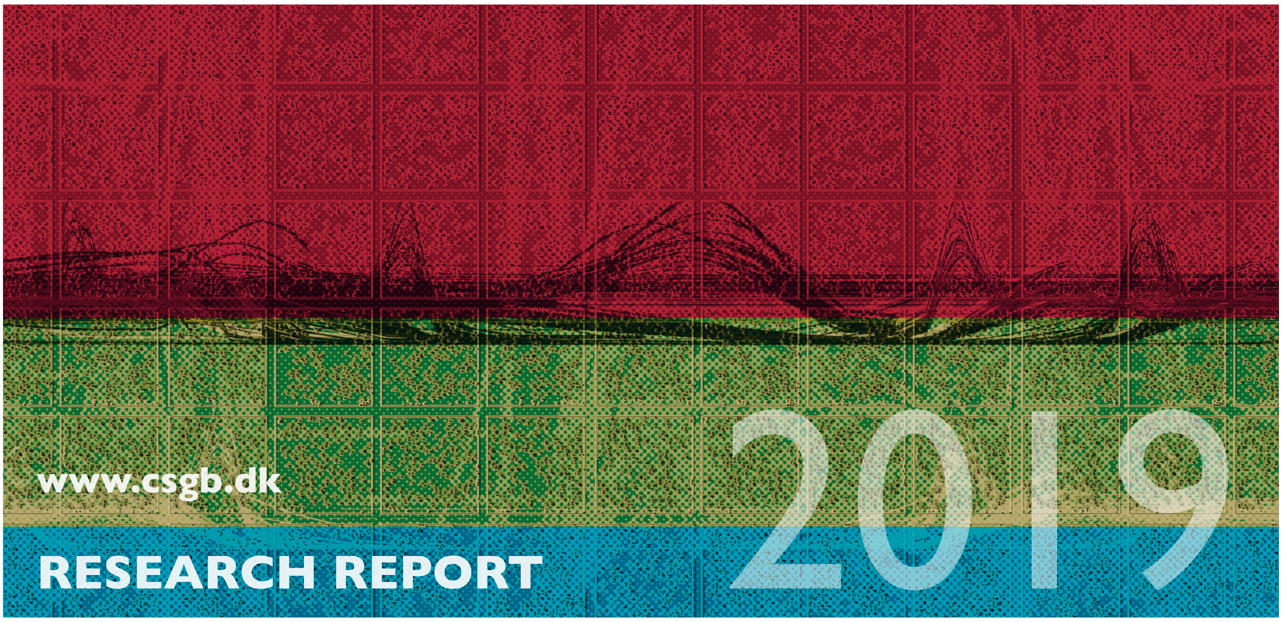
CENTRE FOR **STOCHASTIC GEOMETRY**
AND ADVANCED **BIOIMAGING**



Mads Stehr and Markus Kiderlen

Asymptotic Variance of Newton-Cotes Quadratures based on Randomized Sampling Points

No. 02, February 2019

Asymptotic Variance of Newton-Cotes Quadratures based on Randomized Sampling Points

Mads Stehr and Markus Kiderlen

Centre for Stochastic Geometry and Advanced Bioimaging (CSGB),
Department of Mathematics, Aarhus University, {mads.stehr,kiderlen}@math.au.dk

Abstract

In this paper we consider the problem of numerical integration when sampling nodes are random, and we suggest to use Newton-Cotes quadrature rules to exploit smoothness properties of the integrand. In previous papers it was shown that a Riemann sum approach can cause a severe variance inflation when the sampling points are not equidistant. However, under some integrability conditions on the typical point-distance, we show that Newton-Cotes quadratures based on a stationary point process in $\mathbb{R}$ yield unbiased estimators for the integral and that the aforementioned variance inflation can be avoided if a Newton-Cotes quadrature of sufficiently high order is applied. In a stereological application, this corresponds to the estimation of volume of a compact object from area measurements on parallel sections.

Keywords: Point processes, Cavalieri estimator, randomized Newton-Cotes quadrature, numerical integration, asymptotic variance bounds.

1 Introduction and main results

Before turning to the main subject of Newton-Cotes quadratures based on random nodes, we describe a geometric application that was the original motivation for this work. This application will also be taken up at the end of the paper to illustrate our findings in stereological examples. It is well known that, using Cavalieri's principle, the volume of a d-dimensional solid can be approximated from $(d-1)$-dimensional volume measurements on parallel hyperplanes intersecting the object with equidistant spacing. In particular, the volume of a compact object $Y \subset \mathbb{R}^3$ can be approximated from sections with equidistant and parallel planes positioned along some fixed direction $\nu \in S^2$, if the area of each intersection profile is accessible; see [2, Chap. 7]. Formally, if $f(x)$ is the area of the intersection of Y with the plane positioned at a signed distance $x \in \mathbb{R}$ from the origin along ν, the *classical Cavalieri* *estimator* approximates the integral $\int f \mathrm{d}x = \mathrm{Vol}(Y)$ of the *measurement function* $f : \mathbb{R} \to \mathbb{R}$ by a Riemann sum

$$\hat{V}(f) = t \sum_{x \in X} f(tx) \quad \text{for} \quad t > 0, \tag{1.1}$$

where $X = U + \mathbb{Z}$ is a regular standard grid in $\mathbb{R}$ shifted with $U \in \mathbb{R}$. The variable $t > 0$ rescales this grid, and we will in particular be interested in the behaviour of $\hat{V}(f)$ when t approaches 0. In design-based sampling the set of nodes is randomized by choosing U uniform in the interval $(0, 1)$. This choice of U turns X into a stationary point process (i.e. a random locally finite collection of points in $\mathbb{R}$ with a translation invariant distribution) and implies that the random variable $\hat{V}(f)$ is unbiased for $\int f \mathrm{d}x$. When the points in X are not exactly equidistant, but when the average distance between consecutive points is 1, a natural idea is to simply approximate $\int f \mathrm{d}x$ as if the points were equidistant, yielding the *generalized Cavalieri estimator*,

$$\hat{V}_0(f) = t \sum_{x \in X} f(tx), \tag{1.2}$$

which looks formally like (1.1), but allows for point processes X with non-equidistant points. The estimator (1.2) is unbiased when X is stationary. However, as remarked in [3] and [12], the variance of the generalized Cavalieri estimator may be substantially larger than the variance in the equidistant case. The purpose of the present paper is to show and quantify that this increase in variance can be reduced by using Newton-Cotes quadrature approximations of higher order instead of the crude sum (1.2).

The key problem is the statistical analysis of Newton-Cotes quadratures based on randomized nodes. We therefore treat this problem in full generality and will return to the stereological question at the end of the paper in a simulation study in section 7. Throughout the following we assume that an integrable function $f : \mathbb{R} \to \mathbb{R}$ with compact support is given, and that it can be evaluated at the points of a simple stationary point process $X \subset \mathbb{R}$; see sections 2 and 3 for details. From now on all point processes are assumed to be simple, stationary and with positive and finite intensity. An introduction to the theory of point processes can be found in [9, Chap. 3].

Realizations of X are (almost surely) locally finite collections of distinct points in $\mathbb{R}$ such that the convex hull $\mathrm{Conv}(X)$ of X is $\mathbb{R}$. We therefore recall the definition of the *n-th Newton-Cotes estimator* $\hat{V}_n(f)$, $n \in \mathbb{N}$, from [6] for a fixed realization of X. On the interval from a point $x_0 \in X$ to its n'th 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 $x_1 < \cdots < x_{n-1}$ 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 with respect to the starting point chosen. $\hat{V}_n(f)$ thus becomes a weighted average of f over all points in X, $\hat{V}_n(f) = \sum_{x \in X} \alpha(x; X) f(x)$, where the weights satisfy $\alpha(tx; tX) = t\alpha(x; X)$ for all $x \in X$ and $t > 0$; see Definition 2.1 for details. We will see in Remark 4.1 that $\alpha(x; X) = 1$ when $X = U + \mathbb{Z}$ is the translated standard grid, and therefore, Newton-Cotes estimators of any order applied to the scaled translated standard grid $tX = t(U+\mathbb{Z})$ yield the classical Cavalieri estimator (1.1), that is, $\hat{V}_n(f) = \hat{V}(f)$ for any $n \in \mathbb{N}$.

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 1.1.

$$\mathbb{E}^0 h_1^j < \infty \qquad \textit{for all } j \in \mathbb{Z}. \tag{1.3}$$

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. [9, sec. 3.3]) and h_1 is the lag between 0 and its right neighbour in X. This assumption 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 1.2. *Let $n \in \mathbb{N}$ be given and assume that X is a stationary point process such that* (1.3) *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.

