Random Matrix Theory in a nutshell

Part I: Determinantal Point Processes

Manuela Girotti

based on M. Girotti's PhD thesis and
A. Kuijlaars' notes from Les Houches Winter School 2012

Contents

In these notes we review the main concepts about Determinantal Point Processes.

Determinantal point processes are of considerable current interest in Probability theory and Mathematical Physics. They were first introduced by Macchi ([9]) and they arise naturally in Random Matrix theory, non-intersecting paths, certain combinatorial and stochastic growth models and representation theory of large groups.

Classical references about determinantal processes are the papers by Hough *et al.* [2], Johansson [4], König [6] and Soshnikov [10].

1 Point Processes

Consider a random collection of points on the real line. A configuration $\mathcal{X}$ is a subset of $\mathbb{R}$ that locally contains a finite number of points, i.e. $\#(\mathcal{X} \cap [a,b]) < +\infty$ for every bounded interval $[a,b] \subset \mathbb{R}$.

1

Definition 1. A (locally finite) **point process** $\mathbb{P}$ on $\mathbb{R}$ is a probability measure on the space of all configurations of points $\{\mathcal{X}\}$.

Definition 2. $\mathbb{P}$ is an n-point process if $\mathbb{P}(\#\mathcal{X} = n) = 1$.

Loosely speaking, given a point process on $\mathbb{R}$, it is possible to evaluate the probability of any given configuration.

Example. If $P(x_1, \ldots, x_n)$ is a probability function on $\mathbb{R}^n$ (with respect to the Lebesgue measure) such that it is invariant under permutations

$$P(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = P(x_1, \ldots, x_n) \qquad \forall \sigma \in S_n, \tag{1}$$

then P defines naturally a point process.

The mapping

$$A \mapsto \mathbb{E}[\#(\mathcal{X} \cap A)], \tag{2}$$

which assigns to a Borel set A the expected value of the number of points in A under the configuration $\mathcal{X}$, is a measure on $\mathbb{R}$. Assuming there exists a density ρ_1 with respect to the Lebesgue measure (1-**point correlation function**), we have

$$\mathbb{E}[\#(\mathcal{X} \cap A)] = \int_A \rho_1(x) \mathrm{d}x \tag{3}$$

and $\rho_1(x)dx$ represents the probability to have a point in the infinitesimal interval $[x, x + \mathrm{d}x]$.

More generally, a k-**point correlation function** ρ_k (if it exists) is a function of k variables such that for distinct points

$$\rho_k(x_1, \ldots, x_k) \mathrm{d}x_1 \ldots \mathrm{d}x_k \tag{4}$$

is the probability to have a point in each infinitesimal interval $[x_j, x_j + \mathrm{d}x_j]$, $j = 1, \ldots, k$. Thus, given disjoints sets $A_1, \ldots, A_k$, we have

$$\mathbb{E}\left[\prod_{j=1}^k \#(\mathcal{X} \cap A_j)\right] = \int_{A_1} \ldots \int_{A_k} \rho_k(x_1, \ldots, x_k) \mathrm{d}x_1 \ldots \mathrm{d}x_k \tag{5}$$

i.e. the expected number of k-tuples $(x_1, \ldots, x_k) \in \mathcal{X}^k$ such that $x_j \in A_j$ for every j. In case the A_j's are not disjoint it is still possible to define the quantity above, with little modifications. For example, if $A_j = A$ for every j, then

$$\mathbb{E}\left[\left(\binom{\#\mathcal{X} \cap A}{k}\right)\right] = \frac{1}{k!} \int_A \ldots \int_A \rho_k(x_1, \ldots, x_k) \mathrm{d}x_1 \ldots \mathrm{d}x_k \tag{6}$$

is the expected number of ordered k-tuples $(x_1, \ldots, x_k)$ such that $x_1 < x_2 < \ldots < x_k$ and $x_j \in A$ for every $j = 1, \ldots, k$.

Example (cont'd). If $P(x_1, \ldots, x_n)$ is a probability density function on $\mathbb{R}^n$, invariant under permutations of coordinates, then we can build an n-point process on $\mathbb{R}$ with correlation functions

$$\rho_k(x_1, \ldots, x_k) := \frac{n!}{(n-k)!} \int_{\mathbb{R}^{n-k}} P(x_1, \ldots, x_n) \mathrm{d}x_{k+1} \ldots \mathrm{d}x_n. \tag{7}$$

The problem of existence and uniqueness of a random point field defined by its correlation functions was studied by Lenard in [7] and [8].

2 Determinantal Point Processes

Definition 3. A point process with correlation functions ρ_k is **determinantal** if there exists a kernel $K(x, y)$ such that for every k and every $x_1, \ldots, x_k$ we have

$$\rho_k(x_1, \ldots, x_k) = \det[K(x_i, x_j)]_{i,j=1}^{k}. \tag{8}$$

The kernel K is called correlation kernel of the determinantal point process.

