Image Anal Stereol 2025;44:197-208 doi: 10.5566/ias.3810
Original Research Paper

RANDOMIZED QUADRATURE WITH PERIODIC KERNELS: APPLICATIONS TO CAVALIERI VOLUME ESTIMATION

FRANCISCO-JAVIER SOTO

Department of Computer Science and Statistics, Escuela Técnica Superior de Ingeniería Informática, Universidad Rey Juan Carlos, Calle Tulipán s/n, Móstoles, Madrid, España
e-mail: franciscojavier.soto@urjc.es
(Received November 12, 2025; accepted November 16, 2025)

ABSTRACT

This paper studies randomized algorithms for unbiased numerical integration of d-dimensional periodic functions using kernel-based quadrature rules, with particular emphasis on rules induced by periodic radial basis function (RBF) kernels. The integration points are either deterministically generated or locally perturbed and then randomly shifted, introducing structured randomness into the scheme. The analysis builds on tools from the theory of reproducing kernel Hilbert spaces (RKHS) and Sobolev interpolation. It is shown that the resulting estimators achieve optimal variance decay rates, effectively capturing the smoothness of the integrand even when the assumed regularity is overestimated. The work is motivated by Cavalieri volume estimation, a classical problem in stereology. The theoretical results generalize this framework to higher dimensions and provide a Fourier-based perspective on smoothness, yielding a flexible and mathematically grounded alternative for randomized quadrature with periodic structure.

Keywords: Cavalieri volume estimation, kernel quadrature, radial basis functions, reproducing kernel Hilbert spaces, stereological methods.

INTRODUCTION

In this work, we focus on estimating the integral of a real, compactly supported function f, denoted by $V(f)$:

$$V(f) = \int_{\mathbb{R}^d} f(\mathbf{x})\, d\mathbf{x}, \tag{1}$$

using a weighted sum of function evaluations at points that are randomly translated modulo a fixed period (periodic structure). The original motivation for this problem is described below.

Let $n \in \mathbb{N}$. The volume of a bounded n-dimensional object can be calculated using Cavalieri's principle, which states that the volume is determined by integrating the measurements of its $(n-1)$-dimensional sections along a fixed axis. This method can be extended to higher dimensions by considering $(n-d)$-dimensional sections for integers $1 \leq d < n$, leading to integrals over d-dimensional spaces. While this problem has been widely studied in the one-dimensional case ($d = 1$) due to its applications in volume estimation in stereology—see, e.g., Cruz-Orive (2024) and references therein—its generalization to higher dimensions is less well understood and has received comparatively little attention; see, however, the related high-dimensional works of Janáček (2006; 2008); Janacek and Jirak (2019).

In practical terms, stereological applications typically estimate the volume of a solid $Y \subset \mathbb{R}^3$ by taking planar sections of Y, intersecting it with planes orthogonal to a fixed axis and measuring the resulting areas. Here, in Eq. (1), the function $f(x)$ represents the area of intersection of the solid Y with a plane positioned at $x \in \mathbb{R}$, and it is zero outside a bounded interval; this is the classical Cavalieri setting. However, several practically relevant examples involve multiple variables. For instance, two-dimensional integrals (i.e., $d = 2$), modeled by fiber lengths (i.e., sections along lines), can be found in electron microscopy (Ziegel *et al.*, 2010). Additionally, other widely used formulations for volume estimation in stereology consider periodic two-dimensional integrals, such as in the case of the nucleator (González-Villa *et al.*, 2017; Pausinger *et al.*, 2019).

The most natural method for sectioning physical objects into slices (i.e., $d = 1$) is by using cuts of uniform thickness. The classical Cavalieri estimator of the integral in Eq. (1) is constructed as a Riemann sum based on evaluations of f at systematically sampled equidistant points with thickness $t > 0$:

$$\hat{V}(f) = t \sum_{k \in \mathbb{Z}} f(tk + U), \tag{2}$$

where U is a uniformly distributed random shift within the period $[0,t)$, as described in, for example, Baddeley and Jensen (2004, Chapter 7). While this method is widely used in stereological problems and works well under ideal conditions, it faces limitations when sampling points are not equidistant (Baddeley *et al.*, 2006; Ziegel *et al.*, 2010; 2011), particularly due to local random errors arising from models with different patterns. Recently, significant improvements have been made using Newton-Cotes estimators (Kiderlen and Dorph-Petersen, 2017; Stehr and Kiderlen, 2020a;b; Stehr *et al.*, 2022), which employ Newton-Cotes quadrature rules to achieve optimal variance decay rates, even under very general sampling conditions. However, generalizing these results to multidimensional settings ($d \geq 1$) remains challenging due to increased computational complexity and the absence of a general framework for extending Newton–Cotes rules themselves to integrals in several variables.

An additional challenge is the smoothness assumption of f, which plays a crucial role in the variance behavior of estimators of $V(f)$. In 1D, classical stereology uses $(m,1)$-piecewise smoothness: f is $(m-1)$ times continuously differentiable, and the m-th and $(m+1)$-st derivatives exist and are continuous except at finitely many points with finite jumps. While the classical Cavalieri estimator has been extended to handle fractional smoothness (García-Fiñana and Cruz-Orive, 2004; García-Fiñana, 2006), and Newton–Cotes estimators have broadened the analysis for integer-order smoothness (Stehr and Kiderlen, 2020b; 2025), the two frameworks remain somewhat disjoint—each tied to either fractional or integer notions of smoothness. This separation limits their general applicability.

A recent related work (Soto, 2025) has shown, in the classical 1D Cavalieri setting, that a Fourier-decay smoothness condition subsumes those one-dimensional notions and is essentially optimal: an algebraic variance-decay assumption implies the corresponding Fourier-decay rate (a matching converse result). Building on that insight, we consider a unified higher-dimensional smoothness framework based on the decay of Fourier coefficients, leveraging the well-known link between Fourier decay and function regularity.

Concretely, we estimate the parameter in Eq. (1) using randomized kernel-based quadrature rules grounded in Reproducing Kernel Hilbert Space (RKHS) theory. These techniques have been successfully applied in various contexts, including Bayesian quadrature (O'Hagan, 1991; Briol *et al.*, 2015; 2019) and quasi-Monte Carlo integration (Novak and Woźniakowski, 2012; Dick *et al.*, 2013; 2014). By leveraging kernels to capture the smoothness properties of the function $f(x)$, these methods accelerate convergence compared to Monte Carlo integration, even in misspecified settings (Dick, 2007; 2008; Fuselier *et al.*, 2014; Oates and Girolami, 2016; Kanagawa *et al.*, 2016). Specifically, we investigate scenarios with overestimated smoothness within (isotropic) Sobolev spaces, as discussed in previous works on spheres (Fuselier *et al.*, 2014) and on standard Euclidean spaces using a generic approach (Kanagawa *et al.*, 2020). However, our work differs in that we develop a framework specifically tailored to practitioners of stereology, focusing on kernel-based quadrature rules defined on tori, motivated by periodic models such as the Cavalieri estimator in Eq. (2), which is a function with period t.

It is worth noting that this is not the first time an RKHS-based approach has been used with stereological applications in mind: Pausinger *et al.* (2019) employed Korobov spaces to analyse the variance of the nucleator under different point designs, such as lattice rules and optimal point configurations. Moreover, from a more geometric perspective, Janacek and Jirak (2019) extended the classical Kendall–Hlawka–Matheron variance formula to isotropic uniform systematic sampling with periodic grids induced by point lattices in $\mathbb{R}^d$, focusing on boundary effects for sets of finite perimeter. In contrast, our strategy takes a different approach to the problem, focusing on Sobolev-space error estimates and on Bayesian adaptive weights in the construction of the quadrature rules.

PRELIMINARIES AND BASIC NOTATION

We begin by recalling well-established geometric concepts related to interpolation errors in Riemannian manifolds, see, for instance, Narcowich *et al.* (2002; 2006); Fuselier and Wright (2012); Kanagawa *et al.* (2020). Let T be a positive real number, and let X be a finite set of points in $[0,T]^d$, which may be either deterministic or stochastic. Using boldface for vectors and italics for their coordinates, we define the following measures:

- Separation radius:

$$q_{X,T} = \frac{1}{2} \min_{\mathbf{x} \neq \mathbf{y} \in X} D_T^d(\mathbf{x},\mathbf{y}),$$

- Fill distance:

$$h_{X,T} = \sup_{\mathbf{z} \in [0,T]^d} \min_{\mathbf{x} \in X} D_T^d(\mathbf{z},\mathbf{x}),$$

– Mesh ratio:
$$\rho_{X,T} = \frac{h_{X,T}}{q_{X,T}},$$

where $D_T^d(\mathbf{x},\mathbf{y})$ denotes the the geodesic distance between $\mathbf{x}$ and $\mathbf{y}$ within $[0,T]^d$, viewed as a (flat) torus:

$$D_T^d(\mathbf{x},\mathbf{y}) = \left(\sum_{i=1}^{d}\left(\min\left(|x_i-y_i|, T-|x_i-y_i|\right)^2\right)\right)^{1/2}.$$

In the following, we show the value of the aforementioned measures for examples based on point sets derived from periodic one-dimensional models, originally introduced in the context of stereological applications (Ziegel *et al.*, 2010; 2011; Stehr and Kiderlen, 2020a; Stehr *et al.*, 2022).

