

# STEREOLOGICAL ESTIMATION OF SURFACE AREA FROM DIGITAL IMAGES

JOHANNA ZIEGEL<sup>1</sup> AND MARKUS KIDERLEN<sup>2</sup>

<sup>1</sup>Department of Mathematics and Statistics, The University of Melbourne, Victoria 3010, Australia; <sup>2</sup>Department of Mathematical Sciences, University of Aarhus, Ny Munkegade Build. 1530, DK-8000 Aarhus C, Denmark  
e-mail: jziegel@unimelb.edu.au, kiderlen@imf.au.dk  
(Accepted May 12, 2010)

## ABSTRACT

A sampling design of local stereology is combined with a method from digital stereology to yield a novel estimator of surface area based on counts of configurations observed in a digitization of an isotropic 2-dimensional slice with thickness  $s$ . As a tool, a result of the second author and J. Rataj on infinitesimal increase of volumes of morphological transforms is refined and used. The proposed surface area estimator is asymptotically unbiased in the case of sets contained in the ball centred at the origin with radius  $s$  and in the case of balls centred at the origin with unknown radius. For general shapes bounds for the asymptotic expected relative worst case error are given. A simulation example is discussed for surface area estimation based on  $2 \times 2 \times 2$ -configurations.

Keywords: configurations, digital stereology, local stereology, surface area.

## INTRODUCTION

Algorithms that determine the surface area of a three-dimensional object from a binary image can be grouped into global and local methods (Klette and Rosenfeld, 2004, p. 302). Global methods usually require a digital approximation of the surface or calculation of the surface normals. Local methods estimate the contributions to surface area of small voxel blocks in the binary image independently of the image outside the given block, and add up these contributions to estimate the total surface area. As local methods can be implemented using a linear filter, they are extremely efficient. Lindblad (2005) suggested a local estimator based on voxel blocks of size  $2 \times 2 \times 2$  inspired by the common marching cubes algorithm, but without an explicit reconstruction of the surface. There are  $2^8 = 256$  possible combinations of foreground and background voxels in a  $2 \times 2 \times 2$  block. They will be called configurations in the following. Two of these configurations contain only background or only foreground voxels and do therefore not contribute to the surface area. The remaining configurations, called boundary configurations, can be grouped into 14 classes, due to symmetry. Thus, local estimators of the surface of a set  $X$  in three-dimensional space are of the form

$$\hat{S}(X) = \sum_{i=1}^{254} \lambda_i N_i, \quad (1)$$

where  $\lambda_i$  is the contribution and  $N_i$  is the total number of occurrences of configuration  $i$ . Several natural

choices of the weights  $\lambda_1, \dots, \lambda_{254}$  can be found in the literature (Lindblad and Nyström, 2002; Lindblad, 2005; Schladitz *et al.*, 2006; Ziegel and Kiderlen, 2010).

Assuming that the set  $X$  was randomly translated before digitizing, Ziegel and Kiderlen (2010) found optimal weights in the sense that the asymptotic average error is minimized among all estimators of the form Eq. 1; see also Gutkowski *et al.* (2004) for a weaker result. Both papers are based on an asymptotic formula (Kiderlen and Rataj, 2006), where asymptotic refers to increasing resolution of the image. Ziegel and Kiderlen (2010) have shown that only 102 of the 254 configuration types can arise in the case where  $X$  is bounded by a planar surface, and it is only these that contribute to the surface area measure asymptotically; see also (Lindblad, 2005). These 102 configurations are therefore called *informative* configurations.

The estimator Eq. 1 can in principle also be applied to stacks of binary images of horizontal planar sections of  $X$ , where configurations are composed of voxels in two neighbouring section planes. A possible application is the estimation of the surface area of particles (e.g. cells) in confocal microscopy, where a stack of focal planes is digitized and analysed. However, due to optical effects, section planes that are hitting a particle close to one of the two horizontal touching planes yield only an extremely blurred image of the section profile. In traditional (continuous) stereology, this problem is solved by considering a random isotropic slice centred at a reference point of the particle under consideration. If the thickness  $s > 0$

of the slice is chosen small enough compared to the size of the particle, the blurring effect will not be expressed within the slice. An unbiased surface area estimator of Horvitz-Thompson type is then obtained by weighting the surface area contribution of each infinitesimal surface patch in the slice with its inverse sampling probability. This sampling probability is a function of the distance of the patch from the reference point, and depends on the thickness  $2s$  of the slice (Jensen, 1998).

It is the purpose of this paper to show that the two concepts from digital and local stereology described above can be combined. We will obtain a surface area estimator based on a stack of planar parallel digital images in an isotropic local slice, where the planes of the batch are parallel to the slice. It is intuitively clear that an estimator will no longer be a weighted sum of configuration counts like in Eq. 1. Instead, each observed configuration must additionally be weighted according to its inverse sampling probability. However, one cannot expect to obtain an unbiased estimator this way. We will propose a set of weight functions such that the surface area estimator is asymptotically unbiased in the cases where  $X$  is a ball centred at the reference point or  $X$  is contained in such a ball of radius  $s$ ; see Corollaries 8 and 9. In all other cases we can determine explicit bounds for the asymptotic worst case error; see Ineq. 11 and Proposition 10.

In the next section basic notations and concepts are introduced together with a slight generalization of the asymptotic formula in Kiderlen and Rataj (2006) to *weighted* volumes of morphological transforms. This is the basis for the main theoretical result, Theorem 3, which describes the asymptotic mean behaviour of an estimator based on weighted configuration counts in an isotropic slice. In the penultimate section, Theorem 3 is used to establish a surface area estimator based on weighted counts of  $m$  different configurations in a digitization of an isotropic slice section of  $X$ . We determine estimates for the asymptotic relative mean error, and show that these can be improved, if  $X$  is known to be contained in a ball centred at the reference point with known radius; see Proposition 10. Up to this point, the results hold for digitizations on general lattices (with possibly different resolutions along the different axes, and with not necessarily orthogonal axes) and for voxel blocks, which may be larger than  $2 \times 2 \times 2$  configurations. In the final section we specialize these results to the scaled standard lattice  $\mathbb{L} = t\mathbb{Z}^3$ , define an estimator based on the 102 informative  $2 \times 2 \times 2$ -configurations, and compare its performance in a simulation example with the theoretical asymptotic results.

## PRELIMINARIES

By  $\mathbb{S}^{d-1}$  we denote the unit sphere in  $d$ -dimensional real coordinate space  $\mathbb{R}^d$ . The standard scalar product on  $\mathbb{R}^d$  is  $\langle \cdot, \cdot \rangle$  with associated norm  $\|\cdot\|$ . By a  $k$ -subspace we mean a  $k$ -dimensional linear subspace of  $\mathbb{R}^d$ . Let  $A, B \subset \mathbb{R}^d$ . The reflection of  $A$  at the origin is denoted by  $\check{A} = \{-x \mid x \in A\}$ , its complement by  $A^c = \mathbb{R}^d \setminus A$ , and its topological boundary by  $\partial A$ . We write  $A \oplus B = \{a + b \mid a \in A, b \in B\}$  for the Minkowski sum of  $A$  and  $B$ , and  $A \ominus B = \{x \in \mathbb{R}^d \mid x + \check{B} \subset A\}$  for the dilatate of  $A$  by  $\check{B}$ . The positive part of a real valued function  $f$  is denoted by  $f^+ = \max(f, 0)$ . The support function of a convex body  $K$  in  $\mathbb{R}^d$  is denoted by  $h(K, \cdot)$ . We use this notion also for compact sets  $A$ ,  $A \neq \emptyset$ , defining  $h(A, \cdot) = h(\text{conv}(A), \cdot)$ , where  $\text{conv}(A)$  is the convex hull of  $A$ . The exoskeleton  $\text{exo}(A)$  of a closed set  $A$  is the set of all  $z \in A^c$ , which do not have a unique nearest point in  $A$ . The set  $\text{exo}(A)$  is measurable and has Lebesgue measure zero (Fremlin, 1997).

A closed set  $X \subset \mathbb{R}^d$  is gentle if for  $\mathcal{H}^{d-1}$ -almost all  $x \in \partial X$  there are two non-degenerate open balls touching in  $x$  such that one of them is contained in  $X$  and the other in  $X^c$ , and if also  $\mathcal{H}^{d-1}(N(\partial X \cap B \times S^{d-1})) < \infty$  for all bounded Borel sets  $B \subset \mathbb{R}^d$ .

Here  $\mathcal{H}^k$  is the  $k$ -dimensional Hausdorff measure in  $\mathbb{R}^d$  and  $N(A)$  is the reduced normal bundle of  $A$  (Kiderlen and Rataj, 2006). The class of gentle sets is rather large. It contains for instance all convex bodies (compact convex subsets of  $\mathbb{R}^d$ ) with interior points, all topologically regular sets in the convex ring (the family of finite unions of convex bodies), and certain unions of sets of positive reach.

