Asymptotic Variance of Newton–Cotes Quadratures Based on Randomized Sampling Points

Stehr, Mads ; Kiderlen, Markus

Document Version
Accepted author manuscript

Published in:
Advances in Applied Probability

DOI:
10.1017/apr.2020.41

Publication date:
2020

License
Unspecified

Citation for published version (APA):
Stehr, M., & Kiderlen, M. (2020). Asymptotic Variance of Newton–Cotes Quadratures Based on Randomized Sampling Points. *Advances in Applied Probability*, *52*(4), 1284-1307. https://doi.org/10.1017/apr.2020.41

Applied Probability Trust (13 October 2020)

ASYMPTOTIC VARIANCE OF NEWTON-COTES QUADRATURES BASED ON RANDOMIZED SAMPLING POINTS

MADS STEHR,* *Aarhus University*
MARKUS KIDERLEN,* *Aarhus University*

Abstract

We consider the problem of numerical integration when the sampling nodes form a stationary point process on the real line. In previous papers it was argued that a naïve Riemann sum approach can cause a severe variance inflation when the sampling points are not equidistant. We show that this inflation can be avoided using a higher order Newton-Cotes quadrature rule which exploits smoothness properties of the integrand. Under mild assumptions, the resulting estimator is unbiased and its variance asymptotically obeys a power law as a function of the mean point distance. If the Newton-Cotes rule is of sufficiently high order, the exponent of this law turns out to only depend on the point process through its mean point distance. We illustrate our findings with the stereological estimation of the volume of a compact object, suggesting alternatives to the well-established Cavalieri estimator.

Keywords: Point process; stationary stochastic process; randomized Newton-Cotes quadrature; numerical integration; asymptotic variance bounds; renewal process

2010 Mathematics Subject Classification: Primary 65D30; 60G55
Secondary 60G10; 60K05; 65D32

1. Introduction

Let $f : \mathbb{R} \to \mathbb{R}$ be an integrable function with compact support. We intend to approximate the integral of f based on its values at finitely many random sampling points. If $X \subset \mathbb{R}$ is a stationary point process with intensity $1/t > 0$, the random variable

$$\hat{V}_0(f) = t \sum_{x \in X} f(x), \tag{1.1}$$

is unbiased for $\int f(x)\mathrm{d}x$ due to Campbell's theorem; see, e.g. [10, Chap. 3]. The simplest situation is that of *equidistant* sampling points. In this case, we can write $X = t(U + \mathbb{Z})$, where U is a uniform random variable in the interval $(0, 1)$, and the estimator becomes

$$\hat{V}(f) = t \sum_{k \in \mathbb{Z}} f\big(t(U + k)\big). \tag{1.2}$$

The estimator (1.2) corresponds to systematic sampling with randomized start. Its variance behaviour is well understood; see e.g. [7]. It was remarked in [2] and [14] that the variance of (1.1) can be substantially larger in the non-equidistant case. The purpose of the present paper, following an idea in [6], is to show and quantify that this

* Postal address: Centre for Stochastic Geometry and Advanced Bioimoaging (CSGB), Department of Mathematics, Aarhus University, Ny Munkegade 118, 8000 Aarhus C, Denmark

variance increase can be reduced – essentially to the level of the equidistant case – using Newton-Cotes quadrature approximations of sufficiently high order $n \in \mathbb{N}$ instead of the crude sum (1.1).

The resulting Newton-Cotes estimator $\hat{V}_n(f)$ is unbiased under mild assumptions; see Theorem 2.1. To analyze the variance of $\hat{V}_n(f)$, the refined Euler-MacLaurin theory in [7] appears no longer to be sufficient, and we therefore extend the classical Peano kernel theorem ([13, Theorem 3.2.3]) to locally finite, not necessarily equidistant sets of nodes in Theorem 2.2. This allows us to give explicit variance bounds in Theorem 2.3 depending on the smoothness of the measurement function f. These bounds follow a power law as functions of t. Interestingly, if the order of the Newton-Cotes estimator is large enough (compared to the smoothness of f), the exponent of this power law coincides with the exponent in the equidistant case – independently of the covariance structure of X. However, if the Newton-Cotes order is too small, the exponent may be worse than in the equidistant case, and it may depend on X. We introduce the notion of strongly n-admissible point processes (see Definition 2.1 for details) and show that the exponent for the variance bound is better for such point processes when n is not large enough; see (2.7). In Theorems 2.4 and 2.5 this general theory is applied to particular point process models: a model with i.i.d. perturbations of the equidistant case and a renewal process. In both cases, explicit variance expansions are derived for Newton-Cotes estimators of order $n = 1$ showing in particular that the power law exponents in Theorem 2.3 in general can not be improved.

The paper is organized as follows. The main results, as outlined above, are stated rigorously in the next section. In Section 3 more relevant notation is introduced, the nth order Newton-Cotes estimator is formally derived, and the refined Peano kernel theorem is proven. In Section 4 we derive integrability statements which will be of relevance when proving the main results, Theorems 2.1 and 2.3, in Sections 5 and 6, respectively. In Section 7 we show that point processes from the perturbed and cumulative model (renewal process) are strongly n-admissible for all $n \in \mathbb{N}$, and we derive the exact variance expressions presented in Theorems 2.4 and 2.5. Section 8 applies our findings to the stereological problem of volume estimation of a compact set in $\mathbb{R}^3$ and contains a simulation study. Conclusions and ideas for future work can be found in Section 9.

2. Main results

As for the estimators (1.1) and (1.2), throughout this paper we consider a point process X with intensity $1/t$ and we apply Newton-Cotes quadratures to functions evaluated at the points of X.

Let $n \in \mathbb{N}$ be given. We recall the definition of the *nth order Newton-Cotes estimator* $\hat{V}_n(f)$ from [6] for a fixed realization of X. On the interval from a point $x_0 \in X$ to its nth right neighbour in X, say x_n, the function f is approximated by a polynomial of degree at most $n \in \mathbb{N}$ passing through the points $\{x_j, f(x_j)\}_{j=0}^n$, where $x_1 < \ldots < x_{n-1}$ are the ordered points in $X \cap (x_0, x_n)$. $\hat{V}_n(f)$ is then an average of the integral of the concatenation of such approximations (composite rule) with respect to the starting point chosen. The estimator $\hat{V}_n(f)$ turns out to be a weighted average of f over all points in X,

$$\hat{V}_n(f) = \sum_{x \in X} \alpha(x; X) f(x), \tag{2.1}$$

where the weights satisfy $\alpha(x; X) = t\alpha(x/t; X/t)$ for all $x \in X$, where (when considered random) X/t is of unit intensity; see (3.4) for details. We will see in Remark 5.1 that $\alpha(x; X) = t$ when $X = t(U + \mathbb{Z})$ is an equidistant process, and therefore, Newton-Cotes estimators of any order coincide with (1.2) in the equidistant case.

When applying the estimator on randomized sampling points, we work under the general assumption that a typical distance between two consecutive points has finite positive and negative moments of all orders:

Assumption 2.1.

$$\mathbb{E}^0 h_1^j < \infty \qquad \textit{for all } j \in \mathbb{Z}. \tag{2.2}$$

Here $\mathbb{E}^0$ is the expectation under the Palm-distribution of X, that is, the distribution of X given that $0 \in X$ (see e.g. [10, sec. 3.3]), and h_1 is the lag between 0 and its right neighbour in X. Note that (2.2) holds for X if and only if it holds for aX for all $a > 0$. Assumption 2.1 is certainly not necessary for the results to hold for a given n, but finding a necessary and sufficient condition appears to be quite technical. Our first result shows the unbiasedness of $\hat{V}_n(f)$.

Theorem 2.1. *Let $n \in \mathbb{N}$ and $t > 0$ be given and assume that X is a stationary point process such that Assumption 2.1 is satisfied. Then $\hat{V}_n(f)$ is unbiased:*

$$\mathbb{E}\hat{V}_n(f) = \int_{\mathbb{R}} f(x)\mathrm{d}x$$

for all integrable and real-valued functions f with compact support.

This will be shown in Section 5, where we also argue that Assumption 4.1 below, which is weaker than Assumption 2.1, is sufficient to ensure unbiasedness. We also remark that unbiasedness is known to hold for $n = 1$ without integrability conditions and for $n = 2$ under a condition weaker than Assumption 4.1; see [6, Ex. 1 and Cor. 3].

Like in the case of classical quadrature, high order quadrature is reducing the discretization error when the measurement function is smooth. We adopt a smoothness condition which is in widespread use in stereological applications. For $m, p \in \mathbb{N}_0 \cup \{\infty\}$, we say that a measurable function f with compact support is (m, p)-piecewise smooth if it is in $C^k(\mathbb{R})$ for $k = \max\{m - 1, 0\}$, and all derivatives up to order $m + p$ exist and are continuous except in at most finitely many points, where they may have finite jumps. Hence, if f is (m, p)-piecewise smooth, m is the smallest order of derivative of f which may have jumps; see e.g. [7] for details on such functions. For our results to hold, we require that $p \geq 1$, however, the exact value of p is otherwise irrelevant. We therefore state all results for $(m, 1)$-piecewise smooth functions. We say that a function f is *exactly $(m, 1)$-piecewise smooth* if it is $(m, 1)$-piecewise smooth with discontinuous mth derivative. We let $D_{f^{(m)}}$ denote the finite set of discontinuity points of $f^{(m)}$, with

$$a \mapsto J_{f^{(m)}}(a) = \lim_{x \to a^+} f^{(m)}(x) - \lim_{x \to a^-} f^{(m)}(x)$$

denoting the corresponding jump-function.

Our second result expresses the discretization error

$$R^{(n)}(f) = \hat{V}_n(f) - \int_{\mathbb{R}} f(x)\mathrm{d}x \tag{2.3}$$

in terms of higher order derivatives of f. We state it for a realization of X, that is, we consider X as a deterministic, locally finite set of distinct points with convex hull $\mathrm{Conv}(X) = \mathbb{R}$.

Theorem 2.2. (Refined Peano kernel theorem for Newton-Cotes estimation.)
Let $n \in \mathbb{N}$ be fixed. Given X and $m \leq n$ there exists a function K_m such that

$$R^{(n)}(f) = \int_{\mathbb{R}} f^{(m+1)}(r) K_m(r) \mathrm{d}r + \sum_{a \in D_{f^{(m)}}} J_{f^{(m)}}(a) K_m(a)$$

for all $(m,1)$-piecewise smooth functions $f : \mathbb{R} \to \mathbb{R}$.

Remark 2.1. The function K_m will be called the *mth Peano kernel*. It is a piecewise polynomial of order at most $m+1$ with coefficients given in terms of X. The mth Peano kernel is explicitly given by (3.7), below. It is shown in Lemma 4.1 that for a stationary point process X satisfying Assumption 2.1, K_m is a stationary stochastic process on the real line with finite absolute moments of any (positive) order. In particular, the mean $\mathbb{E}K_m(0) = \mathbb{E}K_m(r)$ and the covariance function $H_m(s) = \mathrm{Cov}(K_m(r), K_m(s+r))$ are both finite and independent of $r \in \mathbb{R}$.

As initially considered and shown in [9], the variance of (1.2) in the equidistant case depends on jumps of high order derivatives of the measurement function; see also [7, Chapter 5]. This is outlined in the following for comparison with the general case. Let $*$ denote the convolution operator and let the reflection $\check{f}$ of f be defined as $\check{f}(x) = f(-x)$. When the measurement function f is $(m,1)$-piecewise smooth, it can be shown [7, Corollary 5.8] that the so-called *covariogram* $g = f * \check{f}$ of f is $(2m+1,1)$-piecewise smooth. When f is exactly $(m,1)$-piecewise smooth, one usually decomposes the variance of $\hat{V}(f)$ as

$$\mathbb{V}\mathrm{ar}(\hat{V}(f)) = \mathbb{V}\mathrm{ar}_E(\hat{V}(f)) + Z(t) + o(t^{2m+2}) \tag{2.4}$$

when $t \downarrow 0$. The *Zitterbewegung* $Z(t)$, which is of order t^{2m+2}, is a finite sum of terms oscillating around 0, $o(t^{2m+2})$ is a low-order remainder and the *extension term*

$$\mathbb{V}\mathrm{ar}_E(\hat{V}(f)) = t^{2m+2} g^{(2m+1)}(0^+) c_m \tag{2.5}$$

explains the overall trend of the variance. Here $c_m = -\frac{2B_{2m+2}}{(2m+2)!} \neq 0$, where B_k is the kth Bernoulli number (see Section 6 below), and as such c_m does not depend on t or the function f, other than through its order of smoothness.

Motivated by stereological applications, and adopting the naming from [14], we will mainly work with two classes of point process models. Both models are defined as scalings of unit-intensity processes.

Example 2.1. (*Perturbed model.*) A stationary point process X with intensity $1/t$ is from the perturbed model if it is derived from equidistant points by having i.i.d. perturbations tE_k, $k \in \mathbb{Z}$, of every point, i.e. $X = \{t(U + k + E_k)\}_{k \in \mathbb{Z}}$; see Section 7.1. Note that the perturbations may have a degenerate distribution concentrated at 0, and hence the equidistant model is a particular instance of the perturbed model.

Example 2.2. (*Model with cumulative errors.*) A stationary point process X with intensity $1/t$ is from the model with cumulative errors if $X = tX^u$, where X^u is a unit-intensity two-sided stationary renewal process on the real line with holding times ω_i, $i \in \mathbb{Z}$. In particular, the holding times $\{t\omega_i\}$ between two consecutive points of X form an i.i.d. sequence; see Section 7.2.

