$. It is clear from this expression that $\zeta \to 0$ as $\alpha \to 0$, and, since $2\alpha/(9 + 16\alpha)$ achieves a maximum value of $2/25$ for all $\alpha \in (0, 1]$, that $\zeta \in (0, 2/25]$. □

Remark 9.2. We can now specify more precisely the regimes of θ provided by the proof of Theorem 7.4 where the tail exponent transitions from $3/2$ to α, as mentioned in Remark 7.11. That is, for $j = [\![1, n]\!]$ with n as in (9.3) and β_j and ζ_j as in (9.2), it holds for $\theta \in [r^{\zeta_{j+1}}, \zeta_j]$ that

$$\mathbb{P}\left(X_r - \mathbb{E}[X_r] \geq \theta r^{1/3}\right) \leq \exp\left(-c\theta^{\beta_j}\right) \quad \text{for} \quad \theta \in [r^{\zeta_{j+1}}, r^{\zeta_j}).$$

It remains to prove Proposition 9.1. A roadmap for the proof is as follows.

1. As indicated immediately before the statement of the proposition, to achieve a stochastic domination of the geodesic weight by a sum, we specify a grid-based discretization of the geodesic, and Lemma 9.3 bounds the cardinality of the number of possible discretizations.

2. Lemma 9.4 provides an improved tail bound (compared to the hypothesis (9.1)) for the weight of a given discretization, using the bootstrapping idea of looking at smaller scales. This makes use of Lemma 9.5, which takes the point-to-point tail available from (9.1) and gives an interval-to-interval bound with the same tail.
3. When Lemmas 9.3 and 9.4 are in hand, the proof of Proposition 9.1 will be completed by taking a union bound.

We address each of the above three steps in turn in the next three subsections.

Step 1: The discretization scheme

We will define a grid $\mathbb{G}^z$ of intervals through which any geodesic from $(1,1)$ to $(r-z, r+z)$, on the event that it is typical, must necessarily pass through; see Figure 9.1.

We recall from Section 7.3 that "width" refers to measurement along the anti-diagonal and "height" to measurement along the diagonal. For $k \in \mathbb{N}$ to be set, the width of a cell in the grid will be $(r/k)^{2/3}$, and the height r/k. The number of cells in a column of the grid is k, and the number of cells in a row is $M = 2\theta^{3/4\alpha}k^{2/3}$ as we want the width of $\mathbb{G}^z$ to be $2\theta^{3/4\alpha}r^{2/3}$. The width of $\mathbb{G}^z$ is set to this value because, by Proposition 7.15 on the probability of any geodesic having large transversal fluctuations, $\mathbb{P}(\mathrm{TF}(\Gamma_r^z) > \theta^{3/4\alpha}r^{2/3}) \le \exp(-c\theta^{3/2})$; note that this is smaller than the bound we are aiming to prove in Proposition 9.1 and so we may essentially ignore the event that any geodesic exits the grid.

We now move to the formal definition. We assume k is small enough that $(r/k)^{2/3} \ge 1$, i.e., $k \le r$ (as the minimum separation of points in $\mathbb{Z}^2$ is 1). The grid $\mathbb{G}^z$ consists of intervals $\mathcal{L}_{ij}^z$ as follows:

$$\mathbb{G}^z = \left\{\mathcal{L}_{ij}^z : i \in [\![0,k]\!], j \in [\![0,M]\!]\right\},$$

where M is defined as

$$M = 2 \cdot \lceil \theta^{\frac{3}{4\alpha}} k^{2/3} \rceil. \tag{9.4}$$

Let $v_i = \lfloor ir/k \rfloor$ and $h_{i,j}^z = \lfloor iz/k + (\theta^{\frac{3}{4\alpha}} - jk^{-2/3})r^{2/3} \rfloor$. For $i \in [\![0,k]\!]$ and $j \in [\![0,M]\!]$, the line segment $\mathcal{L}_{ij}^z$ will connect the points

$$\left(v_i - h_{i,j}^z, v_i + h_{i,j}^z\right) \quad \text{and} \quad \left(v_i - h_{i,j+1}^z, v_i + h_{i,j+1}^z\right).$$

In words, the grid $\mathbb{G}^z$ is contained in the rectangle $\{|y - \frac{r+z}{r-z} \cdot x| \le \theta^{\frac{3}{4\alpha}} r^{2/3}, 0 \le x+y \le 2r\}$. Grid lines along the anti-diagonal will be called $\mathbb{G}_i^z$, i.e., for $i \in [\![0,k]\!]$,

$$\mathbb{G}_i^z = \left\{\mathcal{L}_{ij}^z : j \in [\![0,M]\!]\right\}.$$

We call $\mathcal{L}^z = (L_0, \dots, L_k)$ a discretization, where $L_i \in \mathbb{G}_i^z$ is an interval on the i^{th} grid line. We impose that L_0 and L_k are the intervals whose midpoints are $(1,1)$ and $(r-z, r+z)$ respectively.

Lemma 9.3. *The set of discretizations has size at most* $\exp\left\{k\left(\log k + \frac{3}{4\alpha}\log\theta + \log 2\right)\right\}$.

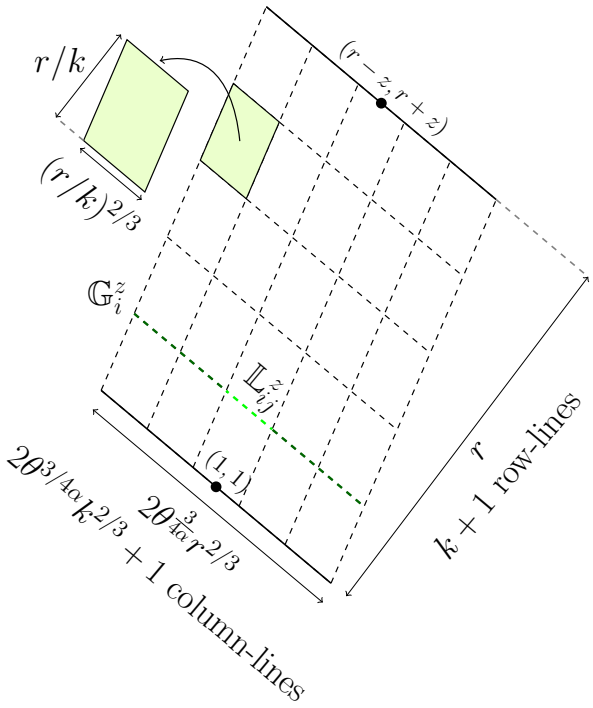


Figure 9.1: The grid utilized for the discretization in Step 1 of the proof of Proposition 9.1. Note that measurements are made along the antidiagonal and diagonal only, with the diagonal chosen over the line with the slope of the left or right boundary of the grid. The lower boundary of the grid $\mathbb{G}^z$ is centered at $(1,1)$ and the upper boundary at $(r-z, r+z)$. From each grid line $\mathbb{G}_i^z$, one interval L_i is picked to form a discretization $\mathcal{L}^z = (L_0, \dots, L_k)$ with the constraint that L_0 is fixed to be the interval on $\mathbb{G}_0^z$ whose midpoint is $(1,1)$ and L_k to be the interval on $\mathbb{G}_k^z$ whose midpoint is $(r-z, r+z)$. On the high probability event that all geodesics passes through the grid, its weight is upper bounded by the maximum, over all discretizations $\mathcal{L}^z$, of the sum of interval-to-interval weights of the intervals in $\mathcal{L}^z$. These weights are independent and have fluctuations of scale $(r/k)^{1/3}$, which allows us to use the idea of bootstrapping.

Proof. This follows from the observation that there are $M = 2\theta^{\frac{3}{4\alpha}}k^{2/3} \le 2\theta^{\frac{3}{4\alpha}}k$ intervals on each grid line $\mathbb{G}_i^z$, and there are $k-1$ grid lines in total where there is a choice of interval (as the intervals from $\mathbb{G}_0^z$ and $\mathbb{G}_k^z$ are fixed), giving $(2\theta^{\frac{3}{4\alpha}}k^{2/3})^{k-1}$ discretizations. □

For a given discretization $\mathcal{L}^z = (L_0, \dots, L_k)$, let $X_{\mathcal{L}^z}$ be the maximum weight of all paths which pass through all intervals of $\mathcal{L}^z$. The discretization described above implies that, on the event that $\mathrm{TF}(\Gamma_r^z) \le \theta^{\frac{3}{4\alpha}}r^{2/3}$,

$$X_r^z \le \max_{\mathcal{L}^z} X_{\mathcal{L}^z},$$

where the maximization is over all discretizations $\mathcal{L}^z$. So to prove Proposition 9.1, we need a tail bound on $X_{\mathcal{L}^z}$ for a fixed discretization $\mathcal{L}^z$; this is Step 2 and is done in the next subsection, where the hypothesis (9.1) and bootstrapping are used to provide an improved tail bound on $X_{\mathcal{L}^z}$.

Step 2: An improved tail bound on $X_{\mathcal{L}^z}$

Because θ is a global parameter which affects the set of discretizations, we will use the symbol t as in (9.5) ahead to denote the scaled deviation when considering the weight associated to a fixed discretization, though we will eventually set $t = \theta$. The following lemma uses the idea of moving to lower scales to obtain an improved tail bound for $X_{\mathcal{L}^z}$ for a fixed discretization $\mathcal{L}^z$.

Lemma 9.4. *Under the hypotheses of Proposition 9.1 there exist $c' = c'(c, \alpha, \beta, j) > 0$, $\delta = \delta(c, \beta, j, \theta_0) > 0$, and $t_0 = t_0(c, \beta, j)$ such that the following holds. Let $t > t_0$, $r > r_0$, $2^6 \le k \le \min(\delta t^{3/2}, r_0^{-1} r)$, $\theta \ge \theta_0$, and $z \in [-r, r]$ be such that $|z| \le r^{5/6}$ and $(r/k)^{5/6} > 4\theta^{3/4\alpha} r^{2/3}$. Let $\mathcal{L}^z = (L_0, \dots, L_k)$ be a fixed discretization. Then*

$$P\left(X_{\mathcal{L}^z} > \mu r - \lambda_{j+1}\frac{Gz^2}{r} + tr^{1/3}\right) \le \exp\left(-c' t^\beta k^{\beta/3}\right) + k \cdot \exp\left(-c'(r/k)^{\alpha\zeta}\right), \tag{9.5}$$

with the second term interpreted as zero if $\zeta = \infty$.

The basic tool in the proof of Lemma 9.4 is to bound $X_{\mathcal{L}^z}$ by the sum of the interval-to-interval weights defined by the intervals in $\mathcal{L}^z$. So given a point-to-point upper tail bound, as in the hypothesis of Proposition 9.1, we will first need to obtain an upper tail bound for interval-to-interval weights.

We define the relevant intervals to state the interval-to-interval bound next. For r fixed, and $|w| \le r^{5/6}$, let $\mathbb{L}_{\text{low}}$ be the line segment joining $(-r^{2/3}, r^{2/3})$ and $(r^{2/3}, -r^{2/3})$ and let $\mathbb{L}_{\text{up}}$ be the line segment joining $(r - w - r^{2/3}, r + w + r^{2/3})$ and $(r - w + r^{2/3}, r + w - r^{2/3})$. Thus w is the midpoint displacement of the intervals, and note that their height difference is r. Define Z by

$$Z = X_{\mathbb{L}_{\text{low}}, \mathbb{L}_{\text{up}}}.$$

The content of the next lemma is a tail bound on Z.

Lemma 9.5. *Suppose (9.1) holds as in Proposition 9.1. Then there exist $\tilde{c} = \tilde{c}(c, j)$, $\tilde{t}_0 = \tilde{t}_0(\theta_0, j)$, and $\tilde{r}_0 = \tilde{r}_0(r_0, j)$ such that, for $r > \tilde{r}_0$, $|w| \le r^{5/6}$, and $t > \tilde{t}_0$*

$$\mathbb{P}\left(Z > \mu r - \lambda_{j+1}\frac{Gw^2}{r} + tr^{1/3}\right) \le \begin{cases} \exp\left(-\tilde{c} t^\beta\right) & \tilde{t}_0 < t < r^\zeta \\ \exp\left(-\tilde{c} t^\alpha\right) & t \ge r^\zeta. \end{cases} \tag{9.6}$$

We note that the hypothesis (9.1) of Proposition 9.1 is a point-to-point tail bound from $\mu r - \lambda_j G z^2 / r$, whereas the conclusion of Lemma 9.5 has the weaker λ_{j+1} in place of λ_j (recall $\lambda_j = 1/2 + 2^{-j}$). This reduction in the coefficient of the parabolic term is the previously mentioned relaxation which allows the bootstrap to proceed to the next iteration.

The proof of Lemma 9.5 relies on the geometric idea of stepping back from the two intervals and considering a proxy point-to-point weight. Similar arguments have appeared in the literature previously (see e.g., [BSS14]), but for completeness we give a self-contained proof of Lemma 9.5 in Appendix A. However, we highlight the main idea in Figure 9.2, where we also say a few words on why it is difficult to avoid the relaxation in the parabolic loss.

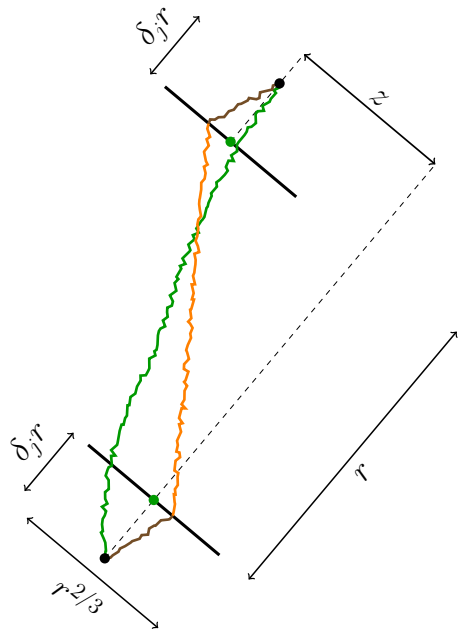


Figure 9.2: The argument for Lemma 9.5. The two black intervals have midpoint separation of z in the antidiagonal direction. The orange path is the heaviest path between the two intervals (so has weight Z), and the brown paths are geodesics connecting the black points to the endpoints of the red path. The green path is a geodesic between the two black points. With positive probability the two brown paths each have weight greater than $\mu\delta_j r - \frac{1}{3}\theta r^{1/3}$, and so, on the intersection of those events with $\{Z > \mu r - \lambda_{j+1} G z^2/r + \theta r^{1/3}\}$, it holds that the green path has weight at least $\mu(1+2\delta_j)r - \lambda_{j+1} G z^2/r + \frac{1}{3}\theta r^{1/3}$. We choose δ_j such that the parabolic term in this expression is $\lambda_j G z^2/(1+2\delta_j)r$ and apply the point-to-point bound we have. It is because the antidiagonal separation between each pair of black and green points is zero that we have a decrease in the parabolic term. If we make this separation proportional to z, then there is no decrease in the parabolic term, but for large z the gradient of the limit shape from Assumption 2 causes issues. This can be more carefully handled if we instead consider the supremum of *fluctuations* of point-to-point weights from their expectation, and we will have need to do this on one occasion in the appendix.

Proof of Lemma 9.4. Observe the following stochastic domination

$$X_{\mathcal{L}^z} \preceq \sum_{i=1}^{k} Z_i,$$

where Z_i are independent random variables distributed as the weight of the best path from L_{i-1} to L_i. Apart from possible rounding, because Z_i and Z_{i-1} are independent versions of weights which overlap on the interval L_{i-1}, it is possible that the linear term in Z_i is $\mu r/k + O(1)$ rather than $\mu r/k$. We handle this discrepancy by absorbing it into the term $tr^{1/3}$ of Lemma 9.5, which is the only situation where it arises, without further comment.

We note that the diagonal separation between the sides of Z_i is r/k, instead of r as in the definition of Z. We denote the anti-diagonal displacement of the midpoints of the corresponding intervals of Z_i by z_i. We want to eventually apply Lemma 8.2 to $\sum Z_i$, appropriately centred, with its input tail bound (8.3) provided by Lemma 9.5. To reach a form of the probability where Lemma 8.2 is applicable, we observe that

$$\begin{aligned}
\mathbb{P}\left(X_{\mathcal{L}^z} > \mu r - \lambda_{j+1}\frac{Gz^2}{r} + tr^{1/3}\right) &\le \mathbb{P}\left(\sum_{i=1}^{k} Z_i \ge \mu r - \lambda_{j+1}\frac{Gz^2}{r} + tr^{1/3}\right) \\
&= \mathbb{P}\left(\sum_{i=1}^{k}(Z_i - \nu_i) \ge \mu r - \lambda_{j+1}\frac{Gz^2}{r} - \sum_{i=1}^{k}\nu_i + tr^{1/3}\right) \\
&\le \mathbb{P}\left(\sum_{i=1}^{k}(Z_i - \nu_i) \ge tr^{1/3}\right), \qquad (9.7)
\end{aligned}$$

where $\nu_i = \mu r/k - \lambda_{j+1} G z_i^2 k/r$. The choice of ν_i is dictated by the desire to apply (9.6) with r replaced by r/k. All the steps before the last inequality are straightforward consequences of definitions. To see the last inequality, note that since $\sum z_i = z$, the Cauchy-Schwarz inequality implies that $\sum \nu_i$ is smaller than $\mu r - \lambda_{j+1} G z^2/r$.

We will soon apply Lemma 9.5, which will yield a tail bound for $Z_i - \nu_i$; the tail bound is (9.6) with r replaced by r/k. However, this tail bound has two regimes with different exponents, α and β, while the basic concentration result we seek to apply, i.e., Theorem 8.2, assumes the same tail exponent throughout.

Thus to have variables that have the larger exponent β in the entire tail, we will apply a simple truncation on Z_i: define

$$\overline{Z}_i = \begin{cases} Z_i & \text{if } Z_i - \nu_i \le \left(\frac{r}{k}\right)^{\zeta+1/3} \\ \nu_i & \text{if } Z_i - \nu_i > \left(\frac{r}{k}\right)^{\zeta+1/3}. \end{cases}$$

Now following (9.7), we get

$$\mathbb{P}\left(X_{\mathcal{L}^z} > \mu r - \lambda_{j+1}\frac{Gz^2}{r} + tr^{1/3}\right) \le \mathbb{P}\left(\sum_{i=1}^{k}(\overline{Z}_i - \nu_i) \ge tr^{1/3}\right) + \mathbb{P}\left(\bigcup_{i=1}^{k}\left\{Z_i - \nu_i > (r/k)^{\zeta+1/3}\right\}\right). \tag{9.8}$$

We will apply the concentration bound Lemma 8.2 to bound the first term. We first want to apply Lemma 9.5, with r/k in place of r, in order to get the tail bound on each individual $Z_i - \nu_i$, which will act as input for Lemma 8.2. Two hypotheses of Lemma 9.5, namely (9.1) and that $r/k > r_0$, are available here by the hypotheses of Lemma 9.4.

But Lemma 9.5 has the additional hypothesis that the anti-diagonal displacement $|w|$ is at most $(r/k)^{5/6}$, which must also be checked. The verification of this follows from the hypothesis in Lemma 9.4 that $(r/k)^{5/6} > 4\theta^{3/4\alpha}r^{2/3}$, as the maximum anti-diagonal displacement possible in a single row of the grid is at most $2\theta^{3/4\alpha}r^{2/3} + |z|/k$, where the first term is the grid width $2\theta^{3/4\alpha}r^{2/3}$, and the second term is the shift caused by the overall slope of the grid. Now since $|z| \le r^{5/6}$ and $k \ge 2^6$, we see that $|z|/k$ is at most $\frac{1}{2}(r/k)^{5/6}$, and some simple algebra completes the verification.

Thus, applying Lemma 9.5 with r/k in place of r, we use the first case of (9.6) (since Z_i has been appropriately truncated to give $\overline{Z}_i$) as the input tail bound with exponent β on $\overline{Z}_i - \nu_i$ required for Lemma 8.2. Finally, with $c_1 = c_1(c, \beta, \theta_0, j)$ as in the statement of the latter, let $\delta = \min(1, (2c_1)^{-3/2})$ where recall we have the hypothesis that $k \le \min(\delta t^{3/2}, r_0^{-1}r)$; δ depends on c, β, j and θ_0. With this preparation, we see

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{k}(\overline{Z}_i - \nu_i) \ge tr^{1/3}\right) &= \mathbb{P}\left(\sum_{i=1}^{k}(\overline{Z}_i - \nu_i)(r/k)^{-1/3} \ge (tk^{1/3} - kc_1) + kc_1\right) \\
&\le \mathbb{P}\left(\sum_{i=1}^{k}(\overline{Z}_i - \nu_i)(r/k)^{-1/3} \ge \frac{1}{2}tk^{1/3} + kc_1\right) \\
&\qquad \text{[since by hypothesis } k \le (2c_1)^{-3/2}t^{3/2}\text{]} \\
&\le \begin{cases} 2\exp\left(-\tilde{c}t^2k^{-1/3}\right) & 0 \le tk^{1/3} < k^{1/(2-\beta)} \\ 2\exp\left(-\tilde{c}t^\beta k^{\beta/3}\right) & tk^{1/3} \ge k^{1/(2-\beta)}; \end{cases} \qquad \text{[by Lemma 8.2]}
\end{aligned}$$

we have applied Lemma 8.2 with $tk^{1/3}$ in place of t and $\alpha = \beta$. Here $\tilde{c}$ is a function of c (as given in the hypothesis (9.1)), β, and j. We now claim that the second case of the last display dictates the fluctuation behavior under our hypotheses. To see this, note that since $\beta \le 1$, $k^{1/(2-\beta)} \le k$. Thus the first case in the last display holds only if $k > t^{3/2}$ while by hypothesis $k \le t^{3/2}$ since $\delta \le 1$. Further, since $k \ge 1$, we may set the lower bound t_0 on t high enough that $\exp(-\frac{1}{2}\tilde{c}t^\beta) \le 1/2$ so as to absorb the pre-factor of 2 in the last display; t_0 depends on $\tilde{c}$ and β. We have hence bounded the first term of (9.8).

To bound the second term when $\zeta < \infty$, we take a union bound and apply Lemma 9.5, where the latter's hypotheses are satisfied by the same reasoning as used above in the application for the first term. This yields that the second term of (9.8) is bounded by $k \cdot \exp\left(-\tilde{c}(r/k)^{\alpha\zeta}\right)$, using the second case of (9.6) with r/k in place of r. Here $\tilde{c}$ is a function of c, α, and j. When $\zeta = \infty$, the second term of (9.8) is clearly zero.

Returning to (9.8) with these two bounds completes the proof of Lemma 9.4, taking $c' = \tilde{c}$. □

Step 3: Handling all the discretizations

With the improved tail bound for a fixed discretization provided by Lemma 9.4, we can implement Step 3 and complete the proof of Proposition 9.1, essentially via a union bound.

Proof of Proposition 9.1. Recall that θ_0' is the lower bound on θ under which the conclusions of Proposition 9.1 must be shown to hold, and that we have the freedom to set it. We will increase its value as needed as the proof proceeds. We will be explicit about the dependencies θ_0' takes on at each such time. We start with $\theta_0' = e$ so that $\log\theta \geq 1$. Also, in this proof, c is reserved for the constant in the point-to-point tail hypothesis (9.1).

Lemma 9.3 says that the entropy from the union bound we will soon perform will be $\exp\{\Theta(k\log k + k\log\theta)\}$, which needs to be counteracted by the bound from Lemma 9.4. Anticipating this we take, in Lemma 9.4,

$$t = \theta \qquad \text{and} \qquad k = \varepsilon \cdot \theta^{\frac{3\beta}{3-\beta}} (\log\theta)^{-\frac{3}{3-\beta}}, \tag{9.9}$$

for $\varepsilon = \varepsilon(c, \alpha, \beta, j) \in (0, 1)$ a sufficiently small constant, to be set shortly. At this point we will ensure that the hypotheses of Lemma 9.4 hold. We set θ_0' larger if needed so that it is at least t_0 as in Lemma 9.4, so that the value of t above satisfies $t > t_0$. Additionally we have to verify that, with δ as provided by Lemma 9.4,

- $4\theta^{3/4\alpha} r^{2/3} < (r/k)^{5/6}$.
- $k \in [\![2^6, \min(\delta t^{3/2}, r_0^{-1} r)]\!]$

For the first condition, the fact that $k \leq \theta^{3/2}$ (since $\varepsilon, \beta \leq 1$ and $\log\theta \geq 1$), and some algebraic manipulation, implies that it is sufficient if $\theta \leq \frac{1}{4} r^{(2\alpha)/(9+15\alpha)}$; to avoid carrying forward the factor of 4, we instead reduce the exponent of r to absorb it and impose that

$$\theta \leq r^{\frac{2\alpha}{9+16\alpha}}; \tag{9.10}$$

this implies $\theta \leq \frac{1}{4} r^{(2\alpha)/(9+15\alpha)}$ (and hence the first condition above) when r_0', which is the lower bound on r that we are free to set, is large enough. The value of r_0' depends only on α.

For the second condition, note that $2\alpha/(9+16\alpha) < 2/3$, and that $\beta \leq 1$ implies $3\beta/(3-\beta) \leq 3/2$. Combining this latter inequality with the value (9.9) of k, and that $\theta \leq r^{2/3}$ from (9.10), ensures that $k \in [\![2^6, \min(\delta\theta^{3/2}, r_0^{-1} r)]\!]$ by setting θ_0' large enough, depending on β, δ, and ε; so the second condition holds.

Thus applying Lemma 9.4 with values of t and k as in (9.9) we obtain that, for $\theta_0' < \theta < r^{2\alpha/(9+16\alpha)}$,

$$\mathbb{P}\left(X_{\mathcal{L}^z} > \mu r - \lambda_{j+1}\frac{Gz^2}{r} + \theta r^{1/3}\right) \leq \exp\left(-c' \cdot \varepsilon^{\beta/3}\theta^{\frac{3\beta}{3-\beta}}(\log\theta)^{-\frac{\beta}{3-\beta}}\right) + \theta^{\frac{3\beta}{3-\beta}}\exp\left(-c' r^{\alpha\zeta}\theta^{-\frac{3\alpha\beta\zeta}{3-\beta}}\right), \tag{9.11}$$

with the second term equal to zero if $\zeta = \infty$, and with c' as in Lemma 9.4; thus c' depends on c, α, β, and j. In substituting k in the second term of (9.5) we have used that $k \leq \theta^{3\beta/(3-\beta)}$ since $\varepsilon < 1$ and $\log\theta \geq 1$. When $\zeta < \infty$, we would like the exponential factor of the second term to be smaller than the first term; i.e., it is sufficient if

$$r^{\alpha\zeta}\theta^{-\frac{3\alpha\beta\zeta}{3-\beta}} \geq \theta^{\frac{3\beta}{3-\beta}}.$$

We will soon absorb the polynomial-in-θ factor in the second term by reducing the constant c. Simple algebraic manipulations show that the inequality of the last display is implied by the condition

$$\theta \leq r^{\bar\zeta} \quad \text{with} \quad \bar\zeta = \frac{\alpha\zeta}{1+\alpha\zeta}\cdot\frac{3-\beta}{3\beta}.$$

For simplicity, we impose this last condition on θ even when $\zeta = \infty$, even though in this case the second term of (9.11) is zero (and so smaller than the first term) for all θ up to $r^{2\alpha/(9+16\alpha)}$ (when $\zeta = \infty$, we interpret the first factor of $\bar\zeta$ to be one, i.e., $\bar\zeta = (3-\beta)/3\beta$).

To handle both the condition in the last display and (9.10), we impose $\theta_0' < \theta < r^{\zeta'}$, with

$$\zeta' = \min\left(\bar\zeta, \frac{2\alpha}{9+16\alpha}\right).$$

So far we have shown that, for $r > r_0'$ and $\theta_0' < \theta < r^{\zeta'}$,

$$\mathbb{P}\left(X_{\mathcal{L}^z} > \mu r - \lambda_{j+1}\frac{Gz^2}{r} + \theta r^{1/3}\right) \leq 2\exp\left(-\tfrac{1}{2}c' \cdot \varepsilon^{\beta/3}\theta^{\frac{3\beta}{3-\beta}}(\log\theta)^{-\frac{\beta}{3-\beta}}\right); \tag{9.12}$$

where, for all $\theta > \theta_0'$, the $\theta^{\frac{3\beta}{3-\beta}}$ polynomial factor coming from the second term of (9.11) has been absorbed by the reduction of c' to $c'/2$. To do this we may also need to increase the value of θ_0'; this choice of θ_0' can be made depending only on c' since we only need $\theta^{3\beta/(3-\beta)}\exp(-c'\theta^{3\beta/(3-\beta)}) \leq \exp(-0.5c'\theta^{3\beta/(3-\beta)})$ and the same function of θ is in the exponent and as the polynomial-factor.

Now we observe that on the event that any geodesic stays within the grid $\mathbb{G}^z$, X_r^z is dominated by $\max_{\mathcal{L}^z} X_{\mathcal{L}^z}$. This yields

$$\mathbb{P}\left(X_r^z > \mu r - \lambda_{j+1}\frac{Gz^2}{r} + \theta r^{1/3}\right) \leq \mathbb{P}\left(\max_{\mathcal{L}^z} X_{\mathcal{L}^z} > \mu r - \lambda_{j+1}\frac{Gz^2}{r} + \theta r^{1/3}\right) + \mathbb{P}\left(\mathrm{TF}(\Gamma_r^z) > \theta^{\frac{3}{4\alpha}} r^{2/3}\right). \tag{9.13}$$

The second term is bounded by $\exp(-c'\theta^{3/2})$ by Proposition 7.15 for all θ such that $\theta^{3/4\alpha} > s_0$, with s_0 an absolute constant as given in the statement of the corollary. We increase θ' if needed to meet this condition; this increase can be done in a way that depends only on s_0 as since $\alpha \leq 1$, it is sufficient if $\theta_0 \geq s_0^{4/3}$.

We want to bound the first term of(9.13) by a union bound over all discretizations $\mathcal{L}^z$. First we bound the cardinality of the set of discretizations using Lemma 9.3. Note that the definition of k in (9.9) implies that $\log k \leq \frac{3\beta}{3-\beta} \log\theta$ as $\varepsilon < 1$. Lemma 9.3 asserts that the set of discretizations has cardinality at most $\exp\{k(\log k + \frac{3}{4\alpha}\log\theta + \log 2)\}$. The just mentioned bound on $\log k$ and the value of k from (9.9) shows that this cardinality is at most

$$\exp\left(\tilde{c}\varepsilon\theta^{\frac{3\beta}{3-\beta}}(\log\theta)^{-\frac{3}{3-\beta}+1}\right) = \exp\left(\tilde{c}\varepsilon\theta^{\frac{3\beta}{3-\beta}}(\log\theta)^{-\frac{\beta}{3-\beta}}\right),$$

with $\tilde{c}$ a constant which depends on only α and β. Given this and the bound in (9.12), we apply a union bound. This yields that, for $\theta_0' < \theta < r^{\zeta'}$, the first term of (9.13) is at most

$$2\exp\left(-\tfrac{1}{2}c' \cdot \varepsilon^{\beta/3}\theta^{\frac{3\beta}{3-\beta}}(\log\theta)^{-\frac{\beta}{3-\beta}} + \tilde{c} \cdot \varepsilon\theta^{\frac{3\beta}{3-\beta}}(\log\theta)^{-\frac{\beta}{3-\beta}}\right).$$

Now since $\beta \leq 1$, for sufficiently small ε it holds that $\tilde{c}\varepsilon \leq \frac{1}{4}c' \cdot \varepsilon^{\beta/3}$, and we fix ε to such a value; note that ε does not depend on θ and only on c, α, β, and j. This can be seen since ε depends on $\tilde{c}$ and c', which respectively depend on α and β only, and c, α, β, and j.

For this value of ε and for $\theta_0' < \theta < r^{\zeta'}$, the previous display is bounded above by

$$\exp\left(-\tfrac{1}{4}c' \cdot \varepsilon^{\beta/3}\theta^{\frac{3\beta}{3-\beta}}(\log\theta)^{-\frac{\beta}{3-\beta}}\right).$$

Putting this bound into (9.13) completes the proof of Proposition 9.1 for $\theta_0' < \theta < r^{\zeta'}$ after relabeling c' in its statement by $\frac{1}{4}c' \cdot \varepsilon^{\beta/3}$. For when $\theta > r^{\zeta'}$, the hypothesis (9.1) provides the bound when $r > r_0$, which we ensure by raising r_0' (if necessary) to be at least r_0. This completes the proof of Proposition 9.1 by relabeling c' in its statement to be less than c if needed. □

9.2 Lower bound on upper tail

We prove the lower bound on the upper tail, i.e., Theorem 7.5.

Proof of Theorem 7.5. Assumption 2 implies that $\mathbb{P}(X_r \geq \mu r + \theta r^{1/3}) \leq \mathbb{P}\left(X_r \geq \mathbb{E}[X_r] + \theta r^{1/3}\right)$, and we prove the stronger bound that $\mathbb{P}(X_r \geq \mu r + \theta r^{1/3}) \geq \exp(-c\theta^{3/2})$ for appropriate θ.