Remark 4. *The correlation kernel is not unique. If K is a correlation kernel, then the conjugation of K with any positive function $h(\cdot)$ gives an equivalent correlation kernel*

$$\widetilde{K}(x, y) := h(x)K(x, y)h(y)^{-1} \tag{9}$$

describing the same point process.

Determinantal processes became quite common as a model describing (random) points that tend to exclude one another. Indeed, it is easy to see that in a determinantal process there is a repulsion between nearby points and from this feature comes the denomination of DPP as Fermionic point process (such name is mostly found in physics literature).

Examples of determinantal processes can be constructed thanks to the following result. We refer to [10] for a thorough exposition.

Theorem 5. *Consider a kernel K with the following properties:*

- *trace-class:* $\operatorname{Tr} K = \int_{\mathbb{R}} K(x, x)\mathrm{d}x = n < +\infty;$

- *positivity:* $\det[K(x_i, x_j)]_{i,j=1}^{n}$ *is non-negative for every $x_1, \ldots, x_n \in \mathbb{R}$;*

- *reproducing property:* $\forall \; x, y \in \mathbb{R}$

$$K(x, y) = \int_{\mathbb{R}} K(x, s)K(s, y)\mathrm{d}s. \tag{10}$$

Then,

$$P(x_1, \ldots, x_n) := \frac{1}{n!} \det[K(x_i, x_j)]_{i,j=1}^{n} \tag{11}$$

is a probability measure on $\mathbb{R}^n$, invariant under coordinates permutations. The associated point process is a determinantal point process with K as correlation kernel.

3 Gap probability

In a determinantal process all information is contained in the correlation kernel and all quantities of interest can be expressed in terms of K. In particular, given a Borel set A, we are interested in the so called gap probability, i.e. the probability to find no points in A.

Consider a point process on $\mathbb{R}$ with correlation function ρ_k and let A be a Borel set such that, with probability 1, there are only finitely may points in A (for example, A is bounded). Denote by $p_A(n)$ the probability that there are exactly n points in A. If there are n points in A, then the number of ordered k-tuples in A is $\binom{n}{k}$.

Also recall that

$$\mathbb{E}\left[\left(\binom{\#\mathcal{X}\cap A}{k}\right)\right] = \frac{1}{k!}\int_A \cdots \int_A \rho_k(x_1,\ldots,x_k)\mathrm{d}x_1\ldots\mathrm{d}x_k \tag{12}$$

is the expected number of ordered k-tuples $(x_1,\ldots,x_k)$ such that $x_1 < x_2 < \ldots < x_k$ and $x_j \in A$ for every $j = 1,\ldots,k$.

Therefore, the following equality holds

$$\frac{1}{k!}\int_{A^k}\rho_k(x_1,\ldots,x_k)\mathrm{d}x_1\ldots\mathrm{d}x_k = \sum_{n=k}^{\infty}\binom{n}{k}p_A(n). \tag{13}$$

Assume the following alternating series is absolutely convergent, then

$$\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}\int_{A^k}\rho_k(x_1,\ldots,x_k)\mathrm{d}x_1\ldots\mathrm{d}x_k = \sum_{k=0}^{\infty}\sum_{n=k}^{\infty}(-1)^k\binom{n}{k}p_A(n)$$

$$= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}(-1)^k\binom{n}{k}\right)p_A(n). \tag{14}$$

On the other hand,

$$\sum_{k=0}^{\infty}(-1)^k\binom{n}{k} = \begin{cases}1 & \text{if } n = 0 \\ 0 & \text{if } n \neq 0\end{cases} \tag{15}$$

therefore,

$$p_A(0) = \sum_{k=0}^{\infty}\frac{(-1)^k}{k!}\int_{A^k}\rho_k(x_1,\ldots,x_k)\mathrm{d}x_1\ldots\mathrm{d}x_k, \tag{16}$$

where we call $p_A(0)$ **gap probability**, i.e. the probability to find no points in A. In particular, when a point process is determinantal, we have

$$p_A(0) = \sum_{k=0}^{\infty}\frac{(-1)^k}{k!}\int_{A^k}\det\left[K(x_i,x_j)\right]_{i,j=1}^k \mathrm{d}x_1\ldots\mathrm{d}x_k, \tag{17}$$

which is clearly the Fredholm determinant

$$\det\left(\mathrm{Id} - \mathbf{K}\Big|_A\right) \tag{18}$$

of the (trace class) integral operator

$$\mathbf{K}\Big|_A : L^2(A) \to L^2(A)$$

$$f(x) \mapsto \mathbf{K}[f](x) = \int_{\mathbb{R}}K(x,y)f(y)\mathrm{d}y. \tag{19}$$

It is actually possible to prove a more general result, which reduces to the one above when considering zero particles.

