

UNIVERSALITY OF LOCAL STATISTICS FOR NONCOLLIDING RANDOM WALKS

BY VADIM GORIN^{*,‡,1} AND LEONID PETROV^{†,‡,2}

Massachusetts Institute of Technology^{}, University of Virginia[†] and IITP[‡]*

We consider the N -particle noncolliding Bernoulli random walk—a discrete time Markov process in \mathbb{Z}^N obtained from a collection of N independent simple random walks with steps $\in \{0, 1\}$ by conditioning that they never collide. We study the asymptotic behavior of local statistics of this process started from an arbitrary initial configuration on short times $T \ll N$ as $N \rightarrow +\infty$. We show that if the particle density of the initial configuration is bounded away from 0 and 1 down to scales $D \ll T$ in a neighborhood of size $Q \gg T$ of some location x (i.e., x is in the “bulk”), and the initial configuration is balanced in a certain sense, then the space-time local statistics at x are asymptotically governed by the extended discrete sine process (which can be identified with a translation invariant ergodic Gibbs measure on lozenge tilings of the plane). We also establish similar results for certain types of random initial data. Our proofs are based on a detailed analysis of the determinantal correlation kernel for the noncolliding Bernoulli random walk.

The noncolliding Bernoulli random walk is a discrete analogue of the $\beta = 2$ Dyson Brownian motion whose local statistics are universality governed by the continuous sine process. Our results parallel the ones in the continuous case. In addition, we naturally include situations with inhomogeneous local particle density on scale T , which nontrivially affects parameters of the limiting extended sine process, and in a particular case leads to a new behavior.

CONTENTS

1. Introduction	2687
Comparison with Dyson Brownian motion	2689
Method	2690
Outline	2690
2. Main results	2690
2.1. Determinantal structure	2690
2.2. Extended discrete sine kernel	2692

Received September 2016; revised May 2018.

¹Supported in part by NSF Grants DMS-1407562 and DMS-1664619 and by the Sloan Research Fellowship.

²Supported in part by the NSF grant DMS-1664617, and by NSF Grant PHY11-25915 through the participation in the KITP program “New approaches to non-equilibrium and random systems: KPZ integrability, universality, applications and experiments.”

MSC2010 subject classifications. Primary 60G55; secondary 60C05, 60B20.

Key words and phrases. Noncolliding random walks, determinantal point processes, bulk universality, discrete sine process, steepest descent method, Dyson’s conjecture, homogenization.

2.3. Bulk limit theorems 2694

2.4. Applications 2698

Notation 2701

3. From lozenge tilings to noncolliding walks: Proof of Theorem 2.1 2701

3.1. Random lozenge tilings 2701

3.2. Noncolliding paths of length L 2702

3.3. Limit $L \rightarrow \infty$ of noncolliding paths 2704

3.4. Limit $L \rightarrow \infty$ in the kernel 2706

4. Setup of the asymptotic analysis 2710

4.1. A change of variables 2710

4.2. Definition of the function $S(z)$ 2711

5. Existence of nonreal critical points: Proof of Proposition 2.6 2712

6. Asymptotics of the kernel: Proofs of Theorems 2.7, 2.10, 2.11 and 2.12 2719

6.1. Behavior of $\text{Im } S(z)$ and $\text{Re } S(z)$ 2720

6.2. Steepest descent/ascent contours in a large rectangle 2726

6.3. Completing the proof of Theorem 2.7 2728

6.4. Convergent initial data and proofs of Theorems 2.10, 2.11 and 2.12 2732

7. Applications: Proofs of Theorems 2.13, 2.15 and Proposition 2.16 2736

7.1. Discretization of a continuous profile: Proof of Theorem 2.13 2736

7.2. Random initial configuration 2737

7.3. Bernoulli initial data: Proof of Theorem 2.15 2739

7.4. Sine process initial data: Proof of proposition 2.16 2740

Appendix A: Determinantal kernels for other noncolliding processes 2742

A.1. Noncolliding Poisson random walk 2742

A.2. Dyson Brownian motion 2744

Appendix B: Representation-theoretic interpretation of noncolliding walks 2747

Acknowledgments 2750

References 2750

1. Introduction. Our main object is a discrete time Markov chain in the N -dimensional lattice \mathbb{Z}^N which is called the *noncolliding Bernoulli random walk*. At $N = 1$, by a single-particle chain, we mean the simple random walk on \mathbb{Z} which at each time step jumps by 1 in the positive direction with probability $\beta \in (0, 1)$, or stays put with the complementary probability $1 - \beta$. For $N > 1$, we consider N independent identical particles on \mathbb{Z} evolving according to the single-particle chain, and condition them to never collide (i.e., never occupy the same location of \mathbb{Z} at the same time). Note that the condition has probability zero and, therefore, needs to be defined through a limit procedure which is performed in, for example, [46] (based on a classical result of [40]). The result is a time and space homogeneous Markov chain $\vec{X}(t)$ living in the Weyl chamber

$$(1.1) \quad \mathbb{W}^N = \{(x_1, \dots, x_N) \in \mathbb{Z}^N : x_1 < x_2 < \dots < x_N\},$$

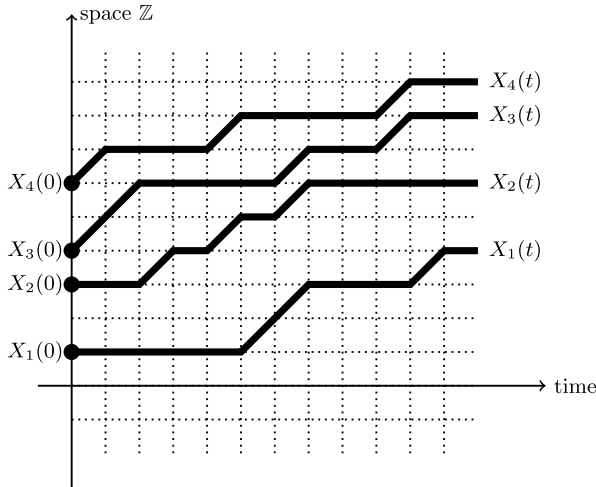


FIG. 1. *Noncolliding Bernoulli random walk of $N = 4$ particles started from the configuration $\vec{X}(0) = (1, 3, 4, 6) \in \mathbb{W}^4$.*

with transition probabilities

$$\begin{aligned}
 & \mathbb{P}(\vec{X}(t + 1) = \vec{x}' \mid \vec{X}(t) = \vec{x}) \\
 (1.2) \quad & = \begin{cases} \frac{V(\vec{x}')}{V(\vec{x})} \beta^{|\vec{x}'| - |\vec{x}|} (1 - \beta)^{N - |\vec{x}'| + |\vec{x}|} & \text{if } x'_i - x_i \in \{0, 1\} \text{ for all } i, \\ 0 & \text{otherwise,} \end{cases}
 \end{aligned}$$

where for $\vec{x} = (x_1, \dots, x_N)$ we denote $|\vec{x}| = x_1 + \dots + x_N$ and

$$(1.3) \quad V(\vec{x}) = \prod_{1 \leq j < i \leq N} (x_i - x_j).$$

We refer to Figure 1 for an illustration.

If instead of the simple random walk we start from the Brownian motion, then the same conditioning would lead to the celebrated ($\beta = 2$) *Dyson Brownian motion*, which plays a prominent role in the random matrix theory [1, 25, 30].³ Therefore, the noncolliding Bernoulli random walk can be viewed as a discrete version of the Dyson Brownian motion. There exists also an intermediate *semi-discrete* version related to the Poisson process; see Appendix A.

³The $\beta = 2$ Dyson Brownian motion is naturally associated to the Gaussian Unitary Ensemble (GUE)—one of the most classical objects in the study of random matrices. There is an extension of the ensemble and of the Dyson Brownian motion depending on a continuous parameter $\beta > 0$ (sometimes referred to as the inverse temperature). The noncolliding Brownian motions correspond to $\beta = 2$. The random-matrix parameter β has no connection to our β of Bernoulli random walk. It is inevitable for us to use both betas, as the latter β is rooted in the traditional notation in the asymptotic representation theory.

On the other hand, the noncolliding Bernoulli random walk can be coupled with a $(2 + 1)$ -dimensional interacting particle system with local push/block interactions [6]. The latter is linked to the totally asymmetric simple exclusion process and its relatives and to random lozenge and domino tilings. We refer to [6, 7, 13, 51] for details.

From yet another side, fixed time distributions of $\vec{X}(t)$ can be identified with coefficients in decompositions of tensor products of certain representations of unitary groups, which are of interest in the asymptotic representation theory; see Appendix B for details.

We concentrate on the local (“bulk”) limits of $\vec{X}(T)$ as both N and T tend to infinity. More precisely, we assume that $T \ll N$, which implies that the *global* profile $\vec{X}(T)$ is almost indistinguishable from the initial condition $\vec{X}(0)$. On the other hand, our main results, Theorems 2.7 and 2.10, show that under mild conditions on $\vec{X}(0)$ (see Assumptions 1, 2 in Section 2.3), the *local* characteristics of $\vec{X}(T)$ (such as, e.g., the asymptotic distribution of the distance between two adjacent particles) become universal: they depend on two real parameters which are computed by explicit formulas involving $\vec{X}(0)$.

In more detail, we prove that the one-dimensional point process describing the particles of $\vec{X}(T)$ as $T, N \rightarrow \infty$ converges to the *discrete sine process* of [10] (we recall its definition in Section 2.2). The two-dimensional point process describing the behavior of $\{\vec{X}(T + t)\}_t$ (where t is kept finite as $T \rightarrow \infty$) asymptotically becomes the *extended discrete sine process* [55], which can also be identified with a translation invariant ergodic Gibbs measure on lozenge tilings of the plane [43, 60].

As far as we know, in the *discrete setting* general results on the universal appearance of the discrete sine process were not available previously, and we are aware only of [31] where a related theorem is proven for random lozenge tilings. However, for *specific examples* (in our context this would mean considering very special initial conditions $\vec{X}(0)$ rather than general ones; note that the existing literature was mostly dealing with *other*, yet related discrete random systems) the appearance of the extended discrete sine process was observed by many authors; cf. [2, 9, 10, 16, 33, 38, 55, 56]. We expect that our results on local behavior of the noncolliding Bernoulli random walk can serve as a step toward establishing more general bulk universality results in discrete random systems.

Comparison with Dyson Brownian motion. In the continuous setting, the *universal* appearance of the *continuous* sine kernel process in bulk local limits of the Dyson Brownian motion is relatively well understood. It was first conjectured by Dyson [25] in the early 1960s that the universal statistics should already appear after *short* times. The first mathematical results in this direction were developed much later in [37], where the universal local behavior on *large* times was proven (using the contour integral formulas of [17] as an important ingredient). For the strongest results in this direction, see [59] and references therein.

The rigorous treatment of the short times is even more recent, with the best results appearing in [29, 47]. It should be noted that these results include cases other than the GUE ($\beta = 2$) one, and do not rely on explicit formulas specific to $\beta = 2$.

A detailed understanding of the bulk local behavior of the Dyson Brownian motion became a crucial step toward establishing bulk universality of generalized Wigner matrices and other random matrix ensembles; see [15, 29, 47], references therein, and the review [30]. See also [63] for an alternative approach to bulk universality of random matrices.

From this point of view, our results are parallel to the $\beta = 2$ Dyson Brownian motion developments as we prove a discrete analogue of the Dyson's conjecture. We also provide a generalization in a different direction and study the case when the local density of particles is not restricted to be the Lebesgue measure (as was usually assumed in the study of the Dyson Brownian motion), but can be quite general instead. This leads to new phenomena; see the end of Section 2.4 for one example.

Method. On the technical side, our approach starts from the double contour integral representation for the correlation kernel for the determinantal point process of uniformly random Gelfand–Tsetlin patterns of [56] (see also [23, 49]). We find a limit transition which turns these random Gelfand–Tsetlin patterns into $\vec{X}(t)$, and leads to formulas for the correlation functions of the latter process. These formulas are then analyzed using the steepest descent method. For this, we develop arguments working for general initial conditions $\vec{X}(0)$ rather than specific ones, and this requires a significant technical effort.

Outline. In Section 2, we formulate our main results and discuss their applications. In Section 3, we show how the noncolliding Bernoulli random walk can be obtained via a limit transition from the ensemble of uniformly random lozenge tilings of certain polygons. This leads to a double contour integral expression for the correlation kernel of the noncolliding Bernoulli random walk. Sections 4, 5 and 6 form the main technical part of the work and are devoted to the asymptotic analysis of the correlation kernel and of the noncolliding Bernoulli random walk. In Section 7, we prove the remaining statements from Section 2 which deal with various applications of our main bulk limit theorems.

We discuss degenerations of our kernel to the kernels for noncolliding Poisson processes and for the Dyson Brownian motion with arbitrary initial configurations in Appendix A. In Appendix B, we explain a representation-theoretic interpretation of discrete-space noncolliding random walks, and formulate a more general conjecture.

2. Main results.

2.1. *Determinantal structure.* Our first result is a formula for the determinantal correlation kernel of the noncolliding Bernoulli random walk. Recall that a par-

ticle dynamics $\vec{X}(t)$ is said to be a (dynamically) determinantal point process if its space-time correlations are given by determinants of a certain kernel $K(t, x; s, y)$:

$$\begin{aligned} &\mathbb{P}(\text{the particle configuration } \vec{X}(t_i) \text{ on } \mathbb{Z} \text{ contains the point } y_i \\ &\quad \text{for all } i = 1, \dots, k) \\ (2.1) \quad &= \det[K(t_\alpha, y_\alpha; t_\beta, y_\beta)]_{\alpha, \beta=1}^k, \end{aligned}$$

for any collection of pairwise distinct space-time points $(t_i, y_i) \in \mathbb{Z}_{\geq 0} \times \mathbb{Z}$, $i = 1, \dots, k$. In particular, when all t_i are the same (and are equal to t), we get a determinantal point process on \mathbb{Z} with the kernel $K_t(x; y) = K(t, x; t, y)$. General details on determinantal point processes can be found in, for example, the surveys [5, 34, 62].

THEOREM 2.1. *The noncolliding Bernoulli random walk with parameter $\beta \in (0, 1)$ started from any initial configuration $\vec{X}(0) = \vec{a} = (a_1 < \dots < a_N) \in \mathbb{W}^N$ is determinantal in the sense of (2.1), and its correlation kernel has the following form for $x_{1,2} \in \mathbb{Z}$, $t_{1,2} \in \mathbb{Z}_{\geq 1}$:⁴*

$$\begin{aligned} &K_{\vec{a}; \beta}^{\text{Bernoulli}}(t_1, x_1; t_2, x_2) \\ &= \mathbf{1}_{x_1 \geq x_2} \mathbf{1}_{t_1 > t_2} (-1)^{x_1 - x_2 + 1} \binom{t_1 - t_2}{x_1 - x_2} \\ &\quad + \frac{t_1!}{(t_2 - 1)! (2\pi \mathbf{i})^2} \int_{x_2 - t_2 + \frac{1}{2} - \mathbf{i}\infty}^{x_2 - t_2 + \frac{1}{2} + \mathbf{i}\infty} dz \oint_{\text{all } w \text{ poles}} dw \frac{(z - x_2 + 1)_{t_2 - 1}}{(w - x_1)_{t_1 + 1}} \\ (2.2) \quad &\times \frac{1}{w - z} \frac{\sin(\pi w)}{\sin(\pi z)} \left(\frac{1 - \beta}{\beta}\right)^{w - z} \prod_{r=1}^N \frac{z - a_r}{w - a_r}. \end{aligned}$$

The z integration contour is the straight vertical line $\text{Re } z = x_2 - t_2 + \frac{1}{2}$ traversed upward, and the w contour is a positively (counterclockwise) oriented circle or a union of two circles (this depends on the ordering of t_1, x_1, t_2 and x_2) encircling all the w poles $\{x_1 - t_1, x_1 - t_1 + 1, \dots, x_1 - 1, x_1\} \cap \{a_1, \dots, a_N\}$ of the integrand (except $w = z$); see Figure 2.

REMARK 2.2. When the N -point noncolliding Bernoulli random walk starts from the densely packed configuration $\vec{a} = (0, 1, 2, \dots, N - 1)$, the distribution of the N -point configuration $\vec{X}(t) \subset \mathbb{Z}$ at any time $t \in \mathbb{Z}_{\geq 0}$ is the Krawtchouk

⁴Throughout the text, $\mathbf{1}_A$ stands for the indicator function of an event A , $\mathbf{i} = \sqrt{-1}$, and we will employ the Pochhammer symbols $(z)_k = \frac{\Gamma(z+k)}{\Gamma(z)} = z(z+1) \cdots (z+k-1)$ and the binomial coefficients $\binom{k}{a} = (-1)^a \frac{(-k)_a}{a!}$, where $k, a \in \mathbb{Z}_{\geq 0}$.

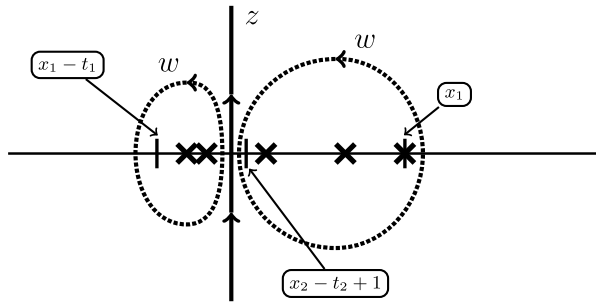


FIG. 2. Integration contours in (2.2). The poles are highlighted by crosses.

orthogonal polynomial ensemble [46]. Orthogonal polynomial ensembles are determinantal, and their correlation kernels are expressed through the corresponding univariate orthogonal polynomials—in our case, the Krawtchouk polynomials.⁵ This correlation kernel is explicit enough to be suitable for asymptotic analysis; see, for example, [35, 36]. The corresponding time-dependent kernel as in (2.1) is also explicitly known; it is the extended Krawtchouk kernel [38]. Theorem 2.1 generalizes these results to an arbitrary initial configuration $\vec{X}(0) = \vec{a} \in \mathbb{W}^N$.

We prove Theorem 2.1 in Section 3 below. In Appendix A, we also discuss two limits of the noncolliding Bernoulli random walk and the kernel (2.2):

- Noncolliding Poisson random walk—independent Poisson processes conditioned to never collide. This limit is obtained by rescaling time from discrete to continuous, and sending $\beta \rightarrow 0$.
- Dyson Brownian motion—independent Brownian motions conditioned to never collide. This process (introduced in [25]) is a diffusion limit of the noncolliding Bernoulli random walk. The correlation kernel for the Dyson Brownian motion started from an arbitrary initial configuration was first obtained in [37] (see also [59]). When the Dyson Brownian motion starts from a special initial condition $(0, 0, \dots, 0)$, its determinantal correlation kernel can be expressed through the Hermite orthogonal polynomials [48, 50].

2.2. *Extended discrete sine kernel.* Let us now discuss the point process describing the local asymptotic behavior the noncolliding Bernoulli random walk.

DEFINITION 2.3. By the *extended discrete sine process*⁶ of slope $u \in \mathbb{C}$, $\text{Im}(u) > 0$, we mean the determinantal point process on $\mathbb{Z} \times \mathbb{Z}$ with the correlation

⁵See Section 1.10 in [44] for definitions and basic properties of the Krawtchouk orthogonal polynomials, and [45] for a survey of orthogonal polynomial ensembles.

⁶Since in this paper we only discuss convergence to the discrete sine process and do not deal with its continuous counterpart, we often drop the word “discrete.”

kernel:

$$(2.3) \quad K_u(t, x; s, y) = -\frac{1}{2\pi i} \int_{\tilde{u}}^u (1-z)^{t-s} z^{-(x-y)-1} dz,$$

where the integration path crosses $(0, 1)$ for $t > s$ and $(-\infty, 0)$ for $t \leq s$.

The extended sine process was first introduced in [55] in relation to the bulk limit of random lozenge tilings (equivalently, 3D Young diagrams). In that paper, the kernel (2.3) was called the *incomplete beta kernel*. When $t = s$, the kernel (2.3) simplifies, after conjugation by the function $x \mapsto (-1)^x |u|^{-x}$,⁷ to the discrete sine kernel of [10]:

$$(2.4) \quad K_\phi^{\text{sine}}(x; y) = \frac{\sin(\phi(x-y))}{\pi(x-y)}, \quad x, y \in \mathbb{Z}, \phi = \pi - \arg(u),$$

with the agreement that $K_\phi^{\text{sine}}(x, x) = \phi/\pi$. The quantity $\phi/\pi \in (0, 1)$ is the density along the x direction of particles in the random configuration from the (extended) discrete sine process.

There exist other extensions of the discrete sine kernel (2.4); see [4, 11, 14]. In Appendix A.1, we briefly discuss the noncolliding Poisson random walk whose local statistics should be universally governed by an extension of (2.4) other than (2.3).

REMARK 2.4. The extended sine kernel was introduced in [55] in terms of *complementary* (to $\mathbb{Z} \times \mathbb{Z}$) configurations. The relation between kernels describing a configuration and its complement is ([10], Appendix A.3)

$$K^{\text{complement}}(t, x; s, y) = \mathbf{1}_{x=y} \mathbf{1}_{t=s} - K(t, x; s, y).$$

In the case of the extended sine kernel, the above delta function can be incorporated inside the contour integral by dragging for $t = s$ the contour through zero and picking the residue of $z^{-(x-y)-1}$, which is exactly $\mathbf{1}_{x=y}$.

The extended sine process is translation invariant in both directions (t and x), and it describes asymptotic bulk distribution of discrete two-dimensional determinantal point processes when both dimensions stay discrete in the limit. A characterization of the measure determined by K_z as a unique translation invariant Gibbs measure of a given complex slope was obtained in [43, 60].

⁷Transformations of a correlation kernel of the form $K(x, y) \mapsto \frac{f(x)}{f(y)} K(x, y)$ (where f is nowhere zero) not changing the correlation functions are sometimes referred to as the gauge transformations.

2.3. *Bulk limit theorems.* Here, we formulate our main asymptotic results—an approximation of the correlation kernel of the noncolliding Bernoulli random walk by the extended sine kernel (2.3), and a corresponding bulk local limit theorem.

Assume that N (the number of noncolliding particles) is our main parameter going to infinity, and that the time scale $T(N) \ll N$, $T(N) \rightarrow +\infty$ is fixed.⁸ For each $N = 1, 2, \dots$, we also fix an initial condition $\vec{X}(0) = \mathfrak{A}(N) = (a_1(N) < a_2(N) < \dots < a_N(N))$. We will often omit the dependence on N and simply write T (meaning $T(N)$) and a_i (meaning $a_i(N)$, $1 \leq i \leq N$), etc., when it leads to no confusion.

In what follows, we are describing the behavior of $\vec{X}(T + t)$ near the point $x = 0$. Since the definition of the noncolliding Bernoulli random walk is translation invariant, one can readily extract similar results on the behavior near an arbitrary point $x = x(N)$ by shifting $\mathfrak{A}(N)$ appropriately.

The following two assumptions will be imposed on $\mathfrak{A}(N)$ throughout the text.

ASSUMPTION 1 (Local density). There exist scales $D = D(N)$ satisfying $D(N) \ll T(N)$ and $Q = Q(N)$ satisfying $T(N) \ll Q(N) \ll N$, and absolute constants⁹ $0 < \rho_\bullet \leq \rho^\bullet < 1$, such that in every segment of length $D(N)$ inside $[-Q(N), Q(N)] \subset \mathbb{R}$ there are at least $\rho_\bullet D(N)$ and at most $\rho^\bullet D(N)$ points of the initial configuration $\mathfrak{A}(N)$.

ASSUMPTION 2 (Intermediate scales). For all $\delta > 0$, $R > 0$ and N large enough, one has

$$(2.5) \quad \left| \sum_{i: RT(N) \leq |a_i(N)| \leq \delta N} \frac{1}{a_i(N)} \right| \leq A_{R,\delta},$$

where $A_{R,\delta} > 0$ are absolute constants.

REMARK 2.5. Note that if (2.5) holds for some $\delta_0 > 0$, $R_0 > 0$, then it holds for all other $\delta > 0$, $R > 0$ because the difference of the sums in the left-hand side of (2.5) can be bounded by a part of the harmonic series (between $R_0 T$ and RT , etc.), which is bounded by a constant independent of N .

Both Assumptions 1 and 2 serve the same goal: we want to guarantee that the average density of particles in $\vec{X}(T)$ near $x = 0$ is bounded away from 0 and from 1, as otherwise the universal local behavior might fail. Assumption 1 simply bounds the density of the initial configuration $\mathfrak{A}(N)$, while Assumption 2 requires that the “densities” of the configuration $\mathfrak{A}(N)$ far (at scales from $RT(N)$ to δN) to the right and to the left of 0 (our point of observation) are “comparable.”

⁸Throughout the paper by $A(N) \ll B(N)$, we mean that $\lim_{N \rightarrow +\infty} A(N)/B(N) = 0$, and similarly for \gg .

⁹Here and below by an absolute constant, we mean a certain constant which does not depend on N , $T(N)$, or the initial configuration $\mathfrak{A}(N)$.

We do not impose any other requirements on particles far away from $x = 0$. We remark that one can easily cook up (e.g., using an intuition coming from the study of frozen boundaries in random tiling models) a situation in which Assumption 1 fails, and yet the average density of particles in $\bar{X}(T)$ close to 0 is still bounded away from 0 and 1, and the universal local behavior holds. In other words, Assumption 1 is not necessary. On the other hand, Assumption 2 is close to being necessary; see the discussion after Theorem 2.12 below. Overall, our Assumptions 1 and 2 are simple to state and straightforward to check in applications, while a full necessary and sufficient condition would be much more technical and involved. We refer to detailed analysis in models of random tilings performed in [21, 23, 24, 31, 56].

The following function of $z \in \mathbb{C}$ will play a prominent role in our asymptotic analysis:

$$(2.6) \quad S'(z) = \sum_{r=1}^N \frac{1}{Tz - a_r} - \text{p.v.} \sum_{j \in \mathcal{L}_T} \frac{1}{Tz - j} - \log(\beta^{-1} - 1),$$

where

$$(2.7) \quad \mathcal{L}_T = \mathcal{L}_{T(N)} = \{\dots, -T - 2, -T - 1, -T\} \cup \{0, 1, 2, \dots\},$$

and the infinite sum should be understood as its principal value, that is,

$$(2.8) \quad \text{p.v.} \sum_{j \in \mathcal{L}_T} \frac{1}{Tz - j} = \lim_{M \rightarrow \infty} \sum_{\substack{j \in \mathcal{L}_T \\ |j| < M}} \frac{1}{Tz - j}.$$

We can alternatively write

$$(2.9) \quad S'(z) = \sum_{r=1}^N \frac{1}{Tz - a_r} + \sum_{i=1}^{T-1} \frac{1}{Tz + i} - \pi \cot(\pi Tz) - \log(\beta^{-1} - 1)$$

using the fact that

$$(2.10) \quad \pi \cot(\pi z) = \text{p.v.} \sum_{k \in \mathbb{Z}} \frac{1}{z - k},$$

which follows from the Euler’s product formula for the sine function

$$(2.11) \quad \sin(\pi z) = \pi z \prod_{k=1}^{\infty} \left(1 - \frac{z^2}{k^2}\right).$$

PROPOSITION 2.6. *Under Assumptions 1 and 2 there exists N_0 such that for all $N > N_0$ the equation $S'(z) = 0$ has a unique root $z_c = z_c(N)$ in the upper half-plane (i.e., satisfying $\text{Im}(z_c) > 0$). Moreover, there exists a compact set $\mathcal{Z} \subset \{z \in \mathbb{C} : \text{Im}(z) > 0\}$ such that $z_c \in \mathcal{Z}$ for all $N > N_0$. The set \mathcal{Z} and the constant N_0 may depend on constants in Assumptions 1 and 2 but not on the choice of $\mathfrak{A}(N)$.*

THEOREM 2.7. *Under Assumptions 1 and 2, for any fixed $t_1, x_1, t_2, x_2 \in \mathbb{Z}$ we have*

$$K_{\mathfrak{A}(N);\beta}^{\text{Bernoulli}}(t_1 + T(N), x_1; t_2 + T(N), x_2) = K_{z_c(N)/(z_c(N)+1)}(t_1, x_1; t_2, x_2) + o(1)$$

as $N \rightarrow \infty$, where $z_c(N)$ is the unique root provided by Proposition 2.6. The remainder $o(1)$ admits a tending to 0 bound which may depend on constants in Assumptions 1 and 2 but not on the choice of $\mathfrak{A}(N)$.

