Estimation of surface area and surface area measure of three-dimensional sets from digitizations

Johanna Ziegel and Markus Kiderlen

ESTIMATION OF SURFACE AREA AND SURFACE AREA MEASURE OF THREE-DIMENSIONAL SETS FROM DIGITIZATIONS

JOHANNA ZIEGEL[†] AND MARKUS KIDERLEN[∗,‡]

ABSTRACT. A local method for estimating surface area and surface area measure of three-dimensional objects from discrete binary images is presented. A weight is assigned to each $2 \times 2 \times 2$ configuration of voxels and the total surface area of an object is given by summation of the local area contributions. The method is based on an exact asymptotic result that holds for increasing resolution of the digitization. It states that the number of occurrences of a $2 \times 2 \times 2$ configuration is asymptotically proportional to an integral of its "h-function" with respect to the surface area measure of the object. We find explicit representations for these h-functions. Analyzing them in detail, we determine weights that lead to an asymptotic worst case error for surface area estimation of less than 4%. We show that this worst case error is the best possible. Exploiting the local nature of the asymptotic result, we also establish two parametric estimators for the surface area measure. The latter allow to quantify anisotropy of the object under consideration. Simulation studies illustrate the validity of the estimation procedure also for finite, but sufficiently high resolution.

1. INTRODUCTION

Already Gauss used the fact that the area of a planar object is approximately proportional to the number of lattice points contained in it. This observation, extended to three dimensions, is the basis of modern volume approximations based on voxel counts in digital black-and-white images of real-world structures. To approximate the surface area, counts of individual voxels are no longer sufficient. Instead, one determines the number of occurrences of certain patterns – so-called *boundary configurations* – of black and white voxels in $2 \times 2 \times 2$ cubes. Then, a weight is assigned to each configuration of voxels and the total surface area is obtained by a summation of the local area contributions. As the choice of these weights is not unique, there is a number of competing choices, all motivated by geometric arguments. In [11] the weights are chosen as the areas of isosurfaces originating from a marching cubes representation. The weights in [13] are determined by discretizing an integral geometric relation (the Crofton formula) and using a digital version of the Euler-Poincaré characteristic in lattice-line sections. Under the assumption that the object is randomly translated and rotated before digitization, the weights derived in [10] are optimal, in the sense that the obtained estimator is an MVUE (minimal variance unbiased estimator). In the present paper we will also work with randomized positions of the object relative to the main axis of the digitization, but

only allow for random "uniform" translations. Under this assumption there appears to be no unbiased estimator based on weighted configuration counts.

The purpose of this work is to present another estimator of surface area (giving a new set of weights), which has a better worst case behavior than the known ones when the resolution is high. Furthermore, we can also estimate a local version of the surface area, the *surface area measure*. Its normalization (also called the *rose of normal directions*) is the distribution of the outer unit normal of the object in a typical boundary point. If, for instance, the object is a ball, then this measure is proportional to the uniform distribution on the set of unit vectors. The deviation of the surface area measure from uniformity can therefore be used to quantify anisotropy of an object. Both estimators are based on a very general asymptotic result in [6] that was only recently shown. *Asymptotics* is understood here with respect to increasing resolution, or, equivalently, with lattice distance $t > 0$ of the digitization converging to zero. Let N_t be the number of occurrences of the boundary configuration (B, W) of black B and white W voxels in the digitized picture of an object Z. The asymptotic result states that N_t behaves in mean like ct^{-2} as $t \to 0$. Interestingly, the constant c can be given explicitly as an integral

$$c = \int_{S^2} h_{(B,W)}(n) S_2(Z, dn) \tag{1}$$

over the unit sphere S^2 of $\mathbb{R}^3$. Here, $h_{(B,W)}$ is the so-called *h-function* depending only on the configuration (B, W) under consideration. The measure $S_2(Z, \cdot)$ is the surface area measure of Z and is independent of the configuration. The total mass $S(Z) = \int_{S^2} 1 \; S_2(Z, dn)$ is the surface area of Z. Hence, any weighted sum of h-functions, approximating the constant function one, can be used to estimate surface area from configuration counts. This is the key idea to construct a set of weight vectors such that the maximal relative error of the derived surface area estimator is asymptotically minimal. The best possible bound is 3.98%. Simulations suggest that this worst case bound even holds for finite (but reasonably high) resolution $1/t$. We can show that the estimators of surface area proposed in [10] and [13] both have an asymptotic worst case error of 7.3%. The weights proposed in [11] have an asymptotic worst case error of 12.8%. Their weight vectors are substantially different from ours and from each other. The analysis of linear dependence between h-functions in Proposition 3.4 enables us to explain these differences to a large extend. Essentially there are only five types of configurations that contribute to the asymptotic worst case error. The coefficients for two of these types are approximately equal in all approaches, while for the other three they differ substantially. The linear dependence of certain h-functions implies that some weight can be shifted between the latter three types without changing the asymptotic worst case error.

As configuration counts allow to approximate certain integrals with respect to the surface area measure, it is natural to use these counts also for surface area measure estimation. As there are only finitely many $2 \times 2 \times 2$ configurations, yielding finitely many integrals of the form (1) with different integrands, we will have to choose a model with finitely many parameters. This leads to the question to find the number of degrees of freedom that can possibly be determined by the integrals over all h-functions of $2 \times 2 \times 2$ configurations. In other words, we ask for the dimension of the subspace $H_{2 \times 2 \times 2}$, spanned by all h-functions, in the Banach space $C(S^2)$ of all continuous functions on the unit sphere S^2. Proposition 3.3 shows that

$\dim(H_{2\times2\times2}) = 50$. We then suggest two simple models for the surface area measure with 50 degrees of freedom. The first model allows for discrete measures with a prescribed support, the second model has a piecewise constant density. The first model has the advantage of statistical simplicity: we can derive explicit formulae for the model parameters depending on the configuration counts, and show that the derived estimator is asymptotically unbiased, i.e. the mean estimator converges (weakly) to the true surface area measure, as the resolution increases. The second model is interesting because it includes the isotropic case, hence it can more easily be used to detect and quantify deviations from isotropy. The two models are a natural extension of the two-dimensional models suggested in [4] and refine an existing approach of [2] allowing for 26 degrees of freedom. Note that estimation of the surface area measure from counts of *point pairs* (in the 26-neighborhood) instead of counts of $2 \times 2 \times 2$ configurations only determines 13 degrees of freedom. The new methods clearly yield refined estimators, but build on the same lateral resolution.

