Monte Carlo integration of non-differentiable functions on $[0,1]^\iota$, $\iota = 1, \ldots, d$, using a single determinantal point pattern defined on $[0,1]^d$ *

Jean-François Coeurjolly

Laboratoire Jean Kuntzmann
Université Grenoble Alpes
38402 Saint Martin d'Hères, France
e-mail: jean-francois.coeurjolly@univ-grenoble-alpes.fr

Adrien Mazoyer

Institut de Mathématiques de Toulouse
118 route de Narbonne
F-31062 Toulouse Cedex 9, France
e-mail: adrien.mazoyer@math.univ-toulouse.fr

Pierre-Olivier Amblard

Gipsa-Lab, CNRS
Université Grenoble Alpes
38402 Saint Martin d'Hères, France
e-mail: pierre-olivier.amblard@gipsa-lab.grenoble-inp.fr

Abstract: This paper concerns the use of a particular class of determinantal point processes (DPP), a class of repulsive spatial point processes, for Monte Carlo integration. Let $d \geq 1$, $I \subseteq \overline{d} = \{1, \ldots, d\}$ with $\iota = |I|$. Using a single set of N quadrature points $\{u_1, \ldots, u_N\}$ defined, once for all, in dimension d from the realization of a specific DPP, we investigate "minimal" assumptions on the integrand in order to obtain unbiased Monte Carlo estimates of $\mu(f_I) = \int_{[0,1]^\iota} f_I(u) \mathrm{d}u$ for any known ι-dimensional integrable function on $[0,1]^\iota$. In particular, we show that the resulting estimator has variance with order $N^{-1-(2s\wedge 1)/d}$ when the integrand belongs to some Sobolev space with regularity $s > 0$. When $s > 1/2$ (which includes a large class of non-differentiable functions), the variance is asymptotically explicit and the estimator is shown to satisfy a Central Limit Theorem.

Introduction

The paper investigates Monte-Carlo evaluation of the integral $\mu(f) = \int_{[0,1]^\iota} f(u)\mathrm{d}u$ for a *known* ι-dimensional integrable function on $[0,1]^\iota$ using *a single set* of N quadrature points $\{u_1, \ldots, u_N\}$ defined once for all in dimension $d \geq \iota$. The same set of nodes, defined in $[0,1]^d$, may therefore be used to estimate *a finite number of different integrals*, and therefore this set does not exploit the form of f, (locations of its possible singularities, *etc*).

Such an approach may be of importance in any application where repeatedly integrating a function over different subsets is needed, (*e.g.* in sensitivity analysis using experimental designs) or where calculating marginals is required (*e.g.* for evaluating the marginal likelihood in some parametric statistical models). The first example is emblematic of our motivation.

In the context of computer experiments (see for example [30, Chapter 5]), complex phenomena are simulated using a mathematical model to replace the process which generates the data. Usually, the model depends on a large number of parameters (inputs). An objective of the experiments is to quantify the influence of the variability of the inputs on the variable of interest. An experiment consists in running simulations, where each simulation represents a possible combination of the inputs. It is impossible in practice to consider all possible configurations, the number of simulations being limited. Therefore, the design of experiments, *i.e.* the choice of combinations of inputs, is of great importance. Under a lack of information on how inputs are linked to outputs, a strategy is to spread chosen inputs to cover as much as possible all the input space. This technique is called *space-filling design*. It can be summarized by generating N points in a given space which regularly cover this space. Latin hypercubes [20, 23], low discrepancy sequences (see *e.g.* [12, 33]) are standard methods to generate designs. The goal of computer experiments is not only to examine the influence of all the inputs on an output of interest, but also the influence of a subset of these inputs, or also the influence of a particular combination of subsets of these inputs. Since computer experiments may be very expensive in terms of computation load and/or storage capacity, the regularity of the coverage of the designs should be conserved when the initial configuration is projected onto lower dimensional spaces. This would allow to use the initial configuration to study the influence of subsets of inputs for example with the same efficiency. Furthermore, integrating a function over different subsets of its variables may be of importance in this context, for example in sensitivity analysis [30, Chapter 7]. When integrating explicitly is impossible, numerical methods are used. We turn our attention here to Monte Carlo integration.

Monte Carlo integration has a long history and it is not the aim of this paper to make a detailed bibliography. We refer the interested reader by an extensive treatment and bibliography to the electronic book by Owen [26]. Let us however cite a few methods keeping in mind what we mean by "minimal" assumptions in the situation $\iota = d$. Crude Monte Carlo methods and importance sampling methods (see *e.g.* [28]) consist in using i.i.d. nodes $\{u_1, \ldots, u_N\}$ with a so-called proposal density (the uniform density in the usual situation

in which the target distribution is also uniform). Under some $L^2([0,1]^d)$ type assumption on the integrand, the resulting estimator denoted by $\widehat{\mu}_N(f)$ has a variance proportional to N^{-1} and satisfies a Central Limit Theorem. When d is large, Monte Carlo Markov Chains (MCMC) methods, where the set of quadrature points is the realization of a particular Markov chain, are usually preferred. When $f \in L^1([0,1]^d)$, the variance of $\widehat{\mu}_N(f)$ is still of order N^{-1} and satisfies a CLT (see *e.g.* [9]). To improve the rate of convergence, the price to pay is to require some regularity assumptions on f. Many methods exist in the literature: grid-based stratified methods [11], possibly combined with antithetic sampling (see Owen [26, Chapter 10]), Quasi Monte Carlo and randomized versions, scrambled nets [8, 24, 25, 26, 4], *etc*. For example, a version of scrambled nets with antithetic sampling can lead to an estimator with variance $\mathcal{O}(N^{-3-2/d}\log(N)^{d-1})$ if, to simplify Owen's assumption [25], f is d times continuously differentiable on $[0,1]^d$. Additional assumptions on the scrambled net are required to obtain a CLT. Grid-based stratified methods which are maybe the first simple alternative to ordinary Monte Carlo methods require that f is continuously differentiable on $[0,1]^d$ and yield an estimator satisfying a CLT with variance asymptotically proportional to $N^{-1-2/d}$. Let us also mention that Owen [24] showed that a version of scrambled net has a variance $o(N^{-1})$ under the sole assumption that $f \in L^2([0,1]^d)$ but the rate is not explicit until strong regularity assumptions are made on f.

In a recent work, Bardenet and Hardy [3] proposed another alternative by defining the nodes as the realization of a repulsive point pattern, and in particular a determinantal point pattern. The class of determinantal point processes (DPPs for short) has received a growing attention in the last decades (see *e.g.* Soshnikov [34], Shirai and Takahashi [32], Hough et al. [15], Lavancier, Møller and Rubak [17], Decreusefond et al. [7]), thanks to its very appealing properties in particular in terms of tractability and exact simulation. Bardenet and Hardy [3] have defined an Orthogonal Polynomial Ensemble, which is a particular inhomogeneous DPP whose kernel is defined through orthonormal polynomials. Under the assumption that f is continuously differentiable and compactly supported in $B' \subset [0,1]^d$ the authors obtained an estimator with variance equivalent to an explicit constant times $N^{-1-1/d}$.

In this paper, we investigate a different DPP. To be more explicit, we consider the most natural kernel, called the Dirichlet kernel in this paper, which is based on the Fourier decomposition of a rectangular subset of N indices of $\mathbb{Z}^d$. It is a projection DPP that is a point process which produces almost surely N points. It has the advantage to lead to a homogeneous DPP pattern, an interesting characteristic as we want the pattern to be used to estimate any integral without taking advantage of the integrand f. A second advantage is that the marginals are fully characterized and explicit which means that marginals can efficiently be used to estimate $\widehat{\mu}_N(f_I)$ for any $I \subset \overline{d} = \{1, \ldots, d\}$. Last but not least, our main result Theorem 3.1, shows that the resulting estimator $\widehat{\mu}_N(f_I)$ has asymptotic variance proportional to $N^{-1-(2s\wedge 1)/d}$ for any $f_I \in \mathcal{H}^s([0,1]^\iota)$ where $\mathcal{H}^s([0,1]^\iota)$ is some Sobolev space with regularity $s \geq 0$, see (3.4) for more details. We remind that for periodic functions $L^2([0,1]^\iota) = \mathcal{H}^0([0,1]^\iota)$ and if f_I

is periodic and continuously differentiable then $f_I \in \mathcal{H}^1([0,1]^\iota)$. In particular, our result states that when $s > 1/2$ (thus potentially for non-differentiable functions), the variance is asymptotically equivalent to an explicit constant times $N^{-1-1/d}$. In this case, we also obtain a central limit theorem for the estimator (assuming in addition that the integrand is bounded). As a summary, the estimator proposed has the characteristic to exhibit a variance that decreases faster than the ordinary Monte Carlo as soon as $s > 0$. The decay is slower than methods such as grid-based methods, scrambled nets, *etc*, but require much less regularity assumptions and can be applied to any ι-dimensional function, $\iota = 1, \ldots, d$.

The paper is organized as follows. Section 1 contains a background on spatial point processes, generalities on the projection of spatial point processes and DPPs. We also outline the interest of repulsive point processes and in particular DPPs for Monte Carlo integration. Section 2 introduces the Dirichlet DPP and exposes some of its properties. Our main result, Theorem 3.1, is presented in Section 3. It details convergence results for Monte Carlo integration based on the realization of a Dirichlet DPP. A multivariate version of this CLT is also proposed. Section 4 discusses computational aspects for the simulation of Dirichlet DPPs and contains a simulation study which illustrates our results. In Section 5, we perform a deeper numerical comparison of the sampling used here with several existing methods among which Bardenet&Hardy's, stratified sampling, maximinLHS, *etc*. Finally, all proofs of the results are postponed to Appendix B.

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, see for example [22] and references therein for measure theoretical details, whose realization is of the form $\{x^{(1)}, \ldots, x^{(k)}\} \in B^k$ 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 $\rho_{\mathbf{X}}^{(k)} : B^k \to \mathbb{R}^+$, $k \in \mathbb{N} \setminus \{0\}$. By Campbell theorem, see *e.g.* [22], $\rho_{\mathbf{X}}^{(k)}$ is characterized by the following integral representation: for any non-negative measurable function $h : B^k \to \mathbb{R}^+$

$$\mathbb{E}\left[\sum_{x^{(1)},\ldots,x^{(k)} \in \mathbf{X}}^{\neq} h\left(x^{(1)}, \ldots, x^{(k)}\right) \right] \tag{1.1}$$
$$= \int_{B^k} \rho_{\mathbf{X}}^{(k)} \left(x^{(1)}, \ldots, x^{(k)}\right) h\left(x^{(1)}, \ldots, x^{(k)}\right) \mathrm{d}x^{(1)} \ldots \mathrm{d}x^{(k)}$$

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

$$g_{\mathbf{X}}(x^{(1)}, x^{(2)}) = \frac{\rho_{\mathbf{X}}^{(2)}(x^{(1)}, x^{(2)})}{\rho_{\mathbf{X}}(x^{(1)})\rho_{\mathbf{X}}(x^{(2)})} \tag{1.2}$$

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, $\rho_{\mathbf{X}}^{(k)}\big(x^{(1)}, \ldots, x^{(k)}\big)$ 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 $x \in B$ by convention. 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 $g_{\mathbf{X}}(x, y) < 1$ for any $x, y \in B$ (see e.g. [16, Section 6.5]). Finally, a point process $\mathbf{X}$ with constant intensity function on B is said to be homogeneous.

1.2. Projection of a spatial point process

In this work, we sometimes consider projection of spatial point processes. By projection, we mean that we keep a given number of coordinates from the original spatial point process. Such a framework requires that the original point process $\mathbf{X}$ must be defined on a compact set $B \subset \mathbb{R}^d$: otherwise, the configuration of points of the projected point processes may not form a locally finite configuration, as also noticed in the two-dimensional case in [1, p. 17].

This section presents a few notation in this context. Let $I \subseteq \overline{d} := \{1, \ldots, d\}$ with cardinality $|I| = \iota$. Let $B_1, \ldots, B_d$ compact sets of $\mathbb{R}$ and $B = B_1 \times \cdots \times B_d$, denote by B_I its orthogonal projection onto $\mathbb{R}^\iota$. In particular $B_I = \prod_{i \in I} B_i$ with $B = B_{\overline{d}}$. We denote by P_I the orthogonal projection of $\mathbb{R}^d$ onto $\mathbb{R}^\iota$. To ease the reading, we let $B_\ell = B_0$ for $\ell = 1, \ldots, d$ and even to fix ideas let $B_0 = [0,1]$. Thus, $B_I = B_0^\iota = [0,1]^\iota$. For any $x \in B$, we let $x_I = P_I x$ and for a point process $\mathbf{X}$ defined on B, the projected point process $\mathbf{X}_I = P_I \mathbf{X}$ is then defined on B_I. Intensity functions and Laplace functionals for $P_I \mathbf{X}$ can be derived from the corresponding functions and functionals from $\mathbf{X}$, see [19] for more details and Section 2.2 for the particular point process considered in this paper.

1.3. Determinantal point processes

In this section, the class of continuous DPPs is introduced. Again, 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 $K : B \times B \to \mathbb{C}$ if for any $k \geq 1$ its kth order intensity function is given by

$$\rho_{\mathbf{X}}^{(k)} \left(x^{(1)}, \ldots, x^{(k)} \right) = \det \left[K \left(x^{(i)}, x^{(j)} \right) \right]_{i,j=1}^{k} \tag{1.3}$$

and we simply denote by $\mathbf{X} \sim \mathrm{DPP}_B(K)$. Note that K needs to be non-negative definite to ensure $\rho_{\mathbf{X}}^{(k)} \geqslant 0$. Our results rely on the spectral decomposition of K, see (1.5). Therefore, we assume that K is a continuous covariance function. The intensity of $\mathbf{X}$ is given by $\rho_{\mathbf{X}}(x) = K(x, x)$ and its pcf by

$$g_{\mathbf{X}}(x, y) = 1 - \frac{|K(x, y)|^2}{K(x, x)K(y, y)}. \tag{1.4}$$

The popularity of DPPs relies mainly upon (1.3)-(1.4): all moments of $\mathbf{X}$ are explicit and since K is Hermitian, $g_{\mathbf{X}}(x, y) < 1$ for any $x, y \in B$. The kernel K defines an integral operator $\mathcal{K}$ (see *e.g.* [6]) defined for any $f \in L^2(B)$ by

$$\mathcal{K}(f)(x) = \int_B K(x, y)f(y)\mathrm{d}y, \quad x \in B.$$

From Mercer's Theorem [27, Sec. 98], K admits the following spectral decomposition for any $x, y \in B$

$$K(x, y) = \sum_{j \in \mathbb{N}} \lambda_j \phi_j(x)\overline{\phi_j(y)} \tag{1.5}$$

where $\{\phi_j\}_{j \in \mathbb{N}}$ are eigenfunctions associated to $\mathcal{K}$ and form an orthonormal basis of $L^2(B)$, and where $\{\lambda_j\}_{j \in \mathbb{N}}$ are the eigenvalues of $\mathcal{K}$ satisfying $\lambda_j \geq 0$ for any $j \in \mathbb{N}$. We abuse notation in the sequel and refer λ_j's to as the eigenvalues of K.

The existence of a DPP on B with kernel K is ensured if its eigenvalues satisfy $\lambda_j \leqslant 1$ for any $j \in \mathbb{N}$, see *e.g.* [15, Theorem 4.5.5.]. Eigenvalues and eigenfunctions are indexed here by $\mathbb{N}$ in (1.5), but other countable sets could be considered. In particular, the d-dimensional Fourier basis is indexed by $\mathbb{Z}^d$. A DPP $\mathbf{X} \sim \mathrm{DPP}_B(K)$ is said to be homogeneous if K is the restriction on $B \times B$ of a kernel C defined on $\mathbb{R}^d \times \mathbb{R}^d$ which satisfies $C(x, y) = C(o, x - y)$ for any $x, y \in \mathbb{R}^d$ where o is the origin in $\mathbb{R}^d$. In that case, we will refer to K as a stationary kernel and will use the abusive notation $K(x, y) \equiv K(x - y)$.

A kernel K such that $\lambda_j \in \{0, 1\}$ for $j \in \mathbb{N}$ is called a "projection kernel" and the corresponding DPP a "projection DPP". The number of points in B of such a point process is almost surely constant and equal to the number of non-zero eigenvalues of K (see e.g. [17]).

1.4. Why are DPPs interesting for Monte Carlo integration?

The repulsive nature of DPPs can be exploited to generate quadrature points that explore nicely the input space. To see this, let $B \subset \mathbb{R}^d$ be a bounded set,

$f \in L^2(B)$ and $\mathbf{Y}$ a homogeneous point process on B with intensity parameter $\rho_{\mathbf{Y}}$ and pair correlation function $g_{\mathbf{Y}}$ (a similar result would hold in the inhomogeneous case). Campbell's Theorem (1.1) ensures that the estimator

$$\widehat{\mu}(f) = \frac{1}{\rho_{\mathbf{Y}}} \sum_{u \in \mathbf{Y}} f(u) \tag{1.6}$$

is an unbiased estimator of $\mu(f) = \int_B f(u) \mathrm{d}u$ with variance

$$\operatorname{Var}\left[\widehat{\mu}(f)\right] = \frac{1}{\rho_{\mathbf{Y}}} \int_B f(u)^2 \mathrm{d}u + \int_{B^2} (g_{\mathbf{Y}}(u, v) - 1) f(u) f(v) \mathrm{d}u \mathrm{d}v. \tag{1.7}$$

If f is non-negative (or non-positive), Equation (1.7) suggests that using a point processes satisfying $g_{\mathbf{Y}} < 1$ makes the variance smaller than the first term which turns out to be the variance under the Poisson case. It is worth noting that the use of a DPP for this task does not require any sign assumption for f. Indeed, given the fact that $1 - g_{\mathbf{Y}}(u, v) = |K(u, v)|^2 / \rho_{\mathbf{Y}}^2$ and from Mercer's decomposition (1.5), we obtain

$$\operatorname{Var}\left[\widehat{\mu}(f)\right] = \frac{1}{\rho_{\mathbf{Y}}} \int_B f(u)^2 \mathrm{d}u - \frac{1}{\rho_{\mathbf{Y}}^2} \sum_{j,k \in \mathbb{N}} \lambda_j \lambda_k \left| \int_B f(u) \phi_j(u) \overline{\phi_k}(u) \mathrm{d}u \right|^2 \tag{1.8}$$

and so the second term is always negative.

The use of a general DPP, *i.e.* with random number of points, does not seem to be of great interest: thinking $\rho_{\mathbf{Y}} = N$ as a number of points, we claim that the rate of convergence remains the same as in the independent case. This claim is based on the previous empirical study done by [19] which tends to show that the empirical variances of Monte-Carlo integral estimates based on quadrature points from a DPP with, for instance, a Gaussian kernel decrease with rate N^{-1}.

Therefore it seems natural to focus on the subclass of projections DPPs, *i.e.* a class for which the number of points is almost surely constant. Bardenet and Hardy [3] have proposed such an approach using an *ad-hoc* Orthogonal Polynomial Ensemble with N points. This class is a particular inhomogeneous projection DPP on B. As already outlined in the introduction, under the assumption that $f \in C^1(B)$ and that f is compactly supported on some bounded $B' \subset B$, it is proved that a central limit theorem holds for the integral estimator with variance decreasing as $N^{-1-1/d}$. We propose a similar approach, based on a realization of an (N, d)-Dirichlet DPP $\mathbf{X}$. The (N, d)-Dirichlet DPP, detailed in the next section, is a projection DPP based on the Fourier basis. Unlike, the point process proposed by [3], this DPP has the advantage to be homogeneous and its projections $\mathbf{X}_I$ for $I \subseteq \{1, \ldots, d\}$ are fully characterized (see Section 2.2). Finally, an advantage of our approach is that we do not require that f is continuously differentiable. We only assume that f belongs to some Sobolev space with low regularity parameter, see Section 3 and in particular Theorem 3.1 for more details.

