Quantifying repulsiveness of determinantal point processes

Christophe Ange Napoléon Biscio, Frédéric Lavancier

Quantifying repulsiveness of determinantal point processes

Christophe Ange Napoléon Biscio[1] and Frédéric Lavancier[1]

[1] Laboratoire de Mathématiques Jean Leray, University of Nantes, France

Abstract

Determinantal point processes (DPPs) have recently proved to be a useful class of models in several areas of statistics, including spatial statistics, statistical learning and telecommunications networks. They are models for repulsive (or regular, or inhibitive) point processes, in the sense that nearby points of the process tend to repel each other. We consider two ways to quantify the repulsiveness of a point process, both based on its second order properties, and we address the question of how repulsive a stationary DPP can be. We determine the most repulsive stationary DPP, when the intensity is fixed, and we investigate repulsiveness in the subclass of R-dependent stationary DPPs (for a given $R > 0$), i.e. stationary DPPs with R-compactly supported kernels. Finally, in both the general case and the R-dependent case, we present some new parametric families of stationary DPPs that can cover a large range of DPPs, from the homogeneous Poisson process (which induces no interaction) to the most repulsive DPP.

Keywords: pair correlation function, R-dependent point process, covariance function, compactly supported covariance function.

1 Introduction

Determinantal point processes (DPPs) were introduced in their general form by O. Macchi in 1975 [26] to model fermions in quantum mechanics, though some specific DPPs appeared much earlier in random matrix theory. DPPs actually arise in many fields of probability and have deserved a lot of attention from a theoretical point of view, see for instance [18] and [32].

DPPs are repulsive (or regular, or inhibitive) point processes, meaning that nearby points of the process tend to repeal each other (this concept will be clearly described in the following). This property is adapted to many statistical problems where DPPs have been recently used, for instance in telecommunication to model the locations of network nodes [6, 27] and in statistical learning to construct a dictionary of diverse sets [22]. Other examples arising from biology, ecology, forestry are studied in [24] and its associated on-line supplementary file.

The growing interest for DPPs in the statistical community is due to that their moments are explicitly known, parametric families can easily been considered, their density on any compact set admits a closed form expression making likelihood inference feasible and they can be simulated easily and quickly. Section 2 summarizes some of these properties and we refer to [24] for a detailed presentation. These features make the class of DPPs a competitive alternative to the usual class of models for repulsiveness, namely the Gibbs point processes. In contrast, for Gibbs point processes, no closed form expression is available for the moments, the likelihood involves an intractable normalizing constant and their simulation requires Markov Chain Monte Carlo methods.

However, DPPs can not model all kinds of repulsive point patterns. For instance, as deduced from Section 3, stationary DPPs can not involve a hardcore distance between points, contrary to the Matérn's hardcore point processes, the RSA (random sequential absorption) model and hardcore Gibbs models, see [19, Section 6.5]. In this paper, we address the question of how repulsive a stationary DPP can be. We also investigate the repulsiveness in the subclass of R-dependent stationary DPPs, i.e. stationary DPPs with R-compactly supported kernels, that are of special interest for statistical inference in high dimension, see Section 4. In both cases, we present in Section 5 some parametric families of stationary DPPs that cover a large range of DPPs, from the homogeneous Poisson process to the most repulsive DPP.

To quantify the repulsiveness of a stationary point process, we consider its second-order properties. Let X be a stationary point process in $\mathbb{R}^d$ with intensity (i.e. expected number of points per unit volume) $\rho > 0$ and second order intensity function $\rho^{(2)}(x, y)$. Denoting dx an infinitesimal region around x and $|dx|$ its Lebesgue measure, $\rho|dx|$ may be interpreted as the probability that X has a point in dx. For $x \neq y$, $\rho^{(2)}(x, y)|dx||dy|$ may be viewed as the probability that X has a point in dx and another point in dy. A formal definition is given in Section 2. Note that $\rho^{(2)}(x, y)$ only depends on $y - x$ because of our stationarity assumption.

In spatial statistics, the second order properties of X are generally studied through the pair correlation function (in short pcf), defined for any $x \in \mathbb{R}^d$ by

$$g(x) = \frac{\rho^{(2)}(0, x)}{\rho^2}.$$

Note that x in $g(x)$ is to be interpreted as the difference between two points of X. Since $\rho^{(2)}$ is unique up to a set of Lebesgue measure zero (see [4]), so is g. As it is implicitly done in the literature, see [19, 34], we choose the version of g with as few discontinuity points as possible. It is commonly accepted, see for example [34], that if $g(x) = 1$ then there is no interaction between two points separated by x, whereas there is attraction if $g(x) > 1$ and repulsiveness if $g(x) < 1$.

Following this remark, we introduce below a way to compare the *global repulsiveness* of two stationary point processes with the same intensity.

Definition 1.1. *Let X and Y be two stationary point processes with the same intensity ρ and respective pair correlation function g_X and g_Y. Assuming that both $(1 - g_X)$ and $(1 - g_Y)$ are integrable, we say that X is globally more repulsive than Y if $\int(1 - g_X) \geq \int(1 - g_Y)$.*

The quantity $\int(1-g)$ is already considered in the on-line supplementary material of [24] as a measure for repulsiveness. It can be justified in several ways. First, it is a natural geometrical method to quantify the distance from g to 1 (corresponding to no interaction), where the area between g and 1 contributes positively to the measure of repulsiveness when $g < 1$ and negatively if $g > 1$. Second, as explained in the on-line supplementary file of [24], denoting P the law of X and $P_o^!$ its reduced Palm distribution, $\rho\int(1-g)$ corresponds to the limit, when $r \to \infty$, of the difference between the expected number of points within distance r from the origin under P and under $P_o^!$. Recall that $P_o^!$ can be interpreted as the distribution of X conditioned to have a point at the origin. In close relation, denoting K and K_0 the Ripley's K-functions of X and of the homogeneous Poisson process with intensity ρ, respectively, $\int(1-g) = \lim_{r\to\infty}(K_0(r) - K(r))$, see [28, Definition 4.6]. Third, the variance of the number of points of X in a compact set D is $Var(X(D)) = \rho|D| - \rho^2\int_{D^2}(1 - g(y-x))dxdy$, see [19]. Thus, the intensity ρ being fixed, maximizing $\int(1-g)$ is equivalent to minimize $Var(X(D))/|D|$ when $D \to \mathbb{R}^d$, provided D and g are sufficiently regular to apply the mean value theorem. Finally, it is worth mentioning that for any stationary point processes, we have $\int(1-g) \leq 1/\rho$, see [23, Equation (2.5)].

Additional criteria could be introduced to quantify the global repulsiveness of a point process, relying for instance on $\int(1-g)^p$ for $p > 0$, or involving higher moments of the point process through the joint intensities of order $k > 2$ (see Definition 2.1). However the theoretical study becomes more challenging in these cases and we do not consider these extensions.

In practice, repulsiveness is often interpreted in a local sense. This is the case for hardcore point processes, where a minimal distance δ is imposed between points and so $g(x) = 0$ whenever $|x| < \delta$ where for a vector x, $|x|$ denotes its euclidean norm. As already mentioned, a DPP can not involve any hardcore distance, but we may want its pcf to satisfy $g(0) = 0$ and stay as close as possible to 0 near the origin. This leads to the following criteria to compare the *local repulsiveness* of two point processes. We denote by ∇g and Δg the gradient and the Laplacian of g, respectively.

Definition 1.2. *Let X and Y be two stationary point processes with the same intensity ρ and respective pair correlation function g_X and g_Y. Assuming that g_X is twice differentiable at 0 with $g_X(0) = 0$, we say that X is more locally repulsive than Y if either $g_Y(0) > 0$, or g_Y is not twice differentiable at 0, or g_Y is twice differentiable at 0 with $g_Y(0) = 0$ and $\Delta g_Y(0) \geq \Delta g_X(0)$.*

As suggested by this definition, a stationary point process is said to be locally repulsive if its pcf is twice differentiable at 0 with $g(0) = 0$. In this case $\nabla g(0) = 0$ because $g(x) = g(-x)$. Therefore to compare the behavior of two such pcfs near the origin, specifically the curvatures of their graphs near the origin, the Laplacian operator is involved in Definition 1.2. As an example, a stationary hardcore process is locally more repulsive than any other stationary point process because $g(0) = 0$ and $\Delta g(0) = 0$ in this case. On the other hand, a concave pcf is not differentiable at the origin and for this reason the associated point process is less locally repulsive than any stationary process with a twice differentiable pcf that vanishes at the origin.

We show in Section 3 that Definitions 1.1 and 1.2 agree in the natural choice of what can be considered as the most repulsive DPP. As a result, a realization of the latter on $[-5,5]^2$ is represented in Figure 1 (d) when $\rho = 1$. For comparison, letting $\rho = 1$ for all plots, Figure 1 shows: (a) the homogeneous Poisson process, which is a situation without any interaction; (b)-(c) two DPPs with intermediate repulsiveness, namely DPPs with kernels (5.1) where $\sigma = 0$ and $\alpha = 0.2, 0.4$ respectively, as presented in Section 5.1; (e) the type II Matérn's hardcore process with hardcore radius $\frac{1}{\sqrt{\pi}}$. Notice that $\frac{1}{\sqrt{\pi}}$ is the maximal hardcore radius that a type II Matérn hardcore process with unit intensity can reach, see [19, Section 6.5]. These models are sorted from (a) to (e) by their ascending repulsiveness in the sense of Definition 1.2. This is clearly apparent in Figure 1 (f), where their theoretical pcfs are represented as radial functions (all aforementioned models being isotropic). Figure 1 illustrates that even if stationary DPPs cannot be as repulsive as hardcore point processes, which may be an important limitation in practice, they nonetheless cover a rather large variety of repulsiveness from (a) to (d) in Figure 1.

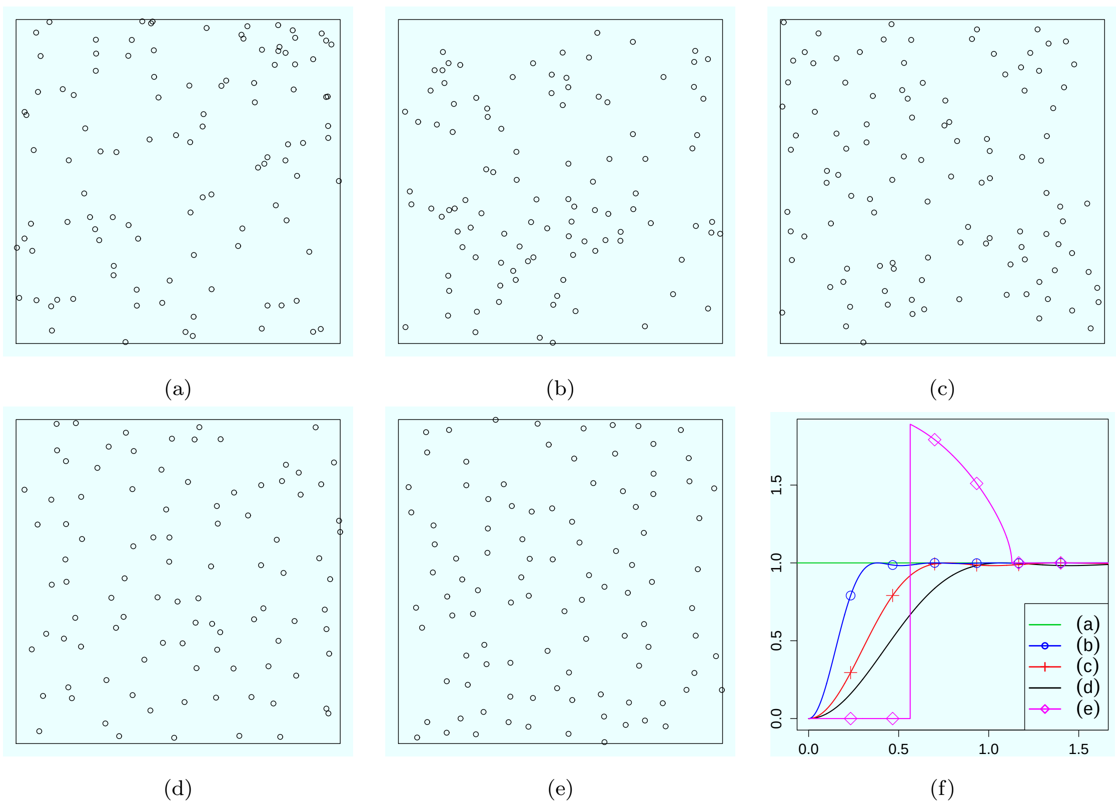


Figure 1: Realizations on $[-5,5]^2$ of (a) the homogeneous Poisson process, (b)-(d) DPPs with kernels (5.1) where $\sigma = 0$ and $\alpha = 0.2, 0.4, \frac{1}{\sqrt{\pi}}$, (e) the type II Matérn's hardcore process with hardcore radius $\frac{1}{\sqrt{\pi}}$. (f) Their associated theoretical pcfs. The intensity is $\rho = 1$ for all models and (d) represents the most repulsive stationary DPP in this case.