In section 4 we present a weaker assumption than (1.3) which is sufficient to ensure the unbiasedness of $\hat{V}_n(f)$, see Assumption 3.1. However, the unbiasedness is known to hold for $n = 1$ without integrability conditions and for $n = 2$ under a weaker condition than the one presented in (1.3); [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 $C^{m-1}(\mathbb{R})$ (only for $m \geq 1$) 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 only 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 m'th 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. The classical Euler-McLaurin formulae comparing the integral of a function with a Riemann sum approximation are insufficient when assessing the error of $\hat{V}_n(f)$, as the sampling points need not be equidistant. However, applying a refined partial integration formula for piecewise smooth functions yields a refined version of the classical Peano error representation theorem ([11, Theorem 3.2.3]) adopted to such functions. This results expresses the discretization error

$$R^{(n)}(f) = \hat{V}_n(f) - \int_{\mathbb{R}} f(x)\mathrm{d}x \tag{1.4}$$

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 $\mathrm{Conv}(X) = \mathbb{R}$.

Theorem 1.3 (Refined Peano error representation 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*

$$R^{(n)}(f) = \int_{\mathbb{R}} f^{(m+1)}(r)K_m(r)\mathrm{d}r + \sum_{a \in D_{f^{(m)}}} J_{f^{(m)}}(a)K_m(a) \tag{1.5}$$

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

Remark 1.4. The function K_m will be called the *m'th Peano kernel.* It is a piecewise polynomial of order at most $m+1$ with coefficients given in terms of X. The m'th Peano kernel is explicitly given by (2.7), below. It is shown in Lemma 3.2 that for a stationary point process X satisfying (1.3), K_m is a stationary stochastic process on the real line with positive moments of all orders. 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 $r \in \mathbb{R}$.

When applying the classical Cavalieri estimator with $tX = t(U + \mathbb{Z})$, exact variance expressions have been found in e.g. [2, Section 13.2]. Let $*$ denote the convolution operator and let the reflection $\check{f}$ 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* $g = f * \check{f}$ is $(2m+1,1)$-piecewise smooth. When f is exactly $(m,1)$-piecewise smooth, one usually decomposes the variance of $\hat{V}(f)$ as

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

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*

$$\mathbb{V}\mathrm{ar}_E(\hat{V}(f)) = t^{2m+2}g^{(2m+1)}(0^+)c_m \tag{1.6}$$

explains the overall trend of the variance. Here $c_m = -\frac{2B_{2m+2}}{(2m+2)!} \neq 0$, where B_k is the k'th Bernoulli number (see section 5 below), and as such c_m does not depend on t or the function f, other than through its order of smoothness. When applying the generalized Cavalieri estimator to the following two specific non-equidistant point models, which are very realistic from a practical point of view, it is shown in [12] that a variance inflation is present.

Example 1.5 (Perturbed model). A stationary point process is from the perturbed model if it is derived from equidistant points by having i.i.d. perturbations E_i, $i \in \mathbb{Z}$, of every point; see subsection 6.1. From a stereological perspective this model is reasonable when cutting an object into slabs using a fixed set of blades which are equidistant but due to e.g. softness of material exhibit some variation from exact equidistance.

Example 1.6 (Model with cumulative errors). A stationary point process X is from the model with cumulative errors if the intercepts ω_i, $i \in \mathbb{Z}$, between two consecutive points of X form an i.i.d. sequence; see subsection 6.2. This model is reasonable when the cutting of an object into slabs works like a meat slicer.

The classical stationary point grid $U + \mathbb{Z}$ is a special case of both the perturbed model and the model with cumulative errors; let $E_1 \sim \delta_0$ and $\omega_1 \sim \delta_1$, respectively, with δ being the Dirac measure. If X is from the perturbed model (with not necessarily vanishing perturbations E_i), the generalized Cavalieri estimator satisfies $\mathbb{V}\mathrm{ar}(\hat{V}_0(f)) = t^2 c' + Z_0(t) + o(t^2)$ when $m = 0$ and $\mathbb{V}\mathrm{ar}(\hat{V}_0(f)) = t^3 c'' + o(t^3)$ when $m \geq 1$ as $t \downarrow 0$. This was shown in [12, 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 latter term in [12, Eq. (A3)] is of order $o(t^{2m+2})$. Hence, the variance of the generalized Cavalieri estimator has a slower rate of decrease than the classical estimator for all $m \geq 1$. 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 [12, Prop. 2]. In contrast to this, the main result of the present paper is the following: Using a Newton-Cotes approximation of at least the same order as the order of smoothness of the measurement function, the variance decreases at the same rate as the classical estimator as long as the randomized sampling points are stationary and (1.3) is satisfied. When the order of smoothness of the measurement function exceeds the order of the estimator, a faster decrease of the variance is obtained in the case where the point process is strongly admissible; a property satisfied for the perturbed model and a fortiori for the equidistant model, and for the cumulative model under certain exponential moment assumptions; see Lemmas 6.1 and 6.3, respectively.

Definition 1.7 (Admissible point process)**.** *Let X be a stationary point process satisfying* (1.3)*. Let $n \in \mathbb{N}$ be given and let H_n be the covariance function of K_n. Then X is called strongly n-admissible if $\int_0^t H_n(s)\mathrm{d}s$ is uniformly bounded in $t \geq 0$. X is called weakly n-admissible if $\lim_{t\to\infty} \frac{1}{t}\int_0^t H_n(s)\mathrm{d}s = 0$.*

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 $n \in \mathbb{N}$.

Theorem 1.8. *Let $n \in \mathbb{N}$ be given and assume that X is a stationary point process such that* (1.3) *holds. If $t > 0$ and f is $(m, 1)$-piecewise smooth, then*

$$\hat{V}_{n,t}(f) = \sum_{x \in tX} \alpha(x; tX) f(x)$$

is unbiased for $\int f \mathrm{d}x$. If $m \leq n$ its variance obeys

$$\mathbb{V}\mathrm{ar}(\hat{V}_{n,t}(f)) \leq ct^{2m+2} \tag{1.7}$$

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

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

$$\mathbb{V}\mathrm{ar}(\hat{V}_{n,t}(f)) \leq c' t^{2n+3} \tag{1.8}$$

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 (1.7) is optimal. This is also true in the case $m = n$ if X is weakly n-admissible; see Remark 5.2.

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 cumulative model. In the perturbed case, the rate of decrease of the upper bound in (1.8) is optimal if the perturbations E_i are non-degenerate.

Theorem 1.9. *Let X be from the perturbed model and assume that the measurement function f is exactly $(m,1)$-piecewise smooth. Let $g = f * \check{f}$ be its covariogram and let μ_k be the k-th moment of the perturbations E_i. Then, for $t \downarrow 0$,*

$$\mathbb{V}\mathrm{ar}(\hat{V}_{1,t}(f)) = -t^2 g'(0^+)(\mu_2 + \tfrac{1}{6}) + Z_0(t) + o(t^2), \qquad \textit{for } m = 0, \tag{1.9}$$

$$\mathbb{V}\mathrm{ar}(\hat{V}_{1,t}(f)) = t^4 g^{(3)}(0^+)\tfrac{1}{12}(2\mu_2 + 2\mu_4 + \tfrac{1}{30}) + Z_1(t) + o(t^4), \quad \textit{for } m = 1, \tag{1.10}$$

$$\mathbb{V}\mathrm{ar}(\hat{V}_{1,t}(f)) = t^5 g^{(4)}(0)\tfrac{1}{8}(2\mu_4 + \mu_2\mu_4 - \mu_2^3 - \mu_3^2) + o(t^5), \qquad \textit{for } m \geq 2, \tag{1.11}$$

where the Zitterbewegung $Z_m(t)$ is given by (6.7). It is of order t^{2m+2}, and it is a finite sum of terms oscillating around 0. Moreover, if E_i has density with a finite number of finite jumps, for $m \geq 2$ the remainder $o(t^5)$ is explicitly given by

$$\begin{aligned} t^6 g^{(5)}(0^+)\tfrac{1}{720}\big(&-34\mu_2 - 90\mu_2^2 + 110\mu_4 + 180\mu_2\mu_4 \\ &- 180\mu_2^3 - 170\mu_3^2 + 8\mu_6 - \tfrac{1}{21}\big) + Z_2(t) + o(t^6). \end{aligned} \tag{1.12}$$

It is worth noticing that the Zitterbewegung appearing in the variance decomposition of the classical Cavalieri estimator is not present in the decomposition of Theorem 1.9 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 of the classical Cavalieri estimator is $\mathbb{V}\mathrm{ar}_E \hat{V}(f) = -t^2 g'(0^+)\frac{1}{6}$, $\mathbb{V}\mathrm{ar}_E \hat{V}(f) = t^4 g^{(3)}(0^+)\frac{1}{12}\frac{1}{30}$ 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 dominating behaviour of the trapezoidal estimator with perturbed sampling can come arbitrarily close to the dominating behaviour of the classical estimator if the errors E_i are sufficiently small. Similarly, the dominating behaviour with cumulative sampling can come arbitrarily close to that of the classical estimator if the increments are sufficiently close to 1.