2. The (N, d)-Dirichlet DPP and its projections

2.1. The (N, d)-Dirichlet DPP

Let us consider the Fourier basis in $B = [0,1]^d$ defined for any $j \in \mathbb{Z}^d$, $x \in B$ by $\phi_j^{(d)}(x) = \mathrm{e}^{2\mathrm{i}\pi j^\top x}$. Given a vector of d positive integers $n = (n_i)_{i=1\ldots d}$, we construct the following kernel

$$K(x, y) = \sum_{j \in E_N} \phi_j^{(d)}(x)\overline{\phi_j^{(d)}(y)} = \sum_{j \in E_N} \mathrm{e}^{2\mathrm{i}\pi j^\top (x-y)} \tag{2.1}$$

where

$$E_N = E_1 \times \cdots \times E_d \quad \text{and for } i = 1 \ldots d, \quad E_i = \{0, 1, \ldots, n_i - 1\}. \tag{2.2}$$

Thus, E_N is the rectangular subset of $\mathbb{Z}^d$ with cardinality N which identifies eigenvalues that are all equal to 1. Due to the invariance by translation of the Fourier basis, the kernel (2.1) is a homogeneous kernel. This construction implies that for any $x, y \in [0,1]^d$, $K(x - y) = \prod_{i=1}^d K_i(x_i - y_i)$ where the K_i's are one-dimensional stationary kernels defined for any $x_i, y_i \in [0,1]$ by $K_i(x_i - y_i) = \sum_{j=0}^{n_i-1} \mathrm{e}^{2\mathrm{i}\pi j(x_i-y_i)}$. We point out that defining E_i as a block of successive n_i frequencies (*e.g.* frequencies centered around 0), would lead to the same DPP [15, Remark 4, p.48]. In particular, if n_i is odd, we could consider $E_i = \{-\lfloor n_i/2 \rfloor, \ldots, \lfloor n_i/2 \rfloor\}$, which leads to the standard Dirichlet kernel, see *e.g.* [37]. This justifies the name Dirichlet DPP for this stochastic process.

Such a DPP, which produces almost surely $N = \prod_{i=1}^d n_i$ points in B, will be referred to as an (N, d)-Dirichlet DPP. We could wonder why we impose E_N to be a rectangular subset of $\mathbb{Z}^d$ instead of, for instance, the graded lexicographic order used by Bardenet and Hardy [3]. As seen in the next section, the rectangular nature of E_N allows us to characterize the distribution of $\mathbf{X}_I$ for any $I \subseteq \overline{d}$.

Let us add that, when $d = 1$, the kernel K corresponds to the Fourier approximation [17] of the one-dimensional sine-kernel $\sin(\pi n_i t)/\pi t$ ($t \in (0, 1)$), which takes its origins in the joint distribution of the eigenvalues (called the Weyl measure) of a unitary matrix. Asymptotic results involving one-dimensional linear functionals from the sine-kernel appear in several papers, see e.g. [35]. The present paper provides therefore an extension to the d-dimensional case and a more thorough treatment of the statistical application to Monte Carlo integration.

The (N, d)-Dirichlet DPP is a homogeneous point process producing exactly N points. Since we do not want to take any advantage from the function to be integrated (and we could potentially be interested in estimating several integrals), it is very natural to use a homogeneous model. We come back to this question in Section 6.

2.2. Projections of an (N, d)-Dirichlet kernel

An (N, d)-Dirichlet kernel (2.1) can be written as the product of d one-dimensional kernels. More precisely, for any $I \subset \overline{d}$, by denoting $N_I = \prod_{i \in I} n_i$ and $N_{I^c} = N/N_I$, the (N, d)-Dirichlet kernel can always be written as

$$K(x - y) = K_I(x_I - y_I)K_{I^c}(x_{I^c} - y_{I^c}) \tag{2.3}$$

where K_I (resp. K_{I^c}) is the (N_I, ι)-Dirichlet kernel (resp. $(N_{I^c}, d - \iota)$-Dirichlet kernel). Projected point processes $\mathbf{X}_I$ from models with kernels satisfying (2.3) have been studied and characterized in [19]. In particular, for an (N, d)-Dirichlet DPP we have the following result.

Proposition 2.1. *Let* $\mathbf{X}$ *be an* (N, d)*-Dirichlet DPP on* B*, let* $I \subseteq \{1, \dots, d\}$*, then* $\mathbf{X}_I$ *is an* $(-1/N_{I^c})$*-DPP on* B_I *with kernel* $N_{I^c}K_I$*, i.e.* $\mathbf{X}_I \sim (-1/N_{I^c})$*-* $\mathrm{DPP}_{B_I}(N_{I^c}K_I)$*. In particular,* $\mathbf{X}_I$ *has (obviously)* N *points and* k*-th order intensity*

$$\rho_{\mathbf{X}_I}\left(x^{(1)}, \dots, x^{(k)}\right) = \det_{-1/N_{I^c}}\left[N_{I^c}K_I\left(x^{(i)}, x^{(j)}\right)\right]_{i,j=1}^k$$

for any pairwise distinct $x^{(1)}, \dots, x^{(k)} \in B_I$*. Its pcf is therefore given, for any pairwise distinct* $x, y \in B_I$*, by*

$$g_{\mathbf{X}_I}(x, y) = 1 - \frac{|K_I(x, y)|^2}{NN_I}. \tag{2.4}$$

In the above result, the notation $\det_\alpha$ stands for an α determinant, see e.g. [32] for details on such quantities and for general properties of (α)-DPPs. Proposition 2.1 therefore proposes a full characterization of the distribution of $\mathbf{X}_I$. This result is not directly used in our paper (except in Appendix B where we propose an alternative proof to our main result). However, its main consequence (for the paper) is that the pcf of $\mathbf{X}_I$ is bounded by 1 (see (2.4)). Therefore, for any I, $\mathbf{X}_I$ remains in the class of repulsive point patterns.

We point out that the DPP proposed by [3] does not satisfy the general assumptions of [19]. Therefore for this DPP, the distribution of $\mathbf{X}_I$ is not explicit and it is unclear whether $\mathbf{X}_I$ remains repulsive or not.

3. Numerical integration with Dirichlet kernel

3.1. Objective

In this section, we study the use of specific DPPs for Monte Carlo integration. To this end, we use notation introduced in Section 1.2. Our objective is to estimate any ι-dimensional integral, for $1 \leq \iota \leq d$, using a Monte Carlo approach and using the same quadrature points. More precisely, let $d \geq 1$, $I \subset \overline{d} = \{1, \dots, d\}$ with cardinality $\iota = |I|$ and let $f_I : B_I \to \mathbb{R}$ be a measurable function on

$B_I = [0,1]^\iota$, such that $f_I \in L^2(B_I)$. More assumptions on f_I will be given later. We intend to estimate

$$\mu(f_I) = \int_{B_I} f_I(u)\mathrm{d}u$$

using the projection onto B_I of $\mathbf{X}$ an (N, d)-Dirichlet DPP on B. In particular, we estimate $\mu(f_I)$ by

$$\widehat{\mu}_N(f_I) = \frac{1}{N} \sum_{u \in \mathbf{X}_I} f_I(u) = \frac{1}{N} \sum_{j=1}^{N} f_I((u_j)_I) \tag{3.1}$$

where $\mathbf{X} = \{u_1, \ldots, u_N\}$ is an (N, d)-Dirichlet DPP on B and where we remind the notation $(u_j)_I = P_I u_j$ for any $u_j \in B$. In the following, we study asymptotic properties for $\widehat{\mu}_N(f_I)$.

In this paper, we have chosen to focus on integrals on $[0,1]^\iota$ for simplicity. It can straightforwardly be extended to rectangles. Indeed, let $a = (a_i)_{i=1,\ldots,d}, b = (b_i)_{i=1,\ldots,d} \in \mathbb{R}^d$ such that $a_i < b_i$, $i = 1, \ldots, d$, let $R = [a_1,b_1] \times \cdots \times [a_d,b_d]$. Denoting $R_I = \prod_{i \in I} [a_i,b_i]$, an estimate of $\int_{R_I} f_I(u)\mathrm{d}u$ where now $f_I \in L^1(R_I)$ is simply given by

$$\widehat{\mu}_N(f_I) = \prod_{i \in I}(b_i - a_i) \left\{ \frac{1}{N} \sum_{i=1}^{N} f_I\left(a_I + (b_I - a_I)(u_j)_I\right) \right\}. \tag{3.2}$$

We introduce two additional fundamental pieces of notation induced by the choice of the Fourier basis and used in our results. Let $f_I \in L^2(B_I)$. The notation $\widehat{f}_I(j)$ for $j \in \mathbb{Z}^\iota$ stands for the jth Fourier coefficient, *i.e.*

$$\widehat{f}_I(j) = \int_{[0,1]^\iota} f_I(u)\mathrm{e}^{-2\mathrm{i}\pi j^\top u}\mathrm{d}u. \tag{3.3}$$

Finally, we define the space $\mathcal{H}^s(B_I)$ as the isotropic (with respect to the sup norm) Sobolev space with index $s \geq 0$ of square integrable periodic functions by

$$\mathcal{H}^s(B_I) = \left\{ f_I \in L^2_{\mathrm{per}}(B_I), : \sum_{j \in \mathbb{Z}^\iota} (1 + \|j\|_\infty)^{2s} |\widehat{f}_I(j)|^2 < \infty \right\} \tag{3.4}$$

where $L^2_{\mathrm{per}}(B_I)$ is the set of square integrable periodic functions. By f_I periodic, we mean that for any $i \in I$

$$f_I(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_\iota) = f_I(x_1, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_\iota).$$

Note that, if f_I does not satisfy this condition, we can for example consider the function $g_I(x_1, \ldots, x_\iota) = f_I(|x_1|, \ldots, |x_\iota|)$ which will satisfy this periodic condition on $[-1, 1]^\iota$, and $\widehat{\mu}_N(f_I) = 2^{-\iota}\widehat{\mu}_N(g_I)$.

In the introduction, we mention that several Monte-Carlo integration methods (stratified-based methods, scrambled nets, *etc*) require some (strong) regularity assumptions of the integrands. In connection with the Sobolev space $\mathcal{H}^s(B_I)$, we remind that (see *e.g.* Robinson [29]) for any $\varepsilon > 0$

$$\widetilde{\mathcal{H}}^{3/2+\varepsilon}(B_I) \ \subset \ C^1_{\mathrm{per}}(B_I) \ \subset \ \mathcal{H}^1(B_I). \tag{3.5}$$

where

$$C^1_{\mathrm{per}}(B_I) = \left\{ f_I \text{ periodic}, f_I \in C^1(B_I) \right\}$$
$$\widetilde{\mathcal{H}}^{3/2+\varepsilon}(B_I) = \left\{ f_I \in L^2_{\mathrm{per}}(B_I) : \widehat{f}_I(j) = \mathcal{O}(\|j\|_\infty^{-2-\varepsilon}) \right\}, \quad \varepsilon > 0.$$

(Of course, $\mathcal{H}^{3/2+\varepsilon}(B_I)$ and $\widetilde{\mathcal{H}}^{3/2+\varepsilon}(B_I)$ are very similar). In the same way, any function f_I periodic and d times continuously differentiable belongs to $\mathcal{H}^d(B_I)$.

Our main results expressed by Theorem 3.1 and Corollary 3.1, have the interest to focus on functions $f_I \in \mathcal{H}^s(B_I)$ with small regularity parameter $s > 0$, thus, according to (3.5), potentially to non-differentiable functions. In particular, we show that a CLT holds as soon as $s > 1/2$.

Finally, let us justify why we focus on periodic functions on B_I. For a non periodic function which is at least twice continuously differentiable it is a known fact that $|\widehat{f}_I(j)| = \mathcal{O}(\|j\|_\infty^{-1})$. So, the summability condition in (3.4) for such a smooth function would be fulfilled only for $s < 1/2$, a situation where no CLT is available (even if f_I is infinitely differentiable).

3.2. Case $\iota = d$

We first consider the d-dimensional case. Let $f \in L^2_{\mathrm{per}}(B)$ with $B = [0,1]^d$. The first result shows how the variance given in (1.8) relates to the Fourier coefficients of f in the case of an (N, d)-Dirichlet DPP.

Proposition 3.1. *Let $\mathbf{X}$ be an (N,d)-Dirichlet DPP and $f \in L^2_{\mathrm{per}}(B)$. For any $j \in \mathbb{Z}^d$, we let $c_j(f)$ denote the jth Fourier coefficient of f. Then, $\widehat{\mu}_N(f)$ given by (3.1) is an unbiased estimator of $\mu(f) = \int_B f(u)\mathrm{d}u$ with variance given by*

$$\mathrm{Var}\left[\widehat{\mu}_N(f)\right] = \frac{1}{N} \sum_{j \in \mathbb{Z}^d} \left|\widehat{f}(j)\right|^2 - \frac{1}{N^2} \sum_{j,k \in E_N} \left|\widehat{f}(j-k)\right|^2 \tag{3.6}$$

$$= \frac{1}{N} \sum_{j \in \mathbb{Z}^d} \left|\widehat{f}(j)\right|^2 - \frac{1}{N^2} \sum_{j \in F_N} \left[\prod_{i=1}^{d} (n_i - |j_i|)\right] \left|\widehat{f}(j)\right|^2 \tag{3.7}$$

where $F_N = \{j \in \mathbb{Z}^d : |j_i| \leq n_i - 1, \ i = 1, \ldots, d\}$.

This simple form of the variance of $\widehat{\mu}_N(f)$ invites us to study its asymptotic behavior as $N \to \infty$. Given the fact that $N = \prod_i n_i$, we require a specific asymptotic. For $i = 1, \ldots, d$, we assume that $(N_\nu)_{\nu \geq 1}$ and $(n_{i,\nu})_{\nu \geq 1}$ are integer

sequences indexed by some $\nu \geq 1$. We assume that each sequence $n_{i,\nu}$ tends to ∞ as $\nu \to \infty$ and that there exist $\kappa_i > 0$ for $i = 1, \ldots, d$ such that

$$\lim_{\nu \to \infty} n_{i,\nu} N_\nu^{-1/d} = \kappa_i. \tag{3.8}$$

For the sake of conciseness, we skip the dependence in ν. Similarly, when we write $N \to \infty$, we implicitly assume (3.8).

We can now obtain the following asymptotic behavior of the variance of $\widehat{\mu}_N(f)$ and for some values of s a central limit theorem. Regarding this asymptotic normality, when $d = 1$, as mentioned in Section 2, **X** corresponds, up to a normalization, to the joint distribution of the eigenvalues of a unitary matrix. Linear statistics for such a DPP have been deeply studied in the literature, see e.g. [35] and the references therein. When $d > 1$, Theorem 3.1 (iii) is therefore original, and relies upon Soshnikov [36, Theorem 1].

Theorem 3.1. *Consider the asymptotic framework* (3.8) *and assume that* $f \in \mathcal{H}^s(B)$. *Then, we have the following statements.*

(i) If $s \in (0,1/2)$, *then as* $N \to \infty$

$$\mathrm{Var}\left[\widehat{\mu}_N(f)\right] = \mathcal{O}\left(N^{-1-\frac{2s}{d}}\right).$$

(ii) If $s \geq 1/2$ *for* $d = 1$ *or* $s > 1/2$ *for* $d > 1$, *then*

$$\lim_{N \to \infty} N^{1+1/d}\mathrm{Var}\left[\widehat{\mu}_N(f)\right] = \sigma^2(f) = \sum_{j \in \mathbb{Z}^d} \left(\sum_{i=1}^{d} \frac{|j_i|}{\kappa_i}\right)\left|\widehat{f}(j)\right|^2 \tag{3.9}$$

where $\widehat{f}(j)$ *is given by* (3.3).

(iii) If $s \geq 1/2$ *for* $d = 1$ *or* $s > 1/2$ *for* $d > 1$ *and, if in addition,* $\|f\|_\infty < \infty$, *then as* $N \to \infty$,

$$\sqrt{N^{1+1/d}}\left(\widehat{\mu}_N(f) - \mu(f)\right) \to N(0, \sigma^2(f)) \tag{3.10}$$

in distribution.

Let us rephrase Theorem 3.1 (i): if $f \in \mathcal{H}^s(B)$ for some $s > 0$, then necessarily $\mathrm{Var}[\widehat{\mu}_N(f)] = o(N^{-1})$. That is, the variance of the estimator proposed decreases to 0 faster than the standard Monte Carlo estimator, as soon as $\left|\widehat{f}(j)\right| = \mathcal{O}(\|j\|_\infty^{-\varepsilon})$ for some $\varepsilon > 0$, which is a very weak assumption.

Theorem 3.1 (ii) shows that the variance $\sigma^2(f)$ can be approximated by $\widehat{\sigma}_N^2(f)$ given by

$$\widehat{\sigma}_N^2(f) = N^{1/d} \sum_{j \in F_N} \left(\sum_{i=1}^{d} \frac{|j_i|}{n_i}\right)\left|\widehat{f}(j)\right|^2 \tag{3.11}$$

which has the interest to avoid the constants κ_i defined in (3.8). We point out that (3.11) does not provide an estimate of $\sigma^2(f)$. We come back to this question in Section 6.

3.3. Discussion on $\sigma^2(f)$

In this section, we would like to discuss how the value of $\sigma^2(f)$ varies with d and how this variance compares to standard Monte-Carlo method and the stratified-based Monte-Carlo method (which are important alternatives for which a central limit theorem is available). For the sake of simplicity, we set in this section $\kappa_i = 1$. For $f \in \mathcal{H}^s([0,1]^d)$ with $s \geq 0$, we denote by

$$\|f\|^2_{\mathcal{H}^s([0,1]^d)} = \sum_{j \in \mathbb{Z}^d} \|j\|^{2s}_{2s} |\hat{f}(j)|^2$$

where $\|j\|^p_p = \sum_{l=1}^d |j_l|^p$. With such notation and recalling that $\kappa_i = 1$ here we get $\sigma^2(f) = \|f\|^2_{\mathcal{H}^{1/2}([0,1]^d)}$. The standard Monte-Carlo approach (resp. the stratified-based Monte-Carlo method, see [26]) assumes that $f \in L^2([0,1]^d)$ (resp. f is continuously differentiable on $[0,1]^d$). The corresponding estimator has variance descreasing as N^{-1} (resp. $N^{-1-2/d}$) with asymptotic constants given by

$$\sigma^2_{\mathrm{std}}(f) = \|f\|^2_{L^2([0,1]^d)} - \mu(f)^2 = \|f - \mu(f)\|^2_{\mathcal{H}^0([0,1]^d)}$$
$$\sigma^2_{\mathrm{strat}}(f) = \| \, \|\nabla f\|_2 \, \|^2_{L^2([0,1]^d)} = 4\pi^2 \|f\|^2_{\mathcal{H}^1([0,1]^d)}.$$

The latter expression ensues from Parseval's identity and standard expression of Fourier coefficients for continuously differentiable functions.

To compare these three constants in terms of d, we consider two particular cases for f:

(i) f as a product of one-dimensional functions: $\exists f_0 \in L^2([0,1])$ such that

$$f(x) = \prod_{i=1}^d f_0(x_i), \qquad x \in [0,1]^d. \tag{3.12}$$

(ii) f as a sum of one-dimensional functions: $\exists f_0 \in L^2([0,1])$ such that

$$f(x) = \sum_{i=1}^d f_0(x_i), \qquad x \in [0,1]^d. \tag{3.13}$$

We are now able to present the following result.

Proposition 3.2.
(i) Let f be defined as (3.12) *and satisfying the appropriate assumptions of the considered method, then*

$$\sigma^2(f) = d \, \|f_0\|^2_{\mathcal{H}^{1/2}([0,1])} \, \|f_0\|^{2d-2}_{\mathcal{H}^0([0,1])}$$
$$\sigma^2_{\mathrm{std}}(f) = \|f_0\|^{2d}_{\mathcal{H}^0([0,1])} - \mu(f_0)^{2d}$$
$$\sigma^2_{\mathrm{strat}}(f) = 4\pi^2 d \, \|f_0\|^2_{\mathcal{H}^1([0,1])} \, \|f_0\|^{2d-2}_{\mathcal{H}^0([0,1])}.$$

(ii) Let f be defined as (3.13) *and satisfying the appropriate assumptions of the considered method, then*

$$\sigma^2(f) = d\,\sigma^2(f_0) = d\,\|f_0\|_{\mathcal{H}^{1/2}([0,1])}^2$$
$$\sigma_{\text{std}}^2(f) = d\sigma_{\text{std}}^2(f_0) = d\,\|f_0 - \mu(f_0)\|_{\mathcal{H}^0([0,1])}^2$$
$$\sigma_{\text{strat}}^2(f) = d\,\sigma_{\text{strat}}^2(f_0) = 4\pi^2 d\,\|f_0\|_{\mathcal{H}^1([0,1])}^2.$$

Proposition 3.2 shows that when f is given as the product (resp. sum) of one-dimensional functions, then $\sigma^2(f)$ varies exponentially (resp. linearly) with the dimension d. This result points out that the constants for the standard Monte-Carlo and stratified-based Monte-Carlo methods vary along the same lines.

3.4. Case $I \subset \overline{d}$

We now consider the situation where we estimate $\mu(f_I)$ (on B_I) based on $\{u_1, \ldots, u_N\}$ which is an (N, d)-Dirichlet DPP on B. In this section, we naturally assume that $d > 1$. The interest of the contruction of the Dirichlet-DPP is revealed by Corollary 3.1 which, briefly, states that Theorem 3.1 can be applied to functions of the form $f_I^\uparrow(x) = f_I(x_I)\mathbb{1}\,[x_{I^c} \in B_{I^c}]$, $x \in B$.

