Adv. Appl. Prob. (SGSA) **46,** 35–58 (2014)
Printed in Northern Ireland

LOCAL DIGITAL ESTIMATORS OF INTRINSIC VOLUMES FOR BOOLEAN MODELS AND IN THE DESIGN-BASED SETTING

ANNE MARIE SVANE,* *Aarhus University*

Abstract

In order to estimate the specific intrinsic volumes of a planar Boolean model from a binary image, we consider local digital algorithms based on weighted sums of 2×2 configuration counts. For Boolean models with balls as grains, explicit formulas for the bias of such algorithms are derived, resulting in a set of linear equations that the weights must satisfy in order to minimize the bias in high resolution. These results generalize to larger classes of random sets, as well as to the design-based situation, where a fixed set is observed on a stationary isotropic lattice. Finally, the formulas for the bias obtained for Boolean models are applied to existing algorithms in order to compare their accuracy.

Keywords: Digitization in 2D; intrinsic volume; local estimator; configuration; Boolean model; design-based digitization

2010 Mathematics Subject Classification: Primary 94A08
Secondary 60D05

1. Introduction

Let $X \subseteq \mathbb{R}^2$ be a compact subset of the plane. Suppose that we are given a digital image of X, i.e. the only information about X available to us is the set $X \cap \mathbb{L}$ where $\mathbb{L} \subseteq \mathbb{R}^2$ is a square lattice. In the language of signal processing, we are thus using an *ideal sampler* to obtain a sample of the characteristic function of X at all the points of $\mathbb{L}$. In image analysis terms, $\mathbb{L}$ can be interpreted as the set of all pixel midpoints and the digitization $X \cap \mathbb{L}$ contains the same information about X as the commonly used Gauss digitization [9, p. 56]. From this binary representation of X, we would like to recover certain geometric properties of X. The quantities we are interested in are the so-called intrinsic volumes V_i. In the plane, these are simply the volume $V_2(X)$, the boundary length $2V_1(X)$, and the Euler characteristic $V_0(X)$. See [13, Chapter 4] for the definition when X is polyconvex.

In this paper, we exclusively consider local digital estimators based on 2×2 configuration counts in a square lattice. Motivated by the additivity of intrinsic volumes, these are defined as follows. The plane is divided into a disjoint union of square cells with vertices in $\mathbb{L}$. For each 2×2 cell in the lattice, each vertex may belong to either X or $\mathbb{R}^2 \setminus X$, yielding $2^4 = 16$ different possible configurations. Each cell contributes to the estimator for $V_i(X)$ with a certain weight depending only on the configuration. Thus, the estimator becomes a weighted sum of the configuration counts. The weights can in principle be chosen freely. Algorithms of this type are desirable as they are simple and efficiently implementable based on linearly filtering the image.

Received 7 August 2012; revision received 28 February 2013.
* Postal address: Department of Mathematics, Aarhus University, 8000 Aarhus C, Denmark.
Email address: anne.marie.svane@gmail.com

https://doi.org/10.1239/aap/1396360102 Published online by Cambridge University Press

One way of testing the quality of local algorithms is by simulations on a fixed test set for various high resolutions, see e.g. [9, Section 10.3.4]. In contrast, we shall follow Ohser *et al.* [12], where the algorithms are applied to a standard model from stochastic geometry, namely the Boolean model. But, rather than testing a known algorithm, we let the weights be arbitrary and derive conditions on the weights such that the bias of the estimator is minimal for high resolutions.

If the grains are almost surely balls, a Steiner-type result for finite sets shown in [6] yields a general formula for the estimator from which the asymptotic behaviour can be derived. The main result is that a local estimator is asymptotically unbiased if and only if the weights satisfy certain linear equations. Moreover, we obtain formulas for the approximate bias in high resolution. These results are stated in Theorems 4.1 and 4.2, below.

Local estimators are introduced in Section 2. This is specialized to Boolean models in Section 3 and the computations are performed in Section 4.

In Section 5, the main theorems are generalized to a larger class of Boolean models where the grains allow a ball of radius $\varepsilon > 0$ to slide freely. A formula by Kiderlen and Vedel Jensen [8] also yields an immediate generalization of the first-order results to general standard random sets; see Section 6.

We then turn to the design-based situation where a deterministic set X is observed on a randomly translated and rotated lattice. Under certain conditions on X, we obtain a generalization of the main theorems for Boolean models. This is done for the boundary length in Section 7, using a result of Kiderlen and Rataj [7], and for the Euler characteristic in Section 8 by a refinement of their approach.

In the literature, various algorithms for computing intrinsic volumes are suggested. The obtained formulas allow for a computation of the bias in high resolution and hence a comparison of the commonly used algorithms. This is the content of the last section of the paper, Section 9.

2. Local digital estimators

Let $\mathbb{Z}^2$ be the standard lattice in $\mathbb{R}^2$. Let C denote the unit square $[0, 1] \times [0, 1]$ in $\mathbb{R}^2$ and let C_0 be the set of vertices in C. We enumerate the elements of C_0 as follows: $x_0 = (0, 0)$, $x_1 = (1, 0)$, $x_2 = (0, 1)$, and $x_3 = (1, 1)$. A configuration is a subset $\xi \subseteq C_0$. We denote the 16 possible configurations by ξ_l, $l = 0, \dots, 15$, where the configuration ξ is assigned the index

$$l = \sum_{i=0}^{3} 2^i \, \mathbf{1}_{x_i \in \xi} \, .$$

Here, $\mathbf{1}_{x_i \in \xi}$ is the indicator function.

More generally, we shall consider an orthogonal lattice $a\mathbb{L} = aR_v(\mathbb{Z}^2 + c)$, where $c \in C$ is a translation vector, R_v is the rotation by the angle $v \in [0, 2\pi]$, and $a > 0$ is the lattice distance. The configuration ξ_l is then understood to be the corresponding transformation $aR_v(\xi_l + c)$ of the configuration $\xi_l \subseteq \mathbb{Z}^2$.

The elements of ξ_l are referred to as the *foreground* or *black pixels* and will also sometimes be denoted by B_l, while the points in the complement $W_l = C_0 \setminus \xi_l = \xi_{15-l}$ are referred to as the *background* or *white pixels*.

The 16 possible configurations are divided into six equivalence classes under rigid motions. These are denoted by η_j for $j = 1, \dots, 6$. These are defined in Table 1. The number d_j is the number of elements in the equivalence class η_j.

https://doi.org/10.1239/aap/1396360102 Published online by Cambridge University Press

TABLE 1: Configuration classes.

j	η_j	d_j	Description	Example
1	$\{\xi_0\}$	1	four white vertices	$\begin{bmatrix} \circ & \circ \\ \circ & \circ \end{bmatrix}$
2	$\{\xi_1, \xi_2, \xi_4, \xi_8\}$	4	three white and one black vertices	$\begin{bmatrix} \circ & \circ \\ \bullet & \circ \end{bmatrix}$
3	$\{\xi_3, \xi_5, \xi_{10}, \xi_{12}\}$	4	two adjacent white and two black vertices	$\begin{bmatrix} \circ & \circ \\ \bullet & \bullet \end{bmatrix}$
4	$\{\xi_6, \xi_9\}$	2	two opposite white and two black vertices	$\begin{bmatrix} \circ & \bullet \\ \bullet & \circ \end{bmatrix}$
5	$\{\xi_7, \xi_{11}, \xi_{13}, \xi_{14}\}$	4	one white and three black vertices	$\begin{bmatrix} \bullet & \circ \\ \bullet & \bullet \end{bmatrix}$
6	$\{\xi_{15}\}$	1	four black vertices	$\begin{bmatrix} \bullet & \bullet \\ \bullet & \bullet \end{bmatrix}$

Now let $X \subseteq \mathbb{R}^2$ be a compact set. Suppose that we observe X on the lattice $a\mathbb{L}$. Based on the set $X \cap a\mathbb{L}$, we want to estimate the intrinsic volumes V_i introduced in Section 1.

In order for the V_i to be well defined and for the digitization $X \cap a\mathbb{L}$ to carry enough information about X, we require that X is sufficiently 'nice'. The notion of a gentle set is introduced in Section 7 when dealing with V_1. This includes all topologically regular polyconvex sets. When we work with V_0, X will be assumed to be either a compact topologically regular polyconvex set or a compact full-dimensional C^2 manifold. A set is called *topologically regular* if it coincides with the closure of its interior.

Our approach is to consider a local algorithm based on the observations of X on the 2×2 cells of $a\mathbb{L}$. By additivity of the intrinsic volumes, $V_i(X)$ is a sum of contributions from each lattice cell $z + aR_v(C)$ for $z \in a\mathbb{L}$. We estimate this by a certain weight $w^{(i)}(a, z)$, depending only on the information we have about the cell, i.e. the configuration

$$X \cap (z + aR_v(C_0)) - (z - c) = (X - (z - c)) \cap \xi_{15}.$$

Recall here that $\xi_{15} = aR_v(C_0 + c)$ is the set of vertices in the unit cell of $a\mathbb{L}$.

Since V_i is invariant under rigid motions, we should like the estimator to satisfy $\hat{V}_i(X) = \hat{V}_i(MX)$, for any rigid motion M preserving $a\mathbb{L}$. Thus, $w^{(i)}(a, z)$ should depend only on the equivalence class η_j of $(X - (z - c)) \cap \xi_{15}$ under rigid motions.

As V_i is homogeneous of degree i, i.e. $V_i(aX) = a^i V_i(X)$, the estimator should also satisfy

$$\hat{V}_i(aX \cap a\mathbb{L}) = a^i \hat{V}_i(X \cap \mathbb{L}).$$

We therefore assume that $w^{(i)}(a, z) = a^i w_j^{(i)}$, where $w_j^{(i)} \in \mathbb{R}$ are constants.

Consequently, we are led to consider estimators of the form

$$\hat{V}_i(X) = a^i \sum_{j=1}^{6} w_j^{(i)} N_j,$$

where N_j is the number of occurrences of the configuration class η_j, i.e.

$$N_j = \sum_{z \in a\mathbb{L}} \mathbf{1}_{(X-(z-c)) \cap \xi_{15} \in \eta_j}.$$

https://doi.org/10.1239/aap/1396360102 Published online by Cambridge University Press

It is also natural to require the estimators to be compatible with interchanging background and foreground as follows:

$$\hat{V}_1(X) = \hat{V}_1(\mathbb{R}^2 \setminus X), \tag{2.1}$$

$$\hat{V}_0(X) = -\hat{V}_0(\mathbb{R}^2 \setminus X). \tag{2.2}$$

The reason for the first condition is that interchanging foreground and background does not change the boundary. The second condition is natural because the Euler characteristic satisfies

$$V_0(X) = -V_0(\overline{\mathbb{R}^2 \setminus X})$$

for both topologically regular compact polyconvex sets (see [11]) and compact 2-manifolds with boundary (where $\overline{A}$ denotes the closure of $A \subseteq \mathbb{R}^2$).

3. The 2D Boolean model

Throughout this paper, a Boolean model Ξ will mean a stationary isotropic Boolean model in the plane with compact convex grains and intensity γ. That is,

$$\Xi = \bigcup_i (x_i + K_i),$$

where $\{x_1, x_2, \dots\}$ is a stationary Poisson process in $\mathbb{R}^2$ with intensity γ and $K_1, K_2, \dots$ is a sequence of independent and identically distributed random compact convex sets in $\mathbb{R}^2$ with rotation invariant distribution $\mathbb{Q}$ satisfying $\mathbb{E}V_i(K) < \infty$ for $i = 0, 1, 2$. See, for example, [14] for more details.

