



HAL
open science

Projections of determinantal point processes

Adrien Mazoyer, Jean-François Coeurjolly, Pierre-Olivier Amblard

► **To cite this version:**

Adrien Mazoyer, Jean-François Coeurjolly, Pierre-Olivier Amblard. Projections of determinantal point processes. *Spatial Statistics*, 2020, 38, pp.100437. 10.1016/j.spasta.2020.100437 . hal-02990859

HAL Id: hal-02990859

<https://hal.science/hal-02990859>

Submitted on 5 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Projections of determinantal point processes

Adrien Mazoyer^a, Jean-François Coeurjolly^{a,b,*}, Pierre-Olivier Amblard^c

^a*Department of Mathematics, Universit du Qubec Montral
Montreal (Quebec) H2X 3Y7, Canada*

^b*Laboratoire LJK, Universit Grenoble Alpes
38401 Saint Martin d'Hres, France*

^c*Gipsa-Lab, CNRS, Universit Grenoble Alpes
38402 Saint Martin dHres, France*

Abstract

Let $\mathbf{x} = \{x^{(1)}, \dots, x^{(n)}\}$ be a space filling-design of n points defined in $[0,1]^d$. In computer experiments, an important property sought for \mathbf{x} is a nice coverage of $[0,1]^d$. This property could be desirable as well as for any projection of \mathbf{x} onto $[0,1]^\iota$ for $\iota < d$. Thus we expect that $\mathbf{x}_I = \{x_I^{(1)}, \dots, x_I^{(n)}\}$, which represents the design \mathbf{x} with coordinates associated to any index set $I \subseteq \{1, \dots, d\}$, remains regular in $[0,1]^\iota$ where ι is the cardinality of I . This paper examines the conservation of nice coverage by projection using spatial point processes, and more specifically using the class of determinantal point processes. We provide necessary conditions on the kernel defining these processes, ensuring that the projected point process \mathbf{X}_I is repulsive, in the sense that its pair correlation function is uniformly bounded by 1, for all $I \subseteq \{1, \dots, d\}$. We present a few examples, compare them using a new normalized version of Ripley's function. Finally, we illustrate the interest of this research for Monte-Carlo integration.

Keywords: spatial point processes, pair correlation function, Ripley's function, space-filling design.

2000 MSC: 60G55, 62K99

*Corresponding author

Email addresses: mazoyer.adrien@courrier.uqam.ca (Adrien Mazoyer),
coeurjolly.jean-francois@uqam.ca (Jean-François Coeurjolly),
pierre-olivier.amblard@gipsa-lab.grenoble-inp.fr (Pierre-Olivier Amblard)

Introduction

Space-filling designs, e.g. Latin hypercubes (McKay et al., 1979; Owen, 1992), low discrepancy sequences (e.g. Halton, 1964; Sobol, 1967), are popular methods in computer experiments. These computational methods are becoming unavoidable to simulate complex phenomena (e.g. Santner et al., 2013, Chapter 5). A space-filling design corresponds to a set $\mathbf{x} = \{x^{(1)}, \dots, x^{(n)}\}$ of n points generated in a bounded domain, for instance $[0,1]^d$ in the following. Usually, the dimension d represents the number of factors (or covariates) on which the numerical code depends. The k th coordinates of points from \mathbf{x} then represent the values of the k th factor. Intuitively, points issued from a space filling-design tend to regularly cover the domain $[0,1]^d$. The quality of this coverage can be a priori evaluated by standard criteria such as maximin distance or L^2 discrepancy (see e.g. Owen, 2013).

Frequently in computer experiments, some factors are a posteriori found to be inactive (see Sun et al., 2019, and references therein). If the experiment is to be performed again, an inactive factor must be discarded to avoid numerical errors and to decrease complexity. But if k factors are discarded, the experimental space-filling design should be done again, this time on $[0,1]^{d-k}$. This induces a new complexity and is expensive. A cheaper strategy is to keep the first space-filling design, but use its projection onto $[0,1]^{d-k}$ by discarding the adequate factors (or coordinates). However, the projected points should provide a good coverage of $[0,1]^{d-k}$. Therefore, an additional property of the initial space-filling design should be the conservation of the “nice coverage” property for any subsets of the coordinates.

This additional property has already been considered in the literature for low discrepancy type designs (see e.g. Sun et al., 2019). Our work in contrast considers spatial point processes as experimental designs. For the question we address, we set $\mathbf{x}_I = \{x_I^{(1)}, \dots, x_I^{(n)}\}$ to be the design obtained from the design \mathbf{x} by keeping the factors (coordinates) indexed by the index set $I \subseteq \{1, \dots, d\}$. For example, when $I = \{1, \dots, d-1\}$, \mathbf{x}_I corresponds to the set \mathbf{x} where the d th coordinate of each point is discarded. We let \mathbf{X} to be the spatial point process generating \mathbf{x} and \mathbf{X}_I the process generating \mathbf{x}_I . In the spatial statistics literature, the pair correlation function (denoted by g) is the most standard way for characterizing the pairwise dependence between points, see e.g. Møller and Waagepetersen (2004): $g_{\mathbf{x}}(x, y)$ measures the probability to observe a pair of distinct points at (x, y) , normalized by the same probability under the Poisson case, i.e. under the situation where there

is no interaction between points (see Section 1 for a more formal definition of the pair correlation function). Møller and Waagepetersen (2004); Illian et al. (2008) qualify as regular or repulsive a point process for which $g_{\mathbf{X}} < 1$, i.e. $g_{\mathbf{X}}(x, y) < 1$ for all $x, y \in [0, 1]^d$. Thanks to repulsiveness, points of a repulsive point process tend to cover more regularly the space than a Poisson point process does. For the application to computer experiments which motivates our study, we intend to develop point process models which are repulsive in all directions, i.e. spatial point processes \mathbf{X} such that $g_{\mathbf{X}_I} < 1$ for all $I \subseteq \{1, \dots, d\}$.

Several classes of spatial point processes are able to generate regular patterns. Among them, Matérn hard-core processes (Teichmann et al., 2013), Gibbs point processes (Møller and Waagepetersen, 2004; Dereudre, 2019) or determinantal point processes (Lavancier et al., 2015) are appealing for many applications. In particular Gibbs point processes have been considered to generate space-filling designs by Franco et al. (2008); Dupuy et al. (2015). The authors build a specific Gibbs model by parameterizing its Papangelou conditional intensity as a Strauss hard-core model with constraints on the marginals. The resulting patterns look regular and the points cover regularly the space. However, Gibbs point processes have the drawback of not having their moments available in a closed form. In particular, the intensity as well as the pair correlation function (see e.g. Dereudre, 2019) are not available analytically. Even worse, using Monte-Carlo simulations, Illian et al. (2008) shows that the pair correlation function of Strauss hard-core models are not uniformly bounded by 1. Although Matérn hard-core processes are more tractable, their pair correlation function suffer from the same problem (see Teichmann et al., 2013).

Determinantal point processes (DPPs for short) have been introduced by Macchi (1975) as “fermion” processes to model the position of particles that repel each other. This class of processes is known for very appealing properties, in particular for its tractability: explicit expressions for the intensity functions are available. Therefore, a growing attention has been paid to DPPs from a theoretical point of view (e.g. Soshnikov, 2000; Shirai and Takahashi, 2003; Hough et al., 2009; Decreusefond et al., 2016), and more recently in the statistics community (Lavancier et al., 2015; Bardenet and Hardy, 2020). In particular, one of the main characteristics of a DPP is that, by construction, its pair correlation function is uniformly bounded by 1. DPPs are defined through a kernel $K : B \times B \rightarrow \mathbb{C}$ which characterizes the distribution of \mathbf{X} and thus which characterizes also its moments. The main result of this paper

concerns necessary conditions (expressed by Assumption ($H[I]$)) on the form of kernel K to ensure that the projected pattern \mathbf{X}_I remains repulsive, i.e. such that $g_{\mathbf{X}_I} < 1$.

The paper is organized as follows. Section 1 contains a brief background on spatial point processes and in particular on DPPs. Section 2 deals with the statistical description of the projected point process \mathbf{X}_I . In particular we provide a closed form for the pair correlation $g_{\mathbf{X}_I}$ when \mathbf{X} is a DPP defined on $[0,1]^d$ with kernel K satisfying a separability assumption. Examples of models satisfying this separability condition are presented and discussed in Section 3. They are compared using an original summary statistic, defined as a normalized version of Ripley's function (see e.g. Møller and Waagepetersen, 2004) based on the sup norm. We illustrate in Section 4 the interest of the models developed in this research. To mimic situations which occur in computer experiments, we consider the Monte-Carlo integration for $\int_{[0,1]^\iota} f_I(u)du$, for any function $f_I : [0,1]^\iota \rightarrow \mathbb{R}$ and any $I \subseteq \{1, \dots, d\}$ with cardinality $\iota = 1, \dots, d$. We demonstrate that the single initial design defined on $[0,1]^d$ and its projections can be used to achieve this task efficiently. Proofs of our results are postponed to appendices.

1. Background and notation

1.1. Spatial point processes

A spatial point process \mathbf{X} defined on a Borel set $B \subseteq \mathbb{R}^d$ is a locally finite measure on B , (for measure theoretical details, see e.g. Møller and Waagepetersen (2004) and references therein) whose realization is of the form $\{x^{(1)}, \dots, x^{(k)}\} \in B^k$ where k is the realization of a random variable and the $x^{(i)}$'s represent the events. We assume that \mathbf{X} is simple meaning that two events cannot occur at the same location. Thus, \mathbf{X} is viewed as a locally finite random set.

