

ESTIMATION OF VOLUME USING THE NUCLEATOR AND LATTICE POINTS

DOMINGO GOMEZ-PEREZ¹, JAVIER GONZALEZ-VILLA² AND FLORIAN PAUSINGER^{✉,3}

¹Department of Mathematics, Statistics and Computer Science. University of Cantabria, Avda. de los Castros, s/n. Santander, 39005, Spain.; ²Department of Applied Mathematics and Computer Science. University of Cantabria, Avda. de los Castros, s/n. Santander, 39005, Spain.; ³School of Mathematics & Physics, Queen's University Belfast, BT7 1NN, Belfast, United Kingdom.

e-mail: domingo.gomez@unican.es, javiergv@gmail.com, f.pausinger@qub.ac.uk

(Received September 19, 2018; revised May 28, 2019; accepted May 28, 2019)

ABSTRACT

The nucleator is a method to estimate the volume of a particle, *i.e.*, a compact subset of \mathbb{R}^3 , which is widely used in Stereology. It is based on geometric sampling and known to be unbiased. However, the prediction of the variance of this estimator is non-trivial and depends on the underlying sampling scheme.

We propose well established tools from quasi-Monte Carlo integration to address this problem. In particular, we show how the theory of reproducing kernel Hilbert spaces can be used for variance prediction and how the variance of estimators based on the nucleator idea can be reduced using lattice (or lattice-like) points. We illustrate and test our results on various examples.

Keywords: nucleator, Quasi-Monte Carlo integration, stereology, variance prediction.

INTRODUCTION

Stereology is the science of inferring geometric information, such as volume or surface area, of a three-dimensional object from sections with two-dimensional planes or one-dimensional lines. Stereological methods are based on geometric sampling and the resulting estimators are usually assessed via their predicted variance. In contrast to what can be achieved in computer simulations, sample sizes are generally very small in the context of microscopy due to costs or other biological and technical limitations. This motivates a very careful analysis of the variance of the involved estimators when the sample size is small.

We propose well established tools from quasi-Monte Carlo (QMC) integration to tackle difficult problems in Stereology. Quasi-Monte Carlo integration specialises on the construction and use of deterministic point sets to minimise approximation errors in numerical integration. In particular, it is well known that the regular grid is a bad choice in terms of worst case errors. However, using lattice point sets can significantly improve the situation. In our numerical experiments we show that switching to point sets from QMC can significantly improve related variance estimates in a stereological context. Our paper is meant as a proof of concept which aims to stimulate further research in both communities. For this reason we try to keep our paper as simple as possible. While classical stereological methods can indeed benefit

from low discrepancy sampling schemes, it turns out that accurate variance prediction leads to an interesting problem in the theory of reproducing kernel Hilbert spaces which goes well beyond the scope of our paper.

The main aim of our paper is to analyse the variance of a volume estimator based on the nucleator as defined in Eq. 3 in Section “The nucleator”. This problem was previously studied in (Cruz-Orive and Gual-Arnau, 2000; Gual-Arnau and Cruz-Orive, 2002; González-Villa *et al.*, 2017), where the authors obtained a variance predictor for the special case that $\mathbf{p}^0, \dots, \mathbf{p}^{N-1}$ are points on a regular grid. We generalize this result and show how to obtain variance predictors for arbitrary point sets using the theory of integration in reproducing kernel Hilbert spaces. In a second step, we focus on the particular case of lattice point sets.

We remark that since we study functions and integrals over \mathbb{S}^2 we could also directly work with reproducing kernel Hilbert spaces on the sphere and use point sets that have been designed and optimised for the sphere. Brauchart and coauthors developed this machinery in a series of papers (Aistleitner *et al.*, 2012; Brauchart and Dick, 2013; Brauchart *et al.*, 2014; 2015). We did not use these results to keep our exposition as basic as possible, but strongly encourage their use in future investigations.

In Section “The nucleator” we introduce the nucleator as well as the main volume estimator we aim to study. Section “Volume Estimator and its variance” contains important properties of this estimator; *i.e.*, we

show that the estimator Eq. 3 is unbiased and derive an expression for its variance in terms of the covariogram. This expression is then used in Section “Variance as worst case error” to form the connection to numerical integration and (worst case) integration errors. We model the variance in terms of a reproducing kernel in Section “Model for variance”, and investigate the modeled variance in Section “Variance prediction for lattice point sets” in the special case when the points \mathbf{p}^n are lattice points. Finally, we conclude this paper with numerical results in Section “Simulations and Conclusions”.

THE NUCLEATOR

Following the exposition in Gual-Arnau and Cruz-Orive (2002) and González-Villa *et al.* (2017) we call a compact subset Y of \mathbb{R}^3 a *particle* and consider its volume $\text{vol}(Y)$. The nucleator (Gundersen, 1988) is a particular method to estimate the volume of a particle, which is widely used in stereology. For our purposes we refine the definition of particle a little. We call a compact subset Y of \mathbb{R}^3 a *nice particle* if it has the following properties:

- (P1) The boundary, ∂Y , of the compact set Y is a smoothly embedded 2-manifold.
- (P2) There exists a smooth local parametrisation Φ_p for every point $p \in \partial Y$.
- (P3) For almost all pivotal points $O \in \mathbb{R}^3$ all rays emanating from O intersect ∂Y only in finitely many points. We call such points O *generic pivotal points*.

These assumptions on Y are reasonable in the context of our intended application; such sets have positive reach (see, *e.g.*, Thäle, 2008, Proposition 14), but can still have a complicated global structure and, therefore, provide enough freedom to model many real world particles.