REMARK 2.8. Since all probabilities describing the local behavior of $\vec{X}(T + t)$ near $x = 0$ (with t kept finite) are expressed through $K_{\mathfrak{A}(N);\beta}^{\text{Bernoulli}}(t_1 + T, x_1; t_2 + T, x_2)$ via (2.1), Theorem 2.7 means that as $N \rightarrow +\infty$, locally near $x = 0$ the distribution of $\{\vec{X}(T + t)\}_t$ becomes indistinguishable from the one corresponding to the extended sine process.

If $\mathfrak{A}(N)$ depends on N in a regular way, then Theorem 2.7 leads to a convergence statement.

DEFINITION 2.9. We say that a sequence $\mu_k, k = 1, 2, \dots$ of σ -finite measures on \mathbb{R} vaguely converges to μ , if for any continuous function f with compact support, we have

$$\lim_{k \rightarrow \infty} \int_{-\infty}^{+\infty} f(x) \mu_k(dx) = \int_{-\infty}^{+\infty} f(x) \mu(dx).$$

Let us also denote

$$(2.12) \quad d_N(\mathbb{R}) = \sum_{i: |a_i(N)| \geq RT(N)} \frac{1}{a_i(N)}.$$

Assumption 2 is equivalent to the boundedness of the $d_N(\mathbb{R})$'s for fixed \mathbb{R} , uniformly in N .

THEOREM 2.10. *Suppose that a sequence $\mathfrak{A}(N), N = 1, 2, \dots$ is such that Assumptions 1 and 2 hold, there exists a σ -finite measure μ_{loc} for which vaguely*

$$\lim_{N \rightarrow +\infty} \frac{1}{T(N)} \sum_{i=1}^N \delta_{a_i(N)/T(N)} = \mu_{\text{loc}},$$

and there exists a limit $d(\mathbb{R}) = \lim_{N \rightarrow +\infty} d_N(\mathbb{R})$.

Then the point process describing $\{\vec{X}(T(N) + t)\}_t$ near $x = 0$ converges in distribution to the extended sine process of complex slope u_* , in the sense that for each $t_1, x_1, t_2, x_2 \in \mathbb{Z}$ we have

$$\lim_{N \rightarrow \infty} K_{\mathfrak{A}(N);\beta}^{\text{Bernoulli}}(t_1 + T(N), x_1; t_2 + T(N), x_2) = K_{u_*}(t_1, x_1; t_2, x_2),$$

where $u = u_* \in \mathbb{C}$ is a unique root in the upper half-plane of the equation¹⁰

$$(2.13) \quad \int_{-\infty}^{+\infty} \left(\frac{1-u}{u-(1-u)v} + \frac{\mathbf{1}_{|v|>R}}{v} \right) \mu_{\text{loc}}(dv) - \log u = d(R) + \log(\beta^{-1} - 1) - i\pi.$$

In fact, the additional hypotheses in Theorem 2.10 as compared to Theorem 2.7 are not too restrictive; see Remark 6.10 below.

The condition that the quantities (2.12) converge can sometimes be not easy to verify, and the determination of the limit $d(R)$ could be even harder. Let us present a sufficient condition and a way to compute $d(R)$ which involves the global profile $\vec{X}(0)$.

THEOREM 2.11. *Suppose that a sequence $\mathfrak{A}(N)$, $N = 1, 2, \dots$ is such that Assumptions 1 and 2 hold and, moreover, $\lim_{\delta \rightarrow 0} \mathbf{A}_{R,\delta} = 0$ in Assumption 2. Next, let there exist a σ -finite measure μ_{loc} and a probability measure μ_{glob} for which vaguely*

$$(2.14) \quad \lim_{N \rightarrow +\infty} \frac{1}{T(N)} \sum_{i=1}^N \delta_{a_i(N)/T(N)} = \mu_{\text{loc}},$$

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{i=1}^N \delta_{a_i(N)/N} = \mu_{\text{glob}},$$

and the principal value integral $\text{p.v.} \int_{-\infty}^{\infty} v^{-1} \mu_{\text{glob}}(dv)$ exists. Then the quantities $d_N(R)$ converge to this integral (so $d(R) = d$ is independent of R), and the conclusion of Theorem 2.10 holds with

$$(2.15) \quad d = \text{p.v.} \int_{-\infty}^{\infty} \frac{\mu_{\text{glob}}(dv)}{v}.$$

In many applications μ_{loc} is a multiple of the Lebesgue measure on \mathbb{R} , in which case u_* is more explicit.

THEOREM 2.12. *Assume that Theorem 2.10 holds with μ_{loc} being $q \in (0, 1)$ times the Lebesgue measure on \mathbb{R} . Then $d(R) = d$ does not depend on R , and the complex slope of the limiting extended sine process is given by*

$$(2.16) \quad u_* = \frac{\beta e^{-d}}{1 - \beta} e^{i\pi(1-q)}.$$

¹⁰The equation and the root u_* do not depend on $R > 0$.

Let us make some remarks about the elegant formula (2.16). First, the (same-time) density of particles under the limiting extended sine process is $1 - \arg(u_*)/\pi = q$, as it should be. In particular, this density does not depend on the “speed” β of the noncolliding random walk, or on the parameter d capturing the effect of the global profile.

To isolate the effect of the global profile, observe that the second parameter $|u_*|$ of the extended sine kernel can be rewritten as

$$(2.17) \quad |u_*| = \frac{\beta_{\text{eff}}}{1 - \beta_{\text{eff}}}, \quad \beta_{\text{eff}} = \frac{1}{1 + e^d(\beta^{-1} - 1)}.$$

That is, for fixed q the bulk local distribution is the same as if the parameter β was replaced by β_{eff} , while the contribution from the global profile (encoded by d) was not present.

The quantity β_{eff} increases in β and decreases in d . The dependence on d can be interpreted as an effect of repulsion. For example, having much more particles of the initial configuration to the right of 0 than to the left corresponds to larger values of d , which leads to a decrease in β_{eff} .

Moreover, if d is very large or very small, then β_{eff} is close to 0 or 1, respectively, leading to an almost deterministic behavior of the noncolliding paths in the bulk local limit. This suggests that our Assumption 2 is close to being necessary for the universal local bulk behavior. Namely, if it is violated, then $d = \pm\infty$ along a subsequence $\{N_k\}$, and so the local bulk distribution is not described by the universal extended sine kernel. However, we will not pursue this analysis further.

2.4. Applications. Let us give several examples which demonstrate that Assumptions 1 and 2 are checkable in applications. The first example deals with an arbitrary smooth deterministic initial configuration $\bar{X}(0)$.

THEOREM 2.13. *Take a twice continuously differentiable function f on $[-\frac{1}{2}, \frac{1}{2}]$ such that $f'(x) > 1$ for all $x \in [-\frac{1}{2}, \frac{1}{2}]$, and $f(-\frac{1}{2}) < 0 < f(\frac{1}{2})$. Let $\chi \in (-\frac{1}{2}, \frac{1}{2})$ be the unique point where $f(\chi) = 0$. Assume for simplicity that N is odd, and let the initial configuration of the noncolliding Bernoulli random walk be¹¹*

$$(2.18) \quad a_i(N) = \lfloor Nf(i/N) \rfloor, \\ i = -\frac{N-1}{2}, -\frac{N-1}{2} + 1, \dots, \frac{N-1}{2} - 1, \frac{N-1}{2}.$$

Fix any $0 < \eta < 1$, and let $T(N) = \lfloor N^\eta \rfloor$. Then Theorem 2.12 is applicable, where μ_{loc} is the Lebesgue measure on \mathbb{R} times $q = 1/f'(\chi)$, and d has the form

$$(2.19) \quad d = \text{p.v.} \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{dx}{f(x)}.$$

¹¹Throughout the text, $\lfloor \cdot \rfloor$ denotes the integer part.

REMARK 2.14. In the situation of Theorem 2.13, the global probability measure exists and has the form $\mu_{\text{glob}}(dv) = dv/f'(f^{-1}(v))$. The expression (2.15) for d is equivalent to (2.19) via a change of variables.

The next two examples deal with *random* initial configurations $\vec{X}(0)$.

THEOREM 2.15. Fix $0 < p < 1$ and $0 < \alpha < 1$. For $M = 1, 2, \dots$, consider a particle configuration on $\{-\lfloor M(1 - \alpha) \rfloor, -\lfloor M(1 - \alpha) \rfloor + 1, \dots, \lfloor M\alpha \rfloor - 1, \lfloor M\alpha \rfloor\}$ obtained by putting a particle at each location with probability p independently of all others. Let \mathbf{N} be the (random) number of particles in this configuration, and $\mathfrak{A}(\mathbf{N})$ denote the configuration itself. By $\vec{X}(t)$, denote the non-colliding Bernoulli random walk started from $\mathfrak{A}(\mathbf{N})$. Choose $0 < \eta < 1$ and set $T(M) = \lfloor M^\eta \rfloor$.

Then the point process $\{\vec{X}(T(M) + t)\}_t$ converges near $x = 0$ to the extended sine process as in Theorem 2.10, where μ_{loc} is p times the Lebesgue measure on \mathbb{R} , and $d(\mathbb{R}) = d = p \log(\frac{\alpha}{1-\alpha})$. That is, the complex slope of the limiting extended sine process is

$$u_* = \frac{\beta}{1 - \beta} \left(\frac{1 - \alpha}{\alpha}\right)^p e^{i\pi(1-p)}.$$

PROPOSITION 2.16. Fix $0 < \phi < \pi$ and $0 < \alpha < 1$. For $M = 1, 2, \dots$, let the initial particle configuration of the noncolliding Bernoulli random walk be obtained by restricting the configuration of the discrete sine process of density ϕ/π (i.e., with the correlation kernel K_ϕ^{sine} given by (2.4)) to $\{-\lfloor M(1 - \alpha) \rfloor, -\lfloor M(1 - \alpha) \rfloor + 1, \dots, \lfloor M\alpha \rfloor - 1, \lfloor M\alpha \rfloor\}$.

With other notation the same as in Theorem 2.15, the point process $\{\vec{X}(T(M) + t)\}_t$ describing the configuration of the noncolliding walk started from the sine process initial configuration converges near $x = 0$ to the extended sine process as in Theorem 2.10, where μ_{loc} is $\frac{\phi}{\pi}$ times the Lebesgue measure on \mathbb{R} , and $d(\mathbb{R}) = d = \frac{\phi}{\pi} \log(\frac{\alpha}{1-\alpha})$, so

$$u_* = \frac{\beta}{1 - \beta} \left(\frac{1 - \alpha}{\alpha}\right)^{\frac{\phi}{\pi}} e^{i(\pi-\phi)},$$

is the complex slope of the limiting extended sine process.

REMARK 2.17. Proposition 2.16 is formulated for time $T(M) = \lfloor M^\eta \rfloor$ going to infinity. However, it is probably true even for $T(M) = 0$ because the initial configuration is already close to the same-time configuration of the extended sine process. Here, we do not pursue in this direction.

For the last example, let us consider an initial configuration $\vec{X}(0)$ for which μ_{loc} differs from the Lebesgue measure.

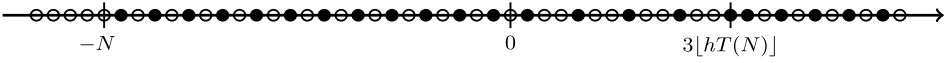


FIG. 3. Initial condition in Proposition 2.18. Black dots mean particles.

PROPOSITION 2.18. Fix two parameters $0 < \eta < 1$ and $h > 0$. Set $T(N) = \lfloor N^\eta \rfloor$. For $N = 1, 2, \dots$, let $\mathfrak{A}(N)$ be the N -particle configuration defined by the following three conditions (cf. Figure 3):

- There are $\lfloor N/2 \rfloor$ particles to the left from the origin and they occupy every second lattice site;
- $\lfloor hT(N) \rfloor$ particles adjacent to the origin from the right occupy every third lattice site;
- The remaining particles are to the right from $3\lfloor hT(N) \rfloor$ and they occupy every second lattice site.

Then the point process $\{\vec{\mathbf{X}}(T(M) + t)\}_t$ converges near $x = 0$ to the extended sine process as in Theorem 2.10. The complex slope u_* of the limiting extended sine process is a unique point in the upper half-plane satisfying

$$(2.20) \quad \sqrt[6]{1 - 3h \frac{1 - u_*}{u_*}} = i u_* (1 - \beta^{-1}),$$

where the 6th degree root is understood in the sense of the principal branch.

The behavior of u_* given by (2.20) as a function of h is quite interesting. When h is small, $u_* \approx i \frac{\beta}{1-\beta}$, matching Theorem 2.12. On the other hand, as $h \rightarrow +\infty$, u_* goes to infinity in the direction $+i\infty$. This leads to a behavior which seems new. Namely, as $h \rightarrow +\infty$, the same-time local configuration around zero is still governed by the discrete sine kernel (2.4) (with $\phi = \pi/2$), while the time-dependent extension of this process is *deterministic*: at each discrete time step each particle always goes to the right by 1 and does not stay put. Heuristically, for very large h the repelling force coming from the higher density region to the left of the origin is so large that this creates a deterministic flow of particles.

It is likely that with a proper time rescaling one can find a more delicate $h \rightarrow +\infty$ limit in which the paths make rare jumps, as in the classical transition from the random walk to the Poisson process. We will not pursue this direction here.

REMARK 2.19. In this observation, we first took the large N limit, and then a degeneration in the parameter h . This is the reason why the same-time distribution stays universal as $h \rightarrow +\infty$. We believe that if instead $h = h(N)$ goes to infinity in a certain way, then the limiting local configuration around zero would become *completely deterministic*: particles would occupy every other site, and at each time step go to the right by 1. This combined limit does not follow from Proposition 2.18, and we will not consider it further.

Notation. Throughout the paper, C stands for positive constants whose values may change from line to line. These constants might depend on the parameters of the model (and our assumptions about them), but not on variables going to zero or infinity.

3. From lozenge tilings to noncolliding walks: Proof of Theorem 2.1.

3.1. *Random lozenge tilings.* Consider uniformly random tilings (by lozenges of three types) of polygons drawn on the triangular lattice; see Figure 4, left. The asymptotic behavior of such tilings in various regimes has been studied in [8, 20, 32, 33, 41, 42, 52, 56, 58].¹²

We will employ a result of [56] (see also [23]) on the determinantal structure of uniformly random lozenge tilings of polygons such as in Figure 4, left. That is, consider a trapezoid $\mathcal{T}_{N,L}$ of height $L \in \mathbb{Z}_{\geq 1}$ with vertices $(\frac{1}{2}, 0)$, $(\frac{1}{2}, L)$, $(\frac{1}{2} - N, L)$ and $(\frac{1}{2} - N - L, 0)$. Fix $(y_1 < \dots < y_L) \in \{0, -1, \dots, -N - L + 1\}$ and put lozenges of type \blacklozenge at each of the y_i 's cutting L small triangles at the base of the trapezoid. We will consider tilings of the resulting figure by lozenges of three types. The assumption that the y_i 's belong to $\{0, -1, \dots, -N - L + 1\}$ (and hence are all negative) is not essential since the whole situation is translation invariant. However, this will be convenient in Section 3.3 when discussing the limit as $L \rightarrow +\infty$.

REMARK 3.1. Putting the lozenges \blacklozenge at the y_i 's fixes locations of some other of the lozenges of the same type (the darker ones in Figure 4, left). In this way, the

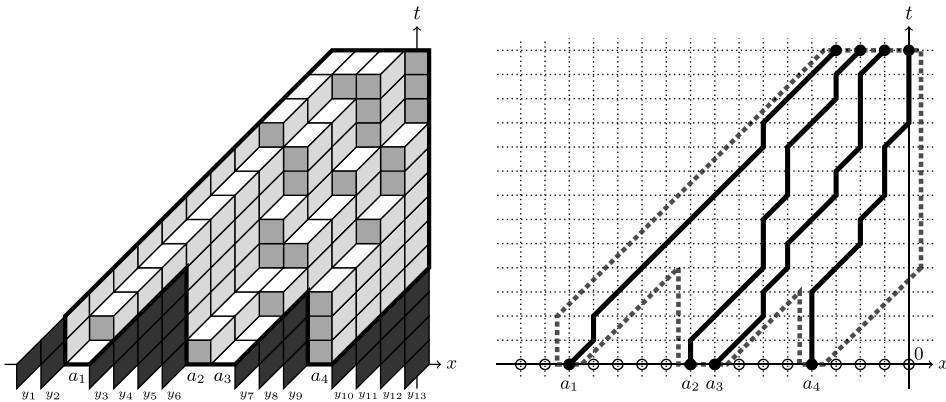


FIG. 4. A lozenge tiling of (left) and its bijective encoding as a collection of noncolliding paths (right). The trapezoid $\mathcal{T}_{4,13}$ has height $L = 13$, and the number of paths ($N = 4$) is the same as the length of the top side of $\mathcal{T}_{4,13}$.

¹²Note that we are using an affine transform of the regular triangular lattice, thus our picture differs from some of the references cited. This is done for a better coordinate notation in our situation.

tiling that we are describing can be alternatively interpreted at a lozenge tiling of a certain polygon.

The total number of such tilings is equal to (e.g., see Section 2 in [13])

$$(3.1) \quad Z_{\{y_1, \dots, y_L\}}^{N,L} = \prod_{1 \leq i < j \leq L} \frac{y_j - y_i}{j - i}.$$

We interpret the uniformly random tiling as a random particle configuration by looking at centers of the lozenges of type \blacklozenge (there are $L(L + 1)/2$ lozenges in total). The centers have integer coordinates $(x, t) \in \mathbb{Z}_{\leq 0} \times \mathbb{Z}_{\geq 0}$.

THEOREM 3.2. *The uniformly random tiling described above gives rise to a determinantal point process, that is,*

$$\begin{aligned} & \mathbb{P}(\text{there are lozenges } \blacklozenge \text{ at locations } (x_i, t_i), i = 1, \dots, k) \\ &= \det[K_{N; \{y_1, \dots, y_L\}}^{\text{tilings}}(t_\alpha, x_\alpha; t_\beta, x_\beta)]_{\alpha, \beta=1}^k, \end{aligned}$$

with the correlation kernel

$$(3.2) \quad \begin{aligned} & K_{N; \{y_1, \dots, y_L\}}^{\text{tilings}}(t_1, x_1; t_2, x_2) \\ &= -\mathbf{1}_{t_1 < t_2} \mathbf{1}_{x_2 \leq x_1} \frac{(x_1 - x_2 + 1)_{t_2 - t_1 - 1}}{(t_2 - t_1 - 1)!} \\ & \quad + \frac{t_1!}{(t_2 - 1)!} \frac{1}{(2\pi i)^2} \oint_{c(x_2 - t_2 + 1)} dz \oint_{C(\infty)} dw \frac{(z - x_2 + 1)_{t_2 - 1}}{(w - x_1)_{t_1 + 1}} \\ & \quad \times \frac{1}{w - z} \prod_{r=1}^L \frac{w - y_r}{z - y_r}. \end{aligned}$$

Here, $x_1, x_2 \in \mathbb{Z}, 0 \leq t_1 \leq L - 1, 1 \leq t_2 \leq L - 1$, and the integration contours look as in Figure 5:

- The z contour $c(x_2 - t_2 + 1)$ is positively oriented and encircles the points $x_2 - t_2 + 1, x_2 - t_2 + 2, \dots, y_L$, and does not encircle $x_2 - t_2, x_2 - t_2 - 1, \dots$;
- The w contour $C(\infty)$ is positively oriented and encircles $c(x_2 - t_2 + 1)$ and is sufficiently large so that to include all w poles $x_1, x_1 - 1, \dots, x_1 - t_1$.

This is the same kernel as in [56], Theorem 5.1, up to the change of coordinates $t_{1,2} = L - n_{1,2}$.

3.2. Noncolliding paths of length L . Let us keep L and y_1, \dots, y_L fixed and consider another interpretation of a lozenge tiling in terms of noncolliding paths as

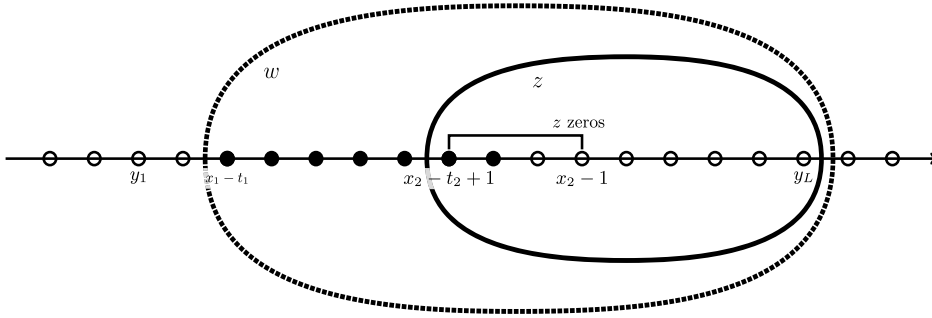


FIG. 5. Integration contours for the kernel $K_{N;\{y_1, \dots, y_L\}}^{\text{tilings}}$ (3.2). The poles at $w = x_1 - t_1, x_1 - t_1 + 1, \dots, x_1 - 1, x_1$ are highlighted in black.

in Figure 4, right. There are N such paths; they trace lozenges of two types other than \blacklozenge , start at points

$$(3.3) \quad \vec{a} = (a_1 < \dots < a_N) = \{0, -1, \dots, -N - L + 1\} \setminus \{y_1, \dots, y_L\}$$

on the $t = 0$ horizontal, and end at points $-N + 1, \dots, -1, 0$, respectively, on the $t = L$ horizontal. Denote the path configuration at height t by $X_1^{(L)}(t) < \dots < X_N^{(L)}(t)$. Thus, we obtain a random path ensemble $\{\vec{X}^{(L)}(t)\}_{t=0}^L$ corresponding to our random tiling.

The correlation kernel for the paths $\vec{X}^{(L)}$ (whose configuration is complementary to the configuration of the \blacklozenge lozenges) can be obtained from the one for the tilings (3.2) by a particle-hole involution (e.g., see [10], Appendix A.3):

$$\begin{aligned} \mathbb{P}(\text{paths } \vec{X}^{(L)} \text{ pass through points } (x_i, t_i), i = 1, \dots, k) \\ = \det[K_{L;\vec{a}}^{\text{paths}}(t_\alpha, x_\alpha; t_\beta, x_\beta)]_{\alpha, \beta=1}^k \end{aligned}$$

with

$$(3.4) \quad K_{L;\vec{a}}^{\text{paths}}(t_1, x_1; t_2, x_2) = \mathbf{1}_{x_1=x_2} \mathbf{1}_{t_1=t_2} - K_{N;\{y_1, \dots, y_L\}}^{\text{tilings}}(t_1, x_1; t_2, x_2),$$

as long as the pairwise distinct points $(x_i, t_i), i = 1, \dots, k$, are inside the trapezoid $\mathcal{T}_{N,L}$ defined in Section 3.1 above.

REMARK 3.3. The reason for the last restriction (that the observation points (x_i, t_i) belong to $\mathcal{T}_{N,L}$) is because the particle-hole involution of the configuration of lozenges gives rise not only to the N noncolliding paths $\vec{X}^{(L)}$ but also to infinitely many trivial paths connecting $(j, 0)$ to $(j, L), j \geq 1$, and $(j, 0)$ to $(j + L, L), j \leq -N - L$. These paths correspond to a unique way of extending the lozenge tiling of our polygon to the infinite horizontal strip of height L with L small triangles added at the bottom. Therefore, to capture the correlation structure of the nontrivial paths $\vec{X}^{(L)}$, the points (x_i, t_i) should be inside $\mathcal{T}_{N,L}$.

REMARK 3.4. When the starting configuration \vec{a} is the densely packed one $(0, 1, \dots, N - 1)$, the polygon which is tiled reduces to the hexagon. In this case, the distribution of the Markov process $\{\vec{X}^{(L)}(t)\}_{t=0}^L$ is described by the extended Hahn kernel expressed through the Hahn orthogonal polynomials; see [39]. Asymptotic analysis of the noncolliding paths in the hexagon utilizing this representation of the kernel was performed in [2, 33].

3.3. *Limit $L \rightarrow \infty$ of noncolliding paths.* We will now perform a limit transition of our uniformly random tilings to the noncolliding Bernoulli random walks. Fix $\beta \in (0, 1)$, $N \in \mathbb{Z}_{\geq 1}$, and $\vec{a} = (a_1 < \dots < a_N) \in \mathbb{W}^N$. Start the path ensemble $\{\vec{X}^{(L)}(t)\}_{t=0}^L$ from the points

$$(3.5) \quad \vec{a}^{(L)} = (a_1 - \lfloor \beta L \rfloor, \dots, a_N - \lfloor \beta L \rfloor).$$

Clearly, for L large enough we have

$$-N - L + 1 < a_i - \lfloor \beta L \rfloor < 0$$

for all i ; cf. (3.3). Thus, for any fixed $\vec{a} \in \mathbb{W}^N$ the uniformly random tiling and the path ensemble $\{\vec{X}^{(L)}(t)\}_{t=0}^L$ are well defined for large L .

The above shifting of the initial configuration of $\vec{X}^{(L)}$ forces the N noncolliding paths to have asymptotic speed β . This leads to the noncolliding Bernoulli random walk with parameter β :

PROPOSITION 3.5. *As $L \rightarrow \infty$, all finite-dimensional distributions of the path ensemble*

$$\{\vec{X}^{(L)}(t) + \lfloor \beta L \rfloor\}_{t=0}^L$$

converge to those of the noncolliding Bernoulli random walk $\vec{X}(t)$ (defined in the Introduction) started from the configuration \vec{a} .

PROOF. Because the random lozenge tiling used to construct the path ensemble $\vec{X}^{(L)}$ is picked uniformly, $\{\vec{X}^{(L)}(t)\}_{t=0}^L$ can be viewed as a Markov process (with time bounded by L). Indeed, the uniformity ensures that the past and the future are conditionally independent given the present configuration ([61], Chapter I.12).

Observe that each conditional probability

$$(3.6) \quad \mathbb{P}(\vec{X}^{(L)}(t + 1) = \vec{b}' \mid \vec{X}^{(L)}(t) = \vec{b}), \quad \vec{b}, \vec{b}' \in \mathbb{W}^N$$

is simply the ratio of the number of lozenge tilings of a polygon of height $L - (t + 1)$ with the bottom boundary determined by \vec{b}' similar to (3.3), and the number of tilings of a polygon of height $L - t$ with the bottom boundary corresponding to \vec{b} . It suffices to show that the transition probabilities (3.6) converge to the transition probabilities (1.2) of the noncolliding Bernoulli random walk.