<p style="text-align:center">*   *   *</p>

We described the new approaches for subsets of three-dimensional space $\mathbb{R}^3$, but as many of the concepts also hold in d-dimensional space, we will present them in this general setting, where appropriate. Similarly, we develop parts of the theory for $n \times n \times n$ configurations, which are natural extensions of the $2 \times 2 \times 2$ case in $\mathbb{R}^3$. It should also be mentioned that asymptotic results always depend on certain regularity conditions of the underlying structure Z. Conditions in the literature are either not made explicit or rather strong, like polygonality or morphological openness and closeness with respect to suitable line segments. As we are working with asymptotic results, the conditions on the set Z in the present paper are very weak, we assume that Z is *gentle*; see Section 2 for a definition. Of course, the validity of the above worst case errors depends on the discretization model chosen. We formulate all results with respect to the common Gauss digitization model. As the underlying asymptotic identity from [6] also holds for the volume-threshold digitization, an extension to this setting is immediate.

Note that the surface area estimators in the present paper are all based on configuration counts and do not require a time consuming reconstruction of the actual object surface. This is the reason why the implementation is straightforward and the algorithms have an excellent time performance. Alternatives are a number of multigrid convergent estimators; see [7] for an overview. These estimators converge to the true surface area as lattice resolution increases. They are, however, computationally more intensive, as they require an explicit approximation of the boundary of the object, which apparently cannot be achieved by local methods. Examples in this spirit are [8], based on global polyhedrization techniques, and [1], where a discrete approximation of the normal vector field is integrated.

<p style="text-align:center">*   *   *</p>

The paper is organized as follows. The main results are presented in Sections 4 and 5. Section 2 introduces notation, the most important definitions and some basic results about convex polytopes, which will be needed in the sequel. In Section 3 we start with a general analysis of the h-function of an $n \times n \times n$ configuration (B, W). Instead of representing it as the difference of two support functions of fairly complicated polytopes, we suggest a compact form as a minimum over the scalar products with only few, so-called *visible* points of the set $B + \breve{W}$. This is a generalization of [4, Corollary 1]. The set of all informative $n \times n \times n$ configurations

can be determined by applying an algorithm suggested in [2]. For each of these configurations, Algorithm 1 in Section 3.1 determines all visible points and thus allows to express its h-function in a compact form, which gives insight into the structure of the h-function. Based on this, we identify linear dependencies between h-functions (Proposition 3.4), and subsequently determine the dimension of $H_{2 \times 2 \times 2}$ (Proposition 3.3). Estimation of surface area is treated in Section 4. Section 5 presents the two mentioned estimators for surface area measure.

2. Basic results and Notation

By S^{d-1} we denote the unit sphere in $\mathbb{R}^d$. For two points $x, y \in \mathbb{R}^d$ let $\langle x, y \rangle$ be the standard scalar product and $[xy] := \{tx + (1-t)y \mid t \in [0,1]\}$ be the line segment with endpoints x and y. 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 $A^c := \mathbb{R}^d \backslash A$ and its topological boundary by ∂A. We write $A + B := \{a + b \mid a \in A, b \in B\}$ for the Minkowski sum of A and B. Let H^+ and H^- be the two closed halfspaces that are bounded by a hyperplane H of $\mathbb{R}^d$. Recall that $A, B \subset \mathbb{R}^d$ are *separated by H*, if $A \subset H^+$ and $B \subset H^-$, or vice versa. They are *strictly* separated by H, if, in addition, $A \cup B$ does not hit H. The positive part of a real valued function f is denoted by $f^+ := \max(f, 0)$. Furthermore, we need some notation from convex geometry. Recall that a (convex) *polytope* is the convex hull of finitely many points in $\mathbb{R}^d$. The *support function* of a *convex body* K (nonempty compact convex set) in $\mathbb{R}^d$ is given by $h(K, \cdot) = \sup_{x \in K} \langle x, \cdot \rangle$. We use this notion also for compact sets $A \subset \mathbb{R}^d$, $A \neq \emptyset$, defining $h(A, \cdot) := h(\text{conv } A, \cdot)$, where conv A denotes the convex hull of A. The support set of the convex body K in direction $u \in S^{d-1}$ is given by $K(u) := \{x \in K \mid \langle x, u \rangle = h(K, u)\}$.

Let $Z \subset \mathbb{R}^d$ be the set that represents the real-world structure to be analyzed. Throughout the paper we assume that $Z \subset \mathbb{R}^d$ is compact and *gentle*. This means that the compact set Z satisfies

(i) $\mathcal{H}^{d-1}(N(\partial Z)) < \infty$,
(ii) for $\mathcal{H}^{d-1}$-almost all $z \in \partial Z$ there are two non-degenerate open balls touching in z such that one of them is contained in Z and the other in Z^c.

