Title
Limit Theorems for Random Walk Local Time, Bootstrap Percolation and Permutation Statistics

Author
Lee, Sangchul

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Limit Theorems for Random Walk Local Time, Bootstrap Percolation and Permutation Statistics

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Mathematics

by

Sangchul Lee

2019

ABSTRACT OF THE DISSERTATION

Limit Theorems for Random Walk Local Time,

Bootstrap Percolation and Permutation Statistics

by

Sangchul Lee

Doctor of Philosophy in Mathematics

University of California, Los Angeles, 2019

Professor Marek Biskup, Co-Chair

Professor Georg Menz, Co-Chair

Limit theorems are established in three different contexts. The first one concerns exceptional points of the simple random walk in planar lattice domains approximating a given bounded continuum domain $D \subseteq \mathbb{R}^2$ with wired boundary conditions; the walk is run for a time proportional to the expected cover time. The sets of suitably defined thick, thin, light and avoided points are shown to be asymptotically distributed according to a log-normal multiple of the zero-average Liouville Quantum Gravity measure in D.

The second area of interest concerns the scaling limit of 2-neighbor polluted bootstrap percolation on $\mathbb{Z}^2$. Here each site is initially independently declared polluted with probability q, occupied with probability p, and vacant otherwise. At each step, each vacant site becomes occupied by contact with 2 or more occupied neighbors. It is shown that in the limit when $p, q \downarrow 0$ with $q/p^2 \to \lambda \in [0, \infty)$, the regime of small λ results in asymptotic full occupancy while the regime of large λ results in asymptotic full vacancy of the terminal

configuration. The proof is based on an identification of a continuum percolation model of "blocking contours" where these regimes correspond to absence and presence of ordinary percolation, respectively.

The last area of interest concerns the number of descents and peaks in a given conjugacy class of a random permutation of n elements. Asymptotic normality is proved proved in the limit $n \to \infty$ for suitably scaled versions of these quantities by establishing a uniform estimate on their moment generating functions.

The dissertation of Sangchul Lee is approved.

Thomas M. Liggett

Robijn Bruinsma

Marek Biskup, Committee Co-Chair

Georg Menz, Committee Co-Chair

University of California, Los Angeles

2019

Contents

ACKNOWLEDGEMENTS

First, I am extremely grateful to my advisor Professor Marek Biskup for introducing me to the wonderful world of probability theory and providing valuable advice on countless occasions. Thanks should also go to Professor Thomas M. Liggett for having been a role model and for serving on my PhD committee. I also wish to thank Professor Jason Fulman for encouraging me throughout our joint project and to Professors Robijn Bruinsma and Georg Menz for making the time to serve on my PhD committee. Also, I want to thank Hanbaek Lyu for discussions on the sharpness of the percolation threshold in the polluted bootstrap percolation.

Huge thanks go to Gene B. Kim and Sangjin Lee, whose friendship has been crucial in my life as an international student. Thanks go also to my friends Younghak Kwon, Dohyun Kwon, Wonjun Lee, and Bumsu Kim for sharing the experience of getting a PhD. Last but not the least, I would like to thank my family for their trust, support and encouragement.

Chapters 2 and 4 are based on completed research papers. Chapter 2 is a *verbatim* transcript of the preprint [2] by Abe, Biskup, and myself, which is a continuation of an earlier paper [1] by Abe and Biskup. My primary contribution in this work is the proof of the time conversion that resolves the pinning of the local time involved. The scaling limit of level sets of zero-average DGFF, the conversion to the case of arbitrary starting point and for discrete-time random walk and the control of the local structure have largely relied on the prior experience of my coauthors.

Section 4.1 in Chapter 4 is a reprinted version of the work [64] by Kim and myself. As one of the authors of this Elsevier article, I retain the right to include this paper in a thesis or dissertation, provided it is not published commercially. Section 4.2 is in turn a *verbatim* copy of the work [50] by Fulman, Kim, and myself. In both works, Fulman provided insight into exact generating functions that serve as the starting point of these works, Kim performed combinatoric computations, and I formulated the problem in probabilistic language and

carried out the proof thereof.

Some of the work reported on in this thesis has been done with financial support from the National Science Foundation grants DMS-1407558 and DMS-1712632.

VITA Fall 2019

SANGCHUL LEE

EDUCATION

Ph.D. candidate in Mathematics, Department of Mathematics, University of California Los Angeles, CA. Major field: Probability. Research interest: probability, statistical physics, analytic combinatorics. Advisor: Marek Biskup. (Aug. 2013 – present)

Bachelor of Art in Mathematics, Seoul National University. (Mar. 2006 – Feb. 2012)

PUBLICATIONS AND PREPRINTS

G. Kim, and S. Lee, Central limit theorems for descents in conjugacy classes of S_n, *J. Combin. Theory Ser. A* (2018).

Y. Abe, M. Biskup, and S. Lee Exceptional points of discrete-time random walks in planar domains, preprint (2019)

J. Fulman, G. Kim, and S. Lee, T. Kyle Petersen, On the joint distribution of descents and signs of permutations, submitted (2019).

J. Fulman, G. Kim, and S. Lee, Central limit theorem for peaks of a random permutation in a fixed conjugacy class of S_n, submitted (2019).

G. Kim, and S. Lee, A central limit theorem for descents and major indices in fixed conjugacy classes of S_n, preprint (2018).

PRESENTATIONS

(Nov. 2019) *Exceptional points of discrete-time random walks in planar domains*, Probability and Statistics Seminar, University of Southern California

(Mar. 2015) *Derivation of the Hartree equation from the N-body Schrödinger equation*, Analysis Seminar, University of Los Angeles

AWARDS AND FELLOWSHIPS

(2013–2018) Doctoral Study Abroad Program Fellowship of Korea Foundation for Advanced Studies

(2013–201) Radcliffe Dee Fellowship College Funds

(2012) Domestic Scholarship of Kwanjeong Educational Foundation

(2006–2011) National Science and Engineering Undergraduate Scholarship of Korea Student Aid Foundation

Chapter 1

Introduction

The limit theory has been a rich source of study in probability theory from the very beginning. The classical subjects of the limit theory are the central limit theorem, the Poisson convergence theorem and extreme value theory, with all of these appearing in a number of contexts and with numerous extensions and generalizations. In this thesis, we cover three topics in the limit theory, namely the structure of exceptional points of planar random walks, the pollution sensitivity of bootstrap percolation and the central limit theorem for permutation statistics in a conjugacy class. In all of these we start with a discrete structure and derive, via a scaling limit approach, a continuum limit object that admits an independent characterization.

1.1 Exceptional points of planar random walk

The *local time* of is an additive functional naturally associated with a random walk on a graph that is proportional to the total amount of time the walk spent at each given site. The local time is not only an indispensable tool for studying excursion lengths of the walk, but it also serves as a subject of independent interest. A natural question in this regard is the structure of the local-time configuration and its statistical properties. In the limit of large times, the additive structure implies that the typical behavior of the local time is governed by a central limit theorem. As it turns out, the limit process is the *discrete Gaussian free* *field* (DGFF), a subject of much independent interest in recent years.

A great deal of attention has been paid to exceptional points of the local times. For the simple random walk (SRW) on $\mathbb{Z}$, the behavior of the maximum local time is well-studied, starting from Kesten's work [62] on the law of iterated logarithm and culminating by a strong approximation via Brownian motion (see Révész [84] and references therein). For the SRW on $\mathbb{Z}^d$ with $d \geq 3$, Erdős and Taylor [44] analyzed both the maximum and the level sets of the local time by taking advantage of the transience of the walk.

The case of $d = 2$ where the walk is "barely" recurrent presents the most interesting questions. In the same work of Erdős and Taylor [44], the order of the maximum local time of the SRW on $\mathbb{Z}^2$ was first examined. More precisely, let $T_n(x)$ denote the number of visits to x by the walk run for time n. Then they showed that

$$\frac{1}{4\pi} \leq \liminf_{n\to\infty} \frac{\max_{x\in\mathbb{Z}^2} T_n(x)}{(\log n)^2} \leq \limsup_{n\to\infty} \frac{\max_{x\in\mathbb{Z}^2} T_n(x)}{(\log n)^2} \leq \frac{1}{\pi} \qquad \text{a.s.} \tag{1.1.1}$$

and conjectured that the limit exists with the value $1/\pi$. The resolution of this conjecture was given only four decades later by Dembo, Peres, Rosen and Zeitouni [30] through a strong approximation by Brownian motion and a multifractal analysis of the Brownian occupation measure. In the same paper, the number of so called α-thick points (where the local time T_n exceeds the α-multiple of the maximum $\frac{1}{\pi}(\log n)^2$) was shown to be asymptotically $n^{1-\alpha+o(1)}$ for each $\alpha \in (0, 1]$. These studies have recently been picked up by Jego [58], who extended these results to a more general setting and Okada [74], who examined the most-frequently visited sites on the inner boundary of the range of SRW in $\mathbb{Z}^2$.

In recent years, *isomorphism theorems* started to play an important role in the study of random fields. Generally speaking, these are statements of distribution identities that connect random fields to random walks, thereby allowing to reformulate field-theoretic questions to those of random walks and vice versa. Such isomorphisms have been successfully applied to the study of cover times and thick points of SRW [1],[38],[58], four-dimensional self avoiding walks [9],[21], ϕ^4-field theory [23],[22],[47], and random walk loop soups [67],[93]. In the case of local time, Dynkin [42] described a general connection between local times and Gaussian free fields, which bears his name. A version of this connection called the *Second Ray-Knight Theorem*, which first appeared in Eisenbaum, Kaspi, Marcus, Rosen and Shi [43], then provides a very precise distributional identity involving the local time and two copies of DGFF.

The Second Ray-Knight Theorem has proved to be very useful in various precise studies of two-dimensional random walks. For instance, Ding, Lee, and Peres [38] used it to show that the cover time of a SRW in a finite, connected graph (V, E) is concentrated near $|E|$ times the square of the expected maximum of DGFF in V. Cortines, Louidor, and Saglietti [26] pushed further in this direction and derived the full distributional limit for the cover time on binary trees; this was later revisited by Dembo, Rosen, and Zeitouni [32] with an alternative proof that does not involve isomorphism theorems. Belius and Kistler [11] identified the second order correction of the cover time for Brownian sausage. Then, in the study of local times, Abe employed the Second Ray-Knight Theorem to analyze the extrema and multifractal structure of the torus local times for the SRW run for a multiple of expected cover time, and to identify the structure of local-time local maxima for the SRW on a b-ary tree run for long enough.

Recently, Abe and Biskup [1] investigated the structure of exceptional points of SRW in lattice approximations of a planar domain D and run for the time proportional to the expected cover time. In their work, the Second Ray-Knight Theorem is utilized to relate the exceptional level-set measures of the local time to those of the DGFF. This in turn allowed to invoke the theory of intermediate level sets of DGFF, developed by Biskup and Louidor, [16], to identify the scaling limit of the geometry of exceptional points of local times. This work, however, deviates from the usual setting in that the results are formulated for the local time $\widehat{L}^D$ pinned at the wired boundary ϱ as opposed to the local time of the discrete-time random walk, due to the inherent pinning structure in the Second Ray-Knight Theorem.

In this thesis, we focus on establishing the convergence theorems for the exceptional points of the *discrete-time* random walk started at an *arbitrary point*. In the subsequent subsections, a basic theory on local times and DGFFs pertaining to our work will be reviewed and then the main results and an outline of the proof will be given.

1.1.1 Local times and discrete Gaussian free fields

We begin by reviewing the necessary concepts with the ultimate goal of stating the Second Ray-Knight Theorem and the scaling limit of intermediate level sets of DGFF.

General Theory of local times and DGFFs Let (V, E) be a finite, connected graph and $\varrho \in V$ be a designated vertex. Then we may consider a discrete-time Markov chain $X = (X_n)_{n\geq 0}$ in V which jumps to one of its neighbors with equal probabilities at each step. Denoting $\deg(V) := \sum_{v\in V}\deg(v)$, the function $\deg(\cdot)/\deg(V)$ is the unique invariant distribution of X under which X is reversible. Then the *(discrete) local time* of X at $v \in V$ and time $n \geq 0$ is defined by

$$\ell_n^V(v) := \frac{1}{\deg(v)}\sum_{k=0}^{n}\mathbf{1}_{\{X_k=v\}}, \tag{1.1.2}$$

which measures the total time spent by the walk at v run for time n, normalized by $\deg(v)$. This normalization scheme ensures that the expected increment of the local time during a single excursion is the same at all locations, providing a rationale for this definition.

In order to properly state the Second Ray-Knight Theorem, we now turn continuous-time random walk. Consider the Poisson process $\widetilde{N} = (\widetilde{N}(t))_{t\in[0,\infty)}$ of unit rate, independent of X, and let $\widetilde{X}_t = X_{\widetilde{N}(t)}$. This defines a continuous-time Markov chain $\widetilde{X} = (\widetilde{X}_t)_{t\in[0,\infty)}$ with independent unit-rate exponential holding times. Then, similarly as before, the *(continuous) local time* of $\widetilde{X}$ at $v \in V$ and time $t \geq 0$ is defined by

$$\widetilde{\ell}_t^V(v) := \frac{1}{\deg(v)}\int_0^t \mathbf{1}_{\{\widetilde{X}_s=v\}}\,\mathrm{d}s. \tag{1.1.3}$$

Denote P^v for the joint law of the chain X started at v and $\widetilde{N}$, and let $\widetilde{H}_\varrho$ be the first hitting time of $\widetilde{X}$ to ϱ. Then the *Green function* is defined by

$$G^V(x,y) := E^x\big[\widetilde{\ell}^V_{\widetilde{H}_\varrho}(y)\big]. \tag{1.1.4}$$

The following theorem highlights the role of Green function in the study of local time:

Theorem 1.1.1 (Kac's moment formula)**.** *Let $\widetilde{H}^+_\varrho$ be the first return time of $\widetilde{X}$ to ϱ. Then for any function $f : V \to \mathbb{R}$ with $f(\varrho) = 0$ and $n \in \mathbb{N}_1$,*

$$E^\varrho\big[\langle f, \widetilde{\ell}^V_{\widetilde{H}^+_\varrho}\rangle^n\big] = \frac{n!}{\deg(\varrho)}\langle \mathbf{1}, M_f(G^V M_f)^{n-1}\,\mathbf{1}\rangle \tag{1.1.5}$$

where $\langle\cdot,\cdot\rangle$ is the Euclidean inner product on $\mathbb{R}^V$, M_f is the multiplication operator by f, and $\mathbf{1}$ is the constant function with the value 1

Proof of Theorem 1.1.1 (Main idea). Abbreviate $\widetilde{l} := \widetilde{\ell}^V_{\widetilde{H}^+_\varrho}$ for brevity. Expanding the inner product,

$$E^\varrho\big[\langle f, \widetilde{l}\rangle^n\big] = n! \sum_{x_1,\cdots,x_n} E^\varrho\left[\int_{0<s_1<\cdots<s_n<\widetilde{H}^+_\varrho} \prod_{i=1}^n \frac{f(x_i)}{\deg(x_i)}\,\mathbf{1}_{\{\widetilde{X}_{s_i}=x_i\}}\,\mathrm{d}s_1\cdots\mathrm{d}s_n\right] \tag{1.1.6}$$

holds by symmetry, where the sum runs over all sites $x_1,\cdots,x_n$ in $V\setminus\{\varrho\}$, thanks to the condition $f(\varrho)=0$. Then by applying Markov property to the times $s_n,\cdots,s_1$ and writing $x_0=\varrho$ for convenience, the expectation evaluates to

$$\prod_{i=1}^n E^{x_{i-1}}\left[\int_{0<s_i<\widetilde{H}^+_\varrho} \frac{f(x_i)}{\deg(x_i)}\,\mathbf{1}_{\{\widetilde{X}_{s_i}=x_i\}}\,\mathrm{d}s_i\right] = \prod_{i=1}^n f(x_i)E^{x_{i-1}}\big[\widetilde{l}(x_i)\big] \tag{1.1.7}$$

The desired equality (1.1.5) now follows by noting that $E^\varrho\big[\widetilde{l}(x_1)\big] = \frac{1}{\deg(\varrho)}$ and $E^{x_{i-1}}\big[\widetilde{l}(x_i)\big] = G^V(x_{i-1},x_i)$ for each $i = 2,\cdots,n$. □

Continuing on the discussion of the Green function, G^V turns out to be a positive-definite symmetric operator on $\mathbb{R}^V$ (See Biskup [14, Chapter 1], for instance). This means that G^V can be used as the covariance kernel of a centered Gaussian process $h^V = (h^V_x)_{x\in V}$, i.e.,

$$\mathbb{E}[h^V_x] = 0 \quad \text{and} \quad \mathbb{E}[h^V_x h^V_y] = G^V(x,y), \qquad \forall x,y\in V. \tag{1.1.8}$$

The resulting process h^V is called the *discrete Gaussian free field (DGFF)* in V. Note that h^V is identically zero at ϱ.

Now we move on to defining the correct object for the Second Ray-Knight Theorem, Denote by $\widehat{\tau}_v(t) = \inf\{s \geq 0 : \widetilde{\ell}^V_s(v) \geq t\}$ the inverse local time at v. Then the *local time (pinned at ϱ)* is obtained by reparametrizing $\widetilde{\ell}^V$ by its value at ϱ:

$$\widehat{L}^V_t(v) := \widetilde{\ell}^V_{\widehat{\tau}_\varrho(t)}(v). \tag{1.1.9}$$

It is clear from the definition that $\widehat{L}^V_t(\varrho) = t$. In fact, the normalization ensures that $E^\varrho[\widehat{L}^V_t(v)] = t$ holds for each $v \in V$, see the case $n = 1$ of Theorem 1.1.1. Moreover, the fluctuation of $\widehat{L}^V_t$ around the level t is governed by a central limit theorem:

Proposition 1.1.2. *Let (V, E) be a finite, connected graph and $\varrho \in V$. Let $\widehat{L}^V$ be the pinned local time and h^V be the DGFF, defined as in* (1.1.8) *and* (1.1.9)*, respectively. Then*

$$\frac{\widehat{L}^V_t(\cdot) - t}{\sqrt{2t}} \text{ under } P^\varrho \xrightarrow[t\to\infty]{law} h^V(\cdot). \tag{1.1.10}$$

Proof. Under P^ϱ, the law of $\widehat{L}^V$ restricted to $V \setminus \{\varrho\}$ is identical to the compound Poisson process of rate $\deg(\varrho)$ whose jumps are identically distributed as the excursion $\widetilde{l} := \widetilde{\ell}^V_{\widetilde{H}_\varrho}$ around ϱ. From Kac's moment formula (1.1.5), for any function $f : V \to \mathbb{R}$ with $f(\varrho) = 0$,

$$\begin{aligned} E^\varrho\big[\mathrm{e}^{\langle f, \widehat{L}^V_t\rangle}\big] &= \exp\big\{\deg(\varrho)t\big(E^\varrho\big[\mathrm{e}^{\langle f,\widetilde{l}\rangle}\big] - 1\big)\big\} \\ &= \exp\big\{t\big\langle \mathbf{1}, M_f(1 - G^V M_f)^{-1}\,\mathbf{1}\big\rangle\big\}. \end{aligned} \tag{1.1.11}$$

From this, we get

$$E^\varrho\Big[\mathrm{e}^{\langle f,(\widehat{L}^V_t(\cdot)-t)/\sqrt{2t}\rangle}\Big] = \mathrm{e}^{\frac{1}{2}\langle f, G^V(1-G^V M_{f/\sqrt{2t}})^{-1} f\rangle} \xrightarrow[t\to\infty]{} \mathrm{e}^{\frac{1}{2}\langle f, G^V f\rangle}. \tag{1.1.12}$$

By the continuity theorem for moment generating functions, the desired convergence holds on $V \setminus \{\varrho\}$. This then extends to all of V, since both sides of (1.1.10) vanish at ϱ. □

Proposition 1.1.2 describes the structure of the typical points of $\widehat{L}^V_t$, and thus shifts the attention to the study of *exceptional points* where the value of $\widehat{L}^V_t$ lies well outside of the typical range $t+\mathcal{O}(\sqrt{t})$. Here we will need the aforementioned Second Ray-Knight Theorem that generalizes the connection between the local time $\widehat{L}^V$ and the DGFF h^V to a very precise form of distributional identity, albeit in a slightly indirect manner. The following version was first established by Eisenbaum, Kaspi, Marcus, Rosen and Shi [43] and then improved to an almost-sure identity by Zhai [100].

Theorem 1.1.3 (Second Ray-Knight Theorem)**.** *Let (V,E) be a finite, connected graph with a designated site $\varrho \in V$. Suppose that $\widehat{L}^V$ is the local time and h^V and $\widetilde{h}^V$ are DGFFs in V defined as above. Then for each $t \geq 0$, there exists a coupling of $\widehat{L}^V_t$, h^V and $\widetilde{h}^V$ such that $\widehat{L}^V_t$ and h^V are independent and*

$$\widehat{L}^V_t(v) + \frac{(h^V_v)^2}{2} = \frac{(\widetilde{h}^V_v + \sqrt{2t})^2}{2}, \qquad \forall\, v \in V. \tag{1.1.13}$$

Proof. Here we only prove the distributional identity. Assume that $\widehat{L}^V$ and h^V are independent. Fix a function $f : V \to \mathbb{R}$ with sufficiently small supremum norm so that $y(s) = \mathbb{E}\big[\mathrm{e}^{\frac{1}{2}\langle (h^V+s)^2, f\rangle}\big]$ is finite. Differentiating with respect to s,

$$y'(s) = \mathbb{E}\left[\langle h^V + s, f\rangle \mathrm{e}^{\frac{1}{2}\langle (h^V+s)^2, f\rangle}\right]. \tag{1.1.14}$$

We consider the following modification of the right-hand side of (1.1.14) with one f replaced by another arbitrary function $g : V \to \mathbb{R}$,

$$\mathbb{E}\left[\langle h^V + s, g\rangle \mathrm{e}^{\frac{1}{2}\langle (h^V+s)^2, f\rangle}\right] = s\,\langle \mathbf{1}, g\rangle\, \mathbb{E}\left[\mathrm{e}^{\frac{1}{2}\langle (h^V+s)^2, f\rangle}\right] + \mathbb{E}\left[\langle h^V, g\rangle \mathrm{e}^{\frac{1}{2}\langle (h^V+s)^2, f\rangle}\right]. \tag{1.1.15}$$

In light of the Gaussian integration-by-parts formula [14, Lemma 5.2]

$$\mathrm{Cov}(\phi(X), \chi(X)) = \sum_{i,j} \mathrm{Cov}(X_i, X_j)\mathbb{E}\left[\frac{\partial \phi}{\partial x_i}(X)\frac{\partial \chi}{\partial x_j}(X)\right] \tag{1.1.16}$$

which holds for all *linear* ϕ, all *sub-gaussian* χ, and all multivariate normals X, the second term on the right-hand side of (1.1.15) satisfies

$$\mathbb{E}\left[\langle h^V, g\rangle \mathrm{e}^{\frac{1}{2}\langle (h^V+s)^2, f\rangle}\right] = \mathbb{E}\left[\langle h^V + s, M_f G^V g\rangle \mathrm{e}^{\frac{1}{2}\langle (h^V+s)^2, f\rangle}\right]. \tag{1.1.17}$$

If we choose g so as to satisfy $g - M_f G^V g = f$, or equivalently, $g = (1 - M_f G^V)^{-1} f$, which again exists whenever f is sufficiently small, then (1.1.14), (1.1.15), and (1.1.17) together give

$$y'(s) = s\langle \mathbf{1}\,, (1 - M_f G^V)^{-1} f\rangle y(s). \tag{1.1.18}$$

Solving this differential equation with the initial condition $y(0) = \mathbb{E}\big[\mathrm{e}^{\frac{1}{2}\langle (h^V)^2, f\rangle}\big]$ yields

$$\mathbb{E}\left[\mathrm{e}^{\frac{1}{2}\langle (h^V+s)^2, f\rangle}\right] = y(s) = y(0)\mathrm{e}^{\frac{s^2}{2}\langle \mathbf{1}\,,(1-M_f G^V)^{-1} f\rangle} = \mathbb{E}\left[\mathrm{e}^{\frac{1}{2}\langle (h^V)^2, f\rangle}\right] \mathbb{E}\left[\mathrm{e}^{\langle \widehat{L}^V_{s^2/2}, f\rangle}\right], \tag{1.1.19}$$

where the last step follows from (1.1.11). Replacing s by $\sqrt{2t}$ proves the desired claim. □

Theorem 1.1.3 reveals the time scale at which the expected maximum of the DGFF $M := \mathbb{E}[\sup_{x\in V} \widetilde{h}^V_x]$ and $\sqrt{2t}$ become of the same order. For instance, if we assume that $t \approx \frac{1}{2}\theta M^2$ and that the term $h^V(\cdot)^2/2$ may be ignored from the identity (1.1.13), then the values of $\widehat{L}^V_t$ would lie roughly between $0 \vee (\sqrt{\theta} - 1)M^2$ and $(\sqrt{\theta} + 1)M^2$. In particular, $t \approx \frac{1}{2}M^2$ provides an *ansatz* for the asymptotic form of the expected cover time. As is shown in this thesis, this heuristics can be made rigorous to give a correct picture in many cases of interest, albeit not without a thorough understanding of the role of the DGFF term on the left-hand side of (1.1.13).

Exceptional points of DGFFs in planar domains Motivated by the idea from the previous section, we focus on the special kind of graphs arising as lattice approximations of a planar domain and review the theory of exceptional points of DGFF therein. We begin by clarifying the type of domains and lattice approximations to work with.

Definition 1.1.4. *A subset $D \subseteq \mathbb{R}^2$ is called an* admissible domain *if it is a bounded open subset of $\mathbb{R}^2$ and its boundary ∂D consists of finitely many connected components, each of which having a positive Euclidean diameter.*

The set of all admissible domains is denoted by $\mathfrak{D}$. Then for each $D \in \mathfrak{D}$, an *admissible lattice approximation* of D is a sequence $(D_N)_{N\geq 1}$ of subsets of $\mathbb{Z}^2$ such that, for each $\delta > 0$,

$$\{x \in \mathbb{Z}^2 : d_\infty({}^x\!/\!{}_N, D^c) > \delta\} \subseteq D_N \subseteq \{x \in \mathbb{Z}^2 : d_\infty({}^x\!/\!{}_N, D^c) > N^{-1}\} \tag{1.1.20}$$

holds for all large N, where d_∞ is the ℓ^∞-distance in $\mathbb{R}^2$. In what follows, we fix an admissible domain $D \in \mathfrak{D}$ and its admissible lattice approximations $(D_N)_{N\geq 1}$. We remark that this definition is chosen so as to guarantee that the discrete harmonic measures on D_N converge weakly to the continuum harmonic measure on D, as proved in Biskup and Louidor [17]. Indeed, this property is all that we require for the lattice approximations $(D_N)_{N\geq 1}$.

Now let h^{D_N} be a DGFF in D_N. The result of Bolthausen, Deuschel, and Giacomin [20] shows that the maximum of h^{D_N} is $(2\sqrt{g}+o(1))\log N$ in probability, where (in our normalization) $g = 1/(2\pi)$. Daviaud [28] showed that, for $\lambda \in (0,1)$, the number of λ-thick points of h^{D_N}, defined as the points where h^{D_N} is larger than the λ-multiple of $2\sqrt{g}\log N$, is $N^{2(1-\lambda^2)+o(1)}$ in probability. Recently, Biskup and Louidor [16] studied the geometry of the λ-thick points. To state their result, fix any centering sequence $(\widehat{a}_N)_{N\geq 1}$ satisfying $\widehat{a}_N \sim 2\lambda\sqrt{g}\log N$ for $\lambda \in (0,1)$ and define the point measure η_N^D on $D \times \mathbb{R}$ by

$$\eta_N^D := \frac{1}{K_N}\sum_{x\in D_N} \delta_{x/N} \otimes \delta_{h_x^{D_N}-\widehat{a}_N}, \qquad K_N := \frac{N^2}{\sqrt{\log N}}\mathrm{e}^{-\frac{(\widehat{a}_N)^2}{2g\log N}}. \tag{1.1.21}$$

Then it is shown that there exist a positive constant $\mathfrak{c}(\lambda) > 0$ and a unique a.s.-finite random measure Z_λ^D on D, called the *Liouville Quantum Gravity (LQG) measure*, such that

$$\eta_N^D \xrightarrow[N\to\infty]{\text{law}} \mathfrak{c}(\lambda) Z_\lambda^D(\mathrm{d}x) \otimes \mathrm{e}^{-\alpha\lambda h}\,\mathrm{d}h \tag{1.1.22}$$

in the sense of vague convergence of measures on $\overline{D} \times (-\infty,\infty]$ and $\alpha := 2/\sqrt{g}$. Moreover, $Z_\lambda^D(A) > 0$ a.s. for every non-empty open set $A \subseteq D$.

1.1.2 Key results and outline of the proof

We will now move to the discussion of some of the results on local-time exceptional points presented in this thesis. Fix an admissible domain $D \in \mathfrak{D}$ and its admissible lattice approximation $(D_N)_{N\geq 1}$. Also, we identify D_N with the graph obtained by collapsing the complement of D_N in $\mathbb{Z}^2$ into a single vertex ϱ while retaining all the edges adjacent to D_N. This 'wired boundary' consideration allows to relate the DGFF in D_N to the local time $\widehat{L}^{D_N}$

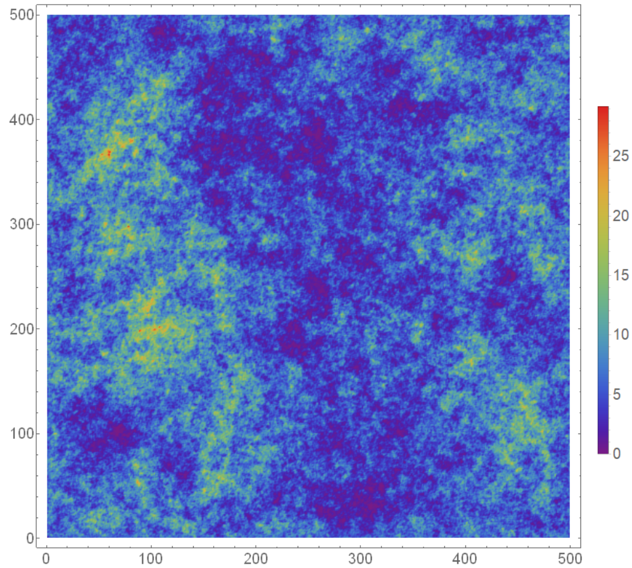


Figure 1.1: A simulation of pinned local time in the 500 by 500 square grid, run for about 1.6-multiple of the expected cover time

pinned at ϱ. In the parametrization by the local time at the "boundary vertex", the expected cover time is asymptotic to $2g(\log N)^2$, which sets a natural time scale for analyzing the exceptional points of the local time. Indeed, fix a parameter $\theta > 0$ and consider any sequence $(t_N)_{N\geq 1}$ satisfying

$$\lim_{N\to\infty} \frac{t_N}{(\log N)^2} = 2g\theta. \tag{1.1.23}$$

Also, define

$$L_t^{D_N}(x) := \ell^{D_N}_{\lfloor t \deg(D_N)\rfloor}(x), \qquad x \in D_N, t \geq 0 \tag{1.1.24}$$

where, abusing our earlier notation, $\deg(D_N) := \sum_{x\in D_N\cup\{\varrho\}} \deg(x)$. Then we have:

Theorem 1.1.5 (Theorem 2.2.1 in Chapter 2)**.** *For any choices of $x_N \in D_N$, the following limits hold in P^{x_N}-probability:*

$$\frac{1}{(\log N)^2} \max_{x\in D_N} L^{D_N}_{t_N}(x) \quad \underset{N\to\infty}{\longrightarrow} \quad 2g\big(\sqrt{\theta}+1\big)^2 \tag{1.1.25}$$

and

$$\frac{1}{(\log N)^2} \min_{x\in D_N} L^{D_N}_{t_N}(x) \quad \underset{N\to\infty}{\longrightarrow} \quad 2g\big[(\sqrt{\theta}-1)\vee 0\big]^2. \tag{1.1.26}$$

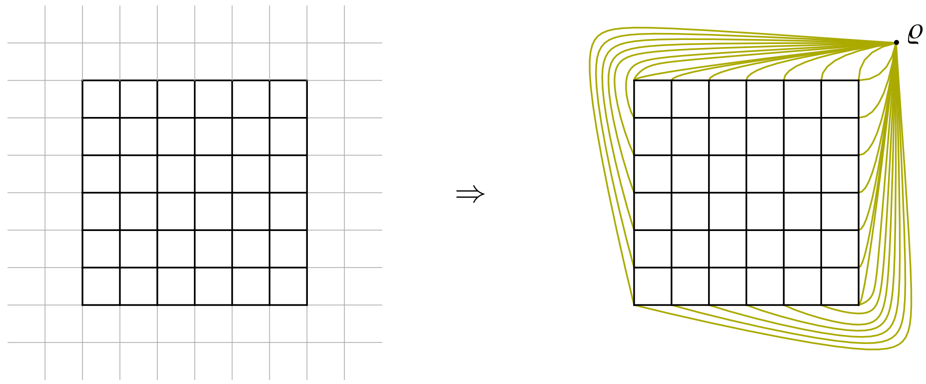


Figure 1.2: Example of the wired boundary condition

For each $\lambda \in (0,1)$, this permits to give a natural definition of λ-*thick* point of $L^{D_N}_{t_N}$ as the point where the value of $L^{D_N}_{t_N}$ exceeds $2g(\sqrt{\theta}+\lambda)^2(\log N)^2$. The structure of λ-think points of $L^{D_N}_{t_N}$ is then encoded via the point measure

$$\zeta^D_N := \frac{1}{W_N}\sum_{x\in D_N} \delta_{x/N}\otimes \delta_{(L^{D_N}_{t_N}(x)-a_N)/\sqrt{2a_N}}, \tag{1.1.27}$$

where $(a_N)_{N\geq 1}$ is any sequence satisfying $a_N \sim 2g(\sqrt{\theta}+\lambda)^2(\log N)^2$, and

$$W_N := \frac{N^2}{\sqrt{\log N}}\mathrm{e}^{-\frac{(\sqrt{2a_N}-\sqrt{2t_N})^2}{2g\log N}}. \tag{1.1.28}$$

To state the main results, let $Z^{D,0}_\lambda$ denote the zero-average Liouville Quantum Gravity measure, which is a unique measure such that

$$Z^{D,0}_\lambda \overset{\text{law}}{=} \mathrm{e}^{\alpha\lambda\mathfrak{d}(x)Y} Z^{D,0}_\lambda(\mathrm{d}x) \tag{1.1.29}$$

holds true with certain deterministic objects $\mathfrak{d}: D\to\mathbb{R}$ and $\sigma^2_D \in (0,\infty)$ – see (2.2.13) and (2.2.15) – and a normal random variable $Y = \mathcal{N}(0,\sigma^2_D)$ independent of $Z^{D,0}_\lambda$. The main theorem on the scaling limit of the λ-thick point measure then reads:

Theorem 1.1.6 (Theorem 2.2.3 in Chapter 2)**.** *Suppose* $(t_N)_{N\geq 1}$ *and* $(a_N)_{N\geq 1}$ *are positive sequences such that, for some* $\theta > 0$ *and some* $\lambda \in (0,1)$*,* (1.1.23) *and*

$$\lim_{N\to\infty}\frac{a_N}{(\log N)^2} = 2g(\sqrt{\theta}+\lambda)^2 \tag{1.1.30}$$

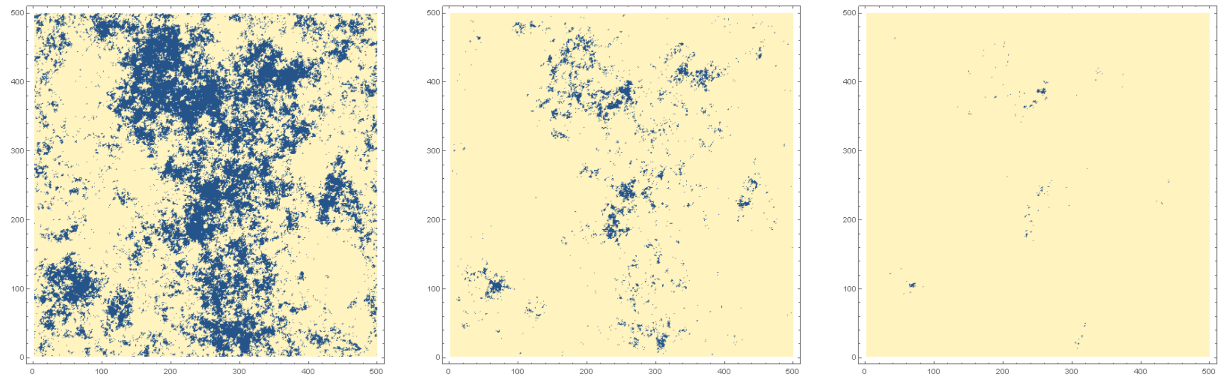


Figure 1.3: λ-thin points corresponding to $\lambda = 0.5$, 0.75, and 1, respectively, on the unit square $D = (0,1)^2$ with $N = 500$ and $\theta \approx 1.6$.

hold true. Then for any $x_N \in D_N$ *and for* X *sampled from* P^{x_N}*, the measures* ζ_N^D *in* (1.1.27) *with* W_N *as in* (1.1.28) *obey*

$$\zeta_N^D \quad \xrightarrow[N\to\infty]{law} \quad \sqrt{\frac{\sqrt{\theta}}{\sqrt{\theta}+\lambda}}\,\mathrm{e}^{-\alpha^2\lambda^2/16}\,\mathfrak{c}(\lambda)\,\mathrm{e}^{\alpha\lambda(\mathfrak{d}(x)-1)Y}\,Z_\lambda^{D,0}(\mathrm{d}x)\otimes \mathrm{e}^{-\alpha\lambda h}\mathrm{d}h \tag{1.1.31}$$

in the sense of vague convergence of measures on $\overline{D}\times(\mathbb{R}\cup\{+\infty\})$*,* $Y \sim \mathcal{N}(0,\sigma_D^2)$ *and* $Z_\lambda^{D,0}$ *are independent, and* $\mathfrak{c}(\lambda)$ *is as in* (1.1.22).

We give a brief sketch of the proof of Theorem 1.1.25. There are two main inputs that are involved in the argument. The first one is the *time conversion*, which allows to relate a deterministic time t_N to the inverse local time $\widehat{\tau}_\varrho(\cdot)$ shifted by the average fluctuation

$$U_N := \frac{1}{|D_N|}\sum_{x\in D_N}\big(\widehat{L}_{t_N}^{D_N}(x) - t_N\big) \tag{1.1.32}$$

of the pinned local time $\widehat{L}^{D_N}$, which is typically of the order $\sqrt{t_N}$ in view of Proposition 1.1.2. Heuristically, the definition of U_N can be rearranged so that $\widehat{\tau}_\varrho(t_N) \approx (t_N + U_N)\deg(D_N)$ holds. Now assuming that U_N is robust under a small perturbation of t_N, we may treat U_N as constant and then replace t_N by $t_N - U_N$ to get $\widehat{\tau}_\varrho\big(t - U_N\big) \approx t_N \deg(D_N)$. This heuristics can be justified at the cost of introducing a negligible error, yielding a statement roughly of the form

$$\widetilde{\ell}_{t_N \deg(D_N)}^{D_N} = \widehat{L}_{t_N - U_N + o(\sqrt{t_N})}^{D_N}. \tag{1.1.33}$$

To handle the random shift U_N in time variable, we need to establish a convergence statement for the λ-thick measure augmented by information on U_N. Following the argument established by Abe and Biskup [1], the relevant result is given by:

$$\begin{aligned}&\frac{1}{W_N}\sum_{x\in D_N}\delta_{x/N}\otimes\delta_{(\widehat{L}^{D_N}_{t_N}(x)-a_N)/\sqrt{a_N}}\otimes\delta_{U_N/\sqrt{2t_N}}\text{ under }P^{\varrho}\\&\qquad\xrightarrow[N\to\infty]{\text{law}}\sqrt{\frac{\sqrt{\theta}}{\sqrt{\theta}+\lambda}}\,\mathfrak{c}(\lambda)\mathrm{e}^{\alpha\lambda\mathfrak{d}(x)Y}Z^{D,0}_{\lambda}(\mathrm{d}x)\otimes\mathrm{e}^{-\alpha\lambda h}\mathrm{d}h\otimes\delta_Y,\end{aligned}\tag{1.1.34}$$

This and the time conversion altogether then permit to replace $\widehat{L}^{D_N}_{t_N}$ in the convergence statement by the continuous local time $\widetilde{L}^{D_N}_{t_N}:=\widetilde{\ell}^{D}_{t_N\deg(D_N)}$ at the cost of introducing an explicit prefactor to the limiting measure.

The next input is the representation of $\widetilde{L}^{D_N}_{t_N}$ via the discrete local time $L^{D_N}_{t_N}$ together with the exponential holding times. Indeed, the exponential holding times of $\widetilde{X}_t=X_{\widetilde{N}(t)}$ may be relabeled, conditional on the sample of X, using a family of independent unit exponential variables $(\tau_j(x))_{j\in\mathbb{N},x\in D_N\cup\{\varrho\}}$, such that the continuous local time admits the representation

$$\widetilde{\ell}^{D_N}_t(x)=\frac{1}{\deg(x)}\sum_{j\geq1}\tau_j(x)\,\mathbf{1}_{\{j\leq\deg(x)\ell^{D_N}_{\widetilde{N}(t)}(x)\}}\qquad\text{whenever}\quad x\neq\widetilde{X}_t.\tag{1.1.35}$$

In light of this, the correction term appearing in the convergence statement along the way of replacing $\widetilde{L}^{D_N}_{t_N}$ by $L^{D_N}_{t_N}$ takes the form

$$\frac{\widetilde{L}^{D_N}_{t_N}(x)-a_N}{\sqrt{2a_N}}-\frac{L^{D_N}_{t_N}(x)-a_N}{\sqrt{2a_N}}=\frac{1}{4\sqrt{2a_N}}\sum_{j\geq1}(\tau_j(x)-1)\,\mathbf{1}_{\{j\leq4L^{D_N}_{t_N}(x)\}}\tag{1.1.36}$$

at all x except at $x=\widetilde{X}_{t_N\deg(D_N)}$ where the walk is (with probability one) in the "middle" of the holding time. For large N, the right-hand may be approximated by a white noise of variance $\frac{1}{8}$ independent of $L^{D_N}_{t_N}$, whose effect can be resolved by invoking the convolution-identity technique introduced in Abe and Biskup's work [1].

The definition and convergence statement for λ-*thin* points is quite similar, with due modifications in the choices of signs and the vague topology. In addition to thick and thin points, we can also study the structure of rarely visited points, called the *light* points, as well as *avoided* points which are the points not visited by the walk. The structure of avoided points can be encoded by a suitably-normalized counting measure of the set $\{x/N \in D : x \in D_N \text{ and } L^{D_N}_{t_N}(x) = 0\}$, which can be shown to converge to

$$\sqrt{2\pi g}\,\mathfrak{c}(\sqrt{\theta})\,\mathrm{e}^{\alpha\sqrt{\theta}(\mathfrak{d}(x)-1)Y}\,Z^{D,0}_{\sqrt{\theta}}(\mathrm{d}x) \tag{1.1.37}$$

in the sense of vague convergence of measures on $\overline{D}$. Moreover, as in [1], the local structure of the exceptional sets can be described as well.

1.2 Pollution-sensitivity of bootstrap percolation

The next subject to be discussed in this thesis is the polluted bootstrap percolation. We start by putting the problem in a general context. Recall that a *cellular automaton* is a discrete-time process on a configuration space S^G, where S is a finite set and G is a connected graph, that evolves in time by local and homogeneous rules. First developed by von Neumann [98] and Ulam [95], cellular automata have been widely studied as reductionist models of many natural phenomena. In recent years, there has been considerable progress on the class of *monotone* cellular automata starting from *random* initial configurations, commonly referred to as *bootstrap percolation (BP)*.

Bootstrap percolation was first introduced and studied by Chalupa, Leath, and Reich [24] as a monotone analogue of Glauber dynamics and was subsequently revisited many times in various contexts. It may as well be considered as a monotone version of the *kinetically constrained model (KCM)*, which was introduced to study liquid-glass transition from a combinatorial perspective. This model is a stochastic cellular automaton with two states, where the rate for changing the state of a site depends on the local configuration near it. The common perspective on KCM is that, while thermodynamically uninteresting, it is rich in dynamic behavior, thus hinting that BP might behave quite differently from typical percolation models.

In the typical setting, BP is a cellular automaton with two states, *vacant* and *occupied*, on a graph (which is typically chosen as either the entire d-dimensional grid $\mathbb{Z}^d$ or a finite box $\Lambda_n := [1,n]^d \cap \mathbb{Z}^d$), with the following features:

(1) The initial configuration is random. More precisely, it is typically given by the product Bernoulli measure so that each site is either initially declared to be occupied with probability $p \in [0,1]$ or vacant otherwise, independently of each other.

(2) The update rule is monotone. It can be represented by a finite collection $\mathcal{U}$ of subsets of $\mathbb{Z}^d \setminus \{0\}$ such that each vacant site v becomes occupied exactly when all the sites in $v + X$ are occupied for some $X \in \mathcal{U}$. Occupied sites remain occupied forever.

Among such models, the most extensively studied is the *r-neighbor BP*, in which a site becomes occupied if it has at least r already-occupied neighbors. Another well-studied model is the *modified BP*, in which a site becomes occupied if it has an occupied neighbor in each coordinate direction.

A primary subject of interest in the study of BP is the *terminal configuration*, which is defined as the limit of the configurations under bootstrap dynamics as time tends to infinity; this limit exists by the monotonicity of the update rule. It turns out, however, that both the r-neighbor BP with $r \leq d$ and the modified BP on the full lattice $\mathbb{Z}^d$ of dimension $d \geq 2$ show only a trivial phase transition. Indeed, in both models, arbitrarily small initial density $p > 0$ results in full occupation of the terminal configuration. This was first proved for the 2-neighbor BP by van Enter [96] and for the modified BP by Schonmann [89].

The absence of "non-percolating" phase shifts focus to the question of finite-size effects. In this regard, it is well-known that BP on finite grids Λ_n demonstrates *metastability*. More precisely, there exists a certain function $\lambda(p,n)$ of p and n such that, as $\lambda(p,q)$ increases, the full occupation probability exhibits a sharp transition from being close to 0 to being close to 1 for sufficiently small p. Aizenman and Lebowitz [4] found 2-neighbor BP to be governed by $\lambda(p,n) = p^{1/(d-1)} \log n$ and successfully attributed this to the formation of *critical droplets* (which are the initially occupied clusters that almost certainly grow to fill the whole grid).

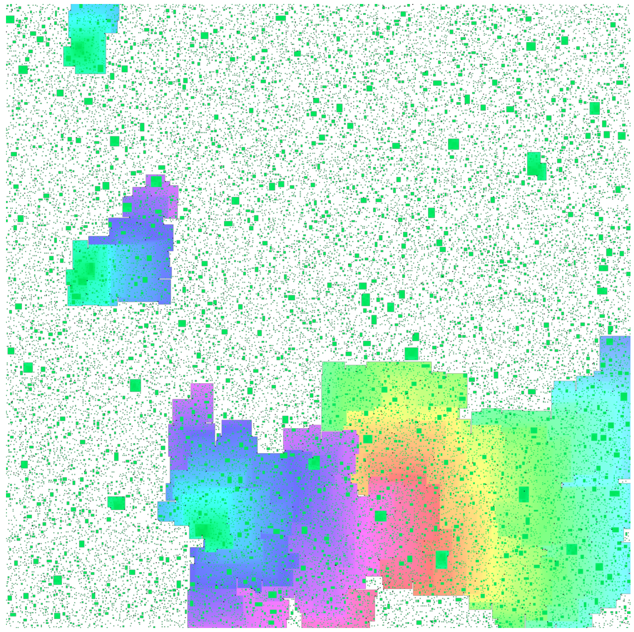


Figure 1.4: Simulation of the PBP with seed probability $p = 0.08$ and pollution density $q = 0.002$ on a square of 750×750 vertices. Occupied sites are colored according to the time of occupation, starting from green at $t = 0$ and in the order of the rainbow colors as time increases.

Then Holroyd [57] established the stated sharp transition as well as identified the exact value of the critical parameter λ_c for both the 2-neighbor and modified BP in $d = 2$, by showing that the full occupation of a square is almost likely to be driven by the scenario that critical droplets expand in all four axis directions. This idea was extended by Balogh, Bollobás, Duminil-Copin, and Morris [8] to the r-neighbor BP in all dimensions $d \geq 2$ and all r with $d \geq r \geq 2$.

The above discussion hints that BP is extremely sensitive to the choice of update rules, as a small modification of the rules may result in completely different long-term macroscopic behavior. In this part of the thesis, we examine yet another type of sensitivity by considering the *polluted bootstrap percolation (PBP)*. This is a generalization of BP introduced by Gravner and McDonald [55] to study the effect of permanent obstacles to growth.

In PBP, each site of $\mathbb{Z}^d$ is independently declared to be polluted with probability q, occupied with probability p, and vacant otherwise, and then the bootstrap dynamics is run on the subgraph of unpolluted sites. In this manner, polluted sites represent obstacles which neither get occupied nor can help other sites get occupied. As before, the primary subject of interest is the structure of the terminal configuration. One of the natural quantities in this regard is the density $\theta(p, q)$ of terminally occupied sites on $\mathbb{Z}^2$. In their work [55], Gravner and McDonald showed that $\theta(p, q)$ for the 2-neighbor PBP exhibits a phase transition in the limit as $p \downarrow 0$ with $\lambda(p, q) := q/p^2$ tending to a constant λ, such that full terminal density occurs when λ is small and vanishing terminal density occurs when λ is large. However, the question of existence of a critical parameter λ_c sharply separating these two regimes has been left open. For the modified PBP, even the correct scaling relationship $\lambda(p, q)$ has not fully been settled and was conjectured to involve a logarithmic correction.

In this part of the thesis, we develop a method for establishing a sharp transition in both the 2-neighbor and modified PBP on $\mathbb{Z}^2$ using the following idea: In both models, one can identify a large-scale *blocking structure* in the *initial* configuration that determines whether the *terminal* configuration will be densely occupied or not. This structure may be chosen to satisfy:

(1) The failure of occupied clusters to grow any further can be traced back to the formation of blocking structures, marking more or less, the occupied/unoccupied interfaces in the terminal configuration.

(2) As $p, q \downarrow 0$ with $\lambda(p, q)$ tending to a constant λ, the blocking structures admit a scaling limit in the form of a continuum percolation model in which λ represents a natural density parameter.

This permits relating the supercritical/subcritical regimes of the continuum model to the subcritical/supercritical regimes of PBP for small p, thereby reducing the question of a sharp transition in PBP to that in the limiting continuum percolation model. Using this idea, it will be shown that there exist two critical parameters $\lambda_1^c \leq \lambda_2^c$, originating from the boundaries of different phases of the associated continuum percolation model, that $\theta(p, q)$ is close to 1 in the regime $\lambda(p, q) < \lambda_1^c$, and is close to 0 in the regime $\lambda(p, q) > \lambda_2^c$, as $p \downarrow 0$. In particular, the blocking structure naturally determines the correct p *versus* q scale. In the modified PBP, the scale factor turns out to be $\lambda(p,q) = (q/p^2)\log(1/p)$, which is already novel and answers the conjecture of Gravner and McDonald [55].

1.2.1 Main results and conjecture

We first provide some definitions and then state the key results. We will mainly focus on the 2-neighbor PBP on $\mathbb{Z}^2$. Write $S = \{\text{vacant}, \text{occupied}, \text{polluted}\}$ for the state space and denote by $\Omega = S^{\mathbb{Z}^2}$ the configuration space for the bootstrap process. For each positive reals p and q satisfying $p + q < 1$, let $\mathbb{P}_{p,q}$ be the law of the initial configuration ω under which each $\omega(x)$ is occupied with probability p, polluted with probability q, and vacant otherwise, independently of all the others. Starting from $\mathcal{B}^0\omega := \omega$, we recursively define $\mathcal{B}^t\omega$ as the outcome of applying the 2-neighbor update rule to all the sites of $\mathcal{B}^{t-1}\omega$ simultaneously, and then define $\mathcal{B}^\infty\omega$ as the pointwise limit of $\mathcal{B}^t\omega$ as $t \to \infty$, which is well-defined by the monotonicity of the update rule.

In order to describe the basic properties of the terminal configuration, we introduce two quantities

$$\begin{aligned} \theta(p,q) &:= \mathbb{P}_{p,q}(0 \text{ is occupied in } \mathcal{B}^\infty\omega), \\ \phi(p,q) &:= \mathbb{P}_{p,q}(\text{there exists an infinite occupied cluster in } \mathcal{B}^\infty\omega). \end{aligned} \tag{1.2.1}$$

Also, define $\lambda(p,q) := q/p^2$ as the scale factor for the 2-neighbor PBP. Then the first result reads:

Theorem 1.2.1. *Under the 2-neighbor PBP, there exist finite, positive constants $\lambda_1 \le \lambda_2$ such that:*

(1) If $\lambda(p,q) < \lambda_1$, then $\phi(p,q) = 1$ for all sufficiently small p and $\theta(p,q) \to 1$ as $p \downarrow 0$.

(2) If $\lambda(p,q) > \lambda_2$, then $\phi(p,q) = 0$ for all sufficiently small p and $\theta(p,q) \to 0$ as $p \downarrow 0$.

Without further information on the critical parameters λ_1 and λ_2, the content of this theorem is identical to the main result of Gravner and McDonald [55]. The improvement in

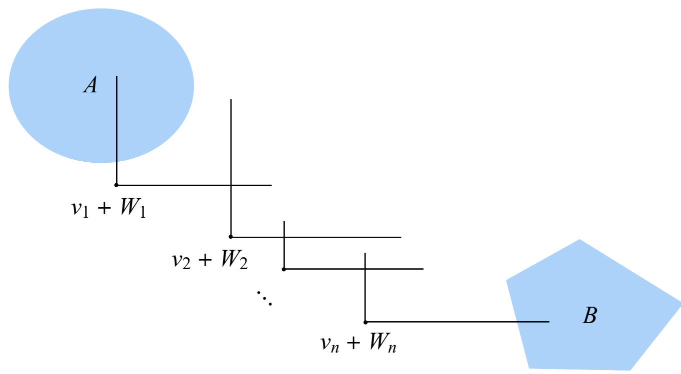


Figure 1.5: Connectivity in the continuum percolation on $\xi_{\lambda,R}$

our work is that both critical parameters are identified with natural percolation thresholds in the continuum percolation model arising as the scaling limit of the *same* blocking structure.

The continuum percolation model will be defined using a marked Poisson point process on $\mathbb{R}^2$, which is a Poisson point process whose every point is associated with a grain that is independently sampled from a probability distribution μ called the *grain distribution*. In the context of 2-neighbor PBP, the grains will be sampled from structures that we call wedges. Here, a *wedge* is the set of the form

$$\mathsf{W}(a,b) := ([0,a] \times \{0\}) \cup (\{0\} \times [0,b]) \tag{1.2.2}$$

with arm lengths $a \geq 0$ and $b \geq 0$. For each $u = (u_i)_{i=1}^6 \in [0,\infty)^6$, the grain is given as the set of the form

$$\mathcal{W}(u) := \{\mathsf{W}(u_1 \wedge u_2, u_4 \wedge u_5), \mathsf{W}(u_2 \wedge u_3, u_4 \wedge u_5), \mathsf{W}(u_1 \wedge u_2, v_5 \wedge v_6)\}. \tag{1.2.3}$$

In our percolation model, the grain distribution μ is such that the grain sampled from μ is identically distributed as $\mathcal{W}((U_i)_{i=1}^6)$, where U_i's are independent Exponential(1) random variables. Now for each $\lambda \geq 0$, consider the marked Poisson point process $\xi_{\lambda,\infty}$ in $\mathbb{R}^2$ of intensity λ and grain distribution μ. Also, for each $R > 0$ we consider the truncated version $\xi_{\lambda,R}$ that is defined as the set of all point/grain pairs $(v, \mathcal{W})$ in $\xi_{\lambda,\infty}$ such that each wedge in $\mathcal{W}$ has arm length not exceeding R. In order to extract a percolation problem from this process, we define the notion of connectivity in $\xi_{\lambda,R}$ as follows: Two subsets A and B of $\mathbb{R}^2$

are connected in $\xi_{\lambda,R}$ if there exist a relabeling $\{A', B'\}$ of the set $\{A, B\}$ and a sequence of point-grain pairs $\{(v_i, \mathcal{W}_i)\}_{i=1}^n \subseteq \xi_{\lambda,R}$, such that

(1) v_{i+1} lies bottom-right of v_i, i.e., $v_{i+1} - v_i \in (0, \infty) \times (-\infty, 0)$, for each $i = 1, \cdots, n-1$.

(2) There exists $W_i \in \mathcal{W}_i$ for each $i = 1, \cdots, n$ such that the sets in every consecutive pair in the sequence $(A', v_1 + W_1, \cdots, v_n + W_n, B')$ intersect. (See Figure 1.5 for an example.)

In such case, we write $A \longleftrightarrow B$ in $\xi_{\lambda,R}$. Note that, in the part (2) of the definition, the sequence $(v_i + W_i)_{i=1}^n$ induces the *zigzag path* of wedges $\cup_{i=1}^n (v_i + W_i)$ that contains a path from A to B.

Denote by B_r the Euclidean ball of radius r centered at the origin and define critical thresholds by

$$\begin{aligned} \lambda_1^{\mathrm{c}} &:= \sup\Big\{\lambda \geq 0 : \limsup_{r\to\infty} \frac{\log \mathbb{P}(\mathsf{B}_1 \longleftrightarrow \partial\mathsf{B}_r \text{ in } \xi_\lambda)}{\log r} < 0\Big\}, \\ \lambda_2^{\mathrm{c}} &:= \inf\Big\{\lambda \geq 0 : \lim_{r\to\infty} \mathbb{P}(\mathsf{B}_1 \longleftrightarrow \partial\mathsf{B}_r \text{ in } \xi_{\lambda,R}) > 0 \text{ for some } R \in (0,\infty)\Big\}. \end{aligned} \tag{1.2.4}$$

The next theorem allows to squeeze the range of λ_1 and λ_2 in the previous theorem by the above critical parameters:

Theorem 1.2.2. *Both critical thresholds λ_1^{c} and λ_2^{c} are positive and finite. Moreover,*

(1) If $\lambda < \lambda_1^{\mathrm{c}}$, then $\phi(p, \lambda p^2) = 1$ for p sufficiently small.

(2) If $\lambda > \lambda_2^{\mathrm{c}}$, then $\phi(p, \lambda p^2) = 0$ for p sufficiently small.

Given that the diameter of grains of ξ_λ has an exponential tail, it is quite reasonable to expect that λ_1^{c} and λ_2^{c} coincide, in analogy with the sharp phase transition known to occur in other well-studied percolation models. Indeed, a sharp transition between the regimes of exponential decay of correlations and long-range order has already been established for a general class of graphs such as d-dimensional lattices (Menshikov [72], Aizenman and Barsky [5]) and quasi-transitive graphs (Antunović and Veselić [7], Duminil-Copin and Tassion [40]). For the Poisson-Boolean percolation on $\mathbb{R}^d$, which corresponds to a marked Poisson point

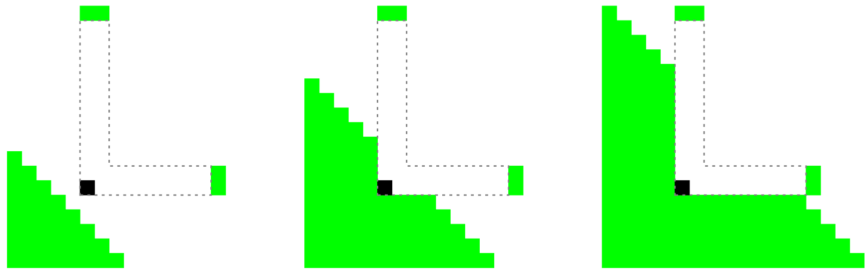


Figure 1.6: Evolution of an occupied half-plane under the 2-neighbor rules when the "front" of occupied sites encounters a NE-wedge. Here, occupied sites are colored green and polluted sites are colored black.

process where the grains are simply Euclidean balls, a similar sharp transition has been proved in the case of bounded radii in Zuev and Sidorenko [102] and Meester, Roy, and Sarkar [71], and for the case of radii with a finite $(5d-3)$-th moment by Duminil-Copin, Raoufi, and Tassion [39]. In another line of generalization, Ziesche [101] established the sharp transition for the continuum percolation with bounded grains of arbitrary shapes. Although the current literature does not settle the sharpness of phase transition for our marked Poisson point process ξ_λ due to the unbounded support of the grain diameter distribution and non-spherical nature of the grains, the stated progress in percolation theory makes future progress on this question quite likely.

1.2.2 Outline of proof

We will sketch the idea of proof. As already noted, a key observation is that the principal mechanism by which the growth is blocked is the formation of a *blocking contour*. To describe this, we introduce some terminology. A *Northeast-wedge (NE-wedge)* is the subset of $\mathbb{Z}^2$ which is the union of horizontal and vertical rectangles of width 2 sharing the common bottom-left corner. A NE-wedge W is called *blocking* in ω if no sites of W are occupied in ω and, writing (x, y) for the bottom-left corner of W, if at least one of (x, y), $(x-1, y)$, or $(x, y-1)$ is polluted in ω. Wedges facing other directions are defined similarly. Figure 1.6 describes an example of a blocking NE-wedge as well as provides motivation for this definition.

It can be proved that the occupied/unoccupied interface in the terminal configuration of the 2-neighbor PBP is predominantly composed of blocking contours (which are blocking wedges arranged in a zigzag way with occasional changes in orientation of the wedges). Since blocking wedges in the terminal configuration are also blocking in the initial configuration, we may instead study the structure of blocking contours in the initial configuration.

Since, for q scaling proportional to p, the typical size of the NE-wedges is comparable with a typical distance between the polluted vertices, 'maximal' blocking NE-wedges under the law $\mathbb{P}_{p,q}$ with $\lambda = \lambda(p,q)$ may be coupled with grains of ξ_λ under the scaling of the lattice by factor p. Then through this coupling, we can translate the subcriticality/supercriticality of ξ_λ to supercriticality/subcriticality of blocking wedges under $\mathbb{P}_{p,q}$. Finally, adapting the technique in Gravner and McDonald [55], we then translate the scarcity/abundance of large blocking contours to the percolation/non-percolation of occupied clusters in the terminal configuration. This leads to the proof of Theorem 1.2.2.

1.3 Asymptotic normality of permutation statistics

The last subject to be discussed in this thesis are limit laws for random permutations. The theory of *permutation statistics* concerns the distributions of various numerical quantities derived from random permutations. One of the best studied examples is that of the descents. Here, a permutation π of the set $[n] = \{1, \cdots, n\}$ is said to have a *descent* at position i if $\pi(i) > \pi(i+1)$. Descents appear in numerous parts of mathematics. Knuth [66] related descents to sorting and runs in permutations, and Diaconis, McGrath, and Pitman [37] studied a model of card shuffling in which descents play a central role. Bayer and Diaconis [10] also used descents and rising sequences to analyze the mixing speed of the most commonly used card shuffling method. Garsia and Gessel [52] found a joint generating function of descents, major indices, and inversions, and Gessel and Reutenauer [53] investigated the

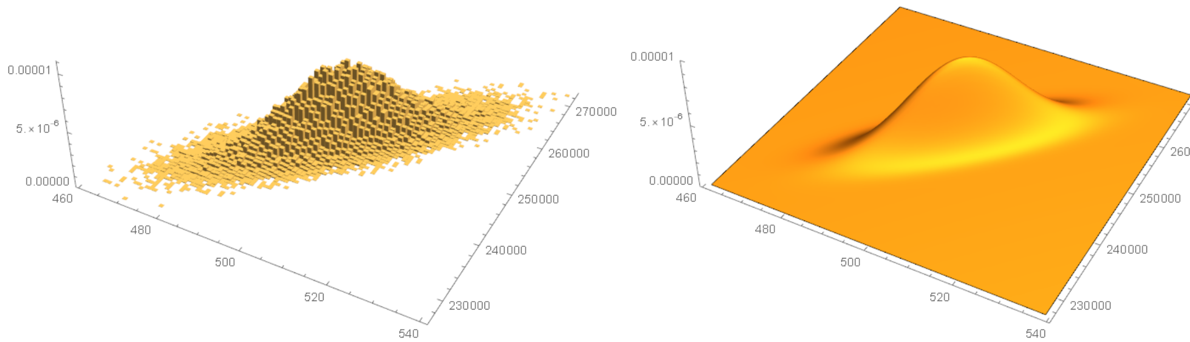


Figure 1.7: Normalized histogram of descents/major-index pairs of 10^5 permutations with cycle type 2^{500} (left) and the PDF of its normal approximation as predicted by the CLT (right).

number of permutations with given cycle structure and descent set using algebraic approach. Petersen has an excellent book [80] on Eulerian numbers.

It is well known that the distribution of descents in a uniformly random permutation S_n is asymptotically normal with mean $(n-1)/2$ and variance $(n+1)/12$. Several proofs are known, each of which utilizing different aspects of descents. Pitman [81] adopted Harper's method, which relies on the real-rootedness of the generating function of descents. David and Barton [10] utilized the method of moments. Tanny [94] argued by observing that the descents are equidistributed to the integer part of the sum of independent uniform random variables on the unit interval. Fulman [49] used Stein's method. Very recently, Özdemir [78] employed Martingale CLT for a suitable insertion process. Asymptotic normality has also been studied for various permutation statistics, including but not restricted to, Mahonian statistics and descents in certain conjugacy classes of S_n [48].

