CrossMark

J. Appl. Prob. **58,** 469–483 (2021)
*doi:*10.1017/jpr.2020.101

COUPLINGS FOR DETERMINANTAL POINT PROCESSES AND THEIR REDUCED PALM DISTRIBUTIONS WITH A VIEW TO QUANTIFYING REPULSIVENESS

JESPER MØLLER,* *Aalborg University*
ELIZA O'REILLY,** *California Institute of Technology*

Abstract

For a determinantal point process (DPP) X with a kernel K whose spectrum is strictly less than one, André Goldman has established a coupling to its reduced Palm process X^u at a point u with $K(u, u) > 0$ so that, almost surely, X^u is obtained by removing a finite number of points from X. We sharpen this result, assuming weaker conditions and establishing that X^u can be obtained by removing at most one point from X, where we specify the distribution of the difference $\xi_u := X \setminus X^u$. This is used to discuss the degree of repulsiveness in DPPs in terms of ξ_u, including Ginibre point processes and other specific parametric models for DPPs.

Keywords: Ginibre point process; isotropic determinantal point process on the sphere; globally most repulsive determinantal point process; projection kernel; stationary determinantal point process in Euclidean space

2010 Mathematics Subject Classification: Primary 60G55; 60J20
Secondary 60D05; 62M30

1. Introduction

Determinantal point processes (DPPs) have been of much interest over many years in mathematical physics and probability theory (see, e.g., [4, 10, 16, 26, 28] and the references therein), and more recently in other areas, including statistics [14, 19], machine learning [12], signal processing [5], and neuroscience [27]. They are models for regularity/inhibition/repulsiveness, but there is a trade-off between repulsiveness and intensity [13, 14] (or see Section 4.1.1). This paper sheds further light on this issue by studying various couplings between a DPP and its reduced Palm distributions. In particular, we relate our results to the definition of most repulsive stationary DPPs on $\mathbb{R}^d$ as specified in [2, 14]. However, our results will be given for DPPs in a more general setting.

Section 2.1.1 provides our general setting for a DPP X defined on a locally compact Polish space Λ and specified by a so-called kernel $K{:}\Lambda \times \Lambda \to \mathbb{C}$ which satisfies certain mild conditions given in Section 2.1.2. Also, for any $u \in \Lambda$ with $K(u, u) > 0$, if X^u follows the reduced

Received 21 March 2019; revision received 14 October 2020.
* Postal address: Department of Mathematical Sciences, Aalborg University, Skjernvej 4A, 9220 Aalborg Øst, Denmark.
** Postal address: Computing and Mathematical Sciences, California Institute of Technology, 1200 E. California Blvd. MC 305-16, Pasadena, CA 91125, USA. Email address: eoreilly@caltech.edu

© The Author(s), 2021. Published by Cambridge University Press on behalf of Applied Probability Trust.

https://doi.org/10.1017/jpr.2020.101 Published online by Cambridge University Press

Palm distribution of X at u—intuitively, this is the conditional distribution of $X \setminus \{u\}$ given that $u \in X$—then X^u is another DPP; Section 2.1.3 provides further details. Furthermore, Section 2.2 discusses Goldman's [8] result that if, for any compact set $S \subseteq \Lambda$, denoting by K_S the restriction of K to $S \times S$, we have that the spectrum of K_S is < 1, then X stochastically dominates X^u and hence, by Strassen's theorem, there exists a coupling so that, almost surely, $X^u \subseteq X$. The difference $\kappa_u := X \setminus X^u$ is a finite point process with a known intensity function. A straightforward calculation also gives that the mean number of points in κ_u is at most 1; see equations (3) and (11). In particular, for a standard Ginibre point process [7], which is a special case of a DPP in the complex plane, Goldman showed that κ_u consists of a single point which follows $N_{\mathbb{C}}(u, 1)$, the complex Gaussian distribution with mean u and unit variance. Section 2.3 then discusses some related coupling results due to Pemantle and Peres [24]. One of their results implies that in the specific case where Λ is finite, there exists a coupling of X and X^u such that $X^u \subseteq X$ and the difference is at most one point. However, apart from these and other special cases, the distribution of κ_u has not been fully characterized.

Section 3 shows that these results can be extended. For a DPP X on any locally compact Polish space Λ, there is a coupling such that $X^u \subseteq X$ almost surely, $\xi_u := X \setminus X^u$ consists of at most one point, and the distribution of ξ_u can be specified. Note that κ_u and ξ_u share the same intensity function, but Goldman did not establish that κ_u consists of at most one point. As in [8] we only verify the existence of our coupling result. We leave it as an open research problem to provide a specific coupling construction or simulation procedure for (X, X^u) (restricted to a compact subset of Λ), hence extending the simulation algorithm for a DPP [9, 13, 14, 19].

Section 4 discusses how our coupling result can be used for describing the repulsiveness in a DPP, including when considering the notion of a globally most repulsive DPP, by which we mean that, for all $u \in \Lambda$ with $K(u, u) > 0$, almost surely ξ_u has one point. For example, if X is a standard Ginibre point process, we obtain a similar result to [8]: X is a globally most repulsive DPP and the point in ξ_u follows $N_{\mathbb{C}}(u, 1)$. In particular, we show that our definitions extend those given in [2, 14] for stationary DPPs on $\mathbb{R}^d$, and by considering the distribution of the point in ξ_u, we demonstrate how the range of repulsion can differ with DPPs that have the same intensity and the same global repulsiveness. Moreover, we consider the cases of a finite set Λ and when we have a stationary DPP defined on $\Lambda = \mathbb{R}^d$. Finally, we compare with globally most repulsive isotropic DPPs on $\mathbb{S}^d$, the d-dimensional unit sphere in $\mathbb{R}^{d+1}$, as studied in [18].

2. Background

This section provides the background material needed in this paper.

2.1. Setting

Below we give the definition of a DPP, specify our assumptions, and recall that the reduced Palm distribution of a DPP is another DPP.

2.1.1. *Definition of a DPP.* Let X be a point process defined on a locally compact Polish space Λ equipped with its Borel σ-algebra $\mathcal{B}$ and a Radon measure ν, which is used as a reference measure in the following. We assume that X is a DPP with kernel K, which by definition means the following. First, X has no multiple points so, dependent on the context, we view X as a random subset of Λ or as a random counting measure, and we let $X(B)$ denote the cardinality

https://doi.org/10.1017/jpr.2020.101 Published online by Cambridge University Press

of $X_B := X \cap B$ for $B \in \mathcal{B}$. Second, K is a complex function defined on $K{:}\Lambda^2 \mapsto \mathbb{C}$. Third, for any $n = 1, 2, \ldots$ and any mutually disjoint bounded sets $B_1, \ldots, B_n \in \mathcal{B}$,

$$\mathbb{E}\left[X\left(B_1\right) \cdots X\left(B_n\right)\right] = \int_{B_1 \times \cdots \times B_n} \det\left\{K\left(u_i, u_j\right)\right\}_{i,j=1}^{n} \mathrm{d}\nu^n\left(u_1, \ldots, u_n\right)$$

is finite, where ν^n denotes the n-fold product measure of ν. This means that X has an nth-order intensity function $\rho(u_1, \ldots, u_n)$ (also sometimes in the literature called an nth-order correlation function) given by the determinant

$$\rho\left(u_1, \ldots, u_n\right) = \det\left\{K\left(u_i, u_j\right)\right\}_{i,j=1}^{n}, \qquad u_1, \ldots, u_n \in \Lambda, \tag{1}$$

and this function is locally integrable. In particular, $\rho(u) = K(u, u)$ is the intensity function of X, and when $B \in \mathcal{B}$ is bounded, almost surely X_B is finite.