If X is from the perturbed model (with non-degenerate perturbations), the variance of (1.1) satisfies $\mathbb{V}\mathrm{ar}(\hat{V}_0(f)) = t^2c' + Z_0(t) + o(t^2)$ when $m = 0$ and $\mathbb{V}\mathrm{ar}(\hat{V}_0(f)) = t^3c'' + o(t^3)$ when $m \geq 1$ as $t \downarrow 0$. This was shown in [14, Prop. 1] apart from the missing Zitterbewegung term $Z_0(t)$ of order t^2 in the first equation, which was omitted there as it was erroneously claimed that the last term in [14, Eq. (A3)] is of order $o(t^{2m+2})$. Hence, for all $m \geq 1$, the rate of decrease of $V_0(f)$ in the non-equidistant case is strictly smaller than in the equisdistant case; cf. (2.5) for the latter. The behaviour is even worse in the model with cumulative errors, as $\mathbb{V}\mathrm{ar}(\hat{V}_0(f)) = tc''' + o(t)$ for all $m \geq 0$; see [14, Prop. 2].

In order to formulate corresponding rates of decrease for Newton-Cotes estimators, we need the notion of an admissible point process. The Peano kernel in the definition of an admissible point process is explicitly given in (3.7) with $m = n$.

Definition 2.1. (*Admissible point process.*) Let X be a stationary point process satisfying Assumption 2.1. For $n \in \mathbb{N}$ let H_n be the covariance function of K_n. Then X is called strongly n-admissible if $\int_0^z H_n(s)\mathrm{d}s$ is uniformly bounded in $z \geq 0$. X is called weakly n-admissible if $\lim_{z\to\infty} \frac{1}{z}\int_0^z H_n(s)\mathrm{d}s = 0$.

From the definition (3.7) of K_m it is easily seen that X is weakly/strongly admissible if and only if aX is weakly/strongly admissible for all constants $a > 0$. Admissibility is closely related to ergodicity properties of the stationary field K_n, and hence to those of X. In fact, if K_n has an exponentially decaying α-mixing coefficient (see, for instance, [5, Subsection 1.3.2] for the definition of this coefficient), then [5, Theorem 3.(1), p. 9] and the fact that $\mathbb{E}K_n(0)^{2+\varepsilon} < \infty$, $\varepsilon > 0$, imply that $H_n(s)$ is exponentially decaying, and hence, X is strongly n-admissible for all $n \in \mathbb{N}$.

The covariance function H_n need not be decaying for X to be strongly n-admissible. When X is from the perturbed model, the covariance function is closely related to Bernoulli functions (see Section 6), which are 1-periodic functions integrating to 0 on an interval of unit length. This is used in Lemma 7.1 to show that X is strongly n-admissible for all $n \in \mathbb{N}$. Concerning the model with cumulative errors, we show in Lemma 7.3 that H_n is indeed exponentially decaying when assuming that the holding times of the process have finite exponential moments. The proof relies on a result from [1] concerning the convergence rate of convolutions of the renewal measure of a pure renewal process.

Theorem 2.3. *Let $n \in \mathbb{N}$ be given and assume that X is a stationary point process with intensity $1/t > 0$ such that Assumption 2.1 holds. If f is $(m, 1)$-piecewise smooth and $k = \min\{m, n\}$, the variance of the estimator* (2.1) *obeys*

$$\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) \leq ct^{2k+2} \tag{2.6}$$

for some constant c, which does not depend on t.

If $m > n$ and X is strongly n-admissible, then

$$\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) \leq c't^{2n+3} \tag{2.7}$$

for some constant c', which does not depend on t.

If f is exactly $(m,1)$-piecewise smooth with $m < n$, the decrease rate in (2.6) is optimal. This is also true in the case $m = n$ if X is weakly n-admissible; see Remark 6.1.

When using the trapezoidal estimator, that is $n = 1$, we have exact expressions of the asymptotic behaviour of the variance when X is from the perturbed model and the model with cumulative errors. In the perturbed case, the rate of decrease of the upper bound in (2.7) is optimal if the perturbations E_i are non-degenerate.

Theorem 2.4. *Let X be from the perturbed model with intensity 1/t, and let μ_k be the kth moment of the perturbations E_i. Assume that the measurement function f is exactly $(m,1)$-piecewise smooth with covariogram $g = f * \check{f}$. Then, for $t \downarrow 0$,*

$$\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = -t^2 g'(0^+)(\mu_2 + \tfrac{1}{6}) + Z_0(t) + o(t^2), \qquad \text{for} \quad m = 0, \tag{2.8}$$

$$\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = t^4 g^{(3)}(0^+)\tfrac{1}{12}(2\mu_2 + 2\mu_4 + \tfrac{1}{30}) + Z_1(t) + o(t^4), \qquad \text{for} \quad m = 1, \tag{2.9}$$

$$\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = t^5 g^{(4)}(0)\tfrac{1}{8}(2\mu_4 + \mu_2\mu_4 - \mu_2^3 - \mu_3^2) + o(t^5), \qquad \text{for} \quad m \geq 2, \tag{2.10}$$

where the Zitterbewegung $Z_m(t)$ is given by (7.7). It is of order t^{2m+2}, and it is a finite sum of terms oscillating around 0. Moreover, if E_i has a density with a finite number of finite jumps and $m \geq 2$, the remainder $o(t^5)$ is explicitly given by

$$\begin{aligned} t^6 g^{(5)}(0^+)\tfrac{1}{720}\big(&-34\mu_2 - 90\mu_2^2 + 110\mu_4 + 180\mu_2\mu_4 \\ &- 180\mu_2^3 - 170\mu_3^2 + 8\mu_6 - \tfrac{1}{21}\big) + Z_2(t) + o(t^6). \end{aligned} \tag{2.11}$$

We compare these findings with the equidistant case. The Zitterbewegung in (2.4) is not present in the decomposition of Theorem 2.4 when $m \geq 2$, or rather it is of lower order and thus part of the low-order remainder. As the Bernoulli numbers satisfy $B_2 = \frac{1}{6}$, $B_4 = -\frac{1}{30}$ and $B_6 = \frac{1}{42}$, the extension term (2.5) becomes $\mathbb{V}\mathrm{ar}_E\hat{V}(f) = -t^2 g'(0^+)\frac{1}{6}$, $\mathbb{V}\mathrm{ar}_E\hat{V}(f) = t^4 g^{(3)}(0^+)\frac{1}{12}\frac{1}{30}$ and $\mathbb{V}\mathrm{ar}_E\hat{V}(f) = -t^6 g^{(5)}(0^+)\frac{1}{21}\frac{1}{720}$ for $m = 0, 1, 2$, respectively. Hence, the extension term of the trapezoidal estimator with perturbed sampling can come arbitrarily close to (2.5) if the errors E_i are sufficiently small. Under the model with cumulative errors, a corresponding statement holds as a consequence of the following result.

Theorem 2.5. *Let X be from the model with cumulative errors with intensity $1/t$ and let the i.i.d. holding times ω_i (of the unit-intensity process) satisfy $\mathbb{E}e^{\eta\omega_1} < \infty$ for some $\eta > 0$. Define ν_k as the kth moment of ω_1. Let the measurement function f be exactly $(m,1)$-piecewise smooth with covariogram $g = f * \check{f}$. Then, for $t \downarrow 0$,*

$$\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = -t^2 g'(0^+)\tfrac{1}{6}\nu_3 + o(t^2), \qquad \text{for} \quad m = 0, \tag{2.12}$$

$$\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = t^4 g^{(3)}(0^+)\tfrac{1}{12}\tfrac{1}{30}(6\nu_5 - 5\nu_3^2) + o(t^4), \qquad \text{for} \quad m = 1\ . \tag{2.13}$$

3. The Peano kernel representation

In this section we consider a locally finite set $X \subset \mathbb{R}$ such that $\mathrm{Conv}(X) = \mathbb{R}$, and an integrable function $f : \mathbb{R} \to \mathbb{R}$ with compact support which is known at all points in X. For any $x \in X$ and $j \in \mathbb{Z}$ we define $s_j(x) = s_j(x; X)$ as the jth successor (predecessor for $j < 0$) of x in X, with $s_0(x) = x$ by definition. Hence, for $j \geq 1$, $s_j(x)$

and $s_{-j}(x)$ are the unique points in $X \cap (x, \infty)$ and $X \cap (-\infty, x)$, respectively, such that $\#(X \cap (x, s_j(x)]) = \#(X \cap [s_{-j}(x), x)) = j$. Note that

$$s_j(x + t; X + t) = s_j(x; X) + t \tag{3.1}$$

for all $t \in \mathbb{R}$. For all $x \in X$ and $j \in \mathbb{Z}$ we let $h_j(x) = h_j(x; X) = s_j(x; X) - s_{j-1}(x; X)$ be the distance from the jth successor (predecessor) of x to its left neighbour in X. By (3.1),

$$h_j(x + t; X + t) = h_j(x; X) \tag{3.2}$$

for all $t \in \mathbb{R}$. We now recall the principle of Newton-Cotes quadrature, adapted to an infinite set of nodes; see [6] for details. On the interval $[x, s_n(x)]$, $x \in X$, the function f is approximated by a polynomial of degree at most $n \in \mathbb{N}$ passing through the points $\{s_j(x), f(s_j(x))\}_{j=0}^n$. The integral of this polynomial on $[x, s_n(x)]$ is

$$I_x^{(n)}(f) = I_x^{(n)}(f; X) = \sum_{j=0}^{n} \beta_j^{(n)}(x) f(s_j(x))$$

where

$$\beta_j^{(n)}(x) = \beta_j^{(n)}(x; X) = \int_x^{s_n(x)} \prod_{\substack{k=0 \\ k \neq j}}^{n} \frac{y - s_k(x)}{s_j(x) - s_k(x)} \mathrm{d}y \tag{3.3}$$

for $x \in X$. The approximation $\hat{V}_n(f) = \frac{1}{n} \sum_{x \in X} I_x^{(n)}(f) = \sum_{x \in X} \alpha(x) f(x)$ is then an average of the sum of the integral-approximations $I_x^{(n)}$ with respect to the starting point chosen. Here

$$\alpha(x) = \alpha(x; X) = \frac{1}{n} \sum_{j=0}^{n} \beta_j^{(n)}(s_{-j}(x)). \tag{3.4}$$

Remark 3.1. From [13, Theorem 2.1.1.1] the integral approximation on an interval $[x, s_n(x)]$ is exact whenever $f = p$ is a polynomial of degree at most n. That is, $R_x^{(n)}(p) = 0$, with the discretization error $R_x^{(n)}$ defined by

$$R_x^{(n)}(f) = R_x^{(n)}(f; X) = I_x^{(n)}(f) - \int_x^{s_n(x)} f(y) \mathrm{d}y, \tag{3.5}$$

$x \in X$.

As shown in the supplementary material [12, **??**], $\beta_j^{(n)}$ is a rational function of point-increments, and (3.2) then implies that

$$\beta_j^{(n)}(x + t; X + t) = \beta_j^{(n)}(x; X) \quad \text{and} \quad \alpha(x + t; X + t) = \alpha(x; X) \tag{3.6}$$

for all $t \in \mathbb{R}$ and $x \in X$.

We are now ready to prove the refined Peano kernel theorem as stated in Theorem 2.2. Given n, X, and $m \in \mathbb{N}_0$, the mth Peano kernel from Theorem 2.2 is defined as

$$K_m(r) = K_m(r; X) = \frac{1}{m!n} \sum_{x \in X} \mathbf{1}_{(x, s_n(x)]}(r) R_x^{(n)} \left((\cdot - r)_+^m \right). \tag{3.7}$$

The mapping $x \mapsto (x-r)_+^m$ should be understood as

$$(x-r)_+^m = \begin{cases} (x-r)^m & \text{for} \quad x > r, \\ 0 & \text{for} \quad x \le r. \end{cases}$$

Hence, K_m is piecewise polynomial of degree at most $m+1$ with coefficients determined by X.

Proof of Theorem 2.2. Fix $n \in \mathbb{N}$ and note that $nR^{(n)}(f) = \sum_{x \in X} R_x^{(n)}(f)$ due to (2.3) and (3.5). For all $x \in X$ and $y \in [x, s_n(x)]$, an induction argument using the refined partial integration formula [7, Lemma 4.1] yields

$$\begin{aligned} f(y^-) = &\sum_{k=0}^{m} \frac{f^{(k)}(x^+)}{k!}(y-x)^k \\ &+ \frac{1}{m!} \sum_{a \in D_{f^{(m)}} \cap (x,y)} J_{f^{(m)}}(a)(y-a)^m + \frac{1}{m!}\int_x^y f^{(m+1)}(t)(y-t)^m \mathrm{d}t, \end{aligned}$$

for all $(m,1)$-piecewise smooth functions f, $m \in \mathbb{N}_0$. We now assume $m \le n$. Using the linearity of $R_x^{(n)}$, the fact that all polynomials of order at most n are integrated exactly, and the fact that $R_x^{(n)}$ commutes with integration, we find that (with all expressions considered as functions of y)

$$\begin{aligned} &m! R_x^{(n)}(f) \\ &= R_x^{(n)}\Big(\sum_{a \in D_{f^{(m)}} \cap (x,y)} J_{f^{(m)}}(a)(y-a)^m + \int_x^y f^{(m+1)}(t)(y-t)^m \mathrm{d}t \Big) \\ &= \sum_{a \in D_{f^{(m)}} \cap (x, s_n(x)]} J_{f^{(m)}}(a) R_x^{(n)}\big((y-a)_+^m\big) + \int_x^{s_n(x)} f^{(m+1)}(t) R_x^{(n)}\big((y-t)_+^m\big)\mathrm{d}t. \end{aligned}$$

Changing the summation order, (3.7) implies that

$$R^{(n)}(f) = \frac{1}{n}\sum_{x \in X} R_x^{(n)}(f) = \sum_{a \in D_{f^{(m)}}} J_{f^{(m)}}(a) K_m(a) + \int_{\mathbb{R}} f^{(m+1)}(t) K_m(t) \mathrm{d}t,$$

as claimed. □

Before proceeding, we state a useful lemma on continuity properties of the Peano kernel. For $r \in \mathbb{R}$ we have

$$K_m(r) = \frac{1}{m!n} \sum_{x \in X} \mathbf{1}_{(x, s_1(x)]}(r) \sum_{i=1-n}^{0} R_{s_i(x)}^{(n)}\left((\cdot - r)_+^m\right).$$

The following result is a simple consequence of this representation and the fact that polynomials of degree at most n are approximated exactly.