At almost all boundary points  $a$  of a gentle set  $X$  there exists a unique outer unit normal  $n(a)$  to  $X$ . Let  $C_{d-1}(X, \cdot)$  be the image measure of  $\mathcal{H}^{d-1}$  on  $\partial X$  under the map  $a \mapsto (a, n(a))$ . The measure  $C_{d-1}(X, \cdot)$  vanishes outside  $N(X)$ . Let  $\xi_{\partial X} : \mathbb{R}^d \setminus \text{exo}(\partial X) \rightarrow \partial X$  denote the metric projection. The following theorem is a generalization of (Kiderlen and Rataj, 2006, Theorem 1).

**Theorem 1.** *Let  $X \subseteq \mathbb{R}^d$  be a closed gentle set,  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  a compactly supported bounded measurable function and  $B, W$  and  $P, Q$  four non-empty compact subsets of  $\mathbb{R}^d$ . Then*

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0^+} \frac{1}{\varepsilon} \int_{[(X \oplus \varepsilon P) \ominus \varepsilon B] \setminus [(X \ominus \varepsilon Q) \oplus \varepsilon W]} f(\xi_{\partial X}(x)) dx \\ &= \int_{N(X)} f(a) (h(P \oplus \check{Q}, n) - h(\check{B} \oplus W, n))^+ \\ & \quad \times C_{d-1}(X; d(a, n)). \quad (2) \end{aligned}$$

*If  $f$  is in addition continuous at all points in  $\partial X$ , then  $f(\xi_{\partial X}(x))$  can be replaced by  $f(x)$  in Eq. 2.*

The proof of Theorem 1 can be found in the appendix. Kiderlen and Rataj (2006) show Theorem 1 in the case where  $f$  is an indicator function. Their result can be generalized to Theorem 1 essentially by applying the monotone convergence theorem.

Let  $x_1, \dots, x_d$  be a basis of  $\mathbb{R}^d$  and let

$$\mathbb{L} = \{n_1x_1 + \dots + n_dx_d \mid n_1, \dots, n_d \in \mathbb{Z}\}$$

be the lattice generated by this basis. A given lattice  $\mathbb{L}$  is generated by infinitely many different bases, but the volume of the fundamental cell  $C_0 = [0, x_1] \oplus \dots \oplus [0, x_d]$  depends only on  $\mathbb{L}$  and not on the basis chosen; see for example (Yap, 2000). This number is denoted by  $\det(\mathbb{L})$ . If  $\xi$  is a uniform random variable in  $C_0$ , then the random lattice  $\xi + \mathbb{L}$  is a stationary random lattice. The distribution of  $\xi + \mathbb{L}$  does not depend on the choice of  $C_0$ . We now define the digitization of a set  $X \subset \mathbb{R}^d$ . The points of  $\xi + \mathbb{L}$  are interpreted as voxel midpoints of a digital image, where each voxel is a translate of  $C_0$ . Often,  $\mathbb{L}$  is the orthogonal standard lattice  $\mathbb{Z}^d$  and the voxels are small cubes. We work essentially with the Gauss digitization model of  $X$  consisting of all voxels having their midpoints in  $X$  (Klette and Rosenfeld, 2004, p. 56). As there is a one-to-one correspondence between the set of voxels on the one hand and  $\xi + \mathbb{L}$  on the other hand, the Gauss digitization is determined by the set  $X \cap (\xi + \mathbb{L})$  of all lattice points in  $X$ . In order to vary the resolution of the digitization, the random lattice is often scaled by a variable  $t > 0$  and  $X \cap t(\xi + \mathbb{L})$  is called the digitization of  $X$  (with resolution  $1/t$ ).

In the following we only consider compact gentle sets  $X$ . Let the function  $f$  be measurable non-negative or integrable. Let  $\xi + \mathbb{L}$  be a stationary random lattice, and let  $B, W \subseteq \mathbb{L}$  be two non-empty finite subsets of  $\mathbb{L}$ . The points in  $B$  represent ‘black’ pixels of the digitization, i.e. points which are contained in the set  $X$ , whereas  $W$  stands for the ‘white’ points of the background. For  $t > 0$  define

$$N_t := \sum_{x \in t(\xi + \mathbb{L})} f(x) \mathbb{1}_{\{x+tB \subseteq X \cap t(\xi + \mathbb{L}), x+tW \subseteq t(\xi + \mathbb{L}) \setminus X\}}.$$

For  $f \equiv 1$  the random variable  $N_t$  counts the number of all translations of the pattern  $(tB, tW)$  in the digitization  $X \cap t(\xi + \mathbb{L})$ . Calculation shows that

$$\mathbb{E}[N_t] = \frac{t^{-d}}{\det(\mathbb{L})} \int_{[X \ominus t\check{B}] \setminus [X \oplus t\check{W}]} f(x) dx. \quad (3)$$

**Corollary 2.** *Let  $X \subseteq \mathbb{R}^d$  be a compact gentle set. Let  $f$  be a locally bounded measurable function, which is*

*continuous on  $\partial X$  and let  $B$  and  $W$  be two non-empty finite subsets of a lattice  $\mathbb{L}$ . Then*

$$\begin{aligned} & \lim_{t \rightarrow 0+} t^{d-1} \det(\mathbb{L}) \mathbb{E}[N_t] \\ &= \int_{N(X)} f(a) (-h(B \oplus \check{W}, n))^+ C_{d-1}(X; d(a, n)). \end{aligned}$$

*Proof.* Let  $C$  be a compact set such that  $[X \ominus t\check{B}] \setminus [X \oplus t\check{W}] \subset C$  for all  $t > 0$  smaller than some fixed  $t_0 > 0$  and  $X \subset C$ . Replacing  $f, P, B, Q$  and  $W$  in Theorem 1 by  $1_C f, \{0\}, \check{B}, \{0\}$  and  $\check{W}$ , respectively, yields the claim.  $\square$

## COMBINING LOCAL AND DIGITAL STEREOLOGY

In the following we restrict ourselves to  $\mathbb{R}^3$ . The results can be generalized to  $\mathbb{R}^d$ ,  $d \geq 4$ , in a straightforward manner. We prefer to present them only in  $\mathbb{R}^3$  in order to keep the notation concise.

Denote the standard basis vectors in  $\mathbb{R}^3$  by  $e_1, e_2, e_3$ . Let  $R$  be a random proper rotation with distribution given by the normalized Haar measure on the rotation group  $SO_3$ . Fix the 2-subspace  $l_0 = \text{span}(e_1, e_2)$ . We define the random 2-subspace  $L = Rl_0$ . It is uniformly distributed in the set  $\mathcal{L}$  of all 2-subspaces of  $\mathbb{R}^3$ . Let  $\mu$  be the distribution of  $L$ . For  $l \in \mathcal{L}$  and  $s > 0$  define  $T_s = T_s(l) = l \oplus B(0, s)$ . The set  $T_s = T_s(L) = R(l_0 \oplus B(0, s)) = L \oplus B(0, s)$  is called a random 2-slice with thickness  $2s$ . It will be clear from the context whether  $T_s$  refers to the deterministic 2-slice  $T_s(l)$  or the random 2-slice  $T_s(L)$ .

The following Theorem 3, which is the main theoretical result of the paper, gives a formula for the asymptotic mean of the weighted number of occurrences of two sets  $B$  and  $W$  (black and white points, respectively), which are specified below. Theorem 3 is an analogous result in a local stereological setup to Kiderlen and Jensen (2003, Theorem 4) in the context of stationary random sets in the plane, and to Gutkowski *et al.* (2004, Theorem 1) and Kiderlen and Rataj (2006, Theorem 5) in the setting of spatial objects, which are digitized by a stationary random lattice.

**Theorem 3.** *Let  $X \subseteq \mathbb{R}^3$  be a compact gentle set. Let  $R$  be a random proper rotation and let  $\xi + \mathbb{L}$  be a stationary random lattice, which is independent of  $R$ . Let  $B, W \subset \mathbb{L}$  be two non-empty finite subsets of the lattice  $\mathbb{L}$  and  $f$  a continuous non-negative function on  $\mathbb{R}^3$ . The weighted sum  $N_t^f$  of occurrences of  $(B, W)$  in*

the digitization of  $X$ , which lie entirely in  $T_s$ , is given by

$$\sum_{\substack{x \in tR(\xi + \mathbb{L}) \\ x+tR(B \cup W) \subseteq T_s}} f(x) \mathbb{1}_{\{x+tRB \subseteq X \cap tR(\xi + \mathbb{L}), x+tRW \subseteq tR(\xi + \mathbb{L}) \setminus X\}},$$

for  $t > 0$ . It satisfies

$$\begin{aligned} \lim_{t \rightarrow 0+} t^2 (\det \mathbb{L}) \mathbb{E}[N_t^f] \\ = \int_{\mathcal{L}} \int_{\partial X \cap T_s} f(a) H(\phi_{a,l}^X) \mathcal{H}^2(da) \mu(dl), \end{aligned}$$

where  $H(\phi)$  is given by

$$\frac{1}{4\pi |\cos \phi|} \int_{S^2 \cap \{|e_3, \cdot\}| = |\sin \phi\|} (-h(B + \check{W}, v))^+ \mathcal{H}^1(dv), \quad (4)$$

and  $\phi_{a,l}^X$  is the angle between  $l$  and the outer normal  $n(a)$  of  $X$  at  $a \in \partial X$ .