To introduce the nucleator, let $O \in \mathbb{R}^3 \setminus Y$ denote a fixed generic pivotal point associated with a nice particle Y and let $u \in \mathbb{S}^2$ be the orientation of a ray L_u emanating from O . In this paper we always assume that a ray L_u passes through O . Each L_u may intersect Y in $m = m(u)$ line segments. We denote the endpoints of these line segments by $x_1(u), \dots, x_{2m}(u)$ and sort them by their Euclidean distance $x_{(1)}(u), \dots, x_{(2m)}(u)$ to O . In the degenerate case, when L_u only touches Y in a point, *i.e.*, when the k th line segment is a point, we think of it as a degenerate line segment in which the endpoints coincide and set $x_{(2k-1)} = x_{(2k)}$. Hence, we

always have an even number of points x_i if we pick O from $\mathbb{R}^3 \setminus Y$. Defining a function $f_{O,Y} : \mathbb{S}^2 \rightarrow \mathbb{R}$ as

$$f_{O,Y}(u) = \frac{1}{3} \sum_{n=1}^{m(u)} (x_{(2n)}(u)^3 - x_{(2n-1)}(u)^3)$$

leads to the general nucleator representation of volume

$$\text{vol}(Y) = \int_{\mathbb{S}^2} f_{O,Y}(u) du. \quad (1)$$

To obtain an unbiased estimator of $\text{vol}(Y)$ from Eq. 1 we first parametrize u by its spherical polar coordinates; *i.e.*, the longitude angle ϕ and the colatitude angle θ . This leads to

$$\text{vol}(Y) = \int_0^{2\pi} \int_0^\pi f_{O,Y}(\phi, \theta) \sin(\theta) d\theta d\phi.$$

This formula works for general compact particles. However, it can be shown that $f_{O,Y}$ is smooth if Y is a nice particle and O is a generic pivotal point. Note that for our later application it is enough that $f_{O,Y}$ is of class C^4 .

Lemma 1. *If Y is a nice particle and O is a corresponding generic pivotal point, then $f_{O,Y}$ is smooth.*

Proof. We write f for $f_{O,Y}$ and remark that (P3) ensures that f is continuous for all u . Now, by the definition of f it is enough to consider one line segment. Let us assume $L_{\phi,\theta}$ intersects ∂Y in x_1 and x_2 . By (P1) ∂Y is a smooth surface, hence there exist open neighborhoods $V_1, V_2 \subset \mathbb{R}^3$ of x_1 and x_2 as well as open sets $U_1, U_2 \subset \mathbb{R}^2$ such that the local parametrisations $\Phi_1 : U_1 \rightarrow V_1 \cap \partial Y$ and $\Phi_2 : U_2 \rightarrow V_2 \cap \partial Y$ are continuously differentiable. These local parametrisations can be used to relate $L_{\phi,\theta}$ to x_1 and x_2 ; *i.e.*, $x_1 = \Phi_1(\phi, \theta)$ and $x_2 = \Phi_2(\phi, \theta)$. Moreover, the transition maps between different local parametrisations are smooth since ∂Y is smooth. Hence, the claim follows from (P2), *i.e.*, from the smoothness of the local parametrisations. \square

It was proposed in Gual-Arnau and Cruz-Orive (2002) to replace the measurement function with

$$F(\phi, \theta) = \frac{f(\phi, \theta) + f(\phi + \pi, \pi - \theta)}{2} = \frac{f(u) + f(-u)}{2},$$

which corresponds to measuring antipodal points in the sphere. This function satisfies $F(\phi + 2k\pi, \theta + \ell\pi) = F(\phi, \theta)$ for all integers k, ℓ , and the variance of F is at most half the variance of f , see Leobacher and Pillichshammer (2014, Exercise 1.10). Next we normalize the domain of the function by making a change of variables,

$$\text{vol}(Y) = 4\pi \int_0^1 \int_0^1 G(\phi, y) dy d\phi, \quad (2)$$

where $G(\phi, y) = F(2\pi\phi, (\cos^{-1}(2y - 1)))$. This function is periodic and has period one in both variables. Given a set of points $\mathbf{p}^0, \dots, \mathbf{p}^{N-1} \in [0, 1]^2$ and a one-periodic function G , we use the estimator

$$\widehat{\text{vol}}(Y) = \frac{4\pi}{N} \sum_{n=0}^{N-1} G((\phi_0, y_0) + \mathbf{p}^n), \quad (3)$$

where (ϕ_0, y_0) is taken uniformly random in $[0, 1]^2$ and points in \mathbb{R}^2 are added componentwise.

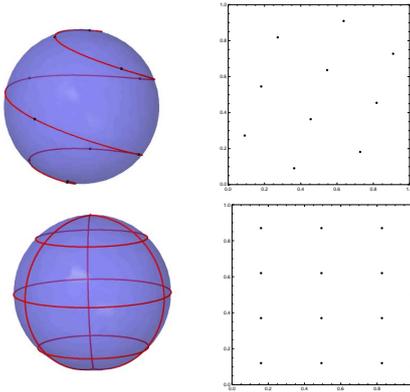


Fig. 1. Sampling with lattice points vs sampling with grid points. On the right we have the point sets in the unit square. To obtain our sampling directions on the sphere (left), we first pick a random point (ϕ_0, y_0) and translate the point set by this vector, before we project it to the sphere. Note that all projected lattice points lie on the red spiral, whereas the projected grid points lie at the center of the triangles resp rectangles on the sphere. Thus, while points from the regular grid tend to lie on very few great circles, this is not the case for lattice points.

VOLUME ESTIMATOR AND ITS VARIANCE

We start this section by showing that $\widehat{\text{vol}}(Y)$ is indeed an unbiased estimator for $\text{vol}(Y)$. Throughout this paper $\{x\}$ denotes the fractional part of $x \in \mathbb{R}$. If $\mathbf{x} \in \mathbb{R}^d$ is a d -dimensional vector, then the fractional part function is applied component-wise.

Lemma 2. For any set of points $\mathbf{p}^0, \dots, \mathbf{p}^{N-1} \in [0, 1]^2$ and any one-periodic function G , if (ϕ_0, y_0) is taken uniformly random in $[0, 1]^2$ the estimator in Eq. 3 is unbiased.

