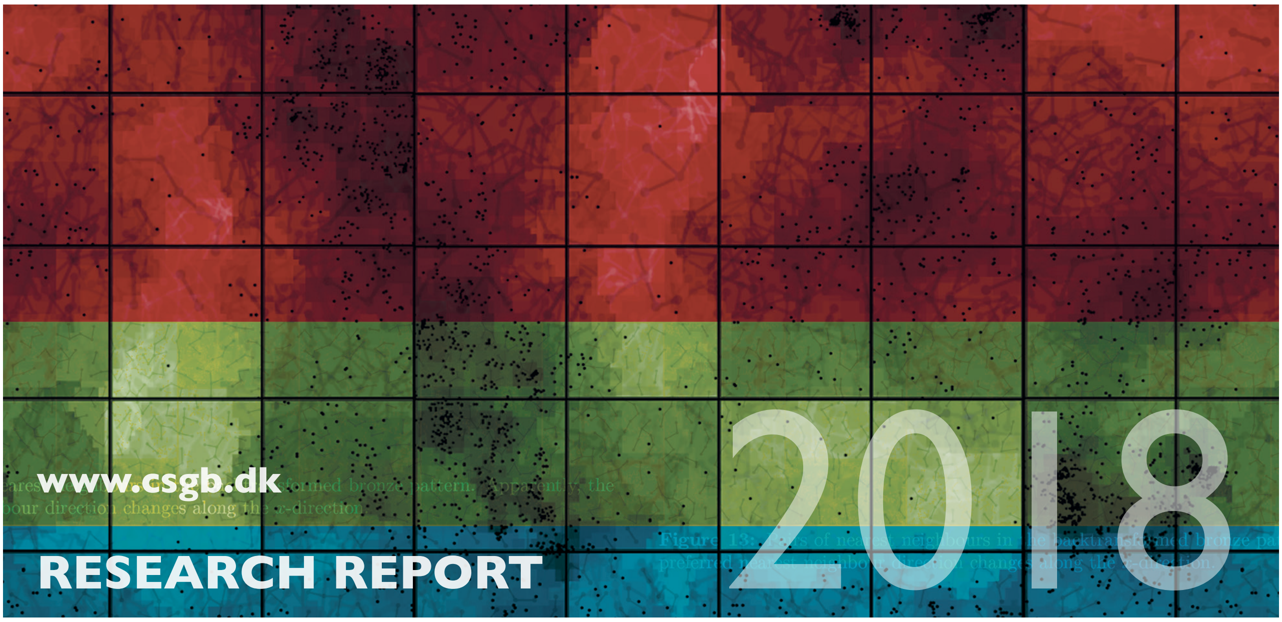
CENTRE FOR **STOCHASTIC GEOMETRY**
AND ADVANCED **BIOIMAGING**



Jesper Møller and Eliza O'Reilly

Couplings for determinantal point processes and their reduced Palm distributions with a view to quantifying repulsiveness

No. 07, June 2018

Couplings for determinantal point processes and their reduced Palm distributions with a view to quantifying repulsiveness

Jesper Møller*,[1] and Eliza O'Reilly**,[2]

[1]Department of Mathematical Sciences, Aalborg University, jm@math.aau.dk

[2]University of Texas at Austin, Department of Mathematics, eoreilly@math.utexas.edu

Abstract

For a determinantal point process 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 in distribution X^u is obtained by removing a finite number of points from X. The intensity function of the difference $X \setminus X^u$ is known, but apart from special cases the distribution of $X \setminus X^u$ is unknown. Considering the restriction X_S of X to any compact set S, we establish a coupling of X_S and its reduced Palm process X_S^u so that the difference is at most one point. Specifically, we assume K restricted to $S \times S$ is either (i) a projection or (ii) has spectrum strictly less than one. In case of (i), we have in distribution that X_S^u is obtained by removing one point from X_S, and we can specify the distribution of this point. In case of (ii), in distribution we obtain X_S^u either by moving one point in X or by removing one point from X_S, and to a certain extent we can describe the distribution of these points. We discuss how Goldman's and our results can be used for quantifying repulsiveness in X.

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

*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.

**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

1 Introduction

Determinantal point processes (DPPs) have been of much interest over the last many years in mathematical physics and probability theory (see e.g. Borodin and Olshanski (2000); Hough et al. (2009); Macchi (1975); Shirai and Takahashi (2003); Soshnikov (2000) and the references therein) and more recently in other areas, including statistics Lavancier et al. (2015); Møller and Rubak (2016), machine learning Kulesza and Taskar (2012), signal processing Deng et al. (2015), and neuroscience Snoek et al. (2013). They are models for regularity/inhibition/repulsiveness, but there is a trade-off between repulsiveness and intensity Lavancier et al. (2014, 2015). This paper sheds further light on this issue by studying various couplings between a DPP and its reduced Palm distributions.

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 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 Goldman (2000) result that if for any compact set $S \subseteq \Lambda$, denoting 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 (in the sense of Section 3.3) and hence by Strassen's theorem there exists a coupling so that $X^u \subseteq X$ in distribution. The difference $\kappa^u := X \setminus X^u$ is a finite point process with a known intensity function which can be used for quantifying repulsiveness in X as discussed in Section 3.1. In particular, for a standard Ginibre point process Ginibre (1965), 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. However, apart from this and other special cases, the distribution of κ^u is unknown.

Considering the restriction X_S of X to any compact set $S \subseteq \Lambda$, we show in Section 3 that more can be said: Section 3.4 concerns the case where K_S is a projection kernel of finite rank n. Then, in fact, X_S consists of n points almost surely, and if $n > 0$ we show there is a coupling so that $X_S^u \subset X$ in distribution, where $X_S \setminus X_S^u$ consists of one point whose distribution can be specified. This result together with the fact that a DPP on a compact set can be viewed as a DPP with a random projection kernel (cf. (Hough et al., 2006, Theorem 7)) is used in Section 3.5 to establish another kind of coupling when the spectrum of K_S is strictly less than one. In this case, it is shown that there exists a point process ξ_S^u consisting of at most one point and a point process η_S^u consisting of one point so that $X_S^u = (X_S \cup \xi_S^u) \setminus \eta_S^u$ in distribution, where $X_S \cap \xi_S^u = \emptyset$ almost surely. Furthermore, we specify the marginal distributions of ξ_S^u and η_S^u, and discuss how these can be used for describing the repulsiveness in X. If for all $u \in \Lambda$ with $K(u,u) > 0$, ξ_S^u tends in distribution to the empty set as S increases to Λ, we call X a most repulsive DPP; both in Section 3.4 and Section 3.5 we discuss this definition in connection to most repulsive stationary DPPs on $\mathbb{R}^d$ as specified in Lavancier et al. (2015); Biscio and Lavancier (2016). For example, if X is a Ginibre point process, we obtain a similar result as in Goldman (2000): X is a most repulsive DPP and the point in η_S^u tends in distribution to $N_{\mathbb{C}}(u,1)$ as S increases to $\mathbb{C}$. Moreover, we compare with most repulsive isotropic DPPs on $\mathbb{S}^d$, the d-dimensional unit sphere in $\mathbb{R}^{d+1}$, as studied in Møller et al. (2018).

Some further results and all proofs of our results in Section 3 are given in Section 4. As in Goldman (2000) we only verify the existence of the various couplings. We leave it as open research problems, for any compact set $S \subseteq \Lambda$,

- to provide specific coupling constructions or simulation procedures for the pair (X_S, X_S^u), including the case in Goldman (2000) and the case considered in the present paper;
- in general, i.e. not only when K_S is a finite rank projection kernel, to find a coupling so that in distribution, $X_S = \zeta_u^S \cup X_u^S$ and $\zeta_u^S \cap X_u^S = \emptyset$, where ζ_u^S contains a single point when $X_S \neq \emptyset$;
- or even better, in general to obtain a coupling so that $X = \zeta_u \cup X_u$ and $\zeta_u \cap X_u = \emptyset$ in distribution, where ζ_u contains a single point when $X \neq \emptyset$.

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 $K : \Lambda^2 \mapsto \mathbb{C}$, which by definition means that 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. Also, if $X(B)$ denotes the cardinality of $X_B := X \cap B$ for $B \in \mathcal{B}$, then the following property has to be satisfied: For any $n = 1, 2, \ldots$ and any mutually disjoint bounded sets $B_1, \ldots, B_n \in \mathcal{B}$,

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

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

$$\rho\left(u_1, \ldots, u_n\right) = \det\left\{K\left(u_i, u_j\right)\right\}_{i,j=1}^n, \qquad u_1, \ldots, u_n \in \Lambda, \tag{2.1}$$

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

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 that there are 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 counting measure, see Kulesza and Taskar (2012); when $\Lambda = \mathbb{R}^d$ and ν is Lebesgue measure, see Hough et al. (2009); Lavancier et al. (2015); and when $\Lambda = \mathbb{S}^d$ (the d-dimensional unit sphere) and ν is surface/Lebesgue measure, see Møller et al. (2018). Examples are also given in Section 3.2.

2.1.2 Assumptions

We always make the following assumptions (a)–(c):

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

(b) $K \in L^2(\Lambda^2)$, the space of square integrable (w.r.t. ν^2) complex functions defined on Λ^2;

(c) K is of 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.

This ensures the existence of a spectral representation for the kernel restricted to any compact set $S \subseteq \Lambda$: By Mercer's theorem, excluding a ν^2-nullset, we can assume that

$$K(u,v) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(u)\overline{\phi_k^S(v)} \qquad u, v \in S, \tag{2.2}$$

where the eigenvalues λ_k^S are real numbers and the eigenfunctions ϕ_k^S constitute an orthonormal basis of $L^2(S)$, cf. (Hough et al., 2009, Section 4.2.1). We denote K restricted to $S \times S$ by K_S. Note that (c) means $\mathrm{E}X(S) = \sum_{k=1}^{\infty} \lambda_k^S < \infty$. Thus X_B is almost surely finite when $B \in \mathcal{B}$ is bounded.

We also always assume that

(d) for any compact set $S \subseteq \Omega$, all eigenvalues satisfy $0 \leq \lambda_k^S \leq 1$.

In fact, under (a)–(c), the existence of the DPP with kernel K is equivalent to (d) (see e.g. (Hough et al., 2009, Theorem 4.5.5)), and the DPP is then unique (Hough et al., 2009, Lemma 4.2.6). If $\Lambda = \mathbb{R}^d$, ν is 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 (d) is equivalent to $0 \leq \hat{K}_0 \leq 1$ (Hough et al., 2009, Proposition 3.1).

We sometimes assume one of the following conditions:

(e) For a given compact set $S \subseteq \Lambda$, K_S is a projection of finite rank n. Thus, without loss of generality, we can assume

$$K(v,w) = \sum_{k=1}^{n} \phi_k^S(v)\overline{\phi_k^S(w)}, \qquad v, w \in S. \tag{2.3}$$

(f) For all compact $S \subseteq \Lambda$, all eigenvalues satisfy that $\lambda_k^S < 1$.

2.1.3 Reduced Palm distributions

Consider an arbitrary point $u \in \Lambda$ with $K(u,u) > 0$. Recall that the reduced Palm distribution of X at u is a point process X^u on Λ with n-th order intensity function

$$\rho^u(u_1, \ldots, u_n) = \rho(u, u_1, \ldots, u_n)/\rho(u).$$

This combined with (2.1) easily shows that X^u is a DPP with kernel

$$K^u(v,w) = K(v,w) - \frac{K(v,u)K(u,w)}{K(u,u)} \qquad v, w \in \Omega, \tag{2.4}$$

see (Shirai and Takahashi, 2003, Theorem 6.5). 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.