Theorem 6 (Theorem 2, [10]). *Consider a determinantal point process with kernel K. For any finite Borel sets $A_j \subseteq \mathbb{R}$, $j = 1, \ldots, n$, the generating function of the probability distribution of the occupation number $\#_{A_j} := \#\{x_i \in A_j\}$ is given by*

$$G(z_1, \ldots, z_n) = \mathbb{E}\left(\prod_{i=1}^{n} z_j^{\#_{A_j}}\right) = \det\left(\mathrm{Id} + \sum_{j=1}^{n}(z_j - 1)K\Big|_{A_j}\right). \tag{20}$$

We refer to [10] for a detailed proof of the Theorem.

Let's break down the formula for some cases of interest (we're using standard results from Probability Theory here), we have that the gap probability (probability of having no points) over a single Borel set $A \subseteq \mathbb{R}$ is

$$\mathbb{P}(\mathcal{X} \cap A = 0) = G(0) = \det\left(\mathrm{Id} + (z-1)K\Big|_A\right)_{z=0} = \det\left(\mathrm{Id} - K\Big|_A\right); \tag{21}$$

the probability of having just one point in the set A is

$$\mathbb{P}(\mathcal{X} \cap A = 1) = \frac{G'(0)}{1!} = \frac{\mathrm{d}}{\mathrm{d}z}\det\left(\mathrm{Id} + (z-1)K\Big|_A\right)_{z=0} \tag{22}$$

and more generally the probability of having k points in A is

$$\mathbb{P}(\mathcal{X} \cap A = k) = \frac{G^{(k)}(0)}{k!} = \frac{1}{k!}\frac{\mathrm{d}^k}{\mathrm{d}z^k}\det\left(\mathrm{Id} + (z-1)K\Big|_A\right)_{z=0}. \tag{23}$$

When considering multiple sets $A_1, \ldots, A_n$, the argument is similar

$$\mathbb{P}\left(\mathcal{X} \cap A_1 = k_1, \mathcal{X} \cap A_2 = k_2, \ldots, \mathcal{X} \cap A_n = k_n\right) = \frac{1}{k_1!}\cdots\frac{1}{k_n!}\frac{\partial^{\sum k_i}}{\partial z_1^{k_1}\ldots\partial z_n^{k_n}}G(z_1, \ldots, z_n)$$

$$= \frac{1}{k_1!}\cdots\frac{1}{k_n!}\frac{\partial^{\sum k_i}}{\partial z_1^{k_1}\ldots\partial z_n^{k_n}}\det\left(\mathrm{Id} + \sum_{j=1}^{n}(z_j - 1)K\Big|_{A_j}\right)_{z_1=\ldots=z_n=0}. \tag{24}$$

4 Examples of Determinantal Point Processes

We will see here a couple of physical/mathematical models that can be effectively described and studied using DPPs. The purpose of this section is simply to give an idea of how ubiquitous DPPs can be.

4.1 Fermi gas

Consider the Schrödinger operator $H = -\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)$ with V a real-valued function, acting on the space $L^2(E)$, E is a separable Hausdorff space (for the sake of simplicity, E will be $\mathbb{R}$ or S^1), and let $\{\varphi_k\}_{k=0}^{\infty}$ be a set of orthonormal eigenfunctions for the operator H. The nth exterior power of H is an operator $\bigwedge^n H := \sum_{i=1}^{n}\left[-\frac{\mathrm{d}^2}{\mathrm{d}x_i^2} + V(x_i)\right]$ acting on $\bigwedge^n L^2(E)$ (the space of antisymmetric

L^2-functions of n variables) and it describes the Fermi gas with n particle, i.e. an ensemble of n fermions.

The ground state of the Fermi gas is given by the so called Slater determinant

$$\psi(x_1, \ldots, x_n) = \frac{1}{\sqrt{n!}} \sum_{\sigma \in S_n} (-1)^\sigma \prod_{i=1}^{n} \varphi_{i-1}(x_{\sigma(i)}) = \frac{1}{\sqrt{n!}} \det \left[\varphi_{i-1}(x_j)\right]_{i,j=1}^{n}. \tag{25}$$

It is known that the absolute value squared of the ground state defines the probability distribution of the particles. Therfore, we have

$$p_n(x_1, \ldots, x_n) = |\psi(x_1, \ldots, x_n)|^2 = \frac{1}{n!} \det \left[K_n(x_i, x_j)\right]_{i,j=1}^{n} \tag{26}$$

$$K_n(x, y) := \sum_{i=0}^{n-1} \varphi_{i-1}(x) \overline{\varphi_{i-1}(y)} \tag{27}$$

and $K_n(x, y)$ is the kernel of the orthogonal projection onto the subspace spanned by the first n eigenfunctions $\{\varphi_j\}$ of H.