Proof. Suppose that (ϕ, y) are uniform random variables in $[0, 1]^2$. To show that the estimator is

unbiased, it is necessary to show that

$$\int_0^1 \int_0^1 G(\phi, y) dy d\phi = \int_0^1 \int_0^1 \left(\frac{1}{N} \sum_{n=0}^{N-1} G((\phi, y) + \mathbf{p}^n) \right) dy d\phi.$$

This case $N = 1$ holds since $(\phi, y) + \mathbf{p}^0 \bmod 1$ is uniformly distributed in $[0, 1]^2$ if and only if (ϕ, y) is uniformly distributed for any \mathbf{p}^0 and because G is one-periodic. The case $N > 1$ can be shown using the case for $N = 1$ together with the linearity of the integral. \square

Next, we show how to obtain an expression for the variance of $\widehat{\text{vol}}(Y)$ in terms of the covariogram of G . The covariogram of a function is an efficient tool to estimate the variance of an estimator and was first proposed in Cruz-Orive and Gual-Arnau (2000) to study the statistical properties of systematic sampling in the circle. Later this theory was extended to the sphere (Gual-Arnau and Cruz-Orive, 2002). The covariogram $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, is defined as,

$$g(h_1, h_2) = \int_0^1 \int_0^1 G(\phi, y) G(\phi - h_1, y - h_2) dy d\phi. \quad (4)$$

We summarize its main properties in the following lemma.

Lemma 3 (Lemma 3.2, (Gual-Arnau and Cruz-Orive, 2002)). Given g , as defined in Eq. 4, the following conditions hold:

- (i) $g(h_1 + k, h_2 + \ell) = g(h_1, h_2)$ for any integers k, ℓ , i.e., the covariogram is a periodic function, with period one.
- (ii) $g(h_1, h_2) = g(-h_1, -h_2)$.
- (iii) $\int_0^1 \int_0^1 g(h_1, h_2) dh_1 dh_2 = \left(\int_0^1 \int_0^1 G(\phi, y) dy d\phi \right)^2$.

To prepare for the next step, we recall that $\mathcal{H}_{2,2,\boldsymbol{\gamma}}$, with $\boldsymbol{\gamma} = (\gamma_1, \gamma_2)$ in which γ_1, γ_2 are positive reals, denotes the weighted Korobov space of smoothness 2 (in dimension 2) as defined in (Leobacher and Pillichshammer, 2014, Section 4.3); for a gentle introduction and further details see also (Novak and Woźniakowski, 2008, Appendix A) in which a different normalization is used for the smoothness parameter. In the following we will simply write $\mathcal{H}_{\gamma_1, \gamma_2}$ for $\mathcal{H}_{2,2,\boldsymbol{\gamma}}$. This space of functions is a reproducing kernel Hilbert space and contains continuous, periodic functions $H : \mathbb{R}^2 \rightarrow \mathbb{R}$.

The Fourier coefficients of such a function H are defined as

$$\hat{H}(h_1, h_2) = \int_0^1 \int_0^1 \exp(-2\pi I(\phi h_1 + y h_2)) H(\phi, y) dy d\phi.$$

Since these coefficients are complex numbers we denote their complex conjugates by $\overline{\hat{H}(h_1, h_2)}$. Now, $\mathcal{H}_{\gamma_1, \gamma_2}$ is the Hilbert space of periodic functions defined in $[0, 1]^2$ with finite norm $\|H\|^2 = \langle H, H \rangle$, where the inner product is given by

$$\langle H_1, H_2 \rangle = \sum_{n_1, n_2 \in \mathbb{Z}} r_{2, \gamma_1, \gamma_2}(n_1, n_2) \hat{H}_1(n_1, n_2) \overline{\hat{H}_2(n_1, n_2)},$$

with

$$r_{2, \gamma_1, \gamma_2}(n_1, n_2) = r_{2, \gamma_1}(n_1) \cdot r_{2, \gamma_2}(n_2),$$

and $r_{\alpha, \gamma}(n) = 1$ if $n = 0$ and $r_{\alpha, \gamma}(n) = |n|^2 / \gamma$ if $n \neq 0$. The inner product defines a norm of functions and this gives a distance between two functions. The vector $\boldsymbol{\gamma}$ contains weights for measuring the distance between functions; *i.e.*, if γ_1 is greater than γ_2 , the first coordinate of the function has more importance when calculating the norm.

This space of functions contains any periodic function with continuous first and second derivative (Leobacher and Pillichshammer, 2014, Proposition 4.16). Note that by Lemma 1 our function G (and therefore also the covariogram g) lies indeed in the space $\mathcal{H}_{\gamma_1, \gamma_2}$. As in (González-Villa *et al.*, 2017; Hinrichs and Oettershagen, 2016), we concentrate on this space.

The following lemma shows the relation between the variance of an estimator and the covariogram.

Lemma 4. *Let $\widehat{\text{vol}}(Y)$ be the estimator defined in Eq. 3. If G is as defined above then*

$$\text{var}(\widehat{\text{vol}}(Y)) = \frac{16\pi^2}{N^2} \sum_{n_1, n_2=0}^{N-1} g(\mathbf{p}^{n_1} - \mathbf{p}^{n_2}) - \text{vol}(Y)^2,$$

where g is the covariogram of G .

