Determinantal Probability

Basic Properties and Conjectures

Russell Lyons*

Abstract. We describe the fundamental constructions and properties of determinantal probability measures and point processes, giving streamlined proofs. We illustrate these with some important examples. We pose several general questions and conjectures.

Mathematics Subject Classification (2010). Primary 60K99, 60G55; Secondary 42C30, 37A15, 37A35, 37A50, 68U99.

Keywords. Random matrices, eigenvalues, orthogonal projections, positive contractions, exterior algebra, stochastic domination, negative association, point processes, mixtures, spanning trees, orthogonal polynomials, completeness, Bernoulli processes.

1. Introduction

Determinantal point processes were originally defined by Macchi [39] in physics. Starting in the 1990s, determinantal probability began to flourish as examples appeared in numerous parts of mathematics [51, 28, 8]. Recently, applications to machine learning have appeared [32].

A discrete determinantal probability measure is one whose elementary cylinder probabilities are given by determinants. More specifically, suppose that E is a finite or countable set and that Q is an $E \times E$ matrix. For a subset $A \subseteq E$, let $Q{\restriction}A$ denote the submatrix of Q whose rows and columns are indexed by A. If $\mathfrak{S}$ is a random subset of E with the property that for all finite $A \subseteq E$, we have

$$\mathbf{P}[A \subseteq \mathfrak{S}] = \det(Q{\restriction}A)\,, \tag{1.1}$$

then we call $\mathbf{P}$ a ***determinantal probability measure***. The inclusion-exclusion principle in combination with (1.1) yields the probability of each elementary cylinder event. Therefore, for every Q, there is at most one probability measure, to be denoted $\mathbf{P}^Q$, on subsets of E that satisfies (1.1). Conversely, it is known (see, e.g., [33]) that there is a determinantal probability measure corresponding to Q if Q is the matrix of a positive contraction on $\ell^2(E)$ (in the standard orthonormal basis).

*Partially supported by NSF grant DMS-1007244.

Technicalities are required even to define the corresponding concept of determinantal point process for E being Euclidean space or a more general space. We present a virtually complete development of their basic properties in a way that minimizes such technicalities by adapting the approach of [33] from the discrete case. In addition, we use an idea of Goldman [21] to deduce properties of the general case from corresponding properties in the discrete case.

Space limitations prevent mention of most of what is known in determinantal probability theory, which pertains largely to the analysis of specific examples. We focus instead on some of the basic properties that hold for all determinantal processes and on some intriguing open questions.

2. Discrete Basics

Let E be a denumerable set.

We identify a subset of E with an element of $\{0,1\}^E = 2^E$ in the usual way. There are several approaches to prove the basic existence results and identities for determinantal probability measures. We sketch the one used by [33]. This depends on understanding first the case where Q is the matrix of an orthogonal projection. It also relies on exterior algebra so that the existence becomes immediate.

Any unit vector v in a Hilbert space with orthonormal basis E gives a probability measure $\mathbf{P}^v$ on E, namely, $\mathbf{P}^v\big(\{e\}\big) := |(v, e)|^2$ for $e \in E$. Applying this simple idea to multivectors instead, we obtain the probability measures $\mathbf{P}^H$ associated to orthogonal projections P_H. We refer to [33] for details not given here.

2.1. Exterior Algebra. Identify E with the standard orthonormal basis of the real or complex Hilbert space $\ell^2(E)$. For $k \geq 1$, let E_k denote a collection of ordered k-element subsets of E such that each k-element subset of E appears exactly once in E_k in some ordering. Define

$$\Lambda^k E := \bigwedge^k \ell^2(E) := \ell^2\Big(\big\{e_1 \wedge \cdots \wedge e_k\,;\ \langle e_1, \ldots, e_k \rangle \in E_k\big\}\Big)\,.$$

If $k > |E|$, then $E_k = \emptyset$ and $\Lambda^k E = \{0\}$. We also define $\Lambda^0 E$ to be the scalar field, $\mathbb{R}$ or $\mathbb{C}$. The elements of $\Lambda^k E$ are called ***multivectors*** of ***rank*** k, or k***-vectors*** for short. We then define the ***exterior*** (or ***wedge***) ***product*** of multivectors in the usual alternating multilinear way: $\bigwedge_{i=1}^k e_{\sigma(i)} = (-1)^\sigma \bigwedge_{i=1}^k e_i$ for any permutation $\sigma \in \mathsf{Sym}(k)$, and

$$\bigwedge_{i=1}^k \sum_{e \in E'} a_i(e) e = \sum_{e_1,\ldots,e_k \in E'} \prod_{j=1}^k a_j(e_j) \bigwedge_{i=1}^k e_i$$

for any scalars $a_i(e)$ ($i \in [1,k]$, $e \in E'$) and any finite $E' \subseteq E$. (Thus, $\bigwedge_{i=1}^k e_i = 0$ unless all e_i are distinct.) The inner product on $\Lambda^k E$ satisfies

$$(u_1 \wedge \cdots \wedge u_k,\, v_1 \wedge \cdots \wedge v_k) = \det\big[(u_i,\, v_j)\big]_{i,j \in [1,k]} \tag{2.1}$$

when u_i and v_j are 1-vectors. (This also shows that the inner product on $\Lambda^k E$ does not depend on the choice of orthonormal basis of $\ell^2(E)$.) We then define the ***exterior*** (or ***Grassmann***) ***algebra*** $\mathrm{Ext}\big(\ell^2(E)\big) := \mathrm{Ext}(E) := \bigoplus_{k\geq 0} \Lambda^k E$, where the summands are declared orthogonal, making it into a Hilbert space. Vectors $u_1, \ldots, u_k \in \ell^2(E)$ are linearly independent iff $u_1 \wedge \cdots \wedge u_k \neq 0$. For a k-element subset $A \subseteq E$ with ordering $\langle e_i \rangle$ in E_k, write $\theta_A := \bigwedge_{i=1}^k e_i$. We also write $\bigwedge_{e\in A} f(e) := \bigwedge_{i=1}^k f(e_i)$ for any function $f\colon E \to \ell^2(E)$.

Although there is an isometric isomorphism

$$u_1 \wedge \cdots \wedge u_k \mapsto \frac{1}{\sqrt{k!}} \sum_{\sigma\in\mathsf{Sym}(k)} (-1)^\sigma u_{\sigma(1)} \otimes \cdots \otimes u_{\sigma(k)} \in \ell^2(E^k)$$

for $u_i \in \ell^2(E)$, this does not simplify matters in the discrete case. It will be very useful in the continuous case later, however.

If H is a closed linear subspace of $\ell^2(E)$, written $H \leq \ell^2(E)$, then we identify $\mathrm{Ext}(H)$ with its inclusion in $\mathrm{Ext}(E)$. That is, $\bigwedge^k H$ is the closure of the linear span of the k-vectors $\{v_1 \wedge \cdots \wedge v_k\,;\ v_1, \ldots, v_k \in H\}$. In particular, if $\dim H = r < \infty$, then $\bigwedge^r H$ is a 1-dimensional subspace of $\mathrm{Ext}(E)$; denote by ω_H a unit multivector in this subspace. Note that ω_H is unique up to a scalar factor of modulus 1; which scalar is chosen will not affect the definitions below. We denote by P_H the orthogonal projection onto H for any $H \leq \ell^2(E)$ or, more generally, $H \leq \mathrm{Ext}(E)$.

Lemma 2.1. *For every closed subspace $H \leq \ell^2(E)$, every $k \geq 1$, and every $u_1, \ldots, u_k \in \ell^2(E)$, we have $P_{\mathrm{Ext}(H)}(u_1 \wedge \cdots \wedge u_k) = (P_H u_1) \wedge \cdots \wedge (P_H u_k)$.*

For $v \in \ell^2(E)$, write $[v]$ for the subspace of scalar multiples of v in $\ell^2(E)$.

2.2. Orthogonal Projections.

Let H be a subspace of $\ell^2(E)$ of dimension $r < \infty$. Define the probability measure $\mathbf{P}^H$ on subsets $B \subseteq E$ by

$$\mathbf{P}^H\big(\{B\}\big) := |\,(\omega_H,\ \theta_B)\,|^2\,. \tag{2.2}$$

Note that this is non-0 only for $|B| = r$. Also, by Lemma 2.1,

$$\mathbf{P}^H\big(\{B\}\big) = \|P_{\mathrm{Ext}(H)}\theta_B\|^2 = \|\bigwedge_{e\in B} P_H e\|^2$$

for $|B| = r$, which is non-0 iff $\langle P_H e\,;\ e \in B\rangle$ are linearly independent. That is, $\mathbf{P}^H\big(\{B\}\big) \neq 0$ iff the projections of the elements of B form a basis of H. Let $\langle v_1, \ldots, v_r\rangle$ be any basis of H. If we use (2.1) and the fact that $\omega_H = c\bigwedge_i v_i$ for some scalar c, then we obtain another formula for $\mathbf{P}^H$:

$$\mathbf{P}^H\big(\{e_1, \ldots, e_r\}\big) = (\det[(v_i,\ e_j)]_{i,j\leq r})^2 / \det[(v_i,\ v_j)]_{i,j\leq r}\,. \tag{2.3}$$

We use $\mathfrak{B}$ to denote a random subset of E arising from a probability measure $\mathbf{P}^H$. To see that (1.1) holds for the matrix of P_H, observe that for $|B| = r$,

$$\mathbf{P}^H[\mathfrak{B} = B] = \big(P_{\mathrm{Ext}(H)}\theta_B,\ \theta_B\big) = \Big(\bigwedge_{e\in B} P_H e,\ \bigwedge_{e\in B} e\Big) = \det[(P_H e,\ f)]_{e,f\in B}$$

by (2.1). This shows that (1.1) holds for $|A| = r$ since $|\mathfrak{B}| = r$ $\mathbf{P}^H$-a.s. The general case is a consequence of multilinearity, which gives the following extension of (1.1). We use the convention that $\theta_\emptyset := 1$ and $\mathbf{u} \wedge 1 := \mathbf{u}$ for any multivector $\mathbf{u}$.

Theorem 2.2. *If A_1 and A_2 are (possibly empty) subsets of a finite set E, then*

$$\mathbf{P}^H[A_1 \subseteq \mathfrak{B}, A_2 \cap \mathfrak{B} = \emptyset] = \big(P_{\mathrm{Ext}(H)}\theta_{A_1} \wedge P_{\mathrm{Ext}(H^\perp)}\theta_{A_2},\ \theta_{A_1} \wedge \theta_{A_2}\big)\,. \tag{2.4}$$

In particular, for every $A \subseteq E$, we have

$$\mathbf{P}^H[A \subseteq \mathfrak{B}] = \|P_{\mathrm{Ext}(H)}\theta_A\|^2\,. \tag{2.5}$$

Corollary 2.3. *If E is finite, then for every subspace $H \le \ell^2(E)$, we have*

$$\forall B \subseteq E \qquad \mathbf{P}^{H^\perp}\big(\{E \setminus B\}\big) = \mathbf{P}^H\big(\{B\}\big)\,. \tag{2.6}$$

These extend to infinite E. In order to define $\mathbf{P}^H$ when H is infinite dimensional, we proceed by finite approximation.

Let $E = \{e_i\,;\ i \ge 1\}$ be infinite. Consider first a finite-dimensional subspace H of $\ell^2(E)$. Define H_k as the image of the orthogonal projection of H onto the span of $\{e_i\,;\ 1 \le i \le k\}$. By considering a basis of H, we see that $P_{H_k} \to P_H$ in the weak operator topology (WOT), i.e., matrix-entrywise, as $k \to \infty$. It is also easy to see that if $r := \dim H$, then $\dim H_k = r$ for all large k and, in fact, $\omega_{H_k} \to \omega_H$ in the usual norm topology. It follows that (2.4) holds for this subspace H and for every finite $A_1, A_2 \subset E$.

Now let H be an infinite-dimensional closed subspace of $\ell^2(E)$. Choose finite-dimensional subspaces $H_k \uparrow H$. It is well known that $P_{H_k} \to P_H$ (WOT). Then

$$\text{for all finite sets } A \quad \det(P_{H_k}\restriction A) \to \det(P_H\restriction A)\,, \tag{2.7}$$

whence $\mathbf{P}^{H_k}$ has a weak* limit that we denote $\mathbf{P}^H$ and that satisfies (2.4).

We also note that for *any* sequence of subspaces H_k, if $P_{H_k} \to P_H$ (WOT), then $\mathbf{P}^{H_k} \to \mathbf{P}^H$ weak* because (2.7) then holds.

2.3. Positive Contractions.

We call Q a ***positive contraction*** if Q is a self-adjoint operator on $\ell^2(E)$ such that for all $u \in \ell^2(E)$, we have $0 \le (Qu, u) \le (u, u)$. A ***projection dilation*** of Q is an orthogonal projection P_H onto a closed subspace $H \le \ell^2(E')$ for some $E' \supseteq E$ such that for all $u \in \ell^2(E)$, we have $Qu = P_{\ell^2(E)}P_H u$, where we regard $\ell^2(E')$ as the orthogonal sum $\ell^2(E) \oplus \ell^2(E' \setminus E)$. In this case, Q is also called the ***compression*** of P_H to $\ell^2(E)$. Choose such a dilation (see (2.16) or (3.9)) and define $\mathbf{P}^Q$ as the law of $\mathfrak{B} \cap E$ when $\mathfrak{B}$ has the law $\mathbf{P}^H$. Then (1.1) for Q is a special case of (1.1) for P_H.