Lemma 3.1. *Fix $n \in \mathbb{N}$ and a locally finite point-set X with $\mathrm{Conv}(X) = \mathbb{R}$. Then, for all $x \in X$ and $m \in \mathbb{N}$, the function K_m is differentiable on $(x, s_1(x))$ with derivative $-K_{m-1}$ and jump*

$$J_{K_m}(x) = \frac{1}{m!n} R_x^{(n)}\big((\,\cdot\, - x)^m\big).$$

In particular, K_m is $(m-1)$-times continuously differentiable for all $1 \leq m \leq n$.

4. Integrability properties

To argue that $\hat{V}_n(f)$ is an unbiased estimator for $\int f(x)\mathrm{d}x$ when applied to randomized sampling points, we recall the notion of the Palm distribution of a stationary point process $X \subset \mathbb{R}$. It can be interpreted as the conditional distribution of X given that $0 \in X$. We denote it by $\mathbb{P}^0$ with the corresponding expectation denoted by $\mathbb{E}^0$. When considering the point process X under its Palm distribution, we will often suppress the dependence on the point $0 \in X$ in the various expression, i.e. under $\mathbb{P}^0$ we for instance write

$$s_i = s_i(0), \qquad h_i = h_i(0), \qquad \beta_j^{(n)} = \beta_j^{(n)}(0)$$

for all $i \in \mathbb{Z}$ and $j = 0, \ldots, n$. In addition we write $\mathbf{h} = (h_1, \ldots, h_n)$ and, for $i \in \mathbb{Z}$, $\mathbf{h}(s_i) = (h_1(s_i), \ldots, h_n(s_i)) = (h_{i+1}, \ldots, h_{i+n})$ under $\mathbb{P}^0$. As mentioned in Section 1, a weaker assumption than (2.2) is sufficient to ensure the unbiasedness of the estimator.

Assumption 4.1. *For a given $n \in \mathbb{N}$ we assume that*

$$\mathbb{E}^0\left[\frac{\mathbf{h}^{\mathbf{m}}}{\mathbf{h}^{\mathbf{m}'}}\right] < \infty \tag{4.1}$$

for all multi-indices $\mathbf{m}, \mathbf{m}' \in \mathbb{N}_0^n$ with $|\mathbf{m}| \in \{n+1, n+2\}$ and $|\mathbf{m}'| = n$, where $|\mathbf{m}| = |(m_1, \ldots, m_n)| = \sum_{k=1}^n m_k$.

Using Hölder's inequality and [6, Eq. (13)], one shows that Assumption 2.1 is stronger than Assumption 4.1. In [12, **??**] of the supplementary material it is shown that the weight $\beta_j^{(n)}(x)$ is a rational function of the point-increments $(h_1(x), \ldots, h_n(x))$, $x \in X$, where the numerator is a homogeneous polynomial of degree $n+1$, and the denominator is a non-vanishing homogeneous polynomial of degree n with non-negative coefficients. From the fact that the Palm distribution is invariant under bijective point shifts [6, Eq. (13)], it is easily seen that $\mathbb{E}^0|\beta_j^{(n)}(s_{-j})| < \infty$ for all $j \in \{0, \cdots, n\}$ when Assumption 4.1 is satisfied, and consequently

$$\mathbb{E}^0|\alpha(0)| < \infty, \tag{4.2}$$

see [12, **??**] of the supplementary material. We conclude that either of the two assumptions is sufficient to guarantee the Palm-integrability of $\alpha(0)$, which will be used in the proof of Theorem 2.1.

To argue for the variance bounds presented in Theorem 2.3 we need higher-order moment and translation invariance properties of the Peano kernel K_m defined in (3.7).

Lemma 4.1. *Let $n \in \mathbb{N}$ be given and assume that X is a stationary point process. Then, for all $m \in \mathbb{N}_0$, K_m is a stationary stochastic process. If Assumption 2.1 holds,*

$K_m(0)$ has finite absolute moments of all (positive) orders. Moreover, if X has intensity γ, K_m satisfies

$$\mathbb{E}K_m(0) = \gamma\mathbb{E}^0 J_{K_{m+1}}(0) = \frac{\gamma}{(m+1)!n}\mathbb{E}^0 R_0^{(n)}\left((\,\cdot\,)^{m+1}\right) \tag{4.3}$$

for all $m \in \mathbb{N}_0$. In particular, $\mathbb{E}K_m(0) = 0$ for all $m < n$.

Proof. Fix $n \in \mathbb{N}$. For any $r, s \in \mathbb{R}$ and any locally finite pointset X, the Peano kernel satisfies

$$K_m(r+s; X) = K_m(r; X-s). \tag{4.4}$$

This follows from the definition of K_m and

$$R_x^{(n)}\left((\cdot - (r+s))_+^m; X\right) = R_{x-s}^{(n)}\left((\cdot - r)_+^m; X-s\right), \qquad x \in X,$$

which in turn is a consequence of (3.1) and (3.6). Due to (4.4) the stationarity of K_m is inherited from the stationarity of the point process X.

We now prove that $K_m(0)$ has finite absolute moments. Let $k \in \mathbb{N}$ be given. For arbitrary $r \in \mathbb{R}$ put $I_r = \{x \in X : r \in (x, s_n(x)]\}$. Using Hölder's inequality and some rather crude upper bounds we obtain from (3.5) and (3.7)

$$\left|K_m^k(0)\right| \leq \sum_{x \in I_0}\left|R_x^{(n)}\left((\cdot)_+^m\right)\right|^k \leq \sum_{x \in I_0}(s_n(x))^{km}\Big(\sum_{j=0}^n|\beta_j^{(n)}(x)| + s_n(x)\Big)^k.$$

By the refined Campbell Theorem [10, Theorem 3.5.3], (3.1) and (3.6) it follows that

$$\mathbb{E}\left|K_m^k(0)\right| \leq \gamma\mathbb{E}^0\int_0^{s_n} x^{km}\Big(\sum_{j=0}^n|\beta_j^{(n)}| + x\Big)^k \mathrm{d}x \leq \gamma\mathbb{E}^0 s_n^{km+1}\Big(\sum_{j=0}^n|\beta_j^{(n)}| + s_n\Big)^k,$$

where γ is the intensity of X. By the supplementary material [12, **??**], Assumption 2.1 and the fact that $s_n = \sum_{j=1}^n h_j$ under $\mathbb{P}^0$, the variables s_n and $\beta_j^{(n)}$ have finite absolute moments of all orders under $\mathbb{P}^0$. This implies that $\mathbb{E}|K_m^k(0)| < \infty$.

Equation (4.3) is a simple consequence of the refined Campbell Theorem [10, Theorem 3.5.3], Lemma 3.1 and [6, Eq. (13)]. □

5. Unbiasedness of Newton-Cotes estimators

Proof of Theorem 2.1. Fix $n \in \mathbb{N}$ and let $X \subseteq \mathbb{R}$ be a stationary point process with finite and positive intensity γ. As α satisfies (3.6) and $\alpha(0)$ is Palm-integrable by (4.2), [6, Theorem 1] can be applied. It states that

$$\mathbb{E}\hat{V}_n(f) = \gamma\mathbb{E}^0[\alpha(0)]\int_{\mathbb{R}} f(x)\mathrm{d}x \tag{5.1}$$

holds for all integrable functions $f : \mathbb{R} \to \mathbb{R}$ with compact support. Hence, if we can show that $\mathbb{E}^0[\alpha(0)] = \gamma^{-1}$, we have shown that $\hat{V}_n(f)$ is unbiased.

For $s \in \mathbb{R}$ reuse the notation I_s from the end of the previous section. When f is an integrable function and $|f| \leq 1$, (3.5) implies

$$\sum_{x \in I_s}|R_x^{(n)}(f)| \leq \sum_{x \in I_s}\Big(\sum_{j=0}^n|\beta_j^{(n)}(x)| + (s_n(x) - x)\Big).$$

The refined Campbell theorem [10, Theorem 3.5.3], (3.1) and (3.6) imply

$$\mathbb{E}\sum_{x\in I_s}|R_x^{(n)}(f)| \le \gamma\mathbb{E}^0\int_{s-s_n}^{s}\Big(\sum_{j=0}^{n}|\beta_j^{(n)}|+s_n\Big)\mathrm{d}x = \gamma\mathbb{E}^0\big[s_n\sum_{j=0}^{n}|\beta_j^{(n)}|+s_n^2\big] < \infty,$$

where the finiteness follows from the supplementary material [12, **??**] and Assumption 4.1, which is weaker than Assumption 2.1. Note that the finite upper bound does not depend on s.

Now let $r > 0$ be given and consider the function $f_r = \mathbf{1}_{[0,r]}$. Recall that

$$R^{(n)}(f_r) = \hat{V}_n(f_r) - \int_{\mathbb{R}} f_r(x)\mathrm{d}x = \frac{1}{n}\sum_{x\in X} R_x^{(n)}(f_r)$$

is the error of the nth Newton-Cotes estimator. The Newton-Cotes approximation on an interval $[x, s_n(x)]$ is exact for all polynomials of degree at most n, and in particular, it is exact for constant functions. Hence, $R_x^{(n)}(f_r) = 0$ whenever $[x, s_n(x)] \cap \{0, r\} = \emptyset$. This implies

$$|\mathbb{E}R^{(n)}(f_r)| \le \mathbb{E}\sum_{x\in I_0}|R_x^{(n)}(f)| + \mathbb{E}\sum_{x\in I_r}|R_x^{(n)}(f)| \le 2C,$$

for some finite $C \in \mathbb{R}$ which is independent of r. Equation (5.1) now implies

$$0 = \lim_{r\to\infty}\tfrac{1}{r}\mathbb{E}R^{(n)}(f_r) = \gamma\mathbb{E}^0\alpha(0) - 1,$$

so $\mathbb{E}^0\alpha(0) = 1/\gamma$ as asserted. □

Remark 5.1. If $X = t\big(U+\mathbb{Z}\big)$ is the equidistant point process, $\alpha(x) = t$ for all $x \in X$. In fact, the Palm version of X is the deterministic set $t\mathbb{Z}$, (3.6) yields

$$\alpha(x;X) = \alpha(x-x;X-x) = \alpha(0;t\mathbb{Z})$$

for all $x \in X$. Hence $\alpha = \alpha(x)$ is deterministic, and $\hat{V}_n(f) = \alpha\sum_{x\in X} f(x)$. Assumption 2.1 is trivially satisfied, so Theorem 2.1 implies the well-known fact that $\hat{V}_n(f)$ is unbiased for $\int f\mathrm{d}x$. This is equivalent to $\alpha = t$.

6. Asymptotic variance behaviour of Newton-Cotes estimators

In this and the following section we derive variance expressions showing the exact dependence on the mean point distance $t > 0$. To this end, we will consider the Peano kernel and its associated covariance function applied to the unit-intensity scaling of the process X: Let the point process of interest X have intensity $1/t$ and define its unit-intensity scaled process by $X^u = X/t$. Hence the Peano kernel $K_m(\,\cdot\,;X^u)$ with respect to X^u remains unchanged when X is rescaled. To avoid intricate notation, we will write $\mathbb{E}_u$, Cov_u, $\mathbb{V}\mathrm{ar}_u$ when expectation, covariance and variance are understood with respect to X^u; for instance $\mathbb{E}_u K_m(r) = \mathbb{E}K_m(r;X^u)$. Similarly, the Palm-distribution of X^u and its expectation are denoted $\mathbb{P}_u^0$ and $\mathbb{E}_u^0$, respectively. Lastly, the covariance function of K_m applied to X^u is denoted by $H_m^u(s) = \mathrm{Cov}_u\big(K_m(s+r), K_m(r)\big)$ for all $s, r \in \mathbb{R}$.

Before proving Theorem 2.3, we recall the variance decomposition of the estimator (1.2) in the equidistant case, as it shows great resemblance to the new non-equidistant set-up. First we introduce the periodic Bernoulli functions P_m, which we define as in [8, Paragraph 297]: Let $(\tilde{P}_m)_{m=0}^{\infty}$ be the sequence of rescaled Bernoulli polynomials, which are defined inductively by $\tilde{P}_0(x) = 1$, $\tilde{P}_1(x) = x - \frac{1}{2}$ and $\tilde{P}'_{m+1} = \tilde{P}_m$, $\tilde{P}_{m+1}(0) = \tilde{P}_{m+1}(1) = \frac{1}{(m+1)!}B_{m+1}$, for $m \in \mathbb{N}$, where B_m is the mth Bernoulli number. This normalization is chosen as in [7] in order to ease comparison with the results there. Then $P_m(x) = \tilde{P}_m(x - \lfloor x \rfloor)$ is the mth Bernoulli polynomial, evaluated at the fractional part of $x \in \mathbb{R}$. Note that P_m is continuous for all $m \neq 1$. When the measurement function f is $(m,1)$-piecewise smooth, the variance decomposes as [7, Chap. 5]

$$\mathbb{V}\mathrm{ar}(\hat{V}(f)) = -t^{2m+2} \sum_{a \in D_{g^{(2m+1)}}} J_{g^{(2m+1)}}(a) P_{2m+2}(\tfrac{a}{t}) + o(t^{2m+2}) \tag{6.1}$$

as $t \downarrow 0$. Here, $g = f * \check{f}$ is the covariogram of f, and the term $o(t^{2m+2})$ can explicitly be given as $-t^{2m+2}\int_{\mathbb{R}} g^{(2m+2)}(s) P_{2m+2}(\frac{s}{t})\mathrm{d}s$. When the point process X is not equidistant, we find a similar variance representation involving the Peano kernels instead of the periodic Bernoulli functions.