Proof. Using the definition of the variance and the fact that the estimator is unbiased, we have that

$$\begin{aligned} \text{var}\left(\frac{1}{N} \sum_{n=0}^{N-1} G((\phi, y) + \mathbf{p}^n)\right) &= \\ \int_0^1 \int_0^1 \frac{1}{N^2} \left(\sum_{n=0}^{N-1} G((\phi, y) + \mathbf{p}^n)\right)^2 dy d\phi &- \frac{\text{vol}(Y)^2}{16\pi^2}. \end{aligned}$$

Now, we expand the first term on the right hand side and use the definition of the covariogram to get

$$\begin{aligned} \int_0^1 \int_0^1 \frac{1}{N^2} \left(\sum_{n=0}^{N-1} G((\phi, y) + \mathbf{p}^n)\right)^2 dy d\phi &= \\ = \frac{1}{N^2} \sum_{n_1, n_2=0}^{N-1} \int_0^1 \int_0^1 G((\phi, y) + \mathbf{p}^{n_1}) \cdot & \\ G((\phi, y) + \mathbf{p}^{n_2}) dy d\phi &= \\ = \frac{1}{N^2} \sum_{n_1, n_2=0}^{N-1} g(\mathbf{p}^{n_1} - \mathbf{p}^{n_2}). & \end{aligned}$$

This finishes the proof. \square

VARIANCE AS WORST CASE ERROR

The goal of quasi-Monte Carlo integration is to provide point sets which can be used to approximate integrals defined on the d -dimensional unit cube via

$$\int_{[0, 1]^d} f(\mathbf{x}) d\mathbf{x} \approx \frac{1}{N} \sum_{n=0}^{N-1} f(\mathbf{p}^n),$$

in which $\mathcal{P} = \{\mathbf{p}^1, \dots, \mathbf{p}^N\}$ are deterministically chosen points in $[0, 1]^d$. The integration error for a particular function f is given as

$$\text{err}(f, \mathcal{P}) := \int_{[0, 1]^d} f(\mathbf{x}) d\mathbf{x} - \frac{1}{N} \sum_{n=0}^{N-1} f(\mathbf{p}^n).$$

A Hilbert space of functions \mathcal{H} in on $[0, 1]^d$ is called a reproducing kernel Hilbert space if there exists a function $K : [0, 1]^{2d} \rightarrow \mathbb{R}$ such that

K1: $K(\mathbf{y}, \cdot)$ belongs to \mathcal{H} ; for all $\mathbf{y} \in [0, 1]^d$;

K2: $\langle f, K(\mathbf{y}, \cdot) \rangle = f(\mathbf{y})$, *i.e.*, the dot product of any function in \mathcal{H} with K gives the evaluation in \mathbf{y} .

The function K is called the (reproducing) kernel of \mathcal{H} .

Assuming f is taken from a reproducing kernel Hilbert space \mathcal{H} , the integration error $\text{err}(f, \mathcal{P})$ as well as the worst case error

$$\text{err}(\mathcal{H}, \mathcal{P}) = \sup_{f \in \mathcal{H}, \|f\| \leq 1} |\text{err}(f, \mathcal{P})| \quad (5)$$

have nice representations in terms of the reproducing kernel K of the space. In particular, it is known (see, *e.g.*, Leobacher and Pillichshammer, 2014, Eq. 3.9) that

$$\text{err}(f, \mathcal{P}) = \langle f, h \rangle, \quad (6)$$

where

$$h(\mathbf{x}) = \int_{[0,1]^d} K(\mathbf{x}, \mathbf{y}) \, d\mathbf{y} - \frac{1}{N} \sum_{n=0}^{N-1} K(\mathbf{x}, \mathbf{p}^n).$$

As explained in Novak and Woźniakowski (2008, Appendix A) the kernel of the space $\mathcal{H}_{\gamma_1, \gamma_2}$ which contains the function G is

$$K_{\gamma_1, \gamma_2}((p_1, p_2), (q_1, q_2)) = (1 + \gamma_1 k(|p_1 - q_1|))(1 + \gamma_2 k(|p_2 - q_2|)), \quad (7)$$

where $k(t) = t^2 - t + 1/6$. Note that k is the Bernoulli polynomial of degree 2 and their properties have been studied in a stereological context in dimension 1 (see, e.g., Cruz-Orive and Gual-Arnau, 2000, Proposition 2.2).

Using this kernel, the properties of reproducing kernels and Eq. 6 we obtain a representation of the variance of our volume estimator which is a generalization of Cruz-Orive and Gual-Arnau (2000, Proposition 2.2).

Lemma 5. *Let $K_{\gamma_1, \gamma_2}((p_1, p_2), (q_1, q_2))$ be as in Eq. 7. If $G \in \mathcal{H}_{\gamma_1, \gamma_2}$ then*

$$\text{var} \left(\frac{1}{4\pi} \widehat{\text{vol}}(Y) \right) = \left\langle g, \frac{1}{N^2} \sum_{n_1, n_2=0}^{N-1} K_{\gamma_1, \gamma_2}(\mathbf{p}^{n_1} - \mathbf{p}^{n_2}, \cdot) - \mathbf{1}_F \right\rangle,$$

where $\mathbf{1}_F$ is the constant function which returns 1 at any point.

Proof. By Eq. 2 and Lemma 3(iii) we know that

$$\begin{aligned} \text{vol}(Y)^2 &= 16\pi^2 \left(\int_0^1 \int_0^1 G(\phi, \mathbf{y}) \, d\mathbf{y} \, d\phi \right)^2 \\ &= 16\pi^2 \int_0^1 \int_0^1 g(h_1, h_2) \, dh_1 \, dh_2. \end{aligned}$$

Thus, by Lemma 4 we see that the *negative* of the variance can be expressed as an integral approximation error. This approximation error can now be rewritten as an inner product using Eq. 6

$$\left\langle g, - \int_{[0,1]^2} K_{\gamma_1, \gamma_2}(\mathbf{y}, \cdot) \, d\mathbf{y} + \frac{1}{N^2} \sum_{n_1, n_2=0}^{N-1} K_{\gamma_1, \gamma_2}(\mathbf{p}^{n_1} - \mathbf{p}^{n_2}, \cdot) \right\rangle.$$

In Leobacher and Pillichshammer (2014, Chapter 4, page 91f) it is shown that for any fixed \mathbf{x}

$$\int_{[0,1]^2} K_{\gamma_1, \gamma_2}(\mathbf{y}, \mathbf{x}) \, d\mathbf{y} = 1,$$

which gives the stated formula. \square

We remark that this result is a generalization of Cruz-Orive and Gual-Arnau (2000, Proposition 2.2), because there are no restrictions on the point set. Of course, the interesting point sets are those with low discrepancy.