Similarly, for any $u_1, \ldots, u_n \in \Lambda$ with $\rho(u_1, \ldots, u_n) > 0$, the reduced Palm distribution of X at $u_1, \ldots, u_n$ has k-th order intensity function

$$\rho^{u_1,\ldots,u_n}(x_1, \ldots, x_k) = \rho(u_1, \ldots, u_n, x_1, \ldots, x_k)/\rho(u_1, \ldots, u_n).$$

Recall Schur's determinant identity:

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(D)\det(A - BD^{-1}C) \tag{2.5}$$

provided D is invertible. Hence by induction and using (2.4)–(2.5), it follows that $X^{u_1,\ldots,u_n}$ is a DPP with kernel

$$K^{u_1,\ldots,u_n}(v,w) = \frac{\det \begin{pmatrix} K(v,w) & K(v,u_1) & \ldots & K(v,u_n) \\ K(u_1,w) & K(u_1,u_1) & \ldots & K(u_1,u_n) \\ \vdots & \vdots & \ddots & \vdots \\ K(u_n,w) & K(u_n,u_1) & \ldots & K(u_n,u_n) \end{pmatrix}}{\det \begin{pmatrix} K(u_1,u_1) & \ldots & K(u_1,u_n) \\ \vdots & \ddots & \vdots \\ K(u_n,u_1) & \ldots & K(u_n,u_n) \end{pmatrix}}, \qquad v, w \in \Lambda, \tag{2.6}$$

cf. (Shirai and Takahashi, 2003, Corollary 6.6). For convenience, set $K^{u_1,\ldots,u_n} = 0$ if $\rho(u_1, \ldots, u_n) = 0$. Note that $\rho(u_1, \ldots, u_n)$, $\rho^{u_1,\ldots,u_n}$, $X^{u_1,\ldots,u_n}$, and $K^{u_1,\ldots,u_n}$ are invariant under permutation of $u_1, \ldots, u_n$, and $X_S^{u_1,\ldots,u_n} := X^{u_1,\ldots,u_n} \cap S$ follows the reduced Palm distribution of X_S at $u_1, \ldots, u_n$.

2.2 Goldman's results

Goldman Goldman (2000) made similar assumptions as in our assumptions (a)-(d), and he assumed condition (f) 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 in distribution $X^u \subseteq X$.

(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

$$K(v,w) = \frac{1}{\pi} \exp\left(v\overline{w} - \frac{|v|^2 + |w|^2}{2} \right), \qquad v, w \in \mathbb{C}. \tag{2.7}$$

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

It follows from (G1) and (2.4) that $\kappa_u := X \setminus X^u$ is a finite point process with intensity function

$$\rho_{\kappa_u}(v) = |K(u,v)|^2 / K(u,u), \qquad v \in \Lambda. \tag{2.8}$$

Note that the standard Ginibre point process is stationary and isotropic with intensity $1/\pi$, but its kernel is only isotropic. In accordance with (G2), combining (2.7) and (2.8), ρ_{κ_u} is immediately seen to be the density of $N_{\mathbb{C}}(u, 1)$.

3 Main results and discussion of repulsiveness in DPPs

From (2.1) 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

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

where the equality holds if and only if the DPP is a Poisson process. This inequality shows that DPPs are indeed repulsive. In this section we try to quantify how repulsive they can be by studying the repulsive effect of a fixed point contained in a DPP.

3.1 A measure of repulsiveness

The results in (G1)-(G2) suggest to use ρ_{κ_u} when quantifying the repulsive effect of having a point of X at u when $K(u,u) > 0$. Note that $\rho_{\kappa_u} = \rho(v) - \rho^u(v)$ is the difference of the intensity functions of X and X^u. Considering (G1), ρ_{κ_u} is the intensity function for the points in X “pushed out" by u under the Palm distribution. 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

$$1 - g(u,v) = |r(u,v)|^2,$$

where $r(v,w) = K(v,w)/\sqrt{K(v,v)K(w,w)}$ is the correlation function obtained from K. Then

$$\rho_{\kappa_u}(v) = \rho(v)(1 - g(u,v)), \tag{3.1}$$

which shows a trade-off between intensity and repulsiveness (see also Lavancier et al. (2015, 2014)). Because of (3.1), when using ρ_{κ_u} to compare repulsiveness in two DPPs, they should share the same intensity function. Taking this into account, small/high values of ρ_{κ_u} correspond to small/high degree of repulsiveness.

As a global measure of repulsiveness in X when having a point of X at u, we suggest

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

This is the expected number of points in κ_u. Apart from a constant (given by the intensity of X), p_u is in agreement with the measure introduced in Lavancier et al. (2015, 2014) for stationary DPPs on $\mathbb{R}^d$; see also Biscio and Lavancier (2016); Baccelli and O'Reilly (2017). The following proposition shows that p_u is a probability which relates to the spectral decomposition of K restricted to a compact set increasing towards Λ.

Proposition 3.1. *Assume (a)–(d) and let $u \in \Lambda$ with $K(u,u) > 0$. Then $0 \le p_u \le 1$. Moreover, $p_u = 1$ if and only if for an increasing sequence of compact sets $S_1 \subseteq S_2 \subseteq \dots$ so that $u \in S_1$ and $\Lambda = \cup_{i=1}^{\infty} S_i$, for all the eigenvalues in the spectral decomposition of K_{S_i} we have that $\lambda_k^{S_i}$ tends to 0 or 1 when $\phi_k^{S_i}(u) \neq 0$, or more precisely*

$$\lim_{i\to\infty} \sum_k \lambda_k^{S_i}(1-\lambda_k^{S_i})|\phi_k^{S_i}(u)|^2 = 0.$$

Consequently, 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 $v, w \in \Lambda$,

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

For short we then say that X is a projection DPP. The standard Ginibre point process given by (2.7) is globally most repulsive but its kernel is not a projection. 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 proof of Proposition 3.1 that

$$p_u = \frac{\sum_k (\lambda_k^{\Lambda})^2 |\phi_k^{\Lambda}(u)|^2}{\sum_k \lambda_k^{\Lambda} |\phi_k^{\Lambda}(u)|^2}. \tag{3.2}$$

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. In Section 3.4, we establish a coupling for any projection DPP on a compact set and with a non-zero kernel so that $X^u \subseteq X$ in distribution and $X \setminus X^u$ consists almost surely of one point whose density then agrees with ρ_{κ_u}.

3.2 Examples

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

3.2.1 DPPs defined on a finite set

Assume $\Lambda = \{1, \ldots, n\}$ is finite and ν is 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.

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 $u \in \{1, \ldots, n\}$, we have $p_u = 1$ and

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

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

The kernel of a Poisson process with intensity function $\rho \leq 1$ and conditioned on having no multiple points is given by a diagonal covariance matrix with diagonal entries $\rho(1), \ldots, \rho(n)$. If $\rho(u) > 0$, then $p_u = \rho(u)$.

This is a much different result as when we consider a Poisson process on a space Λ where the reference measure ν is diffuse: If $\rho(u) > 0$, then $p_u = 0$ and $X = X^u$ almost surely.

3.2.2 Ginibre point processes

From the standard Ginibre point process given by (2.7), other stationary point processes can be obtained. First, independently thinning the process with a retention probability $\beta \in (0, 1]$, and second, multiplying each of the retained points by $\sqrt{\beta}$ gives a new stationary DPP. Third, multiplying the kernel of this new DPP with a parameter $\alpha \in (0, 1/\beta]$, this result in a stationary DPP X with kernel

$$K(v, w) = \frac{\alpha}{\pi} \exp\left(\frac{v\overline{w}}{\beta} - \frac{|v|^2 + |w|^2}{2\beta} \right), \qquad v, w \in \mathbb{C}, \tag{3.3}$$

and if $f_u = \rho_{\kappa_u}/p_u$ denotes the density of a point in $\kappa_u = X \setminus X^u$, we have

$$\rho = \alpha/\pi, \qquad p_u = \alpha\beta, \qquad f_u(v) = \frac{\exp\left(-|v - u|^2/\beta\right)}{\pi\beta} \sim N_{\mathbb{C}}(u, \beta). \tag{3.4}$$

The case where $\alpha = 1$ and $0 < \beta \leq 1$ is mentioned in Goldman's paper Goldman (2000), and the results in (3.4) match those in (Goldman, 2000, Remark 24). Deng et al. (2015) called the DPP with kernel (3.3) the scaled β-Ginibre point process but the bound $\alpha\beta \leq 1$ 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.

3.2.3 DPPs on $\mathbb{R}^d$ with a stationary kernel

Suppose $\Lambda = \mathbb{R}^d$, ν is 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. (Lavancier et al., 2014, Appendix J). 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

$$K(v,w) = \int_{|y|^d \le d\Gamma(d/2)/(2\pi^{1+d/2})} \exp\bigl(2\pi i (v-w)\cdot y\bigr)\,\mathrm{d}y, \qquad v,w \in \mathbb{R}^d, \tag{3.5}$$

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

$$K(v,w) = J_1(2|v-w|)/(\pi|v-w|), \tag{3.6}$$

where J_1 is the Bessel function of order one. We straightforwardly obtain the following proposition, where the moments in (3.7) follow from (DLM, Eq. 10.22.57).

Proposition 3.2. *For the globally most repulsive DPP on $\mathbb{R}^d$ with kernel given by (3.5) and for any $u \in \mathbb{C}$, we have that $\rho_{\kappa_u}(v) = \pi|K(u,v)|^2$ is a probability density function. In particular, for $d = 2$,*

$$\rho_{\kappa_u}(v) = J_1(2|v-u|)^2 / \left(\pi|v-u|^2\right), \qquad v \in \mathbb{R}^2,$$

and the moments of $|Z_u - u|$ with $Z_u \sim \rho_{\kappa_u}$ satisfy

$$\mathrm{E}\left(|Z_u - u|^k\right) = \frac{\Gamma(1+k/2)\Gamma(1-k)}{\Gamma(2-k/2)\Gamma(1-k/2)^2}, \qquad k \in (-2,1), \tag{3.7}$$

and are infinite for $k \ge 1$.

For comparison consider a standard Ginibre point process, where we can define Z_u in a similar way as in Proposition 3.2. 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 k-th 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 (3.6), $|Z_u - u|$ is heavy-tailed and has infinite k-th moments for all $k \ge 1$.

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 3.2.2 (replacing $\sqrt{\beta}$ by $\beta^{1/d}$ when transforming the points in the thinned process), we obtain a new DPP with kernel

$$K_{\mathrm{new}}(v,w) = \alpha K(v/\beta^{1/d}, w/\beta^{1/d}), \qquad v,w \in \mathbb{R}^d,$$

where $\beta \in (0,1]$ and $\alpha \in (0, 1/\beta]$. The existence of this new DPP follows from the fact that the eigenvalues of K_{new} restricted to a compact set $S \subset \mathbb{R}^d$ are given by $\alpha\beta\lambda_k^{S/\beta^{1/d}} \in [0,1]$, where $\lambda_k^{S/\beta^{1/d}}$ is an eigenvalue for $K_{S/\beta^{1/d}}$. For instance, if K is the jinc-like kernel for the globally most repulsive DPP given by (3.6), the new DPP satisfies the same equations for its intensity and the value of p_u as in (3.4). 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 a point in κ_u now becomes

$$f_u(v) = J_1\left(2|v-u|^2/\beta\right) / \left(\pi|v-u|^2/\beta\right). \tag{3.8}$$

The reach of the repulsive effect of the point at u is much different when comparing the densities in (3.4) and (3.8), in particular if β is large.