Theorem 1.10. *Let X be from the model with cumulative errors and assume that $\mathbb{E}e^{\eta\omega_1} < \infty$ for some $\eta > 0$. Let the measurement function f be exactly $(m,1)$-piecewise smooth with covariogram $g = f * \check{f}$, and let ν_k denote the k'th moment of the increments ω_i. Then, for $t \downarrow 0$,*

$$\mathbb{V}\mathrm{ar}(\hat{V}_{1,t}(f)) = -t^2 g'(0^+)\tfrac{1}{6}\nu_3 + o(t^2), \qquad \textit{for } m = 0, \tag{1.13}$$

$$\mathbb{V}\mathrm{ar}(\hat{V}_{1,t}(f)) = t^4 g^{(3)}(0^+)\tfrac{1}{12}\tfrac{1}{30}(6\nu_5 - 5\nu_3^2) + o(t^4), \qquad \textit{for } m = 1. \tag{1.14}$$

The paper is organized as follows. In section 2 relevant notation is introduced and the n'th order Newton-Cotes estimator is formally derived. Moreover, our second main result, Theorem 1.3, is proved. In section 3 we derive integrability statements which will be of relevance when proving the main results, Theorems 1.2 and 1.8, in sections 4 and 5, respectively. In section 6 we prove that point processes from the perturbed and cumulative model are admissible, and we derive the exact variance expressions presented in Theorems 1.9 and 1.10. Section 7 contains a simulation study of volume estimation in $\mathbb{R}^3$ from 2-dimensional area measurements illustrating the results of the paper. Conclusions and ideas for future work are found in section 8.

2 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 j'th 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

$$s_j(x + t; X + t) = s_j(x; X) + t \tag{2.1}$$

for all $t \in \mathbb{R}$. For all $x \in X$ and $j \in \mathbb{Z}$ we let $h_j(x) = h_j(x; X) \coloneqq s_j(x; X) - s_{j-1}(x; X)$ be the distance from the j'th successor (predecessor) of x to its left neighbour in X. By (2.1),

$$h_j(x + t; X + t) = h_j(x; X) \tag{2.2}$$

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

$$\beta_j^{(n)}(x) = \beta_j^{(n)}(x; X) = \int_x^{s_n(x)} \prod_{\substack{k=0 \\ k \neq j}}^{n} \frac{y - s_k(x)}{s_j(x) - s_k(x)} \mathrm{d}y \tag{2.3}$$

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

Definition 2.1 (n'th Newton-Cotes approximation). *Let $n \in \mathbb{N}$ be given. The n'th order Newton-Cotes approximation of $f : \mathbb{R} \to \mathbb{R}$ with nodes in X is given by*

$$\hat{V}_n(f) = \sum_{x \in X} \alpha(x) f(x), \tag{2.4}$$

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

Remark 2.2. From [11, 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

$$R_x^{(n)}(f) = R_x^{(n)}(f; X) = I_x^{(n)}(f) - \int_x^{s_n(x)} f(y)\mathrm{d}y, \tag{2.5}$$

$x \in X$.

As shown in Lemma A.1, $\beta_j^{(n)}$ is a rational function of point-increments, and (2.2) then implies that

$$\beta_j^{(n)}(x + t; X + t) = \beta_j^{(n)}(x; X) \quad \text{and} \quad \alpha(x + t; X + t) = \alpha(x; X) \tag{2.6}$$

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

We are now ready to prove the Peano error representation as stated in Theorem 1.3. Given n, X, and $m \in \mathbb{N}_0$, the m'th Peano kernel from Theorem 1.3 is defined as

$$K_m(r) = K_m(r; X) = \frac{1}{m!n} \sum_{x \in X} \mathbf{1}_{(x, s_n(x)]}(r) R_x^{(n)}\big((\,\cdot\, - r)_+^m\big). \tag{2.7}$$

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

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

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

Proof of Theorem 1.3. Fix $n \in \mathbb{N}$ and note that $nR^{(n)}(f) = \sum_{x \in X} R_x^{(n)}(f)$ due to (1.4) and (2.5). 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

$$\begin{aligned} f(y^-) = {} & \sum_{k=0}^{m} \frac{f^{(k)}(x^+)}{k!}(y - x)^k \\ & + \frac{1}{m!} \sum_{a \in D_{f^{(m)}} \cap (x,y)} J_{f^{(m)}}(a)(y - a)^m + \frac{1}{m!} \int_x^y f^{(m+1)}(t)(y - t)^m \mathrm{d}t, \end{aligned} \tag{2.8}$$

for all $(m, 1)$-piecewise smooth functions f, $m \in \mathbb{N}_0$. We now assume $m \leq n$. 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, (2.7) 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

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

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

Lemma 2.3. *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*

$$J_{K_m}(x) = \frac{1}{m!n} R_x^{(n)}\big((\,\cdot\, - x)^m\big). \tag{2.9}$$

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

3 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 $j = 0, \dots, n$. In addition we write $\mathbf{h} = (h_1, \dots, h_n)$ and for $i \in \mathbb{Z}$, $\mathbf{h}(s_i) = (h_1(s_i), \dots, h_n(s_i)) = (h_{i+1}, \dots, h_{i+n})$ under $\mathbb{P}^0$. As mentioned in section 1 a weaker assumption than (1.3) is sufficient for the estimator to be unbiased.

Assumption 3.1. *For a given $n \in \mathbb{N}$ we assume that*

$$\mathbb{E}^0\left[\frac{\mathbf{h}^{\mathbf{m}}}{\mathbf{h}^{\mathbf{m}'}}\right] < \infty \tag{3.1}$$

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 $|\mathbf{m}| = |(m_1, \dots, m_n)| = \sum_{k=1}^n m_k$.

Using Hölder's inequality and [6, Eq. (13)], one shows that Assumption 1.1 is stronger than Assumption 3.1. In Lemma A.1 of the supplementary material it is shown that the weight $\beta_j^{(n)}(x)$ is a rational function of the point-increments $(h_1(x), \dots, h_n(x))$, $x \in X$, 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 3.1 is satisfied, and consequently

$$\mathbb{E}^0|\alpha(0)| < \infty, \tag{3.2}$$

see Lemma A.2. 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 1.2.

To argue for the variance bounds presented in Theorem 1.8 we need higher-order moment and translation invariance properties of the Peano kernel K_m defined in (2.7).

Lemma 3.2. *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 (1.3) holds, $K_m(0)$ has finite positive moments of all orders. Moreover, if X has intensity γ, K_m satisfies*

$$\mathbb{E}K_m(0) = \mathbb{E}^0 J_{K_{m+1}}(0) = \frac{\gamma}{(m+1)!n}\mathbb{E}^0 R_0^{(n)}\left((\,\cdot\,)^{m+1}\right) \tag{3.3}$$

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 $r, t \in \mathbb{R}$ and any locally finite pointset X, the Peano kernel satisfies

$$K_m(r+t; X) = K_m(r; X-t). \tag{3.4}$$

This follows from the definition of K_m and

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

which in turn is a consequence of (2.1) and (2.6). For any $N \in \mathbb{N}$ and $r_1 < r_2 < \cdots < r_N$ we see from (3.4) that

$$\big(K_m(r_1+t; X), \dots, K_m(r_N+t; X)\big) = \big(K_m(r_1; X-t), \dots, K_m(r_N; X-t)\big),$$

and the stationarity of the stochastic process K_m is therefore inherited from the stationarity of the point process X.

We now prove that $K_m(0)$ has finite positive moments of all orders. 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 (2.5) and (2.7)

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

By the refined Campbell Theorem [9, Theorem 3.5.3], (2.1) and (2.6) it follows that

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

where γ is the intensity of X. By Lemma A.1, Assumption 1.1 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 positive moments of all orders under $\mathbb{P}^0$. This implies that $\mathbb{E}K_m^k(0) < \infty$.

Equation (3.3) is a simple consequence of the refined Campbell Theorem [9, Theorem 3.5.3], Lemma 2.3 and [6, Eq. (13)]. □

4 Unbiasedness of Newton-Cotes estimators

Proof of Theorem 1.2. Fix $n \in \mathbb{N}$ and let X be a simple stationary point process with finite and positive intensity γ. As α satisfies (2.6) and $\alpha(0)$ is Palm-integrable by (3.2), [6, Theorem 1] can be applied. It states that

$$\mathbb{E}\hat{V}_n(f) = \gamma\mathbb{E}^0[\alpha(0)] \int_{\mathbb{R}} f(x)\mathrm{d}x \tag{4.1}$$

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 $t \in \mathbb{R}$ reuse the notation I_t from the end of the previous section. When f is an integrable function and $|f| \le 1$, (2.5) implies

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

The refined Campbell theorem [9, Theorem 3.5.3], (2.1) and (2.6) imply

$$\begin{aligned}\mathbb{E} \sum_{x\in I_t} |R_x^{(n)}(f)| &\le \gamma\mathbb{E}^0 \int_{t-s_n}^t \Big(\sum_{j=0}^n |\beta_j^{(n)}| + s_n \Big)\mathrm{d}x \\ &= \gamma\mathbb{E}^0\Big[s_n \sum_{j=0}^n |\beta_j^{(n)}| + s_n^2 \Big] < \infty,\end{aligned}$$

where the finiteness follows from Lemma A.1 and Assumption 3.1, which is weaker than Assumption 1.1. Note that the finite upper bound does not depend on t.