We recall the definition of a stationary DPP and some related basic results in Section 2. Section 3 is devoted to the study of repulsiveness in stationary DPPs, both in the sense of Definition 1.1 and Definition 1.2. In Section 4, we focus on repulsiveness for the subclass of stationary DPPs with compactly supported kernels. Then, in Section 5, we present three parametric families of DPPs which cover a large range of repulsiveness and have further interesting properties. Section 6 gathers the proofs of our theoretical results.

2 Stationary DPPs

In this section, we review the basic definition and some properties of stationary DPPs. For a detailed presentation, including the non stationary case, we refer to the survey by Hough et al. [18].

Basics of point processes may be found in [4, 5]. Let us recall that a point process X is simple if two points of X never coincide, almost surely. The joint intensities of X are defined as follows.

Definition 2.1. *If it exists, the joint intensity of order k ($k \geq 1$) of a simple point process X is the function $\rho^{(k)} : (\mathbb{R}^d)^k \to \mathbb{R}^+$ such that for any family of mutually disjoint subsets $D_1, \dots, D_k$ in $\mathbb{R}^d$,*

$$E \prod_{i=1}^{k} X(D_i) = \int_{D_1} \dots \int_{D_k} \rho^{(k)}(x_1, \dots, x_k) dx_1 \dots dx_k,$$

where $X(D)$ denotes the number of points of X in D and E is the expectation over the distribution of X.

In the stationary case, $\rho^{(k)}(x_1, \dots, x_k) = \rho^{(k)}(0, x_2 - x_1, \dots, x_k - x_1)$, so that the intensity ρ and the second order intensity function $\rho^{(2)}$ introduced previously become the particular cases associated to $k = 1$ and $k = 2$ respectively.

Definition 2.2. *Let $C : \mathbb{R}^d \to \mathbb{R}$ be a function. A point process X on $\mathbb{R}^d$ is a stationary DPP with kernel C, in short $X \sim DPP(C)$, if for all $k \geq 1$, its joint intensity of order k satisfies the relation*

$$\rho^{(k)}(x_1, \dots x_k) = \det[C](x_1, \dots, x_k)$$

for almost every $(x_1, \dots, x_k) \in (\mathbb{R}^d)^k$, where $[C](x_1, \dots, x_k)$ denotes the matrix with entries $C(x_i - x_j)$, $1 \leq i, j \leq k$.

It is actually possible to consider a complex-valued kernel C, but for simplicity we restrict ourselves to the real case. A first example of stationary DPP is the homogeneous Poisson process with intensity ρ. It corresponds to the kernel

$$C(x) = \rho \mathbf{1}_{\{x=0\}}, \; \forall x \in \mathbb{R}^d. \tag{2.1}$$

However, this example is very particular and represents in some sense the extreme case of a DPP without any interaction, while DPPs are in general repulsive as discussed at the end of this section.

Definition 2.2 does not ensure existence or unicity of $DPP(C)$, but if it exists, then it is unique, see [18]. Concerning existence, a general result, including the non stationary case, was proved by O. Macchi in [26]. It relies on the Mercer representation of C on any compact set. Unfortunately this representation is known only in a few cases, making the conditions impossible to verify in practice for most functions C. Nevertheless, the situation becomes simpler in our stationary framework, where the conditions only involve the Fourier transform of C. We define the Fourier transform of a function $h \in L^1(\mathbb{R}^d)$ as

$$\mathcal{F}(h)(t) = \int_{\mathbb{R}^d} h(x) e^{2i\pi x \cdot t} dx, \quad \forall t \in \mathbb{R}^d. \tag{2.2}$$

By Plancherel's theorem, this definition is extended to $L^2(\mathbb{R}^d)$, see [33]. If C is a covariance function, as assumed in the following, we have $\mathcal{F}\mathcal{F}(C) = C$ so $\mathcal{F}^{-1} = \mathcal{F}$ and from [29, Theorem 1.8.13], $\mathcal{F}(C)$ belongs to $L^1(\mathbb{R}^d)$.

Proposition 2.3 ([24])**.** *Assume C is a symmetric continuous real-valued function in $L^2(\mathbb{R}^d)$. Then $DPP(C)$ exists if and only if $0 \le \mathcal{F}(C) \le 1$.*

In other words, Proposition 2.3 ensures existence of $DPP(C)$ if C is a continuous real-valued covariance function in $L^2(\mathbb{R}^d)$ with $\mathcal{F}(C) \le 1$. Henceforth, we assume the following condition.

Condition $\mathcal{K}(\rho)$. A kernel C is said to verify condition $\mathcal{K}(\rho)$ if C is a symmetric continuous real-valued function in $L^2(\mathbb{R}^d)$ with $C(0) = \rho$ and $0 \le \mathcal{F}(C) \le 1$.

The assumption $0 \le \mathcal{F}(C) \le 1$ is in accordance with Proposition 2.3, while the others assumptions in condition $\mathcal{K}(\rho)$ are satisfied by most statistical models of covariance functions, the main counterexample being (2.1). Standard parametric families of kernels include the Gaussian, the Whittle-Matérn and the generalized Cauchy covariance functions, where the condition $\mathcal{F}(C) \le 1$ implies some restriction on the parameter space, see [24].

From Definition 2.2, all moments of a DPP are explicitly known. In particular, assuming condition $\mathcal{K}(\rho)$, then the intensity of $DPP(C)$ is ρ and denoting g its pcf we have

$$1 - g(x) = \frac{C(x)^2}{\rho^2} \tag{2.3}$$

for almost every $x \in \mathbb{R}^d$. Consequently $g \le 1$, and so we have repulsiveness. Moreover, the study of repulsiveness of stationary DPPs, as defined in Definitions 1.1 and 1.2, reduces to considerations on the kernel C when condition $\mathcal{K}(\rho)$ is assumed.

3 Most repulsive DPPs

We first present the most globally repulsive DPPs, in the sense of Definition 1.1. They are introduced in the on-line supplementary file associated to [24], from which the following proposition is easily deduced.

Proposition 3.1 ([24])**.** *In the sense of Definition 1.1, $DPP(C)$ is the most globally repulsive DPP among all DPPs with kernel satisfying condition $\mathcal{K}(\rho)$ if and only if $\mathcal{F}(C)$ is even and equals almost everywhere an indicator function of a Borel set with volume ρ.*

According to Proposition 3.1, there exists an infinity of choices to the most globally repulsive DPP in the sense of Definition 1.1. This is illustrated in the on-line supplementary material. A natural choice is $DPP(C_B)$ where $\mathcal{F}(C_B)$ is the indicator function of the euclidean ball centered at 0 with volume ρ. In dimension d, this gives $C_B = \mathcal{F}\left(\mathbf{1}_{\{|\cdot|^d \leq \rho\tau^d\}}\right)$ with $\tau = \left\{\Gamma(d/2+1)/\pi^{\frac{d}{2}}\right\}^{\frac{1}{d}}$ and by [16, Appendix B.5],

$$C_B(x) = \frac{\sqrt{\rho\Gamma(\frac{d}{2}+1)}}{\pi^{d/4}} \frac{J_{\frac{d}{2}}\left(2\sqrt{\pi}\Gamma(\frac{d}{2}+1)^{\frac{1}{d}}\rho^{\frac{1}{d}}|x|\right)}{|x|^{\frac{d}{2}}}, \quad \forall x \in \mathbb{R}^d, \tag{3.1}$$

where $J_{\frac{d}{2}}$ is the Bessel function of the first kind. For example, we have

- for $d = 1$, $C_B(x) = \text{sinc}(x) = \frac{\sin(\pi\rho|x|)}{\pi|x|}$,
- for $d = 2$, $C_B(x) = \text{jinc}(x) = \sqrt{\rho}\frac{J_1(2\sqrt{\pi\rho}|x|)}{\sqrt{\pi}|x|}$.

This choice was already favored in [24]. However, there is no indication from Proposition 3.1 to suggest C_B instead of another kernel given by the proposition. This choice becomes clear if we look at the local repulsiveness as defined in Definition 1.2.

Proposition 3.2. *In the sense of Definition 1.2, the most locally repulsive DPP among all DPPs with kernel satisfying condition $\mathcal{K}(\rho)$ is $DPP(C_B)$.*

Thus, from Propositions 3.1 and 3.2, we deduce the following corollary.

Corollary 3.3. *The kernel C_B is the unique kernel C verifying condition $\mathcal{K}(\rho)$ such that $DPP(C)$ is both the most globally and the most locally repulsive DPP among all stationary DPPs with intensity $\rho > 0$.*

Borodin and Serfaty in [3] characterize in dimension $d \leq 2$ the disorder of a point process by its "renormalized energy". In fact, the smaller the renormalized energy, the more repulsive the point process. Theorem 3 in [3] establishes that $DPP(C_B)$ minimizes the renormalized energy among the most globally repulsive stationary DPPs given by Proposition 3.1. This result confirms Corollary 3.3, that the most repulsive stationary DPP, if any has to be chosen, is $DPP(C_B)$. However, except when the DPPs are given by Proposition 3.1, all stationary DPPs have an infinite renormalized energy (see [3, Theorem 1]), which indicates that the renormalized energy is not of practical use to compare the repulsiveness between two arbitrary DPPs.

4 Most repulsive DPPs with compactly supported kernels

In this section, we assume that the kernel C is compactly supported, i.e. there exists $R > 0$ such that $C(x) = 0$ if $|x| > R$. In this case $X \sim DPP(C)$ is an R-dependent point process in the sense that if A and B are two Borel sets in $\mathbb{R}^d$ separated by a distance larger than R, then $X \cap A$ and $X \cap B$ are independent, which is easily verified using Definition 2.2. This situation can be particularly interesting for likelihood inference in presence of a large number of points. Assume we observe $\{x_1, \dots, x_n\}$ on a compact window $W \subset \mathbb{R}^d$, then the likelihood is proportional to $\det[\tilde{C}](x_1, \dots, x_n)$ where $\tilde{C}$ expresses in terms of C and inherits the compactly supported property of C, see [24, 26]. While this determinant is computationally expensive to evaluate if $\tilde{C}$ is not compactly supported and n is large, the situation becomes more convenient in the compactly supported case, as this yields a sparse matrix $[\tilde{C}](x_1, \dots, x_n)$ provided R is small with respect to the size of W. We are thus interested in DPPs with kernels satisfying the following condition.

Condition $\mathcal{K}_c(\rho, R)$. A kernel C or $DPP(C)$ is said to verify condition $\mathcal{K}_c(\rho, R)$ if C verifies condition $\mathcal{K}(\rho)$ and C is compactly supported with range R, i.e. $C(x) = 0$ for $|x| \geq R$.

The following proposition shows that any kernel satisfying condition $\mathcal{K}(\rho)$ can be arbitrarily approximated by kernels verifying $\mathcal{K}_c(\rho, r)$ for r large enough. We define the function h by

$$h(x) = \exp\left(\frac{1}{|x|^2 - 1}\right) \mathbf{1}_{\{|x|<1\}}, \quad \forall x \in \mathbb{R}^d. \tag{4.1}$$

For a function $f \in L^2(\mathbb{R}^d)$, put $\|f\| = \sqrt{\int |f(t)|^2 dt}$ and denote $[f * f]$ the self-convolution product of f.

Proposition 4.1. *Let C be a kernel verifying condition $\mathcal{K}(\rho)$ and h be defined by* (4.1)*. Then, for all $r > 0$, the function C_r defined by*

$$C_r(x) = \frac{1}{\|h\|^2}[h * h]\left(\frac{2x}{r}\right) C(x), \quad \forall x \in \mathbb{R}^d, \tag{4.2}$$

verifies $\mathcal{K}_c(\rho, r)$. Moreover, we have the convergence

$$\lim_{r \to +\infty} C_r = C, \tag{4.3}$$

uniformly on all compact sets.

In particular, by taking $C = C_B$ in Proposition 4.1, it is always possible to find a kernel C_r verifying $\mathcal{K}_c(\rho, r)$ that yields a repulsiveness (local or global) as close as we wish to the repulsiveness of C_B, provided that r is large enough. However, given a maximal range of interaction R, it is clear that the maximal repulsiveness implied by kernels verifying $\mathcal{K}_c(\rho, R)$ can not reach the one of C_B, since the support of C_B is unbounded and $DPP(C_B)$ is the unique most repulsive DPP according to Corollary 3.3. In the following, we study the DPP's repulsiveness when the range R is fixed.