The formula above defines a determinantal process with correlation functions

$$\rho_k^{(n)}(x_1, \ldots, x_k) = \frac{n!}{(n-k)!} \int p_n(x_1, \ldots, x_n) \mathrm{d}x_{k+1} \ldots \mathrm{d}x_n = \det \left[K_n(x_i, x_j)\right]_{i,j=1}^{k}. \tag{28}$$

To give some practical examples, let's focus on two special cases of H. The first case is the harmonic oscillator on the real line $\mathbb{R}$

$$H = -\frac{\mathrm{d}^2}{\mathrm{d}x^2} + x^2; \tag{29}$$

its eigenfunctions are the Weber-Hermite functions

$$\varphi_k(x) = \frac{(-1)^k}{\pi^{\frac{1}{4}}} (2^k k)^{\frac{1}{2}} e^{\frac{x^2}{2}} \frac{\mathrm{d}^k}{\mathrm{d}x^k} e^{-x^2} \tag{30}$$

and the correlation kernel is (using the Christoffel-Darboux formula)

$$K_n(x, y) = \left(\frac{n}{2}\right)^{\frac{1}{2}} \frac{\varphi_n(x)\varphi_{n-1}(y) - \varphi_n(y)\varphi_{n-1}(x)}{x - y}. \tag{31}$$

The second case is the free particle on a circle S^1

$$H = -\frac{\mathrm{d}^2}{\mathrm{d}\theta^2} \tag{32}$$

and its correlation kernel is

$$K_n(\theta, \eta) = \frac{\sin\left(\frac{n}{2}(\theta - \eta)\right)}{2\pi \sin\left(\frac{\theta - \eta}{2}\right)}. \tag{33}$$
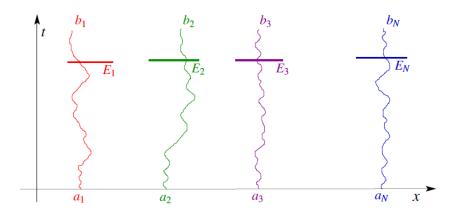
Figure 1: Non-intersecting Brownian paths with fixed starting points $\{a_j\}$.

4.2 Dyson processes

Let $p_t(x; y)$ be the transition probability density from point x to point y at time t of a one-dimensional strong Markov process with continuous sample paths. A classical theorem by S. Karlin and J. McGregor [5] gives a determinantal formula for the probability that a number of paths with given starting and ending positions fall in certain sets at some later time without intersecting in the intermediate time interval (see Figure 1).

Theorem 7 (Karlin, McGregor). *Consider n independent copies $X_1(t), \ldots, X_n(t)$ of a one-dimensional strong Markov process with continuous sample paths, conditioned so that*

$$X_j(0) = a_j \tag{34}$$

for given values $a_1 < a_2 < \ldots < a_n \in \mathbb{R}$. Let $p_t(x, y)$ be the transition probability function of the Markov process and let $E_1, \ldots, E_n \subseteq \mathbb{R}$ be disjoint Borel sets (more precisely, we assume $\sup E_j < \inf E_{j+1}$). Then,

$$\frac{1}{Z_n} \int_{E_1} \cdots \int_{E_n} \det \left[p_t(a_i, x_j) \right]_{i,j=1}^n \, \mathrm{d}x_1 \ldots \mathrm{d}x_n \tag{35}$$

is equal to the probability that each path X_j belongs to the set E_j at time t, without any intersection between paths in the time interval $[0, t]$, for some normalizing constant Z_n.

Sketch of the proof (a heuristic argument). Let $n = 2$, then

$$\frac{1}{Z_2} \int_{E_1} \int_{E_2} p_t(a_1, x_1) p_t(a_2, x_2) - p_t(a_1, x_2) p_t(a_2, x_1) \, \mathrm{d}x_1 \mathrm{d}x_2$$
$$= P\left(X_1(t) \in E_1\right) P\left(X_2(t) \in E_2\right) - P\left(X_1(t) \in E_2\right) P\left(X_2(t) \in E_1\right)$$
$$=: P\left(\mathcal{A}\right) - P\left(\mathcal{B}\right). \tag{36}$$

On the other hand,

$$P\left(\mathcal{A}\right) - P\left(\mathcal{B}\right) = P\left(\mathcal{A}_1\right) + P\left(\mathcal{A}_2\right) - P\left(\mathcal{B}_1\right) - P\left(\mathcal{B}_2\right) \tag{37}$$

where $\mathcal{A}_i$, $\mathcal{B}_i$ represent the following events:

$$\mathcal{A}_1 = \{ X_i(t) \in E_i \text{ respectively and the paths did not intersect} \}$$
$$\mathcal{A}_2 = \{ X_i(t) \in E_i \text{ respectively and the paths did intersect at least once} \}$$
$$\mathcal{B}_1 = \{ X_1(t) \in E_2, X_2(t) \in E_1 \text{ and the paths did not intersect} \}$$
$$\mathcal{B}_2 = \{ X_1(t) \in E_2, X_2(t) \in E_1 \text{ and the paths did intersect at least once} \}$$