Observe that $X_r \geq \sum_{i=0}^{n} X_{r/k}^{(i)}$ where $X_{r/k}^{(i)} = X_{i(r/k,r/k)+(1,0),(i+1)(r/k,r/k)}$. Now by Assumption 4a we have that

$$\mathbb{P}\left(X_{r/k}^{(i)} \geq \mu r/k + C(r/k)^{1/3}\right) \geq \delta$$

for each $i \in [\![0, k-1]\!]$, as long as $r/k > r_0$. Since

$$\left\{ \sum_{i=0}^{k} X_{r/k}^{(i)} \geq \mu r + Ck^{2/3} r^{1/3} \right\} \supseteq \bigcap_{i=0}^{k-1} \left\{ X_{r/k}^{(i)} \geq \mu r/k + C(r/k)^{1/3} \right\},$$

we have

$$\mathbb{P}\left(X_r \geq \mu r + Ck^{2/3} r^{1/3} \right) \geq \delta^k,$$

using the independence of the $X_{r/k}^{(i)}$ across i. Now we set $k = C^{-3/2}\theta^{3/2}$, giving

$$\mathbb{P}\left(X_r \geq \mu r + \theta r^{1/3} \right) \geq \exp(-c\theta^{3/2})$$

for some $c > 0$ and for all θ satisfying $1 \leq C^{-3/2}\theta^{3/2} \leq r/r_0$, which is equivalent to $C \leq \theta \leq Cr_0^{-2/3} \times r^{2/3}$. Thus the proof of Theorem 7.5 is completed by setting $\theta_0 = C$ and $\eta = Cr_0^{-2/3}$. □

9.3 Lower tail and constrained lower tail bounds

In this section we prove Theorems 7.6, 7.7, and 7.13. In fact Theorem 7.13 implies both of the other two, but we prove Theorem 7.6 first separately to aid in exposition.

Upper bound on lower tail

Note that the abstracted bootstrap statement Proposition 8.3 is applicable with $Y_r = -(X_r - \mu r)$ and $Y_{r,i}^{(k)} = -(X_{r/k}^{(i)} - \mu r/k)$, where $X_{r/k}^{(i)}$ is the last passage value from $(i-1)/k \cdot (r, r) + (1, 0)$ to $i/k \cdot (r, r)$ for $i \in [\![1, k]\!]$. Iterating this would yield a lower tail exponent of $3/2$ (a similar argument for the upper tail under *sub*-additivity was outlined in the beginning of Section 7.3) but will not be able to reach the optimal exponent of 3.

Recall from Section 7.3 that our argument relies on a high-probability construction of k disjoint paths with good collective weight, Theorem 7.16. Thus the probability the construction fails is an upper bound on the probability that many disjoint curves have small weight, which in turn bounds the probability that the geodesic has small weight, as we seek.

As outlined before, the construction relies on three inputs: the first is the parabolic curvature on the limit shape, provided by Assumption 2; the second is an exponential upper bound on the lower tail of the maximum weight among all paths constrained to stay within a given parallelogram; and the third is a lower bound on the mean of such weights. Recall that we call such weights "constrained weights". Like the first input, the third input is available to us already, and is the content of (7.8) of Proposition 7.17. So only the second input needs to be attained via bootstrapping.

From here on the argument has two broad steps.

1. Use our assumptions to obtain the exponential bound (in fact, we obtain an exponent of $3/2$) on the constrained weight's lower tail that can be used as an input for the construction of Theorem 7.16. This tail bound is Proposition 9.6. The argument uses bootstrapping as in Proposition 8.3, and applies that proposition iteratively.
2. Relate the lower tail event of X_r to the event of the existence of k disjoint paths constructed in [BGHH20] (Theorem 7.16 here).

We will implement these two steps in turn next, and then, in Chpater 10, we prove Theorem 7.16. We start by specifying some notation for constrained weights.

Recall the notation for parallelograms introduced in Section 7.6, where $U = U_{r,\ell,z}$ is a parallelogram of height r, width $\ell r^{2/3}$, and opposite side midpoints $(1,1)$ and $(r-z, r+z)$. Recall also that X_r^U is the maximum weight over all paths from $(1,1)$ to $(r-z, r+z)$ which are constrained to stay in U.

Proposition 7.17 provides a stretched exponential lower tail for X_r^U from our assumptions. The following upgraded tail obtained via bootstrapping will suffice for our purpose; note that the bound is still not the optimal one stated in Theorem 7.13, which we prove later.

Proposition 9.6. *Let L_1, L_2, and K be such that $L_1 < \ell < L_2$ and $|z| \le Kr^{2/3}$. Under Assumptions 2 and 3, there exist constants r_0, θ_0, and $c > 0$, all depending on only L_1, L_2, K, and α, such that, for $r > r_0$ and $\theta > \theta_0$,*

$$\mathbb{P}\left(X_r^U \le \mu r - \theta r^{1/3}\right) \le \exp(-c\theta^{3/2}).$$

This is the first step outlined above. The proof is similar to that outlined at the beginning of this section for X_r, and involves using bootstrapping for a number of iterations, with the exponent increasing by the end of each iteration to $3/2$ times its value at the start of it. Once the exponent passes 1, a final iteration brings it to $3/2$.

Proof of Proposition 9.6. Consider the k subparallelograms U_i, $i \in [\![1,k]\!]$, where U_i is defined as the parallelogram with height r/k, width $\min(\ell r^{2/3}, (r/k)^{2/3})$, and opposite side midpoints $(r-z, r+z)\cdot(i-1)/k + (1,0)$ and $(r-z, r+z)\cdot i/k$. Let $Y_r = -(X_r^U - \mu r)$ and $Y_{r,i}^{(k)} = -(X_{r/k}^{U_i} - \mu r/k)$.

We want to apply Proposition 8.3 to these variables. By the definition of X_r^U and $X_{r/k}^{U_i}$, we have that $Y_r \le \sum_{i=1}^k Y_{r,i}^{(k)}$ for all $k \le r$. The variables $\{Y_{r,i}^{(k)} : i \in [\![1,k]\!]\}$ are independent for each k as they are defined by the randomness in disjoint parts of the environment, and (7.7) of Proposition 7.17 provides a stretched exponential tail (of exponent $\alpha' = 2\alpha/3$) for each $Y_{r,i}^{(k)}$. Since the constants c, θ_0, and r_0 of Proposition 7.17 depend on only K, L_1, L_2, we obtain from Proposition 8.3 that there exist $\tilde{c}$, $\tilde{r}_0$, and $\tilde{\theta}_0$ (all depending on K, L_1, L_2, and α) such that, for $r > \tilde{r}_0$ and $\tilde{\theta}_0 < \theta < r^{2/3}$,

$$\mathbb{P}\left(X_r^U \le \mu r - \theta r^{1/3}\right) \le \exp(-\tilde{c}\theta^{3\alpha'/2}).$$

Note that if $\mu > 1$, the constraint $\theta < r^{2/3}$ can be extended to $\theta < \mu r^{2/3}$ by reducing the constant c if needed, in a way that depends on only α and μ. Beyond $\mu r^{2/3}$, the probability on the left side of the last display is zero since the vertex weights are non-negative, and so the last displayed inequality actually holds for all $\theta > \tilde{\theta}_0$.

We may iterate the above argument, such that at the end of each iteration the tail exponent is $3/2$ times its value at the beginning, till the tail exponent exceeds 1. Then we may apply the above argument one last time with $\alpha' = 1$, and this completes the proof. Since the finite number of iterations is only a function of α, the proposition follows. □

As mentioned, we will use this proposition to prove Theorem 7.16 in Chpater 10. Here, we next prove Theorem 7.6 using Theorem 7.16.

Proof of Theorem 7.6. Since $\mathbb{E}[X_r] \leq \mu r$ by Assumption 2 (this is also implied directly by the super-additivity of $\{X_r\}_{r\in\mathbb{N}}$), it is sufficient to upper bound the probability $\mathbb{P}\left(X_r \leq \mu r - \theta r^{1/3}\right)$. Let η be as given in Theorem 7.16, and denote the event whose probability is lower bounded there by $E_{m,k,r}$, i.e., $E_{m,k,r}$ is the event that there exist m disjoint paths $\gamma_1, \ldots, \gamma_m$ with prescribed endpoints, $\max_i \mathrm{TF}(\gamma_i) \leq 2mk^{-2/3}r^{1/3}$, and $\sum_{i=1}^m \ell(\gamma_i) \geq \mu rm - C_1 mk^{2/3}r^{1/3}$. Observe that any of these paths γ_i can be extended to a path from $(1,1)$ to (r,r) without decreasing its weight. Now for $\theta \leq C_1\eta^{2/3}r^{2/3}$, set $m = k = C_1^{-3/2}\theta^{3/2}$, and observe that $C_1 mk^{2/3} = m\theta$. Thus,

$$\mathbb{P}\left(X_r \leq \mu r - \theta r^{1/3}\right) \leq \mathbb{P}\left(E^c_{m,k,r}\right) \leq \exp(-cmk) = \exp(-c\theta^3),$$

the second inequality by Theorem 7.16 since the value of k satisfies $k \leq \eta r$; this latter inequality is implied by the condition that $\theta \leq C_1\eta^{2/3}r^{2/3}$. The inequality for $\theta \in [C_1\eta^{2/3}r^{2/3}, \mu r^{2/3}]$ can be handled be reducing the value of c (if $C_1\eta^{2/3} < \mu$), and for $\theta > \mu r^{2/3}$, the probability being bounded is trivially zero. This completes the proof of Theorem 7.6. □

We next present a brief outline of the construction from [BGHH20] before going into the proof of Theorem 7.13, which then also implies Theorem 7.7.

Bounds on constrained lower tail

In this section we will prove Theorem 7.13; this will also imply Theorem 7.7. We start with the short proof of the upper bound of Theorem 7.13, which is a straightforward consequence of Theorem 7.16 and is a refinement of the argument for Theorem 7.6.

Proof of upper bound of Theorem 7.13. We prove a stronger bound with μr in place of $\mathbb{E}[X_r^U]$ (since $\mathbb{E}[X_r^U] \leq \mu r$).

Recall that the width of U is $\ell r^{2/3}$. On the event that $X_r^U \leq \mu r - \theta r^{1/3}$, it follows that any m disjoint paths which lie inside U must have total weight at most $\mu mr - m\theta r^{1/3}$. But for any m which satisfies $2mk^{-2/3}r^{2/3} \leq \ell r^{2/3}$, m disjoint paths which lie inside U with total weight at

least $\mu mr - m\theta r^{1/3}$ are provided by Theorem 7.16 by setting $k = C_1^{-3/2}\theta^{3/2}$, with probability at least $1 - \exp(-cmk) = 1 - \exp(-cm\theta^{3/2})$.

Thus for any m and ℓ satisfying $m \leq k$ and $m \leq \frac{1}{2}\ell k^{2/3}$ we have

$$\mathbb{P}\left(X_r^U \leq \mu r - \theta r^{1/3}\right) \leq \exp(-cm\theta^{3/2}).$$

Taking $m = \min\left(\frac{1}{2}\ell k^{2/3}, k\right)$ with k as above completes the proof if we verify that $m \geq 1$. This follows by setting $C = 2C_1$ in the assumed lower bound $\ell \geq C\theta^{-1}$ and setting $\theta_0 > C_1$: the bound on ℓ and the value of k implies that $\frac{1}{2}\ell k^{2/3} \geq 1$, and the bound on θ and the value of k implies that $k \geq 1$. □

The rest of this section is devoted to assembling the tools to prove, and then proving, the lower bound on the lower tail of Theorem 7.13. To start with, we need a constant lower bound on the lower tail of the point-to-point weight for a range of directions. This is a straightforward consequence of the assumed mean behavior in Assumption 2 and the lower tail bound in Assumptions 3b, and its proof is deferred to the appendix. In fact, we only need $X_r^z < \mu r - Cr^{1/3}$ with positive probability in our application, but we prove a stronger statement with a parabolic loss.

Lemma 9.7. *Let ρ be as given in Assumption 2. Under Assumptions 2 and 3b, there exist $C > 0$ and $\delta > 0$ such that, for $r > r_0$ and $|z| \leq \rho r$,*

$$\mathbb{P}\left(X_r^z < \mu r - \frac{Gz^2}{r} - Cr^{1/3}\right) \geq \delta.$$

The argument for the lower bound of Theorem 7.13, however, will require moving from the above lower bound on the point-to-point lower tail to a similar lower bound on the interval-to-line lower tail. Note that although we had previously encountered interval-to-interval weights, this is the first time in our arguments that we are seeking to bound interval-to-line weights. This is the content of the next lemma. The proof will entail a few steps which we will describe soon. For the precise statement recall that for two sets of vertices A and B in $\mathbb{Z}^2$, $X_{A,B}$ is the maximum weight of all up-right paths starting in A and ending in B.

Lemma 9.8. *Let $I \subseteq \mathbb{Z}^2$ be the interval of lattice points connecting the coordinates $(-r^{2/3}, r^{2/3})$ and $(r^{2/3}, -r^{2/3})$ on the line $x + y = 0$, and let $\mathcal{L}_r \subseteq \mathbb{Z}^2$ be the lattice points on the line $x + y = 2r$. Under Assumptions 3b and 4b, there exist $C', \delta' > 0$, and r_0' such that, for $r > r_0'$,*

$$\mathbb{P}\left(X_{I,\mathcal{L}_r} \leq \mu r - C'r^{1/3}\right) \geq \delta'.$$

Before turning to the proof of Lemma 9.8, we finish the proof of the lower bound of Theorem 7.13 and hence also the proof of Theorem 7.7.

Proof of Theorem 7.7 and lower bound of Theorem 7.13. Recall the lower bound statement of Theorem 7.13 that, for $\theta_0 \le \theta \le \eta r^{2/3}$,

$$\mathbb{P}\left(X_r^U - \mu r \le -\theta r^{1/3})\right) \ge \exp\left(-c_1 \min(\ell\theta^{5/2}, \theta^3)\right);$$

note that Theorem 7.7 is implied by the case that $\ell = r^{1/3}$, since by choosing $\delta < 1$ in the statement of the latter, we assume $\theta \le r^{2/3}$, and hence $\min(\ell\theta^{5/2}, \theta^3) = \theta^3$. We now proceed to proving the bound for general ℓ.

Let k and m be positive integers whose values will be specified shortly. We will define a grid similar to the one in Section 9.1 that was depicted in Figure 9.1, but of width $mk^{-2/3}r^{2/3}$. For $i \in [\![1, k]\!]$ and $j \in [\![1, m]\!]$, let $v_{i,j}$ be the point

$$\left(i\frac{r}{k} - \frac{m}{2}\left(\frac{r}{k}\right)^{2/3},\ i\frac{r}{k} + \frac{m}{2}\left(\frac{r}{k}\right)^{2/3}\right) + j\left(\left(\frac{r}{k}\right)^{2/3},\ -\left(\frac{r}{k}\right)^{2/3}\right),$$

and let $I_{i,j}$ be the interval with endpoints $v_{i,j}$ and $v_{i,j+1}$ (of width $\left(\frac{r}{k}\right)^{2/3}$). As in Section 9.1, we will collectively refer to these intervals as a grid, and so the rows of the grid are indexed by i and the columns by j. Note that the grid lies inside U if $m \le \ell k^{2/3}$ and covers the breadth of U if $m = \ell k^{2/3}$, since the total breadth of the grid is $mk^{-2/3}r^{2/3}$.

This is what dictates the choice of m, although for technical reasons, we set

$$m = \min(\ell k^{2/3}, k), \tag{9.14}$$

where our choice for k later (of order $\theta^{3/2}$) will ensure that indeed for all interesting values of ℓ (i.e., $\ell = O(\theta^{1/2})$), we would have $m = \ell k^{2/3}$.

The idea now is to construct an event on which $X_r^U \le \mu r - \theta r^{1/3}$. Let $X_{I,\mathcal{L}}^{i,j}$ be the maximum weight among all paths with starting point on $I_{i,j}$ and ending point on the line $x + y = 2(i+1)\frac{r}{k}$. The event will be defined by forcing

1. all the $X_{I,\mathcal{L}}^{i,j}$ to be small, i.e., $X_{I,\mathcal{L}}^{i,j} \le \mu(r/k) - C'(r/k)^{1/3}$ for a constant C'; and
2. any path which has transversal fluctuation greater than $k^{1/3}r^{2/3}$ to suffer a parabolic weight loss of order $k^{2/3}r^{1/3}$.

Before proceeding, we let Y_r^k be the maximum weight among all paths Γ from $(1, 1)$ to (r, r) with transversal fluctuation satisfying $\mathrm{TF}(\Gamma) \ge k^{1/3}r^{2/3}$. Thus the second condition above says Y_r^k falls below μr by at least order $k^{2/3}r^{1/3}$.

We claim that, on the event described, $X_r^U \le \mu r - \Omega(k^{2/3}r^{1/3})$. This is due to the following. First, any path within the grid must pass through one of $\{I_{i,j} : j \in [\![1, m]\!]\}$ for every $i \in [\![1, k]\!]$ and so has weight at most $k \cdot \max_{i,j} X_{I,\mathcal{L}}^{i,j} \le \mu r - C'k^{2/3}r^{1/3}$. Second, any path which exits the grid, by our choice of m, either exits U and may be ignored or has transversal fluctuation greater than $k^{1/3}r^{2/3}$ and so suffers a weight loss of at least order $k^{2/3}r^{1/3}$.

Finally, we will show that this event has probability at least $\exp(-cmk)$ (since there are mk values of (i, j) for which $X_{I,\mathcal{L}}^{i,j}$ is made small) and set k to be a multiple of $\theta^{3/2}$.

A more precise form of the above discussion starts with the following inclusion, where c_2 is as in Theorem 7.14 and C' is as in Lemma 9.8:

$$\begin{aligned}\left\{X_r^U \le \mu r - \min(c_2, C')k^{2/3}r^{1/3}\right\} \\ \supseteq \left\{Y_r^k \le \mu r - c_2 k^{2/3} r^{1/3}\right\} \cap \bigcap_{\substack{1\le i\le k\\ 1\le j\le m}} \left\{X_{I,\mathcal{L}}^{i,j} \le \mu r/k - C'(r/k)^{1/3}\right\}.\end{aligned}$$

Note that all the events on the right hand side are decreasing events. Hence, by the FKG inequality,

$$\begin{aligned}\mathbb{P}\left(X_r^U \le \mu r - \min(c_2, C')k^{2/3}r^{1/3}\right) &\ge \mathbb{P}\left(Y_r^k \le \mu r - c_2 k^{2/3} r^{1/3}\right) \\ &\quad \times \prod_{\substack{1\le i\le k\\ 1\le j\le m}} \mathbb{P}\left(X_{I,\mathcal{L}}^{i,j} \le \mu r/k - C'(r/k)^{1/3}\right) \\ &\ge (1 - \exp(-\tilde{c}k^{2\alpha/3})) \cdot (\delta')^{mk}. \qquad (9.15)\end{aligned}$$

The final inequality was obtained by applying Theorem 7.14 with $s = k^{1/3}$ and $t = 0$ to lower bound the first term and Lemma 9.8 (with r/k in place of r) to lower bound the remaining terms. Theorem 7.14 provides an absolute constant s_0 and its application requires $k^{1/3} > s_0$, a condition that will translate into a lower bound on θ after we set the value of k next; Lemma 9.9 requires that $r/k > r_0$, which will translate to an upper bound on θ.

Take $k = (\min(c_2, C'))^{-3/2}\,\theta^{3/2}$ and recall the value of m from (9.14). Note that the assumed lower bound of $\ell \ge C\theta^{-1}$ ensures that $m \ge 1$; we additionally impose that $k \ge s_0^3$ to meet the requirement of Theorem 7.14 mentioned above. We also assume without loss of generality that $s_0 \ge 1$ to encode that k must be at least 1. Thus we obtain from (9.15), for some constant $c_1 > 0$,

$$\mathbb{P}\left(X_r^U \le \mu r - \theta r^{1/3}\right) \ge \exp\left(-c_1 \min(\ell\theta^{5/2}, \theta^3)\right); \qquad (9.16)$$

this holds for every θ which is consistent with $s_0^3 \le k \le r_0^{-1}r$, the latter inequality to ensure that r/k is at least r_0 as obtained from Lemma 9.8. Recalling the value of k, this condition on θ may be written as

$$s_0^2 \cdot \min(c_2, C') \le \theta \le \min(c_2, C')r_0^{-2/3} \cdot r^{2/3}.$$

Recall that Theorem 7.13 must be proven only for $\theta_0 \le \theta \le \eta r^{2/3}$. Thus we may meet the condition of the last display by modifying θ_0 to be greater than $s_0^2 \min(c_2, C')$ and η to be less than $\min(c_2, C')r_0^{-2/3}$, if required.

Now we turn to the final statement in Theorem 7.13 on replacing μr by $\mathbb{E}[X_r^U]$ in (9.16) when ℓ is bounded below by a constant ε. For this, all we require is that $\mathbb{E}[X_r^U] \ge \mu r - Cr^{1/3}$ for $r > r_0$ for a C and r_0 which may depend on ε. This is because with that bound we may absorb the $Cr^{1/3}$ into $\theta r^{1/3}$ by increasing the constant c_1 (which will then depend on ε). Now the required bound on $\mathbb{E}[X_r^U]$ is provided by (7.8) of Proposition 7.17.

This completes the proof of Theorem 7.7 and the lower bound of Theorem 7.13 □

The remaining task is to prove Lemma 9.8, which lower bounds the lower tail for interval-to-line weights. Recall from Figure 9.2 our strategy of obtaining bounds on interval-to-interval weights from similar bounds on point-to-point weights by backing up. Notice that we are presently seeking to lower bound the probability that interval-to-line weights are low; in other words, expressing the line as a union of disjoint intervals, we want to lower bound the probability of the intersection of the decreasing events that all the corresponding interval-to-interval weights are low. This will involve an application of the FKG inequality and the following two lemmas, which lower bound the lower tail of interval-to-interval weights. The first one (Lemma 9.9) treats the case when the anti-diagonal displacement between the two intervals is small. The second one (Lemma 9.10) handles intervals at greater anti-diagonal separation, exploiting the natural parabolic loss in the mean which makes the weights unlikely to be high in this case.

Lemma 9.9. *Let I be the interval connecting the points $(-r^{2/3}, r^{2/3})$ and $(r^{2/3}, -r^{2/3})$, J be the interval connecting the points $(r - z - r^{2/3}, r + z + r^{2/3})$ and $(r - z + r^{2/3}, r + z - r^{2/3})$, and ρ be as in Assumption 2. Under Assumptions 2, 3b, and 4b, there exist C'', $\delta'' > 0$, and r_0'' such that, for $r > r_0''$ and $|z| \leq \rho r - 2r^{2/3}$ (this quantifies the closeness of the intervals in the anti-diagonal direction),*

$$\mathbb{P}\left(X_{I,J} < \mu r - C'' r^{1/3}\right) \geq \delta''.$$

Proof. We first prove a similar statement for intervals of size $\varepsilon r^{2/3}$ for an $\varepsilon > 0$ to be fixed later. This is a crucial first step as it is difficult to directly control the possible gain in weight afforded by allowing the endpoints to vary over an interval of size $2r^{2/3}$; initially using the leeway of making the interval sufficiently small (but on the scale $r^{2/3}$) makes this control achievable. This will be done using a backing up strategy similar to the one illustrated in Figure 9.2 for the interval-to-interval upper tail bound Lemma 9.5.

Let $|w| \leq \rho r$, and let I_ε be the interval joining the points $(-\varepsilon r^{2/3}, \varepsilon r^{2/3})$ and $(\varepsilon r^{2/3}, -\varepsilon r^{2/3})$ and J_ε be the interval joining the points $(r - w - \varepsilon r^{2/3}, r + w + \varepsilon r^{2/3})$ and $(r - w + \varepsilon r^{2/3}, r + w - \varepsilon r^{2/3})$. We will prove that there exist positive C'', δ, and ε, independent of w, such that

$$\mathbb{P}\left(X_{I_\varepsilon, J_\varepsilon} < \mu r - C'' r^{1/3}\right) \geq \delta/2. \tag{9.17}$$

Let $u^* \in I_\varepsilon$ and $v^* \in J_\varepsilon$ be such that $X_{I_\varepsilon, J_\varepsilon} = X_{u^*, v^*}$. Also let $\phi_1 = (-\varepsilon^{3/2} r, -\varepsilon^{3/2} r)$, $\phi_2 = (r - w + \varepsilon^{3/2} r, r + w + \varepsilon^{3/2} r)$ be the backed up points. Then we have the inequality

$$X_{\phi_1, \phi_2} \geq X_{\phi_1, u^* - (1,0)} + X_{I_\varepsilon, J_\varepsilon} + X_{v^* + (1,0), \phi_2}.$$

For M to be fixed and C as in Lemma 9.7, we will consider the constant probability events

$$\begin{aligned}
E_{p \to p} &:= \left\{ X_{\phi_1, \phi_2} \leq \mu(1 + 2\varepsilon^{3/2}) r - G \frac{w^2}{(1 + 2\varepsilon^{3/2} r)} - C r^{1/3} \right\}, \\
E_{p \to \text{int}} &:= \left\{ X_{\phi_1, u^* - (0,1)} > \mu \varepsilon^{3/2} r - M \varepsilon^{1/2} r^{1/3} \right\}, \qquad \text{and} \\
E_{\text{int} \to p} &:= \left\{ X_{v^* + (1,0), \phi_2} > \mu \varepsilon^{3/2} r - M \varepsilon^{1/2} r^{1/3} \right\}.
\end{aligned}$$

On the intersection $E_{p\to p} \cap E_{p\to \text{int}} \cap E_{\text{int}\to p}$ we have

$$X_{I_\varepsilon, J_\varepsilon} < \mu r - (C - 2M\varepsilon^{1/2})r^{1/3}. \tag{9.18}$$

We must lower bound the probability of this intersection. From Lemma 9.7 we see

$$\mathbb{P}\left(E_{p\to p}\right) \geq \delta,$$

since $C < (1+\varepsilon^{3/2})^{1/3}C$. Next, recall that u^* is a vertex of I_ε, which lies on the line $x+y=0$, which is the starting point of a heaviest path from I_ε to J_ε. Thus u^* is independent of the random field below the line $x+y=0$. Now we see that, for large enough M (depending only on δ),

$$\begin{aligned}\mathbb{P}\left(E^c_{p\to\text{int}}\right) &= \mathbb{P}\left(X_{\phi_1, u^*-(0,1)} \leq \mu\varepsilon^{3/2}r - M\varepsilon^{1/2}r^{1/3}\right)\\ &\leq \sup_{u\in I_\varepsilon} \mathbb{P}\left(X_{\phi_1, u-(0,1)} \leq \mu\varepsilon^{3/2}r - M\varepsilon^{1/2}r^{1/3}\right) \leq \frac{\delta}{4},\end{aligned}$$

with the mentioned independence allowing the uniform bound of the second line. The same bound with the same M holds for $\mathbb{P}(E^c_{\text{int}\to p})$ as well; we fix this M. Now we set $\varepsilon > 0$ such that $C - 2M\varepsilon^{1/2} = C/2$. From (9.18) the above yields (9.17) with $C'' = C/2$.

To move from $I_\varepsilon, J_\varepsilon$ to I, J, we let $I_{\varepsilon,i}$ for $i \in [\![1, \varepsilon^{-1}]\!]$ be the intervals of length ε which make up the length one interval I in the obvious way, and similarly for $J_{\varepsilon,j}$ and J. Next observe that

$$\left\{X_{I,J} < \mu r - \tfrac{1}{2}Cr^{1/3}\right\} \supseteq \bigcap_{i,j} \left\{X_{I_{\varepsilon,i}, J_{\varepsilon,j}} < \mu r - \tfrac{1}{2}Cr^{1/3}\right\}. \tag{9.19}$$

Now, the bound (9.17) holds as long as the intervals are of length $\varepsilon r^{2/3}$ and their antidiagonal displacement is at most ρr. The intervals $I_{\varepsilon,i}$ and $J_{\varepsilon,i}$ have this length, and their antidiagonal displacement is at most $|z| + 2r^{2/3}$, where recall z is the antidiagonal displacement between I and J. This occurs, for example, when $I_{\varepsilon,i}$ is the left most subinterval of I and $J_{\varepsilon,j}$ is the right most sub interval of J. But since we have assumed $|z| + 2r^{2/3} \leq \rho r$, the bound (9.17) applies to $X_{I_{\varepsilon,i}, J_{\varepsilon,j}}$, and so the probability of each event in the intersection of (9.19) is at least $\delta/2$.

The intersection of (9.19) is of decreasing events, and so we may invoke the FKG inequality and the just noted probability lower bound to conclude that the probability of the right hand side of (9.19) is at least $(\delta/2)^{\varepsilon^{-2}}$. This completes the proof of Lemma 9.9 with $C'' = C/2$ and $\delta'' = (\delta/2)^{\varepsilon^{-2}}$. □

While the previous lemma provided control when the destination interval is relatively close to (r, r), i.e., have x- and y-coordinates within ρr of r, the next lemma will be used to treat pairs of intervals which have greater separation. Let $I \subseteq \mathbb{Z}^2$ be the interval connecting $(-r^{2/3}, r^{2/3})$ and $(r^{2/3}, r^{2/3})$, and $J \subseteq \mathbb{Z}^2$ be the interval connecting $(r - w - r^{2/3}, r + w + r^{2/3})$ and $(r - w + r^{2/3}, r + w - r^{2/3})$; thus w represents the intervals' antidiagonal displacement, while z will be used as a variable in the hypothesis.

Lemma 9.10. *Suppose there exist $\alpha \in (0,1]$, $\lambda > 0$ and constants $c > 0$, t_0, and r_0 such that, for $t > t_0$, $r > r_0$, and $|z| \le r/2$*

$$\mathbb{P}\left(X_r^z \ge \mu r - \lambda \frac{Gz^2}{r} + tr^{1/3}\right) \le \exp\left(-ct^\alpha\right).$$

Then there exist $\tilde{c} > 0$, $\tilde{t}_0 = \tilde{t}_0(t_0)$, and $\tilde{r}_0 = \tilde{r}_0(r_0)$ such that, for $r > \tilde{r}_0$, $|w| \le |r|$, and $t > \tilde{t}_0$,

$$\mathbb{P}\left(X_{I,J} > \mu r - \frac{\lambda}{3} \cdot \frac{Gw^2}{r} + tr^{1/3}\right) \le \exp\left(-\tilde{c}t^\alpha\right).$$

We impose the condition that $|z| \le r/2$ because the tail bound hypothesis will be provided by Assumption 3a in our application, and the latter requires $|z|$ to be macroscopically away from 0 and r. But note that we allow $|w|$ to be as large as r, as we do need to allow the destination interval to be placed anywhere between the coordinate axes.

Lemma 9.10 will be proved along with Lemma 9.5, as they are both interval-to-interval upper tails, in the appendix via a backing up argument. With the bounds of the previous two lemmas, we can prove the interval-to-line bound in Lemma 9.8 and thus complete the proof of Theorem 7.13.

As earlier, we will construct an event as an intersection of decreasing events which forces $X_{I,\mathcal{L}_r}$ to be small, and use the FKG inequality to lower bound its probability. So, we need to make any path starting in I and ending on the line $\mathcal{L}_r$ have low weight, which we will do by forcing such paths which end on various intervals on $\mathcal{L}_r$ to separately have low weight. When the destination interval is close to the point (r, r), the probability that all such paths have low weight will be lower bounded by Lemma 9.9. When the destination interval is far from (r, r), Lemma 9.10 says that it is highly likely that the paths will have low weight, and it is a matter of checking that the probabilities approach 1 quickly enough that their product is lower bounded by a positive constant.

Proof of Lemma 9.8. For $j \in [\![-r^{1/3}, r^{1/3}]\!]$, let J_j be the interval connecting the points $(r - jr^{2/3}, r + jr^{2/3})$ and $(r - (j+1)r^{2/3}, r + (j+1)r^{2/3})$. We observe that, with C'' as in Lemma 9.9,

$$\left\{X_{I,\mathcal{L}_r} < \mu r - C''r^{1/3}\right\} \supseteq \bigcap_{|j| \le r^{1/3}} \left\{X_{I,J_j} < \mu r - C''r^{1/3}\right\}. \tag{9.20}$$

Assumption 2 says that, for $|z| \le \rho r$, $\mathbb{E}[X_r^z] \le \mu r - Gz^2/r - g_1 r^{1/3}$. But then observe that the concavity of the limit shape implies the existence of a small constant $\lambda \in (0,1)$ such that

$$\mathbb{E}[X_r^z] \le \mu r - \lambda \frac{Gz^2}{r} \tag{9.21}$$

for $|z| \le r$. For this value of λ, Assumption 3a, with $\varepsilon = 1/2$, implies that there exists $c > 0$ such that, for all $r > r_0$ and $|z| \le r/2$,

$$\mathbb{P}\left(X_r^z > \mu r - \lambda \frac{Gz^2}{r} + tr^{1/3}\right) \le \exp(-c\theta^\alpha).$$

With this value of λ and the above bound as input, we will apply Lemma 9.10 with $t = \lambda G j^2/3 - C$. Lemma 9.10 requires $t > \tilde{t}_0$, which implies that $|j|$ must be larger than some j_0. So for $|j| > j_0$, Lemma 9.10 implies that

$$\mathbb{P}\left(X_{I,J_j} < \mu r - C'' r^{1/3}\right) \geq 1 - \exp(-\tilde{c}|j|^{2\alpha}). \tag{9.22}$$

Applying (9.22) and the bound from Lemma 9.9 to (9.20), along with the FKG inequality, gives

$$\begin{aligned}\mathbb{P}\left(X_{I,\mathcal{L}_r} < \mu r - C'' r^{1/3}\right) &\geq \prod_{|j|\leq j_0} \mathbb{P}\left(X_{I,J_j} < \mu r - C'' r^{1/3}\right) \times \prod_{j_0 < |j| \leq r^{1/3}} \mathbb{P}\left(X_{I,J_j} < \mu r - C'' r^{1/3}\right) \\ &\geq (\delta'')^{2j_0+1} \times \prod_{j_0 \leq |j| \leq r^{1/3}} \left(1 - \exp(-\tilde{c}|j|^{\alpha})\right);\end{aligned}$$

we employed Lemma 9.9 for the factors in the first product and (9.22) for those in the second. The proof of Lemma 9.8 is complete by taking $C' = C''$ and δ' equal to the final line of the last display; note that the second product in the last display is bounded below by a positive constant independent of r since $\exp(-\tilde{c}|j|^{\alpha})$ is summable over j for any $\alpha > 0$, and so $\delta' > 0$. □

Chapter 10

Constructing disjoint high weight paths

In this chapter, we construct collections of disjoint paths in the square $[\![n]\!]^2$ that achieve a certain weight with high probability and so prove Theorem 7.16.