The proof of Theorem 3 is based on the following Proposition 4 and Lemmas 5, 6 and can be found in the appendix. The difference of  $N_t^f$  in Theorem 3 and  $\tilde{N}_t^g$  in the following Proposition 4 is that the latter also counts configurations which do not lie entirely in the slice  $T_2$ . These are of course not observable in practice.

**Proposition 4.** *Let  $X \subseteq \mathbb{R}^3$  be a compact gentle set. Let  $R$  be a random proper rotation and let  $\xi + \mathbb{L}$  be a stationary random lattice, which is independent of  $R$ . Let  $B, W \subset \mathbb{L}$  be two non-empty finite subsets of  $\mathbb{L}$ . Let  $g : \mathbb{R}^3 \times \mathcal{L} \rightarrow \mathbb{R}$  be a non-negative bounded measurable function such that  $g(\cdot, l)$  is continuous for all  $l \in \mathcal{L}$ . Then the number*

$$\tilde{N}_t^g = \sum_{x \in tR(\xi + \mathbb{L})} g(x, L) \mathbb{1}_{\left\{ \begin{array}{l} x+tRB \subseteq (X \cap T_s) \cap tR(\xi + \mathbb{L}), \\ x+tRW \subseteq tR(\xi + \mathbb{L}) \setminus (X \cap T_s) \end{array} \right\}},$$

for  $t > 0$  satisfies

$$\lim_{t \rightarrow 0+} t^2 (\det \mathbb{L}) \mathbb{E}[\tilde{N}_t^g] = \mathbb{E}[F_g(R)],$$

where

$$F_g(r) := \int_{N(X \cap T_s)} g(a, rl_0) h(r^{-1}n) C_2(X \cap T_s; d(a, n)) \quad (5)$$

for  $r \in SO_3$  with  $h = (-h(B \oplus \check{W}, \cdot))^+$ .

The proof of Proposition 4 can be found in the appendix. It uses that given the rotation  $R$ , Corollary 2 can be applied to  $\tilde{N}_t^g$ . In the following two Lemmas we derive a formula for the expected value of  $F_g(R)$ .

**Lemma 5.** *Let  $g : \mathbb{R}^3 \times \mathcal{L} \rightarrow \mathbb{R}$  be a non-negative measurable function such that  $g(\cdot, l)$  is continuous for all  $l \in \mathcal{L}$ . Let  $F_g(r)$  be given by Eq. 5. Then the conditional expectation  $\mathbb{E}[F_g(R)|L = l]$  can be expressed as*

$$\begin{aligned} \int_{\partial X \cap T_s} g(a, l) H(\phi_{a,l}^X) \mathcal{H}^2(da) \\ + \int_{X \cap \partial T_s} g(a, l) H(\phi_{a,l}^X) \mathcal{H}^2(da) \end{aligned}$$

for  $\mu$ -almost all  $l \in \mathcal{L}$ , where  $H(\phi)$  is given by Eq. 4.

*Proof.* For  $\mu$ -almost all  $l \in \mathcal{L}$  we have that  $\mathcal{H}^2(\partial X \cap \partial T_s) = 0$ . This implies that there is a unique outer unit normal  $n(a)$  for  $\mathcal{H}^2$ -almost all  $a \in \partial(X \cap T_s)$ . Hence we obtain for  $r \in SO_3$

$$F_g(r) = \int_{\partial(X \cap T_s)} g(a, rl_0) h(r^{-1}n(a)) \mathcal{H}^2(da).$$

There exists a regular version of the conditional distribution of  $R$  given  $L$  (Klenke, 2006, Theorem 8.36). Therefore we can use Fubini's theorem to obtain

$$\begin{aligned} \mathbb{E}[F_g(R)|L = l] \\ = \int \mathbb{1}_{\partial(X \cap T_s)}(a) g(a, l) \mathbb{E}[h(r^{-1}n(a))|L = l] \mathcal{H}^2(da). \end{aligned}$$

Recall that  $\mu$ -almost surely  $\mathbb{1}_{\partial(X \cap T_s)} = \mathbb{1}_{\partial X \cap T_s} + \mathbb{1}_{X \cap \partial T_s}$ . The claim now follows from Lemma 6.  $\square$

**Lemma 6.** *For  $n \in \mathbb{S}^2$  and  $l \in \mathcal{L}$  we have*

$$\mathbb{E}[h(R^{-1}n)|L = l] = H(\phi),$$

where  $\phi$  is the angle between  $n$  and  $l$ , and  $H(\phi)$  is given by Eq. 4.

*Proof.* Fix  $\rho \in SO_3$  such that  $\rho l = l_0$ . Then

$$\begin{aligned} \mathbb{E}[h(R^{-1}n)|L = l] &= \mathbb{E}[h(R^{-1}n)|\rho R l_0 = l_0] \\ &= \mathbb{E}[h(R^{-1}\rho n)|R l_0 = l_0]. \end{aligned}$$

We have  $\langle n, l \rangle = \langle \rho n, l_0 \rangle$ , hence the last conditional expectation in the above equation can be written as the normalized integral over the two small circles  $\subset \mathbb{S}^2$  parallel to  $l_0$  at height  $\pm \sin \phi$  with radius  $\cos \phi$ , where  $\phi$  is the angle between  $n$  and  $l$ .  $\square$

## AN ESTIMATOR FOR SURFACE AREA

In this section we will derive an estimator for surface area using Theorem 3, which is based on a local stereological sampling design.

Let  $X \in \mathbb{R}^3$  be a compact gentle set. Suppose we observe  $X \cap R[(l_0 + B(0, s)) \cap t(\xi + \mathbb{L})]$  for some random proper rotation  $R$ ,  $\xi + \mathbb{L}$  a stationary random lattice, which is independent of  $R$ , and  $s, t > 0$ . Let  $(B_i, W_i)$ ,  $i = 1, \dots, m$  be boundary configurations of  $\mathbb{L}$ , i.e.  $B_i, W_i$  are non-empty, disjoint finite subsets of  $\mathbb{L}$  with  $B_i \cup W_i = C_0 \cap \mathbb{L}$ , where  $C_0$  is a fixed fundamental cell of  $\mathbb{L}$ . We define the following estimator for surface area

$$\hat{S}(X) = t^2(\det \mathbb{L}) \sum_i N_t^i, \quad (6)$$

where  $N_t^i = N_t^{\lambda_i(\|\cdot\|)}$  as defined in Theorem 3 with  $B = B_i, W = W_i$ . The continuous functions  $\lambda_i : [0, \infty) \rightarrow [0, \infty)$  have to be suitably chosen according to the choice of  $(B_i, W_i)$ . We give an example for  $2 \times 2 \times 2$ -configurations in the concluding section.

**Proposition 7.** *We have*

$$\lim_{t \rightarrow 0+} \mathbb{E}[\hat{S}(X)] = \int_{\partial X} \sum_i \lambda_i(\|a\|) g_i(\|a\|, \psi_a^X) \mathcal{H}^2(da), \quad (7)$$

where  $\psi_a^X \in [0, \pi]$  is the angle between  $a$  and the outer normal  $n(a)$  of  $X$  at  $a \in \partial X$ . The function  $g_i(r, \psi)$  for  $\psi \in [0, \pi]$  and  $r \in [0, \infty)$  is given by

$$g_i(r, \psi) = \begin{cases} \int_0^1 H_i(\arcsin(z)) dz, & \text{for } r \leq s, \\ \int_0^1 H_i(\arcsin(z)) G_{\psi, s/r}(z) dz, & \text{for } r > s. \end{cases}$$

With the two sided cut-off function  $x \mapsto x^* = \min\{1, \max\{-1, x\}\}$ , the function  $G_{\psi, q}(z)$  for  $\psi \in (0, \pi)$  and  $q \in (0, 1]$  can be written as

$$G_{\psi, q}(z) = \frac{1}{\pi} (\arccos(\alpha_{\psi, -q}^*(z)) - \arccos(\alpha_{\psi, q}^*(z))),$$

where  $\alpha_{\psi, q}(z) = (q - z \cos \psi) / (\sin \psi \sqrt{1 - z^2})$ . For  $\psi \in \{0, \pi\}$  we have

$$G_{0, q}(z) = G_{\pi, q}(z) = \mathbb{1}_{[0, q]}(z).$$

*Proof.* By Theorem 3 we obtain that

$$\lim_{t \rightarrow 0+} \mathbb{E}[\hat{S}(X)] = \int_{\mathcal{L}} \int_{\partial X \cap T_s} \sum_i \lambda_i(\|a\|) H_i(\phi_{a, l}^X) \mathcal{H}^2(da) \mu(dl), \quad (8)$$

where  $H_i$  is given by Eq. 4 for  $(B, W) = (B_i, W_i)$ . Using Fubini's theorem we can rewrite the right-hand side of Eq. 8 as

$$\int_{\partial X} \sum_i \lambda_i(\|a\|) \int_{\mathcal{L}} \mathbb{1}_{T_s}(a) H_i(\phi_{a, l}^X) \mu(dl) \mathcal{H}^2(da).$$

We have  $\int_{\mathcal{L}} \mathbb{1}_{T_s}(a) H_i(\phi_{a, l}^X) \mu(dl) = g_i(\|a\|, \psi_a^X)$ , which is a tedious but not difficult calculation.  $\square$

Note that for all  $\psi \in [0, \pi]$  and all  $r \in (s, \infty)$

$$\int_0^1 G_{\psi, s/r}(z) dz = \int_{\mathcal{L}} \mathbb{1}_{T_s}(a) \mu(dl) = \frac{s}{r} \quad (9)$$

(Jensen, 1998). We assume from now on that none of the functions  $H_i$ ,  $i = 1, \dots, m$ , is identically zero. This is fulfilled, if and only if  $B_i$  and  $W_i$  can be strictly separated by a hyperplane for all  $i = 1, \dots, m$ , or, in other words,  $(B_i, W_i)$  is an informative configuration.