The specific intrinsic volumes of a Boolean model are defined by

$$\overline{V}_i(\Xi) = \lim_{r\to\infty} \frac{\mathbb{E}V_i(\Xi \cap rW)}{V_2(rW)}, \tag{3.1}$$

where W is any compact convex set with nonempty interior; see [14, Theorem 9.2.1].

Now assume that we observe Ξ on a lattice $a\mathbb{L}$ in a compact convex window W with nonempty interior. By the isotropy assumption, we may as well assume the lattice to be the standard lattice $a\mathbb{Z}^2$. Thus, we observe the set $\Xi \cap a\mathbb{Z}^2 \cap W$.

Let $C_z = z + aC$ be a lattice cell with $z \in a\mathbb{Z}^2$. Write

$$V_{i,z} = V_i(C_z \cap \Xi) - V_i(\partial^+ C_z \cap \Xi),$$

where $\partial^+ C_z = z + a([0, 1] \times \{1\} \cup \{1\} \times [0, 1])$ is the upper-right boundary. Then [14, Theorem 9.2.1] implies that $\mathbb{E}V_{i,z} = a^2\overline{V}_i(\Xi)$. A summation over all lattice cells contained in W yields

$$\overline{V}_i(\Xi) = \sum_{z\in a\mathbb{Z}^2\cap(W\ominus a\check{C})} \frac{\mathbb{E}V_{i,z}}{V_2(C_z)N_0} = \sum_{z\in a\mathbb{Z}^2\cap(W\ominus a\check{C})} \frac{\mathbb{E}V_{i,z}}{a^2N_0}, \tag{3.2}$$

where $\check{C} = \{-x \mid x \in C\}$, $W \ominus a\check{C} = \{x \in \mathbb{R}^2 \mid x + aC \subseteq W\}$, and N_0 is the total number of points in $a\mathbb{Z}^2 \cap (W \ominus a\check{C})$.

As in Section 2, we estimate each contribution $\mathbb{E}V_{i,z}$ by a weight of the form $a^i w_j^{(i)}$ depending on the configuration type η_j. Then (3.2) yields an estimator of the form

$$\hat{V}_i(\Xi) = a^{i-2} \sum_{j=1}^{6} w_j^{(i)} \frac{N_j}{N_0}, \tag{3.3}$$

https://doi.org/10.1239/aap/1396360102 Published online by Cambridge University Press

where $w_j^{(i)} \in \mathbb{R}$ are arbitrary weights and the number of configurations N_j are given by

$$N_j = \sum_{z \in a\mathbb{Z}^2 \cap (W \ominus a\check{C})} \mathbf{1}_{(\Xi - z) \cap \xi_{15} \in \eta_j} . \tag{3.4}$$

Ideally, $\hat{V}_i$ would define an unbiased estimator, i.e. $\mathbb{E}\hat{V}_i(\Xi) = \overline{V}_i(\Xi)$. Generally, this is not possible with finite resolution, i.e. when $a > 0$. Instead, we shall obtain conditions for this to hold asymptotically when the lattice distance tends to zero:

$$\lim_{a \to 0} \mathbb{E}\hat{V}_i(\Xi) = \overline{V}_i(\Xi).$$

The mean value of $\hat{V}_i(\Xi)$ is

$$\mathbb{E}\hat{V}_i(\Xi) = a^{i-2} \sum_{j=1}^{6} w_j^{(i)} \mathbb{E}\left(\frac{N_j}{N_0}\right) = a^{i-2} \sum_{j=1}^{6} w_j^{(i)} \mathbb{P}(\Xi \cap aC_0 \in \eta_j), \tag{3.5}$$

by (3.4) and stationarity of Ξ.

For each ξ_l, there are formulas of the form

$$\mathbb{P}(\Xi \cap aC_0 = \xi_l) = \sum_{k=0}^{15} b_{lk} \mathbb{P}(\xi_k \subseteq \mathbb{R}^2 \setminus \Xi), \tag{3.6}$$

for suitable integers b_{lk}; see also [12]. As Ξ is stationary and isotropic, $\mathbb{P}(\Xi \cap aC_0 = \xi_l)$ and $\mathbb{P}(\xi_k \subseteq \mathbb{R}^2 \setminus \Xi)$ depend only on ξ_l and ξ_k up to rigid motions. Let ξ_{k_i} and ξ_{l_j} be representatives for η_i and η_j, respectively. Then (3.6) reduces to

$$\mathbb{P}(\Xi \cap aC_0 = \xi_{l_j}) = \sum_{i=1}^{6} b'_{ij} \mathbb{P}(\xi_{k_i} \subseteq \mathbb{R}^2 \setminus \Xi), \tag{3.7}$$

where the integer b'_{ij} is the ijth entry of the matrix

$$B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 1 & 0 & -2 & 4 \\ 0 & 0 & 0 & 1 & -1 & 2 \\ 0 & 1 & -2 & -2 & 3 & -4 \\ 1 & -1 & 1 & 1 & -1 & 1 \end{pmatrix}.$$

The right-hand side of (3.7) is now well known, since

$$\mathbb{P}(\xi_k \subseteq \mathbb{R}^2 \setminus \Xi) = \mathrm{e}^{-\gamma \mathbb{E} V_2(\xi_k \oplus K)}, \tag{3.8}$$

where K is a random compact convex set of distribution $\mathbb{Q}$ and $\oplus$ denotes Minkowski addition; see [14]. Thus, we must compute $EV_2(\xi_k \oplus K)$.

If $F_k = \operatorname{conv}(\xi_k)$ denotes the convex hull of ξ_k, an application of the rotational mean value formula, see [14, Theorem 6.1.1], shows that

$$\mathbb{E}V_2(F_k \oplus K) = \mathbb{E}V_2(K) + \frac{2}{\pi} V_1(F_k) \mathbb{E}V_1(K) + V_2(F_k), \tag{3.9}$$

since the grain distribution is isotropic. It remains to compute the error

$$\mathbb{E}V_2(F_k \oplus K) - \mathbb{E}V_2(\xi_k \oplus K). \tag{3.10}$$

4. Boolean models with random balls as grains

We first restrict ourselves to Boolean models where the grains are almost surely (a.s.) balls $B(r)$ of random radius r. For technical reasons, we shall assume throughout this section that there is an $\varepsilon > 0$ such that $r \geq \varepsilon$ a.s.

In [6, Proposition 1], an expression for the error (3.10) was given. Applied to our situation, this becomes a power series in a/r:

$$V_2(F_k \oplus B(r)) - V_2(\xi_k \oplus B(r)) = 2a^2 \sum_{n=1}^{\infty} \frac{(2n-3)!!}{(2n)!!} V_1^{(2n+1)}(a^{-1}\xi_k)\left(\frac{a}{r}\right)^{2n-1}, \tag{4.1}$$

whenever a/r is sufficiently small. Since $a^{-1}\xi_k$ is independent of a, the $V_1^{(2n+1)}(a^{-1}\xi_k)$ are constants. These are called intrinsic power volumes in [6] and are given by

$$V_1^{(m)}(\xi_k) = \frac{1}{m2^{m-1}} \sum_{F \in \mathcal{F}_1(F_k)} \gamma(F_k, F) V_1(F)^m,$$

where $\mathcal{F}_1(F_k)$ is the set of 1-dimensional faces of F_k and $\gamma(F_k, F)$ is the outer angle which in $\mathbb{R}^2$ is just $(\dim(F_k))^{-1}$. See [6] for the definition of the double factorial.

The condition $r \geq \varepsilon$ a.s. ensures that, whenever a is sufficiently small, (4.1) holds a.s. Combining this with (3.9), we obtain a power series expansion

$$\begin{aligned}\mathbb{E}V_2(\xi_k \oplus B(r)) = \mathbb{E}V_2(B(r)) + a\frac{2}{\pi}V_1(a^{-1}F_k)\mathbb{E}V_1(B(r)) + a^2 V_2(a^{-1}F_k) \\ - a^3 V_1^{(3)}(a^{-1}\xi_k)\mathbb{E}(r^{-1}) + O(a^5).\end{aligned}$$

Computing the constants $V_i(a^{-1}F_k)$ and $V_1^{(3)}(a^{-1}\xi_k)$ directly and inserting in the Taylor expansion for the exponential function in (3.8), shows that $\mathbb{P}(\xi_k \subseteq \mathbb{R}^2 \setminus \Xi)$ is given by a power series

$$\begin{aligned} c_1 + \bigg(c_2 + ac_3\frac{\gamma}{\pi}\mathbb{E}V_1(B(r)) + a^2\bigg(c_4\gamma + c_5\Big(\frac{\gamma}{\pi}\mathbb{E}V_1(B(r))\Big)^2\bigg) \\ + a^3\bigg(c_6\gamma\mathbb{E}(r^{-1}) + c_7\frac{\gamma^2}{\pi}\mathbb{E}V_1(B(r)) + c_8\Big(\frac{\gamma}{\pi}\mathbb{E}V_1(B(r))\Big)^3\bigg)\bigg)\mathrm{e}^{-\gamma\mathbb{E}V_2(B(r))} + O(a^4),\end{aligned} \tag{4.2}$$

for a sufficiently small and constants $c_1, \ldots, c_8$ depending on k. If ξ_{k_j} is a representative for η_j, define A to be the matrix with entries a_{mj} and the constant c_m occurring in the formula for $\mathbb{P}(\xi_{k_j} \subseteq \mathbb{R}^2 \setminus \Xi)$ for $j = 1, \ldots, 6$. A direct computation shows that

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & -2 & -2\sqrt{2} & -(2+\sqrt{2}) & -4 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & -1 \\ 0 & 0 & 2 & 4 & 3+2\sqrt{2} & 8 \\ 0 & 0 & \frac{1}{12} & \sqrt{2}/6 & (\sqrt{2}+1)/12 & \frac{1}{6} \\ 0 & 0 & 0 & 0 & (2+\sqrt{2})/2 & 4 \\ 0 & 0 & -\frac{4}{3} & -8\sqrt{2}/3 & -(10+7\sqrt{2})/3 & -\frac{32}{3} \end{pmatrix}.$$

https://doi.org/10.1239/aap/1396360102 Published online by Cambridge University Press

Inserting this in (3.7), we obtain expressions for $\mathbb{P}(\Xi \cap aC_0 = \xi_{l_j})$ of the form (4.2) with constants c_m given by the jth column of AB. Then, by (3.5), $a^{2-i}\mathbb{E}\hat{V}_i(\Xi)$ is also of the form (4.2) with a vector of constants $c^{(i)} = (c_1^{(i)}, \dots, c_8^{(i)})$ given by

$$(c^{(i)})^\top = ABD(w^{(i)})^\top,$$

where $w^{(i)} = (w_1^{(i)}, \dots, w_6^{(i)})$ is the vector of weights and D is the diagonal matrix with jth diagonal entry the number d_j of elements in η_j. Writing this out, we get

$$\begin{aligned}
c_1^{(i)} &= w_6^{(i)},\\
c_2^{(i)} &= w_1^{(i)} - w_6^{(i)},\\
c_3^{(i)} &= 4(-w_1^{(i)} + (2-\sqrt{2})w_2^{(i)} + (-2+2\sqrt{2})w_3^{(i)} + (2-\sqrt{2})w_5^{(i)} - w_6^{(i)}),\\
c_4^{(i)} &= -w_1^{(i)} + 2w_2^{(i)} - 2w_5^{(i)} + w_6^{(i)},\\
c_5^{(i)} &= 4(2w_1^{(i)} + (-5+2\sqrt{2})w_2^{(i)} + (4-4\sqrt{2})w_3^{(i)} + (3-2\sqrt{2})w_4^{(i)}\\
&\qquad + (-7+6\sqrt{2})w_5^{(i)} + (3-2\sqrt{2})w_6^{(i)}),\\
c_6^{(i)} &= \tfrac{1}{6}(w_1^{(i)} + (2\sqrt{2}-2)w_2^{(i)} + (2-4\sqrt{2})w_3^{(i)} + (2\sqrt{2}-2)w_5^{(i)} + w_6^{(i)}),\\
c_7^{(i)} &= 2(2w_1^{(i)} + (-6+\sqrt{2})w_2^{(i)} + (4-2\sqrt{2})w_3^{(i)} + (2-\sqrt{2})w_4^{(i)}\\
&\qquad + (-2+3\sqrt{2})w_5^{(i)} - \sqrt{2}w_6^{(i)}),\\
c_8^{(i)} &= \tfrac{4}{3}(-8w_1^{(i)} + (22-7\sqrt{2})w_2^{(i)} + (-16+14\sqrt{2})w_3^{(i)} + (-6+3\sqrt{2})w_4^{(i)}\\
&\qquad + (10-13\sqrt{2})w_5^{(i)} + (-2+3\sqrt{2})w_6^{(i)}).
\end{aligned} \tag{4.3}$$

Note that $c_8^{(i)} = -16c_6^{(i)} - 2c_7^{(i)}$.