Here $\mathcal{H}^{d-1}$ is the $(d-1)$-dimensional Hausdorff measure in $\mathbb{R}^d$ and $N(\partial Z)$ is the *reduced normal bundle* of ∂Z; for further details see [6]. The class of gentle sets is rather large. It contains for instance all convex bodies 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. By (ii), almost all boundary points of Z have a unique outer unit normal. The *surface area measure $S_{d-1}(Z, \cdot)$ of Z* is the image measure of $\mathcal{H}^{d-1}(\cdot \cap \partial Z)$ under the spherical image map. Hence, for a Borel set $A \subset S^{d-1}$, $S_{d-1}(Z, A)$ is the surface area of the set of all boundary points of Z having an outer normal in A. It follows from [6, Corollary 2.2] that $S_{d-1}(Z, \cdot)$ satisfies the centroid condition

$$\int_{S^{d-1}} n \, S_{d-1}(Z, dn) = 0. \tag{2}$$

A random digitization of Z is obtained as the intersection of Z with the scaled and randomly translated regular lattice $t(U + \mathbb{Z}^d)$, where U is a uniform random variable in the unit cube. This means that we work with the Gauss digitization model and the midpoints of the voxels, instead of thinking of the digital image as a collection of small cubes. We call $\mathbb{Z}^d \cap [0, n-1]^d$ the *n-lattice cube* in $\mathbb{R}^d$. An n^d *configuration*

is a pair (B, W) of disjoint sets B, W such that $B \cup W$ is the n-lattice cube, $n \geq 2$. If both sets B and W are nonempty, (B, W) is called a *boundary* configuration. We will refer to points in B as *black* points (belonging to the object) and to points in W as *white* (background) points. Theorem 5 in [6], already mentioned in the introduction, can now be made precise. For an n^d boundary configuration (B, W) we set

$$h_{(B,W)} := (-h(B + \check{W}, \cdot))^+|_{S^{d-1}}$$

and call it the *h-function* of (B, W). It would also be natural to define $h_{(B,W)}$ as a (positive homogeneous) function on $\mathbb{R}^d$. Whenever we use this extension on $\mathbb{R}^d$, it will be explicitly stated.

Theorem 2.1. *Let $Z \subset \mathbb{R}^d$ be a compact gentle set. Let U be uniformly distributed in the unit cube, $B, W \subset \mathbb{Z}^d$ two non-empty finite subsets of $\mathbb{Z}^d$. Then the number*

$$N_t = \sum_{x \in t(U + \mathbb{Z}^d)} \mathbf{1}_{\{x + tB \subset Z \cap t(U + \mathbb{Z}^d),\, x + tW \subset Z^c \cap t(U + \mathbb{Z}^d)\}}, \quad t > 0,$$

of occurrences of (B, W) in the digitization of Z satisfies

$$\lim_{t \to 0^+} t^{d-1} \mathbb{E} N_t = \int_{S^{d-1}} h_{(B,W)}(n) S_{d-1}(Z, dn). \tag{3}$$

The h-function of a configuration has an intuitive geometrical interpretation, which we describe for $d = 3$. Fix a direction $u \in S^2$. Consider the union S of all planes with normal u that separate B and W such that u points away from B. Then $h_{(B,W)}(u)$ is the width of S in direction u. Of course it can happen that $S = \emptyset$, then $h_{(B,W)}(u) = 0$. We call a configuration *informative*, if it is a boundary configuration and there is a hyperplane which strictly separates B and W. Theorem 2.1 shows that non-informative boundary configurations are asymptotically negligible, as the right hand side of (3) is zero. Hence, we focus on informative configurations for our estimation procedures; see also [10] and the comment in Section 4.2 on this matter. Furthermore, we identify twin configurations for estimation purposes, as they have the same h-function. A configuration (B', W') is called *twin* of an $n \times n \times n$ configuration (B, W), if

$$B' := \rho_n(W), \quad W' := \rho_n(B),$$

where ρ_n denotes the reflection at the point $\left(\frac{n-1}{2}, \frac{n-1}{2}, \frac{n-1}{2}\right) \in \mathbb{R}^3$.

In order to use Theorem 2.1 for estimation purposes, it is necessary to find a simpler form of the right hand side of equation (3). Note that $h(B + \check{W}, \cdot) = h(\mathrm{conv}(B + \check{W}), \cdot)$ and for any convex body K we have

$$(-h(K, \cdot))^+ = h(K \cup \{0\}, \cdot) - h(K, \cdot). \tag{4}$$

The notion of *visibility* can be used to simplify the right hand side of this equation.

Definition 2.1. *Let $K \subset \mathbb{R}^d$ be a convex body with $0 \notin K$. We call $p \in K$ visible (from the origin), if $[0p] \cap K = \{p\}$.*

The use of the term *visible* is motivated by the idea that the visible points of K are exactly those points of K that can be seen by an observer located at the origin. Clearly, any visible point of K is a boundary point of K. If a hyperplane separates the set of visible points and the origin, it also separates K and the origin. Furthermore, if there is a hyperplane through a point $p \in K$, separating K and 0

and not containing 0, then p is visible. The converse is true if K is a polytope, and we will repeatedly work with this characterization of visible points.

Lemma 2.2. *Let K be a polytope with vertices $p_1, \ldots, p_k$ and $0 \notin K$. Possibly after renumbering we may assume that $\{p_1, .., p_l\}$, $l \leq k$, are the visible vertices of K. Then we have*

$$h(K \cup \{0\}, \cdot) - h(K, \cdot) = \min_{i \in \{1, \ldots, l\}} \langle -p_i, \cdot \rangle^+. \tag{5}$$

Proof. As K is a polytope,

$$h(K, u) = \sup_{x \in K} \langle x, u \rangle = \max_{i \in \{1, \ldots k\}} \langle p_i, u \rangle, \qquad u \in \mathbb{R}^d.$$

The convex body $K_0 := \operatorname{conv}(K \cup \{0\})$ is also a polytope, with vertices in $\{p_1, \ldots, p_k, 0\}$. The support function of K_0 therefore has the form

$$h(K_0, \cdot) = \max\{\langle p_1, \cdot \rangle, \ldots, \langle p_k, \cdot \rangle, 0\} = \max\{h(K, u), 0\}.$$