Recall that to prove Theorem 7.16, for some $c > 0$, $c_1 > 0$, $k \leq c_1 n$, and $m \leq k$, we must construct m disjoint paths $\gamma_1, \ldots, \gamma_m$ such that $\sum_{i=1}^{m} \ell(\gamma_i) \geq \mu nm - C_1 m k^{2/3} n^{1/3}$ and $\max_i \mathrm{TF}(\gamma_i) \leq 2mk^{-2/3}n^{1/3}$, with probability at least $1 - \exp(-cmk)$.

We start by giving a high level overview of the high-probability construction.

We have to construct m disjoint paths, each with weight loss at most of order $k^{2/3}r^{1/3}$, in a strip of width $4mk^{-2/3}r^{2/3}$. In the bulk of the environment, this is straightforward: for each curve, we set up order k many parallelograms of width $(r/k)^{2/3}$ and height r/k sequentially and consider the path obtained by concatenating together the heaviest midpoint-to-midpoint path constrained to remain in the corresponding parallelogram; see Figure 10.1b. The weight loss in each parallelogram is on scale, i.e., of order $(r/k)^{1/3}$, and so the total loss across the m curves is $m \cdot k \cdot (r/k)^{1/3} = mk^{2/3}r^{1/3}$. The total transversal fluctuation is of order $m(r/k)^{2/3}$, as required, and it is in this phase of the construction that the transversal fluctuation is maximum.

But the previous description is only possible in the bulk, and if the curves have already been brought to a separation of $(r/k)^{2/3}$. Since the curves start and end at a microscopic separation of 1 at the corners of $[\![1, r]\!]^2$, the difficult part of the construction is there, where the curves must be coaxed apart while not sacrificing too much weight. Here the construction proceeds in a dyadic fashion, doubling the separation between curves as the scale increases, while ensuring that the antidiagonal displacement borne by the curves is not too high, so as to not incur a high weight loss due to parabolic curvature. Again the idea is to construct a sequence of parallelograms for each curve that it is constrained to remain within; see Figure 10.1a. It is to estimate the weight loss in this phase of the construction, where antidiagonal displacement is increasing, that we require a curvature assumption such as Assumption 2, and a calculation shows that the weight loss is again of order $mk^{2/3}r^{1/3}$.

The other two inputs, namely a lower bound on the means of constrained weights and exponential decay of the lower tails of the same, are needed to control the probability that the paths

constructed by concatenating together these constrained paths have the requisite weight; in particular, the mean bound is used to show that the expected weight is correct, while the lower tail bound is used to control the deviation below the mean of the total construction weight after expressing it as a sum of independent subexponential variables and invoking a concentration inequality.

10.1 The construction in outline

The construction leading to Theorem 7.16 may be explained in light of the fact, mentioned in Section 2.6, that the width of the k-geodesic watermelon has order $k^{1/3}n^{2/3}$ (which is proven in [BGHH20]). Indeed, that k watermelon curves coexist in a strip of width $k^{1/3}n^{2/3}$ suggests that, at least around the mid-height $n/2$, adjacent curves will separated on the order of $k^{-2/3}n^{2/3}$. We will demand this separation for the m curves in our construction. The curves will begin near $(1, 1)$ and end near (n, n) at unit-order distance, so we must guide them apart to become separated during their mid-lives.

We will index the life of paths in the square $[\![n]\!]^2$ according to distance along the diagonal interval, indexed so that $[a, b]$ refers to the region between the lines $x + y = 2a$ and $x + y = 2b$. The diagonal interval $[1, n]$ that indexes the whole life of paths in the construction will be divided into five consecutive intervals called *phases* that carry the names take-off, climb, cruise, descent and landing. By the start of the middle, cruise, phase, the sought separation has been obtained, and it will be maintained there. This separation is gained during take-off and climb, and it is undone in a symmetric way during descent and landing.

Take-off is a short but intense phase that takes the curves at unit-order separation on the tarmac to a consecutive separation of order $k^{-1/3}n^{1/3}$ in a duration (or height) of order $k^{2/3}n^{1/3}$. Climb is a longer and gentler phase, of duration roughly $n/3$, in which separation expands geometrically until it reaches the scale $k^{-2/3}n^{2/3}$. Cruise is a stable phase of rough duration $n/3$.

The shortfall in weight of the m paths in Theorem 7.16 relative to the linear term μnm has order $mk^{2/3}n^{1/3}$. The weight shortfall in each phase is the difference in total weight contributed by the curve fragments in the phase and the linear term given by the product of μm and the duration of the phase. The weight shortfall will be shown to have order $mk^{2/3}n^{1/3}$ for each of the five phases.

Take-off is a phase where gaining separation is the only aim. No attempt is made to ensure that the constructed curves have weight, and the trivial lower bound of zero on weight is applied. The weight shortfall is thus at most $\mu m \cdot \Theta(1)k^{2/3}n^{1/3}$.

The climb phase is where a rapid increase in separation is obtained, via a geometrically increasing separation across scales (which depend on k). A calculation is needed to bound above the shortfall by $mk^{2/3}n^{1/3}$ for the climb (and for the descent) phase. The calculation will estimate the loss incurred across the scales using the parabolic loss in weight of Assumption 2, relying on a precise choice of paths which do not incur too large a parabolic loss. The climb and descent phases are at the heart of the construction.

Cruise must maintain consecutive separation of order $k^{-2/3}n^{2/3}$ for a duration of roughly $n/3$. We construct paths that travel through an order of k consecutive boxes of height $k^{-1}n$ and width $k^{-2/3}n^{2/3}$. As the KPZ one-third exponent for energy predicts, each passage of a path across a box incurs weight shortfall of order $k^{-1/3}n^{1/3}$. Cruise weight shortfall is thus of order $m \cdot k \cdot k^{-1/3}n^{1/3} = mk^{2/3}n^{1/3}$.

Now we turn to giving the details of each phase. We adopt a rounding convention for coordinates which are not integers, but, in contrast, we will be ignoring the resulting terms of ± 1. More precisely, all expressions for coordinates of points should be rounded down, but extra terms of ± 1 which thus arise in non-coordinate quantities, such as expected values of weights, will be absorbed into constants without explicit mention.

10.2 The construction in detail

Take-off

In this phase, the m curves will travel from $(1,1), \ldots, (1,m)$ to the line $x+y = 2k^{2/3}n^{1/3}$. Since we will make no non-trivial claim about the weight of take-off curves, we may choose these m curves to be any disjoint upright paths, with the i^{th} starting at $(1,i)$ and ending at

$$\left(k^{2/3}n^{1/3} - \left(\frac{m}{2} - i\right)k^{-1/3}n^{1/3},\ k^{2/3}n^{1/3} + \left(\frac{m}{2} - i\right)k^{-1/3}n^{1/3}\right). \tag{10.1}$$

The next statement suffices to show that these m disjoint paths exist; we omit the straightforward proof.

Lemma 10.1. *Suppose given m starting points $\{(1,i) : i \in [\![m]\!]\}$; and m ending points $\{(x_i, y_i) : i \in [\![m]\!]\}$ on the line $x+y=r$, for some $r \geq m+1$, with $1 \leq x_1 < x_2 < \ldots < x_m \leq m$. Then there exist m disjoint paths such that the i^{th} connects $(1,i)$ to (x_{m-i+1}, y_{m-i+1}).*

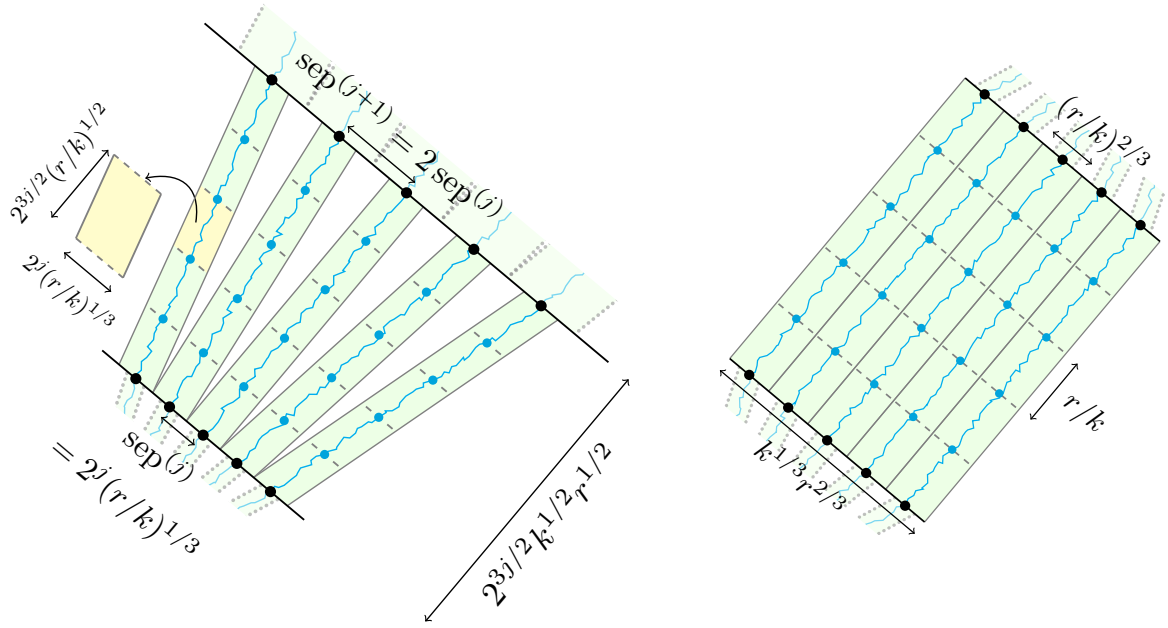
The hypothesis of this lemma that $2k^{2/3}n^{1/3} \geq m+1$ is satisfied when $n \geq 1$ because $m \leq k \leq n$.

Climb

This phase concerns the construction of curves as they pass through diagonal coordinates between $k^{2/3}n^{1/3}$ and a value h that we will specify in (10.6). In Lemma 10.2, we will learn that $cn \leq h \leq n/2$; the climb phase thus has duration $\Theta(n)$, since k is supposed to be at most a small constant multiple of n.

During climb, the order of separation rises from $k^{-1/3}n^{1/3}$ to $k^{-2/3}n^{2/3}$. Separation will increase by a factor $2^{1/k}$ during each of several segments into which the phase will be divided. The number N of segments is chosen to satisfy

$$2^{N/k} \in k^{-1/3}n^{1/3} \cdot [1,2). \tag{10.2}$$



(a) One segment of climb, the second phase

(b) Cruise, the third phase

Figure 10.1: In Panel A is a depiction of climb, between levels j and $j+1$, with the construction between adjoining levels (and thus on smaller and larger scales) in lower opacity. Here $m = k = 5$ and the and each of the green parallelograms is a flight corridor. Observe that any flight corridor between levels j and $j+1$ has width $\mathrm{sep}^{(j)} = 2^{j/k}(n/k)^{1/3}$, which satisfies the relation of width $\approx$ (height)$^{2/3}$ since the height is $3^{-1}2^{3j/2k}(n/k)^{1/2}$. In Panel B is depicted the simpler cruise phase, and how the second phase connects to it on either side in lower opacity. Here h is the distance along the diagonal occupied by the first and second phases on the lower side.

A depiction of one of the segments is shown in Figure 10.1a.

In order that the constructed curves remain disjoint and incur a modest weight shortfall, we will insist that they pass through a system of disjoint parallelograms whose geometry respects KPZ scaling: the width of each parallelogram, and the anti-diagonal offset between its lower and upper sides, will have the order of the two-thirds power of the parallelogram's height.

Segments $[2\ell_{j-1}, 2\ell_j]$ will be indexed by $j \in [\![N]\!]$, with

$$\ell_j := \ell_{j-1} + 3^{-1}2^{3(j-1)/2k}k^{-1/2}n^{1/2}, \tag{10.3}$$

and $\ell_0 = k^{2/3}n^{1/3}$. Climb begins where takeoff ends, at the diagonal coordinate $2\ell_0$: see (10.1). By level j, we mean $\{(x,y) \in \mathbb{Z}^2 : x + y = 2\ell_j\}$. The i^{th} curve will intersect level j at a unique point $(\ell_j - p_{i,k}^{(j)}, \ell_j + p_{i,k}^{(j)})$, where $p_{i,k}^{(j)}$ is inductively defined by

$$p_{i,k}^{(j)} := p_{i,k}^{(j-1)} + \Big(i - \frac{m}{2}\Big)\, 2^{(j-1)/k}\big(2^{1/k} - 1\big)k^{-1/3}n^{1/3} \tag{10.4}$$

from initial data $p_{i,k}^{(0)} := (i - m/2)k^{-1/3}n^{1/3}$ that is chosen consistently with (10.1).

We indicated that separation would increase by a factor of $2^{1/k}$ during each segment in climb from an initial value of $k^{-1/3}n^{1/3}$. This is what our definition ensures: the separation $\mathrm{sep}^{(j)}$ between positions on the j^{th} level is given by

$$\mathrm{sep}^{(j)} := p_{i,k}^{(j)} - p_{i-1,k}^{(j)} = k^{-1/3}n^{1/3} + \sum_{r=0}^{j-1} 2^{r/k}(2^{1/k} - 1)k^{-1/3}n^{1/3} = 2^{j/k}k^{-1/3}n^{1/3}, \tag{10.5}$$

where the second equality is due to repeated applications of (10.4) and since $\mathrm{sep}^{(0)} = p_{i,k}^{(0)} - p_{i-1,k}^{(0)} = k^{-1/3}n^{1/3}$.

The height h that marks the end of the climb phase may now be set:

$$h := \ell_N. \tag{10.6}$$

Flight corridors.

We have indicated that the curves under construction will be forced to pass through certain disjoint parallelograms. The latter regions will be called *flight corridors*. We may consider a flight corridor delimited by levels j and $j + 1$ of width $\mathrm{sep}^{(j)}$.

Each curve will have a single flight corridor through which it passes between levels j and $j + 1$. The corridor corresponding to the i^{th} curve will be denoted $P^{(i,j,s)}$, and we define it next. Consider the planar line segment that runs from

$$\begin{gathered}\left(\ell_j - p_{i,k}^{(j)},\, \ell_j + p_{i,k}^{(j)}\right) + \left(\ell_{j+1} - p_{i,k}^{(j+1)},\, \ell_{j+1} + p_{i,k}^{(j+1)}\right) + (1, 0) \\ \text{to} \\ \left(\ell_j - p_{i,k}^{(j)},\, \ell_j + p_{i,k}^{(j)}\right) + \left(\ell_{j+1} - p_{i,k}^{(j+1)},\, \ell_{j+1} + p_{i,k}^{(j+1)}\right).\end{gathered} \tag{10.7}$$

The flight corridor $P^{(i,j,s)}$ consists of those $(x, y) \in \mathbb{Z}^2$ that are displaced from some element in the just indicated line segment by a vector $(t, -t)$ for some t with absolute value at most $2^{-1}\mathrm{sep}^{(j)} - 1$. Thus the flight corridors associated to the i^{th} curve for different values of i are disjoint. See Figure 10.1. The addition of $(1, 0)$ in the first line of (10.7) is to ensure that the consecutive flight corridors of the same curve are disjoint.

The subpath of the i^{th} curve from level j to $j + 1$ is defined to be a path of maximum weight between the two planar points just displayed that remains in the flight corridor $P^{(i,j,s)}$. Its weight will be denoted by $X_{n,k,2}^{(i,j,s)}$. Note that the level j subpath's ending point in $P^{(i,j,s)}$ and the level $j+1$ subpath's starting point in $P^{(i,j+1,s)}$ are adjacent to each other, which allows us to concatenate them.

The height of $P^{(i,j,s)}$ is $\left(\ell_{j+1} - \ell_j\right) = 3^{-1}2^{3j/2k}k^{-1/2}n^{1/2}$ and its width is $\mathrm{sep}^{(j)} = 2^{j/k}k^{-1/3}n^{1/3}$; so the flight corridor satisfies the relation height $= \frac{1}{3} \cdot \mathrm{width}^{3/2}$.

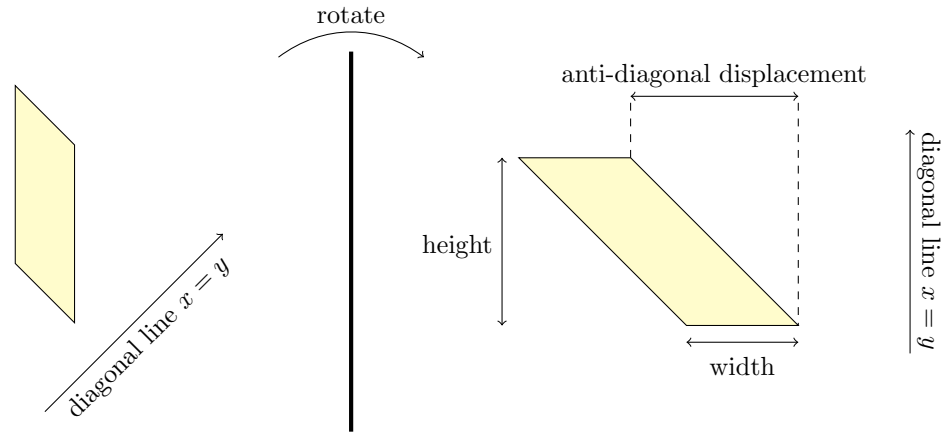


Figure 10.2: In order that the weight loss of the heaviest path constrained to lie in a parallelogram be on-scale, we need the anti-diagonal displacement and width of the parallelogram to both be of order $(\text{height})^{2/3}$. On the left is the parallelogram in the original coordinates, which has been rotated on the right for clarity so that the diagonal $x = y$ is visually vertical.

Next we will consider the anti-diagonal displacement (see Figure 10.2) of the flight corridors. Recall that Proposition 7.17 says that, if $z(\text{height})^{2/3}$ denotes the anti-diagonal displacement, then the weight loss of the heaviest path constrained to connect the midpoints of the opposite sides of $P^{(i,j,s)}$ is of order $z^2(\text{height})^{1/3}$. Thus to ensure a low enough weight loss, it will be important that the anti-diagonal displacement of $P^{(i,j,s)}$ is at most of order $(\text{height})^{2/3}$, i.e., that z is of constant order. We check that this holds next; indeed, it is for this relation to hold that we defined our separation to increase by a constant factor raised to the power of $1/k$, instead of the naively more natural choice of simply a constant factor.

The anti-diagonal displacement of the flight corridor is highest for curves 1 and m. For these paths, this displacement between levels j and $j+1$ is

$$\left|p_{i,k}^{(j+1)} - p_{i,k}^{(j)}\right| = \frac{1}{2} m 2^{j/k}(2^{1/k} - 1)k^{-1/3}n^{1/3} \leq Cm2^{j/k}k^{-4/3}n^{1/3}$$

in view of (10.4), with the inequality following from the bound $2^{1/k} - 1 \leq Ck^{-1}$ for some absolute constant C. As the height gain in $P^{(i,j,s)}$ is $3^{-1}2^{3j/2k}k^{-1/2}n^{1/2}$, and $m \leq k$, the anti-diagonal displacement for each corridors is indeed at most $C \cdot (\text{height})^{2/3}$; the constant C is uniform over i, j, k, and m.

Cruise

Curves enter cruise at separation $k^{-2/3}n^{2/3}$, and must be maintained there for a duration of order n: see Figure 10.1b. More precisely, and recalling that $h = \ell_N$, the i^{th} curve enters cruise at

$$\left(h - \left(\frac{m}{2} - i\right)k^{-2/3}n^{2/3}, h + \left(\frac{m}{2} - i\right)k^{-2/3}n^{2/3}\right).$$

We set $\ell'_0 = h$, and $\ell'_j = \ell'_{j-1} + \frac{n-2h}{k} = h + (n-2h) \cdot \frac{j}{k}$ for $j \in [\![k]\!]$. We will demand that the i^{th} curve visits each point

$$\left(\ell'_j - \left(\frac{m}{2} - i\right) k^{-2/3} n^{2/3}, \ell'_j + \left(\frac{m}{2} - i\right) k^{-2/3} n^{2/3}\right). \tag{10.8}$$

Between consecutive points, this curve will be be constrained to remain in a flight corridor that comprises those vertices in $\mathbb{Z}^2$ that are displaced from the planar line segment that interpolates the pair of points by a vector $(t, -t)$ for some t with absolute value at most $2^{-1} k^{-2/3} n^{2/3}$, which will be made to be disjoint by a unit displacement of the bottom side, as in (10.7). Let $X^{(i,j)}_{n,k,3}$ be the weight of the i^{th} curve in the j^{th} corridor; namely, the maximum weight of paths that interpolate the endpoint pair and remain in the corridor.

Momentarily, we will verify that $h = \Theta(n)$. The new flight corridor thus satisfies the relation

$$\text{width} \in \left[C^{-1} \cdot \text{height}^{2/3}, C \cdot \text{height}^{2/3}\right]$$

for a positive constant C. The corridor has no anti-diagonal displacement.

Lemma 10.2. *There exist positive c_1 and c such that if $k < c_1 n$, then $h = \ell_N$ satisfies $cn \le h \le n/2$ for large enough n.*

Proof. This inference follows by noting from the defintion (10.3) of ℓ_N that

$$h = \ell_0 + \sum_{j=0}^{N-1} 3^{-1} 2^{3j/2k} k^{-1/2} n^{1/2} = k^{2/3} n^{1/3} + \frac{1}{3} \times \frac{2^{3N/2k} - 1}{2^{3/2k} - 1} \times k^{-1/2} n^{1/2}. \tag{10.9}$$

Now, since $2^{3/2k} - 1 = \Theta(k^{-1})$ and $2^{N/k} \in [1, 2) \cdot k^{-1/3} n^{1/3}$ from (10.2), the claim follows after simplification. □

Descent and landing

These two phases are specified symmetrically with climb and take-off.

10.3 Bounding below the expected weight of the construction

We will argue, in two steps, that the constructed curves attain the sought weight $\mu nm - \Theta(mk^{2/3} n^{1/3})$ with high probability. In this section, we will show that, under Assumptions 2 and 3, the expected total curve weight is at least $\mu nm - Cmk^{2/3} n^{1/3}$ for a constant $C > 0$. In the next, the desired bound on curve weight will be obtained by showing that this weight concentrates around its mean due to the independence of contributions from the various flight corridors.

We will control the contribution from a given flight corridor using the lower bound on the expectation of the constrained point-to-point weight from Proposition 7.17. That proposition will permit us to derive the next result, the conclusion of the present section.

Proposition 10.3. *Let X denote the sum of the weights of the m curves in our construction. There exist $n_0 \in \mathbb{N}$, $C > 0$ and $c_1 > 0$ such that, for $n > n_0$, $k < c_1 n$ and $m \leq k$, $\mathbb{E}[X] \geq \mu nm - Cmk^{2/3}n^{1/3}$.*

Proof. The values of C and c may change from line to line but they do not depend on n, m, or k.

We start with the weight of a curve fragment between two levels during climb. That is, we find a lower bound on $\mathbb{E}[X_{n,k,2}^{(i,j,s)}]$, where recall that $X_{n,k,2}^{(i,j,s)}$ the weight of the heaviest path that travels between the points (10.7) without exiting the flight corridor $P^{(i,j,s)}$.

In specifying the climb phase, we noted that the condition in Proposition 7.17 that the anti-diagonal displacement of the flight corridor is of order $r^{2/3}$ holds, where r is the height of the flight corridor. Further, this holds uniformly over the m flight corridors corresponding to the m values of i and the N levels indexed by j.

Note that this implies that the parabolic weight loss term of $\mathbb{E}[X_{n,k,2}^{(i,j,s)}]$ in (7.8) of Proposition 7.17 is of order $r^{1/3}$ and so, by absorbing it into the term $-cr^{1/3}$, does not need to be explicitly calculated. Thus, using Proposition 7.17 with the settings $r = \ell_{j+1} - \ell_j$ and $zr^{2/3} = p_{i,k}^{(j+1)} - p_{i,k}^{(j)}$, we obtain

$$\mathbb{E}[X_{n,k,2}^{(i,j,s)}] \geq \mu\left(\ell_{j+1} - \ell_j\right) - C\left(\ell_{j+1} - \ell_j\right)^{1/3}. \tag{10.10}$$

Here C also includes the weight loss from the parabolic term in (7.8) of Proposition 7.17, since, as noted, z is uniformly bounded.

Recall from (10.3) that $\ell_{j+1} - \ell_j = 3^{-1}2^{3j/2k}k^{-1/2}n^{1/2}$. Simplifying the second term of (10.10) gives

$$C\left(\ell_{j+1} - \ell_j\right)^{1/3} = C \cdot 2^{j/2k}k^{-1/6}n^{1/6}.$$

where we have absorbed the constant factor of $3^{-1/3}$ into the value of C. Combining the preceding two displays and summing over i gives a lower bound on the expected weight of all the curves between levels j and $j+1$:

$$\sum_{i=1}^{m} \mathbb{E}[X_{n,k,2}^{(i,j,s)}] \geq \frac{\mu}{3} \cdot 2^{3j/2k}mk^{-1/2}n^{1/2} - C2^{j/2k}mk^{-1/6}n^{1/6}$$

Summing from $j = 0$ to $N-1$ gives a lower bound for the total weight of the climb phase of

$$\begin{aligned}\sum_{j=0}^{N-1}\sum_{i=1}^{m} \mathbb{E}[X_{n,k,2}^{(i,j,s)}] &\geq \frac{\mu m}{3} \cdot \sum_{j=0}^{N-1} 2^{3j/2k}k^{-1/2}n^{1/2} - C \cdot \sum_{j=0}^{N-1} 2^{j/2k}mk^{-1/6}n^{1/6} \\ &= \mu(h - k^{2/3}n^{1/3}) \cdot m - C \cdot \frac{2^{N/2k} - 1}{2^{1/2k} - 1} \cdot mk^{-1/6}n^{1/6} \qquad (10.11) \\ &= \mu hm - Cmk^{2/3}n^{1/3}, \qquad (10.12)\end{aligned}$$

where we used that $2^{N/2k} = k^{-1/6}n^{1/6}$ from (10.2); that $2^{1/2k} - 1 \geq ck^{-1}$ for some $c > 0$; the expression of h from (10.9); and that $\ell_0 = k^{2/3}n^{1/3}$.

As for cruise, an easy computation from Proposition 7.17 yields that

$$\mathbb{E}[X^{(i,j)}_{n,k,3}] \geq \frac{\mu(n-2h)}{k} - ck^{-1/3}(n-2h)^{1/3} \geq \frac{\mu(n-2h)}{k} - ck^{-1/3}n^{1/3}. \tag{10.13}$$

We write $\overline{X}^{(i,j,s)}_{n,k,2}$ for the descent counterpart of the climb flight corridor weight maximum. The bound on mean valid for climb holds equally for descent. Combining climb, descent and cruise weight bounds with the zero bound for take-off and landing, the total weight X in our construction is seen to satisfy

$$X \geq \sum_{i,j} \left(X^{(i,j,s)}_{n,k,2} + \overline{X}^{(i,j)}_{n,k,2} + X^{(i,j)}_{n,k,3} \right) \tag{10.14}$$

(where the indices that j is summed over differs for the terms corresponding to the climb phase and the cruise phase). Taking expectation, and applying (10.12) and (10.13), we obtain

$$\mathbb{E}\left[\sum_{i,j} \left(X^{(i,j,s)}_{n,k,2} + \overline{X}^{(i,j)}_{n,k,2} + X^{(i,j)}_{n,k,3} \right) \right] \geq \mu nm - Cmk^{2/3}n^{1/3},$$

the conclusion of Proposition 10.3. □

10.4 One-sided concentration of the construction weight

We now know that the expected weight of the construction is correct. To prove Theorem 7.16, we argue that the weight is unlikely to fall much below its mean. We will stochastically dominate the summands in the expression for X from (10.14) with independent exponential random variables of varying parameters, and use the following concentration result from [Jan18] for sums of such variables. We denote by $\mathrm{Exp}(\lambda)$ the exponential distribution with rate λ.

Proposition 10.4 (Theorem 5.1 (i) of [Jan18]). *Let $W = \sum_{i=1}^n W_i$ where $W_i \sim \mathrm{Exp}(a_i)$ are independent. Define*

$$\nu := \mathbb{E}W = \sum_{i=1}^n \mathbb{E}W_i = \sum_{i=1}^n \frac{1}{a_i}, \quad a_* := \min_i a_i.$$

Then for $\lambda \geq 1$,

$$\mathbb{P}(W \geq \lambda\nu) \leq \lambda^{-1} e^{-a_*\nu(\lambda - 1 - \log\lambda)}.$$

Proof of Theorem 7.16. Let $K \geq 0$ and $0 < L_1 \leq L_2$. As before, let $U = U_{r,\ell r^{2/3},zr^{2/3}}$ be a parallelogram with width $\ell r^{2/3}$, height r, and anti-diagonal displacement $zr^{2/3}$, where $|z| \leq K$ and $L_1 \leq \ell \leq L_2$. Let X^U_r denote the weight of the heaviest midpoint-to-midpoint path that lies

in U. By our upgraded-by-bootstrapping tail bound for constrained weighs Proposition 9.6, and $\mathbb{E}[X_r^U] \leq \mathbb{E}[X_r] \leq \mu r$ (which is due to Assumption 2), there exists $c_4 = c_4(L_1, L_2, K) > 0$ such that, for $r > r_0 = r_0(L_1, L_2, K)$ and $\theta > \theta_0 = \theta_0(L_1, L_2, K)$,

$$\mathbb{P}\left(X_r^U - \mathbb{E}[X_r^U] < -\theta r^{1/3}\right) \leq e^{-c_4\theta}\,;$$

or

$$\mathbb{P}\left(X_r^U - \mathbb{E}[X_r^U] < -\theta\right) \leq \exp(-c_4\theta/r^{1/3})\,, \tag{10.15}$$

the latter for θ larger than $\theta_0 r^{1/3}$ and $r > r_0$. Equation (10.15) says that we have the stochastic domination

$$-\left(X_r^U - \mathbb{E}[X_r^U]\right) \leq_{\text{sd}} \text{Exp}(c_4/r^{1/3}) + \theta_0 r^{1/3},$$

where, in an abuse of notation, $\text{Exp}(\lambda)$ denotes a random variable with the exponential distribution of rate λ, and $X \leq_{\text{sd}} Y$ denotes that the distribution of X is stochastically dominated by that of Y. Thus, in our construction, we have a coupling

$$\begin{aligned}
-\left(X_{n,k,2}^{(i,j,s)} - \mathbb{E}[X_{n,k,2}^{(i,j,s)}]\right) &\leq W_2^{(i,j)} + \theta_0 2^{j/2k} k^{-1/6} n^{1/6},\\
-\left(\overline{X}_{n,k,2}^{(i,j)} - \mathbb{E}[\overline{X}_{n,k,2}^{(i,j)}]\right) &\leq \overline{W}_2^{(i,j)} + \theta_0 2^{j/2k} k^{-1/6} n^{1/6}, \quad \text{and}\\
-\left(X_{n,k,3}^{(i,j)} - \mathbb{E}[X_{n,k,3}^{(i,j)}]\right) &\leq W_3^{(i,j)} + \theta_0 k^{-1/3} n^{1/3}.
\end{aligned} \tag{10.16}$$

Here the random variables on the right-hand side are independent and distributed as

$$\begin{aligned}
W_2^{(i,j)} &\sim \text{Exp}\left(c_4 2^{-j/2k} k^{1/6} n^{-1/6}\right),\\
\overline{W}_2^{(i,j)} &\sim \text{Exp}\left(c_4 2^{-j/2k} k^{1/6} n^{-1/6}\right), \quad \text{and}\\
W_3^{(i,j)} &\sim \text{Exp}\left(c_4 k^{1/3} n^{-1/3}\right).
\end{aligned} \tag{10.17}$$

By our construction, K, L_1, and L_2 in Proposition 9.6 may be chosen independently of i and j. Thus, whatever the dependence of c_4 and θ_0 on i and j, they are uniformly bounded away from 0 and ∞ respectively, and so may be assumed to be constant. We set

$$W := \sum_{i,j}\left(W_2^{(i,j)} + \overline{W}_2^{(i,j)}\right) + \sum_{i,j} W_3^{(i,j)}.$$

We wish to use Proposition 10.4 on the sum W of independent exponential random variables, and so we must estimate Proposition 10.4's parameters. We have

$$\mathbb{E}\left[\overline{W}_2^{(i,j)}\right] = \mathbb{E}\left[W_2^{(i,j)}\right] = c_4^{-1} 2^{j/2k} k^{-1/6} n^{1/6} \qquad \text{and} \qquad \mathbb{E}\left[W_3^{(i,j)}\right] = c_4^{-1} k^{-1/3} n^{1/3}.$$

Summing the expressions of the last display over the indices and using that $2^{N/2k} = k^{-1/6} n^{1/6}$ from (10.2) and that $2^{1/k} - 1 = \Theta(k^{-1})$ gives that the total mean ν satisfies

$$\nu := \mathbb{E}[W] \leq 8 c_4^{-1} m k^{2/3} n^{1/3}. \tag{10.18}$$

Noting that the coefficients of θ_0 in (10.16) are the same as the mean of the corresponding exponential random variable up to a factor of c_4, similarly summing these coefficients shows the stochastic domination

$$-(X - \mathbb{E}[X]) \leq W + 8\theta_0 mk^{2/3}n^{1/3}.$$

Using that $2^{N/k} = k^{-1/3}n^{1/3}$ from (10.2), we see that the minimum a^* of the rates of the exponential random variables defined in (10.17) is given by

$$a^* = \min(c_4 2^{-N/2k}k^{1/6}n^{-1/6}, c_4 k^{1/3}n^{-1/3}) = c_4 k^{1/3}n^{-1/3}. \tag{10.19}$$

Let $S = \max(c_4^{-1}, \theta_0)$. By Proposition 10.3, with this result contributing the value of C, and the stochastic domination (10.16),

$$\begin{aligned}\mathbb{P}\Big(X < \mu nm - (C + 20S)mk^{2/3}n^{1/3}\Big) &= \mathbb{P}\left(-(X - \mathbb{E}[X]) > 20Smk^{2/3}n^{1/3}\right) \\ &\leq \mathbb{P}\left(W \geq 20Smk^{2/3}n^{1/3} - 8\theta_0 mk^{2/3}n^{1/3}\right).\end{aligned}$$

Since $20S - 8\theta_0 \geq 12S \geq 12c_4^{-1}$, the latter quantity is bounded using Proposition 10.4, (10.18), and (10.19):

$$\begin{aligned}\mathbb{P}\left(W > 12c_4^{-1}mk^{2/3}n^{1/3}\right) \leq \mathbb{P}\left(W > \frac{12}{8}\nu\right) &\leq \frac{2}{3}\exp\left(-a^*\nu\left(\frac{3}{2} - 1 - \log\frac{3}{2}\right)\right) \\ &\leq \exp\left(-cmk\right).\end{aligned}$$

Numerical evaluation of the exponent shows that we may take $c = 3/4$.