In [14, Theorem 9.1.4], the following formulas for the specific intrinsic volumes, valid for the type of Boolean models we consider, are shown:

$$\overline{V}_2(\Xi) = 1 - \mathrm{e}^{-\gamma \mathbb{E}V_2(K)}, \tag{4.4}$$

$$\overline{V}_1(\Xi) = \gamma \mathbb{E}V_1(K)\mathrm{e}^{-\gamma \mathbb{E}V_2(K)}, \tag{4.5}$$

$$\overline{V}_0(\Xi) = \left(\gamma - \frac{1}{\pi}(\gamma \mathbb{E}V_1(K)^2)\right)\mathrm{e}^{-\gamma \mathbb{E}V_2(K)}. \tag{4.6}$$

These are truncated expressions of the form (4.2) with fixed constants c_m, so the bias of $\mathbb{E}\hat{V}_i(\Xi)$ can be found by comparing coefficients.

First consider $\overline{V}_2(\Xi)$. From (4.2) we see that

$$\lim_{a\to 0} \mathbb{E}\hat{V}_2(\Xi) = c_1^{(2)} + c_2^{(2)}\mathrm{e}^{-\gamma \mathbb{E}V_2(B(r))},$$

so by (4.4) we get an asymptotically unbiased estimator for $\overline{V}_2(\Xi)$ exactly if $c_1^{(2)} = 1$ and $c_2^{(2)} = -1$. Using (4.3), we have the following result.

Proposition 4.1. *The estimator $\hat{V}_2(\Xi)$ is asymptotically unbiased if and only if the weights satisfy $w_1^{(2)} = 0$ and $w_6^{(2)} = 1$.*

It is well known that $\hat{V}_2(\Xi)$ is unbiased, even in finite resolution, with the choice $w^{(2)} = (0, \frac{1}{4}, \frac{1}{2}, \frac{1}{2}, \frac{3}{4}, 1)$, which is the estimator that counts the number of lattice points in X; see e.g. [10, Section 4.1.1].

Next we compare $\mathbb{E}\hat{V}_1(\Xi)$, with (4.5) and obtain the following result.

https://doi.org/10.1239/aap/1396360102 Published online by Cambridge University Press

Theorem 4.1. *The limit* $\lim_{a\to 0} \mathbb{E}\hat{V}_1(\Xi)$ *exists if and only if*

$$w_1^{(1)} = w_6^{(1)} = 0. \tag{4.7}$$

In this case,

$$\lim_{a\to 0} \mathbb{E}\hat{V}_1(\Xi) = \frac{1}{\pi} c_3^{(1)} \bar{V}_1(\Xi).$$

In particular, $\mathbb{E}\hat{V}_1(\Xi)$ *is asymptotically unbiased if and only if the weights satisfy*

$$c_3^{(1)} = 4((2-\sqrt{2})w_2^{(1)} + (-2+2\sqrt{2})w_3^{(1)} + (2-\sqrt{2})w_5^{(1)}) = \pi. \tag{4.8}$$

The bias is

$$a\left(c_4^{(1)}\gamma + c_5^{(1)}\left(\frac{\gamma}{\pi}\mathbb{E}V_1(B(r))\right)^2\right)\mathrm{e}^{-\gamma \mathbb{E}V_2(B(r))} + O(a^2),$$

so the estimator converges as $O(a^2)$ *exactly if the weights satisfy:*

$$w_2^{(1)} - w_5^{(1)} = 0, \tag{4.9}$$

$$(-5+2\sqrt{2})w_2^{(1)} + (4-4\sqrt{2})w_3^{(1)} + (3-2\sqrt{2})w_4^{(1)} + (-7+6\sqrt{2})w_5^{(1)} = 0. \tag{4.10}$$

If these equations are satisfied, the bias is

$$a^2\left(c_6^{(1)}\gamma\mathbb{E}(r^{-1}) + c_7^{(1)}\frac{\gamma^2}{\pi}\mathbb{E}V_1(B(r)) + c_8^{(1)}\left(\frac{\gamma}{\pi}\mathbb{E}V_1(B(r))\right)^3\right) + O(a^3). \tag{4.11}$$

The first condition (4.7) is intuitive, since lattice cells of type η_1 and η_6 will typically not contain any boundary points. Equation (4.9) is also natural since it is exactly the condition (2.2), saying that interchanging foreground and background should not change the estimate. Equation (4.8) is not so obvious. The coefficient in front of $w_j^{(1)}$ in $\frac{1}{8}c_3^{(1)}$ is the asymptotic probability that a lattice square containing a piece of the boundary is of type η_j. Equation (4.10) does not seem to have a simple geometric interpretation. While (4.8) and (4.9) generalize to the design-based setting, see Section 7 and 8, (4.10) seems to be special for the Boolean model and the underlying distribution.

Equations (4.7), (4.8), (4.9), and (4.10) do not determine the weights uniquely. There is still one degree of freedom in the choice. However, this is not enough to remove the a^2 term in (4.11), since the system of linear equations the weights must satisfy becomes overdetermined. The following proposition gives the best possible choice of weights.

Proposition 4.2. *The complete solution to the system of linear equations (4.7), (4.8), (4.9), and (4.10) is*

$$w^{(1)} = \frac{\pi}{16}(0, 1+\sqrt{2}, \sqrt{2}, 12+8\sqrt{2}, 1+\sqrt{2}, 0) + w(0, 1, -\sqrt{2}, -4-4\sqrt{2}, 1, 0),$$

where $w \in \mathbb{R}$ *is arbitrary.*

In general, the best choice of w depends on the intensity γ and the grain distribution $\mathbb{Q}$. Note that negative weights are allowed, even though this does not have an intuitive geometric interpretation.

Finally, for the Euler characteristic, comparing $\mathbb{E}\hat{V}_0(\Xi)$ with (4.6) yields the following result.

https://doi.org/10.1239/aap/1396360102 Published online by Cambridge University Press

Theorem 4.2. *The limit* $\lim_{a\to 0} \mathbb{E}\hat{V}_0(\Xi)$ *exists if and only if*

$$w_1^{(0)} = w_6^{(0)} = 0, \tag{4.12}$$

$$(2-\sqrt{2})w_2^{(0)} + (-2+2\sqrt{2})w_3^{(0)} + (2-\sqrt{2})w_5^{(0)} = 0. \tag{4.13}$$

In this case,

$$\lim_{a\to 0} \mathbb{E}\hat{V}_0(\Xi) = \left(c_4^{(0)}\gamma + c_5^{(0)}\left(\frac{\gamma}{\pi}\mathbb{E}V_1(B(r))\right)^2\right)\mathrm{e}^{-\gamma\mathbb{E}V_2(B(r))},$$

so $\hat{V}_0$ *is asymptotically unbiased if and only if the following two equations are satisfied:*

$$2w_2^{(0)} - 2w_5^{(0)} = 1, \tag{4.14}$$

$$(-5+2\sqrt{2})w_2^{(0)} + (4-4\sqrt{2})w_3^{(0)} + (3-2\sqrt{2})w_4^{(0)} + (-7+6\sqrt{2})w_5^{(0)} = -\frac{\pi}{4}. \tag{4.15}$$

If these equations are satisfied, the bias is

$$a\left(c_6^{(0)}\gamma\mathbb{E}(r^{-1}) + c_7^{(0)}\frac{\gamma^2}{\pi}\mathbb{E}V_1(B(r)) + c_8^{(0)}\left(\frac{\gamma}{\pi}\mathbb{E}V_1(B(r))\right)^3\right) + O(a^2). \tag{4.16}$$

The best possible weights are given by the following result.

Proposition 4.3. *The general solution to the linear equations (4.12), (4.13), (4.14), and (4.15) is*

$$w^{(0)} = \left(0, \frac{1}{2}, -\frac{1}{2\sqrt{2}}, \left(\frac{3}{4} + \frac{1}{\sqrt{2}}\right)(2-\pi), 0, 0\right) + w(0, 1, -\sqrt{2}, -4-4\sqrt{2}, 1, 0),$$

with $w \in \mathbb{R}$ *arbitrary.*

Also here there is one degree of freedom in the choice of weights, which is not enough to annihilate the leading term of (4.16).

Again, (4.12), (4.13), and (4.14) are geometric in the sense that they also show up in the design-based setting, while (4.15) seems to be special for the Boolean model.

Note that $\hat{V}_0$ does not satisfy (2.2), not even asymptotically. For weights satisfying (4.12), we have

$$\hat{V}_0(\Xi) = w_2^{(0)}N_2(\Xi) + w_3^{(0)}N_3(\Xi) + w_4^{(0)}N_4(\Xi) + w_5^{(0)}N_5(\Xi),$$
$$\hat{V}_0(\mathbb{R}^2\setminus\Xi) = w_2^{(0)}N_5(\Xi) + w_3^{(0)}N_3(\Xi) + w_4^{(0)}N_4(\Xi) + w_5^{(0)}N_2(\Xi).$$

Under condition (2.2), we would thus have

$$\begin{aligned}
2\overline{V}_0(\Xi) &= \lim_{a\to 0}(\mathbb{E}\hat{V}_0(\Xi) - \mathbb{E}\hat{V}_0(\mathbb{R}^2\setminus\Xi)) \\
&= \lim_{a\to 0} a^{-2}(w_2^{(0)} - w_5^{(0)})\mathbb{E}(N_2 - N_5) \\
&= (w_2^{(0)} - w_5^{(0)})\left(4\gamma + 4(2-4\sqrt{2})\left(\frac{\gamma}{\pi}\mathbb{E}V_1(B(r))\right)^2\right)\mathrm{e}^{-\gamma\mathbb{E}V_2(B(r))},
\end{aligned}$$

which no choice of weights can satisfy by (4.6).

Equations (4.10) and (4.15) become more important compared to (4.9) and (4.14) when r and γ are large. These are the only equations involving the configuration η_4, which can occur only where two different balls are close.

5. General Boolean models

The case where the grains are random balls generalizes to Boolean models and where the isotropic grain distribution satisfies the following extra condition: there is an $\varepsilon > 0$ such that, for almost all grains K, $B(\varepsilon)$ slides freely inside K, i.e.

$$\text{for all } x \in \partial K : x - \varepsilon n(x) + B(\varepsilon) \subseteq K. \tag{5.1}$$

Here $n(x)$ denotes the (necessarily unique) outward pointing unit normal vector at x. Condition (5.1) is a generalization of the assumption $r \geq \varepsilon$ a.s. in Section 4.