In most cases, the distribution of a point process \mathbf{X} can be described by its intensity functions $\rho_{\mathbf{X}}^{(k)} : B^k \rightarrow \mathbb{R}^+$, $k \in \mathbb{N} \setminus \{0\}$. By Campbell Theorem (e.g. Møller and Waagepetersen, 2004), $\rho_{\mathbf{X}}^{(k)}$ is characterized by the following integral representation: for any non-negative measurable function

$h : B^k \rightarrow \mathbb{R}^+$

$$\begin{aligned} & \mathbb{E} \left[\sum_{x^{(1)}, \dots, x^{(k)} \in \mathbf{X}}^{\neq} h(x^{(1)}, \dots, x^{(k)}) \right] \\ &= \int_{B^k} \rho_{\mathbf{X}}^{(k)}(x^{(1)}, \dots, x^{(k)}) h(x^{(1)}, \dots, x^{(k)}) dx^{(1)} \dots dx^{(k)} \end{aligned} \quad (1)$$

where \neq over the summation means that $x^{(1)}, \dots, x^{(k)}$ are pairwise distinct points. Intuitively, for any pairwise distinct points $x^{(1)}, \dots, x^{(k)} \in B$, $\rho_{\mathbf{X}}^{(k)}(x^{(1)}, \dots, x^{(k)}) dx^{(1)} \dots dx^{(k)}$ is the probability that \mathbf{X} has a point in each of the k infinitesimally small sets around $x^{(1)}, \dots, x^{(k)}$ with volumes $dx^{(1)}, \dots, dx^{(k)}$, respectively. When $k = 1$, this yields the intensity function and we simply denote it by $\rho_{\mathbf{X}} = \rho_{\mathbf{X}}^{(1)}$. The second order intensity $\rho_{\mathbf{X}}^{(2)}$ is used to define the pair correlation function

$$g_{\mathbf{X}}(x^{(1)}, x^{(2)}) = \frac{\rho_{\mathbf{X}}^{(2)}(x^{(1)}, x^{(2)})}{\rho_{\mathbf{X}}(x^{(1)})\rho_{\mathbf{X}}(x^{(2)})} \quad (2)$$

for pairwise distinct $x^{(1)}, x^{(2)} \in B$, and where $g_{\mathbf{X}}(x^{(1)}, x^{(2)})$ is set to 0 if $\rho_{\mathbf{X}}(x^{(1)})$ or $\rho_{\mathbf{X}}(x^{(2)})$ is zero. By convention, $\rho_{\mathbf{X}}^{(k)}(x^{(1)}, \dots, x^{(k)})$ is set to 0 if $x^{(i)} = x^{(j)}$ for some $i \neq j$. Therefore $g_{\mathbf{X}}(x, x)$ is also set to 0 for all $x \in B$. The pair correlation function (pcf for short) can be used to determine the local interaction between points of \mathbf{X} located at x and y : $g_{\mathbf{X}}(x, y) > 1$ characterizes positive correlation between the points; $g_{\mathbf{X}}(x, y) = 1$ means there is no interaction (typically a Poisson point process); $g_{\mathbf{X}}(x, y) < 1$ characterizes negative correlations. A point pattern is often referred to as a repulsive point process, if $g(x, y) < 1$ for any $x, y \in B$ (e.g. [Illian et al., 2008](#), Sec. 6.5).

A point process \mathbf{X} with constant intensity function on B is said to be homogeneous. A pcf with constant intensity is said to be invariant by translation (resp. isotropic) if $\rho_{\mathbf{X}}^{(2)}(x^{(1)}, x^{(2)})$ depends only on $x^{(2)} - x^{(1)}$ (resp. on $\|x^{(2)} - x^{(1)}\|$ for a norm to be defined).

1.2. Determinantal point processes (DPPs)

In this section, the class of continuous DPPs is introduced. We restrict our attention to DPPs defined on a compact set $B \subset \mathbb{R}^d$. A point process \mathbf{X} on B is said to be a DPP on B with kernel $K : B \times B \rightarrow \mathbb{C}$ if for any $k \geq 1$ its k th order intensity function is given by

$$\rho_{\mathbf{X}}^{(k)}(x^{(1)}, \dots, x^{(k)}) = \det [K(x^{(i)}, x^{(j)})]_{i,j=1}^k \quad (3)$$

and we simply denote by $\mathbf{X} \sim \text{DPP}_B(K)$. We assume in this work that K is a continuous covariance function and refer the interested reader to more general situations to [Hough et al. \(2009\)](#). The intensity of \mathbf{X} is given by $\rho_{\mathbf{X}}(x) = K(x, x)$ and its pcf by

$$g_{\mathbf{X}}(x, y) = 1 - \frac{|K(x, y)|^2}{K(x, x)K(y, y)}. \quad (4)$$

The popularity of DPPs relies mainly upon (3)-(4): all moments of \mathbf{X} are explicit and by assumption on K , $g_{\mathbf{X}}(x, y) < 1$ for any $x, y \in B$. From (4) and the continuity of K , it is worth mentioning that $g_{\mathbf{X}}$ is continuous on the diagonal, i.e. $g_{\mathbf{X}}(x, y) \rightarrow 0$ when $y \rightarrow x$ for any $x \in B$.

From Mercer's Theorem ([Riesz and Sz-Nagy, 1990](#), Sec. 98), kernel K admits the following decomposition for any $x, y \in B$

$$K(x, y) = \sum_{j \in \mathcal{N}} \lambda_j \phi_j(x) \overline{\phi_j(y)} \quad (5)$$

where \mathcal{N} is a countable set (e.g. \mathbb{N} , \mathbb{Z} , \mathbb{Z}^d, \dots), $\{\phi_j\}_{j \in \mathcal{N}}$ are eigenfunctions associated to K and form an orthonormal basis of the space of square-integrable functions $L^2(B)$. $\{\lambda_j\}_{j \in \mathcal{N}}$ are the eigenvalues of K . Let us mention that we abuse notation when referring ϕ_j 's and λ_j 's to as eigenfunctions and eigenvalues of K . These should require to introduce the notion of integral operator with kernel K (e.g. [Debnath and Mikusiński, 2005](#)) acting on $L^2(B)$. To simplify the reading, we make the misnomer to consider kernel K instead of the associated integral operator. We define the trace of kernel K on B by

$$\text{tr}_B(K) = \int_B K(x, x) dx = \sum_{j \in \mathcal{N}} \lambda_j.$$

In the following, the kernels we consider are assumed to have finite trace, and are called trace class kernels. The existence of a DPP with kernel K is ensured if K is trace class, and is such that $\lambda_j \leq 1$ for any $j \in \mathcal{N}$ (e.g. [Hough et al., 2009](#), Theorem 4.5.5).

A kernel such that its non-zero eigenvalues are equal to 1 is called a “projection kernel”. In particular, if \mathbf{X} is a “projection DPP”, i.e. $\mathbf{X} \sim \text{DPP}_B(K)$ where K is a “projection kernel”, then the number of points of \mathbf{X} in B , is almost surely constant and equal to the trace of K . Notice that the name “projection kernel” is not related at all with the projection

transformation we are studying here. This terminology seems commonly used though (e.g. [Hough et al., 2006, 2009](#); [McCullagh and Møller, 2006](#); [Lavancier et al., 2015](#)).

The homogeneous case is often considered later. A DPP \mathbf{X} with kernel K is said to be homogeneous, if K is the restriction on $B \times B$ of a kernel \bar{K} defined on $\mathbb{R}^d \times \mathbb{R}^d$ which is stationary, i.e. satisfies

$$\bar{K}(x, y) = \bar{K}(0, x - y), \quad x, y \in \mathbb{R}^d.$$

In that case, we abuse notation, identify K with \bar{K} and refer to K as a stationary kernel. It is worth pointing out that if K admits a Mercer's decomposition with respect to the Fourier basis

$$\phi_j(x) = e^{2i\pi\langle j, x \rangle} \tag{6}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product on \mathbb{R}^d , then K is stationary.

2. Projection of a spatial point process and applications to DPPs

2.1. Projection of a spatial point process

In this work, we consider projection of spatial point processes, i.e. keeping a given number of coordinates from the original spatial point process. Such a framework requires that the original point process \mathbf{X} is defined on a compact set $B \subset \mathbb{R}^d$: otherwise, the configuration of points of the projected point processes may not form locally finite configuration, as also noticed in the two-dimensional case by [Baddeley \(2007, p. 17\)](#).

This section presents a few notation and characterization of projected spatial point processes. Let I be a subset of $\bar{d} = \{1, \dots, d\}$ with cardinality $|I| = \iota$. In the following, we let $B \subset \mathbb{R}^d$ be a compact set, which can be written as $B_1 \times \dots \times B_d$. We denote by B_I the set $B_I = \prod_{i \in I} B_i$ with $B = B_{\bar{d}}$ and by P_I the orthogonal projection of \mathbb{R}^d onto \mathbb{R}^ι . For any point process \mathbf{X} defined on such a compact $B \subset \mathbb{R}^d$, the projected point process $\mathbf{X}_I = P_I \mathbf{X}$ is then defined on B_I . For any $x \in B$, we often use the notation x_I to denote $P_I x$. We sometimes use the notation $\mathbf{X}_{\bar{d}} = \mathbf{X}$ when $I = \bar{d}$. The following Lemma provides a general way to evaluate intensity functions of \mathbf{X}_I .

Lemma 2.1. *Let $I \subset \bar{d}$ and let \mathbf{X} be a spatial point process defined on a compact set $B \subset \mathbb{R}^d$. Then, for any $k \geq 1$ such that $\rho_{\mathbf{X}}^{(k)}$ exists, $\rho_{\mathbf{X}_I}^{(k)}$ is*

well-defined and

$$\begin{aligned} \rho_{\mathbf{X}_I}^{(k)}(x^{(1)}, \dots, x^{(k)}) \\ = \int_{(B_{I^c})^k} \rho_{\mathbf{X}}^{(k)}((x^{(1)}, u^{(1)}), \dots, (x^{(k)}, u^{(k)})) \, du^{(1)} \dots du^{(k)} \end{aligned} \quad (7)$$

for any pairwise distinct $x^{(1)}, \dots, x^{(k)} \in B_I$ where $I^c = \bar{d} \setminus I$.

Lemma 2.1 is obtained by a simple application of Campbell's Theorem. Its proof is provided in Appendix A for the sake of completeness. We now turn to the core of this paper which is the study of projected determinantal point processes.

2.2. Distribution of \mathbf{X}_I when $\mathbf{X} \sim \text{DPP}_B(K)$

According to (7), the k th order intensity function of the projected point process \mathbf{X}_I is given by

$$\begin{aligned} \rho_{\mathbf{X}_I}^{(k)}(x^{(1)}, \dots, x^{(k)}) &= \int_{(B_{I^c})^k} \rho_{\mathbf{X}}^{(k)}((x, u)^{(1)}, \dots, (x, u)^{(k)}) \, du^{(1)} \dots du^{(k)} \\ &= \int_{(B_{I^c})^k} \det [K((x, u)^{(i)}, (x, u)^{(j)})]_{i,j=1}^k \, du^{(1)} \dots du^{(k)} \\ &= \sum_{\sigma \in S_k} (-1)^{k-C(\sigma)} \int_{(B_{I^c})^k} \prod_{i=1}^k K((x, u)^{(i)}, (x, u)^{(\sigma(i))}) \\ &\quad \, du^{(1)} \dots du^{(k)} \end{aligned} \quad (8)$$

where S_k is the symmetric group on $\bar{k} = \{1, \dots, k\}$, $C(\sigma)$ is the number of disjoint cycles of σ , and $(x, y)^{(i)}$ denotes $(x^{(i)}, y^{(i)})$. Without any assumption on kernel K , there is no chance to reduce (8) any further, i.e. to have an explicit form for $\rho_{\mathbf{X}_I}^{(k)}$ and thus $g_{\mathbf{X}_I}$. Therefore, without additional assumption, it is difficult to assess whether $g_{\mathbf{X}_I}$ is smaller than 1 or not. The following assumptions will allow us to solve this problem.

Assumption ($H[I]$) For $I \subseteq \bar{d}$, the kernel K can be written as

$$K(x, y) = K_I(x_I, y_I) K_{I^c}(x_{I^c}, y_{I^c}) \quad (H[I])$$

where $K_I : B_I \times B_I \rightarrow \mathbb{C}$ and $K_{I^c} : B_{I^c} \times B_{I^c} \rightarrow \mathbb{C}$ are two continuous covariance functions.