Example 1 (Equidistant Grid)**.** *Let t and m be positive real and integer numbers, respectively. Consider $T = mt$ and the following lattice $Y = t\mathbb{Z}^d$. Then, we have that $X = Y \cap [0,T)^d$ is a set of m^d points in $[0,T]^d$ with period $[0,t)^d$, satisfying:*

$$q_{X,T} = \frac{t}{2}, \quad h_{X,T} = \frac{\sqrt{d}}{2}t, \quad \rho_{X,T} = \sqrt{d}\,.$$

Example 2 (Perturbed Grid)**.** *Let t and m be positive real and integer numbers, respectively. Consider $T = (m+2)t$ and the stochastically perturbed lattice*

$$Y = \{\,t(\mathbf{k}+E_{\mathbf{k}}) : \mathbf{k} \in \mathbb{Z}^d\,\},$$

where $(E_{\mathbf{k}})_{\mathbf{k}\in\mathbb{Z}^d}$ are i.i.d. random vectors in $\mathbb{R}^d$ with $\|E_{\mathbf{k}}\|_2 \le \frac{\lambda}{2}$ where $0<\lambda<1$. Define

$$X = \{\,t(\mathbf{k}+E_{\mathbf{k}}) \bmod T : \mathbf{k}\in\mathbb{Z}^d,\ t\mathbf{k}\in[0,T)^d\,\},$$

where mod T denotes componentwise reduction to $[0,T)^d$. Then X consists of exactly $(m+2)^d$ points in $[0,T]^d$ and

$$q_{X,T} \ge \tfrac{t}{2}(1-\lambda), \quad h_{X,T} \le \tfrac{t}{2}(\sqrt{d}+\lambda), \quad \rho_{X,T} \le \tfrac{\sqrt{d}+\lambda}{1-\lambda}.$$

Moreover, if U is a uniformly distributed random shift on $[0,T)^d$ and independent of $(E_{\mathbf{k}})$, then $\big((X+U) \bmod T\big) \cap [0,mt)^d$ and $(Y+U)\cap[0,mt)^d$ are identically distributed.

SMOOTHNESS ASSUMPTIONS AND PERIODIC RKHS

Let $L_c^2(\mathbb{R}^d)$ denote the space of square-integrable functions on $\mathbb{R}^d$ with compact essential support, i.e., functions $f \in L^2(\mathbb{R}^d)$ such that

$$S(f) = \mathbb{R}^d \setminus \bigcup\{\,U \subset \mathbb{R}^d : U \text{ open},\ f = 0 \text{ a.e. on } U\,\}$$

is compact (Lieb and Loss, 2001, Section 1.5). [1]

Let $s > d/2$ be a smoothness parameter. We say that $r_s : \mathbb{R}^d \longrightarrow (0,\infty)$ is a rate function if it satisfies the following conditions:

$$\sum_{\mathbf{k}\in\mathbb{Z}^d} \frac{1}{r_s^2\left(\frac{\mathbf{k}}{T}\right)} < \infty, \quad r_s(\mathbf{0}) = 1 \quad \text{for all} \quad T>0.$$

Consider the Fourier-decay-type function space (see, e.g., Korobov (1963); Sloan (1985); Niederreiter (1992)):

$$\varepsilon_{r_s}(\mathbb{R}^d) = \Big\{ f \in L_c^2(\mathbb{R}^d) :\ \exists C>0 \text{ such that } |\hat{f}(\boldsymbol{\xi})| \le \tfrac{C}{r_s(\boldsymbol{\xi})}\ \ \forall \boldsymbol{\xi}\in\mathbb{R}^d \Big\}, \quad (3)$$

where $\hat{f}(\boldsymbol{\xi})$ denotes the Fourier transform of f at $\boldsymbol{\xi} \in \mathbb{R}^d$, defined as:

$$\hat{f}(\boldsymbol{\xi}) = \int_{\mathbb{R}^d} f(\mathbf{x}) e^{-2\pi i\, \boldsymbol{\xi}\cdot\mathbf{x}}\, d\mathbf{x},$$

with $\boldsymbol{\xi}\cdot\mathbf{x}$ denoting the standard dot product.

Let $f \in L_c^2(\mathbb{R}^d)$. Without loss of generality, by translation we may assume $S(f) \subseteq [0,T]^d$ for some $T>0$. The T-periodic extension of f is defined by

$$f_T(\mathbf{x}) = f(\mathbf{x} \bmod T), \quad \mathbf{x}\in\mathbb{R}^d.$$

The T-periodic covariogram of f, denoted by g_T, is defined as the following integral function:

$$g_T(\mathbf{z}) = \int_{[0,T]^d} f_T(\mathbf{x}) f_T(\mathbf{x}+\mathbf{z})\, d\mathbf{x}, \quad \mathbf{z}\in\mathbb{R}^d.$$

We summarize its basic properties in the following lemma (cf. Gual-Arnau and Cruz-Orive (2002, Lemma 3.2) for related arguments). We abuse the notation by writing the Fourier coefficients at $\mathbf{k}\in\mathbb{Z}^d$ of f_T, and of any T-periodic function, as:

$$\hat{f}_T(\mathbf{k}) = \frac{1}{T^d}\int_{[0,T]^d} f_T(\mathbf{x}) e^{-\frac{2\pi i}{T}\mathbf{k}\cdot\mathbf{x}}\, d\mathbf{x}.$$

Lemma 3.1. *Let $f \in L_c^2(\mathbb{R}^d)$ and let g_T be its T-periodic covariogram, then:*

(i) $g_T = f_T * \check{f}_T$ where $\check{f}_T(\mathbf{x}) = f_T(-\mathbf{x})$ and $*$ denotes the convolution on the torus.

(ii) $\hat{g_T}(\mathbf{k}) = T^d\ |\hat{f}_T(\mathbf{k})|^2$.

(iii) $\int_{[0,T]^d} g_T(\mathbf{z})d\mathbf{z} = V(f)^2$.

[1] If f is pointwise-defined, then the essential support coincides with the usual topological support, i.e., $S(f) = \mathrm{cl}\{\mathbf{x}\in\mathbb{R}^d : f(\mathbf{x}) \neq 0\}$.

Proof. By periodicity and the change of variables $\mathbf{y} = \mathbf{x} + \mathbf{z}$, we have

$$g_T(\mathbf{z}) = \int_{[0,T]^d} f_T(\mathbf{y}-\mathbf{z}) f_T(\mathbf{y})\, d\mathbf{y} = (f_T * \check{f}_T)(\mathbf{z}).$$

Since $g_T = f_T * \check{f}_T$ we get (f is real-valued)

$$\hat{g_T} = T^d\, |\hat{f_T}|^2. \tag{4}$$

Lastly, an application of Fubini's theorem shows $\int_{[0,T]^d} g_T(\mathbf{z}) d\mathbf{z} = V(f)^2$. □

We now introduce the RKHS that characterizes the smoothness properties of the T-periodic covariogram.