3.2.4 DPPs on $\mathbb{S}^d$ with an isotropic kernel

Suppose $\Lambda = \mathbb{S}^d$ is the d-dimensional unit sphere, ν is 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 Schoenberg (1942) and by (Møller et al., 2018, 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

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

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

$$\lambda_{\ell,d} = \rho\sigma_d\beta_{\ell,d}/m_{\ell,d}, \qquad \ell = 0, 1, \dots,$$

with multiplicities

$$m_{0,1}, \qquad m_{\ell,1} = 2, \qquad \ell = 1, 2, \dots, \qquad \text{if } d = 1,$$

and

$$m_{\ell,d} = \frac{2\ell+d-1}{d-1}\frac{(\ell+d-2)!}{\ell!(d-2)!}, \qquad \ell = 0, 1, \dots, \qquad \text{if } d \in \{2, 3, \dots\}.$$

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 (3.2), we obtain

$$p_u = \rho\sigma_d \sum_{\ell=0}^{\infty} \beta_{\ell,d}^2/m_{\ell,d}. \tag{3.9}$$

There is a lack of flexible parametric DPP models on the sphere where K_0 is expressible on closed form, see (Møller et al., 2018, 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 (Møller et al., 2018, Section 4.3.2), the sequence

$$\beta_{\ell,2} = (1-\delta)\delta^\ell, \qquad \ell = 0, 1, \dots, \tag{3.10}$$

specifies a geometric distribution and

$$\lambda_{\ell,2} = 4\pi\rho\delta^\ell(1-\delta)/(2\ell+1) \leq \delta^\ell/(2\ell+1), \qquad \ell = 0, 1, \ldots$$

As $\delta \to 0$, then $\lambda_{0,2} \to 4\pi\rho$ and $\lambda_{\ell,2} \to 0$ if $\ell \geq 1$, 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 (3.9) and (3.10) we obtain

$$p_u = 4\pi\rho(1-\delta)/(1+\delta) \leq 1/(1+\delta),$$

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 (Møller et al., 2018, 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.

3.2.5 Remark

The considerations in Section 3.1 and in Section 3.2.1-3.2.4 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, a globally most repulsive stationary point process on $\mathbb{R}^2$ should be of the form $Y = L_Z := \{x + Z : x \in L\}$, 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. Döge et al. (2004); Mase et al. (2001)). 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.

3.3 Coupling and stochastic domination of point processes

Before stating our main results in Section 3.4-3.5 we need to recall what is meant by coupling and stochastic domination of point processes.

The state space for a DPP is the set $\mathcal{N}$ of all locally finite subsets of Λ. This state space is equipped with the σ-algebra $\mathcal{F}$ generated by all events $\{x \in \mathcal{N} \,|\, x(B) = m\}$ for bounded $B \in \mathcal{B}$ and $m = 0, 1, \ldots$, where $x(B)$ is the cardinality of $x \cap B$. An event $\mathcal{A} \in \mathcal{F}$ is called decreasing if $x \in \mathcal{A}$ implies $y \in \mathcal{A}$ whenever $y \subset x$. Let $\mathcal{F}_d \subset \mathcal{F}$ denote the collection of events consisting of finite unions of elementary decreasing events of the form

$$\{x \in \mathcal{N} \,|\, x(A_i) \leq k_i, \ i = 1, \ldots, m\},$$

where m is a positive integer, $k_1, \ldots, k_m$ are non-negative integers, and $A_1, \ldots, A_m \in \mathcal{B}$ are disjoint and bounded.

Consider two point processes (e.g. DPPs) Y and Z with state space $\mathcal{N}$. Then Z is said to stochastically dominate Y if $\mathrm{P}(Z \in \mathcal{A}) \leq \mathrm{P}(Y \in \mathcal{A})$ for every decreasing event $\mathcal{A} \in \mathcal{F}$. By Strassen's theorem Lindvall (1999), this is equivalent to the existence of a coupling between Y and Z, that is, a joint distribution for (Y, Z) so that $Y \subseteq Z$ almost surely; we call this a monotone coupling of Y w.r.t. Z. The following lemma is well-known, where an outline of the proof is given in the Appendix of Goldman (2000).

Lemma 3.1. *The point process Y is stochastically dominated by the point process Z if and only if, for every $\mathcal{A} \in \mathcal{F}_d$,*

$$\mathrm{P}(Z \in \mathcal{A}) \leq \mathrm{P}(Y \in \mathcal{A}).$$

3.4 Coupling for projection DPPs on a compact set

In this section we assume that $S \subseteq \Lambda$ is a given compact set and that K_S is a projection, cf. (2.3). The restriction of the reference measure ν to S is denoted ν_S. The following theorem is our first main result.

Theorem 3.1. *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*

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

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

Theorem 3.1 immediately implies the following special case.

Corollary 3.1. *Assume $S \subseteq \Lambda$ is compact, $u \in S$ with $K(u, u) > 0$, and X is a DPP with a kernel K so that K_S is a projection of finite rank. Then there exists a monotone coupling of X_S^u w.r.t. X_S such that $\eta_S^u := X_S \setminus X_S^u$ consists of one point almost surely and η_S^u has density $|K_S(\cdot, u)|^2/\sqrt{K_S(u, u)}$ w.r.t. ν_S.*

3.5 Coupling for DPPs on a compact set with eigenvalues < 1

Our second main result is Theorem 3.2 below, which applies when a DPP is restricted to a compact set so that its kernel has all eigenvalues < 1. Compared to the case of a projection DPP, our coupling construction for X_S and X_S^u in Theorem 3.2 becomes more complicated as we not only remove but may also move a point in X_S – basically because now in general $K_S(\cdot, u)$ is not an eigenfunction of K_S, and hence K_S^u and $K_S - K_S^u$ are not $L^2(S)$-orthogonal.

3.5.1 The random projection DPP construction

Theorem 3.2 relies on Proposition 3.3 below together with the following fact for a general DPP restricted to a compact set: Let $S \subseteq \Lambda$ be compact, and recall that the DPP X has a kernel K which is assumed to satisfy (a)-(d). Then we construct a point process Y_S as follows. Consider the spectral representation (2.2) for K_S, and let $B_1^S, B_2^S, \ldots$ be independent Bernoulli variables with means $\lambda_1^S, \lambda_2^S, \ldots$, respectively. By (d), with probability one, $\sum_{k=1}^{\infty} B_k^S$ is finite. For integers $n_1, \ldots, n_k$ with $k \geq 0$, define the Bernoulli variable

$$B^S_{n_1,\ldots,n_k} = \left(\prod_{i=1}^{k} B^S_{n_i}\right) \prod_{j \in \mathbb{N} \setminus \{n_1,\ldots,n_k\}} (1 - B^S_{n_j}),$$

where $B^S_{n_1,\ldots,n_k} = B^S_{\emptyset} = \prod_{j=1}^{\infty}(1 - B^S_{n_j})$ if $k = 0$. Conditional on $B^S_{n_1,\ldots,n_k} = 1$, let $Y_S = X_{S,n_1,\ldots,n_k}$ be the projection DPP with kernel

$$K_{S,n_1,\ldots,n_k} = \sum_{i=1}^{k} \phi^S_{n_i}(u)\overline{\phi^S_{n_i}(v)}, \qquad u, v \in S, \tag{3.11}$$

which is equal to 0 if $k = 0$. Then X_S and Y_S are identically distributed, cf. (Hough et al., 2006, Theorem 7). We refer to this doubly stochastic construction of Y_S as the random projection DPP construction. For $S \subset \mathbb{R}^d$ and $S = \mathbb{S}^d$, Lavancier et al. (2014, 2015); Møller and Rubak (2016) provides a simple simulation algorithm based on this construction, but it will not be needed in the following.

By Corollary 3.1, if $k > 0$, there exists a point process $\eta^u_{n_1,\ldots,n_k}$ which is given by a point in $X_{n_1,\ldots,n_k}$ such that in distribution,

$$X^u_{n_1,\ldots,n_k} = X_{n_1,\ldots,n_k} \setminus \eta^u_{n_1,\ldots,n_k}. \tag{3.12}$$

Further, for $n \in \mathbb{N} \setminus \{n_1, \ldots, n_k\}$, by Theorem 3.1, there is a point process $\xi_{n_1,\ldots,n_k;n}$ which consists of one point such that in distribution, $\xi_{n_1,\ldots,n_k;n} \cap X_{n_1,\ldots,n_k} = \emptyset$ and

$$X_{n_1,\ldots,n_k,n} = X_{n_1,\ldots,n_k} \cup \xi_{n_1,\ldots,n_k;n}. \tag{3.13}$$

Finally, there is a joint distribution for $X_{n_1,\ldots,n_k}$, $X_{n_1,\ldots,n_k,n}$, and $X^u_{n_1,\ldots,n_k,n}$: Conditional on $X_{n_1,\ldots,n_k,n}$, let $\xi_{n_1,\ldots,n_k;n}$ and $\eta^u_{n_1,\ldots,n_k,n}$ be independent. Thus, in distribution,

$$X^u_{n_1,\ldots,n_k,n} = \{X_{n_1,\ldots,n_k} \cup \xi_{n_1,\ldots,n_k;n}\} \setminus \eta^u_{n_1,\ldots,n_k,n}. \tag{3.14}$$

Note that $\xi_{n_1,\ldots,n_k;n}$ has density $|\phi_n(\cdot)|^2$ w.r.t. ν_S, and $\eta^u_{n_1,\ldots,n_k}$ conditioned on $X_{n_1,\ldots,n_k}$ has density $|K_{n_1,\ldots,n_k}(\cdot, u)|^2/\sqrt{K_{n_1,\ldots,n_k}(u,u)}$ w.r.t. ν_S. The semicolon is there to emphasize that $\xi_{n_1,\ldots,n_k;n}$ has the same distribution for all permutations of $n_1, \ldots, n_k$ but not for all permutations of $n_1, \ldots, n_k, n$. Also note that we suppress in the notation that $K_{n_1,\ldots,n_k}$, $X_{n_1,\ldots,n_k}$, $X^u_{n_1,\ldots,n_k}$, $\eta^u_{n_1,\ldots,n_k}$, and $\xi_{n_1,\ldots,n_k}$ depend on S and the choice of orthonormal basis functions for the spectral representation of K.

3.5.2 Coupling results

We need the following coupling results for the counts $X(S)$ and $X^u(S)$.

Proposition 3.3. *Assume (a)–(d) and (f) are satisfied, and let $u \in \Lambda$ with $K(u,u) > 0$. For any compact set $S \subseteq \Lambda$ with $u \in S$, there is a coupling of $X(S)$ and $X^u(S)$ so that $0 \leq X(S) - X^u(S) \leq 1$. Specifically,*

$$\mathrm{P}(B^S_{n_1,\dots,n_k} = 1) = \Big(\prod_\ell (1-\lambda^S_\ell)\Big) \prod_{i=1}^k \frac{\lambda^S_{n_i}}{1-\lambda^S_{n_i}}, \tag{3.15}$$

and for $k = 0, 1, \dots$ and $\mathrm{P}(B^S_{n_1,\dots,n_k} = 1) > 0$, we have

$$\mathrm{P}\big(X(S) = X^u(S) = k \,|\, B^S_{n_1,\dots,n_k} = 1\big) = \frac{\sum_{n \notin \{n_1,\dots,n_k\}} \lambda^S_n |\phi^S_n(u)|^2}{K(u,u)} \tag{3.16}$$

and

$$\mathrm{P}\big(X^u(S) = k, X(S) = k+1 \,|\, B^S_{n_1,\dots,n_{k+1}} = 1\big) = \frac{\sum_{i=1}^{k+1} \lambda^S_{n_i} |\phi^S_{n_i}(u)|^2}{K(u,u)}. \tag{3.17}$$

Now, the coupling procedure for X_S and X^u_S is as follows.