First we need a version of (4.1) for grains satisfying (5.1). In the following, $[x, y]$ denotes the closed line segment between $x, y \in \mathbb{R}^2$.

Lemma 5.1. *Let S be a finite set with diameter* $\operatorname{diam} S \leq 2\varepsilon$. *Let K be a convex set satisfying (5.1). Then*

$$V_2(\operatorname{conv} S \oplus K) - V_2(S \oplus K) \leq V_2(\operatorname{conv} S \oplus B(\varepsilon)) - V_2(S \oplus B(\varepsilon)).$$

Proof. After a translation, we may assume that $B(\varepsilon) \subseteq K$. Hence,

$$\operatorname{conv} S \subseteq S \oplus B(\varepsilon) \subseteq S \oplus K.$$

Let F_i, $i \in I$, be the faces of conv S with outward pointing normal vectors u_i. Then we obtain

$$(\operatorname{conv} S \oplus K) \setminus (S \oplus K) = (\operatorname{conv} S \oplus K) \cap (\operatorname{conv} S)^c \setminus (S \oplus K) = \bigcup_{i \in I} (F_i \oplus K^+_{u_i}) \setminus (S \oplus K), \tag{5.2}$$

where $K^+_u = \{z \in K \mid \langle z, u \rangle \geq 0\}$. To show the inclusion $\subseteq$ in the second equality, suppose that $s \in \operatorname{conv} S$ and $c \in K$ with $s + c \notin \operatorname{conv} S$. Then there is a maximal $\lambda \in [0, 1)$ such that $s + \lambda c = f$, where $f \in \partial \operatorname{conv} S$. But if $f \in F_i \setminus S$, then $\langle c, u_i \rangle \geq 0$ and, hence, $s + c = f + (1 - \lambda)c$ belongs to $F_i \oplus K^+_{u_i}$. If $f \in S$, then $s + c \in S \oplus C$.

Let F_i be given and write $u = u_i$. After a translation we may assume that $F_i = [0, x]$ with $x \in B(2\varepsilon)$. Let

$$y \in ([0, x] \oplus K^+_u) \setminus (S \oplus K).$$

Let $l_y = y + \operatorname{span}\{x\}$ be the line parallel to $[0, x]$ containing y. Since K^+_u is convex and $y - \lambda x \in K^+_u$ for some $\lambda \in (0, 1)$, $l_y \cap K^+_u$ is a nonempty line segment $[c_1, c_2]$. Then we have

$$\begin{gathered} y \in l_y \cap ([0, x] \oplus K^+_u) = [c_1, x + c_2], \\ y \notin l_y \cap (\{0, x\} \oplus K^+_u) = [c_1, c_2] \cup [c_1 + x, c_2 + x]. \end{gathered} \tag{5.3}$$

Choose $z \in K^+_u$ such that $n(z) = u$ and let $w = z - \varepsilon u \in K^+_u$ be the center of the touching ball guaranteed by (5.1).

By convexity, $[0, w] \oplus B(\varepsilon) \subseteq C$, so $l_y \cap [0, w] \neq \varnothing$ would imply that

$$|c_1 - c_2| \geq 2\varepsilon \geq |x|,$$

contradicting (5.3). Thus, $\langle w, u \rangle \leq \langle y, u \rangle \leq \langle z, u \rangle$; hence,

$$\varnothing \neq l_y \cap [w, z] \subseteq l_y \cap (w + B(\varepsilon)^+_u) \subseteq [c_1, c_2],$$

showing that

$$y \in ([0, x] \oplus (w + B(\varepsilon)^+_u)) \setminus (S \oplus K) \subseteq ([0, x] \oplus (w + B(\varepsilon)^+_u)) \setminus (S \oplus (w + B(\varepsilon))).$$

Now we may compute

$$\begin{aligned} V_2((\text{conv}\, S \oplus K) \setminus (S \oplus K)) &\le \sum_{i \in I} V_2((F_i \oplus K_{u_i}^+) \setminus (S \oplus K)) \\ &\le \sum_{i \in I} V_2((F_i \oplus B(\varepsilon)_{u_i}^+) \setminus (S \oplus B(\varepsilon))) \\ &= V_2((\text{conv}\, S \oplus B(\varepsilon)) \setminus (S \oplus B(\varepsilon))), \end{aligned}$$

where the last equality uses the fact that, when $K = B(\varepsilon)$, the union in (5.2) is disjoint, since

$$(F_i \oplus B(\varepsilon)_{u_i}^+) \setminus (S \oplus B(\varepsilon)) \subseteq F_i \oplus [0, \varepsilon u_i].$$

Now let ξ_l be a configuration and write $F_l = \text{conv}(\xi_l)$. Then Lemma 5.1 implies the following result.

Corollary 5.1. *Let Ξ be a Boolean model such that, for some $\varepsilon > 0$, the grains satisfy (5.1) a.s. For $\sqrt{2}a < \varepsilon$ and $l = 0, \dots, 15$,*

$$\mathbb{E}V_2(F_l \oplus K) - \mathbb{E}V_2(\xi_l \oplus K) \le a^3 \varepsilon^{-1} V_1^{(3)}(a^{-1}\xi_l) + O(a^5).$$

This allows us to compute $\mathbb{P}(\xi_l \subseteq \mathbb{R}^2 \setminus \Xi)$ using (3.8) and (3.9), but only up to second order, i.e.

$$\begin{aligned} \mathbb{P}(\xi_l \subseteq \mathbb{R}^2 \setminus \Xi) &= \exp\Big(-\gamma\Big(\mathbb{E}V_2(K) + a\frac{2}{\pi}V_1(F_l)\mathbb{E}V_1(K) + a^2 V_2(F_l) + O(a^3)\Big)\Big) \qquad (5.4) \\ &= c_1 + \mathrm{e}^{-\gamma \mathbb{E}V_2(K)}\Big(c_2 + ac_3\frac{\gamma}{\pi}\mathbb{E}V_1(K) + a^2\Big(c_4\gamma + c_5\Big(\frac{\gamma}{\pi}EV_1(K)\Big)^2\Big)\Big) \\ &\quad + O(a^3), \end{aligned}$$

with the same constants c_m as in Section 4, since these depend only on $V_i(a^{-1}F_l)$.

Furthermore, the specific intrinsic volumes were given by (4.4)–(4.6), so, by exactly the same arguments as in Section 4, we obtain the following result.

Theorem 5.1. *Theorems 4.1 and 4.2, except for (4.11) and (4.16), also hold for an isotropic Boolean model with grains satisfying (5.1) a.s.*

Remark 5.1. The term $O(a^3)$ in (5.4) is of the form

$$a^3\Big(c_7\frac{\gamma^2}{\pi}\mathbb{E}V_1(K) + c_8\Big(\frac{\gamma}{\pi}\mathbb{E}V_1(K)\Big)^3\Big) + \gamma\phi(a) + O(a^4),$$

where c_7 and c_8 are as in (4.2) and $0 \le \phi(a) \le c_6\varepsilon^{-1}a^3$ with c_6 as in (4.2).

6. Generalization to standard random sets

As an easy consequence of the well-known results obtained in [8], the first-order results for Boolean models generalize further to isotropic standard random sets. A standard random set Z is a stationary random closed set, such that the realizations are almost all locally polyconvex and Z satisfies the integrability condition

$$\mathbb{E}2^{N(Z \cap B(1))} < \infty,$$

where $N(Z \cap B(1))$ is the minimal number n such that $Z \cap B(1)$ is a union of n convex sets; see also [14, Definition 9.2.1].

The specific intrinsic volumes of a standard random set are defined as in (3.1) and we estimate $\overline{V}_1$ by

$$\hat{V}_1(Z) = a^{-1}\sum_{j=1}^{6} w_j^{(1)}\frac{N_j}{N_0},$$

as in (3.3), where N_j are as in (3.4). Since lower-dimensional parts of Z are usually invisible in the digitization, we assume that Z is a.s. topologically regular.

Theorem 6.1. *Let Z be an isotropic standard random set in the plane, which is a.s. topologically regular. Then $\lim_{a\to 0}\mathbb{E}\hat{V}_1(Z)$ exists if and only if $w_1^{(1)} = w_6^{(1)}$. In this case,*

$$\lim_{a\to 0}\mathbb{E}\hat{V}_1(Z) = \frac{1}{\pi}c_3^{(1)}\overline{V}_1(Z),$$

with $c_3^{(1)}$ as in (4.3). In particular, $\hat{V}_1(Z)$ is asymptotically unbiased exactly if (4.8) holds.

Proof. As in the case of the Boolean model,

$$\mathbb{E}\hat{V}_1(Z) = a^{-1}\sum_{j=1}^{6} w_j^{(1)}\mathbb{P}(Z\cap aC_0 \in \eta_j).$$

First, let ξ_l, $l \neq 0, 15$, be a configuration with $B_l, W_l \neq \emptyset$. Define the support function of a set A by $h(A, n) = \sup\{\langle x, n\rangle \mid x \in A\}$, for $n \in S^1$ and where $\langle\cdot,\cdot\rangle$ is the standard Euclidean inner product. The following formula is shown in [8, Theorem 4]:

$$\lim_{a\to 0} a^{-1}\mathbb{P}(B_l \subseteq Z, W_l \subseteq Z^c) = \int_{S^1}(-h(B_l \oplus \check{W}_l), n)^+ \bar{L}(\mathrm{d}n).$$

Here $x^+ = \max\{x, 0\}$ and $\bar{L}$ is the mean normal measure on S^1, i.e.

$$\bar{L}(A) = \lim_{r\to\infty}\frac{\mathbb{E}S_1(Z\cap B(r); A)}{V_2(B(r))}, \qquad A \in \mathcal{B}(S^1),$$

where $S_1(K;\cdot)$ is the first area measure defined for K polyconvex (see [13, Chapter 4]). In particular, the total measure $\bar{L}(S^1)$ is $2\overline{V}_1(Z)$.

By the isotropy of Z, $\bar{L}$ is rotation invariant, so Tonelli's theorem yields

$$\begin{aligned}\lim_{a\to 0} a^{-1}\mathbb{P}(B_l \subseteq Z, W_l \subseteq Z^c) &= \int_{S^1}(-h(B_l \oplus \check{W}_l, n))^+ \bar{L}(\mathrm{d}n)\\ &= \frac{1}{2\pi}\int_0^{2\pi}\int_{S^1}(-h(B_l \oplus \check{W}_l, R_{-v}n))^+ \bar{L}(\mathrm{d}n)\,\mathrm{d}v\\ &= \frac{1}{2\pi}\int_{S^1}\int_0^{2\pi}(-h(B_l \oplus \check{W}_l, u_v))^+\,\mathrm{d}v\,\mathrm{d}\bar{L},\end{aligned}$$

where $u_v = (\cos v, \sin v)$. The inner integral depends only on the equivalence class η_j containing ξ_l. Thus, we need to compute it for only one representative ξ_{l_j} of each η_j, i.e.

$$\begin{aligned}(-h(B_1 \oplus \check{W}_1, u_v))^+ &= (-h(B_7 \oplus \check{W}_7, u_v))^+ = \max\{|\cos v|, |\sin v|\}\,\mathbf{1}_{v\in[0,\pi/2]},\\ (-h(B_3 \oplus \check{W}_3, u_v))^+ &= (\max\{|\cos v|, |\sin v|\} - \min\{|\cos v|, |\sin v|\})\,\mathbf{1}_{v\in[\pi/4,3\pi/4]},\\ (-h(B_6 \oplus \check{W}_6, u_v))^+ &= 0.\end{aligned}$$

https://doi.org/10.1239/aap/1396360102 Published online by Cambridge University Press

A direct computation now shows that

$$\lim_{a\to 0} a \sum_{j=2}^{5} w_j^{(1)} \mathbb{E}N_j = \sum_{j=2}^{5} w_j^{(1)} d_j \frac{1}{2\pi} \int_{S^1} \int_0^{2\pi} (-h(B_{l_j} \oplus \check{W}_{l_j}, u_v))^+ \, \mathrm{d}v \, \mathrm{d}\bar{L} = \frac{1}{\pi} c_3^{(1)} \overline{V}_1(Z).$$

Finally, it is well known that

$$\lim_{a\to 0} \mathbb{P}(Z \cap aC_0 \in \eta_6) = \overline{V}_2(Z), \qquad \lim_{a\to 0} \mathbb{P}(Z \cap aC_0 \in \eta_1) = 1 - \overline{V}_2(Z),$$

so we must choose $w_1^{(1)} = w_6^{(1)} = 0$ in order for $\lim_{a\to 0} \mathbb{E}\hat{V}_1(Z)$ to exist for all Z.