In the special case where $K(u, v) = 0$ whenever $u \neq v$, the DPP X is just a Poisson process with intensity function $\rho(u)$ conditioned on there being no multiple points in X (if ν is diffuse, it is implicit that there are no multiple points). For other examples when Λ is a countable set and ν is the counting measure, see [12]; when $\Lambda = \mathbb{R}^d$ and ν is the Lebesgue measure, see [10, 14]; and when $\Lambda = \mathbb{S}^d$ (the d-dimensional unit sphere) and ν is the surface/Lebesgue measure, see [18]. Examples are also given in Section 4.2.

From (1) and the fact that the determinant of a complex covariance matrix is less than or equal to the product of its diagonal elements, we obtain that

$$\rho\left(u_1, \ldots, u_n\right) \leq \prod_{i=1}^{n} \rho\left(u_i\right),$$

where the equality holds if and only if X is a Poisson process. Thus, apart from the case of a Poisson process, the counts $X(A)$ and $X(B)$ are negatively correlated whenever $A, B \in \mathcal{B}$ are disjoint.

2.1.2 *Assumptions.* We always make the following assumptions.

Assumption 1. *K is Hermitian; that is, $K(u, v) = \overline{K(v, u)}$ for all $u, v \in \Lambda$.*

Assumption 2. *K is locally square integrable; that is, for any compact set $S \subseteq \Lambda$, the double integral $\int_S \int_S |K(u, v)|^2 \, \mathrm{d}\nu(u) \, \mathrm{d}\nu(v)$ is finite.*

Assumption 3. *K is locally trace class; that is, for any compact set $S \subseteq \Lambda$, the integral $\int_S K(u, u) \, \mathrm{d}\nu(u)$ is finite.*

By Mercer's theorem, excluding a ν^2-nullset, this ensures the existence of a spectral representation for the kernel restricted to any compact set $S \subseteq \Lambda$. Ignoring a ν^2-nullset, we can redefine K on $S \times S$ by

$$K(u, v) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(u) \overline{\phi_k^S(v)}, \qquad u, v \in S, \tag{2}$$

where the eigenvalues λ_k^S are real numbers and the eigenfunctions ϕ_k^S constitute an orthonormal basis of $L^2(S)$; cf. Section 4.2.1 in [10]. Here, for any $B \in \mathcal{B}$, $L^2(B) = L^2(B, \nu)$ is the space of square integrable complex functions with respect to (w.r.t.) ν restricted to B. Note that Assumption 3 means that $\mathbb{E}X(S) = \sum_{k=1}^{\infty} \lambda_k^S < \infty$. Thus, when $B \in \mathcal{B}$ is bounded, almost

https://doi.org/10.1017/jpr.2020.101 Published online by Cambridge University Press

surely X_B is finite. When ν is diffuse, as we are redefining K by (2) we have effectively excluded the special case of the Poisson process (i.e. when K is 0 off the diagonal) because all the eigenvalues in (2) are then 0; however, as shown later, it will still make sense to consider the Poisson process when quantifying repulsiveness in DPPs.

We also always assume the following.

Assumption 4. *For any compact set $S \subseteq \Lambda$, all eigenvalues satisfy $0 \le \lambda_k^S \le 1$.*

In fact, under Assumptions 1–3, the existence of the DPP with kernel K is equivalent to Assumption 4 (see, e.g., [10, Theorem 4.5.5]), and the DPP is then unique [10, Lemma 4.2.6]. If $\Lambda = \mathbb{R}^d$, ν is the Lebesgue measure, and $K(u, v) = K_0(u - v)$ is stationary, where $K_0 \in L^2(\mathbb{R}^d)$ and K_0 is continuous, we denote the Fourier transform of K_0 by $\hat{K}_0$. Then, Assumption 4 is equivalent to $0 \le \hat{K}_0 \le 1$ [10, Proposition 3.1].

Recalling that K_S is the restriction of K to $S \times S$, we sometimes consider one of the following conditions.

Condition 1. *For a given compact set $S \subseteq \Lambda$, K_S is a projection of finite rank n.*

Condition 2. *For all compact $S \subseteq \Lambda$, all eigenvalues satisfy $\lambda_k^S < 1$.*

2.1.3 *Reduced Palm distributions.* Consider an arbitrary point $u \in \Lambda$ with $\rho(u) > 0$. Recall that the reduced Palm distribution of X at u is a point process X^u on Λ with nth-order intensity function $\rho^u(u_1, \dots, u_n) = \rho(u, u_1, \dots, u_n)/\rho(u)$. This, combined with (1), easily shows that X^u is a DPP with kernel [26, Theorem 6.5]

$$K^u(v, w) = K(v, w) - \frac{K(v, u)K(u, w)}{K(u, u)}, \qquad v, w \in \Lambda. \tag{3}$$

For any compact set $S \subseteq \Lambda$, it follows that the restriction $X_S^u := X^u \cap S$ follows the reduced Palm distribution of X_S at u.

2.2. Goldman's results

Goldman [8] made similar assumptions to our Assumptions 1–4, and in addition he assumed Condition 2 throughout his paper. Two of his main results were the following.

(G1) For any $u \in \Lambda$ with $K(u, u) > 0$, there is a coupling of X and X^u so that, almost surely, $X^u \subseteq X$.

(G2) Suppose X is a standard Ginibre point process, that is, the DPP on $\Lambda = \mathbb{C} \equiv \mathbb{R}^2$, with ν being Lebesgue measure, and with kernel

$$K(v, w) = \frac{1}{\pi} \exp\left(v\overline{w} - \frac{|v|^2 + |w|^2}{2}\right), \qquad v, w \in \mathbb{C}. \tag{4}$$

Then, for the coupling in (G1) and any $u \in \mathbb{C}$, $X \setminus X^u$ consists of a single point which follows $N_{\mathbb{C}}(u, 1)$.

It follows from (G1) and (3) that $\kappa_u := X \setminus X^u$ is a finite point process with intensity function

$$\rho_{\kappa_u}(v) = |K(u, v)|^2/K(u, u), \qquad v \in \Lambda. \tag{5}$$

https://doi.org/10.1017/jpr.2020.101 Published online by Cambridge University Press

Note that the standard Ginibre point process is stationary and isotropic with intensity $1/\pi$, but its kernel is not of the form $K(u, v) = K_0(\|u - v\|)$. In accordance with (G2), combining (4) and (5), ρ_{κ_u} is immediately seen to be the density of $N_{\mathbb{C}}(u, 1)$.

2.3. Pemantle and Peres' results

Pemantle and Peres [24] studied probability measures on $\{0, 1\}^n$ satisfying a negative dependence property called the strong Rayleigh property. This class of probability measures was introduced in [3], where it was also shown that determinantal point processes on a finite set satisfy this property. In [24], the authors defined notions called stochastic covering and the stochastic covering property, which can be defined as follows. Letting X and Y be simple point processes, X is said to *stochastically cover* Y if there is a coupling (X,Y) such that $X = Y$ or their difference $X \setminus Y$ is one point. Now, consider a simple point process X on a finite set $\Lambda = \{1, \ldots, n\}$. Then, X is said to have the stochastic covering property if the following holds. If $u \in x \subseteq S \subseteq \Lambda$ and we set $y = x \setminus \{u\}$, then the point process X_{S^c} conditioned on $X_S = x$ is stochastically covered by the point process X_{S^c} conditioned on $X_S = y$. This property implies that, for $u \in \{1, \ldots, n\}$ (letting $S = x = \{u\}$), the point process $X \setminus \{u\}$ conditioned on $u \notin X$ stochastically covers X^u, and in turn that X stochastically covers X^u.