Corollary 3.1. *Let $d > 1$ and $I \subset \{1, \ldots, d\}$ with cardinality $\iota > 0$. Consider the asymptotic framework* (3.8) *and assume that $f_I \in \mathcal{H}^s(B_I)$, then the following statements hold.*

(i) If $s \in (0,1/2)$, then as $N \to \infty$

$$\text{Var}\,[\widehat{\mu}_N(f_I)] = \mathcal{O}\left(N^{-1-\frac{2s}{d}}\right).$$

(ii) If $s > 1/2$, then

$$\lim_{N\to\infty} N^{1+1/d}\text{Var}\,[\widehat{\mu}_N(f_I)] = \sigma^2(f_I) \tag{3.14}$$

where

$$\sigma^2(f_I) = \sum_{j \in \mathbb{Z}^\iota} \left(\sum_{i \in I} \frac{|j_i|}{\kappa_i}\right) \left|\widehat{f}_I(j)\right|^2 \tag{3.15}$$

and where $\widehat{f}_I(j)$ is given by (3.3).
(iii) If $s > 1/2$ and f_I is bounded, then as $N \to \infty$

$$\sqrt{N^{1+1/d}}\,(\,\widehat{\mu}_N(f_I) - \mu(f_I)\,) \to N\left(0, \sigma^2(f_I)\right) \tag{3.16}$$

in distribution.

The asymptotic constant $\sigma^2(f_I)$ can still be approximated by (3.11). The proof of this result is a straightforward consequence of Theorem 3.1. Another approach using the fact that $\mathbf{X}_I$ is distributed as an α-DPP is proposed in Appendix B.

To rephrase Corollary 3.1, we can estimate $\mu(f_I)$ for any ι-dimensional function f_I with the rate of convergence $\sqrt{N^{1+1/d}}$ (if $f_I \in \mathcal{H}^s(B_I)$ with $s > 1/2$), which corresponds to the rate of convergence from the dimension where the points were generated. This is the price to pay to be able to estimate any function in any dimension. Of course, if one knows beforehand that we would like to estimate only, say one-dimensional integrals, one could use the decomposition $N = N \times 1 \times \cdots \times 1$ for the set E_N of eigenvalues, see (2.2). Using such a DPP (in dimension $d > 1$), we claim that any one-dimensional integral could be estimated with a rate of convergence $\sqrt{N^2}$ which would be a gain with respect to $\sqrt{N^{1+1/d}}$. However, the resulting DPP would be catastrophic to estimate $\mu(f_I)$ for any I such that $\iota > 1$.

3.5. Multivariate central limit theorem

We can combine Theorem 3.1 and Corollary 3.1 to obtain a multivariate version of the central limit theorem.

Corollary 3.2. *Let $p \geq 1$. For any $\ell = 1, \ldots, p$, let $I_\ell \subseteq \{1, \ldots, d\}$ with $\iota_\ell = |I_\ell|$ and assume that $f_{I_\ell} \in \mathcal{H}^s(B_{I_\ell})$ with $s \geq 1/2$ if $d = 1$ or $s > 1/2$ if $d > 1$. Let $\widehat{\mu}_{N,p} = \left(\widehat{\mu}_N(f_{I_1}), \ldots, \widehat{\mu}_N(f_{I_p})\right)^\top$ and $\mu_p = \left(\mu(f_{I_1}), \ldots, \mu(f_{I_p})\right)^\top$. Then, under the asymptotic framework (3.8), $\widehat{\mu}_{N,p}$ is an unbiased estimator of μ_p and as $N \to \infty$*

$$\sqrt{N^{1+1/d}}\left(\widehat{\mu}_{N,p} - \mu_p\right) \to N(0, \mathbf{\Sigma}_p)$$

in distribution, where $\mathbf{\Sigma}_p$ is the (p, p) Hermitian matrix with entries

$$(\mathbf{\Sigma}_p)_{\ell\ell'} = \sum_{j \in \mathbb{Z}^d} \left(\sum_{i=1}^d \frac{|j_i|}{\kappa_i}\right) \widehat{f}_{I_\ell}^\uparrow(j) \overline{\widehat{f}_{I_{\ell'}}^\uparrow(j)} \tag{3.17}$$

where $f_{I_\ell}^\uparrow(x) = f_{I_\ell}(x_I)\mathbb{1}\left[x_{I_\ell^c} \in B_{I_\ell^c}\right], x \in B$.

4. Simulation study

We propose now a simulation study to illustrate the results. We first consider the setting of Theorem 3.1 (ii)-(iii) and Corollary 3.1 (ii)-(iii), *i.e.* for the case $s > 1/2$ (when $d > 1$) and and for the situation $s < 1/2$.

4.1. Case $s > 1/2$, illustration of Theorem 3.1 (ii)-(iii)

We consider three different functions with different regularity properties:

- Bump function

$$f_{\text{bump}}(x) = \prod_{i=1}^{d} \frac{\varphi_{\text{bump}}(x_i)}{\int_0^1 \varphi_{\text{bump}}(t)\mathrm{d}t}, \qquad \varphi_{\text{bump}}(t) = \exp\left(-\frac{0.1}{t(1-t)}\right). \quad (4.1)$$

- Sum of cosines

$$f_{\text{mixcos}}(x) = \frac{1}{d}\sum_{i=1}^{d} \frac{\varphi_{\text{mc}}(x_i)}{\int_0^1 \varphi_{\text{mc}}(t)\mathrm{d}t}, \quad \varphi_{\text{mc}}(t) = 0.1\left|\cos(5\pi(t-1/2))\right| + (t-1/2)^2.$$

$$(4.2)$$

- Product of cosines

$$f_{\text{mixcosprod}}(x) = \prod_{i=1}^{d} \frac{\varphi_{\text{mc}}(x_i)}{\int_0^1 \varphi_{\text{mc}}(t)\mathrm{d}t} \qquad (4.3)$$

- Normalized L^γ-norm: let $\gamma > 0$.

$$f_\gamma(x) = \frac{1}{d}\sum_{i=1}^{d} \frac{\varphi_\gamma(x_i)}{\int_0^1 \varphi_\gamma(t)\mathrm{d}t}, \quad \varphi_\gamma(t) = |t-1/2|^\gamma. \qquad (4.4)$$

It is worth mentioning that f_{bump} is infinitely continuously differentiable, so $f_{\text{bump}} \in \mathcal{H}^s$, for any $s > 0$. The function f_{mixcos} is a non-differentiable (with 4^d singularity points) which satisfies $\widehat{f}_{\text{mixcos}}(j) = \mathcal{O}(\|j\|_\infty^{-2})$, so $f_{\text{mixcos}} \in \mathcal{H}^s(B)$ for any $s < 3/2$. $f_{\text{mixcosprod}}$ is also a non-differentiable function, which satisfies $\widehat{f}_{\text{mixcosprod}}(j) = \mathcal{O}(\prod_{i=1}^d \|j_i\|^{-2})$, whereby it can be deduced that $f_{\text{mixcosprod}} \in \mathcal{H}^s(B)$ for any $s < 3/2$. Finally, f_γ is also a non-differentiable square integrable periodic function, which satisfies $\widehat{f}_\gamma(j) = \mathcal{O}(\|j\|_\infty^{-1-\gamma})$. Hence, $f_\gamma \in \mathcal{H}^s(B)$ for any $s < 1/2 + \gamma$. In the following, we consider the cases $\gamma = 0.25$ and $\gamma = 0.75$. These five test functions are depicted for $d = 2$ in Figure 1.

We perform the following experiment. For $N = 50, 100, 150, \ldots,$ $500, 600, \ldots, 1000$ (*i.e.* 15 values for N) and for $d = 1, \ldots, 6, 10$, we generate 2500 realizations of the (N, d)-Dirichlet DPP. The factorization $N = \prod_{i=1}^{d} n_i$ is set such that the fluctuation of the n_i's is minimized. For example, for $N = 100$ and $d = 2$, we set $n = (10, 10)$ while when $d = 6$ we choose $n = (5, 5, 2, 2, 1, 1)$.

We use the simulation algorithm provided by Hough et al. [14]. Basically, it relies upon a Gram-Schmidt orthogonalization of N vectors with dimension N, with a cost of order N^3, and a costly rejection sampling step for the simulation of each point location. Sampling DPPs with large N or when $d > 4$ is very time consuming using the standard R package spatstat [2]. Therefore, we have reimplemented an R package based on C++ functions (available at https://github.com/AdriMaz/rcdpp/). We perform our experiments in very reasonable computing time, with a basic laptop (2,3 GHz Intel Core i5 processor, 8 Go (2133 MHz DDR4) of RAM). For example, sampling a DPP with $N = 1000$ in dimension 6 can be performed within a reasonable time (approximately one minute).
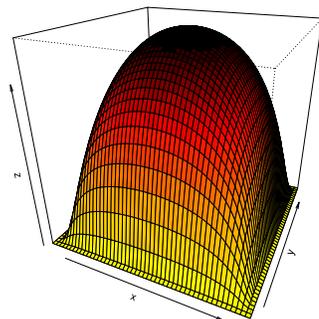
$$f_{\text{bump}}(x)$$
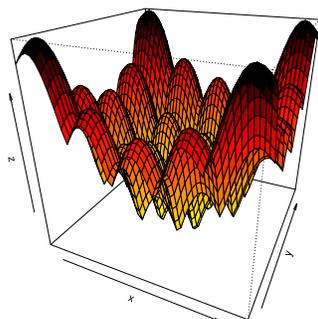
$$f_{\text{mixcos}}(x)$$
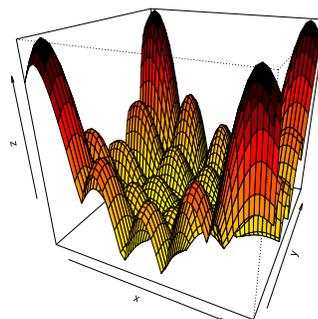
$$f_{\text{mixcosprod}}(x)$$
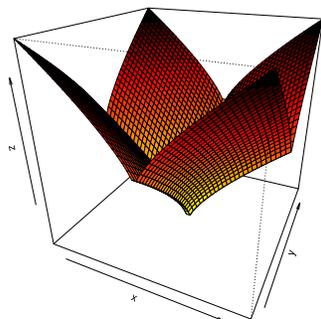
$$f_\gamma(x) \text{ for } \gamma = 0.75$$
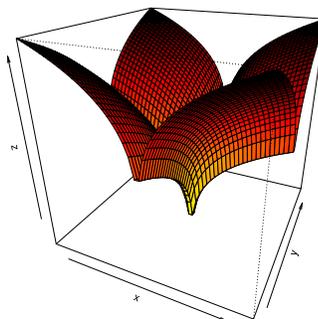
$$f_\gamma(x) \text{ for } \gamma = 0.25$$

FIG 1. *Test functions considered in the simulation study, given by* (4.1)-(4.4) *depicted in dimension* $d = 2$.

For each d, each function and each replication of the point pattern, we evaluate the estimator $\widehat{\mu}_N(f_I)$ given by (3.1) with $I = \overline{d}$. To visualize the rate of convergence of the variance, we perform a linear regression of the logarithm of empirical variances in terms of $\log(N)$. According to Theorem 3.1 (ii), the expected slope is $-1-1/d$. For each test function, d and N, we also test the normality of estimates using the Shapiro-Wilk test [31], after adjusting the p-values using Holm procedure [13] for each function and each d. Results are reported in Figure 2. The fourth columns of the summary tables expose the usual Student-t confidence interval for the slopes. The values of N for which the normality assumptions has been rejected, *i.e.* for which adjusted p-value is smaller than 0.05, are represented by crosses instead of regular dots. We still made the arbitrary choice to keep them when performing the regression lines. We have also arbitrarily translated the curves to focus the interpretation on the slopes. Thus in Figures 2-7 (as well as figures presented in Appendix C) the y-axis has no meaning.

Most of results are in a clear agreement with our theoretical result. For small values of d, the asymptotic normality is not rejected even for small sample size N, and the slope of the logarithm of empirical variances is very close to $-1-1/d$.

For $d = 10$, the theoretical results are hardly recovered. This is related to the values of N considered: the values of the n_i's are very small (several of them are actually equal to 1).

The structure of the integrands affects the quality of empirical results: for functions f_{bump} and $f_{\text{mixcosprod}}$, normality requires higher values of N. The expected slope is also not always included in the confidence intervals. However, this might be related to usual observations when estimating integral of functions defined as (3.12), as illustrated in Section 5.

4.2. Case $s > 1/2$, illustration of Corollary 3.1 (ii)-(iii)

We perform similar experiments. We set the dimension to $d = 6$. For each configuration $\{x^{(1)}, \ldots, x^{(N)}\}$ of $(N, 6)$-Dirichlet DPPs, we evaluate the estimator $\widehat{\mu}_N(f_I)$ given by (3.1) with $I \subset \overline{d}$ and $|I| = \iota = 1, \ldots, 6$. In other words, we use 6-dimensional configurations of points to estimate integrals of ι-dimensional integrands. Let us precise that a single realization $(N, 6)$-Dirichlet is used for each value of ι. However, the directions kept when projecting are chosen randomly. Results are illustrated in Figure 3, as for the previous case. The conclusions are quite similar to the previous case: the points remain nicely aligned along the regression lines and confidence intervals are in agreement with theoretical slopes, all equal to $-1 - 1/6 \approx -1.17$ in this situation. It seems that normality is hardly hinted when a design is projected on low dimensional space: most of the values of N are rejected by the Shapiro-Wilk tests for $\iota = 1, 2$. We also conduct a similar experiment starting with realizations of the Dirichlet DPP in dimension $d = 10$, see Figure 4. Similar comments can be done from this situation.
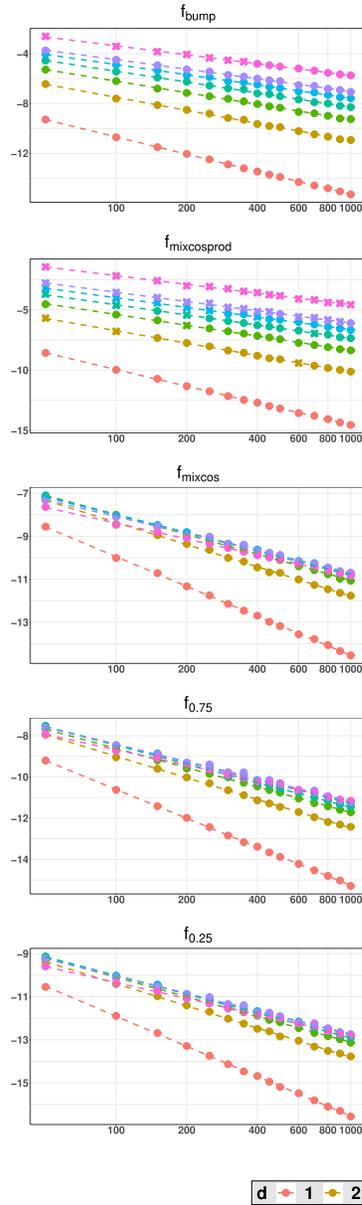
d	Expected	Estimation	Conf. interval
1	−2.00	−2.00	[−2.01 ; −1.98]
2	−1.50	−1.51	[−1.54 ; −1.47]
3	−1.33	−1.34	[−1.37 ; −1.31]
4	−1.25	−1.25	[−1.28 ; −1.23]
5	−1.20	−1.18	[−1.2 ; −1.15]
6	−1.17	−1.11	[−1.13 ; −1.08]
10	−1.10	−1.03	[−1.05 ; −1.02]

d	Expected	Estimation	Conf. interval
1	−2.00	−1.99	[−2.01 ; −1.98]
2	−1.50	−1.48	[−1.5 ; −1.46]
3	−1.33	−1.29	[−1.31 ; −1.27]
4	−1.25	−1.21	[−1.23 ; −1.19]
5	−1.20	−1.15	[−1.18 ; −1.12]
6	−1.17	−1.11	[−1.14 ; −1.08]
10	−1.10	−1.05	[−1.08 ; −1.01]

d	Expected	Estimation	Conf. interval
1	−2.00	−1.99	[−2.01 ; −1.97]
2	−1.50	−1.49	[−1.52 ; −1.46]
3	−1.33	−1.32	[−1.35 ; −1.29]
4	−1.25	−1.23	[−1.25 ; −1.2]
5	−1.20	−1.19	[−1.22 ; −1.15]
6	−1.17	−1.13	[−1.17 ; −1.1]
10	−1.10	−1.06	[−1.08 ; −1.04]

d	Expected	Estimation	Conf. interval
1	−2.00	−2.02	[−2.04 ; −2.01]
2	−1.50	−1.51	[−1.54 ; −1.49]
3	−1.33	−1.35	[−1.37 ; −1.32]
4	−1.25	−1.30	[−1.32 ; −1.28]
5	−1.20	−1.24	[−1.28 ; −1.2]
6	−1.17	−1.20	[−1.25 ; −1.14]
10	−1.10	−1.08	[−1.1 ; −1.06]

d	Expected	Estimation	Conf. interval
1	−2.00	−2.01	[−2.03 ; −1.99]
2	−1.50	−1.48	[−1.5 ; −1.46]
3	−1.33	−1.31	[−1.33 ; −1.29]
4	−1.25	−1.26	[−1.27 ; −1.24]
5	−1.20	−1.19	[−1.22 ; −1.16]
6	−1.17	−1.16	[−1.2 ; −1.12]
10	−1.10	−1.07	[−1.09 ; −1.04]

| d | → 1 | → 2 | → 3 | → 4 | → 5 | → 6 | → 10 |

FIG 2. *Summary of experiments in which integrals of d-dimensional functions are estimated using an (N, d)-Dirichlet DPP. A ● (resp. ×) indicates that the adjusted p-value of the Shapiro-Wilk test is not smaller (resp. smaller) than 5%*
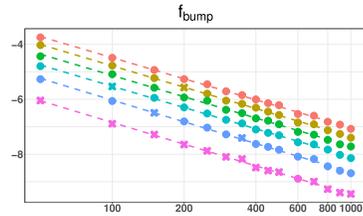
ι	Expected	Estimation	Conf. interval
6	−1.17	−1.11	[−1.13 ; −1.08]
5	−1.17	−1.12	[−1.14 ; −1.1]
4	−1.17	−1.12	[−1.15 ; −1.09]
3	−1.17	−1.12	[−1.14 ; −1.09]
2	−1.17	−1.14	[−1.17 ; −1.11]
1	−1.17	−1.14	[−1.17 ; −1.11]
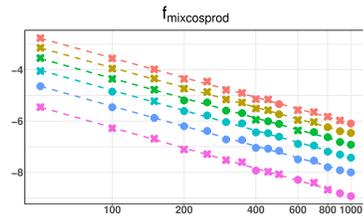
ι	Expected	Estimation	Conf. interval
6	−1.17	−1.11	[−1.14 ; −1.08]
5	−1.17	−1.11	[−1.14 ; −1.09]
4	−1.17	−1.12	[−1.15 ; −1.08]
3	−1.17	−1.13	[−1.17 ; −1.09]
2	−1.17	−1.12	[−1.16 ; −1.09]
1	−1.17	−1.15	[−1.18 ; −1.12]
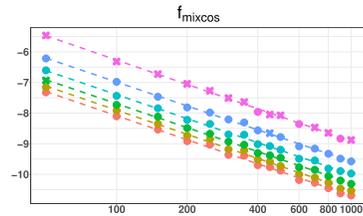
ι	Expected	Estimation	Conf. interval
6	−1.17	−1.13	[−1.17 ; −1.1]
5	−1.17	−1.14	[−1.17 ; −1.1]
4	−1.17	−1.13	[−1.16 ; −1.1]
3	−1.17	−1.13	[−1.17 ; −1.09]
2	−1.17	−1.13	[−1.15 ; −1.1]
1	−1.17	−1.15	[−1.18 ; −1.12]
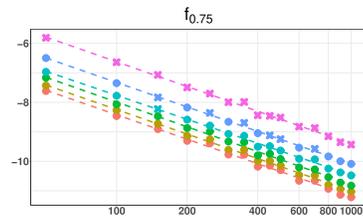
ι	Expected	Estimation	Conf. interval
6	−1.17	−1.20	[−1.25 ; −1.14]
5	−1.17	−1.20	[−1.25 ; −1.15]
4	−1.17	−1.21	[−1.27 ; −1.16]
3	−1.17	−1.17	[−1.21 ; −1.13]
2	−1.17	−1.20	[−1.23 ; −1.16]
1	−1.17	−1.21	[−1.26 ; −1.16]
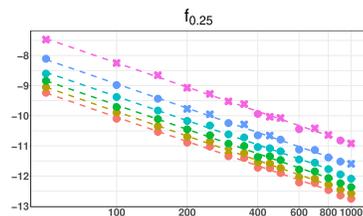
ι	Expected	Estimation	Conf. interval
6	−1.17	−1.16	[−1.2 ; −1.12]
5	−1.17	−1.17	[−1.2 ; −1.13]
4	−1.17	−1.16	[−1.2 ; −1.12]
3	−1.17	−1.16	[−1.19 ; −1.12]
2	−1.17	−1.16	[−1.2 ; −1.12]
1	−1.17	−1.16	[−1.2 ; −1.12]