In this part of the thesis, we prove that various permutation statistics in an arbitrary fixed conjugacy class of S_n is asymptotically normal as $n \to \infty$, with both asymptotic mean and variance depending only on the fixed point density of the cycle type. Specifically, we will cover the joint distribution of descents/major indices and the distribution of peaks. In both cases, the proofs utilize a variant of the Lévy Continuity Theorem for the moment generating function (MGF) to reduce the questions on asymptotic normality, to those on the pointwise convergence of MGFs on some non-empty open sets. A key technical result in this regard is the uniform estimate on the MGFs of normalized statistics. The proof of such estimates is based on remarkably precise control of exact generating functions for permutation statistics and cycle structures, with the prime resource being Gessel and Reutenauer's work [53]. Also, such uniform estimates turn out to be strong enough that they can be applied to more general conjugation-invariant subsets, including interesting cases such as derangements.

1.3.1 The main results and outline of the proof

We briefly demonstrate the statement and the proof in the case of descents in conjugacy classes, which is the content of Section 4.1. Proofs for other permutation statistics are similar in spirit to this case.

Let $n \geq 1$ and let λ be an integer partition of n for which the number of occurrences of natural i is $n_i = n_i(\lambda)$ for each i. Then the conjugacy class $\mathcal{C}_\lambda$ is the set of all permutations having exactly n_i cycles of length i for each i. Let D_λ denote the number of descents in a permutation drawn uniformly at random from $\mathcal{C}_\lambda$. From Gessel and Reutenauer [53], the MGF of D_λ is given by

$$\mathbb{E}\big[\mathrm{e}^{sD_N}\big] = \frac{(1-\mathrm{e}^s)^{n+1}}{n!} \sum_{a \geq 1} \mathrm{e}^{as} \prod_{i=1}^{n} \prod_{j=0}^{n_i - 1} (F_{i,a} + ij), \tag{1.3.1}$$

where $F_{i,a} := \sum_{d|i} \mu(d) a^{i/d}$ and $\mu(\cdot)$ is the Möbius function which takes the value $(-1)^k$ at square-free positive integers with k prime factors and zero otherwise. Differentiating this identity and writing $\alpha_\lambda := n_1(\lambda)/n$ for the density of fixed points, it can be checked that

$$\mathbb{E}[D_\lambda] = \frac{1-\alpha_\lambda^2}{2} n + \mathcal{O}(1) \qquad \text{and} \qquad \mathrm{Var}(D_\lambda) = \frac{1 - 4\alpha_\lambda^3 + 3\alpha_\lambda^4}{12} n + \mathcal{O}(1) \tag{1.3.2}$$

uniformly in λ. This suggests the normalization $W_\lambda := \frac{1}{\sqrt{n}} D_\lambda - \frac{1-\alpha_\lambda^2}{2}\sqrt{n}$. The main technical result of Section 4.1 for descents is stated as follows:

Theorem 1.3.1 (Asymptotic normality of descents)**.** *For each fixed $s \in (-\infty, 0)$, we have*

$$M_{W_\lambda}(s) - \mathrm{e}^{\frac{s^2}{24}(1-4\alpha_\lambda^3+3\alpha_\lambda^4)} \xrightarrow[n\to\infty]{} 0 \tag{1.3.3}$$

uniformly in n and $\lambda \vdash n$.

The main consequence of this theorem is that, along any subsequence of integer partitions λ_n of n with $n \to \infty$ and $\alpha_{\lambda_n} \to \alpha$ for some constant $\alpha \in [0,1]$, W_{λ_n} converge in distribution to the normal distribution with mean zero and variance $\frac{1}{12}(1-4\alpha^3+3\alpha^4)$. Here, the following version of Continuity Theorem for MGF is utilized:

Theorem 1.3.2 (Continuity theorem for MGF)**.** *Suppose that $(X_n)_{n\in\mathbb{N}\cup\{\infty\}}$ is a sequence of random vectors in $\mathbb{R}^d$ such that $M_{X_n}(\xi) := \mathbb{E}[\exp(\xi \cdot X_n)]$ converges pointwise to M_{X_∞} on a non-empty open subset of $\mathbb{R}^d$. Then $X_n \to X_\infty$ in distribution.*

The proof of Theorem 1.3.1 consists of a series of estimations. First, performing some estimations on the egregious coefficient appearing in the MGF of D_N and invoking the idea that $F_{i,a}$ may be regarded as a^i for large a, we get

$$\prod_{i=1}^{n}\prod_{j=0}^{n_i-1}(F_{i,a}+ij) = a^n \mathrm{e}^{\frac{\alpha_1^2 n^2}{2a} - \frac{\alpha_1^3 n^3}{6a^2} + \mathcal{O}(n^{-1/2})} \tag{1.3.4}$$

provided $a > \epsilon n^{3/2}$ for a sufficiently small $\epsilon > 0$. Plugging this to the MGF of W_λ evaluated at $-s$ for $s > 0$ and approximating the sum $\sum_{a\geq 1}$ by the integral $\int_0^\infty \mathrm{d}a$, we expect that

$$\begin{aligned} M_{W_\lambda}(-s) &= \frac{(1-\mathrm{e}^{-s/\sqrt{n}})^{n+1}}{n!}\,\mathrm{e}^{\frac{(1-\alpha_\lambda^2)s\sqrt{n}}{2}} \sum_{a\geq 1} \mathrm{e}^{-\frac{as}{\sqrt{n}}} \prod_{i=1}^{n}\prod_{j=0}^{n_i-1}(F_{i,a}+ij) \\ &\approx \frac{\mathrm{e}^{\frac{s^2}{24}}}{n!}\left(\frac{s}{\sqrt{n}}\right)^{n+1} \int_{\epsilon n^{3/2}}^{\infty} a^n \mathrm{e}^{-\frac{as}{\sqrt{n}} - \frac{\alpha_\lambda^2 s\sqrt{n}}{2} + \frac{\alpha_1^2 n^2}{2a} - \frac{\alpha_1^3 n^3}{6a^2}}\,\mathrm{d}a \end{aligned} \tag{1.3.5}$$

holds, at least heuristically. Then the last integral may be analyzed by the Laplace method. Indeed substituting $as = n^{3/2} + nz$ and manipulating

$$\begin{aligned} &= \frac{\mathrm{e}^{\frac{s^2}{24}} n^{n+\frac{1}{2}} \mathrm{e}^{-n}}{n!} \int_{-(1-\epsilon s)\sqrt{n}}^{\infty} \left(1+\frac{z}{\sqrt{n}}\right)^n \mathrm{e}^{-\sqrt{n}z + \frac{\alpha_\lambda^2 sz}{2(1+z/\sqrt{n})} - \frac{\alpha_\lambda^3 s^2}{6(1+z/\sqrt{n})^3}}\,\mathrm{d}z \\ &\approx \frac{\mathrm{e}^{\frac{s^2}{24}}}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{-\frac{z^2}{2} + \frac{\alpha_\lambda^2 sz}{2} - \frac{\alpha_\lambda^3 s^2}{6}}\,\mathrm{d}z. \end{aligned} \tag{1.3.6}$$

The last integral is then computed as $\mathrm{e}^{(1-4\alpha_\lambda^3+3\alpha_\lambda^4)\frac{s^2}{24}}$, which is the required term appearing in the main theorem. Carefully justifying this heuristics produces a desired uniform convergence, and in fact, a uniform estimate on the difference.

1.4 Outline of the thesis

As mentioned in the acknowledgments, Chapter 2 and 4 are reprints of arXiv postings and published work and so they are included *verbatim*.

Chapter 2 is devoted to the exceptional points of a discrete time simple random walks in lattice approximations of a planar domain run for a constant multiple of the expected local time. In Section 2.1, we briefly discuss the previous works that motivated this project and provide a minimal set of notation and definitions necessary for stating the main results. Section 2.2 contains the main theorems. We first discuss an observation which sets up the time scale to work with and thus provide a logical basis for the definitions of λ-thick/thin points and the associated point measures encoding their structure. Then we digress to the scaling limit for the level sets of zero-average DGFF. The resulting limit measure, called the *zero-average Liouville Quantum Gravity (LQG)* measure, is then used to present the main theorems regarding the convergence of exceptional points for thick/thin/light/avoided points as well as for those augmented by the local structure. The main point is that the LQG measure describing the spatial coordinate of the scaling limit in the previous work [1] is now replaced by the zero-average LQG measure of the same parameters multiplied by a log-normal prefactor. We then conclude this section with remarks on some details and future directions.

The rest of Chapter 2 is devoted to the proofs of the main theorems. In Section 2.3, the scaling limit for the level sets of zero-average DGFF is proved, and as a corollary, the joint scaling limit for the level sets and the average of the DGFF is shown. To make use of this result, Section 2.4 adopts the arguments in [1] and extends the convergence theorems for the exceptional points to those augmented by the normalized fluctuation of the total time of the walk at ϱ. Then in Section 2.5, a time conversion relating the fixed time to the inverse local time shifted by the fluctuation of total time at ϱ is established, and the convergence of exceptional points of the local time of the continuous-time random walk is proved. Section 2.6 controls the effect of making the starting point arbitrary and confirms that this effect is negligible in the limit. In Section 2.7 we make use of the holding-time representation of the local time of the continuous-time random walk to derive the main results without the local structure, which is then resolved in Section 2.8.

Chapter 3 discusses the 2-neighbor *polluted bootstrap percolation (PBP)*. In Section 3.1, we give the definition of the model of interest and state the main results. Section 3.2 examines the polluted bootstrap dynamics and identifies the principal (deterministic) mechanism by which the growth of occupied clusters is blocked. In Section 3.3, the continuum percolation model is introduced as an ideal representation of the scaling limit of blocking structures, and in particular, a useful coupling of the two models are established. Using this, Section 3.4 shows that long blocking contours are very rare in the subcritical regime. In Section 3.5, a key observation on the crossing probabilities in the supercritical regime is established using an analogy with 2-dimensional oriented percolation, and its consequences are translated into the language of polluted bootstrap percolation via a coupling of the two models. Section 3.6 is devoted to the proof of the main results.

Chapter 4 deals with the asymptotic normality of two permutation statistics, descents and peaks, in an arbitrary conjugacy class. Section 4.1 covers the case of the descents. We begin with the historical background and then move to the statement of the main result. In Section 4.1.2, we review the generating functions of the descents in a conjugacy class $\mathcal{C}_\lambda$ of cycle type λ, which is originally represented in terms of the power series in the unit disk, and then derive another representation in terms of the Laurent series in the annulus between the unit circle and the point at infinity. Section 4.1.3 identifies exact formulas for the mean and variance of the descents chosen uniformly at random from $\mathcal{C}_\lambda$ through exact combinatorial computation. Section 4.1.4 uses this computation to set up a proper normalization of the descents in $\mathcal{C}_\lambda$ and then proves the uniform estimate on the moment generating function, which in turn implies the main theorem via Theorem 1.3.2.

Section 4.2 is devoted to the peaks. The organization of this section is almost identical to the previous section with only a small twist. In Section 4.2.1, we begin by briefly reviewing the background of the problem. Then we introduce the generating function of picks in a conjugacy class $\mathcal{C}_\lambda$ of fixed cycle type λ and state the the main results. In Section 4.2.2, we provide a computation of the mean and variance of peaks in the symmetric group S_n and prove a toy result on its asymptotic normality as $n \to \infty$. In Section 4.2.3, we discuss a heuristics that gives a partially correct picture. The ideas contained in it will then be elaborated to yield a uniform estimate on the moment generating functions of (normalized) peaks in $\mathcal{C}_\lambda$.

Finallly, in Chapter 5 we conclude the discussion of each topic in a wider context and suggest possible future directions.

Chapter 2

Exceptional level sets of local times

Given a sequence of lattice approximations $D_N \subset \mathbb{Z}^2$ of a bounded continuum domain $D \subset \mathbb{R}^2$ with the vertices outside D_N fused together into one boundary vertex ϱ, we consider discrete-time simple random walks in $D_N \cup \{\varrho\}$ run for a time proportional to the expected cover time and describe the scaling limit of the exceptional level sets of the thick, thin, light and avoided points. We show that these are distributed, up a spatially-dependent log-normal factor, as the zero-average Liouville Quantum Gravity measures in D. The limit law of the local time configuration at, and nearby, the exceptional points is determined as well. The results extend earlier work by Abe and Biskup [1] who analyzed the continuous-time problem in the parametrization by the local time at ϱ. A novel uniqueness result concerning divisible random measures and, in particular, Gaussian Multiplicative Chaos, is derived as part of the proofs.

2.1 Introduction

This chapter contains a continuation of earlier work by Abe and Biskup [1] who in [1] studied various exceptional level sets associated with the local time of random walks in lattice versions $D_N \subset \mathbb{Z}^2$ of bounded open domains $D \subset \mathbb{R}^2$, at times proportional to the cover time of D_N. The walks in [1] move as the ordinary constant-speed continuous-time simple symmetric random walk on D_N and, upon exit from D_N, reenter D_N through a

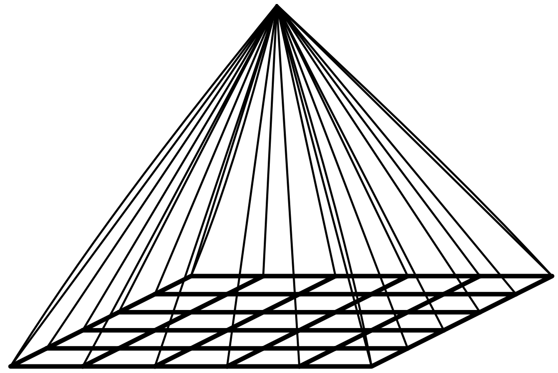


Figure 2.1: The graph $(V \cup \{\varrho\}, E)$ corresponding to D_N being the square of 6×6 vertices and all edges emanating from D_N routed to the boundary vertex ϱ. Note that the graph $(V \cup \{\varrho\}, E)$ is planar whenever $\mathbb{Z}^2 \smallsetminus D_N$ is connected.

uniformly-chosen boundary edge. The re-entrance mechanism is conveniently realized by addition to D_N of a boundary vertex ϱ with all edges emanating out of D_N on $\mathbb{Z}^2$ now ending in ϱ. See Fig. 2.1 for an example.

In [1], the local time was parametrized by the time spent at ϱ. Through the use of the Second Ray-Knight Theorem (Eisenbaum, Kaspi, Marcus, Rosen and Shi [43]) this enabled a connection to the level sets of the Discrete Gaussian Free Field (DGFF) studied earlier by Biskup and Louidor [16]. The goal of the present paper is to extend the results of [1] to the more natural setting of a discrete-time random walk parametrized by its actual time. As we shall see, a close connection to the DGFF still persists, albeit now to that conditioned on vanishing arithmetic mean over D_N. As no version of the Second Ray-Knight Theorem seems available for this specific setting, we have to proceed by suitable, and sometimes tedious, approximations. A key point is to control the fluctuations of the total time of the random walk at a given occupation time of the boundary vertex.

In order to give the precise setting of our problem, we first consider a general finite, unoriented, connected graph $G = (V \cup \{\varrho\}, E)$, where ϱ is a distinguished vertex (not belonging to V). Let X denote a sample path of the simple random walk on G; i.e., a discrete-time Markov chain on $V \cup \{\varrho\}$ with the transition probabilities

$$\mathsf{P}(u,v) := \begin{cases} \frac{1}{\deg(u)}, & \text{if } e := (u,v) \in E, \\ 0, & \text{otherwise,} \end{cases} \tag{2.1.1}$$

where $\deg(u)$ is the degree of u. As usual, we will write P^u to denote the law of X subject to the initial condition $P^u(X_0 = u) = 1$.

Given a path X of the chain, the local time at $v \in V \cup \{\varrho\}$ at time n is then given by

$$\ell_n^V(v) := \frac{1}{\deg(v)} \sum_{k=0}^{n} 1_{\{X_k = v\}}, \quad n \geq 0. \tag{2.1.2}$$

Our aim is to observe the Markov chain at times when most, or even all, of the vertices have already been visited. This requires looking at the chain at times (at least) proportional to the total degree $\deg(V) := \sum_{v \in V \cup \{\varrho\}} \deg(V)$. To simplify our later notations, we thus abbreviate, for any $t > 0$,

$$L_t^V(v) := \ell^V_{\lfloor t \deg(V) \rfloor}(v), \quad v \in V. \tag{2.1.3}$$

In this parametrization, we have $L_t^V(v) = t + o(t)$ with high probability as $t \to \infty$.

Our derivations will make heavy use of the connection between the above Markov chain and an instance of the Discrete Gaussian Free Field (DGFF). Denoting by

$$H_v := \inf\{n \geq 0 \colon X_n = v\} \tag{2.1.4}$$

the first hitting time of vertex v, this DGFF is the centered Gaussian process $\{h_v^V : v \in V\}$ with covariances given by

$$\mathbb{E}\big(h_u^V h_v^V\big) = G^V(u,v) := E^u\big(\ell^V_{H_\varrho}(v)\big), \tag{2.1.5}$$

where $\mathbb{E}$ the expectation with respect to the law of h^V and G^V is the Green function. The field naturally extends to ϱ by $h_\varrho^V = 0$.

We will apply the above to V ranging through a sequence of lattice approximations of a well-behaved continuum domain. The following definitions are taken from [17]:

Definition 2.1.1. *An admissible domain is a bounded open subset of* $\mathbb{R}^2$ *that consists of a finite number of connected components and whose boundary is composed of a finite number of connected sets each of which has positive Euclidean diameter.*

We will write $\mathfrak{D}$ to denote the family of all admissible domains and let $d_\infty(\cdot,\cdot)$ denote the ℓ^∞-distance on $\mathbb{R}^2$. The lattice domains are then assumed to obey:

Definition 2.1.2. *An admissible lattice approximation of* $D \in \mathfrak{D}$ *is a sequence* $\{D_N\}_{N\geq 1}$ *of sets* $D_N \subset \mathbb{Z}^2$ *such that the following holds: There is* $N_0 \in \mathbb{N}$ *such that for all* $N \geq N_0$ *we have*

$$D_N \subseteq \Bigl\{x \in \mathbb{Z}^2 \colon d_\infty\bigl(x/N, \mathbb{R}^2 \smallsetminus D\bigr) > \frac{1}{N}\Bigr\} \tag{2.1.6}$$

and, for any $\delta > 0$ *there is also* $N_1 \in \mathbb{N}$ *such that for all* $N \geq N_1$*,*

$$D_N \supseteq \bigl\{x \in \mathbb{Z}^2 \colon d_\infty(x/N, \mathbb{R}^2 \smallsetminus D) > \delta\bigr\}. \tag{2.1.7}$$

As shown in [17, Appendix A], the conditions (2.1.6–2.1.7) ensure that the discrete harmonic measure on D_N tends, under scaling of space by N, weakly to the harmonic measure on D. This yields a precise asymptotic expansion of the associated Green function; see [14, Chapter 1]. In particular, we have $G^{D_N}(x,x) = g \log N + O(1)$ for

$$g := \frac{1}{2\pi} \tag{2.1.8}$$

whenever x is deep inside D_N. (This is by a factor 4 smaller than the corresponding constant in [14, 17] due to a different normalization of the Green function.)

2.2 Main results

Let us move to discussing our main results. We pick an admissible domain $D \in \mathfrak{D}$ and a sequence of admissible lattice approximation $\{D_N\}_{N\geq 1}$ and consider these fixed throughout the rest of the derivations.

2.2.1 Setting the scales

We begin by setting the scales for the time that the random walk is observed for and determining the range of values taken by the local time:

Theorem 2.2.1. *Let* $\{t_N\}_{N\ge1}$ *be a positive sequence such that, for some* $\theta>0$,

$$\lim_{N\to\infty}\frac{t_N}{(\log N)^2}=2g\theta. \tag{2.2.1}$$

Then for any choices of $x_N\in D_N$*, the following limits hold in* P^{x_N}*-probability:*

$$\frac{1}{(\log N)^2}\max_{x\in D_N}L_{t_N}^{D_N}(x)\;\underset{N\to\infty}{\longrightarrow}\;2g\big(\sqrt{\theta}+1\big)^2 \tag{2.2.2}$$

and

$$\frac{1}{(\log N)^2}\min_{x\in D_N}L_{t_N}^{D_N}(x)\;\underset{N\to\infty}{\longrightarrow}\;2g\big[(\sqrt{\theta}-1)\vee 0\big]^2. \tag{2.2.3}$$

The conclusion (2.2.3) indicates (and our later results on avoided points prove) that the choice $\theta:=1$ identifies the leading order of the *cover time* of D_N — defined as the first time that every vertex of the graph has been visited. The cover time is random but it is typically concentrated (more precisely, whenever the maximal hitting time is much smaller than the expected cover time; see Aldous [6]). The scaling (2.2.1) thus corresponds to the walk run for a θ-multiple of the cover time.

As it turns out, under (2.2.1), the asymptotic $[2g\theta+o(1)](\log N)^2$ marks the value of $L_{t_N}^{D_N}$ at all but a vanishing fraction of the vertices in D_N. In light of (2.2.2–2.2.3), this suggests that we call $x\in D_N$ a λ-*thick point* if (for $\lambda\in[0,1]$)

$$L_{t_N}^{D_N}(x)\ge 2g\big(\sqrt{\theta}+\lambda\big)^2(\log N)^2 \tag{2.2.4}$$

and a λ-*thin point* if (for $\lambda\in[0,\sqrt{\theta})$)

$$L_{t_N}^{D_N}(x)\le 2g\big(\sqrt{\theta}-\lambda\big)^2(\log N)^2. \tag{2.2.5}$$

One of our goals is to describe the scaling limit of the sets of thick and thin points. This is best done via random measures of the form

$$\zeta_N^D:=\frac{1}{W_N}\sum_{x\in D_N}\delta_{x/N}\otimes\delta_{(L_{t_N}^{D_N}(x)-a_N)/\sqrt{2a_N}}\,, \tag{2.2.6}$$

where a_N is a sequence with the asymptotic growth as the right-hand side of (2.2.4–2.2.5) and W_N is a normalizing sequence. The specific choice of the normalization by $\sqrt{2a_N}$ reflects on the natural fluctuations of $L^{D_N}_{t_N}(x)$ (which turn out to be order $\log N$ even between nearest neighbors) and captures best the connection to the corresponding object for the DGFF to be discussed next.

2.2.2 Level sets of zero-average DGFF

Recall that h^{D_N} denotes a sample of the DGFF in D_N. As shown by Bolthausen, Deuschel and Giacomin [20], the maximum of h^{D_N} is asymptotic to $2\sqrt{g}\log N$ and so the λ-thick points are naturally defined as those where the field exceeds $2\lambda\sqrt{g}\log N$. Allowing for sub-leading corrections, these are best captured by the random measure

$$\eta^D_N := \frac{1}{K_N}\sum_{x\in D_N}\delta_{x/N}\otimes\delta_{h^{D_N}_x-\widehat{a}_N}, \tag{2.2.7}$$

where $\{\widehat{a}_N\}$ is a centering sequence with the asymptotic $\widehat{a}_N\sim 2\lambda\sqrt{g}\log N$ and

$$K_N := \frac{N^2}{\sqrt{\log N}}\mathrm{e}^{-\frac{(\widehat{a}_N)^2}{2g\log N}}. \tag{2.2.8}$$

In [16, Theorem 2.1] it was shown that for each $\lambda\in(0,1)$, there is a constant $\mathfrak{c}(\lambda)>0$ (independent of D or the approximating sequence $\{D_N\}_{N\geq 1}$) such that, relative to the topology of vague convergence of measures on $\overline{D}\times(\mathbb{R}\cup\{+\infty\})$,

$$\eta^D_N \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}} \mathfrak{c}(\lambda)\,Z^D_\lambda(\mathrm{d}x)\otimes\mathrm{e}^{-\alpha\lambda h}\mathrm{d}h, \tag{2.2.9}$$

where

$$\alpha := \frac{2}{\sqrt{g}} \tag{2.2.10}$$

and where Z^D_λ is a random a.s.-finite Borel measure in D called the *Liouville Quantum Gravity* (LQG) at parameter λ-times critical. The measure Z^D_λ is normalized so that, for each Borel set $A\subseteq D$,

$$\mathbb{E}Z^D_\lambda(A) = \int_A r^D(x)^{2\lambda^2}\mathrm{d}x, \tag{2.2.11}$$

where r^D is an explicit bounded, continuous function supported on D that, for D simply connected, is the conformal radius; see [16, (2.10)].

As was shown in [1], the measures $\{Z^D_\lambda\colon \lambda\in(0,1)\}$ are quite relevant for the exceptional level sets associated with the continuous-time random walk in the parametrization by the local time spent in the "boundary vertex." Somewhat different measures will arise for the discrete-time random walk. Let $\Pi^D(x,\cdot)$ be the harmonic measure in D defined, e.g., as the exit distribution from D of a Brownian motion started at x. The continuum Green function in D with Dirichlet boundary condition is then given by

$$\widehat{G}^D(x,y) := -g\log|x-y| + g\int_{\partial D}\Pi^D(x,\mathrm{d}z)\log|y-z|. \tag{2.2.12}$$

Writing Leb for the Lebesgue measure on $\mathbb{R}^2$, let $\mathfrak{d}\colon \mathbb{R}^2\to\mathbb{R}$ be defined by

$$\mathfrak{d}(x) := \mathrm{Leb}(D)\frac{\int_D \mathrm{d}y\,\widehat{G}^D(x,y)}{\int_{D\times D}\mathrm{d}z\,\mathrm{d}y\,\widehat{G}^D(z,y)}. \tag{2.2.13}$$

As is readily checked, $\mathfrak{d}$ is bounded and continuous, vanishes outside D and integrates to $\mathrm{Leb}(D)$ over D. (We also have $\mathfrak{d}\ge 0$ because $\widehat{G}^D\ge 0$ and also that the Laplacian of $\mathfrak{d}$ is constant on D but that is of no consequence in the sequel.) See Fig. 2.2. We claim:

Theorem 2.2.2. *For each $\lambda\in(0,1)$ and each $D\in\mathfrak{D}$, there is a unique random measure $Z^{D,0}_\lambda$ on D such that, for any sequence $\{D_N\}_{N\ge1}$ of admissible approximations of D and any centering sequence $\{\widehat{a}_N\}_{N\ge1}$ satisfying $\widehat{a}_N\sim 2\lambda\sqrt{g}\log N$ as $N\to\infty$,*

$$\Bigl(\eta^D_N\Bigm|\sum_{x\in D_N}h^{D_N}_x=0\Bigr)\ \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}}\ \mathfrak{c}(\lambda)\,Z^{D,0}_\lambda(\mathrm{d}x)\otimes\mathrm{e}^{-\alpha\lambda h}\mathrm{d}h, \tag{2.2.14}$$

where $\mathfrak{c}(\lambda)$ is as in (2.2.9)*. Moreover, if Y is a normal random variable with mean zero and variance*

$$\sigma^2_D := \int_{D\times D}\mathrm{d}x\mathrm{d}y\,\widehat{G}^D(x,y), \tag{2.2.15}$$

then the measure from (2.2.9–2.2.11) obeys

$$Y \perp\!\!\!\perp Z^{D,0}_\lambda \quad\Rightarrow\quad Z^D_\lambda(\mathrm{d}x) \overset{\text{law}}{=} \mathrm{e}^{\lambda\alpha\mathfrak{d}(x)Y}\,Z^{D,0}_\lambda(\mathrm{d}x). \tag{2.2.16}$$

The law of $Z^{D,0}_\lambda$ is determined uniquely by (2.2.16)*.*

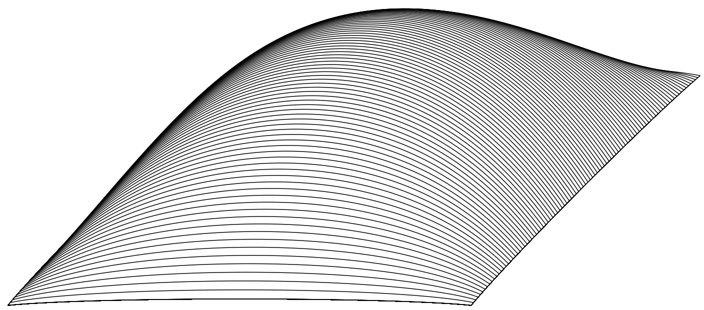


Figure 2.2: A plot of function $\mathfrak{d}$ on $D := (0,1)^2$ obtained by solving the differential equation $-\Delta\mathfrak{d} = \mathrm{Leb}(D)/\sigma_D^2$, where Δ is the Laplacian, with Dirichlet boundary conditions on ∂D.

The existence of a random measure $Z_\lambda^{D,0}$ satisfying (2.2.16) is part of the proof of (2.2.14). The uniqueness of the decomposition (2.2.16) holds quite generally and constitutes the main technical ingredient of the proof; see Theorem 2.3.1 which is of independent interest. The known properties of Z_λ^D (see [16, Theorem 2.3]) imply that $Z_\lambda^{D,0}$ is a.s.-finite and charges every non-empty open subset of D a.s.

2.2.3 Exceptional local-time sets

We are now well equipped to state our results concerning the limits of the random measures (2.2.6) for a given centering sequence $\{a_N\}_{N\geq 1}$ growing as the right-hand sides of (2.2.4–2.2.5) and the normalizing sequence given by

$$W_N := \frac{N^2}{\sqrt{\log N}} \mathrm{e}^{-\frac{(\sqrt{2t_N}-\sqrt{2a_N})^2}{2g\log N}}. \tag{2.2.17}$$

For the thick points we then get:

Theorem 2.2.3 (Thick points). *Suppose $\{t_N\}_{N\geq 1}$ and $\{a_N\}_{N\geq 1}$ are positive sequences such that, for some $\theta > 0$ and some $\lambda \in (0,1)$, (2.2.1) and*

$$\lim_{N\to\infty} \frac{a_N}{(\log N)^2} = 2g(\sqrt{\theta}+\lambda)^2 \tag{2.2.18}$$

hold true. Then for any $x_N \in D_N$ and for X sampled from P^{x_N}, the measures ζ_N^D in (2.2.6) *with W_N as in (2.2.17) obey*

$$\zeta_N^D \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}} \sqrt{\frac{\sqrt{\theta}}{\sqrt{\theta}+\lambda}}\,\mathrm{e}^{-\alpha^2\lambda^2/16}\,\mathfrak{c}(\lambda)\,\mathrm{e}^{\alpha\lambda(\mathfrak{d}(x)-1)Y}Z_\lambda^{D,0}(\mathrm{d}x)\otimes\mathrm{e}^{-\alpha\lambda h}\mathrm{d}h \tag{2.2.19}$$

in the sense of vague convergence of measures on $\overline{D}\times(\mathbb{R}\cup\{+\infty\})$, where $Y=\mathcal{N}(0,\sigma_D^2)$ and $Z_\lambda^{D,0}$ are independent and $\mathfrak{c}(\lambda)$ is as in (2.2.9).

For the thin points, we similarly obtain:

Theorem 2.2.4 (Thin points)**.** *Suppose $\{t_N\}_{N\ge1}$ and $\{a_N\}_{N\ge1}$ are positive sequences such that, for some $\theta>0$ and some $\lambda\in(0,\sqrt{\theta}\wedge1)$, (2.2.1) and*

$$\lim_{N\to\infty}\frac{a_N}{(\log N)^2}=2g(\sqrt{\theta}-\lambda)^2 \tag{2.2.20}$$

hold true. Then for any $x_N\in D_N$ and for X sampled from P^{x_N}, the measures ζ_N^D in (2.2.6) with W_N as in (2.2.17) obey

$$\zeta_N^D \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}} \sqrt{\frac{\sqrt{\theta}}{\sqrt{\theta}-\lambda}}\,\mathrm{e}^{-\alpha^2\lambda^2/16}\,\mathfrak{c}(\lambda)\,\mathrm{e}^{\alpha\lambda(\mathfrak{d}(x)-1)Y}Z_\lambda^{D,0}(\mathrm{d}x)\otimes\mathrm{e}^{+\alpha\lambda h}\mathrm{d}h \tag{2.2.21}$$

in the sense of vague convergence of measures on $\overline{D}\times(\mathbb{R}\cup\{-\infty\})$, where $Y=\mathcal{N}(0,\sigma_D^2)$ and $Z_\lambda^{D,0}$ are independent and $\mathfrak{c}(\lambda)$ is as in (2.2.9).

The limiting spatial distribution of the λ-thick and λ-thin points (as well as the distribution of the total number of these points) is governed by the measure

$$\mathrm{e}^{\alpha\lambda(\mathfrak{d}(x)-1)Y}Z_\lambda^{D,0}(\mathrm{d}x). \tag{2.2.22}$$

In light of (2.2.16), this is somewhere between the zero-average LQG $Z_\lambda^{D,0}$ and the "ordinary" LQG Z_λ^D, which appeared in the limit for the parametrization by the local time at ϱ. The second component of the measure on the right of (2.2.19) and (2.2.21) is exactly as that for the DGFF (2.2.9). This is due to the judicious scaling of the second component of ζ_N^D by $\sqrt{2a_N}$ rather than just $\log N$, as was done in [1].

Apart from the thick and thin points, [1] studied also the sets of points where the local time is order unity, called the *light* points, and the points where the local time vanishes, called the *avoided* points. In both cases, the LQG measure that appears is for parameter $\lambda := \sqrt{\theta}$ (and $\theta \in (0,1)$). The control extends to the discrete-time problem parametrized by the total time as well. We start with the light points:

Theorem 2.2.5 (Light points). *Suppose $\{t_N\}_{N\ge1}$ is a positive sequence such that* (2.2.1) *holds for some $\theta \in (0,1)$. For any $x_N \in D_N$ and for X sampled from P^{x_N}, consider the measure*

$$\vartheta_N^D := \frac{1}{\widehat{W}_N} \sum_{x\in D_N} \delta_{x/N} \otimes \delta_{L_{t_N}^{D_N}(x)}, \tag{2.2.23}$$

where

$$\widehat{W}_N := N^2 \mathrm{e}^{-\frac{t_N}{g\log N}}. \tag{2.2.24}$$

Then, in the sense of vague convergence of measures on $\overline{D} \times [0,\infty)$,

$$\vartheta_N^D \xrightarrow[N\to\infty]{\text{law}} \sqrt{2\pi g}\,\mathfrak{c}(\sqrt{\theta})\ \mathrm{e}^{\alpha\sqrt{\theta}(\mathfrak{d}(x)-1)Y}\, Z_{\sqrt{\theta}}^{D,0}(\mathrm{d}x) \otimes \mu(\mathrm{d}h), \tag{2.2.25}$$

where $\mathfrak{c}(\lambda)$ is as in (2.2.9)*, $Y = \mathcal{N}(0,\sigma_D^2)$ and $Z_{\sqrt{\theta}}^{D,0}$ are independent and $\mu := \sum_{n\ge0} q_n \delta_{n/4}$ for a sequence $\{q_n\colon n\ge 0\}$ of non-negative numbers determined uniquely by*

$$\sum_{n\ge0} q_n (1+s/4)^{-n} = \mathrm{e}^{\frac{\alpha^2\theta}{2s}}, \quad s>0. \tag{2.2.26}$$

That μ is supported on $\frac14\mathbb{N}_0 := \{0, \frac14, \frac12, \frac34, 1, \dots\}$ arises from the normalization in (2.1.2). From (2.2.25) we conclude that the number of the vertices of D_N visited exactly n times during the first

$$[8g\theta + o(1)](\log N)^2 \deg(D_N) \tag{2.2.27}$$

steps of the random walk is thus asymptotic to

$$q_n \left[\sqrt{2\pi g}\,\mathfrak{c}(\sqrt{\theta}) \int_D \mathrm{e}^{\alpha\sqrt{\theta}(\mathfrak{d}(x)-1)Y}\, Z_{\sqrt{\theta}}^{D,0}(\mathrm{d}x)\right] \widehat{W}_N, \tag{2.2.28}$$

jointly for all $n \ge 0$. Noting that $q_0 = 1$, straightforward limit considerations show:

Theorem 2.2.6 (Avoided points). *Suppose $\{t_N\}_{N\ge1}$ is a sequence such that* (2.2.1) *holds for some $\theta \in (0,1)$. For any $x_N \in D_N$ and for X sampled from P^{x_N}, consider the measure*

$$\kappa_N^D := \frac{1}{\widehat{W}_N} \sum_{x\in D_N} 1_{\{L_{t_N}^{D_N}(x)=0\}}\, \delta_{x/N}, \tag{2.2.29}$$

where $\widehat{W}_N$ *is as in* (2.2.24). *Then, in the sense of vague convergence of measures on* $\overline{D}$,

$$\kappa_N^D \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}} \sqrt{2\pi g}\,\mathfrak{c}(\sqrt{\theta})\ \mathrm{e}^{\alpha\sqrt{\theta}(\mathfrak{d}(x)-1)Y}\ Z^{D,0}_{\sqrt{\theta}}(\mathrm{d}x), \tag{2.2.30}$$

where $Y = \mathcal{N}(0,\sigma_D^2)$ *and* $Z^{D,0}_{\sqrt{\theta}}$ *are independent and* $\mathfrak{c}(\lambda)$ *is as in* (2.2.9).

The above theorems will be deduced from the corresponding statements for a continuous-time variant of X observed for a fixed time of order $N^2(\log N)^2$ (see Propositions 2.5.5, 2.5.9, 2.5.10 and 2.5.11). These statements are nearly identical to Theorems 2.2.3–2.2.6 above, respectively, except for the term $\mathrm{e}^{-\alpha^2\lambda^2/16}$ in (2.2.19) and (2.2.21) that arises from the fluctuations of the (continuous-time) local time at points where the discrete-time local time is large, and the measure μ in (2.2.25) which gets replaced (in Proposition 2.5.10) by a continuous, and quite explicit, counterpart.

The fixed-time results for continuous-time random walk will be inferred from the corresponding results in [1] for the parametrization by the local time at ϱ. The main difference is that the measure (2.2.22) gets replaced by the "pure" LQG Z^D_λ.

2.2.4 Local structure

Similarly as in [1], we are also able to control the local structure of the above exceptional sets. For the thick and thin points, this is achieved by considering the measures on $D\times\mathbb{R}\times\mathbb{R}^{\mathbb{Z}^2}$ of the form

$$\zeta_N^{D,\mathrm{loc}} := \frac{1}{W_N}\sum_{x\in D_N} \delta_{x/N}\otimes\delta_{(L^{D_N}_{t_N}(x)-a_N)/\sqrt{2a_N}}\otimes\delta_{\{(L^{D_N}_{t_N}(x)-L^{D_N}_{t_N}(x+z))/\sqrt{2a_N}\colon\, z\in\mathbb{Z}^2\}}, \tag{2.2.31}$$

where the third coordinate captures the "shape" of the local-time configuration near every exceptional point.

In the parametrization by the local time at the boundary vertex, the asymptotic "law" of the third component in (2.2.31) turned out be that of the pinned DGFF (i.e., the DGFF in $\mathbb{Z}^2\smallsetminus\{0\}$) reduced by a multiple of the potential kernel $\mathfrak{a}$. Here we note that, in our normalization, $\mathfrak{a}$ is the unique non-negative function on $\mathbb{Z}^2$ that is discrete harmonic on $\mathbb{Z}^2 \smallsetminus \{0\}$ and obeys $\mathfrak{a}(0) = 0$ and $\mathfrak{a}(x) \sim g \log |x| + O(1)$ as $|x| \to \infty$. The pinned DGFF ϕ then has the covariance structure

$$\mathrm{Cov}(\phi_x, \phi_y) = \mathfrak{a}(x) + \mathfrak{a}(y) - \mathfrak{a}(x-y). \tag{2.2.32}$$

As it turns out, a different (albeit closely related) Gaussian process arises for the discrete-time walk parametrized by its total time:

Theorem 2.2.7 (Local structure of thick/thin points)**.** *For the setting and under the conditions of Theorem 2.2.3, relative to the vague topology of $\overline{D} \times (\mathbb{R} \cup \{+\infty\}) \times \mathbb{R}^{\mathbb{Z}^2}$,*

$$\zeta_N^{D,\mathrm{loc}} \underset{N\to\infty}{\overset{\mathrm{law}}{\longrightarrow}} \zeta^D \otimes \nu_\lambda, \tag{2.2.33}$$

where ζ^D is the measure on the right of (2.2.19) *and ν_λ is the law of $\widetilde{\phi} + \alpha\lambda\mathfrak{a} - \frac{1}{8}\alpha\lambda 1_{\{0\}^{\mathrm{c}}}$, for $\widetilde{\phi}$ a centered Gaussian process on $\mathbb{Z}^2$ with covariances*

$$\mathrm{Cov}(\widetilde{\phi}_x, \widetilde{\phi}_y) = \mathfrak{a}(x) + \mathfrak{a}(y) - \mathfrak{a}(x-y) - \frac{1}{8}\big[1 - \delta_{x,0} - \delta_{y,0} + \delta_{x,y}\big]. \tag{2.2.34}$$

The same statement (relative to the vague topology on $\overline{D} \times (\mathbb{R} \cup \{-\infty\}) \times \mathbb{R}^{\mathbb{Z}^2}$) holds for the setting of Theorem 2.2.4 except that ν_λ is then the law of $\widetilde{\phi} - \alpha\lambda\mathfrak{a} + \frac{1}{8}\alpha\lambda 1_{\{0\}^{\mathrm{c}}}$.

To demonstrate that $\widetilde{\phi}$ is indeed closely related to the pinned DGFF ϕ, we note that, for $\{n_z \colon z \in \mathbb{Z}^2\}$ i.i.d $\mathcal{N}(0, \frac{1}{8})$ that are independent of $\widetilde{\phi}$,

$$\{\phi_z \colon z \in \mathbb{Z}^d\} \overset{\mathrm{law}}{=} \{\widetilde{\phi}_z + n_0 - n_z \colon z \in \mathbb{Z}^2\}. \tag{2.2.35}$$

We will verify this relation, along with the fact that (2.2.34) is positive semidefinite and thus the covariance of a Gaussian process, in Lemma 2.8.4. The i.i.d. normals appear during a conversion from the continuous-time walk to its discrete-time counterpart. They represent the scaling limit of the fluctuations of the local time due to the random (i.i.d. exponential) nature of the jump times.

We will also address the local time structure in the vicinity of the avoided points. This is done by considering the measure on $D \times [0,\infty)^{\mathbb{Z}^2}$ defined by

$$\kappa_N^{D,\mathrm{loc}} := \frac{1}{\widehat{W}_N} \sum_{x \in D_N} 1_{\{L_{t_N}^{D_N}(x)=0\}} \delta_{x/N} \otimes \delta_{\{L_{t_N}^{D_N}(x+z)\colon z \in \mathbb{Z}^2\}}. \tag{2.2.36}$$

For reasons explained earlier, the measure is concentrated on $D \times (\frac{1}{4}\mathbb{N}_0)^{\mathbb{Z}^2}$.

Recall from [1, Theorem 2.8] that, for the continuous-time random walk parametrized by the local time at the boundary vertex and observed at the time corresponding to θ-multiple of the cover time, the limit distribution of the local configuration is described by the law ν_θ^{RI} of the occupation-time field of random-interlacements at level $u := \pi\theta$. This measure was constructed by Rodriguez [87, Theorems 3.3 and 4.2] (see [1, Section 2.6] for a summary of the construction). For the discrete-time random walk parametrized by its total time we get a discrete-time counterpart of ν_θ^{RI}:

Theorem 2.2.8 (Local structure of avoided points)**.** *For each $u > 0$, there is a unique Borel measure $\nu_u^{\mathrm{RI,\,dis}}$ on $[0,\infty)^{\mathbb{Z}^2}$ that is supported on $(\frac{1}{4}\mathbb{N}_0)^{\mathbb{Z}^2}$ and obeys the following: For*

(1) $\{\ell(z)\colon z \in \mathbb{Z}^2\}$ a sample from $\nu_u^{\mathrm{RI,\,dis}}$, and

(2) $\{\tau_{z,j}\colon z \in \mathbb{Z}^2,\, j \geq 1\}$ independent i.i.d. Exponential(1),

we have

$$\nu_u^{\mathrm{RI}} = \text{law of } \biggl\{ \frac{1}{4} \sum_{j=1}^{4\ell(z)} \tau_{z,j}\colon z \in \mathbb{Z}^2 \biggr\}. \tag{2.2.37}$$

For the setting and under the conditions of Theorem 2.2.6, for each $\theta \in (0,1)$ we then have

$$\kappa_N^{D,\mathrm{loc}} \underset{N\to\infty}{\overset{\mathrm{law}}{\longrightarrow}} \kappa^D \otimes \nu_\theta^{\mathrm{RI,\,dis}} \tag{2.2.38}$$

where κ^D is the measure on the right of (2.2.30)*.*

Similarly as in [1], we will not attempt to make statements concerning the local structure of the light points as that would require developing the corresponding extension of the above occupation-time measure to the situation when the local time at the origin does not vanish.

2.2.5 Remarks

We proceed with a couple of remarks. First note that, along with (2.2.3) and the fact that $Z^{D,0}_{\sqrt{\theta}}$ is supported on all of D a.s., Theorem 2.2.6 implies that the cover time is indeed marked by the choice $\theta := 1$. Second, note that an explicit formula for q_n can be extracted from (2.2.26). This is achieved using the identity

$$\mathrm{e}^{x^2/s} = 1 + \int_0^\infty \frac{x}{2\sqrt{t}}\,\mathrm{e}^t\, I_1(x\sqrt{t})\,\mathrm{e}^{-(1+s/4)t}\,\mathrm{d}t, \tag{2.2.39}$$

where $I_1(z) := \sum_{n\ge 0} \frac{1}{n!(n+1)!}(z/2)^{2n+1}$ is a modified Bessel function. Expanding e^t and $\frac{1}{\sqrt{t}} I_1(x\sqrt{t})$ into power series in t and scaling t by $(1+s/4)$ then readily shows

$$q_{n+1} = n!\sum_{j=0}^{n} \frac{(\alpha^2\theta/8)^{j+1}}{j!(j+1)!(n-j)!} \tag{2.2.40}$$

for each $n \ge 0$. See also (2.4.40) for the corresponding formulas in continuous time.

Third, as we will see in the proofs, the random variable Y in the measure (2.2.22) represents the limit of normalized fluctuations of the local time at the boundary vertex for the first $\lfloor t_N \deg(D_N)\rfloor$ steps of the random walk (see Lemma 2.4.2). A key point is that this becomes statistically independent of the level-set statistics in the limit. Incidentally, through (2.2.28), the total mass of the measure (2.2.22) describes the limit law of a normalized total number of uncovered vertices at the time proportional to λ^2-multiple of the cover time.

Fourth, the reader may wonder why we had to include the degree of ϱ into the normalization of the local time (2.1.3) by $\deg(V)$. This is because, although $\deg(\varrho) = o(|D_N|)$ under (2.1.6–2.1.7) (see Lemma 2.5.8), once the ratio of $\deg(\varrho)/|D_N|$ is larger than $1/\log N$ (which can occur under (2.1.6–2.1.7)) removing $\deg(\varrho)$ from the normalization changes the scaling of the normalization constants W_N and $\widehat{W}_N$ with N.

Fifth, as in [1], the above statements deliberately avoid various boundary values of the parameters; i.e., $\lambda = 1$ for the thick points, $\lambda = \sqrt{\theta} \wedge 1$ for the thin points and $\theta = 1$ for the light and avoided points. All of these are closely related to the statistics of nearly-maximal DGFF values, which is different than the regime described in Theorem 2.2.2. While the nearly-maximal DGFF values are now well understood thanks to the work of the Biskup and Louidor [15, 17, 18] and with Biskup, Gufler and Louidor [19], the recent work of Cortines, Louidor and Saglietti [26] shows that the connection between the avoided points at $\theta = 1$ (i.e., the time scale of the cover time) and the DGFF extrema is considerably more subtle.

Sixth, a natural setting for the above problem is the random walk on a lattice torus $(\mathbb{Z}/(N\mathbb{Z}))^2$ started from any given vertex ϱ. As our work in progress shows [2], the scaling of the corresponding measures is then more complicated — and, in particular, the scaling sequences W_N and $\widehat{W}_N$ have to be taken *random*. This is related to the fact that, for random walks of time-length order $N^2(\log N)^2$, the local time at the starting point of the walk exhibits fluctuations of order $(\log N)^{3/2}$ on the torus while these are only of order $\log N$ at the boundary vertex in our planar domains.

Seventh, we note the recent preprints of Jego [60, 59], where measures of the kind (2.2.6) associated with the thick points of planar Brownian motion run until the first exit from a bounded domain are shown to admit a non-trivial scaling limit that is identified with the limit of multiplicative chaos measures associated with the root of the local time. In [59] the limit measure is shown to obey a list of natural properties that characterize it uniquely. It remains to be seen whether the limit measure bears any connection to Gaussian Free Field and/or Liouville Quantum Gravity.

Finally, we note that Dembo, Peres, Rosen and Zeitouni [30, 31] and Okada [76, 77, 75] analyzed the fractal nature and clustering of the sets of thick points and avoided points in the setting of a random walk killed on exit from D_N (for the thick points) and on two-dimensional torus (for the avoided points). In particular, for $0 < \beta < 1$, the growth exponents have been obtained for

$$\#\Big\{(x_1, x_2) \in D_N \times D_N \colon |x_1 - x_2| \le N^\beta,\, \min\{L^{D_N}_{H_\varrho}(x_1), L^{D_N}_{H_\varrho}(x_2)\} \ge s(\log N)^2\Big\} \quad (2.2.41)$$

with $s > 0$ and

$$\#\Big\{(x_1, x_2) \in D_N \times D_N \colon |x_1 - x_2| \le N^\beta,\, \max\{L^{D_N}_{t_N}(x_1), L^{D_N}_{t_N}(x_2)\} = 0\Big\}, \quad (2.2.42)$$

as well as the sets where "min" and "max" are swapped — which amounts to changing from the behavior near a typical point in the level set to a typical point in D_N. These conclusions cannot be gleaned from our results because $N^{-1+\beta}$ vanishes as $N \to \infty$. Notwithstanding, the obtained exponents coincide with those for the DGFF thick points computed by Daviaud [28] and thus affirm the universality of the DGFF.

2.2.6 Outline

The rest of this paper is organized as follows. In Section 2.3 we derive the scaling limit for the level sets of zero-average DGFF. Section 2.4 extends the conclusions of [1] on the local time parametrized by the local time at ϱ to include information on fluctuations of the total time of the walk. This naturally feeds into Section 2.5 where we establish the scaling limit of exceptional points for the local time of the continuous-time random walk in the parametrization of the total time. Section 2.6 then controls the effect of starting the walk at an arbitrary point. In Section 2.7 we then prove our main theorems above concerning the discrete-time walk except for the local behavior, which is deferred to Section 2.8.

2.3 Zero average DGFF level sets

We are now ready to commence the proofs. As our first item of business, we will address Theorem 2.2.2 on the level sets of the zero-average DGFF. Our strategy is to derive the statement from the unconditional convergence (2.2.9). This leads to a convolution identity whose resolution requires a uniqueness statement that pertains to the whole class of Gaussian Multiplicative Chaos measures:

Theorem 2.3.1. *Given a bounded open set $D \subset \mathbb{R}^d$, let M^D and $\widetilde{M}^D$ be two random a.s.-finite Borel measures on D and let Φ be a centered Gaussian field on D independent of M^D*

and $\widetilde{M}^D$ *such that, for some bounded measurable functions* $\mathfrak{h}_k\colon D\to\mathbb{R}$*,*

$$\operatorname{Cov}\bigl(\Phi(x),\Phi(y)\bigr)=\sum_{k=0}^{\infty}\mathfrak{h}_k(x)\mathfrak{h}_k(y),\quad \text{locally uniformly in } x,y\in D. \tag{2.3.1}$$

Then

$$\mathrm{e}^{\Phi(x)}M^D(\mathrm{d}x)\;\overset{\text{law}}{=}\;\mathrm{e}^{\Phi(x)}\widetilde{M}^D(\mathrm{d}x) \tag{2.3.2}$$

implies $M^D\overset{\text{law}}{=}\widetilde{M}^D$.

We remark that for the needs of the present paper it would suffice to treat the case when the sum in (2.3.1) consists of only one non-zero term. However, this still constitutes the bulk of the proof and so we include the more general case as it is interesting in its own right. The result extends (with suitable modifications) even to the case when Φ is a generalized Gaussian Field; the statement thus "reverse engineers" the base measure from the associated Gaussian Multiplicative Chaos. Our setting goes even somewhat beyond that of, e.g., Shamov [91] as we make no moment assumptions on M^D and $\widetilde{M}^D$.

The proof of Theorem 2.3.1 hinges on the following technical observation:

Lemma 2.3.2. *Let* $\mathfrak{h}\colon D\to\mathbb{R}$ *and* $f\colon D\to[0,\infty)$ *be bounded and measurable and let* M^D *be a random a.s.-finite Borel measure on* D*. Let* $Y=\mathcal{N}(0,1)$ *be independent of* M^D*. Define* $\phi\colon\mathbb{R}\times[0,\infty)\to[0,1]$ *by*

$$\phi(\lambda,t):=E\bigl(\mathrm{e}^{-\langle M^D,\mathrm{e}^{\sqrt{t}\,\mathfrak{h}(\cdot)Y-\lambda\mathfrak{h}(\cdot)}f\rangle}\bigr). \tag{2.3.3}$$

Then ϕ *is continuous on its domain and smooth on the interior thereof. Moreover,* ϕ *satisfies the heat equation,*

$$\frac{\partial\phi}{\partial t}=\frac{1}{2}\frac{\partial^2\phi}{\partial\lambda^2},\qquad(\lambda,t)\in\mathbb{R}\times(0,\infty). \tag{2.3.4}$$

Proof. The continuity of ϕ on $\mathbb{R}\times[0,\infty)$ follows by the Bounded Convergence Theorem. Using that $\sqrt{t}Y=\mathcal{N}(0,t)$ and invoking Tonelli's Theorem we get

$$\phi(\lambda,t)=\int\frac{\mathrm{d}y}{\sqrt{2\pi t}}\,\mathrm{e}^{-\frac{(y-\lambda)^2}{2t}}\phi(y,0). \tag{2.3.5}$$

As $y \mapsto \phi(y,0)$ is bounded, ϕ is continuously differentiable on $\mathbb{R} \times (0,t)$. Since the density of $\mathcal{N}(0,t)$ solves the heat equation (2.3.4), the Dominated Convergence Theorem ensures that so does ϕ. $\square$

We are now ready to give:

Proof of Theorem 2.3.1. Let us first assume that Φ takes the form $\mathfrak{h}(x)Y$ for some bounded measurable $\mathfrak{h}\colon D \to \mathbb{R}$ and $Y = \mathcal{N}(0,1)$ independent of M^D and $\widetilde{M}^D$. Assume that

$$e^{\mathfrak{h}(x)Y} M^D(dx) \overset{\text{law}}{=} e^{\mathfrak{h}(x)Y} \widetilde{M}^D(dx). \tag{2.3.6}$$

Given any bounded and measurable $f\colon D \to [0,\infty)$, let $\phi(\lambda,t)$, resp., $\widetilde{\phi}(\lambda,t)$ denote the functions in (2.3.3) with the random measure M^D, resp., $\widetilde{M}^D$. Since also $x \mapsto e^{-\lambda\mathfrak{h}(x)} f(x)$ is non-negative and measurable, from (2.3.6) we then have

$$\phi(\lambda,1) = \widetilde{\phi}(\lambda,1), \quad \lambda \in \mathbb{R}. \tag{2.3.7}$$

In light of Lemma 2.3.2, the difference $\phi - \widetilde{\phi}$ is a bounded solution to the heat equation in $\mathbb{R} \times (0,\infty)$ with a continuous extension to $\mathbb{R} \times [0,\infty)$. A key point is that the heat equation is known to exhibit *backward uniqueness*. More precisely, Seregin and Šverák [90, Theorem 4.1] implies that every bounded solution to (2.3.4) that vanishes at a given positive time vanishes everywhere. Since (2.3.7) implies that $\phi - \widetilde{\phi}$ vanishes at "time" $t = 1$, we have $\phi = \widetilde{\phi}$ on $\mathbb{R} \times [0,\infty)$. From the equality $\phi(0,0) = \widetilde{\phi}(0,0)$ we then infer

$$E\big(e^{-\langle M^D, f\rangle}\big) = E\big(e^{-\langle \widetilde{M}^D, f\rangle}\big). \tag{2.3.8}$$

Since f was arbitrary, the claim thus holds for any Φ of the form $\mathfrak{h}(\cdot)Y$.

To address the general case, we proceed as in Kahane [61] (see [14, Section 5.2] for a review). First note that by (2.3.1) we may write

$$\Phi(x) \overset{\text{law}}{=} \Phi_n(x) + \sum_{k=0}^{n} \mathfrak{h}_k(x) Y_k, \tag{2.3.9}$$

where $(Y_0,\dots,Y_n)$ are i.i.d. standard normal and where Φ_n is an independent centered Gaussian field with covariance

$$\operatorname{Cov}\bigl(\Phi_n(x),\Phi_n(y)\bigr)=\sum_{k=n+1}^{\infty}\mathfrak{h}_k(x)\mathfrak{h}_k(y). \tag{2.3.10}$$

The argument for Φ of the form $\mathfrak{h}(\cdot)Y$ then shows, inductively, that (2.3.2) implies

$$\mathrm{e}^{\Phi_n(x)}M^D(\mathrm{d}x)\ \overset{\text{law}}{=}\ \mathrm{e}^{\Phi_n(x)}\widetilde{M}^D(\mathrm{d}x),\quad n\in\mathbb{N}. \tag{2.3.11}$$

Letting $f\colon D\to[0,\infty)$ be measurable and supported in a compact set $A\subset D$, the assumption of locally-uniform convergence in (2.3.1) implies that, given $\epsilon>0$ there is $n\in\mathbb{N}$ such that $\operatorname{Var}(\Phi_n(x))\le\epsilon$ for all $x\in A$. This also gives $\operatorname{Cov}(\Phi_n(x),\Phi_n(y))\le\epsilon$ for all $x,y\in A$ and so Kahane's convexity inequality along with Jensen's inequality show, for $Y_\epsilon=\mathcal{N}(0,\epsilon)$ independent of M^D and $\widetilde{M}^D$,

$$\begin{aligned}E\bigl(\mathrm{e}^{-\mathrm{e}^{Y_\epsilon}\langle M^D,f\rangle}\bigr)&=E\bigl(\mathrm{e}^{-\mathrm{e}^{\epsilon/2}\mathrm{e}^{Y_\epsilon-\epsilon/2}\langle M^D,f\rangle}\bigr)\\&\overset{\text{Kahane}}{\ge}E\bigl(\mathrm{e}^{-\mathrm{e}^{\epsilon/2}\langle M^D,\mathrm{e}^{\Phi_n(\cdot)-\frac12\operatorname{Var}(\Phi_n(\cdot))}f\rangle}\bigr)\\&\overset{(2.3.11)}{=}E\bigl(\mathrm{e}^{-\mathrm{e}^{\epsilon/2}\langle\widetilde{M}^D,\mathrm{e}^{\Phi_n(\cdot)-\frac12\operatorname{Var}(\Phi_n(\cdot))}f\rangle}\bigr)\overset{\text{Jensen}}{\ge}E\bigl(\mathrm{e}^{-\mathrm{e}^{\epsilon/2}\langle\widetilde{M}^D,f\rangle}\bigr).\end{aligned} \tag{2.3.12}$$

Taking $\epsilon\downarrow0$ and noting that this implies $Y_\epsilon\to0$ in probability then shows, with the help of the Bounded Convergence Theorem,

$$E\bigl(\mathrm{e}^{-\langle M^D,f\rangle}\bigr)\ge E\bigl(\mathrm{e}^{-\langle\widetilde{M}^D,f\rangle}\bigr). \tag{2.3.13}$$

By symmetry, equality must hold for all f as above and so $M^D\overset{\text{law}}{=}\widetilde{M}^D$, as desired. □

Equipped with Theorem 2.3.1, we are ready to give:

Proof of Theorem 2.2.2. Abbreviate

$$Y_N:=\frac{1}{|D_N|}\sum_{x\in D_N}h_x^{D_N}. \tag{2.3.14}$$

Then Y_N is normal with mean zero and variance

$$\operatorname{Var}(Y_N)=\frac{1}{|D_N|^2}\sum_{x,y\in D_N}G^{D_N}(x,y). \tag{2.3.15}$$

Moreover, denoting

$$\mathfrak{d}_N(x) := \frac{|D_N| \sum_{y \in D_N} G^{D_N}(\lfloor xN \rfloor, y)}{\sum_{y,z \in D_N} G^{D_N}(z,y)} \tag{2.3.16}$$

a covariance calculation shows that Y_N is independent of

$$\widehat{h}^{D_N}_x := h^{D_N}_x - \mathfrak{d}_N(x/N) Y_N \tag{2.3.17}$$

which, we note, has zero average over D_N. Hence, if we define the zero-average variant of η^D_N by

$$\eta^{D,0}_N := \frac{1}{K_N} \sum_{x \in D_N} \delta_{x/N} \otimes \delta_{\widehat{h}^{D_N}_x - \widehat{a}_N}, \tag{2.3.18}$$

we have

$$\eta^{D,0}_N \perp\!\!\!\perp Y_N \quad \text{and} \quad \eta^D_N = \eta^{D,0}_N \circ \theta^{-1}_{\mathfrak{d}_N(\cdot) Y_N}, \tag{2.3.19}$$

where $\theta_{s(\cdot)} \colon D \times \mathbb{R} \to D \times \mathbb{R}$ is defined by $\theta_{s(\cdot)}(x,h) := (x, h + s(x))$. The stated independence also shows

$$\Bigl(\eta^D_N \Bigm| \sum_{x \in D_N} h^{D_N}_x = 0 \Bigr) \overset{\text{law}}{=} \eta^{D,0}_N \tag{2.3.20}$$

and so we may and will henceforth focus on the limit of $\eta^{D,0}_N$.

Using the uniform bound $G^{D_N}(x,y) \le g \log \frac{N}{|x-y|+1} + c$ along with

$$G^{D_N}\bigl(\lfloor xN \rfloor, \lfloor yN \rfloor\bigr) \underset{N \to \infty}{\longrightarrow} \widehat{G}^D(x,y), \quad x, y \in D,\ x \ne y, \tag{2.3.21}$$

the Dominated Convergence shows that $\mathrm{Var}(Y_N)$ converges to σ^2_D from (2.2.15). We thus have $Y_N \overset{\text{law}}{\longrightarrow} Y = \mathcal{N}(0, \sigma^2_D)$. In particular, $\{Y_N \colon N \ge 1\}$ is tight and so from the tightness of η^D_N, (2.3.19) and the uniform boundedness of $\mathfrak{d}_N$ we get

$$\{\eta^{D,0}_N \colon N \ge 1\} \text{ is tight.} \tag{2.3.22}$$

Similarly we show that $\mathfrak{d}_N \to \mathfrak{d}$ uniformly on D. (This implies $\mathfrak{d}(x) \ge 0$). Writing the equality in (2.3.19) via Laplace transforms against a test function $f \in C_c(D \times \mathbb{R})$ and invoking (2.2.9), any subsequential limit $\eta^{D,0}$ of $\{\eta^{D,0}_N \colon N \ge 1\}$ thus obeys

$$\eta^{D,0} \circ \theta^{-1}_{\mathfrak{d}(\cdot) Y} \overset{\text{law}}{=} \mathfrak{c}(\lambda) Z^D_\lambda(\mathrm{d}x) \otimes \mathrm{e}^{-\alpha\lambda h} \mathrm{d}h\,, \tag{2.3.23}$$

where $Y = \mathcal{N}(0, \sigma_D^2)$ is such that $Y \perp\!\!\!\perp \eta^{D,0}$ on the left-hand side.

Next we note that we may realize (2.3.23) as an a.s. equality. This is because (2.3.23) implies, for any measurable $A \subseteq D$ and $B \subseteq \mathbb{R}$ with $\mathrm{Leb}(A) > 0$,

$$\frac{\eta^{D,0} \circ \theta^{-1}_{\mathfrak{d}(\cdot)Y}(A \times B)}{\eta^{D,0} \circ \theta^{-1}_{\mathfrak{d}(\cdot)Y}(A \times [0,1])} = \alpha\lambda(1 - \mathrm{e}^{-\alpha\lambda})^{-1} \int_B \mathrm{e}^{-\alpha\lambda h}\,\mathrm{d}h \quad \text{a.s.} \tag{2.3.24}$$

due to the fact that equality in law to a constant implies equality a.e. We conclude that the measure

$$A \mapsto \alpha\lambda[\mathfrak{c}(\lambda)(1 - \mathrm{e}^{-\alpha\lambda})]^{-1}\, \eta^{D,0} \circ \theta^{-1}_{\mathfrak{d}(\cdot)Y}(A \times [0,1]) \tag{2.3.25}$$

is equidistributed to Z^D_λ. Replacing Z^D_λ by this measure then gives us equality a.s.