Clearly $P(\mathcal{B}_1) = 0$. Moreover, consider the event $\mathcal{A}_2$: at the first time when the two path collide, we can interchange the labels. This is a bijection $\Psi : \mathcal{A}_2 \xrightarrow{\sim} \mathcal{B}_2$. Since the process is Markovian and the two particles act independently, we have

$$P(\mathcal{A}_2) = P(\mathcal{B}_2). \tag{38}$$

In conclusion,

$$\frac{1}{Z_2} \int_{E_1} \int_{E_2} \det\left[p_t(a_i, x_j)\right]_{i,j=1,2} \, \mathrm{d}x_1 \mathrm{d}x_2 = P(\mathcal{A}_1). \tag{39}$$

$\square$

However, this is not a determinantal process, since the correlation functions are not expressible in terms of the determinant of a kernel. If we additionally condition the paths to end at time $T > 0$ at some given points $b_1 < \ldots < b_n$, without any intersection between the paths along the whole time interval $[0, T]$, then it can be shown (using an argument again based on the Markov property) that the random positions of the n paths at a given time $t \in [0, T]$ have the joint probability density function

$$\frac{1}{Z_n} \det\left[p_t(a_i, x_j)\right]_{i,j=1}^n \det\left[p_{T-t}(x_i, b_j)\right]_{i,j=1}^n = \frac{1}{Z_n} \det\left[K_n(x_i, x_j)\right]_{i,j=1}^n \tag{40}$$

with a suitable normalizing constant Z_n and kernel

$$K_n(x, y) := \sum_{j=1}^n \phi_j(x)\psi_j(y) \tag{41}$$

$$\phi_j \in \langle p_t(a_1, x), \ldots, p_t(a_n, x) \rangle, \quad \psi_k \in \langle p_{T-t}(x, b_1), \ldots, p_{T-t}(x, b_n) \rangle, \tag{42}$$

$$\int_{\mathbb{R}} \phi_j(x)\psi_k(x)\mathrm{d}x = \delta_{jk}. \tag{43}$$

Remark 8. *The model we just constructed is known in the literature as biorthogonal ensemble. We refer to [1] for a thorough exposition on the subject.*

Of interest is also the confluent case when two or more starting (or ending) points collapse together. For example, in the confluent limit as $a_j \to a$ and $b_j \to b$, for all j's (see Figure 2), applying l'Hôpital rule to (40) gives

$$\frac{1}{\widetilde{Z}_n} \det\left[\frac{\mathrm{d}^{i-1}}{\mathrm{d}a^{i-1}} p_t(a, x_j)\right]_{i,j=1}^n \det\left[\frac{\mathrm{d}^{j-1}}{\mathrm{d}b^{j-1}} p_{T-t}(x_i, b)\right]_{i,j=1}^n \tag{44}$$

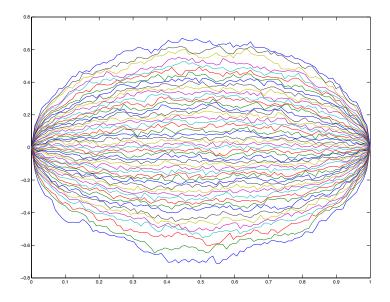which is still a determinantal point process with kernel derived along the same method as in (41)-(43).

Figure 2: Numerical simulation of 50 non-intersecting Brownian paths in the confluent case with one starting and one ending point.

4.3 Domino tilings: the Aztec diamond

An Aztec diamond A_n of size n is the union of all lattice squares $[m, m+1] \times [\ell, \ell+1]$, $m, \ell \in \mathbb{Z}$, that lie inside the region $\{(x,y) | |x| + |y| \le n+1\}$. We are using here the standard coordinate system in the cartesian plane $\mathbb{R}^2$ with the origin being at $(0,0)$ and the axes being $\vec{x} = (1,0)$ and $\vec{y} = (0,1)$.

A domino tile and a domino tiling of a region are what we would expect them to be: a domino is a closed rectangle in $\mathbb{R}^2$ of size 1×2 with corners in $\mathbb{Z}^2$, and a tiling of A_n by dominos is a set of dominos whose interiors are disjoint and whose union is A_n.
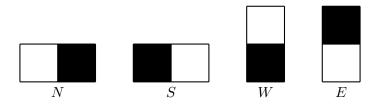

Figure 3: Four types of domino tiles.

We can colour the unit squares in A_n in a checkerboard fashion so that the leftmost square in each row in the top half is white. Depending on how a domino covers the squares we can distinguish four types of dominos (see Figure 3):

1. a horizontal domino is an N-domino if its leftmost square is white and the rightmost is black;

2. a horizontal domino with those colours inverted in an S-domino;

9

3. a vertical domino is a W-domino if its upper square is white and the bottom square is black

4. a vertical domino is an E-domino if the colours are switched.