In comparison with condition $\mathcal{K}(\rho)$, the assumption that C is compactly supported in condition $\mathcal{K}_c(\rho, R)$ makes the optimization problems related to Definitions 1.1-1.2 much more difficult to investigate. As a negative result, we know very little about the most globally repulsive DPP, in the sense of Definition 1.1, under condition $\mathcal{K}_c(\rho, R)$. From relation (2.3), this is equivalent to find a kernel C with maximal L^2-norm under the constraint that C verifies $\mathcal{K}_c(\rho, R)$. Without the constraint $\mathcal{F}(C) \leq 1$, this problem is known as the square-integral Turán problem with range R, see for example [21]. For this less constrained problem, it is known that a solution exists, but no explicit formula is available, cf. [7]. For $d = 1$, it has been proved that the solution is unique and there exists an algorithm to approximate it, see [13]. In this case, numerical approximations show that the solution with range R verifies condition $\mathcal{K}_c(\rho, R)$ only if $R \leq 1.02/\rho$. This gives the most globally repulsive DPP verifying $\mathcal{K}_c(\rho, R)$ in dimension $d = 1$, when $R \leq 1.02/\rho$, albeit without explicit formula. Its pcf is represented in Figure 2. For other values of R, or in dimension $d \geq 2$, no results are available, to the best of our knowledge.

Let us now turn to the investigation of the most locally repulsive DPP, in the sense of Definition 1.2, under condition $\mathcal{K}_c(\rho, R)$. Recall that without the compactly supported constraint of the kernel, we showed in Section 3 that the most locally repulsive DPP, namely $DPP(C_B)$, is also (one of) the most globally repulsive DPP.

For $\nu > 0$, we denote by j_ν the first positive zero of the Bessel function J_ν and by J'_ν the derivative of J_ν. We refer to [1] for a survey about Bessel functions and their zeros. Further, define the constant $M > 0$ by

$$M^d = \frac{2^{d-2} j^2_{\frac{d-2}{2}} \Gamma\left(\frac{d}{2}\right)}{\rho \pi^{\frac{d}{2}}}.$$

We have $M\rho = \pi^2/8 \approx 1.234$ when $d = 1$, $M\rho^{1/2} = j_0/\pi^{1/2} \approx 1.357$ when $d = 2$ and $M\rho^{1/3} = \pi^{1/3} \approx 1.465$ when $d = 3$.

Proposition 4.2. *If $R \leq M$, then, in the sense of Definition 1.2, there exists an unique isotropic kernel C_R such that $DPP(C_R)$ is the most locally repulsive DPP among all DPPs with kernel verifying $\mathcal{K}_c(\rho, R)$. It is given by $C_R = u * u$ where*

$$u(x) = \kappa \, \frac{J_{\frac{d-2}{2}}\left(2 j_{\frac{d-2}{2}} \frac{|x|}{R}\right)}{|x|^{\frac{d-2}{2}}} \, \mathbf{1}_{\left\{|x| < \frac{R}{2}\right\}}, \tag{4.4}$$

with $\kappa^2 = \frac{4\Gamma(d/2)}{\rho \pi^{d/2} R^2} \left(J'_{\frac{d-2}{2}}(j_{\frac{d-2}{2}}) \right)^{-2}$.

In this proposition C_R is only given as a convolution product. Nonetheless, an explicit expression is known in dimension $d = 1$ and $d = 3$, see [8]. On the other hand, the Fourier transform is known in any dimension since $\mathcal{F}(C_R) = \mathcal{F}(u)^2$. We get from the proof in Section 6.3, for all $x \in \mathbb{R}^d$,

$$\mathcal{F}(C_R)(x) = \rho \pi^{d/2} R^d j^2_{\frac{d-2}{2}} \Gamma\left(\frac{d}{2}\right) \left(\frac{J_{\frac{d-2}{2}}(\pi R|x|)}{(\pi R|x|)^{\frac{d-2}{2}} \left(j^2_{\frac{d-2}{2}} - (\pi R|x|)^2 \right)} \right)^2 . \tag{4.5}$$

If $R \geq M$, we have not been able to obtain a closed form expression of the most locally repulsive stationary DPP. However, under some extra regularity assumptions, we can state the following general result about its existence and the form of the solution.

Condition $\mathcal{M}(\rho, R)$. A function u is said to verify condition $\mathcal{M}(\rho, R)$ if $u(x) = 0$ for $|x| > \frac{R}{2}$, u is a radial function and $u \in L^2(\mathbb{R}^d)$ with $\|u\|^2 = \rho$.

Proposition 4.3. *For any $R > 0$, there exists an isotropic kernel C_R such that $DPP(C_R)$ is the most locally repulsive DPP among all DPPs with kernel C verifying $\mathcal{K}_c(\rho, R)$. It can be expressed as $C_R = u * u$ where u satisfies $\mathcal{M}(\rho, R)$. Furthermore, if we assume that $\sup_{x\in\mathbb{R}^d} \mathcal{F}(C)(x) = \mathcal{F}(C)(0)$ and u is twice differentiable on its support, then u is of the form*

$$u(x) = \left(\beta + \gamma \frac{J_{\frac{d-2}{2}}\left(|x|/\alpha\right)}{|x|^{\frac{d-2}{2}}} \right) \mathbf{1}_{\{|x|<\frac{R}{2}\}}, \tag{4.6}$$

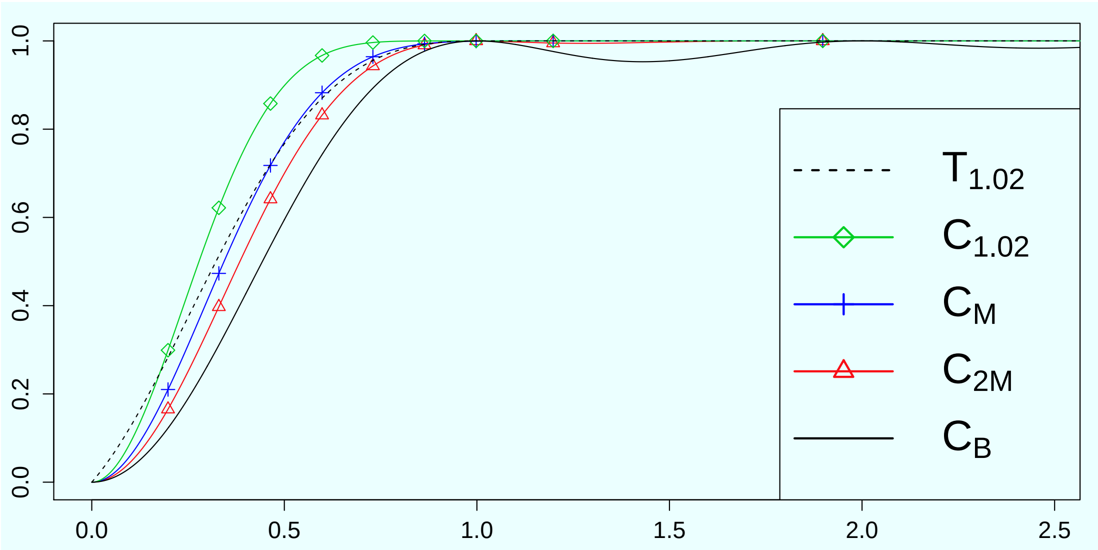
where $\alpha > 0, \beta \geq 0$ and γ are three constants linked by the conditions $\mathcal{M}(\rho, R)$ and $\int_{\mathbb{R}^d} u(x)dx \leq 1$.

In the case $R \leq M$, this proposition is a consequence of Proposition 4.2 where $\beta = 0$, $\alpha = R/(2j_{\frac{d-2}{2}})$ and $\gamma = \kappa$. When $R > M$, it is an open problem to find an explicit expression of the kernel C_R without any extra regularity assumptions. Even in this case, (4.6) only gives the form of the solution and the constants α, β and γ are not explicitly known. In particular the choice $\beta = 0$ does not lead to the most locally repulsive DPP when $R > M$, contrary to the case $R \leq M$. In fact, the condition $\mathcal{M}(\rho, R)$ allows us to express β and γ as functions of α, R and ρ, but then some numerical approximation are needed to find the value of α in (4.6), given R and ρ, such that $DPP(C_R)$ is the most locally repulsive DPP. We detail these relations in Section 5.3, where we start from (4.6) to suggest a new parametric family of compactly supported kernels.

Contrary to what happens in the non compactly supported case of Section 3, the most locally repulsive DPP is not the most globally repulsive DPP under $\mathcal{K}_c(\rho, R)$. This is easily checked in dimension $d = 1$ when $R \leq 1.02/\rho$ implying $R \leq M$: In this case the most globally repulsive DPP under $\mathcal{K}_c(\rho, R)$ is $DPP(T_R)$, where T_R is the solution of the square-integral Turán problem with range R and the most locally repulsive DPP is $DPP(C_R)$ where C_R is given by (4.4). However, according to the results of Section 3 corresponding to $R = \infty$, we expect that $DPP(C_R)$ has a strong global repulsiveness even for moderate values of R. This is confirmed in Figure 2, that shows the pcf of $DPP(C_R)$ when $d = 1$, $\rho = 1$ and $R = 1.02$, $R = M \approx 1.234$ and $R = 2M$, where in this case we take $C_R = u * u$ with u given by (4.6) and the constants are obtained by numerical approximations. The pcfs of $DPP(T_{1.02})$ and

$DPP(C_B)$ are added for sake of comparison. Considering the behavior of the pcf near the origin, note that even if $DPP(T_{1.02})$ is the most globally repulsive DPP under $\mathcal{K}_c(\rho, R)$ when $R \leq 1.02/\rho$, its local repulsiveness is not very strong. On the other hand, $DPP(C_R)$ seems to present strong global repulsiveness for the values of R considered in the figure.

Figure 2: In dimension $d = 1$, comparison between the pcf of $DPP(T_{1.02})$, $DPP(C_B)$ and $DPP(C_R)$ for $R = 1.02, M, 2M$.



5 Parametric families of DPP kernels

A convenient parametric family of kernels $\{C_\theta\}_{\theta \in \Theta}$, where $\Theta \subset \mathbb{R}^q$ for some $q \geq 1$, should ideally:

(a) provide a closed form expression for C_θ, for any θ,

(b) provide a closed form expression for $\mathcal{F}(C_\theta)$, for any θ,

(c) be flexible enough to include a large range of DPPs, going from the Poisson point process to $DPP(C_B)$.

The second property above is needed to check the condition of existence $\mathcal{F}(C_\theta) \leq 1$, but it is also useful for some approximations in practice. Indeed, the algorithm for simulating $DPP(C)$ on a compact set S, as presented in [18], relies on the Mercer representation of C on S, which is rarely known in practice. In [24], this decomposition is simply approximated by the Fourier series of C, where the k-th Fourier coefficients is replaced by $\mathcal{F}(C)(k)$, up to some rescaling. The same approximation is used to compute the likelihood. This Fourier approximation proved to be accurate in most cases, both from a practical and a theoretical point of view, provided ρ is not too small, and to be computationally efficient, see [24].

In addition to (a)-(c), we may also require that C_θ is compactly supported with maximal range R, following the motivation explained in Section 4, in which case the maximal possible repulsiveness is given by $DPP(C_R)$. Or we may require that $\mathcal{F}(C_\theta)$ is compactly supported, in which case the Fourier series mentioned in the previous paragraph becomes a finite sum and no truncation is needed in practice. Note however that C_θ and $\mathcal{F}(C_\theta)$ can not both be compactly supported.

Several standard parametric families of kernels are available, including the well-known Whittle-Matérn and the generalized Cauchy covariance functions, where the condition $\mathcal{F}(C_\theta) \leq 1$ implies some restriction on the parameter space, see [24]. Although they encompass a closed form expression for both C_θ and $\mathcal{F}(C_\theta)$, they are not flexible enough to reach the repulsiveness of $DPP(C_B)$. Another family of parametric kernels is considered in [24], namely the power exponential spectral model, that contains as limiting cases C_B and the Poisson kernel (2.1). For this reason this family is more flexible than the previous ones, but then only $\mathcal{F}(C_\theta)$ is given and no closed expression is available for C_θ. For all these families, none of C_θ and $\mathcal{F}(C_\theta)$ is compactly supported.

Below, we present alternative families of parametric kernels. The first two ones, so-called Bessel-type and Laguerre-Gaussian families, fulfil the three requirements (a)-(c) above and the Bessel-type family has the additional property that the Fourier transform of the kernels is compactly supported. Moreover we introduce new families of compactly supported kernels, inspired by Proposition 4.1 and Proposition 4.3.

5.1 Bessel-type family

For all $\sigma \geq 0,\ \alpha > 0,\ \rho > 0$, we consider the Bessel-type kernel

$$C(x) = \rho\, 2^{\frac{\sigma+d}{2}} \Gamma\left(\frac{\sigma+d+2}{2}\right) \frac{J_{\frac{\sigma+d}{2}}\left(2|\frac{x}{\alpha}|\sqrt{\frac{\sigma+d}{2}}\right)}{\left(2|\frac{x}{\alpha}|\sqrt{\frac{\sigma+d}{2}}\right)^{\frac{\sigma+d}{2}}}, \quad x \in \mathbb{R}^d. \tag{5.1}$$

This positive definite function first appears in [30], where it is called the Poisson function. It has been further studied in [11] and [12], where it is called the Bessel-type function. For obvious reasons, we prefer the second terminology when applied to point processes. For any $x \in \mathbb{R}$, we denote by $x_+ = \max(x, 0)$ its positive part.