Of course, when Q is the orthogonal projection onto a subspace H, then $\mathbf{P}^Q = \mathbf{P}^H$. Basic properties of $\mathbf{P}^Q$ follow from those for orthogonal projections, such as:

Theorem 2.4. *If Q is a positive contraction, then for all finite $A_1, A_2 \subseteq E$,*

$$\mathbf{P}^Q\left[A_1 \subseteq \mathfrak{S}, A_2 \cap \mathfrak{S} = \emptyset\right] = \Big(\bigwedge_{e \in A_1} Qe \wedge \bigwedge_{e \in A_2} (I - Q)e,\ \theta_{A_1} \wedge \theta_{A_2}\Big)\,. \tag{2.8}$$

If (1.1) is given, then (2.8) can be deduced from (1.1) without using our general theory and, in fact, without assuming that the matrix Q is self-adjoint. Indeed, suppose that X is any diagonal matrix. Denote its (e, e)-entry by x_e. Comparing coefficients of x_e shows that (1.1) implies, for finite $A \subseteq E$,

$$\mathbf{E}\Big[\prod_{e\in A}\big(\mathbf{1}_{\{e\in\mathfrak{S}\}} + x_e\big)\Big] = \det\big((Q+X)\restriction A\big)\,. \tag{2.9}$$

Replacing A by $A_1 \cup A_2$ and choosing $x_e := -\mathbf{1}_{A_2}(e)$ gives (2.8). On the other hand, if we substitute $x_e := 1/(z_e - 1)$, then we may rewrite (2.9) as

$$\mathbf{E}\Big[\prod_{e\in A}\big(\mathbf{1}_{\{e\in\mathfrak{S}\}}z_e + \mathbf{1}_{\{e\notin\mathfrak{S}\}}\big)\Big] = \det\big((QZ+I-Q)\restriction A\big)\,, \tag{2.10}$$

where Z is the diagonal matrix of the variables z_e. Let E be finite. Write $z^A := \prod_{e\in A} z_e$ for $A \subseteq E$. Then (2.10) is equivalent to

$$\sum_{A\subseteq E}\mathbf{P}^Q[\mathfrak{S}=A]z^A = \det(I-Q+QZ)\,. \tag{2.11}$$

This is the same as the Laplace transform of $\mathbf{P}^Q$ after a trivial change of variables. When $\|Q\| < 1$, we can write $\det(I-Q+QZ) = \det(I-Q)\det(I+JZ)$ with $J := Q(I-Q)^{-1}$. Thus, for all $A \subseteq E$, we have

$$\mathbf{P}^Q[\mathfrak{S}=A] = \det(I-Q)\det(J\restriction A) = \det(I+J)^{-1}\det(J\restriction A)\,. \tag{2.12}$$

A probability measure $\mathbf{P}$ on 2^E is called ***strongly Rayleigh*** if its generating polynomial $f(z) := \sum_{A\subseteq E}\mathbf{P}[\mathfrak{S}=A]z^A$ satisfies the inequality

$$\frac{\partial f}{\partial z_e}(x)\frac{\partial f}{\partial z_{e'}}(x) \ge \frac{\partial^2 f}{\partial z_e \partial z_{e'}}(x)f(x) \tag{2.13}$$

for all $e \ne e' \in E$ and all real $x \in \mathbb{R}^E$. This property is satisfied by every determinantal probability measure, as was shown by [7], who demonstrated its usefulness in showing other properties, such as negative associations and preservation under symmetric exclusion processes.

For a set $K \subseteq E$, denote by $\mathscr{F}(K)$ the σ-field of events that are measurable with respect to the events $\{e \in \mathfrak{S}\}$ for $e \in K$. Define the ***tail*** σ-field to be the intersection of $\mathscr{F}(E \setminus K)$ over all finite K. We say that a measure $\mathbf{P}$ on 2^E has ***trivial tail*** if every event in the tail σ-field has measure either 0 or 1.

Theorem 2.5 ([33]). *If Q is a positive contraction, then $\mathbf{P}^Q$ has trivial tail.*

For finite E and a positive contraction Q, define the ***entropy*** of $\mathbf{P}^Q$ to be

$$\mathsf{Ent}(Q) := -\sum_{A\subseteq E}\mathbf{P}^Q(\{A\})\log\mathbf{P}^Q(\{A\})\,.$$

Numerical calculation supports the following conjecture [33]:

Conjecture 2.6. *For all positive contractions Q_1 and Q_2, we have*

$$\mathsf{Ent}\big((Q_1+Q_2)/2\big) \ge \big(\mathsf{Ent}(Q_1)+\mathsf{Ent}(Q_2)\big)/2\,. \tag{2.14}$$

2.4. Stochastic Inequalities. Let E be denumerable. A function $f\colon 2^E \to \mathbb{R}$ is called ***increasing*** if for all $A \in 2^E$ and all $e \in E$, we have $f\big(A \cup \{e\}\big) \ge f(A)$. An event is called increasing or ***upwardly closed*** if its indicator is increasing.

Given two probability measures $\mathbf{P}^1$, $\mathbf{P}^2$ on 2^E, we say that $\mathbf{P}^2$ ***stochastically dominates*** $\mathbf{P}^1$ and write $\mathbf{P}^1 \preccurlyeq \mathbf{P}^2$ if for all increasing events $\mathcal{A}$, we have $\mathbf{P}^1(\mathcal{A}) \le \mathbf{P}^2(\mathcal{A})$. This is equivalent to $\int f\, d\mathbf{P}^1 \le \int f\, d\mathbf{P}^2$ for all bounded increasing f.

A ***coupling*** of two probability measures $\mathbf{P}^1$, $\mathbf{P}^2$ on 2^E is a probability measure μ on $2^E \times 2^E$ whose coordinate projections are $\mathbf{P}^1$, $\mathbf{P}^2$; it is ***monotone*** if

$$\mu\big\{(\mathcal{A}_1, \mathcal{A}_2)\,;\ \mathcal{A}_1 \subseteq \mathcal{A}_2\big\} = 1\,.$$

By Strassen's theorem [53], stochastic domination $\mathbf{P}^1 \preccurlyeq \mathbf{P}^2$ is equivalent to the existence of a monotone coupling of $\mathbf{P}^1$ and $\mathbf{P}^2$.

Theorem 2.7 ([33]). *If $H_1 \le H_2 \le \ell^2(E)$, then $\mathbf{P}^{H_1} \preccurlyeq \mathbf{P}^{H_2}$.*

It would be very interesting to find a natural or explicit monotone coupling.

A coupling μ has ***union marginal*** $\mathbf{P}$ if for all events $\mathcal{A} \subseteq 2^E$, we have $\mathbf{P}(\mathcal{A}) = \mu\big\{(A_1, A_2)\,;\ A_1 \cup A_2 \in \mathcal{A}\big\}$.

Question 2.8 ([33]). *Given $H = H_1 \oplus H_2$, is there a coupling of $\mathbf{P}^{H_1}$ and $\mathbf{P}^{H_2}$ with union marginal $\mathbf{P}^H$?*

A positive answer is supported by some numerical calculation. It is easily seen to hold when $H = \ell^2(E)$ by Corollary 2.3.

In the sequel, we write $Q_1 \preceq Q_2$ if $(Q_1 u,\, u) \le (Q_2 u,\, u)$ for all $u \in \ell^2(E)$.

Theorem 2.9 ([33, 7]). *If $0 \preceq Q_1 \preceq Q_2 \preceq I$, then $\mathbf{P}^{Q_1} \preccurlyeq \mathbf{P}^{Q_2}$.*

Proof. By Theorem 2.7, it suffices that there exist orthogonal projections P_1 and P_2 that are dilations of Q_1 and Q_2 such that $P_1 \preceq P_2$. This follows from Naĭmark's dilation theorem [43], which says that any measure whose values are positive operators, whose total mass is I, and which is countably additive in the weak operator topology dilates to a spectral measure. The measure in our case is defined on a 3-point space, with masses Q_1, $Q_2 - Q_1$, and $I - Q_2$, respectively. If we denote the respective dilations by R_1, R_2, and R_3, then we set $P_1 := R_1$ and $P_2 := R_1 + R_2$. ☐

A positive answer in general to Question 2.8 would give the following more general result by compression: If Q_1, Q_2 and $Q_1 + Q_2$ are positive contractions on $\ell^2(E)$, then there is a coupling of $\mathbf{P}^{Q_1}$ and $\mathbf{P}^{Q_2}$ with union marginal $\mathbf{P}^{Q_1+Q_2}$.

It would be very useful to have additional sufficient conditions for stochastic domination: see the end of Subsection 3.8 and Conjecture 5.7. For examples where more is known, see Theorem 5.2.

We shall say that the events in $\mathscr{F}(K)$ are ***measurable with respect to*** K and likewise for functions that are measurable with respect to $\mathscr{F}(K)$. We say that $\mathbf{P}$ has ***negative associations*** if for every pair f_1, f_2 of increasing functions that are measurable with respect to complementary subsets of E,

$$\mathbf{E}[f_1 f_2] \le \mathbf{E}[f_1]\mathbf{E}[f_2]\,. \tag{2.15}$$

Theorem 2.10 ([33]). *If $0 \preceq Q \preceq I$, then $\mathbf{P}^Q$ has negative associations.*

Proof. The details for finite E were given in [33]. For infinite E, let f_1 and f_2 be increasing bounded functions measurable with respect to $\mathscr{F}(A)$ and $\mathscr{F}(E \setminus A)$, respectively. Choose finite $E_n \uparrow E$. The conditional expectations $\mathbf{E}[f_1 \mid \mathscr{F}(A \cap E_n)]$ and $\mathbf{E}[f_2 \mid \mathscr{F}(E_n \setminus A)]$ are increasing functions to which (2.15) applies (because restriction to E_n corresponds to a compression of Q, which is a positive contraction) and which, being martingales, converge to f_1 and f_2 in $L^2(\mathbf{P}^Q)$. $\square$

2.5. Mixtures.

Write $\mathrm{Bern}(p)$ for the distribution of a Bernoulli random variable with expectation p. For $p_k \in [0,1]$, let $\mathrm{Bin}(\langle p_k \rangle)$ be the distribution of a sum of independent $\mathrm{Bern}(p_k)$ random variables. Recall that $[v]$ is the set of scalar multiples of v.

Theorem 2.11 ([1]; Lemma 3.4 of [48]; (2.38) of [49]; [26]). *Let Q be a positive contraction with spectral decomposition $Q = \sum_k \lambda_k P_{[v_k]}$, where $\langle v_k \,;\ k \geq 1 \rangle$ are orthonormal. Let $I_k \sim \mathrm{Bern}(\lambda_k)$ be independent. Let $\mathfrak{H} := \bigoplus_k [I_k v_k]$; thus, $Q = \mathbf{E} P_{\mathfrak{H}}$. Then $\mathbf{P}^Q = \mathbf{E}\mathbf{P}^{\mathfrak{H}}$. Hence, if $\mathfrak{S} \sim \mathbf{P}^Q$, then $|\mathfrak{S}| \sim \mathrm{Bin}(\langle \lambda_k \rangle)$.*

Proof. By Theorem 2.9, it suffices to prove it when only finitely many $\lambda_k \neq 0$. Then by Theorem 2.4, we have $\mathbf{P}^Q[A \subseteq \mathfrak{S}] = \Big(\bigwedge_{e \in A} Qe,\ \theta_A \Big)$ for all $A \subseteq E$. Now

$$\begin{aligned}\bigwedge_{e \in A} Qe = \bigwedge_{e \in A} \sum_k \lambda_k P_{[v_k]} e &= \sum_{j \colon A \to \mathbb{N}} \prod_{e \in A} \lambda_{j(e)} \bigwedge_{e \in A} P_{[v_{j(e)}]} e \\ &= \sum_{j \colon A \rightarrowtail \mathbb{N}} \prod_{e \in A} \lambda_{j(e)} \bigwedge_{e \in A} P_{[v_{j(e)}]} e\end{aligned}$$

because $v \wedge v = 0$ and $P_{[v]} e$ is a multiple of v, so none of the terms where j is not injective contribute. Thus,

$$\begin{aligned}\bigwedge_{e \in A} Qe &= \sum_{j \colon A \rightarrowtail \mathbb{N}} \mathbf{E}\Big[\prod_{e \in A} I_{j(e)} \Big] \bigwedge_{e \in A} P_{[v_{j(e)}]} e = \mathbf{E}\Big[\sum_{j \colon A \rightarrowtail \mathbb{N}} \prod_{e \in A} I_{j(e)} \bigwedge_{e \in A} P_{[v_{j(e)}]} e \Big] \\ &= \mathbf{E}\Big[\sum_{j \colon A \to \mathbb{N}} \prod_{e \in A} I_{j(e)} \bigwedge_{e \in A} P_{[v_{j(e)}]} e \Big] = \mathbf{E} \bigwedge_{e \in A} \sum_k I_k P_{[v_k]} e = \mathbf{E} \bigwedge_{e \in A} P_{\mathfrak{H}} e \,.\end{aligned}$$

We conclude that $\mathbf{P}^Q[A \subseteq \mathfrak{S}] = \mathbf{E}\left(\bigwedge_{e \in A} P_{\mathfrak{H}} e,\ \theta_A \right) = \mathbf{E}\big[\mathbf{P}^{\mathfrak{H}}[A \subseteq \mathfrak{B}] \big]$ by (2.8). $\square$

We sketch another proof: Let E' be disjoint from E with the same cardinality. Choose an orthonormal sequence $\langle v'_k \rangle$ in $\ell^2(E')$. Define

$$H := \bigoplus_k \big[\sqrt{\lambda_k} v_k + \sqrt{1 - \lambda_k} v'_k \big] \leq \ell^2(E \cup E') \,. \tag{2.16}$$

Then Q is the compression of P_H to $\ell^2(E)$. Expanding $\omega_H = \bigwedge_k (\sqrt{\lambda_k} v_k + \sqrt{1 - \lambda_k} v'_k)$ in the obvious way into orthogonal pieces and restricting to E, we obtain the desired equation from (2.2).