Assumption $(H[I])$ implies that K admits the Mercer's decomposition $K(x, y) = \sum_{j \in \mathcal{N}} \lambda_j \phi_j(x) \bar{\phi}_j(y)$, where $\mathcal{N} = \mathcal{N}_I \times \mathcal{N}_{I^c}$, $\lambda_j = \lambda_{j_I}^{(I)} \lambda_{j_{I^c}}^{(I^c)}$, $\phi_j(x) = \phi_{j_I}^{(I)}(x_I) \phi_{j_{I^c}}^{(I^c)}(x_{I^c})$ for $j = (j_I, j_{I^c})$, $x = (x_I, x_{I^c})$. Here, for $\bullet = I, I^c$, $\{\phi_{j_\bullet}^{(\bullet)}\}_{j_\bullet \in \mathcal{N}_\bullet}$ is a set of normalized eigenfunctions of K_\bullet , (and thus an orthonormal basis of $L^2(B_\bullet)$) and λ_{j_\bullet} denote the eigenvalues of K_\bullet .

If K admits a Mercer's decomposition with respect to the Fourier basis such that its eigenvalues satisfy the above separability property, then $(H[I])$ is satisfied. Hence, Fourier basis appears as a natural basis and leads us to consider, the following natural extension of $(H[I])$ that would be assumed for any $I \subseteq \bar{d}$.

Assumption (H') We assume that the kernel K satisfies $(H[I])$ for any $I \subseteq \bar{d}$, is stationary and can be written as the product of d one-dimensional stationary kernels:

$$K(x - y) = \prod_{i=1}^d K_i(x_i - y_i), \quad x, y \in B \quad (H')$$

where for any $i \in \bar{d}$, each $K_i : B_i \times B_i \rightarrow \mathbb{C}$ is a stationary continuous kernel.

Assumption (H'') We will also focus on the particular case where all kernels are identical, i.e. $K_i \equiv K_0$ for all $i \in \bar{d}$:

$$K(x - y) = \prod_{i=1}^d K_0(x_i - y_i), \quad x, y \in B. \quad (H'')$$

Assumption (H''') is well-suited to the situation where we have no a priori information on the projection P_I from the initial point process \mathbf{X} we want to study.

We could remove the stationarity assumption in Assumption (H') . However, as revealed by Sections 3 and 4, stationarity allows us to plot pcfs or Ripley's functions of \mathbf{X}_I for any I . It thus provides a visual interpretation of regularity properties for \mathbf{X}_I . Furthermore, going back to one motivation of this paper, there is a priori no reason to construct a design which favours particular spatial areas. Thus, considering a stationary kernel which ensures that the intensity is constant makes sense.

Theorem 2.1. *Let $I \subseteq \bar{d}$ and $\mathbf{X} \sim \text{DPP}_B(K)$ such that K satisfies $(H[I])$. Then the k th order intensity function of the projected point process \mathbf{X}_I is given by*

$$\rho_{\mathbf{X}_I}^{(k)}(x^{(1)}, \dots, x^{(k)}) = \sum_{\sigma \in \mathcal{S}_k} (-1)^{k-C(\sigma)} \left[\prod_{i=1}^k K_I(x^{(i)}, x^{(\sigma(i))}) \right] \text{tr}_{B_{I^c}}(K_{I^c})^{k-c(\sigma)} \prod_{\varepsilon \in \mathcal{S}(\sigma)} \text{tr}_{B_{I^c}}(K_{I^c}^{(c(\varepsilon))}) \quad (9)$$

where $c(\sigma)$ is the size of the support $\text{supp}(\sigma) = \{i \in \bar{k} \text{ s.t. } \sigma(i) \neq i\}$, $\mathcal{S}(\sigma)$ is the set of disjoint cycles of σ with non-empty support, $C(\sigma)$ is the number of disjoint cycles of σ (including those with empty support) and for a kernel K , $K^{(m)}$ for $m > 1$, stands for the iterated kernel defined by $K^{(m)}(x, y) = \int K^{(m-1)}(x, z)K(z, y)dz$ (with $K^{(1)} = K$). In particular, the intensity of \mathbf{X}_I is given by $\rho_{\mathbf{X}_I}(x) = K_I(x, x)\text{tr}_{B_{I^c}}(K_{I^c})$ and its pcf is given by

$$g_{\mathbf{X}_I}(x, y) = 1 - \frac{\text{tr}_{B_{I^c}}(K_{I^c}^{(2)})}{\text{tr}_{B_{I^c}}(K_{I^c})^2} (1 - g_{\mathbf{Y}^{(I)}}(x, y)) \quad (10)$$

for any pairwise distinct $x, y \in B_I$ and where $\mathbf{Y}^{(I)} \sim \text{DPP}_{B_I}(K_I)$.

We focus in Theorem 2.1 on intensity functions. However, we can prove a full characterization of the distribution of \mathbf{X}_I via its Laplace functional. This is detailed in Appendix C. In particular, Theorem Appendix C.1 shows that \mathbf{X}_I is distributed as an infinite superposition of independent DPPs, each with kernel $\lambda_l^{(I^c)} K_I$. In particular, if K_{I^c} is a projection kernel, \mathbf{X}_I is a finite superposition of $M = \text{tr}_{B_{I^c}}(K_{I^c})$ i.i.d. DPPs with kernel K_I . Such finite superposition corresponds (see Hough et al., 2006; Decreusefond et al., 2016) to the distribution of an α -DPP on B_I with kernel $-\alpha^{-1}K_I$ where $\alpha = -M^{-1}$. α -DPPs (and α -determinants) are introduced by Shirai and Takahashi (2003). And it can indeed be checked from (9) that when K_{I^c} is a projection kernel, the intensity functions of \mathbf{X}_I are given by

$$\begin{aligned} \rho_{\mathbf{X}_I}^{(k)}(x^{(1)}, \dots, x^{(k)}) &= \sum_{\sigma \in \mathcal{S}_k} \alpha^{k-C(\sigma)} \left[\prod_{i=1}^k -\alpha^{-1}K_I(x^{(i)}, x^{(\sigma(i))}) \right] \\ &:= \det_{\alpha} \left[-\alpha^{-1}K_I(x^{(i)}, x^{(j)}) \right]_{i,j=1}^k \end{aligned}$$

Equation (10) is in our opinion the most interesting result of this paper. It reveals the repulsiveness nature of \mathbf{X}_I . Let us examine this in details. Since $\mathbf{Y}^{(I)}$ is a DPP with kernel K_I , it satisfies $0 \leq g_{\mathbf{Y}^{(I)}} \leq 1$, which allows us to rewrite (10) as

$$0 \leq 1 - g_{\mathbf{X}_I}(x, y) = \frac{\text{tr}_{B_{I^c}}(K_{I^c}^{(2)})}{\text{tr}_{B_{I^c}}(K_{I^c})^2} (1 - g_{\mathbf{Y}^{(I)}}(x, y)) \leq \frac{\text{tr}_{B_{I^c}}(K_{I^c}^{(2)})}{\text{tr}_{B_{I^c}}(K_{I^c})^2}. \quad (11)$$

The lower-bound of (11) means that $g_{\mathbf{X}_I} \leq 1$, i.e. \mathbf{X}_I is indeed a repulsive point process on B_I . Furthermore, the upper-bound measures in some sense the loss of repulsion and more precisely, how $g_{\mathbf{X}_I}$ gets closer to 1 which corresponds the pcf of a Poisson point process. To be more precise, let us focus on the particular case (H'). We have in this situation

$$g_{\mathbf{X}_I}(x, y) \geq 1 - \prod_{i \in I^c} \frac{\text{tr}_{B_i}(K_i^{(2)})}{\text{tr}_{B_i}(K_i)^2}.$$

For each $i \in \bar{d}$, $\text{tr}_{B_i}(K_i^{(2)})/\text{tr}_{B_i}(K_i)^2 < 1$. Therefore, when $|I^c| = d - \iota$ is large, $1 - g_{\mathbf{X}_I}$ is bounded by a product of large number of quantities smaller than 1, and thus the pcf of \mathbf{X}_I gets closer and closer to the pcf of a Poisson point process. It is even more obvious when K satisfies (H''). In that case, for any $x, y \in B_I$

$$g_{\mathbf{X}_I}(x, y) \geq 1 - \kappa_0^{d-\iota} \quad \text{where} \quad \kappa_0 = \frac{\text{tr}_{B_0}(K_0^{(2)})}{\text{tr}_{B_0}(K_0)^2}.$$

For example, when $\iota = d - 1$, i.e. when one skips only one coordinate: $g_{\mathbf{X}_I}(x, y) \geq 1 - \kappa_0 > 0$ and this constant is reached when $y \rightarrow x$. Since, $g_{\mathbf{X}}(x, y) \rightarrow 0$ when $y \rightarrow x$, one can clearly measure the loss of repulsion as soon as one skips one coordinate.

3. Examples

In this section, we present particular examples of kernels defined on $B = [0, 1]^d$ and satisfying (H'), thus ensuring that \mathbf{X}_I is repulsive for any $I \subseteq \bar{d}$. Then, these examples are compared for different sets I through their pair correlation function or through a normalized Ripley's function.

All our kernels examples have Mercer’s decomposition defined with respect to the Fourier basis (6), the natural basis which allows (H') to be satisfied.

3.1. Gaussian kernel

The Gaussian kernel (see e.g. Lavancier et al. (2015))

$$K(x, y) = \rho \exp \left(- \left\| \frac{x - y}{\alpha} \right\|^2 \right)$$

where $\|\cdot\|$ denotes the Euclidean norm, is the typical example satisfying (H''), where K_0 is defined for any $x, y \in [0, 1]$ by:

$$K_0(x - y) = \rho^{1/d} \exp \left(- \left(\frac{x - y}{\alpha} \right)^2 \right).$$

The existence of $\mathbf{X} \sim \text{DPP}_B(K)$ is ensured if α is such that $\rho(\alpha\sqrt{\pi})^d \leq 1$. For any $I \subseteq \bar{d}$, the pcf of \mathbf{X}_I is derived from Theorem 2.1: for any pairwise distinct $x, y \in B_I$

$$g_{\mathbf{X}_I}(x, y) = 1 - \kappa_2^{d-\iota} \exp \left(-2 \left\| \frac{x - y}{\alpha} \right\|^2 \right) \quad (12)$$

with

$$\kappa_2 = \frac{\text{tr}_{B_0} \left(K_0^{(2)} \right)}{\text{tr}_{B_0} \left(K_0 \right)^2} \approx \frac{\sum_{j \in \mathbb{Z}} \exp \left(-2(j\alpha\pi)^2 \right)}{\left(\sum_{j \in \mathbb{Z}} \exp \left(-(j\alpha\pi)^2 \right) \right)^2}. \quad (13)$$

This approximation comes from the Fourier approximation of kernel K detailed in Lavancier et al. (2015, Section 4). Note that for all $I \subseteq \bar{d}$ and $x, y \in B_I$, we use with a slight abuse the same notation $\|x - y\|$ for the Euclidean norm in \mathbb{R}^ι .