Once we have (2.3.23) as an a.s. equality, and Z^D_λ thus as a measurable function of $\eta^{D,0}$ and Y, we apply a routine change of variables to get

$$\eta^{D,0} = \mathfrak{c}(\lambda)\, \mathrm{e}^{-\alpha\lambda\mathfrak{d}(x)Y}\, Z^D_\lambda(\mathrm{d}x) \otimes \mathrm{e}^{-\alpha\lambda h}\mathrm{d}h. \tag{2.3.26}$$

Setting

$$Z^{D,0}_\lambda(\mathrm{d}x) := \mathrm{e}^{-\alpha\lambda\mathfrak{d}(x)Y} Z^D_\lambda(\mathrm{d}x) \tag{2.3.27}$$

the independence of $\eta^{D,0}$ of Y shows $Z^{D,0}_\lambda \perp\!\!\!\perp Y$ and thus proves existence of the decomposition (2.2.16). Since the decomposition is unique by Theorem 2.3.1 and the fact that $\mathfrak{d}$ is bounded and continuous, the law of $Z^{D,0}_\lambda$ does not depend on the subsequential limit $\eta^{D,0}$. It follows that all subsequential limits of $\{\eta^{D,0}_N : N \geq 1\}$ are equal in law and so we get the convergence statement (2.2.14) as well. □

Our use of Theorem 2.2.2 will invariably come through:

Corollary 2.3.3. *Under the conditions of Theorem 2.2.2, and for Y_N as in* (2.3.14)*,*

$$\eta^D_N \otimes \delta_{Y_N} \xrightarrow[N\to\infty]{\text{law}} \mathfrak{c}(\lambda)\, \mathrm{e}^{\alpha\lambda\mathfrak{d}(x)Y}\, Z^{D,0}_\lambda(\mathrm{d}x) \otimes \mathrm{e}^{-\alpha\lambda h}\mathrm{d}h \otimes \delta_Y, \tag{2.3.28}$$

where $Y = \mathcal{N}(0, \sigma_D^2)$, for σ_D^2 as in (2.2.15)*, is such that $Y \perp\!\!\!\perp Z^{D,0}_\lambda$.*

Proof. By (2.3.19) and the fact that $Y_n \to Y$ in law and $\mathfrak{d}_N \to \mathfrak{d}$ uniformly shows

$$\eta_N^D \otimes \delta_{Y_N} \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}} \big(\eta^{D,0} \circ \theta^{-1}_{\mathfrak{d}(\cdot)Y}\big) \otimes \delta_Y, \tag{2.3.29}$$

where $\eta^{D,0}$ is as in (2.3.26) and obeys $Y \perp\!\!\!\perp \eta^{D,0}$. Invoking (2.3.27), the claim follows by a routine change of variables. □

2.4 Augmented boundary vertex measures

We will now move to the discussion of local time level sets. Our proofs build on the conclusions derived in [1] for the local time parametrized by its value at the boundary vertex ϱ. In order to transfer these conclusions to the setting of a fixed total time, we will need to control the fluctuations of the total local time at a fixed local time at ϱ. Our first step is thus to augment the results of [1] by information about these fluctuations.

We will again introduce the corresponding quantities on a general finite connected graph with vertex set $V \cup \{\varrho\}$. Consider a joint law of paths X of the discrete-time random walk on $V \cup \{\varrho\}$ and an independent sample $t \mapsto \widetilde{N}(t)$ of a rate-1 Poisson process. The continuous-time walk is then defined as

$$\widetilde{X}_t := X_{\widetilde{N}(t)}, \quad t \geq 0. \tag{2.4.1}$$

The local time naturally associated with $\widetilde{X}$ is given by

$$\widetilde{L}_t^V(u) := \frac{1}{\deg(u)} \int_0^t \mathrm{d}s \, 1_{\{\widetilde{X}_s = u\}}. \tag{2.4.2}$$

Denoting $\hat{\tau}_\varrho(t) := \inf\{s \geq 0 \colon \widetilde{L}_s^V(\varrho) \geq t\}$, the local time parametrized by its value at ϱ is defined as

$$\widehat{L}_t^V(v) := \widetilde{L}^V_{\hat{\tau}_\varrho(t)}(v). \tag{2.4.3}$$

Note that, in particular, we have $\widehat{L}_t^V(\varrho) = t$ for all $t \geq 0$. The same is true about the expected value at any vertex; i.e., $E^\varrho \widehat{L}_t^V(v) = t$ for all $v \in V$.

At a given $t \geq 0$, the total (continuous) local time of the walk is computed by adding $\widehat{L}_t^V(v)$ over all $v \in V \cup \{\varrho\}$. The quantity

$$T(t) := \frac{1}{\sqrt{2t}\,|V|} \sum_{v \in V} \bigl[\,\widehat{L}_t^V(v) - t\bigr] \tag{2.4.4}$$

then denotes the normalized (empirical) fluctuation of the total local time. (Note that $v = \varrho$ can be freely added to the sum as $\widehat{L}_t^V(\varrho) = t$.) To explain the specific choice of the normalization, we recall the following result from Eisenbaum, Kaspi, Marcus, Rosen and Shi [43](with improvements by Zhai [100, Section 5.4]):

Theorem 2.4.1 (Second Ray-Knight Theorem)**.** *For each $t > 0$ there exists a coupling of $\widehat{L}_t^V$ (sampled under P^ϱ) and two copies of the DGFF h^V and $\tilde{h}^V$ such that*

$$h^V \text{ and } \widehat{L}_t^V \text{ are independent} \tag{2.4.5}$$

and

$$\widehat{L}_t^V(u) + \frac{1}{2}(h_u^V)^2 = \frac{1}{2}\bigl(\tilde{h}_u^V + \sqrt{2t}\bigr)^2, \quad u \in V. \tag{2.4.6}$$

Using the stated coupling, we readily compute

$$T(t) = \frac{1}{|V|}\sum_{u \in V} \tilde{h}_u^V + \frac{1}{\sqrt{2t}\,|V|}\sum_{u \in V} \frac{(\tilde{h}_u^V)^2 - (h_u^V)^2}{2}. \tag{2.4.7}$$

Note that the first term is the average of the field $\tilde{h}^V$.

In what follows, the role of V will be taken by the sets D_N and ϱ by the "boundary vertex." We let h^{D_N} be the DGFF on D_N and, given a sequence $\{t_N\}_{N \geq 1}$ and for the continuous-time random walk started at ϱ, let $\tilde{h}^{D_N}$ be the DGFF such that (2.4.5–2.4.6) with $t := t_N$ holds. We then set

$$T_N := \frac{1}{\sqrt{2t_N}\,|D_N|} \sum_{x \in D_N} \bigl[\,\widehat{L}_{t_N}^{D_N}(x) - t_N\bigr] \tag{2.4.8}$$

and denote

$$Y_N := \frac{1}{|D_N|} \sum_{x \in D_N} \tilde{h}_x^{D_N}. \tag{2.4.9}$$

We start by noting:

Lemma 2.4.2. *For any $\{t_N\}_{N\geq 1}$ with $t_N \to \infty$ we have*

$$T_N - Y_N \underset{N\to\infty}{\longrightarrow} 0, \quad \text{in probability.} \tag{2.4.10}$$

In particular,

$$T_N \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}} \mathcal{N}(0,\sigma_D^2), \tag{2.4.11}$$

where σ_D^2 is as in (2.2.15).

Proof. The Wick Pairing Theorem gives

$$\begin{aligned}\text{Var}\Big(\sum_{x\in D_N}(h_x^{D_N})^2\Big) &= \sum_{x,y\in D_N}\text{Cov}\big((h_x^{D_N})^2,(h_y^{D_N})^2\big)\\ &= \sum_{x,y\in D_N} 2\left[E(h_x^{D_N}h_y^{D_N})\right]^2 = 2\sum_{x,y\in D_N}G^{D_N}(x,y)^2.\end{aligned} \tag{2.4.12}$$

The uniform bound $G^{D_N}(x,y) \leq g\log\frac{N}{|x-y|+1} + c$ shows that the double sum on the right is of order $|D_N|^2$. From $t_N \to \infty$ it follows that

$$\frac{1}{\sqrt{2t_N}\,|D_N|}\sum_{x\in D_N}\left[(h_x^{D_N})^2 - \mathbb{E}[(h_x^{D_N})^2]\right] \underset{N\to\infty}{\longrightarrow} 0, \quad \text{in probability.} \tag{2.4.13}$$

Using this along with $\mathbb{E}[(h_x^{D_N})^2] = \mathbb{E}[(\tilde{h}_x^{D_N})^2]$ in (2.4.7), we get (2.4.10). For (2.4.11) we invoke the argument after (2.3.21). □

We are now ready to state and prove convergence theorems for processes associated with exceptional level sets of the boundary vertex local time $\widehat{L}_{t_N}^{D_N}$ augmented by information about T_N. Starting with the thick and thin points, given positive sequences $\{t_N\}_{N\geq 1}$ and $\{a_N\}_{N\geq 1}$, define

$$\widehat{\zeta}_N^D := \frac{1}{W_N}\sum_{x\in D_N}\delta_{x/N}\otimes\delta_{(\widehat{L}_{t_N}^{D_N}(x)-a_N)/\sqrt{2a_N}}\,, \tag{2.4.14}$$

where W_N is as in (2.2.17). For the thick points of $\widehat{L}_{t_N}^{D_N}$, we then have:

Proposition 2.4.3 (Thick points)**.** *Suppose that $\{t_N\}_{N\geq 1}$ and $\{a_N\}_{N\geq 1}$ are such* (2.2.1) *and* (2.2.18) *hold for some $\theta > 0$ and $\lambda \in (0,1)$. Then for X sampled from P^{ϱ}, relative to* *the vague convergence of measures on* $\overline{D}\times(\mathbb{R}\cup\{+\infty\})\times\mathbb{R}$*,*

$$\widehat{\zeta}^D_N\otimes\delta_{T_N}\ \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}}\ \sqrt{\frac{\sqrt{\theta}}{\sqrt{\theta}+\lambda}}\ \mathfrak{c}(\lambda)\,\mathrm{e}^{\alpha\lambda\mathfrak{d}(x)Y}Z^{D,0}_\lambda(\mathrm{d}x)\otimes\mathrm{e}^{-\alpha\lambda h}\mathrm{d}h\otimes\delta_Y(\mathrm{d}t) \tag{2.4.15}$$

where $Y=\mathcal{N}(0,\sigma_D^2)$*, for* σ_D^2 *as in* (2.2.15)*, is such that* $Y\perp\!\!\!\perp Z^{D,0}_\lambda$*.*

Proof. We will rely heavily on the proof of [1, Theorem 2.2] but, due to a different normalization of the second coordinate in (2.4.14) and also the fact that the limit measure is different than in [1], we need to recount the main steps of the proof. Throughout we will assume (for each $N\ge1$ and each $t:=t_N$) a coupling of $\widehat{L}^{D_N}_{t_N}$ and an independent DGFF h^{D_N} to a DGFF $\tilde h^{D_N}$ satisfying (2.4.6).

First, by [1, Corollary 4.2] the measures $\{\widehat{\zeta}^D_N\colon N\ge1\}$ are tight and, by Lemma 2.4.2, the same applies to the enhanced measures $\{\xi_N\colon N\ge1\}$ where

$$\xi_N:=\widehat{\zeta}^D_N\otimes\delta_{T_N}. \tag{2.4.16}$$

Moreover, [1, Lemma 5.3] shows that if $\xi_{N_k}\to\xi$ in law along some increasing sequence $\{N_k\}_{k\ge1}$, then the extended measures

$$\xi_N^{\text{ext}}:=\frac{1}{W_N}\sum_{x\in D_N}\delta_{x/N}\otimes\delta_{(\widehat{L}^{D_N}_{t_N}(x)-a_N)/\sqrt{2a_N}}\otimes\delta_{T_N}\otimes\delta_{h^{D_N}_x/(2a_N)^{1/4}}\,, \tag{2.4.17}$$

where we now normalize the third coordinate differently than in [1], obey

$$\xi^{\text{ext}}_{N_k}\ \underset{k\to\infty}{\overset{\text{law}}{\longrightarrow}}\ \xi\otimes\mathfrak{g} \tag{2.4.18}$$

in which, using (2.2.18), $\mathfrak{g}$ is the law of $\mathcal{N}(0,\frac{1}{\alpha\,(\sqrt{\theta}+\lambda)})$.

Let η^D_N be the process (2.2.7) associated with the field $\tilde h^{D_N}$ and the scale function

$$\widehat{a}_N:=\sqrt{2a_N}-\sqrt{2t_N} \tag{2.4.19}$$

that, by (2.2.1) and (2.2.18), scales as $\widehat{a}_N\sim2\sqrt{g}\,\lambda\log N$ as $N\to\infty$. Let Y_N be the average of $\tilde h^{D_N}$ over D_N; cf (2.4.9). Given $f\in C_c(D\times\mathbb{R}\times\mathbb{R})$, in the assumed coupling of $\widehat{L}^{D_N}_{t_N}$, h^{D_N} and $\tilde h^{D_N}$, the convergence in Lemma 2.4.2 tells us

$$\langle\eta^D_N\otimes\delta_{Y_N},f\rangle=o(1)+\langle\eta^D_N\otimes\delta_{T_N},f\rangle, \tag{2.4.20}$$

where $o(1) \to 0$ as $N \to \infty$ in probability. The calculation in the proof of [1, Lemma 5.4] (enabled by the fact that the field h^{D_N} will be typical at most points contributing to ζ_N^D, as shown in [1, Lemma 5.2]) then gives

$$\langle \eta_N^D \otimes \delta_{T_N}, f\rangle = o(1) + \langle \xi_N^{\rm ext}, f^{\rm ext}\rangle, \tag{2.4.21}$$

where

$$f^{\rm ext}(x,\ell,t,h) := f\bigl(x, \ell + \tfrac{h^2}{2}, t\bigr). \tag{2.4.22}$$

Using Corollary 2.3.3 on the left-hand side of (2.4.20), from (2.4.21) and (2.4.18) and, one more time, [1, Lemma 5.2] we conclude that every subsequential limit ξ of the measures in (2.4.16) satisfies the convolution-type identity

$$\langle \xi, f^{*\mathfrak{g}}\rangle \overset{\rm law}{=} \mathfrak{c}(\lambda) \int \mathrm{e}^{\alpha\lambda\mathfrak{d}(x)Y}\, Z_\lambda^{D,0}(\mathrm{d}x) \otimes \mathrm{e}^{-\alpha\lambda\ell}\mathrm{d}\ell\; f(x,\ell,Y), \tag{2.4.23}$$

where $Y \perp\!\!\!\perp Z_\lambda^{D,0}$ and

$$f^{*\mathfrak{g}}(x,\ell,t) := \int \mathfrak{g}(\mathrm{d}h) f\bigl(x, \ell + \tfrac{h^2}{2}, t\bigr), \tag{2.4.24}$$

jointly for all $f \in C_{\rm c}(D \times \mathbb{R} \times \mathbb{R})$. It remains to "solve" (2.4.23) for ξ.

First we note that the Monotone Convergence Theorem extends (2.4.23) to all f of the form $f(x,\ell,t) := 1_A(x)\tilde f(\ell)1_{(b,\infty)}(t)$, where $\tilde f \in C_{\rm c}(\mathbb{R})$ and where $A \subseteq D$ is non-empty and open. Denoting $\xi_{A,b}(B) := \xi(A \times B \times (b,\infty))$, a calculation then shows

$$\langle \xi, f^{*\mathfrak{g}}\rangle = \langle \xi_{A,b}, \tilde f * \mathfrak{e}\rangle \tag{2.4.25}$$

where

$$\mathfrak{e}(z) := \sqrt{\frac{\beta}{\pi}}\,\frac{\mathrm{e}^{\beta z}}{\sqrt{-z}} 1_{(-\infty,0)}(z) \quad \text{for} \quad \beta := \alpha\bigl(\sqrt{\theta} + \lambda\bigr). \tag{2.4.26}$$

The identity (2.4.23) also implies that $\langle \xi_{A,b}, 1_{[0,\infty)}\rangle < \infty$ a.s. and gives

$$\langle \xi_{A,b}, \tilde f * \mathfrak{e}\rangle = \langle \xi_{A,b}, 1_{[0,\infty)} * \mathfrak{e}\rangle \int \alpha\lambda\, \mathrm{e}^{-\alpha\lambda\ell} \tilde f(\ell)\mathrm{d}\ell, \tag{2.4.27}$$

where the equality now holds pointwise a.s. because once $\langle \xi_{A,b}, 1_{[0,\infty)} * \mathfrak{e}\rangle > 0$ (which is necessary for the left-hand side to be non-zero), the ratio $\langle \xi_{A,b}, \tilde f * \mathfrak{e}\rangle / \langle \xi_{A,b}, 1_{[0,\infty)} * \mathfrak{e}\rangle$ is equal in law, and thus pointwise, to the integral on the right.

Denoting $\mu_\lambda(\mathrm{d}h) := \mathrm{e}^{-\alpha\lambda h}\mathrm{d}h$, a routine change of variables rewrites (2.4.27) as

$$\langle \xi_{A,b}, \tilde f * \mathfrak{e}\rangle = C\langle \mu_\lambda, \tilde f\rangle \tag{2.4.28}$$

where C is a random constant that is finite thanks to $\beta > \alpha\lambda$. By [1, Lemma 5.5], there is at most one Borel measure $\xi_{A,b}$ on $\mathbb{R}$ satisfying (2.4.28) and, in fact, $\xi_{A,b}(\mathrm{d}\ell) = C_{A,b}\mathrm{e}^{-\alpha\lambda\ell}\mathrm{d}\ell$ for some (random) constant $C_{A,b}$. It follows that

$$\xi(\mathrm{d}x\mathrm{d}\ell\mathrm{d}t) = M(\mathrm{d}x\mathrm{d}t) \otimes \mathrm{e}^{-\alpha\lambda\ell}\mathrm{d}\ell, \tag{2.4.29}$$

where, by plugging this in (2.4.23),

$$M(\mathrm{d}x\mathrm{d}t) \overset{\text{law}}{=} \Bigl(\int \mathfrak{g}(\mathrm{d}h)\mathrm{e}^{\alpha\lambda\frac{h^2}{2}}\Bigr)^{-1}\mathfrak{c}(\lambda)\mathrm{e}^{\alpha\lambda\mathfrak{d}(x)Y}\, Z_\lambda^{D,0}(\mathrm{d}x) \otimes \delta_Y(\mathrm{d}t). \tag{2.4.30}$$

The integral equals the root of $(\sqrt{\theta}+\lambda)/\sqrt{\theta}$. The claim follows. □

We proceed with the corresponding result for the thin points:

Proposition 2.4.4 (Thin points)**.** *Suppose that* $\{t_N\}_{N\ge1}$ *and* $\{a_N\}_{N\ge1}$ *are such* (2.2.1) *and* (2.2.20) *hold for some* $\theta > 0$ *and* $\lambda \in (0, \sqrt{\theta}\wedge 1)$. *Then for* X *sampled from* P^ϱ, *relative to the vague convergence of measures on* $\overline{D}\times(\mathbb{R}\cup\{-\infty\})\times\mathbb{R}$,

$$\widehat{\zeta}_N^D \otimes \delta_{T_N} \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}} \sqrt{\frac{\sqrt{\theta}}{\sqrt{\theta}-\lambda}}\; \mathfrak{c}(\lambda)\,\mathrm{e}^{-\alpha\lambda\mathfrak{d}(x)Y} Z_\lambda^{D,0}(\mathrm{d}x) \otimes \mathrm{e}^{+\alpha\lambda h}\mathrm{d}h \otimes \delta_Y(\mathrm{d}t) \tag{2.4.31}$$

where $Y \perp\!\!\!\perp Z_\lambda^{D,0}$ *with* $Y = \mathcal{N}(0,\sigma_D^2)$, *for* σ_D^2 *as in* (2.2.15).

Proof. The proof is very similar to that of Proposition 2.4.3 so we indicate only the needed changes. We will again rely on the coupling of $\widehat{L}_{t_N}^{D_N}$ and two DGFFs h^{D_N} and $\tilde h^{D_N}$ such that (2.4.5–2.4.6) for $t := t_N$ hold. Let η_N^D to denote the process associated with $\tilde h^{D_N}$ and the centering sequence $-\widehat{a}_N$, where

$$\widehat{a}_N := \sqrt{2t_N} - \sqrt{2a_N}. \tag{2.4.32}$$

Note that, under (2.2.1) and (2.2.20) we have $\widehat{a}_N \sim 2\sqrt{g}\lambda\log N$. Writing Y_N for the average of $\tilde h^{D_N}$ over D_N, Corollary 2.3.3 along with the symmetry $h^{D_N} \overset{\text{law}}{=} -h^{D_N}$ ensures

$$\eta_N^D \otimes \delta_{Y_N} \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}} \mathfrak{c}(\lambda)\,\mathrm{e}^{-\lambda\alpha\mathfrak{d}(x)Y}\, Z_\lambda^{D,0}(\mathrm{d}x) \otimes \mathrm{e}^{+\alpha\lambda h}\mathrm{d}h \otimes \delta_Y(\mathrm{d}t), \tag{2.4.33}$$

where $Y=\mathcal{N}(0,\sigma_D^2)$ is independent of $Z_\lambda^{D,0}$.

The argument now proceeds very much like for the thick points. We consider the extended measures (2.4.17), which are tight by [1, Corollary 4.8] and show, with the help of [1, Lemmas 6.1, 6.2] and (2.4.33), that every subsequential limit ξ thereof obeys

$$\langle\xi,f^{*\mathfrak{g}}\rangle \overset{\text{law}}{=} \mathfrak{c}(\lambda)\int \mathrm{e}^{-\alpha\lambda\mathfrak{d}(x)Y}\,Z_\lambda^{D,0}(\mathrm{d}x)\otimes\mathrm{e}^{+\alpha\lambda\ell}\mathrm{d}\ell\; f(x,\ell,Y), \tag{2.4.34}$$

where $f^{*\mathfrak{g}}$ is still defined via (2.4.24) but with

$$\mathfrak{g}:=\text{law of }\mathcal{N}\big(0,\tfrac{1}{\alpha(\sqrt{\theta}-\lambda)}\big). \tag{2.4.35}$$

The identity (2.4.34) readily extends to all f of the form $f(x,\ell,t):=1_A(x)\tilde f(\ell)1_{(-\infty,b)}(t)$, where $\tilde f\in C_{\mathrm{c}}(\mathbb{R})$ and where $A\subseteq D$ is non-empty and open. A calculation then shows (2.4.25) with $\mathfrak{e}$ now defined using $\beta:=\alpha(\sqrt{\theta}-\lambda)$. Proceeding via an analogue of (2.4.27) (with $1_{[0,\infty)}$ replaced by $1_{(-\infty,0]}$), using [1, Lemma 6.4] we then again show

$$\xi(\mathrm{d}x\mathrm{d}\ell\mathrm{d}t)=M(\mathrm{d}x\mathrm{d}t)\otimes\mathrm{e}^{+\alpha\lambda\ell}\mathrm{d}\ell, \tag{2.4.36}$$

where, this time,

$$M(\mathrm{d}x\mathrm{d}t)\overset{\text{law}}{=}\Big(\int\mathfrak{g}(\mathrm{d}h)\mathrm{e}^{-\alpha\lambda\frac{h^2}{2}}\Big)^{-1}\mathfrak{c}(\lambda)\mathrm{e}^{-\alpha\lambda\mathfrak{d}(x)Y}\,Z_\lambda^{D,0}(\mathrm{d}x)\otimes\delta_Y(\mathrm{d}t). \tag{2.4.37}$$

The integral equals the root of $(\sqrt{\theta}-\lambda)/\sqrt{\theta}$. □

Next we move to the discussion of the light and avoided points. Starting with the light points, we define

$$\widehat{\vartheta}_N^D:=\frac{1}{\widehat{W}_N}\sum_{x\in D_N}\delta_{x/N}\otimes\delta_{\widehat{L}_{t_N}^{D_N}(x)}, \tag{2.4.38}$$

where $\widehat{W}_N$ is as in (2.2.24). We then get:

Proposition 2.4.5 (Light points). *Suppose $\{t_N\}_{N\ge1}$ obeys (2.2.1) for some $\theta\in(0,1)$. Then, for the random walk sampled from P^ϱ, in the sense of vague convergence of measures on $\overline{D}\times[0,\infty)\times\mathbb{R}$,*

$$\widehat{\vartheta}_N^D\otimes\delta_{T_N}\;\underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}}\;\sqrt{2\pi g}\,\mathfrak{c}(\sqrt{\theta})\;\mathrm{e}^{-\alpha\sqrt{\theta}\,\mathfrak{d}(x)Y}\,Z_{\sqrt{\theta}}^{D,0}(\mathrm{d}x)\otimes\tilde\mu(\mathrm{d}h)\otimes\delta_Y(\mathrm{d}t), \tag{2.4.39}$$

where $Y = \mathcal{N}(0,\sigma_D^2)$ is independent of $Z^{D,0}_{\sqrt{\theta}}$ and

$$\tilde{\mu}(\mathrm{d}h) := \delta_0(\mathrm{d}h) + \Biggl(\sum_{n=0}^{\infty}\frac{1}{n!(n+1)!}\Bigl(\frac{\alpha^2\theta}{2}\Bigr)^{n+1}h^n\Biggr)1_{(0,\infty)}(h)\,\mathrm{d}h. \tag{2.4.40}$$

Proof. Assuming again the coupling from (2.4.5–2.4.6), we set

$$\xi_N := \widehat{\vartheta}^D_N \otimes \delta_{T_N}. \tag{2.4.41}$$

The family $\{\xi_N\colon N \ge 1\}$ is tight by [1, Corollary 4.6] and so we may consider a subsequential limit ξ thereof. By [1, Lemma 7.1], the extended measure

$$\xi_N^{\text{ext}} := \frac{\sqrt{\log N}}{\widehat{W}_N}\sum_{x\in D_N}\delta_{x/N}\otimes\delta_{\widehat{L}^{D_N}_{t_N}(x)}\otimes\delta_{T_N}\otimes\delta_{h^{D_N}_x}, \tag{2.4.42}$$

then converges to $\xi \otimes \frac{1}{\sqrt{2\pi g}}\mathrm{Leb}$ along the same subsequence. We now pick a test function $f \in C_{\mathrm{c}}(D\times[0,\infty)\times\mathbb{R})$, denote

$$f^{\text{ext}}(x,\ell,t,h) := f\bigl(x,\ell+\tfrac{h^2}{2},t\bigr) \tag{2.4.43}$$

and observe that (2.4.6) implies

$$\sum_{x\in D_N} f^{\text{ext}}\Bigl(x/N, \widehat{L}^{D_N}_{t_N}(x), T_N, h^{D_N}_x\Bigr) = \sum_{x\in D_N} f\Bigl(x/N, \tfrac{1}{2}\bigl(\tilde{h}^{D_N}_x+\sqrt{2t_N}\bigr)^2, T_N\Bigr). \tag{2.4.44}$$

Writing this in terms of the above measures, Lemma 2.4.2 gives

$$\langle \xi_N^{\text{ext}}, f^{\text{ext}}\rangle = o(1) + \bigl\langle \eta^D_N\otimes\delta_{Y_N}, f(\ \cdot\ ,\tfrac{1}{2}|\cdot|^2,\ \cdot\ )\bigr\rangle, \tag{2.4.45}$$

where η^D_N is the DGFF process associated with the scale sequence $\widehat{a}_N := -\sqrt{2t_N}$. As $\widehat{a}_N \sim -2\sqrt{g}\sqrt{\theta}\log N$, from (2.4.33) we get

$$\langle \xi, f^{*\text{Leb}}\rangle \overset{\text{law}}{=} \mathfrak{c}(\sqrt{\theta})\int \mathrm{e}^{-\alpha\sqrt{\theta}\,\mathfrak{d}(x)Y}\,Z^{D,0}_{\sqrt{\theta}}(\mathrm{d}x)\otimes\mathrm{e}^{+\alpha\sqrt{\theta}\,h}\mathrm{d}h\ f\bigl(x,\tfrac{1}{2}h^2,Y\bigr), \tag{2.4.46}$$

where

$$f^{*\text{Leb}}(x,\ell,t) := \frac{1}{\sqrt{2\pi g}}\int \mathrm{d}h\, f\bigl(x,\ell+\tfrac{h^2}{2},t\bigr). \tag{2.4.47}$$

By the Monotone Convergence Theorem, this extends to all f of the form

$$f(x,\ell,t) := 1_A(x)\mathrm{e}^{-s\ell}1_{[0,\infty)}(\ell)1_{[b,\infty)}(t) \tag{2.4.48}$$

for $A\subseteq D$ open, $b\in\mathbb{R}$ and $s>0$. For $\xi_{A,b}(B):=\xi(A\times B\times[b,\infty))$, we then get

$$\int_0^\infty \xi_{A,b}(\mathrm{d}\ell)\mathrm{e}^{-s\ell} \overset{\text{law}}{=} \sqrt{2\pi g}\,\mathfrak{c}(\sqrt{\theta})\Bigl(\int_A \mathrm{e}^{-\alpha\sqrt{\theta}\,\mathfrak{d}(x)Y}\, Z^{D,0}_{\sqrt{\theta}}(\mathrm{d}x)\Bigr)\,\mathrm{e}^{\frac{\alpha^2\theta}{2s}}\,1_{[b,\infty)}(Y). \tag{2.4.49}$$

Since the Laplace transform of a measure, if exists, determines the measure uniquely, this proves that ξ takes the product form

$$\xi \overset{\text{law}}{=} \sqrt{2\pi g}\,\mathfrak{c}(\sqrt{\theta})\,\mathrm{e}^{-\alpha\sqrt{\theta}\,\mathfrak{d}(x)Y}\, Z^{D,0}_{\sqrt{\theta}}(\mathrm{d}x)\otimes\tilde{\mu}(\mathrm{d}\ell)\otimes\delta_Y(\mathrm{d}t) \tag{2.4.50}$$

for some deterministic measure $\tilde{\mu}$ on $[0,\infty)$ with Laplace transform $s\mapsto \mathrm{e}^{\frac{\alpha^2\theta}{2s}}$. A calculation shows that the measure (2.4.40) has this property. □

A direct consequence of our control of the light points is:

Proposition 2.4.6 (Avoided points)**.** *Suppose* $\{t_N\}_{N\ge1}$ *is such that* (2.2.1) *holds for some* $\theta\in(0,1)$ *and let*

$$\widehat{\kappa}^D_N := \frac{1}{\widehat{W}_N}\sum_{x\in D_N} 1_{\{\widehat{L}^{D_N}_{t_N}(x)=0\}}\delta_{x/N}. \tag{2.4.51}$$

Then, for the random walk distributed according to P^ϱ*, in the sense of vague convergence of measures on* $\overline{D}\times\mathbb{R}$*,*

$$\widehat{\kappa}^D_N\otimes\delta_{T_N} \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}} \sqrt{2\pi g}\,\mathfrak{c}(\sqrt{\theta})\;\mathrm{e}^{-\alpha\sqrt{\theta}\,\mathfrak{d}(x)Y}\, Z^{D,0}_{\sqrt{\theta}}(\mathrm{d}x)\otimes\delta_Y(\mathrm{d}t), \tag{2.4.52}$$

where $Y=\mathcal{N}(0,\sigma_D^2)$ *is independent of* $Z^{D,0}_{\sqrt{\theta}}$*.*

Proof. The proof of [1, Theorem 2.5] carries over essentially *verbatim.* □

2.5 Fixed total time

Equipped with the enhanced limit results that include the limit value of suitably-normalized fluctuations of the total local time, we now proceed to derive from these the corresponding conclusions for a fixed total time. We keep working with the random walk started at the boundary vertex ϱ; general starting points will be dealt with in Section 2.6.

2.5.1 Time conversions

The transition from a fixed local time at ϱ to a fixed total time is based on a simple inversion formula. Recall that, in our context,

$$\hat{\tau}_\varrho(t) := \inf\bigl\{s \ge 0 \colon \widetilde{L}_s^{D_N}(\varrho) \ge t\bigr\} \tag{2.5.1}$$

and $\deg(D_N) = \sum_{x\in D_N\cup\{\varrho\}} \deg(x)$. Given a sequence $\{t_N\}_{N\ge1}$ with $t_N \ge 1$, define

$$t_N^\star = \inf\bigl\{t \ge 0 \colon \hat{\tau}_\varrho(t) \ge \deg(D_N)t_N\bigr\}. \tag{2.5.2}$$

This is an inverse of $\hat{\tau}_\varrho$ evaluated at $\deg(D_N)t_N$ and so we expect $\hat{\tau}_\varrho(t_N^\star) \approx \deg(D_N)t_N$. By (2.1.3) and (2.4.3), we should therefore have $\widetilde{L}^{D_N}_{\deg(D_N)t_N}(\cdot) \approx \widehat{L}^{D_N}_{t_N^\star}(\cdot)$. Besides their approximate nature, any use of these identifications are complicated by the appearance of the random time $t_N^\star$ for which we have no better formula than (2.5.2). We will thus base the time conversion on a slightly different (still random) quantity that will turn out to be better adapted to our needs.

Recall the definition of T_N from (2.4.8). We note that this actually coincides with the value of $T_N(t_N)$, where (in accord with (2.4.4)) we set

$$T_N(t) := \frac{U_N(t)}{\sqrt{2t}} \qquad \text{for} \qquad U_N(t) := \frac{1}{|D_N|}\sum_{x\in D_N}\bigl[\,\widehat{L}_t^{D_N}(x) - t\bigr]. \tag{2.5.3}$$

Now let

$$t_N^\circ := t_N - \left(1 - \frac{\deg(\varrho)}{\deg(D_N)}\right)\sqrt{2t_N}\,T_N(t_N). \tag{2.5.4}$$

We then have:

Proposition 2.5.1 (Time conversion)**.** *Fix any sequence $(b_N)_{N\ge1}$ in $(0,\infty)$ such that $b_N \to \infty$ and $b_N/t_N^{1/4} \to 0$ as $N \to \infty$. Then there exist constants $c_1 > 0$ such that*

$$\hat{\tau}_\varrho\bigl(t_N^\circ - b_N t_N^{1/4}\bigr) \le \deg(D_N)t_N \le \hat{\tau}_\varrho\bigl(t_N^\circ + b_N t_N^{1/4}\bigr) \tag{2.5.5}$$

and thus, in particular,

$$\widehat{L}^{D_N}_{t_N^\circ - b_N t_N^{1/4}}(\cdot) \le \widetilde{L}^{D_N}_{\deg(D_N)t_N}(\cdot) \le \widehat{L}^{D_N}_{t_N^\circ + b_N t_N^{1/4}}(\cdot) \tag{2.5.6}$$

hold true with P^ϱ-probability at least $1 - c_1 b_N^{-1}$.

The proof will be split into several intermediate results, some of which will be useful later as well. The first item to note is the "stability" (or slow variation) of the fluctuation of the total local time:

Lemma 2.5.2. *There exists a constant $c_2 > 0$ such that for all $s, t \geq 0$ and all $r > 0$,*

$$P^{\varrho}\left(\sup_{0\leq u\leq t} |U_N(s+u) - U_N(s)| \geq r\right) \leq \frac{c_2 t}{r^2}. \tag{2.5.7}$$

Proof. Note that U_N is a compensated compound Poisson process. In view of stationarity, it suffices to consider the case $s = 0$. Moreover, since U_N is a martingale, Doob's maximal inequality is applicable and hence

$$P^{\varrho}\left(\sup_{0\leq u\leq t} |U_N(u)| \geq r\right) \leq \frac{4\,\mathrm{Var}_{P^{\varrho}}(U_N(t))}{r^2}. \tag{2.5.8}$$

It suffices to show that $\mathrm{Var}_{P^{\varrho}}(U_N(t))$ is bounded by Ct for some $C > 0$. To this end, we note that $t \mapsto (U_N(t) + t)$ is a compound Poisson process with rate $\deg(\varrho)$ and jump size distributed as $\sum_{x\in D_N} \ell(x)/|D_N|$, where $\ell(\cdot)$ is the local time for a single excursion. Hence,

$$\mathrm{Var}_{P^{\varrho}}(U_N(t)) = \mathrm{Var}_{P^{\varrho}}(U_N(t) + t) = \frac{1}{|D_N|^2}\deg(\varrho)t\, E^{\varrho}\left[\left(\sum_{x\in D_N} \ell(x)\right)^2\right]. \tag{2.5.9}$$

The last expectation can be computed via the Kac moment formula,

$$\mathrm{Var}_{P^{\varrho}}(U_N(t)) = \frac{2t}{|D_N|^2}\sum_{x,y\in D_N} G^{D_N}(x,y). \tag{2.5.10}$$

The uniform bound $G^{D_N}(x,y) \leq g\log\frac{N}{|x-y|+1} + c$ shows that the sum is at most a constant times $|D_N|^2$, uniformly in $N \geq 1$. □

The next lemma quantifies the difference between $\hat{\tau}_{\varrho}(t_N^{\star})$ and $\deg(D_N)t_N$:

Lemma 2.5.3. *Let $(b_N)_{N\geq1}$ be as in the statement of Proposition 2.5.1. Then there exists a constant $c_3 > 0$ such that*

$$\left|\frac{\hat{\tau}_{\varrho}(t_N^{\star})}{\deg(D_N)} - t_N\right| \leq b_N \qquad \text{and} \qquad |t_N^{\star} - t_N| < b_N\sqrt{t_N} \tag{2.5.11}$$

hold with P^{ϱ}-probability at least $1 - c_3 b_N^{-2}$.

Proof. Note that $\hat{\tau}_\varrho(t) = \sum_{x\in D_N\cup\{\varrho\}} \deg(x)\widehat{L}_t^{D_N}(x)$. The proof is a straightforward application of Chebyshev's inequality together with some variance estimates. We begin by noting that $\hat{\tau}_\varrho(t_N^\star) - \deg(D_N)t_N$ is the first time to hit ϱ starting from the point $\widetilde{X}_{\deg(D_N)t_N}$. Writing H_ϱ for the first hitting time of ϱ, the Markov property tells

$$E^\varrho\left[(\hat{\tau}_\varrho(t_N^\star) - \deg(D_N)t_N)^2\right] = E^\varrho\left[E^{\widetilde{X}_{\deg(D_N)t_N}}\left[H_\varrho^2\right]\right] \le \max_{x\in D_N} E^x\left[H_\varrho^2\right]. \tag{2.5.12}$$

As in the proof of the previous lemma, applying the Kac moment formula shows

$$E^x\left[H_\varrho^2\right] = 2\sum_{y,z\in D_N} \deg(y)\deg(z)G^{D_N}(x,y)G^{D_N}(y,z) \le c_4|D_N|^2 \tag{2.5.13}$$

for some absolute constant $c_4 > 0$. (This also conforms to the knowledge that the length of a typical excursion on D_N is comparable to the volume of D_N.) Then by the Chebyshev inequality,

$$P^\varrho\left(\left|\frac{\hat{\tau}_\varrho(t_N^\star)}{\deg(D_N)} - t_N\right| \ge b_N\right) \le \frac{c_4|D_N|^2}{(\deg(D_N)b_N)^2} \le \frac{c_4}{16b_N^2}, \tag{2.5.14}$$

where the last step follows from $\deg(D_N) = \deg(\varrho) + 4|D_N|$. Also, by the computation similar to the previous proof, we get

$$E^\varrho\left[\left(\frac{\hat{\tau}_\varrho(t)}{\deg(D_N)} - t\right)^2\right] = \frac{2t}{\deg(D_N)^2}\sum_{x,y\in D_N}\deg(x)\deg(y)G^{D_N}(x,y) \le c_5 t \tag{2.5.15}$$

for some constant $c_5 > 0$. So again, by Chebyshev's inequality,

$$P^\varrho\left(\frac{\hat{\tau}_\varrho(t_N - b_N\sqrt{t_N})}{\deg(D_N)} \ge t_N - b_N\sqrt{t_N}/2\right) \le \frac{c_5(t_N - b_N\sqrt{t_N})}{(b_N\sqrt{t_N}/2)^2} \le \frac{4c_5}{b_N^2} \tag{2.5.16}$$

and likewise

$$P^\varrho\left(\frac{\hat{\tau}_\varrho(t_N + b_N\sqrt{t_N})}{\deg(D_N)} \le t_N + b_N\sqrt{t_N}/2\right) \le \frac{4c_5(1 + b_N t_N^{-1/2})}{b_N^2}. \tag{2.5.17}$$

Combining (2.5.14), (2.5.16), and (2.5.17) we find that there exists a constant $c_3 > 0$, depending only on $(t_N)_{N\ge1}$ and $(b_N)_{N\ge1}$, such that all of

$$\begin{aligned} \hat{\tau}_\varrho(t_N - b_N\sqrt{t_N}) &< \deg(D_N)\big(t_N - b_N\sqrt{t_N}/2\big),\\ \hat{\tau}_\varrho(t_N + b_N\sqrt{t_N}) &> \deg(D_N)\big(t_N + b_N\sqrt{t_N}/2\big),\\ |\hat{\tau}_\varrho(t_N^\star) - \deg(D_N)t_N| &\le \deg(D_N)b_N/2 \end{aligned} \tag{2.5.18}$$

simultaneously hold with P^{ϱ}-probability at least $1 - c_3 b_N^{-2}$. But if all of (2.5.18) hold, then we get

$$\hat{\tau}_{\varrho}(t_N - b_N\sqrt{t_N}) < \hat{\tau}_{\varrho}(t_N^\star) < \hat{\tau}_{\varrho}(t_N + b_N\sqrt{t_N}). \tag{2.5.19}$$

By the monotonicity of $\hat{\tau}_{\varrho}$, these altogether imply (2.5.11) as required. □

Next we will quantify the difference between $t_N^\star$ and t_N°:

Lemma 2.5.4. *Assume $t_N \geq 1$ and let $(b_N)_{N\geq 1}$ be as in the statement of Proposition 2.5.1. Then there exists a constant $c_6 > 0$ such that*

$$|t_N^\star - t_N^\circ| \leq b_N t_N^{1/4} \tag{2.5.20}$$

holds with P^{ϱ}-probability at least $1 - c_6 b_N^{-1}$.

Proof. We note that, by (2.4.3) and the fact that $\deg(x) = 4$ for $x \in D_N$,

$$\begin{aligned} U_N(t) &= \frac{1}{|D_N|}\sum_{x\in D_N}\left(\frac{1}{\deg(x)}\int_0^{\hat{\tau}_{\varrho}(t)} 1_{\{\widetilde{X}_s = x\}}\,\mathrm{d}s - t\right) \\ &= \frac{1}{|D_N|}\left(\frac{1}{4}\left(\hat{\tau}_{\varrho}(t) - \deg(\varrho)t\right) - |D_N|t\right) = \frac{\hat{\tau}_{\varrho}(t) - \deg(D_N)t}{4|D_N|}. \end{aligned} \tag{2.5.21}$$

Rearranging the identity in terms of t, we get

$$t = \frac{\hat{\tau}_{\varrho}(t)}{\deg(D_N)} - \left(1 - \frac{\deg(\varrho)}{\deg(D_N)}\right) U_N(t). \tag{2.5.22}$$

This will be used to prove the desired bound. Plugging $t := t_N^\star$, we notice that the right-hand side of (2.5.22) almost looks like the definition (2.5.4) of t_N°, except that we need t_N in place of $\hat{\tau}_{\varrho}(t_N^\star)/\deg(D_N)$ and $U_N(t_N)$ in place of $U_N(t_N^\star)$. This amounts to estimating their respective differences, and this is where the previous lemmas come handy.

First, we plug $s := t_N - b_N\sqrt{t_N}$ and $t := 2b_N\sqrt{t_N}$ in (2.5.7) to get

$$P^{\varrho}\left(\sup_{|u|\leq b_N\sqrt{t_N}} |U_N(t_N + u) - U_N(t_N)| \geq b_N t_N^{1/4}\right) \leq \frac{8c_2 b_N\sqrt{t_N}}{\left(b_N t_N^{1/4}\right)^2} = \frac{8c_2}{b_N}. \tag{2.5.23}$$

Combining this with Lemma 2.5.3, we can find $c_7 > 0$ such that both (2.5.11) and

$$|U_N(t_N + u) - U_N(t_N)| \leq b_N t_N^{1/4} \qquad \text{for all } |u| \leq b_N\sqrt{t_N} \tag{2.5.24}$$

hold with P^ϱ-probability at least $1-c_7 b_N^{-1}$. Moreover, given (2.5.11) and (2.5.24), we also get $|U_N(t_N^\star)-U_N(t_N)|\le b_N t_N^{1/4}$. Putting this together, we get

$$\begin{aligned}|t_N^\star - t_N^\circ| &\le \left|\frac{\hat\tau_\varrho(t_N^\star)}{\deg(D_N)} - t_N\right| + |U_N(t_N^\star)-U_N(t_N)| \\ &\le b_N\bigl(1+t_N^{1/4}\bigr) \le 2b_N t_N^{1/4}.\end{aligned} \tag{2.5.25}$$

Although this bound is slightly larger than that appearing in the statement, we can repeat all the above argument with $\{b_N/2\}_{N\ge1}$ in place of $\{b_N\}_{N\ge1}$, then the desired claim follows with $c_6 = 2c_7$. □

We are now ready to prove the main statement:

Proof of Proposition 2.5.1. Let $(b_N)_{N\ge1}$ be as in the statement. Then by the definition of $t_N^\star$ and Lemma 2.5.4,

$$\deg(D_N)t_N \le \hat\tau_\varrho(t_N^\star) \le \hat\tau_\varrho\bigl(t_N^\circ + b_N t_N^{1/4}\bigr) \tag{2.5.26}$$

holds with P^ϱ-probability at least $1-\mathcal{O}(b_N^{-1})$. Next, regarding $t_N \mapsto t_N^\star$ and $t_N\mapsto t_N^\circ$ as functions of t_N for each fixed N, Lemma 2.5.3 applied to $(t_N - b_N/4)_{N\ge1}$ and $(b_N/4)_{N\ge1}$ in place of $(t_N)_{N\ge1}$ and $(b_N)_{N\ge1}$, respectively, show that both

$$\deg(D_N)t_N \ge \hat\tau_\varrho\bigl((t_N-b_N/4)^\star\bigr) \ge \deg(D_N)(t_N - b_N/2) \tag{2.5.27}$$

and

$$\bigl|\bigl(t_N - b_N/4\bigr)^\star - t_N\bigr| \le b_N\sqrt{t_N}/2 \tag{2.5.28}$$

are satisfied with P^ϱ-probability at least $1-\mathcal{O}(b_N^{-1})$. Then using (2.5.24) and repeating the argument as in the previous proof, we can bound $(t_N-b_N/4)^\star$ from below by $t_N^\circ - b_N t_N^{1/4}$ again with probability at least $1-\mathcal{O}(b_N^{-1})$. □

2.5.2 Continuous-time exceptional level sets

We are now ready to adapt the convergence theorems for the exceptional level-set measures for the boundary-vertex local times $\widehat{L}^{D_N}$ to those associated with the local time $\widetilde{L}^{D_N}$ of the continuous-time walk $\widetilde{X}$ run for a fixed time of order $N^2(\log N)^2$. We begin by the thick points; the arguments will be readily adapted to the other families of exceptional points as well. Given two positive sequences $\{t_N\}_{N\geq 1}$ and $\{a_N\}_{N\geq 1}$ as before, define

$$\widetilde{\zeta}_N^D = \frac{1}{W_N} \sum_{x\in D_N} \delta_{x/N} \otimes \delta_{(\widetilde{L}^{D_N}_{\deg(D_N)t_N}(x)-a_N)/\sqrt{2a_N}}, \tag{2.5.29}$$

where W_N is the same as in the case of $\widehat{\zeta}_N^D$. Then

Proposition 2.5.5 (Continuous-time thick points)**.** *Under the setting and notation of Theorem 2.2.3 and for the walk started at the "boundary vertex," we have*

$$\widetilde{\zeta}_N^D \xrightarrow[N\to\infty]{\text{law}} \sqrt{\frac{\sqrt{\theta}}{\sqrt{\theta}+\lambda}}\; \mathfrak{c}(\lambda)\, \mathrm{e}^{\alpha\lambda(\mathfrak{d}(x)-1)T}\, Z_\lambda^{D,0}(\mathrm{d}x) \otimes \mathrm{e}^{-\alpha\lambda h}\mathrm{d}h, \tag{2.5.30}$$

where T and $Z_\lambda^{D,0}$ are independent with $T \sim \mathcal{N}(0,\sigma_D^2)$.

The key point is to carefully track the effects of the random time shift $\sqrt{2t_N}\,T_N$ in the quantity t_N° from (2.5.4). Let $\{b_N\}_{N\geq 1}$ be a sequence with $b_N \to \infty$ and $b_N/t_N^{1/4} \to 0$. Consider the event

$$\begin{aligned}\mathcal{E}_N := &\Big\{\hat{\tau}_\varrho\big(t_N^\circ - b_N t_N^{1/4}\big) \leq \deg(D_N)t_N \leq \hat{\tau}_\varrho\big(t_N^\circ + b_N t_N^{1/4}\big)\Big\} \\ &\cap \Big\{\max_{|u|\leq b_N\sqrt{t_N}} |U_N(t_N+u) - U_N(t_N)| \leq b_N t_N^{1/4}\Big\} \cap \{|T_N| \leq b_N\}\,. \end{aligned} \tag{2.5.31}$$

We then have:

Lemma 2.5.6. *There is a constant $c_7 > 0$ such that the following holds for all $N \geq 1$:*

$$P^\varrho(\mathcal{E}_N) \geq 1 - c_7 b_N^{-1} \tag{2.5.32}$$

and

$$\max_{|u|\leq b_N\sqrt{t_N}} |T_N(t_N+u) - T_N| \leq c_7 b_N/t_N^{1/4} \quad \text{on } \mathcal{E}_N. \tag{2.5.33}$$

Proof. The bound (2.5.32) follows from Proposition 2.5.1, Lemma 2.5.2 and the fact that T_N has asymptotically a Gaussian tail. To get (2.5.33), note that for $|u| \leq b_N\sqrt{t_N}$,

$$|T_N(t_N+u) - T_N| \leq \frac{b_N t_N^{1/4}}{\sqrt{2(t_N - b_N\sqrt{t_N})}} + \frac{b_N\,|T_N|}{\sqrt{t_N - b_N\sqrt{t_N}}}. \tag{2.5.34}$$

As $|T_N| \leq b_N$ on $\mathcal{E}_N$ and $\{b_N/t_N^{1/4}\}_{N\geq 1}$ is bounded, this is at most order $b_N/t_N^{1/4}$. □

The argument to follow will be based on dividing the event $\mathcal{E}_N$ depending on the values of T_N. For this we fix an $\epsilon > 0$, and let $\{\rho_k\}_{k\in\mathbb{Z}}$ be a family of continuous functions such that

$$0 \le \rho_k \le \mathbf{1}_{[(k-1)\epsilon,(k+1)\epsilon]} \qquad \text{and} \qquad \sum_{k\in\mathbb{Z}} \rho_k = 1. \tag{2.5.35}$$

We also define two auxilliary time sequences $\{t^+_{N,k}\}_{N\ge1}$ and $\{t^-_{N,k}\}_{N\ge1}$ by

$$\begin{aligned} t^+_{N,k} &= t_N - \Big(1 - \tfrac{\deg(\varrho)}{\deg(D_N)}\Big)\epsilon(k-1)\sqrt{2t_N} + b_N t_N^{1/4}, \\ t^-_{N,k} &= t_N - \Big(1 - \tfrac{\deg(\varrho)}{\deg(D_N)}\Big)\epsilon(k+1)\sqrt{2t_N} - b_N t_N^{1/4}. \end{aligned} \tag{2.5.36}$$

We then have:

Lemma 2.5.7. *For each $M > 0$ there is $N_0 \in \mathbb{N}$ such that for all $N \ge N_0$ and all $k \in \mathbb{Z}$ with $|k| \le M$, the following holds on $\mathcal{E}_N \cap \{T_N \in \operatorname{supp}(\rho_k)\}$:*

$$\big|T_N(t^\pm_{N,k}) - T_N\big| \le c_7 b_N / t_N^{1/4} \tag{2.5.37}$$

and

$$\widehat{L}^{D_N}_{t^-_{N,k}}(\cdot) \le \widetilde{L}^{D_N}_{\deg(D_N)t_N}(\cdot) \le \widehat{L}^{D_N}_{t^+_{N,k}}(\cdot). \tag{2.5.38}$$

Proof. Fix $M > 0$. As $b_N \to \infty$ and $b_N t_N^{-1/4} \to 0$, we can choose $N_0 \in \mathbb{N}$ such that $\epsilon(M+1)\sqrt{2t_N} + b_N t_N^{1/4} \le b_N\sqrt{t_N}$ for all $N \ge N_0$. Then for all $N \ge N_0$,

$$\big|t^\pm_{N,k} - t_N\big| \le b_N\sqrt{t_N}, \quad -M \le k \le M. \tag{2.5.39}$$

The bound (2.5.37) is then implied by (2.5.33).

For (2.5.38) we note that, on $\{T_N \in \operatorname{supp}(\rho_k)\}$ we have $(k-1)\epsilon \le T_N \le (k+1)\epsilon$ and thus also

$$t^-_{N,k} \le t^\circ_N - b_N t_N^{1/4} \le t^\circ_N + b_N t_N^{1/4} \le t^+_{N,k}. \tag{2.5.40}$$

The bound (2.5.38) then follows from the inequalities in (2.5.31) and the monotonicity of $t \mapsto \widehat{L}^{D_N}_t(\cdot)$. □

The inequalities (2.5.38) thus naturally make us consider the level-set measures $\widehat{\zeta}^D_N$ along different choices of time sequences than the base sequence $\{t_N\}_{N\ge1}$. We will explicate the dependence on the time sequence by writing $\widehat{\zeta}^D_N(t'_N)$ whenever it is along $\{t'_N\}_{N\ge1}$ rather than $\{t_N\}_{N\ge1}$, and likewise, we will write $W_N(t'_N)$ for the normalizing constants along $\{t'_N\}_{N\ge1}$. Next we note:

Lemma 2.5.8. *We have* $\deg(\varrho)/\deg(D_N)\to0$ *as* $N\to\infty$. *In particular, for each* $k\in\mathbb{Z}$*,*

$$t^{\pm}_{N,k}\sim 2g\theta(\log N)^2,\quad N\to\infty. \tag{2.5.41}$$

Moreover,

$$\begin{aligned} W_N(t^{+}_{N,k}) &= W_N(t_N)\,\mathrm{e}^{-\alpha\lambda\epsilon(k-1)+o(1)},\\ W_N(t^{-}_{N,k}) &= W_N(t_N)\,\mathrm{e}^{-\alpha\lambda\epsilon(k+1)+o(1)}, \end{aligned} \tag{2.5.42}$$

where $o(1)\to0$ *uniformly in* $k\in\mathbb{Z}$ *with* $|k|\le M$*, for any* $M>0$*.*

Proof. We start by showing $\deg(\varrho)/\deg(D_N)\to0$. For this we note that $\deg(D_N)\ge4|D_N|$ while, for any $\delta>0$ and N sufficiently large, $\deg(\varrho)\le4|D_N\smallsetminus D_N^\delta|$, where $D_N^\delta:=\{x\in D_N\colon d_\infty(x,D_N^{\mathrm{c}})>\delta N\}$. Definition 2.1.2 now ensures

$$\limsup_{N\to\infty}\frac{\deg(\varrho)}{\deg(D_N)}\le\limsup_{N\to\infty}\frac{|D_N\smallsetminus D_N^\delta|}{|D_N|}\le\frac{\mathrm{Leb}(D\smallsetminus D^{2\delta})}{\mathrm{Leb}(D)}, \tag{2.5.43}$$

where $D^\delta:=\{x\in D\colon d_\infty(x,D^{\mathrm{c}})>\delta\}$. As $D^{2\delta}\uparrow D$ as $\delta\downarrow0$, we have $\mathrm{Leb}(D\smallsetminus D^{2\delta})\to0$ as $\delta\downarrow0$.

With $\deg(\varrho)/\deg(D_N)\to0$ settled, the asymptotic (2.5.41) is now checked readily from the definition of $t^{\pm}_{N,k}$. The bounds in (2.5.42) follow similarly from the explicit formula for W_N and some routine estimates. □

We are now ready for:

Proof of Proposition 2.5.5. Let $f\colon\overline{D}\times(\mathbb{R}\cup\{+\infty\})\to[0,\infty)$ be a bounded and continuous function that is non-decreasing in the second coordinate and supported on $\overline{D}\times[b,\infty]$ for some $b\in\mathbb{R}$. Then (2.5.42), (2.5.38) and (2.5.37) show

$$\begin{aligned}\mathrm{e}^{-2\alpha\lambda\epsilon+o(1)}\mathrm{e}^{-\alpha\lambda T_N(t_{N,k}^-)}\langle\widehat{\zeta}_N^D(t_{N,k}^-),f\rangle\le\langle\widetilde{\zeta}_N^D,f\rangle&\\ \le\mathrm{e}^{2\alpha\lambda\epsilon+o(1)}\mathrm{e}^{-\alpha\lambda T_N(t_{N,k}^+)}\langle\widehat{\zeta}_N^D(t_{N,k}^+),f\rangle&\end{aligned}\tag{2.5.44}$$

on $\mathcal{E}_N\cap\{T_N\in\operatorname{supp}(\rho_k)\}$, where $o(1)$ is a deterministic sequence tending to zero uniformly in $k\in\mathbb{Z}$ with $|k|\le M$.

Define the maximal modulus of continuity of $\{\rho_k\colon |k|\le M\}$ by

$$\operatorname{osc}_{M,\epsilon}(r):=\max_{|k|\le M}\sup_{\substack{t,t'\in\mathbb{R}\\|t-t'|\le r}}\bigl|\rho_k(t)-\rho_k(t')\bigr|.\tag{2.5.45}$$

Relying first on the lower bound of (2.5.44), we now estimate

$$\begin{aligned}&E^\varrho\bigl(\mathrm{e}^{-\langle\widetilde{\zeta}_N^D,f\rangle}\bigr)-P^\varrho(\mathcal{E}_N^{\mathrm{c}})-P^\varrho\bigl(|T_N|\ge M/\epsilon\bigr)\\&\qquad\le\sum_{k=-M}^{M}E^\varrho\bigl(\mathrm{e}^{-\langle\widetilde{\zeta}_N^D,f\rangle}\rho_k(T_N)\,\mathbf{1}_{\mathcal{E}_N}\bigr)\\&\qquad\le\sum_{k=-M}^{M}E^\varrho\Bigl(\mathrm{e}^{-\mathrm{e}^{-2\alpha\lambda\epsilon+o(1)}\mathrm{e}^{-\alpha\lambda T_N(t_{N,k}^-)}\langle\widehat{\zeta}_N^D(t_{N,k}^-),f\rangle}\rho_k(T_N)\,\mathbf{1}_{\mathcal{E}_N}\Bigr)\\&\qquad\le(2M+1)\operatorname{osc}_{M,\epsilon}\bigl(c_7b_N/t_N^{1/4}\bigr)\\&\qquad\qquad+\sum_{k=-M}^{M}E^\varrho\Bigl(\mathrm{e}^{-\mathrm{e}^{-2\alpha\lambda\epsilon+o(1)}\mathrm{e}^{-\alpha\lambda T_N(t_{N,k}^-)}\langle\widehat{\zeta}_N^D(t_{N,k}^-),f\rangle}\rho_k\bigl(T_N(t_{N,k}^-)\bigr)\,\mathbf{1}_{\mathcal{E}_N}\Bigr),\end{aligned}\tag{2.5.46}$$

where in the last step we used (2.5.37). The key point is that, dropping the indicator of $\mathcal{E}_N$, the k-th term in the sum is now a continuous function of the process $\widehat{\zeta}_N^D(t_{N,k}^-)$ and the time $T_N(t_{N,k}^-)$. In light of (2.5.41), Proposition 2.4.3 gives

$$E^\varrho\Bigl(\mathrm{e}^{-\mathrm{e}^{-2\alpha\lambda\epsilon+o(1)}\mathrm{e}^{-\alpha\lambda T_N(t_{N,k}^-)}\langle\widehat{\zeta}_N^D(t_{N,k}^-),f\rangle}\rho_k\bigl(T_N(t_{N,k}^-)\bigr)\Bigr)\underset{N\to\infty}{\longrightarrow}E\Bigl(\mathrm{e}^{-\mathrm{e}^{-2\alpha\lambda\epsilon}\mathrm{e}^{-\alpha\lambda T}\langle\widehat{\zeta}^D,f\rangle}\rho_k(T)\Bigr),\tag{2.5.47}$$

where

$$\widehat{\zeta}^D:=\sqrt{\frac{\sqrt{\theta}}{\sqrt{\theta}+\lambda}}\,\mathfrak{c}(\lambda)\,\mathrm{e}^{\alpha\lambda\mathfrak{d}(x)T}Z_\lambda^{D,0}(\mathrm{d}x)\otimes\mathrm{e}^{-\alpha\lambda h}\mathrm{d}h\tag{2.5.48}$$

with $T=\mathcal{N}(0,\sigma_D^2)$ independent of $Z_\lambda^{D,0}$. Dropping the restriction to $|k|\le M$, the $N\to\infty$ *limes superior* of the sum on the extreme right of (2.5.46) is then at most $E(\mathrm{e}^{-\mathrm{e}^{-2\alpha\lambda\epsilon}\mathrm{e}^{-\alpha\lambda T}\langle\widehat{\zeta}^D,f\rangle})$.

Since $\mathrm{osc}_{M,\epsilon}(r) \to 0$ as $r \downarrow 0$, taking $N \to \infty$ followed by $M \to \infty$ and $\epsilon \downarrow 0$ shows

$$\limsup_{N\to\infty} E^{\varrho}\bigl(\mathrm{e}^{-\langle \widetilde{\zeta}_N^D, f\rangle}\bigr) \le E\bigl(\mathrm{e}^{-\mathrm{e}^{-\alpha\lambda T}\langle \widehat{\zeta}^D, f\rangle}\bigr), \tag{2.5.49}$$

where the two "error" terms on the left-hand side of (2.5.46) tend to zero in the stated limits thanks to Lemma 2.5.6 and the Gaussian (asymptotic) tail of T_N.

The argument for a corresponding lower bound is very similar; we need to work with $t_{N,k}^+$ instead of $t_{N,k}^-$ and use explicit estimates to get rid of the indicator $\mathbf{1}_{\mathcal{E}_N}$ and the restriction to the range of k in the sum. As a conclusion, we get

$$\lim_{N\to\infty} E^{\varrho}\bigl(\mathrm{e}^{-\langle \widetilde{\zeta}_N^D, f\rangle}\bigr) = E\bigl(\mathrm{e}^{-\mathrm{e}^{-\alpha\lambda T}\langle \widehat{\zeta}^D, f\rangle}\bigr) \tag{2.5.50}$$

for any function f as above. This is sufficient to give $\widetilde{\zeta}_N^D \xrightarrow{\text{law}} \mathrm{e}^{-\alpha\lambda T}\widehat{\zeta}^D$, as desired. □

For the thin points we now get:

Proposition 2.5.9 (Continous-time thin points)**.** *Under the setting and notation of Theorem 2.2.4 and for the walk started at the "boundary vertex," we have*

$$\widetilde{\zeta}_N^D \;\underset{N\to\infty}{\xrightarrow{\text{law}}}\; \sqrt{\frac{\sqrt{\theta}}{\sqrt{\theta}-\lambda}}\;\mathfrak{c}(\lambda)\,\mathrm{e}^{-\alpha\lambda(\mathfrak{d}(x)-1)T}\,Z_\lambda^{D,0}(\mathrm{d}x)\otimes \mathrm{e}^{+\alpha\lambda h}\mathrm{d}h, \tag{2.5.51}$$

where T and $Z_\lambda^{D,0}$ are independent with $T \sim \mathcal{N}(0,\sigma_D^2)$.

Proof. The argument is similar to that for the thick points: We need to work with compactly-supported, continuous test functions $f\colon \overline{D}\times(\mathbb{R}\cup\{-\infty\}) \to [0,\infty)$ that are non-increasing in the second coordinate. The change in monotonicity effectively swaps the inequalities in (2.5.44) and, due to a sign change in (2.5.42), also that in the exponent of $\mathrm{e}^{-\alpha\lambda T_N(t_{N,k}^{\pm})}$. We also need to rely on Proposition 2.4.4 instead of Proposition 2.4.3. We leave further details to the reader. □

Moving to the light points, we define

$$\widetilde{\vartheta}_N^D := \frac{1}{\widehat{W}_N}\sum_{x\in D_N} \delta_{x/N}\otimes\delta_{\widetilde{L}^{D_N}_{\deg(D_N)t_N}(x)} \tag{2.5.52}$$

and state:

Proposition 2.5.10 (Continuous-time light points)**.** *Under the setting and assumptions of Theorem 2.2.5 and for the walk started at the "boundary vertex," we have*

$$\widetilde{\vartheta}_N^D \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}} \sqrt{2\pi g}\,\mathfrak{c}(\sqrt{\theta})\ \mathrm{e}^{-\alpha\sqrt{\theta}(\mathfrak{d}(x)-1)T}\, Z^{D,0}_{\sqrt{\theta}}(\mathrm{d}x)\otimes\tilde{\mu}(\mathrm{d}h), \tag{2.5.53}$$

where $T = \mathcal{N}(0,\sigma_D^2)$ is independent of $Z^{D,0}_{\sqrt{\theta}}$ and $\tilde{\mu}$ is the measure in (2.4.40)*.*

Proof. Relying on our convention concerning different time sequences, we start by noting

$$\begin{aligned}\widehat{W}_N(t^+_{N,k}) &= \widehat{W}_N(t_N)\,\mathrm{e}^{\alpha\sqrt{\theta}\epsilon(k-1)+o(1)},\\ \widehat{W}_N(t^-_{N,k}) &= \widehat{W}_N(t_N)\,\mathrm{e}^{\alpha\sqrt{\theta}\epsilon(k+1)+o(1)}.\end{aligned} \tag{2.5.54}$$

Given a compactly-supported, continuous function $f\colon \overline{D}\times[0,\infty)\to[0,\infty)$ that is non-increasing in the second coordinate, from (2.5.54), (2.5.38) and (2.5.37) we then have

$$\begin{aligned}\mathrm{e}^{-2\alpha\sqrt{\theta}\epsilon+o(1)}\mathrm{e}^{\alpha\sqrt{\theta}T_N(t^+_{N,k})}\langle\widehat{\vartheta}^D_N(t^+_{N,k}),f\rangle &\le \langle\widetilde{\vartheta}^D_N,f\rangle\\ &\le \mathrm{e}^{2\alpha\sqrt{\theta}\epsilon+o(1)}\mathrm{e}^{\alpha\sqrt{\theta}T_N(t^-_{N,k})}\langle\widehat{\vartheta}^D_N(t^-_{N,k}),f\rangle.\end{aligned} \tag{2.5.55}$$

The rest of the argument for the thick points (with Proposition 2.4.5 instead of Proposition 2.4.3) can now be applied to get

$$\langle\widetilde{\vartheta}^D_N,f\rangle \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}} \mathrm{e}^{+\alpha\sqrt{\theta}T}\langle\widehat{\vartheta}^D,f\rangle, \tag{2.5.56}$$

where

$$\widehat{\vartheta}^D := \sqrt{2\pi g}\,\mathfrak{c}(\sqrt{\theta})\ \mathrm{e}^{-\alpha\sqrt{\theta}\,\mathfrak{d}(x)T}\, Z^{D,0}_{\sqrt{\theta}}(\mathrm{d}x)\otimes\tilde{\mu}(\mathrm{d}h). \tag{2.5.57}$$

The claim now follows by a density argument. □

Finally, for the avoided points we set

$$\widetilde{\kappa}^D_N := \frac{1}{\widehat{W}_N}\sum_{x\in D_N} 1_{\{\widetilde{L}^{D_N}_{\deg(D_N)t_N}(x)=0\}}\,\delta_{x/N} \tag{2.5.58}$$

and state:

Proposition 2.5.11 (Continuous-time avoided points)**.** *Under the setting and assumptions of Theorem 2.2.5 and for the walk started at the "boundary vertex," we have*

$$\widetilde{\kappa}_N^D \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}} \sqrt{2\pi g}\,\mathfrak{c}(\sqrt{\theta})\ \mathrm{e}^{-\alpha\sqrt{\theta}(\mathfrak{d}(x)-1)T}\, Z^{D,0}_{\sqrt{\theta}}(\mathrm{d}x), \tag{2.5.59}$$

where $T = \mathcal{N}(0,\sigma_D^2)$ is independent of $Z^{D,0}_{\sqrt{\theta}}$.