We can also distinguish the four tiles by simply colouring them as red, yellow, blue and green. Two dominos are adjacent if they share an edge of a square, and a domino is adjacent to the boundary if it shares an edge with the boundary of the diamond.

We introduce now some "randomness" in the model. Let T be a tiling of the Aztec diamond A_n and let $v(T)$ be the number of vertical dominos in T (W and E dominos). We define the weight of T by letting vertical dominos have weight $w = \alpha > 0$ and horizontal dominos weight $w = 1$, so that the total weight is $\prod_{j \in T} w_j = \alpha^{v(T)}$. We can then define a probability measure on the set of all possible tilings of A_n (denoted as $\mathcal{T}(A_n)$) by simply normalizing this weight (note that when $\alpha = 1$ we pick the tiling uniformly at random).
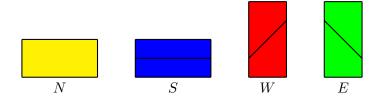

Figure 4: The rules for drawing non-intersecting paths on a domino tiling.

We can establish a one-to-one correspondence between a domino tiling and a family of (discrete) non-intersecting paths in the following way:

1. on an N-domino we draw no path;

2. on a W-domino placed so that it has corners at $(0,0)$ and $(1,2)$ we draw a line from $(0, 1/2)$ to $(1, 3/2)$;

3. on an E-domino in the same position we draw a line from $(0, 3/2)$ to $(1, 1/2)$;

4. on an S-domino, placed so that it has its corners at $(0,0)$ and $(2,1)$, we draw a line from $(0, 1/2)$ to $(2, 1/2)$.

We can see that all the paths generated following these rules form a family of non-intersecting paths from the point $A_r = (-n - 1 + r, -r + 1/2)$ to the point $B_r = (n + 1 - r, -r + 1/2)$ ($\forall\ r = 1, \ldots, n$; see Figure 5)

This model of non-intersecting paths is not yet determinantal. In order to obtain a measure of the form (8), we first need to introduce a new coordinate system: pick the origin to be the point $(-n, -1/2)$ and the axes to be $\vec{e}_1 = (1,1)$ and $\vec{e}_2 = (-1, 1)$.

The coordinate transformation is

$$\begin{cases} x' = x - y - n \\ y' = x + y - \frac{1}{2} \end{cases} \tag{45}$$

The non-intersecting paths now go from $A_j = (0, -j + 1)$ to $B_j = (n + 1 - j, -n)$ ($\forall\ j = 1, \ldots, n$; see Figure 6a) and we have three types of steps: $(1, 0)$, $(0, -1)$ and $(1, -1)$.
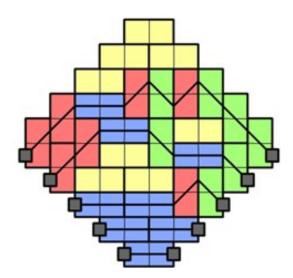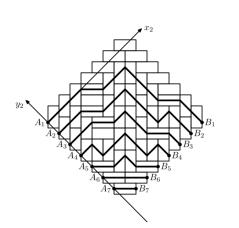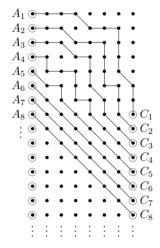
Figure 5: Non-intersecting paths obtained from a domino tiling.

(a) Change of coordinate system.

(b) Extended family of non-intersecting paths.

Figure 6: Family of non-intersecting paths.

The weights on the domino tiling are still $w = \alpha$ for the steps $(1,0)$ and $(0,-1)$ (which corresponded originally to the vertical dominos) and $w = 1$ for the step $(1,-1)$ (the original horizontal domino).

Take $N \geq n$ and extend all the paths from A_j to B_j $(j = 1, \ldots, n)$ to new paths from A_j to points C_j $(j = 1, \ldots, N)$ as shown in Figure 6b. We notice that in doing that, we are only adding steps of type $(1,-1)$, meaning that the total weight for each path doesn't change: in particular, the total weight of a path that goes from A_j to B_j and then to C_j $(j = 1, \ldots, n)$ is still the original weight; the total weight of the new paths from A_ℓ to C_ℓ $(\ell = n+1, \ldots, N)$ is simply 1.

Since this procedure has no effect on the correspondence with domino tilings in the Aztec diamond or the weights, we can just as well consider this extended system of paths.

In order obtain a DPP, we need double the vertical lines and every portion of the path gets doubled according to the rules:

Finally, we shift the paths so that the initial and final points end up at the same height (see Figure 7). The steps from even to odd columns are $(1,0)$ with weight $w = 1$ or $(1,1)$ with weight $w = \alpha$, the steps from odd to even columns are $(1,0)$ with weight $w = 1$, and we also have steps $(0,-1)$ with weight $w = \alpha$ in the even columns. With this choice of weights we still have a weight preserving bijection with the original domino tiling of A_n. The associated particles which we think of as a point process are shown in Figure 7.