This class of examples is of particular interest due to the isotropy property of $g_{\mathbf{X}_I}$. The pcfs $g_{\mathbf{X}_I}$ for different sets I can be represented on the same plot. For $d = 10, 10^2, 10^3, 10^4$, Figure 1 represents the pcfs of a Gaussian DPP \mathbf{X} (solid lines) and its successive projections. The intensity parameter and α are set to $\rho_{\mathbf{X}} = 500$ and $\alpha^{-1} = \rho_{\mathbf{X}}^{1/d} \sqrt{\pi}$. Note that the abscissa corresponds to $\|x - y\|$ for $x, y \in B_I$ for different sets I . Thus the differences should be understood carefully. Figure 1 confirms that the pcf of \mathbf{X}_I is upper-bounded by 1, lower-bounded by $1 - \kappa_2^{d-\iota}$ and gets closer to 1 when ι decreases.

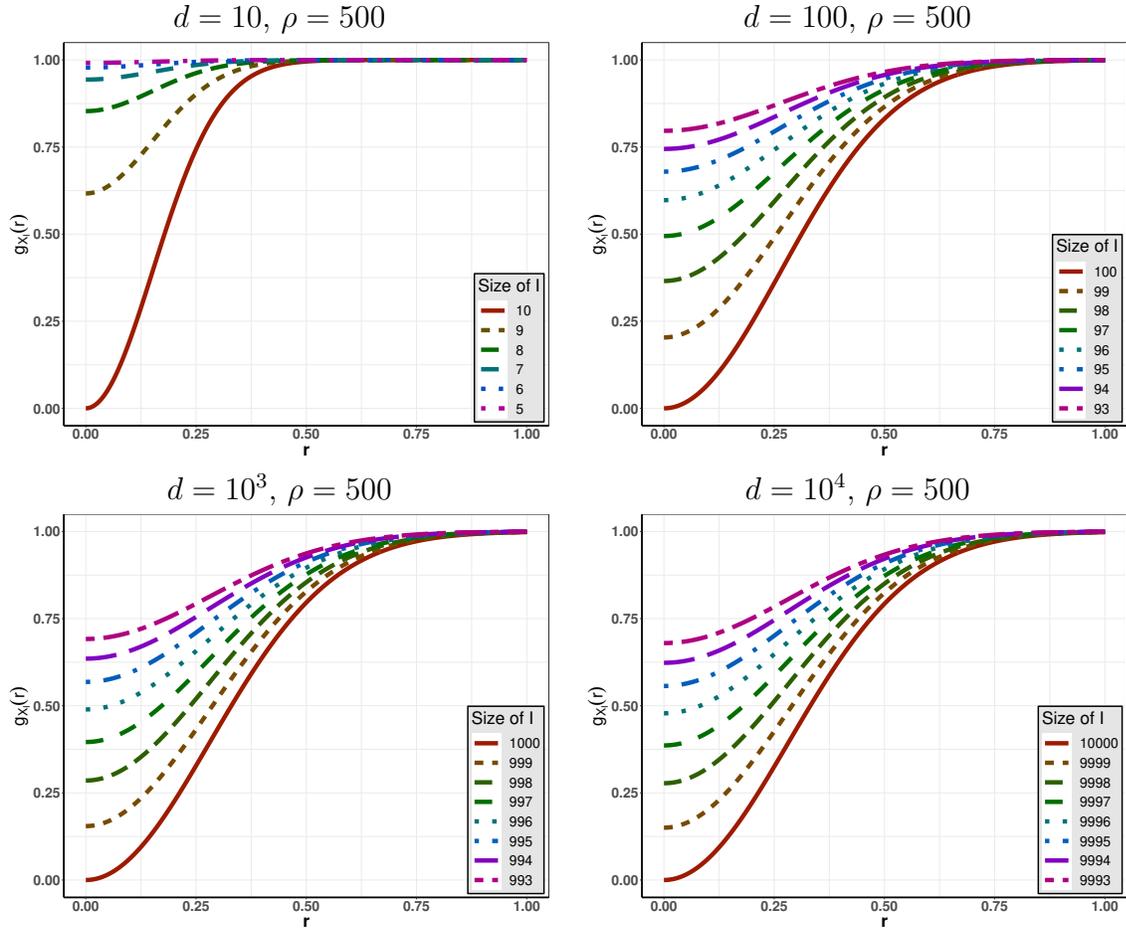


Figure 1: Pair correlation functions of the Gaussian DPP \mathbf{X} (solid lines) with intensity $\rho_{\mathbf{X}} = 500$ and $\alpha^{-1} = \rho_{\mathbf{X}}^{1/d} \sqrt{\pi}$ and its successive projections \mathbf{X}_I ($|I| = d - 1, \dots, 1$; dotted and dashed lines) for $d = 10$ (top left), 10^2 (top right), 10^3 (bottom left) and 10^4 (bottom right).

3.2. L^1 -Exponential kernel

We now consider an exponential kernel, defined with respect to the L^1 -norm instead of the Euclidean norm:

$$K(x - y) = \rho \exp \left(- \left\| \frac{x - y}{\alpha} \right\|_1 \right). \quad (14)$$

The kernel (14) is referred to as the L^1 -Exponential kernel in the following. It also constitutes a natural example as it satisfies (H'') where K_0 is defined for any $x, y \in [0, 1]$ by:

$$K_0(x - y) = \rho^{1/d} \exp \left(- \left| \frac{x - y}{\alpha} \right| \right).$$

The existence of $\mathbf{X} \sim \text{DPP}_B(K)$ is ensured if α is such that $\rho(2\alpha)^d \leq 1$. According to Theorem 2.1, for any $I \subseteq \bar{d}$, the pcf of \mathbf{X}_I is given for any pairwise distinct $x, y \in B_I$ by

$$g_{\mathbf{X}_I}(x, y) = 1 - \kappa_1^{d-\iota} \exp \left(-2 \left\| \frac{x - y}{\alpha} \right\|_1 \right) \quad (15)$$

with

$$\kappa_1 = \frac{\text{tr}_{B_0} \left(K_0^{(2)} \right)}{\text{tr}_{B_0} \left(K_0 \right)^2} \approx \frac{\sum_{j \in \mathbb{Z}} (1 + (2\pi\alpha j)^2)^{-2}}{\left(\sum_{j \in \mathbb{Z}} (1 + (2\pi\alpha j)^2)^{-1} \right)^2} \quad (16)$$

where the approximation corresponds again to the Fourier approximation. For $d = 10, 10^2, 10^3, 10^4$, Figure 2 represents the pcfs of an L^1 -Exponential DPP \mathbf{X} (solid lines) and its successive projections with respect to the L^1 -norm. The intensity parameter and α are set to $\rho_{\mathbf{X}} = 500$ and $\alpha^{-1} = 2\rho_{\mathbf{X}}^{1/d}$. The conclusion drawn from Figure 2 is similar to the one from Figure 1: the pcf of \mathbf{X}_I is upper-bounded by 1, lower-bounded by $1 - \kappa_1^{d-\iota}$ and tends to 1 when ι decreases. We could be tempted to compare Figures 1 and 2 and conclude that the Gaussian DPP seems more repulsive. However, remember that both models are not isotropic with respect to the same norm. We provide in Section 3.4 a summary statistic which allows us to correctly compare these models.

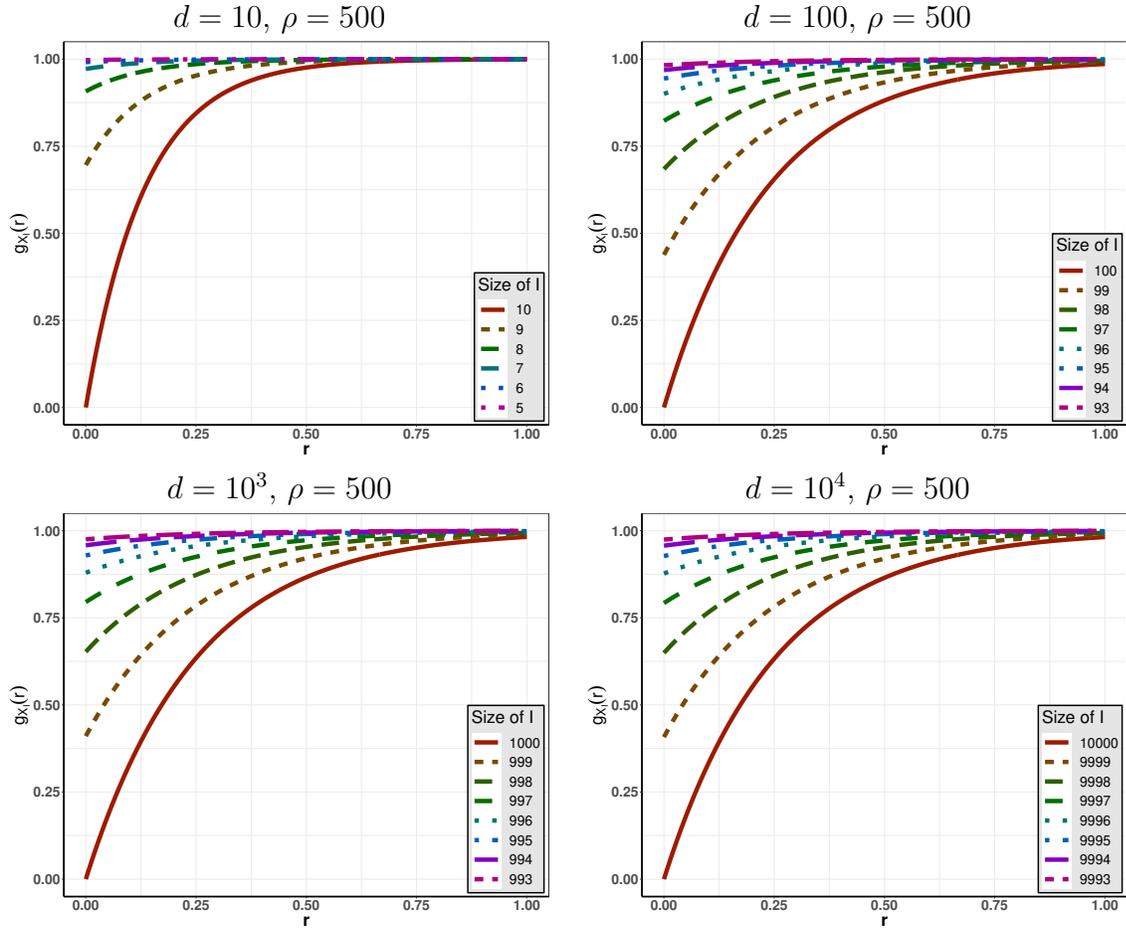


Figure 2: Pair correlation functions of the L^1 -Exponential DPP \mathbf{X} (solid lines) with intensity $\rho_{\mathbf{X}} = 500$ and $\alpha^{-1} = 2\rho_{\mathbf{X}}^{1/d}$ and its successive projections \mathbf{X}_I ($|I| = d - 1, \dots$; dotted and dashed lines) for $d = 10$ (top left), 10^2 (top right), 10^3 (bottom left) and 10^4 (bottom right).