Let us fix $\vec{b}, \vec{b}' \in \mathbb{W}^N$ such that $b'_i - b_i \in \{0, 1\}$ for all i . (Clearly, if these conditions do not hold, then the transition probabilities from \vec{b} to \vec{b}' in both $\vec{X}^{(L)}(t) + \lfloor \beta L \rfloor$ and $\vec{X}(t)$ vanish.) We have

$$\begin{aligned} \mathbb{P}(\vec{X}^{(L)}(t+1) = \vec{b}' - \lfloor \beta L \rfloor \mid \vec{X}^{(L)}(t) = \vec{b} - \lfloor \beta L \rfloor) &= \\ &= \frac{Z_{\{m'_1, \dots, m'_{L-t-1}\}}^{N, L-t-1}}{Z_{\{m_1, \dots, m_{L-t}\}}^{N, L-t}} = \prod_{1 \leq i < j \leq L-t-1} \frac{m'_j - m'_i}{j - i} \prod_{1 \leq i < j \leq L-t} \frac{j - i}{m_j - m_i} \\ &= (L - t - 1)! \frac{\prod_{1 \leq i < j \leq L-t-1} (m'_j - m'_i)}{\prod_{1 \leq i < j \leq L-t} (m_j - m_i)}, \end{aligned}$$

where we used (3.1), and

$$\begin{aligned} &\{m_1, \dots, m_{L-t}\} \\ &= \{0, -1, \dots, -N - L + t + 1\} \setminus \{b_1 - \lfloor \beta L \rfloor, \dots, b_N - \lfloor \beta L \rfloor\}, \\ &\{m'_1, \dots, m'_{L-t-1}\} \\ &= \{0, -1, \dots, -N - L + t + 2\} \setminus \{b'_1 - \lfloor \beta L \rfloor, \dots, b'_N - \lfloor \beta L \rfloor\}. \end{aligned}$$

Above we have assumed that L is large enough so that all polygons are well defined. We can write (using the notation of (1.3))

$$\begin{aligned} &\frac{\prod_{1 \leq i < j \leq L-t-1} (m'_j - m'_i)}{\prod_{1 \leq i < j \leq L-t} (m_j - m_i)} \\ &= \frac{V(-N - L + t + 2, \dots, -1, 0)}{V(-N - L + t + 1, -N - L + t + 2, \dots, -1, 0)} \frac{V(\vec{b})}{V(\vec{b}')} \\ &\quad \times \prod_{i=1}^N \left(\prod_{\substack{-N-L+t+2 \leq j \leq 0 \\ j \notin \vec{b}' - \lfloor \beta L \rfloor}} |b'_i - \lfloor \beta L \rfloor - j|^{-1} \right. \\ &\quad \left. \times \prod_{\substack{-N-L+t+1 \leq j \leq 0 \\ j \notin \vec{b} - \lfloor \beta L \rfloor}} |b_i - \lfloor \beta L \rfloor - j| \right). \end{aligned}$$

We have $\frac{V(-N-L+t+2, \dots, -1, 0)}{V(-N-L+t+1, -N-L+t+2, \dots, -1, 0)} = \frac{1}{(N+L-t-1)!}$. Next, let us insert the products over $i \neq j$ of $|b_i - b_j|/|b'_i - b'_j|$ into the big product in the previous

formula. We obtain

$$\begin{aligned} & \frac{V(\vec{b})}{V(\vec{b}')} \prod_{i=1}^N \left(\prod_{\substack{-N-L+t+2 \leq j \leq 0 \\ j \notin \vec{b}' - \lfloor \beta L \rfloor}} |b'_i - \lfloor \beta L \rfloor - j|^{-1} \right. \\ & \times \left. \prod_{\substack{-N-L+t+1 \leq j \leq 0 \\ j \notin \vec{b} - \lfloor \beta L \rfloor}} |b_i - \lfloor \beta L \rfloor - j| \right) \\ & = \frac{V(\vec{b}')}{V(\vec{b})} \prod_{i=1}^N \prod_{j=-N-L+t+2}^0 \frac{|b_i - \lfloor \beta L \rfloor - j|^{1_{b_i - \lfloor \beta L \rfloor \neq j}}}{|b'_i - \lfloor \beta L \rfloor - j|^{1_{b'_i - \lfloor \beta L \rfloor \neq j}}} \\ & \times \prod_{i=1}^N |b_i - \lfloor \beta L \rfloor + N + L - t - 1|. \end{aligned}$$

Using the well-known asymptotics for the Gamma function ([28], 1.18.(5)),

$$(3.7) \quad \frac{\Gamma(L + \alpha)}{\Gamma(L)} \sim L^\alpha, \quad L \rightarrow +\infty$$

(where α is fixed), let us collect the last product and the factors involving factorials and write

$$\lim_{L \rightarrow +\infty} \frac{(L - t - 1)!}{(N + L - t - 1)!} \prod_{i=1}^N |b_i - \lfloor \beta L \rfloor + N + L - t - 1| = (1 - \beta)^N.$$

Let us turn to the remaining factors. We have to following equivalence as $L \rightarrow +\infty$:

$$\prod_{i=1}^N \prod_{j=-N-L+t+2}^0 \frac{|b_i - \lfloor \beta L \rfloor - j|^{1_{b_i - \lfloor \beta L \rfloor \neq j}}}{|b'_i - \lfloor \beta L \rfloor - j|^{1_{b'_i - \lfloor \beta L \rfloor \neq j}}} \sim \prod_{i=1}^N \frac{\Gamma(\beta L - b_i) \Gamma(b_i + (1 - \beta)L)}{\Gamma(\beta L - b'_i) \Gamma(b'_i + (1 - \beta)L)}.$$

If $b'_i = b_i$, the corresponding term is simply 1, and otherwise it converges to $\beta/(1 - \beta)$ due to (3.7).

Collecting all the terms we see that the transition probabilities of $\vec{X}^{(L)}$ converge to those of the noncolliding random walk. This completes the proof. \square

3.4. *Limit $L \rightarrow \infty$ in the kernel.* Let us now take the $L \rightarrow \infty$ limit in the kernel for the process $\vec{X}^{(L)}(t)$ coming from the uniformly random tilings. The latter kernel is given by (3.2) and (3.4).

PROPOSITION 3.6. *Fix $\beta \in (0, 1)$ and $\vec{a} \in \mathbb{W}^N$, and consider the correlation kernel of the process $\vec{X}^{(L)}(t)$ started from the shifted initial configuration $\vec{a}^{(L)}$*

as in (3.5). Then for any fixed¹³ $t_1 \in \mathbb{Z}_{\geq 0}$, $t_2 \in \mathbb{Z}_{\geq 1}$, and $x_{1,2} \in \mathbb{Z}$ we have the convergence

$$\lim_{L \rightarrow +\infty} K_{L; \vec{a}^{(L)}}^{\text{paths}}(t_1, x_1 - \lfloor \beta L \rfloor; t_2, x_2 - \lfloor \beta L \rfloor) = K_{\vec{a}; \beta}^{\text{Bernoulli}}(t_1, x_1; t_2, x_2),$$

where $K_{\vec{a}; \beta}^{\text{Bernoulli}}$ is given by (2.2).

Proposition 3.6 together with Proposition 3.5 will imply Theorem 2.1.

PROOF OF PROPOSITION 3.6. We first focus on the part of $K_{L; \vec{a}^{(L)}}^{\text{paths}}(t_1, x_1 - \lfloor \beta L \rfloor; t_2, x_2 - \lfloor \beta L \rfloor)$ containing the double contour integral. By substituting the shifted parameters $\vec{a}^{(L)}$ and $x_{1,2} - \lfloor \beta L \rfloor$ into the kernel given by (3.2), (3.4) and at the same time shifting both integration variables z, w by $\lfloor \beta L \rfloor$ turns the double contour integral into

$$\begin{aligned} I^{\text{paths}} &= \frac{t_1!}{(t_2 - 1)!} \frac{1}{(2\pi i)^2} \oint_{c(x_2 - t_2 + 1)} dz \oint_{C(\infty)} dw \frac{(z - x_2 + 1)_{t_2 - 1}}{(w - x_1)_{t_1 + 1}} \\ (3.8) \quad &\times \frac{1}{w - z} \frac{(w - \lfloor \beta L \rfloor)_{N+L}}{(z - \lfloor \beta L \rfloor)_{N+L}} \prod_{r=1}^N \frac{z - a_r}{w - a_r}. \end{aligned}$$

(Note that this integral enters $K_{L; \vec{a}^{(L)}}^{\text{paths}}$ with a negative sign which we ignore for now.) Here, the z contour $c(x_2 - t_2 + 1)$ encircles the points $x_2 - t_2 + 1, x_2 - t_2 + 2, \dots, \lfloor \beta L \rfloor$ and not $x_2 - t_2 - 1, x_2 - t_2 - 2, \dots$, while the w contour $C(\infty)$ encircles $c(x_2 - t_2 + 1)$ and all the w poles of the integrand. For large enough L , these w poles are contained inside the intersection $\{x_1 - t_1, x_1 - t_1 + 1, \dots, x_1 - 1, x_1\} \cap \{a_1, \dots, a_N\}$; see Figure 6.

Let us split the w integration over $C(\infty)$ into integration over two contours: one encircling all the w poles outside the z contour $c(x_2 - t_2 + 1)$ (denote it by $c'(x_2 - t_2)$), and the other one encircling just the z contour $c(x_2 - t_2 + 1)$ (denote it by $c_{\text{out}}(x_2 - t_2 + 1)$). In this second integral, we will drag the w contour inside the z contour at the cost of picking the residue at $w = z$. Denote the resulting w contour by $c_{\text{in}}(x_2 - t_2 + 1)$; see Figure 7. Thus, (3.8) can be rewritten as follows:

$$\begin{aligned} I^{\text{paths}} &= \frac{t_1!}{(t_2 - 1)!} \frac{1}{(2\pi i)^2} \oint_{c(x_2 - t_2 + 1)} dz \oint_{c'(x_2 - t_2) \cup c_{\text{in}}(x_2 - t_2 + 1)} dw \frac{1}{w - z} \\ &\times \frac{(z - x_2 + 1)_{t_2 - 1}}{(w - x_1)_{t_1 + 1}} \frac{(w - \lfloor \beta L \rfloor)_{N+L}}{(z - \lfloor \beta L \rfloor)_{N+L}} \prod_{r=1}^N \frac{z - a_r}{w - a_r} \\ (3.9) \quad &+ \frac{t_1!}{(t_2 - 1)!} \frac{1}{2\pi i} \oint_{c(x_2 - t_2 + 1)} \frac{(z - x_2 + 1)_{t_2 - 1}}{(z - x_1)_{t_1 + 1}} dz. \end{aligned}$$

¹³Clearly, under our scaling the restrictions on the variables in the kernel $K_{L; \vec{a}^{(L)}}^{\text{paths}}$ imposed in Section 3.2 (cf. Remark 3.3) will eventually disappear.

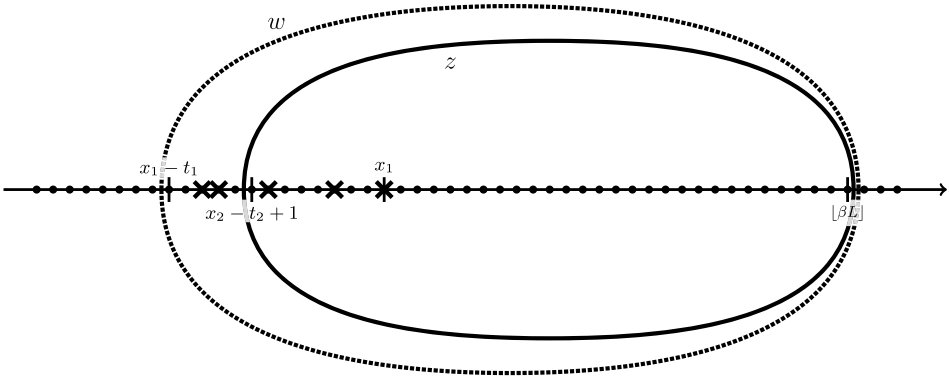


FIG. 6. Integration contours for I^{paths} (3.8). Note that the z contour grows with L but the w poles (highlighted by crosses) do not depend on L .

The single integral in (3.9) can be evaluated using the results of Section 6.2 in [56]. It is equal to

$$(3.10) \quad \mathbf{1}_{x_1 \geq x_2} \frac{(t_2 - t_1)_{x_1 - x_2}}{(x_1 - x_2)!}.$$

In the double contour integral in (3.9), we first note that for L fixed but large enough, due to the presence of the polynomial $(z - \lfloor \beta L \rfloor)_{N+L}$ in the denominator, the integrand decays rapidly as $z \rightarrow \infty$. Thus, the z integration contour can be replaced by the vertical line from $x_2 - t_2 + \frac{1}{2} - i\infty$ to $x_2 - t_2 + \frac{1}{2} + i\infty$ traversed from bottom to top, yielding a new minus sign in front of the double contour integral (cf. Figure 2).

The z integral over the vertical line converges uniformly in L . Indeed, observe that

$$\left| \frac{w + k}{x + iy + k} \right| < \frac{1}{1 + C|y|/|k|}, \quad k \in \mathbb{Z},$$

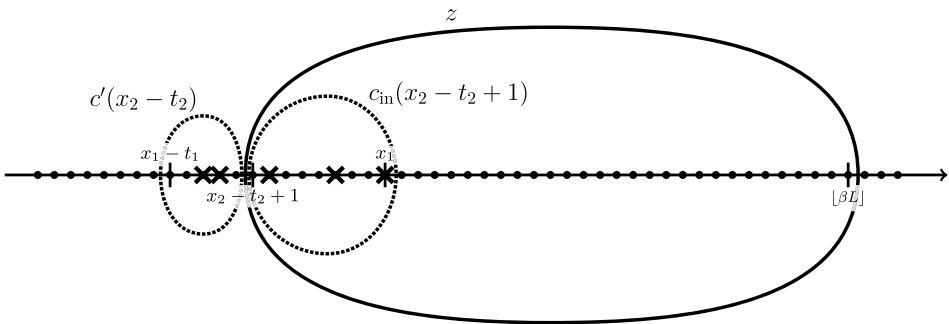


FIG. 7. Integration contours in the double integral in (3.9).

for some $C > 0$, where $z = x + iy$, and C is uniform in w belonging to a bounded contour. For large $|y|$ and $L > L_0$, the product of the above quantities over k as in (3.9) can be bounded by a constant independent of L times a fixed (but arbitrarily large) negative power of $|y|$. Here, we also used the fact that for fixed y the infinite product over k diverges to infinity.

Thus, the integration contours do not depend on L , and we can pass to a pointwise limit as $L \rightarrow \infty$ in the integrand. Since $w, z \notin \mathbb{Z}$ on our contours, we can write

$$\begin{aligned}
 & \frac{(w - \lfloor \beta L \rfloor)_{N+L}}{(z - \lfloor \beta L \rfloor)_{N+L}} \\
 &= \frac{\Gamma(w + L - \lfloor \beta L \rfloor + N)}{\Gamma(z + L - \lfloor \beta L \rfloor + N)} \frac{\Gamma(z - \lfloor \beta L \rfloor)}{\Gamma(w - \lfloor \beta L \rfloor)} \\
 (3.11) \quad &= \frac{\Gamma(w + L - \lfloor \beta L \rfloor + N)}{\Gamma(z + L - \lfloor \beta L \rfloor + N)} \frac{\Gamma(-w + 1 + \lfloor \beta L \rfloor)}{\Gamma(-z + 1 + \lfloor \beta L \rfloor)} \frac{\sin(\pi w)}{\sin(\pi z)},
 \end{aligned}$$

where in the second equality we used

$$(3.12) \quad \Gamma(u) = \frac{\pi}{\sin(\pi u)\Gamma(1 - u)}.$$

Let us employ the Stirling asymptotics for the Gamma function ([28], 1.18.(2)–(3)), which can be formulated as

$$(3.13) \quad \Gamma(L + \alpha) = (1 + O(L^{-1}))\sqrt{2\pi} \exp\left(\left(L + \alpha - \frac{1}{2}\right) \log L - L\right),$$

$L \rightarrow +\infty,$

where $\alpha \in \mathbb{C}$ if fixed and the remainder $O(L^{-1})$ is uniform in α belonging to compact subsets of \mathbb{C} . Thus, continuing (3.11),

$$\frac{(w - \lfloor \beta L \rfloor)_{N+L}}{(z - \lfloor \beta L \rfloor)_{N+L}} = \frac{\sin(\pi w)}{\sin(\pi z)} \left(\frac{1 - \beta}{\beta}\right)^{w-z} (1 + O(L^{-1})).$$

Finally, the summands not involving contour integrals coming from (3.2), (3.4) and (3.10) can be simplified as

$$\begin{aligned}
 & \mathbf{1}_{x_1=x_2} \mathbf{1}_{t_1=t_2} + \mathbf{1}_{t_1 < t_2} \mathbf{1}_{x_2 \leq x_1} \frac{(x_1 - x_2 + 1)_{t_2-t_1-1}}{(t_2 - t_1 - 1)!} - \mathbf{1}_{x_1 \geq x_2} \frac{(t_2 - t_1)_{x_1-x_2}}{(x_1 - x_2)!} \\
 &= \mathbf{1}_{x_1 \geq x_2} \mathbf{1}_{t_1 > t_2} (-1)^{x_1-x_2+1} \binom{t_1 - t_2}{x_1 - x_2}
 \end{aligned}$$

(note that all of them involve only the difference $x_1 - x_2$ which is not affected by the shift by $\lfloor \beta L \rfloor$). This coincides with the summand not containing integrals in (2.2). This completes the proof of Proposition 3.6, and hence of Theorem 2.1. \square

REMARK 3.7. The argument in the above proof implies in particular that the integration in $K_{\vec{a};\beta}^{\text{Bernoulli}}$ (2.2) can be alternatively performed over a shifted contour z . This contour can be shifted as far as to the vertical line traversed from $x_2 - \frac{1}{2} - i\infty$ to $x_2 - \frac{1}{2} + i\infty$. Indeed, the difference between the two expressions is equal to the residue at $z = w$ integrated over a certain part of the w contour; it is the same as the single integral in (3.9) but over a contour which does not contain any poles inside, and thus vanishes.

4. Setup of the asymptotic analysis. Here, we explain the relevance of the function $S'(z)$ defined in (2.6) for the asymptotics of the correlation kernel of the noncolliding Bernoulli random walks.

4.1. *A change of variables.* Changing the variables as $z = t_2z + x_2$, $w = t_2w + x_2$ and employing the shorthand notation

$$(4.1) \quad \Delta t = t_1 - t_2, \quad \Delta x = x_1 - x_2$$

turns the kernel (2.2) of the noncolliding Bernoulli random walk into

$$\begin{aligned}
 & K_{\vec{a};\beta}^{\text{Bernoulli}}(t_1, x_1; t_2, x_2) \\
 &= \mathbf{1}_{\Delta x \geq 0} \mathbf{1}_{\Delta t > 0} (-1)^{\Delta x + 1} \binom{\Delta t}{\Delta x} \\
 &+ \frac{1}{(2\pi i)^2} \int_{-1 + \frac{1}{2}t_2^{-1} - i\infty}^{-1 + \frac{1}{2}t_2^{-1} + i\infty} dz \oint_{\text{all } w \text{ poles}} dw \frac{(t_2 + \Delta t)! \cdot t_2}{(t_2 - 1)!} \\
 &\times \frac{(t_2z + 1)_{t_2 - 1}}{(t_2w - \Delta x)_{t_2 + \Delta t + 1}} \\
 (4.2) \quad &\times \frac{1}{w - z} \frac{\sin(\pi t_2 w)}{\sin(\pi t_2 z)} \left(\frac{1 - \beta}{\beta}\right)^{t_2(w - z)} \prod_{r=1}^N \frac{t_2z + x_2 - a_r}{t_2w + x_2 - a_r}.
 \end{aligned}$$

Here, z is integrated over a vertical line (which crosses the real line to the right of -1), and the w integration contour (a circle or a union of two circles, cf. Figure 2) must encircle all the w poles of the integrand except $w = z$. Note that now these poles all belong to $\{-1 + t_2^{-1}(\Delta x - \Delta t), \dots, t_2^{-1}(\Delta x - 1), t_2^{-1}\Delta x\}$.

From (4.2), we see that by shifting the initial data $\vec{a} \in \mathbb{W}^N$ it is possible to take $x_2 = 0$. Since the initial data is arbitrary and its finite shifts do not change our Assumptions 1 and 2, throughout the sequel without loss of generality we may and will assume that $x_2 = 0$, and so $x_1 = \Delta x \in \mathbb{Z}$ is fixed throughout the analysis. Moreover, since we aim to study the asymptotic behavior of $K_{\mathfrak{A}(N);\beta}^{\text{Bernoulli}}(t_1 + T(N), x_1; t_2 + T(N), 0)$ (cf. Theorem 2.7) and finite shifts in the t parameters can be incorporated into $T = T(N)$, we may also assume that $t_2 = T$ and $t_1 = \Delta t + T$, where $\Delta t \in \mathbb{Z}$ is fixed.

4.2. *Definition of the function S(z).* With the notation explained in Section 4.1, rewrite the integrand in (4.2) (without $1/(w - z)$) as follows:

$$(4.3) \quad \frac{(T + \Delta t)! \cdot T}{(T - 1)!} \frac{(Tz + 1)_{T-1}}{(Tw - \Delta x)_{T+\Delta t+1}} \frac{\sin(\pi Tw)}{\sin(\pi Tz)} \left(\frac{1 - \beta}{\beta}\right)^{T(w-z)} \prod_{r=1}^N \frac{Tz - a_r}{Tw - a_r} \\ = \exp\{T(S(z) - S(w))\} \frac{(T + \Delta t)! \cdot T}{(T - 1)!} \frac{(Tw + 1)_{T-1}}{(Tw - \Delta x)_{T+\Delta t+1}},$$

where

$$(4.4) \quad S(z) = \frac{1}{T} \sum_{r=1}^N \log\left(z - \frac{a_r}{T}\right) + \frac{1}{T} \sum_{i=1}^{T-1} \log\left(z + \frac{i}{T}\right) \\ - \frac{1}{T} \log(\sin(\pi Tz)) - z \log(\beta^{-1} - 1).$$

Let us discuss the choice of branches of the logarithms. Because $S(z)$ is exponentiated in (4.3), different choices of branches lead to the same integrand. However, a certain particular choice makes $S(z)$ holomorphic in the upper half-plane $\mathbb{H} = \{z \in \mathbb{C} : \text{Im } z > 0\}$, which will be convenient in Section 6. Let us restrict our attention to \mathbb{H} , the situation in the lower half-plane is analogous (however, one clearly cannot choose a branch making $S(z)$ holomorphic in the whole complex plane).

The standard branch of the logarithm, denoted by $\log z$, has the cut along the negative real axis, and takes positive real values for real $z > 1$. Let $\log_{\mathbb{H}} z$ denote a branch in the upper half-plane which extends holomorphically to $\mathbb{R} \setminus \{0\}$ and has the cut along the negative imaginary axis:

$$\log_{\mathbb{H}} z = \log(ze^{-i\pi/2}) + i\pi/2.$$

For $z \in \mathbb{H}$, the branches $\log z$ and $\log_{\mathbb{H}} z$ coincide. We will use $\log_{\mathbb{H}}$ for the logarithms of $z - a_r/T$ and $z + i/T$ in (4.4). Next, simply plugging $\sin(\pi Tz)$ into any of these logarithms does not produce a continuous function in \mathbb{H} . Let us use (2.11) instead, and define

$$(4.5) \quad \log(\sin \pi z)_{\mathbb{H}} = \log(\pi z) + \sum_{k=1}^{\infty} (\log(1 + z/k) + \log(1 - z/k)).$$

In the right-hand side, the logarithms are standard, and we mean direct substitution. The series in k converges for any fixed $z \in \mathbb{H}$ because it is bounded by the sum of C/k^2 . One can check that alternatively (4.5) can be written as

$$(4.6) \quad \log(\sin \pi z)_{\mathbb{H}} = \log_{\mathbb{H}}(\sin \pi z) - 2\pi i \left[\frac{1}{2} \text{Re}(z) + \frac{1}{2} \right],$$

where $\log_{\mathbb{H}}(\sin \pi z)$ is the direct substitution. This expression provides a holomorphic continuation of $\log(\sin \pi z)_{\mathbb{H}}$ into $\mathbb{R} \setminus \mathbb{Z}$. From (4.6), it readily follows that

$$(4.7) \quad \log(\sin(\pi(x + iy)))_{\mathbb{H}} = -i\pi(x + iy) + i\pi/2 - \log 2 + o(1), \quad y \rightarrow +\infty,$$

uniformly in $x \in \mathbb{R}$ (the remainder $o(1)$ is periodic in x).

Therefore, the function $S(z)$ takes the form

$$(4.8) \quad S(z) = \frac{1}{T} \sum_{r=1}^N \log_{\mathbb{H}}\left(z - \frac{a_r}{T}\right) + \frac{1}{T} \sum_{i=1}^{T-1} \log_{\mathbb{H}}\left(z + \frac{i}{T}\right) - \frac{1}{T} \log_{\mathbb{H}}(\sin(\pi Tz)) + \frac{2\pi i}{T} \left[\frac{1}{2} \operatorname{Re}(Tz) + \frac{1}{2} \right] - z \log(\beta^{-1} - 1).$$

With these choices of branches it becomes holomorphic in \mathbb{H} , and extends to $z \in \mathbb{R}$ everywhere except the singularities. Recall the notation $\mathcal{L}_T = \mathcal{L}_{T(N)}$ (2.7) and $\mathfrak{A} = \mathfrak{A}(N)$, and denote

$$(4.9) \quad \begin{aligned} \iota_T &= \iota_{T(N)} = \left\{ \dots, -1 - \frac{2}{T}, -1 - \frac{1}{T}, -1 \right\} \cup \left\{ 0, \frac{1}{T}, \frac{2}{T}, \dots \right\}, \\ \mathfrak{a} &= \mathfrak{a}(N) = \left\{ \frac{a_1}{T}, \frac{a_2}{T}, \dots, \frac{a_N}{T} \right\}. \end{aligned}$$

The set of (nonremovable) singularities of $S(z)$ is $\iota_T \Delta \mathfrak{a}$ (the symmetric difference) because of the cancellations in (4.8) with the help of (4.5).

The function $S'(z)$ defined by (2.6) is simply the derivative of $S(z)$. Note that this derivative does not depend on choices of the branches.

We will study the asymptotic behavior of the kernel (4.2) by means of the steepest descent method. That is, we will find *critical points* of the function $S(z)$ (i.e., where $S'(z) = 0$) and deform the contours so that they pass through these critical points and are steepest descent for $\operatorname{Re} S(z)$ (i.e., $\operatorname{Re} S(z)$ on these contours decreases or increases the most). As was first noted in [53], Section 3.2, having a pair of nonreal complex conjugate simple critical points z_c and \bar{z}_c (plus certain properties of the integration contours) leads to the discrete sine kernel.

5. Existence of nonreal critical points: Proof of Proposition 2.6. In this section, we deal with properties of $S'(z)$ (2.6), and prove Proposition 2.6 (stating that the equation $S'(z) = 0$ has a unique root $z_c = z_c(N)$ in \mathbb{H} , and it is uniformly bounded away from the real line and infinity) through a series of lemmas.

LEMMA 5.1. *The equation $S'(z) = 0$ has at most one pair of nonreal complex conjugate roots.*

PROOF. The sum over $|j| < M$ in (2.8) converges, as $M \rightarrow +\infty$, uniformly on compact sets in \mathbb{C} to the corresponding principal value sum (i.e., the left-hand side of (2.8)). Therefore, by Hurwitz’s theorem, for the purpose of counting critical points it is enough to prove that the following equation (approximating $S'(z) = 0$):

$$(5.1) \quad \sum_{r=1}^N \frac{1}{Tz - a_r} - \sum_{j \in \mathcal{L}_T \cap \{-M, \dots, M\}} \frac{1}{Tz - j} = \log(\beta^{-1} - 1),$$

has at most one pair of nonreal complex roots for all large enough M .

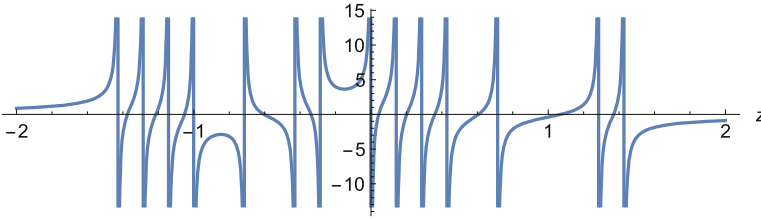


FIG. 8. Plot of the left-hand side of (5.1) as a function of $z \in \mathbb{R}$. The parameters are $\mathfrak{A} = (-5, -3, -2, 4, 6, 7, 8)$, $T = 7$, $M = 10$ and $d = 14$. Note that for $\beta \neq \frac{1}{2}$ equation (5.1) has one extra real root belonging to one of the semi-infinite rays.

Let d be the size of $(\mathfrak{L}_T \Delta \mathfrak{A}) \cap \{-M, \dots, M\}$, this is the number of poles in the left-hand side of (5.1) after canceling out equal terms with opposite signs. Multiplying by the common denominator turns equation (5.1) into a polynomial equation of degree d if $\beta \neq \frac{1}{2}$, and of degree $d - 1$ otherwise (when the logarithm in the right-hand side vanishes).

Let us demonstrate that (5.1) already has at least $d - 3$ real roots. The left-hand side of (5.1) has d poles which divide the real line onto $d - 1$ segments of finite length, plus two semi-infinite rays. These d poles are of two types (see Figure 8 for an example):

- For all $a_i \in \{-T + 1, \dots, -1\}$, the pole comes from the term $\frac{1}{Tz - a_i}$.
- For points of $\{-M, \dots, -T - 1, -T\} \cup \{0, 1, \dots, M\}$ which are not equal to any a_i the pole comes from the term $-\frac{1}{Tz - \ell}$ of the opposite sign.