Proposition 5.1. *Let C be given by* (5.1)*, then its Fourier transform is, for all $x \in \mathbb{R}^d$,*

$$\mathcal{F}(C)(x) = \rho \frac{(2\pi)^{\frac{d}{2}} \alpha^d \Gamma(\frac{\sigma+d+2}{2})}{(\sigma+d)^{\frac{d}{2}} \Gamma(\frac{\sigma+2}{2})} \left(1 - \frac{2\pi^2\alpha^2|x|^2}{\sigma+d}\right)_+^{\frac{\sigma}{2}} \tag{5.2}$$

and $DPP(C)$ exists if and only if $\alpha \leq \alpha_{\max}$ where

$$\alpha_{\max}^d = \frac{(\sigma+d)^{\frac{d}{2}} \Gamma(\frac{\sigma+2}{2})}{\rho (2\pi)^{\frac{d}{2}} \Gamma(\frac{\sigma+d+2}{2})}.$$

In this case, $DPP(C)$ defines a stationary and isotropic DPP with intensity ρ. Moreover, if $\sigma = 0$ and $\alpha = \alpha_{\max}$, then $C = C_B$ where C_B is defined in (3.1)*. In addition, for any $\rho > 0$ and $\alpha > 0$, we have the convergence*

$$\lim_{\sigma \to +\infty} C(x) = \rho\, e^{-\left(\frac{|x|}{\alpha}\right)^2}, \tag{5.3}$$

uniformly on all compact sets.

The Bessel-type family contains C_B as a particular case and the Poisson kernel as a limiting case, when $\alpha \to 0$. Moreover, $\mathcal{F}(C)$ is compactly supported, see (5.2). Figure 3 shows the behavior of the pcf of $DPP(C)$ with respect to σ, while Figure 4 illustrates the convergence result (5.3). The plots in Figure 1 (b)-(d) show some realizations of this model when $\sigma = 0$ and $\alpha = 0.2, 0.4, \alpha_{\max}$, respectively.

Figure 3: Pcf's of $DPP(C)$ where C is given by (5.1), when $d = 2$, $\rho = 1$, $\sigma = 0$ and different values of α. The case $\alpha = \alpha_{\max} = 1/\sqrt{\pi} \approx 0.56$ corresponds to $C = C_B$.

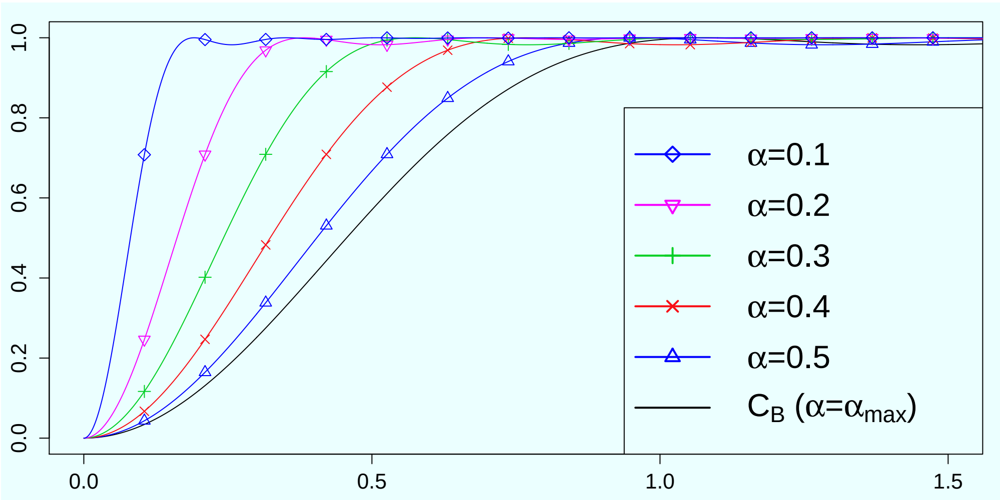
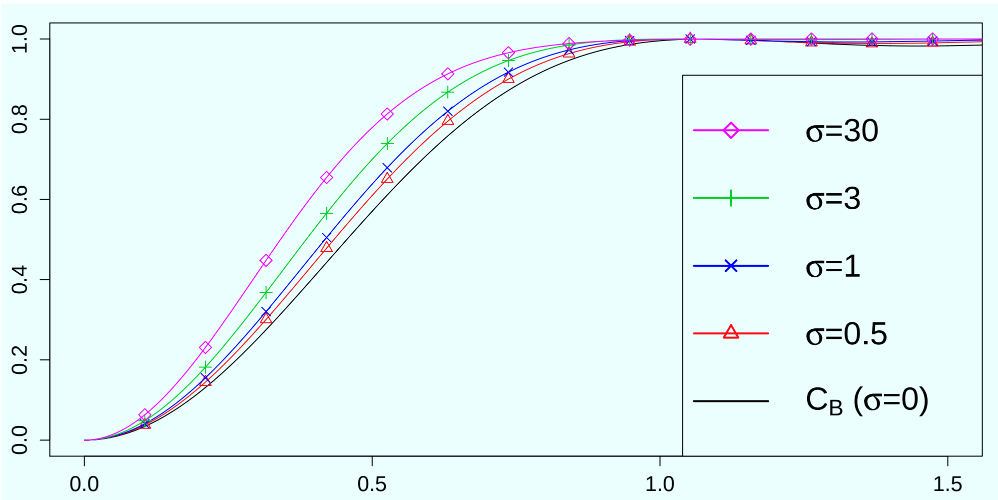


Figure 4: Pcf's of $DPP(C)$ where C is given by (5.1), when $d = 2$, $\rho = 1$, $\alpha = \alpha_{\max}$, and different values of σ. The case $\sigma = 0$ corresponds to $C = C_B$.



5.2 Laguerre-Gaussian family

Let us first recall the definition of the Laguerre polynomials. We denote by $\mathbb{N}$ the set $\{0, 1, 2, \dots\}$ and by $\mathbb{N}^*$ the set $\mathbb{N} \setminus \{0\}$. For integers $0 \le k \le m$ and numbers α, define $\binom{m+\alpha}{k} = \frac{(m+\alpha)\dots(m+\alpha+1-k)}{k}$ if $k > 0$ and $\binom{m+\alpha}{k} = 1$ if $k = 0$.

Definition 5.2. *The Laguerre polynomials are defined for all $m \in \mathbb{N}$ and $\alpha \in \mathbb{R}$ by*

$$L_m^\alpha(x) = \sum_{k=0}^{m} \binom{m+\alpha}{m-k} \frac{(-x)^k}{k!}, \quad \forall x \in \mathbb{R}.$$

For all $m \in \mathbb{N}^*$, $\alpha > 0$, $\rho > 0$ and $x \in \mathbb{R}^d$, we consider the Laguerre-Gaussian function

$$C(x) = \frac{\rho}{\binom{m-1+\frac{d}{2}}{m-1}} L_{m-1}^{\frac{d}{2}} \left(\frac{1}{m} \left| \frac{x}{\alpha} \right|^2 \right) e^{-\frac{1}{m}|\frac{x}{\alpha}|^2}. \tag{5.4}$$

This kernel already appears in the literature, see e.g. [10] for an application in approximation theory. The following proposition summarizes the properties that are relevant for its use as a DPP kernel.

Proposition 5.3. *Let C be given by* (5.4)*, then its Fourier transform is, for all $x \in \mathbb{R}^d$,*

$$\mathcal{F}(C)(x) = \frac{\rho}{\binom{m-1+\frac{d}{2}}{m-1}} \alpha^d (m\pi)^{\frac{d}{2}} e^{-m(\pi\alpha|x|)^2} \sum_{k=0}^{m-1} \frac{(\pi\sqrt{m}|\alpha x|)^{2k}}{k!} \tag{5.5}$$

and $DPP(C)$ exists if and only if $\alpha \leq \alpha_{\max}$ where

$$\alpha_{\max}^d = \frac{\binom{m-1+\frac{d}{2}}{m-1}}{\rho(m\pi)^{\frac{d}{2}}}.$$

In this case, $DPP(C)$ is stationary and isotropic with intensity ρ. Moreover, for any $\rho > 0$ and $\alpha > 0$, we have the convergence

$$\lim_{m \to +\infty} C(x) = \rho \Gamma\left(\frac{d}{2} + 1\right) \frac{J_{\frac{d}{2}}\left(2|\frac{x}{\alpha}|\right)}{|\frac{x}{\alpha}|^{\frac{d}{2}}} \tag{5.6}$$

uniformly on all compact sets. In particular, for $\alpha = \alpha_{\max}$,

$$\lim_{m \to +\infty} C(x) = C_B(x) \tag{5.7}$$

uniformly on all compact sets and where C_B is defined in (3.1)*.*

This family of kernels contains the Gaussian kernel, being the particular case $m = 1$, and includes as limiting cases the Poisson kernel (2.1) (when $\alpha \to 0$) and C_B, in view of (5.7). Some illustrations of this model are provided in the supplementary material, including graphical representations of the pcf and some realizations.

5.3 Families of compactly supported kernels

As suggested by Proposition 4.1, we can consider the following family of compactly supported kernels, parameterized by the range $R > 0$,

$$C_1(x) = \frac{1}{\|h\|^2}[h * h]\left(\frac{2x}{R}\right) C_B(x), \quad \forall x \in \mathbb{R}^d, \tag{5.8}$$

where h is given by (4.1). The Poisson kernel (2.1) and C_B are two limiting cases, when respectively $R \to 0$ and $R \to +\infty$. However this family of kernels has several drawbacks: No closed form expression is available for C_1, nor for $\mathcal{F}(C_1)$. Moreover, when the range R is fixed, $DPP(C_1)$ is not the most repulsive DPP, see Proposition 4.3 and the graphical representations in the supplementary material. This is the reason why we turn to another family of compactly supported kernels.

Following Proposition 4.3, we introduce a new family of compactly supported kernels with range R, given as a convolution product of functions as in (4.6). Specifically, let $R > 0$, $\rho > 0$ and $\alpha > 0$ such that $R/(2\alpha)$ is not a zero of the Bessel function $J_{\frac{d-2}{2}}$ and consider the kernel $C_2 = u * u$ with

$$u(x) = \sqrt{\rho}\ \beta(R,\alpha)\left(1 - \frac{R^{\frac{d}{2}-1}}{2^{\frac{d}{2}-1}J_{\frac{d}{2}-1}(\frac{R}{2\alpha})}\frac{J_{\frac{d}{2}-1}(|\frac{x}{\alpha}|)}{|x|^{\frac{d}{2}-1}}\right)\mathbf{1}_{\{|x|\le\frac{R}{2}\}}, \tag{5.9}$$

where

$$\beta(R,\alpha) = \left[\frac{R^{d-1}\pi^{d/2}}{2^{d-1}\Gamma\left(\frac{d}{2}\right)}\left(\frac{R}{d} - 4\alpha\frac{J_{\frac{d}{2}}(\frac{R}{2\alpha})}{J_{\frac{d}{2}-1}\left(\frac{R}{2\alpha}\right)} + \frac{R}{2}\left(1 - \frac{J_{\frac{d}{2}-2}(\frac{R}{2\alpha})J_{\frac{d}{2}}(\frac{R}{2\alpha})}{J^2_{\frac{d}{2}-1}\left(\frac{R}{2\alpha}\right)}\right)\right)\right]^{-\frac{1}{2}}.$$

Proposition 5.4. *Let $C_2 = u*u$ where u is given by (5.9), then its Fourier transform is $\mathcal{F}(u)^2$ where for all $x \in \mathbb{R}^d$*

$$\mathcal{F}(u)(x) = \sqrt{\rho}\ \beta(R,\alpha)\left(\frac{R}{2|x|}\right)^{\frac{d}{2}-1}\left(\frac{R}{2|x|}J_{\frac{d}{2}}(\pi R|x|)\right.$$
$$\left. + \frac{\pi}{J_{\frac{d-2}{2}}(\frac{R}{2\alpha})}\frac{R\alpha J'_{\frac{d-2}{2}}(\frac{R}{2\alpha})J_{\frac{d-2}{2}}(\pi R|x|) - 2\pi R\alpha^2 J_{\frac{d-2}{2}}(\frac{R}{2\alpha})|x|J'_{\frac{d-2}{2}}(\pi R|x|)}{1 - 4\pi^2|\alpha x|^2}\right).$$

Moreover, $DPP(C_2)$ exists if and only if α is such that $|\mathcal{F}(u)| \le 1$. In this case, $DPP(C_2)$ defines a stationary and isotropic R-dependent DPP with intensity ρ.