Proposition 6.1. *Let $n \in \mathbb{N}$ be given and assume that X is a stationary point process with intensity $1/t$ such that Assumption 2.1 holds. If f is $(m,1)$-piecewise smooth and $k = \min\{m,n\}$, then*

$$\begin{aligned}(-1)^{k+1}\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) = t^{2k+2} &\sum_{a \in D_{g^{(2k+1)}}} J_{g^{(2k+1)}}(a) H_k^u(\tfrac{a}{t}) \\ &+ t^{2k+2} \int_{\mathbb{R}} g^{(2k+2)}(s) H_k^u(\tfrac{s}{t})\mathrm{d}s.\end{aligned} \tag{6.2}$$

If $k = m < n$ or X is weakly n-admissible, the variance behaviour is determined by the first term, as

$$\int_{\mathbb{R}} g^{(2k+2)}(s) H_k^u(\tfrac{s}{t})\mathrm{d}s = o(1)$$

for $t \downarrow 0$.

Proof. The definition of $\alpha(x)$ and elementary calculations give

$$\alpha(x; X) = t\alpha(x/t; X^u)$$

for $x \in X$, so putting $f_t(x) = f(tx)$ we see that

$$\hat{V}_n(f) = t\hat{V}_n(f_t; X^u), \tag{6.3}$$

where the latter estimator is given in terms of the unit-intensity process X^u. As $k \le n$, Theorem 2.2 implies

$$R^{(n)}(f_t; X^u) = \int_{\mathbb{R}} f_t^{(k+1)}(s) K_k(s; X^u)\mathrm{d}s + \sum_{a \in D_{f_t^{(k)}}} J_{f_t^{(k)}}(a) K_k(a; X^u).$$

Using $f_t'(x) = tf'(tx)$ whenever the derivative is defined, we arrive at

$$R^{(n)}(f_t; X^u) = t^k \int_{\mathbb{R}} f^{(k+1)}(s) K_k(\tfrac{s}{t}; X^u) \mathrm{d}s + t^k \sum_{a \in D_{f^{(k)}}} J_{f^{(k)}}(a) K_k(\tfrac{a}{t}; X^u).$$

Hence, using (6.3) and the unbiasedness of $\hat{V}_n$, we get

$$\begin{aligned} \mathbb{V}\mathrm{ar}(\hat{V}_n(f)) &= t^2 \mathbb{V}\mathrm{ar}_u(\hat{V}_n(f_t)) = t^2 \mathbb{E}_u(R^{(n)}(f_t))^2 \\ &= t^{2k+2} \mathbb{E}_u \Big(\int_{\mathbb{R}} f^{(k+1)}(s) K_k(\tfrac{s}{t}) \mathrm{d}s + \sum_{a \in D_{f^{(k)}}} J_{f^{(k)}}(a) K_k(\tfrac{a}{t}) \Big)^2. \end{aligned} \tag{6.4}$$

An application of [7, Prop. 5.7] yields

$$\begin{aligned} f^{(k+1)} * \check{f^{(k+1)}}(x) &= (-1)^{k+1} g^{(2k+2)}(x) \\ &- \sum_{a \in D_{f^{(k)}}} J_{f^{(k)}}(a) f^{(k+1)}(a-x) - \sum_{a \in D_{f^{(k)}}} J_{f^{(k)}}(a) f^{(k+1)}(a+x), \end{aligned} \tag{6.5}$$

and furthermore the jumps of $g^{(2k+1)}$ are given by

$$J_{g^{(2k+1)}}(a) = \sum_{b \in D_{f^{(k)}}} J_{f^{(k)}}(b) J_{f^{(k)}}(b-a);$$

see [7, Eq. (5.12)]. The stationarity and square integrability of K_k from Lemma 4.1 implies that $\mathbb{E}_u K_k(r)$ and $H_k^u(s) = \mathrm{Cov}_u(K_k(r), K_k(s+r))$ are both finite and independent of $r \in \mathbb{R}$. Equation (6.2) now follows by expanding (6.4), applying (6.5), and using the structure of $J_{g^{(2k+1)}}$ together with Fubini's theorem. The latter may be applied due to the square integrability of K_k and the fact that $f^{(k+1)}$ is bounded with compact support.

We now show $\lim_{t \downarrow 0} \int_{\mathbb{R}} g^{(2k+2)}(s) H_k^u(\tfrac{s}{t}) \mathrm{d}s = 0$ if $k = m < n$ or X is weakly n-admissible. The weak admissibility assumption yields

$$\lim_{t \downarrow 0} \int_0^1 H_k^u(\tfrac{s}{t}) \mathrm{d}s = 0 \tag{6.6}$$

for $k = n$. Equation (6.6) also holds for $k = m < n$ without additional assumptions. In fact, for $k < n$ we have $K_{k+1}' = -K_k$ by Lemma 3.1 and thus using Fubini's Theorem,

$$\Big| \int_0^t H_k^u(s) \mathrm{d}s \Big| = \big| \mathrm{Cov}_u\big(K_k(0), K_{k+1}(0)\big) - \mathrm{Cov}_u\big(K_k(0), K_{k+1}(t)\big) \big| \le c < \infty,$$

where Hölder's inequality and the stationarity of the Peano kernels have been used to show that the constant c is independent of t. A substitution allows to derive (6.6) from this.

Now fix $k \le n$ and let $\epsilon > 0$ be given. As $g^{(2k+2)}$ is integrable and bounded, there is a simple function ϕ such that $\phi \le g^{(2k+2)}$ and

$$0 \le \int_{\mathbb{R}} g^{(2k+2)}(s) \mathrm{d}s - \int_{\mathbb{R}} \phi(s) \mathrm{d}s < \frac{\epsilon}{2C},$$

where the finite constant $C > 0$ satisfies $\sup_{s\in\mathbb{R}} |H_k^u(s)| \leq C$. This implies that

$$\left|\int_{\mathbb{R}} g^{(2k+2)}(s)H_k^u(\tfrac{s}{t})\mathrm{d}s - \int_{\mathbb{R}} \phi(s)H_k^u(\tfrac{s}{t})\mathrm{d}s\right| < \frac{\epsilon}{2}.$$

As ϕ is simple, (6.6) implies that $\lim_{t\downarrow 0}\int_{\mathbb{R}} \phi(s)H_k^u(\frac{s}{t})\mathrm{d}s = 0$. We conclude that $|\int_{\mathbb{R}} \phi(s)H_k^u(\frac{s}{t})\mathrm{d}s| < \frac{\epsilon}{2}$ for sufficiently small $t > 0$, and hence

$$\left|\int_{\mathbb{R}} g^{(2k+2)}(s)H_k^u(\tfrac{s}{t})\mathrm{d}s\right| < \epsilon$$

for such small $t > 0$. □

Proof of Theorem 2.3. Recall that $k = \min\{m, n\}$. Due to Lemma 4.1 there exists $C < \infty$ such that $\sup_{s\in\mathbb{R}} |H_k^u(s)| \leq C$, and we immediately see from (6.2) that

$$\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) \leq t^{2k+2}\Big(C\|g^{(2k+2)}\|_\infty \lambda(\operatorname{supp} g) + C\sum_{a\in D_{g^{(2k+1)}}} |J_{g^{(2k+1)}}(a)|\Big),$$

where $\lambda(\operatorname{supp} g) < \infty$ is the Lebesgue measure of the support of g. As g is $(2k+1, 1)$-piecewise smooth by [7, Corollary 5.8], the t-independent constant is finite, and (2.6) therefore follows.

For the stronger result (2.7), note that $m > n$ and hence g is $(2n+3, 1)$-piecewise smooth, and in particular $g^{(2n+2)}$ is continuous. An application of Proposition 6.1 to the $(n, 1)$-piecewise smooth function f and a substitution gives

$$(-1)^{n+1}\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) = t^{2n+3}\int_{\mathbb{R}} g^{(2n+2)}(st)H_n^u(s)\mathrm{d}s. \tag{6.7}$$

Let $b > 0$ satisfy $\operatorname{supp} g \subset [-b, b]$. As $g^{(2n+3)}$ is bounded and measurable, $g^{(2n+2)}$ is absolutely continuous. As H_n^u is bounded and hence integrable on $[-b/t, b/t]$ for any $t > 0$, also the function V given by $V(s) = \int_{-b/t}^{s} H_n^u(y)\mathrm{d}y$ is absolutely continuous on $[-b/t, b/t]$ with derivative H_n^u almost everywhere; see e.g. [4, Section 9.3] for details on absolutely continuous functions. Furthermore, as X and therefore X^u are assumed strongly n-admissible, V is bounded by a t-independent constant C', say. Partial integration for absolutely continuous functions [4, Theorem 9] shows

$$\int_{\mathbb{R}} g^{(2n+2)}(st)H_n^u(s)\mathrm{d}s = \int_{-b/t}^{b/t} g^{(2n+2)}(st)H_n^u(s)\mathrm{d}s = -t\int_{-b/t}^{b/t} g^{(2n+3)}(st)V(s)\mathrm{d}s,$$

where we used that $g^{(2n+2)}$ vanishes at $\pm b$. Returning to (6.7) we find

$$\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) \leq t^{2n+3}t\int_{-b/t}^{b/t} |g^{(2n+3)}(st)||V(s)|\mathrm{d}s \leq t^{2n+3}2b\|g^{(2n+3)}\|_\infty C'.$$

This proves the assertion. □

Remark 6.1. If f is exactly $(m, 1)$-piecewise smooth with $m \leq n$ and (6.6) is satisfied with $k = m$, the variance

$$\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) = (-1)^{m+1}t^{2m+2}\sum_{a\in D_{g^{(2m+1)}}} J_{g^{(2m+1)}}(a)H_m^u(\tfrac{a}{t}) + o(t^{2m+2})$$

is exactly of order t^{2m+2}. This is easily seen by assuming that

$$\sum_{a\in D_{g^{(2m+1)}}} J_{g^{(2m+1)}}(a)H_m^u(\tfrac{a}{t}) \to 0$$

as $t\to 0$, and using that $g^{(2m+1)}$ has a jump at 0. Applying (6.6) yields a contradiction. In particular, the decrease rate in (2.6) is optimal if $m<n$ or X is weakly n-admissible.

7. Variance behaviour under perturbed and cumulative sampling

In this section the general findings will be exemplified and made more explicit for the perturbed model and the model with cumulative errors introduced in Examples 2.1 and 2.2, respectively.

7.1. Perturbed sampling

To construct the perturbed model, we let U be uniform on $(0,1)$ and independent of the sequence of i.i.d. variables $\{E_i\}_{i\in\mathbb{Z}}$, where $|E_i|<\frac{1}{2}$ almost surely and $\mathbb{E}E_1=0$. The perturbed model is the stationary point process $X=\{x_i\}_{i\in\mathbb{Z}}$ for which $x_i = t\big(U+i+E_i\big)$, for all $i\in\mathbb{Z}$. Under its Palm distribution, we have

$$h_k = t\big(1+E_k-E_{k-1}\big)\le 2t, \tag{7.1}$$

$k\in\mathbb{Z}$, so (2.2) is equivalent to

$$\mathbb{E}(1+E_1-E_0)^{-j}<\infty \tag{7.2}$$

for all $j\in\mathbb{N}$. For instance, (7.2) holds if there is $\epsilon>0$ such that $|E_0|\le\frac{1}{2}-\epsilon$ almost surely. For the scaled perturbed model X^u we define the shifted kernel K_m^* by

$$K_m^*(r)=K_m(r+U;X^u)=K_m(r;X^u-U)$$

for $m\in\mathbb{N}_0$. Note that it only depends on the perturbations $\{E_i\}$ and not on the initial uniform translation, and thus it is not (necessarily) a stationary process. However, by the i.i.d. structure of $\{E_i\}$ and the fact that $\beta_j^{(n)}$ is a rational function of point-increments, we see that

$$K_m^*(r)\overset{\mathcal{D}}{=}K_m^*(r+k) \tag{7.3}$$

for all $k\in\mathbb{Z}$. This can be used to show that X^u (and equivalently X) is strongly admissible.

Lemma 7.1. *Let $n\in\mathbb{N}$ be given and assume that X is a stationary point process from the perturbed model such that (7.2) holds. Then, for all $m\in\mathbb{N}_0$ and $r\ge 2n+2$, $H_m^u(r)=H_m^u(r+1)$ and $\int_r^{r+1}H_m^u(s)\mathrm{d}s=0$. In particular, X is strongly n-admissible.*

Proof. Fix $n\in\mathbb{N}$ and $r\ge 2n+2$. For such r, the independence between U and $\{E_i\}$ yields

$$\mathbb{E}_u[K_m(0)K_m(r)]=\int_0^1\mathbb{E}[K_m^*(-u)K_m^*(r-u)]\mathrm{d}u=\int_0^1\mathbb{E}[K_m^*(-u)]\mathbb{E}[K_m^*(r-u)]\mathrm{d}u. \tag{7.4}$$

Equations (7.3) and (7.4) imply that $\mathbb{E}_u[K_m(0)K_m(r)] = \mathbb{E}_u[K_m(0)K_m(r+1)]$ and by stationarity of K_m we conclude that $H_m^u(r) = H_m^u(r+1)$.

Returning to (7.4), we find by Fubini's theorem, a substitution and the stationarity of K_m that

$$\begin{aligned}\int_r^{r+1} & \mathbb{E}_u[K_m(0)K_m(s)]\mathrm{d}s \\ &= \int_0^1 \mathbb{E}[K_m^*(-u)] \int_0^1 \mathbb{E}[K_m^*(r+1-u-s)]\mathrm{d}s\mathrm{d}u \\ &= \int_0^1 \mathbb{E}[K_m^*(-u)]\mathbb{E}_u[K_m(r+1-u)]\mathrm{d}u = (\mathbb{E}_u K_m(0))^2,\end{aligned}$$

which yields the asserted properties of H_m^u. This clearly implies that X^u and equivalently X are strongly n-admissible. □

In order to obtain explicit leading terms in Theorem 2.4 we state in the following a connection between the covariance function H_m^u and certain periodic Bernoulli functions. For our purpose, it is enough to consider $m \in \{0,1\}$, but we also state that the result holds for all m, when no perturbations are present.