Clearly, on a segment between any two poles of the same sign the left-hand side of (5.1) takes all values between $-\infty$ and $+\infty$, and thus equation (5.1) has at least one root on this segment. Among the $d - 1$ segments of finite length, at most two have endpoints which are singularities of different types, and thus the presence of a root there is not guaranteed. Thus, there are at least $d - 3$ real roots. Because the coefficients of (5.1) are real, its nonreal roots come in complex conjugate pairs, and so this equation cannot have more than one such pair of nonreal roots. \square

Lemma 5.1 implies that there is at most one critical point in the upper half-plane. In the rest of this section, we show its existence, and obtain a more precise control on the position of this critical point. The complex equation $S'(z) = 0$ (2.6) is equivalent to a pair of real equations in $z = x + iy$, $x \in \mathbb{R}$, $y \in \mathbb{R}_{>0}$:

$$\begin{aligned}
 0 &= \text{Im } S'(x + iy) \\
 (5.2) \quad &= -\frac{1}{T} \sum_{r=1}^N \frac{y}{y^2 + (x - a_r/T)^2} + \frac{1}{T} \sum_{j \in \mathfrak{L}_T} \frac{y}{y^2 + (x - j/T)^2}, \\
 0 &= \text{Re } S'(x + iy)
 \end{aligned}$$

$$\begin{aligned}
 &= \sum_{r=1}^N \frac{1}{T} \frac{x - a_r/T}{y^2 + (x - a_r/T)^2} \\
 (5.3) \quad &- \text{p.v.} \sum_{j \in \mathcal{L}_T} \frac{1}{T} \frac{x - j/T}{y^2 + (x - j/T)^2} - \log(\beta^{-1} - 1).
 \end{aligned}$$

Note that the infinite sum in (5.2) is absolutely convergent, while in (5.3) we need to use the principal value summation.

We start from (5.2), and rewrite it in a more compact form. For a discrete subset $U \subset \mathbb{R}$, define the atomic measure

$$(5.4) \quad \mathcal{M}_T[U] = \frac{1}{T} \sum_{u \in U} \delta_u$$

(note that it is not necessarily a probability or even a finite measure), and denote by

$$(5.5) \quad \mathcal{C}_y(u) = \frac{y}{\pi(y^2 + u^2)}$$

the Cauchy probability density on \mathbb{R} rescaled by $y > 0$. Using this notation, rewrite (5.2) as

$$(5.6) \quad 0 = \frac{1}{\pi} \text{Im } \mathcal{S}'(x + iy) = -(\mathcal{M}_T[\mathfrak{a}] * \mathcal{C}_y)(x) + (\mathcal{M}_T[\mathfrak{l}_T] * \mathcal{C}_y)(x),$$

where “*” means the usual convolution of measures.

LEMMA 5.2. *Under Assumptions 1 and 2, for each $0 < \delta < 1$ there exists $\varepsilon_0 > 0$ (which may depend on constants in our assumptions but not on the choice of $\mathfrak{A}(N)$), such that for any $0 < \varepsilon < \varepsilon_0$ there is $N_0 \in \mathbb{Z}_{\geq 1}$, and for all $N > N_0$ (see Figure 9):*

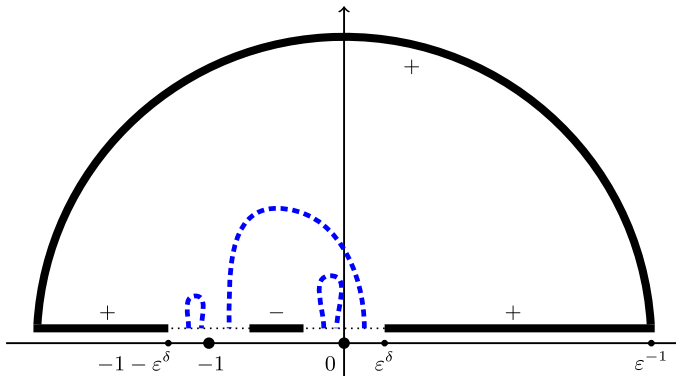


FIG. 9. Signs of $\text{Im } \mathcal{S}'(z)$ along the curves described in Lemma 5.2. Blue dashed curves represent a possible part of the boundary of the set \mathcal{U} inside \mathcal{D} ; see Lemma 5.4 below.

- $\text{Im } S'(x + iy) > 0$ for all (x, y) such that $\sqrt{x^2 + y^2} = \varepsilon^{-1}$ and $y \geq \varepsilon$, and for all $(x, y) = (x, \varepsilon)$, where $x \in [-\varepsilon^{-1}, -1 - \varepsilon^\delta] \cup (\varepsilon^\delta, \varepsilon^{-1}]$;
- $\text{Im } S'(x + iy) < 0$ for all $(x, y) = (x, \varepsilon)$, where $x \in (-1 + \varepsilon^\delta, -\varepsilon^\delta)$.

REMARK 5.3. The presence of δ in this lemma and Lemma 5.4 below is not essential. However, δ is put here to better link these statements with Lemma 5.5 below where δ plays an important role.

PROOF OF LEMMA 5.2. Fix $\varepsilon > 0$. As N (and thus T) grows, the absolutely convergent sum $(\mathcal{M}_T[l_T] * \mathcal{C}_y)(x)$ is a Riemann sum for the corresponding integral, and it approximates the integral uniformly on compact subsets of the upper half-plane (and in particular, for (x, y) in each of the sets described in the hypotheses of the lemma). Thus, for any $c > 0$ there exists N_0 such that for all $N > N_0$,

$$\left| (\mathcal{M}_T[l_T] * \mathcal{C}_y)(x) - \frac{1}{\pi} \left(\int_{-\infty}^{-1} + \int_0^{\infty} \right) \frac{y \, du}{y^2 + (x - u)^2} \right| < c.$$

The integral above can be explicitly evaluated, it is equal to

$$1 + \frac{1}{\pi} \tan^{-1} \left(\frac{x}{y} \right) - \frac{1}{\pi} \tan^{-1} \left(\frac{x + 1}{y} \right).$$

For small y , this expression is close to 1 if $x \in (-\infty, -1) \cup (0, +\infty)$, and close to 0 if $x \in (-1, 0)$. Moreover, for $\sqrt{x^2 + y^2} = \varepsilon^{-1}$ and $y \geq \varepsilon$ this expression is close to 1, too.

Let us now deal with $(\mathcal{M}_T[a] * \mathcal{C}_y)(x)$, which enters (5.2) with a negative sign. We aim to show that this sum is bounded away from 0 and 1, which will imply the claim. Use Assumption 1 and take N so large that $Q > 2T\varepsilon^{-1}$. Throw away summands for which $|a_i| > 2T\varepsilon^{-1}$, and then split the segment $(-2T\varepsilon^{-1}, 2T\varepsilon^{-1})$ into $4T\varepsilon^{-1}/D$ segments of the form $(-2T\varepsilon^{-1} + jD, -2T\varepsilon^{-1} + (j + 1)D)$, each of which contains at least $\rho_\bullet D$ points from the configuration \mathfrak{A} . On each of these segments, replace the summands $\frac{1}{\pi T} \frac{y}{y^2 + (x - a_r/T)^2}$ by $\rho_\bullet D$ times the minimum of $\frac{1}{\pi T} \frac{y}{y^2 + (x - a/T)^2}$ over a belonging to the corresponding segment. This allows us to estimate $(\mathcal{M}_T[a] * \mathcal{C}_y)(x)$ from below by a Riemann sum of the integral

$$\rho_\bullet \int_{-2\varepsilon^{-1}}^{2\varepsilon^{-1}} \frac{1}{\pi} \frac{y \, du}{y^2 + (x - u)^2} = \frac{\rho_\bullet}{\pi} \left[\tan^{-1} \left(\frac{2\varepsilon^{-1} - x}{y} \right) + \tan^{-1} \left(\frac{2\varepsilon^{-1} + x}{y} \right) \right]$$

within error $O(T^{-1}\varepsilon^{-1})$ which goes to zero. For $y = \varepsilon$, the expression in the square brackets is close to π , and for $\sqrt{x^2 + y^2} = \varepsilon^{-1}$ and $y \geq \varepsilon$ it is $\geq \frac{\pi}{2}$. Therefore, $(\mathcal{M}_T[a] * \mathcal{C}_y)(x) \geq \frac{\rho_\bullet}{2}$.

The other estimate is obtained in a similar manner but now we assume that all locations outside $(-2T\varepsilon^{-1}, 2T\varepsilon^{-1})$ are occupied by particles from the configuration \mathfrak{A} . This allows us to write

$$\begin{aligned} (\mathcal{M}_T[\mathbf{a}] * \mathcal{C}_y)(x) &\leq \frac{\rho^\bullet}{\pi} \int_{-2\varepsilon^{-1}}^{2\varepsilon^{-1}} \frac{y \, du}{y^2 + (x - u)^2} \\ &\quad + \int_{\mathbb{R} \setminus (-2\varepsilon^{-1}, 2\varepsilon^{-1})} \frac{1}{\pi} \frac{y \, du}{y^2 + (x - u)^2} + O\left(\frac{1}{T\varepsilon}\right) \\ &= \frac{\rho^\bullet - 1}{\pi} \left[\tan^{-1}\left(\frac{2\varepsilon^{-1} - x}{y}\right) + \tan^{-1}\left(\frac{2\varepsilon^{-1} + x}{y}\right) \right] \\ &\quad + 1 + O\left(\frac{1}{T\varepsilon}\right) \leq \frac{1 + \rho^\bullet}{2} \end{aligned}$$

for large enough N . This completes the proof. \square

LEMMA 5.4. *Under Assumptions 1 and 2, for each $0 < \delta < 1$ there exists $\varepsilon_0 > 0$ (which may depend on constants in our assumptions but not on the choice of $\mathfrak{A}(N)$), such that for any $0 < \varepsilon < \varepsilon_0$ there is $N_0 \in \mathbb{Z}_{\geq 1}$, and for all $N > N_0$ there exists a curve $\gamma = \gamma(N)$ in the upper half-plane with the following properties:*

- For all $z \in \gamma$ we have $\text{Im } S'(z) = 0$, $\text{Im}(z) \geq \varepsilon$, and $|z| < \varepsilon^{-1}$;
- The curve γ starts in the set $\{x + i\varepsilon : -1 - \varepsilon^\delta < x < -1 + \varepsilon^\delta\}$, and ends in the set $\{x + i\varepsilon : -\varepsilon^\delta < x < \varepsilon^\delta\}$.

PROOF. Let $\mathcal{D} = \{x + iy \in \mathbb{C} : y > \varepsilon, \sqrt{x^2 + y^2} < \varepsilon^{-1}\}$, and denote

$$\mathcal{U} = \mathcal{D} \cap \{x + iy \in \mathbb{C} : \text{Im } S'(x + iy) < 0\}.$$

By Lemma 5.2, the part of the boundary of \mathcal{U} which lies inside the interior of \mathcal{D} is a union of several curves whose start and end points belong to

$$\{x + i\varepsilon \in \mathbb{C} : -1 - \varepsilon^\delta < x < -1 + \varepsilon^\delta \text{ or } -\varepsilon^\delta < x < \varepsilon^\delta\};$$

cf. Figure 9.¹⁴

By continuity and Lemma 5.2, on any path from the segment $\{x + i\varepsilon : -1 + \varepsilon^\delta < x < -\varepsilon^\delta\}$ (where $\text{Im } S'(z) < 0$) to the curved boundary of \mathcal{D} (where $\text{Im } S'(z) > 0$) there exists a point where $\text{Im } S'(z) = 0$. Thus, as γ we can take any of the curves forming the boundary of \mathcal{U} inside \mathcal{D} which starts to the left of $-1 + \varepsilon^\delta$, ends to the right of $-\varepsilon^\delta$, and does not intersect the set $\{x + i\varepsilon\}$ except at its endpoints. This implies the claim. \square

¹⁴One can show that these curves do not intersect, that is, that $S''(z)$ cannot vanish where $\text{Im } S'(z) = 0$, but we do not need this fact.

LEMMA 5.5. *Under Assumptions 1 and 2, there exist $0 < \delta < 1$ and $\varepsilon_0 > 0$, (which may depend on constants in our assumptions but not on the choice of $\mathfrak{A}(N)$), such that for each $0 < \varepsilon < \varepsilon_0$ there exists $N_0 \in \mathbb{Z}_{\geq 1}$, and for all $N > N_0$ we have:*

- $\operatorname{Re} S'(x + i\varepsilon) < -1$ for all $-1 - \varepsilon^\delta < x < -1 + \varepsilon^\delta$;
- $\operatorname{Re} S'(x + i\varepsilon) > 1$ for all $-\varepsilon^\delta < x < \varepsilon^\delta$.

PROOF. We will prove only the second claim, as the first one is analogous. We will specify the exact value of ε_0 at the end of the proof, and for now let us just fix arbitrary $\varepsilon < \varepsilon_0 < 1$ and $\delta \in (0, 1)$. In addition, take a large positive real R . If R and N are large enough, then we can restrict the summation in the infinite principal value sum in (5.3) to $j \in \mathfrak{L}_T \cap [-RT, RT]$, so that

$$\left| \text{p.v.} \sum_{j \in \mathfrak{L}_T} \frac{1}{T} \frac{x - j/T}{y^2 + (x - j/T)^2} - \sum_{j \in \mathfrak{L}_T \cap [-RT, RT]} \frac{1}{T} \frac{x - j/T}{y^2 + (x - j/T)^2} \right| < 1,$$

$y = \varepsilon.$

In turn, the sum over $j \in \mathfrak{L}_T \cap [-RT, RT]$ is the Riemann sum for the corresponding integral, so for large N we have

$$\left| \text{p.v.} \sum_{j \in \mathfrak{L}_T} \frac{1}{T} \frac{x - j/T}{y^2 + (x - j/T)^2} - \int_{u \in [-R, -1] \cup [0, R]} \frac{(x - u) du}{y^2 + (x - u)^2} \right| < 2,$$

(5.7) $y = \varepsilon.$

Let us now bound the sum over the configuration $\mathfrak{A}(N)$ in (5.3). For that, we split this sum into three parts:

$$\begin{aligned} & \sum_{i \in \mathfrak{A}(N)} \frac{1}{T} \frac{x - i/T}{y^2 + (x - i/T)^2} \\ &= \sum_{i \in \mathfrak{A}(N) \cap [-RT, RT]} \frac{1}{T} \frac{x - i/T}{y^2 + (x - i/T)^2} \\ & \quad + \sum_{i \in \mathfrak{A}(N) \setminus [-RT, RT]} \frac{1}{T} \left(\frac{x - i/T}{y^2 + (x - i/T)^2} + \frac{T}{i} \right) \\ & \quad - \sum_{i \in \mathfrak{A}(N) \setminus [-RT, RT]} \frac{1}{i}. \end{aligned}$$

(5.8)

The third sum in (5.8) is bounded due to Assumption 2. For the second sum, observe that

$$\frac{1}{T} \left(\frac{x - i/T}{y^2 + (x - i/T)^2} + \frac{T}{i} \right) = \frac{1}{T} \frac{y^2 + x^2 - (i/T)x}{(i/T)(y^2 + (x - i/T)^2)},$$

that is, the second sum over i converges absolutely. Moreover, we can estimate as $N \rightarrow \infty$:

$$\begin{aligned} & \sum_{i \in \mathfrak{A}(N) \setminus [-RT, RT]} \left| \frac{1}{T} \left(\frac{x - i/T}{y^2 + (x - i/T)^2} + \frac{T}{i} \right) \right| \\ & \leq \sum_{i \in \mathbb{Z} \setminus [-RT, RT]} \left| \frac{1}{T} \frac{y^2 + x^2 - (i/T)x}{(i/T)(y^2 + (x - i/T)^2)} \right|, \end{aligned}$$

and the right-hand side is the Riemann sum for the integral

$$\int_{\mathbb{R} \setminus [-R, R]} \left| \frac{y^2 + x^2 - ux}{u(y^2 + (x - u)^2)} \right| du,$$

which is uniformly bounded for (x, y) in our segment (where x is around 0). Thus, the second sum in (5.8) is uniformly bounded by a constant independent of ε .

Finally, for the first sum in (5.8) we use Assumption 1 and approximate sums by integrals similarly to the proof of Lemma 5.2. To get a lower bound, first throw away all nonnegative summands in this sum, and write for the remaining ones

$$(5.9) \quad \sum_{\substack{i \in \mathfrak{A}(N) \cap [-RT, RT] \\ i/T > x}} \frac{1}{T} \frac{x - i/T}{y^2 + (x - i/T)^2} > \frac{1 + \rho^\bullet}{2} \int_x^R \frac{(x - u) du}{y^2 + (x - u)^2},$$

where N is sufficiently large.

Combining all the estimates, we obtain the following bound. For each $\varepsilon > 0$, there exists N_0 such that for all $N > N_0$ we have

$$(5.10) \quad \begin{aligned} \operatorname{Re} S'(x + iy) & > \frac{1 + \rho^\bullet}{2} \int_x^R \frac{(x - u) du}{y^2 + (x - u)^2} \\ & - \int_{u \in [-R, -1] \cup [0, R]} \frac{(x - u) du}{y^2 + (x - u)^2} \\ & - \log(\beta^{-1} - 1) + \text{“error”}, \end{aligned}$$

where “error” is uniform in (x, y) in our segment and is independent of ε . Observe that

$$\int_x^R \frac{(x - u) du}{y^2 + (x - u)^2} = \int_0^{R-x} \frac{-v dv}{v^2 + y^2} > \int_0^{R+1} \frac{-v dv}{v^2 + y^2}$$

and

$$- \int_0^R \frac{(x - u) du}{y^2 + (x - u)^2} = \int_{-x}^{R-x} \frac{v dv}{v^2 + y^2} > \int_{-x}^{R-1} \frac{v dv}{v^2 + y^2}.$$

At the same time,

$$- \int_{-R}^{-1} \frac{(x - u) du}{y^2 + (x - u)^2} = -\frac{1}{2} \log \left(\frac{(R+x)^2 + y^2}{(1+x)^2 + y^2} \right)$$

can be bounded by an absolute constant since both x and y are close to zero. Thus, we can write

$$\begin{aligned} & \frac{1 + \rho^\bullet}{2} \int_x^R \frac{(x - u) du}{y^2 + (x - u)^2} - \int_{u \in [-R, -1] \cup [0, R]} \frac{(x - u) du}{y^2 + (x - u)^2} \\ & > \frac{1 + \rho^\bullet}{2} \int_0^{R+1} \frac{-v dv}{v^2 + y^2} + \int_{-x}^{R-1} \frac{v dv}{v^2 + y^2} + C \\ & = -\frac{1 + \rho^\bullet}{4} \log\left(\frac{(R + 1)^2}{y^2} + 1\right) + \frac{1}{2} \log\left(\frac{(R - 1)^2 + y^2}{x^2 + y^2}\right) + C. \end{aligned}$$

Here and below, in this proof C stands for some real constant which is uniform in x, y and does not depend on ε but may depend on R (but we fixed large R once and for all in the beginning of the proof). The value of C can change from line to line. Since $y = \varepsilon$ is small, we have

$$-\frac{1 + \rho^\bullet}{4} \log\left(\frac{(R + 1)^2}{y^2} + 1\right) = \frac{1 + \rho^\bullet}{2} \log y + C + O(\varepsilon^2).$$

We also have

$$\frac{1}{2} \log\left(\frac{(R - 1)^2 + y^2}{x^2 + y^2}\right) = \frac{1}{2} \log(y^2 + (R - 1)^2) - \log \sqrt{x^2 + y^2} > C - \delta \log y$$

because $x^2 + y^2 < \varepsilon^{2\delta} + \varepsilon^2$, which behaves as $\varepsilon^{2\delta}(1 + o(1)) = y^{2\delta}(1 + o(1))$. When δ is close enough to 1,

$$\frac{1 + \rho^\bullet}{2} \log y - \delta \log y$$

tends to $+\infty$ as $y = \varepsilon \rightarrow 0$, and we are done. \square

PROOF OF PROPOSITION 2.6. Fix $\varepsilon > 0$ and $N_0 \in \mathbb{Z}_{\geq 1}$ depending on ε such that Lemmas 5.2, 5.4 and 5.5 hold (recall that $\varepsilon < \varepsilon_0$, where ε_0 may depend on constants in our assumptions but not on the choice of $\mathfrak{A}(N)$). Consider the curve γ from Lemma 5.4. This is a continuous curve on which $\text{Im } S'(z) = 0$. Furthermore, Lemma 5.5 guarantees that $\text{Re } S'(z)$ has distinct signs at the endpoints of γ . Since $\text{Re } S'(z)$ is a continuous function on γ , we conclude that there exists $z_c \in \gamma$ for which $\text{Re } S'(z_c) = 0$, and so $S'(z_c) = 0$ (as S' depends on N , so does z_c , and this statement is valid for all $N > N_0$). Lemma 5.1 then implies that there are no other critical points in the upper half-plane and, therefore, z_c is the desired unique one. As the compact set \mathcal{Z} capturing z_c take $\{x + iy \in \mathbb{C} : y \geq \varepsilon, \sqrt{x^2 + y^2} \leq \varepsilon^{-1}\}$. \square

6. Asymptotics of the kernel: Proofs of Theorems 2.7, 2.10, 2.11 and 2.12.

In this section, based on the existence of nonreal critical points afforded by Proposition 2.6, we establish the approximation of the correlation kernel (2.2) of the noncolliding Bernoulli random walk by the extended sine kernel, and also the corresponding bulk limit theorems. That is, here we prove the remaining statements from Section 2.3.

6.1. *Behavior of $\text{Im } S(z)$ and $\text{Re } S(z)$.* We aim to describe the steepest descent contours for $\text{Re } S(z)$. For that, we need to analyze the behavior of $\text{Re } S(z)$ and $\text{Im } S(z)$ in various parts of the upper half-plane \mathbb{H} . Recall that we defined $S(z)$ in Section 4.2 so that it is holomorphic in \mathbb{H} and extends to the real axis except the singularities at $\iota_T \Delta \mathfrak{a}$ (all other logarithmic singularities belonging to $\frac{1}{T} \mathbb{Z}$ are removable).

We start by considering the behavior of $\text{Im } S(z)$ close to the real line, and define

$$(6.1) \quad \begin{aligned} b_0^r &= \min(\mathbb{Z}_{\geq 0} \setminus \mathfrak{A}), & b_0^\ell &= \max(\mathfrak{A} \cap \{-T + 1, \dots, -1\}), \\ b_{-1}^\ell &= \max(\mathbb{Z}_{\leq -T} \setminus \mathfrak{A}), & b_{-1}^r &= \min(\mathfrak{A} \cap \{-T + 1, \dots, -1\}). \end{aligned}$$

Clearly,

$$b_{-1}^\ell \leq -T < -T + 1 \leq b_{-1}^r \leq b_0^\ell \leq -1 < 0 \leq b_0^r.$$

LEMMA 6.1. *For $x \in \mathbb{R}$, $x \notin \iota_T \Delta \mathfrak{a}$, the function $\text{Im } S(x)$ is piecewise constant, making jumps at points of $\iota_T \Delta \mathfrak{a}$. It weakly increases for $x \in (-\infty, b_{-1}^\ell/T) \cup (b_0^\ell/T, +\infty)$, and weakly decreases for $x \in (b_{-1}^\ell/T, b_0^\ell/T)$. See Figure 10 for an example.*

PROOF. This is straightforward from the definition of $S(z)$ in Section 4.2 and the observation that $\text{Im}(\log_{\mathbb{H}}(x)) = \pi \mathbf{1}_{x < 0}$, where $x \in \mathbb{R} \setminus \{0\}$. \square

LEMMA 6.2. *Fix $0 < \beta < 1$ and the constants in Assumption 2. There exists $C > 0$ depending only on these choices, and such that for each $T, N = 1, 2, \dots$, $x \in \mathbb{R}$, $y > 0$ we have (note that $y^2 - y \log y$ in the right-hand side is positive for $y > 0$)*

$$(6.2) \quad |\text{Im } S(x + iy) - \text{Im } S(x + i0)| \leq C \cdot \left(y \log(|x| + 1) - y \log y + y^2 + \frac{1}{T} \right).$$

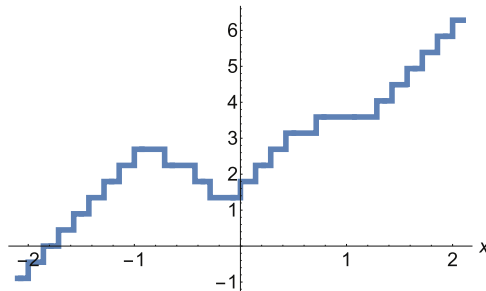


FIG. 10. *Staircase-type plot of $\text{Im } S(x)$ for $x \in \mathbb{R}$, with parameters as in Figure 8. The singularities leading to the down steps are $\{-\frac{6}{7}, \dots, -\frac{1}{7}\} \cap \{-\frac{5}{7}, -\frac{3}{7}, -\frac{2}{7}, \frac{4}{7}, \frac{6}{7}, 1, \frac{8}{7}\}$.*

REMARK 6.3. The value of $\text{Im } \mathbf{S}(x + \mathbf{i}0)$ when $x \in \iota_T \Delta \alpha$ (so that this piecewise linear function makes a jump) can be chosen arbitrarily (as long as Lemma 6.1 holds)—this introduces an error of at most $1/T$ which is included in the right-hand side of (6.2).

PROOF OF LEMMA 6.2. Recall the definition (4.8) of the function \mathbf{S} . Our aim is to obtain a uniform bound on the increment $\text{Im } \mathbf{S}(x + \mathbf{i}y) - \text{Im } \mathbf{S}(x + \mathbf{i}0)$.

We start from the second line in (4.8). For $-z \log(\beta^{-1} - 1)$ the increment is linear in y and fits into the right-hand side of (6.2). For $\frac{2\pi \mathbf{i}}{T} \lfloor \frac{1}{2} \text{Re}(Tz) + \frac{1}{2} \rfloor$, the increment vanishes. For $-\frac{1}{T} \log_{\mathbb{H}}(\sin(\pi Tz))$, the imaginary part of $\log_{\mathbb{H}}(\cdot)$ is bounded, since it is an argument of a complex number. Thus, the increment is bounded by C/T .

We proceed to the first line of (4.8). Let us analyze the first term, $\frac{1}{T} \times \sum_{r=1}^N \log_{\mathbb{H}}(z - \frac{a_r}{T})$. Choose a $\delta > 0$, which will be later set to $\delta = 4y$, and split the sum into

$$\begin{aligned}
 & \frac{1}{T} \sum_{r=1}^N \log_{\mathbb{H}}\left(x + \mathbf{i}y - \frac{a_r}{T}\right) \\
 &= \frac{1}{T} \sum_{1 \leq r \leq N: |x - a_r/T| < \delta} \log_{\mathbb{H}}\left(x + \mathbf{i}y - \frac{a_r}{T}\right) \\
 (6.3) \quad &+ \frac{1}{T} \sum_{1 \leq r \leq N: |x - a_r/T| \geq \delta} \log_{\mathbb{H}}\left(x + \mathbf{i}y - \frac{a_r}{T}\right).
 \end{aligned}$$

The first term in (6.3) has at most $2\delta T$ summands, the increment of each one between the points $x + \mathbf{i}y$ and $x + \mathbf{i}0$ is bounded by a constant. Therefore, the increment of the first term is bounded by $C \cdot \delta = 4Cy$.