In view of the paragraphs after (10.7) and (10.8), flight corridors during climb, descent and cruise lie within the strip around the diagonal of width $2mk^{-2/3}n^{2/3}$; thus, the transversal fluctuations of the constructed curves in these phases also satisfy this bound. By setting $C_1 = C + 20S$, we complete the proof of Theorem 7.16. □

Remark 10.5. The argument just given shows that, for positive constants c and x_0 and $x \geq x_0$,

$$\mathbb{P}\Big(X < \mu nm - (C + x)mk^{2/3}n^{1/3}\Big) \leq \exp\left(-cmkx\right).$$

Part IV

Twin peaks for the KPZ fixed point

Chapter 11

The lower bound on the twin peaks probability

This chapter forms the third pillar of this thesis, which, as discussed in Chapter 2, combines the probabilistic and geometric perspectives to prove a lower bound on the probability of "ε-twin peaks" for the KPZ fixed point. It consists of work done with Ivan Corwin, Alan Hammond, and Konstantin Matetski in [CHHM21].

We have to state the precise version of Theorem 2.9, which was stated informally in Chapter 2. To do so, we first define the assumptions on the initial data we work under.

Assumption 11.1 (Parabolic decay)**.** *The initial data $\mathfrak{h}_0 : \mathbb{R} \to [-\infty, \infty)$ of the KPZ fixed point is a non-random, upper semicontinuous function that is not identically equal to $-\infty$, and for which (a) there exist $\alpha \in \mathbb{R}$ and $\gamma > 0$ such that the bound $\mathfrak{h}_0(y) \leq \alpha - \gamma y^2$ holds for all $y \in \mathbb{R}$ and (b) there exists $\lambda \in \mathbb{R}$ such that $\mathfrak{h}_0(y) = -\infty$ for $y \leq -\lambda$.*

Next we turn to the precise definition of the ε-twin peaks set. For $L > 0$, we denote

$$\mathsf{J}_L := [-L, L]. \tag{11.1}$$

We will estimate the probability that the KPZ fixed point has two near maximizers that differ by at least some fixed $A > 0$. To this end, for a function $f : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ that is bounded from above, and for three values $A, L, \varepsilon > 0$, we specify the set

$$\mathsf{S}^{\varepsilon}_{A,L}(f) := \Big\{(x_1, x_2) \in \mathsf{J}_L^2 : x_2 - x_1 \geq A,\ \mathrm{Max}(f) - f(x_i) \leq \varepsilon \text{ for } i = 1, 2\Big\}. \tag{11.2}$$

Then we define the set

$$\mathsf{TP}^{\varepsilon}_{A,L} := \Big\{f : \mathbb{R} \to \mathbb{R} \cup \{-\infty\} : \mathrm{Max}(f) \in \mathsf{J}_{L^{1/2}},\ \mathsf{S}^{\varepsilon}_{A,L}(f) \neq \emptyset\Big\}. \tag{11.3}$$

Now we may state the precise version of the previously stated Theorem 2.9.

Theorem 11.2. *Let the initial state $\mathfrak{h}_0$ of the KPZ fixed point satisfy Assumption 11.1, and suppose that $A > 0$ and $t > 0$. There exist L_0 and $\eta > 0$ (both depending on A, t, and $\mathfrak{h}_0$) such that, for all $L \geq L_0$ and $\varepsilon \in (0,1)$,*

$$\mathbb{P}_{\mathfrak{h}_0}\left(\mathfrak{h}_t \in \mathsf{TP}^{\varepsilon}_{A,L}\right) \geq \eta\varepsilon. \tag{11.4}$$

Further, η and L_0 can be taken to depend on $t > 0$ in a continuous manner.

Theorem 11.2 offers a lower bound on the probability of the twin peaks' event. The general t version of this result follows from the $t = 1$ case in light of the 1:2:3 invariance of the KPZ fixed point recorded next, and taken from [MQR17, Section 4.3]:

Proposition 11.3 (1:2:3 scaling invariance)**.** *For any $\alpha > 0$, one has $\alpha\mathfrak{h}_{\alpha^{-3}t}(\alpha^{-2}x; \mathfrak{h}_0^{\alpha}) \stackrel{d}{=} \mathfrak{h}_t(x; \mathfrak{h}_0)$ as a process in both x and t, where $\mathfrak{h}_0^{\alpha}(x) := \alpha^{-1}\mathfrak{h}_0(\alpha^2 x)$ and $\mathfrak{h}_t(\cdot; \mathfrak{h}_0)$ denotes the KPZ fixed point started from initial condition $\mathfrak{h}_0$.*

Thus in this section we set $t = 1$ and, for notational convenience, we will denote $\mathfrak{h}_1$ by $\mathfrak{h}$; the claim in Theorem 11.2 that η and L_0 can be taken to depend on t continuously will be handled separately.

In fact, while Assumption 11.1 says that there exist α, γ, and λ such that $\mathfrak{h}_0(y)$ is $-\infty$ for all $y < -\lambda$ and satisfies $\mathfrak{h}_0(y) \leq \alpha - \gamma y^2$ for all $y \in \mathbb{R}$, we will also assume without loss of generality that $\alpha = \lambda = 0$. We may do so because the occurrence of the twin peaks' event $\mathsf{TP}^{\varepsilon}_{A,L}$ is unchanged, provided that L is altered suitably, under arbitrary shifts of $\mathfrak{h}_0$ in the horizontal and vertical directions; i.e., under transformations of the form $\mathfrak{h}_0(\cdot) \mapsto \mathfrak{h}_0(\cdot + z_1) + z_2$ for fixed $z_1, z_2 \in \mathbb{R}$. So we consider the function $\tilde{\mathfrak{h}}_0(y) := \mathfrak{h}_0(y - \lambda) - (\alpha + \gamma\lambda^2)$, which satisfies $\tilde{\mathfrak{h}}_0(y) = -\infty$ for all $y < 0$. For this function, we have that

$$\tilde{\mathfrak{h}}_0(y) \leq \alpha - (\alpha + \gamma\lambda^2) - \gamma(y - \lambda)^2 \leq -\frac{\gamma}{2}y^2, \text{ for all } y \in \mathbb{R},$$

by noting that the arithmetic-geometric mean inequality implies that $2y\lambda \leq y^2/2 + 2\lambda^2$. We will relabel $\gamma/2$ by γ since this quantity differs from γ by a constant multiple that is independent of α and λ. Thus, for $\mathfrak{h}_0$ satisfying Assumption 11.1, we may indeed assume that $\alpha = \lambda = 0$, i.e.,

$$\mathfrak{h}_0(y) \leq -\gamma y^2 \text{ for all } y \in \mathbb{R} \text{ and } \mathfrak{h}_0(y) = -\infty \text{ for } y < 0; \tag{11.5}$$

of course, as mentioned previously, making this simplification may need a modification in the value of L for which we consider the event $\mathsf{TP}^{\varepsilon}_{A,L}$. More precisely, we have that $\mathfrak{h}_t \in \mathsf{TP}^{\varepsilon}_{A,L} \iff \tilde{\mathfrak{h}}_t \in \mathsf{TP}^{\varepsilon}_{A,\tilde{L}}$ where $\tilde{\mathfrak{h}}_t$ is obtained from $\mathfrak{h}_t$ by applying the same horizontal and vertical shift that was applied above to $\mathfrak{h}_0$ to give $\tilde{\mathfrak{h}}_0$ (so $\tilde{\mathfrak{h}}_t$ has the distribution of the KPZ fixed point started from initial condition $\tilde{\mathfrak{h}}_0$ at time t) and $\tilde{L}$ is a function of L, α, γ and λ. This simplification will aid us in some later technical arguments.

Many of the estimates made in this section will depend on $\mathfrak{h}_0$, and it will be helpful to be precise about which aspects of $\mathfrak{h}_0$ are relevant. Thus, in this section the parameter θ will be such that

$$\sup_{|y| \leq \theta} \mathfrak{h}_0(y) \geq -\theta.$$

Estimates will depend on $\mathfrak{h}_0$ only through γ and θ; this is because we assume as above that $\alpha = \lambda = 0$, but note that the transformation described to make this simplification modifies the value of θ.

Instead of working with the KPZ fixed point directly, we work mainly with the prelimiting model of Brownian last passage percolation. By doing so, we gain access to certain important tools: these include the Brownian description given by the distributional identification of the "melon" function of Brownian LPP (to be introduced shortly) with Dyson Brownian motion; and a remarkable deterministic identity (Proposition 11.4) relating last passage values in the Brownian environment with those in the melon-transformed environment. We work with the prelimiting model since the key identity has not been proved at the level of the KPZ fixed point. Once we establish a prelimiting version of Theorem 11.2 (Proposition 11.8 below), we use the convergence of Brownian LPP to the KPZ fixed point (Lemma 11.7) to deduce the theorem. Below, in Figure 11.1, is a diagrammatic representation of the structure of this section.

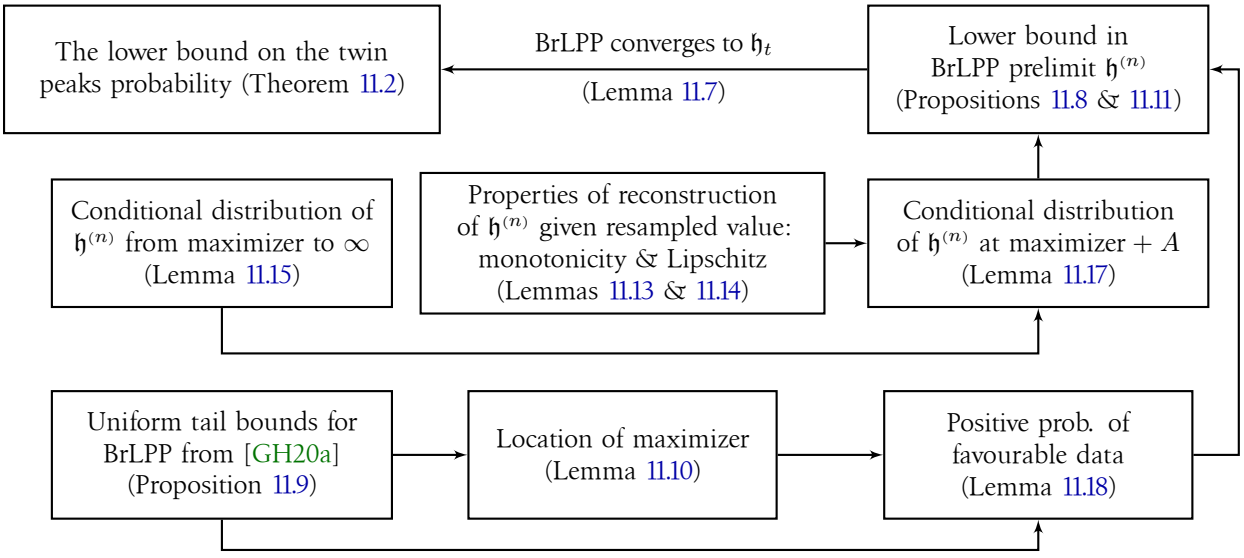


Figure 11.1: Structure of this chapter.

11.1 Preliminaries

The model

We quickly recall the model of semi-discrete LPP that was previously introduced in Chapter 2. We denote the integer interval $\{1, \dots, n\}$ by $[\![n]\!]$. Consider a sequence of continuous functions $f = (f_1, \dots, f_n) : [\![n]\!] \times [0, \infty) \to \mathbb{R}$. We will depict these functions as in Figure 11.2. The functions f_1 through f_n are each indexed by a spatial variable which lies respectively along one

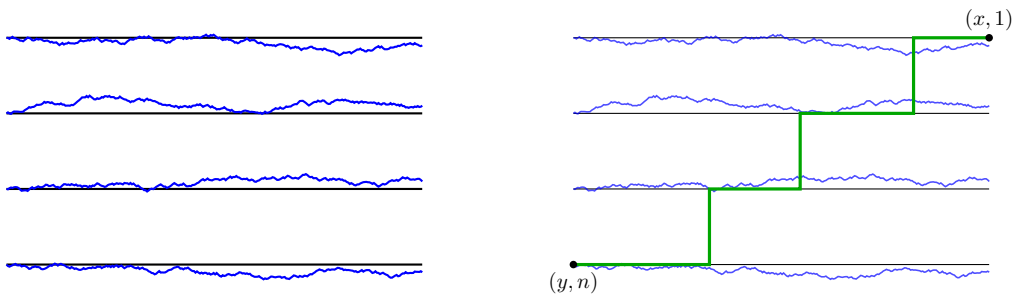


Figure 11.2: Left: A depiction of the environment given by f. The functions f_i corresponding to each line are graphed in red on the corresponding black line for visual clarity; the function values themselves need not be ordered. Right: An upright path γ from (y, n) to $(x, 1)$ is depicted in green; note that in a formal sense the depicted vertical portions are not part of the path. The path's weight is the sum of the increments of f_i along the portion of the i^{th} line γ spends on it.

of n horizontal *lines*, with the top line indexing f_1 and the bottom line indexing f_n. The function values along these lines represent an *environment*.

Let $0 \leq y \leq x$. An upright path γ from (y, n) (i.e., position y on line n) to $(x, 1)$ is a path which starts at (y, n) and moves rightwards, jumping up from one line to the next at certain times until it reaches $(x, 1)$: see Figure 11.2. An upright path is parametrized by its *jump times* $\{t_i\}_{i=1}^{n-1}$ at which it jumps from the $(i+1)^{\text{st}}$ line to the i^{th} line. Define $\Pi_{y,x}^n$ to be the set of upright paths from (y, n) to $(x, 1)$. The *weight* of $\gamma \in \Pi_{y,x}^n$ in f is denoted $f[\gamma]$ and defined by

$$f[\gamma] = \sum_{i=1}^{n-1} \Big(f_i(t_{i-1}) - f_i(t_i) \Big), \tag{11.6}$$

where $\{t_i\}_{i=1}^{n-1}$ are the jump times of γ, with $t_n = y$ and $t_0 = x$. This expression is thus the sum of increments of f along the portions of γ on each line. We define the last passage value in f from (y, n) to $(x, 1)$ by

$$f[(y, n) \to (x, 1)] = \sup_{\gamma \in \Pi_{y,x}^n} f[\gamma]. \tag{11.7}$$

If the set $\Pi_{y,x}^n$ is empty, i.e., if $y > x$, we define the passage value to be $-\infty$. The model of Brownian LPP is specified by taking f to be a collection of n independent standard Brownian motions defined on $[0, \infty)$.

The melon function

We define the weight of a collection of disjoint (except possibly at shared endpoints) upright paths as the sum of the weights of the constituent paths. Then for $j \in [\![n]\!]$, we define $f[(y, n)^j \to (x, 1)^j]$ to be the maximum weight over all collections of j disjoint paths from (y, n) to $(x, 1)$.

With this, we may define the *melon* function $Wf = ((Wf)_1, \dots, (Wf)_n) : [\![n]\!] \times [0, \infty) \to \mathbb{R}$ by

$$(Wf)_j(x) = f[(0,n)^j \to (x,1)^j] - f[(0,n)^{j-1} \to (x,1)^{j-1}], \tag{11.8}$$

for $j \geq 2$ and $(Wf)_1(x) = f[(0,n) \to (x,1)]$. Two important deterministic properties are that the curves of Wf are *ordered*, meaning that $(Wf)_i(\cdot) \geq (Wf)_{i+1}(\cdot)$ for each $i \in [\![n-1]\!]$ (see, for instance, the discussion at the start [DOV18, Section 4] and references given there); and an inference concerning last passage values in the melon environment:

Proposition 11.4 (Special case of Proposition 4.1 of [DOV18]). *Let $f = (f_1, \dots, f_n) : [\![n]\!] \times [0, \infty) \to \mathbb{R}$ be continuous, and let $y \leq x$. Then*

$$f[(y,n) \to (x,1)] = (Wf)[(y,n) \to (x,1)].$$

In particular, this proposition applies when $f = B$, a collection of n Brownian motions, as considered below in Section 11.1. An important technical condition imposed by the definition of Wf, as well as by Proposition 11.4, is that the domain of each line is $[0, \infty)$, rather than $\mathbb{R}$. It is because of this condition that we consider only initial conditions which are $-\infty$ for all small enough arguments; which is to say, this is why we require Assumption 11.1(b).

Scaling limit of Brownian LPP

We will need the convergence of Brownian LPP values to the parabolic Airy sheet, a convergence that holds uniformly on compact sets. Let $B : [\![n]\!] \times \mathbb{R} \to \mathbb{R}$ be a collection of n independent two-sided Brownian motions. Define $\mathcal{S}_n : \mathbb{R}^2 \to \mathbb{R} \cup \{-\infty\}$ by

$$\mathcal{S}_n(y,x) := n^{-1/3} \left(B[(2yn^{2/3}, n) \to (n + 2xn^{2/3}, 1)] - 2n - 2(x-y)n^{2/3} \right); \tag{11.9}$$

the co-domain includes $-\infty$ simply to handle the case that $2yn^{2/3} > n + 2xn^{2/3}$. (Although here we allow y to be negative, the definitions in (11.6) and (11.7) easily generalize.)

Here is a statement of convergence of Brownian LPP to the parabolic Airy sheet. (It is simply to have a cleaner version of this statement that we allowed y to be negative in (11.9).)

Proposition 11.5 (Theorem 1.3 of [DOV18]). *In the topology of uniform convergence on compact sets, we have the convergence in law*

$$\mathcal{S}(y,x) = \lim_{n \to \infty} \mathcal{S}_n(y,x). \tag{11.10}$$

We will generally work with a centred and scaled version of WB. Indeed, let $\mathcal{P}_n = (\mathcal{P}_{n,1}, \dots, \mathcal{P}_{n,n}) : [\![n]\!] \times [-\frac{1}{2}n^{1/3}, \infty) \to \mathbb{R}$ be given by

$$\mathcal{P}_{n,j}(x) = n^{-1/3} \left((WB)_j(n + 2xn^{2/3}) - (2n + 2xn^{2/3}) \right). \tag{11.11}$$

Here, $\mathcal{P}$ indicates "parabolic", as these objects converge to the parabolic Airy line ensemble (though we will not use this fact, as we only require the convergence of Brownian LPP values to the parabolic Airy *sheet* as in Proposition 5.3). Since $(WB)_j(\cdot)$ is ordered, and $(WB)_j(0) = 0$ for $j \in [\![n]\!]$, we see that, for $x \geq -\frac{1}{2}n^{1/3}$,

$$\begin{aligned} \mathcal{P}_{n,1}(x) &= n^{-1/3}\left(B[(0,n) \to (n + 2xn^{2/3}, 1)] - (2n + 2xn^{2/3})\right) \\ &= \mathcal{S}_n(0,x). \end{aligned} \tag{11.12}$$

We used the definition of $WB_{n,1}$ (11.8) for the first equality and (11.9) for the second.

Note also that

$$\mathcal{P}_n[(y,n) \to (x,1)] = n^{-1/3}\left((WB)[(n + 2yn^{2/3}, n) \to (n + 2xn^{2/3}, 1)] - 2(x-y)n^{2/3}\right), \tag{11.13}$$

for all $-\frac{1}{2}n^{1/3} \leq y < x$. We find then that, for $y > 0$ with $x > y - \frac{1}{2}n^{1/3}$,

$$\mathcal{S}_n(y,x) = \mathcal{P}_n[(-\tfrac{1}{2}n^{1/3} + y, n) \to (x,1)] - n^{2/3} \tag{11.14}$$

by comparing (11.13) to the definition (11.9) of $\mathcal{S}_n$, and using Proposition 11.4.

We may now define the prelimiting version of $\mathfrak{h}$, denoted $\mathfrak{h}^{(n)} : [-n^{1/60}, n^{1/60}] \to \mathbb{R} \cup \{-\infty\}$, by

$$\begin{aligned} \mathfrak{h}^{(n)}(x) &= \sup_{0 \leq y \leq n^{1/60}} \Big(\mathfrak{h}_0(y) + \mathcal{S}_n(y,x)\Big) \\ &= \sup_{0 \leq y \leq n^{1/60}} \left(\mathfrak{h}_0(y) + \mathcal{P}_n[(-\tfrac{1}{2}n^{1/3} + y, n) \to (x,1)] - n^{2/3}\right), \end{aligned} \tag{11.15}$$

since by assumption $\mathfrak{h}_0(y) = -\infty$ for $y < 0$. The final equality follows from (11.14). We adopt the upper limit of $n^{1/60}$ on y and $|x|$ in order to meet a technical hypothesis in the application of an upcoming estimate Proposition 11.9 from [GH20a]; note that $n^{1/60} \to \infty$ as $n \to \infty$ and so in the limit y and x can be thought of as respectively taking any non-negative value and any real value. Finally, for given x, a path in the environment defined by $\mathcal{P}_n$ which achieves the supremum implicit in $\mathcal{P}_n[(-\frac{1}{2}n^{1/3} + y, n) \to (x,1)]$ in the last equality of (11.15) is called a *geodesic*.

The next lemma translates the well-known fact that WB can be described as non-intersecting Brownian motions to a similar statement about $\mathcal{P}_n$'s distribution; see also Figure 11.3.

Lemma 11.6. *The distribution of $\mathcal{P}_n : [\![n]\!] \times [-\frac{1}{2}n^{1/3}, \infty) \to \mathbb{R}$ is that of n independent Brownian motions of rate two and of drift $-n^{1/3}$, with common initial value $-n^{2/3}$, and conditioned on mutual non-intersection.*

Proof. We may identify WB with n-level Dyson Brownian motion [OY02], which may be defined as n independent Brownian motions of rate one and zero drift, with common initial value zero, conditioned on mutual non-intersection (the singular conditioning made precise via a suitable limiting procedure or a Doob h-transform); see, for example, [Dys62, Meh91, Gra99]. The expressions for the rate, drift, and initial value in the sought statement follow from the definition (11.11) of $\mathcal{P}_n$. □

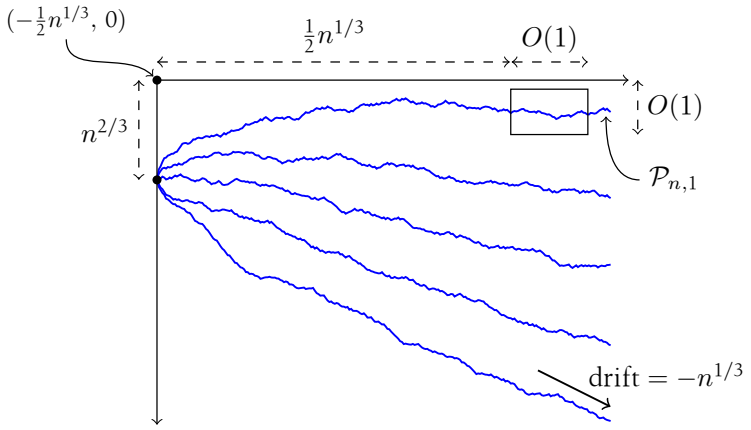


Figure 11.3: A depiction of $\mathcal{P}_n$ from (11.11). The vertical shift by $n^{-2/3}$, drift of $-n^{-1/3}$, and scaling by $n^{-1/3}$ are picked so that the height and fluctuations of $\mathcal{P}_{n,1}$ in a unit order interval are also of unit order, as emphasized by the black box of unit order height and width.

Since we ultimately need to make inferences about $\mathfrak{h}$, we require that $\mathfrak{h}^{(n)} \to \mathfrak{h}$ on compact sets. This is recorded in the next statement, which we will prove in Section 11.2 using the convergence of $\mathcal{S}_n$ to $\mathcal{S}$ from Proposition 11.5.

Lemma 11.7. *Let $\mathfrak{h}_0 : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ satisfy Assumption 11.1. Then we have that $\mathfrak{h}^{(n)} \to \mathfrak{h}$ in distribution in the topology of uniform convergence on compact sets.*

We will prove Theorem 11.2 by deriving the following analogous statement for $\mathfrak{h}^{(n)}$.

Proposition 11.8. *Let $\mathfrak{h}_0 : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ satisfy Assumption 11.1 and suppose that $A > 0$. There exist L_0 and $\eta > 0$ (both depending on γ, θ, and A) such that, for all $L > L_0$, there exists n_0 (depending on γ, θ, and L) so that for all $n > n_0$ and $\varepsilon \in (0, 1)$,*

$$\mathbb{P}\left(\mathfrak{h}^{(n)} \in \mathsf{TP}^{\varepsilon}_{A,L}\right) \geq \eta\varepsilon.$$

Further, L_0 and η may be taken to depend continuously on γ, θ, and A.

We show how Proposition 11.8 implies Theorem 11.2, and then turn to the proof of the former.

Proof of Theorem 11.2. As we have noted, we may take $t = 1$ and use the notation $\mathfrak{h}$ in place of $\mathfrak{h}_1$. Let $L > L_0$ with L_0 as given in Proposition 11.8. By combining the fact that $\mathfrak{h}^{(n)}$ converges to $\mathfrak{h}$ uniformly on compact sets (Lemma 11.7) with the Portmanteau theorem and Proposition 11.8, we see that

$$\mathbb{P}\big(\mathfrak{h} \in \mathsf{TP}^{\varepsilon}_{A,L}\big) \geq \lim_{n\to\infty} \mathbb{P}\left(\mathfrak{h}^{(n)} \in \mathsf{TP}^{\varepsilon}_{A,L}\right) \geq \eta\varepsilon.$$

For general $t > 0$, we must show that η and L_0 can be taken to depend on t continuously. This follows since the KPZ scaling, to move from the ε-twin peaks' event for $t > 0$ to $t = 1$, modifies A, γ, θ, and ε in a manner that depends continuously on t; the dependence of η and L_0 on these quantities in Proposition 11.8 is also continuous. This completes the proof of Theorem 11.2. □

We now turn to the proof of Proposition 11.8. We start with a section on some technical estimates regarding the location of maximizers of $\mathfrak{h}^{(n)}$ as well as of the maximum in its definition (11.15).

11.2 Locations of maximizers

We will need a uniform tail bound for $\mathcal{S}_n(y,x)$ as x and y vary over compact intervals. Such a bound is proved in [GH20a], and we state it here now.

Proposition 11.9 (Proposition 3.15 of [GH20a]). *There exist finite positive constants n_0, K_0, C, and c such that, for $n \geq n_0$, $K_0 \leq K \leq n^{1/30}$, and $0 < R < n^{1/46}$,*

$$\mathbb{P}\left(\sup_{x,y\in[-R,R]} \left| \mathcal{S}_n(y,x) + (y-x)^2 \right| > K \right) \leq CR^2 \exp(-cK^{3/2}).$$

The proof of Proposition 11.9 as given in [GH20a] is not difficult and follows a strategy used earlier in [BSS14] to prove a similar statement in another model of LPP; essentially, one considers a fine discretization of the set of endpoints in $[-R,R]$ and uses known one-point tail bounds and a union bound to get the uniform-over-endpoints statement.

It is to handle the bounds of $n^{1/30}$ and $n^{1/46}$ on K and R that we have imposed that $y \leq n^{1/60}$ in the definition of $\mathfrak{h}^{(n)}$ and restricted $\mathfrak{h}^{(n)}$'s domain to $[-n^{1/60}, n^{1/60}]$.

Next, we show uniform-in-n control over x_0^n, the maximizer of $\mathfrak{h}^{(n)}$.

Lemma 11.10. *Let $\mathfrak{h}_0 : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ satisfy Assumption 11.1 . Let x_0^n be the maximizer of $\mathfrak{h}^{(n)}$ of largest absolute value. Given $\delta > 0$, there exist n_0 and $M < \infty$ (both depending on γ, θ, and δ) such that, for all $n \geq n_0$,*

$$\mathbb{P}\left(|x_0^n| > M\right) \leq \delta.$$

Further, M can be taken to depend on γ, θ, and δ continuously.

Proof. Since by definition $x_0^n \in [-n^{1/60}, n^{1/60}]$, it is enough to prove that

$$\mathbb{P}\left(\sup_{M<|x|<n^{1/60}} \mathfrak{h}^{(n)}(x) > \mathfrak{h}^{(n)}(0) \right) < \delta$$

for large enough M depending only on δ, γ and θ. For any $K > 0$, we may bound above the left-hand side by

$$\mathbb{P}\left(\sup_{M<|x|<n^{1/60}} \mathfrak{h}^{(n)}(x) > \mathfrak{h}^{(n)}(0) \right) \leq \mathbb{P}\left(\sup_{M<|x|<n^{1/60}} \mathfrak{h}^{(n)}(x) > -K \right) + \mathbb{P}\left(\mathfrak{h}^{(n)}(0) < -K\right). \tag{11.16}$$

We will find a K such that both terms are less than $\delta/2$. The second term is easier to bound, and we address it first. Let $y_0 \in [-\theta, \theta]$ be such that $\mathfrak{h}_0(y_0) \geq -2\theta$ and set $K \geq 4\theta$. From the formula (11.15) for $\mathfrak{h}^{(n)}$, we see that

$$\mathbb{P}\left(\mathfrak{h}^{(n)}(0) < -K\right) \leq \mathbb{P}\Big(\mathfrak{h}_0(y_0) + \mathcal{S}_n(y_0, 0) < -K\Big) \leq \mathbb{P}\Big(\mathcal{S}_n(y_0, 0) < -K/2\Big).$$

We can bound this probability using the one-point tail information for $\mathcal{S}_n(y, x)$ from Proposition 11.9. Doing so shows that, for large enough K, the probability is less than $\delta/2$.

Returning to the first term on the right-hand side of (11.16), recall that, by (11.5), there exists $\gamma > 0$ such that $\mathfrak{h}_0(y) \leq -\gamma y^2$ for all $y \in \mathbb{R}$. Then we have from (11.15) that

$$\begin{aligned}\mathbb{P}\left(\sup_{M<|x|<n^{1/60}} \mathfrak{h}^{(n)}(x) > -K\right) &\leq \mathbb{P}\left(\sup_{\substack{M<|x|<n^{1/60}\\ 0\leq y\leq n^{1/60}}} \left(\mathcal{S}_n(y, x) - \gamma y^2\right) > -K\right)\\ &\leq \sum_{i=M}^{\lceil n^{1/60}\rceil} \sum_{j=0}^{\lceil n^{1/60}\rceil} \mathbb{P}\left(\sup_{\substack{|x|\in[i,i+1]\\ y\in[j,j+1]}} \mathcal{S}_n(y, x) > -K + \gamma j^2\right). \qquad (11.17)\end{aligned}$$

Now we want to apply Proposition 11.9 to bound each summand. We see that

$$\begin{aligned}&\mathbb{P}\left(\sup_{\substack{|x|\in[i,i+1]\\ y\in[j,j+1]}} \mathcal{S}_n(y, x) > -K + \gamma j^2\right)\\ &\qquad \leq \mathbb{P}\left(\sup_{\substack{|x|\in[i,i+1]\\ y\in[j,j+1]}} \left(\mathcal{S}_n(y, x) + (x-y)^2\right) > -K + \gamma j^2 + (|i-j|-1)^2\right). \qquad (11.18)\end{aligned}$$

We need a lower bound on the right-hand side of the preceding line's probability, which we record next. We claim that there exists M large enough depending on γ such that, for $i \geq M$ and $j \geq 1$,

$$\gamma j^2 + (|i-j|-1)^2 \geq \frac{\gamma}{4+\gamma}(i^2 + j^2). \qquad (11.19)$$

The proof of this essentially follows by noting that $2ij \leq (1+\gamma/2)j^2 + (1+\gamma/2)^{-1}i^2$ and we omit the further details.