The first proof shows more generally the following: Let Q_0 be a positive contraction. Let $\langle v_k \,;\; k \ge 1\rangle$ be (not necessarily orthogonal) vectors such that $Q_0 + \sum_k P_{[v_k]} \preceq I$. Let I_k be independent Bernoulli random variables with $\mathbf{E} \sum_k I_k < \infty$. Write $\mathfrak{Q} := Q_0 + \sum_k I_k P_{[v_k]}$. Then $\mathbf{P}^{\mathbf{E}\mathfrak{Q}} = \mathbf{E}\mathbf{P}^{\mathfrak{Q}}$. This was observed by Ghosh and Krishnapur (personal communication, 2014).

Note that in the mixture of Theorem 2.11, the distribution of $\langle I_k \,;\; k \ge 1\rangle$ is determinantal corresponding to the diagonal matrix with diagonal $\langle \lambda_k \,;\; k \ge 1\rangle$. Thus, it is natural to wonder whether $\langle I_k \,;\; k \ge 1\rangle$ can be taken to be a general determinantal measure. If such a mixture is not necessarily determinantal, must it be strongly Rayleigh or at least have negative correlations? Here, we say that a probability measure $\mathbf{P}$ on 2^E has ***negative correlations*** if for every pair A, B of finite disjoint subsets of E, we have $\mathbf{P}[A \cup B \subseteq \mathfrak{S}] \le \mathbf{P}[A \subseteq \mathfrak{S}]\mathbf{P}[B \subseteq \mathfrak{S}]$. Note that negative associations is stronger than negative correlations.

2.6. Example: Uniform Spanning Trees and Forests. The most well-known example of a (nontrivial discrete) determinantal probability measure is that where $\mathfrak{S}$ is a uniformly chosen random spanning tree of a finite connected graph $G = (\mathsf{V}, \mathsf{E})$ with $E := \mathsf{E}$. Here, we regard a spanning tree as a set of edges. The fact that (1.1) holds for the uniform spanning tree is due to [12] and is called the Transfer Current Theorem. The case with $|A| = 1$ was shown much earlier by [30], while the case with $|A| = 2$ was first shown by [11]. Write UST_G for the uniform spanning tree measure on G.

To see that UST_G is indeed determinantal, consider the vertex-edge incidence matrix M of G, where each edge is oriented (arbitrarily) and the (x, e)-entry of M equals 1 if x is the head of e, -1 if x is the tail of e, and 0 otherwise. Identifying an edge with its corresponding column of M, we find that a spanning tree is the same as a basis of the column space of M. Given $x \in \mathsf{V}$, define the ***star*** at x to be the x-row of M, regarded as a vector $\star_x$ in the row space, $\bigstar(G) \le \ell^2(\mathsf{E})$. It is easy that the row-rank of M is $|\mathsf{V}| - 1$. Let $x_0 \in \mathsf{V}$ and let $\mathbf{u}$ be the wedge product (in some order) of the stars at all the vertices other than x_0. Thus, $\mathbf{u} = c\,\omega_{\bigstar(G)}$ for some $c \ne 0$. Since spanning trees are bases of the column space of M, we have $\big(\mathbf{u},\ \theta_A\big) \ne 0$ iff A is a spanning tree. That is, the only non-zero coefficients of $\mathbf{u}$ are those in which choosing one edge in each $\star_x$ for $x \ne x_0$ yields a spanning tree; moreover, each spanning tree occurs exactly once since there is exactly one way to choose an edge incident to each $x \ne x_0$ to get a given spanning tree. This means that its coefficient is ± 1. Hence, $\mathbf{P}^{\bigstar(G)}$ is indeed uniform on spanning trees. Simultaneously, this proves the matrix tree theorem that the number of spanning trees equals $\det[(\star_x,\ \star_y)]_{x,y \ne x_0}$, since this determinant is $\|\mathbf{u}\|^2$.

One can define analogues of UST_G on infinite connected graphs [44, 22, 2] by weak limits. For brevity, we simply define them here as determinantal probability measures. Again, all edges of G are oriented arbitrarily. We define $\bigstar(G)$ as the closure of the linear span of the stars. An element of $\ell^2(\mathsf{E})$ that is finitely supported and orthogonal to $\bigstar(G)$ is called a ***cycle***; the closed linear span of the cycles is $\diamondsuit(G)$. The ***wired uniform spanning forest*** is $\mathsf{WSF}_G := \mathbf{P}^{\bigstar(G)}$, while the ***free uniform spanning forest*** is $\mathsf{FSF}_G := \mathbf{P}^{\diamondsuit(G)^\perp}$.

3. Continuous Basics

Our discussion of the "continuous" case includes the discrete case, but the discrete case has the more elementary formulations given earlier.

Let E be a measurable space. As before, E will play the role of the underlying set on which a point process forms a counting measure. While before we implicitly used counting measure on E itself, now we shall have an arbitrary measure μ; it need not be a probability measure. The case of Lebesgue measure on Euclidean space is a common one. The Hilbert spaces of interest will be $L^2(E, \mu)$.

3.1. Symmetrization and Anti-symmetrization.

There may be no natural order in E, so to define, e.g., a probability measure on n points of E, it is natural to use a probability measure on E^n that is symmetric under coordinate changes and that vanishes on the diagonal $\Delta_n(E) := \big\{(x_1, \ldots, x_n) \in E^n \,;\, \exists i \neq j \;\; x_i = x_j\big\}$. Likewise, for exterior algebra, it is more convenient to identify $u_1 \wedge \cdots \wedge u_n$ with

$$\frac{1}{\sqrt{n!}} \sum_{\sigma \in \mathsf{Sym}(n)} (-1)^\sigma u_{\sigma(1)} \otimes \cdots \otimes u_{\sigma(n)} \in L^2(E^n, \mu^n)$$

for $u_i \in L^2(E, \mu)$. Thus, $u_1 \wedge \cdots \wedge u_n$ is identified with the function $(x_1, \ldots, x_n) \mapsto \det[u_i(x_j)]_{i,j \in \{1,\ldots,n\}}/\sqrt{n!}$. Note that

$$\begin{aligned} n!\Big(\bigwedge_{i=1}^n u_i\Big)\Big(\bigwedge_{i=1}^n v_i\Big)(x_1, \ldots, x_n) &= \det[u_i(x_j)] \det[v_i(x_j)] = \det[u_i(x_j)] \det[v_i(x_j)]^T \\ &= \det[u_i(x_j)][v_i(x_j)]^T = \det[K(x_i, x_j)]_{i,j \in \{1,\ldots,n\}} \end{aligned} \tag{3.1}$$

with $K := \sum_{i=1}^n u_i \otimes v_i$. Here, ^{T} denotes transpose.

3.2. Joint Intensities.

Suppose from now on that E is a locally compact Polish space (equivalently, a locally compact second countable Hausdorff space). Let μ be a Radon measure on E, i.e., a Borel measure that is finite on compact sets. Let $\mathcal{N}(E)$ be the set of Radon measures on E with values in $\mathbb{N} \cup \{\infty\}$. We give $\mathcal{N}(E)$ the vague topology generated by the maps $\xi \mapsto \int f \, d\xi$ for continuous f with compact support; then $\mathcal{N}(E)$ is Polish. The corresponding Borel σ-field of $\mathcal{N}(E)$ is generated by the maps $\xi \mapsto \xi(A)$ for Borel $A \subseteq E$. Let $\mathfrak{X}$ be a simple point process on E, i.e., a random variable with values in $\mathcal{N}(E)$ such that $\mathfrak{X}(\{x\}) \in \{0, 1\}$ for all $x \in E$. The power $\mathfrak{X}^k := \mathfrak{X} \otimes \cdots \otimes \mathfrak{X}$ lies in $\mathcal{N}(E^k)$. Thus, $\mathbf{E}[\mathfrak{X}^k]$ is a Borel measure on E^k; the part of it that is concentrated on $E^k \setminus \Delta_k(E)$ is called the ***k*-point intensity measure** of $\mathfrak{X}$. If the intensity measure is absolutely continuous with respect to μ^k, then its Radon-Nikodym derivative ρ_k is called the ***k*-point intensity function** or the ***k*-point correlation function**:

$$\text{for all Borel } A \subseteq E^k \setminus \Delta_k(E) \qquad \mathbf{E}[\mathfrak{X}^k(A)] = \int_A \rho_k \, d\mu^k \,. \tag{3.2}$$

Since the intensity measure vanishes on the diagonal $\Delta_k(E)$, we take ρ_k to vanish on $\Delta_k(E)$. We also take ρ_k to be symmetric under permutations of coordinates. Intensity functions are the continuous analogue of the elementary probabilities (1.1).

Since the sets $\prod_{i=1}^k A_i := A_1 \times \cdots \times A_k$ generate the σ-field on $E^k \setminus \Delta_k(E)$ for pairwise disjoint Borel $A_1, \ldots, A_k \subseteq E$, a measurable function $\rho_k \colon E^k \to [0, \infty)$ is "the" k-point intensity function iff

$$\mathbf{E}\Big[\prod_{i=1}^k \mathfrak{X}(A_i)\Big] = \int_{\prod_{i=1}^k A_i} \rho_k \, d\mu^k \,. \tag{3.3}$$

Since $\mathfrak{X}$ is simple, $\mathfrak{X}^k\big(A^k \setminus D_k(A)\big) = \big(\mathfrak{X}(A)\big)_k$, where $(n)_k := n(n-1)\cdots(n-k+1)$. Since ρ_k vanishes on the diagonal, it follows from (3.2) that for disjoint $A_1, \ldots, A_r$ and non-negative $k_1, \ldots, k_r$ summing to k,

$$\mathbf{E}\Big[\prod_{j=1}^r \big(\mathfrak{X}(A_j)\big)_{k_j}\Big] = \int_{\prod_{j=1}^r A_j^{k_j}} \rho_k \, d\mu^k \,. \tag{3.4}$$

Again, this characterizes ρ_k, even if we use only $r = 1$.

In the special case that $\mathfrak{X}(E) = n$ a.s. for some $n \in \mathbb{Z}^+$, then the definition (3.2) shows that a random ordering of the n points of $\mathfrak{X}$ has density $\rho_n/n!$. More generally, (3.2) shows that for all $k < n$,

$$\text{the density of a random (ordered) } k\text{-tuple of } \mathfrak{X} \text{ is } \rho_k/(n)_k \,, \tag{3.5}$$

whence in this case,

$$\rho_k(x_1, \ldots, x_k) = \frac{1}{(n-k)!} \int_{E^{n-k}} \rho_n(x_1, \ldots, x_n) \, d\mu^{n-k}(x_{k+1}, \ldots, x_n) \,. \tag{3.6}$$

We call $\mathfrak{X}$ ***determinantal*** if for some measurable $K \colon E^2 \to \mathbb{C}$ and all $k \ge 1$, $\rho_k(F) = \det(K{\restriction}F)$ μ^k-a.e. Here, $K{\restriction}(x_1, \ldots, x_k)$ is the matrix $[K(x_i, x_j)]_{i,j \le k}$. In this case, we denote the law of $\mathfrak{X}$ by $\mathbf{P}^K$.

We consider only K that are locally square integrable (i.e., $|K|^2\mu^2$ is Radon), are Hermitian (i.e., $K(y, x) = \overline{K(x, y)}$ for all $x, y \in E$), and are positive semidefinite (i.e., $K{\restriction}F$ is positive semidefinite for all finite F, written $K \succeq 0$). In this case, K defines a positive semidefinite integral operator $(Kf)(x) := \int K(x, y) f(y) \, d\mu(y)$ on functions $f \in L^2(\mu)$ with compact support. For every Borel $A \subseteq E$, we denote by μ_A the measure μ restricted to Borel subsets of A and by K_A the compression of K to A, i.e., $K_A f := (Kf){\restriction}A$ for $f \in L^2(A, \mu_A)$. The operator K is locally trace-class, i.e., for every compact $A \subseteq E$, the compression K_A is trace class, having a spectral decomposition $K_A f = \sum_k \lambda_k^A (f, \phi_k^A) \phi_k^A$, where $\langle \phi_k^A \,;\; k \ge 1 \rangle$ are orthonormal eigenfunctions of K_A with positive summable eigenvalues $\langle \lambda_k^A \,;\; k \ge 1 \rangle$. If A_1 is the set where $\sum_k \lambda_k^A |\phi_k^A|^2 < \infty$, then $\mu(A \setminus A_1) = 0$ and $\sum_k \lambda_k^A \phi_k^A \otimes \overline{\phi_k^A}$ converges on A_1^2, with sum μ_A^2-a.e. equal to K. We normally redefine K on a set of measure 0 to equal this sum. Such a K defines a determinantal point process iff the integral operator K extends to all of $L^2(\mu)$ as a positive contraction [39, 51, 26]. The joint intensities determine uniquely the law of the point process [27, Lemma 4.2.6]. Poisson processes are not determinantal processes, but when μ is continuous, they are distributional limits of determinantal processes.