For the second term in (6.3), we compute the increment directly as

$$(6.4) \quad \text{Im} \frac{1}{T} \sum_{1 \leq r \leq N: |x - a_r/T| \geq \delta} \log_{\mathbb{H}}\left(1 + \frac{\mathbf{i}y}{x - a_r/T}\right).$$

By our choice of δ , $|\frac{\mathbf{i}y}{x - a_r/T}| \leq 1/4$. Therefore, we can Taylor expand each $\log_{\mathbb{H}}(\cdot)$ and bound the absolute value of (6.4) as

$$(6.5) \quad \frac{y}{T} \left| \sum_{1 \leq r \leq N: |x - a_r/T| \geq \delta} \frac{1}{x - a_r/T} \right| + \frac{C \cdot y^2}{T} \sum_{1 \leq r \leq N: |x - a_r/T| \geq \delta} \frac{1}{(x - a_r/T)^2}.$$

The second term in (6.5) is smaller than

$$Cy^2 \int_{|u| > \delta} \frac{1}{u^2} = Cy \cdot \frac{2y}{\delta} < Cy,$$

and, therefore, fits into the right-hand side of (6.2). For the first term in (6.5), we write $\frac{1}{u-v} = \frac{u}{(u-v)v} - \frac{1}{v}$ and bound from above as follows:

$$\begin{aligned}
 & \left| \frac{y}{T} \sum_{\substack{1 \leq r \leq N: \\ |x-a_r/T| \geq \delta, |a_r| > T}} \left(\frac{x}{(x-a_r/T)(a_r/T)} - \frac{T}{a_r} \right) \right| \\
 & + \left| \frac{y}{T} \sum_{\substack{1 \leq r \leq N: \\ |x-a_r/T| \geq \delta, |a_r| \leq T}} \frac{1}{x-a_r/T} \right| \\
 & \leq y \left| \sum_{1 \leq r \leq N: |a_r| > T} \frac{1}{a_r} \right| + y \left| \sum_{1 \leq r \leq N: |x-a_r/T| < \delta, |a_r| > T} \frac{1}{a_r} \right| \\
 & + \left| \frac{y}{T} \sum_{\substack{1 \leq r \leq N: \\ |x-a_r/T| \geq \delta, |a_r| > T}} \frac{x}{(x-a_r/T)(a_r/T)} \right| \\
 (6.6) \quad & + y \left| \sum_{\substack{1 \leq r \leq N: \\ |x-a_r/T| \geq \delta, |a_r| \leq T}} \frac{1}{Tx-a_r} \right|.
 \end{aligned}$$

In the right-hand side of (6.6) the first term is bounded by $C \cdot y$ due to Assumption 2 and Remark 2.5. The second term has at most $2\delta T$ summands, each of which is at most $1/T$. Therefore, the second term in (6.6) is bounded from above by $2y\delta = 4y^2$. For the third term, we can replace the sum over a_r by the sum over all integers j satisfying the inequalities $|x - j/T| \geq \delta, |j| > T$, and then upper bound the sum by the integral to get

$$y \int_{|u-x| \geq \delta, |u| > 1} \frac{x du}{(x-u)u} = y \lim_{M \rightarrow +\infty} \int_{|u-x| \geq \delta, 1 < |u| < M} \left(\frac{1}{u} + \frac{1}{x-u} \right) du.$$

At this point, we need to consider several cases depending on the order of the points $x \pm \delta$ and ± 1 . In all of the cases, the integral evaluates into a combination of the expressions of the form $\log |x \pm \delta|$, $\log \delta$ and $\log |x \pm 1|$. We conclude that this term fits into the form of the right-hand side of (6.2).

For the fourth term on the right-hand side of (6.6), we again replace a_r by all integers and then use $\sum_{n=\theta_1}^{\theta_2 T} n^{-1} \approx \log(\theta_2/\theta_1)$. As a result we get a bound of the form $C(\log y + 1)$, which fits into the right-hand side of (6.2).

We have obtained a uniform bound for the increment of each term in (4.8) except for $\frac{1}{T} \sum_{i=1}^{T-1} \log_{\mathbb{H}}(x + iy + \frac{i}{T})$, and we proceed to bound this term. When x is bounded away from 0 and -1 , the argument is the same as we just had. However, when x is close to 0 or -1 , we need to proceed differently. Let us split the sum into two according to the sign of $x + \frac{i}{T}$. Each of them is analyzed in the same way,

so we will only deal with one. This reduces the problem to bounding

$$(6.7) \quad \frac{1}{T} \sum_{i=1}^{T'-1} \operatorname{Im} \left[\log_{\mathbb{H}} \left(x + \frac{i}{T} + \mathbf{i}y \right) - \log_{\mathbb{H}} \left(x + \frac{i}{T} + \mathbf{i}0 \right) \right],$$

$$x \geq -\frac{1}{T}, T' \leq T.$$

Each term in (6.7) has the form

$$\operatorname{Im} \left[\log_{\mathbb{H}} \left(1 + \frac{\mathbf{i}y}{x + \frac{i}{T}} \right) \right] = \arctan \left(\frac{y}{x + \frac{i}{T}} \right).$$

For $x = -1/T, i = 1$, the corresponding term in (6.7) vanishes. For all other cases, we note that \arctan is monotone and use $\arctan(u) \leq \min(\pi/2, u), u \geq 0$. We thus bound (6.7) by

$$\begin{aligned} \frac{\pi}{2} \cdot \frac{Ty + 1}{T} + \frac{1}{T} \sum_{i=\lceil Ty \rceil + 1}^{T'-1} \left(\frac{Ty}{i - 1} \right) &\leq \frac{\pi}{2} \cdot \frac{Ty + 1}{T} + y \int_T^T \frac{1}{u} du \\ &\leq \frac{\pi}{2} \cdot \frac{Ty + 1}{T} - y \log(\min(y, 1)). \end{aligned}$$

Since the last expression fits into the right-hand side of (6.2), we are done. \square

We now turn to the real part $\operatorname{Re} S(z)$.

LEMMA 6.4. *Under Assumptions 1 and 2 and with constants depending only on these assumptions, the following estimates hold:*

1. *For any $k \in \mathbb{Z}, N = 1, 2, \dots$, and $y > 0$ we have*

$$\left| \frac{\partial}{\partial y} \operatorname{Re} S \left(\frac{k + 1/2}{T} + \mathbf{i}y \right) \right| \leq \pi.$$

2. *There exists $Y > 0, N_0 > 0$, such that for any $|x| < \frac{1}{2} \cdot \frac{Q(N)}{T(N)}, y > Y$, and $N > N_0$ we have*

$$\frac{\partial}{\partial y} \operatorname{Re} S(x + \mathbf{i}y) = -\frac{\partial}{\partial x} \operatorname{Im} S(x + \mathbf{i}y) < -\frac{1}{y}.$$

3. *For each $X > 0$, there exist $C, N_0 > 0$ such that for all $0 < y < 1/2, |x| < X$ and all $N > N_0$ we have*

$$\left| \frac{\partial}{\partial x} \operatorname{Re} S(x + \mathbf{i}y) \right| = \left| \frac{\partial}{\partial y} \operatorname{Im} S(x + \mathbf{i}y) \right| < C \left(\log(y^{-1}) + \frac{1}{y^2 T^2} \right).$$

PROOF. Using (4.5), we see that, apart from the linear term $-(\operatorname{Re} z) \log(\beta^{-1} - 1)$, the function $\operatorname{Re} \mathbf{S}(z)$ is an infinite linear combination (with coefficients ± 1) of shifts of $\frac{1}{T} \log |z|$. We have for any $x, y \in \mathbb{R}$

$$\frac{\partial}{\partial y} \log |x + iy| = \frac{y}{x^2 + y^2}.$$

In particular, for $j \in \mathbb{Z}$,

$$\left| \frac{\partial}{\partial y} \frac{1}{T} \log \left| \frac{k + 1/2}{T} + iy + \frac{j}{T} \right| \right| = \frac{4T|y|}{(2k + 2j + 1)^2 + 4T^2y^2}.$$

Thus, the absolute value of the derivative of $\operatorname{Re} \mathbf{S}$ in the first claim can be bounded in the absolute value by

$$\sum_{j \in \mathbb{Z}} \frac{4T|y|}{(2j + 1)^2 + 4T^2y^2} = \pi \tanh(\pi T|y|) \leq \pi,$$

this is summed with the help of a partial fraction expansion and (2.10), and \tanh is bounded by one. This establishes the first claim.

For the second claim, we need to be more careful with signs. Recalling that $\mathfrak{A}(N)$ is the initial condition and using notation (2.7), we write

$$\begin{aligned} \frac{\partial}{\partial y} \operatorname{Re} \mathbf{S}(x + iy) &= -\frac{1}{T} \sum_{a \in \mathfrak{L}_T \setminus \mathfrak{A}(N)} \frac{y}{(x - a/T)^2 + y^2} \\ (6.8) \quad &+ \frac{1}{T} \sum_{a \in \mathfrak{A}(N) \cap \{1-T, \dots, -1\}} \frac{y}{(x - a/T)^2 + y^2}. \end{aligned}$$

Our aim is to show that in the last sum the first term dominates. Using Assumption 1 and replacing sums by integrals (with multiplicative error at most 2), we upper bound (6.8) by

$$\begin{aligned} &-\frac{1 - \rho^\bullet}{2} \left(\int_{-\mathbf{Q}(N)/T}^{-1} + \int_0^{\mathbf{Q}(N)/T} \right) \frac{y \, dv}{(v - x)^2 + y^2} + 2 \int_{-1}^0 \frac{y \, dv}{(v - x)^2 + y^2} \\ &= -\frac{1 - \rho^\bullet}{2} \left(\int_{-x/y - \mathbf{Q}(N)/(Ty)}^{(x-1)/y} + \int_{x/y}^{-x/y + \mathbf{Q}(N)/(Ty)} \right) \frac{du}{u^2 + 1} \\ &\quad + 2 \int_{-(x+1)/y}^{-x/y} \frac{du}{u^2 + 1} \\ &\leq -\frac{1 - \rho^\bullet}{2} \int_{-\mathbf{Q}(N)/(2Ty)}^{\mathbf{Q}(N)/(2Ty)} \frac{du}{u^2 + 1} + 3 \int_{-(x+1)/y}^{-x/y} \frac{du}{u^2 + 1} \\ (6.9) \quad &\leq -(1 - \rho^\bullet) \arctan\left(\frac{\mathbf{Q}(N)}{2Ty}\right) + \frac{3}{y}. \end{aligned}$$

Considering separately the cases of small and large y , using $Q(N)/T(N) \rightarrow \infty$ as $N \rightarrow \infty$, we see that the last expression is smaller than $-1/y$ for large N , and the second claim is proven. (Notice that $\frac{\partial}{\partial y} \operatorname{Re} S(x + iy) = -\frac{\partial}{\partial x} \operatorname{Im} S(x + iy)$ is the Cauchy–Riemann equation.)

Let us turn to the third claim. Assume that x is fixed. We have

$$\frac{\partial}{\partial x} \frac{1}{T} \log \left| x + iy + \frac{j}{T} \right| = \frac{Tx + j}{y^2 T^2 + (Tx + j)^2},$$

so

$$\begin{aligned} \frac{\partial}{\partial x} \operatorname{Re} S(x + iy) &= \sum_{r=1}^N \frac{Tx - a_r}{y^2 T^2 + (Tx - a_r)^2} + \sum_{i=1}^{T-1} \frac{Tx + i}{y^2 T^2 + (Tx + i)^2} \\ (6.10) \quad &\quad - \log(\beta^{-1} - 1) - \frac{\partial}{\partial x} \frac{1}{T} \log |\sin(\pi T(x + iy))|. \end{aligned}$$

Fix sufficiently large $R > 0$, which might depend on x , but not on y . For the first sum in (6.10) with $|a_r| < RT$, and also for the second sum in (6.10) we upper bound the absolute values of the sums by twice of

$$\begin{aligned} \sum_{j=0}^{2RT} \frac{j+1}{j^2 + y^2 T^2} &\leq \sum_{j=0}^{2RT} \frac{j}{j^2 + y^2 T^2} + \frac{1}{y^2 T^2} + \sum_{j=1}^{\infty} \frac{1}{j^2} \\ &\leq \int_0^{2R} \frac{v}{v^2 + y^2} dv + \frac{1}{y^2 T^2} + C \\ &= \frac{1}{2} \log \left(1 + \frac{4R^2}{y^2} \right) + \frac{1}{y^2 T^2} + C. \end{aligned}$$

Therefore, the contribution of these terms admits the desired bound. Next, for $|a_r| > RT$ in the first sum in (6.10) we have

$$\frac{1}{Tx - a_r} - \frac{Tx - a_r}{y^2 T^2 + (Tx - a_r)^2} = \frac{T^2 y^2}{(Tx - a_r)(T^2 y^2 + (Tx - a_r)^2)},$$

and summing this over $|a_r| > RT$ has order y^2/R^2 , which is bounded. Thus, the contribution from $|a_r| > RT$ in the first sum in (6.10) is the same as if the summands were just $1/(Tx - a_r)$. Observe that

$$\frac{1}{Tx - a_r} + \frac{1}{a_r} = \frac{Tx}{a_r(Tx - a_r)}$$

and the sum of these quantities over $|a_r| > RT$ with large R is bounded (recall that $|x|$ is bounded and R is chosen to be much larger than it). Thus, the sum of the terms with $|a_r| > RT$ in the first sum in (6.10) has the same order as the sum of $1/a_r$ over $|a_r| > RT$, which is bounded by Assumption 2.

Finally, for the last summand in (6.10) we have

$$\left| \frac{\partial}{\partial x} \frac{1}{T} \log \left| \sin(\pi T(x + iy)) \right| \right| = \pi \left| \frac{\cos(\pi T(x + iy))}{\sin(\pi T(x + iy))} \right| \leq C \left(1 + \frac{1}{T^2 y^2} \right),$$

where we used the bound $|\sin(\alpha + i\beta)| \geq C \cdot \min(|\beta|^2, 1)$ for an absolute constant $C > 0$. \square

6.2. *Steepest descent/ascent contours in a large rectangle.* Our next aim is to present a new set of contours for the double contour integral expression of Theorem 2.15. In this section, we explain their geometry in a (sufficiently large) compact subset of the upper half-plane. In discussion of integration contours in the rest of this section, it suffices to argue in the upper half-plane: the contour configuration in the lower half-plane (with a suitable choice of branches of logarithms, cf. Section 4.2) is obtained by reflection with respect to the real line.

Recall the critical point $z_c = z_c(N)$ afforded by Proposition 2.6. We need the following statement which will be proven in Section 6.4 below.

LEMMA 6.5. *Under Assumptions 1 and 2 there exists $C > 0$, such that $C^{-1} < |S''(z)| < C$ for all z in the upper half-plane satisfying $|z - z_c| < C^{-1}$, and all $N = 1, 2, \dots$*

Fix three constants: small $\varepsilon < 0$ and large $R_x, R_y > 0$, which do not depend on N or $\mathfrak{A}(N)$, but might depend on the constants in Assumptions 1, 2. There are four contours $\{z: \text{Im } S(z) = \text{Im } S(z_c)\}$ emanating from the critical point. Let us trace these contours until they leave a rectangle $\mathcal{R} := \{x + iy \in \mathbb{C}: |x| < R_x, \varepsilon < y < R_y\}$ (for some $R_y > \varepsilon > 0, R_x > 0$). Let z_1, z_2, z_3, z_4 be the *escape* points where the contours leave the rectangle.

PROPOSITION 6.6. *There exist $R_y > \varepsilon > 0, R_x > 0$, such that for all large enough N :*

1. *Three escape points, z_1, z_2, z_3 (ordered as $\text{Re } z_1 < \text{Re } z_2 < \text{Re } z_3$) are on the lower side $\text{Im } z = \varepsilon$, and z_4 is on the upper side $\text{Im } z = R_y$.*
2. *The real part $\text{Re } S(z)$ grows along the contours escaping through z_1 and z_3 , and decays along the contours escaping through z_2, z_4 .*
3. *The escape points on the lower sides of the rectangle satisfy*

$$(6.11) \quad \begin{aligned} \text{Re } z_1 &\in (-\infty, -1 + \varepsilon \log^2 \varepsilon), \\ \text{Re } z_2 &\in (-1 - \varepsilon \log^2 \varepsilon, \varepsilon \log^2 \varepsilon), \\ \text{Re } z_3 &\in (-\varepsilon \log^2 \varepsilon, +\infty). \end{aligned}$$

4. $R_y > Y$, where Y is from the second claim of Lemma 6.4.
5. $\varepsilon < \frac{1}{10} \min_{i=1,2,3} |\text{Re } S(z_i) - \text{Re } S(z_c)|$.

PROOF. First, $\operatorname{Re} S(z)$ can not have local extrema on the contours $\{z : \operatorname{Im} S(z) = \operatorname{Im} S(z_c)\}$ inside the rectangle (except at point z_c), as any such extremum would be a new critical point for $S(z)$ contradicting Proposition 2.6. Therefore, $\operatorname{Re} S(z)$ is monotone along these contours. This also means that these contours cannot intersect anywhere in the rectangle except at z_c . Out of these four contours, along two the real part $\operatorname{Re} S(z)$ grows, and along other two it decays. Since the growth/decay types interlace as the contours leave z_c , we conclude that the growth/decay types also interlace along the boundary of the rectangle.

We now fix arbitrary $R_y > Y$, such that $\operatorname{Im}(z_c) < R_y/2$. The second claim of Lemma 6.4 implies that (for large N) $\operatorname{Im} S(z)$ is monotone along the top side $\operatorname{Im} z = R_y$ of the rectangle. Therefore, at most one of the points $z_i, 1 \leq i \leq 4$, can be there.

Next, Lemma 6.1 combined with Assumption 1 implies that there exists $R > 0$ such that $\operatorname{Im} S(x + i0) > R^{-1}x$ for $x > R$ and $\operatorname{Im} S(x + i0) < -R^{-1}x$ for $x < -R$. Thus, Lemma 6.2 implies that we can choose large enough $R_x > R$, such that

$$\operatorname{Im} S(R_x + iy) > 2|\operatorname{Im} S(z_c)|, \quad \operatorname{Im} S(-R_x + iy) < 2|\operatorname{Im} S(z_c)|$$

for all $0 \leq y \leq R_y$. We fix such R_x and notice that this choice implies that $z_i, 1 \leq i \leq 4$ do not belong to the vertical sides of the rectangle.

Thus, either three or four of the points z_i belong to the bottom horizontal side of the rectangle. It remains to specify $\varepsilon > 0$, so that there are exactly three and their positions satisfy (6.11).

By Lemma 6.5, we have a uniform control over the growth/decay of $\operatorname{Re} S(z)$ in a small (but fixed size) neighborhood of z_c . Thus, when the contours reach the boundary of the rectangles, the values of $\operatorname{Re} S(z)$ are separated by a constant. Combining this fact with interlacing of the growth/decay contours and the bound of the third statement of Lemma 6.4 we conclude that there exists $\delta > 0$ such that for each $0 < \varepsilon < 1/2$ and $T > \varepsilon^{-1}$ we have $|z_i - z_j| > \frac{\delta}{\log(1/\varepsilon)}$ for all $i \neq j$.

Next, let $\mathcal{U} \subset \mathbb{R}$ denote the set of points x such that $\operatorname{Im} S(x + 0i) = \operatorname{Im} S(z_c)$. According to Lemma 6.1, \mathcal{U} splits into three disjoint sets $\mathcal{U}_1, \mathcal{U}_2, \mathcal{U}_3$ (some of which might be empty): $\mathcal{U}_1 \subset (-\infty, -1], \mathcal{U}_2 \subset [-1, 0], \mathcal{U}_3 \subset [0, +\infty)$. Using Assumption 1, we see that the diameter of each set \mathcal{U}_i is at most $\frac{D(N)}{T(N)}$ which tends to 0 as $N \rightarrow \infty$.

We further would like to show that z_i is close to $\mathcal{U}_i, i = 1, 2, 3$. For that note that by Assumption 1, the function $x \mapsto \operatorname{Im} S(x + i0)$ has growth bounded away from 0 in the sense that for some $c > 0$ we have

$$\begin{aligned} & |\operatorname{Im}(S(x + i0)) - \operatorname{Im}(S(x' + i0))| \\ & \geq c \cdot |x - x'| \quad \text{if } |x - x'| \geq \frac{D(N)}{T(N)} \text{ and } \begin{cases} x, x' \in [-R_x, -1] & \text{or} \\ x, x' \in [-1, 0] & \text{or} \\ x, x' \in [0, R_x]. \end{cases} \end{aligned}$$

Thus, using Lemma 6.2 we conclude that for fixed $\varepsilon > 0$ and large N, T (compared to that ε) the real part of each point z_i (out of those lying in the bottom horizontal side of the rectangle) should be in $C\varepsilon \log(\varepsilon^{-1})$ -neighborhood of $\mathcal{U}_1 \cup \mathcal{U}_2 \cup \mathcal{U}_3$, where C does not depend on T or N . On the other hand, the diameter of each \mathcal{U}_i is small and $|z_i - z_j| > \frac{\delta}{\log(1/\varepsilon)}$. Since $\frac{\delta}{\log(1/\varepsilon)} \gg C\varepsilon \log(\varepsilon^{-1})$, we conclude that the only possibility is that there are precisely three points z_i on the bottom horizontal side of the rectangle (which means that z_4 is on the upper horizontal side) and each z_i is in $C\varepsilon \log(\varepsilon^{-1})$ neighborhood of \mathcal{U}_i , respectively, for $i = 1, 2, 3$. For small enough ε , we would have $C\varepsilon \log(\varepsilon^{-1}) < \varepsilon \log^2(\varepsilon)$, which completes the proof. \square

6.3. *Completing the proof of Theorem 2.7.* Here, we describe how the new contours in a large rectangle constructed in Section 6.2 should be continued outside the rectangle. (Recall that by symmetry, it suffices to argue in the upper half-plane only.) We then rewrite the correlation kernel $K_{\vec{a};\beta}^{\text{Bernoulli}}$ in terms of these new contours, and complete the proof of Theorem 2.7 on approximation of $K_{\vec{a};\beta}^{\text{Bernoulli}}$ by the extended discrete sine kernel.

Fix $R_y > \varepsilon > 0, R_x > 0$ as in Proposition 6.6. Define the new z contour $\mathcal{C}_z^\varepsilon = \mathcal{C}_z^\varepsilon(N)$ as follows.¹⁵ Inside the rectangle \mathcal{R} it coincides with the union of the steepest descent (for $\text{Re } S$) contours $\{z: \text{Im } S(z) = \text{Im } S(z_c)\}$ escaping through the points z_2 and z_4 . After the point z_4 , we continue the contour vertically so that it escapes to infinity. Inside the ε -neighborhood of the real line, we have to modify the steepest descent contour so that it crosses \mathbb{R} strictly between -1 and 0 . To achieve that, we add to this contour a horizontal segment of the line $\text{Im } z = i\varepsilon$, and then a vertical segment connecting it to the real line such that $\mathcal{C}_z^\varepsilon$ crosses \mathbb{R} at a point of the form $(k + 1/2)/T$ for some $k \in \{-T, \dots, -1\}$ which is close to z_2 within $C \cdot (\varepsilon \log^2 \varepsilon + T^{-1})$. The contour $\mathcal{C}_z^\varepsilon$ is oriented upward.

Next, by $\mathcal{C}_w^\varepsilon = \mathcal{C}_w^\varepsilon(N)$ denote the closed positively oriented contour which inside the rectangle \mathcal{R} coincides with the union of the steepest ascent (for $\text{Re } S$) contours $\{z: \text{Im } S(z) = \text{Im } S(z_c)\}$ escaping through the points z_1 and z_3 . Outside \mathcal{R} we modify the steepest ascent contour so that it encircles $\{-1 + T^{-1}(\Delta x - \Delta t), \dots, T^{-1}(\Delta x - 1), T^{-1} \Delta x\}$, and crosses \mathbb{R} at two points of the form $(k + 1/2)/T, k \in \mathbb{Z}$, close to z_1 and z_3 within $C \cdot (\varepsilon \log^2 \varepsilon + T^{-1})$. This is achieved by adding horizontal and vertical segments similar to $\mathcal{C}_z^\varepsilon$. See Figure 11 for an illustration of the new contours $\mathcal{C}_z^\varepsilon$ and $\mathcal{C}_w^\varepsilon$.

¹⁵This contour, as well as $\mathcal{C}_w^\varepsilon$ defined below, also depends on the constants R_x, R_y and other data in Assumptions 1 and 2. We suppress all this dependence in the notation.

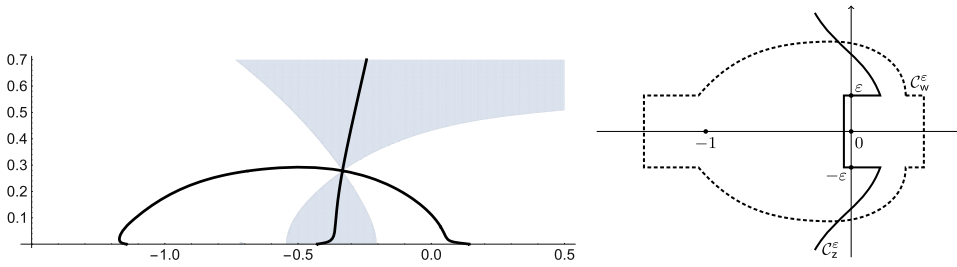


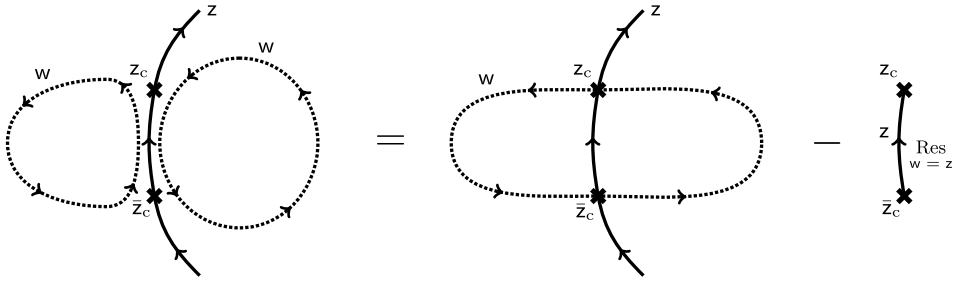
FIG. 11. *Left: Steepest descent/ascent contours $\{z: \text{Im } S(z) = \text{Im } S(z_c)\}$ for the function $\text{Re } S(z)$ in the upper half-plane. Regions where $\text{Re } S(z) < \text{Re } S(z_c)$ are shaded. The parameters \mathfrak{A} and T are as in Figure 8, and $\beta = 0.4$. Right: Modification of the contours in an ε -neighborhood of the real line.*

PROPOSITION 6.7. *With the above definitions and conventions, for any $N > N_0$ the kernel (4.2) can be written as*

$$\begin{aligned}
 &K_{\hat{a}; \beta}^{\text{Bernoulli}}(T + \Delta t, \Delta x; T, 0) \\
 &= K_{z_c/(z_c+1)}(T + \Delta t, \Delta x; T, 0) + \frac{1}{(2\pi i)^2} \int_{C_z^\varepsilon} dz \oint_{C_w^\varepsilon} dw \frac{1}{w - z} \\
 &\quad \times \frac{(T + \Delta t)! \cdot T}{(T - 1)!} \frac{(Tz + 1)_{T-1}}{(Tw - \Delta x)_{T+\Delta t+1}} \frac{\sin(\pi Tw)}{\sin(\pi Tz)} \left(\frac{1 - \beta}{\beta}\right)^{T(w-z)} \\
 (6.12) \quad &\times \prod_{r=1}^N \frac{Tz - a_r}{Tw - a_r} + O(T^{-1}).
 \end{aligned}$$

PROOF. All poles of the integrand in (4.2) are on the real line. This integrand has no poles at $z \in \{b_{-1}^\ell/T, \dots, b_0^r/T\}$ (recall the notation (6.1)), and thus we can drag the point of intersection of the z contour with the real line to the desired location dictated by the contour C_z^ε . Hence we can deform the whole z contour to coincide with C_z^ε without crossing any poles. Next, let us unite the two circles comprising the w contour in (4.2) into the contour C_w^ε intersecting with C_z^ε at the critical points z_c and \bar{z}_c . This leads to an additional integral of the residue at $w = z$ over the arc of C_z^ε from \bar{z}_c to z_c crossing the real line between -1 and 0 ; see Figure 12. This deformation of the w contour does not cross any other w poles of the integrand.¹⁶

¹⁶The desire that these deformations do not cross any real poles is the reason why the contours C_z^ε and C_w^ε should differ from the steepest descent/ascent ones close to the real line.



Double contour integral in (4.2)

FIG. 12. Unifying two circles into a single w contour.