The paths just described connecting $(0, j-1)$ to $(2n, j-1)$ $(1 \leq j \leq N)$ can be thought of as being built up from $2n$ transition steps. We have points $x_1^{(r)}, \ldots, x_N^{(r)}$ on line r which connect to points $x_1^{(r+1)}, \ldots, x_N^{(r+1)}$ on line $r+1$. It follows that the transition weight to go from x on the line r to y on the line $r+1$ is

$$\phi_{2i,2i+1}(x,y) = \begin{cases} \alpha & y - x = 1 \\ 1 & y - x = 0 \\ 0 & \text{otherwise} \end{cases} \tag{46}$$

if $r = 2i$ (even), and

$$\phi_{2i-1,2i}(x,y) = \begin{cases} \alpha^{-(y-x)} & y - x \leq 0 \\ 0 & \text{otherwise} \end{cases} \tag{47}$$

if $r = 2i - 1$ (odd).

At this point, a discrete version of the Karlin-McGregor theorem applies:

Theorem 9 (Lindström, Gessel, Viennot). *The weight of all non-intersecting paths from $x^{(r)} \in \mathbb{Z}^n$ on line r to $x^{(r+1)} \in \mathbb{Z}^n$ on line $r+1$ is*

$$\det\left[\phi_{r,r+1}\left(x_i^{(r)}, x_j^{(r+1)}\right)\right]_{i,j=1,\ldots,n} \tag{48}$$
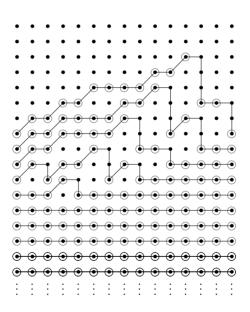
12

Figure 7: Extended family of non-intersecting paths; the circled dots are the particles in the determinantal point process.

For a proof of the theorem, we refer to [4].

The initial and final configurations are fixed $(x_i^{(0)} = x_i^{(2n)} = 1 - i, \, i = 1, \dots, N)$ and the weight of the whole configuration of non-intersecting paths is then the product of all the weights

$$\prod_{r=0}^{2n-1} \det \left[\phi_{r,r+1} \left(x_i^{(r)}, x_j^{(r+1)} \right) \right]_{i,j=1,\dots,n}. \tag{49}$$

Up to a normalization constant, we have a probability measure; even more, it turns out that this probability is determinantal with correlation kernel built from the weights $\phi_{r,r+1}$ and it is called extended Krawtchouk kernel. We do not give here more details about the correlation kernel and we refer to [3] and [4] for a complete description.

As a last remark on this model, we notice that in the limit where n (the dimension of the diamond) goes to infinity, we can observe the four corners (cardinal points) of the diamond to have "frozen region" (given by the corresponding tiles) and a more chaotic behaviour in the interior of the diamond: see Figure 8.

The following remarkable result hold:

Theorem 10 (Artic Circle Theorem; Jockusch, Propp, Shor). *The shape of this central sub-region inside the diamond becomes arbitrarily close to a perfect circle of radius $n/\sqrt{2}$ for all but a negligible proportion of the tilings.*

Lots of fun available at this website: https://sites.uclouvain.be/aztecdiamond/!
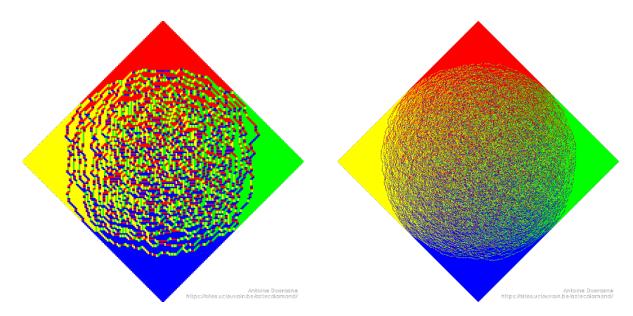
13

Figure 8: Aztec diamond with $n = 100$ (left) $n = 500$ (right).

5 Transformations of Determinantal Point Processes

5.1 Scaling and translations

Let $f \in C^1(\mathbb{R})$ be a bijection, with continuously differentiable inverse $f^{-1} = g \in C^1(\mathbb{R})$. Then, f maps configurations on $\mathbb{R}$ to configurations on $\mathbb{R}$. The push-forward of a point process $\mathbb{P}$ under such mapping is $f(\mathbb{P})$ and it is again a point process.