The choice of u in (5.9) comes from (4.6) where γ has been chosen such that u is continuous at $|x| = R/2$ and where β is deduced from the relation $C_2(0) = \|u\|^2 = \rho$. Given ρ and R, the remaining free parameter in this parametric family becomes α. The restriction that $R/(2\alpha)$ must not be a zero of $J_{\frac{d-2}{2}}$ can be alleviated by setting in these cases $\beta = 0$ in (4.6) and choose γ so that $C_2(0) = \rho$. Then the most locally repulsive DPP (4.4) when $R \le M$ would be part of the parametric family. However, these kernels can be arbitrarily approximated by some kernel given by (5.9) for some value of α, so we do not include these particular values of α in the family above.

The condition $|\mathcal{F}(u)| \leq 1$ on α, given R and ρ, must be checked numerically. In most cases, the maximal value of $\mathcal{F}(u)$ holds at the origin and we simply have to check whether $|\mathcal{F}(u)(0)| \leq 1$. No theoretical results are available to claim the existence of an admissible α, but from our experience, there seems to exist an infinity of admissible α for any R and ρ. Moreover, while the most locally repulsive DPP when $R \leq M$ is known and corresponds to (4.4), the most repulsive DPP when $R > M$ in the above parametric family seems to correspond to the maximal value of α such that $|\mathcal{F}(u)| \leq 1$, denoted $\alpha_{\max}$.

The parametric family given by C_2 is mainly of interest since it covers a large range of repulsive DPPs while the kernels are compactly supported. Moreover, the closed form expression of $\mathcal{F}(C_2)$ is available and this family contains the most locally repulsive DPP with range R, in view of Proposition 4.3, at least when $R \leq M$. Some illustrations are provided in the on-line supplementary material.

6 Proofs

6.1 Proof of Proposition 3.2

As the kernel C_B verifies condition $\mathcal{K}(\rho)$, it defines a DPP with intensity ρ and its associated pcf g_B given by (2.3) vanishes at 0. By the analytic definition of Bessel functions, see [1, Relation (9.1.10)],

$$C_B(x) = \frac{\sqrt{\rho\Gamma(\frac{d}{2}+1)}}{\pi^{d/4}} \sum_{n=0}^{+\infty} \frac{(-1)^n \left(\sqrt{\pi}\Gamma(\frac{d}{2}+1)^{\frac{1}{d}}\rho^{\frac{1}{d}}\right)^{2n}}{2^{2n} n! \Gamma\left(n+1+d/2\right)} |x|^{2n}.$$

Thus C_B is twice differentiable at 0 and by (2.3), so is g_B. By Definition 1.2, any DPP having a pcf g that does not vanish at 0 or is not twice differentiable at 0 is less locally repulsive than $DPP(C_B)$. Consequently we assume in the following of the proof that $g(0) = 0$ and g is twice differentiable at 0. The problem therefore reduces to minimize $\Delta g(0)$ under the constraint that g is the pcf of a DPP with kernel C verifying condition $\mathcal{K}(\rho)$.

According to condition $\mathcal{K}(\rho)$, the Fourier transform of the kernel C is well defined and belongs to $L^1(\mathbb{R}^d)$, as noticed below (2.2). Therefore, we can define the function $f = \frac{\mathcal{F}(C)}{\|\mathcal{F}(C)\|_1}$ where $\|\mathcal{F}(C)\|_1 = \int_{\mathbb{R}^d} |\mathcal{F}(C)(x)| dx$ and consider it as a density function of a random variable $X = (X_1, \cdots, X_d) \in \mathbb{R}^d$. Denote by $\widehat{f}(t) = \mathbb{E}\left(e^{it\cdot X}\right)$ the characteristic function of X. We have

$$\widehat{f}(t) = \frac{C\left(\frac{t}{2\pi}\right)}{\|\mathcal{F}(C)\|_1}, \quad \forall t \in \mathbb{R}^d. \tag{6.1}$$

Thus, $\widehat{f}$ is twice differentiable at 0, so by the usual properties of the characteristic function (see [31]), X has finite second order moments and

$$E(X_i^2) = -\frac{\partial^2 \widehat{f}}{\partial x_i^2}(0) + \left(\frac{\partial \widehat{f}}{\partial x_i}(0)\right)^2, \quad i = 1 \ldots d. \tag{6.2}$$

On the other hand, as already noticed in Section 1, $\nabla g(0) = 0$ and so $\frac{\partial C}{\partial x_i}(0) = 0$ for $i = 1, \ldots, d$. By differentiating both sides of (6.1),

$$\frac{\partial \widehat{f}}{\partial x_i}(0) = \frac{1}{2\pi \|\mathcal{F}(C)\|_1} \frac{\partial C}{\partial x_i}(0) = 0, \quad i = 1 \ldots d \tag{6.3}$$

and

$$\frac{\partial^2 \widehat{f}}{\partial x_i^2}(0) = \frac{1}{4\pi^2 \|\mathcal{F}(C)\|_1} \frac{\partial^2 C}{\partial x_i^2}(0), \quad i = 1 \ldots d. \tag{6.4}$$

Then, by (6.2)-(6.4),

$$E(|X|^2) = E\left(\sum_{i=1}^{d} X_i^2\right) = -\Delta \widehat{f}(0) = -\frac{1}{4\pi^2 \|\mathcal{F}(C)\|_1} \Delta C(0).$$

Moreover,

$$E(|X|^2) = \int_{\mathbb{R}^d} |x|^2 f(x) dx = \int_{\mathbb{R}^d} |x|^2 \frac{\mathcal{F}(C)}{\|\mathcal{F}(C)\|_1}(x) dx.$$

Hence,

$$\Delta C(0) = -4\pi^2 \int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x) dx. \tag{6.5}$$

By (2.3) and since $\nabla C(0) = 0$,

$$\begin{aligned}\Delta g(0) &= \Delta\left(1 - \frac{C^2}{\rho^2}\right)(0) = -\frac{1}{\rho^2}\left(\sum_{i=1}^{d} 2C(0) \frac{\partial^2 C}{\partial x_i^2}(0) + 2\left(\frac{\partial C}{\partial x_i}(0)\right)^2\right) \\ &= -\frac{2}{\rho} \sum_{i=1}^{d} \frac{\partial^2 C}{\partial x_i^2}(0) = -\frac{2}{\rho} \Delta C(0). \end{aligned} \tag{6.6}$$

Finally, we deduce from (6.5) and (6.6) that

$$\Delta g(0) = \frac{8\pi^2}{\rho} \int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x) dx.$$

Thus the two following optimization problems are equivalent.

Problem 1: Minimizing $\Delta g(0)$ under the constraint that g is the pcf of a DPP with kernel C satisfying condition $\mathcal{K}(\rho)$.

Problem 2: Minimizing $\int_{\mathbb{R}} |x|^2 \mathcal{F}(C)(x) dx$ under the constraint that C is a kernel which is twice differentiable at 0 and verifies the condition $\mathcal{K}(\rho)$.

The latter optimization problem is a special case of [25, Theorem 1.14], named bathtub principle, which gives the unique solution $\mathcal{F}(C) = \mathbf{1}_{\{|\cdot|^d \leq \rho \tau^d\}}$ in agreement with (3.1). This completes the proof.

6.2 Proof of Proposition 4.1

Notice that h is symmetric, real-valued, infinitely differentiable and verifies $h(x) = 0$ for $x \geq 1$, see [29, Section 3.2]. Thus, $\|h\|$ is finite and $\|h\| \neq 0$, so C_r is well-defined.

Since $h * h(0) = \|h\|^2$, we have $C_r(0) = \rho$. By product convolution properties, $h * h$ is symmetric, real-valued, infinitely differentiable and compactly supported with range 2. Thus, by (4.2), C_r is symmetric, real-valued, infinitely differentiable and compactly supported with range r. Then, C_r belongs to $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. In particular, $\mathcal{F}(C_r)$ is well-defined pointwise. By well-known properties of the Fourier transform, for all $x \in \mathbb{R}^d$,

$$\mathcal{F}(C_r)(x) = \frac{r^d}{2^d\|h\|^2}\left[\mathcal{F}(h)^2\left(\frac{r}{2}\cdot\right) * \mathcal{F}(C)(\cdot)\right](x). \tag{6.7}$$

Since h is symmetric, $\mathcal{F}(h)$ is real valued, so $\mathcal{F}(h)^2 \geq 0$. Thus, as $\mathcal{F}(C) \geq 0$ by condition $\mathcal{K}(\rho)$, we have $\mathcal{F}(C_r) \geq 0$. Further, since $0 \leq \mathcal{F}(C) \leq 1$,

$$\frac{r^d}{2^d\|h\|^2}\int_{\mathbb{R}^d}\mathcal{F}(h)^2\left(\frac{rt}{2}\right)\mathcal{F}(C)(x-t)dt \leq \frac{r^d}{2^d\|h\|^2}\int_{\mathbb{R}^d}\mathcal{F}(h)^2\left(\frac{rt}{2}\right)dt. \tag{6.8}$$

By the substitution $u = rt/2$ and Parseval's equality, the right-hand side of (6.8) equals 1. Finally, (6.7) and (6.8) give $\mathcal{F}(C_r) \leq 1$, i.e. $0 \leq \mathcal{F}(C_r) \leq 1$.

It remains to show the convergence result (4.3), which reduces to prove that $\frac{1}{\|h\|^2}[h * h]\left(\frac{2}{r}\cdot\right)$ tends to 1 uniformly on all compact set when $r \to \infty$. This follows from $h * h(0) = \|h\|^2$ and the uniform continuity of $h * h$ on every compact set.

6.3 Proof of Proposition 4.2

The proof is based on a theorem from Ehm et al. [8] recalled below with only slight changes in the presentation.

Definition 6.1. *Let H denote the normalized Haar measure on the group $SO(d)$ of rotations in $\mathbb{R}^d$ and let C be a kernel verifying condition $\mathcal{K}_c(\rho, R)$. The radialization of the kernel C is the kernel* $\mathrm{rad}(C)$ *defined by*

$$\mathrm{rad}(C)(x) = \int_{SO(d)} C(j(x))H(dj).$$

Note that for any isotropic kernel C, $C = \mathrm{rad}(C)$. We say that $C_1 = C_2$ up to a radialization if C_1 and C_2 are kernels verifying condition $\mathcal{K}_c(\rho, R)$ and $\mathrm{rad}(C_1) = \mathrm{rad}(C_2)$.

Define $\gamma_d > 0$ by $\gamma_d^2 = \frac{4j_{(d-2)/2}^{d-2}}{\pi^{\frac{d}{2}}\Gamma\left(\frac{d}{2}\right)J_{\frac{d}{2}}^2(j_{(d-2)/2})}$ and set $c_d = \frac{4j_{\frac{d-2}{2}}^2}{4^d\pi^{\frac{d}{2}}\Gamma\left(\frac{d}{2}\right)}$ where $j_{(d-2)/2}$ is introduced before Proposition 4.2.

Theorem 6.2 ([8])**.** *Let Ψ be a twice differentiable characteristic function of a probability density f on $\mathbb{R}^d$ and suppose that $\Psi(x) = 0$ for $|x| \geq 1$. Then*

$$-\Delta\Psi(0) = \int |x|^2 f(x)dx \geq 4j_{(d-2)/2}^2$$

with equality if and only if, up to a radialization, $\Psi = \omega_d * \omega_d$*, where* $\omega_d(x) = \gamma_d \frac{\Gamma\left(\frac{d}{2}\right)}{j_{(d-2)/2}^{(d-2)/2}} \frac{J_{\frac{d-2}{2}}(2j_{\frac{d-2}{2}}|x|)}{|x|^{\frac{d-2}{2}}}$ *for* $|x| \leq \frac{1}{2}$ *and* $\omega_d(x) = 0$ *for* $|x| \geq \frac{1}{2}$*. The corresponding minimum variance density is*

$$f(x) = c_d \Gamma\left(\frac{d}{2}\right)^2 \left(\frac{2^{\frac{d-2}{2}} J_{\frac{d-2}{2}}\left(\frac{|x|}{2}\right)}{\left|\frac{x}{2}\right|^{\frac{d-2}{2}} \left(j_{(d-2)/2}^2 - \left(\frac{|x|}{2}\right)^2 \right)} \right)^2 .$$

According to Definition 1.2 and by the same arguments as in the proof of Proposition 3.2 and (6.6), we seek a kernel C which is twice differentiable at 0 such that $\Delta C(0)$ is maximal among all kernels verifying condition $\mathcal{K}_c(\rho, R)$.

In a first step, we exhibit a candidate for the solution to this optimization problem and in a second step we check that it verifies all required conditions.