Let $T > 0$ and $L^2_{\text{per}}([0,T]^d)$ be the space of T-periodic square-integrable functions on $[0,T]^d$. Consider

$$\mathscr{H}_{r_s}([0,T]^d) = \left\{ u \in L^2_{\text{per}}([0,T]^d) : \|u\|_{r_s} < \infty \right\},$$

where

$$\|u\|^2_{r_s} = \sum_{\mathbf{k}\in\mathbb{Z}^d} r_s^2\left(\frac{\mathbf{k}}{T}\right) |\hat{u}(\mathbf{k})|^2.$$

Endowed with the inner product

$$\langle u, v\rangle_{r_s} = \sum_{\mathbf{k}\in\mathbb{Z}^d} r_s^2\left(\frac{\mathbf{k}}{T}\right) \hat{u}(\mathbf{k})\overline{\hat{v}(\mathbf{k})}, \quad u, v \in \mathscr{H}_{r_s}([0,T]^d),$$

this forms an RKHS with kernel (see, e.g., Krieg *et al.* (2023, Example 4) and (Leobacher and Pillichshammer, 2014, Definition 4.13))

$$K_{r_s}(\mathbf{x},\mathbf{y}) = \sum_{\mathbf{k}\in\mathbb{Z}^d} \frac{e^{\frac{2\pi i}{T}\, \mathbf{k}\cdot(\mathbf{x}-\mathbf{y})}}{r_s^2\left(\frac{\mathbf{k}}{T}\right)}, \quad \mathbf{x},\mathbf{y} \in [0,T]^d.$$

The global integral of the kernel above has a closed form:

$$\int_{[0,T]^d} K_{r_s}(\mathbf{x},\mathbf{y})\, d\mathbf{y} = T^d. \tag{5}$$

In addition, the Fourier coefficient of $K_{r_s}(\cdot,\mathbf{y})$ at $\mathbf{k}$ is given by:

$$\hat{K}_{r_s}(\cdot,\ \mathbf{y})(\mathbf{k}) = \frac{e^{-\frac{2\pi i}{T}\, \mathbf{k}\cdot\mathbf{y}}}{r_s^2\left(\frac{\mathbf{k}}{T}\right)}. \tag{6}$$

Example 3 (Sobolev space)**.** *Let $s > d/2$. The RKHS $\mathscr{H}_{r_s}([0,T]^d)$ is norm equivalent to the T-periodic Sobolev space $H^s_{per}([0,T]^d)$, if the associated rate function r_s satisfies,*

$$\frac{C_1}{(1+\|\boldsymbol{\xi}\|^2)^s} \le \frac{1}{r_s^2(\boldsymbol{\xi})} \le \frac{C_2}{(1+\|\boldsymbol{\xi}\|^2)^s} \tag{7}$$

where C_1, C_2 are positive constants independent of $\boldsymbol{\xi}$ (Narcowich et al.*, 2002; Cobos* et al.*, 2016).*

Lemma 3.2. *If $f \in \varepsilon_{r_s}(\mathbb{R}^d)$, then $g_T \in \mathscr{H}_{r_s}([0,T]^d)$.*

Proof. By Lemma 3.1(ii), $\hat{g_T}(\mathbf{k}) = T^d\, |\hat{f_T}(\mathbf{k})|^2$ and, since $S(f) \subseteq [0,T]^d$, we have $\hat{f_T}(\mathbf{k}) = \frac{1}{T^d}\hat{f}\left(\frac{\mathbf{k}}{T}\right)$. Hence

$$\begin{aligned}\|g_T\|^2_{r_s} &= \sum_{\mathbf{k}\in\mathbb{Z}^d} r_s^2\Big(\frac{\mathbf{k}}{T}\Big)\, |\hat{g_T}(\mathbf{k})|^2 \\ &= \frac{1}{(T^d)^2} \sum_{\mathbf{k}\in\mathbb{Z}^d} r_s^2\Big(\frac{\mathbf{k}}{T}\Big)\, |\hat{f}(\tfrac{\mathbf{k}}{T})|^4.\end{aligned}$$

Since $f \in \varepsilon_{r_s}(\mathbb{R}^d)$, there exists $C > 0$ such that $|\hat{f}(\boldsymbol{\xi})| \le C/r_s(\boldsymbol{\xi})$ for all $\boldsymbol{\xi}$. Thus

$$\|g_T\|^2_{r_s} \le \frac{C^4}{(T^d)^2} \sum_{\mathbf{k}\in\mathbb{Z}^d} \frac{1}{r_s^2\left(\frac{\mathbf{k}}{T}\right)} < \infty,$$

because r_s is a rate function. □

Suppose the rate function r_s satisfies Eq. (7) and that $f \in \varepsilon_{r_s}(\mathbb{R}^d)$. Then, by the previous lemma, the T-periodic covariogram g_T is a function that belongs to an RKHS norm-equivalent to the Sobolev space $H^s_{\text{per}}([0,T]^d)$. This is the scenario we focus on in this work. Below, we briefly outline a well-understood approach for constructing an RKHS equivalent to the periodic Sobolev space on $[0,T]^d$.

Let $\ell \ge 2d$. The d-dimensional torus $([0,T]^d, D^d_T)$ admits a smooth embedding into $\mathbb{R}^\ell$, which allows us to define suitable kernels by restricting a positive definite kernel originally defined on $\mathbb{R}^\ell \times \mathbb{R}^\ell$. Given a smoothness parameter $\tau > (\ell - d)/2$ a fundamental class of kernels in $\mathbb{R}^\ell$ is provided by radial basis functions (RBFs), which depend only on the Euclidean distance:

$$\Phi_\tau(\mathbf{x},\mathbf{y}) = \phi_\tau(\|\mathbf{x}-\mathbf{y}\|), \quad \mathbf{x},\mathbf{y} \in \mathbb{R}^\ell, \tag{8}$$

where $\phi_\tau : [0,\infty) \to \mathbb{R}$ is univariate.

When the Fourier transform of ϕ_τ decays like $(1 + \|\boldsymbol{\xi}\|^2)^{-\tau}$, the RKHS associated with ϕ_τ is norm-equivalent (Wendland, 2004, Section 10) to the standard Sobolev space $H^\tau(\mathbb{R}^\ell)$. If we restrict this RKHS to the smooth embedding of the torus $([0,T]^d, D^d_T)$, we obtain an RKHS that is norm-equivalent (Fuselier and Wright, 2012, Theorem 5) to the Sobolev space

$$H^{\tau-(\ell-d)/2}_{\text{per}}([0,T]^d).$$

We recall that examples of RBFs include the Matérn and Wendland kernels (Matérn, 2013; Wendland, 1995), which are widely used in scattered data approximation and spatial statistics.

By the Moore–Aronszajn theorem (Aronszajn, 1950), it is well known that, for every positive definite kernel $\varphi : [0,T]^d \times [0,T]^d \to \mathbb{R}$, there exists a unique RKHS with φ as reproducing kernel. We denote this space by $\mathscr{H}_\varphi([0,T]^d)$. For instance, $\mathscr{H}_{\Phi_\tau}([0,T]^d)$ denotes the RKHS associated with the kernel Φ_τ in Eq. (8), where, by a slight abuse of notation, we use Φ_τ both for the RBF on $\mathbb{R}^\ell$ and for its restriction to the embedded torus via the fixed smooth embedding.

PERIODIC KERNEL-BASED ESTIMATORS AND VARIANCE

In what follows, we fix $s > d/2$ and $f \in \varepsilon_{r_s}(\mathbb{R}^d)$ with $S(f) \subseteq [0,T]^d$ for some $T > 0$. Additionally, let $X = \{\mathbf{x}_1,\ldots,\mathbf{x}_n\} \subset [0,T]^d$ be a deterministic set of n points, and let U be a random variable uniformly distributed on $[0,T)^d$. To approximate the integral $V(f)$ of f, we define the shift-randomized weighted sum:

$$\hat{V}_T(f) = \sum_{i=1}^{n} w(\mathbf{x}_i,X) f_T(\mathbf{x}_i + U), \tag{9}$$

where $w(\mathbf{x}_i,X) \in \mathbb{R}$ are weights that may depend measurably on the sample location $\mathbf{x}_i$ and all the points of the point set X.

The bias and variance of the estimator $\hat{V}_T(f)$ have closed-form expressions.

Proposition 4.1. *The estimator $\hat{V}_T(f)$ is unbiased if and only if*

$$\sum_{i=1}^{n} w(\mathbf{x}_i,X) \ = T^d. \tag{10}$$

Furthermore, the variance of $\hat{V}_T(f)$ is given by:

$$\operatorname{var}(\hat{V}_T(f)) = \frac{1}{T^d} \ \langle g_T, h_T\rangle_{r_s}, \tag{11}$$

where the function $h_T : [0,T]^d \longrightarrow \mathbb{R}$ is defined as follows :

$$h_T(\cdot) = \sum_{i=1}^{n}\sum_{j=1}^{n} w(\mathbf{x}_i,X) w(\mathbf{x}_j,X) K_{r_s}(\cdot,\ \mathbf{x}_i - \mathbf{x}_j) - \left(\sum_{i=1}^{n} w(\mathbf{x}_i,X)\right)^2.$$

Proof. Firstly, for the unbiasedness, a straightforward calculation yields the following:

$$\begin{aligned}
\mathbb{E}[\hat{V}_T(f)] &= \frac{1}{T^d}\int_{[0,T]^d} \sum_{i=1}^{n} w(\mathbf{x}_i,X) f_T(\mathbf{x}_i + \mathbf{u})\ d\mathbf{u} \\
&= \frac{1}{T^d}\sum_{i=1}^{n} w(\mathbf{x}_i,X) \int_{[0,T]^d} f_T(\mathbf{x}_i + \mathbf{u})\ d\mathbf{u} \\
&= \frac{1}{T^d}\sum_{i=1}^{n} w(\mathbf{x}_i,X) \int_{[0,T]^d + \mathbf{x}_i} f_T(\mathbf{u})\ d\mathbf{u} \\
&= \frac{1}{T^d}\left(\int_{[0,T]^d} f_T(\mathbf{u})\ d\mathbf{u}\right)\sum_{i=1}^{n} w(\mathbf{x}_i,X).
\end{aligned}$$