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 n'th 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 (4.1) 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 4.1. If $X = U + \mathbb{Z}$ is the translated standard grid, $\alpha(x) = 1$ for all $x \in X$. In fact, the Palm version of X is the deterministic set $\mathbb{Z}$, so $h_1(x) = 1$, and (2.6) yields

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

for all $x \in X$. Hence $\alpha = \alpha(x)$ is deterministic, and $\hat{V}_n(f) = \alpha \sum_{x \in X} f(x)$. Assumption 1.1 is trivially satisfied, so Theorem 1.2 implies that $\hat{V}_n(f)$ is unbiased for $\int f \mathrm{d}x$, which is equivalent to $\alpha = 1$.

5 Asymptotic variance behaviour of Newton-Cotes estimators

Before proving Theorem 1.8, we recall the variance decomposition of the classical Cavalieri estimator, as it shows great resemblance to the new non-equidistant setup. 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 m'th 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 m'th 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]

$$\mathbb{V}\mathrm{ar}(\hat{V}(f)) = -t^{2m+2} \sum_{a \in D_{g^{(2m+1)}}} J_{g^{(2m+1)}}(a) P_{2m+2}(\tfrac{a}{t}) + o(t^{2m+2}) \tag{5.1}$$

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 5.1. *Let $n \in \mathbb{N}$ be given and assume that X is a stationary point process such that* (1.3) *holds. If f is $(m,1)$-piecewise smooth with $m \leq n$, then*

$$\begin{aligned}(-1)^{m+1} \mathbb{V}\mathrm{ar}(\hat{V}_{n,t}(f)) = t^{2m+2} &\sum_{a \in D_{g^{(2m+1)}}} J_{g^{(2m+1)}}(a) H_m(\tfrac{a}{t}) \\ &+ t^{2m+2} \int_{\mathbb{R}} g^{(2m+2)}(s) H_m(\tfrac{s}{t}) \mathrm{d}s,\end{aligned} \tag{5.2}$$

where $g = f * \check{f}$ is the covariogram of f. If $m < n$ or X is weakly n-admissible, the variance behaviour is determined by the first term, as

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

for $t \downarrow 0$.

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

$$\alpha(tx; tX) = t\alpha(x; X) = t\alpha(x)$$

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

$$\hat{V}_{n,t}(f) = t\hat{V}_n(f_t). \tag{5.3}$$

As $m \le n$, Theorem 1.3 implies

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

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

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

Hence, using (5.3) and the unbiasedness of $\hat{V}_n$, we get

$$\begin{aligned} \mathbb{V}\mathrm{ar}(\hat{V}_{n,t}(f)) &= \mathbb{V}\mathrm{ar}(t\hat{V}_n(f_t)) = \mathbb{E}(tR^{(n)}(f_t))^2 \\ &= t^{2m+2} \mathbb{E}\Big(\int_{\mathbb{R}} f^{(m+1)}(s) K_m(\tfrac{s}{t}) \mathrm{d}s + \sum_{a \in D_{f^{(m)}}} J_{f^{(m)}}(a) K_m(\tfrac{a}{t}) \Big)^2. \end{aligned} \tag{5.4}$$

An application of [7, Prop. 5.7] yields

$$\begin{aligned} f^{(m+1)} * \check{f^{(m+1)}}(x) &= (-1)^{m+1} g^{(2m+2)}(x) \\ &- \sum_{a \in D_{f^{(m)}}} J_{f^{(m)}}(a) f^{(m+1)}(a - x) - \sum_{a \in D_{f^{(m)}}} J_{f^{(m)}}(a) f^{(m+1)}(a + x), \end{aligned} \tag{5.5}$$

and furthermore the jumps of $g^{(2m+1)}$ are given by $J_{g^{(2m+1)}} = (-1)^{m+1} J_{f^{(m)}} * \check{J}_{f^{(m)}}$, see [7, Eq. (5.12)]. The stationarity and square integrability of K_m from Lemma 3.2 implies that $\mathbb{E}K_m(r)$ and $H_m(s) = \mathrm{Cov}(K_m(r), K_m(s + r))$ are both finite and independent of $r \in \mathbb{R}$. Equation (5.2) now follows by expanding (5.4) and applying (5.5), the structure of $J_{g^{(2m+1)}}$ together with Fubini's theorem which is justified by the square integrability of K_m and the fact that $f^{(m+1)}$ is bounded with compact support.

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

$$\lim_{t \downarrow 0} \int_0^1 H_m(\tfrac{s}{t}) \mathrm{d}s = 0 \tag{5.6}$$

for $m = n$. Equation (5.6) also holds for $m < n$ without additional assumptions. In fact, for $m < n$ we have $K_{m+1}' = -K_m$ by Lemma 2.3 and thus using Fubini's Theorem

$$\left| \int_0^t H_m(s) \mathrm{d}s \right| = \big| \mathrm{Cov}\big(K_m(0), K_{m+1}(0)\big) - \mathrm{Cov}\big(K_m(0), K_{m+1}(t)\big) \big| \le c < \infty \ ,$$

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 (5.6) from this.

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

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

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

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

As ϕ is simple, (5.6) implies that $\lim_{t\downarrow 0}\int_{\mathbb{R}} \phi(s)H_m(\frac{s}{t})\mathrm{d}s = 0$. We conclude that $|\int_{\mathbb{R}} \phi(s)H_m(\frac{s}{t})\mathrm{d}s| < \frac{\epsilon}{2}$ for sufficiently small $t > 0$, and hence

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

for such small $t > 0$. □

Proof of Theorem 1.8. As (1.3) is satisfied for tX, the unbiasedness follows from Theorem 1.2. From Lemma 3.2 there exists $C < \infty$ such that $\sup_{s\in\mathbb{R}} |H_m(s)| \leq C$, and we immediately see from (5.2) that

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

As g is $(2m+1,1)$-piecewise smooth from [7, Corollary 5.8], the t-independent constant is finite, and (1.7) therefore follows.

For the stronger result (1.8), 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 5.1 to the $(n,1)$-piecewise smooth function f and a substitution gives

$$(-1)^{n+1}\mathbb{V}\mathrm{ar}(\hat{V}_{n,t}(f)) = t^{2n+3}\int_{\mathbb{R}} g^{(2n+2)}(st)H_n(s)\mathrm{d}s. \tag{5.7}$$

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

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

where we used that $g^{(2n+2)}$ vanishes at $\pm b$. Returning to (5.7) we find

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

This proves the assertion. □

Remark 5.2. If (5.6) is satisfied, the variance

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

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

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

as $t \to 0$, and using that $g^{(2m+1)}$ is discontinuous and odd, hence, it has a jump at 0. Applying (5.6) yields a contradiction. In particular, the decrease rate in (1.7) is optimal if $m < n$ or X is weakly n-admissible.

6 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 cumulative model introduced in Examples 1.5 and 1.6, respectively.

6.1 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 $x_i = U + i + E_i$, for all $i \in \mathbb{Z}$. Note that it has intensity 1. Under its Palm distribution, we have

$$h_k = 1 + E_k - E_{k-1} \leq 2, \tag{6.1}$$

$k \in \mathbb{Z}$, so (1.3) is equivalent to

$$\mathbb{E}(1 + E_1 - E_0)^{-j} < \infty \tag{6.2}$$

for all $j \in \mathbb{N}$. For instance, (6.2) holds if there is $\epsilon > 0$ such that $|E_0| \leq \frac{1}{2} - \epsilon$ almost surely. For the perturbed model X we define the shifted kernel K_m^* by

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

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

$$K_m^*(r) \overset{\mathcal{D}}{=} K_m^*(r+k) \tag{6.3}$$

for all $k \in \mathbb{Z}$. This can be used to show that X is strongly admissible.

Lemma 6.1. *Let $n \in \mathbb{N}$ be given and assume that X is a stationary point process from the perturbed model such that (6.2) holds. Then, for all $m \in \mathbb{N}_0$ and $r \geq 2n+2$, $H_m(r) = H_m(r+1)$ and $\int_r^{r+1} H_m(s)\mathrm{d}s = 0$. In particular, X is strongly n-admissible.*

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

$$\begin{aligned} \mathbb{E}[K_m(0)K_m(r)] &= \int_0^1 \mathbb{E}[K_m^*(-u)K_m^*(r-u)]\mathrm{d}u \\ &= \int_0^1 \mathbb{E}[K_m^*(-u)]\mathbb{E}[K_m^*(r-u)]\mathrm{d}u. \end{aligned} \tag{6.4}$$

Equations (6.3) and (6.4) imply that $\mathbb{E}[K_m(0)K_m(r)] = \mathbb{E}[K_m(0)K_m(r+1)]$ and by stationarity of K_m we conclude that that $H_m(r) = H_m(r+1)$.