Step 1. We say that a function C verifies $\widetilde{\mathcal{K}_c}(\rho, R)$ if it verifies $\mathcal{K}_c(\rho, R)$ without necessarily verifying $\mathcal{F}(C) \leq 1$. Notice that a function C verifies $\widetilde{\mathcal{K}_c}(\rho, R)$ if and only if the function

$$\Psi(x) = \frac{C(Rx)}{\rho}, \ x \in \mathbb{R}^d, \tag{6.9}$$

verifies $\widetilde{\mathcal{K}_c}(1,1)$. Therefore, we have a one-to-one correspondence between $\widetilde{\mathcal{K}_c}(\rho, R)$ and $\widetilde{\mathcal{K}_c}(1,1)$.

On the other hand, if a function Ψ verifies condition $\widetilde{\mathcal{K}_c}(1,1)$, it is by Bochner's Theorem the characteristic function of a random variable X. Moreover, the function Ψ is continuous and compactly supported, so it is in $L^1(\mathbb{R}^d)$ and the random variable X has a density f, see [31]. Thus, by Theorem 6.2, any function Ψ twice differentiable at 0 and verifying condition $\widetilde{\mathcal{K}_c}(1,1)$ satisfies

$$\Delta\Psi(0) \leq -4j_{(d-2)/2}^2. \tag{6.10}$$

By differentiating both sides of (6.9), we have

$$\Delta\Psi(0) = \frac{R^2}{\rho}\Delta C(0). \tag{6.11}$$

Thus, by (6.10)-(6.11), for any kernel C which is twice differentiable at 0 and verifies $\widetilde{\mathcal{K}_c}(\rho, R)$,

$$\Delta C(0) = \frac{\rho \Delta\Psi(0)}{R^2} \leq -\frac{4\rho j_{(d-2)/2}^2}{R^2}. \tag{6.12}$$

By Theorem 6.2, the equality in (6.12) holds if and only if $\Psi = \omega_d * \omega_d$ and we name C_R the corresponding kernel C given by (6.9).

Step 2. The kernel C_R is the candidate to our optimization problem, however it remains to prove that it verifies condition $\mathcal{K}_c(\rho, R)$. We have seen in *Step 1* that C_R

verifies $\widetilde{\mathcal{K}_c}(\rho, R)$ and is twice differentiable at 0. We must show that $\mathcal{F}(C_R) \leq 1$. By Theorem 6.2, the function $\Psi = \omega_d * \omega_d$ is the characteristic function of a probability density f. Thus, for all $x \in \mathbb{R}^d$,

$$\mathcal{F}(\Psi)(x) = (2\pi)^d f(2\pi x) = (2\pi)^d c_d \Gamma\left(\frac{d}{2}\right)^2 \left(\frac{2^{\frac{d-2}{2}} J_{\frac{d-2}{2}}(|\pi x|)}{|\pi x|^{\frac{d-2}{2}} \left(j^2_{(d-2)/2} - (|\pi x|)^2\right)}\right)^2. \tag{6.13}$$

By (6.9) and the Fourier transform dilatation we thereby obtain (4.5).

Moreover, the Bessel functions are non-negative up to their first non-negative zero so $\omega_d \geq 0$, which implies that $\Psi \geq 0$. Hence by (6.13),

$$\mathcal{F}(\Psi)(x) = \left|\int_{\mathbb{R}^d} \Psi(t) e^{2i\pi x\cdot t} dt\right| \leq \int_{\mathbb{R}^d} \Psi(t) dt = \mathcal{F}(\Psi)(0) = \frac{2^d \pi^d c_d}{j^4_{\frac{d-2}{2}}}. \tag{6.14}$$

Thus, by (6.9) and the Fourier transform dilatation,

$$\mathcal{F}(C_R)(x) \leq \mathcal{F}(C_R)(0) = \frac{2^d R^d \rho \pi^d c_d}{j^4_{\frac{d-2}{2}}} = \frac{R^d}{M^d}. \tag{6.15}$$

Since by hypothesis $R \leq M$, we have $\mathcal{F}(C_R) \leq 1$.

6.4 Proof of Proposition 4.3

According to Definition 1.2 and by the same arguments as in the proof of Proposition 3.2 and (6.6), we seek a kernel C which is twice differentiable at 0 such that $\Delta C(0)$ is maximal among all kernels verifying condition $\mathcal{K}_c(\rho, R)$. By (6.5), this is equivalent to solve the following problem A.

Problem A: Minimize $\int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x)dx$ under the constraints that C is twice differentiable at 0 and verifies $\mathcal{K}_c(\rho, R)$.

The proof of Proposition 4.3 is based on the following three lemmas. In the first lemma, the gradient ∇u has to be considered in the sense of distribution when $u \in L^2(\mathbb{R}^d)$ is not differentiable.

Lemma 6.3. *A kernel C_R is solution to Problem A if and only if there exists a function u such that, up to a radialization, $C_R = u * u$ where u minimizes $\int_{\mathbb{R}^d} |\nabla u(x)|^2 dx$ among all functions u verifying $\mathcal{M}(\rho, R)$ and $\mathcal{F}(u)^2 \leq 1$.*

The existence statement in Proposition 4.3 is given by the following lemma.

Lemma 6.4. *There exists a solution to Problem A.*

By Lemma 6.3, $C_R = u * u$ where u is the solution of the given optimization problem. Then, under the additional constraint $\sup_{x\in\mathbb{R}^d} \mathcal{F}(C)(x) = \mathcal{F}(C)(0)$, we have $\sup_{x\in\mathbb{R}^d} (\mathcal{F}(u)(x))^2 = (\mathcal{F}(u)(0))^2$. Since $\mathcal{F}(u)^2(0) = \left(\int_{\mathbb{R}^d} u(t)dt\right)^2$, the constraint

$\mathcal{F}(u)^2 \leq 1$ in Lemma 6.3 becomes $\left(\int_{\mathbb{R}^d} u(t)dt\right)^2 \leq 1$. Notice that $-u$ is also a solution of the optimization problem. Thus, we can assume without loss of generality that $\int_{\mathbb{R}^d} u(t)dt \geq 0$, so that the constraint $\left(\int_{\mathbb{R}^d} u(t)dt\right)^2 \leq 1$ becomes $\int_{\mathbb{R}^d} u(t)dt \leq 1$. In this situation, the optimization problem addressed in Lemma 6.3 can be solved by variational calculus. However, an explicit form of the solution is available only if we assume that $u \in \mathcal{C}^2(B\left(0, \frac{R}{2}\right))$, meaning that u is twice continuously differentiable on its support. It is given by the following lemma, which completes the proof of Proposition 4.3.

Lemma 6.5. *If a function u minimizes $\int_{\mathbb{R}^d} |\nabla u(x)|^2 dx$ among all functions u verifying $\mathcal{M}(\rho, R)$, $u \in \mathcal{C}^2(B\left(0, \frac{R}{2}\right))$ and $\int_{\mathbb{R}^d} u(x)dx \leq 1$, then u is of the form*

$$u(x) = \left(\beta + \gamma \frac{J_{\frac{d-2}{2}}\left(|x|/\alpha\right)}{|x|^{\frac{d-2}{2}}}\right) \mathbf{1}_{\left\{|x|<\frac{R}{2}\right\}},$$

where $\alpha > 0, \beta \geq 0$ and γ are three constants linked by the conditions $\mathcal{M}(\rho, R)$ and $\int_{\mathbb{R}^d} u(x)dx \leq 1$.

Proof of Lemma 6.3

Let C be a kernel which is twice differentiable at 0 and verifies the condition $\mathcal{K}_c(\rho, R)$. This implies that C is twice differentiable everywhere. Moreover, the quantity $\int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x)dx$ is invariant under radialization of the kernel C, see [8, Relation (44)]. Thus, we can consider C as a radial function. Then, by [8, Theorem 3.8], there exists a countable set A and a sequence of real valued functions $\{u_k\}_{k\in A}$ in $L^2(\mathbb{R}^d)$ such that

$$C(x) = \sum_{k\in A} u_k * u_k(x). \tag{6.16}$$

Further, the convergence of the series is uniform and for each $k \in A$, the support of u_k lies in $B\left(0, \frac{R}{2}\right)$. Thus,

$$\int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x)dx = \int_{\mathbb{R}^d} |x|^2 \sum_{k\in A} \left|\mathcal{F}(u_k)(x)\right|^2 dx = \sum_{k\in A} \sum_{j=1}^{d} \int_{\mathbb{R}^d} \left|x_j \mathcal{F}(u_k)(x)\right|^2 dx \tag{6.17}$$

where x_j denotes the j-th coordinate of the vector x. In addition, we note that $u_k \in L^2(\mathbb{R}^d)$ so $|\cdot| \mathcal{F}(u_k)(\cdot) \in L^2(\mathbb{R}^d)$ by (6.17). Then, by [25, Theorem 7.9], $\nabla u_k \in L^2(\mathbb{R}^d)$ where ∇u_k has to be viewed in the distributional sense and

$$\mathcal{F}\left(\partial_j u_k\right)(x) = 2i\pi x_j \mathcal{F}(u_k)(x). \tag{6.18}$$

Thus, from (6.17)-(6.18) and the Parseval equality,

$$\int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x)dx = \sum_{k\in A} \int_{\mathbb{R}^d} \frac{\left|\nabla u_k(x)\right|^2}{4\pi^2} dx.$$

As every term in the sum above is positive and since this equality holds for every kernel C, the minimum of $\int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x)dx$ is reached if and only if this sum reduces to one term where $u_k = u$. Then we have $C = u * u$ and

$$\int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x)dx = \int_{\mathbb{R}^d} \frac{|\nabla u(x)|^2}{4\pi^2} dx. \tag{6.19}$$

Therefore, minimizing $\int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x)dx$ is equivalent to minimize $\int_{\mathbb{R}^d} |\nabla u(x)|^2 \, dx$. Hence it remains to see what the constraints on the kernel C means for the function u. Since $C = u * u$, where u is one of the function in the decomposition (6.16), u is a so-called real valued Boas-Kac root of C, see [8]. Thus, since C is radial, we have by [8, Theorem 3.1] that u is radial and verifies $u(x) = 0$ for $|x| \geq \frac{R}{2}$. Since C verifies $\mathcal{K}_c(\rho, R)$, we have $C(0) = \rho$ and $0 \leq \mathcal{F}(C) \leq 1$. These constraints are equivalent on u to $\int_{\mathbb{R}^d} u(x)^2 dx = \rho$ and $\mathcal{F}(u)^2 \leq 1$, respectively. Therefore, u verifies condition $\mathcal{M}(\rho, R)$ and $\mathcal{F}(u)^2 \leq 1$.

Proof of Lemma 6.4

According to Lemma 6.3, C_R is a is solution to Problem A if and only if $C_R = u * u$ where u minimizes $\int_{\mathbb{R}^d} |\nabla u(x)|^2 dx$ among all functions u verifying $\mathcal{M}(\rho, R)$ and $\mathcal{F}(u)^2 \leq 1$. We prove the existence of such a minimum u.

Let Ω denote the open euclidean ball $B\left(0, \frac{R}{2}\right)$. Consider the Sobolev space

$$H^1(\Omega) = \left\{ f : \Omega \to \mathbb{R},\ f \in L^2(\Omega),\ \nabla f \in L^2(\Omega) \right\},$$

with the norm $\|f\|_{H^1(\Omega)} = (\|f\|^2 + \|\nabla f\|^2)^{\frac{1}{2}}$. For a review on Sobolev spaces, see for example [9] or [25]. For any $f \in H^1(\Omega)$, we consider its extension to $\mathbb{R}^d$ by setting $f(x) = 0$ if $x \notin \Omega$, so that $f \in L^2(\mathbb{R}^d)$. Let us further denote $\mathcal{E}$ the set of functions $f \in H^1(\Omega)$ verifying $\mathcal{M}(\rho, R)$ and $\mathcal{F}(f)^2 \leq 1$.