Using (11.19) in (11.18) gives

$$\begin{aligned}&\mathbb{P}\left(\sup_{\substack{|x|\in[i,i+1]\\ y\in\gamma^{-1/2}[j,j+1]}} \left(\mathcal{S}_n(y, x) + (x-y)^2\right) > -K + \gamma j^2 + (|i-j|-1)^2\right)\\ &\qquad \leq \mathbb{P}\left(\sup_{\substack{|x|\in[i,i+1]\\ y\in[j,j+1]}} \left(\mathcal{S}_n(y, x) + (x-y)^2\right) > \frac{\gamma}{4+\gamma}\left(j^2 + \frac{1}{2}i^2\right)\right).\end{aligned}$$

Here we assumed that $K \le \frac{\gamma}{2(4+\gamma)} i^2$ for $i \ge M$, which can be assured by supposing that M is large enough. We apply Proposition 11.9 after noting that we are permitted to do so since for $i \le n^{1/60}$ and $j \le n^{1/60}$, we have $i^2, j^2 < n^{1/30}$ and $[j, j+1], [i, i+1] \subseteq [-n^{1/46}, n^{1/46}]$. Thus we see, for positive constants c and C depending on γ,

$$\mathbb{P}\left(\sup_{\substack{|x| \in [i,i+1] \\ y \in [j,j+1]}} \mathcal{S}_n(y,x) > -K + \gamma j^2 \right) \le C \max\{i^2, j^2\} \exp\big\{ -c(i^3 + j^3) \big\}.$$

Returning to the sum in (11.17) and substituting this bound yields, for M large enough,

$$\begin{aligned} \mathbb{P}\left(\sup_{M < |x| < n^{1/60}} \mathfrak{h}^{(n)}(x) > -K \right) &\le \sum_{i=M}^{\lceil n^{1/60} \rceil} \sum_{j=0}^{\lceil n^{1/60} \rceil} C \max\{i^2, j^2\} \exp\big\{ -c(i^3 + j^3) \big\} \\ &\le \sum_{i=M}^{\infty} C' i^2 \exp\{-ci^3\}, \end{aligned}$$

which may be made smaller than $\delta/2$ by choosing M suitably high (which overall depends on γ and θ) and by further assuming that $n > M^{60}$, if required. It may be easily checked that the dependence of M on these quantities is continuous. This completes the proof of Lemma 11.10. □

The proof of Lemma 11.7 on the convergence of $\mathfrak{h}^{(n)}$ to $\mathfrak{h}$ follows similar lines, relying on a bound on the location of the maximizer in the definition of $\mathfrak{h}^{(n)}$ (11.15). We give it now.

Proof of Lemma 11.7. Set $y_n(x)$ equal to $\arg\max_y \big(\mathfrak{h}_0(y) + \mathcal{S}_n(y,x)\big)$ (taking the choice of largest absolute value when it is not unique). Fix $M > 0$. We claim that $(y_n(x))_{n \in \mathbb{N}}$ is uniformly tight for $x \in [-M, M]$. To show this, let $\varepsilon > 0$ be given and let $R \ge M \vee \theta$, so that we may choose $y_0 \in [0, R]$ such that $\mathfrak{h}_0(y_0) \ge -2\theta$. Then we observe that, for every $K > 0$,

$$\begin{aligned} \mathbb{P}\big(y_n(x) > R\big) &\le \mathbb{P}\left(\sup_{R \le y \le n^{1/60}} \big(\mathfrak{h}_0(y) + \mathcal{S}_n(y,x)\big) > \mathfrak{h}_0(y_0) + \mathcal{S}_n(y_0, x) \right) \\ &\le \mathbb{P}\left(\sup_{R \le y \le n^{1/60}} \big(\mathfrak{h}_0(y) + \mathcal{S}_n(y,x)\big) > -K \right) + \mathbb{P}\Big(\mathfrak{h}_0(y_0) + \mathcal{S}_n(y_0, x) < -K\Big). \end{aligned} \tag{11.20}$$

We set $K \ge 4\theta$ large enough that $\mathfrak{h}_0(y_0) \ge -2\theta \ge -K/2$ and $K \ge (y_0 - x)^2$ (for example by setting $K \ge 4R^2$). Then we bound the second term of (11.20) by $CR^2 \exp(-cK^{3/2})$ (uniformly for all $x \in [-M, M]$) by Proposition 11.9. Thus, for all K large enough, the second term is at most $\varepsilon/2$.

Next we bound the first term of (11.20). By a union bound,

$$\begin{aligned}\mathbb{P}\left(\sup_{R\le y\le n^{1/60}}\big(\mathfrak{h}_0(y)+\mathcal{S}_n(y,x)\big)>-K\right) &\le \sum_{j=R}^{\lceil n^{1/60}\rceil}\mathbb{P}\left(\sup_{y\in[j,j+1]}\big(\mathcal{S}_n(y,x)-\gamma y^2\big)>-K\right)\\ &\le \sum_{j=R}^{\lceil n^{1/60}\rceil}\mathbb{P}\left(\sup_{y\in[j,j+1]}\mathcal{S}_n(y,x)>-K+\gamma j^2\right).\end{aligned}$$

Setting R such that $\gamma j^2 > K$ for all $j \ge R$ shows that each summand in the last display is bounded above by $Cj^2\exp(-cj^3)$, uniformly over $x\in[-M,M]$, again by Proposition 11.9. This expression is summable in j. So taking R sufficiently large implies that the sum is bounded above by $\varepsilon/2$. Thus, for such R, we find that, for $x\in[-M,M]$,

$$\mathbb{P}\big(y_n(x)>R\big)\le\varepsilon,$$

so that the claimed uniform tightness is obtained, because $y_n(x)\ge 0$ almost surely by our assumption on $\mathfrak{h}_0$.

That the maximizer sequence has a convergent subsequence, combined with the uniform convergence on compact sets of $\mathcal{S}_n$ to $\mathcal{S}$, implies that $\mathfrak{h}^{(n)}\to\mathfrak{h}$ uniformly on compact sets as well. To see this, fix $M>0$ and let $\mathcal{K}$ be a random compact set such that $y_n(x)\in\mathcal{K}$ for all n and $x\in[-M,M]$. Then simple manipulations show that

$$\sup_{x\in[-M,M]}|\mathfrak{h}^{(n)}(x)-\mathfrak{h}(x)|\le\sup_{\substack{x\in[-M,M]\\ y\in\mathcal{K}}}\Big|\mathcal{S}_n(y,x)-\mathcal{S}(y,x)\Big|\to 0. \qquad \square$$

11.3 The resampling framework

To prove Proposition 11.8, we will prove the following stronger proposition, from which the former immediately follows.

Proposition 11.11. *Let $\mathfrak{h}_0:\mathbb{R}\to\mathbb{R}\cup\{-\infty\}$ satisfy Assumption 11.1 and suppose $A>0$. There exist $\eta>0$ and $L_0>0$ (both depending on γ, θ, and A) such that, for all $L>L_0$, there exists n_0 (depending on γ, θ, and L) such that, for all $n>n_0$ and $\varepsilon\in(0,1)$,*

$$\mathbb{P}\left(\sup_{x\in[x_0^n+A,x_0^n+A+2]}\mathfrak{h}^{(n)}(x)>\mathfrak{h}^{(n)}(x_0^n)-\varepsilon;\ |\mathfrak{h}^{(n)}(x_0^n)|\le L^{1/2};\ |x_0^n|\le L-A-2\right)\ge\eta\varepsilon,$$

where $x_0^n=\arg\max_{|x|\le n^{1/60}}\mathfrak{h}^{(n)}(x)$ and is taken to be the largest (i.e., not necessarily greatest in absolute value) one on the event that it is not unique; we will use shorthand $x_0=x_0^n$ below at times. Further, L_0 and η can be taken to depend on γ, θ, and A continuously.

The strengthening of Proposition 11.11 relative to Proposition 11.8 is that we now assert that it is possible to achieve the twin peaks' event of separation A by moving at most distance two to the right of the maximizer x_0 beyond the imposed distance A.

The proof of Proposition 11.11 follows a Gibbsian resampling argument. (We will recall the Brownian Gibbs property precisely in Section 11.4.) This argument is considerably easier in the case where $\mathfrak{h}_0$ is a narrow wedge; in Section 11.5, we explain how this case works and then give the more general proof of this proposition. To set up the argument, we must first recall that $\mathfrak{h}^{(n)}$ can be expressed in terms of $\mathcal{P}$ via the variational problem in (11.15); and that $\mathcal{P}$ can be expressed in terms of non-intersecting Brownian motions via Lemma 11.6. Roughly put, then, $\mathfrak{h}^{(n)}$ may be expressed in terms of non-intersecting Brownian motions. We will make use of the Gibbs resampling property for these motions, filtered through the variational problem. To do this, we need to define a σ-algebra $\mathcal{F}$ that contains the data which will *not* be resampled. We will study the $\mathcal{F}$-conditional distribution of $\mathfrak{h}^{(n)}$ on $[x_0+A, x_0+A+2]$ and show that, with probability at least $\eta\varepsilon$, an event occurs which implies that $\mathfrak{h}^{(n)} \in \mathsf{TP}^{\varepsilon}_{A,L}$.

To describe $\mathcal{F}$, we need some notation: for a function $f : I \to \mathbb{R}$ and an interval $[a,b] \subseteq I$, the *bridge* of f on $[a,b]$, denoted $f^{[a,b]} : [a,b] \to \mathbb{R}$, is given by

$$f^{[a,b]}(x) = f(x) - \frac{x-a}{b-a} \cdot f(b) - \frac{b-x}{b-a} \cdot f(a); \tag{11.21}$$

this is the function obtained by affinely transforming f so that its values at a and b vanish.

The σ-algebra $\mathcal{F}$ is defined to be the one generated by the following collection of random variables:

1. The maximizer of $\mathfrak{h}^{(n)}$: $x_0 = x_0^n = \arg\max_{x \in [-n^{1/60}, n^{1/60}]} \mathfrak{h}^{(n)}(x)$.
2. The curve data of $\mathcal{P}_n$: $\big\{\mathcal{P}_{n,j}(x) : j \in [\![2,n]\!],\ x \geq -\frac{1}{2}n^{1/3}$ or $j = 1,\ x \notin [x_0 + A, x_0 + A + 2]\big\}$.
3. The side bridge data of the top curve in $[x_0 + A, x_0 + A + 2]$: $\mathcal{P}_{n,1}^{[x_0+A, x_0+A+1]}$ and $\mathcal{P}_{n,1}^{[x_0+A+1, x_0+A+2]}$. (Here $\mathcal{P}_{n,1}^{[x_0+A, x_0+A+1]}$ is the function on $[x_0 + A, x_0 + A + 1]$ defined via (11.21) with $f = \mathcal{P}_{n,1}$, and similarly for $\mathcal{P}_{n,1}^{[x_0+A+1, x_0+A+2]}$.)

Conditionings on similar collections of data have been used in earlier works such as [Ham19a, CHH19], and in Part II of this thesis. There, however, the interval of focus—our $[x_0 + A, x_0 + A + 2]$—is either deterministic or a stopping domain (an analogue of a stopping time suited to the spatial nature of the Brownian Gibbs property used there). This means that the conditional distribution is more easily analysed using standard Markovian properties. Here, x_0 is a random variable which depends on the entirety of $\mathcal{P}_n$ and so is rather non-Markovian. This complicates the analysis considerably; a careful treatment will be provided in Section 11.7, for which we set up some notation and record some observations in the rest of this section.

Conditional on $\mathcal{F}$, the only randomness left in determining $\mathfrak{h}^{(n)}$ is the value of the random variable $Z := \mathcal{P}_{n,1}(x_0 + A + 1)$. Given a value of Z labelled z, and the data of $\mathcal{F}$, we can reconstruct

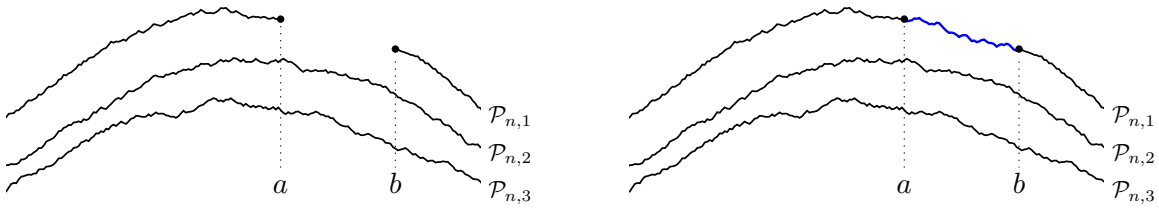


Figure 11.4: A depiction of the Brownian Gibbs resampling procedure. On the left in black is all the data contained in the σ-algebra $\mathcal{F}_{\text{ext}}$. The $\mathcal{F}_{\text{ext}}$-conditional distribution of $\mathcal{P}_{n,1}$ on $[a,b]$ is that of a Brownian bridge (in blue) of rate two from $(a,\mathcal{P}_{n,1}(a))$ to $(b,\mathcal{P}_{n,1}(b))$ conditioned on not intersecting $\mathcal{P}_{n,2}$ on $[a,b]$.

$\mathcal{P}_{n,1}(\cdot)$; when z is distributed according to the correct $\mathcal{F}$-conditional distribution of Z, this reconstruction may be thought of as the $\mathcal{F}$-conditional distribution of $\mathcal{P}_n$. We will denote the reconstruction by $\mathcal{P}^z_{n,1}(\cdot) : [-\frac{1}{2}n^{1/3},\infty) \to \mathbb{R}$. It is given by the formula

$$\mathcal{P}^z_{n,1}(x) = \begin{cases} \mathcal{P}_{n,1}(x) & [l]\text{for } x \in [-\frac{1}{2}n^{1/3},\infty) \\ & \qquad \setminus [x_0+A, x_0+A+2], \\ [l](x_0+A+1-x)\mathcal{P}_{n,1}(x_0+A) & \\ \quad + (x-(x_0+A))z + \mathcal{P}^{[x_0+A+0,x_0+A+1]}_{n,1}(x) & \text{for } x \in (x_0+A, x_0+A+1], \\ (x-(x_0+A+1))\mathcal{P}_{n,1}(x_0+A+2) & \\ \quad + (x_0+A+2-x)z + \mathcal{P}^{[x_0+A+1,x_0+A+2]}_{n,1}(x) & \text{for } x \in [x_0+A+1, x_0+A+2); \end{cases} \tag{11.22}$$

while for $j \geq 2$ and $x \in \mathbb{R}$, $\mathcal{P}^z_{n,j}(x) = \mathcal{P}_{n,j}(x)$. Note that $\mathcal{P}^z_n$ is $\mathcal{F}$-measurable.

11.4 The Brownian Gibbs property

Here, we recall the Brownian Gibbs property of $\mathcal{P}_n$, which was introduced and significantly leveraged in [CH14]. It will be used in the proof of an important upcoming statement, Lemma 11.17, on the $\mathcal{F}$-conditional distribution of Z.

For a fixed interval $[a,b] \subseteq (-\frac{1}{2}n^{1/3},\infty)$, define $\mathcal{F}_{\text{ext}}$ to be the σ-algebra generated by $\{\mathcal{P}_{n,1}(x) : x \in [-\frac{1}{2}n^{1/3},\infty)\setminus[a,b]\}$ and $\{\mathcal{P}_{n,j}(x) : j \in [\![2,n]\!], x \geq [-\frac{1}{2}n^{1/3},\infty)\}$, i.e., the data of everything *external* to $[a,b]$ on the top line. The *Brownian Gibbs* property asserts that the $\mathcal{F}_{\text{ext}}$-conditional distribution of $\mathcal{P}_{n,1}(\cdot)$ on $[a,b]$ is that of a Brownian bridge of rate two from $(a,\mathcal{P}_{n,1}(a))$ to $(b,\mathcal{P}_{n,1}(b))$ which is conditioned not to intersect the second curve $\mathcal{P}_{n,2}(\cdot)$. This can be interpreted as saying that $\mathcal{P}_{n,1}$ can be *resampled* on $[a,b]$ without changing its law by sampling a Brownian bridge with prescribed endpoints and conditioning it to avoid the second curve: see Figure 11.4.

(In fact, the full Brownian Gibbs property gives a similar description of the conditional distribution even under a larger class of conditionings. For example, for $k \in \mathbb{N}$, we may resample the top k curves on $[a, b]$ instead of merely the first; this amounts to a description of the conditional distribution on $[a, b]$ when the data of the top k curves on $[a, b]$ is not included in $\mathcal{F}_{\text{ext}}$. However, in this part of the thesis we will be only making use of the $k = 1$ case.)

Lemma 11.12. *The ensemble $\mathcal{P}_n$ has the Brownian Gibbs property.*

This statement is equivalent to the one that n-level Dyson Brownian motion has the Brownian Gibbs property. Though well-known and used in previous works [CH14, DV18], we were unable to locate a precise proof of this fact in the literature. However, it is fairly straightforward given the fact that an ensemble of Brownian bridges with strictly ordered endpoints conditioned on the (positive probability) event of non-intersection has the Brownian Gibbs property, and we will sketch the proof of Dyson Brownian motion having the Brownian Gibbs property given this fact. That non-intersecting Brownian bridges have the Brownian Gibbs property is very intuitive, but was formally proved only recently, in [DM20].

Proof of Lemma 11.12. As mentioned, this follows from $\mathcal{P}_n$ being an affine transformation of n-level Dyson Brownian motion (Lemma 11.6) and the latter ensemble having the Brownian Gibbs property. We briefly outline how to show that n-level Dyson Brownian motion $\text{DBM}_n : [\![n]\!] \times [0, \infty) \to \mathbb{R}$ has this property.

Let $[a, b] \subseteq (0, \infty)$. We first condition on the σ-algebra generated by $\{\text{DBM}_{n,j}(x) : j \in [\![1, n]\!], x \in [0, \infty) \setminus [a, b]\}$. The Markov property of DBM_n then implies that this conditional distribution depends on only the boundary values $\{\text{DBM}_{n,j}(x) : j \in [\![n]\!], x \in \{a, b\}\}$. Then the conditional distribution is that of n non-intersecting Brownian bridges with the given endpoints, as can be verified by comparing the transition probabilities of this ensemble with that of the conditioned Dyson Brownian motion (which makes use of the Karlin-McGregor formula [KM59] for non-intersecting strong Markov processes). The ensemble of non-intersecting Brownian bridges, quite naturally, has the Brownian Gibbs property [DM20, Lemma 2.13]. □

11.5 An outline of the argument in the narrow-wedge case

Before proving Proposition 11.11, we give an outline of the argument in the simpler narrow-wedge case, under which $\mathfrak{h}_0$ is zero at the origin and $-\infty$ elsewhere.

First observe from (11.15) that, for this initial condition, $\mathfrak{h}^{(n)}(x) = \mathcal{S}_n(0, x) = \mathcal{P}_{n,1}(x)$ for $|x| \leq n^{1/60}$. In particular, $\mathfrak{h}^{(n)}$ is a function of only the top line of $\mathcal{P}_n$, and the same is true for $x_0 = x_0^n$ defined earlier. The collection of curves $\mathcal{P}$ can be expressed in terms of non-intersecting Brownian motions via Lemma 11.6. We will show that, for some η (depending on A), it is with probability at least $\eta\varepsilon$ that the curve $\mathcal{P}_{n,1}$ comes within ε of its maximum $\mathfrak{h}^{(n)}(x_0)$ in the window $[x_0 + A, x_0 + A + 2]$. Our first step, below, will be to identify the $\mathcal{F}$-conditional distribution of $Z = \mathcal{P}_{n,1}(x_0 + A + 1)$. (The event $\mathfrak{h}^{(n)} \in \mathsf{TP}^{\varepsilon}_{A,L}$ also imposes conditions on the location of the

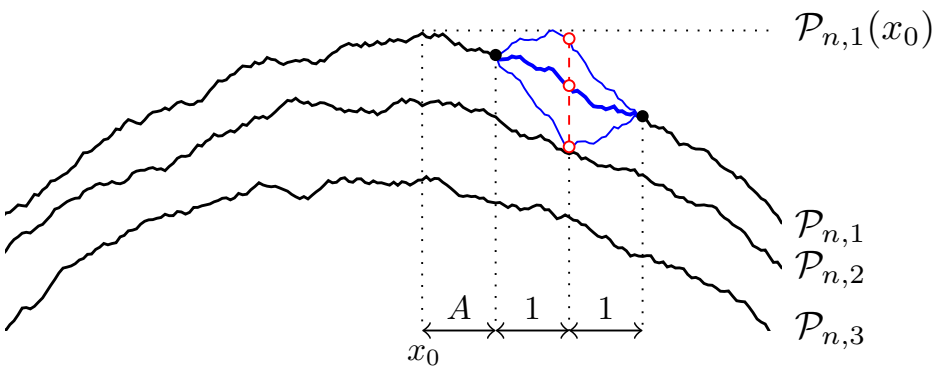


Figure 11.5: A depiction of the resampling in the proof outline for the narrow-wedge case. The sigma-algebra $\mathcal{F}$ contains information about all of the thick curves including the thick blue curve, except that it forgets the location of the red circle whose value is denoted by $Z = \mathcal{P}_{n,1}(x_0 + A + 1)$. The reconstruction $\mathcal{P}^z_{n,1}$ is given by linearly shifting the left and right blue bridges to meet the value z of the red circle as it varies. The thin blue lines here represent two possible reconstructions. The random variable Z is restricted by the fact that the reconstruction must not exceed the global maximum $\mathcal{P}_{n,1}(x_0)$ (here denoted by a horizontal dotted black line) and must not intersect the second curve $\mathcal{P}_{n,2}$. The vertical dashed red line represents the possible range $[\text{Corner}^{\downarrow}, \text{Corner}^{\uparrow}]$ for Z (with the upper and lower red circles corresponding to these bounds).

maximizer and the value of the maximum, but these are more easily handled and not discussed here.)

Step 1: The $\mathcal{F}$-conditional distribution of Z. As mentioned, x_0 is defined in terms of the whole curve $\mathcal{P}_{n,1}(\cdot)$ and so is non-Markovian; in particular, it is not a stopping time. But it is intuitively plausible, based on the definition of x_0 and the Brownian Gibbs property of $\mathcal{P}_n$, that the distribution of $\mathcal{P}_{n,1}(\cdot)$ on $[x_0, n^{1/60}]$ conditional on $\mathcal{P}_{n,1}(\cdot)$ outside of the interval $(x_0, n^{1/60})$ and the lower curves $\mathcal{P}_{n,2}, \mathcal{P}_{n,3}, \ldots$ is of a Brownian bridge (of rate two) between the appropriate endpoints conditioned on (i) not intersecting the lower curve $\mathcal{P}_{n,2}$ and (ii) not exceeding $\mathcal{P}_{n,1}(x_0)$. This intuition is correct and is carefully stated in Lemma 11.15.

This description of $\mathcal{P}_{n,1}$ on $[x_0, n^{1/60}]$ makes it easy to derive the distribution of Z conditional on $\mathcal{F}$. Indeed, when we also condition on the data of $\mathcal{P}_{n,1}$ on $[x_0, x_0 + A]$ and $[x_0 + A + 2, n^{1/60}]$, we see that $\mathcal{P}_{n,1}$ on $[x_0 + A, x_0 + A + 2]$ has the law of Brownian bridge of rate two with endpoints $\mathcal{P}_{n,1}(x_0 + A)$ and $\mathcal{P}_{n,1}(x_0 + A + 2)$ which is conditioned to again (i) not intersect the lower curve and (ii) not exceed $\mathcal{P}_{n,1}(x_0)$. To get from this collection of conditioning data to $\mathcal{F}$, we only have to include the side bridge data $\mathcal{P}^{[x_0+A, x_0+A+1]}_{n,1}$ and $\mathcal{P}^{[x_0+A+1, x_0+A+2]}_{n,1}$; classical decompositions of Brownian bridge then say that the $\mathcal{F}$-conditional distribution of Z is that of a normal random variable of appropriate $\mathcal{F}$-measurable mean and unit variance, conditioned on the reconstruction $\mathcal{P}^Z_{n,1}(\cdot)$ again satisfying (i) and (ii).

We can simplify this description of Z. Essentially, condition (i) places a lower bound on how large

Z can be, while (ii) places an upper bound: see Figure 11.5. To make this rigorous, we observe that the reconstruction $\mathcal{P}^z_{n,1}(x)$ is monotone in z for every x from its formula (11.22). These lower and upper bounds are $\mathcal{F}$-measurable random variables; they are *corners* of the acceptable range of values Z can take, and we label them respectively Corner$^{\downarrow}$ and Corner$^{\uparrow}$. Thus, Z is a normal random variable with explicit $\mathcal{F}$-measurable mean and unit variance, conditioned on lying in the interval [Corner$^{\downarrow}$, Corner$^{\uparrow}$].

Step 2: Finding the sweet spot for Z. It is easy to see that Corner$^{\uparrow}$ is such that, when $z =$ Corner$^{\uparrow}$, $\sup_{x\in[x_0+A,x_0+A+2]} \mathcal{P}^z_{n,1}(x) = \mathcal{P}_{n,1}(x_0)$. With this in hand, the linear—and so certainly Lipschitz—relationship of $\mathcal{P}^z_{n,1}(x)$ with z for every fixed x tells us that reducing z by ε from Corner$^{\uparrow}$ reduces the value of $\sup_{x\in[x_0+A,x_0+A+2]} \mathcal{P}^z_{n,1}(x)$ by an amount of order ε. Thus, to cause $\mathsf{TP}^{\varepsilon}_{A,L}$ to occur, it is enough to have Z fall within a sweet spot interval I_ε of length of order ε with upper endpoint Corner$^{\uparrow}$.

Step 3: The probability of hitting the sweet spot. It remains to bound below the probability that Z falls in I_ε. To do so, we need control over two things: the $\mathcal{F}$-measurable mean of Z, and the value of Corner$^{\uparrow}$. (We can ignore Corner$^{\downarrow}$, i.e., take it to be $-\infty$, as we are only aided in proving a probability lower bound if its value is larger.) For this purpose, we consider a selection of favourable $\mathcal{F}$-measurable data Fav which is defined by demanding bounds on these quantities, as well as on the location of the maximizer and the value of the maximum: see Section 11.8 ahead. We then show that Fav occurs with a probability that is uniformly positive in n. Given control over the mean and Corner$^{\uparrow}$ on a positive probability event, the form of the normal density guarantees that the probability of Z falling in the order ε length interval is at least some constant multiple of ε. This completes the proof outline in the narrow-wedge case.

Complications with general initial data. The narrow-wedge case provided a number of simplifications, the primary being the equality of $\mathfrak{h}^{(n)}$ and the top line of $\mathcal{P}_n$. This had two effects, both in Step 1: we could define x_0 in terms of only $\mathcal{P}_{n,1}(\cdot)$, i.e., without the lower curves (making it simpler to consider that process' distribution on $[x_0, n^{1/60}]$); and we could infer the existence of a valid interval [Corner$^{\downarrow}$, Corner$^{\uparrow}$] for the $\mathcal{F}$-conditional distribution of Z via monotonicity properties of only $\mathcal{P}_n$. Both these aspects will need modification in the general case.

Because we can perform Brownian resamplings only with $\mathcal{P}_n$, we need the representation of $\mathfrak{h}^{(n)}$ in terms of $\mathcal{P}_n$ recorded in the last equality of (11.15), which relies on the identification of LPP values in the original and melon environments cited in Proposition 11.4. Note that $\mathfrak{h}^{(n)}$, and so also x_0, is now defined in terms of all the curves of $\mathcal{P}_n$, not just the first. More specifically, while in the narrow wedge case we could work with the function values of $\mathcal{P}_{n,1}(\cdot)$, in the general case we have to analyse *last passage values through the environment given by $\mathcal{P}_n$*. This is the underlying complication that causes all the others in the general case.

To achieve a description of Z in terms of [Corner$^{\downarrow}$, Corner$^{\uparrow}$] in the general case, we first need a formula for $\mathfrak{h}^{(n),z}$, the reconstruction of $\mathfrak{h}^{(n)}$ when $Z = z$, in terms of $\mathcal{P}^z_n$. This will be recorded shortly in (11.23). Then we need a monotonicity statement about $\mathfrak{h}^{(n),z}(x)$ for fixed x that will allow us to express the condition that z is such that $\sup_{x\in[x_0+A,x_0+A+2]} \mathfrak{h}^{(n),z}(x) \le \mathfrak{h}^{(n)}(x_0)$ as an upper bound on z, just as we did with the monotonicity statement for $\mathcal{P}^z_{n,1}$ above in simplifying

the condition (ii). Such a monotonicity statement is actually not true for $\mathfrak{h}^{(n),z}$, and we circumvent this difficulty by deriving one instead for the weights of individual paths (as opposed to their supremum $\mathfrak{h}^{(n),z}$) in the reconstructed environment. This is Lemma 11.13 recorded ahead. Finally, we need to know that $\mathfrak{h}^{(n),z}(x)$ is Lipschitz in z for each fixed x: see Lemma 11.14 whose argument also proceeds in a pathwise manner.

With these modifications, the proof in the general case proceeds largely along the lines of the narrow-wedge case outlined here. We move to giving the details next, starting with the facts needed to handle the general case's main complications, namely the monotonicity and Lipschitz properties of $\mathfrak{h}^{(n),z}$.

11.6 The reconstructed $\mathfrak{h}^{(n)}$ and its properties

Using (11.15), we provide a formula in terms of $\mathcal{P}_n^z$ for the reconstructed $\mathfrak{h}^{(n)}$, denoted $\mathfrak{h}^{(n),z} : [-n^{1/60}, n^{1/60}] \to \mathbb{R} \cup \{-\infty\}$. It is:

$$\mathfrak{h}^{(n),z}(x) = \sup_{0 \le y \le n^{1/60}} \Big(\mathfrak{h}_0(y) + \mathcal{P}_n^z[(-\tfrac{1}{2}n^{1/3} + y, n) \to (x, 1)] - n^{2/3} \Big). \tag{11.23}$$

Since $\mathcal{P}_n^z$ is $\mathcal{F}$-measurable for all $z \in \mathbb{R}$, so is $\mathfrak{h}^{(n),z}$. As we noted in the preceding section, it is also immediate from the formula of $\mathcal{P}_{n,1}^z(x)$ that, for any $x \in \mathbb{R}$, $\mathcal{P}_{n,1}^z(x)$ is non-decreasing as a function of z. The function $\mathfrak{h}^{(n),z}(x)$ further enjoys a monotonicity property in z that is slightly more complicated, which we record in the next lemma.

Recall that, for an upright path γ, $\mathcal{P}_n^z[\gamma]$ is the weight of γ in the environment $\mathcal{P}_n^z$. Imitating (11.15), let

$$\mathfrak{h}^{(n),z}(\gamma) = \mathfrak{h}_0(y) + \mathcal{P}^z[\gamma] - n^{2/3},$$

where y is the starting coordinate of γ on line n.

Lemma 11.13 (Monotonicity of $\mathfrak{h}^{(n),z}$ in z)**.** *For each upright path γ starting on line n and ending on line 1, the process $z \mapsto \mathfrak{h}^{(n),z}(\gamma)$ is non-increasing almost surely; or constant almost surely; or non-decreasing almost surely. Moreover, it is an almost sure event that:*

1. *if $-n^{1/60} \le x \le x_0 + A$, then $\mathfrak{h}^{(n),z}(x) = \mathfrak{h}^{(n)}(x)$ for all $z \in \mathbb{R}$; and*
2. *if $x \ge x_0 + A + 2$, then $\mathfrak{h}^{(n),z}(x)$ is non-increasing in z.*

Proof. Let u be the coordinate at which γ jumps to the top line (i.e., line 1); and let x be its ending coordinate. Let γ^{u-} be γ restricted to its path before coordinate u, i.e., γ's restriction to the lower $n-1$ lines, indexed by $[\![2, n]\!]$. Then

$$\begin{aligned} \mathfrak{h}^{(n),z}(\gamma) &= \mathfrak{h}^{(n),z}(\gamma^{u-}) + \mathcal{P}_{n,1}^z(x) - \mathcal{P}_{n,1}^z(u) \\ &= \mathfrak{h}^{(n)}(\gamma^{u-}) + \mathcal{P}_{n,1}^z(x) - \mathcal{P}_{n,1}^z(u); \end{aligned} \tag{11.24}$$

the latter equality because the environment of the lower $n-1$ lines of $\mathcal{P}_n^z$ does not depend on z. The claim that $\mathfrak{h}^{(n),z}(\gamma)$ is monotone and the nature of its monotonicity now follow readily by examining the increment $\mathcal{P}_{n,1}^z(x) - \mathcal{P}_{n,1}^z(u)$ from the definitions in (11.22).

Now we move to proving the two numbered claims. For (1), consider the set of paths which end at x. We *claim* that, for such paths γ, $\mathfrak{h}^{(n),z}(\gamma)$ is constant in z; which implies (1). The claim follows by noting that, if $x \le x_0 + A$, then $u \le x_0 + A$; and so $\mathcal{P}_{n,1}^z(x) - \mathcal{P}_{n,1}^z(u)$ does not depend on z from (11.22). This completes the claim by the decomposition (11.24).

A similar argument holds for (2). We *claim* that, for any path γ which ends at x, $\mathfrak{h}^{(n),z}(\gamma)$ is non-increasing in z. To see this, we use the same decomposition as (11.24), and observe that it is enough to prove that $\mathcal{P}_{n,1}^z(x) - \mathcal{P}_{n,1}^z(u) = \mathcal{P}_{n,1}(x) - \mathcal{P}_{n,1}^z(u)$ is non-increasing in z. There are two cases: $u \in [x_0 + A, x_0 + A + 2]$ and $u \notin [x_0 + A, x_0 + A + 2]$. In the first, $\mathcal{P}_{n,1}^z(u)$ is non-decreasing in z; while, in the second, it is constant, as we see from (11.22). This completes the proof of the claim and thus also of Lemma 11.13. □

While a similar monotonicity statement in z as Lemma 11.13(1) and (2) holds for $\mathfrak{h}^{(n),z}(x)$ when $x \in [x_0+A, x_0+A+1]$, there is no counterpart when $x \in [x_0+A+1, x_0+A+2]$. In the notation of the preceding proof, this is because, in the latter case, the type of monotonicity for the weight of a given path ending at x depends on the value of u, the location at which the path jumps to the top line from the second line: for a certain range of u depending on x, the weight of any path with jump point u will be increasing in z; while, for larger u, it will be decreasing. Since $\mathfrak{h}^{(n),z}(x)$ maximizes over all such paths, no monotonicity holds for this quantity. In contrast, notice from the proof of Lemma 11.13 that, for $x \in [-n^{1/60}, x_0 + A]$ or $x \ge x_0 + A + 2$, the weight of any path ending at x has a single form of monotonicity for all possible u. It is in order to handle this absence of z-monotonicity for $\mathfrak{h}^{(n),z}(x)$ when $x \in [x_0 + A + 1, x_0 + A + 2]$ that we proved the first statement of Lemma 11.13, concerning the monotonicity of the weight of *single* paths.