Secondly, the variance of $\hat{V}_T(f)$ can be computed using item (iii) of Lemma 3.1, as follows:

$$\operatorname{var}(\hat{V}_T(f)) = \mathbb{E}[(\hat{V}_T(f))^2] - \left(\frac{1}{T^d}\sum_{i=1}^{n} w(\mathbf{x}_i,X)\right)^2 \left(\int_{[0,T]^d} g_T(\mathbf{z})\, d\mathbf{z}\right).$$

We just have to expand the first term. Applying the definition of expected value we obtain that:

$$\begin{aligned}
\mathbb{E}[(\hat{V}_T(f))^2] &= \frac{1}{T^d}\int_{[0,T]^d} \left(\sum_{i=1}^{n} w(\mathbf{x}_i,X) f_T(\mathbf{x} + \mathbf{z})\right)^2\ d\mathbf{z} \\
&= \frac{1}{T^d}\int_{[0,T]^d} \sum_{i=1}^{n}\sum_{j=1}^{n} w(\mathbf{x}_i,X) w(\mathbf{x}_j,X) f_T(\mathbf{z}+\mathbf{x}_i) f_T(\mathbf{z}+\mathbf{x}_j)\ d\mathbf{z} \\
&= \frac{1}{T^d}\sum_{i=1}^{n}\sum_{j=1}^{n} w(\mathbf{x}_i,X) w(\mathbf{x}_j,X) \int_{[0,T]^d} f_T(\mathbf{z}+\mathbf{x}_i) f_T(\mathbf{z}+\mathbf{x}_j)\ d\mathbf{z} \\
&= \frac{1}{T^d}\sum_{i=1}^{n}\sum_{j=1}^{n} w(\mathbf{x}_i,X) w(\mathbf{x}_j,X) \int_{[0,T]^d} f_T(\mathbf{z}) f_T(\mathbf{z}+(\mathbf{x}_i-\mathbf{x}_j))\ d\mathbf{z} \\
&= \frac{1}{T^d}\sum_{i=1}^{n}\sum_{j=1}^{n} w(\mathbf{x}_i,X) w(\mathbf{x}_j,X) g_T(\mathbf{x}_i - \mathbf{x}_j).
\end{aligned}$$

Note that, by Lemma 3.2, we have $g_T \in \mathscr{H}_{r_s}([0,T]^d)$. The reproducing property of the RKHS implies the following:

$$\left\langle g_T, K_{r_s}(\cdot, \mathbf{x}_i - \mathbf{x}_j)\right\rangle_{r_s} = g_T(\mathbf{x}_i - \mathbf{x}_j),$$

and also,

$$\left\langle g_T, \int_{[0,T]^d} K_{r_s}(\cdot,\mathbf{x})\, d\mathbf{x}\right\rangle_{r_s} = \int_{[0,T]^d} g_T(\mathbf{x})\, d\mathbf{x}.$$

Finally, Eq. (5) gives $\int_{[0,T]^d} K_{r_s}(\cdot,\ \mathbf{x})\ d\mathbf{x} = T^d$ and by using basic properties of the inner product we finish the proof. □

Since Equations (10) and (11) obtained in the previous proposition allow for matrix representation, we adopt the following matrix notation.

We define the vector of weights as a column vector:

$$\mathbf{w} = (w(\mathbf{x}_i,X))_{i=1}^{n},$$

and the kernel matrix as:

$$\mathbf{K}_{r_s} = (K_{r_s}(\mathbf{x}_i, \mathbf{x}_j))_{i,j=1}^n .$$

In addition, we use $\mathbf{1}$ to denote the column vector of ones, and $\mathbf{w}^\top \mathbf{1} = \mathbf{w} \cdot \mathbf{1}$ represents the sum:

$$\sum_{i=1}^n w(\mathbf{x}_i, X),$$

where $\mathbf{w}^\top$ is the transposed vector.

Let $\mathscr{H}_\varphi([0,T]^d)$ be an RKHS with reproducing kernel φ. The worst-case error (WCE) of $\mathscr{H}_\varphi([0,T]^d)$ for a point set $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ and a vector of weights $\mathbf{w}$ is defined as (see, e.g., Sloan (1985); Leobacher and Pillichshammer (2014); Kanagawa *et al.* (2020)):

$$e_X\left(\mathscr{H}_\varphi, \mathbf{w}\right) = \sup_{p \in \mathscr{H}_\varphi, \|p\|_\varphi \leq 1} \left| \int_{[0,T]^d} p(\mathbf{x})\, d\mathbf{x} - \sum_{i=1}^n w(\mathbf{x}_i, X) p(\mathbf{x}_i) \right|.$$

Moreover, the squared WCE is given by:

$$e_X^2\left(\mathscr{H}_\varphi, \mathbf{w}\right) = \sum_{i=1}^n \sum_{j=1}^n w(\mathbf{x}_i, X) w(\mathbf{x}_j, X) \varphi(\mathbf{x}_i, \mathbf{x}_j) - 2\sum_{i=1}^n w(\mathbf{x}_i, X) \int_{[0,T]^d} \varphi(\mathbf{x}, \mathbf{x}_i)\, d\mathbf{x} + \int_{[0,T]^d} \int_{[0,T]^d} \varphi(\mathbf{x}, \mathbf{y})\, d\mathbf{x}\, d\mathbf{y}.$$

Specifically, if we take $\varphi = K_{r_s}$, then using Eq. (5), we can write the squared WCE in the following matrix form:

$$e_X^2\left(\mathscr{H}_{r_s}, \mathbf{w}\right) = \mathbf{w}^\top \mathbf{K}_{r_s} \mathbf{w} - 2T^d\, \mathbf{w}^\top \mathbf{1} + (T^d)^2. \quad (12)$$

Proposition 4.2. *Let $f \in \varepsilon_{r_s}(\mathbb{R}^d)$. Assume that the vector of weights satisfies $\mathbf{w}^\top \mathbf{1} = T^d$. Then, the following inequality holds:*

$$\operatorname{var}(\hat{V}_T(f)) \leq C\, e_X^2(\mathscr{H}_{r_s}, \mathbf{w}), \quad (13)$$

for some constant $C > 0$ independent of X.

Proof. From Eq. (11), we know that the variance of $\hat{V}_T(f)$ can be expressed as:

$$\operatorname{var}(\hat{V}_T(f)) = \frac{1}{T^d} \langle g_T, h_T \rangle_{r_s}.$$

Since $f \in \varepsilon_{r_s}(\mathbb{R}^d)$, and applying Eq. (4), we obtain:

$$\langle g_T, h_T \rangle_{r_s} = \sum_{\mathbf{k} \in \mathbb{Z}^d} r_s^2\left(\frac{\mathbf{k}}{T}\right) \hat{g}_T(\mathbf{k}) \overline{\hat{h}_T(\mathbf{k})} \leq \frac{C^2}{T^d} \sum_{\mathbf{k} \in \mathbb{Z}^d} \overline{\hat{h}_T(\mathbf{k})}.$$

Using the properties of the Fourier transform and complex conjugation $\mathbf{z} \mapsto \bar{\mathbf{z}}$, we have:

$$\overline{\hat{h}_T(\mathbf{k})} = \sum_{i=1}^n \sum_{j=1}^n w(\mathbf{x}_i, X) w(\mathbf{x}_j, X) \overline{\hat{K}_{r_s}(\cdot, \mathbf{x} - \mathbf{y})(\mathbf{k})} - (\mathbf{w}^\top \mathbf{1})^2 1_{\{\mathbf{0}\}}(\mathbf{k}),$$

where $1_{\{\mathbf{0}\}}(\mathbf{k})$ is the indicator function that equals 1 when $\mathbf{k} = \mathbf{0}$ and 0 otherwise. Furthermore, using Eq. (6), we obtain:

$$\sum_{\mathbf{k} \in \mathbb{Z}^d} \overline{\hat{K}_{r_s}(\cdot, \mathbf{x} - \mathbf{y})(\mathbf{k})} = \sum_{\mathbf{k} \in \mathbb{Z}^d} \frac{e^{\frac{2\pi i}{T} \mathbf{k} \cdot (\mathbf{x} - \mathbf{y})}}{r_s^2\left(\frac{\mathbf{k}}{T}\right)} = K_{r_s}(\mathbf{x}, \mathbf{y}).$$

Note that $\overline{\hat{h}_T(\mathbf{0})} = (\mathbf{w}^\top \mathbf{1})^2 - (\mathbf{w}^\top \mathbf{1})^2 = 0$. Therefore, we can write the sum as:

$$\sum_{\mathbf{k} \in \mathbb{Z}^d} \overline{\hat{h}_T(\mathbf{k})} = \mathbf{w}^\top \left(\mathbf{K}_{r_s} - \mathbf{1}\mathbf{1}^\top\right) \mathbf{w} = \mathbf{w}^\top \mathbf{K}_{r_s} \mathbf{w} - (\mathbf{w}^\top \mathbf{1})^2.$$

This expression coincides with Eq. (12) if and only if $\mathbf{w}^\top \mathbf{1} = T^d$, thus completing the proof.