Lemma 7.2. *Let $n = 1$ and let X be from the perturbed model. Then*

$$H_m^u(r) = (-1)^m \mathbb{E}[P_{2m+2}(r + E_1 - E_0)] \tag{7.5}$$

for $m = 0, 1$ and all $|r| \geq 4$.

If $X = t\big(U + \mathbb{Z}\big)$ (hence $X^u = U + \mathbb{Z}$) is the equidistant model, then

$$H_m^u(r) = (-1)^m P_{2m+2}(r) \tag{7.6}$$

for all $n \in \mathbb{N}$, $m \leq n$ and $r \in \mathbb{R}$.

The proof of Lemma 7.2 can be found in the supplementary material of this paper; see [12, **????**]. As a consequence of (7.6), for the equidistant model $X = t\big(U+\mathbb{Z}\big)$, the variance representation (6.2) found using the Peano kernels coincides with the classical variance representation (6.1) found using Euler-McLaurin formulae.

Before turning to the proof of Theorem 2.4, we emphasize that the integrability condition (2.2), or equivalently, condition (7.2), was omitted in the statement of the Theorem as we work with the trapezoidal rule. In fact, the unbiasedness of $\hat{V}_1(f)$ for all stationary point processes and integrable, compactly supported functions f was already remarked in the paragraph following the statement of Theorem 2.1. Due to (7.1), the weights satisfy

$$\beta_0^{(1)}(x) = \beta_1^{(1)}(x) = \tfrac{1}{2}h_1(x) \leq t,$$

$x \in X$, which replaces condition (2.2) in all the arguments in Sections 5 and 6. The assumptions of Proposition 6.1 are thus satisfied.

Proof of Theorem 2.4. Let $m \in \{0,1\}$. The $(2m+1)$st derivative of the covariogram g is an odd function, implying $J_{g^{(2m+1)}}(0) = 2g^{(2m+1)}(0^+)$. As X is strongly admissible, Proposition 6.1 in combination with (7.5) yields the variance decomposition

$$\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = (-1)^{m+1}t^{2m+2}2g^{(2m+1)}(0^+)H_m^u(0) + Z_m(t) + o(t^{2m+2}),$$

where the Zitterbewegung $Z_m(t)$ is given by

$$Z_m(t) = -t^{2m+2} \sum_{a \in D_{g^{(2m+1)}} \setminus \{0\}} J_{g^{(2m+1)}}(a) \mathbb{E}[P_{2m+2}(\tfrac{a}{t} + E_1 - E_0)]. \tag{7.7}$$

The facts that Z_m is a finite sum of terms each oscillating around 0 and that it is of order t^{2m+2} follow from arguments similar to those of [7, Section 5.2] as f is assumed to be exactly $(m, 1)$-piecewise smooth. By the refined Campbell Theorem [10, Theorem 3.5.3] and the facts that $\mathbb{E}_u K_0(0) = 0$ and $\mathbb{E}_u K_1(0) = \frac{1}{2}\mathbb{E}_u^0[R_0^{(1)}((\cdot)^2)]$ by (4.3), we find that $H_m^u(0) = \mathbb{V}\mathrm{ar}_u(K_m(0))$ satisfies

$$H_0^u(0) = \mathbb{E}_u^0 \int_0^{h_1} (\tfrac{1}{2}h_1 - y)^2 \mathrm{d}y = \tfrac{1}{12}\mathbb{E}_u^0 h_1^3 \ , \tag{7.8}$$

$$H_1^u(0) = \mathbb{E}_u^0 \int_0^{h_1} (\tfrac{1}{2}h_1 y - \tfrac{1}{2}y^2)^2 \mathrm{d}y - (\tfrac{1}{12}\mathbb{E}_u^0 h_1^3)^2 = \tfrac{1}{120}\mathbb{E}_u^0 h_1^5 - \tfrac{1}{144}(\mathbb{E}_u^0 h_1^3)^2. \tag{7.9}$$

Using (7.1), it is elementary to conclude (2.8) and (2.9).

Now let $m \geq 2$ be given and define $\tilde{H}_1^u$ by $\tilde{H}_1^u(s) = H_1^u(s) + \mathbb{E}[P_4(s + E_1 - E_0)]$. Due to Lemma 7.2, $\tilde{H}_1^u(s)$ vanishes for $|r| > 4$. Since $g^{(4)}$ is continuous, an application of Proposition 6.1 to the $(1,1)$-piecewise smooth function f, Fubini's theorem and the refined partial integration formula [7, Lemma 4.1] yield

$$\begin{aligned}\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = t^5 \int_{\mathbb{R}} g^{(4)}(st)\tilde{H}_1^u(s)\mathrm{d}s - t^6 \int_{\mathbb{R}} g^{(6)}(s)\mathbb{E}[P_6(\tfrac{s}{t} + E_1 - E_0)]\mathrm{d}s \\ - t^6 \sum_{a \in D_{g^{(5)}}} J_{g^{(5)}}(a)\mathbb{E}[P_6(\tfrac{a}{t} + E_1 - E_0)].\end{aligned} \tag{7.10}$$

As the last two terms in (7.10) are of order $o(t^5)$, we only have to simplify the first term.

For all sufficiently small $t > 0$ and all $s \in \mathbb{R}$ with $|s| \leq 4$ the function $g^{(4)}$ is differentiable on the open interval with endpoints 0 and st, so there is a point ξ_{st} in this interval such that

$$g^{(4)}(st) = g^{(4)}(0) + g^{(5)}(\xi_{st})st$$

by the mean value theorem. Inserting this into the first term of (7.10), and using the fact that $g^{(5)}$ and $\tilde{H}_1^u$ are bounded, yields

$$\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = t^5 g^{(4)}(0) \int_{-4}^{4} \tilde{H}_1^u(s)\mathrm{d}s + o(t^5) \tag{7.11}$$

as $t \downarrow 0$.

Noting that P_4 integrates to 0 on each interval of unit length, another application of Fubini's theorem, the refined partial integration formula [7, Lemma 4.1] and Lemma 3.1 gives

$$\begin{aligned}\int_{-4}^{4} \tilde{H}_1^u(s)\mathrm{d}s &= \int_{-4}^{4} H_1^u(s)\mathrm{d}s \\ &= \mathbb{E}_u\Big(\big[K_2(-4) + \sum_{x \in X^u \cap [-4,4]} J_{K_2}(x) - K_2(4)\big]\big(K_1(0) - \mathbb{E}_u K_1(0)\big)\Big).\end{aligned} \tag{7.12}$$

The arguments that lead to (7.4) in combination with (7.3) where $r = -4$ and $k = 8$ imply $\mathbb{E}_u[K_2(-4)K_1(0)] = \mathbb{E}_u[K_2(4)K_1(0)]$, and the two marginal terms in the last expression of (7.12) cancel. Hence, by the refined Campbell Theorem [10, Theorem 3.5.3] and the translation covariance of J_{K_2},

$$\begin{aligned}\int_{-4}^{4} \tilde{H}_1^u(s)\mathrm{d}s &= \mathbb{E}_u \sum_{x \in X^u \cap [-4,4]} J_{K_2}(x)\big(K_1(0) - \mathbb{E}_u K_1(0)\big) \\ &= \mathbb{E}_u^0 J_{K_2}(0) \int_{-4}^{4} \big(K_1(x) - \mathbb{E}_u K_1(0)\big)\mathrm{d}x = \sum_{j=-3}^{2} \theta_j + Q.\end{aligned} \tag{7.13}$$

where $\theta_j = \mathbb{E}_u^0 J_{K_2}(0) \int_{s_j}^{s_{j+1}} \big(K_1(x) - \mathbb{E}_u K_1(0)\big)\mathrm{d}x$ and

$$Q = \mathbb{E}_u^0 J_{K_2}(0) \int_{-4}^{s_{-3}} \big(K_1(x) - \mathbb{E}_u K_1(0)\big)\mathrm{d}x + \mathbb{E}^0 J_{K_2}(0) \int_{s_3}^{4} \big(K_1(x) - \mathbb{E}_u K_1(0)\big)\mathrm{d}x.$$

Here we have used the fact that $s_3 \leq 4$, $s_5 \geq 4$, $s_{-3} \geq -4$ and $s_{-5} \leq -4$ under $\mathbb{P}_u^0$ (the Palm-distribution of X^u). Using Lemma 4.1, it is seen that $J_{K_2}(0) = (1/12)h_1^3$ and consequently that θ_j evaluates to $\mathbb{E}_u^0(1/144)h_1^3(h_{j+1}^3 - h_{j+1}\mathbb{E}_u^0 h_1^3)$. As h_{j+1} only depends on the perturbations E_{j+1} and E_j, we conclude by independence that $\theta_j = 0$ for all $|j| > 1$. Moreover, a coupling argument shows that $Q = \theta_3 + \theta_4 = 0$. The Palm expectation of $J_{K_2}(0) \int_{-4}^{s_{-3}} \big(K_1(x) - \mathbb{E}_u K_1(0)\big)\mathrm{d}x$ is unchanged when we put $E_{-3} = E_5$, $E_{-4} = E_4$ and $E_{-5} = E_3$. Under this coupling assumption, $s_{-3} = s_5 - 8$, $s_{-4} = s_4 - 8$, $s_{-5} = s_3 - 8$, $h_{-3} = h_5$, and $h_{-4} = h_4$, and hence

$$\begin{aligned}Q &= \mathbb{E}_u^0 J_{K_2}(0) \int_{-4}^{s_5 - 8} \big(K_1(x+8) - \mathbb{E}_u K_1(0)\big)\mathrm{d}x + \mathbb{E}_u^0 J_{K_2}(0) \int_{s_3}^{4} \big(K_1(x) - \mathbb{E}_u K_1(0)\big)\mathrm{d}x \\ &= \mathbb{E}_u^0 J_{K_2}(0) \int_{s_3}^{s_5} \big(K_1(x) - \mathbb{E}_u K_1(0)\big)\mathrm{d}x = \theta_3 + \theta_4.\end{aligned}$$

Summarizing, we obtain from (7.13) that

$$\begin{aligned}\int_{-4}^{4} \tilde{H}_1^u(s)\mathrm{d}s &= \tfrac{1}{144} \sum_{j=-1}^{1} \mathbb{E}_u^0 h_1^3(h_{j+1}^3 - h_{j+1}\mathbb{E}_u^0 h_1^3) \\ &= \tfrac{1}{8}(2\mu_4 + \mu_2\mu_4 - \mu_2^3 - \mu_3^2),\end{aligned}$$

where the last equality follows from lengthy and tedious –but elementary– calculations. Inserting this into (7.11) yields the assertion (2.10).

The expression (2.11) of the remainder is found by different arguments which will be detailed in the upcoming thesis [11]. □

7.2. Cumulative sampling

Before turning to the proof of Theorem 2.5, we state in Lemma 7.3 below that the covariance function of the Peano kernel decreases exponentially, from which admissibility follows.

The unit-intensity scaled cumulative process X^u is a stationary point process with i.i.d. holding times $\{\omega_i\}_{i\in\mathbb{Z}}$. We assume that ω_1 has cumulative distribution function

F with density wrt. Lebesgue measure such that $F(0) = 0$. Moreover, since X^u has intensity 1, the holding times satisfy $\mathbb{E}\omega_i = 1$. To explicitly construct the point process X^u, the first point X_0 of $X^u \cap (0, \infty)$ is chosen with cumulative distribution function G,

$$G(x) = \int_0^x \bar{F}(y)\mathrm{d}y, \qquad x \geq 0,$$

where $\bar{F}(y) = 1 - F(y)$; see eg. [1, Chap. V: Cor. 3.6]. Note that the distribution G has density $\bar{F}$. Given X_0, the last point X_{-1} of $X \cap (-\infty, 0)$ (i.e. largest point) is chosen according to $X_{-1} = X_0 - \omega^*$, where ω^* is the conditional distribution of ω_0 given $\omega_0 > X_0$. This assures that $X_{-1} < 0$, and corrects [14], where ω_0 was used instead of ω^*. Having chosen increments $\{\omega_i\}_{i\neq 0}$ independent of X_{-1}, X_0, and setting $x_0 = X_0$, $x_i = X_0 + \sum_{\ell=1}^{i} \omega_\ell$ and $x_{-i} = X_{-1} + \sum_{\ell=1}^{i-1} \omega_{-\ell}$, for all $i \in \mathbb{N}$, we obtain a realization $X^u = \{x_i\}_{i\in\mathbb{Z}}$ of the cumulative point process. This construction implies that the point interval containing the origin has the length weighted distribution, as expected.

The following lemma is stated in terms of the scaled unit-intensity cumulative process X^u, but it is easily seen that it might as well have been formulated in terms of the process X with intensity $1/t$.

Lemma 7.3. *Let $n \in \mathbb{N}$ be given, and let the unit-intensity process X^u be from the cumulative model such that $\mathbb{E}\mathrm{e}^{\eta\omega_1} < \infty$ for some $\eta > 0$, and such that Assumption 2.1 is satisfied. Then*

$$H_m^u(s) = O(\mathrm{e}^{-\epsilon s}), \qquad s \to \infty, \tag{7.14}$$

for some $\epsilon > 0$. In particular, X^u and X are strongly n-admissible.

Proof. The admissibility claim obviously follows from (7.14).

In the case of the trapezoidal estimator we can state the theorem without the integrability assumption (2.2). This is because finite moments of the Peano kernel only require (2.2) to be true for $j \in \mathbb{N}$. As we assume that the increments have exponential moments, they in particular have finite moments of any positive order, and hence, all integrability results of the Peano kernels apply.

The proof relies on exponential decays in renewal theory, and we refer to [1, Chapter V] for an introduction. Moreover, for fixed $n \in \mathbb{N}$ and all $m \in \mathbb{N}_0$, we will explicitly use the fact that $K_m(s)$ depends on n points of the underlying point process to each side of s.