If the minimum u above exists but $u \notin H^1(\Omega)$, then $\int_\Omega |\nabla u(x)|^2 dx = \infty$, which means that $\mathcal{E}$ is empty, otherwise u would not be the solution of our optimization problem. But $\mathcal{E}$ is not empty, see for instance the functions in Section 5.3, so if u exists, $u \in H^1(\Omega)$. Let $(w_k)_{k\in\mathbb{N}}$ be a minimizing sequence in $\mathcal{E}$, i.e.

$$\int_\Omega |\nabla w_k(x)|^2 \, dx \underset{k\to+\infty}{\longrightarrow} \inf_{v\in\mathcal{E}} \int_\Omega |\nabla v(x)|^2 \, dx, \tag{6.20}$$

where for all k, $w_k \in \mathcal{E}$. By (6.20) and since for all k, $\int_\Omega |w_k(x)|^2 dx = \rho$, the sequence $\{w_k\}$ is bounded in $H^1(\Omega)$. Then, by the Rellich-Kondrachov compactness theorem (see [9]), it follows that, up to a subsequence, $\{w_k\}$ converges in $L^2(\mathbb{R}^d)$ to a certain function $w \in L^2(\mathbb{R}^d)$ verifying

$$\int_\Omega |\nabla w(x)|^2 \, dx = \inf_{v\in\mathcal{E}} \int_\Omega |\nabla v(x)|^2 \, dx. \tag{6.21}$$

We now prove that $w \in \mathcal{E}$, so that $u = w$ is the solution of our optimization problem. First $w \in H^1(\Omega)$ as justified earlier and so $w \in L^2(\mathbb{R}^d)$. Second, as rotations are isometric functions and since any w_k is radial by hypothesis, we have for any $j \in SO(d)$

$$\left\{ \int_{\mathbb{R}^d} |w(x) - w_k(x)|^2 \, dx \to 0 \right\} \Leftrightarrow \left\{ \int_{\mathbb{R}^d} |w(j(x)) - w_k(j(x))|^2 \, dx \to 0 \right\}$$
$$\Leftrightarrow \left\{ \int_{\mathbb{R}^d} |w(j(x)) - w_k(x)|^2 \, dx \to 0 \right\}.$$

Hence, by uniqueness of the limit, the function w is radial and in particular, its Fourier transform is real. Further, since w is the limit in $L^2(\mathbb{R}^d)$ of w_k, w verifies the following properties:

- w is compactly supported in $B\left(0, \frac{R}{2}\right)$, because $w_k \in \mathcal{E}$ for all k.
- $w \in L^2(\mathbb{R}^d)$ by Rellich-Kondrachov theorem.
- $\int_{\mathbb{R}^d} |w(x)|^2 dx = \int_{\mathbb{R}^d} |w_k(x)|^2 dx = \rho$ since a sphere in $L^2(\mathbb{R}^d)$ is closed.

Therefore, w verifies $\mathcal{M}(\rho, R)$. Third, for every k, w_k being compactly supported and in $L^2(\mathbb{R}^d)$, $w_k \in L^1(\mathbb{R}^d)$ so we can consider $\mathcal{F}(w_k)(x)$ for every $x \in \mathbb{R}^d$ and by the Cauchy-Schwartz inequality

$$|\mathcal{F}(w)(x) - \mathcal{F}(w_k)(x)| \leq a \sqrt{\int_{\mathbb{R}^d} |w(t) - w_k(t)|^2 \, dt}, \quad \forall x \in \mathbb{R}^d,$$

where a is a positive constant. Thereby the convergence of w_k to w in $L^2(\mathbb{R}^d)$ implies the pointwise convergence of $\mathcal{F}(w_k)$ to $\mathcal{F}(w)$. Finally, from the relation

$$\mathcal{F}(w_k)(x) \leq 1, \quad \forall x \in \mathbb{R}^d, \quad \forall k \in \mathbb{N},$$

we deduce $\mathcal{F}(w) \leq 1$.

Proof of Lemma 6.5

We denote as before $\Omega = B\left(0, \frac{R}{2}\right)$. The optimization problem in Lemma 6.5 is a variational problem with isoperimetric constraints. By [14, Chapter 2, Theorem 2], every solution must solve

$$\begin{aligned} \Delta u + \lambda_1 u - \frac{\lambda_2}{2} &= 0 \text{ on } \Omega, \\ u &= 0 \text{ on } \partial\Omega. \end{aligned} \tag{6.22}$$

In equation (6.22), λ_1 and λ_2 are the Lagrange multipliers associated to the constraints $\int u^2 = \rho$ and $\int u \leq 1$, respectively. By the Karush-Kuhn-Tucker theorem, see [17, Section VII], $\lambda_2 \geq 0$. Moreover, a solution to the partial differential equation with boundary condition (6.22) is obtained by linear combination of a homogeneous solution and a particular solution. By [9, Section 6.5, Theorem 2], the Laplacian operator $-\Delta$ has only positive eigenvalues. Hence the associated homogeneous equation $\Delta u + \lambda_1 u = 0$ can have a solution only if $\lambda_1 > 0$.

In addition, the function u is radial by hypothesis, so there exists a function $\tilde{u}$ on $\mathbb{R}$ such that $u(x) = \tilde{u}(|x|)$ for all $x \in \mathbb{R}^d$. The partial differential equation (6.22) then becomes

$$\tilde{u}''(t) + \frac{d-1}{t}\tilde{u}'(t) + \lambda_1 \tilde{u}(t) - \frac{\lambda_2}{2} = 0, \quad \forall t \in \left]0, \frac{R}{2}\right[,$$
$$\tilde{u}\left(\frac{R}{2}\right) = 0.$$

As λ_1 is positive, we obtain from [36, Section 4.31, Relations (3) and (4)] that a solution to this equation is of the form

$$\tilde{u}(t) = \left(\frac{\lambda_2}{2\lambda_1} + c_1 \frac{J_{(d-2)/2}(\sqrt{\lambda_1}t)}{t^{(d-2)/2}} + c_2 \frac{Y_{(d-2)/2}(\sqrt{\lambda_1}t)}{t^{(d-2)/2}}\right) \mathbf{1}_{\left\{0<t<\frac{R}{2}\right\}}, \tag{6.23}$$

where $Y_{(d-2)/2}$ denotes the Bessel function of the second kind. By hypothesis, the function u is continuous on Ω and so at 0. Since $Y_{(d-2)/2}$ has a discontinuity at 0, see for example [1], and the remaining terms in (6.23) are continuous, we must have $c_2 = 0$. Then, by renaming the constant c_1 by γ and letting $\alpha = 1/\sqrt{\lambda_1}$, $\beta = \lambda_2/(2\lambda_1)$, we obtain that if u is solution to the optimization problem of Lemma 6.5, then u writes

$$u(x) = \left(\beta + \gamma \frac{J_{(d-2)/2}(|x|/\alpha)}{|x|^{(d-2)/2}}\right) \mathbf{1}_{\{x\in\Omega\}}, \tag{6.24}$$

where $\alpha > 0$ and $\beta \geq 0$.

6.5 Proof of Proposition 5.1

Let C be given by (5.1). According to Proposition 2.3, $DPP(C)$ exists and has intensity ρ if C verifies the condition $\mathcal{K}(\rho)$. By [1, Equation (9.1.7)], we have $C(0) = \rho$. It is immediate that C is a symmetric continuous real-valued function. Since Bessel functions are analytic and by the asymptotic form in [1, (9.2.1)], it is clear that C belongs to $L^2(\mathbb{R}^d)$. It remains to obtain $\mathcal{F}(C)$ and verify the condition $0 \leq \mathcal{F}(C) \leq 1$.

Define

$$p_\sigma(x) = \frac{J_{\frac{\sigma+d}{2}}(|x|)}{|x|^{\frac{\sigma+d}{2}}}, \quad \forall x \in \mathbb{R}^d. \tag{6.25}$$

As p_σ is radial, by [16, Appendix B.5],

$$\mathcal{F}(p_\sigma)(x) = \frac{2\pi}{|x|^{\frac{d-2}{2}}} \int_0^{+\infty} r^{\frac{d}{2}} p_\sigma(r) J_{\frac{d-2}{2}}(2\pi r|x|)dr.$$

By [15, Formula 6.575], we have for $\sigma > -2$

$$\mathcal{F}(p_\sigma)(x) = \frac{2\pi}{|2\pi x|^{\frac{d-2}{2}}} \frac{(1-|2\pi x|^2)_+^{\frac{\sigma}{2}} |2\pi x|^{\frac{d-2}{2}}}{2^{\frac{\sigma}{2}}\Gamma(\frac{\sigma}{2}+1)} = 2^{\frac{d}{2}-\frac{\sigma}{2}} \pi^{\frac{d}{2}} \frac{(1-|2\pi x|^2)_+^{\frac{\sigma}{2}}}{\Gamma(\frac{\sigma+2}{2})}.$$

Since $C(x) = \rho 2^{\frac{\sigma+d}{2}} \Gamma\left(\frac{\sigma+d+2}{2}\right) p_\sigma \left(2\frac{x}{\alpha}\sqrt{\frac{\sigma+d}{2}}\right)$, we obtain (5.2) by dilatation of the Fourier transform.

We have obviously $\mathcal{F}(C) \geq 0$. Since $\sigma \geq 0$, $\mathcal{F}(C)$ attains its maximum at 0. Thus $\mathcal{F}(C) \leq 1$ if and only if

$$\mathcal{F}(C)(0) = \frac{\rho(2\pi)^{\frac{d}{2}}\alpha^d\Gamma(\frac{\sigma+d+2}{2})}{(\sigma+d)^{\frac{d}{2}}\Gamma(\frac{\sigma+2}{2})} \leq 1,$$

which is equivalent to $\alpha^d \leq \frac{(\sigma+d)^{\frac{d}{2}}\Gamma\left(\frac{\sigma+2}{2}\right)}{\rho(2\pi)^{\frac{d}{2}}\Gamma\left(\frac{\sigma+d+2}{2}\right)}$.

Finally, when $\sigma = 0$ and $\alpha = \alpha_{\max}$, $DPP(C)$ exists and a straightforward calculation gives $C = C_B$. The convergence result (5.3) may be found in [11] and is a direct application of [30, Relation (1.8)].

6.6 Proof of Proposition 5.3

Define, for all $m \in \mathbb{N}$,

$$f_m(x) = L_m^{d/2}\left(|x|^2\right) e^{-|x|^2}, \quad \forall x \in \mathbb{R}^d. \tag{6.26}$$

This function is radial, thus by [16, Appendix B.5] we have

$$\mathcal{F}(f_m)(x) = \frac{2\pi}{|x|^{\frac{d-2}{2}}} \int_0^{+\infty} r^{\frac{d}{2}} L_m^{\frac{d}{2}}(r^2) e^{-r^2} J_{\frac{d-2}{2}}(2\pi r|x|) dr.$$

According to [20], we have

$$\begin{aligned}
\mathcal{F}(f_m)(x) &= \frac{2\pi}{|x|^{\frac{d-2}{2}}} \frac{(-1)^m}{2} \left(\frac{|2\pi x|}{2}\right)^{\frac{d-2}{2}} e^{-\frac{|2\pi x|^2}{4}} L_m^{-1-m}\left(\frac{|2\pi x|^2}{4}\right) \\
&= \pi^{\frac{d}{2}}(-1)^m e^{-|\pi x|^2} \sum_{k=0}^m \binom{-1}{m-k} \frac{(-1)^k |\pi x|^{2k}}{k!} \\
&= \pi^{\frac{d}{2}}(-1)^m e^{-|\pi x|^2} \sum_{k=0}^m (-1)^{m-k} \frac{(-1)^k |\pi x|^{2k}}{k!}.
\end{aligned}$$

Therefore,

$$\mathcal{F}(f_m)(x) = \pi^{\frac{d}{2}} e^{-|\pi x|^2} \sum_{k=0}^m \frac{|\pi x|^{2k}}{k!}.$$

As $C(x) = \frac{\rho}{\binom{m-1+\frac{d}{2}}{m-1}} f_{m-1}\left(\frac{1}{\sqrt{m}}\frac{x}{\alpha}\right)$, we obtain (5.5) by dilatation and linearity of the Fourier transform.

Clearly, $\mathcal{F}(C) \geq 0$. Thus we investigate the condition $\mathcal{F}(C) \leq 1$ for the existence of $DPP(C)$. We notice from (5.5) that

$$\mathcal{F}(C)(x) = ae^{-b|x|^2} \sum_{k=0}^{m-1} \frac{b^k |x|^{2k}}{k!}, \tag{6.27}$$

where a and b are positive constants. Since $\mathcal{F}(C)$ depends on the variable x only through its norm, we consider the function h define for all $r \geq 0$ by $h(r) = \mathcal{F}(C)\left((r, 0, \cdots, 0)\right)$, so that for all $x \in \mathbb{R}^d$, $\mathcal{F}(C)(x) = h(|x|)$. For every $r > 0$, h is differentiable at r and a straightforward calculation leads to

$$h'(r) = ae^{-br^2} \left(-2br \sum_{k=0}^{m-1} \frac{b^k r^{2k}}{k!} + \sum_{k=1}^{m-1} 2k \frac{b^k r^{2k-1}}{k!} \right) = -2ae^{-br^2} \frac{b^m r^{2m-1}}{(m-1)!}.$$

Thus, the function h is decreasing on $(0, +\infty)$. Since h is continuous on $\mathbb{R}^+$, its maximum is attained at zero, so for every $x \in \mathbb{R}^d$,

$$\mathcal{F}(C)(x) \leq \mathcal{F}(C)(0) = \frac{\rho\,(m\pi)^{\frac{d}{2}}}{\binom{m-1+\frac{d}{2}}{m-1}} \alpha^d.$$

Hence, $\mathcal{F}(C) \leq 1$ if and only if $\alpha^d \leq \frac{\binom{m-1+\frac{d}{2}}{m-1}}{\rho(m\pi)^{\frac{d}{2}}}$. Moreover C is radial and since $L_{m-1}^{d/2}(0) = \binom{m-1+\frac{d}{2}}{m-1}$, see [1, Relation (22.4.7)], we have $C(0) = \rho$. Therefore, C verifies the condition $\mathcal{K}(\rho)$ and by Proposition 2.3, $DPP(C)$ exists and is stationary with intensity $\rho > 0$.