Proof. Given a continuous $f\colon \overline{D} \to \mathbb{R}$, the identity (2.5.55) applies with $\widetilde{\vartheta}_N^D$, resp., $\widehat{\vartheta}_N^D$ replaced by $\widetilde{\kappa}_N^D$, resp., $\widehat{\kappa}_N^D$. The argument then proceeds as for Proposition 2.5.10. $\square$

2.6 Arbitrary starting points

As our next item of business, we augment the continuous-time conclusions from the previous section to allow the random walk to start at an arbitrary point of D_N. The formal statement is the content of:

Theorem 2.6.1 (Arbitrary starting points)**.** *The statements of Propositions 2.5.5, 2.5.9, 2.5.10 and 2.5.11 apply for random walk starting from an arbitrary point $x_N \in D_N$.*

We will start with the thick points as that is the hardest case. Assume that $\{a_N\}_{N\ge1}$ and $\{t_N\}_{N\ge1}$ satisfy the conditions of Propositions 2.5.5. The integrals of $\{\widetilde{\zeta}_N^D\colon N\ge1\}$ from (2.5.29) against $f \in C_\mathrm{c}(\overline{D}\times(\mathbb{R}\cup\{+\infty\}))$ are tight random variables. Our strategy is to use the strong Markov property after the first hitting of the "boundary vertex." For this let us recall that H_x denotes the first hitting time of vertex x and let θ_t denote the shift on the path space acting as $(\widetilde{X}\circ\theta_t)_s = \widetilde{X}_{t+s}$. We will write $\{(\widetilde{L}^{D_N}\circ\theta_t)_s\colon s\ge0\}$ for the local time process associated with the time-shifted path $\{(\widetilde{X}\circ\theta_t)_s\colon s\ge0\}$. Our first observation is then:

Lemma 2.6.2. *On $\{H_\varrho < t\}$, we have*

$$\widetilde{L}_t^{D_N}(\cdot) = \widetilde{L}_{H_\varrho}^{D_N}(\cdot) + (\widetilde{L}^{D_N}\circ\theta_{H_\varrho})_{t-H_\varrho}(\cdot). \tag{2.6.1}$$

In particular, under the conditions of Proposition 2.5.5, for any $f \in C_{\rm c}(\overline{D} \times (\mathbb{R} \cup \{+\infty\}))$ that is non-decreasing in the second variable and any $x_N \in D_N$,

$$\limsup_{N\to\infty} E^{x_N}\bigl(\mathrm{e}^{-\langle\widetilde{\zeta}_N^D, f\rangle}\bigr) \le \lim_{N\to\infty} E^{\varrho}\bigl(\mathrm{e}^{-\langle\widetilde{\zeta}_N^D, f\rangle}\bigr). \tag{2.6.2}$$

Proof. The relation (2.6.1) is a direct consequence of the additivity of the local time. As to (2.6.2), for f as above and any $m > 0$ with $t_N > m$, dropping the term $\widetilde{L}_{H_\varrho}^{D_N}$ while noting that $W(t_N - m) \ge \mathrm{e}^{-c\frac{m}{\log N}} W(t_N)$ for some $c > 0$ shows

$$\bigl\langle \widetilde{\zeta}_N^{D_N}(t_N), f\bigr\rangle \ge \mathrm{e}^{-c\frac{m}{\log N}} \bigl\langle \widetilde{\zeta}_N^{D_N}(t_N - m), f\bigr\rangle \circ \theta_{H_\varrho} \quad \text{on } \{H_\varrho < m\deg(D_N)\}. \tag{2.6.3}$$

The strong Markov property then gives

$$\begin{aligned} E^{x_N}\bigl(\mathrm{e}^{-\langle\widetilde{\zeta}_N^D(t_N), f\rangle}\bigr) &\le P^{x_N}\bigl(H_\varrho \ge m\deg(D_N)\bigr) + E^{x_N}\bigl(\mathrm{e}^{-\langle\widetilde{\zeta}_N^D(t_N), f\rangle} 1_{\{H_\varrho < m\deg(D_N)\}}\bigr) \\ &\le P^{x_N}\bigl(H_\varrho \ge m\deg(D_N)\bigr) + E^{\varrho}\bigl(\mathrm{e}^{-\mathrm{e}^{-c\frac{m}{\log N}}\langle\widetilde{\zeta}_N^D(t_N - m), f\rangle}\bigr). \end{aligned} \tag{2.6.4}$$

Since the random walk on D_N coincides with the random walk on $\mathbb{Z}^2$ until time H_ϱ, the Central Limit Theorem shows that the probability tends to zero in the limits $N \to \infty$ and $m \to \infty$. The expectation on the right converges by Proposition 2.5.5. $\square$

Our next goal is to prove a complementary bound to (2.6.2) for the *limes inferior*. For this we must control the effect of the first term on the right of (2.6.1). Writing $\{(\widehat{L}^{D_N} \circ \theta_t)_s \colon s \ge 0\}$ for the local time of the process $\widetilde{X} \circ \theta_t$ parametrized at the time spent at the boundary vertex, we then have:

Lemma 2.6.3. *Under the conditions of Proposition 2.5.5, for each $b \in \mathbb{R}$ there is $c > 0$ such that for all $N \ge 1$ and all $x \in D_N$,*

$$\sum_{z \in D_N} P^x\Bigl(\widetilde{L}_{H_\varrho}^{D_N}(z) + (\widehat{L}^{D_N} \circ \theta_{H_\varrho})_{t_N}(z) \ge a_N + b\log N,\, H_z < H_\varrho\Bigr) \le c\frac{W_N}{\log N}. \tag{2.6.5}$$

Proof. Let us for simplicity assume (e.g., by redefining a_N) that $b = 0$. The strong Markov property bounds the probability under the sum by

$$\sum_{m \ge 0} P^x\bigl(H_z < H_\varrho\bigr) P^z\Bigl(\widetilde{L}_{H_\varrho}^{D_N}(z) \ge mG^{D_N}(z,z)\Bigr) P^{\varrho}\Bigl(\widehat{L}_{t_N}^{D_N}(z) \ge a_N - (m+1)G^{D_N}(z,z)\Bigr). \tag{2.6.6}$$

We start by estimating the second term. Denoting $p := P^z(\hat{H}_z < H_\varrho)$ where $\hat{H}_z$ is the first return time to z, we have $\widetilde{L}^{D_N}_{H_\varrho}(z) \overset{\text{law}}{=} \frac{1}{4}\sum_{i=1}^N \tau_i$ for $N :=$ Geometric(p) and $\tau_1, \tau_2, \dots$ i.i.d. Exponential(1) independent of N. For any $q \in (0,1)$, the Chernoff bound gives

$$P\biggl(\sum_{i=1}^N \tau_i > r\biggr) \underset{0\le s<1-p}{\le} \mathrm{e}^{-sr}\frac{1-p}{1-p-s} \underset{s:=q(1-p)}{\le} \frac{1}{1-q}\mathrm{e}^{-rq(1-p)}. \tag{2.6.7}$$

As $1-p = P^z(\hat{H}_z > H_\varrho) = \frac{1}{4G^{D_N}(z,z)}$, we thus get

$$P^z\Bigl(\widetilde{L}^{D_N}_{H_\varrho}(z) \ge mG^{D_N}(z,z)\Bigr) \le \frac{1}{1-q}\mathrm{e}^{-mq}, \quad m \ge 0. \tag{2.6.8}$$

for all $q \in (0,1)$.

Using (2.6.8) in conjunction with the uniform estimate $G^{D_N}(z,z) \le g\log N + c$, we dominate the part of the sum in (2.6.6) for m satisfying $(m+2)G^{D_N}(z,z) \ge a_N - t_N$ by a quantity of order $N^{-2q[(\sqrt{\theta}+\lambda)^2-\theta+o(1)]}$. Recalling that $W_N = N^{2(1-\lambda^2)+o(1)}$, this is $o(W_N N^{-2}/\log N)$ when $1-q>0$ is so small that $q[(\sqrt{\theta}+\lambda)^2-\theta] > \lambda^2$.

In the complementary regime, we have $a_N - (m+2)G^{D_N}(z,z) > t_N$ which permits us to estimate the last term on the right of (2.6.6) via [1, Lemma 4.1] with the choices $a := a_N$, $t := t_N$ and $b := (m+2)G^{D_N}(z,z)$ to get

$$\begin{aligned} &P^z\Bigl(\widetilde{L}^{D_N}_{H_\varrho}(z) \ge mG^{D_N}(z,z)\Bigr)P^\varrho\Bigl(\widehat{L}^{D_N}_{t_N}(z) \ge a_N - (m+1)G^{D_N}(z,z)\Bigr) \\ &\quad\le \frac{1}{1-q}\frac{\sqrt{G^{D_N}(z,z)}}{\sqrt{2a_N - 2(m+1)G^{D_N}(z,z)} - \sqrt{2t_N}}\frac{\sqrt{\log N}}{N^2}W_N\,\mathrm{e}^{-qm+(m+1)\frac{\sqrt{2a_N}-\sqrt{2t_N}}{\sqrt{2a_N}}} \end{aligned} \tag{2.6.9}$$

As $\frac{\sqrt{2a_N}-\sqrt{2t_N}}{\sqrt{2a_N}} \to \frac{\lambda}{\sqrt{\theta}+\lambda}$ as $N\to\infty$, we choose $q \in (\frac{\lambda}{\sqrt{\theta}+\lambda}, 1)$ and proceed as follows: For $(m+1)G^{D_N}(z,z) > \frac{1}{2}(a_N - t_N)$, the prefactor is order $\sqrt{\log N}\,W_N/N^2$ but, thanks to the uniform upper bound on $G^{D_N}(z,z)$, the sum of the exponential terms decays polynomially with N. For m with $(m+1)G^{D_N}(z,z) \le \frac{1}{2}(a_N - t_N)$, the prefactor is order W_N/N^2 and the sum of the exponentials is bounded.

Combining the above estimates, the sum in (2.6.5) is bounded by a quantity of order

$$o\Bigl(\frac{W_N}{\log N}\Bigr) + \frac{W_N}{N^2}\sum_{z\in D_N} P^x\bigl(H_z < H_\varrho\bigr). \tag{2.6.10}$$

Interpreting H_ϱ as the first exit time of the simple random walk on $\mathbb{Z}^2$ from D_N, the sum on the right is non-decreasing in D_N. We may thus assume that D_N is a box of side-length 2^n, for $n = \log_2 N + O(1)$, centered at x. For the probability under the sum we then get, for each $k = 0, \dots, n-1$ and some constant $c > 0$,

$$P^x(H_z < H_\rho) = \frac{G^{D_N}(x,z)}{G^{D_N}(z,z)} \le c\frac{n-k}{n}, \qquad 2^k < |x-z| \le 2^{k+1}. \tag{2.6.11}$$

The sum in (2.6.10) is thus at most of order $1 + \sum_{k=0}^{n} \frac{n-k}{n} 2^{2k}$ which is of order $N^2/\log N$. The claim follows. □

We are now ready to give:

Proof of Theorem 2.6.1, thick points. Consider a non-negative $f \in C_c(\overline{D} \times (\mathbb{R} \cup \{+\infty\})$ that is non-decreasing in the second variable and supported in $\overline{D} \times [b, \infty)$ for some $b \in \mathbb{R}$. Note that $\{H_\varrho < \infty\}$ is a full probability event under P^x. Decomposing the support of ζ_N^D according to whether the point was hit before hitting the boundary vertex or not, the monotonicity of $t \mapsto \widetilde{L}_t^{D_N}$ and the assumed monotonicity of f yield

$$\begin{aligned}\langle \widetilde{\zeta}_N^D, f\rangle &\le \langle \widetilde{\zeta}_N^D, f\rangle \circ \theta_{H_\varrho} \\ &\qquad + \frac{\|f\|_\infty}{W_N} \sum_{z \in D_N} 1_{\{H_z < H_\varrho\}}\, 1_{\{\widetilde{L}_{H_\varrho}^{D_N}(z) + (\widetilde{L}^{D_N}\circ\theta_{H_\varrho})_{t_N \deg(D_N)}(z) \ge a_N + b\sqrt{2a_N}\}}.\end{aligned} \tag{2.6.12}$$

Fix a sequence $b_N \to \infty$ such that $b_N/t_N^{1/4} \to 0$ and let $\mathcal{F}_N$ be the event that the inequalities in (2.5.6) hold. Fix any $m > 0$ and $\epsilon > 0$. Let $\mathcal{G}_N$ be the event that the second term on the right of (2.6.12) is less than ϵ. Then

$$\begin{aligned}E^x\bigl(\mathrm{e}^{-\langle \widetilde{\zeta}_N^D, f\rangle}\bigr) &\ge E^x\Bigl(\mathrm{e}^{-\langle \widetilde{\zeta}_N^D, f\rangle} 1_{\theta_{H_\varrho}^{-1}(\mathcal{F}_N \cap \{T_N \ge -m\})}\Bigr) \\ &\ge \mathrm{e}^{-\epsilon} E^x\Bigl(\mathrm{e}^{-\langle \widetilde{\zeta}_N^D, f\rangle\circ\theta_{H_\varrho}} 1_{\theta_{H_\varrho}^{-1}(\mathcal{F}_N \cap \{T_N \ge -m\})}\Bigr) \\ &\qquad\qquad - P^x\Bigl(\mathcal{G}_N^{\mathrm{c}} \cap \theta_{H_\varrho}^{-1}(\mathcal{F}_N \cap \{T_N \ge -m\})\Bigr).\end{aligned} \tag{2.6.13}$$

As $P^x(H_\varrho < \infty) = 1$, the strong Markov property gives

$$\begin{aligned}E^x\Bigl(\mathrm{e}^{-\langle \widetilde{\zeta}_N^D, f\rangle\circ\theta_{H_\varrho}} 1_{\theta_{H_\varrho}^{-1}(\mathcal{F}_N \cap \{T_N \ge -m\})}\Bigr) &= E^\varrho\Bigl(\mathrm{e}^{-\langle \widetilde{\zeta}_N^D, f\rangle} 1_{\mathcal{F}_N \cap \{T_N \ge -m\}}\Bigr) \\ &\ge E^\varrho\bigl(\mathrm{e}^{-\langle \widetilde{\zeta}_N^D, f\rangle}\bigr) - P^\varrho\bigl((\mathcal{F}_N \cap \{T_N \ge -m\})^{\mathrm{c}}\bigr).\end{aligned} \tag{2.6.14}$$

Proposition 2.5.1 and the fact that $\{T_N\colon N\ge 1\}$ is tight now ensures that the probability on the right tends to zero in the limits $N\to\infty$ and $m\to\infty$.

Concerning the probability on the right of (2.6.13), an inspection of (2.5.4) shows that, on $(\mathcal{F}_N\cap\{T_N\ge -m\})\circ\theta_{H_\varrho}$, we have

$$(\widetilde{L}^{D_N}\circ\theta_{H_\varrho})_{t_N\deg(D_N)}(\cdot)\le(\widehat{L}^{D_N}\circ\theta_{H_\varrho})_{t_N+b_Nt_N^{1/4}+m\sqrt{2t_N}}(\cdot). \tag{2.6.15}$$

By the Markov inequality, the probability in (2.6.13) is thus bounded by $\epsilon^{-1}\|f\|_\infty/W_N(t_N)$ times the sum in Lemma 2.6.3 albeit with t_N replaced by $t_N' := t_N+b_Nt_N^{1/4}+m\sqrt{2t_N}$. As $W_N(t_N')/W_N(t_N)$ is bounded by an m-dependent constant uniformly in N, the probability in (2.6.13) is thus $O(1/\log N)$ uniformly in $x\in D_N$. Taking $N\to\infty$ followed by $m\to\infty$ and $\epsilon\downarrow 0$ shows

$$\liminf_{N\to\infty} E^{x_N}\bigl(\mathrm{e}^{-\langle\widetilde{\zeta}_N^D,f\rangle}\bigr)\ge\lim_{N\to\infty}E^{\varrho}\bigl(\mathrm{e}^{-\langle\widetilde{\zeta}_N^D,f\rangle}\bigr). \tag{2.6.16}$$

Combining with (2.6.2), we then get the desired claim. □

The situation for the thin, light and avoided points is similar albeit simpler. Writing $\widetilde{\xi}_N^D$ for the corresponding continuous-time point measure (parametrized by the total time), as in Lemma 2.6.2, the identity (2.6.1) gives us an easy one-way bound, where the test function f takes values in $\overline{D}\times(\mathbb{R}\cup\{-\infty\})$ for the thin points, $\overline{D}\times[0,\infty)$ for the light points and $\overline{D}$ for the avoided points:

Lemma 2.6.4. *Under the conditions of Proposition 2.5.9, 2.5.10 and 2.5.11, for any any $x_N\in D_N$ and any continuous, compactly-supported, non-negative test function f on the corresponding domain that, for the thin and light points, is non-increasing in the second variable,*

$$\liminf_{N\to\infty} E^{x_N}\bigl(\mathrm{e}^{-\langle\widetilde{\xi}_N^D,f\rangle}\bigr)\ge\lim_{N\to\infty}E^{\varrho}\bigl(\mathrm{e}^{-\langle\widetilde{\xi}_N^D,f\rangle}\bigr). \tag{2.6.17}$$

Proof. Using (2.6.1), on $\{H_\varrho<m\deg(D_N)\}$ we get

$$\bigl\langle\widetilde{\xi}_N^{D_N}(t_N),f\bigr\rangle\le\mathrm{e}^{c\frac{m}{\log N}}\bigl\langle\widetilde{\xi}_N^{D_N}(t_N-m),f\bigr\rangle\circ\theta_{H_\varrho}, \tag{2.6.18}$$

where we now rely on the fact that $t \mapsto W_N(t)$, resp., $t \mapsto \widehat{W}_N(t)$ are non-increasing for t near t_N. The inequalities (2.6.4) then become

$$
\begin{aligned}
E^{x_N}\bigl(\mathrm{e}^{-\langle \widetilde{\xi}_N^D(t_N),f\rangle}\bigr) &\ge E^{x_N}\bigl(\mathrm{e}^{-\langle \widetilde{\xi}_N^D(t_N),f\rangle} 1_{\{H_\varrho < m \deg(D_N)\}}\bigr) \\
&\ge E^\varrho\bigl(\mathrm{e}^{-\mathrm{e}^{c\frac{m}{\log N}}\langle \widetilde{\xi}_N^D(t_N-m),f\rangle}\bigr) - P^{x_N}\bigl(H_\varrho \ge m \deg(D_N)\bigr).
\end{aligned}
\tag{2.6.19}
$$

The claim now follows by taking $N \to \infty$ followed by $m \to \infty$. □

In replacement of Lemma 2.6.3, we then need:

Lemma 2.6.5. *Under the conditions of Proposition 2.5.9, for each $b \in \mathbb{R}$ there is $c > 0$ such that for all $N \ge 1$ and all $x \in D_N$,*

$$
\sum_{z\in D_N} P^x\Bigl((\widehat{L}^{D_N}\circ\theta_{H_\varrho})_{t_N}(z) \le a_N + b\log N,\, H_z < H_\varrho\Bigr) \le c\frac{W_N}{\log N}.
\tag{2.6.20}
$$

Under the conditions of Propositions 2.5.10 and 2.5.11 the same holds with $a_N + b\log N$ replaced by $b \ge 0$ (including, for the avoided points, $b = 0$) and W_N replaced by $\widehat{W}_N$.

Proof. The Strong Markov property and the estimates from [1, Corollary 4.8] bound the probability in (2.6.20) by $P^x(H_z < H_\varrho)$ times

$$
P^\varrho\bigl(\widehat{L}^{D_N}_{t_N}(z) \le a_N + b\log N\bigr) \le c\frac{W_N}{N^2}
\tag{2.6.21}
$$

and so the quantity in (2.6.20) is at most order $W_N N^{-2}\sum_{z\in D_N} P^x(H_z < H_\varrho)$. The argument then concludes as in the proof of Lemma 2.6.3. For the light and avoided points, we instead invoke [1, Corollary 4.6] and proceed analogously. □

With this we get:

Proof of Theorem 2.6.1, thin, light and avoided points. We proceed similarly as for the thick points. First, writing $\widetilde{a}_N := a_N + b\log N$ for the thin points and $\widetilde{a}_N := b$ for the light and (with $b := 0$) avoided points, given a continuous, compactly-supported f that is non-increasing in the second variable, in all three cases of interest we have

$$
\langle \widetilde{\xi}_N^D, f\rangle \ge \langle \widetilde{\xi}_N^D, f\rangle\circ\theta_{H_\varrho} - \frac{\|f\|_\infty}{W_N}\sum_{z\in D_N} 1_{\{H_z<H_\varrho\}}\, 1_{\{(\widetilde{L}^{D_N}\circ\theta_{H_\varrho})_{t_N\deg(D_N)-H_\varrho}(z)\le\widetilde{a}_N\}}.
\tag{2.6.22}
$$

Let $\mathcal{F}_N$ be the event from (2.5.6) with t_N replaced by $t_N - m$. Abusing our earlier notation, given $\epsilon > 0$, let $\mathcal{G}_N$ be the event that the second term (without the minus sign) is at most ϵ. From (2.6.22), we then get

$$\begin{aligned} E^x\bigl(\mathrm{e}^{-\langle \widetilde{\xi}_N^D, f\rangle}\bigr) - P^x\bigl(H_\varrho \geq m \deg(D_N)\bigr) - P^\varrho\bigl((\mathcal{F}_N \cap \{T_N(t_N - m) \leq m\})^{\mathrm{c}}\bigr) \\ \leq E^x\Bigl(\mathrm{e}^{-\langle \widetilde{\xi}_N^D, f\rangle} 1_{\{H_\varrho < m \deg(D_N)\}} 1_{\theta_{H_\varrho}^{-1}(\mathcal{F}_N \cap \{T_N(t_N - m) \leq m\})}\Bigr) \\ \leq \mathrm{e}^\epsilon E^\varrho\bigl(\mathrm{e}^{-\langle \widetilde{\xi}_N^D, f\rangle}\bigr) \\ + P^x\Bigl(\mathcal{G}_N^{\mathrm{c}} \cap \{H_\varrho < m \deg(D_N)\} \cap \theta_{H_\varrho}^{-1}(\mathcal{F}_N \cap \{T_N(t_N - m) \leq m\})\Bigr). \end{aligned} \tag{2.6.23}$$

Thanks to the Central Limit Theorem, the tightness of $\{T_N \colon N \geq 1\}$ and Proposition 2.5.1, the two probabilities on the left-hand side of (2.6.23) tend to zero in the limits $N \to \infty$ and $m \to \infty$, uniformly in $x \in D_N$. For the probability on the right we observe that, on $\{H_\varrho < m \deg(D_N)\} \cap \theta_{H_\varrho}^{-1}(\mathcal{F}_N \cap \{T_N(t_N - m) \leq m\})$, we have

$$(\widetilde{L}^{D_N} \circ \theta_{H_\varrho})_{t_N \deg(D_N) - H_\varrho}(\cdot) \geq (\widehat{L}^{D_N} \circ \theta_{H_\varrho})_{t'_N}(\cdot) \tag{2.6.24}$$

for $t'_N := t_N - m - b_N t_N^{1/4} - m\sqrt{2t_N}$. Lemma 2.6.5 and the Markov inequality then bound the probability by an m-dependent constant times $1/\log N$, uniformly in $x \in D_N$. Combining these observations we thus get

$$\limsup_{N\to\infty} E^{x_N}\bigl(\mathrm{e}^{-\langle \widetilde{\xi}_N^D, f\rangle}\bigr) \leq \lim_{N\to\infty} E^\varrho\bigl(\mathrm{e}^{-\langle \widetilde{\xi}_N^D, f\rangle}\bigr). \tag{2.6.25}$$

In conjunction with Lemma 2.6.4 this proves the claim. □

2.7 Discrete time conclusions

We will now move to the proof of our main results except those on the local structure which are deferred to Section 2.8. Considering, for a moment, a random walk on a general finite, connected graph on $V \cup \{\varrho\}$, recall that the discrete-time local time L_t^V is parametrized by the total number of steps in units of $\deg(V) = \sum_{u \in V \cup \{\varrho\}} \deg(u)$ while its continuous-time counterpart $\widetilde{L}_t^V$ is parametrized by the total time. Both of these are naturally realized on the same probability space through the definition (2.4.1) of $\widetilde{X}$ via the discrete-time walk X and an independent (rate-1) Poisson point process $\widetilde{N}(t)$. A key technical tool in what follows is the following lemma:

Lemma 2.7.1. *There is a family of i.i.d. exponentials $\{\tau_j(v)\colon j \geq 1, v \in V\}$ with parameter 1 independent of X (but not of $\widetilde{N}$) such that*

$$\widetilde{L}_t^V(v) = \frac{1}{\deg(V)} \sum_{j\geq 1} \tau_j(v) 1_{\{j \leq \deg(V) L^V_{\widetilde{N}(t)/\deg(V)}(v)\}}, \quad v \neq \widetilde{X}_t, \tag{2.7.1}$$

holds P^x-a.s. for each $t \geq 0$ and each $x \in V \cup \{\varrho\}$.

Proof. This is a consequence of the standard representation of the wait times of $\widetilde{X}$ by independent exponentials. (In this representation, the process $\widetilde{N}$ is a function of the exponentials and X, albeit independent of X.) Note that the equality (2.7.1) fails at $\widetilde{X}_t$ because the walk is "in-between" jumps there. □

Moving back to the random walk on $D_N \cup \{\varrho\}$, this readily yields:

Lemma 2.7.2. *For each $x \in D_N$, abbreviate*

$$\mathcal{F}_N(x) := \biggl\{ \widetilde{L}^{D_N}_{(t_N-1)\deg(D_N)}(x) \leq \frac{1}{4} \sum_{j\geq 1} \tau_j(x) 1_{\{j \leq 4 L^{D_N}_{t_N}(x)\}} \leq \widetilde{L}^{D_N}_{(t_N+1)\deg(D_N)}(x) \biggr\}. \tag{2.7.2}$$

Then for any $x_N \in D_N$,

$$P^{x_N}\Bigl(\sum_{x\in D_N} 1_{\mathcal{F}_N(x)^{\mathrm{c}}} > 2 \Bigr) \underset{N\to\infty}{\longrightarrow} 0. \tag{2.7.3}$$

Proof. The Central Limit Theorem ensures that $(\widetilde{N}(t) - t)/\sqrt{t}$ tends in law to a standard normal as $t \to \infty$. As $t_N = o(\deg(D_N))$, the inequalities

$$\frac{\widetilde{N}((t_N - 1)\deg(D_N))}{\deg(D_N)} \leq t_N \leq \frac{\widetilde{N}((t_N + 1)\deg(D_N))}{\deg(D_N)} \tag{2.7.4}$$

are satisfied with probability tending to one as $N \to \infty$. Once (2.7.4) is in force, the monotonicity of $t \mapsto L_t^{D_N}$ and (2.7.1) show that the event $\mathcal{F}_N(x)$ occurs at all $x \in D_N$ except perhaps at the position of $\widetilde{X}$ at times $(t_N \pm 1)\deg(D_N)$. □

With these observations in hand, we are now ready to finally present the proofs of our main theorems. The easiest case is that of avoided points:

Proof of Theorem 2.2.6. Note that, whenever $\mathcal{F}_N(x)$ occurs, $\widetilde{L}^{D_N}_{(t_N+1)\deg(D_N)}(x) = 0$ forces $L^{D_N}_{t_N}(x) = 0$ (a.s.), which in turn forces $\widetilde{L}^{D_N}_{(t_N-1)\deg(D_N)}(x) = 0$. For any $f \in C_c(\overline{D})$ with $f \geq 0$, on the event $\sum_{x\in D_N} 1_{\mathcal{F}_N(x)^c} \leq 2$ we thus have

$$\begin{aligned}\frac{\widehat{W}_N(t_N+1)}{\widehat{W}(t_N)}\bigl\langle\widetilde{\kappa}^D_N(t_N+1), f\bigr\rangle - \frac{2}{\widehat{W}_N}\|f\|_\infty \leq \bigl\langle\kappa^D_N, f\bigr\rangle \\ \leq \frac{\widehat{W}_N(t_N-1)}{\widehat{W}(t_N)}\bigl\langle\widetilde{\kappa}^D_N(t_N-1), f\bigr\rangle + \frac{2}{\widehat{W}_N}\|f\|_\infty.\end{aligned} \tag{2.7.5}$$

As $\{t_N \pm 1\}_{N\geq 1}$ have the same leading-order asymptotic as $\{t_N\}_{N\geq 1}$, the random variables $\langle\widetilde{\kappa}^D_N(t_N \pm 1), f\rangle$ have the same weak limit as $\langle\widetilde{\kappa}^D_N, f\rangle$. Since $\widehat{W}_N \to \infty$ and also

$$\frac{\widehat{W}_N(t_N \pm 1)}{\widehat{W}_N(t_N)} \underset{N\to\infty}{\longrightarrow} 1, \tag{2.7.6}$$

the claim follows from Lemma 2.7.2, Proposition 2.5.11 and Theorem 2.6.1. □

Next we tackle the light points:

Proof of Theorem 2.2.5. Denote

$$\overline{L}^{D_N}_{t_N}(x) := \frac{1}{4}\sum_{j\geq 1}\tau_j(x)1_{\{j\leq 4L^{D_N}_{t_N}(x)\}} \tag{2.7.7}$$

and consider the auxiliary point measure

$$\overline{\vartheta}^D_N := \frac{1}{\widehat{W}_N}\sum_{x\in D_N}\delta_{x/N}\otimes\delta_{\overline{L}^{D_N}_{t_N}(x)}. \tag{2.7.8}$$

Thanks to Lemma 2.7.2, on the event $\sum_{x\in D_N} 1_{\mathcal{F}_N(x)^c} \leq 2$, the inequality (2.7.5) holds for any non-negative $f \in C_c(\overline{D}\times[0,\infty))$ that is non-increasing in the second variable and with $\widetilde{\kappa}^D_N$, resp., κ^D_N replaced by $\widetilde{\vartheta}^D_N$, resp., $\overline{\vartheta}^D_N$. As, by Proposition 2.5.10 and Theorem 2.6.1, $\widetilde{\vartheta}^D_N$ tends in law to the measure $\widetilde{\vartheta}^D$ on the right of (2.5.53), we have

$$\langle\overline{\vartheta}^D_N, f\rangle \overset{\text{law}}{\underset{N\to\infty}{\longrightarrow}} \langle\widetilde{\vartheta}^D, f\rangle \tag{2.7.9}$$

for any non-negative $f \in C_{\rm c}(\overline{D} \times [0,\infty))$.

Next we observe that, by that fact that for any $\epsilon > 0$ and any random variable Y taking values in $[0,\epsilon]$,

$$\exp\{-E(Y)\} \le E(\mathrm{e}^{-Y}) \le \exp\{-\mathrm{e}^{-\epsilon} E(Y)\}, \tag{2.7.10}$$

the fact that the random variables $\{\tau_j(x)\colon j \ge 1,\, x \in D_N\}$ are independent of the random walk and independent for different $x \in D_N$ implies

$$E^{x_N}\Bigl(\mathrm{e}^{-E(\langle \overline{\vartheta}^D_N, f\rangle | \sigma(X))}\Bigr) \le E^{x_N}\Bigl(\mathrm{e}^{-\langle \overline{\vartheta}^D_N, f\rangle}\Bigr) \le E^{x_N}\Bigl(\mathrm{e}^{-\mathrm{e}^{-\|f\|_\infty/\widehat{W}_N} E(\langle \overline{\vartheta}^D_N, f\rangle | \sigma(X))}\Bigr) \tag{2.7.11}$$

(see [16, Lemma 3.12]), where the conditional expectation is meaningful because $\langle \overline{\vartheta}^D_N, f\rangle$ is a finite random variable. Defining $f^{*\mathfrak{e}}\colon \overline{D} \times [0,\infty) \to \mathbb{R}$ by

$$f^{*\mathfrak{e}}(x,\ell) := E\biggl[f\Bigl(x, \frac{1}{4}\sum_{j=1}^{\lfloor 4\ell \rfloor} \tau_j\Bigr)\biggr], \tag{2.7.12}$$

where $\{\tau_j\colon j \ge 1\}$ are i.i.d. Exponential(1), we have

$$E^{x_N}\bigl(\langle \overline{\vartheta}^D_N, f\rangle \,\big|\, \sigma(X)\bigr) = \langle \vartheta^D_N, f^{*\mathfrak{e}}\rangle. \tag{2.7.13}$$

Hence we get (under the laws $\{P^{x_N}\colon N \ge 1\}$),

$$\langle \vartheta^D_N, f^{*\mathfrak{e}}\rangle \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}} \langle \widetilde{\vartheta}^D, f\rangle \tag{2.7.14}$$

for any $f \in C_{\rm c}(\overline{D} \times [0,\infty))$.

We now claim that $\{\vartheta^D_N\colon N \ge 1\}$ is tight. For this we pick $M \in \mathbb{N}$, denote $f_M(x,h) := 1_{[0,M]}(h)$ and observe that, for all $n \in \mathbb{N}_0 := \{0,1,2,\dots\}$, we get

$$f^{*\mathfrak{e}}_M(x, n/4) = P\Bigl(\frac{1}{4}\sum_{j=1}^{n} \tau_j \le M\Bigr). \tag{2.7.15}$$

Markov's inequality then shows $f^{*\mathfrak{e}}_{2M}(x,n/4) \ge \frac{1}{2} 1_{[0,M]}(n/4)$ and, therefore,

$$\vartheta^D_N\bigl(\overline{D} \times [0,M]\bigr) \le 2\langle \vartheta^D_N, f^{*\mathfrak{e}}_{2M}\rangle. \tag{2.7.16}$$

The existence of the limit (2.7.14) then implies tightness of $\{\vartheta^D_N(\overline{D} \times [0,M])\colon N \ge 1\}$ for all $M > 0$, and thus tightness of $\{\vartheta^D_N\colon N \ge 1\}$ as well.

The tightness of $\{\vartheta^D_N \colon N \geq 1\}$ permits us to extract a weak subsequential limit ϑ^D along a (strictly) increasing sequence $\{N_k \colon k \geq 1\}$ of naturals. This entails the convergence $\langle \vartheta^D_{N_k}, f\rangle \xrightarrow{\text{law}} \langle \vartheta^D, f\rangle$ for every $f \in C_{\text{c}}(\overline{D} \times [0,\infty))$. We claim that we even have

$$\langle \vartheta^D_{N_k}, f^{*\mathfrak{e}}\rangle \underset{k\to\infty}{\overset{\text{law}}{\longrightarrow}} \langle \vartheta^D, f^{*\mathfrak{e}}\rangle \tag{2.7.17}$$

for every $f \in C_{\text{c}}(\overline{D} \times [0,\infty))$. (This is not automatic because $f^{*\mathfrak{e}}$ is not compactly supported in general.) First we note that straightforward comparisons with the Lebesgue measure show, for each $M > 0$,

$$\lim_{n\to\infty} \frac{P\Bigl(\frac{1}{4}\sum_{j=1}^{4n} \tau_j \leq M\Bigr)}{P\Bigl(\frac{1}{4}\sum_{j=1}^{4n} \tau_j \leq 2M\Bigr)} = 0. \tag{2.7.18}$$

Writing ϵ_n for the ratio of the two probabilities, for f supported in $\overline{D} \times [0, M]$ we have $|f^{*\mathfrak{e}}| \leq \|f\|_\infty f^{*\mathfrak{e}}_M$ and so, by (2.7.15),

$$\bigl|f^{*\mathfrak{e}}(x, n/4)\bigr| \leq \epsilon_n \|f\|_\infty\, f^{*\mathfrak{e}}_{2M}(x,n), \quad n \in \mathbb{N}_0. \tag{2.7.19}$$

It follows that the part of the integral $\langle \vartheta^D_N, f^{*\mathfrak{e}}\rangle$ corresponding to the second coordinate in excess of n is at most $\epsilon_n \|f\|_\infty$ times $\langle \vartheta^D_N, f^{*\mathfrak{e}}_{2M}\rangle$, which is tight by (2.7.14). We can thus approximate $f^{*\mathfrak{e}}$ by a function supported in $\overline{D} \times [0,n]$ and pass to the limit $N \to \infty$ followed by $n \to \infty$. This gives (2.7.17) as desired.

Combining (2.7.14) with (2.7.17) we arrive at the convolution identity

$$\langle \vartheta^D, f^{*\mathfrak{e}}\rangle \overset{\text{law}}{=} \langle \widetilde{\vartheta}^D, f\rangle. \tag{2.7.20}$$

We have proved this (including the absolute convergence of the integral on the left-hand side) for $f \in C_{\text{c}}(\overline{D} \times [0,\infty))$ but the Monotone Convergence Theorem along with the fact that the second coordinate of $\widetilde{\vartheta}^D$ has subexponentially growing density extends this to all $f \in C(\overline{D} \times [0,\infty))$ such that $|f(x,h)| \leq c\mathrm{e}^{-\epsilon h}$ for some $\epsilon, c > 0$. This permits us to consider functions of the form $g_s(x,h) := \tilde{f}(x)\mathrm{e}^{-sh}$ for $s > 0$ and $\tilde{f} \in C(\overline{D})$, for which

$$g^{*\mathfrak{e}}_s(x, n/4) = \tilde{f}(x)(1 + s/4)^{-n}, \quad n \in \mathbb{N}_0. \tag{2.7.21}$$

Since ϑ^D is supported on $\overline{D} \times \frac{1}{4}\mathbb{N}_0$, it makes sense to denote

$$\vartheta^{D,n}(A) := \vartheta^D\bigl(A \times \{n/4\}\bigr). \tag{2.7.22}$$

The identity (2.7.20) then becomes

$$\sum_{n\ge 0}\langle \vartheta^{D,n}, \tilde f\rangle (1+s/4)^{-n} \overset{\text{law}}{=} \langle \widetilde{\vartheta}^D, g_s\rangle. \tag{2.7.23}$$

Assuming $\tilde f > 0$, the explicit form of the right-hand side shows that $\langle \widetilde{\vartheta}^D, g_s\rangle / \langle \widetilde{\vartheta}^D, g_1\rangle$ is well-defined and equal to a non-random quantity — namely, the ratio of two Laplace transforms of $\tilde\mu$. This turns (2.7.23) into the pointwise identity

$$\sum_{n\ge 0}\langle \vartheta^{D,n}, \tilde f\rangle (1+s/4)^{-n} = \frac{\int \tilde\mu(\mathrm{d}h)\mathrm{e}^{-sh}}{\int \tilde\mu(\mathrm{d}h)\mathrm{e}^{-h}}\biggl(\sum_{n\ge 0}\langle \vartheta^{D,n}, \tilde f\rangle (5/4)^{-n}\biggr) \tag{2.7.24}$$

valid, a.s., for each $s > 0$ and (by elementary extensions) all $\tilde f \in C(\overline{D})$. Thanks to the monotonicity of both sides in s and almost-sure continuity in $\tilde f$ of both sides with respect to the supremum norm, the identity actually holds a.s. for all $s > 0$ and all $\tilde f \in C(\overline{D})$ simultaneously.

With (2.7.24) in hand, we are more or less done. Indeed, as the left-hand side is a generating function of the sequence $\{\langle \vartheta^{D,n}, \tilde f\rangle\}_{n\ge 0}$, which determines the sequence uniquely, all $\langle \vartheta^{D,n}, \tilde f\rangle$ must be the same deterministic multiple of the quantity in the large parentheses on the right-hand side. This shows that ϑ^D must be as on the right-hand side of (2.2.25) for some μ of the form $\mu = \sum_{n\ge 0} q_n\delta_{n/4}$ where $\{q_n\}_{n\ge 0}$ is uniquely determined by

$$\sum_{n\ge 0} q_n (1+s/4)^{-n} = \int_0^\infty \tilde\mu(\mathrm{d}h)\mathrm{e}^{-sh}, \quad s > 0. \tag{2.7.25}$$

The Laplace transform of $\tilde\mu$ was calculated in the proof of Proposition 2.4.5. All subsequential limits of $\{\vartheta^D_N \colon N \ge 1\}$ are thus equal in law and so convergence holds. □

Moving to the thick points, we first need a version of (2.7.18):

Lemma 2.7.3. *For $\{\tau_j \colon j \ge 1\}$ be i.i.d. Exponential(1), all $k \in \mathbb{N}$ and all reals $s \ge t \ge 0$,*

$$\frac{P\Bigl(\sum_{j=1}^k (\tau_j - 1) \ge s + t\Bigr)}{P\Bigl(\sum_{j=1}^k (\tau_j - 1) \ge s\Bigr)} \le \mathrm{e}^{-\frac{st}{k+s+t}}. \tag{2.7.26}$$

Proof. Since $\sum_{j=1}^k \tau_j$ has density $\frac{1}{(k-1)!}x^{k-1}\mathrm{e}^{-x}$, the change of variables $y := x+t$ gives

$$\begin{aligned} P\biggl(\sum_{j=1}^k(\tau_j-1)\ge s\biggr) &= \frac{1}{(k-1)!}\int_{x\ge k+s}\mathrm{d}x\, x^{k-1}\mathrm{e}^{-x} \\ &= \mathrm{e}^t\,\frac{1}{(k-1)!}\int_{y\ge k+s+t}\mathrm{d}y\,(y-t)^{k-1}\mathrm{e}^{-y} \\ &\ge \mathrm{e}^t\Bigl(1-\frac{t}{k+s+t}\Bigr)^k P\biggl(\sum_{j=1}^k(\tau_j-1)\ge s+t\biggr). \end{aligned} \tag{2.7.27}$$

Using that $s\ge t$, the prefactor can be written as the exponential of

$$\begin{aligned} t+k\log\Bigl(1-\frac{t}{k+s+t}\Bigr) &= t-k\sum_{n\ge1}\frac{1}{n}\frac{t^n}{(k+s+t)^n} \\ &\ge t-\frac{kt}{k+s+t}-\frac{1}{2}\frac{kt^2}{(k+s+t)^2}\sum_{n\ge0}2^{-n}. \end{aligned} \tag{2.7.28}$$

Noting that right-hand side is no less than $\frac{st}{k+s+t}$, we get the claim. □

A convolution identity that inevitably shows up in the proof also requires:

Lemma 2.7.4. *Suppose ν is a Borel measure on $\mathbb{R}$ such that, for some $\beta\in\mathbb{R}$ and some $\sigma^2>0$ and all $f\in C_\mathrm{c}(\mathbb{R})$,*

$$\int_{\mathbb{R}}\nu(\mathrm{d}h)\,E\bigl[\,f(h+\mathcal{N}(0,\sigma^2))\bigr] = \int_{\mathbb{R}}\mathrm{d}h\;\mathrm{e}^{\beta h}\,f(h) \tag{2.7.29}$$

Then

$$\nu(\mathrm{d}h) = \mathrm{e}^{-\frac{1}{2}\beta^2\sigma^2+\beta h}\mathrm{d}h. \tag{2.7.30}$$

Proof. Consider the measure $\widetilde{\nu}(\mathrm{d}h) := \mathrm{e}^{-\beta h+\frac{1}{2}\beta^2\sigma^2}\nu(\mathrm{d}h)$. Absorbing the exponential term on the right of (2.7.29) into the test function, a calculation shows

$$\int_{\mathbb{R}\times\mathbb{R}}\widetilde{\nu}(\mathrm{d}h)\otimes\frac{\mathrm{d}x}{\sqrt{2\pi\sigma^2}}\;\mathrm{e}^{-\frac{(x-h+\beta\sigma^2)^2}{2\sigma^2}}\;f(x) = \int_{\mathbb{R}}\mathrm{d}h\,f(h) \tag{2.7.31}$$

for all $f\in C_\mathrm{c}(\mathbb{R})$. As $C_\mathrm{c}(\mathbb{R})$ generates all Borel functions in $\mathbb{R}$, we get

$$\frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}}\widetilde{\nu}(\mathrm{d}h)\,\mathrm{e}^{-\frac{(x-h+\beta\sigma^2)^2}{2\sigma^2}} = 1,\quad x\in\mathbb{R}. \tag{2.7.32}$$

This can be interpreted by saying that $\widehat{\nu}(\mathrm{d}h) := \frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{(h-\beta\sigma^2)^2}{2\sigma^2}}\widetilde{\nu}(\mathrm{d}h)$ is a measure such that

$$\int_{\mathbb{R}} \widehat{\nu}(\mathrm{d}h)\mathrm{e}^{-xh} = \mathrm{e}^{-x\beta\sigma^2+x^2\sigma^2/2}, \quad x \in \mathbb{R}. \tag{2.7.33}$$

The right-hand side is the Laplace transform of $\mathcal{N}(\beta\sigma^2,\sigma^2)$ and so, since the Laplace transform of a measure, if exists, determines the measure uniquely, $\widehat{\nu}$ is the law of $\mathcal{N}(\beta\sigma^2,\sigma^2)$. Hence $\widetilde{\nu}$ is the Lebesgue measure, thus proving the claim. □

Proof of Theorem 2.2.3. The proof starts by adapting the argument leading to (2.7.14). Indeed, working again in the coupling of the random walk X and the i.i.d. exponentials $\{\tau_j(x)\colon x \in D_N,\, j \geq 1\}$, let

$$\overline{\zeta}_N^D := \frac{1}{W_N}\sum_{x\in D_N} \delta_{x/N} \otimes \delta_{(\overline{L}_{t_N}^{D_N}(x)-a_N)/\sqrt{2a_N}}\,, \tag{2.7.34}$$

where $\overline{L}_{t_N}^{D_N}(x)$ is the quantity from (2.7.7). Lemmas 2.7.1-2.7.2 along with Proposition 2.5.5, Theorem 2.6.1 and (2.7.10) then show

$$E^{x_N}\big(\langle \overline{\zeta}_N^D, f\rangle \,\big|\, \sigma(X)\big) \;\underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}}\; \langle \widetilde{\zeta}^D, f\rangle \tag{2.7.35}$$

for every $f \in C_{\mathrm{c}}(\overline{D}\times(\mathbb{R}\cup\{+\infty\}))$, where $\widetilde{\zeta}^D$ is the measure on the right of (2.5.30). Writing $\{\tau_j\colon j \geq 1\}$ for generic i.i.d. exponentials with parameter 1 and denoting, with some abuse of earlier notation,

$$f^{N,*\mathfrak{e}}(x,h) := E\biggl[f\Bigl(x, h + \frac{1}{4\sqrt{2a_N}}\sum_{j\geq 1}(\tau_j - 1)1_{\{j\leq 4a_N+4h\sqrt{2a_N}\}}\Bigr)\biggr], \tag{2.7.36}$$

the fact that $L_{t_N}^{D_N}$ takes values in $\frac{1}{4}\mathbb{N}_0$ then shows

$$E^{x_N}\big(\langle \overline{\zeta}_N^D, f\rangle \,\big|\, \sigma(X)\big) \;=\; \langle \zeta_N^D, f^{N,*\mathfrak{e}}\rangle \tag{2.7.37}$$

thus proving

$$\langle \zeta_N^D, f^{N,*\mathfrak{e}}\rangle \;\underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}}\; \langle \widetilde{\zeta}^D, f\rangle \tag{2.7.38}$$

for every $f \in C_{\mathrm{c}}(\overline{D}\times(\mathbb{R}\cup\{+\infty\}))$.

We will now use (2.7.38) to control the behavior of the measures $\{\zeta_N^D\colon N\ge 1\}$. First, writing henceforth $1_{[M,\infty)}$ for the function $(x,h)\mapsto 1_{[M,\infty)}(h)$ we get

$$\bigl(1_{[M,\infty)}\bigr)^{N,*\mathfrak{e}}(x,h)=P\biggl(\sum_{j=1}^k(\tau_j-1)\ge(M-h)4\sqrt{2a_N}\biggr), \tag{2.7.39}$$

where $k:=\lfloor 4a_N+4h\sqrt{2a_N}\rfloor$. Assuming $h\ge 2M$ with $M>0$ large, Markov's inequality along with $E((\tau_j-1)^2)=1$ then gives

$$1-\bigl(1_{[M,\infty)}\bigr)^{N,*\mathfrak{e}}(x,h)\le\frac{4a_N+4h\sqrt{2a_N}}{32a_N(h-M)^2}\le\frac{1}{h^2}+\frac{1}{h\sqrt{2a_N}}. \tag{2.7.40}$$

For M large, the right-hand side is at most $1/2$ thus showing

$$1_{[2M,\infty)}(h)\le 2\bigl(1_{[M,\infty)}\bigr)^{N,*\mathfrak{e}}(x,h). \tag{2.7.41}$$

From (2.7.38) and the fact that $\widetilde{\zeta}^D$ has an exponentially decaying density in the second variable we then get, for each $\epsilon>0$,

$$\lim_{M\to\infty}\limsup_{N\to\infty}P^{x_N}\bigl(\langle\zeta_N^D,1_{[M,\infty)}\rangle>\epsilon\bigr)=0. \tag{2.7.42}$$

This implies tightness of $\{\zeta_N^D\colon N\ge 1\}$ on $\overline{D}\times(\mathbb{R}\cup\{+\infty\})$ along with their asymptotic concentration on $\overline{D}\times\mathbb{R}$. In particular, we may extract a weak subsequential limit ζ^D.

We would like to use the existence of weak subsequential limits to pass to the limit $N\to\infty$ inside the integral on the left-hand side of (2.7.38). For that we need to deal with the fact that the support of $f^{N,*\mathfrak{e}}$ extends to $-\infty$ in the second variable. Pick any $b>0$ and, for any $h<-3b$, invoke Lemma 2.7.3 with the choices $s:=4\sqrt{2a_N}(-2b-h)$, $t:=4\sqrt{2a_N}b$ and k as above to conclude that

$$\bigl(1_{[-b,\infty)}\bigr)^{N,*\mathfrak{e}}(x,h)\le \mathrm{e}^{-\frac{32a_Nb(-2b-h)}{4a_N-4\sqrt{2a_N}b}}\,\bigl(1_{[-2b,\infty)}\bigr)^{N,*\mathfrak{e}}(x,h),\quad h<-3b. \tag{2.7.43}$$

The prefactor decays to zero as $h\to-\infty$ uniformly in $N\ge 1$ and so, plugging this into (2.7.38) and using that $\{\langle\zeta_N^D,(1_{[-2b,\infty)})^{N,*\mathfrak{e}}\rangle\colon N\ge 1\}$ is tight we get, for each bounded, continuous f with $\operatorname{supp}(f)\subseteq\overline{D}\times[b,\infty]$ and each $\epsilon>0$,

$$\lim_{M\to\infty}\limsup_{N\to\infty}P\biggl(\Bigl|\bigl\langle\zeta_N^D,\,f^{N,*\mathfrak{e}}1_{(-\infty,-M]}\bigr\rangle\Bigr|>\epsilon\biggr)=0. \tag{2.7.44}$$

Combining this with (2.7.42), we may truncate the second variable in the integral on the left of (2.7.38) to lie in $[-M,M]$ at the cost of errors that tend to zero in probability as $M\to\infty$. The Central Limit Theorem shows

$$\frac{1}{4\sqrt{2a_N}}\sum_{j\ge1}(\tau_j-1)1_{\{j\le 4a_N\}}\ \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}}\ \mathcal{N}\bigl(0,\tfrac18\bigr) \tag{2.7.45}$$

and a simple estimate based, e.g., on Doob's L^2-martingale inequality to account for the correction $4\sqrt{2a_N}h$ in the number of terms in the sum then gives

$$\lim_{N\to\infty}\ \sup_{h\in[-M,M]}\ \sup_{x\in\overline{D}}\bigl|f^{N,*\mathfrak{e}}(x,h)-f^{*\mathfrak{n}}(x,h)\bigr|=0, \tag{2.7.46}$$

where

$$f^{*\mathfrak{n}}(x,h)=E\Bigl[f\bigl(x,h+\mathcal{N}(0,\tfrac18)\bigr)\Bigr]. \tag{2.7.47}$$

Taking $M\to\infty$ after $N\to\infty$ we then readily conclude that every subsequential weak limit ζ^D of $\{\zeta^D_N\colon N\ge1\}$ satisfies the distributional identity

$$\langle\zeta^D,f^{*\mathfrak{n}}\rangle\ \overset{\text{law}}{=}\ \langle\widetilde{\zeta}^D,f\rangle \tag{2.7.48}$$

for all $f\in C_{\rm c}(\overline{D}\times(\mathbb{R}\cup\{+\infty\}))$. This includes the fact that the integral on the left-hand side converges absolutely for all such f.

We are now more or less done. Indeed, note that the explicit form of $\widetilde{\zeta}^D$ gives, for $\tilde f\in C_{\rm c}(\mathbb{R})$ and $A\subseteq D$ Borel with $\mathrm{Leb}(A)>0$,

$$\frac{\langle\widetilde{\zeta}^D,1_A\otimes\tilde f\rangle}{\langle\widetilde{\zeta}^D,1_A\otimes1_{[0,\infty)}\rangle}=\alpha\lambda\int\mathrm{d}h\,\mathrm{e}^{-\alpha\lambda h}\tilde f(h),\quad \text{a.s.} \tag{2.7.49}$$

The right-hand side is non-random and so (2.7.48) becomes the pointwise equality

$$\bigl\langle\zeta^D,(1_A\otimes\tilde f)^{*\mathfrak{n}}\bigr\rangle=\bigl\langle\zeta^D,(1_A\otimes1_{[0,\infty)})^{*\mathfrak{n}}\bigr\rangle\,\alpha\lambda\int\mathrm{d}h\,\mathrm{e}^{-\alpha\lambda h}\tilde f(h) \tag{2.7.50}$$

for all $\tilde f\in C_{\rm c}(\mathbb{R})$. This shows that, for any $B\subseteq\mathbb{R}$ Borel,

$$\zeta^D(A\times B)=\alpha\lambda\bigl\langle\zeta^D,(1_A\otimes1_{[0,\infty)})^{*\mathfrak{n}}\bigr\rangle\otimes\nu(B), \tag{2.7.51}$$

where ν is a Borel measure on $\mathbb{R}$ that obeys (2.7.29) with $\beta := -\alpha\lambda$ and $\sigma^2 := 1/8$. Lemma 2.7.4 then gives $\nu(\mathrm{d}h) = \mathrm{e}^{-\alpha^2\lambda^2/16-\alpha\lambda h}\,\mathrm{d}h$ and, since the first measure on the right of (2.7.51) has the law of the spatial part of $\widetilde{\zeta}^D$, we get

$$\zeta^D \overset{\text{law}}{=} \mathrm{e}^{-\alpha^2\lambda^2/16}\,\widetilde{\zeta}^D. \tag{2.7.52}$$

The claim follows. □

Finally, we deal with the changes that are required for the thin points:

Proof of Theorem 2.2.4. Following the proof of Theorem 2.2.3, the argument is exactly the same up to (2.7.38), except that now $f \in C_{\mathrm{c}}(\overline{D} \times (\mathbb{R} \cup \{-\infty\}))$. For the tightness, we then need to consider

$$\big(1_{(-\infty,-M]}\big)^{N,*\mathfrak{e}}(x,h) = P\bigg(\sum_{j=1}^{k}(\tau_j - 1) \le -(M+h)4\sqrt{2a_N}\bigg), \tag{2.7.53}$$

where $k := \lfloor 4a_N + 4h\sqrt{2a_N}\rfloor$. For $h \le -2M$ the same estimate as (2.7.40) then shows $1_{(-\infty,-2M]}(h) \le 2(1_{(-\infty,-M]})^{N,*\mathfrak{e}}(x,h)$ and so, for each $\epsilon > 0$, we get

$$\lim_{M\to\infty}\limsup_{N\to\infty} P^{x_N}\big(\langle \zeta_N^D, 1_{(-\infty,-M]}\rangle > \epsilon\big) = 0 \tag{2.7.54}$$

from (2.7.38). For the upper tail, we need a variation on Lemma 2.7.3:

Lemma 2.7.5. *For $\{\tau_j : j \ge 1\}$ i.i.d. Exponential(1), all $k \in \mathbb{N}$ and all $s,t \ge 0$ with $s+t < k$,*

$$\frac{P\Big(\sum_{j=1}^{k}(\tau_j - 1) \le -(s+t)\Big)}{P\Big(\sum_{j=1}^{k}(\tau_j - 1) \le -s\Big)} \le \mathrm{e}^{-\frac{t(s-1)}{k-s}}. \tag{2.7.55}$$

To use this, let $b > 0$ and invoke the choices $s := (h-2b)4\sqrt{2a_N}$, $t := 4b\sqrt{2a_N}$ and k as above while noting that, for N large and $h > 2b$, we have $s + t < k$, to get

$$\big(1_{(-\infty,b]}\big)^{N,*\mathfrak{e}}(x,h) \le \exp\Big\{-\frac{4b\sqrt{2a_N}[(h-2b)4\sqrt{2a_N}-1]}{4a_N + 4h\sqrt{2a_N} - (h-2b)4\sqrt{2a_N}}\Big\}\big(1_{(-\infty,2b]}\big)^{N,*\mathfrak{e}}(x,h). \tag{2.7.56}$$

The exponential prefactor tends to zero as $h \to \infty$ uniformly in N sufficiently large and so, for any bounded and continuous f with $\operatorname{supp}(f) \subseteq \overline{D} \times (-\infty, b]$ and each $\epsilon > 0$,

$$\lim_{M\to\infty}\limsup_{N\to\infty} P\bigg(\Big|\big\langle \zeta_N^D,\, f^{N,*\mathfrak{e}}1_{[M,\infty)}\big\rangle\Big| > \epsilon\bigg) = 0. \tag{2.7.57}$$

This again permits us to truncate the tails and derive (2.7.48) for each $f \in C_c(\overline{D} \times (\mathbb{R} \cup \{-\infty\}))$ and each weak subsequential limit ζ^D of $\{\zeta^D_N \colon N \geq 1\}$. The rest of the proof of Theorem 2.2.3 can be followed literally leading to (2.7.52), as before. $\square$

It remains to give:

Proof of Lemma 2.7.5. The explicit form of the density along with the substitution $y := x+t$ again shows

$$\begin{aligned} P\biggl(\sum_{j=1}^k (\tau_j - 1) \leq -(s+t)\biggr) &= \frac{1}{(k-1)!}\int_{0\leq x\leq k-s-t} \mathrm{d}x\, x^{k-1}\mathrm{e}^{-x} \\ &\leq \mathrm{e}^t \frac{1}{(k-1)!}\int_{t\leq y\leq k-s} \mathrm{d}y\,(y-t)^{k-1}\mathrm{e}^{-y} \\ &\leq \mathrm{e}^t\Bigl(1-\frac{t}{k-s}\Bigr)^{k-1} P\biggl(\sum_{j=1}^k (\tau_j - 1) \leq -s\biggr) \end{aligned} \tag{2.7.58}$$

Using the bound $1-x \leq \mathrm{e}^{-x}$, the prefactor is at most $\mathrm{e}^{-\frac{t(s-1)}{k-s}}$. $\square$

With the help of the above theorems, we can finally settle:

Proof of Theorem 2.2.1. For the local time $\widehat{L}^{D_N}_{t_N}$ parametrized by the time at the boundary vertex and the walk started at ϱ, the statement appears as [1, Theorem 2.1]. The bounds in Proposition 2.5.1 along with the tightness of $\{T_N \colon N \geq 1\}$ then extend the conclusion to $\widehat{L}^{D_N}_{t_N}$ replaced by $\widetilde{L}^{D_N}_{\deg(D_N)t_N}$. Since the random walk started at ϱ visits any given $x_N \in D_N$ in time of order $N^2 \log N$ while the walk started at x_N hits ϱ in time of order N^2 with high probability, shifting t_N by $\pm(\log N)^{3/2}$ and invoking the monotonicity of $t \mapsto \widetilde{L}^{D_N}_t$ extends [1, Theorem 2.1] to arbitrary starting points. The inequalities (2.7.4) then extend it to the discrete-time object $L^{D_N}_{t_N}$ as well. $\square$

2.8 Local structure

The last item to be addressed are the proofs of Theorems 2.2.7 and 2.2.8 dealing with the local structure of the local time field near thick/thin and avoided points, respectively. We will start with the former setting, as it is technically most demanding.

2.8.1 Thick and thin points

We will again carry the argument primarily for the thick points and only comment on the changes for the thin points. Assuming henceforth the setting and notation of Theorem 2.2.3, we start by converting the continuous-time in the boundary-vertex parametrization to that parametrized by the total time.

Proposition 2.8.1. *Let $\overline{\zeta}_N^{D,\mathrm{loc}}$ be given by the same formula as $\zeta_N^{D,\mathrm{loc}}$ in (2.2.31) except with $L_{t_N}^{D_N}(x)$ replaced by $\overline{L}_{t_N}^{D_N}(x)$ from (2.7.7). Then, given an $x_N \in D_N$ for each $N \geq 1$, under P^{x_N},*

$$\overline{\zeta}_N^{D,\mathrm{loc}} \underset{N\to\infty}{\overset{\mathrm{law}}{\longrightarrow}} \widetilde{\zeta}^D \otimes \widehat{\nu}_\lambda, \tag{2.8.1}$$

where $\widetilde{\zeta}^D$ is the measure on the right of (2.5.30) and $\widehat{\nu}_\lambda$ is the law of $\phi + \alpha\lambda\mathfrak{a}$, for ϕ the pinned DGFF; i.e., a centered Gaussian process on $\mathbb{Z}^2$ with covariances (2.2.32).