Returning to (6.4), we find by Fubini's theorem, a substitution and the stationarity of K_m that

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

which yields the asserted properties of H_m. This clearly implies that X is strongly n-admissible. □

In order to obtain explicit leading terms in Theorem 1.9 we state in the following a connection between the covariance function H_m 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 6.2. *Let $n = 1$ and let X be from the perturbed model. Then*

$$H_m(r) = (-1)^m \mathbb{E}[P_{2m+2}(r + E_1 - E_0)] \tag{6.5}$$

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

If $X = U + \mathbb{Z}$ is the shifted standard grid, then

$$H_m(r) = (-1)^m P_{2m+2}(r) \tag{6.6}$$

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

The proof of Lemma 6.2 can be found in the supplementary material of this paper; see Corollaries B.2 and B.3. As a consequence of this result we see that, for the equidistant model $X = U + \mathbb{Z}$, the variance representation (5.2) found using the Peano kernels is the same as the classical variance representation (5.1) found using Euler-McLaurin formulae.

Before turning to the proof of Theorem 1.9, we emphasize that the integrability condition (1.3), or equivalently, condition (6.2), was omitted in the statement of the Theorem as we work with the trapezoidal rule. In fact, the unbiasedness of $\hat{V}_{1,t}(f)$ for all stationary point processes and integrable, compactly supported functions f was already remarked in the paragraph following the statement of Theorem 1.2. Due to (6.1), the weights satisfy

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

$x \in X$, which replaces condition (1.3) in all the arguments in sections 4 and 5. The assumptions of Proposition 5.1 are thus satisfied.

Proof of Theorem 1.9. 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 5.1 in combination with (6.5) yields the variance decomposition

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

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

$$Z_m(t) = -t^{2m+2} \sum_{a \in D_{g^{(2m+1)}} \setminus \{0\}} J_{g^{(2m+1)}}(a)\mathbb{E}[P_{2m+2}(\tfrac{a}{t} + E_1 - E_0)]. \tag{6.7}$$

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 [9, Theorem 3.5.3] and the facts that $\mathbb{E}K_0(0) = 0$ and $\mathbb{E}K_1(0) = \frac{1}{2}\mathbb{E}^0[R_0^{(1)}((\,\cdot\,)^2)]$ by (3.3), we find that $H_m(0) = \mathbb{V}\mathrm{ar}(K_m(0))$ satisfies

$$H_0(0) = \mathbb{E}^0 \int_0^{h_1} (\tfrac{1}{2}h_1 - y)^2 \mathrm{d}y = \tfrac{1}{12}\mathbb{E}^0 h_1^3, \tag{6.8}$$

$$H_1(0) = \mathbb{E}^0 \int_0^{h_1} (\tfrac{1}{2}h_1 y - \tfrac{1}{2}y^2)^2 \mathrm{d}y - (\tfrac{1}{12}\mathbb{E}^0 h_1^3)^2 = \tfrac{1}{120}\mathbb{E}^0 h_1^5 - \tfrac{1}{144}(\mathbb{E}^0 h_1^3)^2. \tag{6.9}$$

Using (6.1), it is elementary to conclude (1.9) and (1.10).

Now let $m \geq 2$ be given and define $\tilde{H}_1$ by $\tilde{H}_1(s) = H_1(s) + \mathbb{E}[P_4(s + E_1 - E_0)]$. Due to Lemma 6.2, $\tilde{H}_1(s)$ vanishes for $|r| > 4$. Since $g^{(4)}$ is continuous, an application of Proposition 5.1 to the $(1, 1)$-piecewise smooth function f, Fubini's theorem and the refined partial integration formula [7, Lemma 4.1] yield

$$\begin{aligned}\mathbb{V}\mathrm{ar}(\hat{V}_{1,t}(f)) = t^5 \int_{\mathbb{R}} g^{(4)}(st)\tilde{H}_1(s)\mathrm{d}s - t^6 \int_{\mathbb{R}} g^{(6)}(s)\mathbb{E}[P_6(\tfrac{s}{t} + E_1 - E_0)]\mathrm{d}s \\ - t^6 \sum_{a \in D_{g^{(5)}}} J_{g^{(5)}}(a)\mathbb{E}[P_6(\tfrac{a}{t} + E_1 - E_0)].\end{aligned} \tag{6.10}$$

As the last two terms in (6.10) 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 (6.10), and using the fact that $g^{(5)}$ and $\tilde{H}_1$ are bounded, yields

$$\mathbb{V}\mathrm{ar}(\hat{V}_{1,t}(f)) = t^5 g^{(4)}(0) \int_{-4}^{4} \tilde{H}_1(s)\mathrm{d}s + o(t^5) \tag{6.11}$$

as $t \downarrow 0$.

Noting that P_4 integrates to 0 on each interval of unit length, we find that

$$\int_{-4}^{4} \tilde{H}_1(s)\mathrm{d}s = \int_{-4}^{4} H_1(s)\mathrm{d}s = \tfrac{1}{8}(2\mu_4 + \mu_2\mu_4 - \mu_2^3 - \mu_3^2),$$

where the last equality follows from rather technical arguments exploiting the independence and identical law of the perturbations; see Appendix B. Inserting this into (6.11) yields the assertion (1.11).

The expression (1.12) of the remainder are found by different arguments which will occur in the upcoming thesis [10]. □

6.2 Cumulative sampling

Before turning to the proof of Theorem 1.10, we state in Lemma 6.3 below that the covariance function of the Peano kernel decreases exponentially, from which admissibility follows. The proof of Lemma 6.3 can be found in the supplementary material of this paper; see Lemma C.2.

The cumulative process X is a stationary point process with i.i.d. increments $\{\omega_i\}_{i\in\mathbb{Z}}$ and with intensity 1. We assume that ω_1 has cumulative distribution function F with density wrt. Lebesgue measure and mean 1 such that $F(0) = 0$. To explicitly construct the point process, the first point X_0 of $X \cap (0, \infty)$ is chosen with cumulative distribution function G,

$$G(x) = \int_0^x (1 - F(y))\mathrm{d}y = \int_0^x \bar{F}(y)\mathrm{d}y, \qquad x \geq 0,$$

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 [12], 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 $X = \{x_i\}_{i\in\mathbb{Z}}$ of the cumulative point process. This construction implies that the point interval containing the origin has the length weighted distribution, as expected.

Lemma 6.3. *Let $n \in \mathbb{N}$ be given, and let X be from the cumulative model such that $\mathbb{E}\mathrm{e}^{\eta\omega_1} < \infty$ for some $\eta > 0$, and such that* (1.3) *is satisfied. Then*

$$H_n(s) = O(\mathrm{e}^{-\epsilon s}), \qquad s \to \infty, \tag{6.12}$$

for some $\epsilon > 0$. In particular, X is strongly n-admissible.

As for the perturbed model, Theorem 1.10 is stated without the integrability assumption (1.3). This is because (1.3) can be relaxed to only be true for $j \in \mathbb{N}$ when using the trapezoidal rule, and finite positive moments are ensured by the exponential moment assumption of the increments. Also, the decrease rate in Lemma 6.3 can be obtained without assuming (1.3) for $n = 1$; see Lemma C.1.

Proof of Theorem 1.10. Let $m \in \{0, 1\}$. As X is strongly admissible, Proposition 5.1 in combination with the decrease rate (6.12) yields the variance decomposition

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

and the result follows using the fact that $\mathbb{E}^0 h_1^j = \nu_j$ in combination with (6.8) and (6.9). $\square$

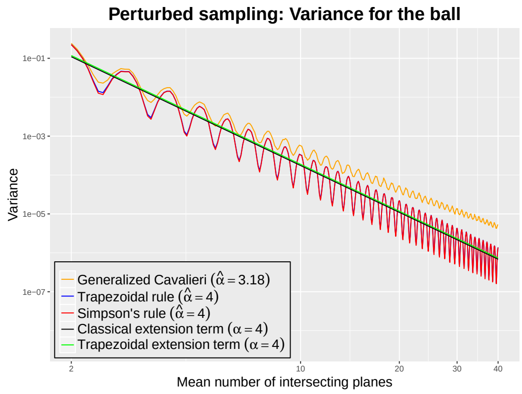
7 A simulated application in stereology

In this section we present Monte Carlo simulations illustrating the results of the paper. If we wish to estimate the volume of the unit ball $B^3 \subset \mathbb{R}^3$ from intersections with 2-dimensional planes, the measurement function (hence area function) becomes

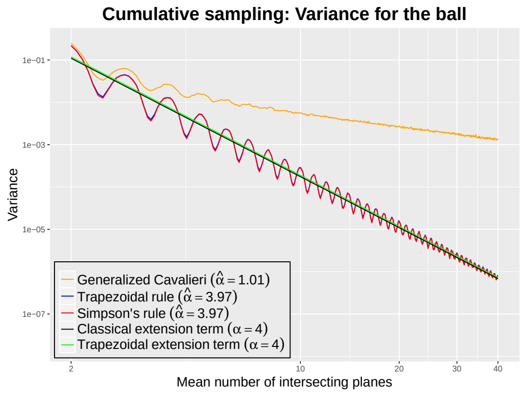
$$f(x) = \mathbf{1}_{[-1,1]}(x)\pi(1 - x^2),$$

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 (1.6). 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 (1.10) and (1.14) 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 $15 \leq 2/t \leq 40$).