MODEL FOR VARIANCE

Our next goal is to find a model for the variance in Lemma 5.

It is well known that by the Moore-Aronszajn theorem (Aronszajn, 1950) any function $H \in \mathcal{H}_{\gamma_1, \gamma_2}$ can be represented as an infinite sum of evaluations of the kernel $K_{\gamma_1, \gamma_2}(\cdot, \mathbf{q})$ at particular points $\mathbf{q} \in [0, 1]^2$. This generic representation of functions $H \in \mathcal{H}_{\gamma_1, \gamma_2}$ as infinite sums of evaluations of the kernel and the fact that $g \in \mathcal{H}_{\gamma_1, \gamma_2}$ motivate the following simple *model* for the covariogram:

$$\tilde{g}(h_1, h_2) = c_0 + c_1 K_{\gamma_1, \gamma_2}((h_1, h_2), (0, 0)). \quad (8)$$

Note that there is no quantified version of the result of Moore-Aronszajn. So the initial and main motivation for this particular model is the hope that it turns out to be useful as well as its simplicity. We will explain one way to estimate the parameters $c_0, c_1, \gamma_1, \gamma_2$ of this model from a single sample in the next section.

Assuming that $g \approx \tilde{g}$ with an unspecified error that we will ignore in the following, we derive a *modelled* variance $\widehat{\text{var}}$ of the estimator defined in Eq. 3 based on the above covariogram model as follows; see also Gual-Arnau and Cruz-Orive (2002, Proposition 4.2) for a different approach to model the variance.

First we use Lemma 5 and the linearity of the inner product to get:

$$\begin{aligned} &\left\langle g, -\mathbf{1}_F + \frac{1}{N^2} \sum_{n_1, n_2=0}^{N-1} K_{\gamma_1, \gamma_2}(\mathbf{p}^{n_1} - \mathbf{p}^{n_2}, \cdot) \right\rangle = \\ &= -\langle g, \mathbf{1}_F \rangle + \frac{1}{N^2} \sum_{n_1, n_2=0}^{N-1} \langle g, K_{\gamma_1, \gamma_2}(\mathbf{p}^{n_1} - \mathbf{p}^{n_2}, \cdot) \rangle. \quad (9) \end{aligned}$$

Next, we replace g by \tilde{g} . Using property K2 of a reproducing kernel Hilbert space, we get

$$-\langle \tilde{g}, \mathbf{1}_F \rangle = -c_0 - c_1,$$

and

$$\begin{aligned} & \langle \tilde{g}, K_{\gamma_1, \gamma_2}(\mathbf{p}^{n_1} - \mathbf{p}^{n_2}, \cdot) \rangle \\ &= c_0 \langle \mathbf{1}_F, K_{\gamma_1, \gamma_2}(\mathbf{p}^{n_1} - \mathbf{p}^{n_2}, \cdot) \rangle + \\ & \quad c_1 \langle K_{\gamma_1, \gamma_2}(\cdot, (0, 0)), K_{\gamma_1, \gamma_2}(\mathbf{p}^{n_1} - \mathbf{p}^{n_2}, \cdot) \rangle \\ &= c_0 + c_1 K_{\gamma_1, \gamma_2}(\mathbf{p}^{n_1} - \mathbf{p}^{n_2}, (0, 0)), \end{aligned}$$

where we used the definition of a reproducing kernel in the last step; see Leobacher and Pillichshammer (2014, Chapter 3, Definition 3.3). We collect all terms to build our modelled variance:

$$\begin{aligned} \overline{\text{var}} \left(\frac{1}{4\pi} \widehat{\text{vol}}(Y) \right) &= \\ & - c_1 + \frac{c_1}{N^2} \sum_{n_1, n_2=0}^{N-1} K_{\gamma_1, \gamma_2}(\mathbf{p}^{n_1} - \mathbf{p}^{n_2}, (0, 0)). \quad (10) \end{aligned}$$

VARIANCE PREDICTION FOR LATTICE POINT SETS

So far we have only considered general sets of N points. In the following we look at a particular construction.

Let $d, N \in \mathbb{N}$, $N \geq 2$ and let $\mathbf{q} \in \mathbb{Z}^d$. The point set $\mathcal{P}(\mathbf{q}, N) = \{\mathbf{q}^0, \dots, \mathbf{q}^{N-1}\}$ with

$$\mathbf{q}^n := \left\{ \frac{n}{N} \mathbf{q} \right\},$$

in which the fractional part function $\{\cdot\}$ is applied component-wise, is called a *lattice point set*. The vector \mathbf{q} is its *generating vector*. Such point sets have been extensively studied; see *e.g.*, (Dick and Pillichshammer, 2010; Leobacher and Pillichshammer, 2014; Niederreiter, 1992; Sloan and Joe, 1994; Wang and Hickernell, 2002) and references therein.

Lattice point sets are additive sets which means that the fractional part of the sum of two points is another point in the set. In the following we write $\widehat{\text{vol}}(Y, \mathcal{P}(\mathbf{q}, N))$ to indicate that we use a lattice point set to generate our estimator; *i.e.*,

$$\widehat{\text{vol}}(Y, \mathcal{P}(\mathbf{q}, N)) = \frac{4\pi}{N} \sum_{n=0}^{N-1} G((\phi_0, y_0) + \mathbf{q}^n)$$

For $\mathcal{P} = \mathcal{P}(\mathbf{q}, N)$, additivity reduces the formula in Eq. 10 to

$$\overline{\text{var}} \left(\frac{1}{4\pi} \widehat{\text{vol}}(Y, \mathcal{P}) \right) = -c_1 + \frac{c_1}{N} \sum_{n=0}^{N-1} K_{\gamma_1, \gamma_2}(\mathbf{q}^n, (0, 0)). \quad (11)$$

We first show that we can always find a good upper bound for the modelled variance based on the covariogram model if we use a particular choice of lattice points. Thus, the variance converges to zero quite fast as the number of points increases. Note that this result holds for fixed but arbitrary parameters $c_0, c_1, \gamma_1, \gamma_2$.