3.3. Dirichlet kernels

The two examples considered so far satisfy (H'') by definition. The next one is a projection kernel which only satisfies (H') . For $\iota = 1, \dots, d$, we let $\{\phi_j^{(\iota)}\}_{j \in \mathbb{Z}^\iota}$ denote the ι -dimensional Fourier basis. We consider d positive integers $(n_i)_{i \in \bar{d}}$ and for $i \in \bar{d}$ the following one-dimensional stationary kernel:

$$K_i(x - y) = \sum_{j \in E_i} \phi_j^{(1)}(x - y), \quad x, y \in [0, 1],$$

where $E_i = \{a_i, a_i + 1, \dots, n_i - 1 + a_i\}$ is a set of n_i consecutive integers and $a_i \in \mathbb{Z}$. Then, we construct a kernel K as

$$K(x - y) = \prod_{i=1}^d K_i(x_i - y_i) = \sum_{j \in E_N} \phi_j^{(d)}(x - y), \quad x, y \in [0, 1]^d$$

where $E_N = \prod_i E_i$. It is worth pointing out that kernel K can be written as

$$K(x - y) = \prod_{i=1}^d \left(\sum_{j=a_i}^{n_i-1+a_i} \phi_j^{(1)}(x_i - y_i) \right) = \phi_a^{(d)}(y - x) \prod_{i=1}^d \left(\sum_{j=0}^{n_i-1} \phi_j^{(1)}(x_i - y_i) \right) \quad (17)$$

where $a = (a_i)_{i \in \bar{d}}$. Therefore, according to Remark (4) from [Hough et al. \(2009, p. 48\)](#), the choice of the E_i 's does not influence the distribution of the DPP with kernel K . Remark that, if the n_i 's are all odd numbers and if we choose $a_i = -\lfloor n_i/2 \rfloor$, the kernel K equals

$$K(x - y) = \prod_{i=1}^d D_{\lfloor \frac{n_i}{2} \rfloor}(x_i - y_i) \quad (18)$$

where D_p is the Dirichlet kernel (e.g. [Zygmund, 2003](#)) with parameter p . That terminology justifies the name Dirichlet kernel for this model. In the general case, and unambiguously we set $a_i = 0$ for any i and thus consider $E_N = \{j \in \mathbb{N}^d : j_i < n_i, i = 1 \dots d\}$

$$K(x - y) = \sum_{j \in E_N} e^{2i\pi \langle j, x - y \rangle}. \quad (19)$$

A DPP on B with kernel given by (19) is referred to as an (N, d) -Dirichlet kernel. From Theorem 2.1, for any $I \subseteq \bar{d}$, the pcf of \mathbf{X}_I is given for any

$x, y \in B_I$ by

$$\begin{aligned} g_{\mathbf{X}_I}(x, y) &= 1 - \frac{1}{N} \sum_{j \in F_{N_I}} \left[\prod_{i \in I} \left(1 - \frac{|j_i|}{n_i} \right) \right] \phi_j^{(\iota)}(x - y) \\ &= 1 - \frac{1}{N} \prod_{i \in I} \sum_{|j| < n_i} \left(1 - \frac{|j|}{n_i} \right) \phi_j^{(1)}(x_i - y_i) \end{aligned} \quad (20)$$

where $F_{N_I} = \{j \in \mathbb{Z}^\iota : |j_i| < n_i, i \in I\}$. The pcf $g_{\mathbf{X}_I}$ is bounded from below by $1 - \prod_{i \in I^c} n_i^{-1}$.

A question remains on the factorization of $N = \prod_{i=1}^d n_i$. We consider the factorization which minimizes the fluctuation of the n_i 's. For instance, when $N = 100$ and $d = 6$, we use $N = 5 \times 5 \times 2 \times 2 \times 1 \times 1$ while for $N = 800$ we use the decomposition $N = 5 \times 5 \times 4 \times 2 \times 2 \times 2$.

The next section provides a summary statistics well-suited to the comparison of the three examples we have so far considered.

3.4. Normalized Ripley's function

Since the Gaussian DPP and L^1 -Exponential DPP are isotropic but with respect to a different norm and since the (N, d) -Dirichlet DPP is even not isotropic, it is hard to compare these different examples. In addition to the pcf, a way of characterizing regularity or repulsion in the literature is obtained by analyzing Ripley's function (e.g. [Møller and Waagepetersen, 2004](#)). This function is not adapted for our framework. However, since all models satisfy (H') , we propose to compare them through a normalized version of Ripley's function based on the sup norm $\|\cdot\|_\infty$.

For a stationary spatial point process \mathbf{X} on $B \subseteq \mathbb{R}^d$, we define the normalized d -dimensional Ripley's function for some $r \geq 0$ by

$$R_{\mathbf{X}}(r) = \frac{\mathbb{E}(N_{\mathbf{X}}(B_{d,\infty}(0, r) \setminus 0) \mid 0 \in \mathbf{X})}{\mathbb{E}(N_{\mathbf{\Pi}}(B_{d,\infty}(0, r) \setminus 0) \mid 0 \in \mathbf{\Pi})} \quad (21)$$

where $B_{d,\infty}(0, r) = \{w \in \mathbb{R}^d : |w_i| \leq r, i = 1, \dots, d\}$ is the d -dimensional ball with norm $\|\cdot\|_\infty$ centered at zero with radius r , $\mathbf{\Pi}$ is a homogeneous Poisson point process on B with intensity ρ and $N_{\mathbf{X}}(A)$ (resp. $N_{\mathbf{\Pi}}(A)$) denotes the number of points of \mathbf{X} (resp. $\mathbf{\Pi}$) in a bounded subset $A \subset \mathbb{R}^d$. Assuming that \mathbf{X} has a pcf, say $g_{\mathbf{X}}$, it is known from the properties of the second factorial

moment that

$$R_{\mathbf{X}}(r) = \frac{\int_{B_{d,\infty}(0,r)} g_{\mathbf{X}}(w)dw}{\int_{B_{d,\infty}(0,r)} g_{\Pi}(w)dw} = (2r)^{-d} \int_{B_{d,\infty}(0,r)} g_{\mathbf{X}}(w)dw. \quad (22)$$

Obviously, under the Poisson case $R_{\mathbf{X}} = 1$ whereas $R_{\mathbf{X}} < 1$ means that \mathbf{X} is repulsive. More precisely, the more $R_{\mathbf{X}} < 1$ the more repulsive \mathbf{X} . We now present the interest of $R_{\mathbf{X}}$ in our context.

Proposition 3.1. *Let $\mathbf{X} \sim \text{DPP}_B(K)$ be a DPP with kernel K satisfying (H'). Then, for any $I \subseteq \bar{d}$*

$$R_{\mathbf{X}_I}(r) = 1 - \left(\prod_{i \in I^c} \frac{\text{tr}_{B_i}(K_i^{(2)})}{\text{tr}_{B_i}(K_i)^2} \right) \left(\prod_{i \in I} \int_0^1 \frac{|K_i(tr)|^2}{K_i(0)^2} dt \right) \quad (23)$$

In particular, if K satisfies (H''):

$$R_{\mathbf{X}_I}(r) = 1 - \kappa_0^{d-\iota} \left(\int_0^1 \frac{|K_0(tr)|^2}{K_0(0)^2} dt \right)^\iota \quad (24)$$

where

$$\kappa_0 = \frac{\text{tr}_{B_0}(K_0^{(2)})}{\text{tr}_{B_0}(K_0)^2}.$$

The proof of this result follows directly from (10) and (22). Focusing on examples presented in the previous sections, we have

$$R_{\mathbf{X}_I}(r) = \begin{cases} 1 - \kappa_2^{d-\iota} \left(\int_0^1 e^{-2t^2 r^2 / \alpha^2} dt \right)^\iota & \text{for a Gaussian DPP,} \\ 1 - \kappa_1^{d-\iota} \left(\int_0^1 e^{-2tr/\alpha} dt \right)^\iota & \text{for an } L^1\text{-Exponential DPP,} \\ 1 - \frac{1}{N} \prod_{i \in I} \sum_{|j| < n_i} \left(1 - \frac{|j|}{n_i} \right) \text{sinc}(2\pi jr) & \text{for an } (N, d)\text{-Dirichlet DPP} \end{cases}$$

where κ_2 and κ_1 are defined by (13) and (16), respectively and sinc is the cardinal sine function.

Figures 3-5 investigate the situation for $d = 6, 10, 100$ respectively. Ripley's functions for point processes \mathbf{X}_I based on the three models exposed in this section are depicted. The intensity is set to $\rho_{\mathbf{X}} = 500$ and $\iota = d - i$ for $i = 0, \dots, 5$. The Gaussian DPP and L^1 -Exponential DPP satisfy (H'') , and so we decide, without loss of generality, to discard the last coordinates to define the projections. Since the (N, d) -Dirichlet DPP satisfies only (H') , the choice of directions has an influence. For this process, Ripley's functions have been computed using a Monte-Carlo approach (based on 10^4 replications): the coordinates to be removed are randomly chosen. The plots for the (N, d) -Dirichlet DPPs represent therefore the empirical mean of Ripley's functions. First and third quartiles are also represented by envelopes to get an idea of the variability. The visual results show that for $\rho_{\mathbf{X}} = 500$, the (N, d) -Dirichlet DPP is the most repulsive among the three models. Moreover, the loss of repulsiveness when projecting turns out to be smaller for (N, d) -Dirichlet DPPs than for the two other DPP models. The envelopes reported for the (N, d) -Dirichlet should be taken with attention. We could be tempted to conclude that the quite high variability observed for $d = 6, 10$, is too important to get practical interesting results. However, Section 4 will discredit this argument.

The (N, d) -Dirichlet DPP is the most repulsive in the situations considered here. However, it is worth mentioning that it may behave very badly according to the value of N . For example, we have observed that the less N has factors the less repulsive the (N, d) -Dirichlet DPP. The values of these factors also affect the repulsiveness of the DPP. In particular, if N is a high prime number, both situations are encountered which yields a disastrous model in terms of repulsion. Figures 3-5 underline that the class of L^1 -Exponential DPP is definitely less interesting than the class of Gaussian DPP. Given an ι , Ripley's function is closer to 1 and the convergence to 1 when ι decreases is faster for L^1 -Exponential DPP. For this reason, the L^1 -Exponential DPP is not considered in the next section.