Let $n \in \mathbb{N}$ and $m \in \mathbb{N}_0$ be given. Let $N = (N(s))_{s\geq 0}$ be a pure renewal process with increments $\{\omega_i\}_{i\in\mathbb{N}}$, and let U be the corresponding renewal measure. Also, let $y_0 = 0$ and $y_i = \sum_{\ell=1}^{i} \omega_\ell$ for $i \in \mathbb{N}$, that is, $X^u \cap (0, \infty) = \{X_0 + y_i\}_{i\in\mathbb{N}_0}$. Then $y_i \sim F^{*i}$ for all $i \in \mathbb{N}$. Define $\psi : [0, \infty) \to \mathbb{R}$ by $\psi(s) = \mathbb{E}[K_m(s; N)\mathbf{1}_{y_{n-1}\leq s}] = \mathbb{E}_u^0[K_m(s)\mathbf{1}_{s_{n-1}\leq s}]$, and let $\tilde{\psi} : [0, \infty) \to \mathbb{R}$ be given by

$$\tilde{\psi}(s) = \mathbb{E}[K_m(s; N)\mathbf{1}_{y_{n-1}\leq s}\mathbf{1}_{y_n > s}] = \mathbb{E}_u^0[K_m(s)\mathbf{1}_{s_{n-1}\leq s}\mathbf{1}_{s_n > s}].$$

Then $\psi(s) = U * \tilde{\psi}(s)$. This can be seen by a renewal argument obtaining the renewal equation $\psi = \tilde{\psi} + \psi * F$, which has the desired solution. Another rather intuitive approach is to condition on the nth to last point of N prior to s happening at time

$x \in [0,s)$, which has probability $U(\mathrm{d}x)\bar{F} * F^{*(n-1)}(s-x)$. Now we initialize a new independent pure renewal process at time x and we obtain, integrating over $[0,s]$,

$$\begin{aligned}\psi(s) &= \int_0^s \mathbb{E}[K_m(s-x;N) \mid y_{n-1} \le s-x,\, y_n > s-x]\bar{F} * F^{*(n-1)}(s-x)U(\mathrm{d}x)\\ &= \int_0^s \tilde{\psi}(s-x)U(\mathrm{d}x)\\ &= U * \tilde{\psi}(s).\end{aligned}$$

The exponential moment assumption implies $\bar{F}(s) = O(\mathrm{e}^{-\eta s})$, which in turn implies that also $\bar{G}(s) = O(\mathrm{e}^{-\eta s})$, as $s \to \infty$. Moreover, $\overline{G * F^{*i}}(s) = O(\mathrm{e}^{-\eta s})$, $s \to \infty$, for all $i \in \mathbb{N}$. We consider

$$\mathbb{E}_u K_m(0)K_m(s) = \mathbb{E}_u K_m(s)K_m(0)\mathbf{1}_{X_0+y_{2n-2}\le s} + \mathbb{E}_u K_m(s)K_m(0)\mathbf{1}_{X_0+y_{2n-2}>s},$$

and an application of Cauchy-Schwarz inequality yields

$$\begin{aligned}\mathbb{E}_u K_m(s)K_m(0)\mathbf{1}_{X_0+y_{2n-2}>s} &\le [\mathbb{E}_u K_m^2(s)K_m^2(0)]^{1/2}\mathbb{P}(X_0+y_{2n-2}>s)^{1/2}\\ &\le C\left(\overline{G * F^{*(2n-2)}}(s)\right)^{1/2}\end{aligned}$$

for some finite C. Hence $\mathbb{E}_u K_m(s)K_m(0)\mathbf{1}_{X_0+y_{2n-2}>s} = O(\mathrm{e}^{-\eta s/2})$ as $s \to \infty$, and (7.14) therefore follows once we show that

$$\mathbb{E}_u K_m(s)K_m(0)\mathbf{1}_{X_0+y_{2n-2}\le s} = (\mathbb{E}_u K_m(0))^2 + O(\mathrm{e}^{-\epsilon s}) \tag{7.15}$$

for some $\epsilon > 0$, as $s \to \infty$. We apply a renewal argument conditioning on the nth arrival in $X^u \cap (0,\infty)$, that is, conditioning on the value of $X_0 + y_{n-1} \sim G * F^{*(n-1)}$, and then initializing a new independent pure renewal process,

$$\begin{aligned}&\mathbb{E}_u K_m(s)K_m(0)\mathbf{1}_{X_0+y_{2n-2}\le s}\\ &\quad= \int_0^s \mathbb{E}_u\big[K_m(0)K_m(s)\mathbf{1}_{X_0+y_{2n-2}\le s} \mid X_0+y_{n-1}=v\big]\big(G * F^{*(n-1)}\big)(\mathrm{d}v)\\ &\quad= \int_0^s \mathbb{E}\big[K_m(s-v;N)\mathbf{1}_{y_{n-1}\le s}\big]\mathbb{E}_u\big[K_m(0) \mid X_0+y_{n-1}=v\big]\big(G * F^{*(n-1)}\big)(\mathrm{d}v)\\ &\quad= \int_0^s \psi(s-v)\mathbb{E}_u\big[K_m(0) \mid X_0+y_{n-1}=v\big]\big(G * F^{*(n-1)}\big)(\mathrm{d}v).\end{aligned}$$

Since $\mathbb{E}_u K_m(0) = \mathbb{E}_u^0 J_{K_{m+1}}(0)$ due to (4.3), an application of Fubini's theorem yields

$$\begin{aligned}\int_0^\infty \tilde{\psi}(s)\mathrm{d}s &= \mathbb{E}\int_{y_{n-1}}^{y_n} K_m(s;N)\mathrm{d}s\\ &= \mathbb{E}K_{m+1}(y_{n-1}^+;N) - \mathbb{E}K_{m+1}(y_n^-;N) = \mathbb{E}_u K_m(0),\end{aligned}$$

and consequently, by [1, Chapter VII: Thm. 2.10(iii)],

$$\psi(s) = U * \tilde{\psi}(s) = \mathbb{E}_u K_m(0) + O(\mathrm{e}^{-\epsilon' s}) \tag{7.16}$$

for some $0 < \epsilon' < \eta$, as $s \to \infty$. Furthermore, by another application of Cauchy-Schwarz inequality, we conclude that

$$\begin{aligned}\int_0^s \mathbb{E}_u\big[K_m(0) \mid X_0 + y_{n-1} = v\big]\big(G * F^{*(n-1)}\big)(\mathrm{d}v) \\ = \mathbb{E}_u K_m(0) - \mathbb{E}_u K_m(0)\mathbf{1}_{X_0+y_{n-1}>s} \\ = \mathbb{E}_u K_m(0) + O(\mathrm{e}^{-\eta s/2})\end{aligned} \tag{7.17}$$

as $s \to \infty$. Combining (7.16) and (7.17) yields (7.15). □

As for the perturbed model, Theorem 2.5 is stated without Assumption 2.1. This is because the strong admissibility and the variance decomposition are satisfied for the trapezoidal rule, when assuming (2.2) for $j \in \mathbb{N}$ only. This relaxed assumption is ensured by the finite exponential moments of the increments.

Proof of Theorem 2.5. Let $m \in \{0, 1\}$. As X is strongly admissible, Proposition 6.1 in combination with the decrease rate (7.14) yields the variance decomposition

$$\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = (-1)^{m+1}t^{2m+2}2g^{(2m+1)}(0^+)H_m^u(0) + o(t^{2m+2}).$$

From (7.8) and (7.9) and the fact that $\mathbb{E}_u^0 h_1^j = \nu_j$, we conclude (2.12) and (2.13). □

8. An application in stereology

To illustrate the general theory, we describe a geometric application that also was the original motivation for this work. In stereology, the volume of a compact object $Y \subset \mathbb{R}^3$ can be approximated from sections with equidistant and parallel planes with joint normal direction ν in the unit sphere S^2, if the area of each intersection profile is accessible; see [3, Chap. 7].

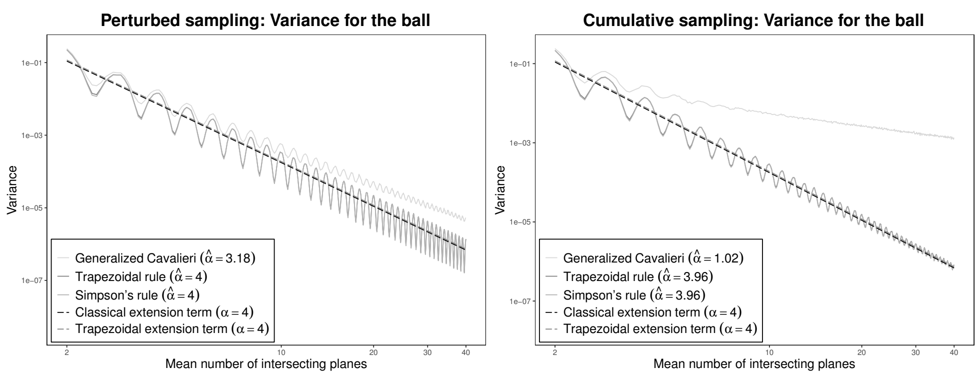
Formally, if $f(x)$ is the area of the intersection of Y with the plane $\{y \in \mathbb{R}^3 : \nu^T y = x\}$ positioned at a signed distance $x \in \mathbb{R}$ from the origin along ν, the integral $\int f \mathrm{d}x$ coincides with the volume of Y by Fubini's theorem. If $f(x)$ is available at all points of the equidistant stationary point process $X = t(U + \mathbb{Z})$, the volume-estimator (1.2) can be used and is called the *(classical) Cavalieri estimator.* When $f(x)$ is known at the points of a stationary point process X with intensity $1/t$, the so-called *(generalized) Cavalieri estimator* (1.1) can be used. However, as outlined above, the generalized Cavalieri estimator does not exploit the smoothness of f and thus has a suboptimal decrease rate as $t \downarrow 0$. We therefore suggest employing Newton-Cotes estimators (of appropriate order) instead.

This section is devoted to Monte Carlo simulations illustrating the advantage of the new estimators when the sampling points are not equidistant. For illustration purposes we start considering the Euclidean unit ball $Y = \{z \in \mathbb{R}^3 : \|z\| \leq 1\}$. In this case, the measurement function is

$$f(x) = \mathbf{1}_{[-1,1]}(x)\pi(1 - x^2),$$

which is a $(1, \infty)$-piecewise smooth function as f' has jumps and is piecewise linear. Applying the classical Cavalieri estimator to such a function yields the extension term $\mathbb{V}\mathrm{ar}_E(\hat{V}(f)) = \frac{\pi^2}{90}t^4$ due to (2.5). Using sampling by the perturbed model or the model with cumulative errors we expect that the generalized Cavalieri estimator decreases at a rate of 3 and 1, respectively, whereas the trapezoidal estimator ($n = 1$)

and Simpson's estimator ($n = 2$) decreases at a rate of 4 in both point-models, an asymptotic behaviour visible in Figure 1 below. It shows the empirical variances of those three estimators based on 2000 Monte Carlo simulations as functions of the mean number of sections, that is $2/t$, with the variance plot including the extension term of the classical Cavalieri estimator and the extension term of the trapezoidal estimator as given by the dominating terms in (2.9) and (2.13) for the perturbed and cumulative model, respectively. The variances in this and the following figures are shown in a double-logarithmic scale with α and $\hat{\alpha}$ being the theoretical and approximate rates of decrease ($\hat{\alpha}$ has been found by the least squares method applied to the datapoints where $15 \leq 2/t \leq 40$).



(a) Perturbed sampling (b) Cumulative sampling

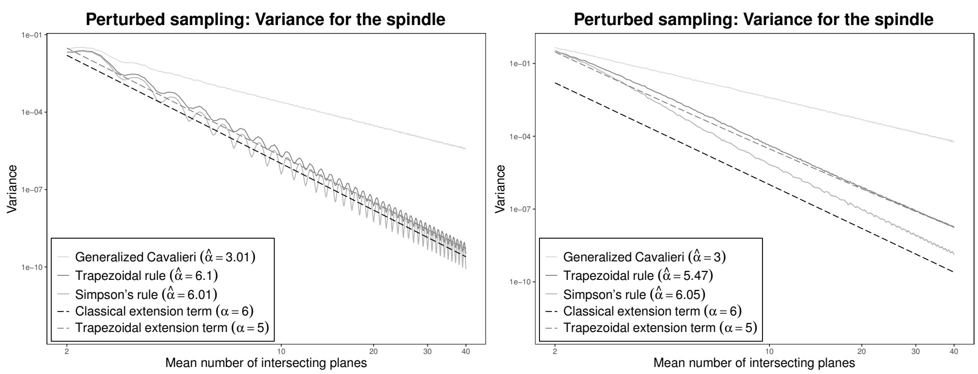
FIGURE 1: Empirical variance for the volume estimation of the unit ball in $\mathbb{R}^3$ based on perturbed sampling with $E_i \sim \text{Unif}((-s, s))$ and sampling with cumulative errors with $\omega_i \sim \text{Unif}((1-c, 1+c))$. We choose s and c such that the average relative deviation (the coefficient of error) of the point-increment from the ideal increment 1 is 5%. In both figures, the graph of the trapezoidal estimator (solid dark grey) is almost completely hidden by the graph of Simpson's estimator (solid grey), and the trapezoidal extension term (dashed dark grey) is almost identical to the classical extension term (dashed black).

The graphs of Figure 1 are characteristic for the behaviour of variances and extensions terms for objects with $(1, 1)$-piecewise smooth measurement functions. For instance ellipsoids, or, more generally, strictly convex bodies lead to the same variance behaviour apart from the facts that intercepts of these curves may be shifted and the Zitterbewegung may differ.