(i) Generate the independent Bernoulli variables $B^S_1, B^S_2, \dots$ Suppose this gives the realization $B^S_{n_1,\dots,n_k} = 1$, with $k \geq 1$, and condition on that realization in the following steps (if $B^S_\emptyset = 1$, then return $X_S = \emptyset$ and $X^u_S = \emptyset$).

(ii) Generate $X_{n_1,\dots,n_k}$ and return $X_S = X_{n_1,\dots,n_k}$.

(iii) Set $X(S) = k$. Generate $X^u(S)$ conditional on $X(S) = k$, cf. (3.16)–(3.17).

(iv) The case $X^u(S) \neq X(S)$: Conditional on $X_{n_1,\dots,n_k}$, generate $\eta^u_{n_1,\dots,n_k}$ in accordance to (3.12). Return $X^u_S = X_{n_1,\dots,n_k} \setminus \eta^u_{n_1,\dots,n_k}$.

(v) The case $X^u(S) = X(S)$: Let $N_{n_1,\dots,n_k}$ be a discrete random variable with distribution

$$\mathrm{P}(N_{n_1,\dots,n_k} = n) = \frac{\lambda^S_n |\phi^S_n(u)|^2}{\sum_{m \notin \{n_1,\dots,n_k\}} \lambda^S_m |\phi^S_m(u)|^2}, \quad n \in \mathbb{N} \setminus \{n_1,\dots,n_k\}. \tag{3.18}$$

Generate first a realization $N_{n_1,\dots,n_k} = n$ and next $\xi_{n_1,\dots,n_k;n}$ in accordance to (3.13). Set $X_{n_1,\dots,n_k,n} = X_{n_1,\dots,n_k} \cup \xi_{n_1,\dots,n_k;n}$ and generate $\eta^u_{n_1,\dots,n_k,n}$, cf. (3.12). Return $X^u_S = X_{n_1,\dots,n_k,n} \setminus \eta^u_{n_1,\dots,n_k,n}$.

Note that we are assuming the following.

- Conditional on $B^S_{n_1,\dots,n_k} = 1$, in (ii) and (iii), $X_{n_1,\dots,n_k}$ and $(X(S), X^u(S))$ are independent.
- Conditional on $B^S_{n_1,\dots,n_k} = 1$, in (v), $N_{n_1,\dots,n_k}$ is independent of $(X_{n_1,\dots,n_k}, X(S), X^u(S))$.
- Conditional on both $B^S_{n_1,\dots,n_k} = 1$, $X_{n_1,\dots,n_k}$, $X^u(S) = X(S)$, and $N_{n_1,\dots,n_k} = n$, in (v), the point in $\xi_{n_1,\dots,n_k;n}$ has density $|\phi^S_n(\cdot)|^2$ w.r.t. ν_S, and this density depends only on n.

- Conditional on both $B^S_{n_1,\dots,n_k} = 1$, $X^u(S) = X(S)$, $N_{n_1,\dots,n_k} = n$, and $X_{n_1,\dots,n_k} \cup \xi_{n_1,\dots,n_k;n}$, in (v), $\eta^u_{n_1,\dots,n_k,n}$ is independent of $\xi_{n_1,\dots,n_k;n}$.

Though $\xi_{n_1,\dots,n_k;n}$ does not depend on u, both ξ^u_S and η^u_S introduced in the following theorem will depend on u, because we let $\xi^u_S = \emptyset$ and $\eta^u_S = \eta^u_{n_1,\dots,n_k}$ if $X^u(S) \neq X(S)$, and let $\xi^u_S = \xi_{n_1,\dots,n_k;n}$ and $\eta^u_S = \eta^u_{n_1,\dots,n_k,n}$ if $X^u(S) = X(S)$.

Theorem 3.2. *Assume (a)–(d) and (f) are satisfied. Let $u \in \Lambda$ with $K(u,u) > 0$. For any compact set $S \subseteq \Lambda$ with $u \in S$, the coupling procedure (i)–(v) correctly returns the DPP X_S with kernel K_S and the DPP X^u_S with kernel K^u. Further, it establishes a coupling such that X^u_S is obtained by either removing or moving a point in X_S, that is, there exist point processes ξ^u_S and η^u_S such that almost surely ξ^u_S is either empty or consists of a single point, η^u_S consists of a single point, and*

$$X^u_S = \{X_S \cup \xi^u_S\} \setminus \eta^u_S, \qquad \xi^u_S \cap X_S = \emptyset, \qquad \eta^u_S \subseteq X_S \cup \xi^u_S. \tag{3.19}$$

For the coupling construction in Theorem 3.2, there is a positive probability that no change happens, i.e. $\eta^u_S = \xi^u_S$ and hence $X_S = X^u_S$. Note that there may be a lack of consistency unless $X_S \neq \emptyset$ implies $X^u_S \subsetneq X_S$: For compact subsets $A \subsetneq B$ of Λ, suppose we claim there is a consistent coupling of (X_A, X^u_A) and (X_B, X^u_B), meaning that $X_A \subseteq X_B$ and $X^u_A \subseteq X^u_B$. Then, if both (X_A, X^u_A) and (X_B, X^u_B) satisfy (3.19), η^u_A becomes the restriction of η^u_B to A, and ξ_A becomes the restriction of ξ_B to A. It is impossible to have $\xi_B \cap A \neq \emptyset$ and $\eta^u_B \cap B \setminus A \neq \emptyset$, because if this does happen, then X^u_A will have one more point than X_A which is impossible for our joint distribution of (X_A, X^u_A). Increasing A to B implies that $\xi_B = \emptyset$, and hence X^u_B is included in X_B, the difference being η^u_B if $X_B \neq \emptyset$. Similarly, assuming Λ is unbounded and there is a coupling as in (3.19) but with S replaced by Λ, unless $X \neq \emptyset$ implies $X^u \subsetneq X$, we face the same problem of inconsistency when restricting X and X^u to a compact subset of Λ.

Despite this inconsistency, Theorem 3.2 provides interesting insight into the repulsive behaviour of DPPs as discussed after establishing the following corollary.

Corollary 3.2. *Under the conditions in Theorem 3.2, we have*

$$p^u_S := \mathrm{P}\left(\xi^u_S = \emptyset\right) = \int_S \frac{|K(u,v)|^2}{K(u,u)}\,\mathrm{d}\nu(v) = \frac{\sum_{n\geq 1} |\lambda^S_n \phi^S_n(u)|^2}{K(u,u)}. \tag{3.20}$$

Further, conditional on $\xi^u_S \neq \emptyset$, ξ^u_S has a density $f_{\xi^u_S}(\cdot \mid \xi^u_S \neq \emptyset)$ w.r.t. ν_S so that

$$(1 - p^u_S)\, f_{\xi^u_S}(v \mid \xi^u_S \neq \emptyset) = \sum_{k=1}^{\infty} \frac{\lambda^S_k(1-\lambda^S_k)|\phi^S_k(u)|^2}{K(u,u)} |\phi^S_k(v)|^2, \qquad v \in S. \tag{3.21}$$

Furthermore, η^u_S has a density w.r.t. ν_S given by

$$f_{\eta^u_S}(v) = \sum_{k=1}^{\infty} \frac{\lambda^S_k(1-\lambda^S_k)|\phi^S_k(u)|^2}{K(u,u)} |\phi^S_k(v)|^2 + \frac{|K(u,v)|^2}{K(u,u)}, \qquad v \in S. \tag{3.22}$$

In general, by the proof of (3.21)–(3.22) in Section 4.6, $f_{\xi_S^u}(v, \xi_S^u \neq \emptyset)$ is different from $f_{\eta_S^u}(v, \xi_S^u \neq \emptyset)$. However, the marginal distributions of ξ_S^u and η_S^u are related as follows. Note that

$$p_S^u = \mathrm{P}(X^u(S) = X(S)) = \mathrm{P}(\xi_S^u = \emptyset),$$

so there is a conditional density $f_S(\cdot \mid \xi_S^u = \emptyset)$ w.r.t. ν_S so that

$$p_S^u f_S(v \mid \xi_S^u = \emptyset) = |K(u,v)|^2/K(u,u), \qquad v \in S.$$

Then, by (3.20)–(3.22),

$$f_{\eta_S^u}(v) = p_S^u f_S(v \mid \xi_S^u = \emptyset) + (1 - p_S^u) f_{\xi_S^u}(v \mid \xi_S^u \neq \emptyset), \tag{3.23}$$

where $f_{\xi_S^u}(\cdot \mid \xi_S^u \neq \emptyset)$ is the conditional density of the point to be added to X_S when obtaining X_S^u and given that $\xi_S^u \neq \emptyset$. Thus, we can view the distribution of η_S^u, the part to be removed from X_S when obtaining X_S^u, as a mixture distribution: With probability p_S^u, η_S^u follows $f_S(\cdot \mid \xi_S^u = \emptyset)$, and else with probability $1 - p_S^u$, η_S^u follows $f_{\xi_S^u}(\cdot \mid \xi_S^u \neq \emptyset)$.

We can quantify repulsiveness in a DPP X by, for any given compact subset $S \subseteq \Lambda$ and any given point $u \in S$ with $K(u,u) > 0$, comparing the marginal distribution of η_S^u with that of ξ_S^u. Referring to the mixture distribution for η_S^u given by (3.23) lead us to consider the case (iv) and to quantify repulsiveness in terms of the intensity function $\kappa^u(v) = |K(u,v)|^2/K(u,u)$ for $v \in \Lambda$, which we studied in Section 3.1–3.2 and which gave rise to the measure p_u. As S increases to Λ and if $K(u,u) > 0$, then p_u is the limiting probability of the event that X_S^u as obtained in (i)–(v) is given by deleting a point in X_S (i.e. the case (iv)), and this point has limiting density $\rho_{\kappa^u}(\cdot)/p_u$.

On the other hand, we can also as in (v) condition on $B^S_{n_1,\dots,n_k} = 1$, $\xi_S^u \neq \emptyset$, and $N_{n_1,\dots,n_k} = n$. Then the point to be added has density $|\phi_n^S(\cdot)|^2$, whilst the point to be taken away has density

$$\frac{|K_{n_1,\dots,n_k,n}(u,\cdot)|^2}{K_{n_1,\dots,n_k,n}(u,u)} = \frac{|\sum_{\ell \in \{n_1,\dots,n_k,n\}} \phi_\ell^S(u)\overline{\phi_\ell^S(\cdot)}|^2}{\sum_{\ell \in \{n_1,\dots,n_k,n\}} |\phi_\ell^S(u)|^2}.$$

Observe that

$$\frac{|K_{n_1,\dots,n_k,n}(u,u)|^2}{K_{n_1,\dots,n_k,n}(u,u)} = \sum_{i=1}^k |\phi_{n_i}^S(u)|^2 + |\phi_n^S(u)|^2 \geq |\phi_n^S(u)|^2.$$

Thus, if the eigenfunctions are continuous, there will be a neighbourhood of u such that the point taken away is more likely to be inside the neighbourhood than the point removed. This provides evidence that conditioning on X_S having a point at u has the effect of pushing a point in X_S farther away.

4 Further results and proofs

In this section we collect our proofs for the main results in Section 3 and we add some technical results.

4.1 Proof of Proposition 3.1

Let $S \subseteq \Lambda$ be compact. As S increases, the integral $\int_S |K(u,v)|^2\,\mathrm{d}\nu(v)$ is non-decreasing. From the spectral representation (2.2) and condition (d) we obtain

$$\begin{aligned}\int_S |K(u,v)|^2\,\mathrm{d}\nu(v) &= \sum_k\sum_\ell \lambda_k^S\lambda_l^S\phi_k^S(u)\overline{\phi_\ell^S(u)}\int_S \overline{\phi_k^S(v)}\phi_\ell^S(v)\,\mathrm{d}\nu(v)\\ &= \sum_k(\lambda_k^S)^2|\phi_k^S(u)|^2 \le \sum_k \lambda_k^S|\phi_k^S(u)|^2 = K(u,u).\end{aligned}$$

Thereby Proposition 3.1 follows.