Lemma 11.14 ($\sup \mathfrak{h}^{(n),z}$ is Lipschitz in z)**.** *It holds almost surely that, for any* $z_1, z_2 \in \mathbb{R}$,

$$\left| \sup_{x \in [x_0+A, x_0+A+2]} \mathfrak{h}^{(n),z_1}(x) - \sup_{x \in [x_0+A, x_0+A+2]} \mathfrak{h}^{(n),z_2}(x) \right| \le 2|z_1 - z_2|.$$

Proof. For convenience of notation, let us define

$$h(z) = \sup_{x \in [x_0+A, x_0+A+2]} \mathfrak{h}^{(n),z}(x).$$

The arguments that we will present hold on the probability one event Ω that, for each $z \in \mathbb{R}$, there exist $x \in [x_0 + A, x_0 + A + 2]$, $y \in [0, n^{1/60}]$ and an upright path $\Gamma^{z,x}$ (for which we use the capital Greek letter to emphasise the path's randomness) ending at x, such that

$$h(z) = \mathfrak{h}_0(y) + \mathcal{P}_n^z[\Gamma^{z,x}] - n^{2/3};$$

that the supremum in the definition of h is achieved uses the compactness of $[x_0 + A, x_0 + A + 2]$.

By symmetry, it is enough to prove that $h(z_1) - h(z_2) \le 2|z_1 - z_2|$. On the event Ω, we see that, for some $y \in [0, n^{1/60}]$, $x \in [x_0 + A, x_0 + A + 2]$, and upright path $\Gamma^{z_1,x}$,

$$h(z_1) = \mathfrak{h}_0(y) + \mathcal{P}_n^{z_1}[\Gamma^{z_1,x}] - n^{2/3},$$
$$h(z_2) \ge \mathfrak{h}_0(y) + \mathcal{P}_n^{z_2}[\Gamma^{z_1,x}] - n^{2/3},$$

and so

$$h(z_1) - h(z_2) \le \mathcal{P}_n^{z_1}[\Gamma^{z_1,x}] - \mathcal{P}_n^{z_2}[\Gamma^{z_1,x}].$$

Let u be the coordinate where $\Gamma^{z_1,x}$ jumps to the top line. Since the environments defined by $\mathcal{P}_n^{z_1}$ and $\mathcal{P}_n^{z_2}$ differ only in the top line, we see that

$$\begin{aligned}\mathcal{P}_n^{z_1}[\Gamma^{z_1,x}] - \mathcal{P}_n^{z_2}[\Gamma^{z_1,x}] &= \big[\mathcal{P}_{n,1}^{z_1}(x) - \mathcal{P}_{n,1}^{z_1}(u)\big] - \big[\mathcal{P}_{n,1}^{z_2}(x) - \mathcal{P}_{n,1}^{z_2}(u)\big] \\ &= \big[\mathcal{P}_{n,1}^{z_1}(x) - \mathcal{P}_{n,1}^{z_2}(x)\big] - \big[\mathcal{P}_{n,1}^{z_1}(u) - \mathcal{P}_{n,1}^{z_2}(u)\big] \\ &\le 2|z_1 - z_2|.\end{aligned}$$

The inequality follows from the definition of $\mathcal{P}_{n,1}^{z}$ in (11.22). The equalities and the bound hold deterministically on Ω. This completes the proof of Lemma 11.14. □

11.7 The $\mathcal{F}$-conditional distribution of Z

We next move towards a description of the $\mathcal{F}$-conditional distribution of Z. First we define the canonical filtration for the top curve, $\mathcal{F}_t^{\text{past}} = \sigma\left(\mathcal{P}_{n,1}(s) : s \in [-\frac{1}{2}n^{1/3}, t]\right)$. We also define a filtration that captures the future of the process by $\mathcal{F}_t^{\text{future}} = \sigma\left(\mathcal{P}_{n,1}(s) : s \ge t\right)$. Certain additional σ-algebras are needed. Let $\mathcal{F}_{\text{lower}}$ be the σ-algebra generated by the lower $n-1$ curves, i.e., $\mathcal{F}_{\text{lower}} = \sigma(\mathcal{P}_{n,j}(x) : x \ge -\frac{1}{2}n^{1/3}, j \in [\![2, n]\!])$. Let $\mathcal{F}_{x_0}^{\text{past}}$ be the σ-algebra generated by all sets of the form

$$F_s \cap \{x_0 > s\}, \tag{11.25}$$

where F_s ranges over all elements of $\mathcal{F}_s^{\text{past}}$ and s ranges over $[-\frac{1}{2}n^{1/3}, \infty)$. This σ-algebra encodes the information known by time x_0. If x_0 were a stopping time, $\mathcal{F}_{x_0}^{\text{past}}$ would coincide with the usual σ-algebra associated with such times. Let $\mathcal{F}_{x_0+A}^{\text{past}}$ be defined similarly to $\mathcal{F}_{x_0}^{\text{past}}$ in (11.25) with $x_0 + A$ replacing x_0. Let $\mathcal{F}_{x_0+A+2}^{\text{future}}$ be defined as the σ-algebra generated by all sets of the form $F_s \cap \{x_0 + A + 2 < s\}$, where F_s ranges over all elements of $\mathcal{F}_s^{\text{future}}$ and s ranges over $[-\frac{1}{2}n^{1/3}, \infty)$.

Finally, let $\mathcal{F}'$ be generated by $\mathcal{F}_{\text{lower}} \cup \mathcal{F}_{x_0}^{\text{past}} \cup \mathcal{F}_{n^{1/60}}^{\text{future}}$, and let $\mathcal{F}''$ be generated by $\mathcal{F}' \cup \mathcal{F}_{x_0+A}^{\text{past}} \cup \mathcal{F}_{x_0+A+2}^{\text{future}}$ (which should be thought of as equalling $\mathcal{F}_{\text{lower}} \cup \mathcal{F}_{x_0+A}^{\text{past}} \cup \mathcal{F}_{x_0+A+2}^{\text{future}}$, since typically $x_0 + A + 2$ will be less than $n^{1/60}$). Observe that $\mathcal{F}$ is the σ-algebra generated by $\mathcal{F}''$ along with the side bridge data on $[x_0 + A, x_0 + A + 1]$ and $[x_0 + A + 1, x_0 + A + 2]$; i.e., $\mathcal{P}_{n,1}^{[x_0+A,x_0+A+1]}$ and $\mathcal{P}_{n,1}^{[x_0+A+1,x_0+A+2]}$. See (11.21) to recall the notation.

We start by describing the $\mathcal{F}'$-conditional distribution of $\mathcal{P}_{n,1}(\cdot)$, which will then be used to give the $\mathcal{F}$-conditional distribution of $Z = \mathcal{P}_{n,1}(x_0 + A + 1)$ in Lemma 11.17. To give the first statement, and with a slight abuse of notation that, we hope, will not cause confusion with the earlier defined $\mathcal{P}_n^z$, define $\mathcal{P}_{n,1}^B(\cdot) : [-\frac{1}{2}n^{1/3}, \infty)$ by

$$\mathcal{P}_{n,1}^B(x) = \begin{cases} \mathcal{P}_{n,1}(x) & \text{for } -\frac{1}{2}n^{1/3} \le x \le x_0 \text{ or } x \ge n^{1/60} \\ B(x) & \text{for } x_0 \le x \le n^{1/60}, \end{cases}$$

where $B : [x_0, n^{1/60}] \to \mathbb{R}$ is a given function with $B(x_0) = \mathcal{P}_{n,1}(x_0)$ and $B(n^{1/60}) = \mathcal{P}_{n,1}(n^{1/60})$; also, let $\mathcal{P}_{n,j}^B(x) = \mathcal{P}_{n,j}(x)$ for $j \in [\![2, n]\!]$ and x in the domain. Define $\mathfrak{h}^{(n),B}$ (with a similar notational abuse) to be the reconstructed $\mathfrak{h}^{(n)}$, given as in (11.23) by

$$\mathfrak{h}^{(n),B}(x) = \sup_{0 \le y \le n^{1/60}} \Big(\mathfrak{h}_0(y) + \mathcal{P}_n^B[(-\tfrac{1}{2}n^{1/3} + y, n) \to (x, 1)] - n^{2/3} \Big).$$

Lemma 11.15. *Conditionally on $\mathcal{F}'$, $\{\mathcal{P}_{n,1}(x) : x \ge -\frac{1}{2}n^{1/3}\}$ has the law of $\mathcal{P}_{n,1}^B$, where $B : [x_0, n^{1/60}] \to \mathbb{R}$ is Brownian bridge of rate two from $(x_0, \mathcal{P}_{n,1}(x_0))$ to $(n^{1/60}, \mathcal{P}_{n,1}(n^{1/60}))$ conditioned on non-intersection of the second curve $\mathcal{P}_{n,2}(\cdot)$ and on $\sup_{x_0 \le x \le n^{1/60}} \mathfrak{h}^{(n),B}(x) \le \mathfrak{h}^{(n)}(x_0)$.*

The proof will mainly rely on [Mil78]. This paper identifies the distribution of a homogeneous strong Markov process $X : [0, \infty) \to E$ on the unbounded interval whose left endpoint is the maximizer of $\Phi(X(t))$, for a given continuous function $\Phi : E \to [-\infty, \infty]$. Here, E is the state space of the Markov process, a set that is supposed compact in [Mil78].

Proof of Lemma 11.15. First we recall that $x_0 \ge -n^{1/60}$ by definition and so $\mathcal{F}^{\text{past}}_{-n^{1/60}} \subseteq \mathcal{F}^{\text{past}}_{x_0}$. Conditional on $\mathcal{F}_{\text{lower}}$, $\mathcal{F}^{\text{past}}_{-n^{1/60}}$ and $\mathcal{F}^{\text{future}}_{n^{1/60}}$, the distribution of $\{\mathcal{P}_{n,1}(x) : -n^{1/60} \le x \le n^{1/60}\}$ is that of a Brownian bridge of rate two on $[-n^{1/60}, n^{1/60}]$ with starting value $\mathcal{P}_{n,1}(-n^{1/60})$ and ending value $\mathcal{P}_{n,1}(n^{1/60})$ conditioned on non-intersection with the second curve $\mathcal{P}_{n,2}(\cdot)$. This is the statement of the Brownian Gibbs property of $\mathcal{P}_n$, Lemma 11.12. In particular, the conditioned process is Markov (and non-homogeneous), and, since Brownian bridge is a strong Markov process and the conditioning event is almost surely of positive probability, the same is true of the conditioned space-time process. (Here we consider the space-time process so as to have a time-homogeneous Markov process: see [RY13, Chapter III, exercise 1.10].)

Consider the process $X : [-n^{1/60}, \infty) \to [-\infty, \infty]^2 \times [-n^{1/60}, \infty]$ defined by $X(t) := (\mathcal{P}_{n,1}(t \wedge n^{1/60}), \mathfrak{h}^{(n)}(t \wedge n^{1/60}), t)$; here $[-\infty, \infty]$ and $[-n^{1/60}, \infty]$ are compactifications of $\mathbb{R}$ and $[-n^{1/60}, \infty)$, and are employed so that the state space of X is compact. We consider X to start at time $-n^{1/60}$, and to be killed at time $n^{1/60}$, so that the maximizer of the second component of X is $x_0^n = \arg\max_{|x| \le n^{1/60}} \mathfrak{h}^{(n)}(x)$. (To be precise, as earlier we will be working with the final maximizer, i.e., the largest one, on the event that there are several. To see that there is such a final maximizer for the process $\mathfrak{h}^{(n)}(t \wedge n^{1/60})$, note that there must be a final one on the interval $[-n^{1/60}, n^{1/60}]$ by continuity of $\mathfrak{h}^{(n)}$, and that, by Lemma 11.16 ahead, the final one is almost surely not $n^{1/60}$.)

We claim that, conditionally on $\mathcal{F}_{\text{lower}}$, $\mathcal{F}^{\text{past}}_{-n^{1/60}}$ and $\mathcal{F}^{\text{future}}_{n^{1/60}}$, X is a homogeneous strong Markov process. To see this, first define the process X' by $X'(t) = \big(\mathcal{P}_{n,1}(t \wedge n^{1/60}), \mathfrak{h}^{(n)}(t \wedge n^{1/60})\big)$. It is enough to prove that X' is a *non-homogeneous* Markov process, as then, by the same trick as used a few lines above, the space-time process $X(t) = (X'(t), t)$ is necessarily a homogeneous Markov process.

To show that X' is Markov under this conditioning, we state a formula that expresses $\mathfrak{h}^{(n)}(t+s)$ in terms of $\mathfrak{h}^{(n)}(t)$ and data contained in $\mathcal{F}_{\text{lower}}$, $\mathcal{F}^{\text{past}}_{-n^{1/60}}$ and $\mathcal{F}^{\text{future}}_{n^{1/60}}$:

$$\mathfrak{h}^{(n)}(t+s) = \max\Big\{\mathfrak{h}^{(n)}(t) + \mathcal{P}_{n,1}(t+s) - \mathcal{P}_{n,1}(t), \\ \sup_{\substack{y \\ u>t}} \Big(\mathfrak{h}_0(y) + \mathcal{P}_n[(-\tfrac{1}{2}n^{1/3} + y, n) \to (u, 2)] + \mathcal{P}_{n,1}(t+s) - \mathcal{P}_{n,1}(u) - n^{2/3}\Big)\Big\}; \tag{11.26}$$

the supremum over y and u is taken over choices such that $0 \le y \le n^{1/60}$ and $y \le -\frac{1}{2}n^{1/3} + u$.

This formula follows by considering the location u that the geodesic for endpoint $t+s$ jumps to the top line. It is the first term that attains the maximum when $u \le t$; and it is the second that does so when $u > t$. In the first case, $\mathfrak{h}^{(n)}(t+s)$ is equal to $\mathfrak{h}^{(n)}(t)$ plus the remaining increment on the top line as the $(t+s)$-geodesic must pass through $(t, 1)$; this is because $\mathfrak{h}^{(n)}(t)$ is the value attained by a similar maximization problem. In the second case, we have rewritten the formula (11.15) for $\mathfrak{h}^{(n)}$ by decomposing the last passage value at $(u, 2)$, which lies on the geodesic by definition.

Observe that, since we have conditioned on the lower $n-1$ curves, $\mathcal{P}_n[(-\frac{1}{2}n^{1/3} + y, n) \to (u, 2)]$ is a deterministic function of y and u. Thus, conditional on $(\mathcal{P}_{n,1}(t), \mathfrak{h}^{(n)}(t))$, $\mathcal{F}_{\text{lower}}$, $\mathcal{F}^{\text{past}}_{-n^{1/60}}$ and $\mathcal{F}^{\text{future}}_{n^{1/60}}$, it holds that $\mathfrak{h}^{(n)}\left((t+s) \wedge n^{1/60}\right)$ is measurable with respect to $\{\mathcal{P}_{n,1}(x) : t \le x \le n^{1/60}\}$, which is conditionally independent of $\mathcal{F}^{\text{past}}_t$ given $\mathcal{P}_{n,1}(t)$ by the Markov property of $\mathcal{P}_{n,1}(\cdot)$. This proves that X' is Markov, and so X is a homogeneous Markov process. (We also used that the canonical filtration of $\mathfrak{h}^{(n)}$ is contained in the filtration generated by $\mathcal{F}^{\text{past}}_t$ and $\mathcal{F}_{\text{lower}}$.) This argument reduced the Markov property of X' to that of $\mathcal{P}_{n,1}$; the reduction also works to show that X' is strong Markov since $\mathcal{P}_{n,1}$ is strong Markov.

Since the state space of X is compact, we may apply the results of [Mil78]. Consider the projection map $\Phi : [-\infty, \infty]^2 \times [-n^{1/60}, \infty] \to [-\infty, \infty]$ given by $(x, y, z) \mapsto y$. Then $x_0 = x_0^n = \arg\max_{t \ge -n^{1/60}} \Phi(X(t))$, the largest one if the argmax is not unique. The main theorem of [Mil78] implies that, conditionally on $\mathcal{F}_{\text{lower}}$, $\mathcal{F}^{\text{past}}_{-n^{1/60}}$ and $\mathcal{F}^{\text{future}}_{n^{1/60}}$, the process $\{X(x_0 + t) : t > 0\}$ is conditionally independent of $\mathcal{F}^{\text{past}}_{x_0}$ given the data $(\mathcal{P}_{n,1}(x_0), \mathfrak{h}^{(n)}(x_0), x_0)$. Further, this process is Markov and has the law of X conditioned on the event that $\Phi(X(t)) \le \Phi(X(x_0))$ for all t.

By projecting to the first coordinate of X, we see that this statement is equivalent to $\{\mathcal{P}_{n,1}(x) : x \ge -\frac{1}{2}n^{1/3}\}$ having the distribution of $\mathcal{P}_n^B$, where $B : [x_0, n^{1/60}] \to \mathbb{R}$ is a Brownian bridge of rate two from $(x_0, \mathcal{P}_{n,1}(x_0))$ to $(n^{1/60}, \mathcal{P}_{n,1}(n^{1/60}))$ conditioned on (i) $B(\cdot) > \mathcal{P}_{n,2}(\cdot)$ on $[x_0, n^{1/60}]$ and (ii) $\sup_{x_0 \le x \le n^{1/60}} \mathfrak{h}^{(n),B}(x) \le \mathfrak{h}^{(n)}(x_0)$. We get an equivalent condition on projecting because

the second component of X is determined by the first along with the lower curve data. This completes the proof of Lemma 11.15. □

The following lemma was used in the proof of Lemma 11.15 and establishes that the distribution of the maximizer of $\mathfrak{h}^{(n)}(\cdot)$ has no atoms.

Lemma 11.16. *Fix z with $-n^{1/60} < z \leq n^{1/60}$. Almost surely, $\mathfrak{h}^{(n)}(z) \neq \sup_{|x| \leq n^{1/60}} \mathfrak{h}^{(n)}(x)$.*

Proof. We use the formula (11.15) for $\mathfrak{h}^{(n)}(z)$ in terms of a last passage problem through $\mathcal{P}_n$. It is a consequence of Lemma 11.12 that $\mathcal{P}_{n,j}(\cdot) - \mathcal{P}_{n,j}(z-1)$ is absolutely continuous with respect to Brownian motion of rate two on $[z-1, z+1]$ for each $j \in [\![n]\!]$ (see for example [CH14, Proposition 4.1]). Thus we have with probability one that z is not a local maximizer of $\mathcal{P}_{n,j}$ for any $j \in [\![n]\!]$, which is the event we work on now. Let Γ be the geodesic associated to $\mathfrak{h}^{(n)}(z)$ implicit in (11.15) and let $J \in [\![n]\!]$ be the index of the line that Γ visits at time z^- (if $J > 1$, this means that the geodesic jumps to line 1 at location z, and so the top line's values do not contribute to the last passage value). Let $\tilde{z}$ belong to the interval of time that Γ spends on line J and be such that $\mathcal{P}_{n,J}(\tilde{z}) > \mathcal{P}_{n,J}(z)$. Now we can consider a modification of Γ that has endpoint $(\tilde{z}, 1)$, which it jumps to from $(\tilde{z}, J)$; this implies that z is not a maximizer of $\mathfrak{h}^{(n)}(\cdot)$ since $\mathfrak{h}^{(n)}(\tilde{z}) > \mathfrak{h}^{(n)}(z)$ and $|\tilde{z}| \leq n^{1/60}$. □

With these preliminaries, we now state what the $\mathcal{F}$-conditional distribution of Z is; recall that $Z = \mathcal{P}_{n,1}(x_0 + A + 1)$.

Lemma 11.17. *There exist $\mathcal{F}$-measurable random variables* $\text{Corner}^{\downarrow}$ *and* $\text{Corner}^{\uparrow}$ *such that the following holds. Conditionally on $\mathcal{F}$, and on the $\mathcal{F}$-measurable event that $x_0 + A + 2 \leq n^{1/60}$, the distribution of Z is a normal random variable with mean $\frac{1}{2}(\mathcal{P}_{n,1}(x_0 + A) + \mathcal{P}_{n,1}(x_0 + A + 2))$ and variance one, conditional on lying inside* $[\text{Corner}^{\downarrow}, \text{Corner}^{\uparrow}]$. *Further, when $z = \text{Corner}^{\uparrow}$,*

$$\sup_{x \in [x_0 + A, x_0 + A + 2]} \mathfrak{h}^{(n),z}(x) = \mathfrak{h}^{(n)}(x_0). \tag{11.27}$$

In the discussion of the narrow-wedge case, the $\text{Corner}^{\uparrow}$ and $\text{Corner}^{\downarrow}$ random variables had a clear interpretation respectively as the largest value of Z such that the reconstruction $\mathcal{P}^Z_{n,1}$ at no point exceeds $\mathcal{P}_{n,1}(x_0)$ and the smallest value of Z such that the reconstruction at no point intersects $\mathcal{P}_{n,2}$; each variable handled one condition. For general initial conditions, these random variables play analogous but slightly different roles. In particular, they are respectively the largest and smallest values of Z such that $\mathfrak{h}^{(n),Z}$ at no point exceeds $\mathfrak{h}^{(n)}(x_0)$ and $\mathcal{P}^Z_{n,1}$ does not intersect $\mathcal{P}_{n,2}$. However, it may not be the case that each variable separately handles one of the conditions: because of a larger class of possible geodesic paths, $\text{Corner}^{\downarrow}$ may also play a role in preventing $\mathfrak{h}^{(n),Z}$ from exceeding $\mathfrak{h}^{(n)}(x_0)$, unlike in the narrow-wedge case. This will be seen in $\text{Corner}^{\downarrow}$'s definition in the proof, to which we turn now.

Proof of Lemma 11.17. From Lemma 11.15, we know that the process $\{\mathcal{P}_{n,1}(x) : x_0 \le x \le n^{1/60}\}$, conditionally on $\mathcal{F}'$, has the law of Brownian bridge of rate two with appropriate endpoints conditioned on not intersecting $\mathcal{P}_{n,2}$ and on $\sup_{x_0 \le x \le n^{1/60}} \mathfrak{h}^{(n)}(x) \le \mathfrak{h}^{(n)}(x_0)$. Let Val (short for "valid") be this conditioning event, i.e.,

$$\mathsf{Val} := \left\{ \mathcal{P}_{n,1}(x) > \mathcal{P}_{n,2}(x) \quad \forall x \in [x_0, n^{1/60}] \right\} \cap \left\{ \sup_{x_0 \le x \le n^{1/60}} \mathfrak{h}^{(n)}(x) \le \mathfrak{h}^{(n)}(x_0) \right\}. \tag{11.28}$$

Recall that $\mathcal{F}''$ is generated by $\mathcal{F}'$ and the additional data of $\mathcal{P}_{n,1}(\cdot)$ on $[x_0, x_0 + A] \cup [x_0 + A + 2, n^{1/60}]$. Note that the σ-algebra $\mathcal{F}$ is generated by $\mathcal{F}''$ along with the additional side bridge data, i.e., $\mathcal{P}_{n,1}^{[x_0+A, x_0+A+1]}$ and $\mathcal{P}_{n,1}^{[x_0+A+1, x_0+A+2]}$.

Recall the following decomposition of a Brownian bridge B of arbitrary endpoints and rate σ^2 on an interval $[a, b]$, with $c \in [a, b]$: conditionally on $B(a)$, $B(b)$, and the side bridges $B^{[a,c]}$ and $B^{[c,b]}$, the distribution of $B(c)$ is that of a normal random variable with mean $\frac{b-c}{b-a}B(a) + \frac{c-a}{b-a}B(b)$ and variance $\sigma^2 \frac{(c-a)(b-c)}{b-a}$. This is because such a Brownian bridge can be decomposed into three independent parts: the side bridges $B^{[a,c]}$ and $B^{[c,b]}$ (both of which have the Brownian bridge law) and the value of $B(c)$ (which is normally distributed as specified).

This decomposition implies the following. By conditioning $\mathcal{P}_{n,1}(\cdot)$ on the side bridge data in addition to $\mathcal{F}''$, and on the $\mathcal{F}$-measurable event that $x_0 + A + 2 \le n^{1/60}$, the $\mathcal{F}$-conditional distribution of $Z = \mathcal{P}_{n,1}(x_0 + A + 1)$ is that of a normal random variable with mean $\frac{1}{2}(\mathcal{P}_{n,1}(x_0 + A) + \mathcal{P}_{n,1}(x_0 + A + 2))$ and variance one (as the Brownian bridge is of rate two), conditioned on Val occurring.

We claim that there exist $\mathcal{F}$-measurable random variables $\mathrm{Corner}^{\downarrow}$ and $\mathrm{Corner}^{\uparrow}$ such that the occurrence of Val is equivalent to Z lying in $[\mathrm{Corner}^{\downarrow}, \mathrm{Corner}^{\uparrow}]$.

We start by focusing on the second event on the right-hand side of (11.28). Consider a fixed upright path γ, and recall the definition of $\mathfrak{h}^{(n),z}(\gamma)$ from (11.23). The second event is equivalent to the event that Z is such that $\mathfrak{h}^{(n),Z}(\gamma) \le \mathfrak{h}^{(n)}(x_0)$ for each upright path γ with endpoint lying in $[x_0, n^{1/60}]$. Now recall that, by Lemma 11.13, with probability one, $\mathfrak{h}^{(n),z}(\gamma)$ is monotone (i.e., non-increasing, non-decreasing, or constant) in z. Thus the condition that $\mathfrak{h}^{(n),z}(\gamma) \le \mathfrak{h}^{(n)}(x_0)$ yields, for each such upright path γ, a condition that z lies in an interval I_γ of the form $(-\infty, \infty)$, $(-\infty, r_\gamma^{\uparrow})$, or $(r_\gamma^{\downarrow}, \infty)$ for some real number $r_\gamma^{\uparrow}$ or $r_\gamma^{\downarrow}$; which form of interval applies depends on the nature of the monotonicity of $\mathfrak{h}^{(n),z}(\gamma)$. Note that $r_\gamma^{\uparrow}$ and $r_\gamma^{\downarrow}$ are $\mathcal{F}$-measurable.

To satisfy the second event in the intersection defining Val in (11.28), Z must lie in the intersection of all of the I_γ as γ varies over the set of upright paths with endpoint in $[x_0, n^{1/60}]$. To satisfy the first event in (11.28), we must also ensure that the value of Z gives non-intersection of $\mathcal{P}_{n,1}^z(\cdot)$ with $\mathcal{P}_{n,2}(\cdot)$. Recall from the definition (11.22) of $\mathcal{P}_{n,1}^z$ that $\mathcal{P}_{n,1}^z(x)$ is non-decreasing in z for all x. Since the $\mathcal{P}_n$ curves are also ordered, it follows that satisfying the first event in the definition of Val is equivalent to Z lying in an infinite ray $I_{\text{lower}} = (r_{\text{lower}}, \infty)$. Further, r_{lower} is an $\mathcal{F}$-measurable random variable.

The idea now is to consider the intersection of I_{lower} and the intervals I_γ corresponding to all paths γ with endpoint in $[x_0, n^{1/60}]$. So, let

$$I = I_{\text{lower}} \cap \bigcap_\gamma I_\gamma.$$

Since we need I to be $\mathcal{F}$-measurable, we take the big intersection over only a countable collection of upright paths γ. More precisely, the intersection is taken over the set of upright paths γ which have start point and endpoint lying in $\mathbb{Q}$ (with the endpoint in $[x_0, n^{1/60}]$) and all of whose jump times from one line to the next occur at values in $\mathbb{Q}$. The continuity of the curves of $\mathcal{P}_n$ implies that this countable dense intersection is sufficient to ensure that $Z \in I$ implies the satisfaction of Val.

In principle, I may be the empty set. But it is not: by the definition of x_0 and I, and since the curves of $\mathcal{P}_n$ obey the non-intersection condition, the value $\mathcal{P}_{n,1}(x_0 + A + 1)$ almost surely lies in I.

We define $\text{Corner}^\downarrow = \inf I$ and $\text{Corner}^\uparrow = \sup I$, which are clearly $\mathcal{F}$-measurable. We note the characterization of $\text{Corner}^\uparrow$ as the largest value of Z which satisfies the second event of Val, i.e., that $\sup_{x_0 \le x \le n^{1/60}} \mathfrak{h}^{(n),z}(x) \le \mathfrak{h}^{(n)}(x_0)$.

We are left with proving the last assertion (11.27). Note that it is immediate from the definition of I that, when $Z = \text{Corner}^\uparrow$, there exists some $x' \neq x_0$ such that $\mathfrak{h}^{(n),z}(x) = \mathfrak{h}^{(n)}(x_0)$. We claim that at least one such x' must lie in $[x_0 + A, x_0 + A + 2]$. Suppose to the contrary that $\sup_{x \in [x_0+A, x_0+A+2]} \mathfrak{h}^{(n),z}(x) < \mathfrak{h}^{(n)}(x_0)$. Consider $\mathfrak{h}^{(n),Z+\varepsilon}$ for small $\varepsilon > 0$. We see from Lemma 11.13 that $\mathfrak{h}^{(n),Z+\varepsilon}(x) \le \mathfrak{h}^{(n),Z}(x) \le \mathfrak{h}^{(n)}(x_0)$ for all $x \notin [x_0 + A, x_0 + A + 2]$. But for all small enough ε, we would still have $\sup_{x \in [x_0+A, x_0+A+2]} \mathfrak{h}^{(n),Z+\varepsilon}(x) < \mathfrak{h}^{(n)}(x_0)$ by Lemma 11.14, contradicting the definition of I via the characterization of $\text{Corner}^\uparrow$ noted in the previous paragraph. This completes the proof of Lemma 11.17. □

11.8 Positive probability favourable data

We next define a $\mathcal{F}$-measurable favourable event $\mathsf{Fav}_{K,L}$, which we will show holds with positive probability. Recall that we are proving Proposition 11.11, which asserts a lower bound on the probability of the twin peaks' event. The argument for this proposition will rely on resampling some randomness, namely $Z = \mathcal{P}_{n,1}(x_0 + A + 1)$, conditionally on the data in $\mathcal{F}$. The role of the favourable event is to specify a class of good $\mathcal{F}$-measurable data under which the resampling can be analysed more easily.

We again adopt the shorthand of x_0 for x_0^n, and let $\mu := \frac{1}{2}\big(\mathcal{P}_{n,1}(x_0 + A) + \mathcal{P}_{n,1}(x_0 + A + 2)\big)$ be the $\mathcal{F}$-measurable mean of the normal random variable in the description of Z's $\mathcal{F}$-conditional distribution from Lemma 11.17. We set

$$\mathsf{Fav}_{K,L} = \mathsf{F}_1 \cap \mathsf{F}_2 \cap \mathsf{F}_3' \cap \mathsf{F}_4',$$

where

$$\mathsf{F}_1 = \Big\{\text{Corner}^{\uparrow} \le 4K\Big\}, \qquad \mathsf{F}_2 = \Big\{\mu \in [-K, K]\Big\},$$
$$\mathsf{F}_3' = \Big\{|\mathfrak{h}^{(n)}(x_0)| \le L^{1/2}\Big\}, \qquad \mathsf{F}_4' = \Big\{|x_0| \le L - A - 2\Big\},$$

(we will shortly define F_3 and F_4 to be modified versions of F_3' and F_4' which will be more convenient to work with).

Let us now say a few words on the form of $\mathsf{Fav}_{K,L}$ and the proof idea of Proposition 11.11. From Lemma 11.17, we see that, when $Z = \text{Corner}^{\uparrow}$, we have that $\sup_{x\in[x_0+A,x_0+A+2]} \mathfrak{h}^{(n)}(x) = \mathfrak{h}^{(n)}(x_0)$. Also, we see from Lemma 11.14 that reducing Z's value from this level affects $\sup_{x\in[x_0+A,x_0+A+2]} \mathfrak{h}^{(n)}(x)$ in a Lipschitz manner. Thus the event in Proposition 11.11 occurs if Z is within order ε of $\text{Corner}^{\uparrow}$.

We know from Lemma 11.17 that, conditionally on $\mathcal{F}$, Z is distributed as a normal random variable with mean μ and variance one conditioned on lying inside $[\text{Corner}^{\downarrow}, \text{Corner}^{\uparrow}]$. Thus to get a good lower bound on the event that Z is close to $\text{Corner}^{\uparrow}$, it is enough to know that, with positive probability, the mean of Z is not too extreme and that the upper limit $\text{Corner}^{\uparrow}$ is not too high. These are the first two events in the intersection defining $\mathsf{Fav}_{K,L}$, and the mentioned positive probability lower bound is the content of the next lemma. The third event handles the first extra condition in the definition of $\mathsf{TP}^{\varepsilon}_{A,L}$ (11.3) on the value of $\mathfrak{h}^{(n)}$'s maximum, while the final event in $\mathsf{Fav}_{K,L}$'s intersection is imposed merely to ensure the second extra condition in (11.3), that twin peaks occurs in the interval $[-L, L]$.

Lemma 11.18. *Let $\mathfrak{h}_0 : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ satisfy Assumption 11.1 and consider $A > 0$. There exist K and L_0 (both depending on γ, θ, and A) such that, for all $L > L_0$, there exists n_0 (depending on γ, θ, and L) so that, for all $n > n_0$,*

$$\mathbb{P}(\mathsf{Fav}_{K,L}) \ge \frac{1}{2}.$$

Further, K and L_0 may be made to depend on γ, θ and A in a continuous manner.