For $u \in \mathbb{R}^d$ define $f(u) := \min_{i \in \{1, \ldots, l\}} \langle -p_i, u \rangle^+$. Let U be the set all $u \in \mathbb{R}^d$ such that there exists a non-visible p_j (i.e. $j > l$) with $h(K, u) = \langle p_j, u \rangle$. For $u \in U$ we have $h(K, u) \geq 0$, hence $h(K_0, u) = h(K, u)$ and $h(K_0, u) - h(K, u) = 0$. Fix one non-visible p_j such that $h(K, u) = \langle p_j, u \rangle$. Suppose that for all $i \in \{1, \ldots, l\}$ we have $\langle p_i, u \rangle < 0$. We derive a contradiction to this assumption implying $f(u) = 0$ and hence that equation (5) holds for all $u \in U$. As p_j is not visible we can find a point $p \in [0p_j] \cap \partial K$ in a face of K, whose vertices are all visible. Hence there are indices $j_1, \ldots, j_t \in \{1, \ldots, l\}$ and $\lambda_{j_1}, \ldots, \lambda_{j_t} \in [0, 1]$ summing up to one such that

$$p = \sum_{n=1}^{t} \lambda_{j_n} p_{j_n}.$$

At least one of the λ_{j_n} is positive, so our assumption gives $\langle p, u \rangle < 0$. On the other hand there is a $\mu \in (0, 1)$ such that $p = \mu p_j$. This yields $\langle p, u \rangle = \mu \langle p_j, u \rangle \geq 0$, which is the desired contradiction.

For u in the complement of U there always exists a visible p_i such that $h(K, u) = \langle p_i, u \rangle$. If $\langle p_i, u \rangle \geq 0$ equation (5) follows immediately. If $\langle p_i, u \rangle < 0$ we have

$$h(K_0, u) - h(K, u) = -h(K, u) = -\langle p_i, u \rangle = \min_{j \in \{1, \ldots, l\}} \langle -p_j, u \rangle = f(u). \qquad \square$$

In the context of h-functions of configurations the polytope K appears as a Minkowski sum of two polytopes L and M. In view of Lemma 2.2 we therefore want to characterize visible vertices of a sum of two polytopes. The characterization of vertices in Lemma 2.3 is based on standard arguments, which need not be repeated here. Lemma 2.4 characterizes visible points of polytopal Minkowski sums.

Lemma 2.3. *Let L and M be polytopes in $\mathbb{R}^d$. Then p is a vertex of $L + M$ if and only if the following two conditions hold.*

(i) *there are vertices l and m of L and M, respectively, such that $p = l + m$,*
(ii) *if $p = l' + m'$ for some $l' \in L$, $m' \in M$ then $l' = l$, $m' = m$.*

Lemma 2.4. *Let L and M be polytopes in $\mathbb{R}^d$ and $p = l + m \in L + M$ with $l \in L$ and $m \in M$. Then p is visible if and only if there is a unit vector u such that*

$$\langle m, u \rangle = h(M, u) < -h(L, u) = -\langle l, u \rangle.$$

Proof. Suppose p is visible, i.e. there is a unit vector u such that $\langle p, u \rangle = h(L + M, u) < 0$. This implies

$$h(L, u) + h(M, u) = h(L + M, u) = \langle l, u \rangle + \langle m, u \rangle < 0.$$

As $\langle l, u \rangle \leq h(L, u)$ and $\langle m, u \rangle \leq h(M, u)$ we obtain equality in both cases and hence the desired result. The other direction is straightforward. □

In the following we restrict our attention to the three-dimensional space ($d = 3$). We suggest an algorithm for finding the *visible vertices of an informative $n \times n \times n$ configuration* (B, W), which are, by definition, the visible vertices of the polytope $\mathrm{conv}(B + \check{W})$. This algorithm is then applied to the special case of $2 \times 2 \times 2$ configurations.

3. VISIBILITY FOR CONFIGURATIONS

3.1. $n \times n \times n$ **configurations.** Consider an informative boundary configuration (B, W) in $\mathbb{R}^3$. Lemmas 2.2, 2.3 and 2.4 imply that

$$h_{(B,W)} = \min_{i \in \{1, \dots, k\}} \langle -p_i, \cdot \rangle^+,$$

where the set $\mathcal{P} = \{p_1, \dots, p_k\}$ of all visible vertices of (B, W) is characterized as follows.

Proposition 3.1. *The point $p = -w + b$, $b \in B$ and $w \in W$, is a visible vertex of (B, W) if and only if*
   (i) *$b \notin \mathrm{conv}(B \setminus \{b\})$ and $w \notin \mathrm{conv}(W \setminus \{w\})$,*
   (ii) *if $p = -w' + b'$ for some $b' \in \mathrm{conv}\, B$, $w' \in \mathrm{conv}\, W$ then $w' = w$, $b' = b$,*
   (iii) *there is a hyperplane separating B and W, containing b but no point of W, and a parallel separating hyperplane containing w but no point of B.*

An *n-lattice plane* is a plane which contains at least three non-collinear points of the n-lattice cube $\mathbb{Z}^3 \cap [0, n-1]^3$. Analogously an *n-lattice line* is a line, which contains two distinct points of the n-lattice cube. According to [2, Proposition 2] an $n \times n \times n$ boundary configuration (B, W) is informative if and only if there is a separating n-lattice plane h such that either
   (a) *h only hits one of the sets B and W, or*
   (b) *there is an n-lattice line $g \subset h$ separating $B \cap h$ and $W \cap h$, only hitting one of them.*

We call a separating n-lattice plane h with one of these properties an *admissible* plane. To determine the set $\mathcal{P}$ of all visible vertices for a given informative configuration (B, W) we have to investigate *all* admissible planes. Note that there are only finitely many.