The proof will rely heavily on the arguments and notation from Sections 2.5–2.7. Throughout, we fix a sequence $\{b_N\}_{N\geq 1}$ such that $b_N \to \infty$ and $b_N/t_N^{1/4} \to 0$. First we condense the ideas underlying Lemmas 2.5.6, 2.5.7 and 2.7.2 into:

Lemma 2.8.2. *Given $\epsilon > 0$, let $\widetilde{t}_{N,k}^{\pm}$ be the quantity from (2.5.36) but with b_N replaced by $3b_N$. Abbreviate*

$$\widetilde{\mathcal{F}}_N(x) := \bigcup_{k\in\mathbb{Z}} \Bigg(\big\{(k-1)\epsilon \leq T_N \circ \theta_{H_\varrho} \leq (k+1)\epsilon\big\} \\ \cap \Big\{ (\widehat{L}^{D_N} \circ \theta_{H_\varrho})_{\widetilde{t}_{N,k}^-}(x) \leq \overline{L}_{t_N}^{D_N}(x) \leq \widetilde{L}_{H_\varrho}^{D_N}(x) + (\widehat{L}^{D_N} \circ \theta_{H_\varrho})_{\widetilde{t}_{N,k}^+}(x) \Big\} \Bigg). \tag{2.8.2}$$

Then for each $b \in \mathbb{R}$ and any choice of $x_N \in D_N$ for each $N \geq 1$,

$$P^{x_N}\Big(\sum_{x\in D_N} 1_{\widetilde{\mathcal{F}}_N(x)^{\mathrm{c}}} > 2 \Big) \underset{N\to\infty}{\longrightarrow} 0. \tag{2.8.3}$$

Proof. The tightness of T_N and $H_\varrho/|D_N|$ allows us to effectively truncate the union in (2.8.2) to $-M \le k \le M$ and assume $H_\varrho \le m\deg(D_N)$. Recall the event $\mathcal{F}_N(x)$ from (2.7.2) and note that on the event

$$\Big\{\sum_{x\in D_N} 1_{\mathcal{F}_N(x)^{\rm c}} \le 2\Big\} \cap \big\{H_\varrho \le m\deg(D_N)\big\}, \tag{2.8.4}$$

we have

$$\begin{aligned} \widetilde{L}^{D_N}_{H_\varrho} + (\widetilde{L}^{D_N}\circ\theta_{H_\varrho})_{(t_N+1)\deg(D_N)}(x) &\ge \overline{L}^{D_N}_{t_N}(x) \\ &\ge \widetilde{L}^{D_N}_{(t_N-1)\deg(D_N)}(x) \ge (\widetilde{L}^{D_N}\circ\theta_{H_\varrho})_{(t_N-m-1)\deg(D_N)}(x) \end{aligned} \tag{2.8.5}$$

at all but at most two $x\in D_N$. Next set $\mathcal{E}_N^+ := \mathcal{E}_N(t_N+1)$ and $\mathcal{E}_N^- := \mathcal{E}_N(t_N-m-1)$, where $\mathcal{E}_N(t_N')$ is the event $\mathcal{E}_N$ from (2.5.31) but for $\{t_N\}$ replaced by $\{t_N'\}$. Recall the notation $(t_N')^\circ$ for the quantity from (2.5.4). On $\theta_{H_\varrho}^{-1}(\mathcal{E}_N^+\cap\mathcal{E}_N^-\cap\{(k-1)\epsilon\le T_N\le(k+1)\epsilon\})$ we then get an analogue of (2.5.40) of the form

$$\big((t_N+1)^\circ + b_N(t_N+1)^{1/4}\big)\circ\theta_{H_\varrho} \le \widetilde{t}^{+}_{N,k}, \tag{2.8.6}$$

$$\big((t_N-m-1)^\circ - b_N(t_N-m-1)^{1/4}\big)\circ\theta_{H_\varrho} \ge \widetilde{t}^{-}_{N,k} \tag{2.8.7}$$

once N is sufficiently large (independent of k). Consequently, the inequalities

$$\widetilde{L}^{D_N}_{H_\varrho}(x) + (\widehat{L}^{D_N}\circ\theta_{H_\varrho})_{\widetilde{t}^{+}_{N,k}}(x) \ge \overline{L}^{D_N}_{t_N}(x) \ge (\widehat{L}^{D_N}\circ\theta_{H_\varrho})_{\widetilde{t}^{-}_{N,k}}(x) \tag{2.8.8}$$

apply on the same event as well. Lemma 2.7.2 shows that (2.8.8) holds at all but two $x\in D_N$ with P^{x_N}-probability tending to one as $N\to\infty$. This proves the claim. □

Lemma 2.8.2 eliminates the need to consider other starting points than ϱ. Next comes the main issue to be dealt with in the proof of Proposition 2.8.1: Since we are after differences of the local time, we cannot rely on monotonicity as we did earlier; instead we have to estimate the variation of $t\mapsto \widehat{L}^{D_N}_t$ over time intervals of length of order $\epsilon\sqrt{2t_N}$. This is the content of:

Lemma 2.8.3. *For all $\delta > 0$, all $b \in \mathbb{R}$ and all $\{t'_N\}_{N\ge1}$ satisfying $t'_N - t_N = O(\log N)$,*

$$\lim_{\epsilon\downarrow 0}\limsup_{N\to\infty}\frac{1}{W_N}\sum_{x\in D_N} P^\varrho\Big(\widehat{L}^{D_N}_{t'_N}(x) \ge a_N + b\log N,\\ \widehat{L}^{D_N}_{t'_N}(x) - \widehat{L}^{D_N}_{t'_N-\epsilon\sqrt{2t_N}}(x) > \delta\sqrt{2t_N}\Big) = 0. \tag{2.8.9}$$

Proof. The proof is based on tail estimates for the local time which will depend, somewhat sensitively, on a choice of a few parameters. Given $\delta > 0$ let $\epsilon_0 > 0$ and $j_0 \in \mathbb{N}$ be such that

$$(\sqrt{\theta} + \lambda)^2 - (1+\epsilon_0)\theta > \lambda^2 \tag{2.8.10}$$

and that, for all integers $j \ge j_0$,

$$(j-\delta)\frac{\sqrt{\delta}-\sqrt{\epsilon_0}}{\sqrt{\delta}} > (j+1)\Big[\epsilon_0 + \frac{\lambda}{\sqrt{\theta}+\lambda}\Big]. \tag{2.8.11}$$

These choices can be made because $(\theta+\lambda)^2 - \theta^2 > \lambda^2$ and $\frac{\lambda}{\sqrt{\theta}+\lambda} < 1$. Assume $\epsilon \in (0, \epsilon_0]$ and abbreviate $t''_N := t'_N - \epsilon\sqrt{2t_N}$ and $\widetilde{a}_N := a_N + b\log N$. Set M to the least integer such that $(M+1)\sqrt{2t_N} \ge \widetilde{a}_N - (1+\epsilon_0)t''_N$.

Using the Markov property of $t \mapsto \widehat{L}^{D_N}_t(x)$, the probability in (2.8.9) is bounded by

$$P^\varrho\Big(\widehat{L}^{D_N}_{t''_N}(x) \ge \widetilde{a}_N - j_0\sqrt{2t_N}\Big) P^\varrho\Big(\widehat{L}^{D_N}_{\epsilon\sqrt{2t_N}}(x) \ge \delta\sqrt{2t_N}\Big)\\ + \sum_{j=j_0}^{M} P^\varrho\Big(\widehat{L}^{D_N}_{t''_N}(x) \ge \widetilde{a}_N - (j+1)\sqrt{2t_N}\Big) P^\varrho\Big(\widehat{L}^{D_N}_{\epsilon\sqrt{2t_N}}(x) \ge j\sqrt{2t_N}\Big)\\ + P^\varrho\Big(\widehat{L}^{D_N}_{\epsilon\sqrt{2t_N}}(x) \ge (M+1)\sqrt{2t_N}\Big). \tag{2.8.12}$$

We now use [1, Lemma 4.1] to bound the individual probabilities on the right-hand side as follows. First, noting that by our choice of M,

$$\sqrt{2\big(\widetilde{a}_N - (M+1)\sqrt{2t_N}\,\big)} - \sqrt{2t''_N} \tag{2.8.13}$$

grows proportionally to $\log N$ as $N \to \infty$, [1, Lemma 4.1] may be used for the choices $a := \widetilde{a}_N - j_0\sqrt{2t_N}$, $t := t''_N$ and $b := 0$. Noting that W_N defined using $\widetilde{a}_N - j_0\sqrt{2t_N}$ and t''_N instead of a_N and t_N is comparable with W_N, the uniform upper bound on $G^{D_N}(x,x)$ then bounds the very first probability in (2.8.12) by a quantity of order W_N/N^2. The Markov inequality shows

$$P^\varrho\Bigl(\widehat{L}^{D_N}_{\epsilon\sqrt{2t_N}}(x) > \delta\sqrt{2a_N}\Bigr) \le \frac{\epsilon\sqrt{2t_N}}{\delta\sqrt{2a_N}} \tag{2.8.14}$$

and so the first term in (2.8.12) is order $\epsilon W_N/N^2$ (with a constant that depends on j_0).

Next we move to the terms under the sum in (2.8.12). Here we use [1, Lemma 4.1] for the choices $a := \widetilde{a}_N$, $t := t''_N$ and $b := -j\sqrt{2t_N}$ to get, for all $j = j_0, \dots, M+1$,

$$P^\varrho\Bigl(\widehat{L}^{D_N}_{t''_N}(x) \ge \widetilde{a}_N - j\sqrt{2t_N}\Bigr) \le c_1\frac{W_N}{N^2}\,\mathrm{e}^{j\frac{\sqrt{2t_N}}{G^{D_N}(x,x)}\frac{\sqrt{2\widetilde{a}_N}-\sqrt{2t''_N}}{\sqrt{2\widetilde{a}_N}}} \tag{2.8.15}$$

for some constant $c_1 \in (0,\infty)$ independent of $N \ge 1$, $j = 0, \dots, M+1$ and $x \in D_N$. For the second probability under the sum in (2.8.12), we apply [1, Lemma 4.1] with the choices $a := \delta\sqrt{2t_N}$, $t := \epsilon\sqrt{2t_N}$ and $b := (j-\delta)\sqrt{2t_N}$ to get

$$P^\varrho\Bigl(\widehat{L}^{D_N}_{\epsilon\sqrt{2t_N}}(x) \ge j\sqrt{2t_N}\Bigr) \le c_2\,\mathrm{e}^{-(j-\delta)\frac{\sqrt{2t_N}}{G^{D_N}(x,x)}\frac{\sqrt{\delta}-\sqrt{\epsilon}}{\sqrt{\delta}}} \tag{2.8.16}$$

for some constant $c_2 \in (0,\infty)$ independent of $N \ge 1$ and $m \ge 1$. Putting (2.8.15) and (2.8.16) together and invoking (2.8.11) along with the uniform upper bound on $G^{D_N}(x,x)$, the sum over $j = j_0, \dots, M$ in (2.8.12) may be performed with the result of order $\mathrm{e}^{-\alpha\sqrt{\theta}j_0\epsilon_0}W_N/N^2$, uniformly in $x \in D_N$.

Finally, for the stand-alone probability in (2.8.12), one more use of [1, Lemma 4.1] with the choices $a := (M+1)\sqrt{2t_N}$, $t := \epsilon\sqrt{2t_N}$ and $b := 0$ yields

$$P^\varrho\Bigl(\widehat{L}^{D_N}_{\epsilon\sqrt{2t_N}}(x) \ge (M+1)\sqrt{2t_N}\Bigr) \le \frac{c_3}{\sqrt{\log N}}\,\mathrm{e}^{-(1-o(1))\frac{(M+1)\sqrt{2t_N}}{G^{D_N}(x,x)}} \tag{2.8.17}$$

for a constant $c_3 \in (0,\infty)$ independent of, and $o(1) \to 0$ uniformly in, $N \ge 1$ and $x \in D_N$. Using the definition of M, the right hand side of (2.8.17) is order $N^{-2[\sqrt{\theta}+\lambda)^2-(1+\epsilon_0)\theta]+o(1)}$ which is $o(W_N/N^2)$ by $W_N = N^{2(1-\lambda^2)+o(1)}$ and (2.8.10), uniformly in $x \in D_N$. The claim follows by taking $N \to \infty$, followed by $\epsilon \downarrow 0$ and $j_0 \to \infty$. $\square$

We are ready to give:

Proof of Proposition 2.8.1. Let $f \in C_{\mathrm{c}}(D\times\mathbb{R}\times\mathbb{R}^{\mathbb{Z}^2})$ be such that $f(x,h,\phi)$ depends only on coordinates $\{\phi_z\colon z\in\Lambda_r(0)\}$ for some $r>0$ and vanishes unless $|h|\le b$ and $\max_{z\in\Lambda_r(0)}|\phi_z|\le b$, for some $b>0$. Given $\epsilon>0$, let $k\in\mathbb{Z}$ be such that $|T_N\circ\theta_{H_\varrho}-k\epsilon|<\epsilon$. Pick $x\in D_N$ and abbreviate

$$f_{N,r}(x,\ell):=f\biggl(x/N,\frac{\ell(x)-a_N}{\sqrt{2a_N}},\Bigl\{\frac{\ell(x)-\ell(x+z)}{\sqrt{2a_N}}\colon z\in\Lambda_r(0)\Bigr\}\biggr). \tag{2.8.18}$$

Introducing the oscillation of f by

$$\operatorname{osc}_f(\delta):=\sup_{x\in D}\ \sup_{\substack{u,v\in\mathbb{R},\\ |u-v|\le\delta}}\ \sup_{\substack{\phi,\widetilde{\phi}\in\mathbb{R}^{\Lambda_r(0)},\\ \max_{z\in\Lambda_r(0)}|\phi_z-\widetilde{\phi}_z|\le 2\delta}}\bigl|f(x,u,\phi)-f(x,v,\widetilde{\phi})\bigr|, \tag{2.8.19}$$

the difference

$$f_{N,r}\bigl(x,\overline{L}^{D_N}_{t_N}\bigr)-f_{N,r}\Bigl(x,(\widehat{L}^{D_N}\circ\theta_{H_\varrho})_{\widetilde{t}^-_{N,k}}\Bigr) \tag{2.8.20}$$

is bounded in absolute value by the sum over $z\in\Lambda_r(x)$ of three terms: $2\|f\|_\infty 1_{\widetilde{\mathcal{F}}_N(z)^{\mathrm{c}}}$,

$$2\|f\|_\infty 1_{\widetilde{\mathcal{F}}_N(z)\cap\{H_z<H_\varrho\}}\Bigl(1_{\{(\widehat{L}^{D_N}\circ\theta_{H_\varrho})_{\widetilde{t}^-_{N,k}}(z)\ge a_N-2b\sqrt{2a_N}\}}+1_{\{\overline{L}^{D_N}_{t_N}(z)\ge a_N-2b\sqrt{2a_N}\}}\Bigr) \tag{2.8.21}$$

and

$$\begin{aligned}1_{\widetilde{\mathcal{F}}_N(z)\cap\{H_z>H_\varrho\}}&\Bigl(\operatorname{osc}_f(\delta)+\|f\|_\infty 1_{\{|\overline{L}^{D_N}_{t_N}(z)-(\widehat{L}^{D_N}\circ\theta_{H_\varrho})_{\widetilde{t}^-_{N,k}}(z)|>\delta\sqrt{2a_N}\}}\Bigr)\\ &\times\Bigl(1_{\{(\widehat{L}^{D_N}\circ\theta_{H_\varrho})_{\widetilde{t}^-_{N,k}}(z)\ge a_N-2b\sqrt{2a_N}\}}+1_{\{\overline{L}^{D_N}_{t_N}(z)\ge a_N-2b\sqrt{2a_N}\}}\Bigr).\end{aligned} \tag{2.8.22}$$

To simplify estimates, introduce the events

$$\mathcal{G}_N(x):=\Bigl\{\widetilde{L}^{D_N}_{H_\varrho}(x)+(\widehat{L}^{D_N}\circ\theta_{H_\varrho})_{\widetilde{t}^+_{N,k}}(x)\ge a_N-2b\sqrt{2a_N}\Bigr\}\cap\{H_x<H_\varrho\} \tag{2.8.23}$$

and

$$\mathcal{H}_N(x):=\Bigl\{\widehat{L}^{D_N}_{\widetilde{t}^+_{N,k}}(x)\ge a_N-2b\sqrt{2a_N}\Bigr\}\cap\Bigl\{\widehat{L}^{D_N}_{\widetilde{t}^+_{N,k}}(x)-\widehat{L}^{D_N}_{\widetilde{t}^-_{N,k}}(x)>\delta\sqrt{2a_N}\Bigr\}. \tag{2.8.24}$$

Then (2.8.21) is bounded by $4\|f\|_\infty 1_{\mathcal{G}_N(z)}$ while (2.8.22) is bounded by

$$2\operatorname{osc}_f(\delta)1_{\{(\widehat{L}^{D_N}\circ\theta_{H_\varrho})_{\widetilde{t}^+_{N,k}}(z)\ge a_N-2b\sqrt{2a_N}\}}+2\|f\|_\infty 1_{\mathcal{H}_N(z)}\circ\theta_{H_\varrho}. \tag{2.8.25}$$

Summarizing these estimates, and writing $\widehat{\zeta}_N^{D,\mathrm{loc}}(t'_N)$ for the measure in (2.2.31) except with L^{D_N} replaced by $\widehat{L}^{D_N}$ and t_N by t'_N, we thus get that, on $\{|T_N\circ\theta_{H_\varrho}-k\epsilon|<\epsilon\}$,

$$\begin{aligned}\Bigl|\langle\overline{\zeta}_N^{D,\mathrm{loc}},f\rangle-\frac{W_N(\widetilde{t}^-_{N,k})}{W_N}\bigl\langle\widehat{\zeta}_N^{D,\mathrm{loc}}(\widetilde{t}^-_{N,k}),f\bigr\rangle\circ\theta_{H_\varrho}\Bigr|&\\ \le 4\|f\|_\infty|\Lambda_r(0)|\frac{1}{W_N}\sum_{x\in D_N}\bigl(1_{\widetilde{\mathcal{F}}_N(x)^{\mathrm{c}}}+1_{\mathcal{G}_N(x)}+1_{\mathcal{H}_N(x)}\circ\theta_{H_\varrho}\bigr)&\\ +2\,\mathrm{osc}_f(\delta)|\Lambda_r(0)|\frac{W_N(\widetilde{t}^+_{N,k})}{W_N}\bigl\langle\widehat{\zeta}_N^{D}(\widetilde{t}^+_{N,k}),1_D\otimes1_{[-2b,\infty)}\bigr\rangle\circ\theta_{H_\varrho}&\qquad(2.8.26)\end{aligned}$$

Using Lemmas 2.8.2, 2.8.3 and 2.6.3, the first term on the right tends to zero in P^{x_N}-probability as $N\to\infty$ and $\epsilon\downarrow0$ for each $\delta>0$. The tightness of $\widehat{\zeta}_N^D$ measures (under P^ϱ) along with the uniform continuity of f ensure that the second term tends to zero in P^{x_N}-probability as $N\to\infty$ and $\delta\downarrow0$.

To finish the proof, note that by [1, Theorem 2.6] and the argument underlying Proposition 2.4.3 we have, under P^ϱ,

$$\widehat{\zeta}_N^{D,\mathrm{loc}}(t'_N)\otimes\delta_{T_N}\ \underset{N\to\infty}{\overset{\mathrm{law}}{\longrightarrow}}\ \widehat{\zeta}^D\otimes\widehat{\nu}_\lambda\otimes\delta_T \qquad (2.8.27)$$

for any sequence $\{t'_N\}_{N\ge1}$ such that $t'_N-t_N=o(t_N)$, where $\widehat{\zeta}^D$ is related to T as in (2.5.48). Since $W_N(\widetilde{t}^-_{N,k})/W_N=(\mathrm{e}^{-\alpha\lambda T_N(\widetilde{t}^-_{N,k})}\circ\theta_{H_\varrho})\mathrm{e}^{O(\epsilon)}$ on $\{|T_N\circ\theta_{H_\varrho}-k\epsilon|<\epsilon\}\cap\mathcal{E}_N^-\circ\theta_{H_\rho}$, from (2.8.26) and the tightness of the random variables $\{T_N\}_{N\ge1}$ and $\{H_\varrho/|D_N|\}_{N\ge1}$ we get, by taking $N\to\infty$ followed by $\delta\downarrow0$, $\epsilon\downarrow0$ and $m\to\infty$, under P^{x_N},

$$\overline{\zeta}_N^{D,\mathrm{loc}}\ \underset{N\to\infty}{\overset{\mathrm{law}}{\longrightarrow}}\ \mathrm{e}^{-\alpha\lambda T}\widehat{\zeta}^D\otimes\widehat{\nu}_\lambda. \qquad (2.8.28)$$

This is the desired claim. □

With Proposition 2.8.1 in hand, we are ready to tackle:

Proof of Theorem 2.2.7, thick points. First observe that the tightness of $\{\zeta_N^D\colon N\ge1\}$ implies tightness of $\{\zeta_N^{D,\mathrm{loc}}\colon N\ge1\}$ and so we may consider subsequential distributional limits $\zeta^{D,\mathrm{loc}}$ of the latter. Using Proposition 2.8.1 in the argument from the proof of Theorem 2.2.3

we conclude that every such subsequential weak limit obeys

$$\langle \zeta^{D,\mathrm{loc}}, f^{*\mathfrak{n}}\rangle \overset{\mathrm{law}}{=} \langle \widetilde{\zeta}^D \otimes \widehat{\nu}_\lambda, f\rangle \tag{2.8.29}$$

for all $f \in C_{\mathrm{c}}(D\times\mathbb{R}\times\mathbb{R}^{\mathbb{Z}^2})$, where

$$f^{*\mathfrak{n}}(x,h,\phi) := E\Big[f\big(x, h+\mathfrak{n}_0, \{\mathfrak{n}_0-\mathfrak{n}_z+\phi_z\colon z\in\mathbb{Z}^2\}\big)\Big], \tag{2.8.30}$$

for $\{\mathfrak{n}_z\colon z\in\mathbb{Z}^2\}$ i.i.d. $\mathcal{N}(0,\frac18)$.

We now proceed similarly as in (2.7.48–2.7.51): Given any $\tilde f \in C_{\mathrm{c}}(\mathbb{R}\times\mathbb{R}^{\mathbb{Z}^2})$ and any Borel $A\subseteq D$ with $\mathrm{Leb}(A)>0$, the explicit form of $\widetilde{\zeta}^{D,\mathrm{loc}}$ gives the pointwise equality

$$\begin{aligned}&\big\langle \zeta^{D,\mathrm{loc}}, (1_A\otimes \tilde f)^{*\mathfrak{n}}\big\rangle \\ &\qquad = \big\langle \zeta^{D,\mathrm{loc}}, (1_A\otimes 1_{[0,\infty)}\otimes 1_{\mathbb{R}^{\mathbb{Z}^2}})^{*\mathfrak{n}}\big\rangle\, \alpha\lambda \int \mathrm{d}h\, \mathrm{e}^{-\alpha\lambda h}\otimes \widehat{\nu}_\lambda(\mathrm{d}\phi)\tilde f(h,\phi).\end{aligned} \tag{2.8.31}$$

Abbreviating $\beta := -\alpha\lambda$, for each A as above, the measure ζ_A on $\mathbb{R}\times\mathbb{R}^{\mathbb{Z}^2}$ defined by

$$\zeta_A(B) := \frac{\zeta^{D,\mathrm{loc}}(A\times B)}{\alpha\lambda\big\langle \zeta^{D,\mathrm{loc}}, (1_A\otimes 1_{[0,\infty)}\otimes 1_{\mathbb{R}^{\mathbb{Z}^2}})^{*\mathfrak{n}}\big\rangle} \tag{2.8.32}$$

then “solves” for μ from the convolution equation

$$\begin{aligned}&\int_{\mathbb{R}\times\mathbb{R}^{\mathbb{Z}^2}} \mu(\mathrm{d}h\mathrm{d}\phi)\, E\big[\, f(h+\mathfrak{n}_0, \{\mathfrak{n}_0-\mathfrak{n}_z+\phi_z\colon z\in\mathbb{Z}^2\})\big] \\ &\qquad\qquad = \int_{\mathbb{R}\times\mathbb{R}^{\mathbb{Z}^2}} \mathrm{d}h\; \mathrm{e}^{\beta h}\otimes\widehat{\nu}_\lambda(\mathrm{d}\phi)\, f(h,\phi)\end{aligned} \tag{2.8.33}$$

for all $f\in C_{\mathrm{c}}(\mathbb{R}\times\mathbb{R}^{\mathbb{Z}^2})$. To solve this equation, we need:

Lemma 2.8.4. *For each $x,y\in\mathbb{Z}^2$, let*

$$\widetilde{C}(x,y) := \mathfrak{a}(x)+\mathfrak{a}(y)-\mathfrak{a}(x-y)-\frac18\big[1-\delta_{x,0}-\delta_{y,0}+\delta_{x,y}\big]. \tag{2.8.34}$$

Then $\widetilde{C}$ is symmetric and positive semidefinite and so there exists a centered Gaussian process $\{\widetilde{\phi}_x\colon x\in\mathbb{Z}^2\}$ with covariance $\widetilde{C}$. This process then satisfies (2.2.35).

Proof. Recall that (in our normalization) $\mathfrak{a}$ solves the equation $\Delta\mathfrak{a} = \delta_0$ and so using Fourier transform techniques we get

$$\mathfrak{a}(x) = \int_{(-\pi,\pi)^2} \frac{\mathrm{d}k}{(2\pi)^2} \frac{1-\mathrm{e}^{-\mathrm{i}k\cdot x}}{\widehat{D}(k)}, \tag{2.8.35}$$

where

$$\widehat{D}(k) := 4\sin(k_1/2)^2 + 4\sin(k_2/2)^2. \tag{2.8.36}$$

Let $v \in \ell^2(\mathbb{Z}^2)$ and denote by $\hat{v}(k) := \sum_{x\in\mathbb{Z}^2} v(x)\mathrm{e}^{\mathrm{i}k\cdot x}$ the Fourier transform of v. A calculation then shows

$$(v, \widetilde{C}v) = \int_{(-\pi,\pi)^2} \frac{\mathrm{d}k}{(2\pi)^2} \left(\frac{1}{\widehat{D}(k)} - \frac{1}{8}\right) \bigl|\hat{v}(0) - \hat{v}(k)\bigr|^2 \tag{2.8.37}$$

Noting that $\widetilde{D}(k) \le 8$, we get that $\widetilde{C}$ is indeed positive semidefinite. We now readily check that $x, y \mapsto \frac{1}{8}[1 - \delta_{x,0} - \delta_{y,0} + \delta_{x,y}]$ is the covariance of $\{n_0 - n_z \colon z \in \mathbb{Z}^2\}$ for $\{n_z \colon z \in \mathbb{Z}^2\}$ i.i.d. $\mathcal{N}(0, \frac{1}{8})$, and so (2.2.35) holds as well. $\square$

The solution of (2.8.33) will require the following extension of Lemma 2.7.4:

Lemma 2.8.5. *Let $\widetilde{\phi}$ be a centered Gaussian process on $\mathbb{Z}^2$ such that, for some $\beta \in \mathbb{R}$ and some $\sigma^2 > 0$, the process $\{\widetilde{\phi}_x + n_0 - n_z \colon z \in \mathbb{Z}^2\}$ with $\{n_z \colon z \in \mathbb{Z}^2\}$ i.i.d. $\mathcal{N}(0, \sigma^2)$ has the law of the pinned DGFF ϕ. Denote*

$$\nu_{\lambda,\beta}(A) := P\Bigl(\widetilde{\phi} + \lambda\alpha\mathfrak{a} + \beta\sigma^2 1_{\mathbb{Z}^2\smallsetminus\{0\}} \in A\Bigr). \tag{2.8.38}$$

Then (2.8.33) *is solved uniquely by*

$$\mu(\mathrm{d}h\mathrm{d}\phi) = \mathrm{e}^{-\frac{1}{2}\beta^2\sigma^2+\beta h}\mathrm{d}h \otimes \nu_{\lambda,\beta}(\mathrm{d}\phi). \tag{2.8.39}$$

Proof. Denote $\widetilde{\mu}(\mathrm{d}h\mathrm{d}\phi) := \mathrm{e}^{\frac{1}{2}\beta^2\sigma^2-\beta h}\mu(\mathrm{d}h\mathrm{d}\phi)$. Pick $\{t_z \colon z \in \mathbb{Z}^2\}$ with finite support and $t_0 = 0$ and, writing $\langle\cdot,\cdot\rangle$ for the inner product in $\ell^2(\mathbb{Z}^2)$, apply (2.8.33) to the test function $h, \phi \mapsto \mathrm{e}^{-\beta h} f(h) \exp\{\langle t, \phi\rangle\}$ with a non-negative $f \in C_{\mathrm{c}}(\mathbb{R})$. (This is permissible in light of the Monotone Convergence Theorem.) Writing x for $h + n_0$ then turns (2.8.39) into

$$\int \widetilde{\mu}(\mathrm{d}h\mathrm{d}\phi) \otimes \mathrm{d}x\, \mathrm{e}^{\langle t,\phi\rangle} \frac{1}{\sqrt{2\pi\sigma^2}}\, \mathrm{e}^{-\frac{1}{2\sigma^2}(x-h)^2} E\bigl(\mathrm{e}^{-\langle t,n\rangle}\bigr)\, \mathrm{e}^{\bar{t}(x-h)} \mathrm{e}^{-\frac{1}{2}\beta^2\sigma^2+\beta h} \mathrm{e}^{-\beta x} f(x)$$
$$= \int \widehat{\nu}_\lambda(\mathrm{d}\phi) \otimes \mathrm{d}x\, \mathrm{e}^{\langle t,\phi\rangle} f(x) \tag{2.8.40}$$

where $\bar{t} := \sum_{z\in\mathbb{Z}^2} t_z$. By assumption we have

$$\{\phi_z\colon z\in\mathbb{Z}^d\} \overset{\text{law}}{=} \{\widetilde{\phi}_z + n_0 - n_z\colon z\in\mathbb{Z}^2\} \tag{2.8.41}$$

and so, in light of $t_0 = 0$,

$$\begin{aligned}\int \widehat{\nu}_\lambda(\mathrm{d}\phi)\mathrm{e}^{\langle t,\phi\rangle} &= \int P(\mathrm{d}\phi)\mathrm{e}^{\langle t,\phi+\alpha\lambda\mathfrak{a}\rangle} \\ &= \int P(\mathrm{d}\widetilde{\phi})E\bigl(\mathrm{e}^{\langle t,\widetilde{\phi}+n_0-n+\alpha\lambda\mathfrak{a}\rangle}\bigr) \\ &= \int \nu_{\lambda,\beta}(\mathrm{d}\widetilde{\phi})\,\mathrm{e}^{\langle t,\widetilde{\phi}\rangle}\,E\bigl(\mathrm{e}^{-\langle t,n\rangle}\bigr)E(\mathrm{e}^{\bar{t}(n_0-\beta\sigma^2)}),\end{aligned} \tag{2.8.42}$$

where the expectation is over $\{n_z\colon z\in\mathbb{Z}^2\}$. Using this in (2.8.40) and cancelling $E\bigl(\mathrm{e}^{-\langle t,n\rangle}\bigr)$ on both sides, the identity $E(\mathrm{e}^{\bar{t}(n_0-\beta\sigma^2)}) = \mathrm{e}^{\frac12\bar{t}^2\sigma^2-\beta\bar{t}\sigma^2}$ along with the fact that functions $f\in C_{\mathrm{c}}(\mathbb{R})$ separate points yield

$$\int \widetilde{\mu}(\mathrm{d}h\mathrm{d}\phi)\,\mathrm{e}^{\langle t,\phi\rangle}\,\frac{1}{\sqrt{2\pi\sigma^2}}\,\mathrm{e}^{-\frac{1}{2\sigma^2}(x-h)^2}\,\mathrm{e}^{\bar{t}(x-h)}\mathrm{e}^{-\beta x}\mathrm{e}^{-\frac12\bar{t}^2\sigma^2+\beta\bar{t}\sigma^2}\mathrm{e}^{-\frac12\beta^2\sigma^2+\beta h} = \int \nu_{\lambda,\beta}(\mathrm{d}\widetilde{\phi})\,\mathrm{e}^{\langle t,\widetilde{\phi}\rangle} \tag{2.8.43}$$

for all $x\in\mathbb{R}$. (Continuity is used to get from Lebesgue a.e. $x\in\mathbb{R}$ to all $x\in\mathbb{R}$.) The five exponentials on the left combine into

$$\mathrm{e}^{-\frac{1}{2\sigma^2}(x-h-\bar{t}\sigma^2)^2-\beta(x-h-\bar{t}\sigma^2)-\frac12\beta^2\sigma^2} = \mathrm{e}^{-\frac{1}{2\sigma^2}(x-h-\bar{t}\sigma^2+\beta\sigma^2)^2}. \tag{2.8.44}$$

Shifting x by $\bar{t}\sigma^2+\beta\sigma^2$ and scaling it by σ^2 shows that $\widehat{\mu}(\mathrm{d}h\mathrm{d}\phi) := \frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{1}{2\sigma^2}h^2}\widetilde{\mu}(\mathrm{d}h\mathrm{d}\phi)$ obeys

$$\int \widehat{\mu}(\mathrm{d}h\mathrm{d}\phi)\,\mathrm{e}^{\langle t,\phi\rangle-xh} = \int \nu_{\lambda,\beta}(\mathrm{d}\widetilde{\phi})\,\mathrm{e}^{\langle t,\widetilde{\phi}\rangle}\mathrm{e}^{\frac12 x^2\sigma^2} \tag{2.8.45}$$

for all $x\in\mathbb{R}$ and all $\{t_z\colon z\in\mathbb{Z}^2\}$ with finite support and $t_0=0$.

The restriction to $t_0=0$ is irrelevant in (2.8.45) since $\nu_{\lambda,\beta}$ is concentrated on $\{\phi\colon \phi_0=0\}$ and, by (2.8.33) so is μ and thus also $\widehat{\mu}$. The right-hand side of (2.8.45) is the Laplace transform of the product of the law of $\mathcal{N}(0,\sigma^2)$ and $\nu_{\lambda,\beta}$. Hence

$$\widetilde{\mu}(\mathrm{d}h\mathrm{d}\phi) = \mathrm{d}h\otimes\nu_{\lambda,\beta}(\mathrm{d}\phi) \tag{2.8.46}$$

and so the claim follows from the definition of $\widetilde{\mu}$. $\square$

Returning to the main line of the proof of Theorem 2.2.7, it remains to observe that the denominator in (2.8.32) has the law of

$$\sqrt{\frac{\sqrt{\theta}}{\sqrt{\theta}+\lambda}}\;\mathfrak{c}(\lambda)\,\mathrm{e}^{\alpha\lambda(\mathfrak{d}(x)-1)Y}\,Z_\lambda^{D,0}(\mathrm{d}x), \tag{2.8.47}$$

for $Y = \mathcal{N}(0,\sigma_D^2)$ independent of $Z_\lambda^{D,0}$. Lemma 2.8.5 with $\beta := -\alpha\lambda$ and $\sigma^2 := \frac{1}{8}$ then yields the claim. $\square$

Moving to the thin points, here we go directly for:

Proof of Theorem 2.2.7, thin points. The proof is considerably simpler because, as a few times earlier, certain key inequalities go in a more favorable direction. Following the argument and the notation from the proof for the thick points, we derive an analogue of (2.8.26) with the events $\mathcal{G}_N(x)$ and $\mathcal{H}_N(x)$ replaced by

$$\widetilde{\mathcal{G}}_N(x) := \Bigl\{(\widehat{L}^{D_N}\circ\theta_{H_\varrho})_{\widetilde{t}^-_{N,k}}(x) \le a_N + 2b\sqrt{2a_N}\Bigr\} \cap \{H_x < H_\varrho\} \tag{2.8.48}$$

and

$$\widetilde{\mathcal{H}}_N(x) := \Bigl\{\widehat{L}^{D_N}_{\widetilde{t}^-_{N,k}}(x) \le a_N + 2b\sqrt{2a_N}\Bigr\} \cap \Bigl\{\widehat{L}^{D_N}_{\widetilde{t}^+_{N,k}}(x) - \widehat{L}^{D_N}_{\widetilde{t}^-_{N,k}}(x) > \delta\sqrt{2a_N}\Bigr\}, \tag{2.8.49}$$

respectively, and $1_{[-2b,\infty)}$ replaced by $1_{(-\infty,2b]}$. The P^{x_N}-probability of event $\widetilde{\mathcal{G}}_N(x)$ is controlled using Lemma 2.6.5. Unlike $\mathcal{H}_N(x)$ which required a non-trivial decomposition in the proof of Lemma 2.8.3, the two events constituting $\widetilde{\mathcal{H}}_N(x)$ can be directly separated using the Markov property of $t \mapsto \widehat{L}^{D_N}_t$. The expected sum over $1_{\widetilde{\mathcal{H}}_N(x)}\circ\theta_{H_\varrho}$ is then shown to be order ϵW_N by (2.8.14) and the fact that $E^\varrho\langle\widehat{\zeta}^D_N(\widetilde{t}^-_{N,k}), 1_{(-\infty,2b]}\rangle$ is bounded in $N \ge 1$. As a consequence, we get that, under P^{x_N},

$$\overline{\zeta}^{D,\mathrm{loc}}_N \;\underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}}\; \widetilde{\zeta}^D \otimes \widehat{\nu}_\lambda, \tag{2.8.50}$$

where $\widetilde{\zeta}^D$ is the measure on the right of (2.2.21) without the term $\mathrm{e}^{-\alpha^2\lambda^2/16}$ and $\widehat{\nu}_\lambda$ is the law of $\phi - \alpha\lambda\mathfrak{a}$. The rest of the argument for the thick points may be followed literally. $\square$

2.8.2 Avoided points

The proof is a variation on the themes encountered in the proof of convergence of the measure associated with the light and avoided points. In particular, since the local time vanishes at the avoided points, we will be able to use monotonicity arguments. The following observation will be useful:

Lemma 2.8.6. *Let μ be a probability measure on $\mathbb{N}^{\mathbb{Z}^2}$ with samples denoted by $\{\hat{n}_z\colon z\in\mathbb{Z}^2\}$. Let $\{\tau_j(x)\colon j\ge 1,\, x\in\mathbb{Z}^2\}$ be i.i.d. Exponential(1), independent of $\{\hat{n}_z\colon z\in\mathbb{Z}^2\}$. Then for any $t\in(-1,\infty)^{\mathbb{Z}^2}$ with finite support,*

$$E\exp\Bigl\{-\sum_{z\in\mathbb{Z}^2}t(z)\sum_{j=1}^{\hat{n}_z}\tau_j(z)\Bigr\}=E\exp\Bigl\{-\sum_{z\in\mathbb{Z}^2}t'(z)\hat{n}_z\Bigr\},\tag{2.8.51}$$

where $t'(z):=\log(1+t(z))$.

Proof. This boils down to a calculation of the Laplace transform of Exponential(1). □

Proof of Theorem 2.2.8. We will establish the existence and uniqueness of the law $\nu_u^{\text{RI,dis}}$ as part of the proof of the convergence. Let $\tilde{f}\in C(\overline{D})$ be non-negative, pick $t\in(0,\infty)^{\mathbb{Z}^2}$ with finite support and consider the test function

$$f_t(x,\phi):=\tilde{f}(x)\mathrm{e}^{-\langle t,\phi\rangle}\tag{2.8.52}$$

where, abusing notation as before, $\langle\cdot,\cdot\rangle$ denotes the canonical inner product in $\ell^2(\mathbb{Z}^2)$. The function $x,h,\phi\mapsto\mathrm{e}^{-hn}f_t(x,\phi)$ is non-increasing in both h and the coordinates of ϕ and so, thanks to Lemma 2.8.2, (2.5.55) applies to f replaced by $\mathrm{e}^{-hn}f_t$ and $\widetilde{\vartheta}_N^D$ by

$$\overline{\vartheta}_N^D:=\frac{1}{\widehat{W}_N}\sum_{x\in D_N}\delta_{x/N}\otimes\delta_{\overline{L}_{t_N}^{D_N}(x)}\otimes\delta_{\{\overline{L}_{t_N}^{D_N}(x+z)\colon z\in\mathbb{Z}^2\}}.\tag{2.8.53}$$

Let $\overline{\kappa}_N^D$ be the measure tracking the local behavior of $\overline{L}_{t_N}^{D_N}(x+z)\colon z\in\mathbb{Z}^2$ around every point x where $\overline{L}_{t_N}^{D_N}(x)=0$ which, we note, is almost surely equivalent to $L_{t_N}^{D_N}(x)=0$. Taking the limits $N\to\infty$ and $n\to\infty$, from [1, Theorem 2.8] we then get, under P^{x_N},

$$\langle\overline{\kappa}_N^{D,\text{loc}},f_t\rangle\ \underset{N\to\infty}{\overset{\text{law}}{\longrightarrow}}\ \langle\widetilde{\kappa}^D\otimes\nu_\theta^{\text{RI}},f_t\rangle,\tag{2.8.54}$$

where $\widetilde{\kappa}^D$ is the law on the right-hand side of (2.5.59).

Next we observe that, by Lemma 2.8.6 and the fact that $4L_{t_n}^{D_N}(x)$ is a natural,

$$E^{\varrho}\big(\langle\overline{\kappa}_N^{D,\mathrm{loc}}, f_t\rangle \,\big|\, \sigma(X)\big) = \langle\kappa_N^{D,\mathrm{loc}}, f_{t'}\rangle \tag{2.8.55}$$

where $t'(z) := 4\log(1+t(z)/4)$. From (2.7.10) and (2.8.54) we then get that every subsequential weak limit $\kappa^{D,\mathrm{loc}}$ of $\{\kappa_N^{D,\mathrm{loc}}\colon N\ge 1\}$ obeys

$$\langle\kappa^{D,\mathrm{loc}}, f_{t'}\rangle \overset{\mathrm{law}}{=} \langle\widetilde{\kappa}^D\otimes\nu_\theta^{\mathrm{RI}}, f_t\rangle \tag{2.8.56}$$

jointly for all $t\in(0,\infty)^{\mathbb{Z}^2}$ with finite support and all $\tilde f\in C(\overline{D})$. Since ν_θ^{RI} is non-random, this is readily turned into the a.s. identity

$$\int \kappa^{D,\mathrm{loc}}(\mathrm{d}x\mathrm{d}\ell)\tilde f(x)\mathrm{e}^{-\langle t',\ell\rangle} = \Big(\int\widetilde{\kappa}^D(\mathrm{d}x)\tilde f(x)\Big)\int\nu_\theta^{\mathrm{RI}}(\mathrm{d}\phi)\mathrm{e}^{-\langle t,\phi\rangle}. \tag{2.8.57}$$

This along with the fact that

$$\mathrm{e}^{-\langle t',\ell\rangle} = E\exp\Big\{-\sum_{z\in\mathbb{Z}^2} t(z)\frac14\sum_{j=1}^{4\ell(z)}\tau_j(z)\Big\} \tag{2.8.58}$$

for $\{\tau_j(z)\colon j\ge 1,\, z\in\mathbb{Z}^2\}$ independent i.i.d. Exponential(1) implies that

$$\kappa^{D,\mathrm{loc}} = \widetilde{\kappa}^D\otimes\nu_\theta^{\mathrm{RI,dis}} \tag{2.8.59}$$

where $\nu_\theta^{\mathrm{RI,dis}}$ is a measure as described in the statement.

This shows that a measure $\nu_u^{\mathrm{RI,dis}}$ exists with the stated properties for all $u\in(0,1)$. Since adding independent samples from this measure for parameters $u\in(0,1)$ and $v\in(0,1)$ gives us a sample from the measure for parameter $u+v$, the existence extends to all $u>0$. The measure is unique by Lemma 2.8.6 and so is thus the distributional limit $\kappa^{D,\mathrm{loc}}$. This completes the proof. $\square$

Chapter 3

Pollution-sensitivity of bootstrap percolation

We study the terminal configuration in polluted bootstrap percolation on $\mathbb{Z}^2$ started from a Bernoulli configuration where, given parameters $p, q \in [0,1]$ with $p + q \leq 1$, a vertex is initially declared occupied with probability p, polluted with probability q and vacant otherwise, independently of other vertices. The update rule turns a vacant site into occupied whenever it has at least two occupied neighbors; polluted vertices stay polluted forever. Setting $q = \lambda p^2$, we define two thresholds, λ_1^{c} and λ_2^{c}, such that, in the limit $p \downarrow 0$, the terminal occupation is asymptotically full when $\lambda < \lambda_1^{\mathrm{c}}$ and asymptotically empty when $\lambda > \lambda_2^{\mathrm{c}}$. This formally reproduces the main result of McDonald and Gravner [55]; the novelty is that the thresholds are now defined using the same continuum (ordinary) percolation model of blocking contours — with λ_1^{c} being the threshold for exponential decay of connectivities and λ_2^{c} for the appearance of an infinite connected component. We expect, although are unable to prove at the moment, that $\lambda_1^{\mathrm{c}} = \lambda_2^{\mathrm{c}}$.

3.1 Setting and main results

We begin by reviewing the settings for the 2-neighbor PBP as described in the introduction. Consider the state space

$$S := \{0 = \text{vacant } (\square),\ 1 = \text{occupied } (\blacksquare),\ 2 = \text{polluted } (\boxtimes)\} \tag{3.1.1}$$

and the space of configurations $\Omega = S^{\mathbb{Z}^2}$. Our bootstrap process will be defined in terms of the dynamics in Ω. Since we are predominantly interested in the occupied sites, for each configuration $\omega \in \Omega$ we define

$$\mathrm{Occ}(\omega) := \{x \in \mathbb{Z}^2 : \omega(x) = 1\}. \tag{3.1.2}$$

For each site $x \in \mathbb{Z}^2$, the *neighborhood* of x is the set $\mathcal{N}(x)$ of nearest-neighbors around x in the square lattice $\mathbb{Z}^2$,

$$\mathcal{N}(x) := \{y \in \mathbb{Z}^2 : \|x - y\|_2 = 1\} = x + \mathcal{N}(0), \tag{3.1.3}$$

where $\|\cdot\|_2$ denotes the Euclidean distance in $\mathbb{R}^2$. Then we define the *evolution operator* $\mathcal{B} : \Omega \to \Omega$ on the space of configurations by

$$(\mathcal{B}\omega)(x) := \begin{cases} 1, & \text{if } \omega(x) = 0 \text{ and } |\mathcal{N}(x) \cap \mathrm{Occ}(\omega)| \geq 2, \\ \omega(x), & \text{otherwise.} \end{cases} \tag{3.1.4}$$

In other words, $\mathcal{B}$ updates each vacant site on ω to an occupied site if it has at least 2 occupied neighbors. The bootstrap process is then defined as the sample path $(\mathcal{B}^t\omega)_{t\in\mathbb{N}_0}$ of iterated evolutions, where $\mathcal{B}^t$ denotes the t-fold composition of $\mathcal{B}$. Noting that each site is updated at most once during these iterations, we can unambiguously define the *terminal configuration* $\mathcal{B}^\infty\omega$ of ω by the following pointwise limit

$$\mathcal{B}^\infty\omega(x) := \lim_{t\to\infty} \mathcal{B}^t\omega(x), \qquad \forall x \in \mathbb{Z}^2. \tag{3.1.5}$$

We refer to ω as the *initial configuration* of the process $(\mathcal{B}^t\omega)_{t\geq 0}$. We remark that $\mathcal{B}$ is measurable relative to the product σ-algebra on Ω. This is because $\{\omega : \mathcal{B}\omega(x) = s\}$ is a cylinder set for each $x \in \mathbb{Z}^2$ and $s \in S$. The same then holds for $\mathcal{B}^t$ for all $t \in \mathbb{N}_0$, and hence also for $\mathcal{B}^\infty$.

With the dynamics at our hands, we are interested in the bootstrap process when the initial configuration is random. In this work, the distribution of initial configurations will be specified by two parameters, p and q, such that all of p, q, and $1-p-q$ lie in $[0,1]$. For each pair of parameters (p,q) in this range, denote by $\mathbb{P}_{p,q}$ the probability law such that, if ω is sampled from $\mathbb{P}_{p,q}$, then $\{\omega(x) : x \in \mathbb{Z}^2\}$ is a family of independent S-valued random variables that are occupied with probability p, polluted with probability q, and vacant with probability $1-p-q$. Then we introduce two quantities $\theta(p,\lambda)$ and $\phi(p,q)$ by

$$\begin{aligned} \theta(p,q) &:= \mathbb{P}_{p,q}(0 \text{ is occupied in } \mathcal{B}^\infty\omega), \\ \phi(p,q) &:= \mathbb{P}_{p,q}(\text{there exists an infinite occupied cluster in } \mathcal{B}^\infty\omega). \end{aligned} \tag{3.1.6}$$

For this setting, Gravner and McDonald [55] proved:

Theorem 3.1.1. *Assume $p,q>0$ with $1-p-q \geq 0$ and let $\lambda(p,q) := q/p^2$. Then there exist finite, positive constants $\lambda_1 \leq \lambda_2$ such that:*

(1) If $\lambda(p,q) < \lambda_1$, then $\phi(p,q) = 1$ for p sufficiently small and $\theta(p,q) \to 1$ as $p \downarrow 0$.

(2) If $\lambda(p,q) > \lambda_2$, $\phi(p,q) = 0$ for p sufficiently small and $\theta(p,q) \to 0$ as $p \downarrow 0$.

This theorem identifies a correct p *versus* q scaling using the scale factor $\lambda(p,q) := q/p^2$ so that the terminal configuration under $\mathbb{P}_{p,q}$ exhibits a non-trivial phase transition as a function of λ in the limit $p,q \downarrow 0$ with $\lambda(p,q)$ fixed to λ. The reference [55] dealt with each regime in Theorem 3.1.1 using a rather different approach making it all but impossible to address the question whether the two regimes are separated by a sharp threshold. The main contribution of this part of the thesis is that we establish the above with constants λ_1 and λ_2 replaced by quantities defined in terms of the same continuum percolation model. This is the content of:

Theorem 3.1.2. *Define λ_1^{c} and λ_2^{c} by* (1.2.4)*. Then we have*

(1) If $\lambda < \lambda_1^c$, then $\phi(p, \lambda p^2) = 1$ for p sufficiently small.

(2) If $\lambda > \lambda_1^c$, then $\phi(p, \lambda p^2) = 0$ for p sufficiently small.

We will give full details of the definition of the continuum percolation model underlying (1.2.4) in Section 3.3.

3.1.1 Notation

Let us start by introducing some notation. For each set $\Lambda \subseteq \mathbb{Z}^2$, its *outer boundary* $\partial\Lambda$ is defined as the set of all sites in $\mathbb{Z}^2 \setminus \Lambda$ that are adjacent to some site of Λ,

$$\partial\Lambda := \{y \in \mathbb{Z}^2 : y \in \mathcal{N}(z) \setminus \Lambda \text{ for some } z \in \Lambda\}. \tag{3.1.7}$$

It is often convenient to consider the set of all integers between two real numbers. Then *integer intervals* are sets of the form $I \cap \mathbb{Z}$ for some interval I in $\mathbb{R}$. In particular, the integer interval from a to b is the set

$$[\![a, b]\!] := [a, b] \cap \mathbb{Z} \tag{3.1.8}$$

We allow a and b to be any real numbers satisfying $a \leq b$, although they are chosen as integers in most applications.

As a 2-dimensional analogue, an *integer rectangle* is a set of the form $\mathsf{R} = [\![a, b]\!] \times [\![c, d]\!]$. If in addition a, b, c, d are integers, the *dimension* of R is the tuple $\dim(\mathsf{R}) := (c-a+1, d-b+1)$. If $\dim(\mathsf{R}) = (m, n)$, then we write $\operatorname{long}(\mathsf{R}) := \max\{m, n\}$ and $\operatorname{short}(\mathsf{R}) := \min\{m, n\}$.

As a special case, the *ℓ^∞-ball* of radius r centered at $x \in \mathbb{Z}^2$ is the set of the form

$$B(x, r) := x + [\![-r, r]\!]^2 = \{y \in \mathbb{Z}^2 : \|y - x\|_\infty \leq r\}, \tag{3.1.9}$$

where $\|\cdot\|_\infty$ is the ℓ^∞-distance. It is clear from the definition that $B(x, r)$ has dimension $(2\lfloor r\rfloor + 1, 2\lfloor r\rfloor + 1)$. Analogously, the r-neighborhood of a set $\Lambda \subseteq \mathbb{Z}^2$ is the set $B(\Lambda, r) := \cup_{x\in\Lambda} B(x, r)$.

For $x \in \mathbb{R}^2$ we write $x^{(1)}$ and $x^{(2)}$ for the coordinates of x, i.e., $x = (x^{(1)}, x^{(2)})$. For any two non-empty subsets A, B of $\mathbb{R}^2$, we define the ℓ^∞-distance between A and B by

$\mathrm{dist}_\infty(A, B) := \inf\{\|x - y\|_\infty : x \in A, y \in B\}$. For any non-empty subset A of $\mathbb{R}^2$, the diameter of A is $\mathrm{diam}_\infty(A) := \sup\{\|x - y\|_\infty : x, y \in A\}$.

If A is a non-empty subset of $\mathbb{R}^2$, then $\mathrm{left}(A) := \inf\{x \in \mathbb{R} : (x, y) \in A \text{ for some } y \in \mathbb{R}\}$ and likewise $\mathrm{bottom}(A) := \inf\{y \in \mathbb{R} : (x, y) \in A \text{ for some } x \in \mathbb{R}\}$.

If I is a non-empty integer interval, then the sequence $\gamma = (x_t)_{t \in I}$ in $\mathbb{Z}^2$ is called an ℓ^∞-path if $\|x_s - x_t\|_\infty = 1$ for each $s, t \in I$ with $|s - t| = 1$. Similarly, γ is called an ℓ^1-path if $\|x_s - x_t\|_1 = 1$ for each $s, t \in I$ with $|s - t| = 1$. In both cases, if in addition $I = [\![a, b]\!]$ is finite and $x_a = x_b$, then γ is called a loop.

3.2 Deterministic input

The proof consists of two parts; one deterministic, dealing with properties of general configurations, and the other random, where the probabilities of various important events are estimated. In this section, we discuss the structure of the boundary of an occupied cluster in any terminal configuration $\mathcal{B}^\infty\omega$. The upshot is that the boundary can be effectively described solely in terms of events involving the initial configuration ω. Naturally, this will make the probabilistic analysis of polluted bootstrap percolation considerably simpler.

Let us start with some definitions. We will conveniently identify $\mathbb{R}^2$ with the complex plane $\mathbb{C}$ and denote the imaginary unit by $\mathrm{i} = \sqrt{-1}$.

Definition 3.2.1 (Defect). *Let $U \subset \mathbb{Z}^2$. Then $x \in U$ is called a* defect *of U in $\omega \in \Omega$ if*

$$\sum_{y \in B(x,2) \cap U} \omega(y) \geq 3. \tag{3.2.1}$$

Alternatively, x is a defect of U in ω if either (i) there exist at least 3 sites of $B(x, 2) \cap U$ which are occupied in ω, or (ii) there exist at least two sites of $B(x, 2) \cap U$ with one polluted and the other non-vacant in ω.

Definition 3.2.2 (Blocking contour). *Let I be either a finite or an infinite interval in $\mathbb{Z}$*

with $|I| \geq 2$. *Then an* ℓ^1*-path* $\gamma = (x_t)_{t\in I}$ *is called a* left-blocking contour *in* $\omega \in \Omega$ *if the following conditions are satisfied:*

(LB1) γ *is* self-avoiding, *i.e.,* $x_s \neq s_t$ *for all distinct* $s, t \in I$.

(LB2) γ *is not occupied in* ω, *i.e.,* $\omega(x_t) \neq 1$ *for all* $t \in I$.

For each t *with* $[\![t-1, t+1]\!] \subseteq I$, *we write* $v_{t+} := x_{t+1} - x_t$ *and* $v_{t-} := x_t - x_{t-1}$. *If* $t = \sup I$ *so that only* v_{t-} *is defined, then set* $v_{t+} := v_{t-}$, *and likewise, if* $t = \inf I$ *then set* $v_{t-} := v_{t+}$. *Then the remaining conditions are:*

(LB3) If $v := v_{t+} = v_{t-}$, *then* $\omega(x_i - \mathrm{i}\, v) \neq 1$, *i.e., the site to the right of* x_t *is unoccupied.*

(LB4) If $v_{t+} = \mathrm{i}\, v_{t-}$, *then either* $\omega(x_t + v_{t+}) \neq 1$ *or* $\omega(x_t + v_{t-}) \neq 1$.

(LB5) If $v_{t+} = -\mathrm{i}\, v_{t-}$, *then at least one of* $\omega(x_t)$, $\omega(x_t - v_{t+})$, *or* $\omega(x_t - v_{t-})$ *is* 2.

If in addition $I = [\![a, a+n-1]\!]$ *is finite and* $x_a = x_{a+n-1}$, *then we call* γ *a* left-blocking loop *of length* n.

Regarding the question of identifying blocking contours in a given configuration, we will find that they essentially arise from percolation of a special kind of configurations called *blocking wedges*. We introduce another set of definitions.

Definition 3.2.3 (Wedges and zigzags)**.** *Let* $\omega \in \Omega$.

(1) An integer rectangle $\mathsf{R} = [\![a, b]\!] \times [\![c, d]\!]$ *is called a* double line *in* ω *if* $\mathrm{short}(\mathsf{R}) = 2$ *and no site of* R *is occupied in* ω. *Also,* R *is called* horizontal *(resp.* vertical*) if* $b - a \geq d - c$ *(resp.* $d - c \geq b - a$*). In the extreme case, a* 2×2 *square is thus both horizontal and vertical.*

(2) For $x \in \mathbb{Z}^2$ *and* $\mathsf{W} \subseteq \mathbb{Z}^2$, *the pair* (x, W) *is called a* northeast-wedge (NE-wedge) *in* ω *if the following conditions are met:*

- *There exist a horizontal double line* R_1 *and a vertical double line* R_2, *sharing the same bottom-left corner* (a, b), *such that* $\mathsf{W} = \mathsf{R}_1 \cup \mathsf{R}_2$
- x *is one of the sites* (a, b), $(a-1, b)$, $(a, b-1)$ *and it is polluted in* ω.

In this case, (a, b) *is called the bottom-left corner of* W. *Other types of wedges (south-*

west, southeast, northwest) are defined in a similar manner.

(3) A subset Z *of* $\mathbb{Z}^2$ *is called a* northeast-zigzag (NE-zigzag) *in* ω *if there exist a sequence* $\{(x_i, \mathsf{W}_i)\}_{i=1}^n$ *of NE-wedges such that* $\mathsf{Z} = \cup_{i=1}^n \mathsf{W}_i$ *and for each* $i = 1, \cdots, n-1$ *the following conditions hold:*

- *The bottom-left corner* c_{i+1} *of* W_{i+1} *lies to the lower right of the bottom-left corner* c_i *of* W_i*. That is,* $c_{i+1} - c_i \in (0, \infty) \times (-\infty, 0)$*.*
- $\mathrm{dist}_\infty(\mathsf{W}_i, \mathsf{W}_{i+1}) \leq 1$.

Roughly speaking, a NE-zigzag is a chain of adjacent NE-wedges that stretches to bottom-right direction. Zigzags for other directions (southwest, southeast, northwest) are defined using the corresponding wedges in a similar manner.

The following observation connects zigzags to blocking contours.

Lemma 3.2.4. *Let* γ *be a left-blocking contour in* $\omega \in \Omega$ *which starts moving leftward, finishes moving upward, and consists entirely of moves pointing leftwards or upwards. Then there exists a blocking NE-zigzag* Z *containing* γ*.*

Proof. Parametrize the left-blocking contour by $\gamma = (x_t)_{t=0}^T$. Then we can find an integer $K \geq 2$ and a sequence $0 = t_0 < t_1 < \cdots < t_K = T$ of times such that γ changes its "direction" exactly at times $t_1, \cdots, t_{K-1}$. The assumption on the initial and final moving directions forces K, the number of changes of direction plus one, to be even. Now for each $k = 1, \cdots, K/2$, we define

$$\begin{aligned} \mathsf{R}_{2k-1} &:= \big[\text{union of all sets } \{x_t, x_t + \mathrm{i}\} \text{ where } t \in [\![t_{2k-2}, t_{2k-1}]\!] \text{ and } \omega(x_t + \mathrm{i}) \neq 1\big], \\ \mathsf{R}_{2k} &:= \big[\text{union of all sets } \{x_t, x_t + 1\} \text{ where } t \in [\![t_{2k-1}, t_{2k}]\!] \text{ and } \omega(x_t + 1) \neq 1\big], \\ \mathsf{W}_k &:= \mathsf{R}_{2k-1} \cup \mathsf{R}_{2k}. \end{aligned} \tag{3.2.2}$$

Then by invoking (LB2)–(LB5), we can check that $\mathsf{W}_1, \cdots, \mathsf{W}_{K/2}$ are NE-blocking wedges such that $\gamma \subseteq Z := \bigcup_{k=1}^{K/2} \mathsf{W}_k$. Indeed, (LB3) ensures that R_{2k-1}'s are horizontal double lines and R_{2k} are vertical double lines with the common lower left corner $x_{t_{2k-1}}$. Also, (LB5) then guarantees that at least one of $x_{t_{2k-1}}$, $x_{t_{2k-1}} - 1$, or $x_{t_{2k-1}} - \mathrm{i}$ are polluted in ω, hence W_k is a NE-blocking wedge. Finally, (LB4) tells that $\mathrm{dist}_\infty(\mathsf{W}_k, \mathsf{W}_{k+1}) \leq 1$. □

Next we identify the geometry of the boundary of a terminally occupied cluster in terms of the initial configuration. The following proposition provides a basic tool for this purpose.

Proposition 3.2.5. *Let $\omega \in \Omega$. Suppose that $\mathcal{C}$ is a bounded, occupied cluster in the terminal configuration $\mathcal{B}^\infty\omega$. Then there exists an ℓ^∞-loop $\gamma = (x_t)_{t=0}^n$ of length $n \geq 4$ such that the following is true:*

(1) γ encloses $\mathcal{C}$, i.e., γ lies in the unique unbounded connected component of $\mathbb{Z}^2 \setminus \mathcal{C}$.

(2) For any integers $s < t$, write $\widehat{\gamma} = (x_{u \bmod n})_{u \in [s,t]}$. Then either $\widehat{\gamma}$ contains a defect of $\mathbb{Z}^2 \setminus \mathcal{C}$ in ω, or $\widehat{\gamma}$ is a blocking path in ω.

Proof. Recall that the *dual lattice* of $\mathbb{Z}^2$ is defined as the graph $(\mathbb{Z}^2)^* := (\frac{1}{2}, \frac{1}{2}) + \mathbb{Z}^2$, where each edge $e = (x, y)$ of $\mathbb{Z}^2$ is identified with its dual edge $e^* = (\frac{1+\mathrm{i}}{2}x + \frac{1-\mathrm{i}}{2}y, \frac{1-\mathrm{i}}{2}x + \frac{1+\mathrm{i}}{2}y)$. Associating with each site x of $\mathbb{Z}^2$ a unit square centered at x, dual-edges may be viewed as sides of such squares.

If $\Delta\mathcal{C}$ denotes the edge-boundary of $\mathcal{C}$, then $\{e^* : e \in \Delta\mathcal{C}\}$ consists of dual loops in $(\mathbb{Z}^2)^*$. Let γ_0^* denote the outermost dual loop oriented in the counter-clockwise direction. We are going to modify γ_0^* through the process similar to loop-erasure. To this end, write $\gamma_0^* = (x_t^*)_{t=0}^T$ and extend this to all of $\mathbb{Z}$ by the periodicity $x_{t+T}^* = x_t^*$. Suppose that $s < t$ are two integers such that

- $\|x_t^* - x_s^*\|_1 = 1$,
- $(x_s^*, x_{s+1}^*, \cdots, x_{t-1}^*, x_t^*, x_s^*)$ encloses a non-empty bounded subset of $\mathbb{Z}^2 \setminus \mathcal{C}$,
- $e_{st}^* := (x_s^*, x_t^*)$ a side of the square centered at a polluted vertex.

Then we delete the segment $(x_{s+1}^*, \cdots, x_{t-1}^*)$ from γ_0^* and join x_s^* directly to x_t^*. A moment's thought reveals that the above condition is simply that the local configuration around e_{st}^* is one of the configurations listed in Figure 3.1 modulo rotations. Since the length of γ_0^* is finite, after a finitely many repeated application of this procedure, we obtain a dual loop γ^*

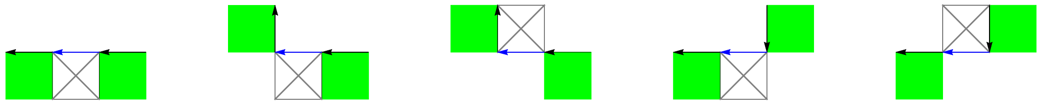


Figure 3.1: Possible local configuration around e^*_{st}.

that cannot be shortened any further. Now let $\mathcal{C}' \subset \mathbb{Z}^2$ be a bounded set having γ^* as its edge-boundary, and let γ_0 be the ℓ^∞-loop arising from tracing the outer boundary $\partial\mathcal{C}'$ in the counter-clockwise direction. Finally, γ is obtained from γ_0 by replacing each "diagonal left-turn" of γ_0 by a "two-step left-turn" as long as the resulting path does not have occupied sites. We claim that $\gamma = (x_t)_{t=0}^T$ satisfies the desired properties. Indeed, the construction ensures that γ is an ℓ^∞-loop enclosing $\mathcal{C}$. Next, assume that $\widehat{\gamma} = (x_u)_{u\in[s,t]}$ contains no defect of $\mathbb{Z}^2 \setminus \mathcal{C}$ in ω. Then each $x \in \widehat{\gamma}$ falls into one of the following classes

(1) $|\mathcal{N}(x) \cap \mathcal{C}| = 1$.

(2) $\omega(x) = 2$ and $\omega(y) = 0$ for all $y \in B(x, 2) \setminus (\mathcal{C} \cup \{x\})$.

(3) One of the four sides of the unit square centered at x arises from loop erasing.

These considerations allow to verify that all the properties (LB1)–(LB5) are satisfied for $\widehat{\gamma}$, and so, the desired conclusion follows. □

3.3 Coupling of blocking structures

In this section we introduce a version of continuum percolation which captures the blocking structure in the limit. Then we relate the connectivity structure of blocking wedges to that of the continuum percolation.

3.3.1 Percolation models

We begin by describing the continuum percolation model of interest. For each $\lambda \geq 0$, denote by $\xi_{\lambda,\infty}$ a sample of the Poisson point process on $\mathbb{R}^2 \times [0, \infty)^6$ with intensity measure $\lambda\,\mathrm{Leb} \otimes \mathrm{Exponential}(1)^{\otimes 6}$. This is equivalent to sampling from the Poisson point process

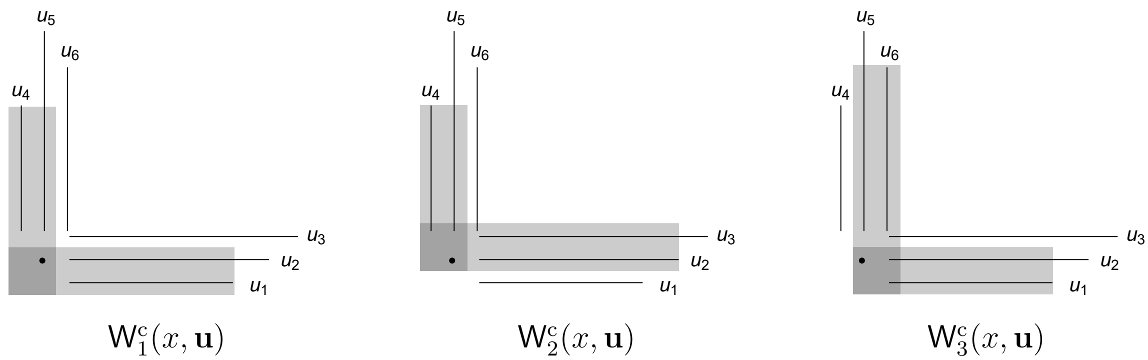


Figure 3.2: Illustrations demonstrating how the wedges $\mathsf{W}_i^{\mathrm{c}}$ are constructed.

on $\mathbb{R}^2$ with intensity λ and decorating each point with six independent Exponential(1) random variables. For the sake of future use, for each $R \in (0, \infty)$ we also consider the truncated version

$$\xi_{\lambda,R} := \{(x, (u_i)_{i=1}^6) \in \xi_{\lambda,\infty} : u_i \leq R \text{ for all } i\}. \tag{3.3.1}$$

The process $\xi_{\lambda,R}$ thus defined is not the actual geometric object that we are going to work with. Rather, it is the *parameter family* for building such an object. The conversion is described as follows: With each parameter $(x, \mathbf{u} = (u_i)_{i=1}^6) \in \mathbb{R}^2 \times [0, \infty)^6$, we associate the sets

$$\begin{aligned} \mathsf{W}_1^{\mathrm{c}}(x, \mathbf{u}) &:= x + ([0, u_1 \wedge u_2] \times \{0\}) \cup (\{0\} \times [0, u_4 \wedge u_5]), \\ \mathsf{W}_2^{\mathrm{c}}(x, \mathbf{u}) &:= x + ([0, u_2 \wedge u_3] \times \{0\}) \cup (\{0\} \times [0, u_4 \wedge u_5]), \\ \mathsf{W}_3^{\mathrm{c}}(x, \mathbf{u}) &:= x + ([0, u_1 \wedge u_2] \times \{0\}) \cup (\{0\} \times [0, u_5 \wedge u_6]). \end{aligned} \tag{3.3.2}$$

The seemingly arbitrary choices of the various pairs from $u_1, \cdots, u_6$ will be explained later through a scaling limit of the wedges.