3.3. Construction. To see that a positive contraction defines a determinantal point process, we first consider K that defines an orthogonal projection onto a finite-dimensional subspace, H. Then $K = \sum_{k=1}^{n} \phi_k \otimes \overline{\phi_k}$ for every orthonormal basis $\langle \phi_k \,;\ k \le n \rangle$ of H and $\omega_H = \bigwedge_{i=1}^{n} \phi_k$ is a unit multivector in the notation of Subsection 2.1. Because of (3.1), we have

$$\frac{1}{n!} \int \det[K(x_i, x_j)]_{i,j \le n} \, d\mu^n(x_1, \ldots, x_n) = \Big\| \bigwedge_{k=1}^{n} \phi_k \Big\|^2 = 1 \,, \tag{3.7}$$

i.e., $\det[K(x_i, x_j)]/n!$ is a density with respect to μ^n. Although in the discrete case, the absolute squared coefficients of $\bigwedge_{k=1}^{n} \phi_k$ give the elementary probabilities, now coefficients are replaced by a function whose absolute square gives a probability density. As noted already, (3.7) means that $F \mapsto \det(K {\restriction} F)$ is the n-point intensity function. In order to show that this density gives a determinantal process with kernel K, we use the Cauchy-Binet formula, which may be stated as follows: For $k \times n$ matrices $a = [a_{i,j}]$ and $b = [b_{i,j}]$ with $a^J := [a_{i,j}]_{\substack{i \le k \\ j \in J}}$, we have

$$\det\big([a_{i,j}][b_{i,j}]^T\big) = \sum_{|J|=k} \det a^J \cdot \det b^J = \sum_{\substack{\sigma,\tau \in \mathsf{Sym}(k,n) \\ \mathrm{im}(\sigma)=\mathrm{im}(\tau)}} (-1)^\sigma (-1)^\tau \prod_{i=1}^{k} a_{i,\sigma(i)} b_{i,\tau(i)} \,,$$

where $\mathrm{im}(\sigma)$ denotes the image of σ and the sums extend over all pairs of injections

$$\sigma, \tau \colon \{1, 2, \ldots, k\} \rightarrowtail \{1, 2, \ldots, n\} \,.$$

Here, the sign $(-1)^\sigma$ of σ is defined in the usual way by the parity of the number of pairs $i < j$ for which $\sigma(i) > \sigma(j)$. We have

$$\begin{aligned}
\rho_k(x_1, \ldots, x_k) &= \frac{1}{(n-k)!} \int_{E^{n-k}} \det[K(x_i, x_j)] \, d\mu^{n-k}(x_{k+1}, \ldots, x_n) \\
&= \frac{1}{(n-k)!} \int_{E^{n-k}} \sum_{\sigma \in \mathsf{Sym}(n)} (-1)^\sigma \prod_{i=1}^{n} \phi_{\sigma(i)}(x_i) \cdot \\
&\qquad \cdot \sum_{\tau \in \mathsf{Sym}(n)} (-1)^\tau \prod_{i=1}^{n} \overline{\phi_{\tau(i)}(x_i)} \, d\mu^{n-k}(x_{k+1}, \ldots, x_n) \\
&= \sum_{\substack{\sigma,\tau \in \mathsf{Sym}(k,n) \\ \mathrm{im}(\sigma)=\mathrm{im}(\tau)}} (-1)^\sigma (-1)^\tau \prod_{i=1}^{k} \phi_{\sigma(i)}(x_i) \overline{\phi_{\tau(i)}(x_i)} \\
&= \det\big(K {\restriction} (x_1, \ldots, x_k)\big) \,.
\end{aligned} \tag{3.8}$$

Here, the first equality uses (3.6), the second equality uses (3.1), the third equality uses the fact that $\int_E \phi_{\sigma(i)}(x_i) \overline{\phi_{\tau(i)}(x_i)} \, d\mu(x_i)$ is 1 or 0 according as $\sigma(i) = \tau(i)$ or not, and the fourth equality uses Cauchy-Binet. Note that a factor of $(n-k)!$ arises because for every pair of injections $\sigma_1, \tau_1 \in \mathsf{Sym}(k,n)$ with equal image, there are $(n-k)!$ extensions of them to permutations $\sigma, \tau \in \mathsf{Sym}(n)$ with $\sigma(i) = \tau(i)$ for all $i > k$; in this case, $(-1)^\sigma(-1)^\tau = (-1)^{\sigma_1}(-1)^{\tau_1}$. We write $\mathbf{P}^H$ for the law of the associated point process on E.

Lemma 3.1. *Let $\mathfrak{X}_n \sim \mathbf{P}^{K_n}$ with $K_n(x,x) \le f(x)$ for some $f \in L^1_{\mathrm{loc}}(E,\mu)$. Then $\{\mathbf{P}^{K_n}\,;\, n \ge 1\}$ is tight and every weak limit point of $\mathfrak{X}_n$ is simple.*

Proof. By using the kernel $K_n(x,y)/\sqrt{f(x)f(y)}$ with respect to the measure $f\mu$, we may assume that $f \equiv 1$. Tightness follows from

$$m\mathbf{P}[\mathfrak{X}_n(A) \ge m] \le \mathbf{E}[\mathfrak{X}_n(A)] = \int_A K_n(x,x)\,d\mu(x)\,.$$

For the rest, we may assume that E is compact and $\mu(E) = 1$. Let $\mathfrak{X}$ be a limit point of $\mathfrak{X}_n$. Let μ_{d} be the atomic part of μ and $\mu_{\mathrm{c}} := \mu - \mu_{\mathrm{d}}$. Choose $m \ge 1$ and partition E into sets $A_1, \ldots, A_m$ with $\mu_{\mathrm{c}}(A_i) \le 1/m$. Let A be such that $\mu_{\mathrm{d}}(E \setminus A) = 0$ and $\mu_{\mathrm{c}}(A) = 0$. Let U be open such that $A \subseteq U$ and $\mu_{\mathrm{c}}(U) < 1/m$. Then

$$\begin{aligned}\mathbf{P}[\mathfrak{X} \text{ is not simple}] &\le \limsup_n \big(\mathbf{P}[\mathfrak{X}_n(U \setminus A) \ge 1] + \mathbf{P}[\exists i \;\; \mathfrak{X}_n(A_i) \ge 2]\big) \\ &\le \limsup_n \big(\mathbf{E}[\mathfrak{X}_n(U \setminus A)] + \sum_i \mathbf{E}[(\mathfrak{X}_n(A_i))_2]\big) \\ &\le \mu_{\mathrm{c}}(U) + \sum_i \mu(A_i)^2 < 2/m\,. \end{aligned}$$

□

Now, given any locally trace-class orthogonal projection K onto H, choose finite-dimensional subspaces $H_n \uparrow H$ with corresponding projections K_n. Clearly $K_n(x,y) \to K(x,y)$ μ^2-a.e. and $K_n(x,x) \le K(x,x)$ μ-a.e. Thus, the joint intensity functions converge a.e. By dominated convergence, if $A \subset E^k \setminus \Delta_k(E)$ is relatively compact and Borel, then $\mathbf{E}^{H_n}[\mathfrak{X}(A)] \to \int_A \det(K{\restriction}F)\,d\mu^k(F)$. By uniform exponential moments of $\mathfrak{X}(A)$ [27, proof of Lemma 4.2.6], it follows that all weak limit points of $\mathbf{P}^{H_n}$ are equal, and hence, by Lemma 3.1, define $\mathbf{P}^H$ with kernel K. (In Subsection 3.7, we shall see that $\langle \mathbf{P}^{H_n}\,;\, n \ge 1\rangle$ is stochastically increasing.)

Finally, let K be any locally trace-class positive contraction. Define the orthogonal projection on $L^2(E,\mu) \oplus L^2(E,\mu)$ whose block matrix is

$$\begin{pmatrix} K & \sqrt{K(I-K)} \\ \sqrt{K(I-K)} & I-K \end{pmatrix}. \tag{3.9}$$

Take an isometric isomorphism of $L^2(E,\mu)$ to $\ell^2(E')$ for some denumerable set E' and interpret the above as an orthogonal projection K' on $L^2(E,\mu) \oplus \ell^2(E')$. Then K' is clearly locally trace-class and K is the compression of K' to E. Thus, we define $\mathbf{P}^K$ by intersecting samples of $\mathbf{P}^{K'}$ with E. We remark that by writing K' as a limit of increasing finite-rank projections that we then compress, we see that $\mathbf{P}^K$ may be defined as a limit of determinantal processes corresponding to increasing finite-rank positive contractions.

Conjecture 3.2. *If K is a locally trace-class positive contraction, then $\mathbf{P}^K$ has trivial tail in that every event in $\bigcap_{\text{compact } A \subset E} \mathscr{F}(E \setminus A)$ is trivial.*

3.4. Mixtures. Rather than using compressions as in the last paragraph above, an alternative approach to defining $\mathbf{P}^K$ uses mixtures and starts from finite-rank projections, as in Subsection 2.5. This approach is due to [26]. Consider first a finite-rank $K := \sum_{j=1}^n \lambda_j \phi_j \otimes \overline{\phi_j}$. Let $I_j \sim \mathrm{Bern}(\lambda_j)$ be independent. Let $\mathfrak{H} := \bigoplus_j [I_j \phi_j]$; thus, $K = \mathbf{E} P_{\mathfrak{H}}$. We claim that $\mathbf{P}^K := \mathbf{E}\mathbf{P}^{\mathfrak{H}}$ is determinantal with kernel K. Indeed, it is clearly a simple point process. Write $\Phi_J := \bigwedge_{j \in J} I_j \phi_j$, $\psi_j := \sqrt{\lambda_j}\phi_j$, and $\psi_J := \bigwedge_{j\in J} \psi_j$. Let $F \in E^k$. Combining Cauchy-Binet with (3.1) yields $\det\big(K{\restriction}F\big) = k! \sum_{|J|=k} |\psi_J(F)|^2$. Similarly, the joint intensities of $\mathbf{E}\mathbf{P}^{\mathfrak{H}}$ are the expectations of the joint intensities of $\mathbf{P}^{\mathfrak{H}}$, which equal

$$\mathbf{E}\big[\det(P_{\mathfrak{H}}{\restriction}F)\big] = \mathbf{E}\Big[k! \sum_{|J|=k} |\Phi_J(F)|^2\Big] = \det\big(K{\restriction}F\big)\,.$$

Essentially the same works for trace-class $K = \sum_{j=1}^\infty \lambda_j \phi_j \otimes \overline{\phi_j}$; we need merely take, in the last step, a limit in the above equation as $n \to \infty$ for $K_n := \sum_{j=1}^n \lambda_j \phi_j \otimes \overline{\phi_j}$, since all terms are non-negative and $K_n \to K$ a.e.

Given this construction of $\mathbf{P}^K$ for trace-class K, one can then construct $\mathbf{P}^K$ for a general locally trace-class positive contraction by defining its restriction to each relatively compact set A via the trace-class compression K_A.

As noted by [26], a consequence of the mixture representation is a CLT due originally to [52]:

Theorem 3.3. *Let K_n be trace-class positive contractions on spaces $L^2(E_n, \mu_n)$. Let $\mathfrak{X}_n \sim \mathbf{P}^{K_n}$ and write $|\mathfrak{X}_n| := \mathfrak{X}_n(E_n)$. If $\mathrm{Var}\big(|\mathfrak{X}_n|\big) \to \infty$ as $n \to \infty$, then $\langle |\mathfrak{X}_n|\,;\ n \ge 1\rangle$ obeys a CLT.*

3.5. Simulation. In order to simulate $\mathbf{P}^K$ when K is a trace-class positive contraction, it suffices, by taking a mixture as above, to see how to simulate $\mathfrak{X} \sim \mathbf{P}^H$ when $n := \dim H < \infty$. The following algorithm [26, Algo. 18] gives a uniform random ordering of $\mathfrak{X}$ as $\langle X_1, \ldots, X_n\rangle$. Since $\mathbf{E}\big[\mathfrak{X}(E)\big] = n$, the measure $\mathbf{E}[\mathfrak{X}]/n = n^{-1}K(x,x)\,d\mu(x)$ is a probability measure on E. Select a point X_1 at random from that measure. If $n = 1$, then we are done. If not, then let H_1 be the orthogonal complement in H of the function $K_{X_1} := \sum_{k=1}^n \overline{\phi_k(X_1)}\phi_k \in H$, where $\langle \phi_k\,;\ k \le n\rangle$ is an orthonormal basis for H. Then $\dim H_1 = n-1$ and we may repeat the above for H_1 to get the next point, X_2, then $H_2 := H_1 \cap K_{X_2}^\perp$, etc. The conditional density of X_{k+1} given $X_1, \ldots, X_k$ is $(n-k)^{-1}\det\big(K{\restriction}(x, X_1, \ldots, X_k)\big)/\det\big(K{\restriction}(X_1, \ldots, X_k)\big)$ by (3.5), i.e., $(n-k)^{-1}$ times the squared distance from K_x to the linear span of $K_{X_1}, \ldots, K_{X_k}$. It can help for rejection sampling to note that this is at most $(n-k)^{-1}K(x,x)$. One can also sample faster by noting that the conditional distribution of X_{k+1} is the same as that of $\mathbf{P}^{\mathfrak{v}}$, where $\mathfrak{v}$ is a uniformly random vector on the unit sphere of H_k.