Algorithm 1 (Visible Vertices): Let (B, W) be an informative $n \times n \times n$ configuration. Start by setting $\mathcal{P} = \emptyset$.
   (α) Choose an admissible plane h not treated before,
   (β.1) if h satisfies (a) above and u is its unit normal pointing away from B then consider the two support sets $B' = (\mathrm{conv}\, B)(u)$ and $W' = (\mathrm{conv}\, W)(-u)$. Insert all vertices of $\mathrm{conv}(B' + \check{W}')$ in $\mathcal{P}$,

(β.2) if h satisfies (b) and v is the unit normal of g in h pointing away from $B \cap h$
then consider $B' = (\mathrm{conv}(B \cap h))(v)$ and $W' = (\mathrm{conv}(W \cap h))(-v)$. Insert
all vertices of $\mathrm{conv}(B' + \check{W}')$ in $\mathcal{P}$.

(γ) repeat (α) and (β) until all admissible planes are treated.

The procedure in (β.2) is inspired by the analogous approach in $\mathbb{R}^2$ in [3].

Proposition 3.2. *The output set $\mathcal{P}$ of Algorithm 1 only contains visible vertices of*
(B, W). For $n = 2$, $\mathcal{P}$ contains all visible vertices of (B, W).

Proof. First we prove that only visible vertices are collected in $\mathcal{P}$. Let h be an
admissible plane with normal u pointing away from B. Suppose h only hits one of
the sets B and W, i.e. we apply step (β.1). Let p be a vertex of $\mathrm{conv}(B' + \check{W}')$. By
Lemma 2.3 there are unique vertices $b' \in B'$ and $w' \in W'$, respectively, such that
$p = b' - w'$. Note that by construction b' and w' are also vertices of $\mathrm{conv}\, B$ and
$\mathrm{conv}\, W$, respectively, and

$$\langle b', u \rangle = h(B, u) < \langle w', u \rangle = -\langle w', -u \rangle = -h(W, -u) = -h(\check{W}, u),$$

hence p is visible. Let $\bar{b} \in B$, $\bar{w} \in W$ be vertices of $\mathrm{conv}\, B$, $\mathrm{conv}\, W$ respectively
such that $p = \bar{b} - \bar{w}$. Then $h(B, u) + h(W, -u) = h(B + \check{W}, u) = \langle p, u \rangle = \langle \bar{b}, u \rangle + \langle \bar{w}, -u \rangle$. As $\langle \bar{b}, u \rangle \leq h(B, u)$ and $\langle \bar{w}, -u \rangle \leq h(W, -u)$ we obtain $\langle \bar{b}, u \rangle = h(B, u)$ and
$\langle \bar{w}, -u \rangle = h(W, -u)$, which means $\bar{b} \in B'$ and $\bar{w} \in W'$, and thus $b' = \bar{b}$ and $w' = \bar{w}$.
Hence, points collected in (β.1) are (visible) vertices of (B, W) by Lemma 2.3.

Suppose now that h hits both sets B and W, and there is an n-lattice line $g \subset h$,
separating $B \cap h$ and $W \cap h$, only hitting one of them. So we are applying step (β.2).
Without loss of generality we assume that $B' \subset g$. Let $g' \subset h$ be a line parallel to
g such that $W' \subset g'$. Note that g' is not necessarily an n-lattice line, as $|W'| = 1$
is possible. Rotating h around g without penetrating one of the points of $B \cup W$
yields a plane h_1, which contains B' and separates B and W. Let u' be its normal,
pointing away from B. Now we have $B' = (\mathrm{conv}\, B)(u')$ and $W' = (\mathrm{conv}\, W)(-u')$.
The same argument as in the first paragraph yields that all points in $\mathrm{conv}(B' + \check{W}')$
are visible vertices.

It remains to show that all visible vertices are collected in $\mathcal{P}$ when $n = 2$. Let
$p = b - w$, $b \in B$, $w \in W$ be a visible vertex of (B, W). Without loss of generality
we may assume $|B| \leq |W|$.

If $|B| = 1$, $B = \{b\}$ is a vertex of the unit cube $C = [0, 1]^2$. The point w can
only be one of the three vertices of C that are endpoints of edges starting in b.
Otherwise, Proposition 3.1.(iii) would be violated. Let h be one of the axis-parallel
planes containing $\{b, w\}$. Then h is a 2-lattice plane and p is collected in (β.2) when
this plane h is considered.

For $B = |2|$, $\mathrm{conv}\, B$ is an edge of C. With similar arguments as above, w cannot
be one of the vertices v_1, v_2 of the most distant parallel edge. Hence, p is collected
in step (β.1), when h is the plane spanned by the four vertices of C not contained
in $B \cup \{v_1, v_2\}$.

For $|B| = 3$, we may, without loss of generality, assume $B = \{(0, 0, 0), (0, 1, 0),$
$(0, 0, 1)\}$. If we assume $b = (0, 0, 0)$ then the only plane separating B and W contain-
ing b is the plane spanned by B. This plane contains also $(0, 1, 1) \in W$, so Propo-
sition 3.1.(iii) leads to a contradiction and we have $b \neq (0, 0, 0)$. If $b = (0, 1, 0)$ we
have $w \in \{(0, 1, 1), (1, 0, 0), (1, 0, 1)\}$ as the other points violate Proposition 3.1.(ii).
Actually, $w \neq (1, 0, 1)$, as $h = \{x \in \mathbb{R} \mid \langle x, (1, 1, 0) \rangle = 1\}$ is the only separating

plane containing $(1, 0, 1) \in W$, but it also contains $b = (0, 1, 0) \in B$. Therefore p is collected in step $(\beta.2)$ when h is considered. The remaining case $b = (0, 0, 1)$ is treated the same way.