Proposition 11. *If K is the kernels of a DPP $\mathbb{P}$, then $f(\mathbb{P})$ is also a DPP with kernel*

$$\widehat{K}(x, y) = \sqrt{g'(x)g'(y)}K(g(x), g(y)). \tag{50}$$

In particular, if f is a linear function $f(x) = \alpha\,(x - x^*)$, $\alpha > 0$ and $x^* \in \mathbb{R}$, then the transformed point process $f(\mathbb{P})$ is a scaled and translated version of the original point process $\mathbb{P}$. The correlation kernel K changes to

$$\widehat{K}(x, y) = \frac{1}{\alpha}K\left(x^* + \frac{x}{\alpha}, x^* + \frac{y}{\alpha}\right). \tag{51}$$

5.2 Thinning

Given a point process $\mathbb{P}$, fix $x^* \in R$ and condition on the event that $x^* \in \mathcal{X}$; then, remove this point from the configuration. The resulting process $\mathcal{X}\backslash\{x^*\}$ is a new point process.

Proposition 12. *Let K be the correlation kernel of a DPP and let x^* be a point such that $0 < K(x^*, x^*) < +\infty$. Then, the point process obtained by removing x^* is determinantal with kernel*

$$\widehat{K}(x, y) = K(x, y) - \frac{K(x, x^*)K(x^*, y)}{K(x^*, x^*)}. \tag{52}$$

14

5.3 Limits (and universality)

Suppose that for each n we can construct a (finite) determinantal point process $\mathbb{P}_n$ with correlation kernel K_n. If the sequence of kernels $\{K_n\}$ converge in some sense to a limit kernel K as $n \to \infty$, one can expect that also the point processes $\mathbb{P}_n$ will converge to a new determinantal point process $\mathbb{P}$ with correlation kernel K.

This is indeed the case provided some mild assumptions.

Proposition 13. *Let $\mathbb{P}$ and $\mathbb{P}_n$ be determinantal point processes with kernels K and K_n respectively. Let K_n converge pointwise to K*

$$\lim_{n\to\infty} K_n(x,y) = K(x,y) \tag{53}$$

uniformly in x, y over compact subsets of $\mathbb{R}$. Then, the point processes $\mathbb{P}_n$ converge to $\mathbb{P}$ weakly.

Remark 14. *The condition of uniform convergence on compact sets may be relaxed.*

Suppose we have a sequence of kernels K_n and a fixed reference point x_*. Before taking the limit, we first perform a centering and rescaling of the form

$$x \mapsto Cn^\gamma(x - x_*) \tag{54}$$

with suitable values of C, $\gamma > 0$. Then in many cases of interest the rescaled kernels have a limit

$$\lim_{n\to\infty} \frac{1}{Cn^\gamma} K_n\left(x_* + \frac{x}{Cn^\gamma}, x_* + \frac{y}{Cn^\gamma}\right) = K(x,y) \tag{55}$$

Therefore, the scaling limit K is a kernel that corresponds to a determinantal point process with an infinite number of points.

The physical meaning of this scaling and limiting procedure is the following: as the number of points tends to infinity, one is interested in the local (microscopic) behaviour of the system in specific points of the domain where the particles may lie, upon suitable rescaling: for example, in an infinitesimal neighbourhood entirely contained in the domain (the so-called "bulk") or in an infinitesimal neighbourhood only including the left-most or right-most particles on the line (the so-called "edge").

In many different situations the same scaling limit K may appear. The phenomenon is known as **universality** in Random Matrix Theory (as we will see soon). Instances of limiting kernels are the sine kernel

$$K_{\text{sine}} = \frac{\sin \pi(x-y)}{\pi(x-y)} \tag{56}$$

and the Airy kernel

$$K_{\text{Ai}} = \frac{\text{Ai}(x)\text{Ai}'(y) - \text{Ai}'(x)\text{Ai}(y)}{x-y}, \tag{57}$$

where Ai is the Airy function.

References

[1] A. Borodin. Biorthogonal ensembles. *Nucl. Phys. B*, 536:704–732, 1999.

[2] J.B. Hough, M. Krishnapur, Y. Peres, and B. Virag. Determinantal processes and independence. *Prob. Surveys*, 3:206–229, 2006.

[3] K. Johansson. The artic circle boundary and the Airy process. *Ann. Prob.*, 33(1):1–30, 2005.

[4] K. Johansson. Random matrices and determinantal processes. In *Mathematical Statistical Physics*, volume LXXXIII, chapter 1, pages 1–55. Elsevier, 1st edition, 2006.

[5] S. Karlin and J. McGregor. Coincidence probabilities. *Pacific J. Math.*, 9(4):1141 – 1164, 1959.

[6] W. König. Orthogonal polynomial ensembles in probability theory. *Prob. Surveys*, 2:385–447, 2005.

[7] A. Lenard. Correlation functions and the uniqueness of the state in classical statistical mechanics. *Comm. Math. Phys.*, 30:35–44, 1973.

[8] A. Lenard. States of classical statistical mechanical system of infinitely many particles, I-II. *Arch. Rational Mech. Anal.*, 59:219–256, 1975.

[9] O. Macchi. The coincidence approach to stochastic point processes. *Adv. Appl. Prob.*, 7:83–122, 1975.

[10] A. Soshnikov. Determinantal random point fields. *Russian Mathematical Surveys*, 55:923 – 975, 2000.