Proposition 2.2 in [24] states that, for a probability measure on $\{0, 1\}^n$, the strong Rayleigh property implies the stochastic covering property, and thus determinantal point processes on $\Lambda = \{1, \ldots, n\}$ satisfy this property. The authors discuss extensions to the case where Λ is continuous, and in particular they extend their Proposition 2.3 to this case. However, a generalization of their Proposition 2.2 to the case of continuous determinantal point processes does not appear in the most recent version [24]. As kindly pointed out by a referee, in the first version of this paper [23], the authors claim X stochastically covers X^u in the continuous case as well, and the main idea of our proof of Theorem 1 below is outlined. However, their justification is not complete for our general setting.

3. Main result

The theorem below is our main result, which sharpens Goldman's result (G1) in that we do not assume Condition 2 and we establish a coupling so that X stochastically covers X^u. It also sharpens Pemantle and Peres' result since it holds for a general locally compact Polish space Λ. In addition, we completely describe the distribution of the difference $X \setminus X^u$. In the proof of the theorem we use basic results and definitions for operators on the Hilbert space $\mathcal{L}^2(\Lambda)$; see. e.g., [20, 22]. An outline of the proof is as follows. First, we dilate the operator associated to the DPP X to a projection operator on the union of two copies of Λ. Second, we use the existence of a coupling for projection operators in Lemma 1. Finally, we compress back down to Λ to obtain the desired coupling.

We use the following special result established under Condition 1 and where ν_S denotes the restriction of the reference measure ν to a compact set $S \subseteq \Lambda$.

Lemma 1. *Assume $S \subseteq \Lambda$ is compact and let $\{\phi_k^S\}_{k=1}^n$ be an orthonormal set of functions in $L^2(S)$ with $1 \leq n < \infty$. Let X and Y be DPPs with kernels K and L, respectively, so that*

$$K(v, w) = \sum_{k=1}^{n} \phi_k^S(v)\overline{\phi_k^S(w)}, \qquad L(v, w) = \sum_{k=1}^{n-1} \phi_k^S(v)\overline{\phi_k^S(w)}, \qquad v, w \in S$$

https://doi.org/10.1017/jpr.2020.101 Published online by Cambridge University Press

(setting $L(v, w) = 0$ if $n = 1$). Then there exists a monotone coupling of Y_S w.r.t. X_S such that, almost surely, $Y_S \subset X_S$, $\eta_S := X_S \setminus Y_S$ consists of one point, and the point in η_S has density $|\phi_n^S(\cdot)|^2$ w.r.t. ν_S.

Proof. Observe that K and L are the kernels of finite-dimensional projections, a special case of trace-class positive contractions, and the difference,

$$K(v, w) - L(v, w) = \phi_n^S(v)\overline{\phi_n^S(w)}, \qquad v, w \in S,$$

is a positive definite kernel. Thus, by [15, Theorem 3.8], X_S stochastically dominates Y_S. Therefore, there is a coupling such that, almost surely, $Y_S \subseteq X_S$. As Y_S has cardinality one less than X_S, almost surely $\eta_S := X_S \setminus Y_S$ consists of one point. Finally, for any Borel set $A \subseteq S$,

$$\mathbb{P}(\eta_S \cap A \neq \emptyset) = \mathbb{E}\left[1_{\{X(A)-Y(A)=1\}}\right] = \mathbb{E}[X(A)] - \mathbb{E}[Y(A)] = \int_A |\phi_n^S(v)|^2 \, \mathrm{d}\nu(v). \qquad \square$$

Denote by $\|\cdot\|_2$ the usual norm on $\mathcal{L}^2(\Lambda)$ w.r.t. ν.

Theorem 1. *Let X be a DPP on Λ with kernel K satisfying Assumptions 1–4. For any $u \in \Lambda$ with $K(u, u) > 0$, there exists a coupling of X and X^u such that, almost surely, $X^u \subseteq X$ and $\xi_u := X \setminus X^u$ consists of at most one point. We have*

$$p_u := \mathbb{P}(\xi_u \neq \emptyset) = \frac{1}{K(u, u)} \int |K(u, v)|^2 \, \mathrm{d}\nu(v), \tag{6}$$

and, conditioned on $\xi_u \neq \emptyset$, the point in ξ_u has density

$$f_u(v) := |K(u, v)|^2 / \|K(u, \cdot)\|_2^2, \qquad v \in \Lambda, \tag{7}$$

w.r.t. ν.

Compared to Goldman's result (G1), we also have that $p_u = \mathbb{P}(\kappa_u \neq \emptyset)$ and f_u is the conditional density of a point in κ_u given that $\kappa_u \neq \emptyset$, cf. (5)–(7).

Proof. We begin by describing the procedure given in Lyons' paper [15, Section 3.3] for dilating a locally trace class operator to a locally trace class orthogonal projection. Denote by $\mathcal{K}$ the locally trace class operator on $\mathcal{L}^2(\Lambda)$ with kernel K. Consider the dilation of $\mathcal{K}$ given by

$$\mathcal{Q} := \begin{bmatrix} \mathcal{K} & \mathcal{L} \\ \mathcal{L} & \mathcal{I} - \mathcal{K} \end{bmatrix},$$

where $\mathcal{L} := \sqrt{\mathcal{K}(\mathcal{I} - \mathcal{K})}$. Then, since $\mathcal{Q} = \mathcal{Q}^2$, $\mathcal{Q}$ is an orthogonal projection on $L^2(\Lambda, \nu) \oplus L^2(\Lambda_0, \nu)$, where Λ_0 is a disjoint identical copy of Λ. If Λ is discrete, then Q is clearly locally trace class, since any compact set of a discrete space is finite. If Λ is not discrete, consider the operator

$$\mathcal{Q}' := \begin{bmatrix} \mathcal{I} & 0 \\ 0 & \mathcal{U} \end{bmatrix}^* \mathcal{Q} \begin{bmatrix} \mathcal{I} & 0 \\ 0 & \mathcal{U} \end{bmatrix} = \begin{bmatrix} \mathcal{K} & \mathcal{L}\mathcal{U} \\ \mathcal{U}^*\mathcal{L} & \mathcal{U}^*(\mathcal{I} - \mathcal{K})\mathcal{U} \end{bmatrix},$$

where $\mathcal{U}$ is a unitary operator from $\ell^2(\Lambda_0')$ to $L^2(\Lambda_0, \nu)$ for some countably infinite space Λ_0'. The operator $\mathcal{U}$ exists since any two infinite-dimensional separable Hilbert spaces are unitarily equivalent. The operator $\mathcal{Q}'$ is an orthogonal projection on $L^2(\Lambda, \nu) \oplus \ell^2(\Lambda_0')$, and $\mathcal{K}$ is the

https://doi.org/10.1017/jpr.2020.101 Published online by Cambridge University Press

compression of $\mathcal{Q}'$ to Λ. Further, $\mathcal{Q}'$ is also locally trace class, because $\mathcal{K}$ is locally trace class on $L^2(\Lambda, \nu)$ by assumption, and all operators on $\ell^2(\Lambda_0')$ are locally of trace class since Λ_0' is discrete. Thus, $\mathcal{Q}'$ defines a projection DPP Y_Q on the union $\Lambda \cup \Lambda_0'$.