If $|B| = 4$, conv B is either a facet of C, in which case p is collected in $(\beta.1)$ when the plane h spanned by B is considered; or it is a tetrahedron of the form conv $B = \text{conv}\{(0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1)\}$, say. There is no separating plane containing $(0, 0, 0)$ or $(1, 1, 1)$. Hence $b \neq (0, 0, 0)$ and $w \neq (1, 1, 1)$. For all other b and w the point p is collected in $(\beta.1)$ when $h = \{x \in \mathbb{R}^3 \mid \langle x, (1, 1, 1) \rangle = 1\}$ is considered. $\square$

Remark. We conjecture that Algorithm 1 also yields all visible vertices for $n > 2$, but we are currently not able to give a proof.

3.2. **Closer analysis of $2 \times 2 \times 2$ configurations.** The goal of this section is to understand the form and interplay of the h-functions of informative $2 \times 2 \times 2$ configurations. The purpose of this analysis is twofold. On the one hand we want to determine the degrees of freedom of a model for the surface area measure $S_2(Z, \cdot)$ that can possibly be determined by the integrals over the h-functions; see Proposition 3.3. On the other hand we need precise information about the supports and maxima of the h-functions in order to construct the actual models for the surface area measure. At the same time this information is essential to determine the weights for the surface area estimator. We state two important consequences of the following analysis as propositions and postpone their proofs to the end of this section. Recall that $H_{2 \times 2 \times 2} \subset C(S^2)$ is the linear subspace spanned by all h-functions derived from informative $2 \times 2 \times 2$ configurations.

Proposition 3.3. $\dim(H_{2 \times 2 \times 2}) = 50$.

To state the second proposition we classify $2 \times 2 \times 2$ configurations. When speaking of configurations in the following, we are always referring to $2 \times 2 \times 2$ configurations. A list of all informative configurations can be found in [2, Appendix B]. The numbers in the following definition refer to this list.

Definition 3.1. An informative configuration is called a *configuration of type*

   one, if it has exactly one black point or exactly one white point
      (e.g. configurations no. 1 and 127),
   two, if it has exactly two black points or exactly two white points
      (e.g. configurations no. 3 and 63),
 three, if it has exactly three black points or exactly three white points
      (e.g. configurations no. 7 and 31),
  four, if it has exactly four black and four white points, which are affinely dependent
      (e.g. configuration no. 15),
  five, if it has exactly four black and four white points, which are affinely independent (e.g. configuration no. 23).

Proposition 3.4. *For each configuration (B^1, W^1) of type one there exists a configuration (B^5, W^5) of type five and three configurations $(B^3_{j_1}, W^3_{j_1})$, $(B^3_{j_2}, W^3_{j_2})$, $(B^3_{j_3}, W^3_{j_3})$ of type three such that*

$$h_{(B^1, W^1)} = h_{(B^5, W^5)} + h_{(B^3_{j_1}, W^3_{j_1})} + h_{(B^3_{j_2}, W^3_{j_2})} + h_{(B^3_{j_3}, W^3_{j_3})} \tag{6}$$

on S^2.

Application of Algorithm 1 in Section 3.1 shows that for configurations of type one and five there are three visible vertices and for configurations of type two, three and four there are four of them. For a configuration (B, W) let $\{p_1, \ldots, p_k\} \subset \mathbb{R}^3$, $k \in \{3, 4\}$, be the set of these visible vertices. By definition of the function $h_{(B,W)}$, equation (4) and Lemma 2.2 we have

$$h_{(B,W)}(u) = \min\{\langle -p_1, u \rangle^+, \ldots, \langle -p_k, u \rangle^+\}$$

for all $u \in S^2$. The support of the function $\mathbb{R}^3 \to \mathbb{R}$, $u \mapsto \langle -p_i, u \rangle^+$ is a closed half-space with outer normal p_i. Hence, the support of the homogeneous extension of $h_{(B,W)}$ to $\mathbb{R}^3$ is a closed polyhedral cone with three or four facets, and the support $\operatorname{supp}(h_{(B,W)})$ of $h_{(B,W)}$ is either a spherical triangle or a spherical quadrilateral.

As all configurations of one type are just reflections or rotations by multiples of $\pi/2$ of one another or can be obtained through the twin operation, it suffices to know the structure of the support and the maximum of the h-function of one configuration of each type. We will analyze in detail the h-function of a configuration of type three. As the analysis for the other types is analogous we only give the results in Table 1.

Remark. Identifying twins we have 58 classes of informative configurations of five different types, hence there is the need for some systematic notation. We will number the 58 classes starting with configurations of type five, then four, three, two and one. The function $\tau : \{1, \ldots, 58\} \to \{1, \ldots, 5\}$ assigns to each index the type of configuration it is associated to. We write $(B_j^{\tau(j)}, W_j^{\tau(j)})$ for a configuration of class j of type $\tau(j)$. If in some context only the type is of importance the lower index will be omitted. If on the other hand we are only interested in the class, we omit the upper index. The supports of h-functions are spherical polygons. There are three types of vectors which will be crucial in describing these supports. First there are visible vertices $p_1, \ldots, p_k$ of each configuration. They are the outer facet normals of the positive cone spanned by the support. For a configuration (B^i, W^i) of type $i = \tau(j)$ we will denote them by $p_1^i, \ldots, p_k^i$, $k = k_i \in \{3, 4\}$. If there is the need to talk of more than one configuration of type i at the same time, we will use a second upper index to indicate this, i.e. for a second configuration of type $i = \tau(j')$ we use $p_1^{i,j'}, \ldots, p_k^{i,j'}$. Secondly, the vertices of the spherical support polygon will be needed and denoted by $v_1^i, \ldots, v_k^i$. Finally, we denote the maximum of $h_{(B_j^i, W_j^i)}$ by m_j^i.