□

Now, consider the vector of weights $\mathbf{w}$ that minimizes $e_X^2\left(\mathscr{H}_{r_s}, \mathbf{w}\right)$ given in Eq. (12). We denote this optimal vector by $\mathbf{w}_{0,s}$, which satisfies the following linear system:

$$\mathbf{K}_{r_s} \mathbf{w}_{0,s} = T^d\, \mathbf{1}.$$

These weights are used to define Bayesian quadrature rules (Kanagawa *et al.*, 2016; 2020), and are known as Bayesian weights. Moreover, since the kernel matrix $\mathbf{K}_{r_s}$ is symmetric positive definite, we have the following relation:

$$\mathbf{w}_{0,s}^\top \mathbf{1} = T^d\, \mathbf{1}^\top \mathbf{K}_{r_s}^{-1} \mathbf{1} > 0.$$

Thus, the vector of optimal unbiased weights $\mathbf{w}_{u,s}$ is defined as (see, e.g., Karvonen *et al.* (2018, Section 2.3), where they were called normalised Bayesian cubature):

$$\mathbf{w}_{u,s} = T^d\, \frac{\mathbf{w}_{0,s}}{\mathbf{w}_{0,s}^\top \mathbf{1}}. \quad (14)$$

This vector minimizes the squared WCE in Eq. (12), subject to the constraint $\mathbf{w}^\top \mathbf{1} = T^d$.

Let φ be a positive definite kernel on $[0,T]^d \times [0,T]^d$. If φ is translation-invariant on the torus, i.e., $\varphi(\mathbf{x},\mathbf{y}) = \psi(\mathbf{x}-\mathbf{y})$ for some T-periodic ψ, and its global mean satisfies

$$\int_{[0,T]^d} \varphi(\mathbf{x},\mathbf{y})\, d\mathbf{x} \;=\; T^d \quad \text{for all } \mathbf{y} \in [0,T]^d, \qquad (15)$$

then the (φ-)optimal unbiased weights are exactly those in Eq. (14), with φ used in place of K_{r_s}. This applies, for example, to translation-invariant periodic kernels obtained by restricting an RBF to a smooth embedding of the torus into $\mathbb{R}^\ell$ (as discussed at the end of Section 3) and scaling it appropriately so that Eq. (15) holds.

Furthermore, let $p \in \mathscr{H}_\varphi([0,T]^d)$. We define the ($\varphi$-)interpolant of p at the point set X as the function:

$$I_X p\,(\cdot) = \sum_{i=1}^n c(\mathbf{x}_i, X)\varphi(\cdot,\, \mathbf{x}_i).$$

The vector of coefficients $\mathbf{c}$ is determined by solving the linear system:

$$\mathbf{K}_\varphi \mathbf{c} = \mathbf{p}, \qquad (16)$$

where $\mathbf{K}_\varphi$ is the kernel matrix induced by φ on X, and $\mathbf{p}$ is the vector of function values $p(\mathbf{x}_i)$ at the points of X. When the interpolant is constructed with an RBF kernel, we refer to $I_X p$ as the RBF interpolant. The proof of the following theorem relies on well-known results for RBF interpolation on smooth embedded submanifolds.

Theorem 4.3. *Let $s' \geq s > d/2$. Suppose $f \in \varepsilon_{r_s}(\mathbb{R}^d)$, where r_s satisfies Eq. (7). Set $\ell \geq 2d$ and let $\Phi_{\tau'}$ be an RBF kernel as in Eq. (8) with $\tau' - (\ell - d)/2 = s'$. Let $\mathbf{w}_{u,s'}$ denote the corresponding vector of optimal unbiased weights. Then, for $h_{X,T} \leq 1$, there exists some constant $C > 0$ independent of X such that*

$$\operatorname{var}(\hat{V}_T(f)) \leq C\, h_{X,T}^{2s} \rho_{X,T}^{2(s'-s)}. \qquad (17)$$

Proof. We set without loss of generality $\ell = 2d$. Let $\mathbf{w}_{0,s'}$ denote the $\Phi_{\tau'}$-Bayesian weights computed using the kernel defined by the RBF $\Phi_{\tau'}$ as in Eq. (8), where $\tau' - d/2 = s'$.

We fix $p \in \mathscr{H}_{\Phi_\tau}([0,T]^d)$, with $\tau - d/2 = s$. It can be shown that a quadrature rule using Bayesian weights is identical to the integral of its corresponding interpolant (see, e.g., the proof of Kanagawa *et al.* (2020, Proposition 1)). The following holds:

$$\sum_{i=1}^n w_{0,s'}(\mathbf{x}_i, X) p(\mathbf{x}_i) = \int_{[0,T]^d} I_X p(\mathbf{x})\, d\mathbf{x},$$

where $I_X p$ is the RBF interpolant of p computed by solving Eq. (16) with a kernel matrix $\mathbf{K}_{\Phi_{\tau'}}$ of smoothness $\tau' = s' + d/2$. Thus, the following quadrature error for the function p:

$$e_X\left(p, \mathbf{w}_{0,s'}\right) = \left| \int_{[0,T]^d} p(\mathbf{x})\, d\mathbf{x} - \sum_{i=1}^n w_{0,s'}(\mathbf{x}_i, X) p(\mathbf{x}_i) \right|$$

can be bounded by the following L^2-norm:

$$\begin{aligned} e_X\left(p, \mathbf{w}_{0,s'}\right) &\leq \|p - I_X p\|_{L^1} \\ &\leq (T^d)^{1/2}\, \|p - I_X p\|_{L^2}, \end{aligned}$$

where the last inequality follows from Hölder's inequality on $[0,T]^d$. Since $I_X p$ is an RBF-interpolant, we can apply the result from Fuselier and Wright (2012, Theorem 17) (taking $\mu = 0$, $q = 2$ there). Thus, we obtain

$$\|p - I_X p\|_{L^2} \leq C_0\, h_{X,T}^s\, \rho_{X,T}^{s'-s}\, \|p\|_{H^s_{\text{per}}}.$$

Moreover, since $\mathscr{H}_{\Phi_\tau}([0,T]^d)$ is norm-equivalent to $H^s_{\text{per}}([0,T]^d)$, we can ensure that

$$\|p - I_X p\|_{L^2} \leq C_1\, h_{X,T}^s\, \rho_{X,T}^{s'-s}\, \|p\|_{\mathscr{H}_{\Phi_\tau}}$$

and

$$e_X\left(p, \mathbf{w}_{0,s'}\right) \leq (T^d)^{1/2}\, C_1\, h_{X,T}^s\, \rho_{X,T}^{s'-s}\, \|p\|_{\mathscr{H}_{\Phi_\tau}}.$$

Thus, we have

$$e_X^2\left(\mathscr{H}_{\Phi_\tau}, \mathbf{w}_{0,s'}\right) \leq T^d\, C_1^2\, h_{X,T}^{2s}\, \rho_{X,T}^{2(s'-s)}. \qquad (18)$$

Now, we consider the vector of optimal unbiased weights $\mathbf{w}_{u,s'} = T^d\, \mathbf{w}_{0,s'} / \mathbf{w}_{0,s'}^\top \mathbf{1}$ (see Eq. (14)). By a straightforward calculation, $e_X^2\left(\mathscr{H}_{\Phi_\tau}, \mathbf{w}_{u,s'}\right)$ can be written as $e_X^2\left(\mathscr{H}_{\Phi_\tau}, \mathbf{w}_{0,s'}\right)$ plus an error term:

$$e_X^2\left(\mathscr{H}_{\Phi_\tau}, \mathbf{w}_{u,s'}\right) = e_X^2\left(\mathscr{H}_{\Phi_\tau}, \mathbf{w}_{0,s'}\right) + \left(\frac{T^d - \mathbf{w}_{0,s'}^\top \mathbf{1}}{\mathbf{w}_{0,s'}^\top \mathbf{1}}\right) \cdot \left(\left(\frac{T^d + \mathbf{w}_{0,s'}^\top \mathbf{1}}{\mathbf{w}_{0,s'}^\top \mathbf{1}}\right) \mathbf{w}_{0,s'}^\top \mathbf{K}_{\Phi_\tau} \mathbf{w}_{0,s'} - 2T^d\, \mathbf{w}_{0,s'}^\top \mathbf{1}\right). \qquad (19)$$