First, assume that Λ is compact. Then, the kernel of the operator $\mathcal{K}$ satisfies

$$K(v, w) = \sum_{k\geq 1} \lambda_k^\Lambda \phi_k^\Lambda(v)\overline{\phi_k^\Lambda(w)}, \qquad v, w \in \Lambda,$$

where $\{\phi_k^\Lambda\}$ is an orthonormal basis for $L^2(\Lambda)$, $\lambda_k^\Lambda \in [0, 1]$ for all k, and $\sum_{k\geq 1} \lambda_k^\Lambda < \infty$. Also, the kernel for the operator $\mathcal{L}$ is then given by

$$L(v, w) = \sum_{k\geq 1} \sqrt{\lambda_k^\Lambda(1 - \lambda_k^\Lambda)}\phi_k^\Lambda(v)\overline{\phi_k^\Lambda(w)}.$$

Note that

$$\mathcal{L}(L(\cdot, u))(w) = \int_\Lambda L(w, v)L(v, u)\, \mathrm{d}\nu(v) = \sum_{k\geq 1} \lambda_k^\Lambda(1 - \lambda_k^\Lambda)\phi_k^\Lambda(w)\overline{\phi_k^\Lambda(u)},$$

and

$$\mathcal{K}(K(\cdot, u))(w) = \int_\Lambda K(w, v)K(v, u)\, \mathrm{d}\nu(v) = \sum_{k\geq 1} \left(\lambda_k^\Lambda\right)^2 \phi_k^\Lambda(w)\overline{\phi_k^\Lambda(u)}.$$

Hence, $\mathcal{K}(K(\cdot, u)) + \mathcal{L}(L(\cdot, u)) = K(\cdot, u)$. Also,

$$\mathcal{L}(K(\cdot, u))(w) = \int_\Lambda L(w, v)K(v, u)\, \mathrm{d}\nu(v) = \sum_{k\geq 1} \lambda_k^\Lambda \sqrt{\lambda_k^\Lambda(1 - \lambda_k^\Lambda)}\phi_k^\Lambda(w)\overline{\phi_k^\Lambda(u)}$$

and

$$\mathcal{K}(L(\cdot, u))(w) = \int_\Lambda K(w, v)L(v, u)\, \mathrm{d}\nu(v) = \sum_{k\geq 1} \lambda_k^\Lambda \sqrt{\lambda_k^\Lambda(1 - \lambda_k^\Lambda)}\phi_k^\Lambda(w)\overline{\phi_k^\Lambda(u)},$$

and so $\mathcal{L}(K(\cdot, u)) = \mathcal{K}(L(\cdot, u))$. Consequently, for fixed $u \in \Lambda$,

$$\psi_u(\cdot) := \begin{bmatrix} \dfrac{K(\cdot, u)}{\sqrt{K(u, u)}} \\ \mathcal{U}^* \left(\dfrac{L(\cdot, u)}{\sqrt{K(u, u)}} \right) \end{bmatrix}$$

is an eigenvector of the operator $\mathcal{Q}'$. Indeed, since $\mathcal{U}\mathcal{U}^* = \mathcal{I}$ by that fact that $\mathcal{U}$ is unitary,

$$\begin{aligned} \mathcal{Q}'(\psi_u(\cdot)) &= \begin{bmatrix} \mathcal{I} & 0 \\ 0 & \mathcal{U} \end{bmatrix}^* Q \begin{bmatrix} \dfrac{K(\cdot, u)}{\sqrt{K(u, u)}} \\ (\mathcal{U}\mathcal{U}^*) \left(\dfrac{L(\cdot, u)}{\sqrt{K(u, u)}} \right) \end{bmatrix} \\ &= \begin{bmatrix} \mathcal{I} & 0 \\ 0 & \mathcal{U}^* \end{bmatrix} \begin{bmatrix} \dfrac{\mathcal{K}(K(\cdot, u))}{\sqrt{K(u, u)}} + \dfrac{\mathcal{L}(L(\cdot, u))}{\sqrt{K(u, u)}} \\ \dfrac{\mathcal{L}(K(\cdot, u))}{\sqrt{K(u, u)}} + \dfrac{(\mathcal{I} - \mathcal{K})(L(\cdot, u))}{\sqrt{K(u, u)}} \end{bmatrix} = \begin{bmatrix} \dfrac{K(\cdot, u)}{\sqrt{K(u, u)}} \\ \mathcal{U}^* \left(\dfrac{L(\cdot, u)}{\sqrt{K(u, u)}} \right) \end{bmatrix} = \psi_u(\cdot). \end{aligned}$$

https://doi.org/10.1017/jpr.2020.101 Published online by Cambridge University Press

Then, we can define the projection $\mathcal{Q}'_u := \mathcal{Q}' - P_{\psi_u}$, where P_{ψ_u} is the projection operator on $L^2(\Lambda, \nu) \oplus \ell^2(\Lambda'_0)$ onto the span of ψ_u. This projection operator is also locally trace class since it is the difference of locally trace class operators. Then we can define the projection DPP Y^u_Q on $\Lambda \cup \Lambda'_0$ associated with $\mathcal{Q}'_u$. If $\mathcal{Q}'$ has finite rank, then $\mathcal{Q}'$ and $\mathcal{Q}'_u$ have corresponding kernels

$$\mathcal{Q}' = \sum_{k=0}^{n} q_k q_k^\top \quad \text{and} \quad \mathcal{Q}'_u = \sum_{k=1}^{n} q_k q_k^\top,$$

where $n < \infty$, $\{q_k\}_{k=1}^n$ is an orthonormal set, and $q_0 := \psi_u$. Applying Lemma 1 then gives the result.

Now, assume $\mathcal{Q}'$ projects onto an infinite-dimensional subspace of $L^2(\Lambda, \nu) \oplus \ell^2(\Lambda'_0)$ and let $\{q_k\}_{k=0}^\infty$ be an orthonormal basis for the range of $\mathcal{Q}'$, where $q_0 := \psi_u$. For each positive integer M, define the finite-dimensional projection kernels

$$\mathcal{Q}'_M = \sum_{k=0}^{M} q_k q_k^\top \quad \text{and} \quad \mathcal{Q}'_{M,u} = \sum_{k=1}^{M} q_k q_k^\top,$$

and let Y_{Q_M} and $Y^u_{Q_M}$ be the corresponding projection DPPs. By Lemma 1, there is a coupling of Y_{Q_M} and $Y^u_{Q_M}$ such that, almost surely, $Y_{Q_M} \supset Y^u_{Q_M}$, where $\xi^u_{Q_M} := Y_{Q_M} \setminus Y^u_{Q_M}$ consists of one point which has density $|\psi_u(\cdot)|^2$. By the same argument as in the proof of [8, Lemma 20], the sequences Y_{Q_M} and $Y^u_{Q_M}$ are tight and converge in distribution to Y_Q and Y^u_Q, respectively, as $M \to \infty$. Also, the sequence $(Y^u_{Q_M}, \xi^u_{Q_M})_M$ is tight, and thus a subsequence converges in distribution to (Y^u_Q, ξ^u_Q), where ξ^u_Q consists of one point with density $|\psi_u(\cdot)|^2$, and $Y^u_Q \cup \xi^u_Q$ is equal in distribution to Y_Q.