4. Numerical illustrations

In this section, we illustrate the interest of projected DPP models by simulation experiments. For some $d \geq 1$ and $I \subseteq \bar{d}$, the problem we consider is to estimate using a Monte-Carlo approach, an integral of the form

$$\mu(f_I) = \int_{[0,1]^d} f_I(u) du$$

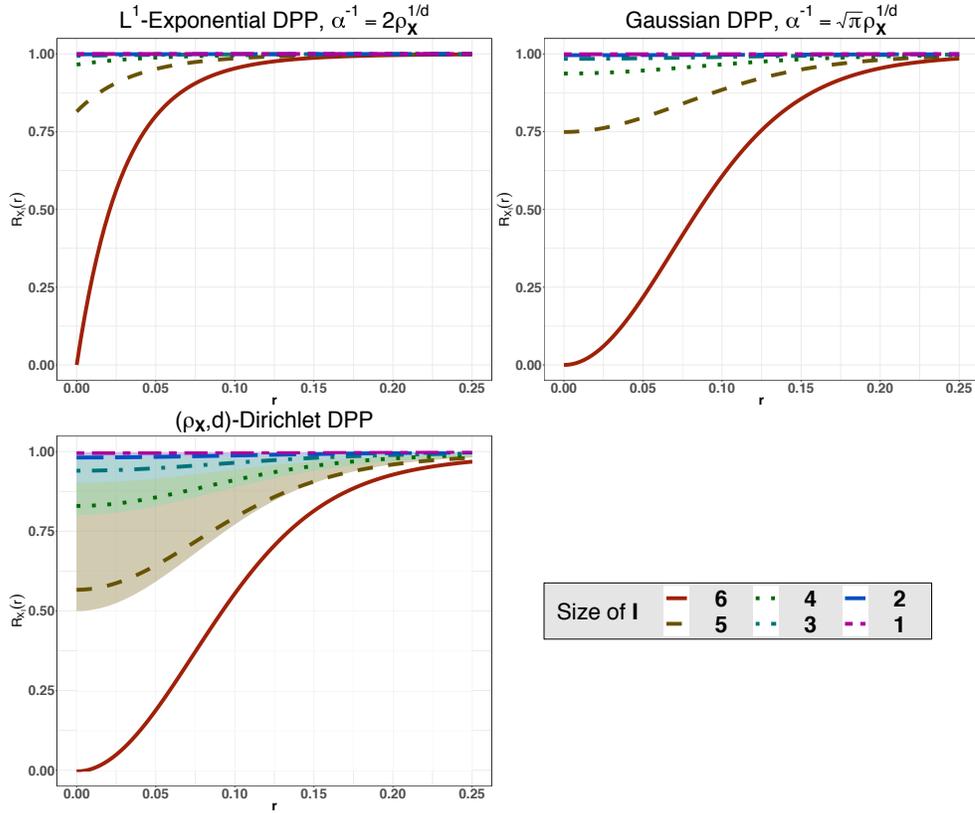


Figure 3: Ripley's functions (see (22)) of a 6-dimensional DPP \mathbf{X} with intensity $\rho_{\mathbf{X}}$ (solid lines) and its successive projections \mathbf{X}_I ($|I| = d - 1, \dots$; dotted and dashed lines) for the L^1 -Exponential DPP (top-left), Gaussian DPP (top right) and the $(N, 6)$ -Dirichlet DPP (bottom left). For the Dirichlet case, coordinates to be removed are chosen randomly (10^4 replications): dotted and dashed lines represent empirical means while first and third quartiles are represented by envelopes.

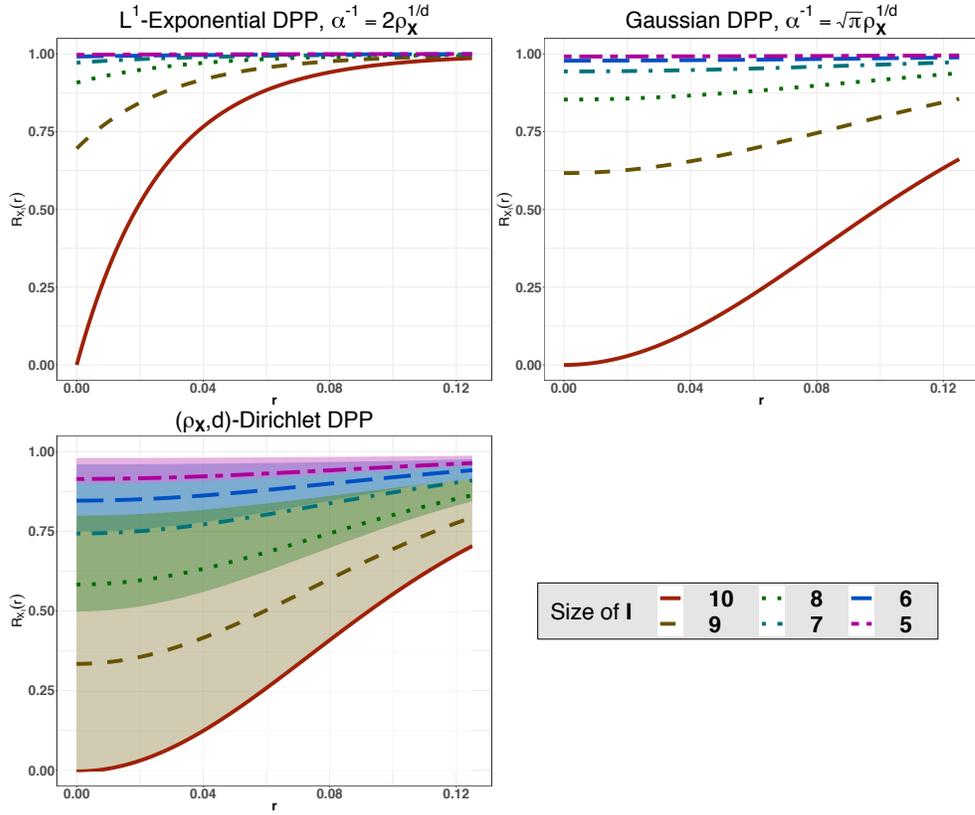


Figure 4: Ripley's functions (see (22)) of a 10-dimensional DPP \mathbf{X} with intensity $\rho_{\mathbf{X}}$ (solid lines) and its successive projections \mathbf{X}_I ($|I| = d - 1, \dots$; dotted and dashed lines) for the L^1 -Exponential DPP (top-left), Gaussian DPP (top right) and the $(N, 10)$ -Dirichlet DPP (bottom left). For the Dirichlet case, coordinates to be removed are chosen randomly (10^4 replications): dotted and dashed lines represent empirical means while first and third quartiles are represented by envelopes.

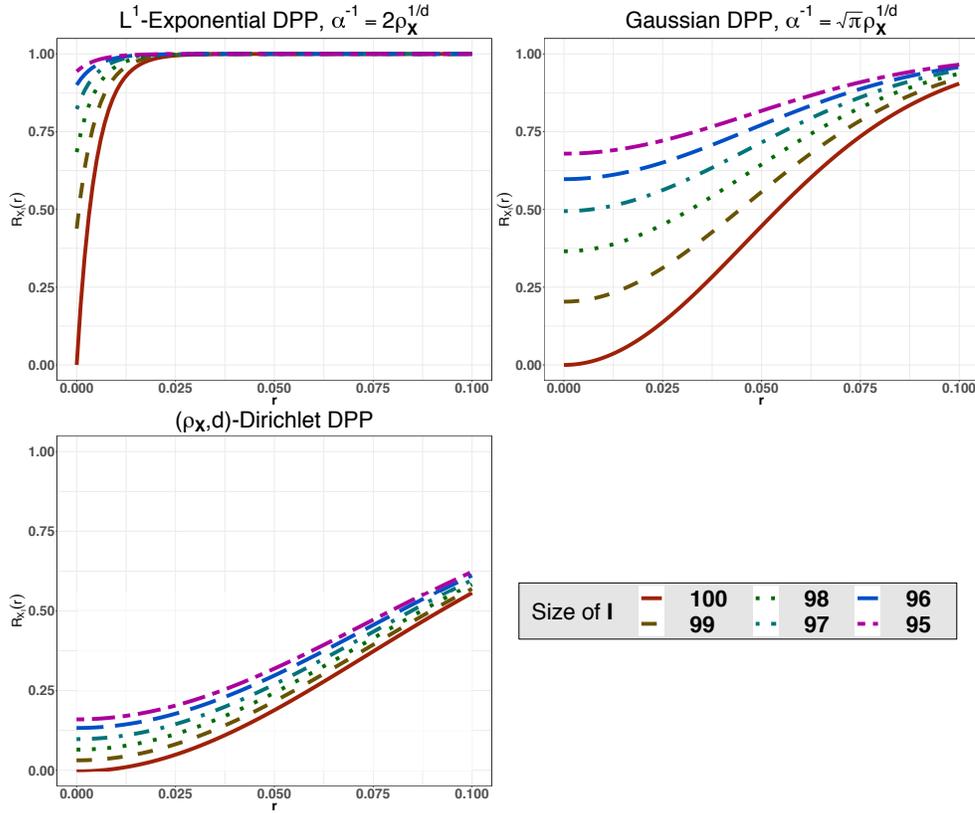


Figure 5: Ripley's functions (see (22)) of a 100-dimensional DPP \mathbf{X} with intensity $\rho_{\mathbf{X}}$ (solid lines) and its successive projections \mathbf{X}_I ($|I| = d - 1, \dots$; dotted and dashed lines) for the L^1 -Exponential DPP (top-left), Gaussian DPP (top right) and the $(N, 100)$ -Dirichlet DPP (bottom left). For the Dirichlet case, coordinates to be removed are chosen randomly (10^4 replications): dotted and dashed lines represent empirical means while first and third quartiles are represented by envelopes.

where $f_I : [0,1]^t \rightarrow \mathbb{R}^+$ is a t -dimensional function. A standard way for achieving this task (which includes the uniform sampling design) is to define a point process, say \mathbf{Z}_I , on $[0,1]^t$ and estimate $\mu(f_I)$ using the unbiased estimator

$$\hat{\mu}_{\mathbf{Z}_I}(f_I) = \rho_{\mathbf{Z}_I}^{-1} \sum_{u \in \mathbf{Z}_I} f_I(u). \quad (25)$$

Given I and f_I , this problem has been widely considered in the literature (e.g. [Robert and Casella, 2004](#); [Delyon and Portier, 2016](#)). In particular, [Bardenet and Hardy \(2020\)](#) have constructed an ad-hoc DPP on $[0,1]^t$ and provided very interesting asymptotic results. In this section, we investigate another aspect. We consider the problem not only for one but various integrals, defined for different subsets $I \subseteq \bar{d}$ and based on a single realization of a point process defined on $[0,1]^d$. This problem, for which investigated models are definitely meaningful, mimics problems encountered in computer experiments where the spatial design is initially defined on \mathbb{R}^d but later used with a few coordinates discarded (e.g. [Woods and Lewis, 2006](#); [Kleijnen, 2017](#)).

To do this, we therefore consider a spatial point process \mathbf{X} (and in particular DPP models developed in the previous section) and we estimate $\mu(f_I)$ by (25) with $\mathbf{Z}_I = \mathbf{X}_I$ where \mathbf{X}_I is the projected point pattern of \mathbf{X} on $[0,1]^t$. The interest of our models lies in the following equation which evaluates $\text{Var}(\hat{\mu}_{\mathbf{X}_I}(f_I))$. Using Campbell Theorem (1)

$$\begin{aligned} \text{Var}(\hat{\mu}_{\mathbf{X}_I}(f_I)) &= \rho_{\mathbf{X}_I}^{-1} \int_{[0,1]^t} f_I(u)^2 du \\ &+ \int_{[0,1]^t} \int_{[0,1]^t} (g_{\mathbf{X}_I}(u, v) - 1) f_I(u) f_I(v) du dv. \end{aligned} \quad (26)$$

As soon as $g_{\mathbf{X}_I} < 1$, the variance is smaller than the first integral which turns out to be the variance under the Poisson case. In this section, we intend to verify this property with models considered in this paper.

In the following, we let $d = 6$ and, following [Bardenet and Hardy \(2020, Sec. 3\)](#), we consider for any $I \subseteq \bar{6}$ the ‘‘bump’’ test function

$$f_I(u) = \exp\left(-\sum_{i \in I} \frac{1}{1 - 4(u_i - 1/2)^2}\right), \quad u \in [0,1]^t. \quad (27)$$

Three type of models are investigated: a homogeneous Poisson point process (which serves as a reference), a Gaussian DPP, and an $(N, 6)$ -Dirichlet DPP.