Lemma 6. *For coprime q and N , let $\mathcal{P} = \mathcal{P}(\mathbf{q}, N)$ be the lattice point set with generating vector \mathbf{q} . Then, for every N , there exists a choice for \mathbf{q} such that the modelled variance of the estimator based on the covariogram model Eq. 8 satisfies*

$$\overline{\text{var}} \left(\frac{1}{4\pi} \widehat{\text{vol}}(Y, \mathcal{P}) \right) \in \mathcal{O} \left(\frac{(\log N)^2 \log \log N}{N^2} \right),$$

in particular for $\gamma_1 = \gamma_2 = 1$, we have

$$\begin{aligned} \overline{\text{var}} \left(\frac{1}{4\pi} \widehat{\text{vol}}(Y, \mathcal{P}) \right) &\leq c_1 \left(\left(1 + \frac{\pi^2}{3} \right)^2 \right) \cdot \\ & \quad \frac{1 + 100(1 + \log N)^2 \log \log N}{N^2}. \end{aligned}$$

Proof. The proof is an immediate consequence of the following results: Leobacher and Pillichshammer (2014, Lemma 4.20) and Larcher (1986, Theorem 3), using the known bounds for the sum of divisors of a positive integer and the value of the Euler constant. \square

We finish this section with an outline of how to estimate the parameters c_0, c_1, γ_1 and γ_2 that appear in our model for the covariogram in Eq. 8 from a single sample $\{(\phi_0, y_0) + \mathbf{q}^n\}$ with $\mathbf{q} = (1, q)$ and $q \neq 1$; see also Gual-Arnau and Cruz-Orive (2002, Section 5). Note that the restriction to $\mathbf{q} = (1, q)$ has no effect on the result in Lemma 6. An unbiased estimator for the covariogram $g(h_1, h_2)$ as defined in Eq. 4 is given by

$$\begin{aligned} \hat{g}(\mathbf{h}, \mathcal{P}) &= \\ & \frac{1}{N} \sum_{n=0}^{N-1} G((\phi_0, y_0) + \mathbf{q}^n) G((\phi_0, y_0) + \mathbf{q}^n - (h_1, h_2)). \quad (12) \end{aligned}$$

Setting $N = x \cdot q + y$, we choose the four points $\mathbf{p}_1 = (0, 0)$, $\mathbf{p}_2 = (1, q)$, $\mathbf{p}_3 = (x + 1, q - y)$ and $\mathbf{p}_4 = (x + 2, 2q - y)$ (see Fig. 2) and evaluate Eq. 12 at these different points to get four equations:

$$c_0 + c_1 K_{\gamma_1, \gamma_2}(\mathbf{p}_i, (0, 0)) = \hat{g}(\mathbf{p}_i, \mathcal{P}).$$

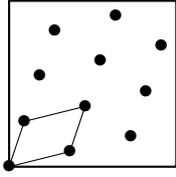


Fig. 2. The lattice point set $\mathcal{P}((1,3), 11)$ such that $N = x \cdot q + y = 3 \cdot 3 + 2$. The highlighted rectangle is thus spanned by the vectors $(1,3)$ and $(4,1)$.

Solving these equations for c_0, c_1, γ_1 and γ_2 gives the desired estimates for the parameters. We remark that we model the space $\mathcal{H}_{\gamma_1, \gamma_2}$, finding γ_1, γ_2 .

SIMULATIONS AND CONCLUSIONS

We use the same two objects as in González-Villa *et al.* (2017) to test the volume estimator and the variance predictor; *i.e.*, a Gaussian particle introduced in (Gual-Arnau and Cruz-Orive, 2002) and a computer reconstruction of a rat brain; see Fig. 3. The Gaussian particle is very smooth whereas the computer reconstruction of the rat brain has edges (and as such may not fully qualify as a nice particle). We use the following point sets in the unit square to generate direction vectors $u \in \mathbb{S}^2$ for the rays in our volume estimation:

- 1) *Lattice points* – for a given N we choose the generating vector $(1, q)$, with $1 \leq q \leq N - 1$, such that q, N are coprime and the resulting lattice has the smallest discrepancy among all possible such lattices with N points; for details we refer to (Pausinger and Topuzoğlu, 2018). Note that discrepancy is a standard tool from uniform distribution theory to assess the distribution quality of a point set used in numerical integration.
- 2) *Optimal points* – for a given N these point sets minimize the worst case integration error Eq. 5 in our function space as shown by (Hinrichs and Oettershagen, 2016).
- 3) *Systematic Points (4,4)* – Points on a regular 4×4 grid which were previously introduced and used by (González-Villa *et al.*, 2017; Gual-Arnau and Cruz-Orive, 2002). We consider a set of 16 points to compare favourably with the other point sets.

EXPERIMENT 1: PERFORMANCE OF VOLUME ESTIMATOR

In the first set of computer experiments, we compare the performance of the different estimators

based on the nucleator idea. We sample $M = 1800$ uniform random points $(\phi_0, y_0) \in [0, 1]^2$ and estimate the volume for each of these seeds and each of the above point sets. Similar to Hinrichs and Oettershagen (2016) we consider point sets with $4 \leq N \leq 16$ points, which are also realistic sample sizes in the context of stereology. The empirical mean $\mathbb{E}_e(\widehat{\text{vol}}(Y))$ and variance $\text{var}_e(\widehat{\text{vol}}(Y))$ of these experiments are computed as follows

$$\mathbb{E}_e(\widehat{\text{vol}}(Y)) = \frac{1}{M} \sum_{i=1}^M \widehat{\text{vol}}(Y, \mathcal{P}_i),$$

$$\text{var}_e(\widehat{\text{vol}}(Y)) = \frac{1}{M} \sum_{i=1}^M \left(\widehat{\text{vol}}(Y, \mathcal{P}_i) - \mathbb{E}_e(\widehat{\text{vol}}(Y)) \right)^2,$$

in which $\widehat{\text{vol}}(Y, \mathcal{P}_i)$ denotes the value of the estimator using the i th point set with $1 \leq i \leq M$. In the stereological literature, the quality of quantitative estimates obtained from design-based stereological methods is often investigated via the *square coefficient of error* CE^2 defined as

$$CE_e^2(\widehat{\text{vol}}(Y)) = \frac{\text{var}_e(\widehat{\text{vol}}(Y))}{\mathbb{E}_e(\widehat{\text{vol}}(Y))^2}.$$

Fixing N , we calculate CE_e^2 for each run of $M = 1800$ experiments and present the results in Fig. 3.