FIG 3. *Summary of experiments in which integrals of ι-dimensional functions are estimated by projecting a single $(N, 6)$-Dirichlet DPP $(\iota = 1 \ldots 6)$. A ● (resp. ×) indicates that the adjusted p-value of the Shapiro-Wilk test is not smaller (resp. smaller) than 5%.*

ι	Expected	Estimation	Conf. interval
10	−1.1	−1.03	[−1.05 ; −1.02]
6	−1.1	−1.05	[−1.07 ; −1.03]
5	−1.1	−1.06	[−1.08 ; −1.04]
4	−1.1	−1.04	[−1.06 ; −1.02]
3	−1.1	−1.05	[−1.07 ; −1.03]
2	−1.1	−1.08	[−1.11 ; −1.05]
1	−1.1	−1.06	[−1.08 ; −1.04]

ι	Expected	Estimation	Conf. interval
10	−1.1	−1.05	[−1.08 ; −1.01]
6	−1.1	−1.05	[−1.07 ; −1.02]
5	−1.1	−1.04	[−1.07 ; −1.02]
4	−1.1	−1.05	[−1.06 ; −1.03]
3	−1.1	−1.05	[−1.08 ; −1.03]
2	−1.1	−1.06	[−1.08 ; −1.04]
1	−1.1	−1.06	[−1.09 ; −1.03]

ι	Expected	Estimation	Conf. interval
10	−1.1	−1.06	[−1.08 ; −1.04]
6	−1.1	−1.07	[−1.08 ; −1.05]
5	−1.1	−1.05	[−1.07 ; −1.04]
4	−1.1	−1.06	[−1.08 ; −1.04]
3	−1.1	−1.07	[−1.08 ; −1.05]
2	−1.1	−1.05	[−1.07 ; −1.03]
1	−1.1	−1.06	[−1.09 ; −1.04]

ι	Expected	Estimation	Conf. interval
10	−1.1	−1.08	[−1.1 ; −1.06]
6	−1.1	−1.08	[−1.1 ; −1.05]
5	−1.1	−1.08	[−1.1 ; −1.06]
4	−1.1	−1.10	[−1.12 ; −1.07]
3	−1.1	−1.09	[−1.11 ; −1.07]
2	−1.1	−1.08	[−1.11 ; −1.05]
1	−1.1	−1.08	[−1.11 ; −1.05]

ι	Expected	Estimation	Conf. interval
10	−1.1	−1.07	[−1.09 ; −1.04]
6	−1.1	−1.06	[−1.09 ; −1.03]
5	−1.1	−1.07	[−1.09 ; −1.04]
4	−1.1	−1.08	[−1.1 ; −1.06]
3	−1.1	−1.06	[−1.08 ; −1.03]
2	−1.1	−1.05	[−1.08 ; −1.03]
1	−1.1	−1.09	[−1.11 ; −1.07]

FIG 4. *Summary of experiments in which integrals of ι-dimensional functions are estimated by projecting a single $(N, 10)$-Dirichlet DPP ($\iota = 1 \ldots 6, 10$). A $\bullet$ (resp. $\times$) indicates that the adjusted p-value of the Shapiro-Wilk test is not smaller (resp. smaller) than 5%.*

γ	s	Expected	Estimation
0.625	0.125	−1.25	−1.4
0.75	0.25	−1.5	−1.57
0.875	0.375	−1.75	−1.74
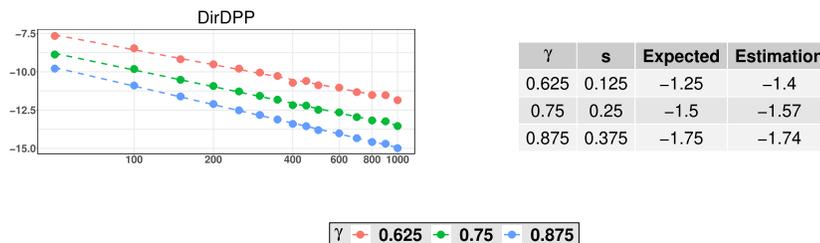
FIG 5. *Summary of experiments in which integrals of function* (4.5) *are estimated using* $(N, 1)$-*Dirichlet DPP.*

4.3. Case $s < 1/2$

We now intend to illustrate Theorem 3.1 (i) for $d = 1$ and $s < 1/2$. We consider the following one-dimensional function defined on [0,1]:

$$h_\gamma(x) = \sum_{j \geqslant 1} \frac{\cos(2\pi j(x - 1/2))}{2\pi j^\gamma}. \tag{4.5}$$

It is clear that $h_\gamma \in L^2_{\mathrm{per}}([0,1])$ if $\gamma > 1/2$, and in such a case, $h_\gamma \in \mathcal{H}^s([0,1])$ for $s < \gamma - 1/2$.

We repeat experiments presented in Section 4.1, but only for the case $d = 1$: getting an accurate evaluation of (4.5) becomes really expensive for higher dimension. We consider three values for γ: 0.625, 0.75 and 0.875. Empirical results are depicted in Figure 5. As expected, the normality is rejected for all values of N (for more readability we keep the dot representation for the points). Surprisingly, the results remains very satisfactory. When $\gamma = 0.75, 0.875$, the points remain well-aligned and the estimated slopes are in good agreement with the theoretical $-1 - 2s/d$. The case $\gamma = 0.625$ is less convincing but still very satisfactory, since $h_{0.625}$ is a very irregular function.

5. A comparative numerical study.

We perform similar experiments with several designs: crude Monte-Carlo with uniform proposal, the DPP proposed by Bardenet and Hardy [3] based on Legendre polynomials available in `DPPy` toolbox [10], stratified sampling [26], *maximin* Latin Hypercube Design of `lhs R` package and Sobol and Halton sequences (`randtoolbox R` package). The Dirichlet DPP and these different designs will respectively be denoted by `DirDPP`, `crude MC`, `BHDPP`, `stratified MC`, `maximinLHS`, `Sobol`, `Halton` in the following.

We first compare the `DirDPP` and `BHDPP` designs since they seem very close in terms of construction. Experiments presented in Section 4 (in the case $s > 1/2$) for the `DirDPP` are conducted for the `BHDPP` design (only for $d = 1, \ldots, 6$ to save time). Thus, Figures 6-8 should be compared to Figures 2, 3 and 5. Regarding

Figures 2 and 6, as expected, the results are quite similar for the f_{bump} function, in terms of slope and confidence interval. However, we observe that for $d > 4$, fluctuations around regression lines as well as confidence intervals are larger for the BHDPP design than for the DirDPP design. Moreover, Gaussianity seems to be reached less often. Surprisingly, the BHDPP design gives satisfactory results even for non-differentiable integrands (including for function (4.5) according to Figure 8), as long as $d < 4$. But results deteriorate for $d \geq 4$. The estimated slopes are even larger than -1 which corresponds to crude MC designs. Bardenet and Hardy [3] did not consider the problem of estimating ι-dimensional integrals. Their point process was not designed to address such a problem, and this is indeed revealed by Figure 7. When projections of a 6-dimensional BHDPP design is used to estimate ι-dimensional integrals with $\iota = 1, \ldots, 5$, the estimated slopes are at best close to -1 and at worse larger than -1. In all cases, the estimated slopes for the BHDPP design are far from $-1 - 1/6$ which on the contrary is reached by the DirDPP design.

Next, we compare the DirDPP design with more standard alternatives. For each design (crude MC, strat MC, maximinLHS, Sobol, Halton), we have reproduced exactly the same experiment which has led to Figures 2-4, except that due to the nature of the maximinLHS, Sobol, Halton designs, we have computed logarithms of the empirical mean squared errors (MSE) in terms of $\log(N)$. Of course, plotting the MSE for the DirDPP design instead of its variance in Figures 2-4 would have led to the same conclusions, since estimates are unbiased (which turns out to be also the case for crude MC, strat MC and BHDPP designs). Moreover, similar figures summarizing empirical bias and variance have also been produced and are available on demand.

Figure 9 is a tentative summary of Figures 2-4 and Figures 10-24 presented in Appendix C for the sake of completeness. It reports confidence intervals for the slope estimates of linear regressions of the logarithms of the MSE in terms of $\log(N)$ for all designs and tests functions. All these figures are quite complex to comment. Our goal is not to point out which method is best for a particular function, or a particular sample size N, *etc.* Instead, we are willing to investigate if a method is "stable" (in terms of rate of convergence, or over a set of test functions, *etc*) in order to estimate d-dimensional integrals or ι-dimensional integrals based on d-dimensional designs.

Crude MC designs react as expected with a slope of -1 in all situations.

Stratified MC designs seem to be quite competitive with respect to DirDPP designs, when one estimates d-dimensional integrals, even for non-differentiable integrands. Nevertheless, we observe that the rate of convergence is not exactly the same over the set of tests functions (*e.g.* -3 for the bump function and -2.5 for $f_{0.25}$ when $d = 1$). Properties of stratified MC designs are significantly deteriorated when we estimate ι-dimensional integrals (see Figures 16 and 21). Estimates tend to converge but with a lower rate of convergence than estimates based on the DirDPP design.

MaximinLHS designs are quite impressive when $d = 1, \ldots, 6, 10$ for functions like $f_{\text{mixcos}}, f_{0.75}$ and $f_{0.25}$ but present some huge failures for functions f_{bump} and $f_{\text{mixcosprod}}$. We have no explanation for the difficulties of this design to estimate

f_{bump}
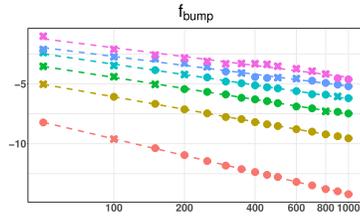
d	Expected	Estimation	Conf. interval
1	−2.00	−2.00	[−2.02 ; −1.99]
2	−1.50	−1.51	[−1.52 ; −1.5]
3	−1.33	−1.33	[−1.36 ; −1.3]
4	−1.25	−1.25	[−1.3 ; −1.21]
5	−1.20	−1.07	[−1.17 ; −0.981]
6	−1.17	−1.07	[−1.19 ; −0.956]

$f_{mixcosprod}$
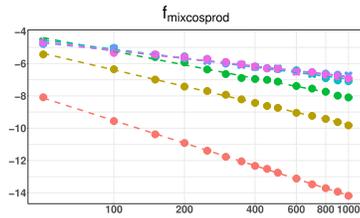
d	Expected	Estimation	Conf. interval
1	−2.00	−2.010	[−2.03 ; −2]
2	−1.50	−1.470	[−1.49 ; −1.44]
3	−1.33	−1.210	[−1.29 ; −1.13]
4	−1.25	−0.803	[−0.876 ; −0.729]
5	−1.20	−0.693	[−0.735 ; −0.651]
6	−1.17	−0.725	[−0.775 ; −0.675]

f_{mixcos}
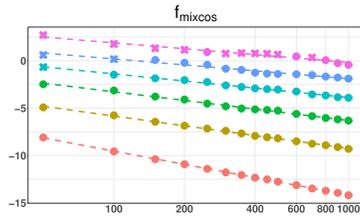
d	Expected	Estimation	Conf. interval
1	−2.00	−2.010	[−2.03 ; −2]
2	−1.50	−1.480	[−1.52 ; −1.45]
3	−1.33	−1.310	[−1.37 ; −1.26]
4	−1.25	−1.080	[−1.13 ; −1.03]
5	−1.20	−0.904	[−1.01 ; −0.8]
6	−1.17	−0.893	[−1.01 ; −0.771]

$f_{0.75}$
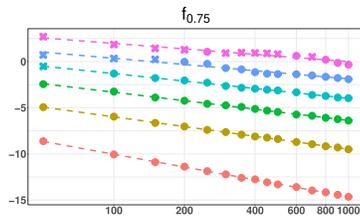
d	Expected	Estimation	Conf. interval
1	−2.00	−2.000	[−2.03 ; −1.98]
2	−1.50	−1.520	[−1.54 ; −1.5]
3	−1.33	−1.330	[−1.35 ; −1.3]
4	−1.25	−1.160	[−1.2 ; −1.13]
5	−1.20	−0.968	[−1.09 ; −0.844]
6	−1.17	−0.855	[−0.983 ; −0.727]

$f_{0.25}$
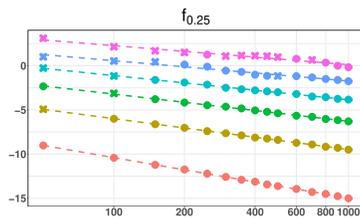
| d	Expected	Estimation	Conf. interval
1	−2.00	−1.990	[−2.01 ; −1.97]
2	−1.50	−1.520	[−1.53 ; −1.5]
3	−1.33	−1.340	[−1.37 ; −1.31]
4	−1.25	−1.190	[−1.22 ; −1.16]
5	−1.20	−0.998	[−1.11 ; −0.884]
6	−1.17	−0.936	[−1.07 ; −0.805]

| d | 1 | 2 | 3 | 4 | 5 | 6 |

FIG 6. *Summary of experiments in which integrals of d-dimensional functions are estimated using the* BHDPP *design. A ● (resp. ×) indicates that the adjusted p-value of the Shapiro-Wilk test is not smaller (resp. smaller) than 5%*

| ι	Expected	Estimation	Conf. interval
6	−1.17	−1.07	[−1.19 ; −0.956]
5	NA	−1.06	[−1.18 ; −0.949]
4	NA	−1.06	[−1.18 ; −0.95]
3	NA	−1.06	[−1.18 ; −0.943]
2	NA	−1.04	[−1.15 ; −0.922]
1	NA	−1.03	[−1.15 ; −0.908]

ι	Expected	Estimation	Conf. interval
6	−1.17	−0.725	[−0.775 ; −0.675]
5	NA	−0.767	[−0.812 ; −0.723]
4	NA	−0.824	[−0.863 ; −0.785]
3	NA	−0.874	[−0.924 ; −0.825]
2	NA	−0.912	[−0.981 ; −0.843]
1	NA	−0.953	[−1.05 ; −0.852]

ι	Expected	Estimation	Conf. interval
6	−1.17	−0.893	[−1.01 ; −0.771]
5	NA	−0.897	[−1.02 ; −0.776]
4	NA	−0.896	[−1.01 ; −0.777]
3	NA	−0.911	[−1.03 ; −0.792]
2	NA	−0.918	[−1.03 ; −0.806]
1	NA	−0.950	[−1.04 ; −0.859]

ι	Expected	Estimation	Conf. interval
6	−1.17	−0.855	[−0.983 ; −0.727]
5	NA	−0.861	[−0.988 ; −0.733]
4	NA	−0.867	[−0.994 ; −0.74]
3	NA	−0.870	[−0.995 ; −0.746]
2	NA	−0.882	[−0.997 ; −0.767]
1	NA	−0.927	[−1.03 ; −0.821]

ι	Expected	Estimation	Conf. interval
6	−1.17	−0.936	[−1.07 ; −0.805]
5	NA	−0.936	[−1.07 ; −0.806]
4	NA	−0.938	[−1.07 ; −0.809]
3	NA	−0.938	[−1.07 ; −0.809]
2	NA	−0.943	[−1.07 ; −0.815]
1	NA	−0.951	[−1.08 ; −0.826]

FIG 7. *Summary of experiments in which integrals of ι-dimensional functions are estimated by projecting a single 6-dimensionnal* BHDPP *design (ι = 1 . . . 6). A ● (resp. ×) indicates that the adjusted p-value of the Shapiro-Wilk test is not smaller (resp. smaller) than 5%.*

γ	s	Expected	Estimation
0.625	0.125	NA	−1.29
0.75	0.25	NA	−1.51
0.875	0.375	NA	−1.71

FIG 8. *Summary of experiments in which integrals of function* (4.5) *are estimated using the BH design.*

integrals of functions of the form (3.12). Similar behaviours are observed when we estimate ι-dimensional integrals (see Figures 17 and 22). This apparent lack of convergence for the MSE is essentially explained by the fact that the empirical bias does not converge with N.

Sobol and Halton designs exhibit similar behaviours. Looking at Figures 13 and 14, we observe that these designs lead to very small MSE (often much smaller than the ones obtained with other methods). However, what is striking when we examine Figure 9 is the huge variability of slope estimates. Indeed, in Figures 13 and 14, the (logarithms of) MSE have a high variance in particular for functions $f_{\mathrm{mixcos}}, f_{0.75}$ and $f_{0.25}$ (*i.e.* for functions of the form (3.13)). For these functions, results are difficult to explain as the dimension does not seem to affect the rate of convergence. Empirical results are even more variable when we are interested in estimating ι-dimensional integrals (see Figures 18-19 and Figures 23-24). There is no clear rate of convergence over the set of test functions.

Overall, from this (clearly not exhaustive) empirical study, we do not claim that the DirDPP design outperforms other designs investigated. For a specific sample size N, a specific function, a specific dimension, we may find a much better method. However, the DirDPP design is the only one (with the baseline crude MC design) which exhibits a clear and expected rate of convergence for any function, any dimension d or any "sub-dimension" ι.

6. Conclusion

In this paper, we build a specific class of repulsive point process and use its realization as quadrature points to estimate integrals. The resulting Monte-Carlo estimator is unbiased with a variance scaling as $\mathcal{O}(N^{-1-(2s\wedge1)/d})$ when the integrand belongs to $\mathcal{H}^s([0,1]^d)$. Our methodology and results have the interest to be applied to non differentiable functions, an assumption which is often considered for other methods such as grid-based methods, RQMC, scrambled nets [26]. We also show that the initial configuration of points can be used to estimate ι-dimensional integrands ($\iota = 1, \ldots, d$) with the same efficiency than in dimension d. To complete this work on a theoretical side, it could be interesting to inves-
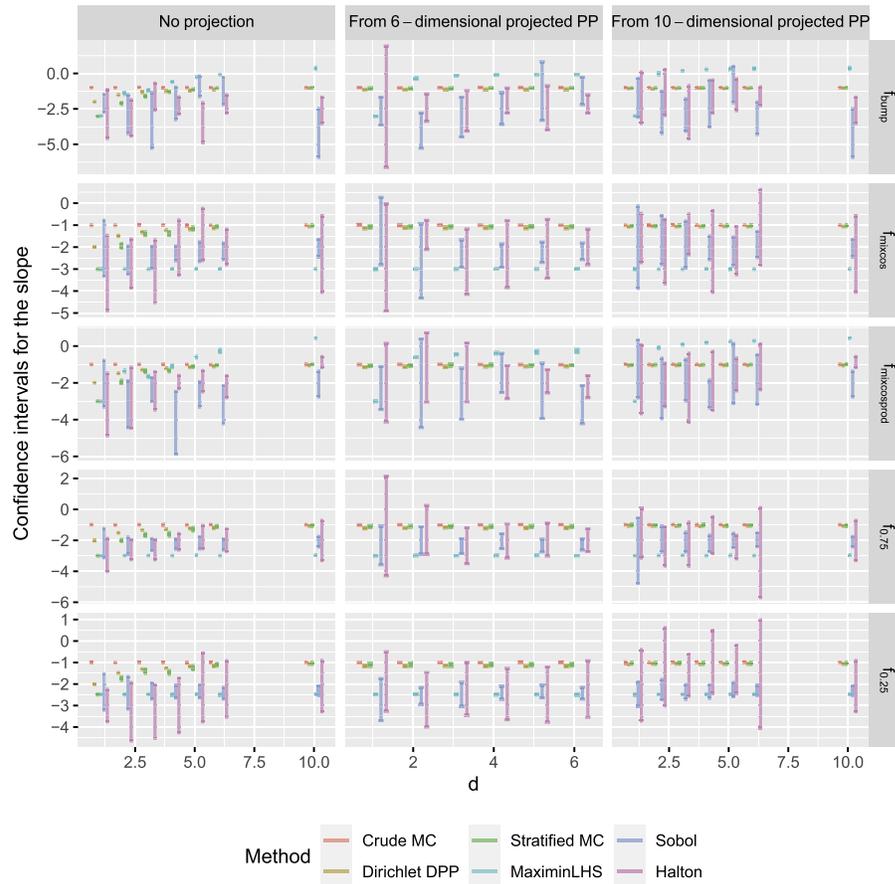
FIG 9. *Confidence intervals for the slope estimates of linear regressions of the logarithms of the MSE in terms of* $\log(N)$ *when we estimate d dimensional integrals and designs generated in dimension d (left) or ι-dimensional integrals based on designs generated in dimension* $d = 6$ *or* 10. *Such experiments are done for the five test functions given by* (4.1)-(4.4) *and for different experimental designs.*

tigate the rate of convergence to normality in some metric (*e.g.* Wassertstein or Kolmogorov-Smirnov). We have also left behind an important practical question: how to estimate $\sigma^2(f)$ or its approximation given by (3.11) and the matrix $\mathbf{\Sigma}_p$ given by (3.17)? Obviously, we do not want to assume known the Fourier expansion of f nor to use external simulations to estimate such quantities which would make the current work definitely useless. Estimating $\sigma^2(f)$ or $\mathbf{\Sigma}_p$ using only the realization of the Dirichlet DPP appears to be highly challenging. Due to the intrinsic nature of this projection DPP, strategies like bootstrap, subsampling or leave-one out type procedures did not lead to a consistent estimate of the asymptotic variance. This problem is definitely an interesting perspective.