Simulations of DPPs can be realized using R package `spatstat`. However, using this package leads to performance issues (in particular in terms of memory) when simulating DPPs with high intensity and/or high dimension. Therefore, we have implemented the simulation algorithms in C++ and have made them usable with R. The codes are available on GitHub (<https://github.com/AdriMaz/rcdpp/>).

Figure 6 reports empirical variances of estimates of $\mu(f_I)$ based on $m = 10^4$ replications of each model, in terms of $\rho_{\mathbf{X}}$ where $\rho_{\mathbf{X}} = 200, 400, 600, 800, 1000$. We consider all possible projections, i.e. $\iota = 6, 5, 4, 3, 2, 1$. For the Poisson case, note that \mathbf{X}_I has the same distribution as a homogeneous Poisson point process (with the same intensity) defined on B_I . For the Gaussian DPP, the parameter α is set to $\alpha^{-1} = \sqrt{\pi}\rho^{1/6}$. When $\iota < d$, the coordinates to be discarded are chosen randomly. This has no influence for the Poisson, Gaussian DPP since these models satisfy Assumption (H'') but is important for the $(N, 6)$ -Dirichlet DPP.

Figure 6 illustrates the interest of this research. It is clear that whatever the dimension of the function to integrate, i.e. whatever $\iota = 6, \dots, 1$, the empirical variance of Monte-Carlo estimates using one single realization of a spatial point process defined in dimension d , is always smaller than in the independent case. The (N, d) -Dirichlet model outperforms the Gaussian DPP for any $I \subseteq \bar{d}$ as already observed from a theoretical point of view in the previous section. The general result of this paper states that a projected DPP seems less and less repulsive after successive projections. However, It is interesting to observe that this fact does not affect that much the properties of Monte-Carlo integration estimates.

Conclusion

The objective of this paper is to explore properties of projections of a DPP \mathbf{X} with kernel K and defined on a compact set B of \mathbb{R}^d . For any $I \subset \{1, \dots, d\}$, our general conclusion is that the projection \mathbf{X}_I remains repulsive when kernel K is separable, in the sense that $g_{\mathbf{X}_I} < 1$ uniformly for any $I \subset \{1, \dots, d\}$. In particular if kernel K is a projection kernel, \mathbf{X}_I falls in the class of α -DPPs (with $\alpha = -1/n, n \in \mathbb{N}$). We have proposed a few examples of such separable kernels and compared them using an original summary statistic based on a normalized version of the Ripley's function defined with the sup norm. We have finally illustrated this paper for Monte-Carlo integration problems when the problem is to estimate integrals over

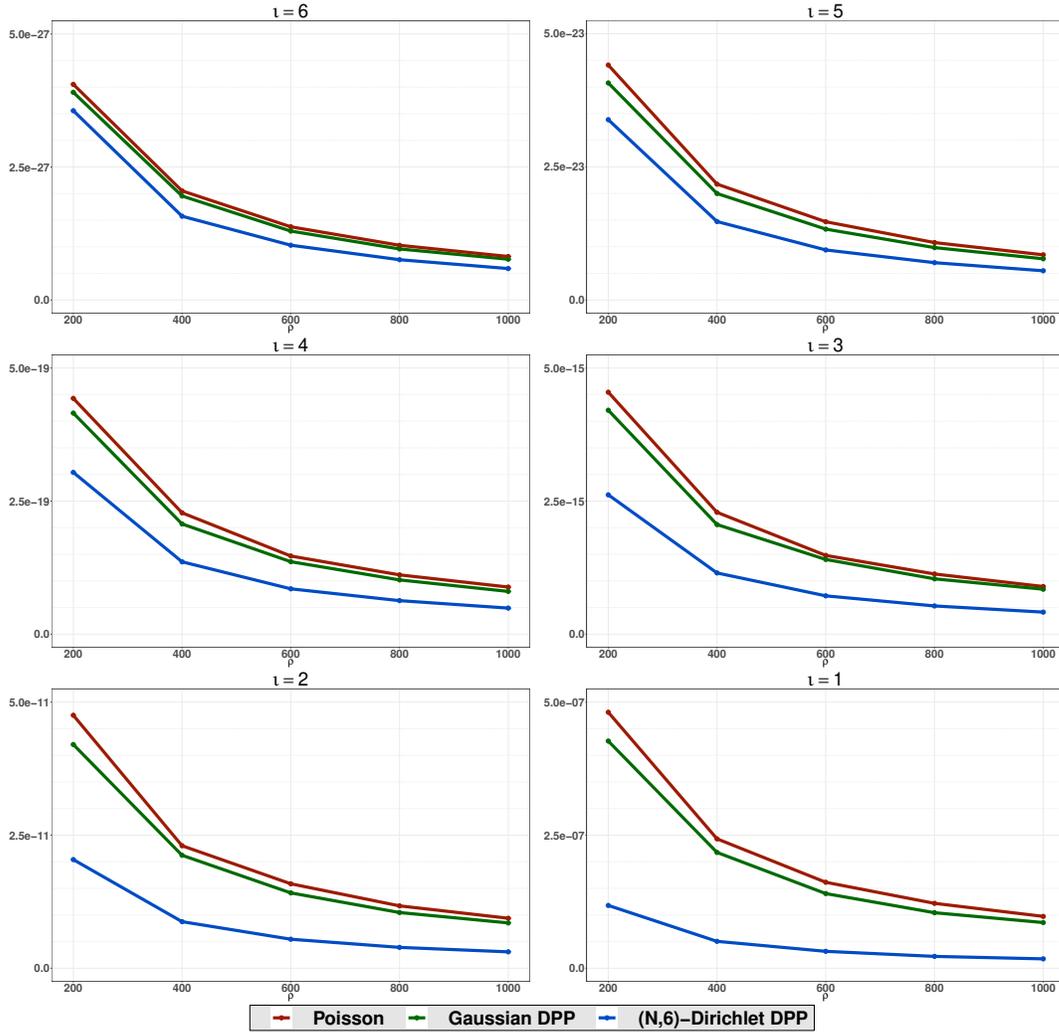


Figure 6: Empirical variances of Monte-Carlo integral estimates of the form (25) for the function (27) using Poisson process (red lines), Gaussian DPP (green lines) and Dirichlet DPP (blue lines) for $l = |I| = 6, \dots, 1$, based on 10^4 replications of a 6-dimensional point processes with intensity $\rho_{\mathbf{X}} = 200, 400, 600, 800, 1000$. When $l < 6$, coordinates to be removed are chosen randomly.

a compact set B_I of an ι -dimensional function for any $1 \leq \iota \leq d$, using the same quadrature points defined in B . To be fully relevant, comparisons with designs built from other classes of point processes (e.g. Gibbs, Matérn, Multivariate OP Ensembles), more standard designs (e.g. Halton, Sobol, Quasi-Monte-Carlo), in which other test functions with different properties (e.g. less regular or non-compactly supported) would be considered, should be performed. We leave this for a future research.

Acknowledgements

The authors would like to thank Frédéric Lavancier and Arnaud Poinas for fruitful discussions, and the associate editor and reviewers for valuable suggestions and comments. The research of J.-F. Coeurjolly and A. Mazoyer is supported by the Natural Sciences and Engineering Research Council. P.-O. Amblard is partially funded by Grenoble Data Institute (ANR-15-IDEX-02) and LIA CNRS/Melbourne Univ Geodesic.

References

- Baddeley, A., 2007. Spatial Point Processes and their Applications. Springer, Berlin, Heidelberg. pp. 1–75.
- Bardenet, R., Hardy, A., 2020. Monte Carlo with determinantal point processes. *Ann. Appl. Probab.* (to appear).
- Debnath, L., Mikusiński, P., 2005. Introduction to Hilbert spaces with applications. 3rd ed., Elsevier Academic Press.
- Decreusefond, L., Flint, I., Privault, N., Torris, G., 2016. Determinantal point processes. Springer International Publishing, Cham. pp. 311–342.
- Delyon, B., Portier, F., 2016. Integral approximation by kernel smoothing. *Bernoulli* 22, 2177–2208.
- Dereudre, D., 2019. Introduction to the theory of Gibbs point processes, in: *Stochastic Geometry*. Springer, pp. 181–229.
- Dupuy, D., Helbert, C., Franco, J., 2015. Dicedesign and diceeval: Two R packages for design and analysis of computer experiments. *J. Stat. Softw.* 65, 1–38.

- Franco, J., Bay, X., Dupuy, D., Corre, B., 2008. Planification d'expériences numériques à partir du processus ponctuel de Strauss. URL: <https://hal.archives-ouvertes.fr/hal-00260701>. working paper or preprint.
- Halton, J.H., 1964. Algorithm 247: Radical-inverse quasi-random point sequence. *Communications ACM* 7, 701–702.
- Hough, J., Krishnapur, M., Peres, Y., Virág, B., 2006. Determinantal processes and independence. *Probability Surveys* 3, 206–229.
- Hough, J., Krishnapur, M., Peres, Y., Virág, B., 2009. Zeros of Gaussian Analytic Functions and Determinantal Point Processes. American Mathematical Society.
- Illian, J., Penttinen, A., Stoyan, H., Stoyan, D., 2008. Statistical Analysis and Modelling of Spatial Point Patterns. *Statistics in Practice*, Wiley, Chichester.
- Kleijnen, J., 2017. Design and analysis of simulation experiments: tutorial. Springer, New York. pp. 135–158.
- Lavancier, F., Møller, J., Rubak, E., 2015. Determinantal point process models and statistical inference: Extended version. *J. Roy. Stat. Soc. B* 77, 853–877.
- Macchi, O., 1975. The coincidence approach to stochastic processes. *Adv. Appl. Probab.* 7, 83–122.
- McCullagh, P., Møller, J., 2006. The permanental process. *Adv. Appl. Probab.* 38, 873–88.
- McKay, M., Beckman, R.J., Conover, W., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21, 239–245.
- Møller, J., Waagepetersen, R., 2004. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton.
- Owen, A., 1992. Orthogonal arrays for computer experiments, integration and visualization. *Stat. Sinica* 2, 439–452.
- Owen, A., 2013. *Monte Carlo theory, methods and examples*.

- Riesz, F., Sz-Nagy, B., 1990. *Functional Analysis*. Dover Publications, New York.
- Robert, C., Casella, G., 2004. *Monte Carlo Statistical Methods*. Springer-Verlag, New York.
- Santner, T., Williams, B., Notz, W., 2013. *The design and analysis of computer experiments*. Springer Science & Business Media.
- Shirai, T., Takahashi, Y., 2003. Random point fields associated with certain fredholm determinants i. Fermion, Poisson and boson point processes. *J. Funct. Anal.* 205, 414–463.
- Sobol, I., 1967. On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Comp. Math. Math+* 7, 86 – 112.
- Soshnikov, A., 2000. Determinantal random point fields. *Russ. Math. Surv+* 55, 923–975.
- Sun, F., Wang, Y., Xu, H., 2019. Uniform projection designs. *Ann. Stat.* 47, 641–661.
- Teichmann, J., Ballani, F., van den Boogaart, K., 2013. Generalizations of Matérn’s hard-core point processes. *Spat. Stat.-Neth.* 3, 33–53.
- Woods, D., Lewis, S., 2006. *Design of experiments for screening*. Springer, New York. pp. 1143–1185.
- Zygmund, A., 2003. *Trigonometric Series (Volumes I & II combined)*. 3rd ed., Cambridge University Press.