3.6. Transference Principle. Note that if $N_1, \ldots, N_r$ are bounded $\mathbb{N}$-valued random variables, then the function $(k_1, \ldots, k_r) \mapsto \mathbf{E}\Big[\prod_{j=1}^r \big(N_j\big)_{k_j}\Big]$ determines the joint distribution of $\langle N_j\,;\ j \le r\rangle$ since it gives the derivatives at $(1, 1, \ldots, 1)$ of the probability generating function $(s_1, \ldots, s_r) \mapsto \mathbf{E}\Big[\prod_{j=1}^r s_j^{N_j}\Big]$.

Let us re-examine (3.4) in the context of a finite-rank $K = \sum_{i=1}^n \lambda_i \phi_i \otimes \overline{\phi_i}$. Given disjoint $A_1, \ldots, A_r \subseteq E$ and non-negative $k_1, \ldots, k_r$ summing to k, it will be convenient to write $\kappa(j) := \min\big\{m \ge 1\,;\ j \le \sum_{\ell=1}^m k_\ell\big\}$ for $j \le k$. We have by Cauchy-Binet

$$\begin{aligned}
\mathbf{E}^K\Big[\prod_{\ell=1}^r \big(\mathfrak{X}(A_\ell)\big)_{k_\ell}\Big] &= \int_{\prod_{\ell=1}^r A_\ell^{k_\ell}} \rho_k \, d\mu^k = \int_{\prod_{\ell=1}^r A_\ell^{k_\ell}} \det(K{\restriction}(x_1, \ldots, x_k)) \prod_{j=1}^k d\mu(x_j) \\
&= \int_{\prod_{\ell=1}^r A_\ell^{k_\ell}} \sum_{\substack{\sigma,\tau \in \mathsf{Sym}(k,n) \\ \mathrm{im}(\sigma)=\mathrm{im}(\tau)}} (-1)^\sigma (-1)^\tau \prod_{j=1}^k \lambda_{\sigma(j)} \phi_{\sigma(j)}(x_j) \overline{\phi_{\tau(j)}(x_j)} \prod_{j=1}^k d\mu(x_j) \\
&= \sum_{\substack{\sigma,\tau \in \mathsf{Sym}(k,n) \\ \mathrm{im}(\sigma)=\mathrm{im}(\tau)}} (-1)^\sigma (-1)^\tau \prod_{j=1}^k \int_{A_{\kappa(j)}} \lambda_{\sigma(j)} \phi_{\sigma(j)}(x_j) \overline{\phi_{\tau(j)}(x_j)} \, d\mu(x_j) \\
&= \sum_{\substack{\sigma,\tau \in \mathsf{Sym}(k,n) \\ \mathrm{im}(\sigma)=\mathrm{im}(\tau)}} (-1)^\sigma (-1)^\tau \lambda^{\mathrm{im}(\sigma)} \prod_{j=1}^k \big(\mathbf{1}_{A_{\kappa(j)}} \phi_{\sigma(j)},\ \overline{\phi_{\tau(j)}}\big) \\
&= \sum_{\sigma \in \mathsf{Sym}(k,n)} (-1)^\sigma \lambda^{\mathrm{im}(\sigma)} \det\Big[\Big(\mathbf{1}_{A_{\kappa(j)}} \phi_{\sigma(j)},\ \overline{\phi_\ell}\Big)\Big]_{\substack{j \le k \\ \ell \in \mathrm{im}(\sigma)}} .
\end{aligned}$$

As an immediate consequence of this formula, we obtain the following important principle of Goldman [21, Proposition 12] that allows one to infer properties of continuous determinantal point processes from corresponding properties of discrete determinantal probability measures:

Theorem 3.4. *Let (E, μ) and (F, ν) be two Radon measure spaces on locally compact Polish sets. Let $\langle A_i\rangle$ be pairwise disjoint Borel subsets of E and $\langle B_i\rangle$ be pairwise disjoint Borel subsets of F. Let $\lambda_k \in [0, 1]$ with $\sum_k \lambda_k < \infty$. Let $\langle \phi_k\rangle$ be orthonormal in $L^2(E, \mu)$ and $\langle \psi_k\rangle$ be orthonormal in $L^2(F, \nu)$. Let $K := \sum_k \lambda_k \phi_k \otimes \overline{\phi_k}$ and $L := \sum_k \lambda_k \psi_k \otimes \overline{\psi_k}$. If $\big(\mathbf{1}_{A_i}\phi_j,\ \phi_k\big) = \big(\mathbf{1}_{B_i}\psi_j,\ \psi_k\big)$ for all i, j, k, then the $\mathbf{P}^K$-distribution of $\langle \mathfrak{X}(A_i)\rangle$ equals the $\mathbf{P}^L$-distribution of $\langle \mathfrak{X}(B_i)\rangle$.*

Proof. When only finitely many $\lambda_k \ne 0$, this follows from our previous calculation. The general case follows from weak convergence of the processes corresponding to the partial sums, as in the paragraph following Lemma 3.1. ☐

This permits us to compare to discrete measures via [21, Lemma 16]:

Lemma 3.5. *Let μ be a Radon measure on a locally compact Polish space, E. Let $\langle A_i\rangle$ be pairwise disjoint Borel subsets of E. Let $\phi_k \in L^2(E, \mu)$ for $k \ge 1$. Then there exists a denumerable set F, pairwise disjoint subsets $\langle B_i\rangle$ of F, and $v_k \in \ell^2(F)$ such that $(\phi_j,\ \phi_k) = (v_j,\ v_k)$ and $\big(\mathbf{1}_{A_i}\phi_j,\ \phi_k\big) = \big(\mathbf{1}_{B_i}v_j,\ v_k\big)$ for all i, j, k.*

Proof. Without loss of generality, we may assume that $\bigcup_i A_i = E$. For each i, fix an orthonormal basis $\langle w_{i,j}\,;\ j < n_i\rangle$ for the subspace of $L^2(E,\mu)$ spanned by $\{\mathbf{1}_{A_i}\phi_j\}$. Here, $n_i \in \mathbb{N} \cup \{\infty\}$. Define $B_i := \{(i,j)\,;\ j < n_i\}$ and $F := \bigcup_i B_i$. Let T be the isometric isomorphism from the span of $\{w_{i,j}\,;\ i \ge 1, j < n_i\}$ to $\ell^2(F)$ that sends $w_{i,j}$ to $\mathbf{1}_{\{(i,j)\}}$. Defining $v_k := T(\phi_k)$ yields the desired vectors. □

3.7. Stochastic Inequalities. We now show how the discrete models of Subsection 3.6 allow us to obtain the analogues of the stochastic inequalities known to hold for discrete determinantal probability measures.

For a Borel set $A \subseteq E$, let $\mathscr{F}(A)$ denote the σ-field on $\mathcal{N}(E)$ generated by the functions $\xi \mapsto \xi(B)$ for Borel $B \subseteq A$. We say that a function that is measurable with respect to $\mathscr{F}(A)$ is, more simply, measurable with respect to A. The obvious partial order on $\mathcal{N}(E)$ allows us to define what it means for a function $f \colon \mathcal{N}(E) \to \mathbb{R}$ to be ***increasing***. As in the discrete case, we say that $\mathbf{P}$ has ***negative associations*** if $\mathbf{E}[f_1 f_2] \le \mathbf{E}[f_1]\mathbf{E}[f_2]$ for every pair f_1, f_2 of bounded increasing functions that are measurable with respect to complementary subsets of E. An event is increasing if its indicator is increasing. Then $\mathbf{P}$ has negative associations iff

$$\mathbf{P}(\mathcal{A}_1 \cap \mathcal{A}_2) \le \mathbf{P}(\mathcal{A}_1)\mathbf{P}(\mathcal{A}_2) \tag{3.10}$$

for every pair $\mathcal{A}_1$, $\mathcal{A}_2$ of increasing events that are measurable with respect to complementary subsets of E.

We also say that $\mathbf{P}_1$ ***is stochastically dominated by*** $\mathbf{P}_2$ and write $\mathbf{P}_1 \preccurlyeq \mathbf{P}_2$ if $\mathbf{P}_1(\mathcal{A}) \le \mathbf{P}_2(\mathcal{A})$ for every increasing event $\mathcal{A}$.

Call an event ***elementary increasing*** if it has the form $\{\xi\,;\ \xi(B) \ge k\}$, where B is a relatively compact Borel set and $k \in \mathbb{N}$. Write $\mathscr{U}(A)$ for the closure under finite unions and intersections of the collection of elementary increasing events with $B \subseteq A$; the notation $\mathscr{U}$ is chosen for "upwardly closed". Note that every event in $\mathscr{U}(A)$ is measurable with respect to some finite collection of functions $\xi \mapsto \xi(B_i)$ for pairwise *disjoint* relatively compact Borel $B_i \subseteq A$. Write $\overline{\mathscr{U}(A)}$ for the closure of $\mathscr{U}(A)$ under monotone limits, i.e., under unions of increasing sequences and under intersections of decreasing sequences; these events are also increasing. This is the same as the closure of $\mathscr{U}(A)$ under countable unions and intersections.

Lemma 3.6. *Let A be a Borel subset of a locally compact Polish space, E. Then $\overline{\mathscr{U}(A)}$ is exactly the class of increasing Borel sets in $\mathscr{F}(A)$.*

We give a proof at the end of this subsection. First, we derive two consequences. A weaker version (negative correlations of elementary increasing events) of the initial one is due to [20].

Theorem 3.7. *Let μ be a Radon measure on a locally compact Polish space, E. Let K be a locally trace-class positive contraction on $L^2(E,\mu)$. Then $\mathbf{P}^K$ has negative associations.*

Proof. Let $A \subset E$ be Borel. Let $\mathcal{A}_1 \in \mathscr{U}(A)$ and $\mathcal{A}_2 \in \mathscr{U}(E \setminus A)$. Then $\mathcal{A}_1, \mathcal{A}_2 \in \mathscr{F}(B)$ for some compact B by definition of $\mathscr{U}(\cdot)$. We claim that (3.10) holds for $\mathcal{A}_1$, $\mathcal{A}_2$, and $\mathbf{P} = \mathbf{P}^{K_B}$, i.e., for $\mathbf{P} = \mathbf{P}^K$.

Now $\mathcal{A}_1$ is measurable with respect to a finite number of count functions $\xi \mapsto \xi(B_i)$ for some disjoint $B_i \subseteq A \cap B$ $(1 \le i \le n)$ and likewise $\mathcal{A}_2$ is measurable with respect to a finite number of functions $\xi \mapsto \xi(C_i)$ for some disjoint $C_i \subseteq B \backslash A$ $(1 \le i \le n)$. Thus, there are functions g_1 and g_2 such that $\mathbf{1}_{\mathcal{A}_1}(\xi) = g_1\big(\xi(B_1), \ldots, \xi(B_n)\big)$ and $\mathbf{1}_{\mathcal{A}_2}(\xi) = g_2\big(\xi(C_1), \ldots, \xi(C_n)\big)$. By Theorem 3.4 and Lemma 3.5, there is some discrete determinantal probability measure $\mathbf{P}^Q$ on some denumerable set F and pairwise disjoint sets $B'_i, C'_i \subseteq F$ such that the joint $\mathbf{P}^{K_B}$-distribution of all $\mathfrak{X}(B_i)$ and $\mathfrak{X}(C_i)$ is equal to the joint $\mathbf{P}^Q$-distribution of all $\mathfrak{X}(B'_i)$ and $\mathfrak{X}(C'_i)$. Define the corresponding events $\mathcal{A}'_i$ by $\mathbf{1}_{\mathcal{A}'_1}(\xi) = g_1\big(\xi(B'_1), \ldots, \xi(B'_n)\big)$ and $\mathbf{1}_{\mathcal{A}'_2}(\xi) = g_2\big(\xi(C'_1), \ldots, \xi(C'_n)\big)$. Since $\mathcal{A}'_i$ depend on disjoint subsets of F, Theorem 2.10 gives that $\mathbf{P}^Q(\mathcal{A}'_1 \cap \mathcal{A}'_2) \le \mathbf{P}^Q(\mathcal{A}'_1)\mathbf{P}^Q(\mathcal{A}'_2)$. This is the same as (3.10) by Theorem 3.4.

The same (3.10) clearly then holds in the less restrictive setting $\mathcal{A}_i \in \overline{\mathscr{U}(A)}$ by taking monotone limits. Lemma 3.6 completes the proof. □

Theorem 3.8 (Theorem 3 of [21]). *Suppose that K_1 and K_2 are two locally trace-class positive contractions such that $K_1 \preceq K_2$. Then $\mathbf{P}^{K_1} \preccurlyeq \mathbf{P}^{K_2}$.*

Proof. It suffices to show that $\mathbf{P}^{K_1}(\mathcal{A}) \le \mathbf{P}^{K_2}(\mathcal{A})$ for every $\mathcal{A} \in \mathscr{U}(E)$. Again, it suffices to assume that K_i are trace class. Lemma 3.5 applied to all eigenfunctions of K_1 and K_2 yields a denumerable F and two positive contractions K'_i on $\ell^2(F)$, together with an event $\mathcal{A}'$, such that $\mathbf{P}^{K'_i}(\mathcal{A}') = \mathbf{P}^{K_i}(\mathcal{A})$ for $i = 1, 2$. Furthermore, by construction, every function in $\ell^2(F)$ is the image of a function in $L^2(E)$ under the isometric isomorphism T used to prove Lemma 3.5, whence $K'_1 \preceq K'_2$. Therefore Theorem 2.9 yields $\mathbf{P}^{K'_1}(\mathcal{A}') \le \mathbf{P}^{K'_2}(\mathcal{A}')$, as desired. □

Again, it would be very interesting to have a natural monotone coupling of $\mathbf{P}^{K_1}$ with $\mathbf{P}^{K_2}$. For some examples where this would be desirable, see Subsection 3.8.