The following two corollaries to Proposition 7 show that the weight functions  $\lambda_i$  can be chosen to yield an asymptotically unbiased estimator  $\hat{S}(X)$  in the case where  $X$  is sufficiently small compared to  $s$  or when  $X$  is a ball centred at the origin with unknown radius.

**Corollary 8.** *Let  $X \subseteq B(0, s)$ . Set  $\lambda_i(r) = a_i g_i(0, 0)^{-1}$  with coefficients  $a_1, \dots, a_m \in \mathbb{R}$  that are summing up to one. Then the estimator  $\hat{S}(X)$  as given in Eq. 6 is an asymptotically unbiased estimator of the surface area  $S(X)$ .*

*Proof.* By definition  $H_i \geq 0$ . The assumption that  $H_i \neq 0$  yields  $g_i(0, 0) > 0$ . As  $g_i(r, \psi) = g_i(0, 0)$  for all  $(r, \psi) \in [0, s] \times [0, \pi]$ , Eq. 7 implies the claim.  $\square$

**Corollary 9.** *Let  $X$  be a ball centred at the origin with unknown radius. Suppose that the sets  $B_i, W_i$  are such that  $g_i(r, 0) > 0$  for all  $r \in [0, \infty)$  and  $i = 1, \dots, m$ . Choosing  $\lambda_i(r) = a_i g_i(r, 0)^{-1}$ , where  $\sum_{i=1}^m a_i = 1$ , yields an asymptotically unbiased estimator  $\hat{S}(X)$  of  $S(X)$ .*

*Proof.* If  $X$  is a ball centred at the origin we have  $\psi_a^X = 0$  for all  $a \in \partial X$ . The claim follows from Eq. 7.  $\square$

The condition  $g_i(r, 0) > 0$  in Corollary 9 is equivalent to requiring that the support of  $H_i$  contains an interval  $[0, \varepsilon)$  for some  $\varepsilon > 0$ .

For general shapes we cannot expect to obtain an unbiased estimator of the form Eq. 7. A suitable choice of  $\lambda_i$  for  $r > s$  will strongly depend on the choice of the pairs  $(B_i, W_i)$ . In the sequel we propose a method to choose the weight functions  $\lambda_i$  and show how the asymptotic relative worst case error can be determined

in this case. Suppose we can determine coefficients  $\mu_i \geq 0$  such that for all  $z \in [0, 1]$  we have

$$\sum_{i=1}^m \mu_i H_i(\arcsin(z)) \approx 1,$$

then by Eq. 9 we obtain for all  $\psi \in [0, \pi]$  that  $\sum_{i=1}^m \mu_i g_i(r, \psi) \approx f(r)^{-1}$ , where  $f(r) = \max\{1, r/s\}$ . The function  $f(\|a\|)^{-1}$  is the probability that  $a$  is contained in the random 2-slice  $T_s$  (Jensen, 1998). Setting  $\lambda_i(r) = \mu_i f(r)$ , we obtain by Proposition 7, that

$$\lim_{t \rightarrow 0^+} \mathbb{E}[\hat{S}(X)] \approx S(X).$$

We suggest to chose  $(\mu_1, \dots, \mu_m)$  within the set  $\mathcal{S} \subseteq [0, \infty)^m$  of all  $(\mu_1, \dots, \mu_m)$  such that there exists  $(a_1, \dots, a_m) \in [0, 1]^m$  with  $\sum_{i=1}^m a_i = 1$  and  $\mu_i = a_i g_i(0, 0)^{-1}$ . This guarantees that the estimator is asymptotically unbiased for sets  $X \subseteq B(0, s)$  by Corollary 8.

In the remainder of this section we show how to determine the asymptotic relative worst case error for given coefficients  $(\mu_1, \dots, \mu_m) \in \mathcal{S}$ . It is immediate from Eq. 8, that if

$$1 - v_1^m \leq \sum_{i=1}^m \mu_i H_i(\arcsin(z)) \leq 1 + v_1^M \quad (10)$$

for some  $v_1^m, v_1^M > 0$  and for all  $z \in [0, 1]$ , then

$$1 - v_1^m \leq \frac{\lim_{t \rightarrow 0^+} \mathbb{E}[\hat{S}(X)]}{S(X)} \leq 1 + v_1^M. \quad (11)$$

This error bound is independent of the size and shape of  $X$ . If we know that  $X \subseteq B(0, R)$  for some  $R > 0$ , then the worst case error is typically smaller and one can determine a bound using the following proposition. Note that the function  $f$  occurring here is the inverse probability that a point of distance  $r$  from the origin is contained in  $T_s$ .

**Proposition 10.** *Suppose that  $X \subseteq B(0, R)$ . Let  $(\mu_1, \dots, \mu_m) \in \mathcal{S}$  and set*

$$\lambda_i(r) = \mu_i f(r),$$

where  $f(r) = \max\{1, r/s\}$ . Let  $\varepsilon > 0$ , and let  $L$  and  $M(r)$  be given as in Lemma 11 below. Define  $r_k = (1 + \varepsilon/(2L))^k s$  for  $k \in \mathbb{N}$ . Let  $n$  be minimal such that  $r_n \geq R$ . For each  $k = 0, \dots, n$  let  $0 = \psi_{k0} < \psi_{k1} < \dots < \psi_{kn_k} = \pi$  be a partition of  $[0, \pi]$ , such that  $|\psi_{k,l+1} - \psi_{k,l}| \leq \varepsilon/(2M(r_k))$  for all  $l = 0, \dots, n_k$ . Set

$$v_2^M = \max_{k,l} \sum_{i=1}^m \lambda_i(r_k) g_i(r_k, \psi_{kl}) - 1,$$

$$v_2^m = 1 - \min_{k,l} \sum_{i=1}^m \lambda_i(r_k) g_i(r_k, \psi_{kl}).$$

Then

$$1 - v_2^m - \varepsilon \leq \frac{\lim_{t \rightarrow 0^+} \mathbb{E}[\hat{S}(X)]}{S(X)} \leq 1 + v_2^M + \varepsilon. \quad (12)$$

For a proof of Proposition 10 we make use of a Lipschitz result for the integrand in Eq. 7.

**Lemma 11.** *The function  $(r/s) \sum_{i=1}^m \mu_i g_i(r, \psi)$  for  $r > s$  is Lipschitz continuous with respect to  $\psi \in [0, \pi]$  with Lipschitz constant*

$$M(r) = \frac{1}{\pi} \sqrt{\frac{r^2}{s^2} - 1} \left\| \frac{d}{dz} (H(\arcsin(z))) \right\|_1,$$

where  $H = \sum_i \mu_i H_i$  and  $\|\cdot\|_1$  denotes the  $L^1$ -norm on  $[0, 1]$ . It is also Lipschitz continuous with respect to  $r \in [r_0, \infty)$ ,  $r_0 \geq s$ , uniformly in  $\psi$ , with Lipschitz constant  $(1/r_0)L$ , where

$$L = \|H(\arcsin(\cdot))\|_\infty + H\left(\frac{\pi}{2}\right) + 2 \left\| \frac{d}{dz} (H(\arcsin(z))) \right\|_1.$$

The proof of Lemma 11 can be found in the appendix.