For comparison, we therefore give another example, where the measurement function exhibits a higher order of smoothness. The measurement function

$$f(x) = \mathbf{1}_{[-1,1]}(x) \frac{\pi}{2}(1 + \cos \pi x),$$

is obtained from a spindle shaped body of revolution, if all section planes are orthogonal to the rotation axis. The corresponding convex body is illustrated in [6, Fig. 4]. The measurement function f is $(2, \infty)$-piecewise smooth. Using this measurement function, the extension term of the classical estimator is $\mathbb{V}\text{ar}_E(\hat{V}(f)) = \frac{\pi^6}{60480} t^6$. Figure 2 shows empirical variances based on perturbed sampling with the two extension terms included, where the extension term of the trapezoidal estimator is given as the sum of the dominating terms in (2.10) and (2.11). In Figure 2a we use small perturbations to illustrates the fact that the dominating term in (2.10) can be made arbitrarily small. Hence, a decrease rate of 6 for the variance of the trapezoidal estimator can be a good approximation with small perturbations, as the trapezoidal extension term is approximately given by $1.7 \cdot 10^{-4}t^5 + 3.0 \cdot 10^{-2}t^6$ here. Even when we consider $100 \leq 2/t \leq 200$, we only obtain an approximate decrease rate of $\hat{\alpha} = 5.66$. For comparison, Figure 2b gives a better illustration of the actual asymptotic rate of decrease which corresponds to the bound from Theorem 2.3, that is, $\alpha = 5$. Here we use larger perturbations, which in turn gives an approximate trapezoidal extension term of $0.043t^5 + 0.25t^6$. Increasing the number of intersecting planes to $2/t \leq 100$ the actual rate is even more apparent, as we here obtain an approximate decrease rate of $\hat{\alpha} = 5.19$.



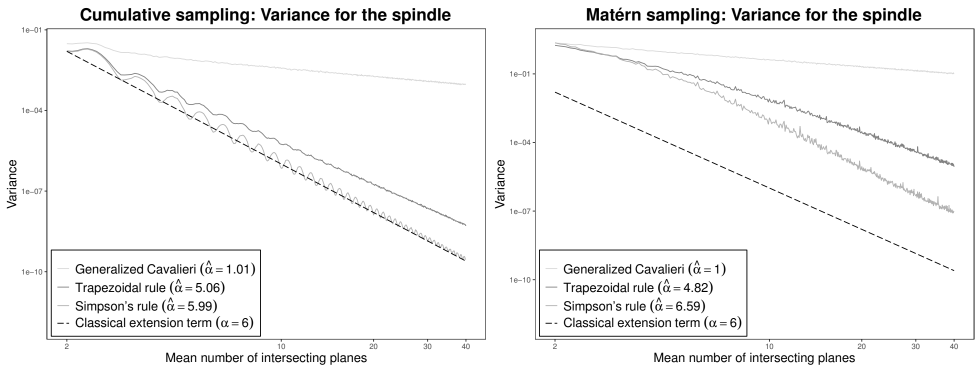
(a) 5% CE (small perturbations) (b) 20% CE (large perturbations)

FIGURE 2: Empirical variance for the volume estimation of a spindle shaped body of revolution in $\mathbb{R}^3$ based on perturbed sampling with $E_i \sim \mathrm{Unif}((-s, s))$. We choose s such that the coefficient of error (CE) of the point-increments are 5% (left) and 20% (right).

The last two simulations are meant to illustrate the findings in Theorem 2.3 for point process models where we do not have explicit formulae for the extension term. The first is the already discussed model with accumulated errors. To illustrate the wide range of point process models to which our results apply, we also simulated from the Matérn hard core process of type II; see [10, sec. 3.5 pp. 93-94], which satisfies the strong integrability assumption (2.2). The empirical variances for the aforementioned spindle shaped body are depicted in Figure 3. It is worth noticing that the variance of the trapezoidal estimator under the Matérn model seem to satisfy the strong bound of Theorem 2.3, that is (2.7). Increasing the number of intersecting planes to $2/t \leq 100$ the result is more clear, as we find approximate decrease rates of $\hat{\alpha} = 4.94$ and $\hat{\alpha} = 6.07$ for the trapezoidal estimator and Simpson's estimator, respectively.

9. Conclusions and future work

Estimating integrals based on known randomized sampling points with unequal increments, we have shown that higher order Newton-Cotes quadratures are to be



(a) Cumulative sampling (b) Matérn sampling

FIGURE 3: Empirical variance for the volume estimation of a spindle shaped body in $\mathbb{R}^3$ based on sampling with cumulative errors with $\omega_i \sim \text{Unif}((1-c, 1+c))$ and sampling with a Matérn hard core process of type II with intensity 1 and a hard core distance of 0.4. The value of c is chosen such that the coefficient of error of the increment is 5%.

preferred over naïve Riemann sums, as they are unbiased and have a faster decrease in variance for decreasing average point-increment. In particular, if the measurement function is exactly $(n, 1)$-piecewise smooth, applying nth order Newton-Cotes estimation yields an upper bound of the variance decreasing at the same rate as the variance based on equidistant sampling, that is, a rate of $2n+2$. Applying nth order estimation to a function with smoothness of order, say, $m > n$, the variance has an upper bound with a rate of decrease of $2n + 2$ in the general case, whereas the bound decreases at the rate $2n + 3$ if the set of sampling points is *strongly n-admissible*. We have shown that point processes from the perturbed and cumulative models are strongly admissible and thus the strong bound holds in these cases. Based on a simulation study of the trapezoidal estimator it appears that also sampling from Matérn's hard core model of the second kind satisfies the strong bound. From a practical point of view the trapezoidal estimator is very interesting as the unbiasedness does not require any integrability conditions of the underlying sampling model. Applying this estimator to perturbed and cumulative sampling we have found asymptotic variance expressions, with an overall trend arbitrarily close to the trend of the equidistant case if the perturbations are small and the increments are close to 1, respectively. This asymptotic trend can be calculated if only the derivatives of the covariogram of the measurement function is known at 0, and if moments of the perturbations and increments, respectively, can be computed. This observation allows in principle to estimate the extension term of the variance from measurements of sampling positions and sampled areas in analogy to established methods in the classical, equidistant case. We intend to carry out this program in a future study.

It is an open question if the variance bounds in Theorem 2.3 are optimal in all cases. As the rate of decrease in (2.6) is optimal if the model is weakly admissible or the order of the estimator exceeds the order of smoothness of the measurement function, we expect that the rate in (2.6) is optimal for any stationary point process satisfying the assumptions of the theorem. Similarly we know that the bound presented in (2.7) yields the optimal decay-rate when $n = 1$ under the perturbed model (assuming non-degenerate perturbations), and thus it is of interest to investigate whether this is the case for all n in perturbed sampling and in general for any admissible point process with unequal increments.

Acknowledgements

This work was supported by the Centre for Stochastic Geometry and Advanced Bioimaging, funded by a grant from the Villum Foundation.

References

[1] Asmussen, S. (2003). *Applied Probability and Queues.* Applications of Mathematics: Stochastic Modelling and Applied Probability. Springer, New York.

[2] Baddeley, A., Dorph-Petersen, K.-A. and Jensen, E.B.V. (2006). A note on the stereological implications of irregular spacing of sections. *J. Microsc* **222,** 177–181.

[3] Baddeley, A. and Jensen, E.B.V. (2004). *Stereology for Statisticians.* Chapman & Hall/CRC Monographs on Statistics & Applied Probability. CRC Press, Boca Raton.

[4] De Barra, G. (1981). *Measure Theory and Integration.* New Age International, New Delhi.

[5] Doukhan, P. (1994). *Mixing: Properties and Examples.* Lecture Notes in Statistics. Springer, New York.

[6] Kiderlen, M. and Dorph-Petersen, K.-A. (2017). The Cavalieri estimator with unequal section spacing revisited. *Image Anal. Stereol.* **36,** 135–141.

[7] Kiêu, K. (1997). *Three Lectures on Systematic Geometric Sampling.* Technical Report, Aarhus University.

[8] Knopp, K. (1947). *Theorie und Anwendung der unendlichen Reihen.* Die Grundlehren der mathematischen Wissenschaften. Springer, Berlin.

[9] Matheron, G. (1965). *Les variables régionalisées et leur estimation.* Masson et Cie, Paris.

[10] Schneider, R. and Weil, W. (2008). *Stochastic and Integral Geometry.* Springer, Heidelberg.

[11] Stehr, M. (In preperation). Stereology and spatio-temporal models. *PhD thesis.* Aarhus University.

[12] Stehr, M. and Kiderlen, M. (2019). Asymptotic variance of Newton-Cotes quadratures based on randomized sampling points. Supplementary material.

[13] Stoer, J. and Bulirsch, R. (1993). *Introduction to Numerical Analysis.* Texts in Applied Mathematics. Springer, New York.

[14] Ziegel, J., Baddeley, A., Dorph-Petersen, K.-A. and Jensen, E.B.V. (2010). Systematic sampling with errors in sample locations. *Biometrika* **97,** 1–13.

Applied Probability Trust (27 May 2020)

SUPPLEMENTARY MATERIAL: ASYMPTOTIC VARIANCE OF NEWTON-COTES QUADRATURES BASED ON RANDOMIZED SAMPLING POINTS

MADS STEHR,* *AND*

MARKUS KIDERLEN,* *Aarhus University*

SM1. Integrability properties

As mentioned in Section 4 the weight $\beta_j^{(n)}(x)$ is particularly simple.

Lemma SM1.1. *For all $n \in \mathbb{N}$, $x \in X$ and $j = 0, \ldots, n$, the weight $\beta_j^{(n)}(x)$ is a rational function of point-increments,*

$$\beta_j^{(n)}(x) = \frac{q_j^{(n)}(h_1(x), \ldots, h_n(x))}{p_j^{(n)}(h_1(x), \ldots, h_n(x))}$$

where $q_j^{(n)} : (0, \infty)^n \to \mathbb{R}$ is a homogeneous polynomial of degree $n+1$, and $p_j^{(n)} : (0, \infty)^n \to \mathbb{R}$ is a non-vanishing homogeneous polynomial of degree n with non-negative coefficients.

Proof. Fix $x \in X$, $n \in \mathbb{N}$ and $j \in \{0, \ldots, n\}$, and consider $\beta_j^{(n)}(x)$ as defined by (3.3). Recall that points in X are distinct and therefore all point-increments are strictly positive. At first we note that the denominator of the integrand in (3.3) is constant with each term in the product satisfying

$$s_j(x) - s_k(x) = \begin{cases} \sum_{\ell=k+1}^{j} h_\ell(x) & \text{for } j > k, \\ -\sum_{\ell=j+1}^{k} h_\ell(x) & \text{for } j < k, \end{cases}$$

and hence

$$\prod_{\substack{k=0 \\ k \neq j}}^{n} (s_j(x) - s_k(x)) = (-1)^{n-j} p_j^{(n)}(h_1(x), \ldots, h_n(x)),$$

* Postal address: Centre for Stochastic Geometry and Advanced Bioimoaging (CSGB), Department of Mathematics, Aarhus University, Ny Munkegade 118, 8000 Aarhus C, Denmark.

where $p_j^{(n)} : (0,\infty)^n \to \mathbb{R}$ is the polynomial defined by

$$p_j^{(n)}(y_1,\dots,y_n) = \Big(\prod_{k=0}^{j-1}\sum_{\ell=k+1}^{j} y_\ell\Big)\Big(\prod_{k=j+1}^{n}\sum_{\ell=j+1}^{k} y_\ell\Big).$$

The definition of $p_j^{(n)}$ implies that it is non-vanishing with non-negative coefficients and that $p_j^{(n)}(\lambda y_1,\dots,\lambda y_n) = \lambda^n\, p_j^{(n)}(y_1,\dots,y_n)$ for any $\lambda \in (0,\infty)$. That is, it is homogeneous of degree n.

With the abbreviation $\tilde{s}_k(x) = s_k(x) - x = \sum_{\ell=1}^{k} h_\ell(x)$, a substitution yields

$$\int_x^{s_n(x)} \prod_{\substack{k=0\\k\neq j}}^{n} (y - s_k(x))\,\mathrm{d}y = \int_0^{\tilde{s}_n(x)} \prod_{\substack{k=0\\k\neq j}}^{n} (y - \tilde{s}_k(x))\,\mathrm{d}y,$$

for $k \geq 0$. The right side of this equation is a polynomial of degree at most $n+1$ in $(\tilde{s}_0(x),\dots,\tilde{s}_n(x))$, as all its derivatives of order $n+2$ vanish. We therefore can define the polynomial $q_j^{(n)} : (0,\infty)^n \to \mathbb{R}$ by

$$q_j^{(n)}(h_1(x),\dots,h_n(x)) = (-1)^{n-j}\int_0^{\tilde{s}_n(x)} \prod_{\substack{k=0\\k\neq j}}^{n} (y - \tilde{s}_k(x))\,\mathrm{d}y.$$

A substitution argument shows that the right side is homogeneous of degree $n+1$ as a function of $(\tilde{s}_0(x),\dots,\tilde{s}_n(x))$ and thus also as a function of $(h_1(x),\dots,h_n(x))$. This shows the assertion. □

Assuming either Assumption 2.1 or Assumption 4.1, this representation ensures the Palm integrability of $\alpha(0)$, which is used in the proof of Theorem 2.1.

Lemma SM1.2. *Fix $n \in \mathbb{N}$. If X is a stationary point process such that* (4.1) *is satisfied, then*

$$\mathbb{E}^0|\beta_j^{(n)}(s_{-j})| < \infty \tag{SM1.1}$$

for all $j = 0,\dots,n$, and consequently $\mathbb{E}^0|\alpha(0)| < \infty$.