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



(a) Perturbed sampling



(b) Cumulative sampling

Figure 1: Empirical variance for the volume estimation of the unit ball B^3 in $\mathbb{R}^3$ based on perturbed sampling with perturbations $E_1 \sim \mathrm{Unif}((-s, s))$ and sampling with cumulative errors with increments $\omega_1 \sim \mathrm{Unif}((1-c, 1+c))$. 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 (blue) is almost completely hidden by the graph of Simpson's estimator (red), and the trapezoidal extension term (green) is almost identical to the classical extension term (black).

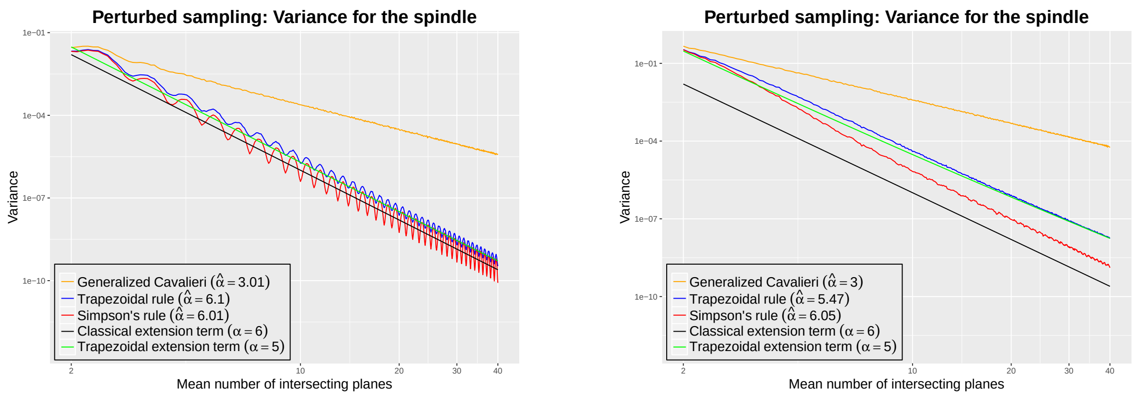
variance behaviour apart from the facts that intercepts of these curves may be shifted and the Zitterbewegung may differ.

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 $\mathbb{V}\mathrm{ar}_E(\hat{V}(f)) = \frac{\pi^6}{60480}t^6$. 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 (1.11) and (1.12). In Figure 2a we use small perturbations to illustrates the fact that the dominating term in (1.11) 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 1.8, 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 $\hat{\alpha} = 5.19$.

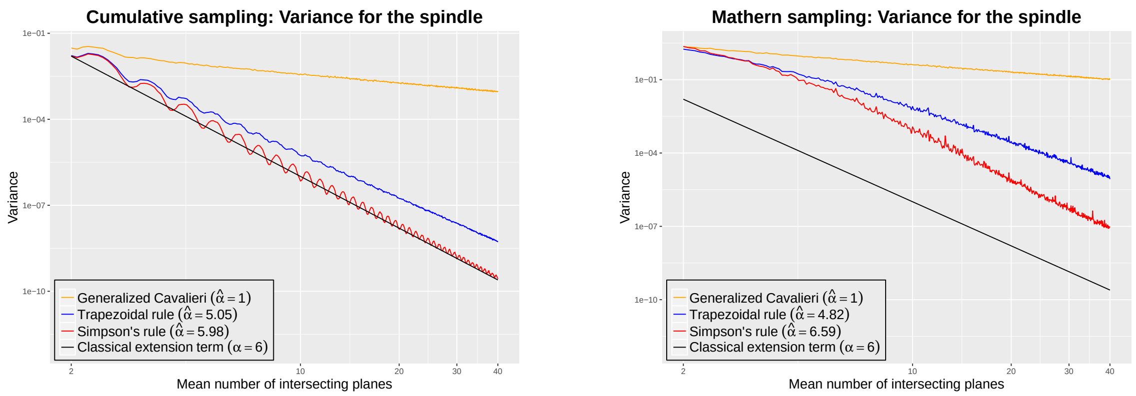
The last two simulations are meant to illustrate the findings in Theorem 1.8 for point process models where we do not have explicit formulae for the extension term.



(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 perturbations $E_1 \sim \text{Unif}((-s, s))$. We choose s such that the coefficient of error (CE) of the point-increments are 5% (left) and 20%, (right).

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 [9, sec. 3.5 pp. 93-94], which satisfies the strong integrability assumption (1.3). 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 1.8, that is (1.8). 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.



(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 increments $\omega_1 \sim \text{Unif}((1-c, 1+c))$ and sampling with a Matérn hard core process of type II with intensity 1 and a hard core distance of 0.4. c is chosen such that the coefficient of error of the increment is 5%.

8 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 preferred over the generalized Cavalieri method, 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 n'th order Newton-Cotes estimation yields an upper bound of the variance decreasing at the same rate as the variance of the classical Cavalieri estimator based on equidistant sampling, that is, a rate of $2n + 2$. Applying n'th 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 are *strongly n-admissible*. We have shown that point processes from the perturbed and cumulative models are strongly admissible and as such the strong bound holds in these cases. Based on a simulation study of the trapezoidal estimator it seems that also sampling from the 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 classical estimator based on equidistant sampling 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 positive 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 1.8 are optimal in all cases. As the rate of decrease in (1.7) 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 (1.7) is optimal for any stationary point process satisfying the assumptions of the theorem. Similarly we know that the bound presented in (1.8) yields the optimal decay-rate when $n = 1$ under the perturbed model (assuming non-degenerate perturbations), and as such 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.

Acknowledgments

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

References

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

[2] A.J. Baddeley and Jensen, E.B.V. *Stereology for Statisticians*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. CRC Press, 2004.

[3] A.J. Baddeley, K-A. Dorph-Petersen, and Jensen, E.B.V. A note on the stereological implications of irregular spacing of sections. *J. Microsc*, 222(3):177–181, 2006.

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

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

[6] M. Kiderlen and K-A. Dorph-Petersen. The Cavalieri estimator with unequal section spacing revisited. *Image Anal. Stereol.*, 36(2):135–141, 2017.

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

[8] K. Knopp. *Theorie und Anwendung der Unendlichen Reihen*. Die Grundlehren der mathematischen Wissenschaften. Springer, Berlin & Heidelberg, 1947.

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

[10] Mads Stehr. *Stereology and Spatio-Temporal Models*. PhD thesis, Aarhus University, In preperation.

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

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

Appendix: Supplement

A Integrability properties

As mentioned in section 3 the weight $\beta_j^{(n)}(x)$ is particularly simple, as given in the lemma below.

Lemma A.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 $j \in \{0,\dots,n\}$, and consider $\beta_j^{(n)}(x)$ as defined by (2.3). 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 (2.3) 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

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

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

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

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. $\square$

Assuming either Assumption 1.1 or Assumption 3.1, this representation ensures the Palm integrability of $\alpha(0)$, which is used in the proof of Theorem 1.2.

Lemma A.2. *Fix $n \in \mathbb{N}$. If X is a stationary point process such that* (3.1) *is satisfied, then*

$$\mathbb{E}^0|\beta_j^{(n)}(s_{-j})| < \infty \tag{A.2}$$

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

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

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

where $\mathbf{m}'_0$ is a specific multi-index such that $a_{\mathbf{m}'_0}^{(n,j)} > 0$ which exists by Lemma A.1. By linearity (A.2) is satisfied whenever

$$\mathbb{E}^0\left[\frac{\mathbf{h}(s_{-j})^{\mathbf{m}}}{\mathbf{h}(s_{-j})^{\mathbf{m}'}}\right] = \mathbb{E}^0\left[\frac{\mathbf{h}^{\mathbf{m}}}{\mathbf{h}^{\mathbf{m}'}}\right] < \infty, \tag{A.3}$$

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 [6, Eq. (13)]. The right side of (A.3) is finite by Assumption 3.1. □

B 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 $X = \{U + E_k + k\}_{k\in\mathbb{Z}}$ from the perturbed model. Note that the equidistant model is obtained with degenerate perturbations concentrated at 0. As in section 6 we work with the shifted kernel, K_m^*, defined by

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

where $X^* = X - U = \{E_k + k\}_{k\in\mathbb{Z}}$ is the shifted process. From (6.3) 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 B.1. *Let $n \in \mathbb{N}$ be given and let X be from the perturbed model such that* (1.3) *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*

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

where

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

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

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

for $m \leq n$ and all $|r| \geq 2n+2$. If the perturbations are degenerate, that is X is the equidistant model, (B.2) *is true for all $r \in \mathbb{R}$.*

Proof. By (6.3) 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 then

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

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

$$A_i(r) = \sum_{\ell=1-n}^{0} R^{(n)}_{x_{i+\ell}}((\,\cdot\, - r)^0_+) = \sum_{\ell=1-n}^{0}\sum_{j=0}^{n} \beta_j^{(n)}(x_{i+\ell})\, \mathbf{1}_{\ell+j\ge1} - \sum_{\ell=1}^{n}(x_{i+\ell} - r).$$