Lemma 3.6 will follow from this folklore variant of a theorem of Dyck [16]:

Theorem 3.9. *Let X be a Polish space on which $\le$ is a partial ordering that is closed in $X \times X$. Let $\mathscr{U}$ be a collection of open increasing sets that generates the Borel subsets of X. Let $\mathscr{U}^*$ be the closure of $\mathscr{U}$ under countable intersections and countable unions. Suppose that for all $x, y \in X$, either $x \le y$ or there is $U \in \mathscr{U}$ and an open set $V \subset X$ such that $x \in U$, $y \in V$, and $U \cap V = \emptyset$. Then $\mathscr{U}^*$ equals the class of increasing Borel sets.*

Proof. Obviously every set in $\mathscr{U}^*$ is Borel and increasing. To show the converse, we prove a variant of Lusin's separation theorem. Namely, we show that if $W_1 \subset X$ is increasing and analytic (with respect to the paving of closed sets, as usual) and if $W_2 \subset X$ is analytic with $W_1 \cap W_2 = \emptyset$, then there exists $U \in \mathscr{U}^*$ such that $W_1 \subseteq U$ and $U \cap W_2 = \emptyset$. Taking W_1 to be Borel and $W_2 := X \setminus W_1$ forces $U = W_1$ and gives the desired conclusion.

To prove this separation property, we first show a stronger conclusion in a special case: Suppose that $A_1, A_2 \subset X$ are compact such that A_1 is contained in an increasing set W_1 that is disjoint from A_2; then there exists an open $U \in \mathscr{U}^*$ and an open V such that $A_1 \subseteq U$, $A_2 \subseteq V$, and $U \cap V = \emptyset$. Indeed, since W_1 is increasing, for every $(x, y) \in A_1 \times A_2$, we do *not* have that $x \le y$, whence by hypothesis, there exist $U_{x,y} \in \mathscr{U}$ and an open $V_{x,y}$ with $x \in U_{x,y}$, $y \in V_{x,y}$, and $U_{x,y} \cap V_{x,y} = \emptyset$. Because A_2 is compact, for each $x \in A_1$, we may choose $y_1, \ldots, y_n \in A_2$ such that $A_2 \subseteq V_x := \bigcup_{i=1}^n V_{x,y_i}$. Define $U_x := \bigcap_{i=1}^n U_{x,y_i}$. Then U_x is open, contains x, and is disjoint from V_x, whence compactness of A_1 ensures the existence of $x_1, \ldots, x_m \in A_1$ with $A_1 \subseteq U := \bigcup_{j=1}^m U_{x_j} \in \mathscr{U}^*$. Then $V := \bigcap_{j=1}^m V_{x_j}$ is open, contains A_2, and is disjoint from U, as desired.

To prove the general case, let π_1 and π_2 be the two coordinate projections on $X^2 = X \times X$. Define $I(A) = I\big(\pi_1(A) \times \pi_2(A)\big)$ for $A \subseteq X^2$ to be 0 if there exists $U \in \mathscr{U}^*$ such that $\pi_1(A) \subseteq U$ and $U \cap \pi_2(A) = \emptyset$; and to be 1 otherwise.

We claim that I is a capacity in the sense of [29, (30.1)]. It is obvious that $I(A) \le I(B)$ if $A \subseteq B$ and it is simple to check that if $A_1 \subseteq A_2 \subseteq \cdots$, then $\lim_{n\to\infty} I(A_n) = I\big(\bigcup_n A_n\big)$. Suppose for the final property that A is compact and $I(A) = 0$; we must find an open $B \supseteq A$ for which $I(B) = 0$. There exists some $W_1 \in \mathscr{U}^*$ with $\pi_1(A) \subseteq W_1$ and $W_1 \cap \pi_2(A) = \emptyset$. Then the result of the second paragraph yields sets U and V that give $B := U \times V$ as desired.

Now let W_1 and W_2 be as in the first paragraph. If $A \subseteq W_1 \times W_2$ is compact, then setting $A_i := \pi_i(A)$ and applying the second paragraph shows that $I(A) = 0$. Thus, by the Choquet capacitability theorem [29, (30.13)], $I(W_1 \times W_2) = 0$. □

Proof of Lemma 3.6. Clearly every set in $\overline{\mathscr{U}(A)}$ is increasing and in $\mathscr{F}(A)$. For the converse, endow A with a metric so that it becomes locally compact Polish while preserving its class of relatively compact sets and its Borel σ-field: Choose a denumerable partition of A into relatively compact sets A_i and make each one compact and of diameter at most 1; make the distance between x and y be 1 if x and y belong to different A_i. Let $X := \mathcal{N}(A)$ with the vague topology and let $\mathscr{U}$ be the class of elementary increasing events defined with respect to (relatively compact) sets $B \subseteq A$ that are open for this new metric. Apply Theorem 3.9. Since $\mathscr{U}^* \subseteq \overline{\mathscr{U}(A)}$, the result follows. □

3.8. Example: Orthogonal Polynomial Ensembles.

Natural examples of determinantal point processes arise from orthogonal polynomials with respect to a probability measure μ on $\mathbb{C}$. Assume that μ has infinite support and finite moments of all orders. Let K_n denote the orthogonal projection of $L^2(\mathbb{C}, \mu)$ onto the linear span Poly_n of the functions $\{1, z, z^2, \ldots, z^{n-1}\}$. There exist unique (up to signum) polynomials ϕ_k of degree k such that for every n, $\langle \phi_k\,;\ 0 \le k < n\rangle$ is an orthonormal basis of Poly_n. By elementary row operations, we see that for variables $(z_1, \ldots, z_n)$, the map $(z_1, \ldots, z_n) \mapsto \det[\phi_i(z_j)]_{i,j\le n}$ is a Vandermonde polynomial up to a constant factor, whence

$$\det(K_n{\restriction}\{z_1, \ldots, z_n\}) = \det[\phi_i(z_j)][\phi_i(z_j)]^* = c_n \prod_{1\le i<j\le n} |z_i - z_j|^2$$

for some constant c_n. Therefore, the density of $\mathbf{P}^{K_n}$ (with points randomly ordered) with respect to μ^n is given by $c_n/n!$ times the square of a Vandermonde determinant.

Classical examples include the following:

OPE1. If μ is Gaussian measure on $\mathbb{R}$, i.e., $d\mu(x) = (2\pi)^{-1/2} e^{-x^2/2}\,dx$, then ϕ_k are the Hermite polynomials, $c_n = \big(\prod_{j=1}^{n-1} j!\big)^{-1}$, and $\mathbf{P}^{K_n}$ is the law of the ***Gaussian unitary ensemble***, which is the set of eigenvalues of $(\mathfrak{M} + \mathfrak{M}^*)/\sqrt{2}$, where $\mathfrak{M}$ is an $n \times n$ matrix whose entries are independent standard complex Gaussian. (A standard complex Gaussian random variable is the same as a standard Gaussian vector in $\mathbb{R}^2$ divided by $\sqrt{2}$ in order that the complex variance equal 1. Its density is $\pi^{-1} e^{-|z|^2}$ with respect to Lebesgue measure on $\mathbb{C}$.) This is due to Wigner; see [40].

OPE2. If μ is unit Lebesgue measure on the unit circle $\{z;\ |z| = 1\}$, then $\phi_k(z) = z^k$, so $c_n = 1$, and $\mathbf{P}^{K_n}$ is the law of the ***circular unitary ensemble***, which is the set of eigenvalues of a random matrix whose distribution is Haar measure on the set of $n \times n$ unitary matrices. This ensemble was introduced by Dyson, but the law of the eigenvalues is due to Weyl; see [27].

OPE3. If μ is standard Gaussian measure on $\mathbb{C}$, then $\phi_k(z) = z^k/\sqrt{k!}$, $c_n = \big(\prod_{j=1}^{n-1} j!\big)^{-1}$, and $\mathbf{P}^{K_n}$ is the law of the n***th (complex) Ginibre process***, which is the set of eigenvalues of an $n \times n$ matrix whose entries are independent standard complex Gaussian. This is due to Ginibre; see [27].

OPE4. If μ is unit Lebesgue measure on the unit disk $\mathbb{D} := \{z;\ |z| < 1\}$, then $\phi_k(z) = \sqrt{k+1}\,z^k$, so $c_n = n!$, and the limit of $\mathbf{P}^{K_n}$ is the law of the zero set of the random power series whose coefficients are independent standard complex Gaussian, which converges in the unit disk a.s. This is due to Peres and Virág [45].

OPE5. If μ has density $z \mapsto n\pi^{-1}\big(1 + |z|^2\big)^{-n-1}$ with respect to Lebesgue measure on $\mathbb{C}$, then $\phi_k(z) = \sqrt{\binom{n-1}{k}}z^k$ for $k < n$, so $c_n = \prod_{j=1}^{n-1}\binom{n-1}{j}$, and $\mathbf{P}^{K_n}$ is the law of the n***th spherical ensemble***, which is the set of eigenvalues of $\mathfrak{M}_1^{-1}\mathfrak{M}_2$ when $\mathfrak{M}_i$ are independent $n \times n$ matrices whose entries are independent standard complex Gaussian. (Here, we are limited to Poly_n since the larger spaces do not lie in $L^2(\mu)$.) This is due to Krishnapur [31]; see [27]. The process was studied earlier by [13] and [18], but without observing the connection to eigenvalues. Inverting stereographic projection, we identify this process with one whose density with respect to Lebesgue measure on the unit sphere in $\mathbb{R}^3$ is proportional to $\prod_{1 \le i < j \le n} \|\mathbf{v}_i - \mathbf{v}_j\|^2$.

For additional information on such processes, see [50, 23, 47, 17]. For an extension to complex manifolds, see [3, 4, 5].

By Theorem 3.8, the processes $\mathbf{P}^{K_n}$ stochastically increase in n for each of the examples above except the last. It would be interesting to see natural monotone couplings. Perhaps the last example also increases stochastically in n.

The ***Ginibre process*** is the limit of the nth Ginibre processes as $n \to \infty$; it has the kernel $e^{z\bar{w}}$ with respect to standard Gaussian measure on $\mathbb{C}$. This process is invariant under all isometries of $\mathbb{C}$. For each of the plane, sphere, and hyperbolic disk, there is only a 1-parameter family of determinantal point processes having a kernel $K(z, w)$ that is holomorphic in z and in $\bar{w}$ and whose law is isometry invariant [31, Theorem 3.0.5]. For the sphere, that family has already been given above; the parameter is a positive integer. For the other two families, the parameter is a positive real number, α. In the case of the plane, the processes are related simply by homotheties, $M_\alpha \colon z \mapsto z/\alpha$. The push-forward of the Ginibre process with respect to $M_{\sqrt{\alpha}}$ has kernel $e^{\alpha z\bar{w}}$ with respect to the measure $\alpha\pi^{-1}e^{-\alpha|z|^2}d\mu(z)$, where μ is Lebesgue measure on $\mathbb{C}$. Do these processes increase stochastically in α, like Poisson processes do? In the hyperbolic disk, the processes have kernel $\alpha(1 - z\bar{w})^{-\alpha-1}$ with respect to the measure $\pi^{-1}(1 - |z|^2)^{\alpha-1}d\mu(z)$, where μ is Lebesgue measure on $\mathbb{D}$. (We fix a branch of $(1 - z)^{-\alpha-1}$ for $z \in \mathbb{D}$.) These give orthogonal projections onto the generalized Bergman spaces. The case $\alpha = 1$ is that of the limiting OPE4 above. Do these processes stochastically increase in α?

4. Completeness

Recall that when H is a finite-dimensional subspace of $\ell^2(E)$, the measure $\mathbf{P}^H$ is supported by those subsets $B \subseteq E$ that project to a basis of H under P_H. Similarly, when K is the kernel of a finite-rank orthogonal projection onto $H \subset L^2(E, \mu)$, define the functions $K_x := K(\cdot, x) = \sum_{k \ge 1} \overline{\phi_k(x)} \phi_k \in H$. Then the measure $\mathbf{P}^K$ is supported by those ξ such that $\langle K_x \,;\; x \in \xi \rangle$ is a basis of H, since $K(x, y) = (K_y, K_x)$. Here, $x \in \xi$ means that $\xi(\{x\}) = 1$.

The question of extending this to infinite-dimensional H turns out to be very interesting. A basis of a finite-dimensional vector space is a minimal spanning set. Although $P_H \mathfrak{B}$ is $\mathbf{P}^H$-a.s. linearly independent, minimality does not hold in general, even for the wired spanning forest of a tree, as shown by the examples in [24]. See also Corollary 4.5. However, the other half of being a basis does hold in the discrete case and is open in the continuous case.

4.1. Discrete Completeness.

Let $[V]$ be the closed linear span of $V \subseteq \ell^2(E)$.