It remains to prove the convergence results (5.6) and (5.7). An immediate application of [35, Theorem 8.1.3] gives the convergence (5.6), see also [2, Proposition 1]. Moreover,

$$\lim_{m \to +\infty} \alpha_{\max} = \frac{1}{\sqrt{\pi}\Gamma\left(\frac{d}{2}+1\right)^{\frac{1}{d}} \rho^{\frac{1}{d}}}. \tag{6.28}$$

Hence, by (6.28) and (5.6), we obtain the convergence (5.7).

6.7 Proof of Proposition 5.4

By the discussion in Section 4, $DPP(C)$ exists and is an R-dependent DPP with intensity ρ if C verifies $\mathcal{K}_c(\rho, R)$. Since $u \in L^2(\mathbb{R}^d)$, the kernel C is continuous by [25, Theorem 2.20]. Moreover, $u(x) = 0$ for $|x| > \frac{R}{2}$, so by product convolution properties, $C(x) = 0$ for $|x| > R$. Hence C belongs to $L^2(\mathbb{R}^d)$. Since u is radial, so is C. It remains to verify that $0 \leq \mathcal{F}(C) \leq 1$ and $C(0) = \rho$.

By product convolution properties, we have $C(0) = \int_{\mathbb{R}^d} u(x)^2 dx$. From the definition of u in (5.9), we have

$$\begin{aligned}
&\frac{\int_{\mathbb{R}^d} u^2(x)dx}{\rho\beta(R,\alpha)^2}\\
&= \int_{\mathbb{R}^d}\left(1-2\left(\frac{R}{2}\right)^{\frac{d}{2}-1}\frac{J_{\frac{d-2}{2}}(\left|\frac{x}{\alpha}\right|)}{J_{\frac{d-2}{2}}(\frac{R}{2\alpha})|x|^{\frac{d-2}{2}}}+\left(\frac{R}{2}\right)^{d-2}\frac{J^2_{\frac{d-2}{2}}(\left|\frac{x}{\alpha}\right|)}{J^2_{\frac{d-2}{2}}(\frac{R}{2\alpha})|x|^{d-2}}\right)\mathbf{1}_{\{|x|\le\frac{R}{2}\}}dx\\
&= \frac{2\pi^{\frac{d}{2}}}{\Gamma(d/2)}\int_0^{\frac{R}{2}}\left(r^{d-1}-2\left(\frac{R}{2}\right)^{\frac{d-2}{2}}\frac{J_{\frac{d-2}{2}}(\frac{r}{\alpha})}{J_{\frac{d-2}{2}}(\frac{R}{2\alpha})}r^{\frac{d}{2}}+\left(\frac{R}{2}\right)^{d-2}\frac{J^2_{\frac{d-2}{2}}(\frac{r}{\alpha})}{J^2_{\frac{d-2}{2}}(\frac{R}{2\alpha})}r\right)dr.
\end{aligned}$$

By properties of Bessel functions, see [1], we notice that for all $b\in\mathbb{R}$, a primitive of $xJ^2_{\frac{d-2}{2}}(bx)$ is given by $\frac{x^2}{2}\left(J^2_{\frac{d-2}{2}}(xb)-J_{\frac{d}{2}-2}(xb)J_{\frac{d}{2}}(xb)\right)$. It follows from [16, Appendix B.3] that

$$\begin{aligned}
\frac{\int_{\mathbb{R}^d} u^2(x)dx}{\rho\beta(R,\alpha)^2} &= \frac{\pi^{d/2}R^d}{d\,\Gamma(\frac{d}{2})2^{d-1}}-4\left(\frac{R}{2}\right)^{d-1}\frac{\pi^{\frac{d}{2}}\alpha}{\Gamma(\frac{d}{2})}\frac{J_{\frac{d}{2}}(\frac{R}{2\alpha})}{J_{\frac{d}{2}-1}(\frac{R}{2\alpha})}\\
&\quad+\left(\frac{R}{2}\right)^{d}\frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})}\left(1-\frac{J_{\frac{d}{2}-2}(\frac{R}{2\alpha})J_{\frac{d}{2}}(\frac{R}{2\alpha})}{J^2_{\frac{d}{2}-1}(\frac{R}{2\alpha})}\right).
\end{aligned}$$

Thus, by the definition of $\beta(R,\alpha)$, we obtain that $\int_{\mathbb{R}^d} u(x)^2dx=\rho$.

We now calculate $\mathcal{F}(C)$. We have $\mathcal{F}(C)=\mathcal{F}(u)^2$. Since u is radial, $\mathcal{F}(u)$ is real valued and so $\mathcal{F}(C)\ge 0$. In addition, we have by [16, Appendix B.5] and (5.9),

$$\begin{aligned}
\mathcal{F}(u)(x) = \sqrt{\rho}\beta(R,\alpha)\frac{2\pi}{|x|^{\frac{d-2}{2}}}&\left(\int_0^{\frac{R}{2}} r^{\frac{d}{2}}J_{\frac{d-2}{2}}(2\pi r|x|)dr\right.\\
&\left.-\frac{R^{\frac{d}{2}-1}}{2^{\frac{d}{2}-1}J_{\frac{d}{2}-1}(\frac{R}{2\alpha})}\int_0^{\frac{R}{2}} rJ_{\frac{d-2}{2}}\left(\frac{r}{\alpha}\right)J_{\frac{d-2}{2}}(2\pi r|x|)dr\right).
\end{aligned}$$

Since $\alpha>0$, we have by [16, Appendix B.3] and [15, Formula 6.521],

$$\begin{aligned}
\mathcal{F}(u)(x) &= \sqrt{\rho}\beta(R,\alpha)\frac{2\pi}{|x|^{\frac{d-2}{2}}}\left(\frac{R^{\frac{d}{2}}}{\pi 2^{\frac{d}{2}+1}}\frac{J_{\frac{d}{2}}(\pi R|x|)}{|x|}\right.\\
&\left.+\frac{R^{\frac{d}{2}-1}}{2^{\frac{d}{2}}J_{\frac{d}{2}-1}(\frac{R}{2\alpha})}\frac{R\alpha J'_{\frac{d-2}{2}}(\frac{R}{2\alpha})J_{\frac{d-2}{2}}(\pi R|x|)-2\pi\alpha^2RJ_{\frac{d-2}{2}}(\frac{R}{2\alpha})|x|J'_{\frac{d-2}{2}}(\pi R|x|)}{1-4\pi^2|\alpha x|^2}\right)
\end{aligned}\tag{6.29}$$

from which we deduce the Fourier transform of u in Proposition 5.4. Therefore, if α is such that $\mathcal{F}(u)^2\le 1$, then $\mathcal{F}(C)\le 1$ and so C verifies $\mathcal{K}_c(\rho,R)$.

Acknowledgments

We are grateful to Jean-François Coeurjolly for illuminating comments.

References

[1] Abramowitz, M., and Stegun, I. A. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, vol. 55 of *National Bureau of Standards Applied Mathematics Series.* U.S. Government Printing Office, Washington, D.C., 1964.

[2] Álvarez-Nodarse, R., and Moreno-Balcázar, J. J. Asymptotic properties of generalized Laguerre orthogonal polynomials. *Indagationes Mathematicae. New Series 15*, 2 (2004), 151–165.

[3] Borodin, A., and Serfaty, S. Renormalized energy concentration in random matrices. *Communications in Mathematical Physics 320*, 1 (2013), 199–244.

[4] Daley, D. J., and Vere-Jones, D. *An Introduction to the Theory of Point Processes, Vol. I, Elementary Theory and Methods*, 2 ed. Springer, 2003.

[5] Daley, D. J., and Vere-Jones, D. *An Introduction to the Theory of Point Processes, Vol. II, General Theory and Structure*, 2 ed. Springer, 2008.

[6] Deng, N., Zhou, W., and Haenggi, M. The Ginibre point process as a model for wireless networks with repulsion. *to appear in IEEE Transactions on Wireless Communications. doi: 10.1109/TWC.2014.2332335* (2014).

[7] Domar, Y. An extremal problem for positive definite functions. *Journal of Mathematical Analysis and Applications 52* (1975), 56–63.

[8] Ehm, W., Gneiting, T., and Richards, D. Convolution roots of radial positive definite functions with compact support. *Transactions of the American Mathematical Society 356*, 11 (2004), 4655–4685.

[9] Evans, L. *Partial Differential Equations*, 2 ed., vol. 19 of *Graduate Studies in Mathematics.* American Mathematical Society, 2010.

[10] Fasshauer, G. E., and Zhang, J. G. Iterated approximate moving least squares approximation. In *Advances in Meshfree Techniques.* Springer, 2007, pp. 221–239.

[11] Flyer, N. Exact polynomial reproduction for oscillatory radial basis functions on infinite lattices. *Computers & Mathematics with Applications. 51*, 8 (2006), 1199–1208.

[12] Fornberg, B., Larsson, E., and Wright, G. A new class of oscillatory radial basis functions. *Computers & Mathematics with Applications 51*, 8 (2006), 1209–1222.

[13] Garsia, A., Rodemich, E., and Rumsey, H. On some extremal positive definite functions. *Journal of Mathematics and Mechanics 18*, 9 (1969), 805–834.

[14] Giaquinta, M., and Hildebrandt, S. *Calculus of Variations I.* Springer, 1996.

[15] Gradshteyn, I. S., and Ryzhik, I. M. *Table of Integrals, Series, and Products*, seventh ed. Elsevier/Academic Press, Amsterdam, 2007.

[16] Grafakos, L. *Classical Fourier Analysis.* Graduate Texts in Mathematics. Springer, 2008.

[17] Hiriart-Urruty, J.-B., and Lemaréchal, C. *Convex Analysis and Minimization Algorithms: Part 1: Fundamentals*, vol. 305. Springer, 1993.

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

[19] Illian, J., Penttinen, A., Stoyan, H., and Stoyan, D. *Statistical Analysis and Modelling of Spatial Point Patterns.* John Wiley and Sons, Chichester, 2008.

[20] Kölbig, K. S., and Scherb, H. On a Hankel transform integral containing an exponential function and two Laguerre polynomials. *Journal of Computational and Applied Mathematics 71*, 2 (1996), 357–363.

[21] Kolountzakis, M., and Révész, S. On a problem of Turán about positive definite functions. *Proceedings of the American Mathematical Society 131*, 11 (2003), 3423–3430.

[22] Kulesza, A., and Taskar, B. Determinantal point processes for machine learning. *Foundations and Trends in Machine Learning 5* (2012), 123–286.

[23] Kuna, T., Lebowitz, J. L., and Speer, E. R. Realizability of point processes. *Journal of Statistical Physics 129*, 3 (2007), 417–439.

[24] Lavancier, F., Møller, J., and Rubak, E. Statistical aspects of determinantal point processes. *To appear in Journal of the Royal Statistical Society, series B* (2015).

[25] Lieb, E. H., and Loss, M. *Analysis*, second ed., vol. 14 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2001.

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

[27] Miyoshi, N., and Shirai, T. A cellular network model with Ginibre configured base stations. *Advances in Applied Probability 46*, 3 (2014), 832–845.

[28] Møller, J., and Waagepetersen, R. P. *Statistical Inference and Simulation for Spatial Point Processes*, vol. 100 of *Monographs on Statistics and Applied Probability.* Chapman & Hall/CRC, Boca Raton, FL, 2004.

[29] Sasvári, Z. *Multivariate Characteristic and Correlation Functions*, vol. 50. Walter de Gruyter, 2013.

[30] Schoenberg, I. J. Metric spaces and completely monotone functions. *Annals of Mathematics. Second Series 39*, 4 (1938), 811–841.

[31] Shiryaev, A. *Probability.* Graduate Texts in Mathematics. Springer, 1995.

[32] Soshnikov, A. Determinantal random point fields. *Russian Mathematical Surveys 55* (2000), 923–975.

[33] Stein, E., and Weiss, G. *Introduction to Fourier Analysis on Euclidean Spaces (PMS-32)*, vol. 1. Princeton university press, 1971.

[34] Stoyan, D., Kendall, W. S., and Mecke, J. *Stochastic Geometry and Its Applications.* Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Ltd., Chichester, 1987.

[35] Szegö, G. *Orthogonal Polynomials*, vol. 23. AMS Bookstore, 1992.

[36] Watson, G. N. *A Treatise on the Theory of Bessel Functions.* Cambridge university press, 1995.