7. Boundary length in the design-based setting

Instead of considering random sets observed on a fixed lattice, we now turn to the design-based setting where we sample a deterministic compact set $X \subseteq \mathbb{R}^2$ with a stationary isotropic random lattice, by which we mean that $\mathbb{L}$ is the random set $\mathbb{L}(c, v) = R_v(\mathbb{Z}^2 + c)$, where $v \in [0, 2\pi]$ and $c \in C$ are mutually independent uniform random variables.

We first consider estimators for the boundary length $2V_1$, as this is a fairly easy consequence of [7, Theorem 5]. Based on the random set $X \cap a\mathbb{L}$, we consider an estimator of the form

$$\hat{V}_1(X) = a \sum_{j=1}^{6} w_j^{(1)} N_j(X \cap a\mathbb{L}),$$

as described in Section 2 and study the asymptotic behaviour of $\mathbb{E}\hat{V}_1(X)$.

We first need some conditions on X. A compact set $X \subseteq \mathbb{R}^2$ is called gentle, see [7], if the following two conditions hold.

(i) $\mathcal{H}^1(\mathcal{N}(\partial X)) < \infty$.

(ii) For $\mathcal{H}^1$-almost all $x \in \partial X$, there exist two balls B_i and B_o with nonempty interior, both containing x, and such that $B_i \subseteq X$ and $\operatorname{int}(B_o) \subseteq \mathbb{R}^2 \setminus X$.

Here and in the following $\mathcal{H}^d$ denotes the d-dimensional Hausdorff measure and $\mathcal{N}(\partial X)$ is the reduced normal bundle

$$\mathcal{N}(\partial X) = \{(x, n) \in \partial X \times S^1 \mid \text{ there exists } t > 0\colon \text{for all } y \in \partial X\colon |tn| < |tn + x - y|\}.$$

Theorem 7.1. *Let $X \subseteq \mathbb{R}^2$ be a compact gentle set and $\mathbb{L}$ a stationary isotropic random lattice. Then $\lim_{a\to 0} \mathbb{E}\hat{V}_1(X)$ exists if and only if $w_6^{(1)} = w_1^{(1)} = 0$. In this case,*

$$\lim_{a\to 0} \mathbb{E}\hat{V}_1(X) = \frac{1}{\pi} c_3^{(1)} V_1(X),$$

with $c_3^{(1)}$ as in (4.3). In particular, $\hat{V}_1(X)$ is asymptotically unbiased if and only if $w^{(1)}$ satisfies (4.8).

In Section 8 we shall see that under stronger conditions on X, the convergence is actually $O(a)$ and the weights can be chosen so that it is even $O(a^2)$.

Theorem 5 of [7] is shown only for a uniformly translated lattice, whereas we assume isotropy as well. Thus, we need the following lemma.

https://doi.org/10.1239/aap/1396360102 Published online by Cambridge University Press

Lemma 7.1. *For any compact gentle set X there is an $\varepsilon > 0$ such that, for any square lattice $\mathbb{L}$ with unit grid distance,*

$$N_j(X \cap a\mathbb{L}) \leq a^{-1}(1 + 4\sqrt{2}V_1(X))$$

for all $a < \varepsilon$ and $j = 2, \dots, 5$.

Proof. If $(z + aR_vC_0) \cap \partial X$ is a configuration of type $j \neq 1, 6$, for some $z \in a\mathbb{L}$, then $(z + aR_vC) \cap \partial X \neq \varnothing$; hence, $z + aR_vC \subseteq \partial X \oplus B(\sqrt{2}a)$. Thus,

$$N_j(X \cap a\mathbb{L}) \leq a^{-2}V_2(\partial X \oplus B(\sqrt{2}a)).$$

Now, [7, Theorem 1] with $P = B(\sqrt{2}a)$ and $Q = B(ar)$ shows that

$$\lim_{a \to 0} a^{-1}V_2(X \oplus B(\sqrt{2}a) \setminus X \ominus B(ar)) = (\sqrt{2} + r)2V_1(\partial X).$$

Letting $r = \sqrt{2} \pm \varepsilon$ for $\varepsilon \to 0$ yields

$$\lim_{a \to 0} a^{-1}V_2(\partial X \oplus B(\sqrt{2}a)) = 4\sqrt{2}V_1(X).$$

In particular, $a^{-1}V_2(\partial X \oplus B(\sqrt{2}a)) - 4\sqrt{2}V_1(X) \leq 1$, for all a sufficiently small.

Proof of Theorem 7.1. Since X is compact, N_1 is infinite, so $w_1^{(1)}$ must equal zero in order for the estimator to be well defined. Moreover, $\lim_{a\to 0} a^2N_6 = V_2(X)$. Thus, aN_6 diverges when $a \to 0$, while all other aN_j remain bounded by Lemma 7.1. Hence, $w_6^{(1)} = 0$ is necessary for $\lim_{a\to 0} \mathbb{E}\hat{V}_1(X)$ to exist.

By Lemma 7.1, $aN_l(X \cap a\mathbb{L}(v, c))$ is uniformly bounded, so, using the Lebesgue theorem of dominated convergence, we obtain

$$\begin{aligned}
\lim_{a \to 0} a\mathbb{E}N_l(X \cap a\mathbb{L}(v, c)) &= \lim_{a \to 0} a\frac{1}{2\pi}\int_0^{2\pi}\int_C N_l(X \cap a\mathbb{L}(v, c))\, \mathrm{d}c\, \mathrm{d}v \\
&= \frac{1}{2\pi}\int_0^{2\pi}\lim_{a \to 0} a\int_C N_l(X \cap a\mathbb{L}(v, c))\, \mathrm{d}c\, \mathrm{d}v \\
&= \frac{1}{2\pi}\int_{S^1}\int_0^{2\pi}(-h(R_v(B_l) \oplus R_v(\check{W}_l), n))^+\, \mathrm{d}v S_1(X; \mathrm{d}n) \\
&= \frac{1}{2\pi}\int_{S^1}\int_0^{2\pi}(-h(B_l \oplus \check{W}_l, R_{-v}n))^+\, \mathrm{d}v S_1(X; \mathrm{d}n),
\end{aligned}$$

where the third equality is [7, Theorem 5]. The remaining computations are as in the proof of Theorem 6.1, since $S_1(X; S^1) = 2V_1(X)$.

Note how the isotropy of the lattice was crucial in the proof. This corresponds to the isotropy requirement for the Boolean model.

8. Euler characteristic in the design-based setting

We remain in the design-based setting of Section 7 and consider the estimation of the Euler characteristic and the higher-order behaviour of boundary length estimators. For this, we need some stronger boundary conditions on X. For instance, Jürgen Kampf has shown in a yet unpublished paper (see [5]) that without the isotropy of the lattice, there are no local estimators for V_0 that are asymptotically unbiased for all polyconvex sets. On the other hand, it is well known that there exists a local algorithm for V_0, which is asymptotically unbiased on the class of so-called r-regular sets; see, for example, the discussion in [15]. We shall assume throughout this section that X is a compact full-dimensional C^2 manifold, which is slightly stronger than r-regularity.

The estimator for the Euler characteristic was defined in Section 2 as

$$\hat{V}_0(X) = \sum_{j=1}^{6} w_j^{(0)} N_j(X \cap a\mathbb{L}).$$

Note that $\hat{V}_1(X) = a\hat{V}_0(X)$ if $w_j^{(1)} = w_j^{(0)}$. To treat both estimators, we sometimes just write $w_j^{(i)}$ for the weights. As noted in Section 7, we must choose $w_1^{(i)} = 0$ in order for $\hat{V}_i$ to be well defined and $w_6^{(i)} = 0$ to make $a^{1-i}\mathbb{E}\hat{V}_i(X)$ asymptotically convergent. Hence, we assume $w_1^{(i)} = w_6^{(i)} = 0$ throughout this section.

We now present our main result.

Theorem 8.1. *Assume that $X \subseteq \mathbb{R}^2$ is a compact 2-dimensional C^2 submanifold with boundary. Then*

$$\lim_{a\to 0}\Big(\mathbb{E}\hat{V}_0(X) - a^{-1}\lim_{a\to 0} a\mathbb{E}\hat{V}_0(X)\Big) = c_4^{(0)} V_0(X),$$

with $c_4^{(0)}$ as in (4.3). Thus, $\lim_{a\to 0}\mathbb{E}\hat{V}_0(X)$ exists if and only if the weights satisfy (4.13) and $\hat{V}_0(X)$ is asymptotically unbiased if and only if (4.14) holds. In this case, $\mathbb{E}\hat{V}_0(X)$ satisfies (2.2) asymptotically.

Moreover, $\mathbb{E}\hat{V}_1(X)$ converges as $O(a)$, and if (4.9) is satisfied, even as $o(a)$. In this case, $\hat{V}_1(X)$ satisfies (2.1).

Theorem 8.1 generalizes (4.9) and (4.14) to the design-based setting. However, (4.10) and (4.15) do not appear. These involve the configuration η_4, which cannot occur when the boundary is C^2 and a is sufficiently small.

For the proof, we must compute

$$\sum_{j=2}^{5} w_j^{(i)}\mathbb{E}N_j = \sum_{j=2}^{5} w_j^{(i)} \frac{1}{2\pi}\int_0^{2\pi}\int_C N_j(X \cap a\mathbb{L}(c, v))\,\mathrm{d}c\,\mathrm{d}v.$$

We follow the same approach as in [7]. The idea is that

$$N_j(X \cap a\mathbb{L}(c, v)) = \sum_{l\colon \xi_l\in\eta_j}\sum_{z\in a\mathbb{L}(c,v)} \mathbf{1}_{\{z+aR_v(B_l)\subseteq X\}}\mathbf{1}_{\{z+aR_v(W_l)\subseteq \mathbb{R}^2\setminus X\}}.$$

Integrating over all $c \in C$, we obtain

$$\int_C N_j(X \cap a\mathbb{L}(c, v))\,\mathrm{d}c = a^{-2}\sum_{l\colon \xi_l\in\eta_j}\int_{\mathbb{R}^2} f_l(z, v)\mathcal{H}^2(\mathrm{d}z), \tag{8.1}$$

where f_l denotes the indicator function

$$f_l(z, v) = \mathbf{1}_{\{z+aR_v(B_l)\subseteq X\}}\mathbf{1}_{\{z+aR_v(W_l)\subseteq \mathbb{R}^2\setminus X\}}. \tag{8.2}$$

By the assumptions on X, there is a unique outward pointing normal vector $n(x)$ at x. Since ∂X is an embedded C^2 submanifold, the tubular neighbourhood theorem ensures that there is an $\varepsilon > 0$ such that all points in $\partial X \oplus B(\varepsilon)$ have a unique closest point in ∂X. For $\sqrt{2}a < \varepsilon$, the support of f_l is contained in $\partial X \oplus B(\varepsilon)$.