Consider the configuration (B^3, W^3) of type three as given in Table 1. An application of Algorithm 1 in Section 3.1 yields the visible vertices

$$p_1^3 := (-1, 0, 1), \ p_2^3 := (-1, 1, 0), \ p_3^3 := (0, -1, 0) \ \text{ and } p_4^3 := (0, 0, -1)$$

of this configuration. The support of $h_{(B^3, W^3)}$ is a spherical quadrilateral with vertices

$$v_1^3 := \frac{1}{\sqrt{3}}(1, 1, 1), \ v_2^3 := \frac{1}{\sqrt{2}}(1, 0, 1), \ v_3^3 := \frac{1}{\sqrt{2}}(1, 1, 0), \ v_4^3 := (1, 0, 0).$$

In order to determine the maximum m^3 of $h_{(B^3, W^3)}$ we first consider the function $h' : S^2 \to \mathbb{R}$, $u \mapsto \min\{\langle -p_1^3, u \rangle^+, \langle -p_2^3, u \rangle^+, \langle -p_3^3, u \rangle^+\}$. The maximum is attained at points $u \in S^2$ where $\langle -p_1^3, u \rangle = \langle -p_2^3, u \rangle = \langle -p_3^3, u \rangle > 0$. There is a unique solution to this system of equations; the maximum $1/\sqrt{6}$ is attained in the direction $m^3 := \frac{1}{\sqrt{6}}(2, 1, 1)$. As $\langle -p_4^3, \cdot \rangle^+$ takes the same value in m^3, this is the maximum of $h_{(B^3, W^3)}$. Roughly, each p_l^3, $l \in \{1, 2, 3, 4\}$, contributes an arc in the boundary of $\operatorname{supp}(h_{(B^3, W^3)})$ with endpoints v_i^3, v_j^3, where $i, j \in \{1, 2, 3, 4\}$. On the spherical

Type i	$p_1^i, \ldots, p_k^i$	$v_1^i, \ldots, v_k^i$	m^i	CP	Number
1	$(-1,0,0)$ $(0,-1,0)$ $(0,0,-1)$	$(1,0,0)$ $(0,1,0)$ $(0,0,1)$	$\frac{1}{\sqrt{3}}(1,1,1)$	yes	8
2	$(-1,0,-1)$ $(0,-1,-1)$ $(-1,0,1)$ $(0,-1,1)$	$\frac{1}{\sqrt{3}}(1,1,-1)$ $(0,1,0)$ $(1,0,0)$ $\frac{1}{\sqrt{3}}(1,1,1)$	$\frac{1}{\sqrt{2}}(1,1,0)$	yes	12
3	$(-1,0,1)$ $(-1,1,0)$ $(0,-1,0)$ $(0,0,-1)$	$\frac{1}{\sqrt{3}}(1,1,1)$ $\frac{1}{\sqrt{2}}(1,0,1)$ $\frac{1}{\sqrt{2}}(1,1,0)$ $(1,0,0)$	$\frac{1}{\sqrt{6}}(2,1,1)$	yes	24
4	$(-1,-1,-1)$ $(-1,1,-1)$ $(-1,-1,1)$ $(-1,1,1)$	$\frac{1}{\sqrt{2}}(1,0,-1)$ $\frac{1}{\sqrt{2}}(1,-1,0)$ $\frac{1}{\sqrt{2}}(1,1,0)$ $\frac{1}{\sqrt{2}}(1,0,1)$	$\frac{1}{\sqrt{6}}(2,1,1)$	no	6
5	$(1,-1,-1)$ $(-1,1,-1)$ $(-1,-1,1)$	$\frac{1}{\sqrt{2}}(0,1,1)$ $\frac{1}{\sqrt{2}}(1,0,1)$ $\frac{1}{\sqrt{2}}(1,1,0)$	$\frac{1}{\sqrt{3}}(1,1,1)$	no	8

TABLE 1. Support and maxima of the h-functions for a typical example of each configuration type. A type of configurations has the *covering property (CP)* if the union of all supports of the h-functions of this type covers S^2. In the column labeled *Number* we give the number of configurations of each type identifying twins.

triangle with vertices m^3, v_i^3, v_j^3 the function $h_{(B^3, W^3)}$ coincides with $\langle -p_l^3, \cdot \rangle$. For an illustration see Figure 1. Figure 2 helps to clarify the relative positions of the supports of a configuration of type one, three and five. The support of the configuration of type three is depicted in dark gray, the support of the configuration of type five in light gray and the support of the configuration of type one is not shaded. There are two (five if we also consider their twins) more configurations of type three, such that the supports of their h-functions lie inside the support of the considered configuration of type one. The twins with three black points all have a black point at $(0,0,0)$; see Figure 3. Hence, in total, there are 24 configurations of type three, which can be grouped (into eight groups) such that the supports of the h-functions of the three configurations in each group cover the support of a configuration of type one.

Now we can give the proof of Proposition 3.4.

Proof of Proposition 3.4. Suppose we are given the configuration (B^1, W^1) as in Table 1. Then we choose (B^5, W^5) to be the configuration of type five as given in Table 1 and $(B^3_{j_1}, W^3_{j_1})$, $(B^3_{j_2}, W^3_{j_2})$, $(B^3_{j_3}, W^3_{j_3})$ as given in Figure 3. The support of
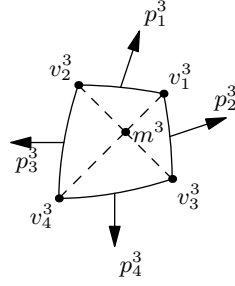
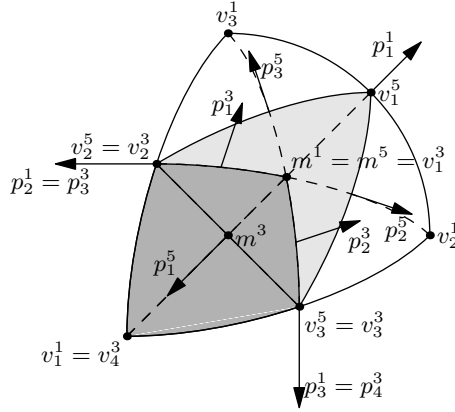FIGURE 1. Support of configuration of type three.