Appendix A. Proof of Lemma 2.1

Proof. For any non-negative measurable function $h_I : B_I^k \rightarrow \mathbb{R}^+$, we have using Campbell Theorem (1)

$$\begin{aligned}
& \int_{B_I^k} h_I(x_I^{(1)}, \dots, x_I^{(k)}) \rho_{\mathbf{X}_I}^{(k)}(x_I^{(1)}, \dots, x_I^{(k)}) dx_I^{(1)} \dots dx_I^{(k)} \\
&= \mathbb{E} \left[\sum_{x_I^{(1)}, \dots, x_I^{(k)} \in \mathbf{X}_I}^{\neq} h_I(x_I^{(1)}, \dots, x_I^{(k)}) \right] \\
&= \mathbb{E} \left[\sum_{x^{(1)}, \dots, x^{(k)} \in \mathbf{X}}^{\neq} (h_I \circ P_I)(x^{(1)}, \dots, x^{(k)}) \right] \\
&= \int_{B_I^k} h_I(x_I^{(1)}, \dots, x_I^{(k)}) \left\{ \int_{(B_{I^c})^k} \rho_{\mathbf{X}}^{(k)}((x^{(1)}, u^{(1)}), \dots, (x^{(k)}, u^{(k)})) \right. \\
&\quad \left. du^{(1)} \dots du^{(k)} \right\} dx_I^{(1)} \dots dx_I^{(k)}
\end{aligned}$$

whereby we deduce (7) by identification. \square

Appendix B. Proof of Theorem 2.1

Proof. Let us write (8) under ($H[I]$).

$$\begin{aligned}
\rho_{\mathbf{X}_I}^{(k)}(x^{(1)}, \dots, x^{(k)}) &= \sum_{\sigma \in S_k} (-1)^{k-C(\sigma)} \int_{(B_{I^c})^k} \prod_{i=1}^k K((x, u)^{(i)}, (x, u)^{(\sigma(i))}) du^{(1)} \dots du^{(k)} \\
&= \sum_{\sigma \in S_k} (-1)^{k-C(\sigma)} \prod_{i=1}^k K_I(x_I^{(i)}, x_I^{(\sigma(i))}) \\
&\quad \times \int_{(B_{I^c})^k} \prod_{i=1}^k K_{I^c}(u^{(i)}, u^{(\sigma(i))}) du^{(1)} \dots du^{(k)}.
\end{aligned} \tag{B.1}$$

For any $\sigma \in S_k$ let us denote by $\text{supp}(\sigma)$ its support:

$$\text{supp}(\sigma) = \{i \in \bar{k} \text{ s.t. } \sigma(i) \neq i\},$$

by $c(\sigma)$ the number of elements of $\text{supp}(\sigma)$, by $\mathcal{S}(\sigma)$ the set of disjoint cycles of σ with non-empty support and by $C(\sigma)$ the number of disjoint cycles of σ (including those with empty support). Consider the case where $C(\sigma) = 1$ (i.e. σ is a circular permutation of \bar{k}). Then the integral part in (B.1) can be written as

$$\begin{aligned}
& \int_{(B_{I^c})^k} \prod_{i=1}^k K_{I^c}(u^{(i)}, u^{(\sigma(i))}) \, du^{(1)} \dots du^{(k)} \\
&= \int_{(B_{I^c})^k} K_{I^c}(u^{(1)}, u^{(\sigma(1))}) \dots K_{I^c}(u^{(\sigma(1))}, u^{(\sigma^2(1))}) \dots \\
&\quad \dots K_{I^c}(u^{(k)}, u^{(\sigma(k))}) \, du^{(1)} \dots du^{(\sigma(1))} \dots du^{(k)} \\
&= \int_{(B_{I^c})^{k-1}} K_{I^c}^{(2)}(u^{(1)}, u^{(\sigma^2(1))}) \dots \\
&\quad \dots K_{I^c}(u^{(\sigma(1)-1)}, u^{(\sigma(\sigma(1)-1))}) K_{I^c}(u^{(\sigma(1)+1)}, u^{(\sigma(\sigma(1)+1))}) \dots \\
&\quad \dots K_{I^c}(u^{(k)}, u^{(\sigma(k))}) \, du^{(1)} \dots du^{(\sigma(1)-1)} du^{(\sigma(1)+1)} \dots du^{(k)} \\
&\quad \vdots \\
&= \int_{(B_{I^c})^2} K_{I^c}^{(k-1)}(u^{(1)}, u^{(\sigma^{k-1}(1))}) K_{I^c}(u^{(\sigma^{k-1}(1))}, u^{(\sigma^k(1))}) \, du^{(1)} du^{(\sigma^{k-1}(1))} \\
&= \text{tr}_{B_{I^c}} \left(K_{I^c}^{(k)} \right).
\end{aligned}$$

Assume now that $C(\sigma) > 1$. Then σ can be written as

$$\sigma = \left(\begin{array}{c} \odot \\ \varepsilon \end{array} \right)_{\varepsilon \in \mathcal{S}(\sigma)} \odot \mathbf{i}_k(\sigma), \tag{B.2}$$

where $\mathbf{i}_k(\sigma)$ is the identity on $\bar{k} \setminus \text{supp}(\sigma)$, and \odot denotes the permutation product. Observe that (B.2) implies

$$C(\sigma) = \#\mathcal{S}(\sigma) + k - c(\sigma). \tag{B.3}$$

If $1 \in \text{supp}(\sigma)$, there is only one permutation $\varsigma \in \mathcal{S}(\sigma)$ such that $1 \in \text{supp}(\varsigma)$. Therefore:

$$\begin{aligned} & \int_{(B_{I^c})^k} \prod_{i=1}^k K_{I^c} (u^{(i)}, u^{(\sigma(i))}) \, du^{(1)} \dots du^{(k)} \\ &= \text{tr}_{B_{I^c}} \left(K_{I^c}^{(c(\iota))} \right) \int_{(B_{I^c})^{k-c(\varsigma)}} \prod_{i \in \bar{k} \setminus \text{supp}(\varsigma)} K_{I^c} (u^{(i)}, u^{(\sigma(i))}) \, du^{(1)} \dots du^{(k)} \end{aligned}$$

Denote by α the minimum of $\bar{k} \setminus \text{supp}(\varsigma)$. As above there is only one permutation $\varsigma' \in \mathcal{S}(\sigma)$ such that $\alpha \in \text{supp}(\varsigma')$. Then:

$$\begin{aligned} & \int_{(B_{I^c})^k} \prod_{i=1}^k K_{I^c} (u^{(i)}, u^{(\sigma(i))}) \, du^{(1)} \dots du^{(k)} \\ &= \text{tr}_{B_{I^c}} \left(K_{I^c}^{(c(\varsigma))} \right) \int_{(B_{I^c})^{k-c(\varsigma)}} \prod_{i \in \bar{k} \setminus \text{supp}(\varsigma)} K_j (u^{(i)}, u^{(\sigma(i))}) \, du^{(1)} \dots du^{(k)} \\ &= \text{tr}_{B_{I^c}} \left(K_{I^c}^{(c(\varsigma))} \right) \text{tr}_{B_j} \left(K_{I^c}^{(c(\varsigma'))} \right) \\ & \quad \times \int_{(B_{I^c})^{k-(c(\varsigma)+c(\varsigma'))}} \prod_{i \in \bar{k} \setminus (\text{supp}(\varsigma) \cup \text{supp}(\varsigma'))} K_{I^c} (u^{(i)}, u^{(\sigma(i))}) \, du^{(1)} \dots du^{(k)}. \end{aligned}$$

Therefore, one gets by induction:

$$\begin{aligned} & \int_{(B_{I^c})^k} \prod_{i=1}^k K_{I^c} (u^{(i)}, u^{(\sigma(i))}) \, du^{(1)} \dots du^{(k)} \\ &= \left[\prod_{\varepsilon \in \mathcal{S}(\sigma)} \text{tr}_{B_{I^c}} \left(K_{I^c}^{(c(\varepsilon))} \right) \right] \int_{(B_{I^c})^{k-c(\sigma)}} \prod_{i \in \bar{k} \setminus \text{supp}(\sigma)} K_{I^c} (u^{(i)}, u^{(\sigma(i))}) \, du^{(1)} \dots du^{(k)}. \end{aligned} \tag{B.4}$$

Plugging (B.4) into (B.1) leads to (9). □

Appendix C. Laplace functionals

Lemma Appendix C.1. *Let $I \subset \bar{d}$ and let \mathbf{X} be a spatial point process defined on a compact set $B \subset \mathbb{R}^d$. Then, for any Borel function $h_I : B_I \rightarrow \mathbb{R}^+$*

$$\mathcal{L}_{\mathbf{X}_I}(h_I) = \mathcal{L}_{\mathbf{X}}(h_I \circ P_I). \quad (\text{C.1})$$

Proof. Equation (C.1) follows arguments similar to the ones used in the proof of Lemma 2.1.

$$\mathcal{L}_{\mathbf{X}_I}(h_I) = \mathbb{E} \left[\prod_{y \in \mathbf{X}_I} e^{-h_I(y)} \right] = \mathbb{E} \left[\prod_{x \in \mathbf{X}} e^{-h_I(x_I)} \right] = \mathcal{L}_{\mathbf{X}}(h_I \circ P_I)$$

□

Theorem Appendix C.1. *Let $I \subseteq \bar{d}$ and $\mathbf{X} \sim \text{DPP}_B(K)$ such that K satisfies (H[I]). The Laplace functional of the projected point process \mathbf{X}_I is given for any Borel function $h_I : B_I \rightarrow \mathbb{R}^+$ by:*

$$\mathcal{L}_{\mathbf{X}_I}(h_I) = \prod_{l \in \mathcal{N}_{I^c}} \exp \left\{ - \sum_{k \geq 1} \frac{\text{tr}_{B_I} \left(K_{\lambda_l^{(I^c)}}^{(k)} K_{I, h_I} \right)}{k} \right\} \quad (\text{C.2})$$

$$= \exp \left\{ - \sum_{k \geq 1} \frac{\text{tr}_{B_{I^c}} \left(K_{I^c}^{(k)} \right) \text{tr}_{B_I} \left(K_{I, h_I}^{(k)} \right)}{k} \right\} \quad (\text{C.3})$$

where $K_{I, h_I} : B_I \times B_I \rightarrow \mathbb{C}$ is the kernel defined by

$$K_{I, h_I}(x, y) = \sqrt{1 - e^{-h_I(x)}} K_I(x, y) \sqrt{1 - e^{-h_I(y)}}.$$

Proof. The proof is straightforward and follows from Lemma Appendix C.1, assumption (H[I]) and the definition of K_{I, h_I} . □