The expression coming from the residue of the integrand at $w = z$ behaves as

$$\begin{aligned}
 & -\frac{1}{2\pi i} \frac{(T + \Delta t)! \cdot T}{(T - 1)!} \frac{(Tw + 1)_{T-1}}{(Tw - \Delta x)_{T+\Delta t+1}} \\
 & = -\frac{1}{2\pi i} (1 + O(T^{-1})) T^{\Delta t+2} \\
 & \quad \times \frac{\Gamma(Tw + T)\Gamma(Tw - \Delta x)}{\Gamma(Tw + 1)\Gamma(Tw - \Delta x + T + \Delta t + 1)} \\
 (6.13) \quad & = -\frac{1}{2\pi i} (1 + O(T^{-1})) w^{-\Delta x-1} (1 + w)^{-\Delta t + \Delta x - 1},
 \end{aligned}$$

where we used (3.13), and the asymptotic expression is valid for all $w \in \mathcal{C}_w^\varepsilon$ (for real $w < -1$ one should apply (3.12) to all four gamma functions before using (3.13)).

The right-hand side of (6.13) above has singularities at $w = -1$ and $w = 0$, and the arc of the contour $\mathcal{C}_z^\varepsilon$ between the critical points \bar{z}_c and z_c crosses $(-1, 0)$. The integral of the error $O(T^{-1})$ in (6.13) over the arc from \bar{z}_c to z_c is bounded by $O(|z_c|/T)$, which is $O(1/T)$ because z_c belongs to a compact set \mathcal{Z} .

Let us now identify the extended sine kernel (2.3) in the remaining terms outside the double contour integral over $\mathcal{C}_z^\varepsilon$ and $\mathcal{C}_w^\varepsilon$. Observe that for $\Delta t \geq 0$ we have

$$\text{Res}_{w=0} (w^{-\Delta x-1} (1 + w)^{-\Delta t + \Delta x - 1}) = -(-1)^{\Delta x+1} \binom{\Delta t}{\Delta x} \mathbf{1}_{\Delta x \geq 0}.$$

By dragging the integration arc through 0 for $\Delta t > 0$, we obtain

$$\begin{aligned}
 & \mathbf{1}_{\Delta x \geq 0} \mathbf{1}_{\Delta t > 0} (-1)^{\Delta x+1} \binom{\Delta t}{\Delta x} - \frac{1}{2\pi i} \int_{z_c}^{\bar{z}_c} w^{-\Delta x-1} (1 + w)^{-\Delta t + \Delta x - 1} dw \\
 & = -\frac{1}{2\pi i} \int_{z_c}^{\bar{z}_c} w^{-\Delta x-1} (1 + w)^{-\Delta t + \Delta x - 1} dw,
 \end{aligned}$$

where in the right-hand side the arc crosses $(-1, 0)$ for $\Delta t \leq 0$ and $(0, +\infty)$ for $\Delta t > 0$. Changing the variables in the right-hand side as $w = \frac{z}{1-z}$ (so $z = \frac{w}{1+w}$) turns the above integral into $K_{z_c/(z_c+1)}(T + \Delta t, \Delta x; T, 0)$ (2.3), as desired. \square

To complete the proof of Theorem 2.7 it remains to show that the double contour integral in (6.12) is negligible as $N \rightarrow +\infty$. It has the form (cf. (4.3), (4.8))

$$(6.14) \quad \frac{1}{(2\pi i)^2} \int_{C_z^\varepsilon} dz \oint_{C_w^\varepsilon} dw \frac{1}{w-z} \exp\{T(S(z) - S(w))\} \\ \times \frac{(T + \Delta t)! \cdot T}{(T - 1)!} \frac{(Tw + 1)_{T-1}}{(Tw - \Delta x)_{T+\Delta t+1}}.$$

We need the following statement which we prove later in Section 6.4.

LEMMA 6.8. *Under Assumptions 1 and 2 the length of C_w^ε is bounded uniformly in N .*

This fact together with (6.13) implies that that the parts in (6.14) outside the exponent are bounded by a constant depending on $\Delta x, \Delta t$.

For z and w in a fixed small neighborhood of the critical point $z_c = z_c(N)$ which is bounded away from \mathbb{R} , we can Taylor expand the function $S(z)$. Because the second derivative of S is nonzero by Lemma 6.5, this leads to a convergent integral times $T^{-\frac{1}{2}}$ which goes to zero. This is a standard part of the steepest descent analysis, and we refer to, for example, [53], Section 3, for details.

Consider the situation when z and w are outside of this neighborhood of z_c . On the parts of the contours C_z^ε and C_w^ε inside the rectangle \mathcal{R} , we have the steepest descent/ascent properties. Together with Lemma 6.5, they imply that outside a sufficiently small neighborhood of z_c and for a sufficiently small fixed $\delta > 0$ (both depend only on the constants in Assumptions 1 and 2):

$$\operatorname{Re} S(z) - \operatorname{Re} S(z_c) < -\delta, \\ \operatorname{Re} S(w) - \operatorname{Re} S(z_c) > \delta.$$

Along the part of the z contour escaping to infinity $\operatorname{Re} S(z)$ cannot increase due to the second claim of Lemma 6.4.

Let us consider the possible change of $\operatorname{Re} S$ along C_z^ε and C_w^ε close to the real line. The vertical segments crossing the real line at points $(k + 1/2)/T, k \in \mathbb{Z}$, have length 2ε , and due to the first claim of Lemma 6.4 we see that the change of $\operatorname{Re} S$ is of order ε . The horizontal segments have length $C \cdot (\varepsilon \log^2 \varepsilon + T^{-1})$ for some $C > 0$ independent of N or ε (but C might depend on $\Delta x, \Delta t$). Using the third claim of Lemma 6.4, we can upper bound the absolute value of the change of $\operatorname{Re} S$ along the horizontal parts of the contours by a constant times

$$\left(-\log \varepsilon + \frac{1}{T^2 \varepsilon^2}\right) \left(\varepsilon \log^2 \varepsilon + \frac{1}{T}\right) = \frac{1}{\varepsilon^2 T^3} + \frac{\log^2 \varepsilon}{T^2 \varepsilon} - \frac{\log \varepsilon}{T} - \varepsilon \log^3 \varepsilon.$$

This can be made much smaller than δ : first choose ε that the forth term is small, and then choose N (thus $T(N)$) large enough so that the first three terms are also small. Therefore, the whole double contour integral (6.14) is negligible in the limit. This completes the proof of Theorem 2.7.

6.4. *Convergent initial data and proofs of Theorems 2.10, 2.11 and 2.12.* In this subsection, we present proofs of Theorems 2.10, 2.11 and 2.12 describing the convergence of the point processes to the extended sine process under suitable additional assumptions. Moreover, using similar arguments we prove Lemmas 6.5 and 6.8 which were formulated in the previous two subsections. These lemmas are not directly involved in the proofs of Theorems 2.10, 2.11 and 2.12.

In addition to Assumptions 1 and 2, let

$$(6.15) \quad \lim_{N \rightarrow +\infty} \frac{1}{T(N)} \sum_{i=1}^N \delta_{a_i(N)/T(N)} = \mu_{\text{loc}},$$

where μ_{loc} is a σ -finite measure, and the limit is understood according to Definition 2.9. By Assumption 1, μ_{loc} has a density (with respect to the Lebesgue measure) which is between ρ_\bullet and ρ^\bullet .

Let us also assume that the quantities (2.12) have a limit $d(\mathbb{R}) = \lim_{N \rightarrow +\infty} d_N(\mathbb{R})$. Then the meromorphic function $S'(z)$ (2.6) has a limit as $N \rightarrow +\infty$.

LEMMA 6.9. *Under the above assumptions and notation, we have $\lim_{N \rightarrow +\infty} S'(z) = S'_*(z)$ for all $z \in \mathbb{H}$, where*

$$(6.16) \quad \begin{aligned} S'_*(z) &= \int_{-\infty}^{\infty} \left(\frac{1}{z-v} + \frac{\mathbf{1}_{|v|>R}}{v} \right) \mu_{\text{loc}}(dv) - d(\mathbb{R}) \\ &+ \log(z+1) - \log z + i\pi - \log(\beta^{-1} - 1), \end{aligned}$$

and $R > 0$ is arbitrary (the limit does not depend on R). The convergence is uniform in z belonging to compact subsets of \mathbb{H} .

PROOF. Fix $z \in \mathbb{H}$. Let us use formula (2.9) for $S'(z)$. First, we have

$$-\pi \cot(\pi Tz) = -i\pi \frac{e^{i\pi Tz} + e^{-i\pi Tz}}{e^{i\pi Tz} - e^{-i\pi Tz}} \rightarrow i\pi,$$

because $e^{-i\pi Tz}$ dominates for $z \in \mathbb{H}$.

Next, the sum over $i = 1, \dots, T-1$ approximates the corresponding Riemann integral:

$$\sum_{i=1}^{T-1} \frac{1}{Tz+i} = \frac{1}{T} \sum_{i=1}^{T-1} \frac{1}{z+i/T} \rightarrow \int_0^1 \frac{dv}{z+v} = \log(z+1) - \log z,$$

and the convergence is uniform over z in compact subsets of \mathbb{H} .

Finally, recall the definition of the atomic measure (5.4), and note that (6.15) means that the measures $\mathcal{M}_T[\mathbf{a}]$ vaguely converge to μ_{loc} . The remaining N -dependent part of (2.9) can be written as

$$\frac{1}{T} \sum_{r=1}^N \frac{1}{z - a_r/T} = \int_{-\infty}^{+\infty} \frac{\mathcal{M}_T[\mathbf{a}](dv)}{z - v}.$$

Since the function $1/(z - v)$ in v does not have compact support, and its integral with respect to the Lebesgue measure diverges at infinity, one cannot directly apply (6.15) to the integral above. Here, we need a regularization afforded by the convergence of the constants $d_N(\mathbf{R})$ (2.12). Namely, take any $\mathbf{R} > 0$ and write

$$(6.17) \quad \int_{-\infty}^{+\infty} \frac{\mathcal{M}_T[\mathbf{a}](dv)}{z - v} = \int_{-\infty}^{+\infty} \left(\frac{1}{z - v} + \frac{\mathbf{1}_{|v| > \mathbf{R}}}{v} \right) \mathcal{M}_T[\mathbf{a}](dv) - d_N(\mathbf{R})$$

(this expression does not depend on \mathbf{R}). Now the function under the integral decays as v^{-2} at infinity, and so is Lebesgue integrable. Since the density of μ_{loc} is bounded, by restricting the integration to $[-M, M]$ for large M and applying (6.15), we conclude that the integral converges to the corresponding integral with respect to μ_{loc} . The uniformity of each convergence above is evident, so this completes the proof. \square

We are now in a good position to prove the helpful Lemmas 6.5 and 6.8 formulated previously: their proofs are similar to each other and to that of Lemma 6.9.

PROOF OF LEMMAS 6.5 AND 6.8. We will prove both statements simultaneously. If the contrary to one of the lemmas holds, then there exists a subsequence $\{N_k\}$ along which the local measures converge in the sense of (6.15) (consider measures of segments with rational endpoints and choose a diagonal subsequence), the constants $d_{N_k}(\mathbf{R})$ converge to $d(\mathbf{R})$, but:

- (Lemma 6.5) The derivatives $\mathbf{S}''(z_{N_k})$ converge to zero or infinity along a subsequence z_{N_k} belonging to a compact subset of the upper half-plane. Further choosing a subsequence of $\{N_k\}$ we may assume that along this subsequence the z_{N_k} 's converge to some point \tilde{z} is in the upper half-plane.
- (Lemma 6.8) The length of the contour $C_w^e(N_k)$ grows to infinity.

To simplify notation, let us use the sequence $\{N\}$ instead of $\{N_k\}$ in the rest of the proof.

Let us now show that there are constants c_N such that there exists $\lim_{N \rightarrow +\infty} (\mathbf{S}(z) - c_N)$, call it $\mathbf{S}_*(z)$, uniformly in z belonging to bounded subsets of \mathbb{H} . Moreover, \mathbf{S}_* is holomorphic in \mathbb{H} , and its derivative is given by (6.16).

Fix $z \in \mathbb{H}$. Recall that $S(z)$ is given by (4.8), and let us consider the five summands in that formula separately. First, observe that

$$\begin{aligned}
 &-\frac{1}{T} \log_{\mathbb{H}}(\sin(\pi Tz)) + \frac{2\pi i}{T} \left[\frac{1}{2} \operatorname{Re}(Tz) + \frac{1}{2} \right] - z \log(\beta^{-1} - 1) \\
 &\quad \rightarrow i\pi z - z \log(\beta^{-1} - 1)
 \end{aligned}$$

due to (4.6), (4.7). Next, the second sum in (4.8) approximates a convergent Riemann integral:

$$\frac{1}{T} \sum_{i=1}^{T-1} \log_{\mathbb{H}}\left(z + \frac{i}{T}\right) \rightarrow \int_0^1 \log(z + v) dv = (z + 1) \log(z + 1) - z \log z - 1$$

(in \mathbb{H} the branches \log and $\log_{\mathbb{H}}$ coincide). The first sum in (4.8) can be rewritten as

$$\begin{aligned}
 &\frac{1}{T} \sum_{r=1}^N \log_{\mathbb{H}}\left(z - \frac{a_r}{T}\right) \\
 &= \int_{-\infty}^{+\infty} \log(z - v) \mathcal{M}_T[\mathbf{a}](dv) \\
 &= \int_{-\infty}^{+\infty} \left(\log(z - v) - \log(\mathbf{i} - v) + (z - \mathbf{i}) \frac{\mathbf{1}_{|v|>R}}{v} \right) \mathcal{M}_T[\mathbf{a}](dv) \\
 &\quad + \int_{-\infty}^{+\infty} \log(\mathbf{i} - v) \mathcal{M}_T[\mathbf{a}](dv) - (z - \mathbf{i}) d_N(\mathbb{R}).
 \end{aligned}$$

The integrand $\log(z - v) - \log(\mathbf{i} - v) + (z - \mathbf{i}) \frac{\mathbf{1}_{|v|>R}}{v}$ decays as v^{-2} at infinity, so the first integral in the right-hand side converges as $N \rightarrow +\infty$ to the same integral over μ_{loc} . The second integral in the right-hand side does not depend on z , call it c_N . The third summand converges to $-(z - \mathbf{i})d(\mathbb{R})$. Thus, we have the convergence $S(z) - c_N \rightarrow S_*(z)$ to a holomorphic function, uniformly in z in compact subsets in \mathbb{H} . One can check that the derivative of S_* is (6.16).

Once the existence of the uniform limit $S_*(z) = \lim_{N \rightarrow +\infty} (S(z) - c_N)$ is established, we continue with separate arguments:

- (Lemma 6.5) We have $S''_*(\tilde{z}) = 0$ or ∞ and $\tilde{z} \in \mathbb{H}$. The second case is not possible since S_* is holomorphic in \mathbb{H} . If $S''_*(\tilde{z}) = 0$ then by Hurwitz’s theorem for all sufficiently large N there exist two complex critical points of $S - c_N$ (equivalently, of S) in the upper half-plane. This is impossible by Lemma 5.1. So in either case we get a contradiction.
- (Lemma 6.8) Observe that the length of the part of the contour C_W^ε close to the real line is bounded. From the convergence of $S(z) - c_N$, it follows that away from the real line the contour C_W^ε approximates the corresponding contour for $S_*(z)$, and the latter has finite length. We get a contradiction, too.

This proves both desired statements. \square

REMARK 6.10. The previous argument shows that the additional hypotheses of Theorem 2.10, namely, convergence of measures (6.15) and convergence $d(\mathbb{R}) = \lim_{N \rightarrow +\infty} d_N(\mathbb{R})$, are not too restrictive.

PROOF OF THEOREM 2.10. By Hurwitz’s theorem, the critical points $z_c(N) \in \mathbb{H}$ of S' converge, as $N \rightarrow +\infty$, to a critical point $z_* \in \mathbb{H}$, which belongs to the compact set \mathcal{Z} from Theorem 2.7. Applying the latter, we get the desired convergence. Equation (2.13) is simply $S'_*(z) = 0$ under a change of variables $z = u/(1 - u)$, which maps the upper half-plane onto itself. \square

As Proposition 2.18 is a particular case of Theorem 2.10, let us give its proof here.

PROOF OF PROPOSITION 2.18. Under the hypotheses of Proposition 2.18, Assumptions 1 and 2 clearly hold and, moreover, $d_N(\mathbb{R})$ is close to zero for $\mathbb{R} > 3$. Therefore, the function $S'_*(z)$ (6.16) looks as

$$\begin{aligned} S'_*(z) &= \frac{1}{2} \int_{-3h}^0 \frac{dv}{z - v} + \frac{1}{3} \int_0^{3h} \frac{dv}{z - v} + \frac{1}{2} \text{p.v.} \int_{|v|>3h} \frac{dv}{z - v} \\ &\quad + \log(z + 1) - \log z + i\pi - \log(\beta^{-1} - 1) \\ &= \frac{i\pi}{2} - \frac{1}{6} \log z + \frac{1}{6} \log(z - 3h) \\ &\quad + \log(z + 1) - \log z - \log(\beta^{-1} - 1). \end{aligned}$$

After the substitution $z = u/(1 - u)$ this leads to the equation (2.20) for the complex slope. \square

PROOF OF THEOREM 2.11. Fix an arbitrary $\varepsilon > 0$. Choose and fix $\delta > 0$ so small that $A_{\mathbb{R},\delta} < \varepsilon/3$ and that

$$\left| \int_{|v|>\delta} \frac{\mu_{\text{glob}}(dv)}{v} - \text{p.v.} \int_{-\infty}^{\infty} \frac{\mu_{\text{glob}}(dv)}{v} \right| < \frac{\varepsilon}{3}.$$

This approximation is possible because μ_{glob} is a probability measure, so $1/v$ is integrable at infinity, and thus the only singularity is at zero. Next, let N be so large that

$$\left| \sum_{i: |a_i(N)|>\delta N} \frac{1}{a_i(N)} - \int_{|u|>\delta} \frac{\mu_{\text{glob}}(dv)}{v} \right| < \frac{\varepsilon}{3}.$$

This is possible because the sum above is the same as the integral of $\mathbf{1}_{|v|>\delta}/v$ with respect to the atomic measure $\mu_{\text{glob}}^N = \frac{1}{N} \sum_{i=1}^N \delta_{a_i(N)/N}$ converging vaguely to μ_{glob} . Here, the vague convergence implies the convergence on the function

$\mathbf{1}_{|v|>\delta}/v$ because we are dealing with probability measures and so can cut away the tails at infinity. These estimates imply that

$$d_N(\mathbf{R}) = \sum_{i: \mathbf{R}T \leq |a_i(N)| \leq \delta N} \frac{1}{a_i(N)} + \int_{|v|>\delta} \frac{\mu_{\text{glob}}^N(dv)}{v}$$

is close to p.v. $\int_{-\infty}^{\infty} v^{-1} \mu_{\text{glob}}(dv)$ within ε , and so an application of Theorem 2.10 gives the result. \square

PROOF OF THEOREM 2.12. When μ_{loc} is a multiple of the Lebesgue measure, the integral in (6.16) can be explicitly computed:

$$\begin{aligned} & \int_{-\infty}^{\infty} \left(\frac{1}{z-v} + \frac{\mathbf{1}_{|v|>\mathbf{R}}}{v} \right) dv \\ &= \int_{-\mathbf{R}}^{\mathbf{R}} \frac{dv}{z-v} + \lim_{M \rightarrow +\infty} \int_{\mathbf{R} < |v| < M} \left(\frac{1}{z-v} + \frac{1}{v} \right) dv \\ &= \log(z + \mathbf{R}) - \log(z - \mathbf{R}) \\ &\quad + \lim_{M \rightarrow +\infty} (\log(z + M) - \log(z - M) + \log(z - \mathbf{R}) - \log(z + \mathbf{R})) \\ &= \lim_{M \rightarrow +\infty} (\log(z + M) - \log(z - M)), \end{aligned}$$

where the branches of all the logarithms above are standard. The last limit is equal to $-\mathbf{i}\pi$ because $\arg(z + M) \rightarrow 0$ while $\arg(z - M) \rightarrow \pi$. This immediately leads to the desired formula for u_* in Theorem 2.12. \square

REMARK 6.11. When μ_{loc} is a multiple of the Lebesgue measure, the above computation shows that the integral in (6.16) is independent of \mathbf{R} , and hence $d(\mathbf{R})$ is, too. This agrees with the fact that the difference $d_N(\mathbf{R}) - d_N(\mathbf{R}')$ (where, say, $\mathbf{R} > \mathbf{R}'$) is equal to the integral of $\mathcal{M}_T[\mathbf{a}](dv)/v$ over $\mathbf{R}' < |v| < \mathbf{R}$, and thus vanishes as $N \rightarrow +\infty$.

7. Applications: Proofs of Theorems 2.13, 2.15 and Proposition 2.16.

7.1. *Discretization of a continuous profile: Proof of Theorem 2.13.* Let us show that the initial data (2.18) defined using a twice continuously differentiable function f satisfies Assumptions 1 and 2, as well as additional hypotheses of Theorem 2.10.

Since $1 < f'(x) < +\infty$ for all $x \in [-\frac{1}{2}, \frac{1}{2}]$, Assumption 1 holds with

$$\rho_{\bullet} = \frac{1}{2} \inf_{x \in [-\frac{1}{2}, \frac{1}{2}]} \frac{1}{f'(x)}, \quad \rho^{\bullet} = \frac{1}{2} + \frac{1}{2} \sup_{x \in [-\frac{1}{2}, \frac{1}{2}]} \frac{1}{f'(x)}$$

on scales, say, $\mathbf{D}(N) = \lfloor N^{\eta/2} \rfloor$ and $\mathbf{Q}(N) = \lfloor N^{(1+\eta)/2} \rfloor$.

Next, the local and global measures as in (2.14) exist, μ_{loc} is the Lebesgue measure on \mathbb{R} times $q = 1/f'(\chi)$ (recall that $f(\chi) = 0$), and μ_{glob} has the density

$$\mu_{\text{glob}}(dv) = \frac{dv}{f'(f^{-1}(v))}.$$

Because $f' > 1$, the principal value integral

$$(7.1) \quad \text{p.v.} \int_{-\infty}^{+\infty} \frac{dv}{vf'(f^{-1}(v))} = \text{p.v.} \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{dx}{f(x)}$$

also exists.

Let us now consider the quantities $d_N(\mathbb{R})$ (2.12). Replace the condition $|a_i| \geq RT$ on i in (2.12) by $|i - N\chi| > RT/f'(\chi)$. The difference between the two sums can be estimated by a part of the harmonic series between RT and $RT + CT^2/N$, which is negligible. Thus,

$$(7.2) \quad \begin{aligned} \lim_{N \rightarrow +\infty} d_N(\mathbb{R}f'(\chi)) &= \lim_{N \rightarrow +\infty} \sum_{i: |i - N\chi| > RT} \frac{1}{[Nf(i/N)]} \\ &= \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{i: |i/N - \chi| > RT/N} \frac{1}{f(i/N)}. \end{aligned}$$

Taylor expand $f(i/N) = (i/N - \chi)f'(\chi) + (i/N - \chi)^2r$, where $|r|$ is uniformly bounded (here we use that f is twice continuously differentiable). Observe that the sum of

$$\frac{1}{Nf(i/N)} - \frac{1}{(i - N\chi)f'(\chi)} = -\frac{r}{N} \cdot \frac{1}{f'(\chi)^2 + rf'(\chi)(i/N - \chi)}$$

over i such that $RT/N < i/N - \chi < \delta$ (i.e., the one-sided sum) is bounded for sufficiently small $\delta > 0$, and goes to zero as $\delta \rightarrow 0$, and similarly for $-\delta < i/N - \chi < -RT/N$. The sum of $1/((i - N\chi)f'(\chi))$ over $|i - N\chi| > RT$ (i.e., the symmetric sum) is negligible for large N . Thus, the last sum in (7.2) over $|i/N - \chi| > RT/N$ is close to the same sum over $|i/N - \chi| > \delta$ for small δ , and the latter approximates the principal value integral (7.1). This implies Assumption 2 and the property that the $d_N(\mathbb{R})$'s converge as $N \rightarrow +\infty$. This completes the proof of Theorem 2.13.

7.2. *Random initial configuration.* Let us now consider the noncolliding Bernoulli random walk started from a random initial configuration. We assume that this random configuration belongs to $\{-M, -M + 1, \dots, M\}$, where $M \rightarrow +\infty$ is our main large parameter. Denote by

$$(7.3) \quad \mathbb{W}(M) = \bigcup_{k=0}^{2M+1} \{\vec{x} \in \mathbb{W}^k : -M \leq x_1 < \dots < x_k \leq M\}$$

the space of possible initial configurations (cf. (1.1)). The law of the initial configuration will be denoted by \mathbf{P}^M , and the configuration itself by $\mathfrak{A}(\mathbf{N})$ (here \mathbf{N} can be random). Let $\vec{\mathbf{X}}(t)$ stand for the noncolliding Bernoulli random walk started from $\mathfrak{A}(\mathbf{N})$.

When does $\vec{\mathbf{X}}(t)$ satisfy an annealed¹⁷ bulk limit theorem similar to the ones formulated in Section 2.3? Informally speaking, this happens when the walk started from a fixed “typical” configuration (with respect to \mathbf{P}^M) satisfies a bulk limit theorem with a constant complex slope u_* (i.e., independent of the randomness coming from the random initial data). Let us formalize this understanding; cf. [22, 31] for recent annealed limit theorems for uniformly random tilings.

PROPOSITION 7.1. *Choose and fix a time scale $T = T(M) \ll M$, $T(M) \rightarrow +\infty$. Suppose that there exist subsets $\mathbb{W}_{\text{reg}}(M) \subset \mathbb{W}(M)$, $M = 1, 2, \dots$, such that:*

1. $\lim_{M \rightarrow +\infty} \mathbf{P}^M(\mathbb{W}_{\text{reg}}(M)) = 1$;
2. *For any fixed sequence of (nonrandom) initial configurations $\mathfrak{A}(N_M) \in \mathbb{W}_{\text{reg}}(M)$ the bulk limit theorem near $x = 0$ (i.e., the conclusion of Theorem 2.10) holds for a complex slope u_* independent of this sequence $\mathfrak{A}(N_M)$.*

Then as $M \rightarrow +\infty$ the point process describing $\{\vec{\mathbf{X}}(T(M) + t)\}_t$ near $x = 0$ converges in distribution to the extended sine process of the complex slope u_ .*

In particular, hypotheses of Proposition 7.1 imply that the random number \mathbf{N} of particles in the initial configuration goes to infinity in probability (with respect to the \mathbf{P}^M 's).

PROOF OF PROPOSITION 7.1. Fix an event of the form

$F = \{\text{the configuration } \vec{\mathbf{X}}(T + t_i) \text{ on } \mathbb{Z} \text{ contains the point } y_i \text{ for all } i = 1, \dots, k\}$, where $k = 1, 2, \dots$, and $t_i, y_i \in \mathbb{Z}$ (cf. (2.1)). Such events generate the σ -algebra describing the configuration $\{\vec{\mathbf{X}}(T + t)\}_t$ near $x = 0$.

Let $\mathbb{P}_{\mathfrak{A}(N)}$ stand for the law of the noncolliding Bernoulli random walk started from the initial configuration $\mathfrak{A}(N)$. We need to show that

$$(7.4) \quad \lim_{M \rightarrow +\infty} \mathbf{E}_{\mathbf{P}^M}[\mathbb{P}_{\mathfrak{A}(\mathbf{N})}(F)] = \det[K_{u_*}(t\alpha, y\alpha; t\beta, y\beta)]_{\alpha, \beta=1}^k,$$

where K_{u_*} is the extended sine kernel (2.3), and $\mathbf{E}_{\mathbf{P}^M}$ denotes the expectation with respect to \mathbf{P}^M . We have

$$\mathbf{E}_{\mathbf{P}^M}[\mathbb{P}_{\mathfrak{A}(\mathbf{N})}(F)] = \mathbf{E}_{\mathbf{P}^M}[\mathbb{P}_{\mathfrak{A}(\mathbf{N})}(F)\mathbf{1}_{\mathbb{W}_{\text{reg}}(M)}] + \mathbf{E}_{\mathbf{P}^M}[\mathbb{P}_{\mathfrak{A}(\mathbf{N})}(F)\mathbf{1}_{\mathbb{W}(M) \setminus \mathbb{W}_{\text{reg}}(M)}].$$

¹⁷That is, with respect to the combined randomness coming from the initial configuration and from the random walk itself.