In addition, we have $\Phi_{\tau'}(\mathbf{x},\mathbf{y}) = \phi_{\tau'}(\mathbf{x}-\mathbf{y})$. Hence, it holds (we abuse notation for the embedding)

$$\phi_{\tau'}(\mathbf{z}) = \sum_{\mathbf{k} \in \mathbb{Z}^d} \hat{\phi_{\tau'T}}(\mathbf{k})\, e^{\frac{2\pi i}{T} \mathbf{k} \cdot \mathbf{z}},\ \hat{\phi_{\tau'T}}(\mathbf{k}) \geq 0,\ \hat{\phi_{\tau'T}}(\mathbf{0}) = 1$$

by Eq. (15). Moreover, there exist constants $K_1, K_2 > 0$ such that

$$K_1\, (1 + \|\mathbf{k}\|^2/T^2)^{-s'} \;\leq\; \hat{\phi_{\tau'T}}(\mathbf{k}) \;\leq\; K_2\, (1 + \|\mathbf{k}\|^2/T^2)^{-s'},$$

(see Eq. (7) with $\boldsymbol{\xi} = \mathbf{k}/T$ and s replaced by s'), so that

$$\alpha_{s'} \;=\; \sum_{\mathbf{k}\in\mathbb{Z}^d} \hat{\phi_{\tau' T}}(\mathbf{k}) \;<\; \infty.$$

Define $\mathbf{w}_\beta = \beta\,\mathbf{1}$ with $\beta > 0$. Then the squared WCE in $\mathscr{H}_{\Phi_{\tau'}}([0,T]^d)$ satisfies

$$e_X^2\left(\mathscr{H}_{\Phi_{\tau'}}, \mathbf{w}_\beta\right) \;\le\; (n^2\,\alpha_{s'})\,\beta^2 \;-\; (2nT^d)\,\beta \;+\; (T^d)^2, \tag{20}$$

where n is the size of X. If we take $\beta = \frac{T^d}{n\alpha_{s'}}$, then the value $(T^d)^2\left(1-\frac{1}{\alpha_{s'}}\right)$ appears on the second side of Eq. (20). Consequently, we obtain the following:

$$\begin{aligned} T^d\left(T^d - \mathbf{w}_{0,s'}^\top\mathbf{1}\right) &= e_X^2\left(\mathscr{H}_{\Phi_{\tau'}}, \mathbf{w}_{0,s'}\right) \\ &\le e_X^2\left(\mathscr{H}_{\Phi_{\tau'}}, \mathbf{w}_\beta\right) \\ &\le (T^d)^2\left(1-\frac{1}{\alpha_{s'}}\right). \end{aligned}$$

Therefore, we get this lower bound for $\mathbf{w}_{0,s'}^\top\mathbf{1}$:

$$\frac{T^d}{\alpha_{s'}} \le \mathbf{w}_{0,s'}^\top\mathbf{1}, \tag{21}$$

and we can ensure the existence of a positive constant $C_2 > 0$ such that

$$\begin{aligned} \frac{T^d - \mathbf{w}_{0,s'}^\top\mathbf{1}}{\mathbf{w}_{0,s'}^\top\mathbf{1}} &= \frac{T^d(T^d - \mathbf{w}_{0,s'}^\top\mathbf{1})}{T^d\,\mathbf{w}_{0,s'}^\top\mathbf{1}} \\ &= \frac{e_X^2\left(\mathscr{H}_{\Phi_{\tau'}}, \mathbf{w}_{0,s'}\right)}{T^d\,\mathbf{w}_{0,s'}^\top\mathbf{1}} \\ &\le \frac{\alpha_{s'}}{(T^d)^2}\, e_X^2\left(\mathscr{H}_{\Phi_{\tau'}}, \mathbf{w}_{0,s'}\right) \\ &\le \frac{\alpha_{s'}}{(T^d)^2} C_2\, h_{X,T}^{2s'}, \end{aligned} \tag{22}$$

where the first inequality is thanks to Eq. (21) and the second inequality follows from applying the same arguments as at the beginning of the proof when the RKHS has the same smoothness as the chosen Bayesian weights, and using the result from Fuselier and Wright (2012, Corollary 13).

Note that, applying Equations (18), (21) and (22) in Eq. (19), we obtain

$$\begin{aligned} e_X^2\left(\mathscr{H}_{\Phi_\tau}, \mathbf{w}_{u,s'}\right) &\le T^d C_1^2\, h_{X,T}^{2s}\, \rho_{X,T}^{2(s'-s)} + \\ &\quad C_2\, h_{X,T}^{2s'}\left(\frac{2\alpha_{s'}^2}{(T^d)^2}\,\mathbf{w}_{0,s'}^\top \mathbf{K}_{\Phi_\tau}\mathbf{w}_{0,s'} - 2\right), \end{aligned}$$

because, using $\mathbf{w}_{0,s'}^\top\mathbf{1} \le T^d$ and the lower bound $\mathbf{w}_{0,s'}^\top\mathbf{1} \ge T^d/\alpha_{s'}$ from Eq. (21), we have

$$\frac{T^d + \mathbf{w}_{0,s'}^\top\mathbf{1}}{\mathbf{w}_{0,s'}^\top\mathbf{1}} \le 2\alpha_{s'} \quad \text{and} \quad -2T^d\,\mathbf{w}_{0,s'}^\top\mathbf{1} \le -\frac{2(T^d)^2}{\alpha_{s'}}.$$

Moreover, from

$$\mathbf{w}_{0,s'}^\top \mathbf{K}_{\Phi_\tau}\mathbf{w}_{0,s'} = e_X^2\left(\mathscr{H}_{\Phi_\tau}, \mathbf{w}_{0,s'}\right) + 2T^d\,\mathbf{w}_{0,s'}^\top\mathbf{1} - (T^d)^2$$

and the same reasoning, we deduce

$$\mathbf{w}_{0,s'}^\top \mathbf{K}_{\Phi_\tau}\mathbf{w}_{0,s'} \le T^d C_1^2\, h_{X,T}^{2s}\, \rho_{X,T}^{2(s'-s)} + (T^d)^2.$$

Since $h_{X,T} \le 1$ and worst-case errors of norm-equivalent RKHSs are equivalent up to constants, it follows that

$$\begin{aligned} e_X^2\left(\mathscr{H}_{r_s}, \mathbf{w}_{u,s'}\right) &\le C_3\, e_X^2\left(\mathscr{H}_{\Phi_\tau}, \mathbf{w}_{u,s'}\right) \\ &\le C_4\, h_{X,T}^{2s}\, \rho_{X,T}^{2(s'-s)}, \end{aligned}$$

and one application of Eq. (13) finishes the proof.

□

As a result, the following corollary is an immediate consequence of applying the above theorem to the equidistant and perturbed grids.

Corollary 4.4. *Let X be either the equidistant grid or the perturbed grid (as in Examples 1 and 2). Assume $t > 0$ is small enough so that the corresponding fill distance satisfies $h_{X,T} \le 1$. Define*

$$T_* = \inf\{L > 0 : S(f) \subseteq [0,L]^d\}$$

and set $m = \lceil T_*/t \rceil$. Consider $T = mt$ for the equidistant grid and $T = (m+2)t$ for the perturbed grid. Under the assumptions of Theorem 4.3, the estimator $\hat{V}_T(f)$ (with the $\Phi_{\tau'}$-optimal unbiased weights) is unbiased, and its variance satisfies

$$\mathrm{var}(\hat{V}_T(f)) \;\le\; C\,t^{2s}, \tag{23}$$

where $C > 0$ is a constant independent of t.

Proof. From Examples 1 and 2 we know that the fill distance is of order t and the mesh ratio remains bounded independently of t. Thus, Theorem 4.3 yields the claimed rate. Although the constant in Eq. (17) may depend on T, note that $T \in [T_*, T_* + 1]$ for the equidistant grid (and analogously for the perturbed grid). Therefore T remains uniformly bounded in terms of $S(f)$, independently of t.