Proof. From Lemma SM1.1 we find real constants $\{c_{\mathbf{m}}^{(n,j)}\}$ and non-negative constants $\{a_{\mathbf{m}'}^{(n,j)}\}$ such that

$$|\beta_j^{(n)}(s_{-j})| = \frac{\Big|\sum\limits_{\substack{\mathbf{m}\in\mathbb{N}_0^n\\|\mathbf{m}|=n+1}} c_{\mathbf{m}}^{(n,j)}\,\mathbf{h}(s_{-j})^{\mathbf{m}}\Big|}{\sum\limits_{\substack{\mathbf{m}'\in\mathbb{N}_0^n\\|\mathbf{m}'|=n}} a_{\mathbf{m}'}^{(n,j)}\,\mathbf{h}(s_{-j})^{\mathbf{m}'}} \leq \sum_{\substack{\mathbf{m}\in\mathbb{N}_0^n\\|\mathbf{m}|=n+1}} \frac{|c_{\mathbf{m}}^{(n,j)}|\,\mathbf{h}(s_{-j})^{\mathbf{m}}}{a_{\mathbf{m}_0'}^{(n,j)}\,\mathbf{h}(s_{-j})^{\mathbf{m}_0'}},$$

where $\mathbf{m}_0'$ is a multi-index such that $a^{(n,j)}_{\mathbf{m}_0'} > 0$ which exists by Lemma SM1.1. By linearity (SM1.1) is satisfied whenever

$$\mathbb{E}^0\left[\frac{\mathbf{h}(s_{-j})^{\mathbf{m}}}{\mathbf{h}(s_{-j})^{\mathbf{m}'}}\right] = \mathbb{E}^0\left[\frac{\mathbf{h}^{\mathbf{m}}}{\mathbf{h}^{\mathbf{m}'}}\right] < \infty, \tag{SM1.2}$$

for all multi-index $\mathbf{m}$ and $\mathbf{m}'$ in $\mathbb{N}_0^n$ with $|\mathbf{m}| = n+1$ and $|\mathbf{m}'| = n$, where the equality is a consequence of the fact that the Palm distribution is invariant under bijective point shifts; see [1, Eq. (13)]. The right side of (SM1.2) is finite by Assumption 4.1. □

SM2. Peano kernels, Bernoulli functions and variance in perturbed sampling

In this section we consider the relation between the Peano kernels K_m and the Bernoulli functions P_m when we sample $X^u = \{U + E_k + k\}_{k\in\mathbb{Z}}$ from the perturbed model (recall that X^u is scaled to have unit-intensity). Note that the unit-intensity equidistant model is obtained with degenerate perturbations concentrated at 0. As in Section 7 we work with the shifted kernel, K_m^*, defined by

$$K_m^*(r) = K_m(r+U) = K_m(r; X^*),$$

where $X^* = X^u - U = \{E_k + k\}_{k\in\mathbb{Z}}$ is the shifted process. From (7.3) the shifted kernel is periodic in law with period 1. Recall that the 1st Bernoulli function is given by $P_1(r) = \tilde{P}_1(r - \lfloor r \rfloor)$, with $\tilde{P}_1(r) = r - \frac{1}{2}$ for $r \in \mathbb{R}$.

Lemma SM2.1. *Let $n \in \mathbb{N}$ be given and let X^u be a unit-intensity process from the perturbed model such that (2.2) is satisfied. Let $X^* = \{x_k\}_{k\in\mathbb{Z}}$, with $x_k = E_k + k$, be its shifted process. For all $r \in \mathbb{R}$, K_0^* satisfies*

$$\mathbb{E}K_0^*(r) = -\mathbb{E}P_1(E_0 - r) + \mathbb{E}\Big[\frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(x_0)j\Big] - \frac{n}{2} + Q(r), \tag{SM2.1}$$

where

$$Q(r) = \begin{cases} \mathbb{E}\mathbf{1}_{E_0 \geq r}[\frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(x_{-j}) - 1] & \textit{for} \quad r < \frac{1}{2}, \\ \mathbb{E}\mathbf{1}_{E_0 \geq r-1}[\frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(x_{-j}) - 1] & \textit{for} \quad r \geq \frac{1}{2}. \end{cases}$$

Furthermore, if $\mathbb{E}K_0^*(r) = -\mathbb{E}P_1(E_0 - r)$ for all $r \in \mathbb{R}$, then

$$H_m^u(r) = (-1)^m\mathbb{E}[P_{2m+2}(r + E_1 - E_0)] \tag{SM2.2}$$

for $m \le n$ and all $|r| \ge 2n+2$. If the perturbations are degenerate, that is X^u is the unit-intensity equidistant model, (SM2.2) *is true for all $r \in \mathbb{R}$.*

Proof. By (7.3) it is enough to consider $r \in [0,1)$. Let $n \in \mathbb{N}$ and $r \in [0,1)$ be given. Recall that

$$nK_0^*(r) = \sum_{i\in\mathbb{Z}} \mathbf{1}_{x_i<r\le x_{i+1}} \sum_{\ell=1-n}^{0} R^{(n)}_{x_{i+\ell}}((\cdot - r)^0_+).$$

Only the summands with $i=-1,0,1$ can be non-zero, and thus

$$nK_0^*(r) = \mathbf{1}_{E_0\ge r}A_{-1}(r) + \mathbf{1}_{E_0<r}\mathbf{1}_{E_1\ge r-1}A_0(r) + \mathbf{1}_{E_1<r-1}A_1(r),$$

where, for $i=-1,0,1$,

$$A_i(r) = \sum_{\ell=1-n}^{0} R^{(n)}_{x_{i+\ell}}((\cdot - r)^0_+) = \sum_{\ell=1-n}^{0}\sum_{j=0}^{n} \beta_j^{(n)}(x_{i+\ell})\mathbf{1}_{\ell+j\ge 1} - \sum_{\ell=1}^{n}(x_{i+\ell}-r).$$

We let q_0 and q_1 be the i.i.d. variables defined by $q_0 = (E_0 - r) - \lfloor E_0 - r\rfloor$ and $q_1 = (E_1 - r) - \lfloor E_1 - r\rfloor$. We will consider the cases $r < \frac{1}{2}$ and $r \ge \frac{1}{2}$ separately.

Let $r<\frac{1}{2}$ be given. As $E_1 \ge r-1$ the kernel simplifies as

$$nK_0^*(r) = \mathbf{1}_{E_0\ge r}A_{-1}(r) + \mathbf{1}_{E_0<r}A_0(r).$$

Note that $q_0 = E_0 - r$ when $E_0 \ge r$, and $q_0 = E_0 - r + 1$ when $E_0 < r$. Using the independence of the perturbations, $\mathbb{E}E_i = 0$, and the representation of the second power sum, we find

$$\mathbb{E}\mathbf{1}_{E_0\ge r}A_{-1}(r) = \mathbb{E}\mathbf{1}_{E_0\ge r}\Big(\sum_{j=0}^{n}\sum_{\ell=1-j}^{0} \beta_j^{(n)}(x_{\ell-1}) - n\tilde{P}_1(q_0) + (n-1)E_0 - \frac{n^2}{2}\Big),$$
$$\mathbb{E}\mathbf{1}_{E_0<r}A_0(r) = \mathbb{E}\mathbf{1}_{E_0<r}\Big(\sum_{j=0}^{n}\sum_{\ell=1-j}^{0} \beta_j^{(n)}(x_\ell) - n\tilde{P}_1(q_0) + nE_0 - \frac{n^2}{2}\Big).$$

Constant functions are approximated exactly, and hence $\sum_{j=0}^{n}\beta_j^{(n)}(x_0) = x_n - x_0 =$

$E_n - E_0 + n$. An index-shift in the former term above then implies

$$\begin{aligned}\mathbb{E}K_0^*(r) &= -\mathbb{E}\tilde{P}_1(q_0) + \mathbb{E}\Big[\frac{1}{n}\sum_{j=0}^{n}\sum_{\ell=1-j}^{0}\beta_j^{(n)}(x_\ell)\Big] \\ &\quad - \frac{n}{2} + \mathbb{E}\mathbf{1}_{E_0\geq r}\Big[\frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(x_{-j}) - 1\Big] \\ &= -\mathbb{E}P_1(E_0 - r) + \mathbb{E}\Big[\frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(x_0)j\Big] \\ &\quad - \frac{n}{2} + \mathbb{E}\mathbf{1}_{E_0\geq r}\Big[\frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(x_{-j}) - 1\Big],\end{aligned}$$

where the last equality follows as $\beta_j^{(n)}(x_\ell)$ equals $\beta_j^{(n)}(x_0)$ in law, as they are rational functions of identically distributed increments.

Now let $r \geq \frac{1}{2}$ be given. Then $E_0 < r$ and the kernel simplifies as

$$nK_0^*(r) = \mathbf{1}_{E_1\geq r-1}A_0(r) + \mathbf{1}_{E_1<r-1}A_1(r).$$

Note that $q_1 = E_1 - r + 1$ when $E_1 \geq r - 1$, and $q_1 = E_0 - r + 2$ when $E_1 < r - 1$. With similar arguments as above we find that

$$\begin{aligned}\mathbb{E}\mathbf{1}_{E_1\geq r-1}A_0(r) &= \mathbb{E}\mathbf{1}_{E_1\geq r-1}\Big(\sum_{j=0}^{n}\sum_{\ell=1-j}^{0}\beta_j^{(n)}(x_\ell) - n\tilde{P}_1(q_1) + (n-1)E_1 - \frac{n^2}{2}\Big), \\ \mathbb{E}\mathbf{1}_{E_1<r-1}A_1(r) &= \mathbb{E}\mathbf{1}_{E_1<r-1}\Big(\sum_{j=0}^{n}\sum_{\ell=1-j}^{0}\beta_j^{(n)}(x_{\ell+1}) - n\tilde{P}_1(q_1) + nE_1 - \frac{n^2}{2}\Big).\end{aligned}$$

By the i.i.d. property of the perturbations and the exact arguments as above we conclude that

$$\begin{aligned}\mathbb{E}K_0^*(r) &= -\mathbb{E}P_1(E_0 - r) + \mathbb{E}\Big[\frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(x_0)j\Big] \\ &\quad - \frac{n}{2} + \mathbb{E}\mathbf{1}_{E_0\geq r-1}\Big[\frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(x_{-j}) - 1\Big]\end{aligned}$$

when $r \geq \frac{1}{2}$. This proves the first part of the lemma.

To show (SM2.2), we note that

$$\mathbb{E}K_m^*(r) - \mathbb{E}_u K_m(0) = -\mathbb{E}P_{m+1}(E_0 - r) \tag{SM2.3}$$

for all $r \in \mathbb{R}$. This is seen by induction using Fubini's theorem, the relations $P_m' = P_{m-1}$ and $K_m' = -K_{m-1}$, the fact that $\mathbb{E}_u K_m(0) = 0$ for all $m < n$ (see Lemma 4.1), and the continuity properties of the kernels and polynomials. For $|r| \geq 2n+2$, the perturbations in $K_m(r;X^u) = K_m^*(r-U)$ and $K_m(0;X^u) = K_m^*(-U)$ are independent. With $\mathbb{E}_U$, $\mathbb{E}_{X^*}$ and $\mathbb{E}_{E_0,E_1}$ denoting the expectations with respect to the given variables, (SM2.3) and independence then implies

$$\begin{aligned} H_m^u(r) &= \mathbb{E}_U \mathbb{E}_{X^*}[K_m^*(r-U) - \mathbb{E}_u K_m(0)] \mathbb{E}_{X^*}[K_m^*(-U) - \mathbb{E}_u K_m(0)] \\ &= \mathbb{E}_{E_0,E_1} \mathbb{E}_U [P_{m+1}(U + E_0 - r) P_{m+1}(U + E_1)] \\ &= (-1)^m \mathbb{E}[P_{2m+2}(r + E_1 - E_0)], \end{aligned} \tag{SM2.4}$$

where the last equality is shown in the proof of [2, Prop. 5.2]. This shows (SM2.2). If the model has degenerate perturbations concentrated at 0, (SM2.4) is true for all $r \in \mathbb{R}$ with $X^* = \mathbb{Z}$ deterministic. This concludes the proof. □

Corollary SM2.1. *Let $n \in \mathbb{N}$ be given. If $X^u = U + \mathbb{Z}$ is the unit-intensity equidistant model, then*

$$H_m^u(r) = (-1)^m P_{2m+2}(r)$$

for $m \leq n$ and all $r \in \mathbb{R}$.

Proof. Fix $n \in \mathbb{N}$. Note that $X^* = \mathbb{Z}$ and therefore it is deterministic. From Lemma SM2.1 it suffices to show that $K_0^*(r) = -P_1(-r)$ for $r \in [0,1)$. Also, the weights $\beta_j^{(n)}(x)$ do not depend on $x \in X^*$, and we therefore denote the common weights by $\beta_j^{(n)}$. As polynomials of degree 1 are approximated exactly, we find that

$$\frac{1}{n} \sum_{j=0}^{n} \beta_j^{(n)} = 1 \qquad \text{and} \qquad \frac{1}{n} \sum_{j=0}^{n} \beta_j^{(n)} j = \frac{n}{2}.$$

Returning to (SM2.1), we conclude that $K_0^*(r) = -P_1(-r)$. □

Corollary SM2.2. *Let $n = 1$. If X^u is from the unit-intensity perturbed model, then*

$$H_m^u(r) = (-1)^m \mathbb{E}[P_{2m+2}(r + E_1 - E_0)]$$

for $m \in \{0,1\}$ and all $|r| \geq 4$.

Proof. Since $\beta_0^{(1)}(x) = \beta_1^{(1)}(x) = \frac{1}{2} h_1(x)$, $x \in X^u$, holds for all point processes X^u, it is easily seen that $\mathbb{E}\beta_1^{(n)}(E_0) = \frac{1}{2}$ and $Q(r) = 0$. The result follows from (SM2.1). □

Acknowledgements

This work was supported by the Centre for Stochastic Geometry and Advanced Bioimaging, funded by a grant from the Villum Foundation.

References

[1] Kiderlen, M. and Dorph-Petersen, K.-A. (2017). The Cavalieri estimator with unequal section spacing revisited. *Image Anal. Stereol.* **36,** 135–141.

[2] Kiêu, K. (1997). *Three Lectures on Systematic Geometric Sampling.* Technical Report, Aarhus University.