*Proof of Proposition 10.* Let  $\psi \in [0, \pi]$ . Then with  $l \in \{0, \dots, n_k - 1\}$  such that  $\psi \in [\psi_{kl}, \psi_{k,l+1}]$  we obtain

$$\begin{aligned} & \sum_{i=1}^m \lambda_i(r_k) g_i(r_k, \psi) \\ &= \sum_{i=1}^m \lambda_i(r_k) g_i(r_k, \psi) - \sum_{i=1}^m \lambda_i(r_k) g_i(r_k, \psi_{kl}) \\ & \quad + \sum_{i=1}^m \lambda_i(r_k) g_i(r_k, \psi_{kl}) \\ & \leq M_i(r_k) |\psi - \psi_{kl}| + 1 + v_2^M \\ & \leq 1 + v_2^M + \frac{\varepsilon}{2}. \end{aligned}$$

Hence for  $r \in [s, R]$  with  $k$  such that  $r \in [r_k, r_{k+1})$  we have

$$\begin{aligned} & \sum_{i=1}^m \lambda_i(r) g_i(r, \psi) \\ &= \sum_{i=1}^m \lambda_i(r) g_i(r, \psi) - \sum_{i=1}^m \lambda_i(r_k) g_i(r_k, \psi) \\ & \quad + \sum_{i=1}^m \lambda_i(r_k) g_i(r_k, \psi) \\ & \leq \frac{L}{r_k} (r - r_k) + 1 + v_2^M + \frac{\varepsilon}{2} \\ & \leq 1 + v_2^M + \varepsilon. \end{aligned}$$

By the same arguments one obtains the analogous lower bound for  $\sum_{i=1}^m \lambda_i(r)g_i(r, \psi)$ . Using Eq. 7 this yields Ineq. 12.  $\square$

## COEFFICIENTS FOR $2 \times 2 \times 2$ -CONFIGURATIONS

Recall that a configuration of size  $2 \times 2 \times 2$  is a pair  $(B, W)$  of non-empty disjoint subsets of  $\mathbb{L} = \mathbb{Z}^3$ , such that  $B \cup W = \mathbb{Z}^3 \cap [0, 1]^3$ . It is informative, if there is a hyperplane, which strictly separates  $B$  and  $W$ . In this section we want to investigate the surface area estimator Eq. 6 in the case, where  $(B_i, W_i)$  runs through the family of all 102 informative configurations. These configurations are thoroughly investigated in Ziegel and Kiderlen, (2010). In particular the functions  $(-h(B_i + \check{W}_i))^+$  are explicitly given, hence we can numerically determine the functions  $H_i$ . As in Ziegel and Kiderlen (2010) we classify the informative configurations into five types, depending on the number and position of black points  $B$  or white points  $W$ . A configuration of type one has exactly one black point or exactly one white point, a configuration of type two has exactly two black points or exactly two white points, and a configuration of type three has exactly three black points or exactly three white points. Configurations of type four and five have exactly four white and four black points, which are affinely dependent in the case of type four, and affinely independent in the case of type five.

For configurations of type one, all functions  $H_i$  are identical and we denote them by  $H^1$ . The function  $H^1(\arcsin(z))$  for  $z \in [0, 1]$  is shown in Fig. 1. All functions  $H_i(\arcsin(z))$ ,  $z \in [-1, 1]$ , are symmetric with respect to the origin, which is why we only display them for values  $z \in [0, 1]$ . For configurations of type two, three and four there are two different functions  $H_i$  occurring per type. We denote them by  $H^{2,1}, H^{2,2}, H^{3,1}, H^{3,2}$  and  $H^{4,1}, H^{4,2}$  respectively; see Figs. 2, 3 and 4. For configurations of type five all functions  $H_i$  coincide and we denote them by  $H^5$ , which is displayed in Fig. 5. Fig. 6 shows all functions  $H_i$  scaled by the number of their occurrence amongst all functions  $H_i$  induced by informative configurations. The numbers of occurrence of the functions  $H^1, H^{2,1}, H^{2,2}, H^{3,1}, H^{3,2}, H^{4,1}, H^{4,2}$  and  $H^5$  are given in the last column of Table 2.

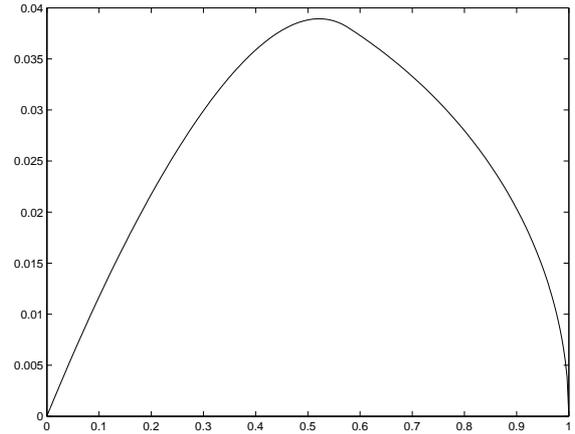


Fig. 1. Plot of the function  $H_i(\arcsin(\cdot))$  for a configuration of type one. There are 16 configurations of type one (not identifying twins).

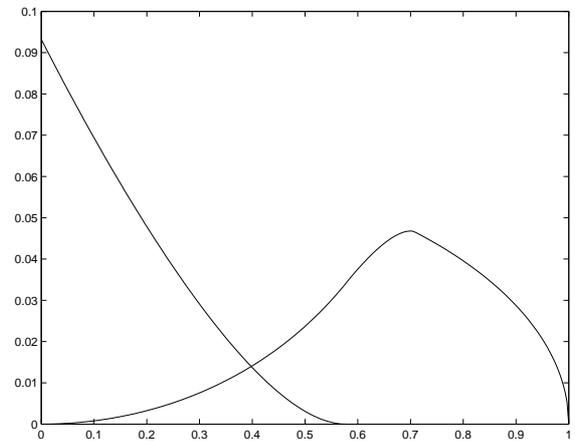


Fig. 2. Plot of the functions  $H_i(\arcsin(\cdot))$  that occur for configurations of type two. There are eight configurations of type two with  $H_i = H^{2,1}$  (left curve) and 16 configurations with  $H_i = H^{2,2}$  (right curve).

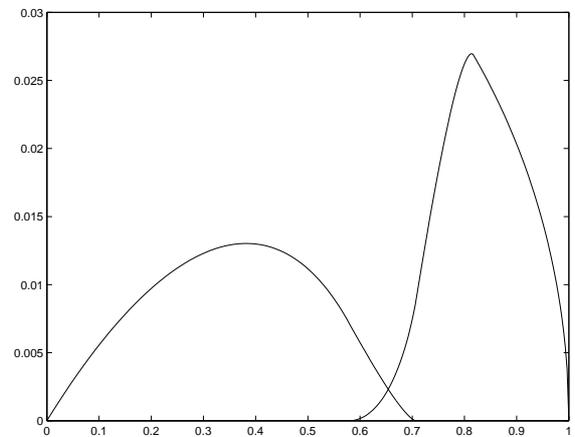


Fig. 3. Plot of the functions  $H_i(\arcsin(\cdot))$  that occur for configurations of type three. There are 32 configurations of type three with  $H_i = H^{3,1}$  (left curve) and 16 configurations with  $H_i = H^{3,2}$  (right curve).

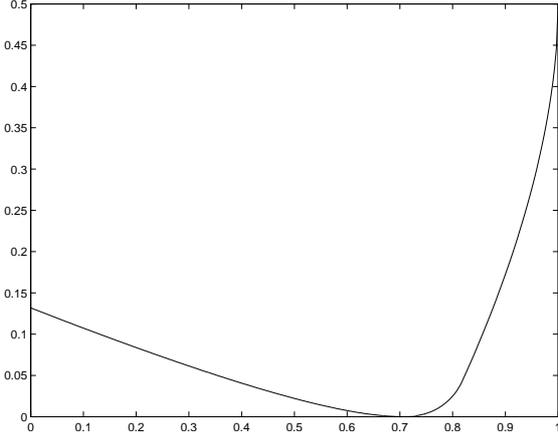


Fig. 4. Plot of the functions  $H_i(\arcsin(\cdot))$  that occur for configurations of type four. There are eight configurations of type four with  $H_i = H^{4,1}$  (curve for  $z \in [0, 0.7]$ , zero otherwise) and four configurations with  $H_i = H^{4,2}$  (curve for  $z \in [0.7, 1]$ , zero otherwise).

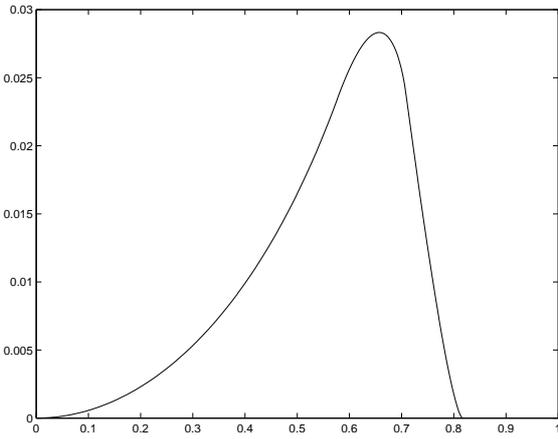


Fig. 5. Plot of the function  $H_i(\arcsin(\cdot))$  for a configuration of type five. There are 16 configurations of type five.