4.2 Some results for projection DPPs

Lemma 4.1. *Suppose $S \subseteq \Lambda$ is compact, n is a non-negative integer, X is a simple point process on S consisting of n points almost surely, and $\mathcal{A} \in \mathcal{F}_d$. Then there exists a finite partition $S = \bigcup_{i=1}^m B_i$ such that each $B_i \in \mathcal{B}$ and $\mathrm{P}(X \in \mathcal{A})$ can be expressed as a finite sum, up to the sign, of terms of the form*

$$\mathrm{P}\Big(\bigcap_{i=1}^m \{X(B_i) = k_i\}\Big),$$

where $\sum_{i=1}^m k_i = n$.

Proof. For simplicity, let $\mathcal{A} \in \mathcal{F}_d$ be such that $x \in \mathcal{A}$ if and only if $\bigcup_{i=1}^r \{x(A_i) \le k_i\}$. A more complicated but similar argument as given below applies for general events in $\mathcal{F}_d$.

By the inclusion-exclusion principle,

$$\begin{aligned}\mathrm{P}\Big(\bigcup_{i=1}^r \{X(A_i) \le k_i\}\Big) &= (-1)^0\sum_i \mathrm{P}(X(A_i) \le k_i)\\ &\quad + (-1)^1\sum_{i<j}\mathrm{P}(\{X(A_i) \le k_i\} \cap \{X(A_j) \le k_j\})\\ &\quad + \cdots + (-1)^{r-1}\mathrm{P}\Big(\bigcap_{i=1}^r \{X(A_i) \le k_i\}\Big).\end{aligned}$$

Set $A = \bigcup_{i=1}^r A_i$. Let $\{B_j\}_{j=1}^{m-1}$ be a collection of disjoint Borel sets such that $\bigcup_{j=1}^{m-1} B_j = A$ and for every A_i, there exists $I_i \subseteq \{1, \dots m-1\}$ such that $A_i = \cup_{j \in I_i} B_j$. Then, for any $I \subseteq \{1, \dots, n\}$,

$$\begin{aligned}&\mathrm{P}\Big(\bigcap_{i \in I} \{X(A_i) \le k_i\}\Big)\\ &\quad = \sum_{\substack{(\ell_1,\dots,\ell_{m-1}):\\ \sum_{j\in I_i}\ell_j \le k_i\,\forall i \in I}} \mathrm{P}\Big(\bigcap_{j=1}^{m-1} \{X(B_j) = \ell_j\} \cap \{X(S \setminus A) = n - \ell\}\Big),\end{aligned}$$

where in each term, $\ell = \sum_{i=1}^{m-1} \ell_i$, and the term is a probability of X being in an elementary increasing event. Thus the conclusion holds. □

Assume S is finite, ν is counting measure, and X is a projection DPP with kernel Q. Then the matrix $\{Q(u,v)\}_{u,v\in S}$ is complex, Hermitian, has eigenvalues 0 or 1, and its rank is equal to the cardinality of X almost surely. So Q can be identified with the linear subspace $H \subseteq \mathbb{C}^S$ spanned by the columns of Q. We also write X^H for X. The following is a relevant result taken from Russell Lyon's paper (Lyons, 2003, Theorem 6.2) but using our terminology and notation.

Lemma 4.2. *Let S be finite, let $H_1 \subset H_2$ be linear subspaces of $\mathbb{C}^S$. Then X^{H_2} stochastically dominates X^{H_1}.*

4.3 Proof of Theorem 3.1

By Lemma 3.1, for the existence of a monotone coupling of Y w.r.t. X, it suffices to show that for any event $\mathcal{A} \in \mathcal{F}_d$,

$$\mathrm{P}(X \in \mathcal{A}) \leq \mathrm{P}(Y \in \mathcal{A}). \tag{4.1}$$

In (Goldman, 2000, Theorem 3), Goldman proved this in the case where all eigenvalues of the associated kernels K and L are strictly less than one and the difference kernel $K - L$ is positive definite. The proof can be modified as follows for the case of projection DPPs, where all eigenvalues are exactly one.

So fix an event $\mathcal{A} \in \mathcal{F}_d$. By Lemma 4.1, there exists a finite partition of S into Borel sets, $S = \bigcup_{i=1}^m B_i$, such that $\mathrm{P}(X \in \mathcal{A})$ can be expressed as a finite sum, up to the sign, of terms of the form

$$\mathrm{P}\Big(\bigcap_{i=1}^m \{X(B_i) = k_i\}\Big), \tag{4.2}$$

where $n = \sum_{i=1}^m k_i$. Similarly, $\mathrm{P}(Y \in \mathcal{A})$ can be expressed as a finite sum, up to the sign, of terms of the form

$$\mathrm{P}\Big(\bigcap_{i=1}^m \{Y(B_i) = k_i\}\Big),$$

where $\sum_{i=1}^m k_i = n - 1$. For $i = 1, \ldots, m$, define the subspaces

$$V_i = \mathrm{span}\{\phi_k^S 1_{B_i}\}_{k=1}^n$$

and let $n_i = \dim(V_i)$. Then, there exist orthonormal vectors

$$z^k = (z_1^k, \ldots, z_M^k) \in \prod_{i=1}^M \mathbb{C}^{n_i}, \qquad k = 1, \ldots, n,$$

such that for all integers $1 \leq i \leq m$, $1 \leq k \leq n$, and $1 \leq \ell \leq n$,

$$\sum_{j=1}^{n_i} z_i^k(j)\overline{z_i^\ell(j)} = \int_{B_i} \phi_k^S(v)\overline{\phi_\ell^S(v)}\,\mathrm{d}\nu(v), \tag{4.3}$$

where $z_i^k = (z_i^k(1), \dots, z_i^k(n_i))$. Indeed, let $\{e_j^i\}_{j=1}^{n_i}$ be an orthonormal basis of $V_i \subset L^2(S)$. Then

$$\phi_k^S 1_{B_i} = \sum_{j=1}^{n_i} \alpha_{i,j}^k e_j^i \qquad \text{for some } \{\alpha_{i,j}^k\} \subset \mathbb{C}.$$

Letting $z_i^k(j) = \alpha_{i,j}^k$ gives (4.3).

Let $E_i = \{(i,j) \,|\, j \in \{1, \dots, n_i\}\}$ and define the finite space $E = \bigcup_{i=1}^m E_i$ with counting measure λ_E. Define DPPs X_E and Y_E on E with projection kernels

$$K_E((i_1, j_1), (i_2, j_2)) = \sum_{k=1}^{n} z_{i_1}^k(j_1) \overline{z_{i_2}^k(j_2)}$$

and

$$L_E((i_1, j_1), (i_2, j_2)) = \sum_{k=1}^{n-1} z_{i_1}^k(j_1) \overline{z_{i_2}^k(j_2)},$$

respectively. Then $X_E = X_E^{H_1}$ and $Y_E = Y_E^{H_2}$, where

$$H_1 = \operatorname{span}\{z^1, \dots, z^n\}, \qquad H_2 = \operatorname{span}\{z^1, \dots, z^{n-1}\} \subset H_1.$$

Thus, by Lemma 4.2, X_E stochastically dominates Y_E, meaning that for any decreasing event $\mathcal{A}_E \subseteq 2^E$,

$$\mathrm{P}(X_E \in \mathcal{A}_E) \le \mathrm{P}(Y_E \in \mathcal{A}_E). \tag{4.4}$$

Now, let $\sigma = (\sigma_1, \dots, \sigma_n)$ be a permutation of $(1, \dots, n)$. Note that

$$\det\left(K(x_i, x_j)\right)_{i,j=1}^n = \prod_{i=1}^n K(x_i, x_{\sigma(i)}) = \sum_{k_1, \dots, k_n = 1}^{n} \prod_{i=1}^n \phi_{k_i}^S(x_i) \overline{\phi_{k_{\sigma^{-1}(i)}}^S(x_i)}$$

and

$$\det\left(K_E((i_\ell, j_\ell), (i_m, j_m))\right)_{\ell,m=1}^n = \prod_{\ell=1}^n K_E((i_\ell, j_\ell), (i_{\sigma(\ell)}, j_{\sigma(\ell)}))$$
$$= \sum_{k_1, \dots, k_n = 1}^{n} \prod_{\ell=1}^n z^{k_\ell}(i_\ell, j_\ell) \overline{z^{k_{\sigma^{-1}(\ell)}}(i_\ell, j_\ell)},$$

where $z^k(i,j) = z_i^k(j)$. By (4.3), for sets B_i and integers $k_i \ge 0$ as in (4.2),

$$\int_{\prod_{i=0}^M B_i^{k_i}} \prod_{i=1}^n \phi_{k_i}^S(x_i) \overline{\phi_{k^S_{\sigma^{-1}(i)}}(x_i)} \, \mathrm{d}\nu(x_i)$$
$$= \int_{\prod_{i=1}^m E_i^{k_i}} \prod_{\ell=1}^n z^{k_\ell}(i_\ell, j_\ell) \overline{z^{k_{\sigma^{-1}(\ell)}}(i_\ell, j_\ell)} d\lambda_E((i_\ell, j_\ell)), \tag{4.5}$$

where $n = \sum_{i=1}^m k_i$. Expanding the determinants below and using (4.5) gives

$$\int_{\prod_{i=1}^m B_i^{k_i}} \det(K(x_i, x_j))_{i,j=1}^n \prod_{i=1}^n \mathrm{d}\nu(x_i)$$
$$= \int_{\prod_{i=0}^M E_i^{k_i}} \det(K_E((i_\ell, j_\ell), (i_k, j_k))_{\ell,k=1}^n \prod_{\ell=1}^n \mathrm{d}\lambda_E((i_\ell, j_\ell)),$$

and so

$$\mathrm{P}\Big(\bigcap_{i=1}^{m}\{X(B_i)=k_i\}\Big)=\mathrm{P}\Big(\bigcap_{i=1}^{m}\{X_E(E_i)=k_i\}\Big).$$

Similarly,

$$\mathrm{P}\Big(\bigcap_{i=1}^{m}\{Y(B_i)=k_i\}\Big)=\mathrm{P}\Big(\bigcap_{i=1}^{m}\{Y_E(E_i)=k_i\}\Big),$$

where here $n-1=\sum_{i=1}^{m}k_i$.

Consequently,

$$\mathrm{P}(X_E\in\mathcal{A}_E)=\mathrm{P}(X\in\mathcal{A})\quad\text{and}\quad\mathrm{P}(Y_E\in\mathcal{A}_E)=\mathrm{P}(Y\in\mathcal{A})$$

for a decreasing event $\mathcal{A}_E$. Thus, by (4.4),

$$\mathrm{P}(X\in\mathcal{A})\le\mathrm{P}(Y\in\mathcal{A}),$$

and hence (4.1) is verified. Therefore, there is a coupling such that $Y\subseteq X$ almost surely. As Y has cardinality one less than X, $\eta:=Y\setminus X$ consists of one point almost surely, and for any $A\subseteq S$,

$$\mathbb{P}(\eta\in A)=\mathbb{E}\left[1_{\{Y(A)-X(A)=1\}}\right]=\mathbb{E}[Y(A)]-\mathbb{E}[X(A)]=\int_A|\phi_n^S(\cdot)|^2\,\mathrm{d}x.$$

Thus, the result follows.