We consider in this paper integrals on $[0,1]^d$ (or rectangles). Addressing the

similar question for more complex domains or for improper integrals is of great interest. Similarly, considering integrands defined on manifolds such as torus or spheres is definitely an interesting perspective. We could, for example, take advantage of the works by [21] which presents several spherical-valued DPP models.

Crude Monte-Carlo or stratified Monte-Carlo designs can be improved by using antithetic or local antithetic methods, see Owen [25, 26]. An idea of such improvements relies on the fact that $U \sim U(0,1)$ and $1 - U$ have the same distribution. Such an idea could also be applied to our design: This is again an advantage of the homogeneity of the Dirichlet DPP. We have not deeply studied these extensions but leave them for future works. To give some flavor however, we can easily prove that the estimator

$$\widehat{\mu}'_N(f_I) = \frac{1}{2N} \sum_{j=1}^{N} \left\{ f_I((u_j)_I) + f_I((1 - u_j)_I) \right\},$$

where $1 - u = (1 - u_1, \ldots, 1 - u_d)^\top$ for any $u \in [0, 1]^d$, satisfies Theorem 3.1 and Corollary 3.1, with the same rate of convergence, the same assumptions on f_I. However, the asymptotic constant is given by (3.15) where $|\hat{f}_I(j)|^2$ is replaced by $\text{Re}(\hat{f}_I(j))^2$: This necessarily leads to a variance reduction.

In this paper we did not exploit the form of the integrand f. "Is it possible to improve the rate of convergence by designing an appropriate DPP which exploits the form of f?" seems to be a difficult but again interesting question. We believe this is not feasible with the Dirichlet DPP proposed in this paper but the question remains open.

Another natural extension would be to consider a general target product measure m on $[0, 1]^d$, instead of the uniform distribution. The naive approach would be obviously to estimate $\int_{[0,1]^\iota} f_I m_I \mathrm{d}x$ with Dirichlet DPP. But, as Bardenet and Hardy [3] did, the measure m could be used to build a dedicated DPP. According to the proofs in our work, it seems reasonable to think that Theorem 3.1 and Corollary 3.1 could be extended to any product measure m such that an orthonormal basis $\{\varphi_k(\cdot)\}_k$ of $L^2([0, 1]^d, m)$ satisfies a separability property $\varphi_k(x)\varphi_l(x) = \varphi_{k+l}(x)$.

One of the limitation of the methodology lies in the simulation of a continuous DPP. Current algorithms have a computational cost $\mathcal{O}(N^3)$ and are based on chain rules, where each point location is sampled with a rejection method, with evaluations of the acceptance ratio can be very costly. As the number of generated points increases, passing the rejection step becomes harder, and this step makes the sampler even more costly. A personal communication from [18] shows that in dimension one this rejection step can be significantly improved with a proposal adapted to the Fourier kernel. This result lets us hope that this proposal could be extended to the d-dimensional Dirichlet DPP. Finally, it is worth reminding the reader that we use homogeneous point patterns which can be used to estimate any integral. Therefore the 2500 replications from the Dirichlet DPP, for $N = 50, 100, \ldots, 1000$ and $d = 1, \ldots, 6$ can be used to evaluate any integral and are freely available upon request.

Appendix A: Proofs

A.1. Proof of Proposition 3.1

Proof. Since $\mathbf{X}$ is a projection kernel, in (1.8), $\lambda_j = 0$ or 1, $\rho_{\mathbf{Y}} = N$. Moreover, the trick is that the Fourier basis satisfies $\phi_j(u)\overline{\phi_k(u)} = \phi_{j-k}(u)$ for any $j, k \in \mathbb{Z}^d$ and $u \in B$. These facts reduce (1.8) to

$$\mathrm{Var}\left[\widehat{\mu}_N(f)\right] = \frac{1}{N} \int_B f(u)^2 \mathrm{d}u - \sum_{j,k \in E_N} |\widehat{f}(j-k)|^2, \tag{A.1}$$

whereby (3.6) is deduced using Parseval's identity. Let us focus now on the second term of the right-hand side of (3.6). When $d = 1$, it is quite well-known that

$$\sum_{j,k=0}^{N-1} \left|\widehat{f}(j-k)\right|^2 = \sum_{|l| \leqslant N-1} (N - |l|) \left|\widehat{f}(l))\right|^2. \tag{A.2}$$

The result is easily deduced using the fact that E_N is a rectangular subset of $\mathbb{Z}^d$. $\qquad\square$

A.2. Proof of Theorem 3.1

Proof. Let $i = 1, \ldots, d$ and $j_i \in \mathbb{Z}$ with $|j_i| \leq n_i - 1$. Then,

$$\begin{aligned}
\prod_{i=1}^{d} (n_i - |j_i|) &= N \prod_{i=1}^{d} \left(1 - \frac{|j_i|}{n_i}\right) \\
&= N \sum_{k=0}^{d} (-1)^k \sigma_k \left(\frac{|j_1|}{n_1}, \ldots, \frac{|j_d|}{n_d}\right) \\
&= N \left(1 - \sum_{i=1}^{d} \frac{|j_i|}{n_i} + \sum_{1 \leqslant i_1 < i_2 < d} \frac{|j_{i_1} j_{i_2}|}{n_{i_1} n_{i_2}} + \cdots + (-1)^d \prod_{i=1}^{d} \frac{|j_i|}{n_i}\right)
\end{aligned}$$

where, for $k = 1 \ldots d$, $\sigma_k(x_1, \ldots, x_d)$ denotes the k-th elementary symmetric polynomials in d variables:

$$\sigma_k(x_1, \ldots, x_d) = \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant d} x_{i_1} \ldots x_{i_k}$$

For $u, v \in \mathbb{R}^d$ (with $u_\ell \neq 0$, $\ell = 1, \ldots, d$), we define $1/u$ (resp. v/u) by $1/u = (1/u_1, \ldots, 1/u_d)^\top$ (resp. $v/u = (v_1/u_1, \ldots, v_d/u_d)^\top$). Similarly, $\|u\|_\infty$ stands for $\max_\ell |u_\ell|$. It is clear that for any $k \geq 2$ and $j \in F_N = \{j \in \mathbb{Z}^d : |j_i| \leq n_i - 1, \ i = 1, \ldots, d\}$

$$\sigma_k \left(\frac{|j_1|}{n_1}, \ldots, \frac{|j_d|}{n_d}\right) = \mathcal{O}\left(\left\|\frac{j}{n}\right\|_\infty^2\right).$$

Therefore, (3.7) can be rewritten as

$$\mathrm{Var}\left[\widehat{\mu}_N(f)\right] = \frac{1}{N} \sum_{j \in F_N^c} \left|\widehat{f}(j)\right|^2 + \frac{1}{N} \sum_{j \in F_N} \left\|\frac{j}{n}\right\|_1 \left|\widehat{f}(j)\right|^2$$
$$+ \frac{1}{N} \sum_{j \in F_N} R(j/n) \left|\widehat{f}(j)\right|^2 \qquad (A.3)$$

where

$$R(j/n) = \begin{cases} 0 & \text{if } d = 1 \\ \mathcal{O}\left(\left\|\frac{j}{n}\right\|_\infty^2\right) & \text{if } d > 1. \end{cases}$$

Under the asymptotic framework (3.8), we have that as $N \to \infty$

$$\mathrm{Var}\left[\widehat{\mu}_N(f)\right] \sim \frac{1}{N} \sum_{j \in F_N^c} \left|\widehat{f}(j)\right|^2 + \frac{1}{N^{1+1/d}} \sum_{j \in F_N} \left\|\frac{j}{\kappa}\right\|_1 \left|\widehat{f}(j)\right|^2$$
$$+ \frac{1}{N^{1+2/d}} \sum_{j \in F_N} R(j/\kappa) \left|\widehat{f}(j)\right|^2. \qquad (A.4)$$

In the following, τ denotes a positive generic constant which may vary from line to line.

(i) Let $s \in (0, 1/2)$ and $j \in F_N$,

$$\left\|\frac{j}{\kappa}\right\|_1 = \left\|\frac{j}{\kappa}\right\|_1^{2s} \left\|\frac{j}{\kappa}\right\|_1^{1-2s} \leq \tau N^{\frac{1-2s}{d}} (1 + \|j\|_\infty)^{2s}. \qquad (A.5)$$

Hence,

$$\frac{1}{N^{1+1/d}} \sum_{j \in F_N} \left\|\frac{j}{\kappa}\right\|_1 \left|\widehat{f}(j)\right|^2 \leq \frac{\tau}{N^{1+2s/d}} \sum_{j \in F_N} (1 + \|j\|_\infty)^{2s} \left|\widehat{f}(j)\right|^2$$
$$= \mathcal{O}\left(N^{-1-2s/d}\right). \qquad (A.6)$$

Similarly,

$$\frac{1}{N^{1+2/d}} \sum_{j \in F_N} R(j/\kappa) \left|\widehat{f}(j)\right|^2 \leq \frac{\tau N^{(2-2s)/d}}{N^{1+2/d}} \sum_{j \in F_N} (1 + \|j\|_\infty)^{2s} \left|\widehat{f}(j)\right|^2$$
$$= \mathcal{O}\left(N^{-1-2s/d}\right) \qquad (A.7)$$

and

$$\frac{1}{N} \sum_{j \in F_N^c} \left|\widehat{f}(j)\right|^2 \leq \frac{\tau}{N^{1+2s/d}} \sum_{j \in F_N^c} (1 + \|j\|_\infty)^{2s} \left|\widehat{f}(j)\right|^2$$
$$= o\left(N^{-1-2s/d}\right). \qquad (A.8)$$

Combining (A.6)-(A.8) with (A.4) leads to the result.

(ii) We proceed similarly for the case $s > 1/2$. Let us focus on the case $d > 1$ as the other one is easily deduced. Since $f \in \mathcal{H}^s(B)$ for $s > 1/2$, we have that as $N \to \infty$

$$\frac{1}{N^{1+1/d}} \sum_{j \in F_N} \left\| \frac{j}{\kappa} \right\|_1 \left| \widehat{f}(j) \right|^2 \sim \frac{1}{N^{1+1/d}} \sum_{j \in \mathbb{Z}^d} \left\| \frac{j}{\kappa} \right\|_1 \left| \widehat{f}(j) \right|^2. \tag{A.9}$$

We also have that for any $j \in F_N$, $R(j/\kappa) \leq \tau(1 + \|j\|_\infty)^{\min(2s,2)}$ which leads to

$$\frac{1}{N^{1+2/d}} \sum_{j \in F_N} R(j/\kappa) \leq \frac{\tau}{N^{1+2/d}} \sum_{j \in F_N} (1 + \|j\|_\infty)^{\min(2s,2)} \left| \widehat{f}(j) \right|^2$$

$$= \mathcal{O}\left(N^{-1-2/d} \right). \tag{A.10}$$

Finally,

$$\frac{1}{N} \sum_{j \in F_N^c} \left| \widehat{f}(j) \right|^2 \leq \frac{\tau}{N^{1+2s/d}} \sum_{j \in F_N^c} (1 + \|j\|_\infty)^{2s} \left| \widehat{f}(j) \right|^2$$

$$= \mathcal{O}\left(N^{-1-2s/d} \right) = o\left(N^{-1-1/d} \right). \tag{A.11}$$

Combining (A.9)-(A.11) with (A.4) leads again to the result.

(iii) We consider again the case $d > 1$. To achieve this step, we apply Soshnikov [36, Theorem 1] to the sequence of random variables $S_N(f) = \sum_{j=1}^N f(u_j)$. First, from (3.9)

$$\mathrm{Var}(S_N(f)) \sim N^2 \mathrm{Var}(\widehat{\mu}_N(f)) \sim N^{1-1/d} \sigma^2(f) \to \infty \tag{A.12}$$

as $N \to \infty$ if $d > 1$. Second,

$$\|f\|_\infty = \mathcal{O}(1) = o\left(N^{\tau(1-1/d)} \right) \tag{A.13}$$

for any $\tau > 0$. Third, $\mathbb{E}(S_N(|f|)) = \mathcal{O}(N)$ and for $\delta = (1 + \delta')/(1 - 1/d)$ for some $\delta' > 0$, we have $\mathrm{Var}(S_N(f))^\delta = \mathcal{O}(N^{\delta(1-1/d)}) = \mathcal{O}(N^{1+\delta'})$, which implies that

$$\mathbb{E}(S_N(|f|)) = \mathcal{O}(\mathrm{Var}(S_N(f))^\delta). \tag{A.14}$$

Equations (A.12)-(A.14) are the key-ingredients of Soshnikov [36, Theorem 1], which proves that

$$\frac{S_N(f) - N\mu(f)}{\sqrt{\mathrm{Var}(S_N(f))}} \to N(0,1)$$

as $N \to \infty$ and yields the result. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

A.3. Proof of Proposition 3.2

The proof of this result is a direct consequence of the expression of $\sigma^2(f), \sigma^2_{\text{std}}(f)$ and $\sigma^2_{\text{strat}}(f)$ and the following lemma.

Lemma A.1.
(i) Let f be defined as (3.12) then for any $j \in \mathbb{Z}^d$, $\hat{f}(j) = \prod_{k=1}^d \hat{f}_0(j_k)$, whereby we deduce that for any $s \geq 0$

$$|\hat{f}(j)|^2 = \prod_{k=1}^d |\hat{f}_0(j_k)|^2 \quad and \quad \|f\|^2_{\mathcal{H}^s([0,1]^d)} = d\,\|f_0\|^2_{\mathcal{H}^s([0,1])}\|f_0\|^{2d-2}_{\mathcal{H}^0([0,1])}.$$

(ii) Let f be defined as (3.13) then for any $j \in \mathbb{Z}^d$

$$\hat{f}(j) = \sum_{k=1}^d \hat{f}_0(j_k)\delta_{0,j_{-k}}$$

where for $u, v \in \mathbb{Z}^m$

$$\delta_{u,v} = \left\{ \begin{array}{ll} 1 & if\ u_1 = v_1, \ldots, u_m = v_m \\ 0 & otherwise \end{array} \right.$$

and $j_{-k} = (j_1, \ldots, j_{k-1}, j_{k+1}, \ldots, j_d)$. Hence, for any $s \geq 0$

$$|\hat{f}(j)|^2 = \sum_{k=1}^d |\hat{f}_0(j_k)|^2\delta_{0,j_{-k}} \quad and \quad \|f\|^2_{\mathcal{H}^s([0,1]^d)} = d\|f_0\|^2_{\mathcal{H}^s([0,1])}.$$

Proof. (i) The expressions for $\hat{f}(j)$ and $|\hat{f}(j)|^2$ are straightforward. From this we derive

$$\|f\|^2_{\mathcal{H}^s([0,1]^d)} = \sum_{j_1 \in \mathbb{Z}} \cdots \sum_{j_d \in \mathbb{Z}} \{|j_1|^{2s} + \cdots + |j_d|^{2s}\} \prod_{k=1}^d |\hat{f}_0(j_k)|^2$$

$$= d \sum_{j_1 \in \mathbb{Z}} |j_1|^{2s}|\hat{f}_0(j_1)|^2 \sum_{k=2}^d \sum_{j_k \in \mathbb{Z}} |\hat{f}_0(j_k)|^2$$

$$= d\,\|f_0\|^2_{\mathcal{H}^s([0,1])}\|f_0\|^{2(d-1)}_{L^2([0,1])}.$$

(ii) Omitted. $\qquad\qquad\square$

A.4. Proof of Corollary 3.1

Proof. Let $f_I^\uparrow : \mathbb{R}^d \to \mathbb{R}$ be the d-dimensional measurable function given by $f_I^\uparrow(x) = f_I(x_I)\mathbb{1}\,[x_{I^c} \in B_{I^c}]$. Then

$$\widehat{\mu}_N(f_I^\uparrow) = \frac{1}{N} \sum_{u \in \mathbf{X}} h(u) = \frac{1}{N} \sum_{u \in \mathbf{X}_I} f_I(u) = \widehat{\mu}_N(f_I).$$

Since $f_I \in \mathcal{H}^s(B_I)$, it is straightforward to see that $f_I^\uparrow \in \mathcal{H}^s(B)$. In particular

$$\sum_{j \in \mathbb{Z}^d} (1 + \|j\|_\infty)^{2s} \left| \widehat{f_I^\uparrow}(j) \right|^2 = \sum_{j \in \mathbb{Z}^d} (1 + \|j\|_\infty)^{2s} \mathbb{1}\left[j_{I^c} = \mathbf{0}_{d-\iota} \right] \left| \widehat{f_I}(j_I) \right|^2$$

$$= \sum_{j \in \mathbb{Z}^\iota}^d (1 + \|j\|_\infty)^{2s} \left| \widehat{f_I}(j) \right|^2$$

where for some $p \geq 1$, $\mathbf{0}_p$ is the zero vector in $\mathbb{R}^p$. Corollary 3.1 is deduced by applying Theorem 3.1 to the function $f_I^\uparrow$. $\square$

A.5. Proof of Corollary 3.2

Proof. Let us first note, as in the proof of Corollary 3.1 that $f_{I_\ell} \in \mathcal{H}^s(B_{I_\ell})$ is equivalent to say that $f_{I_\ell^\uparrow} \in \mathcal{H}^s(B)$. Moreover, following the proof of Theorem 3.1(i), it is shown that as $N \to \infty$

$$N^{1+1/d} \, \mathrm{Cov}\left(\widehat{\mu}_N(f_{I_\ell}^\uparrow), \widehat{\mu}_N(f_{I_{\ell'}}^\uparrow) \right) \to (\mathbf{\Sigma}_p)_{\ell\ell'} \, .$$

Now, we follow Cramèr-Wold device: let $a \in \mathbb{R}^p$, and let $Z_a = a^\top \left(\widehat{\mu}_{N,p} - \mu_p \right)$. Then, $Z_a = \widehat{\mu}_N(g_a) - \mu(g_a)$ with $g_a(u) = \sum_{\ell=1}^p a_\ell f_{I_\ell}^\uparrow(u)$. The result is therefore deduced since $g_a \in \mathcal{H}^s(B)$, $N^{1+1/d}\mathrm{Var}(\widehat{\mu}_N(g_a)) \to a^\top \mathbf{\Sigma}_p a$ and Theorem 3.1(ii) can be applied to g_a, that is as $N \to \infty$

$$\sqrt{N^{1+1/d}} \, a^\top \left(\widehat{\mu}_{N,p} - \mu_p \right) \to N(0, a^\top \mathbf{\Sigma}_p a).$$

in distribution. $\square$

Appendix B: Alternative proof of Theorem 3.1

Here we propose an alternative proof of Corollary 3.1 (ii)-(iii) based on the characterization of the projected point process $\mathbf{X}_I$.

Proof. (ii) From Proposition 2.1 and in particular the characterization of (α)-DPPs as the union of independent particular DPPs (see Hough et al. [14]), we have that

$$\widehat{\mu}_N(f_I) = \frac{1}{N_{I^c}} \sum_{j=1}^{N_{I^c}} \widehat{\mu}_{N_I,j}(f_I) \tag{B.1}$$

where for $j = 1, \ldots, N_{I^c}$

$$\widehat{\mu}_{N_I,j}(f_I) = \frac{1}{N_I} \sum_{u \in \mathbf{Y}_j} f_I(v)$$

and where $\mathbf{Y}_1, \ldots, \mathbf{Y}_{N_I,j}$ are iid (N_I, ι)-Dirichlet DPPs, that is $\widehat{\mu}_{N_I,j}$ is nothing else than an average of unbiased estimators of $\mu(f_I)$ based on an (N_I, ι)-Dirichlet DPP for which Theorem 3.1 can now be applied.