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

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

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

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

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

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

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

To show (B.2), we note that

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

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 $\mathbb{E}K_m(0) = 0$ for all $m < n$ (see Lemma 3.2), and the continuity properties of the kernels and polynomials. For $|r| \geq 2n + 2$, the perturbations in $K_m(r) = K_m^*(r - U)$ and $K_m(0) = K_m^*(-U)$ 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, (B.3) and independence then implies

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

where the last equality is shown in the proof of [7, Prop. 5.2]. This shows (B.2). If the model has degenerate perturbations concentrated at 0, (B.4) is true for all $r \in \mathbb{R}$ with $X^* = \mathbb{Z}$ deterministic. This concludes the proof. □

Corollary B.2. *Let $n \in \mathbb{N}$ be given. If $X = U + \mathbb{Z}$ is the equidistant model, then*

$$H_m(r) = (-1)^m P_{2m+2}(r) \tag{B.5}$$

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

Proof. Fix $n \in \mathbb{N}$. Note that $X^* = \mathbb{Z}$ and hence it is deterministic. From Lemma B.1 it therefore suffices to show that $K_0^*(r) = -P_1(-r)$ for $r \in [0, 1)$. Also, the weights $\beta_j^{(n)}(x)$ does not depend on $x \in X^*$, and we therefore let the common weights be denoted $\beta_j^{(n)}$. 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 (B.1), we conclude that $K_0^*(r) = -P_1(-r)$. □

Corollary B.3. *Let $n = 1$. If X is the perturbed model, then*

$$H_m(r) = (-1)^m \mathbb{E}[P_{2m+2}(r + E_1 - E_0)] \tag{B.6}$$

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

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

Lastly we show the asymptotic variance expression (1.11) fully, with the beginning of the proof given already.

Proof of (1.11). Let $m \geq 2$ be given and define $\tilde{H}_1$ by $\tilde{H}_1(s) = H_1(s) + \mathbb{E}[P_4(s + E_1 - E_0)]$. Due to Lemma 6.2, $\tilde{H}_1(s)$ vanishes for $|r| > 4$. Since $g^{(4)}$ is continuous, an application of Proposition 5.1 to the $(1, 1)$-piecewise smooth function f, Fubini's theorem and the refined partial integration formula [7, Lemma 4.1] yield

$$\begin{aligned}\mathbb{V}\mathrm{ar}(\hat{V}_{1,t}(f)) = t^5 \int_{\mathbb{R}} g^{(4)}(st)\tilde{H}_1(s)\mathrm{d}s - t^6 \int_{\mathbb{R}} g^{(6)}(s)\mathbb{E}[P_6(\tfrac{s}{t} + E_1 - E_0)]\mathrm{d}s \\ - t^6 \sum_{a \in D_{g^{(5)}}} J_{g^{(5)}}(a)\mathbb{E}[P_6(\tfrac{a}{t} + E_1 - E_0)].\end{aligned} \tag{B.7}$$

As the last two terms in (B.7) 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 (B.7), and using the fact that $g^{(5)}$ and $\tilde{H}_1$ are bounded, yields

$$\mathbb{V}\mathrm{ar}(\hat{V}_{1,t}(f)) = t^5 g^{(4)}(0) \int_{-4}^{4} \tilde{H}_1(s)\mathrm{d}s + o(t^5) \tag{B.8}$$

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 2.3 gives

$$\int_{-4}^{4}\tilde{H}_1(s)\mathrm{d}s = \int_{-4}^{4}H_1(s)\mathrm{d}s$$
$$= \mathbb{E}\Big(\big[K_2(-4) + \sum_{x\in X\cap[-4,4]} J_{K_2}(x) - K_2(4)\big]\big(K_1(0)-\mathbb{E}K_1(0)\big)\Big). \quad \text{(B.9)}$$

The arguments that lead to (6.4) in combination with (6.3) where $r=-4$ and $k=8$ imply $\mathbb{E}[K_2(-4)K_1(0)] = \mathbb{E}[K_2(4)K_1(0)]$, and the two marginal terms in the last expression of (B.9) cancel. Hence,

$$\int_{-4}^{4}\tilde{H}_1(s)\mathrm{d}s = \mathbb{E}\sum_{x\in X\cap[-4,4]} J_{K_2}(x)\big(K_1(0)-\mathbb{E}K_1(0)\big)$$
$$= \mathbb{E}\sum_{y\in X}\mathbf{1}_{(y,s_1(y)]}(0)\sum_{j=-\infty}^{\infty}1_{[-4,4]}(s_j(y))J_{K_2}(s_j(y))\big(K_1(0)-\mathbb{E}K_1(0)\big).$$

Applying the refined Campbell Theorem [9, Theorem 3.5.3] and using the fact that the summation over all successors of y can be replaced for the Palm version of the point process by a summation over all successors of 0 and has only finitely many nontrivial terms, we arrive at

$$\int_{-4}^{4}\tilde{H}_1(s)\mathrm{d}s = \sum_{j=-\infty}^{\infty}\mathbb{E}^0\theta_j \quad \text{(B.10)}$$

with $\theta_j = J_{K_2}(s_j)\int_0^{s_1}\mathbf{1}_{[-4,4]}(s_j-y)\big(K_1(y)-\mathbb{E}K_1(0)\big)\mathrm{d}y$.

We now show that all summands in (B.10) with $|j|>1$ can be omitted, and that the remaining terms can be expressed in terms of moments of E_0. Using $s_j = j+E_j-E_0$, we see that the indicator in the definition of θ_j is constant 1 when $j\in I_1=\{-2,-1,0,1,2,3\}$ and it is constant 0 when $j\in I_0=\mathbb{Z}\setminus\{-4,\dots,3,5\}$. For $j\in I_1$ the integral can be shown to evaluate to $(1/12)(h_1^3-h_1\mathbb{E}^0h_1^3)$. As $J_{K_2}(s_j) = (1/2)R^{(1)}_{s_j}\big((\,\cdot\,)^2\big) = (1/12)h_{j+1}^3$ only depends on the perturbations E_{j+1} and E_j, it is stochastically independent of the integral when $|j|>1$. In particular, $\theta_j = \frac{1}{144}h_{j+1}^3(h_1^3-h_1\mathbb{E}^0h_1^3)$ for $j\in I_1$, and

$$\int_{-4}^{4}\tilde{H}_1(s)\mathrm{d}s = \mathbb{E}^0\theta{-1}+\mathbb{E}^0\theta_0+\mathbb{E}^0\theta_1+Q, \quad \text{(B.11)}$$

where

$$Q = \sum_{j\in\mathbb{Z}\setminus(I_0\cup I_1)}\mathbb{E}^0\theta_j = \mathbb{E}^0\sum_{j\in\{-4,-3,4,5\}}\theta_j.$$

A coupling argument shows that Q is zero. The Palm expectation of θ_{-4} and θ_{-3} are unchanged when we put $E_{-3}=E_5$, $E_{-4}=E_4$ and $E_{-5}=E_3$. Under this coupling assumption, $s_{-4}=s_4-8$, $s_{-3}=s_5-8$, $h_{-4}=h_4$, and $h_{-3}=h_5$, and hence

$$\mathbb{E}^0\theta_{-3}+\mathbb{E}^0\theta_5 = \frac{1}{12}\mathbb{E}^0h_6^3\int_0^{s_1}\Big(\mathbf{1}_{[-4,4]}((s_5-y)-8)$$
$$+\mathbf{1}_{[-4,4]}(s_5-y)\Big)\big(K_1(y)-\mathbb{E}K_1(0)\big)\mathrm{d}y = 0,$$

as this sum of indicators is almost surely 1 and the independence property from above can be exploited once again. Similar arguments show $\mathbb{E}^0\theta_{-4} + \mathbb{E}^0\theta_4 = 0$, so $Q = 0$. Summarizing, we obtain from (B.11)

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

where the last equality follows from lengthy and tedious – but elementary – calculations. Inserting this into (B.8) yields the assertion (1.11). □

C Admissibility of the cumulative model

In this section we show that the covariance function of the Peano kernel decrease exponentially when we sample from the cumulative model under certain integrability conditions; we refer to the beginning of subsection 6.2 for the construction of the process with increments $\{\omega_i\}_{i\in\mathbb{Z}}$ and increment distribution F. Also, recall the overall assumption that the typical increment has finite moments of all orders.

$$\mathbb{E}\omega_1^j < \infty \tag{C.1}$$

for all $j \in \mathbb{Z}$. This ensures that the Peano kernel has finite positive moments of any order. We will explicitly use that, for fixed $n \in \mathbb{N}$, $K_n(s)$ depends on n points in X to each side of s.

We show the assertion for $n = 1$ and arbitrary $n \in \mathbb{N}$. The proof of the latter is a generalization of the former, as we need to consider multiple points simultaneously, hence, the case $n = 1$ serves as an introduction to the ideas behind the general proof.

Lemma C.1. *Let $n = 1$ and let X be from the cumulative model such that $\mathbb{E}\mathrm{e}^{\eta\omega_1} < \infty$ for some $\eta > 0$. Then*

$$H_1(s) = O(\mathrm{e}^{-\epsilon s}), \qquad s \to \infty, \tag{C.2}$$

for some $\epsilon > 0$. In particular, X is strongly 1-admissible.