As in the proof of [7, Theorem 1], we apply [4, Theorem 2.1] to compute (8.1). In the case of C^2 manifolds, this reduces to the Weyl tube formula,

$$\int_{\mathbb{R}^2} f_l(z, v)\mathcal{H}^2(\mathrm{d}z) = \int_{\partial X}\int_{-\varepsilon}^{\varepsilon} t f_l(x + tn, v)k(x)\,\mathrm{d}t\,\mathcal{H}^1(\mathrm{d}x) + \int_{\partial X}\int_{-\varepsilon}^{\varepsilon} f_l(x + tn, v)\,\mathrm{d}t\,\mathcal{H}^1(\mathrm{d}x), \tag{8.3}$$

where $k(x)$ is the signed curvature at x.

The main part of the proof of Theorem 8.1 is contained in Lemmas 8.3 and 8.4, below, handling each of the two integrals in (8.3). Before proving these, we show two technical lemmas. The first is a standard differential geometric description of ∂X.

In the following, $\tau(x)$ denotes the unit tangent vector at x, chosen so that $\{\tau(x), n(x)\}$ are positively oriented.

Lemma 8.1. *Let $X \subseteq \mathbb{R}^2$ be a C^2 submanifold with boundary. For some $\delta < 0$, there is a well-defined C^1 function $l: [-2\delta, 2\delta] \times \partial X \to \mathbb{R}$, such that $l(r, x)$ is the signed length of the line segment parallel to $n(x)$ from $x + r\tau(x)$ to ∂X. The sign is chosen such that $x + r\tau(x) + l(r, x)n(x) \in \partial X$.*

The function $r^{-2}l(br, x)$ is bounded for $(b, r, x) \in [-2, 2] \times [-\delta, \delta] \setminus \{0\} \times \partial X$ and

$$\lim_{r\to 0} r^{-2}l(br, x) = -\tfrac{1}{2}b^2k(x).$$

Proof. By the assumptions on X, there are finitely many isometric C^2 parametrizations of the form $\alpha: (a - 2\mu, b + 2\mu) \to \partial X$ such that the sets $\alpha([a, b])$ cover ∂X. For any $t \in (a-2\mu, b+2\mu)$, we write $n(t) = n(\alpha(t))$ for short. There are unique functions $l, r: (-\mu, \mu) \times (a - \mu, b + \mu) \to \mathbb{R}$ such that, for any $(s, t) \in (-\mu, \mu) \times (a - \mu, b + \mu)$,

$$\alpha(s + t) - \alpha(t) = r(s, t)\alpha'(t) + l(s, t)n(t),$$

where

$$r(s, t) = \langle \alpha(s + t) - \alpha(t), \alpha'(t)\rangle, \qquad l(s, t) = \langle \alpha(s + t) - \alpha(t), n(t)\rangle.$$

In particular, note that both functions are C^1 and, as functions of s, they are even C^2. In an open neighbourhood of $[a, b] \times 0$, $(\partial/\partial s)r(s, t) > 0$. By the inverse function theorem applied to $(r(s, t), t)$, there is a δ such that the inverse $s(r, t)$ is defined and is C^1 on $(-3\delta, 3\delta) \times [a, b]$. In fact, $r \mapsto s(r, t)$ is C^2 as it is the inverse of $s \mapsto r(s, t)$. Then $l(s(r, t), t)$ is the distance from $\alpha(t) + r\alpha'(t)$ to $\alpha(s(r, t) + t)$. If $3\delta < \varepsilon$, this is the boundary point on the line parallel to $n(t)$ closest to $\alpha(t) + r\alpha'(t)$.

By the mean value theorem,

$$\begin{aligned}\frac{l(s(br, t), t)}{r} &= b\frac{\partial}{\partial s}l(s, t)\bigg|_{s=s(br_0,t)}\frac{\partial}{\partial r}s(r, t)\bigg|_{r=br_0},\\ \frac{l(s(br, t), t)}{r^2} &= b^2\frac{r_0}{r}\frac{\partial^2}{\partial s^2}l(s, t)\bigg|_{s=s(br_1,t)}\frac{\partial}{\partial r}s(r, t)\bigg|_{r=br_0}\frac{\partial}{\partial r}s(r, t)\bigg|_{r=br_1},\end{aligned} \tag{8.4}$$

for some $0 \le |r_1| \le |r_0| \le |r|$. The continuity of $(\partial/\partial s)l$, $(\partial^2/\partial s^2)l$, and $(\partial/\partial r)s$ on $[-2\delta, 2\delta] \times [a, b]$ implies that (8.4) is bounded on $[-2, 2] \times [-\delta, \delta] \setminus \{0\} \times [a, b]$.

Finally, since $l(s(0,t),t) = 0$ and $(\partial/\partial s) l(s,t)|_{s=0} = 0$, we obtain

$$\lim_{r\to 0} \frac{l(s(br,t),t)}{r} = \frac{\partial}{\partial r} l(s(br,x))\bigg|_{r=0} = 0,$$

$$\lim_{r\to 0} \frac{l(s(br,t),t)}{r^2} = \frac{1}{2}\frac{\partial^2}{\partial r^2} l(s(br,x))\bigg|_{r=0} = \frac{1}{2} b^2 \langle \alpha''(t), n(t)\rangle = -\frac{1}{2} b^2 k(\alpha(t)),$$

proving the last claim.

Before proving the next lemmas, we introduce some notation. Let $v \in [0, 2\pi]$ and $x \in \partial X$. Let $v_0, \dots, v_3$ be the elements of $R_v(C_0)$ ordered such that $s_i \geq s_{i+1}$, where $s_i = \langle v_i, n(x)\rangle$. Let $b_i = \langle v_i, \tau(x)\rangle$. Note that the ordering of the v_i depends only on $R_{-v} n \in S^1$, and that S^1 is divided into eight arcs of length $\pi/4$ on each of which the ordering of the $R_v(C_0)$ is constant as a function of $R_{-v} n \in S^1$. The s_i and b_i can be computed explicitly as a function of $R_{-v} n \in S^1$. Though used in the explicit calculations below, these values have been omitted.

Define

$$t_i = -a s_i + l(b_i a, x).$$

The t_i are constructed such that, for $t \in [-\varepsilon, \varepsilon]$,

$$x + t n(x) + a v_i \in X \quad \text{if and only if} \quad t \leq t_i. \tag{8.5}$$

Let t_i' be a reordering of the t_i such that $t_i' \leq t_{i+1}'$ and let v_i' be the corresponding ordering of the v_i. This ordering depends on both x, v, and a. Since t_i may not equal t_i', we need the following lemma, ensuring that this does not happen too often.

Lemma 8.2. *There is a constant M such that, for all $x \in \partial X$ and a sufficiently small,*

$$a^{-1}\mathcal{H}^1(v \in [0, 2\pi] \mid \text{there exists } i\colon v_i \neq v_i') \leq M.$$

Furthermore, there is a constant M' such that

$$|t_i - t_i'| \leq 4 \sup\{|l(ba, x)| \mid (b, x) \in [-\sqrt{2}, \sqrt{2}] \times \partial X\} \leq M' a^2.$$

Proof. Let $v \in [0, 2\pi]$ and $x \in \partial X$ is given. If $v_i \neq v_i'$ then, in particular, there is a $j_1 < j_2$ with $t_{j_1} > t_{j_2}$. But then

$$0 \leq t_{j_1} - t_{j_2} = a(s_{j_2} - s_{j_1}) + l(b_{j_1} a, x) - l(b_{j_2} a, x); \tag{8.6}$$

hence,

$$0 \leq a(s_{j_1} - s_{j_2}) \leq l(b_{j_1} a, x) - l(b_{j_2} a, x) \leq C a^2,$$

for some uniform constant C, according to Lemma 8.1.

But then

$$0 \leq \cos(\theta(x, v)) \leq \langle (v_{j_1} - v_{j_2}), n(x)\rangle \leq C a,$$

where $\theta(x, v)$ is the angle from $n(x)$ to $v_{j_1} - v_{j_2}$. Thus, $\theta(x, v) = \theta(x, 0) + v$ must lie in $\cos^{-1}([0, Ca])$. But

$$\mathcal{H}^1(v \in [0, 2\pi] \mid \theta(x, v) \in \cos^{-1}([0, Ca])) = \mathcal{H}^1(\cos^{-1}([0, Ca]) \cap [0, 2\pi]) \leq C' a,$$

and there are only six possible combinations of j_1 and j_2, so

$$a^{-1}\mathcal{H}^1(v \in [0, 2\pi] \mid \text{there exists } i\colon v_i \neq v_i') \leq a^{-1} 6 \mathcal{H}^1(\cos^{-1}([0, Ca]) \cap [0, 2\pi]) \leq 6C'.$$

https://doi.org/10.1239/aap/1396360102 Published online by Cambridge University Press

Suppose that $t_i < t_i' = t_j$. If $j < i$, the last claim of the lemma follows from Lemma 8.1 and (8.6) as $a(s_{j_2} - s_{j_1})$ is negative. If $i < j$, there must be a $k < i$ with $t_j < t_k$. Then

$$|t_i - t_i'| \le |t_i - t_k| + |t_k - t_j| \le 4 \sup\{|l(ba, x)| \mid (b, x) \in [-\sqrt{2}, \sqrt{2}] \times \partial X\},$$

by a double application of (8.6). The case $t_i > t_i'$ can be treated in a similar way.

We are now ready to state and prove our two main lemmas.

Lemma 8.3. *With f_l as in (8.2),*

$$\lim_{a\to 0} a^{-2} \sum_{l\,:\,\xi_l \in \eta_j} \frac{1}{2\pi} \int_0^{2\pi} \int_{\partial X} \int_{-\varepsilon}^{\varepsilon} t f_l(x + tn, v) k(x)\, \mathrm{d}t\, \mathcal{H}^1(\mathrm{d}x)\, \mathrm{d}v = \begin{cases} V_0(X), & j = 2, \\ 0, & j = 3, 4, \\ -V_0(X), & j = 5. \end{cases}$$

Proof. For $x \in \partial X$ fixed, let

$$I_j(x, v) = \sum_{l\,:\,\xi_l \in \eta_j} \int_{-\varepsilon}^{\varepsilon} t f_l(x + tn, v)\, \mathrm{d}t.$$

For $\sqrt{2}a < \varepsilon$, configurations of type η_4 can never occur, so $(x + tn + aR_v(C_0)) \cap X$ corresponds to a configuration of type η_1 for $t < t_3'$, η_2 for $t \in (t_2', t_3']$, η_3 for $t \in (t_1', t_2']$, η_5 for $t \in (t_0', t_1']$, and η_6 for $t \le t_0'$, according to (8.5).