The projection operator P_{ψ_u} has kernel $\psi_u \psi_u^\top$ and the compression of P_{ψ_u} to Λ is the integral operator with kernel

$$\frac{K(v, u)K(u, w)}{K(u, u)}.$$

Then, since the compression of $\mathcal{Q}'$ to Λ is the operator $\mathcal{K}$, the compression of $\mathcal{Q}'_u$ to Λ is the integral operator $\mathcal{K}^u$ with kernel

$$K^u(v, w) = K(v, w) - \frac{K(v, u)K(u, w)}{K(u, u)}.$$

This gives that $Y_Q \cap \Lambda$ has the same distribution as X and $Y^u_Q \cap \Lambda$ has the same distribution as X^u. Thus, almost surely, $X = X^u \cup \xi_u$, where $\xi_u := \xi^u_Q \cap \Lambda$ and X^u are disjoint. Therefore, we have a coupling of X and X^u, where almost surely $X^u \subseteq X$ and the difference is at most one point. The probability of $\xi_u \neq \emptyset$ is the probability that ξ^u_Q is in Λ, and the density of ξ^u_Q restricted to Λ is

$$f_{\xi^u_Q}(v)\mathbf{1}_{\{v \in \Lambda\}} = \frac{|K(v, u)|^2}{K(u, u)}$$

w.r.t. ν. Hence,

$$\mathbb{P}(\xi^u \neq \emptyset) = \mathbb{P}(\xi^u_Q \in \Lambda) = \int \frac{|K(v, u)|^2}{K(u, u)} \, \mathrm{d}\nu(v),$$

and the density of ξ_u conditioned on $\xi_u \neq \emptyset$ is $f_u(v) = |K(v, u)|^2 / \|K(\cdot, u)\|_2^2$ w.r.t. ν.

https://doi.org/10.1017/jpr.2020.101 Published online by Cambridge University Press

Second, if Λ is not assumed to be compact, consider a sequence of compact sets $S_n \subset \Lambda$ such that $\cup_{n=1}^{\infty} S_n = \Lambda$ and $S_n \subseteq S_{n+1}$ for $n = 1, 2, \ldots$. For each n, using the result above with Λ replaced by S_n, there exists a coupling of $(X_{S_n}, X^u_{S_n})$ where, almost surely, $X_{S_n} = X^u_{S_n} \cup \xi^u_{S_n}$, $\xi^u_{S_n} = X_{S_n} \setminus X^u_{S_n}$ consists of at most one point,

$$\mathbb{P}\left(\xi^u_{S_n} \neq \emptyset\right) = \int_{S_n} \frac{|K(v,u)|^2}{K(u,u)}\, \mathrm{d}\nu(v), \tag{8}$$

and, conditioned on $\xi^u_{S_n} \neq \emptyset$, the density of the point in $\xi^u_{S_n}$ is

$$f_{u,S_n}(v) = |K(v,u)|^2 / \int_{S_n} |K(w,u)|^2\, \mathrm{d}\nu(w) \tag{9}$$

w.r.t. ν_{S_n}. For consistency, let $T_1 = S_1$ and generate a realization $(y_{T_1}, y^u_{T_1})$ of $(Y_{T_1}, Y^u_{T_1}) := (X_{S_1}, X^u_{S_1})$, and for $n = 2, 3, \ldots$, let $T_n = S_n \setminus S_{n-1}$ and generate a realization $(y_{T_n}, y^u_{T_n})$ of $(Y_{T_n}, Y^u_{T_n})$ which follows the conditional distribution of $(X_{S_n} \setminus S_{n-1}, X^u_{S_n} \setminus S_{n-1})$ given that $(X_{S_n} \cap S_{n-1}, X^u_{S_n} \cap S_{n-1}) = (\cup_{i=1}^{n-1} y_{T_i}, \cup_{i=1}^{n-1} y^u_{T_i})$. Then, (X, X^u) is distributed as $(Y, Y^u) := (\cup_{n=1}^{\infty} Y_{T_n}, \cup_{n=1}^{\infty} Y^u_{T_n})$ and, almost surely, for $n = 2, 3, \ldots$, $Y_{T_{n-1}} \setminus Y^u_{T_{n-1}} \neq \emptyset$ implies that $Y_{T_n} \setminus Y^u_{T_n} = Y_{T_{n+1}} \setminus Y^u_{T_{n+1}} = \cdots = \emptyset$, and so $\xi_u := Y \setminus Y^u$ consists of at most one point. The probability that ξ_u is non-empty is, by (8),

$$\mathbb{P}(\xi_u \neq \emptyset) = \lim_{n\to\infty} \int_{S_n} \frac{|K(v,u)|^2}{K(u,u)}\, \mathrm{d}\nu(v),$$

and hence, by monotone convergence, we obtain (6). Finally, (7) is obtained in a similar way using (9). □

4. Describing repulsiveness in DPPs

In this section we use the probability p_u to quantify how repulsive a DPP can be, and we use the density f_u from Theorem 1 to describe the repulsive effect of a fixed point contained in a DPP. As mentioned in Section 4.1, in the case of stationary DPPs on $\mathbb{R}^d$, p_u turns out to agree with a measure for repulsiveness studied in [1, 2, 13, 14], but we are not aware of references where f_u has been considered when discussing repulsiveness in DPPs. Examples of p_u and f_u for specific models of DPPs are given in Section 4.2.

Note that X^u is the point process where there is a 'ghost point' at u that is affecting the remaining points. Using this coupling of X^u and X, it is clear that the repulsive effect of a point at location u is characterized by the difference between X^u and the original DPP X, where there is no repulsion coming from the location u. Further, as X and X^u have intensity functions $\rho(\cdot)$ and $\rho(u,\cdot)/\rho(u)$, respectively, $\xi_u = X \setminus X^u$ has intensity function $\rho_u(v) := |K(v,u)|^2/K(u,u)$, $v \in \Lambda$. This is the intensity function for the points in X 'pushed out' by u under the Palm distribution. It makes also sense to consider ρ_u as the intensity function of $X \setminus X^u$ when ν is diffuse and X is a Poisson process, because then $X = X^u$ and $\rho_u(v) = 0$ for $v \neq u$.

4.1. A measure of repulsiveness

Setting $0/0 = 0$, recall that the pair correlation function of X is defined by $g(v,w) = \rho(v,w)/(\rho(v)\rho(w))$ for $v, w \in \Lambda$, so it satisfies $1 - g(v,w) = |r(v,w)|^2$, $v, w \in \Lambda$, where $r(v,w) = K(v,w)/\sqrt{K(v,v)K(w,w)}$ is the correlation function obtained from K. Note that

$$\rho_u(v) = \rho(v)(1 - g(u,v)), \qquad v \in \Lambda. \tag{10}$$

https://doi.org/10.1017/jpr.2020.101 Published online by Cambridge University Press

4.1.1 *Defining a global measure of repulsiveness.* As a global measure of repulsiveness in X when having a point of X at u, we suggest the probability of $\xi_u \neq \emptyset$, that is,

$$p_u = \int \rho_u(v)\,\mathrm{d}v(v) = \int |K(u,v)|^2/K(u,u)\,\mathrm{d}v(v).$$

By (10), there is a trade-off between intensity and repulsiveness: if p_u is fixed, we cannot both increase ρ and decrease g. Therefore, when using p_u as a measure to compare repulsiveness in two DPPs, they should share the same intensity function ρ. Then, small/large values of p_u correspond to a small/large degree of repulsiveness. For a stationary DPP X on $\mathbb{R}^d$, p_u agrees with the measure for repulsiveness in DPPs introduced in [13, 14]; see also [1, 2]. Indeed, this measure is very specific for DPPs, as discussed later in Section 4.2.5.