□

To conclude this section, we note the following. For any translation-invariant periodic kernel

$$\varphi(\mathbf{x},\mathbf{y}) = \psi(\mathbf{x}-\mathbf{y})$$

satisfying Eq. (15), the kernel matrix $\mathbf{K}_\varphi$ on an equidistant grid is (block-)circulant, hence $\mathbf{1}$ is an eigenvector:

$$\mathbf{K}_\varphi \mathbf{1} = \lambda\,\mathbf{1}, \quad \lambda = \sum_{j=1}^{n} \psi(\mathbf{x}_j - \mathbf{x}_1)$$

Therefore $\mathbf{K}_\varphi^{-1}\mathbf{1} = (1/\lambda)\mathbf{1}$, and the φ-optimal unbiased weights (see Eq. (14)) reduce to

$$\mathbf{w}_{u,s} = \frac{T^d\,\mathbf{K}_\varphi^{-1}\mathbf{1}}{\mathbf{1}^\top \mathbf{K}_\varphi^{-1}\mathbf{1}} = \frac{T^d\,\frac{1}{\lambda}\mathbf{1}}{\frac{n}{\lambda}} = \frac{T^d}{n}\mathbf{1} = t^d\,\mathbf{1},$$

where $T = mt,\ n = m^d$. Thus, for the equidistant grid $X = t\mathbb{Z}^d \cap [0,T)^d$,

$$\hat{V}_T(f) = \sum_{i=1}^{n} t^d f_T(\mathbf{x}_i + U),$$

which coincides with the (multivariate) classical Cavalieri estimator. Therefore, our results extend the classical asymptotic theory in stereology to the function space $\varepsilon_{r_s}(\mathbb{R}^d)$.

NUMERICAL RESULTS

This section presents simulation experiments designed to empirically evaluate the theoretical results obtained. The focus is on the following one-dimensional function ($d = 1$), which has been previously studied in the context of the classical Cavalieri estimator (see, e.g., García-Fiñana and Cruz-Orive (2004)):

$$f_a(x) = \begin{cases} \pi(1-x^2)^a, & \text{if } x \in [-1,1], \\ 0, & \text{otherwise.} \end{cases} \tag{24}$$

The parameter a takes the following values: $a = 0.25, 0.75, 1.25, 1.75, 2.25, 2.75$. The Fourier transform of this function exhibits an algebraic decay with the following exponents: 1.25, 1.75, 2.25, 2.75, 3.25, 3.75, respectively.

The goal is to verify the decay behavior predicted by Eq. (23) for this function at different values of the parameter a. To achieve this, we estimate the empirical variance with respect to the thickness t, which ranges from 1 to 0.05, corresponding to an average number of intersection slices ranging from 2 to 40. The discretization step for t is set to 0.001. The empirical variance is computed using 2000 Monte Carlo simulations, and a regression line is fitted to estimate the variance decay rate on a logarithmic scale, using the least squares method within the specific range $20 \leq 1/t \leq 40$.

The point process used in the experiments follows a perturbed model, implemented via independent uniform jitters. In this model, the relative standard deviation of the point increments, with respect to the nominal step t, is fixed at 10%. The estimator $\hat{V}_T(f)$ for $V(f)$ is constructed using

$$T = \left(\left\lceil \frac{2}{t} \right\rceil + 2 \right) t,$$

and the optimal weight vector $\mathbf{w}_{u,s}$ is obtained using Wendland kernels with parameters $(3,0), (3,1), (3,2)$ (see, for example, Zhu (2012, Table 4.1)), rescaled so that condition Eq. (15) holds. These RBF kernels define RKHSs that are norm-equivalent to the standard Sobolev spaces $H^\tau(\mathbb{R}^3)$ with $\tau = 2,3,4$, respectively (Zhu, 2012, Proposition 3.3). Restricting them to the circle $S^1 \subset \mathbb{R}^3$ yields RKHSs that are norm-equivalent to $H^s_{\text{per}}([0,T])$ with $s = 1,2,3$, respectively. Additionally, to benchmark $\hat{V}_T$, we also consider the Cavalieri estimator (see Eq. (2)), which, when applied to non-equidistant sampling locations, is referred to as the generalized Cavalieri estimator.

The numerical results can be seen in Fig. 1. We observe that the generalized Cavalieri estimator, as is known for the integer smoothness case (see, e.g., Ziegel *et al.* (2010); Kiderlen and Dorph-Petersen (2017)), exhibits variance inflation, except for $a = 0.25$, where the variance decay is close to the theoretical value of 2.5. For the lower smoothness kernel-based estimator ($s = 1$), the optimal variance ratio is achieved up to approximately $a = 1.75$. In this case, for the two largest values of a, significant variance inflation is also observed, corresponding to the highest smoothness cases of the function. The higher smoothness kernel-based estimators ($s = 2,3$) perform almost identically, showing a variance decay that is consistently close to the theoretical expectation. In addition, we note that using a kernel-based estimator with overestimated smoothness did not lead to worse results, while using an estimator with underestimated smoothness, although showing some adaptation—as previously observed in other experiments (Bach, 2017; Kanagawa *et al.*, 2020)—does not guarantee optimal decay in all cases.

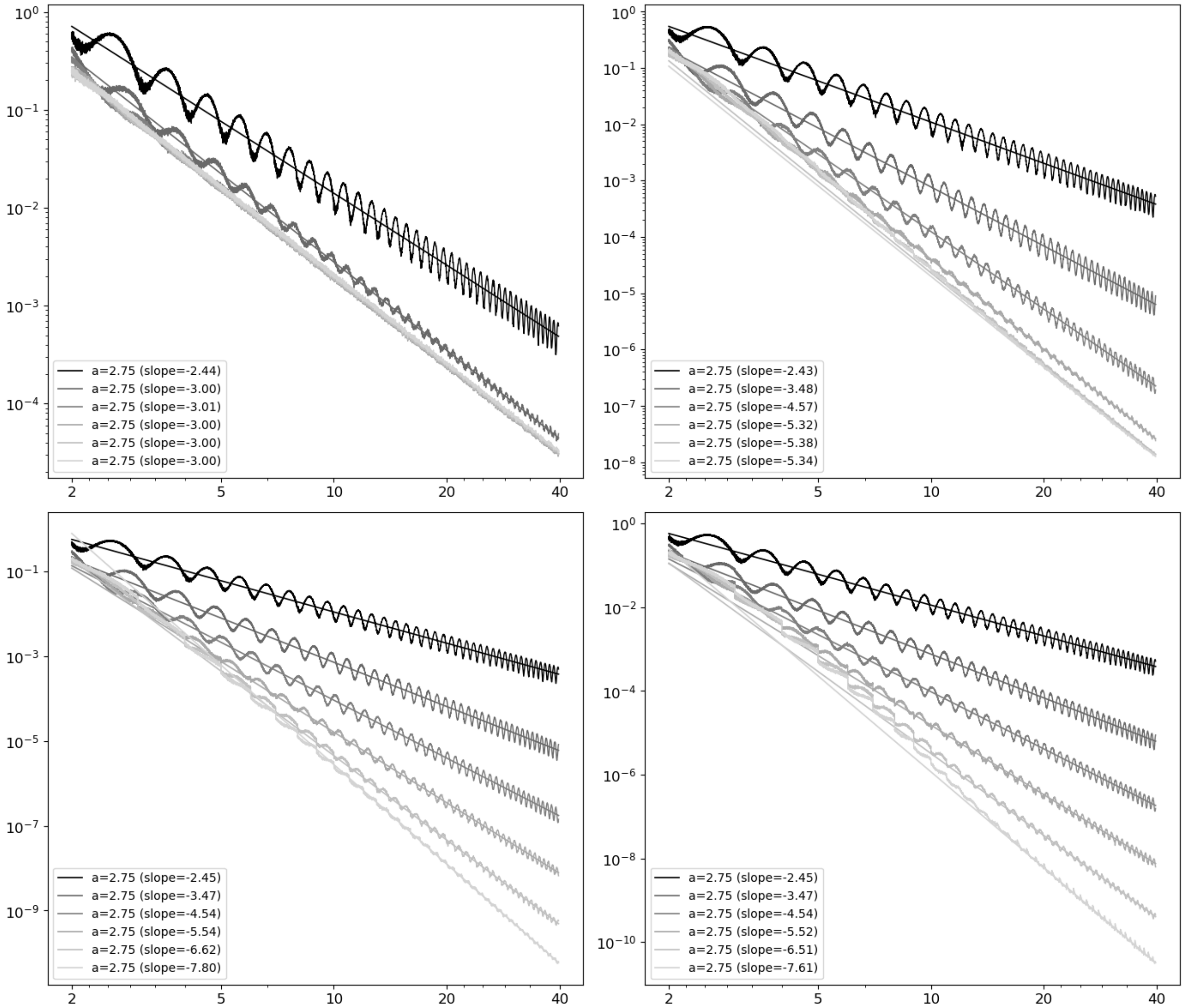

Fig. 1. *Empirical variance (y-axis) of the generalized Cavalieri estimator (top left) and of kernel-based estimators with optimal weights* $\mathbf{w}_{u,s}$ *for smoothness levels* $s = 1,2,3$ *(top right, bottom left, bottom right), for the test function* $f_a(x)$ *defined in Eq.* (24)*, whose Fourier transform decays as* $O(|\xi|^{-(a+1)})$*. Least-squares log–log fits to the empirical variances are shown as lines, and the estimated decay rate (slope) is reported in the legend for each value of a. The x-axis shows the mean number of cuts* $1/t$*, where the sampling errors* tE_i *are drawn from a uniform distribution* $U(-\varepsilon,\varepsilon)$*, with* ε *chosen so that the relative standard deviation of the perturbed increments is* 10%.