4.4 Proof of Proposition 3.3

Equation (3.15) follows immediately from the independence of the Bernoulli variables $B_1^S,B_2^S,\ldots$ We have

$$\mathrm{P}\left(X(S)=0\right)=\prod_{\ell}\left(1-\lambda_\ell^S\right)$$

and by (Goldman, 2000, Equation (34)),

$$\mathrm{P}\left(X^u(S)=k\right)=\frac{\mathrm{P}\left(X(S)=0\right)}{K(u,u)}\sum_n|\phi_n^S(u)|^2\frac{\lambda_n^S}{1-\lambda_n^S}\sum_{\substack{n_1<\ldots<n_k:\\ n\notin\{n_1,\ldots,n_k\}}}\prod_{\ell=1}^{k}\frac{\lambda_\ell^S}{1-\lambda_\ell^S}.$$

Switching the sums above and observing that

$$\lambda_n^S/\left(1-\lambda_n^S\right)=\lambda_n^S+\left(\lambda_n^S\right)^2/\left(1-\lambda_n^S\right)$$

gives

$$
\begin{aligned}
\mathrm{P}(X^u(S)=k) &= \frac{\prod_\ell (1-\lambda_\ell^S)}{K(u,u)} \sum_{n_1<\cdots<n_k} \prod_{i=1}^k \frac{\lambda_{n_i}^S}{1-\lambda_{n_i}^S} \sum_{n\notin\{n_1,\dots,n_k\}} \frac{\lambda_n^S}{1-\lambda_n^S} |\phi_n^S(u)|^2 \\
&= \frac{\prod_\ell (1-\lambda_\ell^S)}{K(u,u)} \Bigg(\sum_{n_1<\cdots<n_k} \prod_{i=1}^k \frac{\lambda_{n_i}^S}{1-\lambda_{n_i}^S} \sum_{n\notin\{n_1,\dots,n_k\}} \lambda_n^S |\phi_n^S(u)|^2 \\
&\qquad\qquad + \sum_{n_1<\cdots<n_k} \prod_{i=1}^k \frac{\lambda_{n_i}^S}{1-\lambda_{n_i}^S} \sum_{n\notin\{n_1,\dots,n_k\}} \frac{(\lambda_n^S)^2}{1-\lambda_n^S} |\phi_n^S(u)|^2 \Bigg) \\
&= \sum_{n_1<\cdots<n_k} \mathrm{P}(B^S_{n_1,\dots,n_k}=1) \frac{\sum_{n\notin\{n_1,\dots,n_k\}} \lambda_n^S |\phi_n^S(u)|^2}{K(u,u)} \\
&\quad + \sum_{n_1<\cdots<n_{k+1}} \mathrm{P}(B^S_{n_1,\dots,n_{k+1}}=1) \frac{\sum_{i=1}^{k+1} \lambda_{n_i}^S |\phi_{n_i}^S(u)|^2}{K(u,u)}.
\end{aligned}
$$

By the random projection DPP construction, $B^S_{n_1,\dots,n_k}=1$ implies $X(S)=k$, and so (3.16)–(3.17) follow.

4.5 Proof of Theorem 3.2

By the random projection DPP construction, the procedure (i)–(v) gives the correct marginal for X_S, and it establishes the coupling in (3.19), so it only remains to prove that it generates the correct marginal distribution of X_S^u. Note that under the assumptions in Theorem 3.2, $\lambda_n^S<1$ for $n=1,2,\dots$, so X_S has a density f_S and X_S^u has a density f_S^u w.r.t. $\exp(\nu(S))$ times the probability measure for the Poisson process on S with intensity measure ν_S, see Macchi (1975) and (Shirai and Takahashi, 2003, Lemma 3.4). Defining

$$\tilde K(v,w)=\sum_{k=1}^\infty \frac{\lambda_k^S}{1-\lambda_k^S}\phi_k^S(v)\overline{\phi_k^S(w)}, \qquad v,w\in S, \tag{4.6}$$

then for multiple distinct points $x_1,\dots,x_n\in S$,

$$f_S(\{x_1,\dots,x_n\})=\Big(\prod_{k=1}^\infty(1-\lambda_k^S)\Big)\det\{\tilde K(x_i,x_j)\}_{i,j=1}^n$$

and

$$f_S^u(\{x_1,\dots,x_n\})=f_S(\{x_1,\dots,x_n,u\})/K(u,u).$$

First, assume for an integer $M>1$, $\lambda_n^S=0$ whenever $n>M$, so

$$K(v,w)=\sum_{n=1}^M \lambda_n^S\phi_n^S(v)\overline{\phi_n^S(w)}, \qquad v,w\in S.$$

Then X_S has at most M points. For $k = 0, \dots, M-1$ and $\{x_1, \dots, x_k\} \subseteq S$,

$$f_S^u(\{x_1, \dots, x_k\})K(u,u) / \prod_{\ell=1}^{M} \left(1 - \lambda_\ell^S\right) = \det(AB),$$

where A is the $(k+1) \times M$ matrix with ijth entry $\frac{\lambda_j^S}{1-\lambda_j^S} \phi_j^S(x_i)$ and B is the $M \times (k+1)$ matrix with ijth entry $\overline{\phi_i^S(x_j)}$, where in both cases, $x_{k+1} = u$. Define the sum notation $\sum_{(n_i)_1^k}^M := \sum_{1 \le n_1 < \cdots < n_k \le M}$, that is, the sum over all ordered subsets of size k in $\{1, \dots, M\}$. Then the Cauchy-Binet formula gives

$$\begin{aligned}
&f_S^u(\{x_1, \dots, x_k\})K(u,u) \Big/ \prod_{\ell=1}^{M} (1 - \lambda_\ell^S) \\
&\quad = \sum_{(n_i)_1^{k+1}}^{M} \Big(\prod_{i=1}^{k+1} \frac{\lambda_{n_i}^S}{1 - \lambda_{n_i}^S} \Big) \det(B^*) \det(B) \\
&\quad = \sum_{(n_i)_1^{k+1}}^{M} \Big(\prod_{i=1}^{k+1} \frac{\lambda_{n_i}^S}{1 - \lambda_{n_i}^S} \Big) f_{n_1, \dots, n_{k+1}}(\{x_1, \dots, x_k, u\}),
\end{aligned}$$

where $f_{n_1,\dots,n_{k+1}}$ is the density of the DPP on S with projection kernel $K_{n_1,\dots,n_{k+1}}$. Further, since

$$f_{n_1,\dots,n_{k+1}}^u(\{x_1, \dots, x_k\}) = f_{n_1,\dots,n_{k+1}}(\{x_1, \dots, x_k, u\}) / K_{n_1,\dots,n_{k+1}}(u,u),$$

we have

$$\begin{aligned}
&\sum_{(n_i)_1^{k+1}}^{M} \Big(\prod_{j=1}^{k+1} \frac{\lambda_{n_j}^S}{1 - \lambda_{n_j}^S} \Big) f_{n_1,\dots,n_{k+1}}(\{x_1, \dots, x_k, u\}) \\
&\quad = \sum_{(n_i)_1^{k+1}}^{M} \Big(\prod_{j=1}^{k+1} \frac{\lambda_{n_j}^S}{1 - \lambda_{n_j}^S} \Big) \sum_{i=1}^{k+1} |\phi_{n_i}^S(u)|^2 f_{n_1,\dots,n_{k+1}}^u(\{x_1, \dots, x_k\}) \\
&\quad = \sum_{(n_i)_1^{k+1}}^{M} \Big(\prod_{j=1}^{k+1} \frac{\lambda_{n_j}^S}{1 - \lambda_{n_j}^S} \Big) \sum_{i=1}^{k+1} (1 - \lambda_{n_i}^S) |\phi_{n_i}^S(u)|^2 f_{n_1,\dots,n_{k+1}}^u(\{x_1, \dots, x_k\}) \\
&\qquad + \sum_{(n_i)_1^{k+1}}^{M} \Big(\prod_{j=1}^{k+1} \frac{\lambda_{n_j}^S}{1 - \lambda_{n_j}^S} \Big) \sum_{i=1}^{k+1} \lambda_{n_i}^S |\phi_{n_i}^S(u)|^2 f_{n_1,\dots,n_{k+1}}^u(\{x_1, \dots, x_k\}) \\
&\quad = \sum_{(n_i)_1^{k+1}}^{M} \sum_{i=1}^{k+1} \Big(\prod_{\substack{j=1,\dots,k+1: \\ j \ne i}} \frac{\lambda_{n_j}^S}{1 - \lambda_{n_j}^S} \Big) \lambda_{n_i}^S |\phi_{n_i}^S(u)|^2 f_{n_1,\dots,n_{k+1}}^u(\{x_1, \dots, x_k\}) \qquad (4.7) \\
&\qquad + \sum_{(n_i)_1^{k+1}}^{M} \Big(\prod_{j=1}^{k+1} \frac{\lambda_{n_j}^S}{1 - \lambda_{n_j}^S} \Big) \sum_{i=1}^{k+1} \lambda_{n_i}^S |\phi_{n_i}^S(u)|^2 f_{n_1,\dots,n_{k+1}}^u(\{x_1, \dots, x_k\}),
\end{aligned}$$

where the term in (4.7) is interpreted as zero when $k = 0$. Therefore,

$$\begin{aligned}
&f_S^u(\{x_1, ..., x_k\}) \Big/ \prod_{\ell=1}^{M}(1-\lambda_\ell^S) \\
&= \sum_{(n_i)_1^k}^{M} \Big(\prod_{j=1}^{k} \frac{\lambda_{n_j}^S}{1-\lambda_{n_j}^S}\Big) \frac{\sum_{\substack{n=1,\dots,M:\\ n\notin\{n_1,\dots,n_k\}}} \lambda_n^S |\phi_n^S(u)|^2 f^u_{n_1,\dots,n_k,n}(\{x_1,\dots,x_k\})}{K(u,u)} \\
&\quad + \sum_{(n_i)_1^{k+1}}^{M} \Big(\prod_{j=1}^{k+1} \frac{\lambda_{n_j}^S}{1-\lambda_{n_j}^S}\Big) \frac{\sum_{i=1}^{k+1} \lambda_{n_i}^S |\phi_{n_i}^S(u)|^2}{K(u,u)} f^u_{n_1,\dots,n_{k+1}}(\{x_1,\dots,x_k\}). \qquad (4.8)
\end{aligned}$$

Second, letting $M \to \infty$, (4.8) gives

$$\begin{aligned}
&f_S^u(\{x_1, ..., x_k\}) \\
&= \prod_{\ell\geq 1}\left(1-\lambda_\ell^S\right) \sum_{(n_i)_1^k} \prod_{j=1}^{k} \frac{\lambda_{n_j}^S}{1-\lambda_{n_j}^S} \sum_{\substack{n\geq 1:\\ n\notin\{n_i\}_{i=1}^k}} \frac{\lambda_n^S|\phi_n^S(u)|^2}{K(u,u)} f^u_{n_1,\dots,n_k,n}(\{x_1,\dots,x_k\}) \\
&\quad + \prod_{\ell\geq 1}\left(1-\lambda_\ell^S\right) \\
&\quad \cdot \sum_{(n_i)_1^{k+1}} \Big(\prod_{j=1}^{k+1} \frac{\lambda_{n_j}^S}{1-\lambda_{n_j}^S}\Big) \frac{\sum_{i=1}^{k+1} \lambda_{n_i}^S |\phi_{n_i}^S(u)|^2}{K(u,u)} f^u_{n_1,\dots,n_{k+1}}(\{x_1,\dots,x_k\}) \\
&= \sum_{(n_i)_1^k} \mathrm{P}\left(B^S_{n_1,\dots,n_k} = 1\right) \sum_{\substack{n\geq 1:\\ n\notin\{n_i\}_{i=1}^k}} \frac{\lambda_n^S|\phi_n^S(u)|^2}{K(u,u)} f^u_{n_1,\dots,n_k,n}(\{x_1,\dots,x_k\}) \qquad (4.9) \\
&\quad + \sum_{(n_i)_1^{k+1}} \mathrm{P}(B^S_{n_1,\dots,n_{k+1}} = 1) \frac{\sum_{i=1}^{k+1} \lambda_{n_i}^S |\phi_{n_i}^S(u)|^2}{K(u,u)} f^u_{n_1,\dots,n_{k+1}}(\{x_1,\dots,x_k\}), \qquad (4.10)
\end{aligned}$$

where the term (4.9) is interpreted as zero if $k = 0$. Further, by (3.16) and (3.18), for any $k \in \mathbb{N}$,

$$\begin{aligned}
\frac{\lambda_n^S|\phi_n^S(u)|^2}{K(u,u)} &= \mathrm{P}\left(X(S) = X^u(S) = k \mid B^S_{n_1,\dots,n_k} = 1\right) \\
&\quad \cdot \mathrm{P}\left(N_{n_1,\dots,n_k} = n \mid B^S_{n_1,\dots,n_k} = 1, X(S) = X^u(S) = k\right),
\end{aligned}$$

whilst $\sum_{i=1}^{k+1} \lambda_{n_i}^S |\phi_{n_i}^S(u)|^2 / K(u,u)$ is given by (3.17). Combining this with (4.9)–