4.1.2 *Definition of globally most repulsive DPPs.* If $p_u = 1$ for all $u \in \Lambda$ with $K(u,u) > 0$, we say that X is a globally most repulsive DPP. This is the case if K is a projection, that is, for all $v, w \in \Lambda$,

$$K(v,w) = \int K(v,y)K(y,w)\,\mathrm{d}v(y).$$

For short, we then say that X is a projection DPP. The standard Ginibre point process given by (4) is globally most repulsive, and its kernel is indeed a projection; this follows from a straightforward calculation using that $(v,w) \to \exp(v\overline{w})$ is the reproducing kernel of the Bargmann–Fock space equipped with the standard complex Gaussian measure. At the other end, if ν is diffuse and X is a Poisson process with intensity function ρ, then $p_u = 0$ for all $u \in \Lambda$ with $\rho(u) > 0$, and so X is a globally least repulsive DPP.

If Λ is compact, then it follows from the spectral representation (2) and Assumption 4 that

$$\begin{aligned}\int_S |K(u,v)|^2\,\mathrm{d}v(v) &= \sum_k \sum_\ell \lambda_k^S \lambda_l^S \phi_k^S(u)\overline{\phi_\ell^S(u)} \int_S \overline{\phi_k^S(v)}\phi_\ell^S(v)\,\mathrm{d}v(v) \\ &= \sum_k \left(\lambda_k^S\right)^2 |\phi_k^S(u)|^2 \leq \sum_k \lambda_k^S |\phi_k^S(u)|^2 = K(u,u), \end{aligned} \tag{11}$$

and so

$$p_u = \frac{\sum_k \left(\lambda_k^\Lambda\right)^2 |\phi_k^\Lambda(u)|^2}{\sum_k \lambda_k^\Lambda |\phi_k^\Lambda(u)|^2}. \tag{12}$$

Consequently, in this case, projection DPPs are the only globally most repulsive DPPs. Such a process has a fixed number of points which agrees with the rank of the kernel.

4.2. Examples

This section shows specific examples of our measure p_u and the distribution of a point in ξ_u.

4.2.1 *DPPs defined on a finite set.* Assume $\Lambda = \{1, \ldots, n\}$ is finite and ν is the counting measure; this is the simplest situation. Then $L^2(\Lambda) \equiv \mathbb{C}^n$, the class of possible kernels for DPPs corresponds to the class of $n \times n$ complex covariance matrices with all eigenvalues ≤ 1, and the eigenfunctions simply correspond to normalized eigenvectors for such matrices. For simplicity we only consider projection DPPs and Poisson processes below, but other examples of DPPs on finite sets include uniform spanning trees [10, Example 14] and finite DPPs converging to the continuous Airy process on the complex plane [11].

https://doi.org/10.1017/jpr.2020.101 Published online by Cambridge University Press

The projection DPPs are given by complex projection matrices, ranging between the degenerated cases where $X = \emptyset$ and $X = \Lambda$. For example, consider the projection kernel of rank two given by $K(v, w) = \frac{1}{n} + t_v\overline{t_w}$, where $\sum_{i=1}^n t_i = 0$ and $\sum_{i=1}^n |t_i|^2 = 1$. For any $u \in \{1, \dots, n\}$, we have $p_u = 1$ and

$$\rho_u(v) = \frac{|\frac{1}{n} + t_u\overline{t_v}|^2}{\frac{1}{n} + |t_u|^2}, \qquad v \in \{1, \dots, n\},$$

is a probability mass function. This shows the repulsive effect of having a point of X at u; in particular, $\rho_u(v)$ has a global maximum point at $v = u$.

The kernel of a Poisson process with intensity function $\rho \le 1$ and conditioned on having no multiple points is given by a diagonal covariance matrix with diagonal entries $\rho(1), \dots, \rho(n)$. If $\rho(u) > 0$, then $p_u = \rho(u)$. This is a much different result than when we consider a Poisson process X on a space Λ where the reference measure ν is diffuse: if $\rho(u) > 0$, then $p_u = 0$ and, almost surely, $X = X^u$.

4.2.2 *Ginibre point processes.* From the standard Ginibre point process given by (4), other stationary point processes can be obtained. Independently thinning the process with a retention probability $\alpha\beta$, where $\beta > 0$ and $\alpha \in (0, 1/\beta]$, and multiplying each of the retained points by $\sqrt{\beta}$ gives a new stationary DPP with kernel

$$K(v, w) = \frac{\alpha}{\pi} \exp\left(\frac{v\overline{w}}{\beta} - \frac{|v|^2 + |w|^2}{2\beta}\right), \qquad v, w \in \mathbb{C}. \tag{13}$$

We have

$$\rho = \alpha/\pi, \qquad p_u = \alpha\beta, \qquad f_u(v) = \frac{\exp\left(-|v - u|^2/\beta\right)}{\pi\beta} \sim N_{\mathbb{C}}(u, \beta). \tag{14}$$

The case where $\alpha = 1$ and $0 < \beta \le 1$ is mentioned in Goldman's paper [8], and the results in (14) match those in [8, Remark 24]. Deng *et al.* [5] called the DPP with kernel (13) the scaled β-Ginibre point process, but the bound $\alpha\beta \le 1$ was not noticed. For any fixed value of $\rho > 0$, as the value of β increases to its maximum $\min\{1, 1/(\pi\rho)\}$, the more repulsive the process becomes, whilst as β decreases to 0, in the limit a Poisson process with intensity ρ is obtained.

4.2.3 *DPPs on $\mathbb{R}^d$ with a stationary kernel.* Suppose $\Lambda = \mathbb{R}^d$, ν is the Lebesgue measure, and $K(u, v) = K_0(u - v)$ is stationary, where $K_0 \in L^2(\mathbb{R}^d)$ and K_0 is continuous. Then it follows from Parseval's identity that $p_u = 1$ if and only if $\hat{K}_0$ is an indicator function whose integral agrees with the intensity of X, cf. Appendix J in [13]. A natural choice for the support of this indicator function is a ball centred at the origin in $\mathbb{R}^d$, and if (as in the standard Ginibre point process) we let the intensity be $1/\pi$, then the globally most repulsive DPP has a stationary and isotropic kernel given by

$$K(v, w) = \int_{|y|^d \le d\Gamma(d/2)/(2\pi^{1+d/2})} \exp(2\pi i(v - w) \cdot y)\ \mathrm{d}y, \qquad v, w \in \mathbb{R}^d, \tag{15}$$

where $x \cdot y$ denotes the usual inner product for $x, y \in \mathbb{R}^d$, and $|y|$ is the usual Euclidean distance. For instance, for $d = 1$ this kernel is the sinc function, and for $d = 2$ it is the jinc-like function

$$K(v, w) = J_1(2|v - w|)/(\pi|v - w|), \tag{16}$$

https://doi.org/10.1017/jpr.2020.101 Published online by Cambridge University Press

where J_1 is the Bessel function of order one. We straightforwardly obtain the following proposition, where the moments in (17) follow from [21, Eq. 10.22.57].