Now we equip $\xi_{\lambda,R}$ with a percolation structure. This amounts to designating a notion of connectivity on $\xi_{\lambda,R}$. We will do this by introducing a notion of connectivity for collections of general planar grains. Consider the set $\Pi^{\mathrm{c}} := \mathbb{R}^2 \times \mathcal{P}(\mathbb{R}^2)$ of all pairs of the form (x, V) where $x \in \mathbb{R}^2$ and $\mathsf{V} \subseteq \mathbb{R}^2$. We turn Π^{c} into a hypergraph by declaring that a subset $H \subseteq \Pi^{\mathrm{c}}$

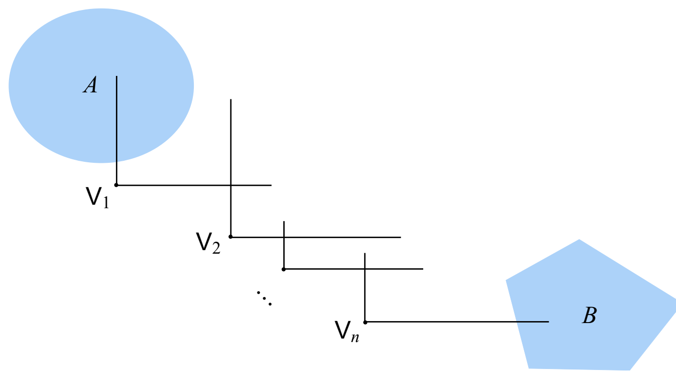


Figure 3.3: Connectivity in the continuum percolation

is a hyperedge of Π^{c} if H is finite and can be enumerated as $H = \{(x_k, \mathsf{V}_k)\}_{k=1}^n$ so that all of

$$\begin{gathered} x_{k+1} \neq x_k, \qquad \mathsf{V}_k \cap \mathsf{V}_{k+1} \neq \varnothing, \\ \mathrm{left}(\mathsf{V}_{k+1}) \geq \mathrm{left}(\mathsf{V}_k), \qquad \text{and} \qquad \mathrm{bottom}(\mathsf{V}_{k+1}) \leq \mathrm{bottom}(\mathsf{V}_k). \end{gathered} \tag{3.3.3}$$

hold true for each $k = 1, \cdots, n-1$. Then the connectivity on each parameter family $\zeta \subseteq \mathbb{R}^2 \times [0, \infty)^6$ is defined as the induced hyper-subgraph

$$G(\zeta) := \{(x, \mathsf{W}_i^{\mathrm{c}}(x, \mathbf{u})) : \text{ for some } (x, \mathbf{u}) \in \zeta \text{ and } i \in \{1, 2, 3\}\} \tag{3.3.4}$$

of Π. In other words, hyperedges in $G(\zeta)$ are sequences of the form $\{(x_k, \mathsf{V}_k = \mathsf{W}_{i_k}^{\mathrm{c}}(x_k, \mathbf{u}_k))\}_{k=1}^n$ for some $(x_k, \mathbf{u}_k) \in \zeta$ and $i_k \in \{1, 2, 3\}$ such that (3.3.3) holds true for each $k < n$. Then two subsets A and B of $\mathbb{R}^2$ are said to be *connected in* ζ and denoted by $A \overset{\zeta}{\longleftrightarrow} B$, if there exists a hyperedge $\{(x_k, \mathsf{V}_k)\}_{k=1}^n$ in $G(\zeta)$ such that $\cup_{k=1}^n \mathsf{V}_k$ intersects both A and B. If in addition Λ is a subset of $\mathbb{R}^2$, then we write $A \xleftrightarrow{\zeta \text{ on } \Lambda} B$ if A and B are connected in $\zeta \cap (\Lambda \times [0, \infty)^6)$.

The connectivity in the discrete counterpart can be defined similarly. Let $\Pi^{\mathrm{d}} = \mathbb{Z}^2 \times \mathcal{P}(\mathbb{Z}^2)$ be the set of all pairs of the form (x, V) where $x \in \mathbb{Z}^2$ and $\mathsf{V} \subseteq \mathbb{Z}^2$. Similarly as before, we turn Π^{d} into a hypergraph by declaring that a subset $H \subseteq \Pi^{\mathrm{d}}$ is a hyperedge of Π^{d} if and only if H is a finite set with an enumeration $H = \{(x_k, \mathsf{V}_k)\}_{k=1}^n$ so that all of

$$\begin{gathered} x_{k+1} \neq x_k, \qquad \mathrm{dist}_\infty(\mathsf{V}_k, \mathsf{V}_{k+1}) \leq 1, \\ \mathrm{left}(\mathsf{V}_{k+1}) \geq \mathrm{left}(\mathsf{V}_k), \qquad \text{and} \qquad \mathrm{bottom}(\mathsf{V}_{k+1}) \leq \mathrm{bottom}(\mathsf{V}_k). \end{gathered} \tag{3.3.5}$$

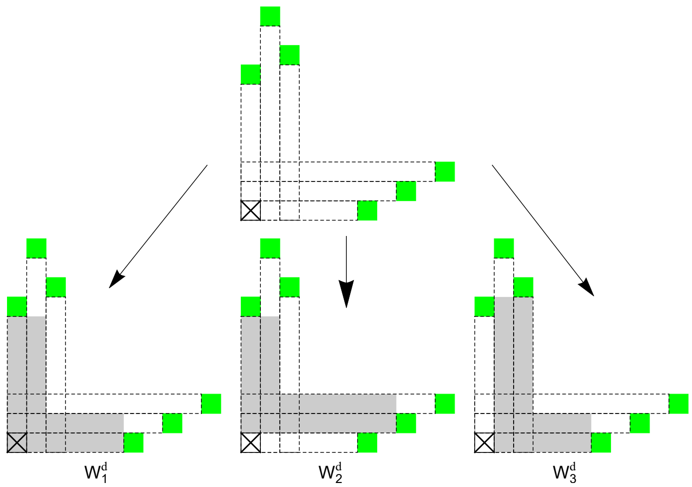


Figure 3.4: Maximal wedges discoverable around a given polluted point.

hold true for each $k < n$. Notice that $\mathrm{dist}_\infty(A, B) \leq 1$ holds for subsets A and B of $\mathbb{Z}^2$ if and only if either A and B intersect or A is "in close contact" with B. Then, given a configuration $\omega \in \Omega$ and a set $\Lambda \subseteq \mathbb{Z}^2$, define

$$G(\omega, \Lambda) := \{(x, \mathsf{W}) : \text{maximal NE-wedge in } \omega \text{ with } x \in \Lambda \text{ and } \mathsf{W} \subseteq \Lambda\}. \tag{3.3.6}$$

Here, the maximality means that there is no NE-wedge (x, W') in ω with $\mathsf{W} \subsetneq \mathsf{W}' \subseteq \Lambda$. As an induced hyper-subgraph of Π^{d}, $G(\omega, \Lambda)$ is also a hypergraph. We remark that the union of all grains in a hyperedge of $G(\omega, \Lambda)$ is a zigzag in ω. Then two subsets A and B of $\mathbb{Z}^2$ are said to be *connected by* $G(\omega, \Lambda)$ and denoted by $A \xleftrightarrow{\omega \text{ on } \Lambda} B$, if there exists a hyperedge $\{(x_k, \mathsf{V}_k)\}_{k=1}^n$ in $G(\omega, \Lambda)$ such that the zigzag $\mathsf{Z} = \cup_{k=1}^n \mathsf{V}_k$ satisfies $\mathrm{dist}_\infty(\mathsf{Z}, A) \leq 1$ and $\mathrm{dist}_\infty(\mathsf{Z}, B) \leq 1$. Also, if $\Lambda = \mathbb{Z}^2$, then we suppress Λ from the notation.

Finally, in order to systematically describe the structure of maximal NE-wedges that can be discovered around each polluted site, for each $x \in \mathbb{Z}^2$ and $\mathbf{u} = (u_i)_{i=1}^6 \in ([0, \infty) \cap \mathbb{Z})^6$, we define

$$\begin{aligned}
\mathsf{W}^{\mathrm{d}}_1(x, \mathbf{u}) &:= x + ([\![0, u_1 \wedge u_2]\!] \times [\![0, 1]\!]) \cup ([\![0, 1]\!] \times [\![0, u_4 \wedge u_5]\!]), \\
\mathsf{W}^{\mathrm{d}}_2(x, \mathbf{u}) &:= x + ([\![0, u_2 \wedge u_3]\!] \times [\![1, 2]\!]) \cup ([\![0, 1]\!] \times [\![1, u_4 \wedge u_5]\!]), \\
\mathsf{W}^{\mathrm{d}}_3(x, \mathbf{u}) &:= x + ([\![1, u_1 \wedge u_2]\!] \times [\![0, 1]\!]) \cup ([\![1, 2]\!] \times [\![0, u_5 \wedge u_6]\!]).
\end{aligned} \tag{3.3.7}$$

Figure 3.4 provides some motivation for these definitions.

3.3.2 Construction of the coupling

We will now move on to discussing the coupling between the discrete and continuum percolation model. Let $q := \lambda p^2$ and introduce the following auxiliary quantities

$$\lambda_* = \lambda_*(p) := -\frac{\log(1-q)}{p^2} \qquad \text{and} \qquad \mu = \mu(p) := -\frac{1}{p}\log\left(1 - \frac{p}{1-q}\right). \tag{3.3.8}$$

A computation shows that both $\lambda_*(\cdot)$ and $\mu(\cdot)$ are increasing near 0 and

$$\lim_{p\downarrow 0} \lambda_*(p) = \lambda \qquad \text{and} \qquad \lim_{p\downarrow 0} \mu(p) = 1. \tag{3.3.9}$$

Now let $\mathbf{P}$ be a probability law under which $\widetilde{\xi}_p$ is a Poisson point process on $\mathbb{R}^2 \times [0,\infty)^6$ with intensity measure $\lambda_* \,\mathrm{Leb} \otimes \mathrm{Exponential}(\mu)^{\otimes 6}$ and $\widetilde{\omega}_0$ is a random configuration, independent of $\widetilde{\xi}_p$, such that $\widetilde{\omega}_0(x)$ for each $x \in \mathbb{Z}^2$ is a Bernoulli random variable with parameter $p/(1-q)$ independently of all the others. We enumerate points of $\widetilde{\xi}_p$ in a measurable way, for instance, in the increasing order of the distance of their "spatial component" to the origin, to write $\widetilde{\xi}_p = \{(X_k, (U_{k,i})_{i=1}^6)\}_{k=1}^\infty$ almost surely. Then starting from $\widetilde{\omega}_0$, we recursively define sequences of configurations $(\widetilde{\omega}_{k,0})_{k\in\mathbb{N}_1}$ and $(\widetilde{\omega}_k)_{k\in\mathbb{N}_1}$ as follows:

$$\widetilde{\omega}_{k,0}(x) := \begin{cases} 0, & \text{if } x \in \lfloor \frac{1}{p} X_k \rfloor + [\![0, \lfloor \frac{1}{p} U_{k,i} \rfloor - 1]\!] \times \{i-1\} \text{ for some } i \in \{1,2,3\}, \\ 1, & \text{if } x = \lfloor \frac{1}{p} X_k \rfloor + (\lfloor \frac{1}{p} U_{k,i} \rfloor, i-1) \text{ for some } i \in \{1,2,3\}, \\ \widetilde{\omega}_{k-1}(x), & \text{otherwise.} \end{cases} \tag{3.3.10}$$

and

$$\widetilde{\omega}_{k}(x) := \begin{cases} 0, & \text{if } x \in \lfloor \frac{1}{p} X_k \rfloor + \{i-4\} \times [\![0, \lfloor \frac{1}{p} U_{k,i} \rfloor - 1]\!] \text{ for some } i \in \{4,5,6\}, \\ 1, & \text{if } x = \lfloor \frac{1}{p} X_k \rfloor + (i-4, \lfloor \frac{1}{p} U_{k,i} \rfloor) \text{ for some } i \in \{4,5,6\}, \\ \widetilde{\omega}_{k,0}(x), & \text{otherwise.} \end{cases} \tag{3.3.11}$$

Here, interpreting the "points" of $\widetilde{\xi}_p$ as wedges, $\widetilde{\omega}_{k,0}$ arises by updating the configuration $\widetilde{\omega}_{k-1}$ at x to vacant if, after suitable discretization, it falls onto the horizontal arm of the k-th wedge and to occupied if it lands on the tip thereof. Vertical arms of the k-th wedge are examined instead for the definition of $\widetilde{\omega}_k$. Although some blocking wedges constructed at a given site may be set multiple times, the probability that this happens will be shown to be negligible as $p \downarrow 0$.

Finally, define the configuration $\widetilde{\omega}$ by

$$\widetilde{\omega}(x) := \begin{cases} 2, & \text{if } x = \lfloor \frac{1}{p} X_k \rfloor \text{ for some } k, \\ \lim_{k\to\infty} \widetilde{\omega}_k(x), & \text{otherwise.} \end{cases} \tag{3.3.12}$$

We show that $\widetilde{\omega}$ is indeed well-defined and satisfies:

Lemma 3.3.1. *Let* $\mathbf{P}$ *be the coupling of* $\widetilde{\xi}_p$ *and* $\widetilde{\omega}_0$ *described as above. Then* $\widetilde{\omega}$ *under* $\mathbf{P}$*, defined by* (3.3.12)*, has the same law as* ω *under* $\mathbb{P}_{p,q}$*.*

Proof. That $\widetilde{\omega}$ is well-defined is part of the proof. Noting that $\lfloor \frac{1}{p} U_{k,i} \rfloor$ are independent random variables with geometric distribution started at 0 with parameter $1-\mathrm{e}^{-p\mu} = p/(1-q)$, each $\widetilde{\omega}_k$ has the same product Bernoulli distribution as $\widetilde{\omega}_0$ by the Markov property. Moreover, for each site $x \in \mathbb{Z}^2$ the expected number of updates occurring at x is bounded by

$$\begin{aligned} &\mathbf{E}\left[\sum_{k\geq 1}\left(\sum_{i=1}^{3} \mathbf{1}_{\left\{x\in\lfloor\frac{1}{p}X_k\rfloor+[0,\lfloor\frac{1}{p}U_{k,i}\rfloor]\times\{i-1\}\right\}}\right) + \sum_{k\geq 1}\left(\sum_{i=4}^{6} \mathbf{1}_{\left\{x\in\lfloor\frac{1}{p}X_k\rfloor+\{i-4\}\times[0,\lfloor\frac{1}{p}U_{k,i}\rfloor]\right\}}\right)\right] \\ &\quad = 6\sum_{l\geq 0}\mathbf{E}\left[\sum_{k\geq 1}\mathbf{1}_{\left\{\lfloor\frac{1}{p}X_k\rfloor=x-(l,0)\right\}\cap\left\{\lfloor\frac{1}{p}U_{k,1}\rfloor\geq l\right\}}\right] \\ &\quad = 6\sum_{l\geq 0} p^2\lambda_* \cdot \left(1-\frac{p}{1-q}\right)^l = 6p(1-q)\lambda_* < \infty. \end{aligned} \tag{3.3.13}$$

This shows that the pointwise limit of $\widetilde{\omega}_k$ as $k \to \infty$ exists almost surely and has the same product Bernoulli law as $\widetilde{\omega}_0$. Also, $\widetilde{\omega}(x)$'s for different $x \in \mathbb{Z}^2$'s are mutually independent and satisfy

$$\mathbf{P}(\widetilde{\omega}(x) = 2) = \mathbf{P}\left(\widetilde{\xi}_p \cap (px + [0,p)^2) \times [0,\infty)^6 \neq \varnothing\right) = 1 - \mathrm{e}^{-p^2\lambda_*} = q \tag{3.3.14}$$

and thus $\mathbf{P}(\widetilde{\omega}(x) = 1) = p$ and $\mathbf{P}(\widetilde{\omega}(x) = 0) = 1 - p - q$. □

Now we fix the scale parameters

$$\alpha \in (0, \tfrac{1}{5}), \qquad \beta \in [3, \infty), \qquad b \in (0, \infty). \tag{3.3.15}$$

Set $L := \lfloor bp^{-1-\alpha} \rfloor$ and $\Lambda := B(0, L)$. We are going to check that the wedges in $G(\widetilde{\omega}, \Lambda)$ arise exactly from the wedges in $G(\widetilde{\xi}_p)$ with high probability. To this end, we introduce events which encode various exceptional cases. Define events $\mathcal{E}_1$ and $\mathcal{E}_2$ by

$$\begin{aligned} \mathcal{E}_1 &:= \{\exists x \in \Lambda \text{ s.t. } \widetilde{\omega}(x) = 2 \text{ and } x + [\![0,2]\!]^2 \not\subseteq \Lambda\}, \\ \mathcal{E}_2 &:= \{\exists x \in \Lambda \text{ s.t. } \widetilde{\omega}(x) = 2 \text{ and } \exists y \in (x + [\![0,2]\!]^2) \cap \Lambda \text{ s.t. } \widetilde{\omega}(y) = 1\}. \end{aligned} \tag{3.3.16}$$

These two events are designed to control the situation where a polluted site x in Λ is too close either to the boundary of Λ or to some initially occupied sites, this would prevent wedges to form at x. There is another scenario that we want to avoid, where some maximal wedges in the local configuration $\widetilde{\omega}_k|_\Lambda$ are destroyed during the recursive definition (3.3.10)–(3.3.11) when some of their arms become either occupied or polluted. To describe this scenario, define

$$\begin{aligned} C_k &:= \lfloor \tfrac{1}{p} X_k \rfloor + ([\![-\beta,\beta]\!] \times \mathbb{Z}) \cup (\mathbb{Z} \times [\![\beta,\beta]\!]), \\ D_k &:= \lfloor \tfrac{1}{p} X_k \rfloor + \{\mathbf{0}\} \cup \Big\{ \big(\lfloor \tfrac{1}{p} U_{k,i+1} \rfloor, i\big), \big(i, \lfloor \tfrac{1}{p} U_{k,i+4} \rfloor\big) : i = 0,1,2 \Big\}. \end{aligned} \tag{3.3.17}$$

The set C_k encodes the cross of width $2\beta + 1$ around the point $\lfloor \frac{1}{p} X_k \rfloor$ and D_k consists of the polluted site $\lfloor \frac{1}{p} X_k \rfloor$ along with the occupied endpoints at the six Geometric(p) "arms" defining the wedges in (3.3.7). Then set

$$\mathcal{E}_3 := \{\exists k \neq l \text{ s.t. } \lfloor \tfrac{1}{p} X_k \rfloor \in \Lambda,\ \lfloor \tfrac{1}{p} X_l \rfloor \in \Lambda, \text{ and } D_k \cap C_l \neq \varnothing\} \tag{3.3.18}$$

for the event that these two objects intersect for two distinct polluted sites. Finally, we define

$$\mathcal{E}_4 := \{\exists k \text{ s.t. } \lfloor \tfrac{1}{p} X_k \rfloor \in \Lambda \text{ and } U_{k,i} \leq \beta p \text{ for some } i\} \tag{3.3.19}$$

to capture the situations where some of $U_{k,i}$'s are too short. The following lemma shows that all these events are indeed negligible in the limit $p \downarrow 0$.

Lemma 3.3.2. *Fix $\lambda \in (0,\infty)$ and assume* (3.3.15). *Let $\mathcal{E}_1, \cdots, \mathcal{E}_4$ be as in* (3.3.16), (3.3.18)*, and* (3.3.19)*. Then there exists a constant $C = C(\lambda,\alpha,\beta,b) \in (0,\infty)$ such that*

$$\mathbf{P}(\mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3 \cup \mathcal{E}_4) \leq C p^{1-5\alpha} \tag{3.3.20}$$

holds true for all sufficiently small p.

Proof. As before, we write $L := \lfloor bp^{-1-\alpha} \rfloor$ so that $\Lambda = [\![-L, L]\!]^2$. We begin by handling the case of $\mathcal{E}_1$. If $x \in \Lambda$ satisfy $x + [\![0,2]\!]^2 \nsubseteq \Lambda$, then it must lie either in $\mathsf{R}_1 := [\![-L, L]\!] \times [\![L-1, L]\!]$ or in $\mathsf{R}_2 := [\![L-1, L]\!] \times [\![-L, L]\!]$. Noting that $|\mathsf{R}_1 \cup \mathsf{R}_2| = 8L$, we get

$$\begin{aligned} \mathbf{P}(\mathcal{E}_1) &= 1 - \mathbf{P}(\widetilde{\omega}(x) \neq 2 \text{ for all } x \in \mathsf{R}_1 \cup \mathsf{R}_2) \\ &= 1 - (1-q)^{8L} \leq \frac{8Lq}{1-q} = \mathcal{O}(p^{1-\alpha}) \end{aligned} \tag{3.3.21}$$

where the inequality $1 - (1-x)^y \leq \frac{yx}{1-x}$ which holds true for all $x \in [0,1)$ and $y \geq 0$ is used in the intermediate steps. Next, taking union bound and noting that $|\Lambda| = (2L+1)^2$ and $|x + [\![0,2]\!]^2| = 9$,

$$\mathbf{P}(\mathcal{E}_2) \leq (2L+1)^2 q(1-(1-p)^8) = \mathcal{O}(p^{1-2\alpha}). \tag{3.3.22}$$

For $\mathcal{E}_3$, note that this event depends only on $\widetilde{\xi}_p$. Let N be the number of Poisson points $(x, \mathbf{u}) \in \widetilde{\xi}_p$ such that $x \in [-pL, p(L+1))^2$. (This is equivalent to saying that $\lfloor \frac{1}{p} x \rfloor \in \Lambda$.) Then we may re-enumerate $\widetilde{\xi}_p = \{(X_k, (U_{k,i})_{i=1}^6)\}_{k=1}^\infty$ via a random bijection $\sigma : \mathbb{N}_1 \to \mathbb{N}_1$ such that, for each $n \in \mathbb{N}_1$, $(X_{\sigma(k)})_{k=1}^n$ given $\{N = n\}$ are independent and uniformly distributed on $[-pL, p(L+1))^2$. Also, partitioning the event $\mathcal{E}_3$ according to the value of N and the smallest value of k for which the condition in the definition (3.3.18) of $\mathcal{E}_3$ holds true, union bound yields

$$\begin{aligned} \mathbf{P}(\mathcal{E}_3) &\leq \sum_{n=1}^\infty \mathbf{P}(N=n) \sum_{\substack{k,l \in [1,n] \\ k \neq l}} \mathbf{P}\left(D_{\sigma(k)} \cap C_{\sigma(j)} \neq \varnothing \,\middle|\, N = n \right) \\ &= \sum_{n=1}^\infty n(n-1) \mathbf{P}(N=n) \mathbf{P}\left(D_{\sigma(1)} \cap C_{\sigma(2)} \neq \varnothing \,\middle|\, N = n \right). \end{aligned} \tag{3.3.23}$$

Using the symmetry, we may bound the conditional probability in the last line of (3.3.23) by

$$\begin{aligned} &\mathbf{P}\left(\lfloor \tfrac{1}{p} X_{\sigma(1)} \rfloor \in C_{\sigma(2)} \,\middle|\, N = n \right) \\ &\quad + 6\mathbf{P}\left(\lfloor \tfrac{1}{p} X_{\sigma(1)} \rfloor \notin C_{\sigma(2)} \text{ but } \lfloor \tfrac{1}{p} X_{\sigma(1)} \rfloor + \left(\lfloor \tfrac{1}{p} U_{\sigma(1),1} \rfloor, 0 \right) \in C_{\sigma(2)} \,\middle|\, N = n \right) \end{aligned} \tag{3.3.24}$$

Noting that $\text{Exponential}(\mu) = \mu e^{-\mu x} \mathbf{1}_{[0,\infty)}(x) \, \mathrm{d}x$ is dominated by $\mu \, \mathrm{d}x$, (3.3.24) is further bounded from above by $\mathcal{O}(L^{-1})$. Plugging this back to (3.3.23) and invoking the fact that

N has the Poisson distribution with rate $\lambda_* p^2(2L+1)^2 = \mathcal{O}(p^{-2\alpha})$,

$$\mathbf{P}(\mathcal{E}_2) = \mathcal{O}(L^{-1}\mathbf{E}[N^3]) = \mathcal{O}\left(p^{1-5\alpha}\right). \tag{3.3.25}$$

Finally, using the same setting and invoking (3.3.9),

$$\mathbf{P}(\mathcal{E}_4) \leq 6\mathbf{E}[N^2]\mathbf{P}\left(U_{\sigma(1),1} \leq \beta p\right) = \mathcal{O}(p^{1-4\alpha}). \tag{3.3.26}$$

Combining (3.3.21)–(3.3.26) altogether, we get (3.3.20) as desired. □

3.4 Crossing probabilities in the subcritical regime

We are now in the position to prove that

Proposition 3.4.1. *Let $\lambda < \lambda_1^{\mathrm{c}}$ and $\alpha \in (0, \frac{1}{5})$. Then for each $0 < a < b < \infty$, we get*

$$\lim_{p \downarrow 0} \mathbb{P}_{p,\lambda p^2}\left(B(0, ap^{-1-\alpha}) \overset{\omega}{\longleftrightarrow} \mathbb{Z}^2 \setminus B(0, bp^{-1-\alpha})\right) = 0. \tag{3.4.1}$$

The proof is based on domination of the probability in the statement by that for a similar event in the continuum model. To make this comparison easier, we will actually consider the continuum model with parameter $\lambda' \in (\lambda, \lambda_1^{\mathrm{c}})$ and then apply a scaling argument.

Proof. Fix $\lambda' \in (\lambda, \lambda_1^{\mathrm{c}})$ and set

$$r = r(p) := \sqrt{\lambda_*(p)/\lambda'}. \tag{3.4.2}$$

Then by comparing the intensity measures, we find that

$$\widetilde{\xi}_p \overset{\text{law}}{=} \left\{\left(\tfrac{1}{r}x, \left(\tfrac{1}{\mu}u_i\right)_{i=1}^6\right) : \left(x, (u_i)_{i=1}^6\right) \in \xi_{\lambda',\infty}\right\}. \tag{3.4.3}$$

In light of this, we may realize (3.4.3) as a pointwise equality. Also, let

$$\beta = \frac{1}{\sqrt{\lambda'/\lambda} - 1} + 3. \tag{3.4.4}$$

In order to understand the reason behind the choice (3.4.4), note that

$$\left\lfloor \tfrac{1}{rp}x \right\rfloor + \left\lfloor \tfrac{1}{\mu p}u \right\rfloor \geq \left\lfloor \tfrac{1}{rp}y \right\rfloor \text{ and } u \geq \frac{\mu r p}{\mu - r} \quad \Rightarrow \quad u \geq y - x \tag{3.4.5}$$

holds. Then β in (3.4.4) is chosen so that $\beta > \frac{\mu r}{\mu - r}$ holds for all sufficiently small p.

Let $L = \lfloor bp^{-1-\alpha} \rfloor$ and $\Lambda = [\![-L, L]\!]^2$ be as in the proof of Lemma 3.3.2. Now write $\{(\widetilde{X}_k, (\widetilde{U}_{k,i})_{i=1}^6)\}_{k=1}^N$ for an enumeration of the point process

$$\eta := \xi_{\lambda',\infty} \cap \left([-prL, pr(L+1))^2 \times [0,\infty)^6 \right). \tag{3.4.6}$$

Then, on the complement of $\mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3 \cup \mathcal{E}_4$, the correspondence

$$\mathsf{W}_i^{\mathrm{c}}\big(\widetilde{X}_k, (\widetilde{U}_{k,i})_{i=1}^6\big) \quad \longleftrightarrow \quad \mathsf{W}_i^{\mathrm{d}}\Big(\left\lfloor \tfrac{1}{pr}\widetilde{X}_k \right\rfloor, \big(\left\lfloor \tfrac{1}{p\mu}\widetilde{U}_{k,i} \right\rfloor - 1\big)_{i=1}^6\Big) \cap \Lambda \tag{3.4.7}$$

gives rise to a bijection between the hypergraphs $G(\eta)$ and $G(\widetilde{\omega}, \Lambda)$. (Recall that $G(\widetilde{\omega}, \Lambda)$ is the hypergraph of all maximal NE-wedges in the local configuration $\widetilde{\omega}|_\Lambda$.) Then (3.4.5) ensures that each hyperedge of $G(\widetilde{\omega}, \Lambda)$ also lies in $G(\eta)$. In particular, still assuming that $\cap_{i=1}^4 \mathcal{E}_i^c$ holds, if $B(0, ap^{-1-\alpha})$ and $\mathbb{Z}^2 \setminus B(0, bp^{-1-\alpha})$ are joined by a hyperedge of $G(\widetilde{\omega}, \Lambda)$, then it follows that $[-arp^{-\alpha}, arp^{-\alpha}]^2$ and $\mathbb{R}^2 \setminus [-prL, pr(L+1))^2$ are joined by a hyperedge of $G(\eta)$. From this, we obtain

$$\begin{aligned} &\mathbb{P}_{p,\lambda p^2}\Big(B(0, ap^{-1-\alpha}) \stackrel{\omega}{\longleftrightarrow} \mathbb{Z}^2 \setminus B(0, bp^{-1-\alpha})\Big) \\ &\quad \leq \mathbb{P}\Big([-arp^{-\alpha}, arp^{-\alpha}]^2 \stackrel{\xi_{\lambda',\infty}}{\longleftrightarrow} \mathbb{R}^2 \setminus [-prL, pr(L+1))^2\Big) + \mathbf{P}(\mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3 \cup \mathcal{E}_4) \end{aligned} \tag{3.4.8}$$

Invoking the union bound, the first term on the right-hand side of (3.4.8) can be bounded from above by

$$(2arp^{-\alpha})^2 \mathbb{P}\Big([0,1]^2 \stackrel{\xi_{\lambda',\infty}}{\longleftrightarrow} \mathbb{R}^2 \setminus (-\tfrac{b-a}{3}p^{-\alpha}, \tfrac{b-a}{3}p^{-\alpha})^2\Big), \tag{3.4.9}$$

which decays at least exponentially fast in $p^{-\alpha}$ by the assumption $\lambda' < \lambda_1^{\mathrm{c}}$. Moreover, by Lemma 3.3.2, the second term on the right-hand side of (3.4.8) is bounded by $Cp^{1-3\alpha}$. Therefore the desired claim follows. $\square$

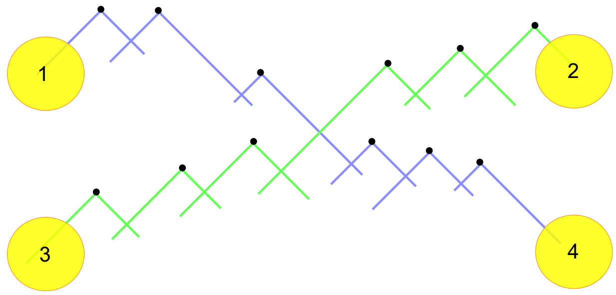


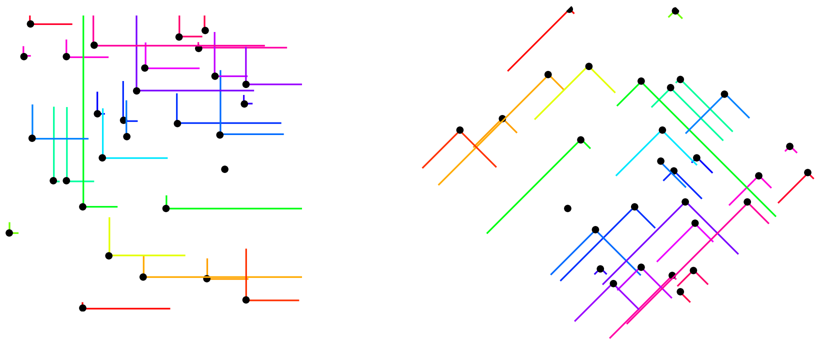
Figure 3.5: Intersecting zigzags.

3.5 Blocking contours in the supercritical regime

The aim of this section is to supply the necessary ingredients for building blocking loops in the supercritical regime of the continuum percolation model. Two main obstructions are standing in the way.

First, the coupling scheme described in Subsection 3.3.2 only allows to safely translate the connectivity of the continuum percolation model into that of the blocking wedges in bootstrap percolation in a box of scale $p^{-6/5+\epsilon}$. Although the exponent $\frac{6}{5}$ is not necessarily optimal, any attempts to find a coupling that preserves the connectivity of wedges in both worlds must fail beyond the spatial scale $p^{-3/2}$. Indeed, there we begin to witness that some polluted sites are no longer associated to well-defined blocking wedges with high probability. On the other hand, in order to build blocking loops with non-vanishing probability as $p \downarrow 0$, we have to work at the spatial scale at least as large as $p^{-3/2}$, since this enables some rare configurations that are needed for the argument. This issue can be circumvented if we can stitch blocking paths together to build larger ones.

Then we are faced with another issue, which is that joining two zigzag paths is not always possible, even if the paths intersect geometrically. In Figure 3.5, we are given two zigzags $1 \leftrightarrow 4$ and $2 \leftrightarrow 3$ that intersect in the middle. A moment's thought shows that 1 and 2 are always connected in any such configuration, while, 3 and 4 are not. This poses a serious restriction on joining zigzags to build a larger one. A considerable portion of this section is devoted to overcoming these issues.



$\xi_{\lambda,R}$ $\mathcal{R}_{5\pi/4}(\xi_{\lambda,R})$

Figure 3.6: Visualization of a sample of $\xi_{\lambda,R}$ and its rotation $\mathcal{R}_{5\pi/4}(\xi_{\lambda,R})$.

We first establish a series of percolation statements for $\xi_{\lambda,R}$ in the supercritical regime which are not immediate from the definition of the critical parameter λ_2^c. We are particularly interested in the box-crossing probabilities. For convenience, we define

$$\lambda_2^c(R) := \inf\Big\{\lambda \geq 0 : \lim_{r\to\infty} \mathbb{P}(\mathsf{B}_1 \overset{\xi_{\lambda,R}}{\longleftrightarrow} \partial\mathsf{B}_r) > 0\Big\}. \tag{3.5.1}$$

Comparing this with the definition (1.2.4), we have $\lambda_2^c \leq \lambda_2^c(R)$ for all $R \in (0,\infty)$ and

$$\lambda_2^c = \lim_{R\to\infty} \lambda_2^c(R). \tag{3.5.2}$$

By the nature of the definition (3.3.3), the percolation model on $\xi_{\lambda,R}$ behaves much like a 2-dimensional oriented percolation, or more generally, a 2-dimensional contact process. In such models, the infinite occupied cluster $\mathcal{C}(0)$ at the origin has a well-defined "opening", in the sense that $\mathcal{C}(0)$ scales to a deterministic cone. This motivates to introduce an analogous quantity for $\xi_{\lambda,R}$.

When doing so, it is convenient to consider a suitably rotated version of $\xi_{\lambda,R}$. Let $\mathcal{R}_\theta$ denote the counter-clockwise rotation by θ radian and define $\mathcal{R}_\theta(\xi_{\lambda,R}) := \{(\mathcal{R}_\theta x, \mathbf{u}) : (x,\mathbf{u}) \in \xi_{\lambda,R}\}$. The connectivity on $\mathcal{R}_\theta(\xi_{\lambda,R})$ is inherited from that of $\xi_{\lambda,R}$, or equivalently, $H = \{(x'_k, \mathbf{u}_k)\}_{k=1}^n$ is a hyperedge of $\mathcal{R}_\theta(\xi_{\lambda,R})$ exactly when $\{(\mathcal{R}_\theta^{-1} x'_k, \mathbf{u}_k)\}_{k=1}^n$ is a hyperedge of $\xi_{\lambda,R}$, where $\mathcal{R}_\theta^{-1} = \mathcal{R}_{-\theta}$ is the inverse operator. We will predominantly work with the choice $\theta = 5\pi/4$, so we will suppress the angle $5\pi/4$ from the notation and simply write $\mathcal{R} = \mathcal{R}_{5\pi/4}$ whenever no ambiguity arises.

Returning to the business of developing oriented percolation theory for $\xi_{\lambda,R}$, for each $t > 0$, we define

$$r_{0,t} := \sup\Big\{r : \{0\} \times (-\infty, 0] \xleftrightarrow{\mathcal{R}(\xi_{\lambda,R}) \text{ on } [0,t)\times\mathbb{R}} \{(t, r)\}\Big\}. \tag{3.5.3}$$

This marks the position of the top-most point on the vertical line $\{t\} \times \mathbb{R}$ that is connected to negative part $\{0\} \times (-\infty, 0]$ of the y-axis in the local configuration $\mathcal{R}(\xi_{\lambda,R}) \cap ([0, t) \times \mathbb{R} \times [0,\infty)^6)$. Then we have:

Lemma 3.5.1. *There exists a constant $\alpha(\lambda, R) \in [-\infty, 1)$ such that, for any sequence $(t_n)_{n=1}^{\infty}$ of positive real numbers satisfying $t_n \to \infty$,*

$$\lim_{n\to\infty} \frac{r_{0,t_n}}{t_n} = \alpha(\lambda, R) \qquad a.s. \tag{3.5.4}$$

Proof. Define $(r_{s,t})_{0<s<t}$ by

$$r_{s,t} := \sup\Big\{r : \{s\} \times (-\infty, r_{0,s} + \sqrt{2}R] \xleftrightarrow{\mathcal{R}(\xi_{\lambda,R}) \text{ on } [s,t)\times\mathbb{R}} \{(t, r_{0,s} + \sqrt{2}R + r)\}\Big\}. \tag{3.5.5}$$

Then we verify that, for each fixed $t > 0$, the following properties hold:

(1) $\{r_{(k-1)nt,knt}\}_{k=1}^{\infty}$ are independent and identically distributed for each $n \geq 1$,

(2) The distribution of $\{r_{nt,(n+k)t}\}_{k=0}^{\infty}$ does not depend on n,

(3) $r_{0,t} \leq r_{0,s} + r_{s,t} + \sqrt{2}R$ for all $0 \leq s \leq t$,

(4) $\mathbb{E}(r_{0,t} \vee 0) < t$ for all $t > 0$.

All the items except (3) follow readily from the properties of the homogeneous Poisson point process and the geometry of the grains. Also, (3) follows from the fact any hyperedge H joining $\{0\} \times (-\infty, 0]$ to $\{t\} \times \mathbb{R}$ must pass through $\{s\} \times (-\infty, r_{0,s} + \sqrt{2}R]$. The extra factor $\sqrt{2}R$ cannot be dropped since H may bypass the line $\{s\} \times (-\infty, r_{0,s}]$ due to the non-convexity of the wedges. Then by Liggett's subadditive ergodic theorem [69], $r_{0,nt}/nt$ converges to a constant almost surely. Finally, this improves to the limit (3.5.4) by using the part (3) again. □

Now we study the properties of $\alpha(\lambda, R)$.

Lemma 3.5.2. *Let $\alpha(\lambda, R)$ be defined by the statement of Lemma 3.5.1. Then*

(1) If $\lambda' > \lambda$, then $\alpha(\lambda', R) > \alpha(\lambda, R)$.

(2) If $\lambda > \lambda_2^c(R)$, then $\alpha(\lambda, R) > 0$.

Proof. For (1), the standard argument shows that $\xi_{\lambda,R}$ and $\xi_{\lambda',R}$ can be coupled to satisfy $\xi_{\lambda,R} \subseteq \xi_{\lambda',R}$, which implies $\alpha(\lambda', R) \geq \alpha(\lambda, R)$. In order to show that this inequality is strict, we construct another kind of coupling which we call *compression coupling.* Write $c := \lambda/\lambda' \in (0, 1)$ and define $\mathcal{T} : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$\mathcal{T}(x, y) := \left(\tfrac{1}{2}(x + y) + \tfrac{1}{2}c(x - y), \tfrac{1}{2}(x + y) - \tfrac{1}{2}c(x - y)\right). \tag{3.5.6}$$

This is a linear transformation which compresses the entire plane in the direction $(1, -1)$, which is best explained by the identity

$$(\mathcal{R}^{-1} \circ \mathcal{T} \circ \mathcal{R})(x, y) = (x, cy). \tag{3.5.7}$$

Then comparing the intensity measures, together with $\det(\mathcal{T}) = c = \lambda/\lambda'$, shows that

$$\mathcal{R}(\xi_{\lambda',R}) \overset{\text{law}}{=} \{(\mathcal{T}(x), \mathbf{u}) : (x, \mathbf{u}) \in \mathcal{R}(\xi_{\lambda,R})\}. \tag{3.5.8}$$

Using this, we may couple $\mathcal{R}(\xi_{\lambda,R})$ to $\mathcal{R}(\xi_{\lambda',R})$ so that (3.5.8) is a pointwise equality. Then by (3.5.7), any hyperedge of $\mathcal{R}(\xi_{\lambda,R})$ is mapped by $\mathcal{T}$ to a hyperedge of $\mathcal{R}(\xi_{\lambda',R})$. In particular, if $\{0\} \times (-\infty, 0]$ is connected to the singleton $\{(t, r)\}$ in $\mathcal{R}(\xi_{\lambda,R})$, then the line $\mathcal{T}(\{0\} \times (-\infty, 0])$ is connected to $\{\mathcal{T}(t, r)\}$ in $\mathcal{R}(\xi_{\lambda',R})$. Moreover, writing $(t', r') = \mathcal{T}(t, r)$, we get

$$\frac{r'}{t'} = \frac{(1 - c) + (1 + c)(r/t)}{(1 + c) + (1 - c)(r/t)}. \tag{3.5.9}$$

Now plugging $r := r_{0,t}$ to (3.5.9) and letting $t \to \infty$,

$$\alpha(\lambda', R) \geq \frac{(1 - c) + (1 + c)\alpha(\lambda, R)}{(1 + c) + (1 - c)\alpha(\lambda, R)} = \alpha(\lambda, R) + \frac{(1 - c)(1 - \alpha(\lambda, R)^2)}{(1 - c)\alpha(\lambda, R) + (1 + c)}. \tag{3.5.10}$$

This proves that α is indeed strictly increasing in the first variable.

For (2), assume that $\lambda > \lambda_2^c(R)$ holds. Then by the strict monotonicity of α in the first variable, it suffices to prove $\alpha(\lambda, R) \geq 0$. Define

$$\mathcal{E} := \bigcap_{n \in \mathbb{N}_1} \{\mathsf{B}_1 \xleftrightarrow{\mathcal{R}(\xi_{\lambda,R})} \partial \mathsf{B}_n\} \tag{3.5.11}$$

and note that $\mathbb{P}(\mathcal{E}) > 0$ by the assumption. The standard argument as in Durrett [41] shows that, if we define

$$\widetilde{r}_t := \sup\bigl\{r : \mathsf{B}_1 \xleftrightarrow{\mathcal{R}(\xi_{\lambda,R})} \{(t,r)\}\bigr\}, \tag{3.5.12}$$

then for any positive sequence $t_n \to \infty$,

$$\lim_{n\to\infty} \frac{\widetilde{r}_{t_n}}{t_n} = \alpha(\lambda, R) \qquad \text{a.s. on } \mathcal{E}. \tag{3.5.13}$$

Now assume otherwise that $\alpha(\lambda, R) < 0$. Let $\mathcal{F}_{1,t}$ and $\mathcal{F}_{2,t}$ be events defined by

$$\begin{aligned} \mathcal{F}_{1,t} &:= \bigl\{\mathsf{B}_1 \xleftrightarrow{\mathcal{R}(\xi_{\lambda,R})} \{t\} \times (0, -2t\alpha(\lambda, R))\bigr\}, \\ \mathcal{F}_{2,t} &:= \bigl\{(t,0) + \mathsf{B}_1 \xleftrightarrow{\mathcal{R}(\xi_{\lambda,R})} \{0\} \times (0, -2t\alpha(\lambda, R))\bigr\}. \end{aligned} \tag{3.5.14}$$

By the the symmetry of the law of $\mathcal{R}(\xi_{\lambda,R})$ under horizontal reflection, we have $\mathbb{P}(\mathcal{F}_{1,t}) = \mathbb{P}(\mathcal{F}_{2,t})$. Moreover, given $\mathcal{F}_{1,t} \cap \mathcal{F}_{2,t}$, we can find a hyperedge of $\mathcal{R}(\xi_{\lambda,R})$ joining B_1 to $(t,0)+\mathsf{B}_1$. So by the FKG inequality,

$$\mathbb{P}\bigl(\mathsf{B}_1 \xleftrightarrow{\mathcal{R}(\xi_{\lambda,R})} (t,0) + \mathsf{B}_1\bigr) \geq \mathbb{P}(\mathcal{F}_{1,t} \cap \mathcal{F}_{2,t}) \geq \mathbb{P}(\mathcal{F}_{1,t})^2. \tag{3.5.15}$$

Now, in light of (3.5.13), we know that $\liminf_{t\to\infty} \mathbb{P}(\mathcal{F}_{1,t}) \geq \mathbb{P}(\mathcal{E}) > 0$. However, this contradicts the assumption $\alpha(\lambda, R) < 0$, Therefore the desired conclusion follows. □

We will make frequent use of the crossing probabilities in long parallelograms. For each $L, M, \alpha \in (0, \infty)$, define $\mathsf{P}_{L,M,\alpha}$ as the parallelogram given by

$$\mathsf{P}_{L,M,\alpha} := \{(x,y) : x \in [0,L] \text{ and } y \in [\alpha x, \alpha x + M]\}. \tag{3.5.16}$$

Also, for each non-empty compact subset $K \subset \mathbb{R}^2$, we define

$$\begin{aligned} \texttt{left}(\mathsf{K}) &:= \{(x,y) \in K : x \leq x' \text{ for all } (x',y') \in K\}, \\ \texttt{right}(\mathsf{K}) &:= \{(x,y) \in K : x \geq x' \text{ for all } (x',y') \in K\}. \end{aligned} \tag{3.5.17}$$

In particular, $\mathtt{left}(\mathsf{P}_{L,M,\alpha}) = \{0\} \times [0, M]$ and $\mathtt{right}(\mathsf{P}_{L,M,\alpha}) = \{L\} \times [\alpha L, \alpha L + M]$. Then the first result concerning the crossing probability is:

Lemma 3.5.3. *Let $\lambda > \lambda_2^{\mathrm{c}}(R)$. Then for $\alpha = \alpha(\lambda, R)$ and for any $\delta > 0$,*

$$\lim_{L\to\infty} \mathbb{P}\Big(\mathtt{left}(\mathsf{P}_{L,\delta L,\alpha}) \xleftrightarrow{\mathcal{R}(\xi_{\lambda,R}) \text{ on } \mathsf{P}_{L,\delta L,\alpha}} \mathtt{right}(\mathsf{P}_{L,\delta L,\alpha})\Big) = 1. \tag{3.5.18}$$

Proof. This follows from the standard argument using large deviation principle, as in Durrett [41, Section 7, 9]. □

Lemma 3.5.4. *Let $\lambda > \lambda_2^{\mathrm{c}}(R)$. Then there exist $\delta_0 > 0$ and $0 < \beta_0 < \beta_1 < \beta_2 \leq \alpha(\lambda, R)$ such that the following assertions hold: For each $\delta \in (0, \delta_0)$, write*

$$\mathsf{Q}_1 := \mathsf{P}_{L/2,\delta L,\beta_2} \cup \left(\left(\tfrac{L}{2}, \tfrac{\beta_2 L}{2}\right) + \mathsf{P}_{L/2,\delta L,\beta_0}\right) \qquad \textit{and} \qquad \mathsf{Q}_2 := \mathsf{P}_{L/2,\delta L,\beta_1}. \tag{3.5.19}$$

Then for both $i = 1, 2$, we have

$$\lim_{L\to\infty} \mathbb{P}\Big(\mathtt{left}(\mathsf{Q}_i) \xleftrightarrow{\mathcal{R}(\xi_{\lambda,R}) \text{ on } \mathsf{Q}_i} \mathtt{right}(\mathsf{Q}_i)\Big) = 1 \tag{3.5.20}$$

Proof. The proof utilizes the idea of compression coupling as demonstrated in the proof of Lemma 3.5.2. Fix $\lambda_0 \in (\lambda_2^{\mathrm{c}}(R), \lambda)$ and let $\beta_0 = \alpha(\lambda_0, R)$, $c_2 = \lambda_0/\lambda$, and $c_1 \in (c_2, 1)$. Define β_1 and β_2 by

$$\beta_i := \frac{(1 - c_i) + (1 + c_i)\alpha(\lambda_0, R)}{(1 + c_i) + (1 - c_i)\alpha(\lambda_0, R)}, \qquad i \in \{1, 2\}. \tag{3.5.21}$$

Then by (3.5.10), we have $0 < \beta_0 < \beta_1 < \beta_2 \leq \alpha(\lambda, R)$. Using this, set $\delta_0 := \frac{1-\beta_0}{2}$. Then we prove that the assertion of the lemma holds with this choice of parameters. Indeed, for each $i \in \{1, 2\}$, let $\mathcal{T}_i : \mathbb{R}^2 \to \mathbb{R}^2$ be the function defined by

$$\mathcal{T}_i(x, y) := \begin{cases} \left(\frac{1}{2}(x + y) + \frac{1}{2}c_i(x - y), \frac{1}{2}(x + y) - \frac{1}{2}c_i(x - y)\right), & \text{if } y > x, \\ (x, y), & \text{otherwise.} \end{cases} \tag{3.5.22}$$

Arguing similarly as in the proof of Lemma 3.5.2, for each $i \in \{1, 2\}$, we may couple $\mathcal{R}(\xi_{\lambda_0,R})$ and $\mathcal{R}(\xi_{\lambda,R})$ so that

$$\{(\mathcal{T}_i(x), \mathbf{u}) : (x, \mathbf{u}) \in \mathcal{R}(\xi_{\lambda_0,R})\} \subseteq \mathcal{R}(\xi_{\lambda,R}) \tag{3.5.23}$$

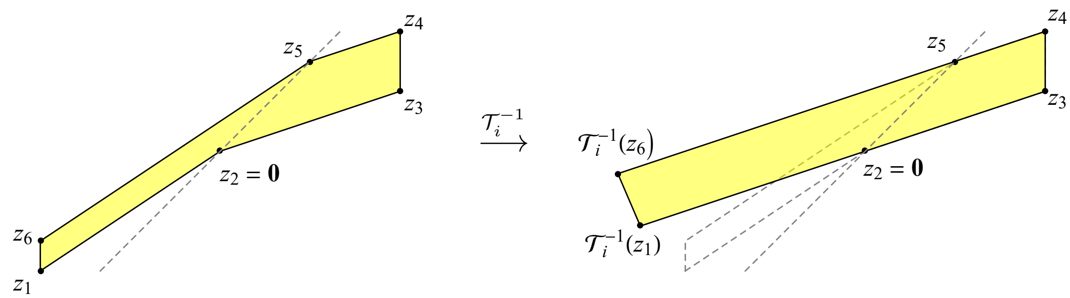


Figure 3.7: Polygon with vertices $z_1, \cdots, z_6$ given by (3.5.24) and its inverse image under $\mathcal{T}_i$.

holds true. In particular, under this coupling, any hyperedge of $\mathcal{R}(\lambda_{\lambda_0,R})$ is mapped by $\mathcal{T}_i$ to a hyperedge of $\mathcal{R}(\xi_{\lambda,R})$. To make use of this observation, we fix $i \in \{1,2\}$ and consider the points

$$\begin{aligned} z_1 &:= \left(-\tfrac{L}{2}, -\tfrac{\beta_i L}{2}\right), \qquad z_2 := \mathbf{0}, \qquad z_3 := \left(\tfrac{L}{2}, \tfrac{\beta_0 L}{2}\right), \\ z_4 &:= z_3 + (0, \delta L), \qquad z_5 := \left(\tfrac{\delta L}{1-\beta_0}, \tfrac{\delta L}{1-\beta_0}\right), \qquad z_6 := z_1 + \left(0, \tfrac{1-\beta_i}{1-\beta_0}\delta L\right). \end{aligned} \tag{3.5.24}$$

Here, the condition $\delta < \delta_0$ ensures that $\frac{\delta L}{1-\beta_0} < \frac{L}{2}$, and so, z_5 lies to the left of z_4. Then by a direct computation, we check that the following holds true:

- All of $\mathcal{T}_i^{-1}(z_1)$, $\mathcal{T}_i^{-1}(z_2)$, and $\mathcal{T}_i^{-1}(z_3)$ lie on the line $y = \beta_0 x$.
- All of $\mathcal{T}_i^{-1}(z_4)$, $\mathcal{T}_i^{-1}(z_5)$, and $\mathcal{T}_i^{-1}(z_6)$ lie on the line $y = \beta_0 x + \delta L$.

So, if Q_i' denotes the polygon with vertices $z_1, \cdots, z_6$ as in Figure 3.7, then the inverse image $\mathcal{T}_i^{-1}(\mathsf{Q}_i')$ is a polygon with four corners $\mathcal{T}_i^{-1}(z_1)$, z_3, z_4, and $\mathcal{T}_i^{-1}(z_6)$. By invoking the coupling (3.5.23) and writing

$$\mathsf{L}_i := [\text{line segment joining } \mathcal{T}_i^{-1}(z_1) \text{ to } \mathcal{T}_i^{-1}(z_6)], \tag{3.5.25}$$

we get

$$\begin{aligned} &\mathbb{P}\Big(\mathtt{left}(\mathsf{Q}_i') \xleftrightarrow{\mathcal{R}(\xi_{\lambda,R}) \text{ on } \mathsf{Q}_i'} \mathtt{right}(\mathsf{Q}_i')\Big) \\ &\qquad \geq \mathbb{P}\Big(\mathsf{L}_i \xleftrightarrow{\mathcal{R}(\xi_{\lambda_0,R}) \text{ on } \mathcal{T}_i^{-1}(\mathsf{Q}_i')} \mathtt{right}\big(\mathcal{T}_i^{-1}(\mathsf{Q}_i')\big)\Big). \end{aligned} \tag{3.5.26}$$

Note that $\mathcal{T}_i^{-1}(\mathsf{Q}_i)$ lies between two lines $\{-aL\} \times \mathbb{R}$ and $\{\frac{L}{2}\} \times \mathbb{R}$, where $a := \frac{1+\beta_i}{2(1+\beta_0)} + \frac{\beta_i - \beta_0}{1-\beta_0^2}$. Then by setting $b := a + \frac{1}{2}$, we may further bound (3.5.26) from below by

$$\mathbb{P}\Big(\mathtt{left}(\mathsf{P}_{bL,\delta L,\beta_0}) \xleftrightarrow{\mathcal{R}(\xi_{\lambda,R}) \text{ on } \mathsf{P}_{bL,\delta L,\beta_0}} \mathtt{right}(\mathsf{P}_{bL,\delta L,\beta_0})\Big), \tag{3.5.27}$$

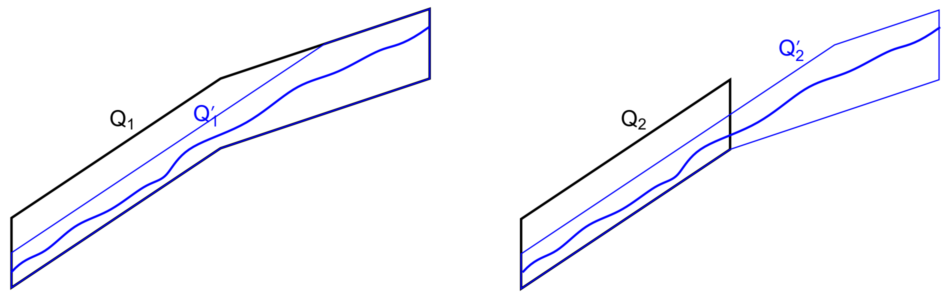


Figure 3.8: A left-right crossing of Q_i' inducing that of (a translate of) Q_i.

which tends to 1 as $L \to \infty$ by Lemma 3.5.3. Finally, the desired conclusion follows by noting that, for each $i \in \{1, 2\}$, the existence of a left-right crossing in Q_i' implies the existence of a crossing of a translate of Q_i, see Figure 3.8. $\square$

Corollary 3.5.5. *Let $\lambda > \lambda_2^c(R)$. Then there exist $\delta_1 > 0$ and $0 < \gamma_0 < \gamma_1 < \gamma_2 < 1$ such that the following assertion holds: For each δ with $0 < \delta < \delta_1$ and $L > 0$, set*

$$\begin{aligned}\mathsf{Q}_L := &\left\{(-\gamma_3 t + u, t) : 0 \le t \le \tfrac{L}{2} \textit{ and } |u| \le \delta L\right\}\\ &\cup \left\{(-\gamma_1 t + u, t) : -\tfrac{L}{2} \le t \le 0 \textit{ and } |u| \le \delta L\right\}\\ &\cup \left\{(-2\delta L - \gamma_2 t + u, t) : |t| \le \tfrac{L}{2} \textit{ and } |u| \le \delta L\right\}\end{aligned} \tag{3.5.28}$$

and

$$\begin{aligned}\mathtt{top}_1(\mathsf{Q}_L) &:= \left[-\tfrac{\gamma_3 L}{2} - \delta L, -\tfrac{\gamma_3 L}{2} + \delta L\right] \times \left\{\tfrac{L}{2}\right\}\\ \mathtt{top}_2(\mathsf{Q}_L) &:= \left[-\tfrac{\gamma_2 L}{2} - 3\delta L, -\tfrac{\gamma_2 L}{2} - \delta L\right] \times \left\{\tfrac{L}{2}\right\}\\ \mathtt{bottom}_1(\mathsf{Q}_L) &:= \left[\tfrac{\gamma_1 L}{2} - \delta L, \tfrac{\gamma_1 L}{2} + \delta L\right] \times \left\{-\tfrac{L}{2}\right\}\\ \mathtt{bottom}_2(\mathsf{Q}_L) &:= \left[\tfrac{\gamma_2 L}{2} - 3\delta L, \tfrac{\gamma_2 L}{2} - \delta L\right] \times \left\{-\tfrac{L}{2}\right\}.\end{aligned} \tag{3.5.29}$$

Then we have

$$\lim_{L\to\infty} \mathbb{P}\left(\bigcap_{i,j\in\{1,2\}} \left\{\mathtt{top}_i(\mathsf{Q}_L) \xleftrightarrow{\mathcal{R}_\pi(\xi_{\lambda,R}) \textit{ on } \mathsf{Q}_L} \mathtt{bottom}_j(\mathsf{Q}_L)\right\}\right) = 1. \tag{3.5.30}$$

Proof. Let β_i's be as in Lemma 3.5.4 and set $\gamma_{i+1} := \frac{1-\beta_i}{1+\beta_i}$. Then by rotating Q_L counter-clockwise by $\frac{\pi}{4}$ radian and flipping horizontally, the desired claim follows from Lemma 3.5.4 and the FKG inequality. $\square$

Now we are ready to establish:

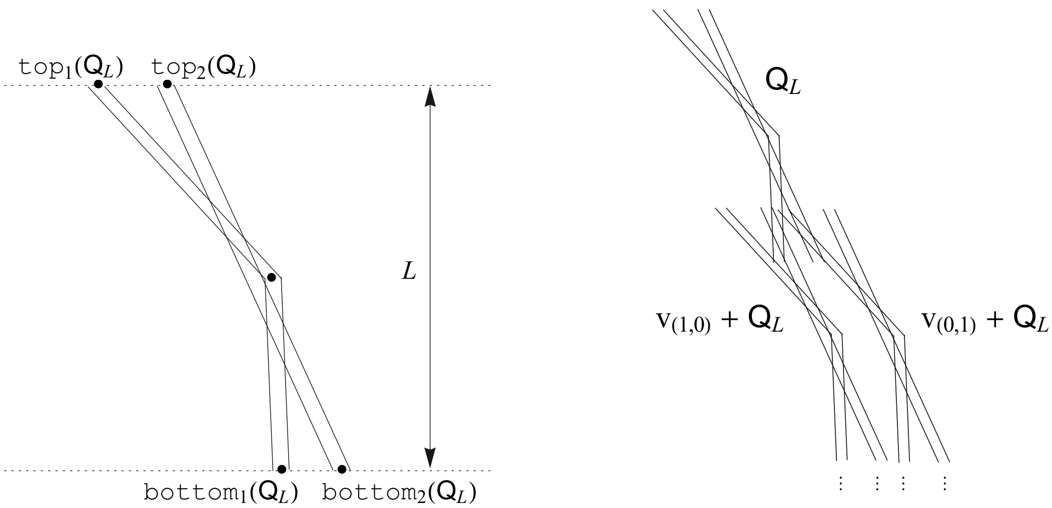


Figure 3.9: (Left) The set Q_L in Corollary 3.5.5. (Right) Part of the binary tree of rescaled sites.

Proposition 3.5.6. *Let $\lambda > \lambda_2^{\mathrm{c}}(R)$, $d > 1$, and $\epsilon > 0$. Then there exist $L > 0$ and $0 < a < b < 1$ such that the following holds with $\mathbb{P}_{p,\lambda p^2}$-probability at least $1-\epsilon$:*

(1) There exists a blocking SW-zigzag Z joining $\{0\} \times [\![p^{-d} - \frac{L}{p}, p^{-d}]\!]$ to $[\![ap^{-d}, bp^{-d}]\!] \times \{0\}$ in ω.

(2) For any two SE-wedges W and W' in Z, no site of $B(\mathsf{W} \cap \mathsf{W}', 100)$ is occupied in ω.

(3) Z does not intersect $B(0, \frac{a}{2}p^{-d})$.

Proof. Fix $\lambda' \in (\lambda_2^{\mathrm{c}}(R), \lambda)$ and let $\gamma_1, \gamma_2, \gamma_3$ be as in the statement of Corollary 3.5.5 with λ' in place of λ. Choose $\delta \in (0, \delta_1)$ so that $\delta < \frac{1}{16}(\gamma_3 - \gamma_2) \wedge (\gamma_2 - \gamma_1)$. For each $p \in (0,1)$ and $L > 0$, define the "discretized" version of (3.5.28) and (3.5.29) by

$$\mathsf{Q}_L^{\mathrm{d}} := \{x \in \mathbb{Z}^2 : px \in \mathsf{Q}_L\} \tag{3.5.31}$$

and

$$\begin{aligned}
\mathtt{top}_1(\mathsf{Q}_L^{\mathrm{d}}) &:= [\![-\tfrac{\gamma_3 L}{2p} - \tfrac{\delta L}{p}, -\tfrac{\gamma_3 L}{2p} + \tfrac{\delta L}{p}]\!] \times \{\lfloor \tfrac{L}{2p} \rfloor\} \\
\mathtt{top}_2(\mathsf{Q}_L^{\mathrm{d}}) &:= [\![-\tfrac{\gamma_2 L}{2p} - \tfrac{3\delta L}{p}, -\tfrac{\gamma_2 L}{2p} - \tfrac{\delta L}{p}]\!] \times \{\lfloor \tfrac{L}{2p} \rfloor\} \\
\mathtt{bottom}_1(\mathsf{Q}_L^{\mathrm{d}}) &:= [\![\tfrac{\gamma_1 L}{2p} - \tfrac{\delta L}{p}, \tfrac{\gamma_1 L}{2p} + \tfrac{\delta L}{p}]\!] \times \{-\lfloor \tfrac{L}{2p} \rfloor\} \\
\mathtt{bottom}_2(\mathsf{Q}_L^{\mathrm{d}}) &:= [\![\tfrac{\gamma_2 L}{2p} - \tfrac{3\delta L}{p}, \tfrac{\gamma_2 L}{2p} - \tfrac{\delta L}{p}]\!] \times \{-\lfloor \tfrac{L}{2p} \rfloor\}.
\end{aligned} \tag{3.5.32}$$

Also, for each $m, n \in \mathbb{Z}$, define

$$\begin{aligned} \mathrm{v}_{(m,n)} := m\big(\lfloor \tfrac{1}{4p}(\gamma_1 + 2\gamma_2 - 8\delta)L\rfloor, -\lfloor \tfrac{3}{4p}L\rfloor\big) \\ + n\big(\lfloor \tfrac{1}{4p}(\gamma_2 + 2\gamma_3 - 16\delta)L\rfloor, -\lfloor \tfrac{3}{4p}L\rfloor\big) \end{aligned} \tag{3.5.33}$$

Now let $\mathbb{L} := \mathbb{N}_0^2$. We turn $\mathbb{L}$ into an oriented graph by declaring that the edge-set of $\mathbb{L}$ consists of all edges of the form $(m, n) \to (m+1, n)$ or $(m, n) \to (m, n+1)$ for any $(m, n) \in \mathbb{L}$. We say that the site $x \in \mathbb{L}$ is *rescaled open* if the event

$$\begin{aligned} \mathcal{F}_x := \bigcap_{i,j\in\{1,2\}} \Big\{ \mathtt{top}_i(\mathrm{v}_x + \mathsf{Q}_L^{\mathrm{d}}) \xleftrightarrow{\omega \text{ on } \mathrm{v}_x + \mathsf{Q}_L^{\mathrm{d}}} \mathtt{bottom}_j(\mathrm{v}_x + \mathsf{Q}_L^{\mathrm{d}}) \Big\} \\ \cap \left\{ \begin{array}{c} \text{for any hyperedge } ((x, \mathsf{W}), (x', \mathsf{W}')) \text{ of } G(\omega, \mathrm{v}_x + \mathsf{Q}_L^{\mathrm{d}}), \\ \text{no site of } B(\mathsf{W} \cap \mathsf{W}', 100) \text{ is occupied in } \omega \end{array} \right\} \end{aligned} \tag{3.5.34}$$

holds. Then we remark that, δ and v_x's are chosen in such a way that, if $(x_0, \cdots, x_n)$ is a rescaled open path in $\mathbb{L}$, then there exists a hyperedge of $G(\omega, \mathbb{Z}^2)$ joining $\mathtt{top}_i(\mathrm{v}_{x_0} + \mathsf{Q}_L^{\mathrm{d}})$ to $\mathtt{bottom}_j(\mathrm{v}_{x_0} + \mathsf{Q}_L^{\mathrm{d}})$ for each $i, j \in \{1, 2\}$. Following the same argument as in the proof of Proposition 3.4.1, together with the estimate (3.3.25) and Lemma 3.5.5, for each $\epsilon > 0$ we can find large enough $L > 0$ so that

$$\liminf_{p\downarrow 0} \inf_{x\in\mathbb{L}} \mathbb{P}_{p,\lambda p^2}(\mathcal{F}_x) > 1 - \epsilon. \tag{3.5.35}$$

occurs. Also, there exists $K \in \mathbb{N}_1$ such that states of rescaled sites with ℓ^∞-distance larger than K are independent. Invoking stochastic domination of Liggett, Schonmann and Stacey [68] by a product Bernoulli measure with parameter sufficiently close to 1, we find that there exists an infinite rescaled-open path from 0 with probability close to 1. These altogether imply the desired claim. □

3.6 Proof of the main result

In this section, we establish Theorem 3.1.2.