(4.10), it follows that

$$
\begin{aligned}
&f_S^u(\{x_1,\dots,x_k\})\\
&\quad= \sum_{(n_i)_1^k} \mathrm{P}(B^S_{n_1,\dots,n_k}=1)\\
&\qquad\quad \cdot \sum_{\substack{n\geq 1:\\ n\notin\{n_1,\dots,n_k\}}} \mathrm{P}(N_{n_1,\dots,n_k}=n, X(S)=X^u(S)=k \mid B^S_{n_1,\dots,n_k}=1)\\
&\qquad\qquad\qquad \cdot f^u_{n_1,\dots,n_k,n}(\{x_1,\dots,x_k\})\\
&\qquad + \sum_{(n_i)_1^{k+1}} \mathrm{P}(B^S_{n_1,\dots,n_{k+1}}=1)\mathrm{P}\big(X^u(S)=k, X(S)=k+1 \mid B^S_{n_1,\dots,n_{k+1}}=1\big)\\
&\qquad\qquad\qquad \cdot f^u_{n_1,\dots,n_{k+1}}(\{x_1,\dots,x_k\}).
\end{aligned}
$$

Hence, by (3.12)–(3.14), the output in (iv)–(v) generates the correct marginal distribution of X_S^u.

4.6 Proof of Corollary 3.2

As

$$p_S^u = \mathrm{P}\left(X(S) - X^u(S) = 1\right)$$

we obtain immediately the first result in (3.20), and thereby the second result using the spectral representation of K restricted to $S \times S$.

As in step (v), assume that we have conditioned on $B^S_{n_1,\dots,n_k}=1$, $X(S)=X^u(S)$, and $N=n$. Then, as noticed just before Theorem 3.2, $\xi_S^u = \xi_{n_1,\dots,n_k;n}$ (the "added point") has density $|\phi_n^S(\cdot)|^2$ w.r.t. ν_S. If we also condition on $X_{n_1,\dots,n_k,n}$, which is given by $X_{n_1,\dots,n_k}\cup\xi_S^u$, the distribution of $\eta_S^u=\eta_{n_1,\dots,n_k,n}\in X_{n_1,\dots,n_k,n}$ (the "deleted point") is independent of ξ_S^u, and η_S^u has density

$$f_{\eta_S^u}(v \mid B^S_{n_1,\dots,n_k}=1, X^u(S)=X(S)=k, N=n) = \frac{|K_{n_1,\dots,n_k,n}(u,v)|^2}{K_{n_1,\dots,n_k,n}(u,u)}$$

w.r.t. ν_S.

Consequently, the density of η_S^u when $X(S)=X^u(S)$ is

$$
\begin{aligned}
&f_{\eta_S^u}(v, X(S)=X^u(S))\\
&\quad= \sum_{k\geq 0}\sum_{(n_i)_1^k} f_\eta(v, B^S_{n_1,\dots,n_k}=1, X(S)=X^u(S)=k, N=n)\\
&\quad= \sum_{k\geq 0}\sum_{(n_i)_1^k}\sum_{n\notin\{n_i\}_{i=1}^k} \mathrm{P}(B^S_{n_1,\dots n_k}=1)\mathrm{P}(X(S)=X^u(S)=k \mid B^S_{n_1,\dots n_k}=1)\\
&\qquad\quad \cdot \mathrm{P}(N=n \mid B^S_{n_1,\dots n_k}=1, X(S)=X^u(S)=k)\frac{|K_{n_1,\dots,n_k,n}(u,v)|^2}{K_{n_1,\dots,n_k,n}(u,u)},
\end{aligned}
$$

where $\sum_{(n_i)_1^k}\sum_{n\notin\{n_i\}_{i=1}^k}$ is interpreted as $\sum_{n=1}^\infty$ when $k=0$. Setting $\prod_{i=1}^k \cdots = 1$

when $k = 0$, we have

$$
\begin{aligned}
\mathrm{P}(B^S_{n_1,\dots n_k} = 1) &= \Big(\prod_{\ell\ge 1}(1-\lambda^S_\ell)\Big)\prod_{i=1}^k \frac{\lambda^S_{n_i}}{1-\lambda^S_{n_i}},\\
\mathrm{P}(X(S) = X^u(S) = k \,|\, B^S_{n_1,\dots n_k} = 1) &= \frac{\sum_{m\notin\{n_1,\dots,n_k\}} \lambda^S_m |\phi^S_m(u)|^2}{K(u,u)},\\
\mathrm{P}(N = n \,|\, B^S_{n_1,\dots n_k} = 1, X(S) = X^u(S) = k) &= \frac{\lambda^S_n |\phi^S_n(u)|^2}{\sum_{m\notin\{n_1,\dots,n_k\}} \lambda^S_m |\phi^S_m(u)|^2},
\end{aligned}
$$

so

$$
\begin{aligned}
&f_{\eta^u_S}(v, X(S) = X^u(S))\Big/\prod_{\ell\ge1}(1-\lambda^S_\ell)\\
&\quad= \sum_{k\ge0}\sum_{(n_i)_1^k}\Big(\prod_{i=1}^k \frac{\lambda^S_{n_i}}{1-\lambda^S_{n_i}}\Big)\sum_{n\ge1, n\notin\{n_1,\dots,n_k\}} \frac{\lambda^S_n|\phi^S_n(u)|^2}{K(u,u)}\frac{|K_{n_1,\dots,n_k,n}(u,v)|^2}{K_{n_1,\dots,n_k,n}(u,u)}\\
&\quad= \sum_{k\ge0}\sum_{(n_i)_1^k}\sum_{n\notin\{n_i\}_{i=1}^k}\Big(\prod_{i=1}^k \frac{\lambda^S_{n_i}}{1-\lambda^S_{n_i}}\Big)\frac{\lambda^S_n}{1-\lambda^S_n}\frac{(1-\lambda^S_n)|\phi^S_n(u)|^2}{K(u,u)}\frac{|K_{n_1,\dots,n_k,n}(u,v)|^2}{K_{n_1,\dots,n_k,n}(u,u)}\\
&\quad= \sum_{k\ge0}\sum_{(n_i)_1^{k+1}}\Big(\prod_{i=1}^{k+1} \frac{\lambda^S_{n_i}}{1-\lambda^S_{n_i}}\Big)\sum_{i=1}^{k+1}\frac{(1-\lambda^S_{n_i})|\phi^S_{n_i}(u)|^2}{K(u,u)}\frac{|K_{n_1,\dots,n_{k+1}}(u,v)|^2}{K_{n_1,\dots,n_{k+1}}(u,u)}.
\end{aligned}
$$

Similarly, the density of η^u_S when $X(S) \neq X^u(S)$ is given by

$$
\begin{aligned}
&f_{\eta^u_S}(v, X(S) \neq X^u(S))\Big/\prod_{\ell\ge1}(1-\lambda^S_\ell)\\
&\quad= \sum_{k\ge0}\sum_{(n_i)_1^{k+1}}\Big(\prod_{i=1}^{k+1} \frac{\lambda^S_{n_i}}{1-\lambda^S_{n_i}}\Big)\sum_{i=1}^{k+1}\frac{\lambda^S_{n_i}|\phi^S_{n_i}(u)|^2}{K(u,u)}\frac{|K_{n_1,\dots,n_{k+1}}(u,v)|^2}{K_{n_1,\dots,n_{k+1}}(u,u)}.
\end{aligned}
$$

Thus, the density of η^u_S is given by

$$
\begin{aligned}
f_{\eta^u_S}(v) &= f_\eta(v, X(S) = X^u(S)) + f_\eta(v, X(S) \neq X^u(S))\\
&= \Big(\prod_{\ell\ge1}(1-\lambda^S_\ell)\Big)\sum_{k\ge0}\sum_{(n_i)_1^{k+1}}\Big(\prod_{i=1}^{k+1} \frac{\lambda^S_{n_i}}{1-\lambda^S_{n_i}}\Big)\sum_{i=1}^{k+1}\frac{|\phi^S_{n_i}(u)|^2}{K(u,u)}\frac{|K_{n_1,\dots,n_{k+1}}(u,v)|^2}{K_{n_1,\dots,n_{k+1}}(u,u)}\\
&= \Big(\prod_{\ell\ge1}(1-\lambda^S_\ell)\Big)\sum_{k\ge0}\sum_{(n_i)_1^{k+1}}\Big(\prod_{i=1}^{k+1} \frac{\lambda^S_{n_i}}{1-\lambda^S_{n_i}}\Big)\frac{|K_{n_1,\dots,n_{k+1}}(u,v)|^2}{K(u,u)},
\end{aligned}
$$

where the last equality follows from the fact that $K_{n_1,\dots,n_{k+1}}(u,u) = \sum_{i=1}^{k+1}|\phi^S_{n_i}(u)|^2$.