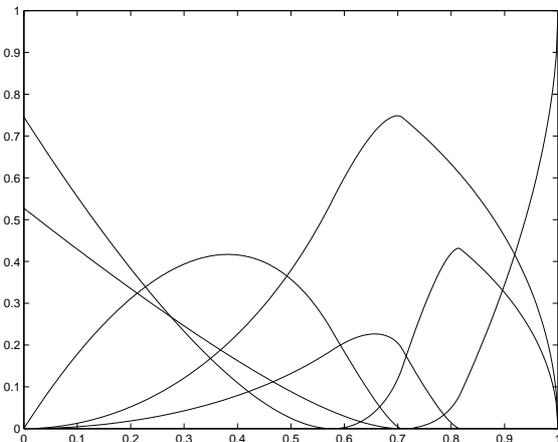


Fig. 6. Plot of all functions  $H_i(\arcsin(\cdot))$  that occur for  $2 \times 2 \times 2$ -configurations scaled according to their number of occurrence.

We define  $\eta_i = g_i(0,0)^{-1}$ . The values  $\eta_i$  for informative configurations are given in Table 1, where  $\eta_i = \eta^1$  whenever  $(B_i, W_i)$  is a configuration of type one, and analogously in the other cases.

Table 1. Values of  $\eta_i$  for informative configurations.

$H_i$	$\eta_i = \left( \int_0^1 H_i(\arcsin(z)) dz \right)^{-1}$
$H^1$	$\eta^1 = 38.119$
$H^{2,1}$	$\eta^{2,1} = 48.179$
$H^{2,2}$	$\eta^{2,2} = 48.179$
$H^{3,1}$	$\eta^{3,1} = 171.187$
$H^{3,2}$	$\eta^{3,2} = 171.187$
$H^{4,1}$	$\eta^{4,1} = 25.856$
$H^{4,2}$	$\eta^{4,2} = 25.856$
$H^5$	$\eta^5 = 114.825$

We have seen that the estimator in Eq. 6 with

$$\lambda_i(r) = \mu_i f(r) \quad (13)$$

is unbiased for  $S(X)$  if  $X$  is a subset of  $B(0, s)$ , whenever  $(\mu_1, \dots, \mu_{102}) \in \mathcal{S}$ . In Ziegel and Kiderlen, 2010) a one-parameter family of coefficients  $(\mu'_1(u), \dots, \mu'_{102}(u))$ ,  $u \in [0, 1]$  was derived that minimizes

$$\max_{n \in \mathcal{S}^2} \left| \sum_{i=1}^{102} \mu'_i(u) (-h(B_i + \check{W}_i, n))^+ - 1 \right| \quad (14)$$

for each  $u \in [0, 1]$ . Adapting the coefficients  $(\mu'_1(u), \dots, \mu'_{102}(u))$  to yield an asymptotically unbiased estimator for spherical shapes, we obtain the family of coefficients  $(\mu_1(u), \dots, \mu_{102}(u))$ ,  $u \in [0, 1]$  given in Table 2; for details see Ziegel and Kiderlen (2010). It turns out that  $(\mu_1(0), \dots, \mu_{102}(0)) \in \mathcal{S}$  or in other words

$$\sum_{i=1}^{102} \frac{\mu_i(0)}{\eta_i} = 1.$$

We therefore suggest to set  $\mu_i = \mu_i(0)$  and  $\lambda_i$  as in Eq. 13. Hence we obtain the estimator

$$\hat{S}(X) = t^2 \sum_{i=1}^{102} \mu_i(0) W_t^i, \quad (15)$$

which has the same formal structure as the estimator in equation Eq. 1. However, the ordinary configuration counts  $N_i$  in the classical estimator are replaced by the weighted configuration counts  $W_t^i$  given by

$$\sum_{\substack{x \in tR(\xi + \mathbb{Z}^3) \\ x + tR(BUW) \subseteq T_s}} f(\|x\|) \mathbb{1}_{\left\{ \begin{array}{l} x + tRB \subseteq X \cap tR(\xi + \mathbb{Z}^3), \\ x + tRW \subseteq tR(\xi + \mathbb{Z}^3) \setminus X \end{array} \right\}},$$

where  $f(r) = \max\{1, r/s\}$  is the inverse sampling probability. The fusion of digital and local stereology

becomes apparent in the structure of this estimator: While the weights  $\mu_i(0)$  stem from the ordinary digital surface area estimation, the weighting of the individual configurations in  $W_i^f$  with  $f$  reflects the local design. With this choice we obtain

$$0.9539 \leq \max_{z \in [0,1]} \sum_i \mu_i(0) H_i(\arcsin(z)) \leq 1.0225, \quad (16)$$

using the definition Eq. 4 of the  $H_i$ , hence  $v_1^m = 0.0461$ ,  $v_1^M = 0.0225$  fulfill Eq. 10. Therefore, by Ineq. 11, the upper and lower bound in Ineq. 16 are bounds for the asymptotic relative worst case error of  $\hat{S}(X)$ , as given in Eq. 15, independent of the shape and size of  $X$ . In the next paragraph we illustrate the application of Proposition 10 with a simulation example. This yields better error bounds but requires information about the size and preferably also the shape of  $X$ .

Table 2. *One-parameter family of coefficients  $\mu_i$  for  $u \in [0, 1]$ . The number in the last column indicates the number of occurrences of the function  $H^1, H^{2,1}, H^{2,2}, H^{3,1}, H^{3,2}, H^{4,1}, H^{4,2}, H^5$ , respectively amongst all functions  $H_i$  for informative  $2 \times 2 \times 2$ -configurations.*

$H_i$	$\mu_i(u)$ ,	$u \in [0, 1]$	
$H^1$	$\mu^1(u)$	$= (1.652/2)u$	16
$H^{2,1}$	$\mu^{2,1}(u)$	$= 0.675$	8
$H^{2,2}$	$\mu^{2,2}(u)$	$= 0.675$	16
$H^{3,1}$	$\mu^{3,1}(u)$	$= 1.168 - (1.652/4)u$	32
$H^{3,2}$	$\mu^{3,2}(u)$	$= 1.168 - (1.652/2)u$	16
$H^{4,1}$	$\mu^{4,1}(u)$	$= 0.954$	4
$H^{4,2}$	$\mu^{4,2}(u)$	$= 0.954$	2
$H^5$	$\mu^5(u)$	$= 1.652(1 - u)$	8

We consider a cylinder with radius 1 and height 2 centred at 0 which is contained in  $B(0, \sqrt{2})$ . We observe an isotropic slice of thickness  $s = 1$ . For a lattice distance of  $t = 0.055$  we obtain a mean estimated surface area of 18.115 with variance 1.146 in 1000 Monte Carlo simulations. This corresponds to a mean relative error of 3.8%. For  $t = 0.020$  the mean estimated surface area in 1000 simulations is 18.553 with variance 1.205 and mean relative error 1.6%. The asymptotic relative mean error for sets  $X$  with  $X \subseteq B(0, \sqrt{2})$  and  $\psi_a^X \in [0, \pi/4]$  for all  $a \in X$  is less than 1.2%. We determined this value numerically using Proposition 10. We obtain  $\|(d/dz)H(\arcsin(z))\|_1 = 5.7002$ , where  $H = \sum_i \mu_i(0)H_i$  and  $L = 13.3768$  as defined in Lemma 11. Note that the information  $\psi_a^X \in [0, \pi/4]$  reduces the asymptotic relative mean error as it is then sufficient to work with partitions  $0 = \psi_{k0} < \psi_{k1} < \dots < \psi_{kn_k} = \pi/4$  in Proposition 10. With  $\varepsilon = 0.0030$  we obtain  $v_2^M = 0.0042$  and  $v_2^m = 0.0082$ .

In summary this simulation example indicates that the bias of the simple linear surface area estimator Eq. 6 with weight functions Eq. 13 is of reasonable magnitude. It appears that the theoretical asymptotic error bounds are often too optimistic, and should only be used in the case where good to very good resolution images are available.

## ACKNOWLEDGEMENTS

The first author was generously supported by RiskLab Switzerland. The second author was supported by the Danish Council of Strategic Research.

## APPENDIX

In this appendix we present the proofs of Theorems 1 and 3, Proposition 4 and Lemma 11.