The second summand goes to zero by hypothesis 1, and the first summand can be estimated as

$$(7.5) \quad \mathbb{P}_{\mathfrak{A}_{\min}(N_{\min})}(F) \leq \mathbf{E}_{\mathbf{P}^M} [\mathbb{P}_{\mathfrak{A}(\mathbf{N})}(F) \mathbf{1}_{\mathbb{W}_{\text{reg}}(M)}] \leq \mathbb{P}_{\mathfrak{A}_{\max}(N_{\max})}(F),$$

where $\mathfrak{A}_{\max}(N_{\max})$ is the configuration which maximizes $\mathbb{P}_{\mathfrak{A}}(F)$ over all $\mathfrak{A} \in \mathbb{W}_{\text{reg}}(M)$ (it exists because this is a finite set), and similarly for $\mathfrak{A}_{\min}(N_{\min})$. Because both minimizing and maximizing configurations belong to $\mathbb{W}_{\text{reg}}(M)$, both bounds in (7.5) converge to the right-hand side of (7.4) by hypothesis 2, and so the desired convergence holds. \square

7.3. *Bernoulli initial data: Proof of Theorem 2.15.* Let the parameters p and α be as in Theorem 2.15, the time scale be $T(M) = \lfloor M^\eta \rfloor$, and the initial particle configuration on $\{-\lfloor M(1 - \alpha) \rfloor, -\lfloor M(1 - \alpha) \rfloor + 1, \dots, \lfloor M\alpha \rfloor - 1, \lfloor M\alpha \rfloor\}$ be obtained by putting a particle at each location with probability p independently of all others.

We will construct a subset $\mathbb{W}_{\text{reg}}(M) \subset \mathbb{W}(M)$ (7.3) satisfying Proposition 7.1 by imposing two conditions (of asymptotic \mathbf{P}^M -probability 1) on the configuration. First, fix $0 < \delta < \min(1 - \eta, \eta/7)$, and take scales $D(M) = \lfloor M^{\eta - \delta} \rfloor$, $Q(M) = \lfloor M^{\eta + \delta} \rfloor$. Denote

$$\mathbb{W}_{\text{reg},1}(M) = \left\{ \vec{x} \in \mathbb{W}(M) \text{ such that in every segment of length } \begin{matrix} D(M) \text{ inside } [-Q(M), Q(M)] \text{ the number of} \\ \text{points in the configuration } \vec{x} \text{ is between } pD - D^{2/3} \\ \text{and } pD + D^{2/3} \end{matrix} \right\}.$$

Since the expected number of points in one of the segments of length D is pD and the variance is of order D , by the Chebyshev inequality the probability that in one of such segments the (random) number of points is not between $pD - D^{2/3}$ and $pD + D^{2/3}$ can be bounded from above by a constant times $D^{-1/3}$. The number of segments of length D inside $[-Q, Q]$ is of order Q/D , and so

$$1 - \mathbf{P}^M(\mathbb{W}_{\text{reg},1}(M)) \leq C \cdot D^{-\frac{1}{3}} \frac{Q}{D} = C \cdot M^{-\frac{4}{3}(\eta - \delta) + \eta + \delta} = C \cdot M^{\frac{1}{3}(7\delta - \eta)} \rightarrow 0.$$

Clearly, configurations in $\mathbb{W}_{\text{reg},1}(M)$ satisfy Assumption 1. Moreover, for these configurations the local density of particles at 0 vaguely converges as $M \rightarrow +\infty$ to p times the Lebesgue measure on \mathbb{R} .

Second, recall $d_M(\mathbf{R})$ defined by (2.12) as a sum of $1/a_i$ over $|a_i| \geq RT(M)$, and interpret it as a sum of independent random variables δ_i/i over all $i \in \{-\lfloor M(1 - \alpha) \rfloor, \dots, \lfloor M\alpha \rfloor\}$, where δ_i is the indicator of the event that there is a point of the configuration at the location i . We have

$$\mathbf{E}_{\mathbf{P}^M}(d_M(\mathbf{R})) = p \sum_{\substack{j: |j| \geq RT(M) \\ -\lfloor M(1 - \alpha) \rfloor \leq j \leq \lfloor M\alpha \rfloor}} \frac{1}{j},$$

$$\mathbf{Var}_{\mathbf{P}^M}(\mathbf{d}_M(\mathbf{R})) = p(1 - p) \sum_{\substack{j: |j| \geq \mathbf{R}T(M) \\ -\lfloor M(1-\alpha) \rfloor \leq j \leq \lfloor M\alpha \rfloor}} \frac{1}{j^2}.$$

We see that the expectation approximates an integral

$$\mathbf{E}_{\mathbf{P}^M}(\mathbf{d}_M(\mathbf{R})) = \text{p.v.} \int_{-(1-\alpha)}^{\alpha} \frac{p}{v} dv + O(M^{-1}) = p \log\left(\frac{\alpha}{1-\alpha}\right) + O(M^{-1}),$$

and $\mathbf{Var}_{\mathbf{P}^M}(\mathbf{d}_M(\mathbf{R})) = O(T(M)^{-1})$, so the random variable $\mathbf{d}_M(\mathbf{R})$ converges as $M \rightarrow +\infty$ to the constant $\mathbf{d} = p \log\left(\frac{\alpha}{1-\alpha}\right)$. Thus, if we define for some fixed $\mathbf{R} > 0$:

$$\mathbb{W}_{\text{reg},2}(M) = \left\{ \vec{x} \in \mathbb{W}(M) : \left| p \log\left(\frac{\alpha}{1-\alpha}\right) - \sum_{i: |x_i| \geq \mathbf{R}T(M)} \frac{1}{x_i} \right| < M^{-\eta/3} \right\},$$

then by the Chebyshev inequality we have $\mathbf{P}^M(\mathbb{W}_{\text{reg},2}(M)) \rightarrow 1$. Configurations from the sets $\mathbb{W}_{\text{reg},2}(M)$ satisfy Assumption 2, and have $\mathbf{d}_M(\mathbf{R}) \rightarrow \mathbf{d}$.

Defining $\mathbb{W}_{\text{reg}}(M) = \mathbb{W}_{\text{reg},1}(M) \cap \mathbb{W}_{\text{reg},2}(M)$, we see that the \mathbf{P}^M -probabilities of these sets go to 1, while any sequence of configurations from $\mathbb{W}_{\text{reg}}(M)$ satisfies the hypotheses of Theorem 2.10, and hence the bulk limit theorem near $x = 0$. Thus, applying Proposition 7.1 we see that Theorem 2.15 is established.

7.4. *Sine process initial data: Proof of proposition 2.16.* Let the parameters ϕ and α be as in Proposition 2.16, the time scale be $T(M) = \lfloor M^\eta \rfloor$, and the initial particle configuration be obtained by restricting the configuration of the discrete sine process of density ϕ/π to $\{-\lfloor M(1-\alpha) \rfloor, -\lfloor M(1-\alpha) \rfloor + 1, \dots, \lfloor M\alpha \rfloor - 1, \lfloor M\alpha \rfloor\}$. We will use the same scales $\mathbf{D}(M), \mathbf{Q}(M)$ and sets $\mathbb{W}_{\text{reg},1}(M), \mathbb{W}_{\text{reg},2}(M)$ as in the Bernoulli case in Section 7.3 with p replaced by ϕ/π . This ensures that the second hypothesis of Proposition 7.1 holds for $\mathbb{W}_{\text{reg}}(M) = \mathbb{W}_{\text{reg},1}(M) \cap \mathbb{W}_{\text{reg},2}(M)$. To establish Proposition 2.16, it remains to show that

$$\lim_{M \rightarrow +\infty} \mathbf{P}^M(\mathbb{W}_{\text{reg},1}(M)) = \lim_{M \rightarrow +\infty} \mathbf{P}^M(\mathbb{W}_{\text{reg},2}(M)) = 1,$$

where now \mathbf{P}^M stands for the law of the initial configuration under the restriction of the discrete sine process. For $\mathbb{W}_{\text{reg},1}(M)$, observe that the variance under \mathbf{P}^M of the number of points in a segment of length \mathbf{D} (say, $\{1, \dots, \mathbf{D}\}$, since the sine process is translation invariant) can be estimated as

$$\begin{aligned} \mathbf{Var}_{\mathbf{P}^M} \left(\sum_{i=1}^{\mathbf{D}} \delta_i \right) &= \frac{\phi}{\pi} \mathbf{D} - \left(\frac{\phi}{\pi} \mathbf{D} \right)^2 + 2 \sum_{1 \leq i < j \leq \mathbf{D}} \mathbf{E}_{\mathbf{P}^M}(\delta_i \delta_j) \\ &= \mathbf{D} \frac{\phi}{\pi} \left(1 - \frac{\phi}{\pi} \right) - 2 \sum_{1 \leq i < j \leq \mathbf{D}} \left(\frac{\sin(\phi(i-j))}{\pi(i-j)} \right)^2 \\ (7.6) \qquad &\leq \mathbf{D} \frac{\phi}{\pi} \left(1 - \frac{\phi}{\pi} \right), \end{aligned}$$

where δ_i is as in Section 7.3. This variance does not exceed the one in the Bernoulli case, and so using the argument from Section 7.3 we conclude that $\mathbf{P}^M(\mathbb{W}_{\text{reg},1}(M)) \rightarrow 1$.

REMARK 7.2. In fact, the variance in the left-hand side of (7.6) grows as $O(\log D)$ (see [6], Lemma 4.6, [19], Section 4.2), but we do not need this for our proof.

For $\mathbb{W}_{\text{reg},2}(M)$ consider the random variable $d_M(\mathbf{R})$ (2.12). Arguing as in Section 7.3, we see that its expectation converges to $\frac{\phi}{\pi} \log(\frac{\alpha}{1-\alpha})$. Let us estimate its variance. Observe that for any subset $B \subset \mathbb{Z} \setminus \{0\}$ one has

$$(7.7) \quad \text{Var}_{\mathbf{P}^M} \left(\sum_{i \in B} \frac{\delta_i}{i} \right) = \frac{\phi}{\pi} \left(1 - \frac{\phi}{\pi} \right) \sum_{i \in B} \frac{1}{i^2} - 2 \sum_{i, j \in B: i < j} \frac{1}{ij} \left(\frac{\sin(\phi(i-j))}{\pi(i-j)} \right)^2.$$

Apply this with

$$B = \{-\lfloor M(1-\alpha) \rfloor, \dots, -RT-1, -RT, \dots, RT, RT+1, \dots, \lfloor M\alpha \rfloor - 1, \lfloor M\alpha \rfloor\}$$

for some fixed $R \in \mathbb{Z}_{\geq 1}$. We see that as $M \rightarrow +\infty$, the first sum in (7.7) decays as $O(T^{-1})$. Throwing away the pairs (i, j) of the same sign from the second sum, we can bound the second sum in (7.7) by a constant times

$$(7.8) \quad \begin{aligned} \sum_{k, j \geq RT} \frac{1}{kj(k+j)^2} &\leq \sum_{j \geq RT} \left(\sum_{k \geq 1} \frac{1}{kj(k+j)^2} \right) \\ &\leq \sum_{j \geq RT} \frac{C_1 + C_2 \log j}{j^2} = O(T^{-1+\epsilon}) \end{aligned}$$

for some $C_{1,2} > 0$ and an arbitrary small $\epsilon > 0$. Thus, the variance of $d_M(\mathbf{R})$ decays as $O(T^{-1+\epsilon})$, and so by the Chebyshev inequality we have $\mathbf{P}^M(\mathbb{W}_{\text{reg},2}(M)) \rightarrow 1$. Applying Proposition 7.1, we see that Proposition 2.16 holds.

REMARK 7.3. One can say that the constant $d = \frac{\phi}{\pi} \log(\frac{\alpha}{1-\alpha})$ in Proposition 2.16 corresponds via (2.15) to the global probability measure μ_{glob} which is the uniform measure on the segment $[-(1-\alpha)\frac{\pi}{\phi}, \alpha\frac{\pi}{\phi}]$. Indeed, this μ_{glob} is a limit as in (2.14) of random atomic measures corresponding to the sine process initial data, where as N one should take the random number of particles \mathbf{N} (it is concentrated around $\frac{\phi}{\pi}M$). Similarly, the constant $p \log(\frac{\alpha}{1-\alpha})$ in Theorem 2.15 corresponds to μ_{glob} being the uniform measure on $[-(1-\alpha)p^{-1}, \alpha p^{-1}]$.

REMARK 7.4. The proof of Proposition 2.16 carries over from the Bernoulli case modulo two estimates of the variance (7.6) and (7.7)–(7.8), which are rather

straightforward for the sine process. Thus, the bulk limit theorem should hold for rather general random initial data, but we will not formulate any other results in this direction.

APPENDIX A: DETERMINANTAL KERNELS FOR OTHER NONCOLLIDING PROCESSES

A.1. Noncolliding Poisson random walk. Taking the limit as $\beta \rightarrow 0$ and scaling to the continuous time as $t = \lfloor \beta^{-1} \tau \rfloor$, $\tau \in \mathbb{R}_{\geq 0}$, turns the noncolliding Bernoulli random walk into the noncolliding Poisson random walk—the continuous time dynamics of N independent speed 1 Poisson particles conditioned to never collide [46]. This Markov chain $\vec{X}(\tau)$ on \mathbb{W}^N has jump rates (cf. (1.2))

$$\mathbb{P}(\vec{X}(\tau + d\tau) = \vec{x}' \mid \vec{X}(\tau) = \vec{x}) = \begin{cases} \frac{V(\vec{x}')}{V(\vec{x})} d\tau + O(d\tau^2) & x'_i = x_i + 1 \text{ for some} \\ & i, \text{ and } x'_j = x_j \text{ for} \\ & j \neq i, \\ 1 - N d\tau + O(d\tau^2) & \vec{x}' = \vec{x}, \\ 0 & \text{otherwise,} \end{cases}$$

where $\vec{x}, \vec{x}' \in \mathbb{W}^N$ are arbitrary.¹⁸ The noncolliding Poisson random walk is also sometimes referred to as the Charlier process [6] due to the fact that if it starts from the densely packed initial configuration $(0, 1, \dots, N - 1)$, then its (fixed time) distribution is the Charlier orthogonal polynomial ensemble (cf. Remark 2.2).

THEOREM A.1. *The noncolliding Poisson random walk $\vec{X}(\tau)$ started from an arbitrary initial configuration $\vec{a} \in \mathbb{W}^N$ is determinantal in the sense of (2.1), with the kernel*

$$\begin{aligned} & K_{\vec{a}}^{\text{Poisson}}(\tau_1, x_1; \tau_2, x_2) \\ &= -\mathbf{1}_{x_1 \geq x_2} \mathbf{1}_{\tau_1 > \tau_2} \frac{(\tau_1 - \tau_2)^{x_1 - x_2}}{(x_1 - x_2)!} \\ &\quad - \frac{1}{(2\pi i)^2} \int_{x_2 - \frac{1}{2} - i\infty}^{x_2 - \frac{1}{2} + i\infty} dz \oint_{\text{all } w \text{ poles}} dw \frac{1}{w - z} \frac{\Gamma(x_2 - z)}{\Gamma(x_1 - w + 1)} \\ \text{(A.1)} \quad & \times \tau_1^{x_1 - w} \tau_2^{z - x_2} \prod_{r=1}^N \frac{z - a_r}{w - a_r}, \end{aligned}$$

¹⁸In particular, this implies $\sum_{i=1}^N V(\vec{x} + e_i) = NV(\vec{x})$, where e_i is the i th basis vector $(0, \dots, 0, 1, 0, \dots, 0)$.

where $x_{1,2} \in \mathbb{Z}$, $\tau_{1,2} > 0$, the z integration contour is a vertical line $\text{Re } z = x_2 - \frac{1}{2}$ traversed upwards, and the w contour is a positively oriented circle or a union of two circles encircling all the w poles $\{\dots, x_1 - 1, x_1\} \cap \{a_1, \dots, a_N\}$ of the integrand except $w = z$.

PROOF. We will obtain $K_{\bar{a}}^{\text{Poisson}}$ from $K_{\bar{a};\beta}^{\text{Bernoulli}}$ (2.2) via the $\beta \rightarrow 0$ limit described above. Employing Remark 3.7, write $K_{\bar{a};\beta}^{\text{Bernoulli}}$ as

$$\begin{aligned}
 & K_{\bar{a};\beta}^{\text{Bernoulli}}(t_1, x_1; t_2, x_2) \\
 &= \mathbf{1}_{x_1 \geq x_2} \mathbf{1}_{t_1 > t_2} (-1)^{x_1 - x_2 + 1} \binom{t_1 - t_2}{x_1 - x_2} + \frac{t_1!}{(t_2 - 1)!} \frac{1}{(2\pi i)^2} \\
 &\quad \times \int_{x_2 - \frac{1}{2} - i\infty}^{x_2 - \frac{1}{2} + i\infty} dz \oint_{\text{all } w \text{ poles}} dw \frac{(z - x_2 + 1)_{t_2 - 1}}{(w - x_1)_{t_1 + 1}} \\
 \text{(A.2)} \quad &\quad \times \frac{1}{w - z} \frac{\sin(\pi w)}{\sin(\pi z)} \left(\frac{1 - \beta}{\beta}\right)^{w - z} \prod_{r=1}^N \frac{z - a_r}{w - a_r}.
 \end{aligned}$$

Here, the w contour (a circle or a union of two circles) can be taken to encircle the points $\{a_1, a_2, \dots, a_N\}$, which contain all the w poles except $w = z$. (Indeed, for $w = a_i$ to be a pole, it must additionally satisfy $(a_i - x_1)_{t_1 + 1} = 0$ to not cancel with the zero coming from $\sin(\pi w)$.) Therefore, the integration contours do not depend on $t_{1,2}$, and we can take the Poisson rescaling of (A.2), that is, $\beta \rightarrow 0$ and $t_{1,2} = \lfloor \beta^{-1} \tau_{1,2} \rfloor$ with $\tau_1 \geq 0$, $\tau_2 > 0$.

The t -dependent part of the first summand in (A.2) scales as

$$\begin{aligned}
 & \mathbf{1}_{\tau_1 > \tau_2} \binom{\lfloor \beta^{-1} \tau_1 \rfloor - \lfloor \beta^{-1} \tau_2 \rfloor}{x_1 - x_2} \\
 &= \mathbf{1}_{\tau_1 > \tau_2} \frac{\Gamma(\lfloor \beta^{-1} \tau_1 \rfloor - \lfloor \beta^{-1} \tau_2 \rfloor + 1)}{\Gamma(\lfloor \beta^{-1} \tau_1 \rfloor - \lfloor \beta^{-1} \tau_2 \rfloor - x_1 + x_2 + 1)(x_1 - x_2)!} \\
 &\sim \mathbf{1}_{\tau_1 > \tau_2} \frac{(\tau_1 - \tau_2)^{x_1 - x_2}}{(x_1 - x_2)!} \beta^{-(x_1 - x_2)},
 \end{aligned}$$

where we used (3.7). Similarly, for the part of the integrand depending on β and $t_{1,2}$ we have

$$\begin{aligned}
 & \frac{t_1!}{(t_2 - 1)!} \frac{(z - x_2 + 1)_{t_2 - 1}}{(w - x_1)_{t_1 + 1}} \left(\frac{1 - \beta}{\beta}\right)^{w - z} \frac{\sin(\pi w)}{\sin(\pi z)} \\
 &= \frac{\Gamma(t_1 + 1)}{\Gamma(w - x_1 + t_1 + 1)} \frac{\Gamma(z - x_2 + t_2)}{\Gamma(t_2)} \frac{\Gamma(w - x_1)}{\Gamma(z - x_2 + 1)}
 \end{aligned}$$

$$\begin{aligned} & \times \left(\frac{1 - \beta}{\beta} \right)^{w-z} \frac{\sin(\pi w)}{\sin(\pi z)} \\ & \sim \frac{\Gamma(x_2 - z)}{\Gamma(x_1 - w + 1)} (-1)^{x_1 - x_2 + 1} \tau_1^{x_1 - w} \tau_2^{z - x_2} \beta^{-(x_1 - x_2)}. \end{aligned}$$

Note that $(1 - \beta)^{w-z} \rightarrow 1$ as $\beta \rightarrow 0$, and in the last step we used (3.12). This implies

$$\lim_{\beta \rightarrow 0} ((-\beta)^{x_1 - x_2} K_{\tilde{a}; \beta}^{\text{Bernoulli}}(\lfloor \beta^{-1} \tau_1 \rfloor, x_1; \lfloor \beta^{-1} \tau_1 \rfloor, x_2)) = K_{\tilde{a}}^{\text{Poisson}}(\tau_1, x_1; \tau_2, x_2),$$

where $K_{\tilde{a}}^{\text{Poisson}}$ is given by (A.1). Because the multiplication by $(-\beta)^{x_1 - x_2}$ does not change the correlation functions (cf. footnote 7), this completes the proof. \square

The correlation kernel of Theorem A.1 appears to be new. By analogy with the results in Section 2.3, we believe that the local statistics of the noncolliding Poisson random walk are universally described by an extension of the discrete sine kernel with the continuous time parameter. This extension first appeared in [11], see also [4, 14] for a general discussion of extensions of the discrete sine kernel. We will not pursue this in the present paper.

A.2. Dyson Brownian motion. A diffusion scaling brings the kernel $K_{\tilde{a}; \beta}^{\text{Bernoulli}}$ (2.2) to the kernel of the Dyson Brownian motion. We will use the following scaling (where $M \rightarrow +\infty$):

$$\begin{aligned} t_{1,2} &= \lfloor M \tau_{1,2} \rfloor, \\ x_{1,2} &= \lfloor \beta M \tau_{1,2} + \xi_{1,2} \sqrt{\beta(1 - \beta)} \sqrt{M} \rfloor, \\ a_i &= \lfloor \alpha_i \sqrt{\beta(1 - \beta)} \sqrt{M} \rfloor, \end{aligned}$$

where $(\alpha_1 \leq \dots \leq \alpha_N) \in \mathbb{R}^N$ are the rescaled starting points (by agreement, when some of the α_i 's coincide, the corresponding a_i 's differ by 1, so that the discrete noncolliding Bernoulli random walk is well defined).

THEOREM A.2. *Under the above scaling and up to a gauge transformation as in footnote 7, the kernel $(M\beta(1 - \beta))^{\frac{1}{2}} K_{\tilde{a}; \beta}^{\text{Bernoulli}}$ converges to the following kernel:*

$$\begin{aligned} & K_{\tilde{a}; \beta}^{\text{DBM}}(\tau_1, \xi_1; \tau_2, \xi_2) \\ &= -\frac{\mathbf{1}_{\tau_1 > \tau_2}}{\sqrt{2\pi \Delta \tau}} \exp\left\{-\frac{(\Delta \xi)^2}{2\Delta \tau}\right\} \\ & \quad - \frac{1}{(2\pi \mathbf{i})^2 \sqrt{\tau_1 \tau_2}} \int_{c - \mathbf{i}\infty}^{c + \mathbf{i}\infty} dz \oint_{\text{all } w \text{ poles}} dw \frac{1}{w - z} \\ & \quad \times \exp\left\{\frac{\tau_1(z - \xi_2)^2 - \tau_2(w - \xi_1)^2}{2\tau_1 \tau_2}\right\} \prod_{r=1}^N \frac{z - \alpha_r}{w - \alpha_r}, \end{aligned} \tag{A.3}$$

where $\xi_{1,2} \in \mathbb{R}$, $\tau_{1,2} > 0$, and we use the notation $\Delta\xi = \xi_1 - \xi_2$, $\Delta\tau = \tau_1 - \tau_2$. The z contour is a vertical line which lies to the left of all the α_r 's (i.e., $c < \alpha_1$), and the w contour is a positively oriented circle encircling all the α_r 's.

The multiplication by $(M\beta(1 - \beta))^{\frac{1}{2}}$ corresponds to the rescaling of the space from discrete to continuous (the correlation kernel should be viewed as a kernel of an integral operator). Note also that the kernel (A.3) can similarly be obtained as a diffusion limit of the Poisson kernel (A.1), but we will not perform this computation.

PROOF OF THEOREM A.2. Let us denote $\sigma = \sqrt{\beta(1 - \beta)}$ to shorten the notation. Changing the variables as $z = \tilde{z}\sigma\sqrt{M}$, $w = \tilde{w}\sigma\sqrt{M}$, and renaming back to z, w , we can rewrite (2.2) as

$$\begin{aligned}
 & K_{\tilde{a};\beta}^{\text{Bernoulli}}(t_1, x_1; t_2, x_2) \\
 &= \mathbf{1}_{x_1 \geq x_2} \mathbf{1}_{t_1 > t_2} (-1)^{x_1 - x_2 + 1} \binom{t_1 - t_2}{x_1 - x_2} \\
 &+ \sigma\sqrt{M} \frac{t_1!}{(t_2 - 1)!} \frac{1}{(2\pi\mathbf{i})^2} \int_{(x_2 - t_2 + \frac{1}{2})/(\sigma\sqrt{M}) - i\infty}^{(x_2 - t_2 + \frac{1}{2})/(\sigma\sqrt{M}) + i\infty} dz \\
 &\times \oint_{\text{all } w \text{ poles}} dw \frac{(z\sigma\sqrt{M} - x_2 + 1)_{t_2 - 1}}{(w\sigma\sqrt{M} - x_1)_{t_1 + 1}} \\
 &\times \frac{1}{w - z} \frac{\sin(\pi w\sigma\sqrt{M})}{\sin(\pi z\sigma\sqrt{M})} \left(\frac{1 - \beta}{\beta}\right)^{(w - z)\sigma\sqrt{M}} \\
 &\times \prod_{r=1}^N \frac{z - \lfloor \alpha_r \sigma\sqrt{M} \rfloor / (\sigma\sqrt{M})}{w - \lfloor \alpha_r \sigma\sqrt{M} \rfloor / (\sigma\sqrt{M})}.
 \end{aligned}
 \tag{A.4}$$

The w contour encircles all the w poles of the integrand except $w = z$, which are close to the points $(\alpha_1 \leq \dots \leq \alpha_N)$. For large M , the z contour will get shifted further to the left; the computation below shows that there will be no poles crossed while doing this.

Denote

$$\begin{aligned}
 c_{1,2} &= t_{1,2} - M\tau_{1,2}, \\
 d_{1,2} &= x_{1,2} - \beta M\tau_{1,2} - \xi_{1,2}\sigma\sqrt{M},
 \end{aligned}$$

these are numbers between -1 and 0 . Let us first consider asymptotics of the nonintegral summand. For large M , the two indicators reduce simply to $\mathbf{1}_{\tau_1 > \tau_2}$,

and the binomial coefficient has the asymptotics

$$\begin{aligned} \binom{t_1 - t_2}{x_1 - x_2} &= \frac{\Gamma(t_1 - t_2 + 1)}{\Gamma(x_1 - x_2 + 1)\Gamma(t_1 - t_2 - (x_1 - x_2) + 1)} \\ &= \frac{1/\sqrt{M}}{\sqrt{2\pi\sigma\Delta\tau}} \left(\frac{1-\beta}{\beta}\right)^{\sigma\sqrt{M}\Delta\xi} \\ &\quad \times e^{-M\Delta\tau(\beta\log\beta+(1-\beta)\log(1-\beta))} \beta^{d_2-d_1} (1-\beta)^{c_2-c_1+d_1-d_2} \\ &\quad \times e^{-(\Delta\xi)^2/(2\Delta\tau)} (1 + O(1/\sqrt{M})), \end{aligned}$$

where we used (3.13).