In particular, using Theorem 3.1 (ii), we have

$$\mathrm{Var}(\widehat{\mu}_N(f_I)) = \frac{1}{N_{I^c}}\mathrm{Var}(\widehat{\mu}_{N_I,1}(f_I)) \sim \frac{1}{N_{I^c}} \frac{1}{(N_I)^{1+1/\iota}}\varsigma^2(f_I)$$

as $N_I \to \infty$, with

$$\varsigma^2(f_I) = \sum_{j\in\mathbb{Z}^\iota} \left(\sum_{i\in I} \frac{|j_i|}{\gamma_i}\right)\left|\widehat{f}_I(j)\right|^2$$

where for any $i \in I$

$$\gamma_i = \lim_{N_I\to\infty} n_i N_I^{-1/\iota}.$$

Now, since

$$\gamma_i \underset{N_I\to\infty}{\sim} \kappa_i N^{1/d} N_I^{-1/\iota} \tag{B.2}$$

where κ_i are given by (3.8), we deduce that

$$\mathrm{Var}(\widehat{\mu}_N(f_I)) \sim \frac{1}{N^{1+1/d}} \sum_{j\in\mathbb{Z}^\iota} \sigma^2(f_I)$$

as $N_I \to \infty$ and where $\sigma^2(f_I)$ is given by (3.15), which yields the result.

(iii) We can observe from (B.1) that

$$\widehat{\mu}_N(f_I) - \mu(f_I) = \frac{1}{N_{I^c}} \sum_{j=1}^{N_{I^c}} (\widehat{\mu}_{N_I,j}(f_I) - \mu(f_I))$$

From Theorem 3.1 (ii) and for any $j \in I$

$$Z_{N_I,j} = \sqrt{N_I^{1+1/\iota}}\frac{\widehat{\mu}_{N_I,j}(f_I) - \mu(f_I)}{\varsigma(f_I)} \to N(0,1)$$

in distribution. Then we apply Lindeberg-Feller theorem (see *e.g.* Billingsley [5]) to establish that as $N_{I^c} \to \infty$

$$\frac{1}{\sqrt{N_{I^c}}} \sum_{j=1}^{N_I^c} Z_{N_I,j} \to N(0,1)$$

in distribution. Therefore, in distribution as $N \to \infty$

$$\sqrt{\frac{N_I^{1+1/\iota}}{N_{I^c}}} \; \frac{\widehat{\mu}_N(f_I) - \mu(f_I)}{\varsigma(f_I)} \; \to \; N(0,1)$$

whereby we deduce the result thanks to (B.2). $\qquad\qquad\qquad\qquad\square$

Appendix C: Simulation results

Figures 10-24 display simulation results for the logarithms of the empirical of MSE in terms of $\log(N)$ of estimates of d-dimensional integrals ($d = 1, 2, \ldots, 6$, 10, Figures 10-14) or ι-dimensional integrals using point projected point patterns generated in dimension $d = 6$ (Figures 15-19) or $d = 10$ (Figures 20-24). Results are presented for the five test functions given by (4.1)-(4.4) and for the `crude MC`, `stratified MC`, `maximinLHS`, `Sobol` and `Halton` designs, see Section 5 for more details.

Acknowledgments

f_{bump}
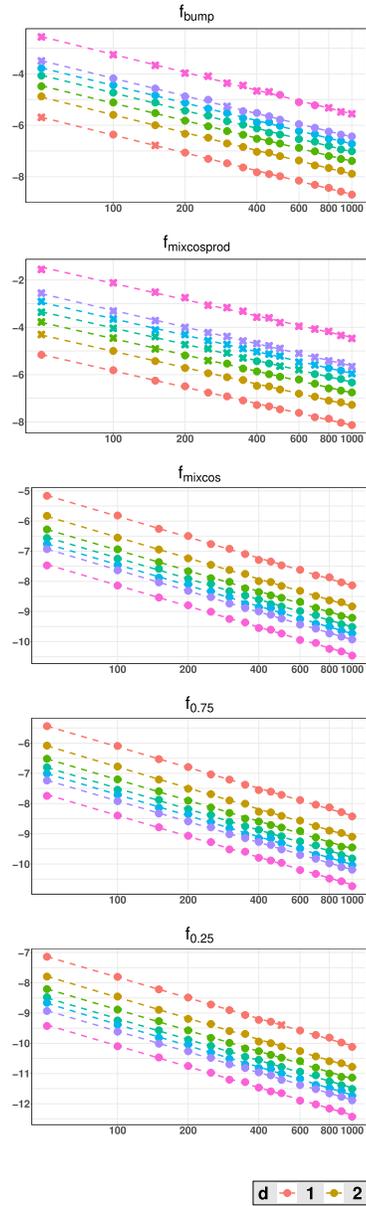
d	Expected	Estimation	Conf. interval
1	−1	−1.000	[−1.02 ; −0.986]
2	−1	−1.000	[−1.02 ; −0.982]
3	−1	−0.980	[−0.992 ; −0.967]
4	−1	−0.994	[−1.01 ; −0.977]
5	−1	−0.986	[−1.01 ; −0.967]
6	−1	−0.990	[−1.01 ; −0.973]
10	−1	−1.000	[−1.03 ; −0.977]

$f_{mixcosprod}$

d	Expected	Estimation	Conf. interval
1	−1	−0.998	[−1.02 ; −0.978]
2	−1	−0.995	[−1.01 ; −0.975]
3	−1	−0.994	[−1.01 ; −0.976]
4	−1	−0.992	[−1.01 ; −0.976]
5	−1	−1.000	[−1.02 ; −0.984]
6	−1	−1.020	[−1.03 ; −1]
10	−1	−0.992	[−1.03 ; −0.954]

f_{mixcos}

d	Expected	Estimation	Conf. interval
1	−1	−0.998	[−1.02 ; −0.978]
2	−1	−0.998	[−1.02 ; −0.98]
3	−1	−0.986	[−1 ; −0.968]
4	−1	−0.992	[−1.01 ; −0.974]
5	−1	−0.994	[−1.02 ; −0.973]
6	−1	−1.000	[−1.01 ; −0.994]
10	−1	−0.999	[−1.02 ; −0.981]

$f_{0.75}$

d	Expected	Estimation	Conf. interval
1	−1	−0.989	[−1.01 ; −0.971]
2	−1	−1.010	[−1.03 ; −0.985]
3	−1	−0.992	[−1.01 ; −0.969]
4	−1	−0.994	[−1.01 ; −0.975]
5	−1	−1.000	[−1.03 ; −0.981]
6	−1	−0.988	[−0.999 ; −0.976]
10	−1	−0.995	[−1.01 ; −0.975]

$f_{0.25}$

d	Expected	Estimation	Conf. interval
1	−1	−0.985	[−1 ; −0.969]
2	−1	−1.000	[−1.02 ; −0.983]
3	−1	−0.992	[−1.01 ; −0.969]
4	−1	−0.992	[−1.01 ; −0.972]
5	−1	−1.010	[−1.03 ; −0.992]
6	−1	−0.985	[−0.996 ; −0.973]
10	−1	−0.995	[−1.01 ; −0.974]

d	1	2	3	4	5	6	10

FIG 10. *Summary of experiments in which integrals of d-dimensional functions are estimated using a* `crude MC` *design. A ● (resp. ×) indicates that the adjusted p-value of the Shapiro-Wilk test is not smaller (resp. smaller) than 5%*
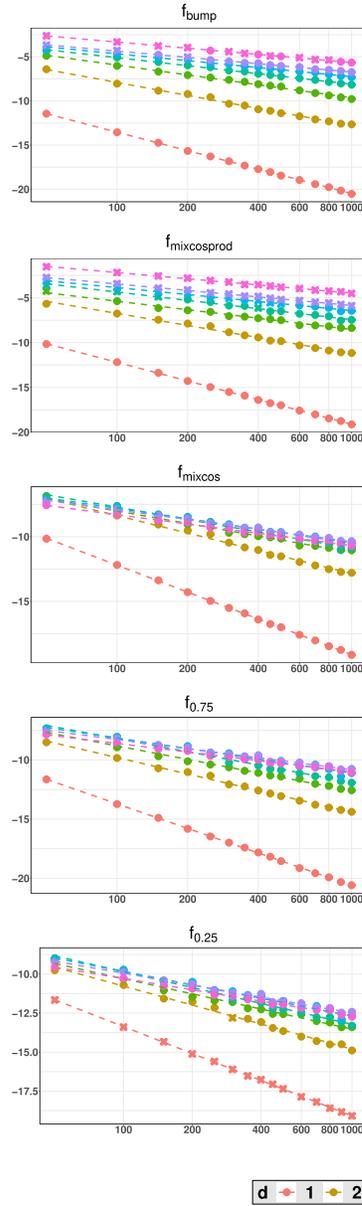
d	Expected	Estimation	Conf. interval
1	−3.00	−3.02	[−3.04 ; −3]
2	−2.00	−2.11	[−2.19 ; −2.04]
3	−1.67	−1.61	[−1.68 ; −1.54]
4	−1.50	−1.31	[−1.35 ; −1.26]
5	−1.40	−1.13	[−1.17 ; −1.09]
6	−1.33	−1.03	[−1.06 ; −1]
10	−1.20	−1.01	[−1.03 ; −0.982]

d	Expected	Estimation	Conf. interval
1	−3.00	−3.000	[−3.02 ; −2.98]
2	−2.00	−1.940	[−2.04 ; −1.84]
3	−1.67	−1.370	[−1.45 ; −1.29]
4	−1.50	−1.340	[−1.41 ; −1.26]
5	−1.40	−1.120	[−1.17 ; −1.06]
6	−1.33	−1.040	[−1.07 ; −1.02]
10	−1.20	−0.996	[−1.02 ; −0.97]

d	Expected	Estimation	Conf. interval
1	−3.00	−3.00	[−3.02 ; −2.98]
2	−2.00	−1.94	[−2.05 ; −1.84]
3	−1.67	−1.39	[−1.47 ; −1.3]
4	−1.50	−1.38	[−1.47 ; −1.29]
5	−1.40	−1.17	[−1.23 ; −1.1]
6	−1.33	−1.07	[−1.12 ; −1.02]
10	−1.20	−1.04	[−1.07 ; −0.997]

d	Expected	Estimation	Conf. interval
1	−3.00	−2.99	[−3.02 ; −2.97]
2	−2.00	−2.01	[−2.08 ; −1.93]
3	−1.67	−1.65	[−1.74 ; −1.56]
4	−1.50	−1.61	[−1.73 ; −1.49]
5	−1.40	−1.29	[−1.41 ; −1.16]
6	−1.33	−1.11	[−1.18 ; −1.04]
10	−1.20	−1.05	[−1.09 ; −0.997]

d	Expected	Estimation	Conf. interval
1	−3.00	−2.49	[−2.51 ; −2.46]
2	−2.00	−1.75	[−1.84 ; −1.66]
3	−1.67	−1.40	[−1.51 ; −1.3]
4	−1.50	−1.46	[−1.56 ; −1.37]
5	−1.40	−1.24	[−1.34 ; −1.13]
6	−1.33	−1.11	[−1.17 ; −1.04]
10	−1.20	−1.04	[−1.09 ; −0.995]

FIG 11. *Summary of experiments in which integrals of d-dimensional functions are estimated using a* `stratified MC` *design. A • (resp. ×) indicates that the adjusted p-value of the Shapiro-Wilk test is not smaller (resp. smaller) than 5%*
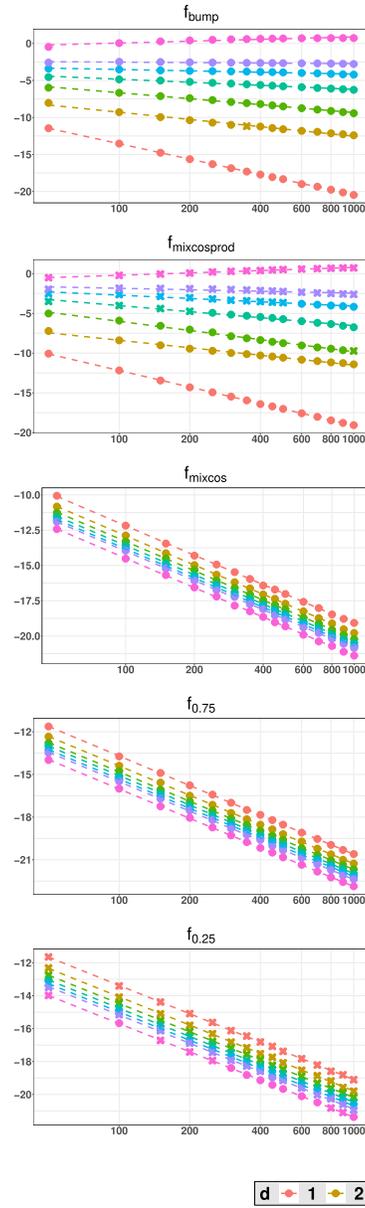
d	Expected	Estimation	Conf. interval
1	NA	−3.0000	[−3.02 ; −2.98]
2	NA	−1.4100	[−1.49 ; −1.32]
3	NA	−1.1600	[−1.21 ; −1.12]
4	NA	−0.5940	[−0.633 ; −0.555]
5	NA	−0.2680	[−0.3 ; −0.237]
6	NA	−0.0868	[−0.117 ; −0.0561]
10	NA	0.3610	[0.281 ; 0.441]

d	Expected	Estimation	Conf. interval
1	NA	−3.010	[−3.03 ; −2.99]
2	NA	−1.360	[−1.42 ; −1.29]
3	NA	−1.660	[−1.72 ; −1.6]
4	NA	−1.100	[−1.18 ; −1.02]
5	NA	−0.603	[−0.66 ; −0.545]
6	NA	−0.266	[−0.358 ; −0.174]
10	NA	0.426	[0.409 ; 0.443]

d	Expected	Estimation	Conf. interval
1	NA	−3.01	[−3.03 ; −2.99]
2	NA	−2.99	[−3.01 ; −2.98]
3	NA	−3.00	[−3.02 ; −2.98]
4	NA	−2.99	[−3.01 ; −2.97]
5	NA	−2.98	[−3 ; −2.97]
6	NA	−2.99	[−3.01 ; −2.97]
10	NA	−3.00	[−3.02 ; −2.98]

d	Expected	Estimation	Conf. interval
1	NA	−3.00	[−3.01 ; −2.99]
2	NA	−2.99	[−3 ; −2.97]
3	NA	−2.98	[−3 ; −2.96]
4	NA	−2.99	[−3.01 ; −2.97]
5	NA	−2.97	[−2.99 ; −2.96]
6	NA	−2.99	[−3.01 ; −2.97]
10	NA	−2.98	[−2.99 ; −2.97]

d	Expected	Estimation	Conf. interval
1	NA	−2.48	[−2.5 ; −2.47]
2	NA	−2.48	[−2.51 ; −2.46]
3	NA	−2.48	[−2.49 ; −2.46]
4	NA	−2.48	[−2.5 ; −2.45]
5	NA	−2.48	[−2.49 ; −2.46]
6	NA	−2.48	[−2.51 ; −2.45]
10	NA	−2.47	[−2.48 ; −2.45]

FIG 12. *Summary of experiments in which integrals of d-dimensional functions are estimated using a* `maximinLHS` *design. A ● (resp. ×) indicates that the adjusted p-value of the Shapiro-Wilk test is not smaller (resp. smaller) than 5%*
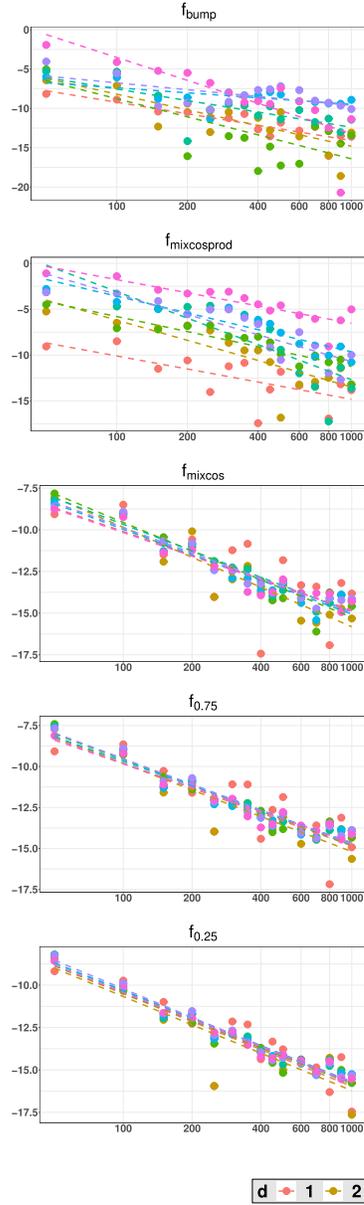
d	Expected	Estimation	Conf. interval
1	NA	−2.110	[−2.74 ; −1.48]
2	NA	−2.870	[−4.25 ; −1.5]
3	NA	−3.320	[−5.28 ; −1.35]
4	NA	−2.110	[−3.27 ; −0.961]
5	NA	−0.916	[−1.65 ; −0.18]
6	NA	−1.240	[−2.21 ; −0.282]
10	NA	−4.200	[−5.93 ; −2.47]

d	Expected	Estimation	Conf. interval
1	NA	−2.05	[−3.32 ; −0.778]
2	NA	−3.16	[−4.44 ; −1.89]
3	NA	−2.36	[−3.01 ; −1.72]
4	NA	−4.17	[−5.88 ; −2.47]
5	NA	−2.63	[−3.32 ; −1.94]
6	NA	−3.19	[−4.22 ; −2.15]
10	NA	−2.07	[−2.77 ; −1.36]

d	Expected	Estimation	Conf. interval
1	NA	−2.05	[−3.32 ; −0.778]
2	NA	−2.59	[−3.23 ; −1.95]
3	NA	−2.48	[−2.98 ; −1.97]
4	NA	−2.29	[−2.63 ; −1.95]
5	NA	−2.21	[−2.66 ; −1.76]
6	NA	−2.19	[−2.57 ; −1.81]
10	NA	−2.04	[−2.43 ; −1.65]

d	Expected	Estimation	Conf. interval
1	NA	−2.18	[−3.13 ; −1.23]
2	NA	−2.34	[−2.89 ; −1.78]
3	NA	−2.27	[−2.65 ; −1.89]
4	NA	−2.25	[−2.58 ; −1.91]
5	NA	−2.16	[−2.56 ; −1.77]
6	NA	−2.27	[−2.63 ; −1.91]
10	NA	−2.09	[−2.43 ; −1.75]

d	Expected	Estimation	Conf. interval
1	NA	−2.37	[−3.22 ; −1.52]
2	NA	−2.41	[−3.16 ; −1.66]
3	NA	−2.33	[−2.7 ; −1.95]
4	NA	−2.35	[−2.65 ; −2.06]
5	NA	−2.34	[−2.64 ; −2.04]
6	NA	−2.43	[−2.7 ; −2.15]
10	NA	−2.31	[−2.54 ; −2.08]