Proof. We start by specifying some further good events. Let $y_0 \in [-\theta, \theta]$ be some fixed real number such that $\mathfrak{h}_0(y_0) \ge -2\theta$. For $M > 0$ to be specified later, define

$$\mathsf{F}_3 = \big\{|\mathfrak{h}^{(n)}(x_0)| \le K\big\}, \qquad \mathsf{F}_4 = \big\{|x_0| \le M\big\},$$
$$\mathsf{F}_5 = \Big\{\mathcal{S}_n(y_0, -M) \ge -K/2\Big\}, \qquad \mathsf{F}_6 = \Big\{\mathcal{P}_{n,1}(-M) \le K/2\Big\}.$$

Note that $\mathsf{Fav}_{K,L} \supseteq \mathsf{F}_1 \cap \mathsf{F}_2 \cap \mathsf{F}_3 \cap \mathsf{F}_4$ when $L \ge M + A + 2$ and $L^{1/2} \ge K$. We set L_0 high enough that both conditions on L are met whenever $L \ge L_0$ (when K and M are set, which will be done in a way not depending on L).

We will show that $\mathbb{P}\left(\mathsf{Fav}^c_{K,L}\right) \le \frac{1}{2}$ for large enough K and L by showing the stronger statement that, for $\delta = 1/10$ and appropriate choices of K, L, and M, we have

$$\mathbb{P}\left(\bigcup_{i=1}^{6} \mathsf{F}_i^c\right) \le 5\delta = \frac{1}{2}.$$

Bounding $\mathbb{P}(\mathsf{F}_4^c)$: This is a simple application of Lemma 11.10, which yields $M = M(\gamma, \theta)$ such that $\mathbb{P}(\mathsf{F}_4^c) \le \delta = \frac{1}{10}$ for $n \ge n_0 = n_0(\gamma, \theta)$. We fix the value of M obtained here for the rest of the proof.

Bounding $\mathbb{P}(\mathsf{F}_2^c \cap \mathsf{F}_4)$: On the event F_4 we have $|x_0| \le M$. Also, for large enough K depending only on $\delta = 1/10$, A, and M,

$$\mathbb{P}\left(\inf_{|x| \le M+A+2} \mathcal{P}_{n,1}(x) < -K \right) \le \delta/2 \quad \text{and} \quad \mathbb{P}\left(\sup_{|x| \le M+A+2} \mathcal{P}_{n,1}(x) > K \right) \le \delta/2;$$

this is implied by Proposition 11.9 after recalling from (11.12) that $\mathcal{P}_{n,1}(x) = \mathcal{S}_n(0, x)$. Thus $\mathbb{P}(\mathsf{F}_2^c \cap \mathsf{F}_4)$ is at most δ, since $\mu = \frac{1}{2}(\mathcal{P}_{n,1}(x_0 + A) + \mathcal{P}_{n,1}(x_0 + A + 2))$ is bounded above and below on F_4 by $\sup_{x \in [-M-A-2, M+A+2]} \mathcal{P}_{n,1}(x)$ and $\inf_{x \in [-M-A-2, M+A+2]} \mathcal{P}_{n,1}(x)$.

Bounding $\mathbb{P}(\mathsf{F}_5^c)$ *and* $\mathbb{P}(\mathsf{F}_6^c)$: These correspond respectively to lower and upper tails on one-point last passage values, i.e., on $\mathcal{S}_n(y, x)$ for fixed y and x, because $\mathcal{P}_{n,1}(-M) = \mathcal{S}_n(0, -M)$ in view of (11.12). Thus, we obtain $\mathbb{P}(\mathsf{F}_5^c) \le \delta$ and $\mathbb{P}(\mathsf{F}_6^c) \le \delta$ by applying Proposition 11.9 in a closed interval of unit length around the starting and ending points and by setting K high enough, depending on θ.

Bounding $\mathbb{P}(\mathsf{F}_3^c \cap \mathsf{F}_4 \cap \mathsf{F}_5)$: We first bound the probability that $\mathfrak{h}^{(n)}(x_0) \ge K$ on F_4. Recall that, by assumption, $\mathfrak{h}_0(y) \le -\gamma y^2$ for all $y \in \mathbb{R}$. Note that, on F_4, $\mathfrak{h}^{(n)}(x_0) = \sup_{|x| \le M} \mathfrak{h}^{(n)}(x)$. Thus,

$$\sup_{|x| \le M} \mathfrak{h}^{(n)}(x) = \sup_{\substack{|x| \le M \\ 0 \le y \le n^{1/60}}} \left(\mathfrak{h}_0(y) + \mathcal{S}_n(y, x) \right) \le \sup_{\substack{|x| \le M \\ 0 \le y \le n^{1/60}}} \left(\mathcal{S}_n(y, x) - \gamma y^2 \right).$$

By a union bound, we can bound $\mathbb{P}\left(\sup_{|x| \le M} \mathfrak{h}^{(n)}(x) > K \right)$ above by

$$\sum_{j=0}^{\lceil n^{1/60} \rceil} \mathbb{P}\left(\sup_{\substack{|x| \le M \\ y \in [j, j+1]}} \left(\mathcal{S}_n(y, x) - \gamma y^2 \right) > K \right) \le \sum_{j=0}^{\lceil n^{1/60} \rceil} \mathbb{P}\left(\sup_{\substack{|x| \le M \\ y \in [j, j+1]}} \mathcal{S}_n(y, x) > K + \gamma j^2 \right)$$
$$\le \sum_{j=0}^{\lceil n^{1/60} \rceil} C \max(M^2, j^2) \exp\left(-c(K^{3/2} + j^3) \right),$$

the final inequality by an application of Proposition 11.9. The final expression can be made less than $\delta = \frac{1}{10}$ by raising K appropriately (depending on M and γ) if needed. Doing so, we learn that $\mathbb{P}(\{\mathfrak{h}^{(n)}(x_0) \ge K\} \cap \mathsf{F}_4) \le \delta$.

Next we bound the probability that $\mathfrak{h}^{(n)}(x_0) \le -K$ on F_5. Recall that $y_0 \in [-\theta, \theta]$ is such that $\mathfrak{h}_0(y_0) \ge -2\theta$, and increase K if needed so that $\mathfrak{h}_0(y_0) \ge -K/3$. Since x_0 is the maximizer of $\mathfrak{h}^{(n)}$, and (11.15) holds, we see that

$$\mathfrak{h}^{(n)}(x_0) \ge \mathfrak{h}_0(y_0) + \mathcal{S}_n(y_0, -M) \ge -5K/6 > -K,$$

the last inequality holding on F_5. Thus $\mathbb{P}(\{\mathfrak{h}^{(n)}(x_0) \le -K\} \cap F_5) = 0$. Overall we have shown that $\mathbb{P}(\mathsf{F}_3^c \cap \mathsf{F}_4 \cap \mathsf{F}_5) \le \delta$.

Bounding $\mathbb{P}(\mathsf{F}_1^c \cap \mathsf{F}_3 \cap \mathsf{F}_4 \cap \mathsf{F}_5 \cap \mathsf{F}_6)$: Recall that we have set y_0 and K such that $\mathfrak{h}_0(y_0) \ge -K/3 > -K$. Observe that $\mathrm{Corner}^{\uparrow} > 4K$ implies that there is a value of z in $[4K, \infty)$ such that $\mathfrak{h}^{(n)}(x_0) = \sup_{x \in [x_0+A, x_0+A+2]} \mathfrak{h}^{(n),z}(x)$ by Lemma 11.17. But we will now show that if $z \ge 4K$, then, on the event $\bigcap_{i=3}^6 \mathsf{F}_i$, we have that $\mathfrak{h}^{(n),z}(x_0 + A + 1) > \mathfrak{h}^{(n)}(x_0)$; this is a contradiction and so the probability we are bounding must be zero. We use the formula for $\mathfrak{h}^{(n),z}$ from (11.23), and the formula (11.14) relating $\mathcal{S}_n$ and LPP values through $\mathcal{P}_n$. Indeed, if $z \ge 4K$ and the event $\bigcap_{i=3}^6 \mathsf{F}_i$ holds, then

$$\begin{aligned}
\mathfrak{h}^{(n),z}(x_0 + A + 1) &= \sup_{0 \le y \le n^{1/60}} \left(\mathfrak{h}_0(y) + \mathcal{P}_n^z[(-\tfrac{1}{2}n^{1/3} + y, n) \to (x_0 + A + 1, 1)] - n^{2/3} \right) \\
&\ge -K + \left(\mathcal{P}_n^z[(-\tfrac{1}{2}n^{1/3} + y_0, n) \to (-M, 1)] - n^{2/3} \right) \\
&\quad + \mathcal{P}_n^z[(-M, 1) \to (x_0 + A + 1, 1)] \\
&= -K + \mathcal{S}_n(y_0, -M) + z - \mathcal{P}_{n,1}(-M) \\
&\ge 3K + \mathcal{S}_n(y_0, -M) - \mathcal{P}_{n,1}(-M) \ge 2K.
\end{aligned}$$

The first inequality bounded the supremum by the choice of $y = y_0$ and used our assumption that $\mathfrak{h}_0(y_0) > -K$; the penultimate inequality used the assumption that $z \ge 4K$; and the final inequality used the bounds that hold on $\mathsf{F}_5 \cap \mathsf{F}_6$. The conclusion $\mathfrak{h}^{(n),z}(x_0 + A + 1) \ge 2K$ contradicts $\mathfrak{h}^{(n)}(x_0) \le K$, which holds on F_4, since, on this event, $x_0 + A + 1 \in [-M, M]$. Thus the probability we are bounding is zero.

Overall we have shown that $\mathbb{P}\big(\bigcup_{i=1}^6 \mathsf{F}_i^c\big) \le 5\delta = 1/2$. It may be easily checked that the setting of K and L_0 can be made to depend on γ, θ, and A continuously, completing the proof of Lemma 11.18. □

11.9 Performing the resampling: the proof of Proposition 11.11

In this proof, we will need a monotonicity property of conditional probabilities of the normal distribution. The proof is a straightforward calculation that we omit here, but details are available in [CHH19, Lemma 5.15].

Lemma 11.19. *Fix $r > 0$, $m \in \mathbb{R}$, and $\sigma^2 > 0$, and let X be distributed as $N(m, \sigma^2)$. Then the quantity $\mathbb{P}(X \ge s - r \mid X \le s)$ is a strictly decreasing function of $s \in \mathbb{R}$.*

Proof of Proposition 11.11. We fix K and L_0 as given by Lemma 11.18. For any given $L > L_0$, we have that $\mathbb{P}(\mathsf{Fav}_{K,L}) \ge 1/2$.

Recall $x_0 = x_0^n = \arg\max_{|x| \le n^{1/60}} \mathfrak{h}^{(n)}(x)$. We have that

$$\mathbb{P}\left(\sup_{x \in [x_0+A, x_0+A+2]} \mathfrak{h}^{(n)}(x) > \mathfrak{h}^{(n)}(x_0) - \varepsilon;\ |\mathfrak{h}^{(n)}(x_0)| \le L^{1/2};\ |x_0| \le L - A - 2 \right)$$

$$= \mathbb{E}\left[\mathbb{P}\left(\sup_{x\in[x_0+A,x_0+A+2]} \mathfrak{h}^{(n)}(x) > \mathfrak{h}^{(n)}(x_0) - \varepsilon \;\middle|\; \mathcal{F}\right) \mathbb{1}_{|\mathfrak{h}^{(n)}(x_0)|\le L^{1/2}\,|x_0|\le L-A-2}\right].$$

We have to bound below the inner conditional probability. Set $\mathsf{Fav} = \mathsf{Fav}_{K,L}$ for notational convenience. Recall that the occurrence of Fav implies that $|x_0| \le L-A-2$ and $|\mathfrak{h}^{(n)}(x_0)| \le L^{1/2}$. We claim that, on Fav, the event of the inner conditional probability is implied by

$$Z := \mathcal{P}_{n,1}(x_0 + A + 1) \in \left[\mathrm{Corner}^{\downarrow} \vee (\mathrm{Corner}^{\uparrow} - \varepsilon/2),\ \mathrm{Corner}^{\uparrow}\right] =: I_\varepsilon. \tag{11.29}$$

The validity of the claim follows from two facts. The first is that $\sup_{x\in[x_0+A,x_0+A+2]} \mathfrak{h}^{(n),z}(x) = \mathfrak{h}^{(n)}(x_0)$ when $z = \mathrm{Corner}^{\uparrow}$ (from Lemma 11.17); and the second is that, almost surely for all $z \in \mathbb{R}$, $\left|\sup_{x\in[x_0+A,x_0+A+2]} \mathfrak{h}^{(n),z_1}(x) - \sup_{x\in[x_0+A,x_0+A+2]} \mathfrak{h}^{(n),z_2}(x)\right|$ is at most $2|z_1 - z_2|$ (from Lemma 11.14). We apply this to $z_1 = \mathrm{Corner}^{\uparrow}$ and $z_2 \in I_\varepsilon$.

With this preparation, we see that

$$\begin{aligned}
&\mathbb{P}\left(\sup_{x\in[x_0+A,x_0+A+2]} \mathfrak{h}^{(n)}(x) > \mathfrak{h}^{(n)}(x_0) - \varepsilon \;\middle|\; \mathcal{F}\right) \mathbb{1}_{|\mathfrak{h}^{(n)}(x_0)|\le L^{1/2},\,|x_0|\le L-A-2} \\
&\qquad\qquad \ge \mathbb{P}\left(\sup_{x\in[x_0+A,x_0+A+2]} \mathfrak{h}^{(n)}(x) > \mathfrak{h}^{(n)}(x_0) - \varepsilon \;\middle|\; \mathcal{F}\right) \cdot \mathbb{1}_{\mathsf{Fav}} \\
&\qquad\qquad \ge \mathbb{P}\big(Z \in I_\varepsilon \;\big|\; \mathcal{F}\big) \cdot \mathbb{1}_{\mathsf{Fav}}.
\end{aligned} \tag{11.30}$$

Recall now from Lemma 11.17 that Z is distributed as a normal random variable with mean $\mu = \frac{1}{2}(\mathcal{P}_{n,1}(x_0+A)+\mathcal{P}_{n,1}(x_0+A+2))$ and variance one, conditioned on lying inside $[\mathrm{Corner}^{\downarrow}, \mathrm{Corner}^{\uparrow}]$. Observe from (11.29) that I_ε is one of two intervals: $[\mathrm{Corner}^{\downarrow}, \mathrm{Corner}^{\uparrow}]$ or $[\mathrm{Corner}^{\uparrow}-\varepsilon/2, \mathrm{Corner}^{\uparrow}]$. In the first case, the conditional probability in (11.30) equals one. We show now that, in the second case, the conditional probability is bounded below by $\eta\varepsilon$ for some constant $\eta > 0$.

We let N be a standard normal random variable with mean zero and variance one. Then, on the event $\mathsf{Fav} \cap \{\mathrm{Corner}^{\downarrow} < \mathrm{Corner}^{\uparrow} - \varepsilon/2\}$, we have that $|\mu| \le K$ and $\mathrm{Corner}^{\uparrow} \le 4K$, which implies that, on the same event,

$$\begin{aligned}
\mathbb{P}\big(Z \in I_\varepsilon \;\big|\; \mathcal{F}\big) &= \mathbb{P}\left(N + \mu \in [\mathrm{Corner}^{\uparrow} - \varepsilon/2, \mathrm{Corner}^{\uparrow}] \;\middle|\; N + \mu \in [\mathrm{Corner}^{\downarrow}, \mathrm{Corner}^{\uparrow}], \mathcal{F}\right) \\
&= \frac{\mathbb{P}\left(N + \mu \in [\mathrm{Corner}^{\uparrow} - \varepsilon/2, \mathrm{Corner}^{\uparrow}] \;\middle|\; \mathcal{F}\right)}{\mathbb{P}\left(N + \mu \in [\mathrm{Corner}^{\downarrow}, \mathrm{Corner}^{\uparrow}] \;\middle|\; \mathcal{F}\right)} \\
&\ge \mathbb{P}\left(N + \mu \ge \mathrm{Corner}^{\uparrow} - \varepsilon/2 \;\middle|\; N + \mu \le \mathrm{Corner}^{\uparrow}, \mathcal{F}\right) \\
&\ge \mathbb{P}\left(N + \mu \ge 4K - \varepsilon/2 \;\middle|\; N + \mu \le 4K, \mathcal{F}\right) \\
&\ge \mathbb{P}\big(N + \mu \in [4K - \varepsilon/2, 4K] \;\big|\; \mathcal{F}\big).
\end{aligned}$$

For the first equality, we interpret $\mathbb{P}(Z \in \cdot \mid \mathcal{F})$ as a regular conditional distribution, which exists as Z takes values in $\mathbb{R}$ (see [Kal02, Theorem 6.3]); conditioning on an event E is then understood by the usual equality of conditional probability with a ratio of probabilities, i.e., $\mathbb{P}(\cdot \mid E, \mathcal{F}) = \mathbb{P}(\cdot \cap E \mid \mathcal{F})/\mathbb{P}(E \mid \mathcal{F})$. Then the first equality follows from the characterization of Z's law recalled above after (11.30). The second equality can be seen by noting that we are working on the event that $\mathrm{Corner}^{\downarrow} < \mathrm{Corner}^{\uparrow} - \varepsilon/2$, and some simple manipulations of the probabilities in the ratio gives the third line, i.e., the first inequality. The penultimate inequality used the monotonicity property of normal random variables recorded in Lemma 11.19 and that $\mathrm{Corner}^{\uparrow} \leq 4K$ on Fav. Now the form of the normal density gives that the final expression is bounded below by $\eta\varepsilon \cdot \mathbb{1}_{\mathsf{Fav}}$ for some $\eta > 0$ depending only on K, since $|\mu| \leq K$; further, this dependence is clearly continuous in K.

Substituting into (11.30) this bound, as well as the earlier bound of one in the case that $\mathrm{Corner}^{\downarrow} \geq \mathrm{Corner}^{\uparrow} - \varepsilon/2$, gives that

$$\mathbb{P}\left(\sup_{x \in [x_0 + A, x_0 + A + 2]} \mathfrak{h}^{(n)}(x) > \mathfrak{h}^{(n)}(x_0) - \varepsilon;\ |\mathfrak{h}^{(n)}(x_0)| \leq L;\ |x_0| \leq L - A - 2 \right) \geq \eta\varepsilon \cdot \mathbb{P}\big(\mathsf{Fav}_{K,L}\big) \geq \frac{1}{2}\eta\varepsilon,$$

in view of K and L being such that $\mathbb{P}(\mathsf{Fav}_{K,L}) \geq 1/2$. Relabelling η completes the proof of Proposition 11.11. □

Appendix A

Proofs of basic bootstrapping tools

In this appendix we explain how to obtain the first, second, and fourth tools of Section 7.6, i.e., Theorem 7.14 and Propositions 7.15 and 7.17, and provide the outstanding proofs of Lemmas 9.5, 9.10, and 9.7 from the main text. The third tool was already explained in Section 9.3.

The proofs of the first and fourth tools follow verbatim from corresponding results in [BGHH20] by replacing the upper and lower tails used there with Assumption 3; the parabolic curvature assumption there is provided by our Assumption 2. In particular, Theorem 7.14 follows from [BGHH20, Theorem 3.3] and Proposition 7.17 from [BGHH20, Proposition 3.7].

The proof of the second tool, Proposition 7.15, will be addressed in Section A, after we next provide the outstanding proofs of Lemmas 9.5, 9.10, and 9.7 from the main text.

We start with the proofs of Lemmas 9.5 and 9.10 on an upper tail bound of the interval-to-interval weight; this largely follows the proof of [BGHH20, Proposition 3.5]. The strategy is to back up from the intervals appropriately and consider a point-to-point weight for which we have tail bounds by hypothesis; this strategy was illustrated in Figure 9.2 and in the proof of Lemma 9.9.

Proofs of Lemmas 9.5 and 9.10. We prove Lemma 9.5 and indicate at the end the modifications for Lemma 9.10. We set $\lambda = \lambda_j$ and $\lambda' = \lambda_{j+1}$ to avoid confusion later when we describe the modifications for Lemma 9.10. Note that $\lambda' < \lambda$.

By considering the event that Z is large and two events defined in terms of the environment outside of U, we find a point-to-point path which has large length. To define these events, first define points ϕ_{low} and ϕ_{up} on either side of the lower and upper intervals as follows, where $\delta = \frac{1}{2}\left(\frac{\lambda}{\lambda'} - 1\right) > 0$:

$$\begin{aligned} \phi_{\text{low}} &:= (-\delta r, -\delta r) \\ \phi_{\text{up}} &:= ((1+\delta)r - w, (1+\delta)r + w)\,. \end{aligned}$$

Let u^* and v^* be the points on A and B where the suprema in the definition of Z are attained, and let the events E_{low} and E_{up} be defined as

$$E_{\text{low}} = \left\{ X_{\phi_{\text{low}}, u^* - (1,0)} > \mu\delta r - \frac{t}{3} r^{1/3} \right\} \quad \text{and} \quad E_{\text{up}} = \left\{ X_{v^* + (1,0), \phi_{\text{up}}} > \mu\delta r - \frac{t}{3} r^{1/3} \right\}.$$

Let $\tilde{r} = \frac{\lambda}{\lambda'} r = (1+2\delta)r$, and observe that the diagonal distance of ϕ_{low} and ϕ_{up} is $\tilde{r}$. Also note

$$X_{\phi_{\text{low}},\phi_{\text{up}}} \geq X_{\phi_{\text{low}},u*-(1,0)} + Z + X_{v^*+(1,0),\phi_{\text{up}}}.$$

Then we have the following:

$$\begin{aligned}
\mathbb{P}\left(Z > \mu r - \lambda' \frac{Gw^2}{r} + tr^{1/3}, E_{\text{low}}, E_{\text{up}}\right) &\leq \mathbb{P}\left(X_{\phi_{\text{low}},\phi_{\text{up}}} \geq \mu(1+2\delta)r - \lambda' \frac{Gw^2}{r} + \frac{t}{3} r^{1/3}\right) \qquad \text{(A.1)} \\
&= \mathbb{P}\left(X_{\phi_{\text{low}},\phi_{\text{up}}} \geq \mu\tilde{r} - \lambda \frac{Gw^2}{\tilde{r}} + \frac{t}{3} \cdot \left(\frac{\lambda'}{\lambda}\right)^{1/3} \cdot (\tilde{r})^{1/3}\right) \\
&\leq \begin{cases} \exp(-\tilde{c}t^{\beta}) & t_0 < t < r^{\zeta} \\ \exp(-\tilde{c}t^{\alpha}) & t \geq r^{\zeta}. \end{cases}
\end{aligned}$$

The final inequality uses the hypothesis (9.1) on the point-to-point tail, which is applicable since the antidiagonal separation of ϕ_{low} and ϕ_{up} is w while the diagonal separation is $(1+2\delta)r$, and clearly $|w| \leq r^{5/6}$ implies $|w| \leq (1+2\delta)^{5/6} r^{5/6}$. We applied (9.1) with $\theta = t(\lambda'/\lambda)^{1/3}/3$, which is required to be greater than θ_0. This translates to $t \geq t_0$ for a t_0 depending on θ_0 and λ'/λ. Similarly, we absorbed the λ'/λ dependency in the tail into the value of $\tilde{c}$, which thus depends on the original tail coefficient c in (9.1) and λ'/λ.

Let us denote conditioning on the environment U by the notation $\mathbb{P}(\cdot \mid U)$. By this we mean we condition on the weights of vertices interior to U as well as those on the lower side A, but not those on the upper side B. Then we see

$$\begin{aligned}
&\mathbb{P}\left(Z > \mu r - \lambda' \frac{Gw^2}{r} + tr^{1/3}, E_{\text{low}}, E_{\text{up}} \mid U\right) \\
&\qquad = \mathbb{P}\left(Z > \mu r - \lambda' \frac{Gw^2}{r} + tr^{1/3} \mid U\right) \cdot \mathbb{P}\left(E_{\text{low}} \mid U\right) \cdot \mathbb{P}\left(E_{\text{up}} \mid U\right).
\end{aligned}$$

So with (A.1), all we need is a lower bound on $\mathbb{P}\left(E_{\text{low}} \mid U\right)$ and $\mathbb{P}\left(E_{\text{up}} \mid U\right)$. This is straightforward using independence of the environment below and above U from U:

$$\mathbb{P}\left(E_{\text{lower}}^c \mid U\right) \leq \sup_{u \in A} \mathbb{P}\left(X_{\phi_{\text{low}},u} \leq \mu\delta r - \frac{t}{3} r^{1/3}\right) \leq \frac{1}{2} \qquad \text{(A.2)}$$

for large enough t (independent of δ) and r (depending on δ), using Assumption 3b. A similar upper bound holds for $\mathbb{P}\left(E_{\text{upper}}^c \mid U\right)$. Together this gives

$$\mathbb{P}\left(Z > \mu r - \lambda' \frac{Gw^2}{r} + tr^{1/3}, E_{\text{low}}, E_{\text{up}} \mid U\right) \geq \frac{1}{4} \cdot \mathbb{P}\left(Z > \mu r - \lambda' \frac{Gw^2}{r} + tr^{1/3}\right),$$

and taking expectation on both sides, combined with (A.1), gives Lemma 9.5. The fact that λ'/λ depends only on j and the previously mentioned dependencies gives the claimed dependencies of $\tilde{t}_0, \tilde{r}_0$, and $\tilde{c}$.

To prove Lemma 9.10, we take $\delta = 1$, which is equivalent to $\lambda' = \lambda/3$. Then in (A.1) the final bound is done with the hypothesized bound on $X_{\phi_{\text{low}},\phi_{\text{up}}}$, i.e.,

$$\mathbb{P}\left(X_{\phi_{\text{low}},\phi_{\text{up}}} > \mu\tilde{r} - \lambda\frac{Gw^2}{\tilde{r}} + tr^{1/3}\right) \le \exp(-\tilde{c}t^{\alpha}).$$

Applying this bound requires $|w| \le \tilde{r}/2$. Since $|w| \le r$ and $\tilde{r} = \lambda r/\lambda' = 3r$, this is valid. □

Next we prove Lemma 9.7, on a constant probability lower bound on the lower tail, based on Assumptions 2 and 3b.

Proof of Lemma 9.7. Let $\tilde{X}_r^z = X_r^z - \mu r + Gz^2/r$. We know from Assumption 2 that $\mathbb{E}[\tilde{X}_r^z] \le -g_2 r^{1/3}$. Let E be the event

$$E = E(\theta) = \left\{X_r^z < \mu r - \frac{Gz^2}{r} - \theta r^{1/3}\right\},$$

so that $\mathbb{P}(E) \le \exp(-c\theta^{\alpha})$ for $\theta > \theta_0$, by Assumption 3b.

Observe that $-\tilde{X}_r^z \mathbb{1}_E$ is a positive random variable and so, by Assumption 3b,

$$\begin{aligned}
\mathbb{E}[-\tilde{X}_r^z \mathbb{1}_E] &= r^{1/3}\int_0^{\infty} \mathbb{P}\left(\tilde{X}_r^z \mathbb{1}_E < -tr^{1/3}\right)\,\mathrm{d}t \\
&= r^{1/3}\left[\theta\cdot\mathbb{P}\left(X_r^z < \mu r - \frac{Gz^2}{r} - \theta r^{1/3}\right) + \int_{\theta}^{\infty}\mathbb{P}\left(X_r^z < \mu r - \frac{Gz^2}{r} - tr^{1/3}\right)\,\mathrm{d}t\right] \\
&\le r^{1/3}\left[\theta\exp(-c\theta^{\alpha}) + \int_{\theta}^{\infty}\exp(-ct^{\alpha})\,\mathrm{d}t\right];
\end{aligned}$$

this may be made smaller than $0.5 g_2 r^{1/3}$ by taking θ large enough. We now set θ to such a value. We also have $\mathbb{E}[\tilde{X}_r^z] = \mathbb{E}[\tilde{X}_r^z(\mathbb{1}_E + \mathbb{1}_{E^c})]$. Combining this, the above lower bound on $\mathbb{E}[\tilde{X}_r^z \mathbb{1}_E]$, and the upper bound on $\mathbb{E}[X_r^z]$, gives that

$$\mathbb{E}[\tilde{X}_r^z \mathbb{1}_{E^c}] \le -\frac{1}{2} g_2 r^{1/3}. \tag{A.3}$$

The fact that $\tilde{X}_r^z \mathbb{1}_E^c$ is supported on $[-\theta r^{1/3}, \infty)$ implies that

$$\mathbb{P}\left(X_r^z \mathbb{1}_{E^c} < \mu r - \frac{Gz^2}{r} - \frac{1}{4} g_2 r^{1/3}\right) \ge \frac{g_2}{4\theta}; \tag{A.4}$$

this follows from (A.3) and since

$$\mathbb{E}[\tilde{X}_r^z \mathbb{1}_E^c] \ge -\theta\cdot\mathbb{P}\left(\tilde{X}_r^z \mathbb{1}_E^c < -\frac{1}{4} g_2 r^{1/3}\right) - \frac{1}{4} g_2 r^{1/3}\mathbb{P}\left(\tilde{X}_r^z \mathbb{1}_E^c \ge -\frac{1}{4} g_2 r^{1/3}\right)$$

$$\geq -\theta \cdot \mathbb{P}\left(\tilde{X}_r^z \mathbb{1}_E^c < -\frac{1}{4} g_2 r^{1/3}\right) - \frac{1}{4} g_2 r^{1/3}.$$

Since $\tilde{X}_r^z \mathbb{1}_E < 0$, it follows that $\tilde{X}_r^z \leq \tilde{X}_r^z \mathbb{1}_{E^c}$, so (A.4) gives a lower bound on the lower tail of X_r^z, as desired, with $C = \frac{1}{4} g_2$ and $\delta = g_2/4\theta$. □

Proof of transversal fluctuation bound, Proposition 7.15

In this section we prove Proposition 7.15 on the tail (with exponent 2α) of the transversal fluctuation of the geodesic path on scale $r^{2/3}$; we closely follow the proof of Theorem 11.1 of the preprint [BSS14], but adapted to our setting and assumptions. We give the argument for the left-most geodesic Γ_r^z from $(1,1)$ to $(r-z, r+z)$; the argument is symmetric for the right-most geodesic. (Note that these are well-defined by the planarity and the weight-maximizing properties of all geodesics).

We start with a similar bound at the midpoint of the geodesic, which needs some notation. For $x \in [\![1, r]\!]$, let $\Gamma_r^z(x)$ be the unique point y such that $(x-y, x+y) \in \Gamma_r^z$.

Proposition A.1. *Under Assumption 2 and 3, there exist $c = c(\alpha) > 0$, r_0, and s_0 such that, for $r > r_0$, $s > s_0$, and $|z| \leq r^{5/6}$,*

$$\mathbb{P}\left(|\Gamma_r^z(r/2)| > z/2 + s r^{2/3}\right) \leq 2\exp(-c s^{2\alpha}).$$

To prove this we will need a bound on the maximum, $\widetilde{Z}$, of fluctuations of the point-to-point weight as the endpoint varies over an interval, i.e.,

$$\widetilde{Z} = \sup_{v \in \mathbb{L}_{\text{up}}} \Big(X_v - \mathbb{E}[X_v]\Big),$$

where $\mathbb{L}_{\text{up}}$ is the interval of width $2r^{2/3}$ around $(r-w, r+w)$. Note that this is not the same as the point-to-interval weight.

Lemma A.2. *Let $K > 0$ and $|w| \leq K r^{5/6}$. Under Assumptions 2 and 3, there exist $c > 0$, $\theta_0 = \theta_0(K)$, and r_0, such that, for $\theta > \theta_0$ and $r > r_0$,*

$$\mathbb{P}\left(\widetilde{Z} > \theta r^{1/3}\right) \leq \exp(-c\theta^{\alpha}).$$

Proof. The proof is very similar to that of Lemma 9.5 above.

We take $\phi_{\text{up}} = (2(r-w), 2(r+w))$ to be the backed up point. Let $v^* \in \mathbb{L}_{\text{up}}$ be the maximizing point in the definition of $\widetilde{Z}$. For clarity, define the lower and upper mean weight functions M_{low} and M_{up} by $M_{\text{low}}(v) = \mathbb{E}[X_v]$ and $M_{\text{up}}(v) = \mathbb{E}[X_{v,\phi_{\text{up}}}]$; this is to use the unambiguous notation $M_{\text{low}}(v^*)$ (which is a function of v^*) instead of $\mathbb{E}[X_{v^*}]$. We also define

$$E_{\text{up}} = \left\{ X_{v^*+(1,0),\phi_{\text{up}}} - M_{\text{up}}(v^* + (1,0)) > -\frac{\theta}{2} r^{1/3} \right\}.$$

Now observe that

$$X_{v^*} - M_{\text{low}}(v^*) + X_{v^*+(1,0),\phi_{\text{up}}} - M_{\text{up}}(v^* + (1,0)) \\ \leq X_{\phi_{\text{up}}} - \inf_{v \in \mathbb{L}_{\text{up}}} \left(M_{\text{low}}(v) + M_{\text{up}}(v + (1,0)) \right). \tag{A.5}$$

We want to replace the infimum on the right hand side by $\mathbb{E}[X_{\phi_{\text{up}}}]$. The latter is at most $2\mu r - 2Gw^2/r$. We need to show that the infimum term is at least something which is within $O(r^{1/3})$ of this expression. For this we do the following calculation using Assumption 2. Parametrize $v \in \mathbb{L}_{\text{up}}$ as $(r - w - tr^{2/3}, r + w + tr^{2/3})$ for $t \in [-1, 1]$. Then, for all $t \in [-1, 1]$,

$$\begin{aligned} M_{\text{low}}(v) + M_{\text{up}}(v + (1,0)) &\geq \left[\mu r - \frac{G(w + tr^{2/3})^2}{r} - H\frac{(w + tr^{2/3})^4}{r^3} \right] \\ &\quad + \left[\mu r - \frac{G(w - tr^{2/3})^2}{r} - H\frac{(w - tr^{2/3})^4}{r^3} \right] \\ &\geq 2\mu r - \frac{2Gw^2}{r} - 2Gt^2 r^{1/3} - 32HK^4 r^{1/3}, \end{aligned}$$

the last inequality since $|w \pm tr^{2/3}| \leq 2Kr^{5/6}$. Since $t \in [-1, 1]$, $2Gt^2 r^{1/3} \leq 2Gr^{1/3}$, and so the right hand side of (A.5) is at most $X_{\phi_{\text{up}}} - \mathbb{E}[X_{\phi_{\text{up}}}] + \frac{\theta}{4} r^{1/3}$ for all large enough θ (depending on K). Thus, recalling the definition of E_{up},

$$\mathbb{P}\left(\widetilde{Z} > \theta r^{1/3}, E_{\text{up}} \right) \leq \mathbb{P}\left(X_{\phi_{\text{up}}} - \mathbb{E}[X_{\phi_{\text{up}}}] > \frac{\theta}{4} r^{1/3}) \right) \leq \exp(-c\theta^{\alpha}).$$

We now claim that E_{up} has probability at least $1/2$; since E_{up} is independent of $\widetilde{Z}$, this will imply that $\mathbb{P}(\widetilde{Z} > \theta r^{1/3}) \leq 2\exp(-c\theta^{\alpha})$. The proof of the claim is straightforward using the independence of u^* with the environment above $\mathbb{L}_{\text{up}}$ and Assumption 3b, for

$$\mathbb{P}\left(E_{\text{up}}^c \right) \leq \sup_{v \in \mathbb{L}_{\text{up}}} \mathbb{P}\left(X_{v+(1,0)} - M_{\text{up}}(v + (1,0)) \leq -\frac{\theta}{2} r^{1/3} \right) \leq 1/2,$$

for all θ larger than an absolute constant. □

Proof of Proposition A.1. We will prove the bound for the event that $\Gamma_r^z(r/2) > z/2 + sr^{2/3}$, as the event that it is less than $-z/2 - sr^{2/3}$ is symmetric.