CONCLUSIONS

We analyzed the variance of kernel-based quadrature rules with randomly shifted periodic point sets. We derived variance bounds with optimal decay rates that match the integrand's smoothness, even under smoothness misspecification. Our results contribute to the theory of Bayesian quadrature in misspecified settings by treating isotropic smoothness on the d-dimensional torus with randomized point sets, extending related work.

From a more geometric perspective, our results are complementary to asymptotic variance formulas for isotropic uniform systematic sampling with periodic grids in $\mathbb{R}^d$ (Janacek and Jirak, 2019). In that line of work, the leading variance term is asymptotically expressed as a boundary measure of the set (perimeter or surface area) times a grid-dependent constant given by a lattice sum, whereas here it is driven by Fourier/Sobolev smoothness and the geometry of the underlying point sets (fill distance, mesh ratio).

Furthermore, we advance the stereological theory of Cavalieri volume estimation: our kernel-based estimator improves on the generalized Cavalieri estimator for non-equidistant point models (e.g., perturbed grids). Compared with recently proposed Newton–Cotes estimators, our approach accommodates fractional (non-integer) smoothness and extends to higher-dimensional settings, whereas the Newton–Cotes theory currently covers broader point-process classes in one dimension. Preliminary experiments suggest that our kernel-based estimators perform well beyond the settings analyzed here, opening interesting avenues for future work.

ACKNOWLEDGMENT

The author is grateful to Markus Kiderlen for valuable discussions and insightful suggestions during a stay at Aarhus University. Soto also acknowledges support from the 'PREDOCT2022-006' program at Universidad Rey Juan Carlos. In addition, he was partially supported by grant PID2023-151238OA-I00, financed by MICIU/AEI/10.13039/501100011033 and by EU, ERDF.

REFERENCES

Aronszajn N (1950). Theory of reproducing kernels. Trans Am Math Soc 68:337–404.

Bach F (2017). On the equivalence between kernel quadrature rules and random feature expansions. J Mach Learn Res 18:1–38.

Baddeley A, Dorph-Petersen KA, Vedel Jensen EB (2006). A note on the stereological implications of irregular spacing of sections. J Microsc 222:177–81.

Baddeley A, Jensen EBV (2004). Stereology for statisticians. Chapman and Hall/CRC.

Briol FX, Oates C, Girolami M, Osborne MA (2015). Frank-Wolfe Bayesian quadrature: Probabilistic integration with theoretical guarantees. Adv Neural Inf Process Syst 28.

Briol FX, Oates CJ, Girolami M, Osborne MA, Sejdinovic D (2019). Probabilistic integration. Stat Sci 34:1–22.

Cobos F, Kühn T, Sickel W (2016). Optimal approximation of multivariate periodic Sobolev functions in the sup-norm. J Funct Anal 270:4196–212.

Cruz-Orive LM (2024). Stereology: Theory and Applications. Cham: Springer.

Dick J (2007). Explicit constructions of quasi-Monte Carlo rules for the numerical integration of high-dimensional periodic functions. SIAM J Numer Anal 45:2141–76.

Dick J (2008). Walsh spaces containing smooth functions and quasi–Monte Carlo rules of arbitrary high order. SIAM J Numer Anal 46:1519–53.

Dick J, Kuo FY, Sloan IH (2013). High-dimensional integration: the quasi-Monte Carlo way. Acta Numer 22:133–288.

Dick J, Nuyens D, Pillichshammer F (2014). Lattice rules for nonperiodic smooth integrands. Numer Math 126:259–91.

Fuselier E, Hangelbroek T, Narcowich FJ, Ward JD, Wright GB (2014). Kernel based quadrature on spheres and other homogeneous spaces. Numer Math 127:57–92.

Fuselier E, Wright GB (2012). Scattered data interpolation on embedded submanifolds with restricted positive definite kernels: Sobolev error estimates. SIAM J Numer Anal 50:1753–76.

García-Fiñana M (2006). Confidence intervals in Cavalieri sampling. J Microsc 222:146–57.

García-Fiñana M, Cruz-Orive LM (2004). Improved variance prediction for systematic sampling on R. Statistics 38:243–72.

González-Villa J, Cruz M, Cruz-Orive LM (2017). On the precision of the nucleator. Image Anal Stereol 36:121–32.

Gual-Arnau X, Cruz-Orive LM (2002). Variance prediction for pseudosystematic sampling on the sphere. Adv Appl Probab 34:469–83.

Janáček J (2006). Variance of periodic measure of bounded set with random position. Comment Math Univ Carolin 47:443–55.

Janáček J (2008). Asymptotics of variance of the lattice point count. Czechoslovak Math J 58:751–8.

Janacek J, Jirak D (2019). Variance of the isotropic uniform systematic sampling. Image Anal Stereol 38:261–7.

Kanagawa M, Sriperumbudur BK, Fukumizu K (2016). Convergence guarantees for kernel-based quadrature rules in misspecified settings. Adv Neural Inf Process Syst 29.

Kanagawa M, Sriperumbudur BK, Fukumizu K (2020). Convergence analysis of deterministic kernel-based quadrature rules in misspecified settings. Found Comput Math 20:155–94.

Karvonen T, Oates CJ, Särkkä S (2018). A Bayes-Sard cubature method. Adv Neural Inf Process Syst 31.

Kiderlen M, Dorph-Petersen KA (2017). The Cavalieri estimator with unequal section spacing revisited. Image Anal Stereol 36:133–9.

Korobov NM (1963). Number-theoretic methods in approximate analysis.

Krieg D, Pozharska K, Ullrich M, Ullrich T (2023). Sampling recovery in L_2 and other norms. arXiv preprint arXiv230507539 .

Leobacher G, Pillichshammer F (2014). Introduction to quasi-Monte Carlo integration and applications. Springer.

Lieb EH, Loss M (2001). Analysis, vol. 14. American Mathematical Society.

Matérn B (2013). Spatial variation, vol. 36. Springer.

Narcowich FJ, Schaback R, Ward JD (2002). Approximation in Sobolev spaces by kernel expansions. J Approx Theory 114:70–83.

Narcowich FJ, Ward JD, Wendland H (2006). Sobolev error estimates and a Bernstein inequality for scattered data interpolation via radial basis functions. Constr Approx 24:175–86.

Niederreiter H (1992). Random Number Generation and Quasi-Monte Carlo Methods. SIAM.

Novak E, Woźniakowski H (2012). Tractability of Multivariate Problems. European Mathematical Society.

Oates C, Girolami M (2016). Control functionals for quasi-Monte Carlo integration. In: Artificial Intelligence and Statistics. PMLR.

O'Hagan A (1991). Bayes–Hermite quadrature. J Stat Plan Inference 29:245–60.

Pausinger F, Gomez-Perez D, Gonzalez-Villa J (2019). Estimation of volume using the nucleator and lattice points. Image Anal Stereol 38:141–50.

Sloan IH (1985). Lattice methods for multiple integration. J Comput Appl Math 12:131–43.

Soto FJ (2025). Sharp unified smoothness theory for Cavalieri estimation via fourier decay. Axioms 14:786.

Stehr M, Kiderlen M (2020a). Asymptotic variance of newton–cotes quadratures based on randomized sampling points. Adv Appl Probab 52:1284–307.

Stehr M, Kiderlen M (2020b). Improving the Cavalieri estimator under non-equidistant sampling and dropouts. Image Anal Stereol 39:197–212.

Stehr M, Kiderlen M (2025). Confidence intervals for Newton–Cotes quadratures based on stationary point processes. Stat Methods Appl :1–29.

Stehr M, Kiderlen M, Dorph-Petersen KA (2022). Improving Cavalieri volume estimation based on non-equidistant planar sections: The trapezoidal estimator. J Microsc 288:40–53.

Wendland H (1995). Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. Adv Comput Math 4:389–96.

Wendland H (2004). Scattered data approximation, vol. 17. Cambridge University Press.

Zhu SX (2012). Compactly supported radial basis functions: How and why? Tech. Rep. OCCAM Preprint 12/57, Oxford Centre for Collaborative Applied Mathematics, University of Oxford. Eprints Archive, Oxford, UK.

Ziegel J, Baddeley A, Dorph-Petersen KA, Jensen EBV (2010). Systematic sampling with errors in sample locations. Biometrika 97:1–13.

Ziegel J, Vedel Jensen EB, Dorph-Petersen KA (2011). Variance estimation for generalized Cavalieri estimators. Biometrika 98:187–98.