FIG 13. *Summary of experiments in which integrals of d-dimensional functions are estimated using a* `Sobol` *design.*
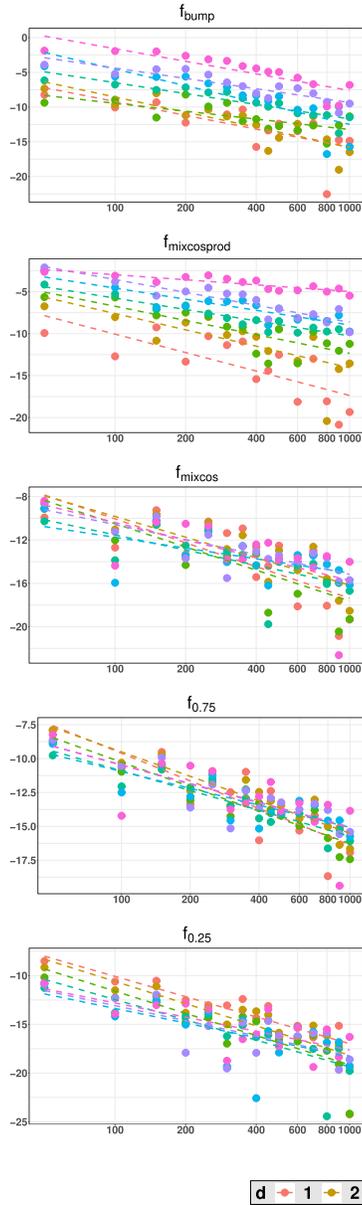
d	Expected	Estimation	Conf. interval
1	NA	−2.87	[−4.6 ; −1.14]
2	NA	−3.14	[−4.42 ; −1.85]
3	NA	−1.64	[−2.6 ; −0.679]
4	NA	−2.26	[−2.88 ; −1.65]
5	NA	−3.46	[−4.86 ; −2.05]
6	NA	−2.15	[−2.8 ; −1.5]
10	NA	−2.61	[−3.55 ; −1.67]

d	Expected	Estimation	Conf. interval
1	NA	−3.18	[−4.89 ; −1.46]
2	NA	−2.80	[−4.47 ; −1.14]
3	NA	−2.44	[−3.48 ; −1.4]
4	NA	−1.95	[−2.32 ; −1.59]
5	NA	−1.91	[−2.49 ; −1.33]
6	NA	−2.21	[−2.81 ; −1.61]
10	NA	−0.87	[−1.17 ; −0.572]

d	Expected	Estimation	Conf. interval
1	NA	−3.18	[−4.89 ; −1.46]
2	NA	−2.75	[−3.87 ; −1.63]
3	NA	−3.10	[−4.53 ; −1.66]
4	NA	−2.02	[−3.28 ; −0.758]
5	NA	−1.43	[−2.62 ; −0.237]
6	NA	−1.99	[−2.81 ; −1.18]
10	NA	−2.32	[−4.06 ; −0.587]

d	Expected	Estimation	Conf. interval
1	NA	−2.94	[−4.02 ; −1.87]
2	NA	−2.61	[−3.25 ; −1.96]
3	NA	−2.61	[−3.26 ; −1.96]
4	NA	−2.10	[−2.63 ; −1.56]
5	NA	−1.79	[−2.55 ; −1.04]
6	NA	−1.99	[−2.72 ; −1.26]
10	NA	−2.01	[−3.31 ; −0.711]

d	Expected	Estimation	Conf. interval
1	NA	−2.99	[−3.75 ; −2.24]
2	NA	−3.29	[−4.65 ; −1.93]
3	NA	−3.28	[−4.53 ; −2.04]
4	NA	−2.99	[−4.27 ; −1.72]
5	NA	−2.15	[−3.76 ; −0.547]
6	NA	−2.23	[−3.55 ; −0.918]
10	NA	−2.09	[−3.29 ; −0.901]

FIG 14. *Summary of experiments in which integrals of d-dimensional functions are estimated using a* `Halton` *design.*

ι	Expected	Estimation	Conf. interval
6	−1	−0.990	[−1.01 ; −0.973]
5	−1	−0.991	[−1.01 ; −0.971]
4	−1	−0.985	[−0.998 ; −0.971]
3	−1	−1.000	[−1.02 ; −0.982]
2	−1	−1.020	[−1.03 ; −0.999]
1	−1	−0.995	[−1.01 ; −0.978]

ι	Expected	Estimation	Conf. interval
6	−1	−1.020	[−1.03 ; −1]
5	−1	−1.010	[−1.03 ; −0.99]
4	−1	−1.000	[−1.02 ; −0.989]
3	−1	−0.997	[−1.01 ; −0.98]
2	−1	−0.996	[−1.02 ; −0.976]
1	−1	−1.010	[−1.03 ; −0.982]

ι	Expected	Estimation	Conf. interval
6	−1	−1.000	[−1.01 ; −0.994]
5	−1	−0.999	[−1.01 ; −0.986]
4	−1	−1.000	[−1.02 ; −0.985]
3	−1	−0.990	[−1.01 ; −0.971]
2	−1	−1.010	[−1.03 ; −0.989]
1	−1	−0.981	[−0.996 ; −0.966]

ι	Expected	Estimation	Conf. interval
6	−1	−0.988	[−0.999 ; −0.976]
5	−1	−0.987	[−1 ; −0.971]
4	−1	−1.000	[−1.01 ; −0.987]
3	−1	−1.010	[−1.02 ; −0.998]
2	−1	−0.997	[−1.01 ; −0.98]
1	−1	−1.010	[−1.03 ; −0.991]

ι	Expected	Estimation	Conf. interval
6	−1	−0.985	[−0.996 ; −0.973]
5	−1	−0.978	[−0.995 ; −0.962]
4	−1	−1.010	[−1.02 ; −0.993]
3	−1	−1.000	[−1.02 ; −0.981]
2	−1	−1.000	[−1.02 ; −0.985]
1	−1	−1.000	[−1.02 ; −0.983]

FIG 15. *Summary of experiments in which integrals of ι-dimensional functions are estimated by projecting a single 6-dimensional* crude MC *design ($\iota = 1\ldots6$). A $\bullet$ (resp. $\times$) indicates that the adjusted p-value of the Shapiro-Wilk test is not smaller (resp. smaller) than 5%.*

ι	Expected	Estimation	Conf. interval
6	−1.33	−1.03	[−1.06 ; −1]
5	NA	−1.04	[−1.07 ; −1.01]
4	NA	−1.03	[−1.06 ; −0.999]
3	NA	−1.05	[−1.09 ; −1.01]
2	NA	−1.06	[−1.11 ; −1.02]
1	NA	−1.05	[−1.09 ; −1.01]

ι	Expected	Estimation	Conf. interval
6	−1.33	−1.04	[−1.07 ; −1.02]
5	NA	−1.04	[−1.08 ; −1.01]
4	NA	−1.05	[−1.09 ; −1]
3	NA	−1.04	[−1.08 ; −0.999]
2	NA	−1.06	[−1.11 ; −1.02]
1	NA	−1.07	[−1.12 ; −1.02]

ι	Expected	Estimation	Conf. interval
6	−1.33	−1.07	[−1.12 ; −1.02]
5	NA	−1.07	[−1.11 ; −1.02]
4	NA	−1.08	[−1.13 ; −1.03]
3	NA	−1.08	[−1.13 ; −1.02]
2	NA	−1.08	[−1.13 ; −1.03]
1	NA	−1.08	[−1.14 ; −1.02]

ι	Expected	Estimation	Conf. interval
6	−1.33	−1.11	[−1.18 ; −1.04]
5	NA	−1.11	[−1.18 ; −1.04]
4	NA	−1.10	[−1.18 ; −1.03]
3	NA	−1.11	[−1.18 ; −1.05]
2	NA	−1.10	[−1.17 ; −1.03]
1	NA	−1.11	[−1.18 ; −1.04]

ι	Expected	Estimation	Conf. interval
6	−1.33	−1.11	[−1.17 ; −1.04]
5	NA	−1.10	[−1.17 ; −1.04]
4	NA	−1.10	[−1.17 ; −1.03]
3	NA	−1.11	[−1.19 ; −1.04]
2	NA	−1.10	[−1.18 ; −1.02]
1	NA	−1.12	[−1.19 ; −1.05]

FIG 16. *Summary of experiments in which integrals of ι-dimensional functions are estimated by projecting a single 6-dimensional* `stratified MC` *design ($\iota = 1 \dots 6$). A • (resp. ×) indicates that the adjusted p-value of the Shapiro-Wilk test is not smaller (resp. smaller) than 5%.*

f_{bump}
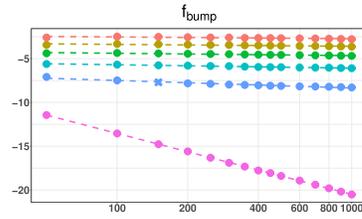
ι	Expected	Estimation	Conf. interval
6	NA	−0.0868	[−0.117 ; −0.0561]
5	NA	−0.0922	[−0.123 ; −0.0614]
4	NA	−0.1170	[−0.137 ; −0.0955]
3	NA	−0.1680	[−0.18 ; −0.156]
2	NA	−0.3730	[−0.417 ; −0.33]
1	NA	−3.0200	[−3.04 ; −3]

$f_{mixcosprod}$
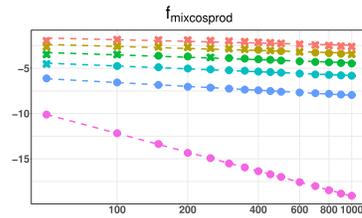
ι	Expected	Estimation	Conf. interval
6	NA	−0.266	[−0.358 ; −0.174]
5	NA	−0.322	[−0.399 ; −0.246]
4	NA	−0.378	[−0.434 ; −0.322]
3	NA	−0.457	[−0.486 ; −0.428]
2	NA	−0.612	[−0.628 ; −0.595]
1	NA	−3.010	[−3.03 ; −2.99]

f_{mixcos}
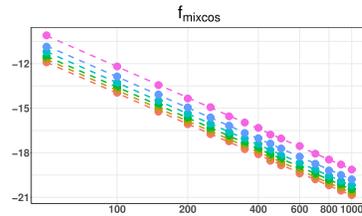
ι	Expected	Estimation	Conf. interval
6	NA	−2.99	[−3.01 ; −2.97]
5	NA	−3.00	[−3.02 ; −2.98]
4	NA	−2.99	[−3 ; −2.98]
3	NA	−3.01	[−3.02 ; −2.99]
2	NA	−3.00	[−3.02 ; −2.98]
1	NA	−3.01	[−3.03 ; −2.98]

$f_{0.75}$
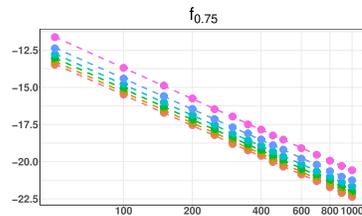
ι	Expected	Estimation	Conf. interval
6	NA	−2.99	[−3.01 ; −2.97]
5	NA	−2.98	[−3 ; −2.96]
4	NA	−2.99	[−3 ; −2.97]
3	NA	−2.99	[−3.01 ; −2.97]
2	NA	−2.98	[−3.01 ; −2.96]
1	NA	−3.00	[−3.02 ; −2.99]

$f_{0.25}$
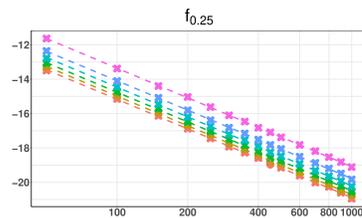
ι	Expected	Estimation	Conf. interval
6	NA	−2.48	[−2.51 ; −2.45]
5	NA	−2.49	[−2.51 ; −2.47]
4	NA	−2.49	[−2.52 ; −2.46]
3	NA	−2.49	[−2.51 ; −2.47]
2	NA	−2.47	[−2.49 ; −2.45]
1	NA	−2.49	[−2.5 ; −2.47]

ι ● 6 ● 5 ● 4 ● 3 ● 2 ● 1

FIG 17. *Summary of experiments in which integrals of ι-dimensional functions are estimated by projecting a single 6-dimensional* `maximinLHS` *design (ι = 1 . . . 6). A ● (resp. ×) indicates that the adjusted p-value of the Shapiro-Wilk test is not smaller (resp. smaller) than 5%.*

f_{bump}

| ι	Expected	Estimation	Conf. interval
6	NA	−1.24	[−2.21 ; −0.282]
5	NA	−1.24	[−3.31 ; 0.828]
4	NA	−2.50	[−3.6 ; −1.39]
3	NA	−3.08	[−4.48 ; −1.68]
2	NA	−4.03	[−5.28 ; −2.79]
1	NA	−2.65	[−3.63 ; −1.68]

$f_{mixcosprod}$

ι	Expected	Estimation	Conf. interval
6	NA	−3.19	[−4.22 ; −2.15]
5	NA	−2.43	[−3.93 ; −0.921]
4	NA	−1.46	[−2.52 ; −0.397]
3	NA	−2.58	[−3.98 ; −1.19]
2	NA	−2.02	[−4.43 ; 0.385]
1	NA	−2.28	[−3.45 ; −1.11]

f_{mixcos}

ι	Expected	Estimation	Conf. interval
6	NA	−2.19	[−2.57 ; −1.81]
5	NA	−2.23	[−2.7 ; −1.77]
4	NA	−2.39	[−2.93 ; −1.85]
3	NA	−2.30	[−2.93 ; −1.66]
2	NA	−2.64	[−4.32 ; −0.957]
1	NA	−1.27	[−2.81 ; 0.265]

$f_{0.75}$

ι	Expected	Estimation	Conf. interval
6	NA	−2.27	[−2.63 ; −1.91]
5	NA	−2.34	[−2.75 ; −1.92]
4	NA	−2.05	[−2.53 ; −1.58]
3	NA	−2.38	[−2.86 ; −1.9]
2	NA	−2.00	[−2.88 ; −1.12]
1	NA	−2.34	[−3.58 ; −1.09]

$f_{0.25}$

ι	Expected	Estimation	Conf. interval
6	NA	−2.43	[−2.7 ; −2.15]
5	NA	−2.35	[−2.64 ; −2.05]
4	NA	−2.42	[−2.72 ; −2.11]
3	NA	−2.47	[−3.03 ; −1.91]
2	NA	−2.55	[−2.94 ; −2.16]
1	NA	−2.73	[−3.7 ; −1.75]

ι ● 6 ● 5 ● 4 ● 3 ● 2 ● 1

FIG 18. *Summary of experiments in which integrals of ι-dimensional functions are estimated by projecting a single 6-dimensional* Sobol *design ($\iota = 1 \ldots 6$).*

ι	Expected	Estimation	Conf. interval
6	NA	−2.15	[−2.8 ; −1.5]
5	NA	−2.42	[−3.98 ; −0.865]
4	NA	−1.91	[−2.79 ; −1.03]
3	NA	−2.64	[−4.07 ; −1.21]
2	NA	−2.42	[−3.38 ; −1.45]
1	NA	−2.35	[−6.63 ; 1.93]

ι	Expected	Estimation	Conf. interval
6	NA	−2.21	[−2.81 ; −1.61]
5	NA	−1.92	[−2.55 ; −1.29]
4	NA	−1.97	[−2.86 ; −1.07]
3	NA	−1.43	[−3.03 ; 0.167]
2	NA	−1.16	[−3.04 ; 0.718]
1	NA	−1.99	[−4.13 ; 0.152]

ι	Expected	Estimation	Conf. interval
6	NA	−1.99	[−2.81 ; −1.18]
5	NA	−2.08	[−3.41 ; −0.738]
4	NA	−2.33	[−3.85 ; −0.797]
3	NA	−2.66	[−4.15 ; −1.17]
2	NA	−1.44	[−2.11 ; −0.78]
1	NA	−2.46	[−4.9 ; −0.0276]

ι	Expected	Estimation	Conf. interval
6	NA	−1.99	[−2.72 ; −1.26]
5	NA	−1.96	[−3.02 ; −0.888]
4	NA	−2.05	[−3.15 ; −0.948]
3	NA	−2.35	[−3.51 ; −1.19]
2	NA	−1.31	[−2.88 ; 0.251]
1	NA	−1.07	[−4.29 ; 2.14]

| ι	Expected	Estimation	Conf. interval
6	NA	−2.23	[−3.55 ; −0.918]
5	NA	−2.49	[−3.78 ; −1.21]
4	NA	−2.46	[−3.66 ; −1.27]
3	NA	−2.22	[−3.44 ; −0.994]
2	NA	−2.74	[−4 ; −1.47]
1	NA	−1.88	[−3.27 ; −0.497]

Fig 19. *Summary of experiments in which integrals of ι-dimensional functions are estimated by projecting a single 6-dimensional* `Halton` *design ($\iota = 1 \ldots 6$).*

f_{bump}
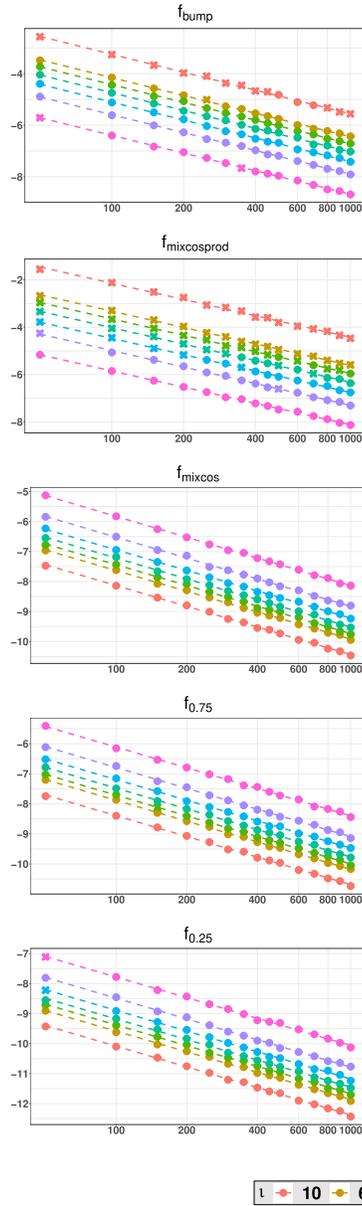
ι	Expected	Estimation	Conf. interval
10	−1	−1.000	[−1.03 ; −0.977]
6	−1	−0.996	[−1.02 ; −0.976]
5	−1	−0.997	[−1.01 ; −0.981]
4	−1	−0.996	[−1.02 ; −0.975]
3	−1	−1.000	[−1.02 ; −0.986]
2	−1	−1.000	[−1.02 ; −0.988]
1	−1	−0.991	[−1.01 ; −0.974]

$f_{\mathrm{mixcosprod}}$

ι	Expected	Estimation	Conf. interval
10	−1	−0.992	[−1.03 ; −0.954]
6	−1	−0.996	[−1.01 ; −0.978]
5	−1	−0.999	[−1.02 ; −0.983]
4	−1	−1.000	[−1.02 ; −0.985]
3	−1	−0.992	[−1.01 ; −0.974]
2	−1	−1.000	[−1.03 ; −0.983]
1	−1	−0.986	[−1 ; −0.969]

f_{mixcos}

ι	Expected	Estimation	Conf. interval
10	−1	−0.999	[−1.02 ; −0.981]
6	−1	−0.993	[−1.01 ; −0.976]
5	−1	−0.998	[−1.02 ; −0.976]
4	−1	−1.000	[−1.02 ; −0.98]
3	−1	−0.988	[−1 ; −0.972]
2	−1	−1.000	[−1.03 ; −0.979]
1	−1	−1.000	[−1.02 ; −0.988]

$f_{0.75}$

ι	Expected	Estimation	Conf. interval
10	−1	−0.995	[−1.01 ; −0.975]
6	−1	−0.996	[−1.01 ; −0.982]
5	−1	−0.993	[−1.01 ; −0.978]
4	−1	−0.994	[−1.01 ; −0.977]
3	−1	−0.989	[−1.01 ; −0.969]
2	−1	−0.999	[−1.02 ; −0.977]
1	−1	−0.992	[−1.01 ; −0.971]

$f_{0.25}$

ι	Expected	Estimation	Conf. interval
10	−1	−0.995	[−1.01 ; −0.974]
6	−1	−0.989	[−1.01 ; −0.967]
5	−1	−0.993	[−1.01 ; −0.974]
4	−1	−0.987	[−1.01 ; −0.965]
3	−1	−0.993	[−1.01 ; −0.971]
2	−1	−0.998	[−1.02 ; −0.976]
1	−1	−0.990	[−1.02 ; −0.965]

ι 10   6   5   4   3   2   1

FIG 20. *Summary of experiments in which integrals of ι-dimensional functions are estimated by projecting a single 10-dimensional* `crude MC` *design ($\iota = 1 \ldots 6, 10$). A ● (resp. ×) indicates that the adjusted p-value of the Shapiro-Wilk test is not smaller (resp. smaller) than 5%.*
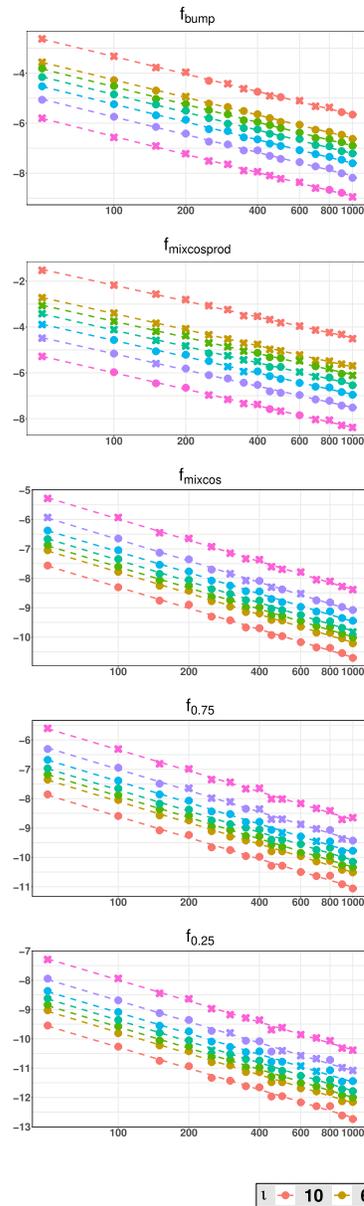
ι	Expected	Estimation	Conf. interval
10	−1.2	−1.01	[−1.03 ; −0.982]
6	NA	−1.01	[−1.03 ; −0.993]
5	NA	−1.03	[−1.05 ; −1]
4	NA	−1.02	[−1.05 ; −0.997]
3	NA	−1.02	[−1.05 ; −0.997]
2	NA	−1.02	[−1.06 ; −0.991]
1	NA	−1.04	[−1.06 ; −1.02]