This first experiment reveals two interesting facts. First, refining the sampling scheme, *i.e.*, replacing grid points by lattice or lattice-like points as in Hinrichs and Oettershagen (2016) considerably reduces the variance. In fact, it seems that we can obtain a similar precision with only half as many points, whereas we vastly improve the precision when using the same number of points. Second, it is interesting to see that our results illustrate previous results by Hinrichs and Oettershagen (2016, Table 1). Whenever the optimal point set is not a lattice, there is a spike in our graphs. We would like to highlight the case $N = 6$; here the choice of lattices is particularly limited since the only possible generating vectors are $(1, 1)$ and $(1, 5)$. On the other hand, note that the *optimal points* in Hinrichs and Oettershagen (2016) are designed to minimize the worst case error in the function space, whereas our experiments only look at a particular function in this space. Thus, it is of course not surprising that lattice point sets can also achieve better results for this particular function than the so called optimal points.

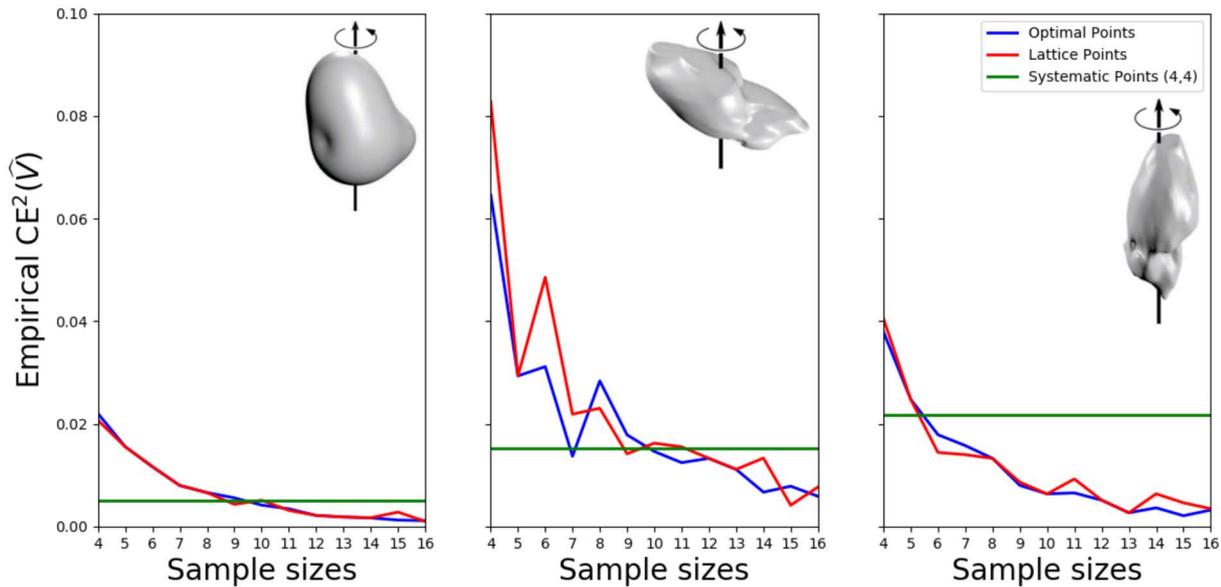


Fig. 3. Results of the simulations for the Gaussian particle (left) and the reconstructed rat brain with transversal (middle) and longitudinal (right) vertical axes through the center of mass. The graphs show the corresponding coefficient of error for estimators based on different point sets.

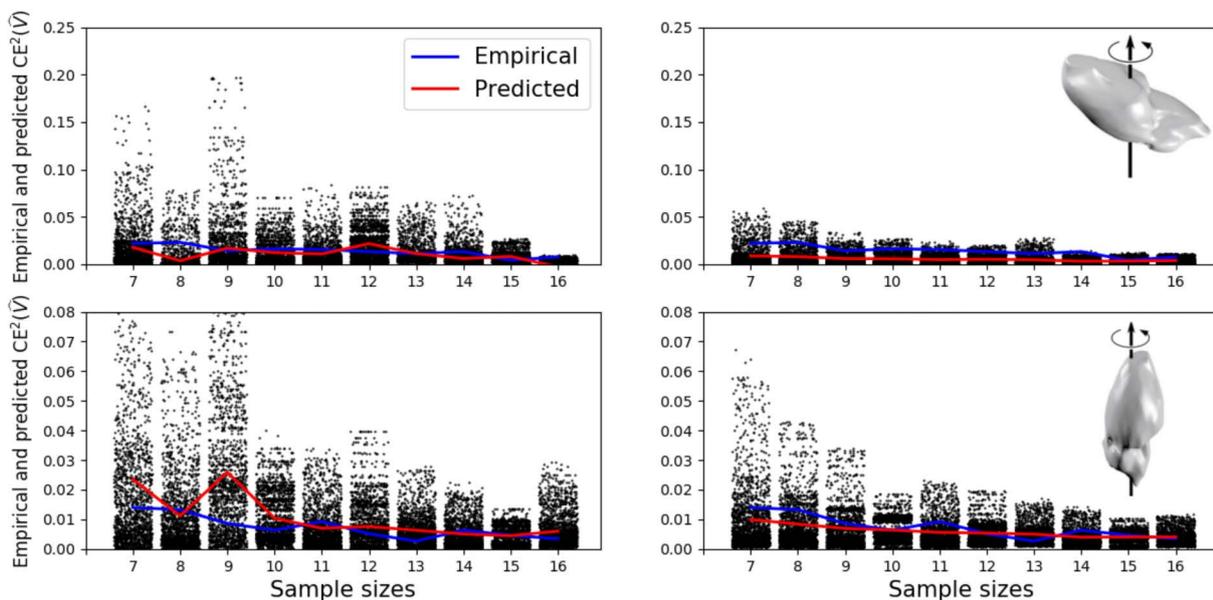


Fig. 4. Each black dot corresponds to the predicted variance of one of our experiments and the red curve is the average of all black dots whereas the blue line is the empirical variance of the volume estimators. The two plots on the left show the results for our new variance predictor, while the two plots on the right show the results for the adapted variance predictor from Gual-Arnau and Cruz-Orive (2002).