*Proof of Theorem 1.* Let  $C \subseteq \mathbb{R}^d$  be a bounded Borel set. Then Eq. 2 holds for  $f = 1_C$  by Kiderlen and Rataj (2006, Theorem 1). It is immediate that Eq. 2 also holds for compactly supported measurable step functions. For a non-negative compactly supported bounded measurable function  $f$ , let  $(f_k)_{k \in \mathbb{N}}, (g_k)_{k \in \mathbb{N}}$  be sequences of step functions such that  $f_k \uparrow f$  and  $g_k \downarrow f$  and  $f_k \geq 0$ . Let  $M_\varepsilon := [(X \oplus \varepsilon P) \ominus \varepsilon B] \setminus [(X \ominus \varepsilon Q) \oplus \varepsilon W]$ . Then we obtain with  $h(\cdot) = (h(P \oplus Q, \cdot) - h(B \oplus W, \cdot))^+$  that

$$\begin{aligned} & \int_{N(X)} f_k(a) h(n) C_{d-1}(X; d(a, n)) \\ & \leq \limsup_{\varepsilon \rightarrow 0^+} \frac{1}{\varepsilon} \int_{M_\varepsilon} f(\xi_{\partial X}(x)) dx \\ & \leq \int_{N(X)} g_k(a) h(n) C_{d-1}(X; d(a, n)) . \end{aligned} \quad (17)$$

Using the monotone convergence theorem we obtain

$$\begin{aligned} & \limsup_{\varepsilon \rightarrow 0^+} \frac{1}{\varepsilon} \int_{[(X \oplus \varepsilon P) \ominus \varepsilon B] \setminus [(X \ominus \varepsilon Q) \oplus \varepsilon W]} f(\xi_{\partial X}(x)) dx \\ & = \int_{N(X)} f(a) h(n) C_{d-1}(X; d(a, n)) . \end{aligned}$$

Note that for applying the monotone convergence theorem to the right-hand side of Ineq. 17 we can assume that  $\cup_{k \in \mathbb{N}} \text{supp}(g_k)$  is compact and that the sequence  $(g_k)_{k \in \mathbb{N}}$  is uniformly bounded. The same argument holds if we take  $\liminf_{\varepsilon \rightarrow 0^+}$  in Ineq. 17 and hence the claim is shown for  $f \geq 0$ . For general  $f = f^+ - f^-$  we can treat  $f^+$  and  $f^-$  separately to obtain Eq. 2. If  $f$  is in addition continuous in all points of  $\partial X$ , we obtain uniform continuity in the following sense.

For each  $\eta > 0$  there exists a  $\delta > 0$  such that for all  $x \in \partial X \cap \text{supp}(f)$  and  $y \in \mathbb{R}^d$  with  $\|x - y\| \leq \delta$  it follows that  $\|f(x) - f(y)\| < \eta$ . Furthermore  $x \in M_\varepsilon$  implies  $\|x - \xi_{\partial X}(x)\| \leq \varepsilon R$ , where  $R = 2 \max\{\|y\| \mid y \in (W \oplus Q) \cup (B \oplus P)\}$ . This implies the second claim.  $\square$

*Proof of Proposition 4.* For  $\mu$ -almost all  $l \in \mathcal{L}$  we have that  $\mathcal{H}^2(\partial X \cap \partial T_s) = 0$ , which can be seen using Fubini's theorem. This implies that the set  $X \cap T_s$  is compact gentle for  $\mu$ -almost all  $l \in \mathcal{L}$ . Applying Corollary 2 we obtain that  $t^2 \mathbb{E}[\tilde{N}_t^g | R = r] \rightarrow F_g(r)$  for  $t \rightarrow 0+$  pointwise for almost all  $r \in SO_3$  as  $(-h(rB \oplus r\check{W}, n))^+ = h(r^{-1}n)$ . We claim that the conditional expectation  $t^2 \mathbb{E}[\tilde{N}_t^g | R]$  is uniformly bounded for  $t \leq 1$ , hence Lebesgue's dominated convergence theorem yields the assertion. In fact Eq. 3 implies

$$\mathbb{E}[\tilde{N}_t^g | R = r] = t^{-3} \int_{[(X \cap T_s) \ominus tr\check{B}] \setminus [(X \cap T_s) \oplus tr\check{W}]} g(x, rl_0) dx.$$

By assumption there is a constant  $C > 0$  such that  $|g| \leq C$ . If  $x \in [(X \cap T_s) \ominus tr\check{B}] \setminus [(X \cap T_s) \oplus tr\check{W}]$ , then  $\text{dist}(x, \partial(X \cap T_s)) \leq tC'$ , where  $C' > 0$  is a constant depending only on  $(B, W)$ . Hence we obtain

$$\begin{aligned} |t^2 \mathbb{E}[\tilde{N}_t^g | R]| &= |t^{-1} \int_{[(X \cap T_s) \ominus tr\check{B}] \setminus [(X \cap T_s) \oplus tr\check{W}]} g(x, L) dx| \\ &\leq Ct^{-1} \mathcal{H}^3(\partial(X \cap T_s) \oplus B(0, tC')) \\ &\leq Ct^{-1} \mathcal{H}^3(\partial X \oplus B(0, tC')) \\ &\quad + Ct^{-1} \mathcal{H}^3(\partial(T_s \cap B(0, \text{diam} X)) \oplus B(0, tC')). \end{aligned}$$

Applying (Kiderlen and Rataj, 2006, Proposition 4), which is derived from a far-reaching generalization of Steiner's formula (Hug *et al.*, 2004), we obtain

$$\begin{aligned} t^{-1} \mathcal{H}^3(\partial X \oplus B(0, tC')) &= \sum_{i=1}^3 i \kappa_i \int_{N(\partial X)} \int_0^{\delta(\partial X; a, n)} s^{i-1} \mathbb{1}_{\partial X \oplus B(0, tC')}(a + sn) \\ &\quad \times ds \mu_{d-i}(\partial X; d(a, n)) \\ &\leq t^{-1} \sum_{i=1}^3 i \kappa_i \int_{N(\partial X)} \int_0^{tC'} s^{i-1} ds |\mu_{d-i}|(\partial X; d(a, n)) \\ &\leq t^{-1} \sum_{i=1}^3 \kappa_i (tC')^i |\mu_{d-i}|(\partial X; \partial X), \end{aligned}$$

where  $\mu_i(\partial X; \cdot)$  are the support measures of  $\partial X$ , and  $\delta(\partial X; a, n) = \inf\{t \geq 0 \mid a + tn \in \text{exo}(\partial X)\}$  is the reach function of  $\partial X$  at  $(a, n)$ . The support measures have locally finite total variation as  $X$  is gentle and hence the compactness of  $X$  yields boundedness of the last term in the above inequality for  $t \leq 1$ . The same argument can also be applied to  $t^{-1} \mathcal{H}^3(\partial(T_s \cap B(0, \text{diam} X)) \oplus B(0, tC'))$  as  $T_s \cap B(0, tC')$  is compact gentle.  $\square$

*Proof of Theorem 3.* Let  $0 < \varepsilon < s/2$ . For  $l \in \mathcal{L}$ , define the continuous function  $\chi_{s, \varepsilon}^l : \mathbb{R}^3 \rightarrow [0, 1]$  as a smoothed version of  $\mathbb{1}_{T_s}$  such that  $\chi_{s, \varepsilon}^l(x) = 0$ , if  $x \in (T_s)^c$  and  $\chi_{s, \varepsilon}^l(x) = 1$ , if  $x \in T_{s-\varepsilon}$ . Substituting  $g(x, l) = f(x) \chi_{s-\varepsilon, \varepsilon}^l(x)$  in Lemma 5 we obtain

$$\begin{aligned} \mathbb{E}[F_{f \chi_{s-\varepsilon, \varepsilon}^l}^l(R) | L = l] &= \int_{\partial X \cap T_s} f(a) \chi_{s-\varepsilon, \varepsilon}^l(a) H(\phi_{a, l}^X) \mathcal{H}^2(da). \end{aligned}$$

The right hand side converges pointwise in  $l$  to

$$\int_{\partial X \cap T_s} f(a) H(\phi_{a, l}^X) \mathcal{H}^2(da),$$

as  $\varepsilon \rightarrow 0$ . It is also bounded independently of  $\varepsilon$  and  $l$ , hence by the dominated convergence theorem  $\mathbb{E}[F_{f \chi_{s-\varepsilon, \varepsilon}^l}^l(R)]$  converges to

$$\int_{\mathcal{L}} \int_{\partial X \cap T_s} f(a) H(\phi_{a, l}^X) \mathcal{H}^2(da) \mu(dl). \quad (18)$$

Proposition 4 yields that

$$\lim_{\varepsilon \rightarrow 0} \lim_{t \rightarrow 0+} t^2 (\det L) \mathbb{E} \left[ \tilde{N}_t^f \chi_{s-\varepsilon, \varepsilon}^L \right]$$

is also given by Eq. 18. It remains to show that

$$\lim_{\varepsilon \rightarrow 0} \lim_{t \rightarrow 0+} \mathbb{E} \left[ \left| \tilde{N}_t^f \chi_{s-\varepsilon, \varepsilon}^L - N_t^f \right| \right] = 0.$$