Proposition 4.1. *For the globally most repulsive DPP on $\mathbb{R}^d$ with kernel given by (15) and for any $u \in \mathbb{C}$, $\rho_u(v) = \pi |K(u, v)|^2$ is a probability density function. In particular, for $d = 2$, $\rho_u(v) = J_1(2|v-u|)^2 / \left(\pi |v-u|^2\right)$, $v \in \mathbb{R}^2$, and the moments of $|Z_u - u|$ with $Z_u \sim \rho_u$ satisfy*

$$\mathbb{E}\left(|Z_u - u|^k\right) = \frac{\Gamma(1+k/2)\Gamma(1-k)}{\Gamma(2-k/2)\Gamma(1-k/2)^2}, \qquad k \in (-2, 1), \tag{17}$$

and are infinite for $k \geq 1$.

For comparison consider a standard Ginibre point process, where we can define Z_u in a similar way to Proposition 4.1. In both cases, $|Z_u - u|$ is independent of $(Z_u - u)/|Z_u - u|$, which is uniformly distributed on the unit circle. However, the distribution of $|Z_u - u|$ is very different in the two cases. For the standard Ginibre point process, $|Z_u - u|^2$ is exponentially distributed and $|Z_u - u|$ has a finite kth moment for all $k > -2$ given by $\Gamma(1 + k/2)/(\pi\rho)^{k/2}$, whilst for the DPP on $\mathbb{R}^2$ with jinc-like kernel (16), $|Z_u - u|$ is heavy-tailed and has infinite kth moments for all $k \geq 1$.

For any DPP X with kernel K and defined on $\mathbb{R}^d$, using independent thinning and scale transformation procedures similar to those in Section 4.2.2 (replacing $\sqrt{\beta}$ by $\beta^{1/d}$ when transforming the points in the thinned process), we obtain a new DPP with kernel $K_{\text{new}}(v, w) = \alpha K(v/\beta^{1/d}, w/\beta^{1/d})$, $v, w \in \mathbb{R}^d$, where $\beta \in (0, 1]$ and $\alpha \in (0, 1/\beta]$. For instance, if K is the jinc-like kernel for the globally most repulsive DPP given by (16), the new DPP satisfies the same equations for its intensity ρ and its probability p_u as in (14). Hence, if ρ and β are the same for this new DPP and the scaled β-Ginibre point process, the two DPPs are equally repulsive in terms of p_u. However, the probability density function for the point in ξ_u conditioned on $\xi_u \neq \emptyset$ now becomes

$$f_u(v) = J_1\left(2|v-u|^2/\beta\right) / \left(\pi |v-u|^2/\beta\right). \tag{18}$$

The reach of the repulsive effect of the point at u is much different when comparing the densities in (14) and (18), particularly if β is large. See Figure 1 for a comparison of the densities (14) and (18) for $\beta = 1$.

4.2.4 *DPPs on $\mathbb{S}^d$ with an isotropic kernel.* Suppose $\Lambda = \mathbb{S}^d$ is the d-dimensional unit sphere, ν is the Lebesgue measure, and $K(v, w) = K_0(v \cdot w)$ is isotropic for all $v, w \in \mathbb{S}^d$. Then the DPP with kernel K is isotropic, and $\rho = K_0(1)$ and p_u do not depend on the choice of $u \in \Lambda$. By a classical result of Schoenberg [25] and by [18, Theorem 4.1], we have the following. The normalized eigenfunctions will be complex spherical harmonic functions, and K_0 will be real and of the form

$$K_0(t) = \rho \sum_{\ell=0}^{\infty} \beta_{\ell,d} \frac{C_\ell^{\left(\frac{d-1}{2}\right)}(t)}{C_\ell^{\left(\frac{d-1}{2}\right)}(1)}, \qquad -1 \leq t \leq 1,$$

where $C_\ell^{\left(\frac{d-1}{2}\right)}$ is a Gegenbauer polynomial of degree ℓ and the sequence $\beta_{0,d}, \beta_{1,d}, \ldots$ is a probability mass function. Further, letting $\sigma_d = \nu(\mathbb{S}^d) = 2\pi^{(d+1)/2} / \Gamma((d+1)/2)$, the

https://doi.org/10.1017/jpr.2020.101 Published online by Cambridge University Press

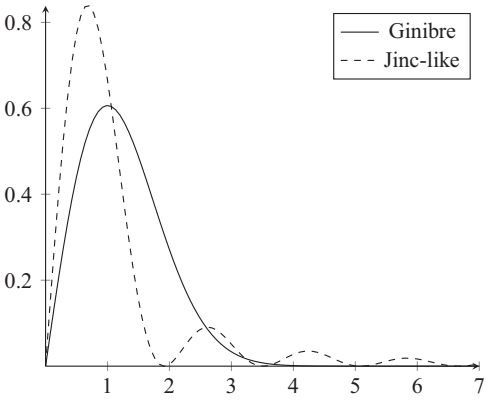


FIGURE 1: The densities of $|Z_0|$ for $Z_0 := X \backslash X^0$ for two globally most repulsive DPPs.

eigenvalues of K are $\lambda_{\ell,d} = \rho\sigma_d\beta_{\ell,d}/m_{\ell,d}$, $\ell = 0, 1, \ldots$, with multiplicities

$$m_{\ell,d} = \begin{cases} m_{0,1}, & m_{\ell,1} = 2, \quad \ell = 1, 2, \ldots, \quad \text{if } d = 1, \\ \dfrac{2\ell + d - 1}{d - 1}\dfrac{(\ell + d - 2)!}{\ell!(d - 2)!}, & \ell = 0, 1, \ldots, \quad \text{if } d \in \{2, 3, \ldots\}. \end{cases}$$

So, the DPP exists if and only if $\rho \le \inf_{\ell:\, \beta_{\ell,d} > 0} m_{\ell,d}/(\sigma_d\beta_{\ell,d})$. Now, applying (12), we obtain

$$p_u = \rho\sigma_d \sum_{\ell=0}^{\infty} \beta_{\ell,d}^2/m_{\ell,d}. \tag{19}$$

There is a lack of flexible parametric DPP models on the sphere where K_0 is expressible in closed form [18, Section 4.3]. For instance, let $d = 2$ and consider the special case of the multiquadric model given by

$$K_0(t) = \rho\frac{1 - \delta}{\sqrt{1 + \delta^2 - 2\delta t}}, \qquad -1 \le t \le 1,$$

with $\delta \in (0, 1)$ a parameter and $0 < \rho \le 1/(4\pi(1 - \delta))$. Then, as shown in [18, Section 4.3.2], the sequence

$$\beta_{\ell,2} = (1 - \delta)\delta^\ell, \qquad \ell = 0, 1, \ldots, \tag{20}$$

specifies a geometric distribution and $\lambda_{\ell,2} = 4\pi\rho\delta^\ell(1 - \delta)/(2\ell + 1) \le \delta^\ell/(2\ell + 1)$, $\ell = 0, 1, \ldots$ As $\delta \to 0$, then $\lambda_{0,2} \to 4\pi\rho$ and $\lambda_{\ell,2} \to 0$ if $\ell \ge 1$, corresponding to the uninteresting case of a DPP with at most one point if $\rho < 1/(4\pi)$ and with exactly one point if $\rho = 1/(4\pi)$. From (19) and (20) we obtain $p_u = 4\pi\rho(1 - \delta)/(1 + \delta) \le 1/(1 + \delta)$, with this upper bound obtained for the maximal value of $\rho = 1/(4\pi(1 - \delta))$. Therefore, the DPP with the multiquadric kernel is far from being globally most repulsive unless the expected number of points is very small.