EXPERIMENT 2: PERFORMANCE OF VARIANCE PREDICTOR

In our second experiment, we focus only on lattice point sets and compare our variance predictor based on the simple covariogram model Eq. 8 to a previously suggested variance predictor for grid points based on a different model for the covariogram. This variance predictor was introduced in (Gual-Arnau and Cruz-Orive, 2002) and can be adapted to our point sets as follows

$$\text{var}_1(\widehat{\text{vol}}(Y, \mathcal{P})) = \frac{16\pi^2 \cdot (C_0 - C_1)}{6N^2 - 6N}, \quad (13)$$

in which

$$C_j = \sum_{i=0}^{N-1} G((\phi_0, y_0) + \mathbf{q}^i) G((\phi_0, y_0) + \mathbf{q}^{i+j}).$$

This variance predictor is not based on a kernel function and the price for its improved precision is that its computation is more cumbersome in practice and its precise analysis is much more technical and involved.

We calculate for each of our M experiments the variance predicted by Eq. 11 as well as by Eq. 13 and compare the mean of the values of the predicted variance to the empirical variance already computed in Experiment 1 as part of calculating CE_e^2 ; see Fig. 4. The main observation is that while the curves in Fig. 4 agree very well, the values are much more spread for our new variance predictor. Given the simplicity of our covariogram model (Eq. 8), this is not too surprising.

The main question that arises in this context is how to improve our covariogram model (Eq. 8). In theory, a better approximation could be given using first eigen-functions obtained of the spectral decomposition of the kernel, which gives the best approximations, see Hsing and Eubank (2015). Unfortunately, we could not find an easy analytical expression in this case. Note that the Moore-Aronszajn theorem provides a theoretical justification for adding more terms to our model. However, there is no approximate version of this theorem meaning it is not clear how to actually improve this approximation systematically; *i.e.*, at which points we have to evaluate the kernel or how many terms are needed to reach a certain approximation error. Therefore, we conclude with the following questions for future research:

How to improve our covariogram model (Eq. 8) by adding more terms? How to quantify the quality of this approximation of the covariogram? How to choose the additional points at which the kernel function should be evaluated?

REFERENCES

- Aistleitner C, Brauchart J, Dick J (2012). Point sets on the sphere \mathbb{S}^2 with small spherical cap discrepancy. *Discrete Comput Geom* 48:990–1024.
- Aronszajn N (1950). Theory of reproducing kernels. *Trans Am Math Soc* 68:337–404.
- Brauchart J, Dick J, Saff E, Sloan I, Wang Y, Womersley R (2015). Covering of spheres by spherical caps and worst-case error for equal weight cubature in Sobolev spaces. *J Math Anal Appl* 431:782–811.
- Brauchart J, Saff E, Sloan I, Womersley R (2014). QMC designs: optimal order quasi Monte Carlo integration schemes on the sphere. *Math Comput* 83:2821–51.
- Brauchart JS, Dick J (2013). A characterization of Sobolev spaces on the sphere and an extension of Stolarsky’s invariance principle to arbitrary smoothness. *Constr Approx* 38:397–445.
- Cruz-Orive LM, Gual-Arnau X (2000). Systematic Sampling on the Circle and on the Sphere. *Adv Appl Probab* 32:628–47.
- Dick J, Pillichshammer F (2010). *Digital Nets and Sequences. Discrepancy Theory and Quasi-Monte Carlo Integration*. Cambridge: Cambridge University Press.
- 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.
- Gundersen HJG (1988). The nucleator. *J Microsc* 151:3–21.
- Hinrichs A, Oettershagen J (2016). Optimal point sets for quasi-Monte Carlo integration of bivariate periodic functions with bounded mixed derivatives. In: Cools R, Nuyens D, eds. *Monte Carlo and quasi-Monte Carlo methods*. Springer Proc Math Stat 163:385–405.
- Hsing T, Eubank R (2015). *Theoretical foundations of functional data analysis, with an introduction to linear operators*. Hoboken: John Wiley & Sons.
- Larcher G (1986). On the distribution of sequences connected with good lattice points. *Monatsh Math* 101:135–50.
- Leobacher G, Pillichshammer F (2014). *Introduction to Quasi-Monte Carlo Integration and Applications*. Cham: Springer.
- Niederreiter H (1992). *Random Number Generation and Quasi-Monte Carlo Methods*. Philadelphia: SIAM.
- Novak E, Woźniakowski H (2008). *Tractability of Multivariate Problems. Volume I: Linear information*. Zürich: Eur Math Soc.
- Pausinger F, Topuzoğlu A (2018). On the discrepancy of two

- families of permuted van der Corput sequences. *Unif Distrib Theory* 13:47–64.
- Sloan I, Joe S (1994). *Lattice Methods for Multiple Integration*. Oxford: Clarendon Press.
- Thäle C (2008). 50 years sets with positive reach – a survey. *Surv Math Appl* 3:123–65.
- Wang Y, Hickernell FJ (2002). An historical overview of lattice point sets. In: Fang KT, Niederreiter H, Hickernell FJ, eds. *Monte Carlo and quasi-Monte Carlo methods 2000*. Berlin: Springer, 158–67.