Proof. In the case of the trapezoidal estimator, the overall assumption (C.1) can be relaxed to only hold true for $j \in \mathbb{N}$. As we assume that the increments have exponential moment, they in particular have finite moments of any positive order, and hence, all integrability results of the Peano kernels apply.

The proof of (C.2) relies heavily on exponential decays in renewal theory, and we refer to [1, Chapter V] for an introduction to renewal theory. 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, $U(s) = \mathbb{E}N(s)$. If X^0 denotes the Palm version of X, the points generated by N are exactly (in law) the points in $X^0 \cap [0,\infty)$, and hence $U(s) = \mathbb{E}^0\#(X \cap [0,s])$. Let $\psi_1 : [0,\infty) \to \mathbb{R}$ denote the expected Peano kernel driven by the points generated by the pure renewal process N, that is, $\psi_1(s) = \mathbb{E}K_1(s;N) = \mathbb{E}^0 K_1(s;X)$, and define $\tilde{\psi}_1$ by $\tilde{\psi}_1(s) = \mathbb{E}K_1(s;N)\,\mathbf{1}_{\omega_1>s} = \mathbb{E}^0 K_1(s;X)\,\mathbf{1}_{h_1>s}$. Then

$\psi_1(s) = U * \tilde{\psi}_1(s) = \int_0^s \tilde{\psi}_1(s-x)U(\mathrm{d}x)$. This can be seen by a renewal argument obtaining the renewal equation $\psi_1 = \tilde{\psi}_1 + \psi_1 * F$, which has the desired solution. Another, rather intuitive, approach is to condition on the last arrival prior to s happening at time $x \in [0, s)$, which has probability $U(\mathrm{d}x)\bar{F}(s-x)$. By starting a new independent pure renewal process at time x and conditioning on the size of the first jump, the conditional expectation of $K_1(s; N)$ given that the last arrival prior to s happens at x simplifies as

$$\mathbb{E}\big[K_1(s-x; N) \mid \omega_1 > s-x\big] = \tilde{\psi}_1(s-x)/\bar{F}(s-x),$$

and we obtain the expression $\psi_1(s) = U * \tilde{\psi}_1(s)$ by integrating over $[0, s]$.

The exponential moment assumption implies $\bar{F}(x) = O(\mathrm{e}^{-\eta x})$, $x \to \infty$, which in turn implies that also $\bar{G}(x) = O(\mathrm{e}^{-\eta x})$. We consider

$$\mathbb{E}K_1(0)K_1(s) = \mathbb{E}K_1(s)K_1(0)\,\mathbf{1}_{X_0 \le s} + \mathbb{E}K_1(s)K_1(0)\,\mathbf{1}_{X_0 > s},$$

and an application of Cauchy-Schwarz' inequality yields

$$\mathbb{E}K_1(s)K_1(0)\,\mathbf{1}_{X_0 > s} \le [\mathbb{E}K_1^2(s)K_1^2(0)]^{1/2}\mathbb{P}(X_0 > s)^{1/2} \le C\big(\bar{G}(s)\big)^{1/2}$$

for some finite C. Hence $\mathbb{E}K_1(s)K_1(0)\,\mathbf{1}_{X_0 > s} = O(\mathrm{e}^{-\eta s/2})$ as $s \to \infty$, and (C.2) therefore follows once we show that

$$\mathbb{E}K_1(s)K_1(0)\,\mathbf{1}_{X_0 \le s} = (\mathbb{E}K_1(0))^2 + O(\mathrm{e}^{-\epsilon s}) \tag{C.3}$$

for some $\epsilon > 0$, as $s \to \infty$. We apply a renewal argument conditioning on the first arrival in $X \cap (0, \infty)$, that is, conditioning on the value of $X_0 \sim G$, and then initializing a new independent renewal process,

$$\begin{aligned}
\mathbb{E}K_1(s)K_1(0)\,\mathbf{1}_{X_0 \le s} &= \int_0^s \mathbb{E}[K_1(0)K_1(s) \mid X_0 = u]G(\mathrm{d}u) \\
&= \int_0^s \mathbb{E}[K_1(s-u; N)]\mathbb{E}[K_1(0) \mid X_0 = u]G(\mathrm{d}u) \\
&= \int_0^s \psi_1(s-u)\mathbb{E}[K_1(0) \mid X_0 = u]G(\mathrm{d}u).
\end{aligned}$$

Since $\mathbb{E}K_1(0) = \mathbb{E}^0 J_{K_2}(0)$ due to (3.3), an application of Fubini's theorem yields

$$\int_0^\infty \tilde{\psi}_1(s)\mathrm{d}s = \mathbb{E}\int_0^{\omega_1} K_1(s; N)\mathrm{d}s = \mathbb{E}K_2(0^+; N) - \mathbb{E}K_2(\omega_1^-; N) = \mathbb{E}K_1(0),$$

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

$$\psi_1(s) = U * \tilde{\psi}_1(s) = \mathbb{E}K_1(0) + O(\mathrm{e}^{-\epsilon' s}) \tag{C.4}$$

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

$$\begin{aligned}
\int_0^s \mathbb{E}[K_1(0) \mid X_0 = u]G(\mathrm{d}u) &= \mathbb{E}K_1(0) - \mathbb{E}K_1(0)\,\mathbf{1}_{X_0 > s} \\
&= \mathbb{E}K_1(0) + O(\mathrm{e}^{-\eta s/2})
\end{aligned} \tag{C.5}$$

as $s \to \infty$. Combining (C.4) and (C.5) yields (C.3). □

Lemma C.2. *Let $n \in \mathbb{N}$ be given, and let X be the cumulative model such that $\mathbb{E}e^{\eta\omega_1} < \infty$ for some $\eta > 0$, and such that* (C.1) *is satisfied. Then*

$$H_n(s) = O(e^{-\epsilon s}), \qquad s \to \infty, \tag{C.6}$$

for some $\epsilon > 0$. In particular, X is strongly n-admissible.

Proof. We apply arguments similar to those of the proof of Lemma C.1 above, only we have to use renewal arguments conditioning on the n'th point in $X \cap (0, \infty)$, as $K_n(s)$ depends on n points to the left and right of s. As before, we let $N = (N(s))_{s \geq 0}$ be a pure renewal process with increments $\{\omega_i\}_{i \in \mathbb{N}}$, and we 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, $X \cap (0, \infty) = \{X_0 + y_i\}_{i \in \mathbb{N}}$. Then $y_i \sim F^{*i}$ for all $i \in \mathbb{N}$. Define $\psi_n : [0, \infty) \to \mathbb{R}$ by $\psi_n(s) = \mathbb{E}[K_n(s; N) \mathbf{1}_{y_{n-1} \leq s}] = \mathbb{E}^0[K_n(s) \mathbf{1}_{s_{n-1} \leq s}]$, and let $\tilde{\psi}_n : [0, \infty) \to \mathbb{R}$ be given by $\tilde{\psi}_n(s) = \mathbb{E}[K_n(s; N) \mathbf{1}_{y_{n-1} \leq s} \mathbf{1}_{y_n > s}] = \mathbb{E}^0[K_n(s) \mathbf{1}_{s_{n-1} \leq s} \mathbf{1}_{s_n > s}]$. Then $\psi_n(s) = U * \tilde{\psi}_n(s)$. This can be seen by conditioning on the n'th 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 $[0, s]$,

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

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

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

and an application of Cauchy-Schwarz' inequality yields

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

for some finite C. Hence $\mathbb{E}K_n(s)K_n(0) \mathbf{1}_{X_0 + y_{2n-2} > s} = O(e^{-\eta s/2})$ as $s \to \infty$, and (C.6) therefore follows once we show that

$$\mathbb{E}K_n(s)K_n(0) \mathbf{1}_{X_0 + y_{2n-2} \leq s} = (\mathbb{E}K_n(0))^2 + O(e^{-\epsilon s}) \tag{C.7}$$

for some $\epsilon > 0$, as $s \to \infty$. We apply a renewal argument conditioning on the n'th arrival in $X \cap (0, \infty)$, 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,

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

Since $\mathbb{E}K_n(0)=\mathbb{E}^0 J_{K_{n+1}}(0)$ due to (3.3), an application of Fubini's theorem yields

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

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

$$
\psi_n(s)=U*\tilde{\psi}_n(s)=\mathbb{E}K_n(0)+O(\mathrm{e}^{-\epsilon' s}) \tag{C.8}
$$

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

$$
\begin{aligned}
&\int_0^s \mathbb{E}\big[K_n(0)\mid X_0+y_{n-1}=u\big]\big(G*F^{*(n-1)}\big)(\mathrm{d}u)\\
&\quad= \mathbb{E}K_n(0)-\mathbb{E}K_n(0)\,\mathbf{1}_{X_0+y_{n-1}>s}\\
&\quad= \mathbb{E}K_n(0)+O(\mathrm{e}^{-\eta s/2})
\end{aligned} \tag{C.9}
$$

as $s\to\infty$. Combining (C.8) and (C.9) yields (C.7). □