As an example, consider the configuration type η_5. Then we get

$$I_5 = \int_{t_0'}^{t_1'} t\, \mathrm{d}t = \tfrac{1}{2}(t_1'^2 - t_0'^2).$$

By Fubini's theorem we must compute

$$\lim_{a\to 0} a^{-2} \int_{\partial X} \int_0^{2\pi} I_5\, \mathrm{d}v k\, \mathrm{d}\mathcal{H}^1 = \lim_{a\to 0} a^{-2} \int_{\partial X} \int_0^{2\pi} \tfrac{1}{2}(t_1'^2 - t_0'^2)\, \mathrm{d}v k\, \mathrm{d}\mathcal{H}^1.$$

By Lemma 8.2, $\lim_{a\to 0} \mathcal{H}^1(v \in [0, 2\pi] \mid t_i \ne t_i') = 0$ uniformly. Moreover, it follows from Lemma 8.1 that

$$a^{-2} t_i^2 = s_i^2 - 2 s_i a^{-1} l(b_i a, x) + a^{-2} l(b_i a, x)^2$$

is uniformly bounded. Hence, we may replace $t_i'^2$ by t_i^2 in the integral by the Lebesgue theorem of dominated convergence. This also applies to give

$$\begin{aligned} \lim_{a\to 0} a^{-2} \int_{\partial X} \int_0^{2\pi} I_5\, \mathrm{d}v k\, \mathrm{d}\mathcal{H}^1 &= \int_{\partial X} \int_0^{2\pi} \lim_{a\to 0} a^{-2} \tfrac{1}{2}(t_1^2 - t_0^2)\, \mathrm{d}v k\, \mathrm{d}\mathcal{H}^1 \\ &= \int_{\partial X} \int_0^{2\pi} \tfrac{1}{2}(s_1^2 - s_0^2)\, \mathrm{d}v k\, \mathrm{d}\mathcal{H}^1. \end{aligned}$$

The last step used Lemma 8.1.

Substituting $u = R_{-v} n$ and inserting the values of $s_i(u)$, a direct computation shows

$$\begin{aligned} \lim_{a\to 0} a^{-2} \int_{\partial X} \int_0^{2\pi} I_5(x, v)\, \mathrm{d}v k(x) \mathcal{H}^1(\mathrm{d}x) &= \int_{\partial X} \int_{S^1} \tfrac{1}{2}(s_1^2(u) - s_0^2(u))\, \mathrm{d}u k\, \mathrm{d}\mathcal{H}^1 \\ &= -2\pi V_0(X). \end{aligned}$$

The remaining configuration types η_2 and η_3 are treated similarly.

https://doi.org/10.1239/aap/1396360102 Published online by Cambridge University Press

Lemma 8.4. *For $w_j^{(i)} \in \mathbb{R}$ and $c_3^{(i)}$ as in (4.3), the limit*

$$\lim_{a\to 0} a^{-2}\frac{1}{2\pi}\bigg(\sum_{j=2}^{5} w_j^{(i)} \int_{\partial X}\int_0^{2\pi}\int_{-\varepsilon}^{\varepsilon}\sum_{l\colon \xi_l\in\eta_j} f_l(x+tn,v)\,\mathrm{d}t\,\mathrm{d}v\,\mathcal{H}^1(\mathrm{d}x) - 2ac_3^{(i)}V_1(X)\bigg)$$

exists and is equal to $(w_2^{(i)} - w_5^{(i)})V_0(X)$.

Proof. Let $x \in \partial X$ be given and define

$$I_j(x,v) = \sum_{l\colon \xi_l\in\eta_j}\int_{-\varepsilon}^{\varepsilon} f_l(x+tn,v)\,\mathrm{d}t.$$

By the same reasoning as in the proof of Lemma 8.3,

$$I_2 = t_3' - t_2', \qquad I_3 = t_2' - t_1', \qquad I_5 = t_1' - t_0'.$$

As an example, consider η_5. We shall compute

$$\begin{aligned}\lim_{a\to 0} a^{-2}&\int_{\partial X}\int_0^{2\pi}(I_5 + a(s_1 - s_0))\,\mathrm{d}v\,\mathrm{d}\mathcal{H}^1\\ &= \lim_{a\to 0}\int_{\partial X}\int_0^{2\pi}(a^{-2}(t_1'-t_0') + a^{-1}(s_1-s_0))\,\mathrm{d}v\,\mathrm{d}\mathcal{H}^1. \end{aligned}\tag{8.7}$$

Since $a^{-2}|t_i - t_i'| \le M'$ and $\mathcal{H}^1(t_i \ne t_i') < Ma$, by Lemma 8.2 for some uniform constants M and M', we may replace t_i by t_i' in (8.7).

By another application of Lemma 8.1, $a^{-2}t_i + a^{-1}s_i = a^{-2}l(b_i a, x)$ is uniformly bounded. This allows us to apply Lebesgue's theorem to (8.7). In the case of η_5, this yields

$$\begin{aligned}\lim_{a\to 0}\int_{\partial X}\int_0^{2\pi}&(a^{-2}I_5 + a^{-1}(s_1-s_0))\,\mathrm{d}v\,\mathrm{d}\mathcal{H}^1\\ &= \int_{\partial X}\int_0^{2\pi}\lim_{a\to 0}(a^{-2}(t_1'-t_0') + a^{-1}(s_1-s_0))\,\mathrm{d}v\,\mathrm{d}\mathcal{H}^1\\ &= \int_{\partial X}\int_0^{2\pi}\lim_{a\to 0} a^{-2}(l(ab_1,x) - l(ab_0,x))\,\mathrm{d}v\mathcal{H}^1(\mathrm{d}x)\\ &= \int_{\partial X}\int_0^{2\pi}\frac{-k}{2}(b_1^2 - b_0^2)\,\mathrm{d}v\,\mathrm{d}\mathcal{H}^1,\end{aligned}$$

where the last step also follows from Lemma 8.1.

Doing the same for the remaining configurations, a computation shows that

$$\begin{aligned}-\int_{\partial X}\int_0^{2\pi}&\frac{k}{2}(w_2^{(i)}(b_3^2-b_2^2) + w_3^{(i)}(b_2^2-b_1^2) + w_5^{(i)}(b_1^2-b_0^2))\,\mathrm{d}v\,\mathrm{d}\mathcal{H}^1\\ &= \lim_{a\to 0} a^{-2}\int_{\partial X}\int_0^{2\pi}\bigg(\sum_{j=2}^{5} w_j^{(i)} I_j - a(w_2^{(i)}(s_2-s_3) + w_3^{(i)}(s_1-s_2)\\ &\qquad\qquad + w_5^{(i)}(s_0-s_1))\bigg)\,\mathrm{d}v\,\mathrm{d}\mathcal{H}^1 \qquad (8.8)\\ &= \lim_{a\to 0} a^{-2}\bigg(\sum_{j=2}^{5} w_j^{(i)}\int_{\partial X}\int_0^{2\pi} I_j\,\mathrm{d}v\,\mathrm{d}\mathcal{H}^1 - 2ac_3^{(i)}V_1(X)\bigg).\end{aligned}$$

https://doi.org/10.1239/aap/1396360102 Published online by Cambridge University Press

On the other hand, another computation shows that (8.8) is equal to

$$-\int_{\partial X}\frac{k(x)}{2}(-2w_2^{(i)}+2w_5^{(i)})\mathcal{H}^1(\mathrm{d}x)=2\pi V_0(X)(w_2^{(i)}-w_5^{(i)}),$$

from which the claim follows.

Proof of Theorem 8.1. From Lemmas 8.3 and 8.4, it follows that the limit

$$\begin{aligned}\lim_{a\to 0}&\Big(a^{-i}\mathbb{E}\hat{V}_i(X)-a^{-1}\frac{1}{\pi}c_3^{(i)}V_1(X)\Big)\\ &=\lim_{a\to 0}a^{-2}\Big(\sum_{j=2}^{5}w_j^{(i)}\sum_{l\,:\,\xi_l\in\eta_j}\frac{1}{2\pi}\int_0^{2\pi}\Big(\int_{\partial X}\int_{-\varepsilon}^{\varepsilon}tf_l(x+tn,v)k(x)\,\mathrm{d}t\,\mathcal{H}^1(\mathrm{d}x)\\ &\qquad+\int_{\partial X}\int_{-\varepsilon}^{\varepsilon}f_l(x+tn,v)\,\mathrm{d}t\,\mathcal{H}^1(\mathrm{d}x)\Big)\,\mathrm{d}v\\ &\qquad-a\frac{1}{\pi}c_3^{(i)}V_1(X)\Big)\end{aligned}\tag{8.9}$$

exists and is equal to $c_4^{(i)}V_0(X)$.

In the limit, the condition (2.2) is

$$\lim_{a\to 0}\mathbb{E}\hat{V}_0(X)=\lim_{a\to 0}(w_2^{(0)}\mathbb{E}N_2(X)+w_3^{(0)}\mathbb{E}N_3(X)+w_5^{(0)}\mathbb{E}N_5(X))=V_0(X),$$

$$\lim_{a\to 0}\mathbb{E}\hat{V}_0(\mathbb{R}^2\setminus X)=\lim_{a\to 0}(w_2^{(0)}\mathbb{E}N_5(X)+w_3^{(0)}\mathbb{E}N_3(X)+w_5^{(0)}\mathbb{E}N_2(X))=-V_0(X).$$

This is equivalent to

$$\lim_{a\to 0}(w_2^{(0)}\mathbb{E}N_2+w_3^{(0)}\mathbb{E}N_3+w_5^{(0)}\mathbb{E}N_5)=V_0(X),$$

$$\lim_{a\to 0}(w_2^{(0)}-w_5^{(0)})(\mathbb{E}N_2-\mathbb{E}N_5)=2V_0(X).$$

From (8.9) with $w_2^{(0)}=1$, $w_3^{(0)}=w_4^{(0)}=0$, and $w_5^{(0)}=-1$, it follows that

$$\lim_{a\to 0}(\mathbb{E}N_2-\mathbb{E}N_5)=4V_0(X).$$

Thus, (4.14) ensures that (2.2) holds asymptotically.

When ∂X is actually a C^3 manifold, we can get slightly better asymptotic results, as we show in the following theorem.

Theorem 8.2. *Let $X\subseteq\mathbb{R}^2$ be a C^3 full-dimensional submanifold. Assume that the weights defining $\hat{V}_1(X)$ satisfy (4.8) and (4.9) and the weights defining $\hat{V}_0(X)$ satisfy (4.13) and (4.14). Then $\mathbb{E}\hat{V}_1(X)$ and $\mathbb{E}\hat{V}_0(X)$ converge as $O(a^2)$ and $O(a)$, respectively.*

Proof. It is enough to check that $a^{-i-1}(\mathbb{E}\hat{V}_i(X)-\lim_{a\to 0}\mathbb{E}\hat{V}_i(X))$ is bounded. Going through the proofs of Lemmas 8.3 and 8.4, we see that it is enough to show that

$$a^{-3}(t_{i+1}'^2-t_i'^2)-a^{-1}(s_{i+1}^2-s_i^2)\tag{8.10}$$

and

$$a^{-1}\int_0^{2\pi}\Big(a^{-2}(t_{i+1}'-t_i')-a^{-1}(s_i-s_{i+1})+\frac{k}{2}(b_{i+1}^2-b_i^2)\Big)\,\mathrm{d}v\tag{8.11}$$

are uniformly bounded.

The triangle inequality yields

$$|a^{-3}t_i'^2 - a^{-1}s_i^2| \le |a^{-3}t_i^2 - a^{-1}s_i^2| + a^{-3}|t_i'^2 - t_i^2|.$$

The terms

$$|a^{-3}t_i^2 - a^{-1}s_i^2| = |-2s_i a^{-2} l(b_i a, x) + a^{-3} l(b_i a, x)^2|$$

are uniformly bounded by Lemma 8.1. Furthermore,

$$\frac{|t_i'^2 - t_i^2|}{a^3} = \frac{|t_i' + t_i|}{a} \frac{|t_i' - t_i|}{a^2}$$

is bounded by Lemma 8.2. This takes care of (8.10).