Now consider the asymptotics of various parts of the integrand. We have

$$\frac{t_1!}{(t_2 - 1)!} = M(M\tau_1)^{M\tau_1+c_1} (M\tau_2)^{-M\tau_2-c_2} e^{-M\Delta\tau} \sqrt{\tau_1\tau_2} (1 + O(1/M)).$$

We can also write

$$\begin{aligned} &\frac{(z\sigma\sqrt{M} - x_2 + 1)_{t_2-1}}{(w\sigma\sqrt{M} - x_1)_{t_1+1}} \\ &= \frac{\Gamma(z\sigma\sqrt{M} - x_2 + t_2)}{\Gamma(z\sigma\sqrt{M} - x_2 + 1)} \frac{\Gamma(w\sigma\sqrt{M} - x_1)}{\Gamma(w\sigma\sqrt{M} - x_1 + t_1 + 1)} \\ &= (-1)^{x_1-x_2+1} \frac{\Gamma(z\sigma\sqrt{M} - x_2 + t_2)}{\Gamma(w\sigma\sqrt{M} - x_1 + t_1 + 1)} \frac{\Gamma(-z\sigma\sqrt{M} + x_2)}{\Gamma(1 - w\sigma\sqrt{M} + x_1)} \\ &\quad \times \frac{\sin(\pi z\sigma\sqrt{M})}{\sin(\pi w\sigma\sqrt{M})}, \end{aligned}$$

where we used (3.12). This cancels with the existing ratio of the sine functions. Continuing with the asymptotics, we obtain

$$\begin{aligned} &\frac{\Gamma(z\sigma\sqrt{M} - x_2 + t_2)}{\Gamma(w\sigma\sqrt{M} - x_1 + t_1 + 1)} \frac{\Gamma(-z\sigma\sqrt{M} + x_2)}{\Gamma(1 - w\sigma\sqrt{M} + x_1)} \\ &= M^{-2} (M\tau_1)^{-M\tau_1-c_1} (M\tau_2)^{M\tau_2+c_2} e^{M\Delta\tau} \\ &\quad \times e^{-M\Delta\tau(\beta\log\beta+(1-\beta)\log(1-\beta))} \\ &\quad \times \beta^{d_2-d_1} (1-\beta)^{c_2-c_1+d_1-d_2} \left(\frac{1-\beta}{\beta}\right)^{\sigma\sqrt{M}(\Delta\xi+z-w)} \\ &\quad \times \frac{1}{\beta(1-\beta)\tau_1\tau_2} \exp\left\{\frac{\tau_1(z-\xi_2)^2 - \tau_2(w-\xi_1)^2}{2\tau_1\tau_2}\right\} (1 + O(1/\sqrt{M})). \end{aligned}$$

We see that the factor

$$(-1)^{x_1-x_2} e^{-M\Delta\tau(\beta \log \beta + (1-\beta) \log(1-\beta))} \beta^{d_2-d_1} (1-\beta)^{c_2-c_1+d_1-d_2} \left(\frac{1-\beta}{\beta}\right)^{\sigma\sqrt{M}\Delta\xi}$$

appearing in both summands in (A.4) is of the form $f(\tau_1, \xi_1)/f(\tau_2, \xi_2)$, and thus does not affect the correlation functions (cf. footnote 7). The remaining parts of (A.4) multiplied by $\sigma\sqrt{M}$ converge to (A.3), as desired. \square

Since the noncolliding Bernoulli random walks converge under the diffusion scaling to the Dyson Brownian motion [27], the kernel (A.3) is the correlation kernel for the latter process started from the arbitrary initial configuration $(\alpha_1 \leq \dots \leq \alpha_N)$. When $\tau_1 = \tau_2$, the kernel (A.3) turns into the one appeared in [17, 37, 59]. Utilizing (A.3), these papers show that the local statistics of the eigenvalues of the deformed GUE ensemble (equivalently, the distribution of the Dyson Brownian motion started from $(\alpha_1 \leq \dots \leq \alpha_N)$ at a fixed time) are universally governed by the continuous sine kernel.

APPENDIX B: REPRESENTATION-THEORETIC INTERPRETATION OF NONCOLLIDING WALKS

The random matrix analogue of noncolliding Bernoulli or Poisson random walks is the GUE Dyson Brownian motion. The distribution of the Dyson Brownian motion started from an arbitrary initial configuration $(\alpha_1 \leq \dots \leq \alpha_N) \in \mathbb{R}^N$ after time $t > 0$ can also be interpreted as the eigenvalue distribution of the deformed GUE ensemble $A + \sqrt{t}G$, where $A = \text{diag}(\alpha_1, \dots, \alpha_N)$ is a fixed diagonal matrix, and G is an $N \times N$ random matrix from the GUE.¹⁹ Let us discuss a similar interpretation of the noncolliding Bernoulli or Poisson random walks in terms of representation theory of unitary groups.

Irreducible representations of the unitary group $U(N)$ can be parametrized by points of \mathbb{W}^N . Let $\chi_{\vec{x}}(u_1, \dots, u_N)$, $\vec{x} \in \mathbb{W}^N$ denote the corresponding normalized irreducible characters. Here, the u_i 's are eigenvalues of the matrix from $U(N)$, and the characters $\chi_{\vec{x}}$ are normalized in the sense that $\chi_{\vec{x}}(1, \dots, 1) = 1$. These characters are the normalized Schur polynomials:

$$\chi_{\vec{x}}(u_1, \dots, u_N) = \frac{s_\lambda(u_1, \dots, u_N)}{\dim_N \vec{x}},$$

where $\lambda = (\lambda_1 \geq \dots \geq \lambda_N)$ with $\lambda_i = x_{N+1-i} + i - N$ is the highest weight of the representation, and

$$\dim_N \vec{x} = s_\lambda(1, \dots, 1) = \prod_{1 \leq i < j \leq N} \frac{x_j - x_i}{j - i} = \frac{V(\vec{x})}{V(0, 1, \dots, N - 1)}$$

¹⁹There are different normalizations of the GUE random matrices and Dyson Brownian motion in the literature (cf. [1], (2.2.2), [63], Example 2), and here we assume that they agree.

is the dimension of this representation. Details on representations of unitary groups can be found in, for example, [66].

An (abstract) normalized character χ of $U(N)$ is defined as a nonnegative definite continuous function on $U(N)$ which satisfies $\chi(e) = 1$ and $\chi(ab) = \chi(ba)$ for any $a, b \in U(N)$ (i.e., we speak about characters which do not necessarily correspond to actual representations). The set of such characters is convex, and the normalized irreducible characters are its extreme points. Thus, any abstract character can be decomposed into irreducibles as $\chi(u_1, \dots, u_N) = \sum_{\vec{x} \in \mathbb{W}^N} c_{\vec{x}} \chi_{\vec{x}}(u_1, \dots, u_N)$, where the numbers $\{c_{\vec{x}}\}_{\vec{x} \in \mathbb{W}^N}$ are nonnegative and sum to one, hence they define a probability distribution on \mathbb{W}^N .

The product of two normalized characters $\chi^{(1)}$ and $\chi^{(2)}$ is also a normalized character. (If both $\chi^{(1)}$ and $\chi^{(2)}$ correspond to actual representations, then $\chi^{(1)}\chi^{(2)}$ is the normalized character of the tensor product of these representations.) The product $\chi^{(1)}\chi^{(2)}$ then can be decomposed into irreducibles, thus yielding a probability distribution on \mathbb{W}^N .

Fix $\vec{a} \in \mathbb{W}^N$ and take the irreducible normalized character $\chi_{\vec{a}}$ as $\chi^{(1)}$. Let $\chi^{(2)}$ be the restriction to $U(N)$ of a certain extreme character of the infinite-dimensional unitary group $U(\infty)$. We will consider two classes of such characters of $U(N)$ having the form

$$\chi_{\beta;t}(u_1, \dots, u_N) = \prod_{r=1}^N (1 - \beta + \beta u_r)^t \quad \text{or} \quad \chi_{\tau}(u_1, \dots, u_N) = \prod_{r=1}^N e^{\tau(u_r - 1)},$$

where $t \in \mathbb{Z}_{\geq 0}$ and $\tau > 0$. There is a number of papers discussing classification of extreme characters of $U(\infty)$; for example, see [26, 54, 64, 65], and other references in [12] and [57].

PROPOSITION B.1. *The probability weights $\{c_{\vec{x}}\}_{\vec{x} \in \mathbb{W}^N}$ arising from the decomposition*

$$\chi_{\vec{a}}(u_1, \dots, u_N) \chi_{\beta;t}(u_1, \dots, u_N) = \sum_{\vec{x} \in \mathbb{W}^N} c_{\vec{x}} \chi_{\vec{x}}(u_1, \dots, u_N)$$

describe the distribution of the noncolliding Bernoulli random walk with parameter β started from the initial configuration \vec{a} after t steps.

Similarly, the decomposition of $\chi_{\vec{a}}(u_1, \dots, u_N) \chi_{\tau}(u_1, \dots, u_N)$ into irreducibles corresponds to the distribution of the noncolliding Poisson random walk started from the configuration \vec{a} after time τ .

The case $\vec{a} = (0, 1, \dots, N - 1)$ leads to the trivial representation: $\chi_{\vec{a}}(u_1, \dots, u_N) \equiv 1$. In this case, the distribution of the noncolliding Bernoulli or Poisson random walks is related to the decomposition of extreme characters of $U(\infty)$ into irreducibles. Probabilistic properties of the corresponding measures were studied in, for example, [6, 9]. These measures can be regarded as discrete analogues of the GUE eigenvalue distribution.

PROOF OF PROPOSITION B.1. This fact is well known to specialists, but we include its proof for completeness.

It suffices to consider only the Bernoulli $t = 1$ case, because the general t case follows by induction, and the Poisson statement follows by a simple limit transition. The result would follow if we interpret the coefficients c_μ in the decomposition

$$\begin{aligned} & \frac{s_\lambda(u_1, \dots, u_N)}{s_\lambda(1, \dots, 1)} \prod_{r=1}^N (\beta u_r + 1 - \beta) \\ &= \sum_{\mu_1 \geq \dots \geq \mu_N} c_\mu \frac{s_\mu(u_1, \dots, u_N)}{s_\mu(1, \dots, 1)}, \quad \lambda_i = a_{N+1-i} + i - N, \end{aligned}$$

as one-step transition probabilities (1.2). Multiply the above decomposition by $s_\lambda(1, \dots, 1)$ and the Vandermonde in the u_i 's, and expand the determinants in

$$\begin{aligned} & \det[u_i^{\lambda_j + N - j}]_{i,j=1}^N \prod_{r=1}^N (\beta u_r + 1 - \beta) \\ &= \sum_{\mu_1 \geq \dots \geq \mu_N} c_\mu \det[u_i^{\mu_j + N - j}]_{i,j=1}^N \frac{s_\lambda(1, \dots, 1)}{s_\mu(1, \dots, 1)}. \end{aligned}$$

Because of the ordering in λ and μ , it suffices to consider the coefficient by $u_1^{\mu_1 + N - 1} \dots u_N^{\mu_N}$ in $u_1^{\lambda_1 + N - 1} u_2^{\lambda_2 + N - 2} \dots u_N^{\lambda_N}$ multiplied by $\prod_{r=1}^N (\beta u_r + 1 - \beta)$. Clearly, picking βu_r from each r th factor corresponds to the r th particle jumping by one to the right, while picking $(1 - \beta)$ means that this particle stays put. Particle collisions are not allowed because $\mu_1 + N - 1 > \dots > \mu_N$, and the factor $s_\lambda(1, \dots, 1)/s_\mu(1, \dots, 1)$ can be identified with the ratio of the Vandermondes in (1.2).

The Poisson case follows from the above argument in the limit as $\beta \searrow 0$. \square

We see that tensor multiplication of representations (and, more generally, multiplication of normalized characters) is a discrete analogue of the matrix addition. Moreover, multiplying by a suitable extreme character of $U(\infty)$ corresponds to adding a multiple of the GUE matrix. This similarity can be continued further to include the operation of the free convolution—its discrete analogue is the so-called quantized free convolution; see [18].

Therefore, our main universality results in Section 2.3 can be reformulated as bulk universality for tensor products of two representations of $U(N)$, when one of the factors is arbitrary, and the other factor is the specific representation $\chi_{\beta;t}$ (with t large). We conjecture that under mild technical conditions the same bulk universality should hold for tensor products of two arbitrary representations of $U(N)$ (cf. [3] for a progress toward a similar random matrix result). A weaker version of the bulk universality for tensor products of two arbitrary representations is established in [31].

Acknowledgments. We are grateful to Alexei Borodin for valuable comments on an earlier draft of the paper, and to Paul Bourgade for helpful remarks. We are grateful to the anonymous reviewer for valuable comments which helped improve our technical arguments.

REFERENCES

- [1] ANDERSON, G. W., GUIONNET, A. and ZEITOUNI, O. (2010). *An Introduction to Random Matrices. Cambridge Studies in Advanced Mathematics* **118**. Cambridge Univ. Press, Cambridge. MR2760897
- [2] BAIK, J., KRIECHERBAUER, T., MCLAUGHLIN, K. T.-R. and MILLER, P. D. (2007). *Discrete Orthogonal Polynomials: Asymptotics and Applications. Annals of Mathematics Studies* **164**. Princeton Univ. Press, Princeton, NJ. Available at arXiv:math/0310278 [math.CA]. MR2283089
- [3] BAO, Z., ERDŐS, L. and SCHNELLI, K. (2017). Local law of addition of random matrices on optimal scale. *Comm. Math. Phys.* **349** 947–990. Available at arXiv:1509.07080 [math.PR]. MR3602820
- [4] BORODIN, A. (2007). Periodic Schur process and cylindric partitions. *Duke Math. J.* **140** 391–468. Available at arXiv:math/0601019 [math.CO]. MR2362241
- [5] BORODIN, A. (2011). Determinantal point processes. In *The Oxford Handbook of Random Matrix Theory* 231–249. Oxford Univ. Press, Oxford. Available at arXiv:0911.1153 [math.PR]. MR2932631
- [6] BORODIN, A. and FERRARI, P. L. (2014). Anisotropic growth of random surfaces in $2 + 1$ dimensions. *Comm. Math. Phys.* **325** 603–684. Available at arXiv:0804.3035 [math-ph]. MR3148098
- [7] BORODIN, A. and GORIN, V. (2016). Lectures on integrable probability. In *Probability and Statistical Physics in St. Petersburg. Proc. Sympos. Pure Math.* **91** 155–214. Amer. Math. Soc., Providence, RI. Available at arXiv:1212.3351 [math.PR]. MR3526828
- [8] BORODIN, A., GORIN, V. and RAINS, E. M. (2010). q -distributions on boxed plane partitions. *Selecta Math. (N.S.)* **16** 731–789. Available at arXiv:0905.0679 [math-ph]. MR2734330
- [9] BORODIN, A. and KUAN, J. (2008). Asymptotics of Plancherel measures for the infinite-dimensional unitary group. *Adv. Math.* **219** 894–931. Available at arXiv:0712.1848 [math.RT]. MR2442056
- [10] BORODIN, A., OKOUNKOV, A. and OLSHANSKI, G. (2000). Asymptotics of Plancherel measures for symmetric groups. *J. Amer. Math. Soc.* **13** 481–515. Available at arXiv:math/9905032 [math.CO]. MR1758751
- [11] BORODIN, A. and OLSHANSKI, G. (2006). Stochastic dynamics related to Plancherel measure on partitions. In *Representation Theory, Dynamical Systems, and Asymptotic Combinatorics. Amer. Math. Soc. Transl. Ser. 2* **217** 9–21. Amer. Math. Soc., Providence, RI. Available at arXiv:math-ph/0402064. MR2276098
- [12] BORODIN, A. and OLSHANSKI, G. (2012). The boundary of the Gelfand–Tsetlin graph: A new approach. *Adv. Math.* **230** 1738–1779. Available at arXiv:1109.1412 [math.CO]. MR2927353
- [13] BORODIN, A. and PETROV, L. (2014). Integrable probability: From representation theory to Macdonald processes. *Probab. Surv.* **11** 1–58. Available at arXiv:1310.8007 [math.PR]. MR3178541
- [14] BORODIN, A. and SHLOSMA, S. (2010). Gibbs ensembles of nonintersecting paths. *Comm. Math. Phys.* **293** 145–170. Available at arXiv:0804.0564 [math-ph]. MR2563801
- [15] BOURGADE, P., ERDŐS, L., YAU, H.-T. and YIN, J. (2016). Fixed energy universality for generalized Wigner matrices. *Comm. Pure Appl. Math.* **69** 1815–1881. Available at arXiv:1407.5606 [math.PR]. MR3541852

- [16] BREUER, J. and DITS, M. (2012). Nonintersecting paths with a staircase initial condition. *Electron. J. Probab.* **17** no. 60, 24. Available at [arXiv:1105.0388](#) [math.PR]. [MR2959066](#)
- [17] BRÉZIN, E. and HIKAMI, S. (1997). Extension of level-spacing universality. *Phys. Rev. E* (3) **56** 264–269. Available at [arXiv:cond-mat/9702213](#) [cond-mat.mes-hall].
- [18] BUFETOV, A. and GORIN, V. (2015). Representations of classical Lie groups and quantized free convolution. *Geom. Funct. Anal.* **25** 763–814. Available at [arXiv:1311.5780](#) [math.RT]. [MR3361772](#)
- [19] BUFETOV, A. I. (2012). On the Vershik–Kerov conjecture concerning the Shannon–McMillan–Breiman theorem for the Plancherel family of measures on the space of Young diagrams. *Geom. Funct. Anal.* **22** 938–975. Available at [arXiv:1001.4275](#) [math.RT]. [MR2984121](#)
- [20] COHN, H., KENYON, R. and PROPP, J. (2001). A variational principle for domino tilings. *J. Amer. Math. Soc.* **14** 297–346. Available at [arXiv:math/0008220](#) [math.CO]. [MR1815214](#)
- [21] DUSE, E. (2015). On uniformly random discrete interlacing systems. Ph.D. thesis, KTH Royal Institute of Technology, Stockholm.
- [22] DUSE, E., JOHANSSON, K. and METCALFE, A. (2016). The cusp–Airy process. *Electron. J. Probab.* **21** Paper No. 57, 50. Available at [arXiv:1510.02057](#) [math.PR]. [MR3546394](#)
- [23] DUSE, E. and METCALFE, A. (2015). Asymptotic geometry of discrete interlaced patterns: Part I. *Internat. J. Math.* **26** 1550093, 66. Available at [arXiv:1412.6653](#) [math.PR]. [MR3413988](#)
- [24] DUSE, E. and METCALFE, A. (2015). Asymptotic geometry of discrete interlaced patterns: Part II. Preprint. Available at [arXiv:1507.00467](#) [math-ph].
- [25] DYSON, F. J. (1962). A Brownian-motion model for the eigenvalues of a random matrix. *J. Math. Phys.* **3** 1191–1198. [MR0148397](#)
- [26] EDREI, A. (1953). On the generation function of a doubly infinite, totally positive sequence. *Trans. Amer. Math. Soc.* **74** 367–383. [MR0053989](#)
- [27] EICHELSBACHER, P. and KÖNIG, W. (2008). Ordered random walks. *Electron. J. Probab.* **13** 1307–1336. Available at [arXiv:math/0610850](#) [math.PR]. [MR2430709](#)
- [28] ERDÉLYI, A., ed. (1953). *Higher Transcendental Functions*. McGraw–Hill, New York.
- [29] ERDŐS, L. and SCHNELLI, K. (2017). Universality for random matrix flows with time-dependent density. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 1606–1656. Available at [arXiv:1504.00650](#) [math.PR]. [MR3729630](#)
- [30] ERDŐS, L. and YAU, H.-T. (2012). Universality of local spectral statistics of random matrices. *Bull. Amer. Math. Soc. (N.S.)* **49** 377–414. Available at [arXiv:1106.4986](#) [math.PR]. [MR2917064](#)
- [31] GORIN, V. (2017). Bulk universality for random lozenge tilings near straight boundaries and for tensor products. *Comm. Math. Phys.* **354** 317–344. Available at [arXiv:1603.02707](#) [math.PR]. [MR3656520](#)
- [32] GORIN, V. and PANOVA, G. (2015). Asymptotics of symmetric polynomials with applications to statistical mechanics and representation theory. *Ann. Probab.* **43** 3052–3132. Available at [arXiv:1301.0634](#) [math.RT]. [MR3433577](#)
- [33] GORIN, V. E. (2008). Nonintersecting paths and the Hahn orthogonal polynomial ensemble. *Funktional. Anal. i Prilozhen.* **42** 23–44, 96. Available at [arXiv:0708.2349](#) [math.PR]. [MR2454474](#)
- [34] HOUGH, J. B., KRISHNAPUR, M., PERES, Y. and VIRÁG, B. (2006). Determinantal processes and independence. *Probab. Surv.* **3** 206–229. Available at [arXiv:math/0503110](#) [math.PR]. [MR2216966](#)
- [35] JOHANSSON, K. (2000). Shape fluctuations and random matrices. *Comm. Math. Phys.* **209** 437–476. Available at [arXiv:math/9903134](#) [math.CO]. [MR1737991](#)

- [36] JOHANSSON, K. (2001). Discrete orthogonal polynomial ensembles and the Plancherel measure. *Ann. of Math. (2)* **153** 259–296. Available at [arXiv:math/9906120](#) [math.CO]. [MR1826414](#)
- [37] JOHANSSON, K. (2001). Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. *Comm. Math. Phys.* **215** 683–705. Available at [arXiv:math-ph/0006020](#). [MR1810949](#)
- [38] JOHANSSON, K. (2005). The arctic circle boundary and the Airy process. *Ann. Probab.* **33** 1–30. Available at [arXiv:math/0306216](#) [math.PR]. [MR2118857](#)
- [39] JOHANSSON, K. (2005). Non-intersecting, simple, symmetric random walks and the extended Hahn kernel. *Ann. Inst. Fourier (Grenoble)* **55** 2129–2145. Available at [arXiv:math/0409013](#) [math.PR]. [MR2187949](#)
- [40] KARLIN, S. and MCGREGOR, J. (1959). Coincidence probabilities. *Pacific J. Math.* **9** 1141–1164. [MR0114248](#)
- [41] KENYON, R. (2008). Height fluctuations in the honeycomb dimer model. *Comm. Math. Phys.* **281** 675–709. Available at [arXiv:math-ph/0405052](#). [MR2415464](#)
- [42] KENYON, R. and OKOUNKOV, A. (2007). Limit shapes and the complex Burgers equation. *Acta Math.* **199** 263–302. Available at [arXiv:math-ph/0507007](#). [MR2358053](#)
- [43] KENYON, R., OKOUNKOV, A. and SHEFFIELD, S. (2006). Dimers and amoebae. *Ann. of Math. (2)* **163** 1019–1056. Available at [arXiv:math-ph/0311005](#). [MR2215138](#)
- [44] KOEKOEK, R. and SWARTTOUW, R. F. (1996). The Askey-scheme of hypergeometric orthogonal polynomials and its q -analogue. Technical report, Delft Univ. Technology and Free Univ. Amsterdam.
- [45] KÖNIG, W. (2005). Orthogonal polynomial ensembles in probability theory. *Probab. Surv.* **2** 385–447. Available at [arXiv:math/0403090](#) [math.PR]. [MR2203677](#)
- [46] KÖNIG, W., O’CONNELL, N. and ROCH, S. (2002). Non-colliding random walks, tandem queues, and discrete orthogonal polynomial ensembles. *Electron. J. Probab.* **7** no. 5, 24. [MR1887625](#)
- [47] LANDON, B. and YAU, H.-T. (2017). Convergence of local statistics of Dyson Brownian motion. *Comm. Math. Phys.* **355** 949–1000. Available at [arXiv:1504.03605](#) [math.PR]. [MR3687212](#)
- [48] MEHTA, M. L. (2004). *Random Matrices*, 3rd ed. *Pure and Applied Mathematics (Amsterdam)* **142**. Elsevier/Academic Press, Amsterdam. [MR2129906](#)
- [49] METCALFE, A. P. (2013). Universality properties of Gelfand–Tsetlin patterns. *Probab. Theory Related Fields* **155** 303–346. Available at [arXiv:1105.1272](#) [math.PR]. [MR3010400](#)
- [50] NAGAO, T. and FORRESTER, P. J. (1998). Multilevel dynamical correlation functions for Dyson’s Brownian motion model of random matrices. *Phys. Lett. A* **247** 42–46.
- [51] NORDENSTAM, E. (2010). On the shuffling algorithm for domino tilings. *Electron. J. Probab.* **15** 75–95. Available at [arXiv:0802.2592](#) [math.PR]. [MR2578383](#)
- [52] NOVAK, J. (2015). Lozenge tilings and Hurwitz numbers. *J. Stat. Phys.* **161** 509–517. Available at [arXiv:1407.7578](#) [math-ph]. [MR3401028](#)
- [53] OKOUNKOV, A. (2002). Symmetric functions and random partitions. In *Symmetric Functions 2001: Surveys of Developments and Perspectives*. *NATO Sci. Ser. II Math. Phys. Chem.* **74** 223–252. Kluwer Academic, Dordrecht. Available at [arXiv:math/0309074](#) [math.CO]. [MR2059364](#)
- [54] OKOUNKOV, A. and OLSHANSKI, G. (1998). Asymptotics of Jack polynomials as the number of variables goes to infinity. *Int. Math. Res. Not.* **1998** 641–682. Available at [arXiv:q-alg/9709011](#). [MR1636541](#)
- [55] OKOUNKOV, A. and RESHETIKHIN, N. (2003). Correlation function of Schur process with application to local geometry of a random 3-dimensional Young diagram. *J. Amer. Math. Soc.* **16** 581–603. Available at [arXiv:math/0107056](#) [math.CO]. [MR1969205](#)

- [56] PETROV, L. (2014). Asymptotics of random lozenge tilings via Gelfand–Tsetlin schemes. *Probab. Theory Related Fields* **160** 429–487. Available at [arXiv:1202.3901](https://arxiv.org/abs/1202.3901) [math.PR]. [MR3278913](https://doi.org/10.1007/s00220-014-1789-1)
- [57] PETROV, L. (2014). The boundary of the Gelfand–Tsetlin graph: New proof of Borodin–Olshanski’s formula, and its q -analogue. *Mosc. Math. J.* **14** 121–160. Available at [arXiv:1208.3443](https://arxiv.org/abs/1208.3443) [math.CO]. [MR3221949](https://doi.org/10.1007/s12220-014-1789-1)
- [58] PETROV, L. (2015). Asymptotics of uniformly random lozenge tilings of polygons. Gaussian free field. *Ann. Probab.* **43** 1–43. Available at [arXiv:1206.5123](https://arxiv.org/abs/1206.5123) [math.PR]. [MR3298467](https://doi.org/10.1007/s12220-014-1789-1)
- [59] SHCHERBINA, T. (2009). On universality of bulk local regime of the deformed Gaussian unitary ensemble. *Zh. Mat. Fiz. Anal. Geom.* **5** 396–433, 440. Available at [arXiv:0804.2116](https://arxiv.org/abs/0804.2116) [math-ph]. [MR2590774](https://doi.org/10.1007/s12220-014-1789-1)
- [60] SHEFFIELD, S. (2005). Random surfaces. *Astérisque* **304** vi+175. Available at [arXiv:math/0304049](https://arxiv.org/abs/math/0304049) [math.PR]. [MR2251117](https://doi.org/10.1007/s12220-014-1789-1)
- [61] SHIRYAEV, A. N. (1995). *Probability* Springer, New York.
- [62] SOSHIKOV, A. (2000). Determinantal random point fields. *Russian Math. Surveys* **55** 923–975. Available at [arXiv:math/0002099](https://arxiv.org/abs/math/0002099) [math.PR].
- [63] TAO, T. and VU, V. (2014). Random matrices: The universality phenomenon for Wigner ensembles. In *Modern Aspects of Random Matrix Theory. Proc. Sympos. Appl. Math.* **72** 121–172. Amer. Math. Soc., Providence, RI. Available at [arXiv:1202.0068](https://arxiv.org/abs/1202.0068) [math.PR]. [MR3288230](https://doi.org/10.1007/s12220-014-1789-1)
- [64] VERSHIK, A. M. and KEROV, S. V. (1982). Characters and factor-representations of the infinite unitary group. *Dokl. Akad. Nauk SSSR* **267** 272–276. [MR0681202](https://doi.org/10.1007/s12220-014-1789-1)
- [65] VOICULESCU, D. (1976). Représentations factorielles de type II_1 de $U(\infty)$. *J. Math. Pures Appl.* (9) **55** 1–20. [MR0442153](https://doi.org/10.1007/s12220-014-1789-1)
- [66] WEYL, H. (1997). *The Classical Groups: Their Invariants and Representations. Princeton Landmarks in Mathematics.* Princeton Univ. Press, Princeton, NJ. [MR1488158](https://doi.org/10.1007/s12220-014-1789-1)

DEPARTMENT OF MATHEMATICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS 02139
USA
AND
INSTITUTE FOR INFORMATION
TRANSMISSION PROBLEMS
MOSCOW 127051
RUSSIA
E-MAIL: vadicgor@gmail.com

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VIRGINIA 22904
USA
AND
INSTITUTE FOR INFORMATION
TRANSMISSION PROBLEMS
MOSCOW 127051
RUSSIA
E-MAIL: lenia.petrov@gmail.com