Choose  $q > 0$  such that  $B \cup W \subseteq B(0, q)$ . Then for  $t < \varepsilon/q$  we have that  $x \in T_{s-2\varepsilon}$  implies  $\chi_{s-\varepsilon, \varepsilon}^L(x) = 1$  and  $x + t(B \cup W) \subseteq T_s$ . Therefore

$$\begin{aligned} \tilde{N}_t^f \chi_{s-\varepsilon, \varepsilon}^L - N_t^f &= \sum_{\substack{x \in tR(\xi + L) \\ x \in T_{s+\varepsilon} \setminus T_{s-2\varepsilon}}} f(x) (\chi_{s-\varepsilon, \varepsilon}^L(x) - \mathbb{1}_{\{x+tR(B \cup W) \subseteq T_s\}}) I_t(x), \end{aligned}$$

where  $I_t(x) = \mathbb{1}_{\{x+tRB \subseteq X \cap tR(\xi + L), x+tRW \subseteq tR(\xi + L) \setminus X\}}$  and we used that  $\chi_{s-\varepsilon, \varepsilon}^L(x) = 0$  for  $x \in (T_{s-\varepsilon})^c$ . Using Eq. 3 this yields

$$\begin{aligned} t^2 \mathbb{E} \left[ \left| \tilde{N}_t^f \chi_{s-\varepsilon, \varepsilon}^L - N_t^f \right| \mid R = r \right] &\leq t^{-1} \int_{[X \ominus tr\check{B}] \setminus [X \oplus tr\check{W}]} |f(x)| \mathbb{1}_{\partial T_s \oplus B(0, 2\varepsilon)}(x) dx. \end{aligned}$$

Choose a constant  $C$  such that  $|f| \leq C$  on a compact set containing  $[(X \cap T_s) \ominus tr\check{B}]$  for all  $r \in SO_3$  and all  $t \leq 1$ . Then the last expression in the above inequality is bounded by

$$Ct^{-1} \int_{[X \ominus tr\check{B}] \setminus [X \oplus tr\check{W}]} \mathbb{1}_{\partial T_s \oplus B(0, 3\varepsilon)}(\xi_{\partial X}(x)) dx.$$

This integral converges for all  $r \in SO_3$  by Theorem 1 as  $t \rightarrow 0+$

$$\int_{\partial X} (-h(rB \oplus r\check{W}, n(a)))^+ \mathbb{1}_{\partial T_s \oplus B(0, 3\varepsilon)}(a) \mathcal{H}^2(da).$$

The function  $(-h(rB \oplus r\check{W}, \cdot))^+$  is bounded by a constant  $C'$ , independent of  $r$ . Therefore, using the dominated convergence theorem, the limit of the above integral as  $\varepsilon \rightarrow 0$  is bounded by  $C' \mathcal{H}^2(\partial X \cap \partial T_s) = 0$  for almost all  $r$ , which yields the claim.  $\square$

*Proof of Lemma 11.* In order to find a Lipschitz constant with respect to  $\psi$  for

$$\sum_{i=1}^m \mu_i g_i(r, \psi) = \int_0^1 \sum_{i=1}^m \mu_i H_i(\arcsin(z)) G_{\psi, s/r}(z) dz,$$

we use partial integration to rewrite the function for  $r > s$  as

$$H\left(\frac{\pi}{2}\right) \frac{s}{r} - \int_0^1 \frac{d}{dz}(H(\arcsin(z))) \int_0^z G_{\psi, s/r}(x) dx dz, \tag{19}$$

where  $H = \sum_{i=1}^m \mu_i H_i$ . Then  $(\partial/\partial\psi) \sum_{i=1}^m \mu_i g_i(r, \psi)$  is given by

$$- \int_0^1 \frac{d}{dz}(H(\arcsin(z))) \int_0^z \frac{\partial}{\partial\psi} G_{\psi, s/r}(x) dx dz.$$

Let  $q = s/r$  and  $\psi, q, z$  such that  $|\alpha_{\psi, q}(z)| \leq 1$ . Then we obtain

$$\begin{aligned} & \frac{\partial}{\partial\psi} \arccos(\alpha_{\psi, q}(z)) \\ &= \frac{q \cos(\psi) - z}{\sin(\psi) \sqrt{\sin(\psi)^2 - z^2 - q^2 + 2qz \cos(\psi)}}, \end{aligned}$$

and hence

$$\begin{aligned} & \int \frac{\partial}{\partial\psi} \arccos(\alpha_{\psi, q}(z)) dz \\ &= \frac{1}{\sin(\psi)} \sqrt{\sin(\psi)^2 - z^2 - q^2 + 2qz \cos(\psi)}. \end{aligned}$$

The above expression is non-negative and bounded by  $\sqrt{1 - q^2}$ . Therefore  $(\partial/\partial\psi) \sum_{i=1}^m \mu_i g_i(r, \psi)$  is bounded by

$$\begin{aligned} & \int_0^1 \left| \frac{d}{dz}(H(\arcsin(z))) \right| \left| \int_0^z \frac{\partial}{\partial\psi} G_{\psi, s/r}(x) dx \right| dz \\ & \leq \left\| \frac{d}{dz}(H(\arcsin(z))) \right\|_1 \frac{1}{\pi} \sqrt{1 - q^2}, \end{aligned}$$

hence  $M(r)$  is a Lipschitz constant for  $(r/s) \sum_{i=1}^m \mu_i g_i(r, \psi)$  with respect to  $\psi$ .

To find a Lipschitz constant for  $\sum_{i=1}^m \lambda_i(r) g_i(r, \psi)$  with respect to  $r$  we first differentiate with respect to  $r$  and obtain

$$\begin{aligned} & \frac{1}{s} \int_0^1 H(\arcsin(z)) G_{\psi, s/r}(z) dz \\ & + \frac{r}{s} \frac{\partial}{\partial r} \int_0^1 H(z) G_{\psi, s/r}(z) dz. \end{aligned}$$

The first term of the above expression is bounded by  $1/r \|H(\arcsin(\cdot))\|_\infty$ . In order to find a bound for the second term we use partial integration to rewrite  $\sum_{i=1}^m \mu_i g_i(r, \psi)$  for  $r > s$  as in Eq. 19. Then  $(\partial/\partial r) \sum_{i=1}^m \mu_i g_i(r, \psi)$  is given by

$$\begin{aligned} & -H\left(\frac{\pi}{2}\right) \frac{s}{r^2} \\ & - \int_0^1 \frac{d}{dz}(H(\arcsin(z))) \int_0^z \frac{\partial}{\partial r} G_{\psi, s/r}(x) dx dz. \end{aligned}$$

Let  $\psi, r, z$  such that  $|\alpha_{\psi, s/r}(z)| \leq 1$ , then we have

$$\begin{aligned} & \frac{\partial}{\partial r} \arccos(\alpha_{\psi, s/r}(z)) \\ &= \frac{s}{r^2 \sqrt{\sin(\psi)^2 - z^2 - \frac{s^2}{r^2} + 2\frac{s}{r}z \cos(\psi)}}, \end{aligned}$$

and

$$\begin{aligned} & \int \frac{\partial}{\partial r} \arccos(\alpha_{\psi, s/r}(z)) dz \\ &= \frac{s}{r^2} \arctan\left(\frac{z - \frac{s}{r} \cos(\psi)}{\sqrt{\sin(\psi)^2 - z^2 - \frac{s^2}{r^2} + 2\frac{s}{r}z \cos(\psi)}}\right). \end{aligned}$$

The term on the right-hand side of the above equation is bounded in absolute value by

$$\frac{s}{r^2} \frac{\pi}{2}.$$

Therefore, for  $r > r_0$ , the function  $(r/s) \sum_{i=1}^m \mu_i g_i(r, \psi)$  has Lipschitz constant  $(1/r_0)L$  in  $r$  uniformly in  $\psi$ .  $\square$

## REFERENCES

Fremlin DH (1997). Skeletons and central sets. Proc London Math Soc 74:701–20.  
 Gutkowski P, Jensen EBV, Kiderlen M (2004). Directional analysis of digitized three-dimensional images by configuration counts. J Microsc 216:175–85.  
 Hug D, Last G, Weil W (2004). A local Steiner-type formula for general closed sets and applications. Math Z 246:237–72.

- Jensen EBV (1998). *Local stereology*. London: World Scientific.
- Kiderlen M, Jensen EBV (2003). Estimation of the directional measure of planar random sets by digitization. *Adv Appl Prob SGSA* 35:583–602.
- Kiderlen M, Rataj J (2006). On infinitesimal increase of volumes of morphological transforms. *Mathematika* 53:103–27.
- Klenke A (2006). *Wahrscheinlichkeitstheorie*. Berlin: Springer.
- Klette R, Rosenfeld A (2004). *Digital geometry: Geometric methods for digital picture analysis*. San Francisco: Morgan Kaufmann.
- Lindblad J (2005). Surface area estimation of digitized 3D objects using weighted local configurations. *Image Vision Comput* 23:111–22.
- Lindblad J, Nyström I (2002). Surface area estimation of digitized 3D objects using local computations. In: *Discrete Geometry for Computer Imagery: 10th Int Conf, DGCI 2002*. Berlin: Springer.
- Schladitz K, Ohser J, Nagel W (2006). Measuring intrinsic volumes in digital 3D images. In: Kuba A, Nyúl LG, Palágyi K, eds., *13th Int Conf Discrete Geom Computer Imagery*. Heidelberg: Springer.
- Yap CK (2000). *Fundamental problems of algorithmic algebra*. New York: Oxford University Press.
- Ziegel J, Kiderlen M (2010). Estimation of surface area and surface area measure of three-dimensional sets from digitizations. *Image Vision Comput* 28:64–77.