Similarly,

$$\left|a^{-3}t_i' + a^{-2}s_i + a^{-1}\frac{k}{2}b_i^2\right| \le \left|a^{-3}t_i + a^{-2}s_i + a^{-1}\frac{k}{2}b_i^2\right| + a^{-3}|t_i - t_i'|.$$

Again, by Lemma 8.2, $a^{-2}|t_i - t_i'|$ is uniformly bounded by some C; hence,

$$\int_0^{2\pi} a^{-3}|t_i - t_i'| \, \mathrm{d}v \le \int_0^{2\pi} a^{-1} C \, \mathbf{1}_{\{t_i \ne t_i'\}} \, \mathrm{d}v$$

is also uniformly bounded by Lemma 8.2. Finally,

$$a^{-3}t_i + a^{-2}s_i + a^{-1}\frac{k}{2}b_i^2 = a^{-3}l(b_i a, x) + a^{-1}\frac{k}{2}b_i^2.$$

But, by a refinement of Lemma 8.1, $r \mapsto l(r, x)$ is C^3 when ∂X is a C^3 manifold and

$$\frac{l(br, x)}{r^3} + \frac{b^2 k(x)}{2r}$$

is bounded for $(b, r, x) \in [-\sqrt{2}, \sqrt{2}] \times [-\delta, \delta] \setminus \{0\} \times \partial X$. This takes care of (8.11).

9. Classical choices of weights

Recall that, for a stationary isotropic Boolean model Ξ with grain distribution satisfying (5.1) a.s., we found in Theorem 4.1 that

$$\lim_{a \to 0} \mathbb{E}\hat{V}_1(\Xi) = \frac{1}{\pi} c_3^{(1)} \overline{V}_1(\Xi).$$

If $c_3^{(1)} = \pi$, the bias for small values of a is approximately

$$\mathbb{E}\hat{V}_1(\Xi) - \overline{V}_1(\Xi) \approx a\left(c_4^{(1)}\gamma + c_5^{(1)}\left(\frac{\gamma}{\pi}\mathbb{E}V_1(C)\right)^2 \mathrm{e}^{-\gamma \mathbb{E}V_2(C)}\right),$$

with $c_m^{(1)}$ as in (4.3).

In the literature, various local algorithms are used for estimating the boundary length of a planar set. With the formulas above we can compute their asymptotic bias and thus compare their accuracy.

https://doi.org/10.1239/aap/1396360102 Published online by Cambridge University Press

Ohser and Mücklich [10] described an estimator for $\overline{V}_1(\Xi)$ based on a discretized version of the Cauchy projection formula. In the rotation invariant setting, the estimator corresponds to (3.3) with weights

$$w^{(1)} = \left(0, \frac{\pi}{16}\left(1+\frac{\sqrt{2}}{2}\right), \frac{\pi}{16}(1+\sqrt{2}), \frac{\pi}{8}, \frac{\pi}{16}\left(1+\frac{\sqrt{2}}{2}\right), 0\right).$$

Inserting these weights in the equations shows that this estimator satisfies (4.8) and is thus asymptotically unbiased. The weights also satisfy (4.9) but not (4.10). For small values of a, the error is approximately

$$-a\frac{1+\sqrt{2}}{2}\frac{\gamma^2}{\pi}\mathbb{E}V_1(C)^2\mathrm{e}^{-\gamma\mathbb{E}V_2(C)} \approx -1.207a\frac{\gamma^2}{\pi}\mathbb{E}V_1(C)^2\mathrm{e}^{-\gamma\mathbb{E}V_2(C)}.$$

One of the oldest algorithms for estimating the boundary length is suggested by Bieri and Nef [1]. The idea is to approximate the underlying object by a union of squares of side length a, centered at the foreground pixels, and use the boundary length of the approximation as an estimate. This corresponds to a local estimator with weights

$$w^{(1)} = \left(0, \tfrac{1}{2}, \tfrac{1}{2}, 1, \tfrac{1}{2}, 0\right).$$

However, it is well known that, for a compact object X, this is the boundary length of the smallest box containing X; hence, it is a very coarse estimate. The asymptotic mean is $(4/\pi)\overline{V}_1(X)$. Of course, we can correct for the factor $4/\pi$ and consider the weights

$$w^{(1)} = \left(0, \frac{\pi}{8}, \frac{\pi}{8}, \frac{\pi}{4}, \frac{\pi}{8}, 0\right) \tag{9.1}$$

instead. These weights can be justified by the Cauchy formula in [10] using $\theta_1 = \pi/2$. It is also the unique unbiased estimator where all weights are equal, except that configurations of type η_4 are counted with double weight. These weights satisfy (4.8) and (4.9) but not (4.10). The bias for small a is approximately

$$-a\frac{\gamma^2}{\pi}\mathbb{E}V_1(C)^2\mathrm{e}^{-\gamma\mathbb{E}V_2(C)}.$$

The approach of Dorst and Smeulders [2] is to reconstruct the underlying set by an 8-adjacency system and compute the length of the boundary of the reconstructed set, letting vertical and horizontal segments contribute with one weight, and diagonal segments with another weight. The resulting estimators are of the forms

$$w^{(1)} = \left(0, 0, \frac{\theta}{2}, \sqrt{2}\theta, \frac{\sqrt{2}\theta}{2}, 0\right), \qquad w^{(1)} = (0, 0, \alpha, 2\beta, \beta, 0). \tag{9.2}$$

These algorithms are tested only on straight lines in [2] and therefore it was not necessary to assign a value $w_4^{(1)}$. The weights chosen here are such that a diagonal segment coming from a configuration of type η_4 is counted twice.

Dorst and Smeulders [2] list some of the constants frequently used in the literature. The case $\theta = 1$ goes back to Freeman [3]. This yields a biased estimator. But even if the constants are chosen such that the estimator is asymptotically unbiased, all weights of this form have the disadvantage of not satisfying (4.9), which is the most desirable of the two equations (4.9) and (4.10), as it also appears in the design-based setting.

The boundary is also sometimes approximated using a 4- or 6-adjacency graph. However, the same problem with (4.9) arises.

Another classical approach is the marching squares algorithm. This is based on a reconstruction of both foreground and background. The boundary is then approximated by a digital curve lying between these, see e.g. [9, Figure 4.29]. The corresponding weights are

$$w^{(1)} = \left(0, \frac{\sqrt{2}}{4}, \frac{1}{2}, \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{4}, 0\right).$$

This estimator is not asymptotically unbiased either. In fact, the asymptotic mean is

$$(2\sqrt{2} - 2)\frac{4}{\pi}\overline{V}_1(\Xi) \approx 1.0548\overline{V}_1(\Xi).$$

Correcting for this factor, we obtain an asymptotically unbiased estimator satisfying (4.13) with approximate bias for small values of a, i.e.

$$a\frac{\sqrt{2} - 6}{4}\frac{\gamma^2}{\pi}\mathbb{E}V_1(C)^2 \mathrm{e}^{-\gamma \mathbb{E}V_2(C)} \approx -1.146a\frac{\gamma^2}{\pi}\mathbb{E}V_1(C)^2 \mathrm{e}^{-\gamma \mathbb{E}V_2(C)}.$$

Similarly, we can compare the classical estimators for V_0. Ohser and Mücklich [10] suggested an estimator based on the approximation of Ξ by a 6-neighbourhood graph. This results in weights

$$w^{(0)} = \left(0, \tfrac{1}{4}, 0, 0, -\tfrac{1}{4}, 0\right). \tag{9.3}$$

These satisfy (4.13) and (4.14), but not (4.15). Hence, it does not define an asymptotically unbiased estimator for Boolean models, but it does in the design-based setting of Section 8. For Boolean models, the asymptotic bias is

$$\lim_{a\to 0} \mathbb{E}\hat{V}_0 - \overline{V}_0 = \left(\frac{2 - 4\sqrt{2}}{\pi} + 1\right)\frac{\gamma^2}{\pi}\mathbb{E}V_1(C)^2 \mathrm{e}^{-\gamma \mathbb{E}V_2(C)} \approx -0.164\frac{\gamma^2}{\pi}\mathbb{E}V_1(C)^2 \mathrm{e}^{-\gamma \mathbb{E}V_2(C)}.$$

The estimator for the Euler characteristic suggested in [1] corresponds to the weights

$$w^{(0)} = \left(0, \tfrac{1}{4}, 0, -\tfrac{1}{2}, -\tfrac{1}{4}, 0\right).$$

The bias of this estimator is

$$\lim_{a\to 0} \mathbb{E}\hat{V}_0 - \overline{V}_0 = \left(\frac{-4}{\pi} + 1\right)\frac{\gamma^2}{\pi}\mathbb{E}V_1(C)^2 \mathrm{e}^{-\gamma \mathbb{E}V_2(C)} \approx -0.273\tfrac{\gamma^2}{\pi}\mathbb{E}V_1(C)^2 \mathrm{e}^{-\gamma \mathbb{E}V_2(C)},$$

which is slightly worse.

The conclusion is that for Boolean models, the best of the estimators for $\overline{V}_1$ and $\overline{V}_0$ listed here are (9.1) and (9.3), respectively. However, the weights in Propositions 4.2 and 4.3, respectively, give better estimators.

In the design-based setting, all of the classical algorithms listed here, except (9.2), are equally good when assessed by means of the results of the present paper.

Acknowledgements

The author is supported by the Centre for Stochastic Geometry and Advanced Bioimaging, funded by the Villum Foundation. The author is most grateful to Markus Kiderlen for giving me the ideas for this paper and for many helpful suggestions and final proofreading.

References

[1] Bieri, H. and Nef, W. (1984). Algorithms for the Euler characteristic and related additive functionals of digital objects. *Comput. Vision Graphics Image Process.* **28,** 166–175.

[2] Dorst, L. and Smeulders, A. W. M. (1987). Length estimators for digitized contours. *Comput. Vision Graphics Image Process.* **40,** 311–333.

[3] Freeman, H. (1970). Boundary encoding and processing. In *Picture Processing and Psychopictorics*, eds B. S. Lipkin and A. Rosenfeld, Academic Press, New York, pp. 241–266.

[4] Hug, D., Last, G. and Weil, W. (2004). A local Steiner-type formula for general closed sets and applications. *Math. Z.* **246,** 237–272.

[5] Kampf, J. (2012) A limitation of the estimation of intrinsic volumes via pixel configuration counts. WiMa Rep. 144, University of Kaiserslautern.

[6] Kampf, J. and Kiderlen, M. (2013). Large parallel volumes of finite and compact sets in d-dimensional Euclidean space. *Documenta Math.* **18,** 275–295.

[7] Kiderlen, M. and Rataj, J. (2006). On infinitesimal increase of volumes of morphological transforms. *Mathematika* **53,** 103–127.

[8] Kiderlen, M. and Vedel Jensen, E. B. (2003). Estimation of the directional measure of planar random sets by digitization. *Adv. Appl. Prob.* **35,** 583–602.

[9] Klette, R. and Rosenfeld, A. (2004). *Digital Geometry*. Morgan Kaufmann, San Fransisco, CA.

[10] Ohser, J. and Mücklich, F. (2000). *Statistical Analysis of Microstructures in Materials Science*. John Wiley, New York.

[11] Ohser, J., Nagel, W. and Schladitz, K. (2003). The Euler number of discretised sets—surprising results in three dimensions. *Image Anal. Stereol.* **22,** 11–19.

[12] Ohser, J., Nagel, W. and Schladitz, K. (2009). Miles formulae for Boolean models observed on lattices. *Image Anal. Stereol.* **28,** 77–92.

[13] Schneider, R. (1993). *Convex Bodies: The Brunn–Minkowski Theory* (Encyclopedia Math. Appl. **44**). Cambridge University Press.

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

[15] Stelldinger, P. and Köthe, U. (2005). Shape preserving digitization of binary images after blurring. In *Discrete Geometry for Computer Iimagery* (Lecture Notes Comput. Sci. **3429**), Springer, Berlin, pp. 383–391.