ι	Expected	Estimation	Conf. interval
10	−1.2	−0.996	[−1.02 ; −0.97]
6	NA	−1.010	[−1.03 ; −0.985]
5	NA	−1.010	[−1.04 ; −0.98]
4	NA	−1.020	[−1.05 ; −0.994]
3	NA	−1.020	[−1.06 ; −0.989]
2	NA	−1.020	[−1.05 ; −0.985]
1	NA	−1.030	[−1.06 ; −1.01]

ι	Expected	Estimation	Conf. interval
10	−1.2	−1.04	[−1.07 ; −0.997]
6	NA	−1.04	[−1.07 ; −1]
5	NA	−1.03	[−1.07 ; −0.995]
4	NA	−1.05	[−1.08 ; −1.01]
3	NA	−1.03	[−1.07 ; −0.993]
2	NA	−1.05	[−1.08 ; −1.02]
1	NA	−1.04	[−1.07 ; −1.01]

ι	Expected	Estimation	Conf. interval
10	−1.2	−1.05	[−1.09 ; −0.997]
6	NA	−1.05	[−1.09 ; −1]
5	NA	−1.05	[−1.09 ; −1.01]
4	NA	−1.04	[−1.09 ; −0.986]
3	NA	−1.05	[−1.09 ; −1]
2	NA	−1.05	[−1.09 ; −1]
1	NA	−1.03	[−1.08 ; −0.985]

| ι	Expected	Estimation	Conf. interval
10	−1.2	−1.04	[−1.09 ; −0.995]
6	NA	−1.04	[−1.09 ; −0.994]
5	NA	−1.05	[−1.09 ; −0.996]
4	NA	−1.05	[−1.09 ; −1.01]
3	NA	−1.03	[−1.08 ; −0.975]
2	NA	−1.04	[−1.09 ; −0.988]
1	NA	−1.04	[−1.08 ; −0.994]

FIG 21. *Summary of experiments in which integrals of ι-dimensional functions are estimated by projecting a single 10-dimensional* stratified MC *design (ι = 1 . . . 6, 10). A* ● *(resp.* ×*) indicates that the adjusted p-value of the Shapiro-Wilk test is not smaller (resp. smaller) than 5%.*
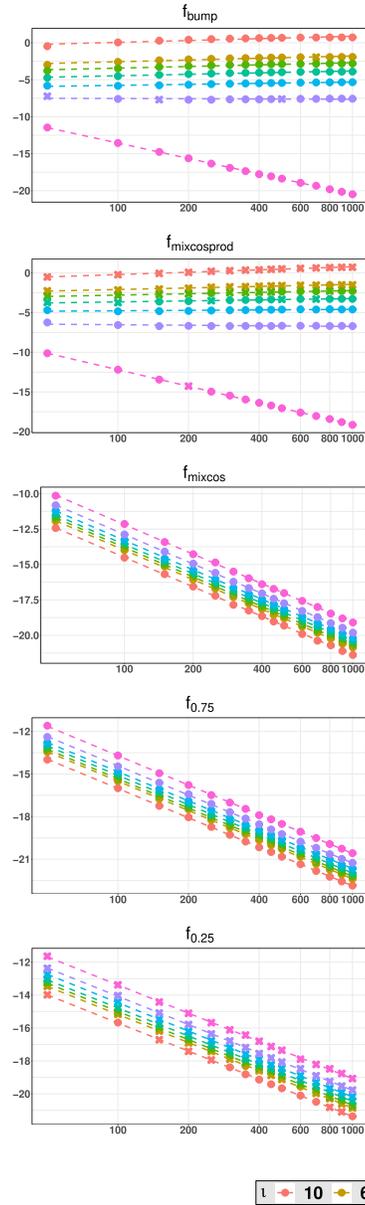
ι	Expected	Estimation	Conf. interval
10	NA	0.3610	[0.281 ; 0.441]
6	NA	0.3380	[0.277 ; 0.398]
5	NA	0.3190	[0.271 ; 0.367]
4	NA	0.2820	[0.248 ; 0.315]
3	NA	0.1940	[0.166 ; 0.222]
2	NA	−0.0413	[−0.117 ; 0.0347]
1	NA	−3.0000	[−3.03 ; −2.98]

ι	Expected	Estimation	Conf. interval
10	NA	0.4260	[0.409 ; 0.443]
6	NA	0.2820	[0.257 ; 0.307]
5	NA	0.2460	[0.224 ; 0.267]
4	NA	0.1830	[0.155 ; 0.211]
3	NA	0.0787	[0.0406 ; 0.117]
2	NA	−0.0943	[−0.153 ; −0.0352]
1	NA	−3.0000	[−3.02 ; −2.99]

ι	Expected	Estimation	Conf. interval
10	NA	−3.00	[−3.02 ; −2.98]
6	NA	−3.00	[−3.02 ; −2.98]
5	NA	−3.01	[−3.03 ; −2.98]
4	NA	−2.99	[−3.01 ; −2.97]
3	NA	−3.00	[−3.02 ; −2.98]
2	NA	−3.01	[−3.02 ; −2.99]
1	NA	−3.01	[−3.03 ; −2.98]

ι	Expected	Estimation	Conf. interval
10	NA	−2.98	[−2.99 ; −2.97]
6	NA	−2.99	[−3.01 ; −2.97]
5	NA	−2.99	[−3 ; −2.97]
4	NA	−2.99	[−3.01 ; −2.97]
3	NA	−2.97	[−2.98 ; −2.95]
2	NA	−2.98	[−3 ; −2.96]
1	NA	−2.99	[−3.01 ; −2.98]

ι	Expected	Estimation	Conf. interval
10	NA	−2.47	[−2.48 ; −2.45]
6	NA	−2.47	[−2.48 ; −2.45]
5	NA	−2.47	[−2.49 ; −2.45]
4	NA	−2.46	[−2.48 ; −2.45]
3	NA	−2.49	[−2.5 ; −2.47]
2	NA	−2.48	[−2.5 ; −2.47]
1	NA	−2.47	[−2.5 ; −2.45]

FIG 22. *Summary of experiments in which integrals of ι-dimensional functions are estimated by projecting a single 10-dimensional* `maximinLHS` *design ($\iota = 1\ldots6, 10$). A ● (resp. ×) indicates that the adjusted p-value of the Shapiro-Wilk test is not smaller (resp. smaller) than 5%.*
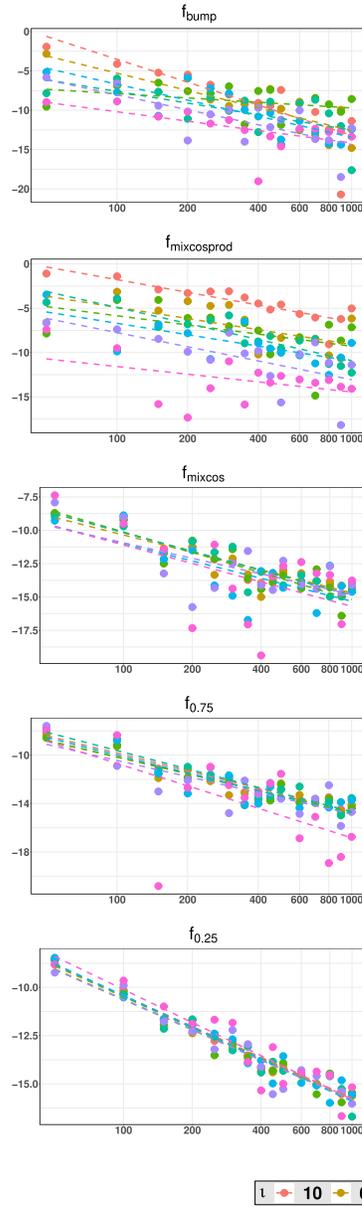
ι	Expected	Estimation	Conf. interval
10	NA	−4.200	[−5.93 ; −2.47]
6	NA	−3.170	[−4.32 ; −2.02]
5	NA	−0.791	[−2.08 ; 0.499]
4	NA	−2.130	[−3.78 ; −0.484]
3	NA	−2.930	[−4.08 ; −1.78]
2	NA	−2.750	[−4.25 ; −1.25]
1	NA	−1.720	[−3.1 ; −0.341]

ι	Expected	Estimation	Conf. interval
10	NA	−2.07	[−2.77 ; −1.36]
6	NA	−1.82	[−3.17 ; −0.469]
5	NA	−1.49	[−3.12 ; 0.141]
4	NA	−2.61	[−3.37 ; −1.84]
3	NA	−1.85	[−2.99 ; −0.716]
2	NA	−2.31	[−3.95 ; −0.66]
1	NA	−1.25	[−2.82 ; 0.331]

ι	Expected	Estimation	Conf. interval
10	NA	−2.04	[−2.43 ; −1.65]
6	NA	−1.89	[−2.48 ; −1.29]
5	NA	−2.20	[−2.86 ; −1.53]
4	NA	−2.03	[−2.58 ; −1.49]
3	NA	−1.88	[−2.94 ; −0.813]
2	NA	−1.68	[−2.81 ; −0.555]
1	NA	−2.00	[−3.88 ; −0.123]

ι	Expected	Estimation	Conf. interval
10	NA	−2.09	[−2.43 ; −1.75]
6	NA	−1.97	[−2.41 ; −1.52]
5	NA	−2.00	[−2.46 ; −1.54]
4	NA	−2.19	[−2.72 ; −1.66]
3	NA	−2.11	[−2.76 ; −1.46]
2	NA	−1.92	[−2.77 ; −1.08]
1	NA	−2.67	[−4.78 ; −0.551]

ι	Expected	Estimation	Conf. interval
10	NA	−2.31	[−2.54 ; −2.08]
6	NA	−2.30	[−2.54 ; −2.06]
5	NA	−2.26	[−2.59 ; −1.93]
4	NA	−2.36	[−2.67 ; −2.04]
3	NA	−2.37	[−2.72 ; −2.03]
2	NA	−2.25	[−2.73 ; −1.78]
1	NA	−2.47	[−3.04 ; −1.9]

ι — 10 — 6 — 5 — 4 — 3 — 2 — 1

FIG 23. *Summary of experiments in which integrals of ι-dimensional functions are estimated by projecting a single 10-dimensional* Sobol *design ($\iota = 1 \ldots 6, 10$).*
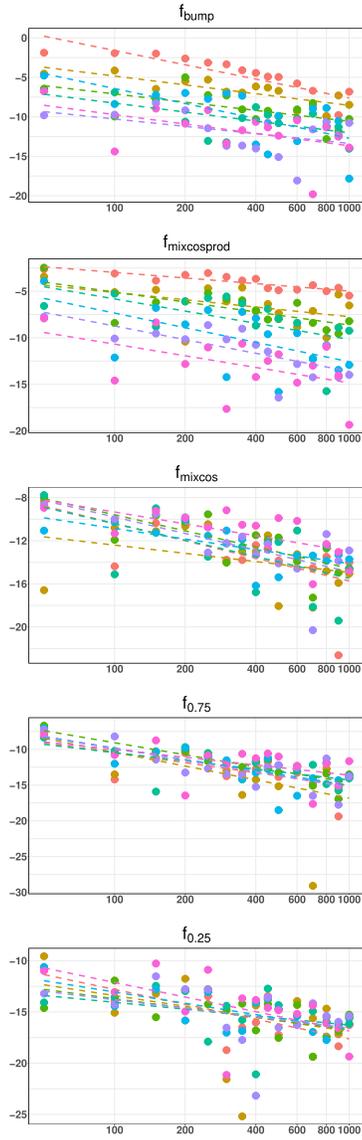
ι	Expected	Estimation	Conf. interval
10	NA	−2.61	[−3.55 ; −1.67]
6	NA	−1.61	[−2.28 ; −0.945]
5	NA	−1.50	[−2.63 ; −0.381]
4	NA	−1.63	[−2.8 ; −0.468]
3	NA	−2.73	[−4.61 ; −0.853]
2	NA	−1.34	[−2.97 ; 0.282]
1	NA	−1.71	[−3.53 ; 0.106]

ι	Expected	Estimation	Conf. interval
10	NA	−0.87	[−1.17 ; −0.572]
6	NA	−1.13	[−2.38 ; 0.116]
5	NA	−1.54	[−2.44 ; −0.646]
4	NA	−1.89	[−3.49 ; −0.287]
3	NA	−2.27	[−4.15 ; −0.386]
2	NA	−2.12	[−3.3 ; −0.928]
1	NA	−1.81	[−3.68 ; 0.0686]

ι	Expected	Estimation	Conf. interval
10	NA	−2.32	[−4.06 ; −0.587]
6	NA	−1.10	[−2.82 ; 0.624]
5	NA	−2.16	[−3.26 ; −1.07]
4	NA	−2.20	[−4.06 ; −0.337]
3	NA	−1.42	[−2.35 ; −0.484]
2	NA	−2.20	[−3.67 ; −0.741]
1	NA	−1.59	[−2.71 ; −0.462]

ι	Expected	Estimation	Conf. interval
10	NA	−2.01	[−3.31 ; −0.711]
6	NA	−2.79	[−5.7 ; 0.115]
5	NA	−2.43	[−3.2 ; −1.67]
4	NA	−1.66	[−2.86 ; −0.472]
3	NA	−2.25	[−3.67 ; −0.817]
2	NA	−2.39	[−3.66 ; −1.13]
1	NA	−1.56	[−3.15 ; 0.0382]

ι	Expected	Estimation	Conf. interval
10	NA	−2.090	[−3.29 ; −0.901]
6	NA	−1.520	[−4.03 ; 0.994]
5	NA	−1.310	[−2.44 ; −0.193]
4	NA	−0.963	[−2.42 ; 0.497]
3	NA	−1.590	[−2.57 ; −0.614]
2	NA	−1.200	[−3.02 ; 0.622]
1	NA	−2.060	[−3.72 ; −0.391]

FIG 24. *Summary of experiments in which integrals of ι-dimensional functions are estimated by projecting a single 10-dimensional* `Halton` *design (ι = 1 . . . 6, 10).*

References

[1] BADDELEY, A. (2007). *Spatial Point Processes and their Applications* In *Stochastic geomery: Lectures given the C.I.M.E. Summer School held in Martina Franca, Italy, September 13–18, 2004* 1–75. Springer, Berlin, Heidelberg. MR2327289

[2] BADDELEY, A., RUBAK, E. and TURNER, R. (2015). *Spatial point patterns: methodology and applications with R.* Chapman and Hall/CRC.

[3] BARDENET, R. and HARDY, A. (2020). Monte Carlo with determinantal point processes. *Ann. Appl. Probab.* **30** 368–417. MR4068314

[4] BASU, K. and MUKHERJEE, R. (2017). Asymptotic normality of scrambled geometric net quadrature. *The Annals of Statistics* **45** 1759–1788. MR3670195

[5] BILLINGSLEY, P. (2008). *Probability and measure.* John Wiley & Sons. MR0534323

[6] DEBNATH, L. and MIKUSIŃSKI, P. (2005). *Introduction to Hilbert spaces with applications*, 3rd ed. Elsevier Academic Press. MR1058201

[7] DECREUSEFOND, L., FLINT, I., PRIVAULT, N. and TORRIS, G. L. (2016). *Determinantal point processes* In *Stochastic Analysis for Poisson Point Processes* 311–342. Springer International Publishing, Cham. MR3585405

[8] DICK, J., KUO, F. Y. and SLOAN, I. H. (2013). High-dimensional integration: the quasi-Monte Carlo way. *Acta Numerica* **22** 133–288. MR3038697

[9] DOUC, R., MOULINES, E. and STOFFER, D. (2014). *Nonlinear time series: Theory, methods and applications with R examples.* Chapman and Hall/CRC. MR3289095

[10] GAUTIER, G., POLITO, G., BARDENET, R. and VALKO, M. DPPy: DPP Sampling with Python. *Journal of Machine Learning Research - Machine Learning Open Source Software (JMLR-MLOSS).* Code at http://github.com/guilgautier/DPPy/ – Documentation at http://dppy.readthedocs.io/.

[11] HABER, S. (1966). A modified Monte-Carlo quadrature. *Mathematics of Computation* **20** 361–368. MR0210285

[12] HALTON, J. H. (1964). Algorithm 247: Radical-inverse Quasi-random Point Sequence. *Commun. ACM* **7** 701–702.

[13] HOLM, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics* 65–70. MR0538597

[14] HOUGH, J. B., KRISHNAPUR, M., PERES, Y. and VIRÁG, B. (2006). Determinantal Processes and Independence. *Probability Surveys* **3** 206-229. MR2216966

[15] HOUGH, J. B., KRISHNAPUR, M., PERES, Y. and VIRÁG, B. (2009). *Zeros of Gaussian Analytic Functions and Determinantal Point Processes.* American Mathematical Soc. MR2552864

[16] ILLIAN, J., PENTTINEN, A., STOYAN, H. and STOYAN, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns. Statistics in Practice.* Wiley, Chichester. MR2384630

[17] LAVANCIER, F., MØLLER, J. and RUBAK, E. (2015). Determinantal point

process models and statistical inference: Extended version. *J. R. Stat. Soc. B* **77** 853–877. MR3382600

[18] LAVANCIER, F. and RUBAK, E. A bound on the conditional probability density function for the Fourier basis. unpublished.

[19] MAZOYER, A., COEURJOLLY, J. F. and AMBLARD, P. O. (2020). Projections of determinantal point processes. *Spat. Stat.* **38** 100437. MR4103294

[20] MCKAY, M. D., BECKMAN, R. J. and CONOVER, W. J. (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. MR0533252

[21] MØLLER, J., NIELSEN, M., PORCU, E. and RUBAK, E. (2018). Determinantal point process models on the sphere. *Bernoulli* **24** 1171–1201. MR3706791

[22] MØLLER, J. and WAAGEPETERSEN, R. P. (2004). *Statistical Inference and Simulation for Spatial Point Processes.* Chapman and Hall/CRC, Boca Raton. MR2004226

[23] OWEN, A. B. (1992). Orthogonal arrays for computer experiments, integration and visualization. *Statisitica Sinica* **2** 439–452. MR1187952

[24] OWEN, A. B. (1997). Scrambled net variance for integrals of smooth functions. *The Annals of Statistics* **25** 1541–1562. MR1463564

[25] OWEN, A. B. (2008). Local antithetic sampling with scrambled nets. *The Annals of Statistics* **36** 2319–2343. MR2458189

[26] OWEN, A. B. (2013). *Monte Carlo theory, methods and examples.*

[27] RIESZ, F. and SZ-NAGY, B. (1990). *Functional Analysis.* Dover Publications, New York. MR1068530

[28] ROBERT, C. P. and CASELLA, G. (2004). *Monte Carlo Statistical Methods.* Springer-Verlag, New York. MR2080278

[29] ROBINSON, J. (2001). Topics in Mathematical Physics (4th year lecture course at Warwick). available at https://homepages.warwick.ac.uk/~masdh/.

[30] SANTNER, T. J., WILLIAMS, B. J. and NOTZ, W. I. (2013). *The design and analysis of computer experiments.* Springer Science & Business Media. MR3887662

[31] SHAPIRO, S. S. and WILK, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika* **52** 591–611. MR0205384

[32] SHIRAI, T. and TAKAHASHI, Y. (2003). Random point fields associated with certain Fredholm determinants I. Fermion, Poisson and boson point processes. *J. Funct. Anal.* **205** 414–463. MR2018415

[33] SOBOL, I. M. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics* **7** 86 - 112. MR0219238

[34] SOSHNIKOV, A. (2000a). Determinantal random point fields. *Russ. Math. Surv+* **55** 923–975. MR1799012

[35] SOSHNIKOV, A. (2000b). The central limit theorem for local linear statistics in classical compact groups and related combinatorial identities. *The Annals of Probability* **28** 1353–1370. MR1797877

[36] SOSHNIKOV, A. (2002). Gaussian limit for determinantal random point

fields. *Annals of Probability* 171–187. MR1894104

[37] Zygmund, A. (2003). *Trigonometric Series (Volumes I & II combined)*, 3$^{\mathrm{rd}}$ ed. Cambridge University Press. MR1963498