For $j \in [\![0, r^{1/3}]\!]$, let I_j be the interval

$$\left(\frac{r}{2} - \frac{z}{2} - sr^{2/3}, \frac{r}{2} + \frac{z}{2} + sr^{2/3} \right) - [j, j+1] \cdot (r^{2/3}, -r^{2/3}).$$

Let A_j be the event that Γ_r^z passes through I_j, for $j \in [\![0, r^{1/3}]\!]$. Observe that

$$\left\{ \Gamma_r^z(r/2) > z/2 + sr^{2/3} \right\} \subseteq \bigcup_{j=0}^{r^{1/3}} A_j. \tag{A.6}$$

We claim that $\mathbb{P}(A_j) \le \exp(-c(s+j)^{2\alpha})$ for each such j; this will imply Proposition A.1 by a union bound which we perform at the end.

Let $Z_j^{(1)} = X_{(1,1),I_j}$ and $Z_j^{(2)} = X_{I_j,(r-z,r+z)}$. Also, let $\widetilde{Z}_j^{(1)} = \sup_{v\in I_j}(X_v - \mathbb{E}[X_v])$, and define $\widetilde{Z}_j^{(2)}$ analogously.

We have to bound the probability of A_j. The basic idea is to show that any path from $(1,1)$ to $(r-z, r+z)$ which passes through I_j suffers a weight loss greater than that which X_r^z typically suffers (which is of order Gz^2/r), and so such paths are not competitive. When j is very large, it is possible to show this even if we do not have the sharp coefficient of G for the parabolic loss; but for smaller values of j, we will need to be very tight with the coefficient of the parabolic loss. So we divide into two cases, depending on the size of j, and first address the case when j is large (in a sense to be specified more precisely shortly). Observe that, for a $c_2 > 0$ to be fixed,

$$\mathbb{P}(A_j) \le \mathbb{P}\left(X_r^z < \mathbb{E}[X_r^z] - c_2(s+j)^2 r^{1/3}\right) + \mathbb{P}\left(Z_j^{(1)} + Z_j^{(2)} > \mathbb{E}[X_r^z] - c_2(s+j)^2 r^{1/3}\right);$$

the first term is bounded by $\exp(-c(s+j)^{2\alpha})$ by Assumption 3b for a c depending on c_2, and we must show a similar bound for the second. Note that the second term is bounded by

$$\mathbb{P}\left(Z_j^{(1)} + Z_j^{(2)} > \mu r - \frac{Gz^2}{r} - Hr^{1/3} - c_2(s+j)^2 r^{1/3}\right), \tag{A.7}$$

using Assumption 2 and since $|z| \le r^{5/6}$.

Recall from (9.21) and Lemma 9.10 that there exists a $\lambda \in (0,1)$ such that, for $|z/2 + (s+j)r^{2/3}| \le r$, and $i = 1$ and 2,

$$\mathbb{P}\left(Z_j^{(i)} > \nu_{i,j} + \theta r^{1/3}\right) \le \exp(-c\theta^\alpha), \tag{A.8}$$

where $\nu_{i,j} = \frac{1}{2}\mu r - \lambda \cdot \frac{G}{r/2} \cdot (\frac{1}{2}z \pm (s+j)r^{2/3})^2$ with the $+$ for $i = 1$ and $-$ for $i = 2$; $\nu_{i,j}$ captures the typical weight of these paths. Note that we are very crude with the parabolic coefficient, but the bound (A.8) holds for all j; and also that we measure the deviation from the same expression $\nu_{i,j}$ (which is obtained by evaluating (9.21) at one endpoint) for all points in the interval. As we will see, comparing the full interval to a single point will not work for the second case of small j.

We want to show that the typical weight $\nu_{1,j} + \nu_{2,j}$ is much lower than $\mu r - Gz^2/r$. Simple algebraic manipulations show that, if $(s+j)r^{2/3} > (\lambda^{-1} - 1)^{1/2} r^{5/6}$ (which is the largeness condition on j defining the first case),

$$\sum_{i=1}^{2} \nu_{i,j} < \mu r - \lambda \frac{Gz^2}{r} - (1-\lambda)Gr^{2/3} - 3\lambda G(s+j)^2 r^{1/3} < \mu r - \frac{Gz^2}{r} - 3\lambda G(s+j)^2 r^{1/3},$$

the final inequality since $|z| \le r^{5/6}$. We have to bound (A.7) with some value of c_2, and we take it to be $2\lambda G$; note that any bound we prove on (A.7) will still be true if we later further lower c_2. The previous displayed bound shows that, for $(s+j)r^{2/3} > (\lambda^{-1} - 1)^{1/2} r^{5/6}$,

$$\mathbb{P}\left(Z_j^{(1)} + Z_j^{(2)} > \mu r - \frac{Gz^2}{r} - Hr^{1/3} - c_2(s+j)^2 r^{1/3}\right)$$

$$\leq \mathbb{P}\left(Z_j^{(1)} + Z_j^{(2)} > \nu_{1,j} + \nu_{2,j} + \tfrac{1}{2}\lambda G(s+j)^2 r^{1/3}\right).$$

In the inequality we absorbed $-Hr^{1/3}$ into the last term by imposing that s is large enough, depending on λ, G, and H. Now by a union bound and (A.8), the last display, and hence (A.7), is bounded by $2\exp(-c(s+j)^{2\alpha})$.

Now we address the other case that $(s+j)r^{2/3} \leq (\lambda^{-1}-1)^{1/2}r^{5/6}$. Thus I_j is close to the interpolating line, and we need a bound on the interval-to-interval weight with a much sharper parabolic term than in the previous case. Here above approach of the first case faces an issue. Since the gradient of Gz^2/r at z is $2Gz/r$, the weight difference across an interval of length $r^{2/3}$ at antidiagonal displacement z is of order $z/r^{1/3}$, which is much larger than the bearable error of $O(r^{1/3})$ when z is, say, $r^{5/6}$; so the crude approach of using the same expression (which we need to be less than $\mu r - Gz^2/r$) for the typical weight of all points in the interval, as we did in the first case, is insufficient—to have a single expression for which a tail bound exists for all points in the interval, we must necessarily include the linear gain of moving across the interval in the expression, and this will force it above $\mu r - Gz^2/r$. So, for this case, we will use Lemma A.2, which avoids the problem by taking the supremum after centering by the point-specific expectation.

Let $X'_v = X_{v,(r-z,r+z)}$. Now we observe

$$\mathbb{P}(A_j) \leq \mathbb{P}\left(X_r^z < \mathbb{E}[X_r^z] - c_2(s+j)^2 r^{1/3}\right) + \mathbb{P}\left(\sup_{v\in I_j}(X_v + X'_v) > \mathbb{E}[X_r^z] - c_2(s+j)^2 r^{1/3}\right);$$

note that $X_v + X'_v$ counts the weight of v twice, but this is acceptable as this sum dominates the weight of the best path through v. The first term is at most $\exp(-c(s+j)^{2\alpha})$ for a $c > 0$ depending on c_2. We bound the second term as follows. First we note that $\mathbb{E}[X_r^z] \geq \mu r - Gz^2/r - Hr^{1/3}$ and that $\sup_{v\in I_j}(\mathbb{E}[X_v + X'_v]) \leq \mu r - Gz^2/r - G(s+j)^2 r^{1/3}$ by a simple calculation with Assumption 2, and so

$$\mathbb{P}\left(\sup_{v\in I_j}(X_v + X'_v) > \mathbb{E}[X_r^z] - c_2(s+j)^2 r^{1/3}\right)$$
$$\leq \mathbb{P}\left(\sup_{v\in I_j}(X_v - \mathbb{E}[X_v] + X'_v - \mathbb{E}[X'_v]) > -Hr^{1/3} + (G-c_2)(s+j)^2 r^{1/3}\right).$$

We lower c_2 (if required) from its earlier value to be less than $G/2$. Now, we need to absorb the $-Hr^{1/3}$ term above into the $(s+j)^2 r^{1/3}$ term, which we can do for $s > s_0$ by setting s_0 large enough depending on G and H. So for such s, by a union bound we see that the previous display is at most

$$\mathbb{P}\left(\sup_{v\in I_j}(X_v - \mathbb{E}[X_v]) > \tfrac{1}{6}G(s+j)^2 r^{1/3}\right) + \mathbb{P}\left(\sup_{v\in I_j}(X'_v - \mathbb{E}[X'_v]) > \tfrac{1}{6}G(s+j)^2 r^{1/3}\right).$$

We bound this by applying Lemma A.2, with $K = (\lambda^{-1} - 1)^{1/2}$ and $\theta = \frac{1}{6}G(s+j)^2$. Recall that the bound of Lemma A.2 holds for $\theta > \theta_0(K)$. Thus we raise s_0 further if necessary so that $(s+j)^2 > \theta_0(K)$ for all $s > s_0$ and $j \geq 0$. Then we see that, for s and j such that $s > s_0$ and $(s+j)r^{2/3} \leq (\lambda^{-1} - 1)r^{5/6}$, the last display is at most $2\exp(-c(s+j)^{2\alpha})$.

Returning to the inclusion (A.6) and the bound of $\exp(-c(s+j)^{2\alpha})$ of $\mathbb{P}(A_j)$ for the two cases, we see that

$$\mathbb{P}\left(\Gamma_r^z(r/2) > z/2 + sr^{2/3}\right) \leq \sum_{j=1}^{r^{1/3}} \exp(-c(s+j)^{2\alpha}) \leq C\exp(-cs^{2\alpha})$$

for some absolute constant $C < \infty$ and $c > 0$ depending on α. Here we used that, if $\alpha \in (0, 1/2)$, then $(s+j)^{2\alpha} \geq 2^{2\alpha-1}(s^{2\alpha} + j^{2\alpha})$, while if $\alpha \geq 1/2$, then $(s+j)^{2\alpha} \geq s^{2\alpha} + j^{2\alpha}$; and finally $\exp(-cj^{2\alpha})$ is summable over j. This completes the proof of Proposition A.1. □

To extend the transversal fluctuation bound from the midpoint (as in Proposition A.1) to anywhere along the geodesic (as in Proposition 7.15), we follow very closely a multiscale argument previously employed in [BSS14, Theorem 11.1] and [BGHH20, Theorem 3.3] for similar purposes. For this reason, we will not write a detailed proof but only outline the idea.

Proof sketch of Proposition 7.15. First, the interpolating line is divided up into dyadic scales, indexed by j. The j^{th} scale consists of $2^j + 1$ anti-diagonal intervals, placed at separation $2^{-j}r$, of length of order $s_j r^{2/3} := \prod_{i=1}^{j}(1 + 2^{-i/3})sr^{2/3}$. By choosing the maximum j for which this is done large enough, it can be shown that, on the event that $\mathrm{TF}(\Gamma_r^z) > sr^{2/3}$, there must be a j such that there is a pair (I_1, I_3) of consecutive intervals on the j^{th} scale, and the interval I_2 of the $(j+1)^{\text{th}}$ scale in between such that the following holds: the geodesic passes through I_1 and I_3, but fluctuates enough that it avoids I_2, say by passing to its left.

Planarity and that the geodesic is a weight-maximising path then implies that the geodesic from the left endpoint of I_1 to that of I_3 is to the left of the geodesic Γ_r^z (this observation is often called geodesic or polymer ordering), and so must have midpoint transversal fluctuation at least of order $(s_{j+1} - s_j)r^{2/3} = 2^{-(j+1)/3}sr^{2/3}$. But since this transversal fluctuation happens across a scale of length $r' = 2^{-j}r$, in scaled coordinates it is of order $2^{j/3}s(r')^{2/3}$. Applying Proposition A.1 says that this probability is at most $\exp(-c2^{2\alpha j/3}s^{2\alpha})$. Now it remains to take a union bound over all the scales and the intervals within each scale. Since the number of intervals in the j^{th} scale is 2^j, and since $2^j \exp(-c2^{2\alpha j/3}s^{2\alpha}) \leq 2^{-j}\exp(-cs^{2\alpha})$ for all $s \geq s_0$ (by setting s_0 large enough) and $j \geq 1$, we obtain the overall probability bound of $\exp(-cs^{2\alpha})$ of Proposition 7.15. □

Bibliography

[ACQ11] Gideon Amir, Ivan Corwin, and Jeremy Quastel. Probability distribution of the free energy of the continuum directed random polymer in 1+ 1 dimensions. *Communications on pure and applied mathematics*, 64(4):466–537, 2011.

[AD14] Antonio Auffinger and Michael Damron. A simplified proof of the relation between scaling exponents in first-passage percolation. *The Annals of Probability*, 42(3):1197–1211, 2014.

[ADH17] Antonio Auffinger, Michael Damron, and Jack Hanson. *50 years of first-passage percolation*, volume 68. American Mathematical Soc., 2017.

[AKQ14] Tom Alberts, Konstantin Khanin, and Jeremy Quastel. The continuum directed random polymer. *Journal of Statistical Physics*, 154(1-2):305–326, 2014.

[Ale97] Kenneth S Alexander. Approximation of subadditive functions and convergence rates in limiting-shape results. *The Annals of Probability*, 25(1):30–55, 1997.

[Ale20] Kenneth S Alexander. Geodesics, bigeodesics, and coalescence in first passage percolation in general dimension. *arXiv preprint arXiv:2001.08736*, 2020.

[AVM05] Mark Adler and Pierre Van Moerbeke. PDEs for the joint distributions of the Dyson, Airy and sine processes. *The Annals of Probability*, 33(4):1326–1361, 2005.

[Bar01] Yu Baryshnikov. GUEs and queues. *Probability Theory and Related Fields*, 119(2):256–274, 2001.

[BDJ99] Jinho Baik, Percy Deift, and Kurt Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *Journal of the American Mathematical Society*, 12(4):1119–1178, 1999.

[BDM+01] Jinho Baik, Percy Deift, Kenneth DT McLaughlin, Peter Miller, and Xin Zhou. Optimal tail estimates for directed last passage site percolation with geometric random variables. *Advances in Theoretical and Mathematical Physics*, 5(6):1–41, 2001.

[BF08] Alexei Borodin and Patrik Ferrari. Large time asymptotics of growth models on space-like paths I: PushASEP. *Electronic Journal of Probability*, 13:1380–1418, 2008.

[BG16] Alexei Borodin and Vadim Gorin. Moments match between the KPZ equation and the Airy point process. *SIGMA. Symmetry, Integrability and Geometry: Methods and Applications*, 12:102, 2016.

[BG18] Riddhipratim Basu and Shirshendu Ganguly. Time correlation exponents in last passage percolation. *arXiv preprint arXiv:1807.09260*, 2018.

[BGH19] Riddhipratim Basu, Shirshendu Ganguly, and Alan Hammond. Fractal geometry of Airy_2 processes coupled via the Airy sheet. *arXiv preprint arXiv:1904.01717*, 2019.

[BGHH20] Riddhipratim Basu, Shirshendu Ganguly, Alan Hammond, and Milind Hegde. Interlacing and scaling exponents for the geodesic watermelon in last passage percolation. *arXiv preprint arXiv:2006.11448*, 2020.

[BGHK19] Riddhipratim Basu, Shirshendu Ganguly, Milind Hegde, and Manjunath Krishnapur. Lower deviations in β-ensembles and law of iterated logarithm in last passage percolation. *Israel Journal of Mathematics*, 2019. To appear.

[BGS17] Riddhipratim Basu, Shirshendu Ganguly, and Allan Sly. Upper tail large deviations in first passage percolation. *arXiv preprint arXiv:1712.01255*, 2017.

[BGS19] Riddhipratim Basu, Shirshendu Ganguly, and Allan Sly. Delocalization of polymers in lower tail large deviation. *Communications in Mathematical Physics*, 370(3):781–806, 2019.

[BGZ19] Riddhipratim Basu, Shirshendu Ganguly, and Lingfu Zhang. Temporal correlation in last passage percolation with flat initial condition via Brownian comparison. *arXiv preprint arXiv:1912.04891*, 2019.

[BH19] Gerandy Brito and Christopher Hoffman. Geodesic rays and exponents in ergodic planar first passage percolation. *arXiv preprint arXiv:1912.06338*, 2019.

[BHS18] Riddhipratim Basu, Christopher Hoffman, and Allan Sly. Nonexistence of bigeodesics in integrable models of last passage percolation. *arXiv preprint arXiv:1811.04908*, 2018.

[BKS03] Itai Benjamini, Gil Kalai, and Oded Schramm. First passage percolation has sublinear distance variance. *The Annals of Probability*, 31(4):1970–1978, 2003.

[BLM13] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford University Press, 2013.

[BOO00] Alexei Borodin, Andrei Okounkov, and Grigori Olshanski. Asymptotics of Plancherel measures for symmetric groups. *Journal of the American Mathematical Society*, 13(3):481–515, 2000.

[BR08] Michel Benaim and Raphaël Rossignol. Exponential concentration for first passage percolation through modified Poincaré inequalities. In *Annales de l'IHP Probabilités et statistiques*, volume 44, pages 544–573, 2008.

[BSS14] Riddhipratim Basu, Vladas Sidoravicius, and Allan Sly. Last passage percolation with a defect line and the solution of the slow bond problem. *arXiv preprint arXiv:1408.3464*, 2014.

[BSS19] Riddhipratim Basu, Sourav Sarkar, and Allan Sly. Coalescence of geodesics in exactly solvable models of last passage percolation. *Journal of Mathematical Physics*, 60(9):093301, 2019.

[CD81] J Theodore Cox and Richard Durrett. Some limit theorems for percolation processes with necessary and sufficient conditions. *Annals of Probability*, 9(4):583–603, 1981.

[CD13] Sourav Chatterjee and Partha S Dey. Central limit theorem for first-passage percolation time across thin cylinders. *Probability Theory and Related Fields*, 156(3-4):613–663, 2013.

[CD18] Ivan Corwin and Evgeni Dimitrov. Transversal fluctuations of the ASEP, stochastic six vertex model, and Hall-Littlewood Gibbsian line ensembles. *Communications in Mathematical Physics*, 363(2):435–501, 2018.

[CG18] Ivan Corwin and Promit Ghosal. KPZ equation tails for general initial data. *arXiv preprint arXiv:1810.07129*, 2018.

[CGH19] Ivan Corwin, Promit Ghosal, and Alan Hammond. KPZ equation correlations in time. *arXiv preprint arXiv:1907.09317*, 2019.

[CH14] Ivan Corwin and Alan Hammond. Brownian Gibbs property for Airy line ensembles. *Inventiones Mathematicae*, 195(2):441–508, 2014.

[CH16] Ivan Corwin and Alan Hammond. KPZ line ensemble. *Probability Theory and Related Fields*, 166(1-2):67–185, 2016.

[Cha13] Sourav Chatterjee. The universal relation between scaling exponents in first-passage percolation. *Annals of Mathematics*, pages 663–697, 2013.

[CHH19] Jacob Calvert, Alan Hammond, and Milind Hegde. Brownian structure in the KPZ fixed point, 2019.

[CHHM21] Ivan Corwin, Alan Hammond, Milind Hegde, and Konstantin Matetski. Exceptional times when the KPZ fixed point violates Johansson's conjecture on maximizer uniqueness. *arXiv preprint arXiv:2101.04205*, 2021.

[CIW19a] Pietro Caputo, Dmitry Ioffe, and Vitali Wachtel. Confinement of Brownian polymers under geometric area tilts. *Electronic Journal of Probability*, 24, 2019.

[CIW19b] Pietro Caputo, Dmitry Ioffe, and Vitali Wachtel. Tightness and line ensembles for Brownian polymers under geometric area tilts. *arXiv preprint arXiv:1906.06533*, 2019.

[Com17] Francis Comets. *Directed polymers in random environments*. Springer, 2017.

[Cor12] Ivan Corwin. The Kardar–Parisi–Zhang equation and universality class. *Random Matrices: Theory and Applications*, 1(01):1130001, 2012.

[CP15] Eric Cator and Leandro P. R. Pimentel. On the local fluctuations of last-passage percolation models. *Stochastic Process. Appl.*, 125(2):538–551, 2015.

[CS14] Ivan Corwin and Xin Sun. Ergodicity of the Airy line ensemble. *Electronic Communications in Probability*, 19, 2014.

[DH14] Michael Damron and Jack Hanson. Busemann functions and infinite geodesics in two-dimensional first-passage percolation. *Communications in Mathematical Physics*, 325(3):917–963, 2014.

[DHS14] Michael Damron, Jack Hanson, and Philippe Sosoe. Subdiffusive concentration in first passage percolation. *Electronic Journal of Probability*, 19, 2014.

[DHS15] Michael Damron, Jack Hanson, and Philippe Sosoe. Sublinear variance in first-passage percolation for general distributions. *Probability Theory and Related Fields*, 163(1-2):223–258, 2015.

[DM20] Evgeni Dimitrov and Konstantin Matetski. Characterization of Brownian Gibbsian line ensembles, 2020.

[DNV19] Duncan Dauvergne, Mihai Nica, and Bálint Virág. Uniform convergence to the Airy line ensemble. *arXiv preprint arXiv:1907.10160*, 2019.

[DOV18] Duncan Dauvergne, Janosch Ortmann, and Bálint Virág. The directed landscape. *arXiv preprint arXiv:1812.00309*, 2018.

[DSV20] Duncan Dauvergne, Sourav Sarkar, and Bálint Virág. Three-halves variation of geodesics in the directed landscape. *arXiv preprint arXiv:2010.12994*, 2020.

[DV18] Duncan Dauvergne and Bálint Virág. Basic properties of the Airy line ensemble. *arXiv preprint arXiv:1812.00311*, 2018.

[Dys62] Freeman J. Dyson. A Brownian-motion model for the eigenvalues of a random matrix. *J. Math. Phys.*, 3:1191–1198, 1962.

[DZ21] Duncan Dauvergne and Lingfu Zhang. Disjoint optimizers and the directed landscape. *arXiv preprint arXiv:2102.00954*, 2021.

[Ede61] Murray Eden. A two-dimensional growth process. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*, volume 4, pages 223–239. Univ of California Press Berkeley, 1961.

[EJS20] Elnur Emrah, Chris Janjigian, and Timo Seppäläinen. Right-tail moderate deviations in the exponential last-passage percolation. *arXiv preprint arXiv:2004.04285*, 2020.

[FO18] Patrik L Ferrari and Alessandra Occelli. Universality of the GOE Tracy-Widom distribution for TASEP with arbitrary particle density. *Electronic Journal of Probability*, 23, 2018.

[FO19] Patrik L Ferrari and Alessandra Occelli. Time-time covariance for last passage percolation with generic initial profile. *Mathematical Physics, Analysis and Geometry*, 22(1):1, 2019.

[FQR13] Gregorio Moreno Flores, Jeremy Quastel, and Daniel Remenik. Endpoint distribution of directed polymers in 1+ 1 dimensions. *Communications in Mathematical Physics*, 317(2):363–380, 2013.

[Gan20] Ujan Gangopadhyay. Fluctuations of transverse increments in two-dimensional first passage percolation. *arXiv preprint arXiv:2011.14686*, 2020.

[GH20a] Shirshendu Ganguly and Alan Hammond. The geometry of near ground states in Gaussian polymer models. *arXiv preprint arXiv:2010.05836*, 2020.

[GH20b] Shirshendu Ganguly and Alan Hammond. Stability and chaos in dynamical last passage percolation. *arXiv preprint arXiv:2010.05837*, 2020.

[GH20c] Shirshendu Ganguly and Milind Hegde. Optimal tail exponents in general last passage percolation via bootstrapping & geodesic geometry. *arXiv preprint arXiv:2007.03594*, 2020.

[GH21] Shirshendu Ganguly and Milind Hegde. Local and global comparisons of the Airy difference profile to Brownian local time. *arXiv preprint arXiv:2103.12029*, 2021.

[GIP12] Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski. Paracontrolled distributions and singular PDEs. *Forum of Mathematics, Pi, to appear*, 2012.

[GJ14] Patrícia Gonçalves and Milton Jara. Nonlinear fluctuations of weakly asymmetric interacting particle systems. *Archive for Rational Mechanics and Analysis*, 212(2):597–644, 2014.

[GP17] Massimiliano Gubinelli and Nicolas Perkowski. KPZ reloaded. *Communications in Mathematical Physics*, 349(1):165–269, 2017.

[Gra99] David J Grabiner. Brownian motion in a Weyl chamber, non-colliding particles, and random matrices. In *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, volume 35, pages 177–204. Elsevier, 1999.

[Gre74] Curtis Greene. An extension of Schensted's theorem. *Advances in Mathematics*, 14(2):254–265, 1974.

[GTW01] Janko Gravner, Craig A Tracy, and Harold Widom. Limit theorems for height fluctuations in a class of discrete space and time growth models. *Journal of Statistical Physics*, 102(5-6):1085–1132, 2001.

[Häg08] Jonas Hägg. Local Gaussian fluctuations in the Airy and discrete PNG processes. *Ann. Probab.*, 36(3):1059–1092, 2008.

[Hai13] Martin Hairer. Solving the KPZ equation. *Annals of Mathematics*, pages 559–664, 2013.

[Ham19a] Alan Hammond. Brownian regularity for the Airy line ensemble, and multi-polymer watermelons in Brownian last passage percolation. *Mem. Amer. Math. Soc., to appear*, 2019+.

[Ham19b] Alan Hammond. Exponents governing the rarity of disjoint polymers in Brownian last passage percolation. *Proc. Lond. Math. Soc., to appear.*, 2019+.

[Ham19c] Alan Hammond. Modulus of continuity of polymer weight profiles in Brownian last passage percolation. *Ann. Probab., to appear.*, 2019+.

[Ham19d] Alan Hammond. A patchwork quilt sewn from Brownian fabric: Regularity of polymer weight profiles in Brownian last passage percolation. *Forum of Mathematics, Pi*, 7:e2, 2019.

[HKPV09] John Ben Hough, Manjunath Krishnapur, Yuval Peres, and Bálint Virág. *Zeros of Gaussian analytic functions and determinantal point processes*, volume 51. American Mathematical Soc., 2009.

[HW65] John M Hammersley and Dominic JA Welsh. First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In *Bernoulli 1713, Bayes 1763, Laplace 1813*, pages 61–110. Springer, 1965.

[Jan18] Svante Janson. Tail bounds for sums of geometric and exponential variables. *Statistics & Probability Letters*, 135:1–6, 2018.

[Joh00] Kurt Johansson. Shape fluctuations and random matrices. *Communications in mathematical physics*, 209(2):437–476, 2000.

[Joh01] Kurt Johansson. Discrete orthogonal polynomial ensembles and the Plancherel measure. *Annals of Mathematics*, pages 259–296, 2001.

[Joh03] Kurt Johansson. Discrete polynuclear growth and determinantal processes. *Communications in Mathematical Physics*, 242(1-2):277–329, 2003.

[Kal02] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[KC18] Arun Kumar Kuchibhotla and Abhishek Chakrabortty. Moving beyond sub-gaussianity in high-dimensional statistics: Applications in covariance estimation and linear regression. *arXiv preprint arXiv:1804.02605*, 2018.

[Kes86] Harry Kesten. Aspects of first passage percolation. In *École d'été de probabilités de Saint Flour XIV-1984*, pages 125–264. Springer, 1986.

[KM59] Samuel Karlin and James McGregor. Coincidence probabilities. *Pacific J. Math.*, 9:1141–1164, 1959.

[KPZ86] Mehran Kardar, Giorgio Parisi, and Yi-Cheng Zhang. Dynamic scaling of growing interfaces. *Physical Review Letters*, 56(9):889, 1986.

[KS98] Ioannis Karatzas and Steven E Shreve. Brownian motion. In *Brownian Motion and Stochastic Calculus*, pages 47–127. Springer, 1998.

[Led18] Michel Ledoux. A law of the iterated logarithm for directed last passage percolation. *Journal of Theoretical Probability*, 31(4):2366–2375, 2018.

[LM01] Matthias Löwe and Franz Merkl. Moderate deviations for longest increasing subsequences: the upper tail. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 54(12):1488–1519, 2001.

[LMR02] Matthias Löwe, Franz Merkl, and Silke Rolles. Moderate deviations for longest increasing subsequences: the lower tail. *Journal of Theoretical Probability*, 15(4):1031–1047, 2002.

[LR10] Michel Ledoux and Brian Rider. Small deviations for beta ensembles. *Electron. J. Probab.*, 15:1319–1343, 2010.

[Mar06] James B Martin. Last-passage percolation with general weight distribution. *Markov Process. Related Fields*, 12(2):273–299, 2006.

[Mat99] Pertti Mattila. *Geometry of sets and measures in Euclidean spaces: fractals and rectifiability*. Number 44. Cambridge University Press, 1999.

[Meh91] Madan Lal Mehta. *Random matrices*. Academic Press Inc., Boston, MA, second edition, 1991.

[Mil78] PW Millar. A path decomposition for Markov processes. *Ann. Probab.*, 6(2):345–348, 1978.

[MQR17] Konstantin Matetski, Jeremy Quastel, and Daniel Remenik. The KPZ fixed point. *arXiv preprint arXiv:1701.00018*, 2017.

[Mue91] Carl Mueller. On the support of solutions to the heat equation with noise. *Stochastics: An International Journal of Probability and Stochastic Processes*, 37(4):225–245, 1991.

[NP95] Charles M Newman and Marcelo ST Piza. Divergence of shape fluctuations in two dimensions. *The Annals of Probability*, pages 977–1005, 1995.

[NQR20] Mihai Nica, Jeremy Quastel, and Daniel Remenik. One-sided reflected Brownian motions and the KPZ fixed point. In *Forum of Mathematics, Sigma*, volume 8. Cambridge University Press, 2020.

[Oko00] Andrei Okounkov. Random matrices and random permutations. *International Mathematics Research Notices*, 2000(20):1043–1095, 2000.

[OY02] Neil O'Connell and Marc Yor. A representation for non-colliding random walks. *Electronic communications in probability*, 7:1–12, 2002.

[Pim14] Leandro P.R. Pimentel. On the location of the maximum of a continuous stochastic process. *Journal of Applied Probability*, 51(1):152–161, 2014.

[PS02] Michael Prähofer and Herbert Spohn. Scale invariance of the PNG droplet and the Airy process. *Journal of Statistical Physics*, 108(5-6):1071–1106, 2002.

[QR13] Jeremy Quastel and Daniel Remenik. Local behavior and hitting probabilities of the Airy_1 process. *Probability Theory and Related Fields*, 157(3-4):605–634, 2013.

[QR14] Jeremy Quastel and Daniel Remenik. Airy processes and variational problems. In *Topics in percolative and disordered systems*, pages 121–171. Springer, 2014.

[QS20] Jeremy Quastel and Sourav Sarkar. Convergence of exclusion processes and KPZ equation to the KPZ fixed point. *arXiv preprint arXiv:2008.06584*, 2020.

[Rom15] Dan Romik. *The surprising mathematics of longest increasing subsequences*. Number 4. Cambridge University Press, 2015.

[RRV11] Jose Ramirez, Brian Rider, and Bálint Virág. Beta ensembles, stochastic Airy spectrum, and a diffusion. *Journal of the American Mathematical Society*, 24(4):919–944, 2011.

[RY13] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293. Springer Science & Business Media, 2013.

[Sep98a] Timo Seppäläinen. Coupling the totally asymmetric simple exclusion process with a moving interface. *Markov Process. Related Fields*, 4(4):593–628, 1998.

[Sep98b] Timo Seppäläinen. Large deviations for increasing sequences on the plane. *Probability theory and related fields*, 112(2):221–244, 1998.

[TW94] Craig A Tracy and Harold Widom. Level-spacing distributions and the Airy kernel. *Communications in Mathematical Physics*, 159(1):151–174, 1994.

[Ver18] Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge University Press, 2018.

[Vir20] Bálint Virág. The heat and the landscape I. *arXiv preprint arXiv:2008.07241*, 2020.

[Wai19] Martin J. Wainwright. *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2019.

[War07] Jon Warren. Dyson's Brownian motions, intertwining and interlacing. *Electronic Journal of Probability*, 12:573–590, 2007.

[Wil91] David Williams. *Probability with martingales*. Cambridge University Press, 1991.

[Wu19] Xuan Wu. Tightness of discrete Gibbsian line ensembles with exponential interaction Hamiltonians. *arXiv preprint arXiv:1909.00946*, 2019.