Proof of Theorem 3.1.2, part (1). Let $\lambda < \lambda_1^c$ and $(\mathcal{B}^t\omega)_{t\in\mathbb{N}_0\cup\{\infty\}}$ be the bootstrap process started with the initial configuration ω sampled from the law $\mathbb{P}_{p,\lambda p^2}$. Fix $\alpha \in (0, \frac{1}{5})$ and let $L = \lfloor p^{-1-\alpha} \rfloor$. Call $x \in \mathbb{Z}^2$ a *rescaled bad site* if at least one of the followings events hold:

$$\begin{aligned} &\mathcal{F}_1(x) := \{B(Lx, 3L) \text{ contains a defect in } \omega\},\\ &\mathcal{F}_2(x) := \{B(Lx, 3L) \text{ contains a double line with two or more polluted sites in } \omega\},\\ &\mathcal{F}_3(x) := \{B(Lx, 3L) \text{ contains a zigzag of length} \geq L \text{ in } \omega\}. \end{aligned} \tag{3.6.1}$$

Otherwise, x called a *rescaled good site.* By the union bound, we immediately bet

$$\mathbb{P}_{p,\lambda p^2}(\mathcal{F}_1(x)) \leq (6L+1)^2 \mathbb{P}_{p,\lambda p^2}(\mathbf{0} \text{ is a defect in } \omega) = \mathcal{O}(p^{1-2\alpha}) \tag{3.6.2}$$

Moreover, using the coupling of Lemma 3.3.1 and noting that $\mathcal{E}_3$ is implied by $\mathcal{F}_2(0)$ with the choice $b = 3$, we get

$$\mathbb{P}_{p,\lambda p^2}(\mathcal{F}_2(x)) = \mathcal{O}(p^{1-3\alpha}). \tag{3.6.3}$$

Finally, Proposition 3.4.1 together with the union bound shows that $\mathbb{P}_{p,\lambda p^2}(\mathcal{F}_3(x)) \to 0$ as $p \downarrow 0$. Combining altogether, we get

$$\lim_{p\downarrow 0} \mathbb{P}_{p,\lambda p^2}(x \text{ is a rescaled bad site}) = 0. \tag{3.6.4}$$

Since $\cup_{i=1}^4 \mathcal{F}_i(x)$ and $\cup_{i=1}^4 \mathcal{F}_i(y)$ are independent if $\|x-y\|_\infty \geq 6$, invoking stochastic domination by a suitable Bernoulli percolation in $\mathbb{Z}^2$ as proved in Liggett, Schonmann, and Stacey [68] shows that the set of rescaled good sites percolates.

Now assume that every site of $B(0, L)$ is initially occupied. If the terminally occupied cluster $\mathcal{C}(0)$ around the origin is finite, then by Proposition 3.2.5, the outer boundary of $\mathcal{C}(0)$ is a subset of $\mathbb{Z}^2 \setminus B(0, L)$ which consists of blocking contours joined by defects in ω, which then implies that there exists a rescaled bad loop in $\mathbb{Z}^2 \setminus B(0, 4)$. Since the probability that such rescaled bad loop exists is strictly less than one by the percolation of rescaled good sites, we deduce that $\mathcal{C}(0)$ must be infinite with positive probability, which is enough to conclude part (1) of Theorem 3.1.2. □

Proof of Theorem 3.1.2, part (2). Let $\lambda > \lambda_2^{\mathrm{c}}$. Then by (3.5.2), there exists $R \in (0, \infty)$ such that $\lambda > \lambda_2^{\mathrm{c}}(R)$. Now let $\epsilon > 0$ and choose $L > 0$ and $0 < a < b < 1$ as in the statement of Proposition 3.5.6 with $d = 6$. Now write $y_t = \lfloor p^{-6} \rfloor - t \lfloor \frac{L}{p} \rfloor$ and run the following algorithm given ω:

- t is initialized with the value 0, i.e., we start with $t = 0$.
- Find the smallest $l_t, r_t \in [\![3, \frac{1}{p}]\!]$ such that $\omega(-l_t, y_t) = 2$ and $\omega(r_t, y_t) = 2$.
- If such l_t and r_t exist, then test whether there exists a site x in any of the rectangles $[\![-l_t, r_t]\!] \times [\![y_t - 1, y_t]\!]$, $[\![-l_t, -l_t + 1]\!] \times [\![y_t - \frac{L}{p}, y_t]\!]$, or $[\![r_t - 1, r_t]\!] \times [\![y_t - \frac{L}{p}, y_t]\!]$ that satisfies $\omega(x) = 1$. If indeed there is no such site x, then halt the algorithm.
- Otherwise, increase the value of t by 1. If $t > p^{-4}$, then halt the algorithm.

Then the $\mathbb{P}_{p,\lambda p^2}$-probability that the algorithm halts before time $\lfloor p^{-4} \rfloor$ is at least as large as $1 - \epsilon$ for any sufficiently small p. Moreover, given $t \leq p^{-4}$ and the values of l_t and r_t, the initial configuration ω restricted to $(\mathbb{Z} \setminus [\![-l_t, r_t]\!]) \times ((-\infty, y_n] \cap \mathbb{Z})$ is not explored, and so, is independent of the states of sites that have been explored by the algorithm. Then by Proposition 3.5.6, there exist

- a blocking SW-zigzag joining $\{r_t\} \times [\![y_t - \frac{L}{p}, y_t]\!]$ to $(r_t, y_t) + [\![ap^{-3}, bp^{-3}]\!]$, and
- a blocking SE-zigzag joining $\{-l_t\} \times [\![y_t - \frac{L}{p}, y_t]\!]$ to $(-l_t, y_t) + [\![-bp^{-3}, -ap^{-3}]\!]$

with $\mathbb{P}_{p,\lambda p^2}$-probability at east $1 - 2\epsilon$. Then by the rotation symmetry of the law of the initial configuration and the FKG inequality, it follows that

$$\mathbb{P}_{p,\lambda p^2}(\exists \text{ an 'outward' blocking loop in } B(0, p^{-6}) \setminus B(0, \tfrac{a}{3} p^{-6})) \geq 1 - 12\epsilon \tag{3.6.5}$$

for all sufficiently small p. From this place, we can employ Gravner and McDonald's argument [55, Lemma 3.3 and thereafter] *mutatis mutandis* to conclude that $\phi(p, \lambda p^2) = 0$ if p is sufficiently small. □

Chapter 4

Asymptotic normality of permutation statistics

4.1 A central limit theorem for descents in conjugacy classes of S_n

The distribution of descent numbers in a fixed conjugacy class of the symmetric group S_n of n elements has been studied, and its moments have been shown to exhibit interesting properties. Fulman [48] proved that the descent numbers of permutations in conjugacy classes with large cycles are asymptotically normal, and Kim [63] proved that the descent numbers of fixed point free involutions are also asymptotically normal. In this chapter, we generalize these results to prove asymptotic normality for descent numbers of permutations in any conjugacy class of S_n.

4.1.1 Introduction

The Eulerian function $A_n(x)$ was first defined by the relation

$$\sum_{a=1}^{\infty} a^n x^a = \frac{A_n(x)}{(1-x)^{n+1}} \tag{4.1.1}$$

by Euler in [45] when he evaluated the zeta function $\zeta(s)$ at negative integers. It turns out that the coefficients of the x^{k+1} term in $A_n(x)$, written $A_{n,k}$ and called *Eulerian numbers*, can be interpreted combinatorially.

Definition 4.1.1. *A permutation $\pi \in S_n$ has a descent at position i if $\pi(i) > \pi(i+1)$, where $i = 1, \ldots, n-1$, and the descent set of π, denoted $Des(\pi)$ is the set of all descents of π. The descent number of π is defined as $d(\pi) := |Des(\pi)|$.*

The results of MacMahon and Riordan, in [70] and [85] respectively, showed that $A_{n,k}$ is the number of permutations in S_n with k descents.

The theory of descents in permutations has been studied thoroughly and is related to many questions. In [65], Knuth connected descents with the theory of sorting and the theory of runs in permutations, and in [37], Diaconis, McGrath, and Pitman studied a model of card shuffling in which descents play a central role. Bayer and Diaconis also used descents and rising sequences to give a simple expression for the chance of any arrangement after any number of shuffles and used this to give sharp bounds on the approach to randomness in [10]. Garsia and Gessel found a generating function for the joint distribution of descents, major index, and inversions in [52], and Gessel and Reutenauer showed that the number of permutations with given cycle structure and descent set is equal to the scalar product of two special characters of the symmetric group in [53]. Diaconis and Graham also explained Peirce's dyslexic principle using descents in [36]. Petersen also has an excellent and very thorough book on Eulerian numbers [80].

It is well known ([12], [46]) that the distribution of $d(\pi)$ in S_n is asymptotically normal with mean $\frac{n+1}{2}$ and variance $\frac{n-1}{12}$. Fulman also used Stein's method to show that the number of descents of a random permutation satisfies a central limit theorem with error rate $n^{-1/2}$ in [49]. In [97], Vatutin proved a central limit theorem for $d(\pi)+d(\pi^{-1})$, where π is a random permutation. Later this result was rediscovered and generalized by Chatterjee and Diaconis in [25].

Using generating functions, Fulman proved the following analogous result in [48] about conjugacy classes with large cycles only:

Theorem 4.1.2. *For every $n \geq 1$, pick a conjugacy class $\mathcal{C}_n$ in S_n, and let $m_i(\mathcal{C}_n)$ be* *the number of i-cycles in $\mathcal{C}_n$. Suppose that for all i, $m_i(\mathcal{C}_n) \to 0$ as $n \to \infty$. Then, the distribution of $d(\pi)$ in $\mathcal{C}_n$ is asymptotically normal with mean $\frac{n-1}{2}$ and variance $\frac{n+1}{12}$.*

Kim also used generating functions, in [63], to prove the following central limit theorem about the conjugacy class of fixed point free involutions:

Theorem 4.1.3. *For every $n \geq 1$ even, let $\mathcal{C}_n$ be the conjugacy class of fixed point free involutions in S_n. Then, the distribution of $d(\pi)$ in $\mathcal{C}_n$ is asymptotically normal with asymptotic mean $\frac{n}{2}$ and asymptotic variance $\frac{n}{12}$.*

After the above result was proved, Diaconis conjectured that there are asymptotic normality results for conjugacy classes that are fixed point free. In this section, we will prove a generalized version of this conjecture that proves asymptotic normality of descents for all conjugacy classes of S_n.

Theorem 4.1.4. *For every $n \geq 1$, pick a conjugacy class $\mathcal{C}_n$ in S_n, and let $m_i(\mathcal{C}_n)$ be the number of i-cycles in $\mathcal{C}_n$. Suppose that $m_1(\mathcal{C}_n)/n \to \alpha$ for some $\alpha \in [0,1]$ as $n \to \infty$. Then, the distribution of $d(\pi)$ in $\mathcal{C}_n$ is asymptotically normal with asymptotic mean $(1-\alpha^2)\frac{n}{2}$ and asymptotic variance $(1-4\alpha^3+3\alpha^4)\frac{n}{12}$.*

The outline is as follows. In Section 4.1.2, we expand the generating function $A_{\mathcal{C}_\lambda}(t)$ at infinity to obtain a series expression that is convergent for $|t| > 1$. In Section 4.1.3, we calculate the asymptotic variance of descent numbers of permutations, chosen uniformly at random, from a conjugacy class $\mathcal{C}_\lambda$, where $\lambda \vdash n$. Finally, in Section 4.1.4, we prove the following main theorem on the moment generating function M_λ of the normalized descent numbers.

Theorem 4.1.5. *Write $\alpha_\lambda = m_1/n$. Then, there exists a function $C : \mathbb{R} \to (0,\infty)$ such that*

$$\left| M_\lambda(s) - \exp\left\{ \frac{s^2}{24}\left(1-4\alpha_\lambda^3+3\alpha_\lambda^4\right)\right\}\right| \leq C(s)\frac{\log^3 n}{\sqrt{n}}$$

for any $n \geq 1$ and for any $\lambda \vdash n$.

We obtain the asymptotic normality as a consequence.

4.1.2 Crossing the singularity

Let $\mathcal{C}_\lambda$ be a conjugacy class of S_n, where $\lambda \vdash n$ and λ consists of the cycle lengths. For each subset $S \subseteq S_n$, define the generating function $A_S(t) = \sum_{\pi \in S} t^{d(\pi)+1}$. In [48], Fulman showed that, if λ has m_i i's,

$$A_{\mathcal{C}_\lambda}(t) = (1-t)^{n+1} \sum_{a=1}^{\infty} t^a \prod_{i=1}^{n} \binom{f_{i,a} + m_i - 1}{m_i}, \tag{4.1.2}$$

where $f_{i,a} = \frac{1}{i} \sum_{d|i} \mu(d) a^{i/d}$ and $\mu(d)$ is the Möbius function. This identity holds as a formal power series, and as an actual convergent series for $|t| < 1$. In [83], Reutenauer showed that $f_{i,a}$ counts the number of primitive circular words of length i from the alphabet $\{1, \ldots, a\}$.

Recall that the moment generating function (MGF) of a random variable X is defined by $M_X(s) = \mathbb{E}[e^{sX}]$. In [63], Kim observed that we can construct $M_n(s)$, the MGF of descents in fixed point free involutions, from (4.1.2), by the relation $M_n(s) = A_{\mathcal{C}_\lambda}(e^s)$. After showing that the MGF converges pointwise to $e^{s^2/24}$, which is the MGF of a normal distribution, the desired central limit theorem followed from the pointwise convergence and the following result of Curtiss from [27].

Theorem 4.1.6. *Suppose we have a sequence $\{X_n\}_{n=1}^{\infty}$ of random variables and there exists $s_0 > 0$ such that each MGF $M_n(s) = \mathbb{E}\left[e^{sX_n}\right]$ converges for $s \in (-s_0, s_0)$. If $M_n(s)$ converges pointwise to some function $M(s)$ for each $s \in (-s_0, s_0)$, then M is the MGF of some random variable X, and X_n converges to X in distribution.*

However, since (4.1.2) is convergent for $|t| < 1$, the above relation for $M_n(s)$ is only convergent for $s < 0$. Fortunately, the descents of fixed point free involutions have a crucial palindromic property, also proven in [63], which implies $M_n(s) = M_n(-s)$, and so, the pointwise convergence follows for all s.

It turns out that descents of other conjugacy classes of S_n do not have this palindromic property. Hence, we need an expression for $A_{\mathcal{C}_\lambda}(t)$, similar to (4.1.2), that converges for $|t| > 1$, in order to deal with MGF for $s < 0$. We claim the following proposition.

Proposition 4.1.7. *Let $\mathcal{C}_\lambda$ be a conjugacy class of S_n. Then, for $|t| > 1$,*

$$A_{\mathcal{C}_\lambda}(t) = (t-1)^{n+1} \sum_{a=1}^{\infty} t^{-a} \left[(-1)^n \prod_{i=1}^{n} \binom{f_{i,-a} + m_i - 1}{m_i} \right]. \tag{4.1.3}$$

In order to prove the proposition, we first prove an analogous statement for $A_n(t) = A_{S_n}(t)$.

Lemma 4.1.8. *For $|t| > 1$,*

$$A_n(t) = (t-1)^{n+1} \sum_{a=1}^{\infty} a^n t^{-a}. \tag{4.1.4}$$

Proof. Recall that $A_n(u) = (1-u)^{n+1} \sum_{a \geq 1} a^n u^a$, which holds both formally and as convergent series for $|u| < 1$. Now assume that $|t| > 1$ and set $u = 1/t$. Using the identity $t^{n+1} A_n(1/t) = A_n(t)$, we get

$$A_n(t) = t^{n+1}(1-1/t)^{n+1} \sum_{a=1}^{\infty} a^n t^{-a} = (t-1)^{n+1} \sum_{a=1}^{\infty} a^n t^{-a}.$$

□

Proof of Proposition 4.1.3. Since $f_{i,a}$ is a polynomial in a, $\binom{f_{i,a}+m_i-1}{m_i}$ is also a polynomial in a. In view of this fact, we can write

$$\prod_{i=1}^{n} \binom{f_{i,a} + m_i - 1}{m_i} = \sum_{k=1}^{n} c_k a^k,$$

and so,

$$A_{\mathcal{C}_\lambda}(t) = \sum_{k=1}^{n} c_k (1-t)^{n-k} A_k(t).$$

Hence, by (4.1.4), for $|t| > 1$, we have

$$\begin{aligned} A_{\mathcal{C}_\lambda}(t) &= \sum_{k=1}^{n} c_k (1-t)^{n-k} (t-1)^{k+1} \sum_{a=1}^{\infty} a^k t^{-a} \\ &= (t-1)^{n+1} \sum_{a=1}^{\infty} t^{-a} \left[(-1)^n \sum_{k=1}^{n} c_k (-a)^k \right] \\ &= (t-1)^{n+1} \sum_{a=1}^{\infty} t^{-a} \left[(-1)^n \prod_{i=1}^{n} \binom{f_{i,-a} + m_i - 1}{m_i} \right] \end{aligned}$$

□

4.1.3 Computation of the asymptotic variance

In [48], Fulman showed that the asymptotic mean of the descent numbers of $\mathcal{C}_\lambda$ is

$$(1-\alpha^2)\frac{n}{2}$$

as $n \to \infty$ and $m_1/n \to \alpha$, by analyzing (4.1.2). Using similar methods, we calculate the asymptotic variance.

Lemma 4.1.9. *The asymptotic variance of the descent numbers of $\mathcal{C}_\lambda$ is*

$$\left(1-4\alpha^3+3\alpha^4\right)\frac{n}{12}$$

as $n \to \infty$ and $m_1/n \to \alpha$.

Proof. From (4.1.2), we see that

$$\frac{A_{\mathcal{C}_\lambda}(t)}{|\mathcal{C}_\lambda|} = \frac{(1-t)^{n+1}}{n!}\sum_{a=0}^{\infty} t^a \prod_{i=1}^{n}\left(\sum_{k=1}^{m_i}\sum_{d_1,\ldots,d_k|i} i^{m_i-k}\begin{bmatrix} m_i \\ k \end{bmatrix} \times \mu\left(\frac{i}{d_1}\right)\cdots\mu\left(\frac{i}{d_k}\right)a^{d_1+\cdots+d_k}\right),$$

where $\left[{n \atop k}\right]$ denotes the Stirling numbers of the first kind. The asymptotic mean in [48] was calculated by noting that $d_1+\cdots+d_k = im_i$ if and only if $k=m_1$ and $d_1=\cdots d_k=i$, and also noting that $d_1+\cdots+d_k = im_i - 1$ if and only if one of the following is true:

(1) $i=2$, $k=m_2$, and $\{d_1,\ldots,d_k\}=\{1,2,\ldots,2\}$ as multisets.

(2) $i=1$, $k=m_1-1$, and $\{d_1,\ldots,d_k\}=\{1,\ldots,1\}$ as multisets.

Similarly, we note that $d_1+\cdots+d_k = im_i-2$ if and only if one of the following is true:

(1) $i=4$, $k=m_4$, and $\{d_1,\ldots,d_k\}=\{2,4,\ldots,4\}$ as multisets.

(2) $i=3$, $k=m_3$, and $\{d_1,\ldots,d_k\}=\{1,3,\ldots,3\}$ as multisets.

(3) $i=2$, $k=m_2$, and $\{d_1,\ldots,d_k\}=\{1,1,2,\ldots,2\}$ as multisets.

(4) $i=2$, $k=m_2-1$, and $\{d_1,\ldots,d_k\}=\{2,\ldots,2\}$ as multisets.

(5) $i=1$, $k=m_1-2$, and $\{d_1,\ldots,d_k\}=\{1,\ldots,1\}$ as multisets.

Hence, it follows that

$$\begin{aligned}\frac{A_{\mathcal{C}_\lambda}(t)}{|\mathcal{C}_\lambda|} &= \frac{A_n(t)}{n!} + \frac{1-t}{n}\frac{A_{n-1}(t)}{(n-1)!}\left(\binom{m_1}{2} - m_2\right) \\ &\quad + \frac{(1-t)^2}{n(n-1)}\frac{A_{n-2}(t)}{(n-2)!}\left(\frac{3m_1-1}{4}\binom{m_1}{3} + 3\binom{m_2}{2} - m_2\binom{m_1}{2} - m_3 - m_4\right) \\ &\quad + (1-t)^3 g(t)\end{aligned}$$

for some polynomial $g(t)$, from which we can calculate the asymptotic variance of descent numbers to be $(1-4\alpha^3+3\alpha^4)\,\frac{n}{12}$. □

4.1.4 Central Limit theorem for descents in conjugacy classes of S_n

Write D_λ for the descent number $d(\pi)$ of a permutation π which is uniformly chosen from the conjugacy class $\mathcal{C}_\lambda$ of S_n. Let us define the normalized random variable W_λ by

$$D_\lambda = \frac{n+1}{2} - \frac{m_1^2}{2n} + \sqrt{n}W_\lambda,$$

and denote by $M_\lambda(s) = \mathbb{E}[e^{sW_\lambda}]$ the MGF of W_λ.

Since we now know the asymptotic mean and variance, we expect that the distribution of W_λ is asymptotically the normal distribution with the zero mean and the variance $\frac{1}{12}(1-4\alpha^3+3\alpha^4)$ along the limit $m_1/n \to \alpha$. This is equivalent to showing that M_λ converges pointwise to the MGF of this normal distribution as $m_1/n \to \alpha$. The following result, Theorem 4.1.5, concerns a uniform estimate on M_λ that serves this purpose.

Theorem 1.5. *Write $\alpha_\lambda = m_1/n$. Then, there exists a function $C : \mathbb{R} \to (0,\infty)$ such that*

$$\left| M_\lambda(s) - \exp\left\{\frac{s^2}{24}\left(1-4\alpha_\lambda^3+3\alpha_\lambda^4\right)\right\}\right| \leq C(s)\frac{\log^3 n}{\sqrt{n}}$$

for any $n \geq 1$ and for any $\lambda \vdash n$.

This section is aimed at proving this theorem. First, we develop a series representation of M_λ along with some preliminary estimates on its coefficients. This reduces the main claim to Proposition 4.1.2, which is then proved.

Series representation of M_λ In [63], estimating the coefficients of (4.1.2) played an important role in the computation. Likewise, we need to provide an estimation on the coefficients of (4.1.3). Let $F_{i,a} = if_{i,a} = \sum_{i|d} \mu(i) a^{d/i}$. We first prove the following lemma about $F_{i,a}$ and $F_{i,-a}$.

Lemma 4.1.1. *Let a and i be positive integers. Then,*

(1) $(-1)^i F_{i,-a} = F_{i,a} + 2F_{\frac{i}{2},a} \mathbf{1}_{\{ord_2(i)=1\}}$, *where $ord_2(i)$ is the largest integer k such that 2^k divides i.*

(2) (upper bound) $0 \le F_{i,a} \le a^i$ *and* $0 \le (-1)^i F_{i,-a} \le a^i + 2a^{\frac{i}{2}}$, *and*

(3) (lower bound) $(-1)^i F_{i,-a} \ge F_{i,a} \ge a^{\frac{i}{2}} \left(a^{\frac{i}{2}} - \frac{i}{2} \right)$.

Proof of Lemma 4.1.1. Part (1) is proven by looking at the definition of $f_{i,a}$. Let us write $i = 2^k q$, where k is a positive integer and q is an odd integer. Then, by the multiplicity of μ, we have

$$F_{i,a} = \sum_{j=0}^{k} \sum_{d|q} \mu(d) \mu\left(2^j\right) a^{2^{k-j} \frac{q}{d}},$$

and

$$F_{i,-a} = \sum_{j=0}^{k} \sum_{d|q} \mu(d) \mu\left(2^j\right) (-a)^{2^{k-j} \frac{q}{d}}.$$

We divide the computation into three cases.

(1) If $k = 0$, $i = q$ is odd, and so,

$$(-1)^i F_{i,-a} = -\sum_{d|q} \mu(d)(-a)^{\frac{q}{d}} = \sum_{d|q} \mu(d) a^{\frac{q}{d}} = F_{i,a}.$$

(2) If $k = 1$, both the terms for $j = 0, 1$ may survive, and

$$\begin{aligned}(-1)^i F_{i,-a} &= \sum_{d|q} \mu(d)(-a)^{\frac{2q}{d}} + \sum_{d|q} \mu(2)\mu(d)(-a)^{\frac{q}{d}} \\ &= \sum_{d|q} \mu(d) a^{\frac{2q}{d}} + \sum_{d|q} \mu(d) a^{\frac{q}{d}} \\ &= \sum_{d|2q} \mu(d) a^{\frac{2q}{d}} + 2\sum_{d|q} \mu(d) a^{\frac{q}{d}} \\ &= F_{i,a} + 2F_{\frac{i}{2},a}.\end{aligned}$$

(3) If $k \geq 2$, we have $(-1)^i = 1$ and $(-a)^{2^{k-j}\frac{q}{d}} = a^{2^{k-j}\frac{q}{d}}$ for $j = 0, 1$ and $d \mid q$. Hence, by comparing the formula for $F_{i,a}$ and $(-1)^i F_{i,-a}$, we see that they coincide.

For part (2), we note that $f_{i,a}$ counts certain types of words, and so, $F_{i,a} = i f_{i,a} \geq 0$. By using Möbius inversion formula, we see that $F_{i,a} \leq \sum_{d|i} F_{i,a} = a^i$. The second inequality follows from part (1) and the first inequality.

For part (3), note that, by parts (1) and (2), we have $(-1)^i F_{i,-a} \geq F_{i,a}$. The other half of the inequality follows by noting that

$$F_{i,a} \geq a^i - \sum_{d|i, d \neq i} a^d \geq a^i - \frac{i}{2} a^{\frac{i}{2}},$$

and so, the lemma is proven. □

In order to utilize both representations (4.1.2) and (4.1.3) simultaneously, we introduce some auxiliary notation as follows. Given a partition $\lambda \vdash n$ and a non-zero real number s, define

$$K_a = \prod_{i=1}^{n} K_a^{(i)}, \qquad K_a^{(i)} = \begin{cases} \prod_{k=0}^{m_i - 1} (F_{i,a} + ik), & \text{if } s < 0 \\ \prod_{k=0}^{m_i - 1} (-1)^i (F_{i,-a} + ik), & \text{if } s > 0 \end{cases}$$

for $1 \leq i \leq n$ and $a \geq 1$. Strictly speaking, both K_a and $K_a^{(i)}$ depend on both s and λ as well. Since s and λ are assumed to be given throughout the computation, however, we suppress them from the notation. Then by (4.1.2) and (4.1.3), we obtain the following concise formula

$$\mathbb{E}[e^{sD_\lambda}] = \frac{A_{\mathcal{C}_\lambda}(e^s)}{|\mathcal{C}_\lambda|} = \left(\frac{e^s - 1}{s} \right)^{n+1} \frac{|s|^{n+1}}{n!} \sum_{a=1}^{\infty} K_a e^{-|s|a}.$$

From this, we find that M_λ is given by

$$\begin{aligned} M_\lambda(s) &= \mathbb{E} \exp \left\{ \frac{s}{\sqrt{n}} \left(D_\lambda - \frac{n+1}{2} + \frac{m_1^2}{2n} \right) \right\} \\ &= \left(\frac{\sinh\left(\frac{s}{2\sqrt{n}}\right)}{\frac{s}{2\sqrt{n}}} \right)^{n+1} \frac{(|s|/\sqrt{n})^{n+1}}{n!} \sum_{a=1}^{\infty} K_a \exp\left\{ -\frac{|s|}{\sqrt{n}} a + \frac{m_1^2 s}{2n^{3/2}} \right\}. \end{aligned}$$

For the sake of simplicity, let us denote

$$L_a = \frac{\left(|s|/\sqrt{n}\right)^{n+1}}{n!} K_a \exp\left\{-\frac{|s|}{\sqrt{n}}a + \frac{m_1^2 s}{2n^{3/2}}\right\}.$$

Note that, for s fixed and $n \to \infty$,

$$\left(\frac{\sinh\left(\frac{s}{2\sqrt{n}}\right)}{\frac{s}{2\sqrt{n}}}\right)^{n+1} = \left(1 + \frac{s^2}{24n} + \mathcal{O}\left(\frac{1}{n^2}\right)\right)^{n+1} = e^{\frac{s^2}{24}} + \mathcal{O}\left(\frac{1}{n}\right),$$

where the implicit bounds depend only on s. In light of this, we have only to prove the following proposition.

Proposition 4.1.2. *Write $\alpha_\lambda = m_1/n$. Then, there exists a function $C : \mathbb{R} \to (0, \infty)$ such that*

$$\left|\sum_{a=1}^{\infty} L_a - \exp\left\{\frac{s^2}{24}\left(-4\alpha_\lambda^3 + 3\alpha_\lambda^4\right)\right\}\right| \leq C(s)\frac{\log^3 n}{\sqrt{n}}.$$

Proof of Proposition 4.1.2 We inspect the sum over two ranges – the small range, where $a \leq \varepsilon n^{3/2}$, and the large range, where $a > \varepsilon n^{3/2}$. Here, ε is a positive real number chosen to satisfy $4\varepsilon e|s| < 1$. This choice will be explained shortly later, but it is important to note that ε depends only on s.

When invoking asymptotic notation $\mathcal{O}(\cdot)$, it is always assumed that implicit bounds depend only possibly on s. This way, we can keep track of uniform estimates. Likewise, we indulge in luxury of changing the meaning of the generic function $C = C(s)$ from line to line, as its exact values are not important to the argument.

Estimation of the small range. If $a \leq \varepsilon n^{3/2}$, we have

$$|K_a| \leq \prod_{i=1}^{n} \prod_{k=0}^{m_i - 1} \left(a^i + 2a^{\frac{i}{2}} + ik\right) \leq \prod_{i=1}^{n} \prod_{k=0}^{m_i - 1} \left(3\varepsilon n^{\frac{3}{2}} + n\right) \leq \max\left\{4n, 4\varepsilon n^{\frac{3}{2}}\right\}^n,$$

where the last inequality follows from bounding each factor $\left(3\varepsilon n^{\frac{3}{2}} + n\right)$ by 4 times the bigger of n and $\varepsilon n^{3/2}$. This induces the following upper bound of L_a.

$$|L_a| \leq \frac{\left(|s|/\sqrt{n}\right)^{n+1}}{n!} \max\left\{4n, 4\varepsilon n^{\frac{3}{2}}\right\}^n e^{\sqrt{n}|s|/2}.$$

Taking a union bound, it follows that

$$\sum_{a\leq \varepsilon n^{3/2}} L_a \leq \varepsilon n^{\frac{3}{2}} \left(\max_{a\leq \varepsilon n^{3/2}} |L_a| \right) \leq \frac{\varepsilon |s| n}{n!} \max \left\{ 4|s|\sqrt{n}, 4\varepsilon |s| n \right\}^n e^{\sqrt{n}|s|/2}.$$

In view of the Stirling's approximation $n! \sim \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}$, this bound decays to 0 at least as exponentially fast as $n \to \infty$ by our choice of ε.

Estimation of the large range. Through this section, we assume that $a > \varepsilon n^{\frac{3}{2}}$. In this range, we first check that $K_a^{(i)}$, for $2 \leq i \leq n$, behaves almost the same as a^{im_i}. More precisely, fix $N_1 = N_1(s)$ so that $\frac{1}{2\varepsilon n^{1/2}} + \frac{1}{\varepsilon^2 n^2} < \frac{1}{2}$ for all $n \geq N_1$, which we assume hereafter. Then, by Lemma 4.1.1, we find that

$$0 < \prod_{k=0}^{m_i - 1} \left(1 - \frac{i}{2a^{i/2}} - \frac{ik}{a^i} \right) \leq \frac{K_a^{(i)}}{a^{im_i}} \leq \prod_{k=0}^{m_i - 1} \left(1 + \frac{2}{a^{i/2}} + \frac{ik}{a^i} \right)$$

for all $2 \leq i \leq n$. Applying the estimate $\log(1+x) = \mathcal{O}(x)$ for $|x| \leq \frac{1}{2}$, we have

$$\begin{aligned} \log \left(\frac{1}{a^{n-m_1}} \prod_{i=2}^{n} K_a^{(i)} \right) &\leq C \sum_{i=2}^{n} \sum_{k=0}^{m_i - 1} \left(\frac{i}{a^{i/2}} + \frac{ik}{a^i} \right) \leq C \sum_{i=2}^{n} \sum_{k=0}^{m_i - 1} \left(\frac{i}{a} + \frac{ik}{a^2} \right) \\ &\leq C \left(\frac{n}{a} + \frac{n^2}{a^2} \right) \leq \frac{C}{\sqrt{n}}, \end{aligned}$$

As for $K_a^{(1)}$, we need to consider both $s > 0$ and $s < 0$ cases. If $s < 0$, then by the expansion $1 + x = \exp\{\log(1+x)\} = \exp\left\{ x - \frac{1}{2}x^2 + \mathcal{O}(x^3) \right\}$ we get

$$\begin{aligned} a^{-m_1} K_a^{(1)} &= \prod_{k=0}^{m_1 - 1} \left(1 + \frac{k}{a} \right) = \exp \left\{ \sum_{k=0}^{m_1 - 1} \left(\frac{k}{a} - \frac{k^2}{2a^2} + \mathcal{O}\left(\frac{k^3}{a^3} \right) \right) \right\} \\ &= \exp \left\{ \frac{m_1^2}{2a} - \frac{m_1^3}{6a^2} + \mathcal{O}\left(\frac{1}{\sqrt{n}} \right) \right\}. \end{aligned}$$

Similarly, for $s > 0$, we get

$$a^{-m_1} K_a^{(1)} = \prod_{k=0}^{m_1 - 1} \left(1 - \frac{k}{a} \right) = \exp \left\{ -\frac{m_1^2}{2a} - \frac{m_1^3}{6a^2} + \mathcal{O}\left(\frac{1}{\sqrt{n}} \right) \right\}.$$

Combining the results, we see that, for $s \neq 0$,

$$K_a = a^n \exp \left\{ -\frac{m_1^2}{2a} \operatorname{sign}(s) - \frac{m_1^3}{6a^2} + \mathcal{O}\left(\frac{1}{\sqrt{n}} \right) \right\},$$

where the implicit constant in $\mathcal{O}\left(\frac{1}{n}\right)$ depends only on ε. From this, it easily follows that, for $x \in [a, a+1]$,

$$K_a \exp\left\{-\frac{|s|}{\sqrt{n}}a\right\} = \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right) x^n \exp\left\{-\frac{m_1^2}{2x}\operatorname{sign}(s) - \frac{m_1^3}{6x^2} - \frac{|s|}{\sqrt{n}}x\right\}$$

and hence,

$$\sum_{a > \varepsilon n^{3/2}} L_a = \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right) \exp\left\{\frac{m_1^2 s}{2n^{3/2}}\right\} \frac{(|s|/\sqrt{n})^{n+1}}{n!}$$
$$\times \int_{\varepsilon n^{3/2}}^{\infty} x^n \exp\left\{-\frac{m_1^2}{2x}\operatorname{sign}(s) - \frac{m_1^3}{6x^2} - \frac{|s|}{\sqrt{n}}x\right\} dx.$$

Applying the substitution $y = \frac{|s|}{\sqrt{n}}x$, followed by $y = n + \sqrt{n}z$, the above integral becomes

$$\begin{aligned}
&\exp\left\{\frac{m_1^2 s}{2n^{3/2}}\right\} \frac{(|s|/\sqrt{n})^{n+1}}{n!} \int_{\varepsilon n^{3/2}}^{\infty} x^n \exp\left\{-\frac{m_1^2}{2x}\operatorname{sign}(s) - \frac{m_1^3}{6x^2} - \frac{|s|}{\sqrt{n}}x\right\} dx \\
&= \exp\left\{\frac{m_1^2 s}{2n^{3/2}}\right\} \frac{1}{n!} \int_{\varepsilon |s| n}^{\infty} y^n \exp\left\{-\frac{m_1^2 s}{2\sqrt{n}y} - \frac{m_1^3 s^2}{6ny^2} - y\right\} dy \\
&= \frac{n^{n+\frac{1}{2}} e^{-n}}{n!} \int_{\mathbb{R}} g_n(z) \exp\left\{\frac{\alpha_\lambda^2 s}{2(1+\frac{z}{\sqrt{n}})} z - \frac{\alpha_\lambda^3 s^2}{6(1+\frac{z}{\sqrt{n}})^2}\right\} dz,
\end{aligned}$$

where $\alpha_\lambda = m_1/n$ and $g_n : \mathbb{R} \to \mathbb{R}$ is defined by

$$g_n(z) = \begin{cases} \left(1 + \frac{z}{\sqrt{n}}\right)^n e^{-\sqrt{n}z}, & \text{if } z > -(1 - \varepsilon|s|)\sqrt{n} \\ 0, & \text{otherwise.} \end{cases}$$

Now we aim at estimating the last integral. First, by the Stirling's approximation we have $\frac{n^{n+\frac{1}{2}} e^{-n}}{n!} = \frac{1}{\sqrt{2\pi}} + \mathcal{O}\left(\frac{1}{n}\right)$. Next, we claim the following lemma.

Lemma 4.1.3. *Let $N \geq 1$ be arbitrary. Then, for any $n \geq N$, we have*

$$g_n(z) \leq \begin{cases} g_N(z), & \text{if } z \geq 0 \\ e^{-\frac{1}{2}z^2}, & \text{if } z < 0 \end{cases}.$$

Proof of Lemma 4.1.3. Consider the function $h(n,z) = \log\left(\left(1 + \frac{z}{\sqrt{n}}\right)^n e^{-\sqrt{n}z}\right)$ on $z > -\sqrt{n}$. By direct computation, we find that

$$\frac{\partial^2 h}{\partial n^2} = \frac{z^3}{4n^{3/2}\left(z + \sqrt{n}\right)^2} \qquad \text{and} \qquad \lim_{n\to\infty} \frac{\partial h}{\partial n} = 0.$$

So, it follows that $\frac{\partial h}{\partial n} \le 0$ if $z \ge 0$, and $\frac{\partial h}{\partial n} \ge 0$ if $z < 0$. Hence, for $n \ge N$ and $z \ge 0$, we obtain $g_n(z) = \exp\{h(n,z)\} \le \exp\{h(N,z)\} = g_N(z)$. Similarly, when $z < 0$ we have $g_n(z) \le \lim_{n' \to \infty} \exp\{h(n', z)\} = e^{-\frac{1}{2}z^2}$. □

Now, pick $N_2 = N_2(s)$ so that $N_2 \ge \max\{N_1, s^2\}$ (recall that we introduced N_1 at the beginning of the estimation in the large range.) Writing $\tilde{g}_n$ for the integrand

$$\tilde{g}_n(z) = g_n(z) \exp\left\{ \frac{\alpha_\lambda^2 s}{2(1 + \frac{z}{\sqrt{n}})} z - \frac{\alpha_\lambda^3 s^2}{6(1 + \frac{z}{\sqrt{n}})^2} \right\},$$

the above Lemma 4.1.3 provides the following bound

$$g_n(z) \le \begin{cases} g_{N_2}(z) e^{\frac{|s|}{2} z} & \text{if } z \ge 0 \\ e^{-\frac{1}{2}z^2 + \frac{|s|}{2} z} & \text{if } z < 0 \end{cases}$$

for all $z \in \mathbb{R}$ and for all $n \ge N_2$. The specific detail of this bound is not important, however, and we only need to note that this decays exponentially fast. To be precise, there exist constants $C > 0$ and $c > 0$, which depend only on s, such that

$$\max\left\{ g_{N_2}(|z|) e^{\frac{|s|}{2}|z|}, e^{-\frac{1}{2}z^2 + \frac{|s|}{2}|z|} \right\} \le C e^{-c|z|}$$

for all $z \in \mathbb{R}$. Now, to estimate the integral of $\tilde{g}_n$, we split this into two parts

$$\int_{\mathbb{R}} \tilde{g}_n(z) = \int_{|z| \le \frac{\log n}{2c}} \tilde{g}_n(z)\, dz + \int_{|z| > \frac{\log n}{2c}} \tilde{g}_n(z)\, dz.$$

The latter integral is easily estimated by direct computation.

$$\int_{|z| > \frac{\log n}{2c}} \tilde{g}_n(z)\, dz \le 2 \int_{\frac{\log n}{2c}}^{\infty} C e^{-cz}\, dz = \frac{2C}{c\sqrt{n}}.$$

For the first integral, we have

$$\begin{aligned} \int_{|z| \le \frac{\log n}{2c}} \tilde{g}_n(z)\, dz &= \int_{|z| \le \frac{\log n}{2c}} \exp\left\{ -\frac{1}{2}z^2 + \frac{\alpha_\lambda^2 s}{2} z - \frac{\alpha_\lambda^3 s^2}{6} + \mathcal{O}\left(\frac{\log^3 n}{\sqrt{n}} \right) \right\} dz \\ &= \left(1 + \mathcal{O}\left(\frac{\log^3 n}{\sqrt{n}} \right) \right) \int_{\mathbb{R}} \exp\left\{ -\frac{1}{2}z^2 + \frac{\alpha_\lambda^2 s}{2} z - \frac{\alpha_\lambda^3 s^2}{6} \right\} dz \\ &\quad + \mathcal{O}\left(\int_{|z| > \frac{\log n}{2c}} C e^{-c|z|}\, dz \right) \\ &= \sqrt{2\pi} \exp\left\{ \frac{s^2}{24} \left(-4\alpha_\lambda^3 + 3\alpha_\lambda^4 \right) \right\} + \mathcal{O}\left(\frac{\log^3 n}{\sqrt{n}} \right). \end{aligned}$$

Combining altogether, we obtain, for $n \geq N_2$,

$$\sum_{a > \varepsilon n^{3/2}} L_a = \exp\left\{ \frac{s^2}{24} \left(-4\alpha_\lambda^3 + 3\alpha_\lambda^4 \right) \right\} + \mathcal{O}\left(\frac{\log^3 n}{\sqrt{n}} \right).$$

Together with the exponential decay of $\sum_{a \leq \varepsilon n^{3/2}} L_a$ proved in the previous section, the desired proposition follows by revealing the implicit bound C and then making it larger, if needed, so that the inequality is also true for $n < N_2$.

This concludes the proof of Proposition 4.1.2 and, in turn, Theorem 4.1.5. Combining Theorem 4.1.5 and Theorem 4.1.6 yields the desired central limit theorem, Theorem 4.1.4.

4.2 A central limit theorem for peaks of a random permutation in a fixed conjugacy class of S_n

The number of peaks of a random permutation is known to be asymptotically normal. We give a new proof of this and prove a central limit theorem for the distribution of peaks in a fixed conjugacy class of the symmetric group. Our technique is to apply "analytic combinatorics" to study a complicated but exact generating function for peaks in a given conjugacy class.

4.2.1 Introduction

We say that a permutation on n symbols has a descent at position i if $\pi(i) > \pi(i+1)$, and we let $d(\pi)$ denote the number of descents of π. For example. the permutation $1\underline{4}32\underline{6}5$ has descents at positions 2 and 5, and has $d(\pi) = 2$. Descents appear in numerous parts of mathematics. For examples, see Knuth [66] for connections of descents with the theory of sorting and the theory of runs in permutations and see Bayer and Diaconis [10] for applications of descents to card shuffling. The number $A(n, k)$ of permutations on n symbols with k descents is called an Eulerian number, and there is an entire book devoted to their study [80].

It is well known that the distribution of descents $d(\pi)$ in S_n is asymptotically normal with mean $(n-1)/2$ and variance $(n+1)/12$, in the sense that $d(\pi)$ has the prescribed mean and variance and its normalization $(d(\pi)-\mathbb{E}[d(\pi)])/\sqrt{\mathrm{Var}(d(\pi))}$ converges in distribution to the standard normal distribution. There are many proofs of this:

(1) Pitman [81] uses real-rootedness of the Eulerian polynomials

$$A_n(t) = \sum_{\pi \in S_n} t^{d(\pi)+1}$$

(2) David and Barton [29] use the method of moments.

(3) Tanny [94] uses the fact that if $U_1, \cdots, U_n$ are independent uniform $[0,1]$ random variables, the for all integers k,

$$\mathbb{P}\left(k \leq \sum_{i=1}^{n} U_i < k+1\right) = A(n,k)/n!$$

(4) Fulman [49] uses Stein's method.

There is also interesting literature on the joint distribution of descents and cycles. Gessel and Reutenauer [53] use symmetric function theory to enumerate permutations with a given cycle structure and descent set, and Diaconis, McGrath, and Pitman [37] interpret this in the context of card shuffling. We regard these exact results as a miracle, and they enable one to write down an exact (but quite complicated) generating function for descents of permutations in a given conjugacy class. These exact generating functions make it possible to prove central limit theorems for the number of descents in fixed conjugacy classes of the symmetric group. Fulman [48] proved a central limit theorem when the conjugacy classes consist of large cycles. Almost twenty years later, Kim [63] proved a central limit for descents in random fixed point free involutions. Quite recently, Kim and Lee [64] proved a central limit theorem for arbitrary conjugacy classes. These results would be very difficult to obtain without exact generating functions.

Given the above discussion, it is natural to ask if there are other permutation statistics for which there is exact information about the joint distribution with cycle structure. In their work on casino shuffling machines, Diaconis, Fulman, and Holmes [34] discovered that there is a lovely exact generating function for the number of peaks of a permutation enumerated according to cycle structure. Let us describe their result. We say that a permutation $\pi \in S_n$ has a peak at position $1 < i < n-1$ if $\pi(i-1) < \pi(i) > \pi(i+1)$, and let $p(\pi)$ be the number of peaks of π. Thus $\pi = 1\underline{4}26\underline{7}53$ has peaks at positions 2 and 5, so that $p(\pi) = 2$. Letting λ be a partition of n with n_i parts of size i, Corollary 3.8 of [34] gives that

$$\sum_{\pi \in \mathcal{C}_\lambda} \left(\frac{4t}{(1+t)^2} \right)^{p(\pi)+1} = 2 \left(\frac{1-t}{1+t} \right)^{n+1} \sum_{a \geq 1} t^a \prod_i [x_i^{n_i}] \left(\frac{1+x_i}{1-x_i} \right)^{f_{a,i}}. \tag{4.2.1}$$

Here, $\mathcal{C}_\lambda$ denotes the elements of S_n of cycle type λ, and $[x_i^{n_i}]g(x_i)$ denotes the coefficient of $x_i^{n_i}$ in the function $g(x_i)$, and

$$f_{a,i} = \frac{1}{2i} \sum_{\substack{d|i \\ d \text{ odd}}} \mu(d)(2a)^{i/d},$$

where μ is the Möbius function of elementary number theory. (The result of [34] actually deals with valleys rather than peaks, but the joint generating function with cycle structure is the same as can be seen by conjugating by the longest permutation $n \cdots 21$). The reader will agree that the generating function (4.2.1) looks hard to deal with (it need not be real-rooted), and our main insight is that we can adapt the methods of Kim and Lee [64] to analyze it.

To close the introduction, we mention that the number of peaks of a permutation is a feature of interest. The paper [34] uses peaks to analyze casino shelf-shuffling machines. The number of peaks is classically used as a test of randomness for time series; see Warren and Seneta [99] and their references, which also include a central limit theorem for the number of peaks for a uniform random permutation. Permutations with no peaks are called unimodal (usually unimodal refers to no valleys but these are equivalent for our purposes), and are of interest in social choice theory through Coombs's "unfolding hypothesis" (see Chapter 6 of [33]). They also appear in dynamical systems and magic tricks (see Chapter 5 of [35]).

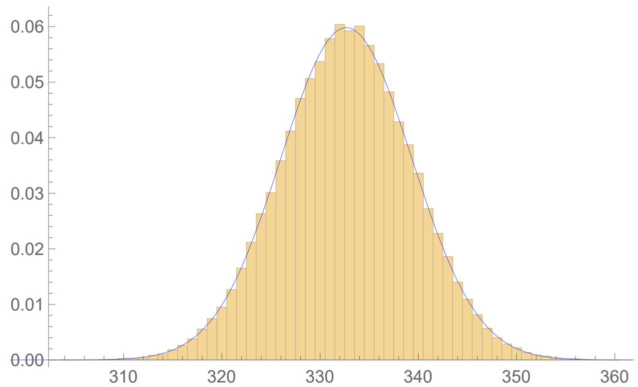


Figure 4.1: Histogram of peaks of 10^5 samples drawn from $\mathcal{C}_{2^{250}4^{125}} \subset S_{1000}$.

Finally, we note that peaks have been widely studied by combinatorialists; see Petersen [79], Stembridge [92], Nyman [73], Schocker [88] and a paper of Billey, Burdzy, and Sagan [13], for a small sample of combinatorial work on peaks.

Main results To motivate the readers, we first demonstrate a numerical simulation result. Figure 4.1 is a histogram of peaks of 10^5 permutations drawn from the conjugacy class $\mathcal{C}_{2^{250}4^{125}} \subset S_{1000}$. The histogram suggests that the peaks of permutations in $\mathcal{C}_{2^{250}4^{125}}$ are normally distributed, and indeed, the p.d.f. of $\mathcal{N}\left(\frac{n-2}{3}, \frac{2(n+1)}{45}\right)$ with $n = 1000$ fits very well. This suggests that the behavior of peaks for a particular conjugacy class is mostly the same as that of peaks for S_n. This does turn out to be true for conjugacy classes with no fixed points, as the following main theorem states that the asymptotic distribution of peaks in conjugacy classes is normal, where the asymptotic mean and variance depend only on the density of fixed points.

Theorem 4.2.1. *Let C_n be a conjugacy class of S_n for each $n \geq 1$. Denote by $\alpha_1(C_n)$ the fraction of fixed points of each element of C_n. Suppose that π_n is chosen uniformly at random from C_n and that $\alpha_1(C_n)$ converges to some $\alpha \in [0,1]$ as $n \to \infty$. Then, as $n \to \infty$,*

$$\frac{p(\pi_n) - \frac{1-\alpha_1(C_n)^3}{3}n}{\sqrt{n}} \quad \textit{converges in distribution to} \quad \mathcal{N}\left(0, \frac{2}{45} + \frac{1}{9}\alpha^3 - \frac{3}{5}\alpha^5 + \frac{4}{9}\alpha^6\right).$$

Our main strategy is to adopt the modified Curtiss' theorem from [64], which relates convergence in distribution of random variables to the pointwise convergence of their moment generating functions on an open set. In this regard, the main theorem is a direct consequence of the following technical theorem:

Theorem 4.2.2. *For each $s > 0$, there exists a universal constant $C = C(s) > 0$, depending only on s, such that the following is true: Let $\mathcal{C}_\lambda \subseteq S_n$ be the conjugacy class of cycle type $\lambda = 1^{n_1}2^{n_2}\cdots$ and π be chosen uniformly at random from $\mathcal{C}_\lambda$. Denote by $\alpha_1 = n_1/n$ the density of fixed points. Then,*

$$\mathbb{E}\big[e^{-sp(\pi)/\sqrt{n}}\big] = \exp\left\{-\frac{1-\alpha_1^3}{3}s\sqrt{n} + \left(\frac{1}{45} + \frac{\alpha_1^3}{18} - \frac{3\alpha_1^5}{10} + \frac{2\alpha_1^6}{9}\right)s^2 + E_{\lambda,s}\right\},$$

where $|E_{\lambda,s}| \leq Cn^{-1/4}$.

This theorem is interesting in its own right, because the uniform estimate allows us to readily extend the scope of the main theorem to a more general class of sequences (C_n). More precisely, the statement of Theorem 4.2.1 readily extends to the case where each C_n is simply a conjugacy-invariant subset of S_n such that every element of C_n has the same number of fixed points. For example, if we consider the set of all elements of S_n with zero fixed points, we would obtain a central limit theorem for peaks of derangements.

4.2.2 Central limit theorem for peaks of a random permutation in S_n

Denoting the peak generating function by

$$W_n(t) = \sum_{\pi \in S_n} t^{p(\pi)+1},$$

it is well known [92, p779] that $A_n(t)$ and $W_n(t)$ are related by the identity

$$W_n\left(\frac{4t}{(1+t)^2}\right) = \left(\frac{2}{1+t}\right)^{n+1} A_n(t). \tag{4.2.2}$$

Our aim in this section is to identify the asymptotic distribution of peaks of a random permutation in S_n using (4.2.2).

Computing mean and variance of peaks in S_n We begin by calculating the derivatives of $A_n(t)$ at 1 up to the fourth order.

Lemma 4.2.3. *We have*

$$\begin{aligned}
A_n^{(0)}(1) &= n!, \\
A_n^{(1)}(1) &= n! \cdot \frac{n+1}{2}\,\mathbf{1}_{\{n\geq 1\}}, \\
A_n^{(2)}(1) &= n! \cdot \frac{3n^2+n-2}{12}\,\mathbf{1}_{\{n\geq 2\}}, \\
A_n^{(3)}(1) &= n! \cdot \frac{n^3-2n^2-n+2}{8}\,\mathbf{1}_{\{n\geq 3\}}, \text{ and} \\
A_n^{(4)}(1) &= n! \cdot \frac{15n^4-90n^3+125n^2+78n-152}{240}\,\mathbf{1}_{\{n\geq 4\}}.
\end{aligned}$$

Proof. It is well known that the Eulerian polynomials satisfy the identity

$$A_n(t) = (1-t)^{n+1}\sum_{a\geq 1} a^n t^a.$$

Recall that the Stirling numbers of the second kind $\left\{{n \atop k}\right\}$ count the number of partitions of an n-element set into k blocks. Plugging the expansion $a^n = \sum_{k=0}^{n} \left\{{n \atop k}\right\} \frac{a!}{(a-k)!}$ into the expression above, we see that

$$\begin{aligned}
A_n(t) &= (1-t)^{n+1}\sum_{a=0}^{\infty}\left(\sum_{k=0}^{n}\left\{{n \atop k}\right\}\frac{a!}{(a-k)!}\right)t^a \\
&= (1-t)^{n+1}\sum_{k=0}^{n}\left\{{n \atop k}\right\}\frac{k!t^k}{(1-t)^{k+1}} \\
&= \sum_{k=0}^{n} k!\left\{{n \atop k}\right\}t^k(1-t)^{n-k} \\
&= \sum_{k=0}^{n}(n-k)!\left\{{n \atop n-k}\right\}t^{n-k}(1-t)^k.
\end{aligned}$$

Now one can compute $A_n^{(p)}(1)$ by plugging the above identity into $A_n^{(p)}(1) = p![s^p]A_n(1+s)$. More specifically, if $p > n$, then $A_n(1+s)$ has degree n, and so, $A_n^{(p)}(1) = 0$. If $p \leq n$, then

$$\begin{aligned}
A_n^{(p)}(1) = p![s^p]A_n(1+s) &= p![s^p]\sum_{k=0}^{n}(-1)^k(n-k)!\left\{{n \atop n-k}\right\}(1+s)^{n-k}s^k \\
&= p!\sum_{k=0}^{p}(-1)^k(n-k)!\left\{{n \atop n-k}\right\}\binom{n-k}{p-k}.
\end{aligned}$$

For each given p, the last sum can be computed by calculating $\left\{{n \atop n-k}\right\}$'s for $k = 0, \cdots, p$. For instance, $\left\{{n \atop n}\right\} = 1$ and $\left\{{n \atop n-1}\right\} = \binom{n}{2}$, and for larger values of k, they can be systematically computed by utilizing the relationship between the Stirling numbers of the second kind and Eulerian numbers of the second kind (see equation (6.43) of [54]). The $\left\{{n \atop n-k}\right\}$'s relevant to us are

$$\begin{aligned}
\left\{{n \atop n-2}\right\} &= 2\binom{n}{4} + \binom{n+1}{4}, \\
\left\{{n \atop n-3}\right\} &= 6\binom{n}{6} + 8\binom{n+1}{6} + \binom{n+2}{6}, \text{ and} \\
\left\{{n \atop n-4}\right\} &= 24\binom{n}{8} + 58\binom{n+1}{8} + 22\binom{n+2}{8} + \binom{n+3}{8}.
\end{aligned}$$

Plugging these back into the formula for $A_n^{(p)}(1)$ provides the desired lemma. □

Next, (4.2.2) relates $W_n^{(p)}(1)$ to the derivatives of $A_n(t)$ up to order $2p$ evaluated at 1. Differentiating both sides of (4.2.2) gives us

$$-\frac{4(t-1)}{(1+t)^3} W_n'\left(\frac{4t}{(1+t)^2}\right) = -\frac{(n+1)2^{n+1}}{(1+t)^{n+2}} A_n(t) + \frac{2^{n+1}}{(1+t)^{n+1}} A_n'(t),$$

and by multiplying $-\frac{(1+t)^3}{4(t-1)}$ to both sides and simplifying, we see that

$$W_n'\left(\frac{4t}{(1+t)^2}\right) = \left(\frac{2}{1+t}\right)^{n-1} \frac{(n+1)A_n(t) - (t+1)A_n'(t)}{t-1}.$$

This formula cannot be evaluated directly at $t = 1$, but we can use L'Hôpital's rule to get

$$\begin{aligned}
W_n'(1) &= \lim_{t\to 1} W_n'\left(\frac{4t}{(1+t)^2}\right) = \lim_{t\to 1} \frac{(n+1)A_n(t) - (t+1)A_n'(t)}{t-1} \\
&= nA_n'(1) - 2A_n''(1) = n! \cdot \frac{n+1}{3}, \qquad \text{if } n \geq 2.
\end{aligned}$$

The last step is a consequence of Lemma 4.2.3. The second derivative $W_n''(1)$ can be computed in similar fashion. By differentiating both sides of (4.2.2) twice and simplifying, we obtain an identity relating W_n'' to the derivatives of A_n:

$$W_n''\left(\frac{4t}{(1+t)^2}\right) = \left(\frac{2}{1+t}\right)^{n-3} \frac{P_n(t)}{(t-1)^3},$$

where $P_n(t)$ is given by

$$P_n(t) = (n+1)(nt-n+2)A(t) - 2(t+1)(nt-n+1)A'(t) + (t-1)(t+1)^2 A''(t).$$

Similarly as before, we find $W_n''(1)$ by using L'Hôpital's rule:

$$\begin{aligned}
W_n''(1) &= \lim_{t\to 1} W_n''\left(\frac{4t}{(1+t)^2}\right) = \lim_{t\to 1}\frac{P_n(t)}{(t-1)^3} = \frac{P_n^{(3)}(1)}{6} \\
&= \frac{(3n^2-9n+6)A_n^{(2)}(1) - (10n-20)A_n^{(3)}(1) + 8A_n^{(4)}(1)}{6} \\
&= n!\cdot\frac{(5n-8)(n+1)}{45}, \qquad \text{if } n \ge 4,
\end{aligned}$$

where the last step follows from Lemma 4.2.3. Finally, since $W_n'(1) = n!\mathbb{E}[p(\pi)+1]$ and $W_n''(1) = n!\mathbb{E}[(p(\pi)+1)p(\pi)]$, we have

$$\mathbb{E}[p(\pi)] = \frac{W_n'(1)}{n!} - 1 = \frac{n-2}{3} \qquad \text{if } n \ge 2,$$

and

$$\mathrm{Var}(p(\pi)) = \frac{W_n''(1)}{n!} + \frac{W_n'(1)}{n!} - \left(\frac{W_n'(1)}{n!}\right)^2 = \frac{2(n+1)}{45} \qquad \text{if } n \ge 4.$$

At this point, it is worth noting (4.2.2) implies that, like $A_n(t)$, $W_n(t)$ has only real roots, and so, by Harper's method [56], we can obtain a central limit theorem for peaks of a random permutation in S_n. In the upcoming section, we give a new proof of this central limit theorem by using analytic combinatorics and will go further to prove a central limit theorem for peaks in arbitrary conjugacy classes of S_n, where the mean and variance depend only on the density of fixed points in the conjugacy classes.

Establishing the asymptotic normality of peaks in S_n Kim and Lee [64] proved the following modification of Curtiss' theorem:

Theorem 4.2.4. *Let X_n be random vectors in $\mathbb{R}^d$ for each $n \in \mathbb{B} \cup \{\infty\}$ and $M_{X_n}(s) = \mathbb{E}[e^{sX_n}]$ be the moment generating function (m.g.f.) of X_n. Suppose that there is a non-empty open subset $U \subseteq \mathbb{R}^d$ such that $\lim_{n\to\infty} M_{X_n}(s) = M_{X_\infty}(s)$ for all $s \in U$. Then, X_n converges in distribution to X_∞.*

This theorem will be used in this subsection to prove a central limit theorem about peaks of permutations chosen, uniformly at random, from S_n, and in Section 4.2.3 to prove an analogous theorem about peaks of permutations chosen, uniformly at random, from arbitrary conjugacy classes, where the asymptotic mean and variance are functions of only α, the density of fixed points in the conjugacy classes.

Theorem 4.2.5. *Let π_n be chosen uniformly at random from S_n. Then $p(\pi_n)$ is asymptotically normal with mean $\frac{n-2}{3}$ and variance $\frac{2(n+1)}{45}$. More precisely, $p(\pi_n)$ has the prescribed mean and variance, and as $n \to \infty$,*

$$\frac{p(\pi_n) - \frac{n-2}{3}}{\sqrt{n}} \quad \text{converges in distribution to} \quad \mathcal{N}\left(0, \tfrac{2}{45}\right).$$

Proof. Let $X_n = \left(p(\pi_n) - \frac{n-2}{3}\right)/\sqrt{n}$ denote the normalized peaks. In view of Theorem 4.2.4, it suffices to show that $M_{X_n}(s)$ converges pointwise to the m.g.f. of $\mathcal{N}\left(0, \frac{2}{45}\right)$ on some open interval. Let $0 < t < 1$. By a simple comparison, it follows that

$$t \cdot \frac{n!}{\log^{n+1}(1/t)} = \int_0^\infty a^n t^{a+1}\,\mathrm{d}a \leq \sum_{a\geq 1} a^n t^a \leq \int_0^\infty a^n t^{a-1}\,\mathrm{d}a = \frac{1}{t}\cdot\frac{n!}{\log^{n+1}(1/t)}.$$

Plugging this into (4.2.2), we obtain

$$\frac{1}{n!}W_n\left(\frac{4t}{(1+t)^2}\right) = \frac{1}{n!}\left(\frac{2(1-t)}{1+t}\right)^{n+1}\left(\sum_{a\geq 1} a^n t^a\right) = e^{\mathcal{O}(\log t)}\left(\frac{2(1-t)}{(1+t)\log(1/t)}\right)^{n+1}.$$

Now, fix $s > 0$ and choose t as the unique solution of $\frac{4t}{(1+t)^2} = e^{-s/\sqrt{n}}$ in the range $(0,1)$, which is given by

$$t = \frac{1-\sqrt{1-e^{-s/\sqrt{n}}}}{1+\sqrt{1-e^{-s/\sqrt{n}}}} = 1 - \frac{2s^{1/2}}{n^{1/4}} + \frac{2s}{n^{1/2}} - \frac{3s^{3/2}}{2n^{3/4}} + \frac{s^2}{n} + \mathcal{O}\left(n^{-5/4}\right), \tag{4.2.3}$$

where the implicit bound of the error term depends only on s. From this expansion, we have both $\log(t) = \mathcal{O}\left(n^{-1/4}\right)$ and $\log\left(\frac{2(1-t)}{(1+t)\log(1/t)}\right) = -\frac{s}{3\sqrt{n}} + \frac{s^2}{45n} + \mathcal{O}(n^{-5/4})$. Plugging these into $M_{X_n}(s)$, we see that

$$M_{X_n}(s) = \frac{1}{n!}W_n\left(e^{-s/\sqrt{n}}\right)e^{\frac{n+1}{3\sqrt{n}}s} = e^{\frac{s^2}{45}+\mathcal{O}\left(n^{-1/4}\right)}.$$

The desired conclusion follows since $e^{s^2/45}$ is the m.g.f. of the $\mathcal{N}\left(0, \frac{2}{45}\right)$. □

4.2.3 Central limit theorem for peaks of a random permutation in a fixed conjugacy class of S_n

Let $\mathcal{C}_\lambda$ denote the set of all permutations of S_n of cycle type $\lambda = 1^{n_1}2^{n_2}\cdots$ of n. Recall that the peak generating function over $\mathcal{C}_\lambda$ has an explicit formula (4.2.1), which involves the quantity $f_{a,i}$ defined in the introduction. Along the proof of the main theorem, it is important to know a precise estimation of $f_{a,i}$. Define $g_{a,i}$ by

$$g_{a,i} := \frac{2i}{(2a)^i} f_{a,i}.$$

The main reason for introducing $g_{a,i}$ is that $f_{a,i}$ is expected to behave much like $(2a)^i/(2i)$, and so, it is necessary to study the relative difference and produce a precise estimate for the difference. The following lemma serves this purpose.