Instead, a flexible parametric model for the eigenvalues $\lambda_{\ell,d}$ is suggested in [18, Section 4.3.4] so that globally most repulsive DPPs as well as Poisson processes are obtained as limiting cases. However, the disadvantage of that model is that we can only numerically calculate ρ and p_u.

https://doi.org/10.1017/jpr.2020.101 Published online by Cambridge University Press

4.2.5 *Remark.* The considerations in Sections 4.1 and 4.2.1–4.2.4 are strictly for DPPs. For example, the intensity function of a Gibbs point process can be both smaller and larger than the intensity function of its Palm distribution at a given point, whilst for a DPP, $\rho \geq \rho^u$. Furthermore, as a candidate for a 'globally most repulsive stationary Gibbs point process on $\mathbb{R}^2$', we may consider $Y = L_Z := \{x + Z : x \in L\}$, where L is the vertex set of a regular triangular lattice (the centres of a honeycomb structure) with one lattice point at the origin, and where Z is a uniformly distributed point in the hexagonal region given by the Voronoi cell of the lattice and centred at the origin (in other words, Y may be considered as the limit of a stationary Gibbs hard core process when the packing fraction of hard discs increases to the maximal value ≈ 0.907; see, e.g., [6, 17]). However, the reduced Palm process at $u \in \mathbb{R}^2$ will be degenerated and given by $Y^u = L_u \setminus \{u\}$, which is a much different situation as compared to DPPs.

Acknowledgements

Jesper Møller was supported in part by The Danish Council for Independent Research | Natural Sciences, grant 7014-00074B, 'Statistics for point processes in space and beyond', and by the 'Centre for Stochastic Geometry and Advanced Bioimaging', funded by grant 8721 from the Villum Foundation. Eliza O'Reilly was supported in part by a grant of the Simons Foundation (#197982 to UT Austin), the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1110007, and the Department of Mathematical Sciences, Aalborg University. We are grateful to Professor Russell Lyons for many useful comments, in particular for changing our focus on a more complicated coupling result established in an earlier version of our paper (briefly, to obtain the reduced Palm process, one point in the DPP was either moved or removed) to the simpler coupling result in Theorem 1, which indeed is more suited for studying repulsiveness in DPPs. Also, he pointed our attention to his paper [15], which is essential in the proof of Theorem 1. Finally, we are grateful to two referees, in particular for pointing our attention to the work by Pemantle and Peres.

References

[1] Baccelli, F. and O'Reilly, E. (2018). Reach of repulsion for determinantal point processes in high dimensions. *J. Appl. Prob.* **55**, 760–788.
[2] Biscio, C. A. N. and Lavancier, F. (2016). Quantifying repulsiveness of determinantal point processes. *Bernoulli* **22**, 2001–2028.
[3] Borcea, J., Brändén, P. and Liggett, T. (2009). Negative dependence and the geometry of polynomials. *J. Am. Math. Soc.* **22**, 521–567.
[4] Borodin, A. and Olshanski, G. (2000). Distributions on partitions, point processes and the hypergeometric kernel. *Commun. Math. Phys.* **211**, 335–358.
[5] Deng, N., Zhou, W. and Haenggi, M. (2015). The Ginibre point process as a model for wireless networks with repulsion. *IEEE Trans. Wireless Commun.* **14**, 107–121.
[6] Döge, G., Mecke, K., Møller, J., Stoyan, D. and Waagepetersen, R. (2004). Grand canonical simulations of hard-disk systems by simulated tempering. *Int. J. Mod. Phys. C* **15**, 129–147.
[7] Ginibre, J. (1965). Statistical ensembles of complex, quaternion, and real matrices. *J. Math. Phys.* **6**, 440–449.
[8] Goldman, A. (2010). The Palm measure and the Voronoi tessellation for the Ginibre process. *Ann. Appl. Prob.* **20**, 90–128.
[9] Hough, J. B., Krishnapur, M., Peres, Y. and Virág, B. (2006). Determinantal processes and independence. *Prob. Surv.* **3**, 206–229.
[10] Hough, J. B., Krishnapur, M., Peres, Y. and Virág, B. (2009). *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*. American Mathematical Society, Providence, RI.
[11] Johansson, K. (2005). The arctic circle boundary and the Airy process. *Ann. Prob.* **33**, 1–30.
[12] Kulesza, A. and Taskar, B. (2012). *Determinantal Point Processes for Machine Learning*. Now Publishers Inc., Hanover, MA.

https://doi.org/10.1017/jpr.2020.101 Published online by Cambridge University Press

[13] LAVANCIER, F., MØLLER, J. AND RUBAK, E. (2014). Determinantal point process models and statistical inference (extended version). Preprint, arXiv:1205.4818.
[14] LAVANCIER, F., MØLLER, J. AND RUBAK, E. (2015). Determinantal point process models and statistical inference. *J. R. Statist. Soc. B* **77**, 853–877.
[15] LYONS, R. (2014). Determinantal probability: Basic probability and conjectures. *Proc. Int. Cong. Math.* **IV**, 137–161.
[16] MACCHI, O. (1975). The coincidence approach to stochastic point processes. *Adv. Appl. Prob.* **7**, 83–122.
[17] MASE, S., MØLLER, J., STOYAN, D., WAAGEPETERSEN, R. AND DÖGE, G. (2001). Packing densities and simulated tempering for hard core Gibbs point processes. *Ann. Inst. Statist. Math.* **53**, 661–680.
[18] MØLLER, J., NIELSEN, M., PORCU, E. AND RUBAK, E. (2018). Determinantal point process models on the sphere. *Bernoulli* **24**, 1171–1201.
[19] MØLLER, J. AND RUBAK, E. (2016). Functional summary statistics for point processes on the sphere with an application to determinantal point processes. *Spatial Stat.* **18**, 4–23.
[20] NAGY, B. S., FOIAS, C., BERCOVICI, H. AND KÉRCHY, L. (2010). *Harmonic Analysis of Operators on Hilbert Space*. Springer, New York.
[21] OLVER, F. W. J. *et al.* (eds) (2017). *NIST Digital Library of Mathematical Functions*. http://dlmf.nist.gov/, release 1.0.14 of 2017-09-18.
[22] PAULSEN, V. (2002). *Completely Bounded Maps and Operator Algebras*. Cambridge University Press, Cambridge.
[23] PEMANTLE, R. AND PERES, Y. (2011). Concentration of Lipschitz functionals of determinantal and other strong Rayleigh measures. Preprint, arXiv:1108.0687v1.
[24] PEMANTLE, R. AND PERES, Y. (2014). Concentration of Lipschitz functionals of determinantal and other strong Rayleigh measures. *Combinatorics Prob. Comput.* **23**, 140–160.
[25] SCHOENBERG, I. J. (1942). Positive definite functions on spheres. *Duke Math. J.* **9**, 96–108.
[26] SHIRAI, T. AND TAKAHASHI, Y. (2003). Random point fields associated with certain Fredholm determinants I: Fermion, Poisson and boson point processes. *J. Func. Anal.* **205**, 414–463.
[27] SNOEK, J., ZEMEL, R. S. AND ADAMS, R. P. (2013). A determinantal point process latent variable model for inhibition in neural spiking data. In *Proc. Advances in Neural Information Processing Systems 26 (NIPS 2013)*, eds I. S. Francis, B. J. F. Manly, and F. C. Lam, pp. 1932–1940.
[28] SOSHNIKOV, A. (2000). Determinantal random point fields. *Russian Math. Surv.* **55**, 923–975.

https://doi.org/10.1017/jpr.2020.101 Published online by Cambridge University Press