Further,

$$
\begin{aligned}
&\sum_{(n_i)_1^{k+1}} \Big(\prod_{i=1}^{k+1} \frac{\lambda_{n_i}^S}{1-\lambda_{n_i}^S}\Big) \frac{|K_{n_1,\dots,n_{k+1}}(u,v)|^2}{K(u,u)} \\
&\quad = \sum_{(n_i)_1^{k+1}} \Big(\prod_{i=1}^{k+1} \frac{\lambda_{n_i}^S}{1-\lambda_{n_i}^S}\Big) \frac{|\sum_{i=1}^{k+1} \phi_{n_i}^S(u)\phi_{n_i}^S(v)|^2}{K(u,u)} \\
&\quad = \sum_{(n_i)_1^{k+1}} \Big(\prod_{i=1}^{k+1} \frac{\lambda_{n_i}^S}{1-\lambda_{n_i}^S}\Big) \sum_{j=1}^{k+1} \phi_{n_j}^S(u)\phi_{n_j}^S(v) \frac{\sum_{i=1}^{k+1} \overline{\phi_{n_i}^S(u)\phi_{n_i}^S(v)}}{K(u,u)} \\
&\quad = \sum_{(n_i)_1^k} \sum_{\substack{n\geq 1: \\ n\notin\{n_1,\dots,n_k\}}} \Big(\prod_{i=1}^{k} \frac{\lambda_{n_i}^S}{1-\lambda_{n_i}^S}\Big) \frac{\lambda_n^S}{1-\lambda_n^S} \big(\phi_n^S(u)\phi_n^S(v)\big) \\
&\qquad\qquad \cdot \frac{\Big(\overline{\phi_n^S(u)\phi_n^S(v)} + \sum_{i=1}^k \overline{\phi_{n_i}^S(u)\phi_{n_i}^S(v)}\Big)}{K(u,u)} \\
&\quad = \frac{1}{K(u,u)} \sum_{n\geq 1} \frac{\lambda_n^S}{1-\lambda_n^S} |\phi_n^S(u)|^2 |\phi_n^S(v)|^2 \sum_{\substack{n_1<\dots<n_k: \\ n_i\neq n, i=1,\dots,k}} \Big(\prod_{i=1}^{k} \frac{\lambda_{n_i}^S}{1-\lambda_{n_i}^S}\Big) \\
&\qquad + \frac{1}{K(u,u)} \sum_{n\geq 1} \frac{\lambda_n^S}{1-\lambda_n^S} \phi_n^S(u)\phi_n^S(v) \\
&\qquad\qquad \cdot \sum_{\substack{n_1<\dots<n_k: \\ n_i\neq n, i=1,\dots,k}} \Big(\prod_{i=1}^{k} \frac{\lambda_{n_i}^S}{1-\lambda_{n_i}^S}\Big) \sum_{i=1}^k \overline{\phi_{n_i}^S(u)\phi_{n_i}^S(v)}.
\end{aligned}
$$

Hence, using the fact that for any multiple distinct $m_1,\dots,m_j \in \mathbb{N}$,

$$
\begin{aligned}
&\sum_{k\geq 0} \sum_{\substack{n_1<\dots<n_k: \\ n_i\notin\{m_1,\dots,m_j\}}} \Big(\prod_{\substack{\ell\geq 1: \\ \ell\neq m_1,\dots,m_j,n_1,\dots,n_k}} (1-\lambda_\ell^S)\Big) \prod_{i=1}^k \lambda_{n_i}^S \\
&\quad = \sum_{k\geq 0} \sum_{\substack{n_1<\dots<n_k: \\ n_i\notin\{m_1,\dots,m_j\}}} \mathrm{P}\Big(\bigcap_{i=1}^k \{B_{n_i}=1\} \cap \bigcap_{\ell\neq m_1,\dots,m_j,n_1,\dots,n_k} \{B_\ell = 0\}\Big) = 1,
\end{aligned}
$$

we obtain

$$
\begin{aligned}
&f_{\eta_S^u}(v) \\
&= \frac{1}{K(u,u)} \sum_{n\geq 1} \lambda_n^S |\phi_n^S(u)\phi_n^S(v)|^2 \sum_{k\geq 0} \sum_{\substack{n_1<\cdots<n_k:\\ n_i\neq n}} \Big(\prod_{\substack{\ell\geq 1:\\ \ell\neq n,n_1,\ldots,n_k}} (1-\lambda_\ell^S) \Big) \prod_{i=1}^k \lambda_{n_i}^S \\
&\quad + \sum_{n\geq 1} \lambda_n^S \phi_n^S(u)\phi_n^S(v) \Big(\prod_{\ell\geq 1,\ell\neq n} (1-\lambda_\ell^S) \Big) \sum_{k\geq 0} \sum_{\substack{n_1<\cdots<n_k:\\ n_i\neq n}} \Big(\prod_{i=1}^k \frac{\lambda_{n_i}^S}{1-\lambda_{n_i}^S} \Big) \\
&\qquad\qquad \cdot \frac{\sum_{i=1}^k \overline{\phi_{n_i}^S(u)\phi_{n_i}^S(v)}}{K(u,u)} \\
&= \frac{\sum_{n\geq 1} \lambda_n^S |\phi_n^S(u)\phi_n^S(v)|^2}{K(u,u)} + \frac{\sum_{n\geq 1} \lambda_n^S \phi_n^S(u)\phi_n^S(v) \sum_{\substack{m\geq 1:\\ m\neq n}} \lambda_m^S \overline{\phi_m^S(u)\phi_m^S(v)}}{K(u,u)} \\
&\qquad\qquad \cdot \sum_{k\geq 0} \sum_{\substack{n_1<\cdots<n_{k-1}:\\ n_i\neq n,m}} \Big(\prod_{\substack{\ell\geq 1:\\ \ell\neq n,m,n_1,\ldots,n_{k-1}}} (1-\lambda_\ell^S) \Big) \prod_{i=1}^{k-1} \lambda_{n_i}^S \\
&= \frac{\sum_{n\geq 1} \lambda_n^S |\phi_n^S(u)\phi_n^S(v)|^2}{K(u,u)} + \frac{\big|\sum_{n\geq 1} \lambda_n^S \phi_n^S(u)\phi_n^S(v)\big|^2}{K(u,u)} - \frac{\sum_{n\geq 1} (\lambda_n^S)^2 |\phi_n^S(u)\phi_n^S(v)|^2}{K(u,u)} \\
&= \frac{\sum_{n\geq 1} \lambda_n^S (1-\lambda_n^S) |\phi_n^S(u)\phi_n^S(v)|^2}{K(u,u)} + \frac{|K(u,v)|^2}{K(u,u)}.
\end{aligned}
$$

Thus (3.22) is verified.

If $X(S) \neq X^u(S)$, then $\xi_S^u = \emptyset$. When $X(S) = X^u(S)$, ξ_S^u consists of a single point with density

$$
\begin{aligned}
&f_{\xi_S^u}(v, X(S) = X^u(S)) \\
&\quad = \sum_{k\geq 0} \sum_{(n_i)_1^k} \sum_{n\notin\{n_i\}_{i=1}^k} \mathrm{P}(B_{n_1,\ldots n_k}^S = 1) \mathrm{P}(X^u(S) = X(S) = k \,|\, B_{n_1,\ldots n_k}^S = 1) \\
&\qquad \cdot \mathrm{P}(N = n \,|\, B_{n_1,\ldots n_k}^S = 1, X^u(S) = X(S) = k) \\
&\qquad \cdot f_{\xi_S^u}(v \,|\, N = n, X^u(S) = X(S) = k, B_{n_1,\ldots,n_k}^S = 1).
\end{aligned}
$$

Then, by (3.15)–(3.16) and step (v), we obtain

$$
\begin{aligned}
&f_{\xi^u_S}(v, X(S) = X^u(S)) \\
&= \sum_{k\geq 0} \sum_{(n_i)_1^k} \Big(\prod_{i=1}^{k} \lambda^S_{n_i} \Big) \prod_{\substack{\ell \geq 1: \\ \ell \notin \{n_i\}_{i=1}^k}} (1 - \lambda^S_\ell) \sum_{\substack{n \geq 1: \\ n \notin \{n_i\}_{i=1}^k}} \frac{\lambda^S_n |\phi^S_n(u)|^2}{K(u,u)} |\phi^S_n(v)|^2 \\
&= \sum_{n\geq 1} \frac{\lambda^S_n |\phi^S_n(u)|^2}{K(u,u)} |\phi^S_n(v)|^2 \sum_{k \geq 0} \sum_{\substack{n_1 < \dots < n_k: \\ n_i \neq n}} \Big(\prod_{i=1}^{k} \lambda^S_{n_i} \Big) \prod_{\ell \geq 1, \ell \notin \{n_1, \dots, n_k\}} (1 - \lambda^S_\ell) \\
&= \sum_{n\geq 1} \frac{\lambda^S_n (1 - \lambda^S_n) |\phi^S_n(u)|^2}{K(u,u)} |\phi^S_n(v)|^2 \sum_{k \geq 0} \sum_{\substack{n_1 < \dots < n_k: \\ n_i \neq n}} \Big(\prod_{i=1}^{k} \lambda^S_{n_i} \Big) \\
&\qquad \cdot \prod_{\ell \geq 1, \ell \notin \{n_1, \dots, n_k, n\}} (1 - \lambda^S_\ell) \\
&= \sum_{n\geq 1} \frac{\lambda^S_n (1 - \lambda^S_n) |\phi^S_n(u)|^2}{K(u,u)} |\phi^S_n(v)|^2.
\end{aligned}
$$

Thereby, (3.21) follows.

References

NIST Digital Library of Mathematical Functions. http://dlmf.nist.gov/, Release 1.0.14 of 2017-09-18. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds.

F. Baccelli and E. O'Reilly. Reach of repulsion for determinantal point processes in high dimensions. Technical report, available on arXiv:1705.03515, 2017.

C. A. N. Biscio and F. Lavancier. Quantifying repulsiveness of determinantal point processes. *Bernoulli*, 22:2001–2028, 2016.

A. Borodin and G. Olshanski. Distributions on partitions, point processes and the hypergeometric kernel. *Communications in Mathematical Physics*, 211:335–358, 2000.

N. Deng, W. Zhou, and M. Haenggi. The Ginibre point process as a model for wireless networks with repulsion. *IEEE Transactions on Wireless Communications*, 14:107–121, 2015.

G. Döge, K. Mecke, J. Møller, D. Stoyan, and R. Waagepetersen. Grand canonical simulations of hard-disk systems by simulated tempering. *International Journal of Modern Physics C*, 15:129–147, 2004.

J. Ginibre. Statistical ensembles of complex, quaternion, and real matrices. *Journal of Mathematical Physics*, 6:440–449, 1965.

A. Goldman. The Palm measure and the Voronoi tessellation for the Ginibre process. *Annals of Applied Probability*, 20:90–128, 2000.

J. B. Hough, M. Krishnapur, Y. Peres, and B. Viràg. Determinantal processes and independence. *Probability Surveys*, 3:206–229, 2006.

J. B. Hough, M. Krishnapur, Y. Peres, and B. Viràg. *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*. American Mathematical Society, Providence, 2009.

A. Kulesza and B. Taskar. *Determinantal Point Processes for Machine Learning*. Now Publishers Inc., Hanover, MA, USA, 2012.

F. Lavancier, J. Møller, and E. Rubak. Determinantal point process models and statistical inference (extended version). Technical report, available on arXiv:1205.4818, 2014.

F. Lavancier, J. Møller, and E. Rubak. Determinantal point process models and statistical inference. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77:853–877, 2015.

T. Lindvall. On strassen's theorem on stochastic domination. *Electronic Communications in Probability*, 4:51–59, 1999.

R. Lyons. Determinantal probability measures. *Publications Mathématiques de l'Institut des Hautes Études Scientifiques*, 98:167–212, 2003.

O. Macchi. The coincidence approach to stochastic point processes. *Advances in Applied Probability*, 7:83–122, 1975.

S. Mase, J. Møller, D. Stoyan, R. Waagepetersen, and G. Döge. Packing densities and simulated tempering for hard core Gibbs point processes. *Annals of the Institute of Statistical Mathematics*, 53:661–680, 2001.

J. Møller and E. Rubak. Functional summary statistics for point processes on the sphere with an application to determinantal point processes. *Spatial Statistics*, 18:4–23, 2016.

J. Møller, M. Nielsen, E. Porcu, and E. Rubak. Determinantal point process models on the sphere. *Bernoulli*, 24:1171–1201, 2018.

I. J. Schoenberg. Positive definite functions on spheres. *Duke Mathematical Journal*, 9: 96–108, 1942.

T. Shirai and Y. Takahashi. Random point fields associated with certain Fredholm determinants I: fermion, Poisson and boson point processes. *Journal of Functional Analysis*, 205:414–463, 2003.

J. Snoek, R. S. Zemel, and R. P. Adams. A determinantal point process latent variable model for inhibition in neural spiking data. In I. S. Francis, B. J. F. Manly, and F. C. Lam, editors, *Advances in Neural Information Processing Systems 26 (NIPS 2013)*, pages 1932–1940. Electronic proceedings from the conference, "Neural Information Processing Systems 2013", 2013.

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