Theorem 4.1 ([33]). *For every $H \le \ell^2(E)$, we have $[P_H \mathfrak{B}] = H$ $\mathbf{P}^H$-a.s.*

We give an application of Theorem 4.1 for $E = \mathbb{Z}$, but it has an analogous statement for every countable abelian group. Let $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ be the unit circle equipped with unit Lebesgue measure. For a measurable function $f \colon \mathbb{T} \to \mathbb{C}$ and $n \in \mathbb{Z}$, the ***Fourier coefficient*** of f at n is $\widehat{f}(n) := \int_{\mathbb{T}} f(t) e^{-2\pi i n t}\, dt$. Let $\widehat{f} \restriction S$ denote the restriction of $\widehat{f}$ to S. If $A \subseteq \mathbb{T}$ is measurable, we say $S \subseteq \mathbb{Z}$ is ***complete for*** A if the set $\{ f \mathbf{1}_A \,;\; f \in L^2(\mathbb{T}),\, \widehat{f} \restriction (\mathbb{Z} \setminus S) \equiv 0 \}$ is dense in $L^2(A)$, where we identify $L^2(A)$ with the set of functions in $L^2(\mathbb{T})$ that vanish outside A. The case where A is an interval is quite classical; see [46] for a review. A crucial role in that case is played by the following notion of density of S.

Definition 4.2. *For an interval $[a, b] \subset \mathbb{R} \setminus \{0\}$, define its* ***aspect***

$$\alpha\big([a, b]\big) := \max\big\{|a|, |b|\big\} / \min\big\{|a|, |b|\big\} .$$

For a discrete $S \subseteq \mathbb{R}$, the ***Beurling-Malliavin density*** *of S, denoted $\mathsf{BM}(S)$, is the supremum of those $D \ge 0$ for which there exist disjoint nonempty intervals $I_n \subset \mathbb{R} \setminus \{0\}$ with $|S \cap I_n| \ge D |I_n|$ for all n and $\sum_{n \ge 1} [\alpha(I_n) - 1]^2 = \infty$.*

Corollary 4.3 ([33]). *Let $A \subset \mathbb{T}$ be Lebesgue measurable with measure $|A|$. Then there is a set of Beurling-Malliavin density $|A|$ in $\mathbb{Z}$ that is complete for A. Indeed, let $\mathbf{P}^A$ be the determinantal probability measure on $2^{\mathbb{Z}}$ corresponding to the Toeplitz matrix $(j, k) \mapsto \widehat{\mathbf{1}_A}(k - j)$. Then $\mathbf{P}^A$-a.e. $S \subset \mathbb{Z}$ is complete for A and has $\mathsf{BM}(S) = |A|$.*

When A is an interval, the celebrated theorem of Beurling and Malliavin [6] says that if S is complete for A, then $\mathsf{BM}(S) \ge |A|$, and that if $\mathsf{BM}(S) > |A|$, then S is complete for A. (This holds for S that are not necessarily sets of integers, but we are concerned in this subsection only with $S \subseteq \mathbb{Z}$.)

Corollary 4.3 can be compared (take $\mathbb{T} \setminus A$ and $\mathbb{Z} \setminus S$) to a theorem of Bourgain and Tzafriri [9], according to which there is a set $S \subset \mathbb{Z}$ of (Schnirelman) density at least $2^{-8}|A|$

such that if $f \in L^2(\mathbb{T})$ and $\widehat{f}$ vanishes off S, then

$$|A|^{-1} \int_A |f(t)|^2 \, dt \geq 2^{-16} \|f\|_2^2 \, .$$

It would be interesting to find a quantitative strengthening of Corollary 4.3 that would encompass this theorem of [9].

The following theorem is equivalent to Theorem 4.1 by duality:

Theorem 4.4 ([33]). *For every $H \leq \ell^2(E)$, we have $\overline{P_{[\mathfrak{B}]} H} = [\mathfrak{B}]$ $\mathbf{P}^H$-a.s.*

As an example, consider the wired spanning forest of a graph, G. Here, $H := \bigstar(G)$. In this case, $H_B := \overline{P_{[B]} \bigstar(G)} = \bigstar(B)$ for $B \subseteq E$. Thus, the conclusion of Theorem 4.4 is that $\mathbf{P}^{H_{\mathfrak{F}}}$, which equals $\mathsf{WSF}_{\mathfrak{F}}$, is concentrated on the singleton $\{\mathfrak{F}\}$ for WSF_G-a.e. $\mathfrak{F}$. This was a conjecture of [2], established by [42].

Corollary 4.5. *For every $H \leq \ell^2(E)$, $\mathbf{P}^H$-a.s. the maps $P_H \colon [\mathfrak{B}] \to H$ and $P_{[\mathfrak{B}]} \colon H \to [\mathfrak{B}]$ are injective with dense image.*

Proof. Both statements are equivalent to $[\mathfrak{B}] \cap H^\perp = \{0\} = H \cap \mathfrak{B}^\perp$, and these are the contents of Theorems 4.1 and 4.4. ☐

4.2. Continuous Completeness.

If K is a locally trace-class orthogonal projection onto H, then for $h \in H$, we have

$$h(x) = (Kh)(x) = \int_E K(x,y) h(y) \, d\mu(y) = \int_E h(y) \overline{K(y,x)} \, d\mu(y) = \big(h,\ K_x\big) \, .$$

In other words, K is a reproducing kernel for H. A subset S of H is called ***complete for*** H if the closed linear span of S equals H; equivalently, the only element of H that is orthogonal to S is 0.

An analogue of Theorem 4.1 was conjectured by Lyons and Peres in 2010:

Conjecture 4.6. *If K is a locally trace-class orthogonal projection onto H, then for $\mathbf{P}^K$-a.e. $\mathfrak{X}$, $[\{K_x \, ; \ x \in \mathfrak{X}\}] = H$, i.e., if $h \in H$ and $h{\restriction}\mathfrak{X} = 0$, then $h \equiv 0$.*

Just as in the discrete case, this appears to be on the critical border for many special instances, as we illustrate for several processes where $E = \mathbb{C}$:

1. Let μ be Lebesgue measure on $\mathbb{R}$ and $K(x,y) := \sin \pi(x-y)/\big(\pi(x-y)\big)$, the ***sine-kernel process***. Denote the Fourier transform on $\mathbb{R}$ by $\widehat{f}(t) := \int_{\mathbb{R}} f(x) e^{-2\pi i t x} \, dx$ for $f \in L^1(\mathbb{R})$, and, by isometric extension, for $f \in L^2(\mathbb{R})$. Write $I := \mathbf{1}_{[-1/2,1/2]}$. Since $K(x,0) = \widehat{I}(x)$, we have $(Kf)(x) = \big(f * \widehat{I}\big)(x) = \widehat{\check{f} I}(x)$, where $\check{f}$ is the inverse Fourier transform of f. Therefore, the induced operator K arises from the orthogonal projection onto the Paley-Wiener space $\{f \in L^2(\mathbb{R}, \mu) \, ; \ \check{f}(t) = 0 \text{ if } |t| > 1/2\}$. The sine-kernel process arises frequently; e.g., it is various scaling limits of the nth Gaussian unitary ensemble "in the bulk" as $n \to \infty$. (A related scaling limit of the GUE is Wigner's semicircle distribution.) We may more easily interpret Conjecture 4.6 for

Fourier transforms of functions in $L^2[-1/2, 1/2]$: It says that for $\mathbf{P}^K$-a.e. $\mathfrak{X}$, the only $h \in L^2[-1/2, 1/2]$ such that $\widehat{h}\restriction\mathfrak{X} = 0$ is $h \equiv 0$. Although the Beurling-Malliavin theorem applies, no information can be deduced because $\mathsf{BM}(\mathfrak{X}) = 1$ a.s. However, Ghosh [20] has proved this case.

2. Let μ be standard Gaussian measure on $\mathbb{C}$ and $K(z, w) := e^{z\bar{w}}$. This is the Ginibre process. It corresponds to orthogonal projection onto the ***Bargmann-Fock space*** $B^2(\mathbb{C})$ consisting of the entire functions that lie in $L^2(\mathbb{C}, \mu)$; this is the space of power series $\sum_{n \ge 0} a_n z^n$ such that $\sum_n n!|a_n|^2 < \infty$. Completeness of a set of elements $\{e^{\lambda z}\,;\ \lambda \in \Lambda\} \subset B^2(\mathbb{C})$ in $B^2(\mathbb{C})$ is equivalent to completeness in $L^2(\mathbb{R})$ (with Lebesgue measure) of the Gabor system of windowed complex exponentials

$$\left\{t \mapsto \exp\left[-i\,\mathrm{Im}\,\lambda t - (t - \mathrm{Re}\,\lambda)^2\right];\ \lambda \in \sqrt{2}\Lambda\right\},$$

which is used in time-frequency analysis of non-band-limited signals. The equivalence is proved using the Bargmann transform

$$f \mapsto \left(z \mapsto \pi^{-1/4} \int_{\mathbb{R}} f(t) \exp\left[\sqrt{2}tz - \frac{z^2}{2} - \frac{t^2}{2}\right] dt\right),$$

which is an isometry from $L^2(\mathbb{R})$ to $B^2(\mathbb{C})$. That the critical density is 1 was shown in various senses going back to von Neumann; see [14]. This case has also been proved by Ghosh [20].

3. Let μ be unit Lebesgue measure on the unit disk $\mathbb{D} := \{z\,;\ |z| < 1\}$ and $K(z, w) := (1 - z\bar{w})^{-2}$. This process is the limiting OPE4 in Subsection 3.8. It corresponds to orthogonal projection onto the ***Bergman space*** $A^2(\mathbb{D})$ consisting of the analytic functions that lie in $L^2(\mathbb{D}, \mu)$. What is known about the zero sets of functions in the Bergman space [15] is insufficient to settle Conjecture 4.6 in this case and it remains open.

The two instances above that have been proved by Ghosh [20] follow from his more general result that Conjecture 4.6 holds whenever μ is continuous and $\mathbf{P}^K$ is ***rigid***, which means that $\mathfrak{X}(B)$ is measurable with respect to the $\mathbf{P}^K$-completion of $\mathscr{F}(E \setminus B)$ for every ball $B \subset E$. The limiting process OPE4 is not rigid [25]. Ghosh and Krishnapur (personal communication, 2014) have shown that $\mathbf{P}^K$ is rigid only if K is an orthogonal projection. It is not sufficient that K be a projection, as the example of the Bergman space shows. A necessary and sufficient condition to be rigid is not known.

Let K be a locally trace-class orthogonal projection onto $H \le L^2(E, \mu)$. For a function f, write f_K for the function $f(x)/\sqrt{K(x, x)}$. Let $\mathfrak{X} \sim \mathbf{P}^K$. Clearly $f_K\restriction\mathfrak{X} \in \ell^2(\mathfrak{X})$ for a.e. $\mathfrak{X}$. Also, for $h \in H$, the function h_K is bounded. A conjecture analogous to Corollary 4.5 is that $\mathfrak{X}$ is a sort of set of interpolation for H in the sense that given any countable dense set $H_0 \subset H$, for a.e. $\mathfrak{X}$, the set $\{h_K\restriction\mathfrak{X}\,;\ h \in H_0\}$ is dense in $\ell^2(\mathfrak{X})$.

One may also ask about completeness for appropriate Poisson point processes.

5. Discrete Invariance

Suppose Γ is a group that acts on E and that K is Γ-invariant, i.e., $K(\gamma x, \gamma y) = K(x, y)$ for all $\gamma \in \Gamma$, $x \in E$, and $y \in E$. (This is equivalent to the operator K being Γ-equivariant.) Then the probability measure $\mathbf{P}^K$ is Γ-invariant. This contact with ergodic theory and other areas of mathematics suggests many interesting questions. Lack of space prevents us from considering more than just a few aspects of the case where E is discrete and from giving all definitions.

5.1. Integer Lattices. Let $E := \Gamma := \mathbb{Z}^d$. In this case, K is invariant iff $K(m, n) = \widehat{f}(n - m)$ for some $f\colon \mathbb{T}^d \to [0, 1]$, where $\widehat{f}(n) := \int_{\mathbb{T}^d} f(t) e^{-2\pi i n \cdot t}\, dt$. We write $\mathbf{P}^f$ in place of $\mathbf{P}^K$. Some results and questions from [37] follow.

Theorem 5.1. *For all f, the process $\mathbf{P}^f$ is isomorphic to a Bernoulli process.*

This was shown in dimension 1 by [49] for those f such that $\sum_{n \ge 1} n |\widehat{f}(n)|^2 < \infty$ by showing that those $\mathbf{P}^f$ are weak Bernoulli (WB), also called "β-mixing" and "absolutely regular". Despite its name, it is known that WB is strictly stronger than Bernoullicity. The precise class of f for which $\mathbf{P}^f$ is WB is not known.

As usual, the ***geometric mean*** of a nonnegative function f is $\mathsf{GM}(f) := \exp \int \log f$.

Theorem 5.2. *For all f, the process $\mathbf{P}^f$ stochastically dominates product measure $\mathbf{P}^{\mathsf{GM}(f)}$ and is stochastically dominated by product measure $\mathbf{P}^{1-\mathsf{GM}(\mathbf{1}-f)}$. These bounds are optimal.*

We conjecture that (Kolmogorov-Sinai) entropy is concave, as would follow from Conjecture 2.6.

Conjecture 5.3. *For all f and g, we have $H\big(\mathbf{P}^{(f+g)/2}\big) \ge \big(H(\mathbf{P}^f) + H(\mathbf{P}^g)\big)/2$.*

Question 5.4. *Let $f\colon \mathbb{T} \to [0, 1]$ be a trigonometric polynomial of degree m. Then $\mathbf{P}^f$ is m-dependent, as are all $(m+1)$-block factors of independent processes. Is $\mathbf{P}^f$ an $(m+1)$-block factor of an i.i.d. process? This is known when $m = 1$ [10].*

5.2. Sofic Groups. Let Γ be a sofic group, a class of groups that includes all finitely generated amenable groups and all finitely generated residually amenable groups. No finitely generated group is known not to be sofic. Let E be Γ or, more generally, a set acted on by Γ with finitely many orbits, such as the edges of a Cayley graph of Γ. The following theorems are from [38].