Lemma 4.2.6. *There exists a universal constant $c_1 > 0$ such that*

$$e^{-c_1(2a)^{-2i/3}} \le g_{a,i} \le e^{c_1(2a)^{-2i/3}}$$

for all $a \ge 1$ and $i \ge 1$. Moreover, we have $e^{-c_1/4a^2} \le g_{a,i} \le e^{c_1/4a^2}$.

Although the intermediate step of the proof will show that the explicit choice $c_1 = 4$ works, we prefer to leave it as a named constant. This is because its value is not important for the argument and its presence will clarify the way we utilize this lemma.

Proof. Recall that $f_{a,i} = \frac{1}{2i}\sum \mu(d)(2a)^{i/d}$, where the sum is over d, the positive odd divisors of i. From this, we see that $g_{a,i} = 1$ when i is either 1 or 2, and so, it suffices to assume that $i \ge 3$. For such $i \ge 3$,

$$(2a)^i\,|g_{a,i} - 1| \le \sum_{\substack{d|i \\ d \text{ odd},\, d\ne 1}} (2a)^{i/d} \le \sum_{k=1}^{\lfloor i/3 \rfloor}(2a)^k = \frac{2a}{2a-1}\big((2a)^{\lfloor i/3\rfloor} - 1\big) \le 2(2a)^{i/3}.$$

Rearranging, it follows that

$$1 - 2(2a)^{-2i/3} \le g_{a,i} \le 1 + 2(2a)^{-2i/3}.$$

Since $a \geq 1$ and $i \geq 3$, we have $2(2a)^{-2i/3} \leq \frac{1}{2}$. Then, applying the inequalities $e^{-2x} \leq 1 - x$ and $1 + x \leq e^{2x}$, which are valid for $0 \leq x \leq \frac{1}{2}$, proves the claim with the choice $c_1 = 4$. The remaining assertion is a simple consequence of the fact that $(2a)^{-2i/3} \leq (2a)^{-2}$ for $i \geq 3$. $\square$

Remark 1. The quantity $f_{a,i}$ is a positive integer. In the special case when a is a power of 2, this follows from Lemma 1.3.16 of [51], which enumerates monic, irreducible, self-conjugate polynomials of degree $2i$ over a finite field of size $2a$.

For general a, the quantity $f_{a,i}$ enumerates what Victor Reiner calls "nowhere-zero primitive twisted necklaces" with values in

$$A = \{+1, -1, +2, -2, \cdots, +a, -a\}$$

having i entries. To define this notion, let the cyclic group C_{2i} act on i-tuples of words $(b_1, \cdots, b_i)$ where the b_k's take values in A, and the generator of C_{2i} acts by

$$g(b_1, \cdots, b_i) = (b_2, \cdots, b_i, -b_1).$$

An orbit P of this action is called a twisted necklace, and P primitive means that the C_{2i} action is free (i.e. no non-trivial group element fixes any vector in the orbit P). Arguing as in the proof of Theorem 4.2 of [82] shows that $f_{a,i}$ does indeed enumerate nowhere-zero primitive twisted necklaces. We thank Victor Reiner for this observation.

Heuristics and main idea We begin by focusing on the product of coefficients appearing in the formula of the peak generating function (4.2.1). More specifically, we seek to find a formula of each coefficient that is more manageable for estimation. Applying the generalized binomial theorem to expand the function, we get

$$\begin{aligned}[x_i^{n_i}] \left(\frac{1 + x_i}{1 - x_i}\right)^{f_{a,i}} &= [x_i^{n_i}] \left((1 + x_i)^{f_{a,i}} (1 - x_i)^{-f_{a,i}}\right) \\ &= \sum_{k=0}^{\infty} \binom{f_{a,i}}{k} \binom{f_{a,i} - 1 + n_i - k}{f_{a,i} - 1} = \frac{(2f_{a,i})^{n_i}}{n_i!} \mathsf{K}_{a,i}, \end{aligned} \tag{4.2.4}$$

where $\mathsf{K}_{a,i}$ is defined by

$$\mathsf{K}_{a,i} = \sum_{\nu=0}^{n_i} \frac{1}{2^{n_i}} \binom{n_i}{\nu} \frac{(f_{a,i} - \nu + n_i - 1)!}{(f_{a,i} - \nu)! f_{a,i}^{n_i - 1}}.$$

To apply (4.2.4), note that the term $t^a \prod_i [x_i^{n_i}] \left((1+x_i)/(1-x_i)\right)^{f_{a,i}}$ in (4.2.1) appears to contribute to the sum meaningfully only when a is comparable to $n^{5/4}$. Also, the $\mathsf{K}_{a,i}$'s are approximately 1 if $f_{a,i}$ is considerably larger than n_i. If all these observations get along, one may argue heuristically that

$$\begin{aligned} \frac{1}{|\mathcal{C}_\lambda|} \sum_{\pi \in \mathcal{C}_\lambda} \left(\frac{4t}{(1+t)^2} \right)^{p(\pi)+1} &\overset{?}{\approx} \left(\frac{\prod_i n_i! i^{n_i}}{n!} \right) \cdot 2 \left(\frac{1-t}{1+t} \right)^{n+1} \int_0^\infty t^x \prod_i \frac{((2x)^i/i)^{n_i}}{n_i!} \, \mathrm{d}x \\ &= \frac{1}{n!} \left(\frac{2(1-t)}{1+t} \right)^{n+1} \int_0^\infty t^x x^n \, \mathrm{d}x \\ &= \left(\frac{2(1-t)}{(1+t)\log(1/t)} \right)^{n+1}. \end{aligned}$$

The final result is the same as what appears in the proof of the asymptotic normality of peaks over S_n. This leads to a naive guess that the peaks over $\mathcal{C}_\lambda$ have asymptotically the same normal distribution as the peaks over S_n. Of course, we must test the validity of this claim. One main concern is that the alleged asymptotic behavior of (4.2.4) may not be valid for small i's. Such phenomenon is already observed in the case of descents [64], where the asymptotic distribution of descents for a fixed cycle type is parametrized by the density of fixed points. And indeed, we will find that corrections are also needed for the peak distribution due to the presence of fixed points. In summary, we need to

- precisely control error terms appearing in various approximations, and
- investigate how the presence of fixed points affects the asymptotic formula for the peak generating function.

From this point forward, let $s > 0$ be a fixed positive real number. Then, t_n is chosen as in (4.2.3), which is the unique solution of the equation $4t_n/(1+t_n)^2 = e^{-s/\sqrt{n}}$ in the interval $(0,1)$. For convenience sake, we indulge in the luxury of abbreviating $t = t_n$ whenever the value of n is clear from the context. As the first step of rigorization, we mimic the heuristic computation without using approximations. Applying (4.2.4) to the peak generating function (4.2.1), we get

$$\begin{aligned}\frac{1}{|\mathcal{C}_\lambda|}\sum_{\pi\in\mathcal{C}_\lambda} e^{-\frac{s}{\sqrt{n}}(p(\pi)+1)} &= \frac{2}{n!}\left(\frac{1-t}{1+t}\right)^{n+1}\sum_{a\geq 1} t^a \prod_{1\leq i\leq n} n_i! i^{n_i}[x_i^{n_i}]\left(\frac{1+x_i}{1-x_i}\right)^{f_{a,i}} \\ &= \frac{2}{n!}\left(\frac{1-t}{1+t}\right)^{n+1}\sum_{a\geq 1} t^a \prod_{1\leq i\leq n} (2a)^{in_i} g_{a,i}^{n_i}\mathsf{K}_{a,i} \\ &= \left(\frac{2(1-t)}{(1+t)\log(1/t)}\right)^{n+1}\left[\frac{\log^{n+1}(1/t)}{n!}\sum_{a\geq 1} a^n t^a \prod_{1\leq i\leq n} g_{a,i}^{n_i}\mathsf{K}_{a,i}\right].\end{aligned}$$

For the sake of conciseness, define $\mathsf{L}_\bullet$ by

$$\mathsf{L}_A := \frac{\log^{n+1}(1/t)}{n!}\sum_{a\in A\cap\mathbb{N}} a^n t^a \prod_{1\leq i\leq n} g_{a,i}^{n_i}\mathsf{K}_{a,i}$$

for all $A\subseteq\mathbb{R}$. Then, the above computation simplifies to

$$\frac{1}{|\mathcal{C}_\lambda|}\sum_{\pi\in\mathcal{C}_\lambda} e^{-\frac{s}{\sqrt{n}}(p(\pi)+1)} = \left(\frac{2(1-t)}{(1+t)\log(1/t)}\right)^{n+1}\mathsf{L}_{[1,\infty)}. \tag{4.2.5}$$

As in the heuristic computation, $\mathsf{L}_\bullet$ will be approximated by its integral analogue. In doing so, it is convenient to split the sum into two parts at a certain threshold. The primary reason is that the aforementioned approximation tends to fail for small a, and so, such case deserves to be handled separately. To describe this threshold, let

$$\delta_0 = \left[\sup_{n\geq 1}\left(n^{1/4}\log(1/t)e^{(c_1/4)+1}\right)\right]^{-1} \tag{4.2.6}$$

and fix any $\delta\in(0,\delta_0)$. In view of (4.2.3), $\log(1/t) = 2\sqrt{s}n^{-1/4} + \mathcal{O}(n^{-3/4})$ for large n. This guarantees that δ_0 is away from 0, and so, the choice of δ does make sense. Then, the sum $\mathsf{L}_{[1,\infty)}$ will be split into $\mathsf{L}_{[1,\delta n^{5/4}]} + \mathsf{L}_{(\delta n^{5/4},\infty)}$, and we will call the former term the *small range* and the latter term the *large range*.

Estimation of small range We will focus on the range $a\leq\delta n^{5/4}$, where δ will be chosen from $(0,\delta_0)$. The main goal in this section is to show that the contribution arising from this range is negligible. The precise statement is as follows.

Lemma 4.2.7. *For each $\delta \in (0, \delta_0)$ and $\rho \in (\delta/\delta_0, 1)$, there exists a constant $c_3 = c_3(\delta, \rho) > 0$, depending only on δ and ρ, such that*

$$\mathsf{L}_{[0,\delta n^{5/4}]} \leq c_3 \rho^{n+1}.$$

We begin by producing a simple upper bound for the product of the $\mathsf{K}_{a,i}$'s.

Lemma 4.2.8. *Let $\delta > 0$. Then, there exists a constant $c_2 = c_2(\delta) > 0$, depending only on δ, such that*

$$\prod_{1 \leq i \leq n} \mathsf{K}_{a,i} \leq \left(\frac{\delta n^{5/4}}{a} \right)^n e^{c_2 n^{3/4}} \tag{4.2.7}$$

whenever $a \leq \delta n^{5/4}$ holds.

Proof. Assume that $a \leq \delta n^{5/4}$. If $0 \leq \nu \leq n_i$, then

$$\frac{(f_{a,i} - \nu + n_i - 1)!}{(f_{a,i} - \nu)! f_{a,i}^{n_i - 1}} = \prod_{k=1}^{n_i - 1} \left(1 + \frac{k - \nu}{f_{a,i}} \right) \leq \left(1 + \frac{n_i}{f_{a,i}} \right)^{n_i}.$$

Plugging this to the definition of $\mathsf{K}_{a,i}$, we obtain $\mathsf{K}_{a,i} \leq (1 + (n_i/f_{a,i}))^{n_i}$. This bound will be further simplified depending on whether $i = 1$ or $i \geq 2$. For brevity, we write

$$r := \frac{\delta n^{5/4}}{a}.$$

By assumption, we have $r \geq 1$. Now, when $i = 1$, plug $f_{a,1} = a$ and proceed as

$$\mathsf{K}_{a,1} \leq \left(1 + \frac{n_1}{a} \right)^{n_1} \leq r^{n_1} \left(1 + \frac{n_1}{ra} \right)^{n_1} \leq r^{n_1} e^{n_1^2/ra} \leq r^{n_1} e^{(1/\delta) n^{3/4}}.$$

In the third and fourth steps, inequalities $1 + x \leq e^x$ and $n_1 \leq n$ are utilized, respectively. Likewise, when $i \geq 2$, we apply Lemma 4.2.6 and proceed as in the previous case to get

$$\mathsf{K}_{a,i} \leq \left(1 + 2e^{c_1} \frac{i n_i}{(2a)^i} \right)^{n_i} \leq r^{i n_i} \left(1 + 2e^{c_1} \frac{i n_i}{(2ra)^i} \right)^{n_i} \leq r^{i n_i} e^{2e^{c_1} i n_i^2/(2ra)^i} \leq r^{i n_i} e^{(e^{c_1}/\delta^2) i n_i / n^{3/2}}.$$

In the last step, the obvious inequality $n_i \leq n$ is used. Combining altogether and utilizing the identity $\sum_{i \geq 2} i n_i = n - n_1$, we see that

$$\prod_{1 \leq i \leq n} \mathsf{K}_{a,i} \leq \left(r^{n_1} e^{(1/\delta) n^{3/4}} \right) \left(r e^{(e^{c_1}/\delta^2)/n^{3/2}} \right)^{n - n_1} \leq r^n e^{c_2 n^{3/4}},$$

where c_2 can be chosen as $c_2 = (1/\delta) + (e^{c_1}/\delta^2)$. $\square$

Proof of Lemma 4.2.7. By Lemmas 4.2.6 and 4.2.8, we see that

$$\begin{aligned}\mathsf{L}_{[0,\delta n^{5/4}]} &\leq \frac{\log^{n+1}(1/t)}{n!}\sum_{1\leq a\leq \delta n^{5/4}}(\delta n^{5/4})^n t^a e^{c_2 n^{3/4}} e^{(c_1/4)n/a^2}\\ &\leq \frac{\log^{n+1}(1/t)}{n!}(\delta n^{5/4})^{n+1} e^{c_2 n^{3/4}} e^{(c_1/4)n}\end{aligned}$$

Here, the last step follows by taking the union bound together with the fact that $t^a \leq 1$. Now, by the definition of δ_0, we have $n^{1/4}\log(1/t)e^{(c_1/4)+1} \leq 1/\delta_0$. Moreover, a quantitative form of the Stirling's formula [86] tells us that $n! \geq \sqrt{2\pi} n^{n+1/2}e^{-n}$, and so,

$$\begin{aligned}\mathsf{L}_{[0,\delta n^{5/4}]} &\leq \frac{1}{(2\pi)^{1/2}n^{n+1/2}e^{-n}}\left(\frac{1}{\delta_0 n^{1/4}e^{(c_1/4)+1}}\right)^{n+1}(\delta n^{5/4})^{n+1}e^{c_2 n^{3/4}}e^{(c_1/4)n}\\ &= \rho^{n+1}\cdot\left(\frac{\delta}{\delta_0\rho}\right)^{n+1}\frac{n^{1/2}e^{c_2 n^{3/4}}}{(2\pi)^{1/2}e^{(c_1/4)+1}}.\end{aligned}$$

If $\rho \in (\delta/\delta_0, 1)$, then the factor $(\delta/\rho\delta_0)^{n+1}n^{1/2}e^{c_2 n^{3/4}}$ is bounded, and hence, the claim follows.

□

Estimation of large range We now turn our attention to the range $a > \delta n^{5/4}$, where we recall that $\delta > 0$ is a fixed number chosen to satisfy (4.2.6). We begin by proving the following lemma, which resolves the contribution of the $\mathsf{K}_{a,i}$'s for $i \geq 2$.

Lemma 4.2.9. *There exists a universal constant $c_4 > 0$ such that*

$$e^{-c_4 n^2/a^2} \leq \prod_{i\geq 2}\mathsf{K}_{a,i} \leq e^{c_4 n^2/a^2}$$

whenever $a \geq \delta n^{5/4} \geq e^{c_1}n$. Here, c_1 is chosen as in Lemma 4.2.6.

Proof. Assume that $a \geq \delta n^{5/4} \geq e^{c_1}n$. When $i \geq 2$, Lemma 4.2.6 gives us that $f_{a,i} \geq e^{-c_1}\frac{(2a)^2}{2i} \geq \frac{2na}{i} \geq 2e^{c_1}n \geq 2n_i$. Now, letting $0 \leq \nu \leq n_1$, we have, as in the beginning of the proof of Lemma 4.2.8,

$$\left(1 - \frac{n_i}{f_{a,i}}\right)^{n_i} \leq \mathsf{K}_{a,i} \leq \left(1 + \frac{n_i}{f_{a,i}}\right)^{n_i}.$$

Since $\frac{n_i}{f_{a,i}} \leq \frac{1}{2}$, we may apply inequalities $-2x \leq \log(1-x)$ and $\log(1+x) \leq 2x$, which are valid for $0 \leq x \leq \frac{1}{2}$, to further simplify the above bounds, which results in

$$-e^{c_1}\frac{in_i^2}{a^2} \leq -\frac{2n_i^2}{f_{a,i}} \leq \log(\mathsf{K}_{a,i}) \leq \frac{2n_i^2}{f_{a,i}} \leq e^{c_1}\frac{in_i^2}{a^2}.$$

Finally, by summing this inequality for $i = 2, \cdots, n$ and utilizing the bound $\sum_i in_i^2 \leq n^2$, the desired conclusion follows with $c_4 = e^{c_1}$. □

Next, we establish a detailed asymptotic expansion of $\mathsf{K}_{a,1}$.

Lemma 4.2.10. *Let $\delta \in (0, \delta_0)$. Then,*

$$\mathsf{K}_{a,1} = \exp\left\{\frac{n_1^3}{12a^2} - \frac{3n_1^5}{160a^4} + \mathcal{O}\left(n^{-1/4}\right)\right\}$$

holds in the range $a \geq \delta n^{5/4} \geq 2n$. Moreover, the implicit bound of the error term depends only on s and δ.

Proof. It is convenient to separate the case of small n_1 from the general argument. Letting $0 \leq \nu \leq n_1$ and using the fact that $1 + x = e^{x+\mathcal{O}(x^2)}$ near $x = 0$, we get

$$\frac{(a-\nu+n_i-1)!}{(a-\nu)!a^{n_i-1}} = \prod_{k=1}^{n_1-1}\left(1+\frac{k-\nu}{a}\right) = \exp\left\{\sum_{k=1}^{n_1-1}\left(\frac{k-\nu}{a} + \mathcal{O}\left(\frac{n_1^2}{a^2}\right)\right)\right\}.$$

So, if N is a random variable having binomial distribution with parameters n_1 and $\frac{1}{2}$, then

$$\mathsf{K}_{a,1} = \mathbb{E}\left[\frac{(a-N+n_i-1)!}{(a-N)!a^{n_i-1}}\right] = e^{\mathcal{O}(n_1^3/a^2)}\mathbb{E}\left[\exp\left\{\frac{n_1-1}{a}\left(\frac{n_1}{2}-N\right)\right\}\right]$$

and

$$\mathbb{E}\left[\exp\left\{\frac{n_1-1}{a}\left(\frac{n_1}{2}-N\right)\right\}\right] = \left[\cosh\left(\frac{n_1-1}{2a}\right)\right]^{n_1} = e^{\mathcal{O}(n_1^3/a^2)},$$

where we utilized the fact that $\cosh(x) = e^{\mathcal{O}(x^2)}$ near $x = 0$. In particular, if we set $\beta = \frac{3}{4}$ and assume that $n_1 \leq n^\beta$, then $n_1^3/a^2 \leq \delta^{-2}n^{-1/4}$, and so, the conclusion of the lemma holds. Again, we prefer to use the named variable β rather than the actual value in order to emphasize how it is employed in each step of the proof.

The previous computation leads our attention to the case $n_1 \geq n^\beta$ with $\beta = \frac{3}{4}$. In such case, we will write

$$\mathsf{K}_{a,1} = \sum_{\nu=0}^{n_1} p(\nu), \quad \text{where} \quad p(\nu) := \frac{1}{2^{n_1}} \binom{n_1}{\nu} \frac{(a+n_1-1-\nu)!}{(a-\nu)!a^{n_1-1}}.$$

We adopt the idea of Laplace's method to estimate $\mathsf{K}_{a,1}$. That said, we will argue by showing that $p(\nu)$ is approximately a gaussian density. Our goal is to establish a rigorous version of this claim and then draw the desired estimate from it.

We first obtain a global upper bound of p. Let $\Gamma(\cdot)$ denote the Gamma function and recall that $n! = \Gamma(n+1)$ for any non-negative integer n. In light of this identity, the definition of $p(\cdot)$ is recast as

$$p(\nu) = \frac{1}{2^{n_1}} \frac{n_1!}{\Gamma(\nu+1)\Gamma(n_1-\nu+1)} \frac{\Gamma(a+n_1-\nu)}{\Gamma(a-\nu+1)a^{n_1-1}}.$$

Utilizing this identity, we extend $p(\nu)$ to an analytic function of ν on $[0, n_1]$. Then adopting the well-known identity $(\log\Gamma(z+1))'' = \sum_{n=1}^{\infty}(n+z)^{-2}$ involving the second derivative of the log-Gamma function [3, Section 6.4], we get

$$\begin{aligned}
(\log p(\nu))'' &= -\sum_{n=1}^{\infty}\left(\frac{1}{(\nu+n)^2} + \frac{1}{(n_1-\nu+n)^2}\right) - \sum_{k=1}^{n_1-1} \frac{1}{(a-\nu+k)^2} \\
&\leq -\sum_{n=1}^{\infty} \frac{2}{(\frac{n_1}{2}+n)^2} \leq -\int_1^{\infty} \frac{2}{(\frac{n_1}{2}+x)^2}\,\mathrm{d}x = -\frac{4}{n_1+2}.
\end{aligned}$$

In particular, $(\log p(\nu))'$ is strictly decreasing on $[0, n_1]$. Moreover, there exists a unique solution $\nu = \tilde{\nu}_0$ of the equation $\log p(\nu+1) - \log p(\nu) = 0$ on $[0, n_1]$ such that

$$\tilde{\nu}_0 := \frac{a+n_1-1-\sqrt{a^2+n_1^2-1}}{2} = \frac{n_1}{2} - \frac{n_1^2}{4a} + \frac{n_1^4}{16a^3} + \mathcal{O}(1) \tag{4.2.8}$$

holds uniformly in a and n_1 in the range $a \geq \delta n_1^{5/4}$. Then, by the mean-value theorem, there exists $\nu_0 \in [\tilde{\nu}_0, \tilde{\nu}_0+1]$ at which $(\log p(\nu))'$ vanishes, and ν_0 is unique by the strict monotonicity. Integrating twice, we get

$$p(\nu) = p(\nu_0)\exp\left\{\int_{\nu_0}^{\nu}(\nu-t)(\log p(t))''\,\mathrm{d}t\right\} \leq p(\nu_0)\exp\left\{-\frac{2}{n_1+2}(\nu-\nu_0)^2\right\}. \tag{4.2.9}$$

Next, we claim that this upper bound is a correct asymptotic formula for $p(\nu)$, which amounts to providing a lower bound similar to (4.2.9). However, one minor issue is that such lower bound cannot generally exist on all of $[0, n_1]$. To circumvent this, we notice that $p(\nu)/p(\nu_0)$ becomes small if $|\nu - \nu_0|$ is sufficiently large compared to $\sqrt{n_1}$. This suggests that we may focus on the range $|\nu - \nu_0| \leq n^{\gamma}\sqrt{n_1}$, where γ is chosen as $\gamma = \frac{\beta}{2} - \frac{1}{4} = \frac{1}{8}$. And in this range, we want to obtain a gaussian lower bound of p. Focusing on the second derivative of $\log p(\nu)$ as before, we obtain

$$\begin{aligned}(\log p(\nu))'' &= -\left(\frac{1}{\nu} + \mathcal{O}\left(\frac{1}{\nu^2}\right) + \frac{1}{n_1 - \nu} + \mathcal{O}\left(\frac{1}{(n_1-\nu)^2}\right)\right) + \mathcal{O}\left(\frac{n_1}{a^2}\right) \\ &= -\frac{n_1}{\nu(n_1-\nu)} + \mathcal{O}\left(n^{-2\beta}\right) + \mathcal{O}\left(n^{-3/2}\right),\end{aligned}$$

where both estimates $\sum_{n=1}^{\infty} \frac{1}{(n+x)^2} = \frac{1}{x} + \mathcal{O}\left(\frac{1}{x^2}\right)$ uniformly in $x > 0$ and $\left|\frac{1}{a-\nu+k}\right| \leq \frac{2}{a}$ are exploited in the first step. To simplify further, we note that

$$\left|\nu - \frac{n_1}{2}\right| \leq |\nu - \nu_0| + \left|\nu_0 - \frac{n_1}{2}\right| \leq n^{\gamma}\sqrt{n_1} + \mathcal{O}\left(\frac{n_1^2}{a}\right) \leq \mathcal{O}\left(\frac{n_1}{n^{1/4}}\right).$$

In the last step, we made use of the bounds $n_1/a = \mathcal{O}(n^{-1/4})$ and $n^{\gamma}/\sqrt{n_1} \leq n^{\gamma-\beta/2} = n^{-1/4}$. So it follows that

$$\frac{n_1}{\nu(n_1-\nu)} = \frac{4}{n_1} \cdot \frac{1}{1 - \left(\frac{\nu-(n_1/2)}{n_1/2}\right)^2} = \frac{4}{n_1}\left(1 + \mathcal{O}\left(n^{-1/2}\right)\right) = \frac{4}{n_1} + \mathcal{O}\left(n^{-\beta-1/2}\right).$$

Plugging this into the asymptotic formula of $(\log p(\nu))''$ and combining all the error terms into a single one, we end up with

$$(\log p(\nu))'' = -\frac{4}{n_1} + \mathcal{O}\left(n^{-5/4}\right).$$

Given this asymptotic formula, we can proceed as in (4.2.9) to obtain

$$p(\nu) = p(\nu_0)\exp\left\{-\frac{2}{n_1}(\nu-\nu_0)^2 + \mathcal{O}\left(n^{2\gamma-5/4}\right)\right\}.$$

From this, we have

$$\begin{aligned}\sum_{\nu:|\nu-\nu_0|\le n^\gamma\sqrt{n_1}}\frac{p(\nu)}{p(\nu_0)} &= e^{\mathcal{O}(n^{-1/4})}\int_{|t|\le n^\gamma\sqrt{n_1}} e^{-\frac{2}{n_1}t^2}\,\mathrm{d}t \\ &= e^{\mathcal{O}\left(n^{-1/4}\right)}\left(\int_{\mathbb{R}} e^{-\frac{2}{n_1}t^2}\,\mathrm{d}t - \int_{|t|>n^\gamma\sqrt{n_1}} e^{-\frac{2}{n_1}t^2}\,\mathrm{d}t\right) \\ &= e^{\mathcal{O}\left(n^{-1/4}\right)}\sqrt{\frac{\pi n_1}{2}} + \mathcal{O}\left(e^{-2n^\gamma}\right)\end{aligned}$$

The first step follows by noting that $-\frac{2}{n_1}(t-\nu_0)^2 = -\frac{2}{n_1}(\nu-\nu_0)^2 + \mathcal{O}\left(n^{\gamma-\beta/2}\right)$ if $|t-\nu|\le 1$ and $\gamma-\beta/2 = -1/4$. Also, in the last step, we utilized the tail estimate $\int_x^\infty e^{-t^2/2}\,\mathrm{d}t < e^{-x^2/2}/x$, which is valid for $x>0$, to produce a stretched-exponential decay. Similar reasoning shows that

$$\sum_{\nu:|\nu-\nu_0|>n^\gamma\sqrt{n_1}}\frac{p(\nu)}{p(\nu_0)} \le \mathcal{O}\left(\int_{|t|>n^\gamma\sqrt{n_1}} e^{-\frac{2}{n_1+2}t^2}\,\mathrm{d}t\right) \le \mathcal{O}\left(e^{-2n^\gamma}\right).$$

Putting all the estimates altogether, we obtain

$$\mathsf{K}_{a,1} = \sqrt{\frac{\pi n_1}{2}}e^{\mathcal{O}\left(n^{-1/4}\right)}p(\nu_0). \tag{4.2.10}$$

In view of (4.2.10), it remains to estimate $p(\nu_0)$. Since $\nu_0-\tilde{\nu}_0 = \mathcal{O}(1)$, it follows ν_0 satisfies the same asymptotic formula as in (4.2.8). Write $\mu = \nu_0 - \frac{n_1}{2}$. We know that $\mu = o(n_1)$, or more precisely, $\mu/n_1 = \mathcal{O}(n^{-1/4})$. Then, by using Stirling's approximation [86]

$$\log(n!) = \left(n+\frac{1}{2}\right)\log n - n + \log\sqrt{2\pi} + \mathcal{O}\left(\frac{1}{n}\right),$$

we obtain

$$\begin{aligned}\log\left[\frac{1}{2^{n_1}}\binom{n_1}{\nu_0}\right] &= -n_1\log 2 + \log(n_1!) - \log\left(\frac{n_1}{2}+\mu\right)! - \log\left(\frac{n_1}{2}-\mu\right)! \\ &= -\left(\frac{n_1}{2}+\mu+\frac{1}{2}\right)\log\left(1+\frac{2\mu}{n_1}\right) - \left(\frac{n_1}{2}-\mu+\frac{1}{2}\right)\log\left(1-\frac{2\mu}{n_1}\right) \\ &\quad + \log 2 - \frac{1}{2}\log n_1 - \log\sqrt{2\pi} + \mathcal{O}\left(\frac{1}{n_1}\right) \\ &= \frac{n_1}{2}\left[\left(\frac{1}{n_1}-1\right)\left(\frac{2\mu}{n_1}\right)^2 + \left(\frac{1}{2n_1}-\frac{1}{6}\right)\left(\frac{2\mu}{n_1}\right)^4 + \mathcal{O}\left(\frac{2\mu}{n_1}\right)^6\right] \\ &\quad + \frac{1}{2}\log\left(\frac{2}{\pi n_1}\right) + \mathcal{O}\left(n^{-\beta}\right).\end{aligned}$$

This can be further simplified by noting that $\frac{\mu}{n_1} = -\frac{n_1}{4a} + \frac{n_1^3}{16a^3} + \mathcal{O}\left(\frac{1}{n^{1/4}n_1}\right) = \mathcal{O}(n^{-1/4})$, and the result is

$$\begin{aligned}\log\left[\frac{1}{2^{n_1}}\binom{n_1}{\nu_0}\right] &= \frac{1}{2}\log\left(\frac{2}{\pi n_1}\right) - \frac{2\mu^2}{n_1} - \frac{4\mu^4}{3n_1^3} + \mathcal{O}\left(n^{-1/4}\right) \\ &= \frac{1}{2}\log\left(\frac{2}{\pi n_1}\right) - \frac{n_1^3}{8a^2} + \frac{11n_1^5}{192a^4} + \mathcal{O}\left(n^{-1/4}\right). \end{aligned} \tag{4.2.11}$$

For the remaining factor, we estimate it as follows.

$$\begin{aligned}\log\left[\frac{(a+n_1-1-\nu_0)!}{(a-\nu_0)!a^{n_1-1}}\right] &= \log\left[\frac{(a+n_1-\nu_0)!}{(a-\nu_0)!a^{n_1}}\right] + \log\left[\frac{a}{a+n_1-\nu_0}\right] \\ &= -n_1 + \left(a+\frac{1}{2}+\frac{n_1}{2}-\mu\right)\log\left(1+\frac{\frac{n_1}{2}-\mu}{a}\right) \\ &\quad - \left(a+\frac{1}{2}-\frac{n_1}{2}-\mu\right)\log\left(1-\frac{\frac{n_1}{2}+\mu}{a}\right) + \mathcal{O}\left(n^{-1/4}\right)\end{aligned}$$

After some painful expansion, we end up with

$$\log\left[\frac{(a+n_1-1-\nu_0)!}{(a-\nu_0)!a^{n_1-1}}\right] = \frac{5n_1^3}{24a^2} - \frac{73n_1^5}{960a^4} + \mathcal{O}\left(n^{-1/4}\right). \tag{4.2.12}$$

Therefore, the conclusion follows by combining (4.2.10), (4.2.11) and (4.2.12) altogether. □

Estimation of the peak generating function

Lemma 4.2.11. *Let $\delta \in (0, \delta_0)$ and write $\alpha_1 = n_1/n$ for the density of fixed points. Then*

$$\mathsf{L}_{(\delta n^{5/4},\infty)} = \exp\left\{\frac{\alpha_1^3}{3}s\sqrt{n} + \left(\frac{\alpha_1^3}{18} - \frac{3\alpha_1^5}{10} + \frac{2\alpha_1^6}{9}\right)s^2 + \mathcal{O}\left(n^{-1/4}\right)\right\}$$

holds in the range $\delta n^{5/4} \geq \max\{e^{c_1}, 2\}n$. Moreover, the implicit bound of the error term depends only on δ and s.

Following Kim and Lee's method [64], we will utilize Laplace's method to approximate the sum by the integral of a certain gaussian density function and show that the relative error due to this approximation can be controlled in an explicit and uniform manner. The following simple lemma is useful for this purpose.

Lemma 4.2.12. *Define $f_n : \mathbb{R} \to \mathbb{R}$ by $f_n(x) = \left(1 + \frac{x}{\sqrt{n}}\right)^n e^{-\sqrt{n}x}\mathbf{1}_{[-\sqrt{n},\infty)}(x)$. Then*

(1) If $x \geq 0$ and $l > n > 0$, then $f_l(x) \leq f_n(x) \leq (2/\sqrt{e})^n e^{-\sqrt{n}x/2}$.

(2) If $x \leq 0$ and $l > n > 0$, then $f_n(x) \leq f_l(x) \leq e^{-x^2/2}$.

(3) $f_n(x) \to e^{-x^2/2}$ pointwise as $n \to \infty$.

The estimation of f_n is a recurring tool in previous works (see Lemma 4.3 of [64] and the proof therein, for instance) and requires only basic calculus computation. Nevertheless, we include the proof for self-containedness.

Proof. Let $h(t,x) = t\log\left(1 + \frac{x}{\sqrt{t}}\right) - \sqrt{t}x$. It is easy to check that

- $x \mapsto h(t,x)$ is concave on $(0,\infty)$ for each $t \in (0,\infty)$,
- $t \mapsto h(t,x)$ is decreasing on $(0,\infty)$ for each $x \geq 0$,
- $t \mapsto h(t,x)$ is increasing on (x^2,∞) for each $x \leq 0$, and
- $h(t,x) \to -x^2/2$ as $t \to \infty$ for each $x \in \mathbb{R}$.

From $f_n(x) = e^{h(n,x)}$, the assertions (2) and (3) follows immediately. Moreover, we may exploit the concavity of $x \mapsto h(t,x)$ to bound $h(t,x) \leq h(t,\sqrt{n}) + \frac{\partial h}{\partial x}(t,\sqrt{n})(x - \sqrt{n})$, which gives (1). □

Now we return to the proof of the main claim of this section.

Proof of Lemma 4.2.11. Assume that $\delta n^{5/4} \geq \max\{e^{c_1}, 2\}n$ holds. Also, recall that $t = t_n$ is the unique solution of $4t_n/(1+t_n)^2 = e^{-s/\sqrt{n}}$. Then, by Lemmas 4.2.6, 4.2.9, and 4.2.10, we have

$$\mathsf{L}_{(\delta n^{5/4},\infty)} = e^{\mathcal{O}(n^{-1/4})} \frac{\log^{n+1}(1/t)}{n!} \sum_{a > \delta n^{5/4}} t^a a^n \exp\left\{\frac{n_1^3}{12a^2} - \frac{3n_1^5}{160a^4}\right\}$$

Next, we approximate the sum in the right-hand side by its integral analogue. If $x \in \mathbb{R}$ and $a > \delta n^{5/4}$ are such that $|x - a| \leq 1$, then

- $t^x = t^a e^{\mathcal{O}(\log t)} = t^a e^{\mathcal{O}(n^{-1/4})}$,
- $x^n = a^n e^{n\log(x/a)} = a^n e^{\mathcal{O}(n/a)} = a^n e^{\mathcal{O}(n^{-1/4})}$, and

- for each $k \geq 0$ given, $\frac{n_1^{k+1}}{x^k} = \frac{n_1^{k+1}}{a^k}\left(1+\mathcal{O}\left(\frac{1}{a}\right)\right)^k = \frac{n_1^{k+1}}{a^k} + \mathcal{O}\left(\frac{n_1}{a}\right)^{k+1} = \frac{n_1^{k+1}}{a^k} + \mathcal{O}(n^{-1/4})$. The implicit error bound now depends on k as well. However, it will be used only for $k = 2$ and $k = 4$, and so, this causes no harm for our objective of retaining error bounds depending only on s and δ.

This allows us to approximate the sum by its integral analogue at the expense of the relative error $e^{\mathcal{O}(n^{-1/4})}$, yielding

$$\mathsf{L}_{(\delta n^{5/4},\infty)} = e^{\mathcal{O}(n^{-1/4})}\frac{\log^{n+1}(1/t)}{n!}\mathsf{J}, \qquad \text{where } \mathsf{J} = \int_{\delta n^{5/4}}^{\infty} t^x x^n \exp\left\{\frac{n_1^3}{12x^2} - \frac{3n_1^5}{160x^4}\right\} \mathrm{d}x \tag{4.2.13}$$

So it remains to estimate J. To this end, we substitute $x = \frac{n}{\log(1/t)}\left(1+\frac{w}{\sqrt{n}}\right)$. For the sake of brevity, we also write $c_5 = c_5(n) = 1 - \delta n^{1/4}\log(1/t)$. Although c_5 depends on n and s, the choice of δ and (4.2.6) tell us that c_5 is uniformly away from 0 and 1, which will be sufficient for our purpose. Then,

$$\begin{aligned}\mathsf{J} = \int_{-c_5\sqrt{n}}^{\infty} \exp\Bigg\{ &\left(n\log\left(\frac{n}{\log(1/t)}\right) - n\right) + \left(n\log\left(1+\frac{w}{\sqrt{n}}\right) - \sqrt{n}w\right) \\ &+ \frac{\alpha_1^3 n\log^2(1/t)}{12\,(1+(w/\sqrt{n}))^2} - \frac{3\alpha_1^5 n\log^4(1/t)}{160\,(1+(w/\sqrt{n}))^4}\Bigg\}\frac{\sqrt{n}}{\log(1/t)}\,\mathrm{d}w.\end{aligned}$$

The first two grouped terms in the exponent of the integrand are easily controlled, as they originated from the 'unperturbed term' $t^x x^n$. So, it suffices to study the effect of the 'perturbation terms'. Taking advantage of the explicit formula of the perturbation term, one may expand

$$\frac{\alpha_1^3 n\log^2(1/t)}{12\,(1+(w/\sqrt{n}))^2} = \frac{\alpha_1^3 n\log^2(1/t)}{12}\left(1 - \frac{(w/\sqrt{n})\,(2+(w/\sqrt{n}))}{(1+(w/\sqrt{n}))^2}\right)$$

Plugging this back in, the integral takes the form

$$\mathsf{J} = \frac{n^{n+1/2}e^{-n}}{\log^{n+1}(1/t)}\exp\left\{\frac{\alpha_1^3}{12}n\log^2(1/t)\right\}\int_{-c_5\sqrt{n}}^{\infty} f_n(w)e^{g_n(w)}\,\mathrm{d}w,$$

where f_n is as in Lemma 4.2.12 and g_n is defined by

$$g_n(w) = -\frac{\alpha_1^3\sqrt{n}\log^2(1/t)w(2+(w/\sqrt{n}))}{12(1+(w/\sqrt{n}))^2} - \frac{3\alpha_1^5 n\log^4(1/t)}{160\,(1+(w/\sqrt{n}))^4}.$$

As mentioned before, c_5 is uniformly away from 1, meaning that $\sup_{n\geq 1} c_5(n) < 1$ holds. Then $g_n(w) \leq 0$ for $w \geq 0$ and $g_n(w) \leq -c_6 w$ for $w \in [-c_5\sqrt{n}, 0]$, where $c_6 > 0$ is a constant depending only on s. Now using the tail estimates in Lemma 4.2.12, we can check that

$$\int\limits_{\substack{w\geq -c_5\sqrt{n}\\ |w|\geq \log n}} f_n(w) e^{g_n(w)}\,\mathrm{d}w = \mathcal{O}\left(n^{-1/4}\right).$$

Moreover, if $|w| \leq \log n$, then using $\log(1/t) = \frac{2\sqrt{s}}{n^{1/4}} + \frac{s^{3/2}}{6n^{3/4}} + \mathcal{O}\left(n^{-5/4}\right)$,

$$f_n(w) = -\frac{w^2}{2} + \mathcal{O}\left(\frac{\log^3 n}{\sqrt{n}}\right), \qquad g_n(w) = -\frac{\alpha_1^3}{3} sw - \frac{3\alpha_1^5}{10}s^2 + \mathcal{O}\left(\frac{\log n}{\sqrt{n}}\right).$$

Plugging this back to $\mathsf{L}_{(\delta n^{5/4},\infty)}$ and utilizing Stirling's formula,

$$\begin{aligned}\mathsf{L}_{(\delta n^{5/4},\infty)} &= \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{\alpha_1^3}{3}s\sqrt{n} + \left(\frac{\alpha_1^3}{18} - \frac{3\alpha_1^5}{10}\right)s^2 + \mathcal{O}(n^{-1/4})\right\} \\ &\quad \times \left(\int_{|w|\leq \log n} \exp\left\{-\frac{w^2}{2} - \frac{\alpha_1^3}{3}sw\right\}\mathrm{d}w + \mathcal{O}\left(n^{-1/4}\right)\right) \\ &= \exp\left\{\frac{\alpha_1^3}{3}s\sqrt{n} + \left(\frac{\alpha_1^3}{18} - \frac{3\alpha_1^5}{10} + \frac{2\alpha_1^6}{9}\right)s^2 + \mathcal{O}\left(n^{-1/4}\right)\right\}\end{aligned}$$

as required. □

With all the ingredients ready, we immediately obtain the proof of Theorem 4.2.2.

Proof of Theorem 4.2.2. In the proof of Theorem 4.2.5, we checked that

$$\left(\frac{2(1-t)}{(1+t)\log(1/t)}\right)^{n+1} = \exp\left\{-\frac{s}{3}\sqrt{n} + \frac{1}{45}s^2 + \mathcal{O}(n^{-1/4})\right\}.$$

Moreover, if we fix $\delta \in (0, \delta_0)$, by Lemma 4.2.7, we can choose $\rho \in (0,1)$, independent of n and λ, so that $\mathsf{L}_{[1,\delta n^{5/4}]} = \mathcal{O}(\rho^n)$. Also, if n is sufficiently large so that $\delta n^{5/4} \geq \max\{e^{c_1}, 2\}n$, Lemma 4.2.11 gives a uniform estimate on $\mathsf{L}_{(\delta n^{5/4},\infty)}$. Finally, if π is chosen uniformly at random from $\mathcal{C}_\lambda$, then

$$\mathbb{E}\left[e^{-sp(\pi)/\sqrt{n}}\right] = e^{s/\sqrt{n}}\left(\frac{2(1-t)}{(1+t)\log(1/t)}\right)^{n+1}\mathsf{L}_{[1,\infty)}.$$

Plugging in all the estimates and taking advantage of the fact that $\mathsf{L}_{[1,\delta n^{5/4}]} = \mathcal{O}(\rho^n)$ can be absorbed into the relative error $\mathcal{O}(n^{-1/4})$, we have

$$\mathbb{E}\left[e^{-sp(\pi)/\sqrt{n}}\right] = \exp\left\{\left(-\frac{s}{3}\sqrt{n} + \frac{1}{45}s^2\right) + \left(\frac{\alpha_1^3}{3}s\sqrt{n} + \left(\frac{\alpha_1^3}{18} - \frac{3\alpha_1^5}{10} + \frac{2\alpha_1^6}{9}\right)s^2\right) + \mathcal{O}\left(n^{-1/4}\right)\right\}.$$

This provides the desired bound for the term $E_{\lambda,s}$ appearing in the statement of Theorem 4.2.2, completing the proof. □

Chapter 5

Conclusions and future plans

We finish this thesis with some concluding remarks and a note on future work.

In the first part of this thesis, we proved the scaling limit of exceptional points of the simple random walk in lattice approximations of a planar domain run for a constant multiple of the expected cover time. This was a continuation of the previous work by Abe and Biskup [1] that first established a precise picture of the structure of exceptional point of the pinned local time, and was aimed at ameliorating several aspects of the previous results that deviate from the natural setting through various conversion techniques.

This work is part of a much larger project that attempts to understand the structure of exceptional points of planar random walks. In this regard, the results of [1] and those reported here still "suffer" from the restriction, to the wired boundary condition, which was introduced in order to facilitate the connection to the current version of extreme value theory of DGFFs in planar domains. A natural question is whether analogous results continue to hold in the case of random walks with more natural boundary conditions, such as periodic or free boundary condition.

In an ongoing project, Abe, Biskup, and I are currently attempting to resolve the torus case. One key difference is that the fluctuation U_N of the total time has a different scaling than in the case of wired boundary, to the point that the exceptional measures need to be normalized by *random* quantities. This is due to the fact that growth of the local time away

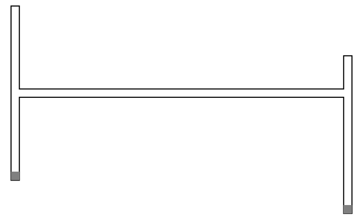


Figure 5.1: An example of a fence. Polluted endpoints are represented by gray squares. The typical length-scale of this structure is $\log(1/p)/p$.

from the pinning point now hinges on much longer but rarer exclusions. This poses an extra difficulty in removing the effect of the pinning, and indeed, it is one of the current challenges in the project.

In the second part of this thesis, we demonstrated a method for identifying a correct scaling relationship between two initial densities p and q in the 2-neighbor polluted bootstrap percolation on $\mathbb{Z}^2$ that separates different percolation regimes. This was done by analyzing the occupied/unoccupied interface in the terminal configuration and then attributing the emergence of different percolation regimes to the formation of *blocking structures* that capture the prominent features of the aforementioned interface. The desired p *versus* q scaling was then determined so that the blocking structure permits a non-trivial scaling limit in the form of a continuum percolation model. This limit model is then used to study the percolation of blocking structures.

One of our future plans is to apply this method to the modified PBP on $\mathbb{Z}^2$. Heuristic computations suggest that the correct p *versus* q scaling is $q = p^2/\log(1/p)$. The extra logarithmic factor comes from the observation that the correct blocking structure in the modified PBP is a blocking contour made of wedges whose arms are themselves percolation clusters of another type of blocking structures called *blocking fences*. A blocking fence is a chain of H-shaped structures, where each H-shaped structure consists of two parallel unoccupied rectangles (called *poles*) of width 1 that are crossed by a perpendicular unoccupied rectangle of width 1 and the endpoints in the same sides of the poles are polluted, see Figure 5.1. This idea has already been used in an unpublished work by the author to show that, if $p, q \downarrow 0$

while the ratio $\lambda(p,q) := q/(p^2/\log(1/p))$ is kept above a sufficiently large constant, then no infinite cluster appears in the terminal configuration. We hope to extend this to the same level of control that we currently have for the ordinary BP model.

In the last part of the thesis, we proved the asymptotic normality for descents and peaks in an arbitrary conjugacy class. The explicit combinatorial formulas for the generating function of these two permutation statistics were essential for predicting the asymptotic mean and variance. In light of the continuity theorem, the question of asymptotic normality was reduced to showing the pointwise convergence of the moment generating functions of rescaled versions of permutation statistics, which we accomplished by obtaining a uniform estimates on the moment generating functions.

One shortcoming of the approach adopted in this work is that the burden of understanding the structure of random permutations is deferred to the known exact formulas, which are algebraic in origin. This thus poses an interesting question of whether these results can be replicated through probabilistic ideas. Especially, finding a plausible explanation on why the asymptotic normality depends only on the density of fixed points in the limit would be a good starting point in this direction.

Bibliography

[1] Y. Abe and M. Biskup. Exceptional points of two-dimensional random walks at multiples of the cover time. *arXiv:1903.04045 [math-ph]*, March 2019. arXiv: 1903.04045.

[2] Y. Abe, M. Biskup, and S. Lee. Exceptional points of discrete-time random walks in planar domains. *arXiv:1911.11810 [math-ph]*, November 2019. arXiv: 1911.11810.

[3] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Courier Corporation, January 1965.

[4] M. Aizenman and J. L. Lebowitz. Metastability effects in bootstrap percolation. *Journal of Physics A: Mathematical and General*, 21(19):3801–3813, October 1988.

[5] M. Aizenman and D. J. Barsky. Sharpness of the phase transition in percolation models. *Communications in Mathematical Physics*, 108(3):489–526, September 1987.

[6] D. J. Aldous. Threshold limits for cover times. *Journal of Theoretical Probability*, 4(1):197–211, January 1991.

[7] T. Antunović and I. Veselić. Sharpness of the phase transition and exponential decay of the subcritical cluster size for percolation on quasi-transitive graphs. *Journal of Statistical Physics*, 130(5):983–1009, March 2008. arXiv: 0707.1089.

[8] J. Balogh, B. Bollobás, H. Duminil-Copin, and R. Morris. The sharp threshold for bootstrap percolation in all dimensions. *Transactions of the American Mathematical Society*, 364(5):2667–2701, 2012.

[9] R. Bauerschmidt, D. C. Brydges, and G. Slade. Logarithmic correction for the susceptibility of the 4-dimensional weakly self-avoiding walk: A renormalisation group analysis. *Communications in Mathematical Physics*, 337(2):817–877, July 2015.

[10] D. Bayer and P. Diaconis. Trailing the dovetail shuffle to its lair. *The Annals of Applied Probability*, 2(2):294–313, May 1992.

[11] D. Belius and N. Kistler. The subleading order of two dimensional cover times. *Probability Theory and Related Fields*, 167(1):461–552, February 2017.

[12] E. A. Bender. Central and local limit theorems applied to asymptotic enumeration. *Journal of Combinatorial Theory, Series A*, 15(1):91–111, July 1973.

[13] S. Billey, K. Burdzy, and B. E. Sagan. Permutations with given peak set. *Journal of Integer Sequences*, 16(6):18, 2013.

[14] M. Biskup. Extrema of the two-dimensional Discrete Gaussian Free Field. *In: M. Barlow and G. Slade (eds.): Random Graphs, Phase Transitions, and the Gaussian Free Field.*, SSPROB 2017. Springer Proceedings in Mathematics & Statistics, 304 (2020) 163–407. Springer, Cham.

[15] M. Biskup and O. Louidor. Extreme local extrema of two-dimensional Discrete Gaussian Free Field. *Communications in Mathematical Physics*, 345(1):271–304, July 2016.

[16] M. Biskup and O. Louidor. On intermediate level sets of two-dimensional discrete Gaussian Free Field. Ann. Inst. Henri Poincaré 55 (2019), no. 4, 1948–1987

[17] M. Biskup and O. Louidor. Conformal symmetries in the extremal process of two-dimensional discrete Gaussian Free Field. Commun. Math. Phys., (to appear)

[18] M. Biskup and O. Louidor. Full extremal process, cluster law and freezing for the two-dimensional discrete Gaussian Free Field. *Advances in Mathematics*, 330:589–687, May 2018.

[19] M. Biskup, O. Louidor, and S. Gulfer. Near-maxima of two-dimensional Discrete Gaussian Free Field and support sets of critical Liouville Quantum Gravity. (in preparation)

[20] E. Bolthausen, J. Deuschel, and G. Giacomin. Entropic repulsion and the maximum of the two-dimensional harmonic. *The Annals of Probability*, 29(4):1670–1692, October 2001.

[21] D. Brydges, S. N. Evans, and J. Z. Imbrie. Self-avoiding walk on a hierarchical lattice in four dimensions. *The Annals of Probability*, 20(1):82–124, 1992.

[22] D. C. Brydges, J. Fröhlich, and A. D. Sokal. The random-walk representation of classical spin systems and correlation inequalities. *Communications in Mathematical Physics*, 91(1):117–139, March 1983.

[23] D. C. Brydges, T. Helmuth, and M. Holmes. The continuous-time lace expansion. *arXiv:1905.09605 [math-ph]*, May 2019. arXiv: 1905.09605.

[24] J. Chalupa, P. L. Leath, and G. R. Reich. Bootstrap percolation on a Bethe lattice. *Journal of Physics C: Solid State Physics*, 12(1):L31–L35, January 1979.

[25] S. Chatterjee and P. Diaconis. A central limit theorem for a new statistic on permutations. *Indian Journal of Pure and Applied Mathematics*, 48(4):561–573, December 2017.

[26] Aser Cortines, Oren Louidor, and Santiago Saglietti. A Scaling Limit for the Cover Time of the Binary Tree. *arXiv:1812.10101 [math]*, January 2019. arXiv: 1812.10101.

[27] J. H. Curtiss. A Note on the Theory of Moment Generating Functions. *The Annals of Mathematical Statistics*, 13(4):430–433, December 1942.

[28] O. Daviaud. Extremes of the discrete two-dimensional Gaussian free field. *The Annals of Probability*, 34(3):962–986, May 2006.

[29] F. N. David and D. E. Barton. *Combinatorial Chance*. Hafner Publishing Company, 1962. Google-Books-ID: FPJLRcErNtAC.

[30] A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni. Thick points for planar Brownian motion and the Erdős-Taylor conjecture on random walk. *Acta Mathematica*, 186(2):239–270, 2001.

[31] A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni. Late points for random walks in two dimensions. *The Annals of Probability*, 34(1):219–263, January 2006.

[32] A. Dembo, J. Rosen, and O. Zeitouni. Limit law for the cover time of a random walk on a binary tree. *arXiv:1906.07276 [math]*, June 2019. arXiv: 1906.07276.

[33] P. Diaconis. *Group representations in probability and statistics*. Institute of Mathematical Statistics, 1988. Google-Books-ID: nfvuAAAAMAAJ.

[34] P. Diaconis, J. Fulman, and S. Holmes. Analysis of casino shelf shuffling machines. *The Annals of Applied Probability*, 23(4):1692–1720, August 2013.

[35] P. Diaconis and R. Graham. *Magical Mathematics: The Mathematical Ideas that Animate Great Magic Tricks*. Princeton University Press, 2012.

[36] P. Diaconis and R. Graham. The Magic of Charles Sanders Peirce. 2017.

[37] P. Diaconis, M. Mc. Grath, and J. Pitman. Riffle shuffles, cycles, and descents. *Combinatorica*, 15(1):11–29, March 1995.

[38] J. Ding, J. R. Lee, and Y. Peres. Cover times, blanket times, and majorizing measures. April 2010.

[39] H. Duminil-Copin, A. Raoufi, and V. Tassion. Subcritical phase of d-dimensional Poisson-Boolean percolation and its vacant set. *arXiv:1805.00695 [math-ph]*, May 2018. arXiv: 1805.00695.

[40] H. Duminil-Copin and V. Tassion. A new proof of the sharpness of the phase transition for Bernoulli percolation and the Ising model. *Communications in Mathematical Physics*, 343(2):725–745, April 2016.

[41] R. Durrett. Oriented percolation in two dimensions. *The Annals of Probability*, 12(4):999–1040, November 1984.

[42] E. B. Dynkin. Gaussian and non-Gaussian random fields associated with Markov processes. *Journal of Functional Analysis*, 55(3):344–376, February 1984.

[43] N. Eisenbaum, H. Kaspi, M. B. Marcus, J. Rosen, and Z. Shi. A Ray-Knight theorem for symmetric Markov processes. *The Annals of Probability*, 28(4):1781–1796, October 2000.

[44] P. Erdős and S. J. Taylor. Some problems concerning the structure of random walk paths. *Acta Mathematica Academiae Scientiarum Hungarica*, 11(1):137–162, March 1963.

[45] L. Euler. Institutiones calculi differentialis cum eius usu in analysi finitorum ac doctrina serierum. 1787.

[46] P. Flajolet and R. Sedgewick. *Analytic combinatorics.* Cambridge University Press, Cambridge, 2009.

[47] J. Fröhlich. On the triviality of $\lambda\phi_d^4$ theories and the approach to the critical point in $d_{(-)} > 4$ dimensions. *Nuclear Physics B*, 200(2):281–296, February 1982.

[48] J. Fulman. The Distribution of Descents in Fixed Conjugacy Classes of the Symmetric Groups. *Journal of Combinatorial Theory, Series A*, 84(2):171–180, November 1998.

[49] J. Fulman. *Stein's method and non-reversible Markov chains.* Institute of Mathematical Statistics, 2004.

[50] J. Fulman, G. B. Kim, and S. Lee. Central limit theorem for peaks of a random permutation in a fixed conjugacy class of S_n. *arXiv:1902.00978 [math]*, February 2019. arXiv: 1902.00978, submitted for publication.

[51] J. Fulman, P. Neumann, and C. Praeger. A generating function approach to the enumeration of matrices in classical groups over finite fields. *Memoirs of the American Mathematical Society*, 830, July 2005.

[52] A. M. Garsia and I. M. Gessel. Permutation statistics and partitions. *Advances in Mathematics*, 31(3):288–305, March 1979.

[53] I. M. Gessel and C. Reutenauer. Counting permutations with given cycle structure and descent set. *Journal of Combinatorial Theory, Series A*, 64(2):189–215, November 1993.

[54] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete mathematics: a foundation for computer science.* Addison-Wesley, Reading, Mass, 2nd ed edition, 1994.

[55] J. Gravner and E. McDonald. Bootstrap percolation in a polluted environment. *Journal of Statistical Physics*, 87(3-4):915–927, May 1997.

[56] L. H. Harper. Stirling behavior is asymptotically normal. *The Annals of Mathematical Statistics*, 38(2):410–414, April 1967.

[57] A. E. Holroyd. Sharp metastability threshold for two-dimensional bootstrap percolation. *Probability Theory and Related Fields*, 125(2):195–224, February 2003.

[58] A. Jego. Thick points of random walk and the Gaussian free field. *arXiv:1809.04369 [math]*, September 2018. arXiv: 1809.04369.

[59] A. Jego. Characterisation of planar Brownian multiplicative chaos. *arXiv:1909.05067 [math]*, October 2019. arXiv: 1909.05067.

[60] A. Jego. Planar Brownian motion and Gaussian multiplicative chaos. *arXiv:1812.01903 [math]*, October 2019. arXiv: 1812.01903.

[61] J. Kahane. Sur le chaos multiplicatif. *Annales mathématiques du Québec*, 9(2):105–150, 1985.

[62] H. Kesten. An iterated logarithm law for local time. *Duke Mathematical Journal*, 32(3):447–456, September 1965.

[63] G. B. Kim. Distribution of descents in matchings. *Annals of Combinatorics*, 23(1):73–87, March 2019.

[64] G. B. Kim and S. Lee. Central limit theorem for descents in conjugacy classes of S_n. *Journal of Combinatorial Theory, Series A*, 169:105123, January 2020. 10.1016/j.jcta.2019.105123.

[65] D. E. Knuth. *The Art of Computer Programming: Volume 1: Fundamental Algorithms*. Pearson Education, 1997.

[66] D. E. Knuth. *The Art of Computer Programming: Volume 3: Sorting and Searching*. Pearson Education, 1998.

[67] Y. L. Jan. Temps local et superchamp. In Jacques Azéma, Marc Yor, and Paul André Meyer, editors, *Séminaire de Probabilités XXI*, volume 1247, pages 176–190. Springer Berlin Heidelberg, Berlin, Heidelberg, 1987.

[68] T. M. Liggett, R. H. Schonmann, and A. M. Stacey. Domination by product measures. *The Annals of Probability*, 25(1):71–95, January 1997.

[69] T. M. Liggett. An Improved Subadditive Ergodic Theorem. *The Annals of Probability*, 13(4):1279–1285, November 1985.

[70] P. A. MacMahon. *Combinatory Analysis*. Courier Corporation, July 2004.

[71] R. Meester, R. Roy, and A. Sarkar. Nonuniversality and continuity of the critical covered volume fraction in continuum percolation. *Journal of Statistical Physics*, 75(1):123–134, April 1994.

[72] M. V. Menshikov, S. A. Molchanov, and A. F. Sidorenko. Percolation theory and some applications. *Journal of Soviet Mathematics*, 42(4):1766–1810, August 1988.

[73] K. L. Nyman. The Peak Algebra of the Symmetric Group. *Journal of Algebraic Combinatorics*, 17(3):309–322, May 2003.

[74] I. Okada. Frequently visited sites of the inner boundary of simple random walk range. *Stochastic Processes and their Applications*, 126(5):1412–1432, May 2016.

[75] I. Okada. Exponent for sets of frequently visited points of a simple random walk in two dimensions. *Latin American Journal of Probability and Mathematical Statistics*, 16:1129, January 2019.

[76] I. Okada. Geometric structures of late points of a two-dimensional simple random walk. *The Annals of Probability*, 47(5):2869–2893, September 2019.

[77] I. Okada. Exponents for the number of pairs of α-favorite points of a simple random walk in $\mathbb{Z}^2$. *Stochastic Processes and their Applications*, 130(1):108–138, January 2020.

[78] A. Y. Özdemir. Martingales and descent statistics. *arXiv:1901.01719 [math]*, January 2019. arXiv: 1901.01719.

[79] T. K. Petersen. Enriched P-partitions and peak algebras. *Advances in Mathematics*, 209(2):561–610, March 2007.

[80] T. K. Petersen. Eulerian numbers. Birkhäuser Advanced Texts Basler Lehrbücher, pages 3–18. Springer New York, New York, NY, 2015.

[81] J. Pitman. Probabilistic Bounds on the Coefficients of Polynomials with Only Real Zeros. *Journal of Combinatorial Theory, Series A*, 77(2):279–303, February 1997.

[82] V. Reiner. Signed Permutation Statistics and Cycle Type. *European Journal of Combinatorics*, 14(6):569–579, November 1993.

[83] C. Reutenauer. *Free Lie Algebras.* Clarendon Press, 1993.

[84] P. Révész. *Random Walk in Random and Non-random Environments*. World Scientific, 2005.

[85] J. Riordan. *An Introduction to Combinatorial Analysis.* Wiley, 1958.

[86] H. Robbins. A Remark on Stirling's Formula. *The American Mathematical Monthly*, 62(1):26, January 1955.

[87] P. Rodriguez. On pinned fields, interlacements, and random walk on $(\mathbb{Z}/N\mathbb{Z})^2$ *Probability Theory and Related Fields*, 173(3):1265–1299, April 2019.

[88] M. Schocker. The peak algebra of the symmetric group revisited. *Advances in Mathematics*, 192:259–309, April 2005.

[89] R. H. Schonmann. On the behavior of some cellular automata related to bootstrap percolation. *The Annals of Probability*, 20(1):174–193, January 1992.

[90] G. Seregin and V. Šverák. The Navier—Stokes Equations and Backward Uniqueness. In: Nonlinear problems in mathematical physics and relat topics, II, 353–366, Int. Math. Ser. (N. Y.), 2, Kluwer/Plenum, New York.

[91] A. Shamov. On Gaussian multiplicative chaos. *Journal of Functional Analysis*, 270(9):3224–3261, May 2016.

[92] J. R. Stembridge. Enriched P-Partitions. *Transactions of the American Mathematical Society*, 349(2):763–788, 1997.

[93] A. Sznitman. *Topics in Occupation Times and Gaussian Free Fields*. European Mathematical Society, 2012.

[94] S. Tanny. A probabilistic interpretation of Eulerian numbers. *Duke Mathematical Journal*, 40(4):717–722, December 1973.

[95] S. Ulam. Random processes and transformations. In *In Sets, Numbers, and Universes*, pages 326–337. MIT Press, 1974.

[96] A. C. D. van Enter. Proof of Straley's argument for bootstrap percolation. *Journal of Statistical Physics*, 48(3-4):943–945, August 1987.

[97] V. A. Vatutin. The numbers of ascending segments in a random permutation and in the inverse to it are asymptotically independent. *Discrete Mathematics and Applications*, 6(1):41–52, 2009.

[98] J. V. Neumann and A. W. Burks. *Theory of self-reproducing automata.* Urbana, University of Illinois Press, 1966.

[99] D. I. Warren and E. Seneta. Peaks and Eulerian Numbers in a Random Sequence. *Journal of Applied Probability*, 33(1):101–114, 1996.

[100] A. Zhai. Exponential concentration of cover times. *Electronic Journal of Probability*, 23, 2018.

[101] S. Ziesche. Sharpness of the phase transition and lower bounds for the critical intensity in continuum percolation on $\mathbb{R}^d$. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 54(2):866–878, May 2018.

[102] S. A. Zuev and A. F. Sidorenko. Continuous models of percolation theory. I. *Theoretical and Mathematical Physics*, 62(1):51–58, January 1985.