Theorem 5.5. *For every Γ-equivariant positive contraction Q on $\ell^2(E)$, the process $\mathbf{P}^Q$ is a $\bar{d}$-limit of finitely dependent (invariant) processes. If Γ is amenable and $E = \Gamma$, then $\mathbf{P}^Q$ is isomorphic to a Bernoulli process.*

Even if $\mathbf{P}^1$ and $\mathbf{P}^2$ are Γ-invariant probability measures on 2^Γ with $\mathbf{P}^1 \preccurlyeq \mathbf{P}^2$, there need not be a Γ-invariant monotone coupling of $\mathbf{P}^1$ and $\mathbf{P}^2$ [41]. The proof of the preceding theorem depends on the next one:

Theorem 5.6. *If Q_1 and Q_2 are two Γ-equivariant positive contractions on $\ell^2(E)$ with $Q_1 \preceq Q_2$, then there exists a Γ-invariant monotone coupling of $\mathbf{P}^{Q_1}$ and $\mathbf{P}^{Q_2}$.*

The proof of Theorem 5.5 also uses the inequality

$$\bar{d}\big(\mathbf{P}^{Q}, \mathbf{P}^{Q'}\big) \le 6 \cdot 3^{2/3} \|Q - Q'\|_1^{1/3}$$

for equivariant positive contractions, Q and Q', where $\|T\|_1 := \operatorname{tr}(T^*T)^{1/2}$ is the Schatten 1-norm. When Q and Q' commute, one can improve this bound to

$$\bar{d}(\mathbf{P}^{Q}, \mathbf{P}^{Q'}) \le \|Q - Q'\|_1 \,.$$

We do not know whether this inequality always holds.

Write $\mathsf{FK}(Q) := \exp \operatorname{tr} \log |Q|$ for the Fuglede-Kadison determinant of Q when Q is a Γ-equivariant operator. The following would extend Theorem 5.2. It is open even for finite groups.

Conjecture 5.7. *For all Γ-equivariant positive contractions Q on $\ell^2(\Gamma)$, the process $\mathbf{P}^Q$ stochastically dominates product measure $\mathbf{P}^{\mathsf{FK}(Q)I}$ and is stochastically dominated by product measure $\mathbf{P}^{I-\mathsf{FK}(I-Q)I}$, and these bounds are optimal.*

5.3. Isoperimetry, Cost, and ℓ^2-Betti Numbers.

It turns out that the expected degree of a vertex in the free uniform spanning forest of a Cayley graph depends only on the group, via its first ℓ^2-Betti number, $\beta_1(\Gamma)$, and not on the generating set used to define the Cayley graph [34]:

Theorem 5.8. *In every Cayley graph G of a group Γ, we have*

$$\mathbf{E}_{\mathsf{FSF}(G)}[\deg_{\mathfrak{F}}(o)] = 2\beta_1(\Gamma) + 2 \,.$$

This is proved using the representation of FSF as a determinantal probability measure. It can be used to give a uniform bound on expansion constants [36]:

Theorem 5.9. *For every finite symmetric generating set S of a group Γ, we have $|SA \setminus A| > 2\beta_1(\Gamma)|A|$ for all finite non-empty $A \subset \Gamma$.*

There are extensions of these results to higher-dimensional CW-complexes and higher ℓ^2-Betti numbers [34].

In unpublished work with D. Gaboriau [35], we have shown the following:

Theorem 5.10. *Let G be a Cayley graph of a finitely generated group Γ and $\epsilon > 0$. Then there exists a Γ-invariant finitely dependent determinantal probability measure $\mathbf{P}^Q$ on $\{0,1\}^{\mathsf{E}(G)}$ that stochastically dominates FSF_G and such that*

$$\mathbf{E}^Q\big[\deg_{\mathfrak{G}}(o)\big] \le \mathbf{E}_{\mathsf{FSF}}\big[\deg_{\mathfrak{F}}(o)\big] + \epsilon \,.$$

In addition, if Γ is sofic, then $\bar{d}(\mathbf{P}^Q, \mathsf{FSF}) \le \epsilon$.

If it could be shown that $\mathbf{P}^Q$, or indeed every invariant finitely dependent probability measure that dominates FSF, yields a connected subgraph a.s., then it would follow that $\beta_1(\Gamma) + 1$ is equal to the cost of Γ, a major open problem of [19].

Acknowledgments

I am grateful to Alekos Kechris for informing me of Theorem 3.9; the proof given seems to be due to Alain Louveau. I thank Norm Levenberg for references.

References

[1] Bapat, R.B., Mixed discriminants and spanning trees, *Sankhyā* Ser. A **54** (1992), no. Special Issue, 49–55, Combinatorial mathematics and applications (Calcutta, 1988).

[2] Benjamini, I., Lyons, R., Peres, Y., and Schramm, O., Uniform spanning forests. *Ann. Probab.* **29** (2001), 1–65.

[3] Berman, R.J., Determinantal point processes and fermions on complex manifolds: Bulk universality. Preprint, http://www.arxiv.org/abs/0811.3341, (2008).

[4] Berman, R.J., Determinantal point processes and fermions on complex manifolds: large deviations and bosonization. Preprint, http://www.arxiv.org/abs/0812.4224, (2008).

[5] Berman, R.J., Sharp asymptotics for Toeplitz determinants. *Int. Math. Res. Not. IMRN* **2012** (2012), 22, 5031–5062.

[6] Beurling, A. and Malliavin, P., On the closure of characters and the zeros of entire functions. *Acta Math.* **118** (1967), 79–93.

[7] Borcea, J., Brändén, P., and Liggett, T.M., Negative dependence and the geometry of polynomials. *J. Amer. Math. Soc.* **22** (2009), 521–567.

[8] Borodin, A., Determinantal point processes. *The Oxford Handbook of Random Matrix Theory*, 231–249, Oxford Univ. Press, Oxford, 2011.

[9] Bourgain, J. and Tzafriri, L., Invertibility of "large" submatrices with applications to the geometry of Banach spaces and harmonic analysis. *Israel J. Math.* **57** (1987), 2, 137–224.

[10] Broman, E., One-dependent trigonometric determinantal processes are two-block-factors. *Ann. Probab.* **33** (2005), 2, 601–609.

[11] Brooks, R.L., Smith, C.A.B., Stone, A.H., and Tutte, W.T., The dissection of rectangles into squares. *Duke Math. J.* **7** (1940), 312–340.

[12] Burton, R.M. and Pemantle, R., Local characteristics, entropy and limit theorems for spanning trees and domino tilings via transfer-impedances. *Ann. Probab.* **21** (1993), 1329–1371.

[13] Caillol, J.M., Exact results for a two-dimensional one-component plasma on a sphere. *J. Physique - LETTRES* **42** (1981), L-245–L-247.

[14] Chistyakov, G., Lyubarskii, Yu. and Pastur, L., On completeness of random exponentials in the Bargmann-Fock space, *J. Math. Phys.* **42** (2001), 8, 3754-3768.

[15] Duren, P. and Schuster, A., *Bergman Spaces*. Mathematical Surveys and Monographs, **100**. American Mathematical Society, Providence, RI, 2004.

[16] Dyck, S., Some applications of positive formulas in descriptive set theory and logic. *Ann. Pure Appl. Logic* **46** (1990), 2, 95–146.

[17] Forrester, P.J., *Log-Gases and Random Matrices*. London Mathematical Society Monographs Series, **34**. Princeton University Press, Princeton, NJ, 2010.

[18] Forrester, P.J., Jancovici, B., and Madore, J., The two-dimensional Coulomb gas on a sphere: exact results. *J. Statist. Phys.* **69** (1992), 1–2, 179–192.

[19] Gaboriau, D., Invariants ℓ^2 de relations d'équivalence et de groupes. *Publ. Math. Inst. Hautes Études Sci.*, **95** (2002), 93–150.

[20] Ghosh, S., Determinantal processes and completeness of random exponentials: the critical case. Preprint, http://www.arxiv.org/abs/1211.2435, (2012).

[21] Goldman, A., The Palm measure and the Voronoi tessellation for the Ginibre process. *Ann. Appl. Probab.* **20** (2010), 1, 90–128.

[22] Häggström, O., Random-cluster measures and uniform spanning trees. *Stochastic Process. Appl.* **59** (1995), 267–275.

[23] Hardy, A., Average characteristic polynomials of determinantal point processes. Preprint, http://www.arxiv.org/abs/1211.6564, (2012).

[24] Heicklen, D. and Lyons, R. Change intolerance in spanning forests. *J. Theoret. Probab.* **16** (2003), 47–58.

[25] Holroyd, A.E. and Soo, T., Insertion and deletion tolerance of point processes. *Electron. J. Probab.* **18** (2013), 74, 24 pp.

[26] Hough, J.B., Krishnapur, M., Peres, Y., and Virág, B., Determinantal processes and independence. *Probab. Surv.* **3** (2006), 206–229.

[27] Hough, J.B., Krishnapur, M., Peres, Y., and Virág, B., *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*. University Lecture Series, **51**. American Mathematical Society, Providence, RI, 2009.

[28] Johansson, K., Random matrices and determinantal processes. *Mathematical Statistical Physics*, 1–55, Elsevier B. V., Amsterdam, 2006.

[29] Kechris, A.S., *Classical Descriptive Set Theory*. Graduate Texts in Mathematics, **156**. Springer-Verlag, New York, 1995.

[30] Kirchhoff, G., Ueber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird. *Ann. Phys. und Chem.* **72** (1847), 497–508.

[31] Krishnapur, M., *Zeros of Random Analytic Functions*. Ph.D. Thesis, U.C. Berkeley, 2006. http://www.arxiv.org/abs/math/0607504.

[32] Kulesza, A. and Taskar, B., Determinantal point processes for machine learning, *Foundations and Trends in Machine Learning*, **5** (2012), 2–3, 123–286. DOI: 10.1561/2200000044.

[33] Lyons, R., Determinantal probability measures. *Publ. Math. Inst. Hautes Études Sci.* **98** (2003), 167–212. Errata, http://pages.iu.edu/~rdlyons/errata/bases.pdf.

[34] Lyons, R., Random complexes and ℓ^2-Betti numbers. *J. Topology Anal.* **1** (2009), 2, 153–175.

[35] Lyons, R. and Gaboriau, D., An approach to the cost vs. ℓ^2-Betti-numbers problem. In preparation.

[36] Lyons, R., Pichot, M., and Vassout, S., Uniform non-amenability, cost, and the first ℓ^2-Betti number. *Groups Geom. Dyn.* **2** (2008), 4, 595–617.

[37] Lyons, R. and Steif, J.E., Stationary determinantal processes: Phase multiplicity, Bernoullicity, entropy, and domination. *Duke Math. J.*, **120** (2003), 3, 515–575.

[38] Lyons, R. and Thom, A., Invariant coupling of determinantal measures on sofic groups. *Ergodic Theory Dynam. Systems*, to appear. Preprint, http://www.arxiv.org/abs/1402.0969, (2014).

[39] Macchi, O., The coincidence approach to stochastic point processes, *Advances in Appl. Probability* **7** (1975), 83–122.

[40] Mehta, M., *Random Matrices*. Third edition. Pure and Applied Mathematics (Amsterdam), **142**. Elsevier/Academic Press, Amsterdam, 2004.

[41] Mester, P., Invariant monotone coupling need not exist. *Ann. Probab.* **41** (2013), 3A, 1180–1190.

[42] Morris, B., The components of the wired spanning forest are recurrent. *Probab. Theory Related Fields* **125** (2003), 259–265.

[43] Paulsen, V., *Completely Bounded Maps and Operator Algebras*. Cambridge Studies in Advanced Mathematics, **78**. Cambridge University Press, Cambridge, 2002.

[44] Pemantle, R., Choosing a spanning tree for the integer lattice uniformly. *Ann. Probab.* **19** (1991), 1559–1574.

[45] Peres, Y. and Virág, B., Zeros of the i.i.d. Gaussian power series: a conformally invariant determinantal process. *Acta Math.* **194** (2005), 1, 1–35.

[46] Redheffer, R., Completeness of sets of complex exponentials. *Advances in Math.* **24** (1977), 1–62.

[47] Rider, B. and Virág, B., Complex determinantal processes and H^1 noise. *Electron. J. Probab.* **12** (2007), 45, 1238–1257.

[48] Shirai, T. and Takahashi, Y., Random point fields associated with certain Fredholm determinants. I. Fermion, Poisson and boson point processes. *J. Funct. Anal.* **205** (2003), 414–463.

[49] Shirai, T. and Takahashi, Y. Random point fields associated with certain Fredholm determinants II: fermion shifts and their ergodic and Gibbs properties. *Ann. Probab.* **31** (2003), 1533–1564.

[50] Simon, B., Weak convergence of CD kernels and applications. *Duke Math. J.* **146** (2009), 2, 305–330.

[51] Soshnikov, A., Determinantal random point fields. *Uspekhi Mat. Nauk* **55** (2000), 107–160.

[52] Soshnikov, A., Gaussian limit for determinantal random point fields. *Ann. Probab.* **30** (2002), 1, 171–187.

[53] Strassen, V., The existence of probability measures with given marginals. *Ann. Math. Statist.* **36** (1965), 423–439.

Dept. of Math., Indiana University, 831 E. 3rd St., Bloomington, IN 47405-7106 USA
E-mail: rdlyons@indiana.edu