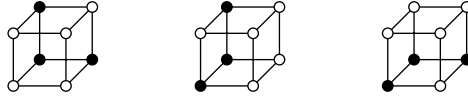FIGURE 2. Support of configurations of type one, three and five.

FIGURE 3. Configurations of type three used in the proof of Proposition 3.4.

$h_{(B^1,W^1)}$ can be split into twelve small spherical triangles such that on each triangle all five functions $h_{(B^1,W^1)}$, $h_{(B^5,W^5)}$, $h_{(B^3_{j_1},W^3_{j_1})}$, $h_{(B^3_{j_2},W^3_{j_2})}$ and $h_{(B^3_{j_3},W^3_{j_3})}$ are linear. By symmetry it suffices to check equation (6) on the two triangles numbered I and II in Figure 4. In triangle I only two of the considered h-functions are not equal to zero; these are $h_{(B^1,W^1)} = \langle -p_2^1, \cdot \rangle$ and $h_{(B^3_{j_1},W^3_{j_1})} = \langle -p_3^{3,j_1}, \cdot \rangle = \langle -p_2^1, \cdot \rangle$. Hence, we obtain the desired identity on triangle I. On triangle II we have $h_{(B^1,W^1)} = \langle -p_2^1, \cdot \rangle$, $h_{(B^5,W^5)} = \langle -p_1^5, \cdot \rangle$, $h_{(B^3_{j_1},W^3_{j_1})} = \langle -p_2^{3,j_1}, \cdot \rangle$ and $h_{(B^3_{j_2},W^3_{j_2})} = h_{(B^3_{j_3},W^3_{j_3})} = 0$. Going back to the respective definitions we see that

$$-p_1^5 + (-p_1^{3,j_1}) = (1,-1,-1) + (-1,0,1) = (0,-1,0) = -p_2^1,$$

hence we obtain the claim. $\qquad\square$

We included Figure 5 to visualize the relative positions of the supports of configurations of type one, two and four. The support of the configuration of type four is depicted in middle gray, the one of type two in light gray.

Summarizing the analysis of the supports of the different h-functions, we obtain a triangulation of the sphere with $12 \cdot 8 = 96$ triangles. We call this triangulation the *support triangulation of S^2 by $2 \times 2 \times 2$ configurations* and each individual triangle
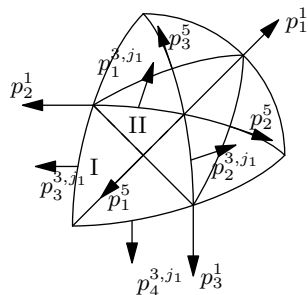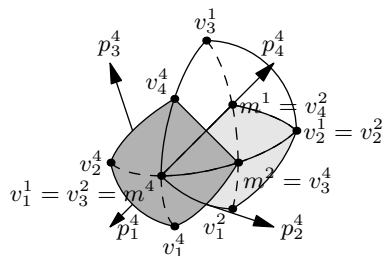
FIGURE 4. Proposition 3.4.

FIGURE 5. Support of configurations of type one, two and four.

a *support triangle*. Any of the h-functions restricted to a support triangle is of the form $h_{(B,W)} = \langle p, \cdot \rangle$ for some suitable $p \in \mathbb{R}^3$.

Proof of Proposition 3.3. Proposition 3.4 shows that the dimension of $H_{2 \times 2 \times 2}$ is at most

$$58 - \#\{ \text{ h-functions arising from configurations of type one } \} = 50.$$

In order to see that the dimension is in fact 50, it suffices to note that the maximum of any h-function of type two, three, four or five is not contained in the interior of the support of any other h-function of type two, three, four or five, i.e. if we evaluate the 50 h-functions of type two, three, four and five at the 50 maxima of the h-functions of type two, three, four and five, we obtain a regular diagonal (50×50)-matrix, which obviously has rank 50. □

4. ESTIMATION OF SURFACE AREA

In this section we discuss the estimation of the surface area of $Z \subset \mathbb{R}^3$ from counts of $2 \times 2 \times 2$ configurations. The relative worst case error of the estimator can be determined asymptotically, independently of the shape of the considered object.

As explained in the introduction any linear combination of h-functions close to one can be used to estimate the surface area of a compact gentle set $Z \subset \mathbb{R}^3$ by configuration counts. For a configuration (B_j, W_j) we abbreviate $h_{(B_j, W_j)}$ by h_j. Let $N_{t,j}$ be the number of occurrences of configurations of class j in the digitized image of Z. Define a vector $n = (n_1, \ldots, n_{58}) \in \{1, 2\}^{58}$ such that $n_j = 1$ if there is only one configuration in class $j \in \{1, \ldots, 58\}$ (the configuration in this class is coinciding with its twin) and $n_j = 2$ otherwise. The estimation procedure works as follows. First we omit all counts of configurations of type one as they are asymptotically redundant. Of course this is a loss of information, but apart

from using these counts for a check on sufficient accuracy it does not seem sensible to incorporate them in the estimation procedure; see Section 4.2. For a vector of weights $\Lambda := (\lambda_1, \ldots, \lambda_{50}) \in \mathbb{R}^{50}$ define $H^\Lambda(u) := \sum_{j=1}^{50} \lambda_j h_j(u)$. By Theorem 2.1

$$\hat{S}(Z) := t^2 \sum_{j=1}^{50} \frac{\lambda_j}{n_j} N_{t,j} \tag{7}$$

is an asymptotically unbiased estimator for $\int_{S^2} H^\Lambda(u) S_2(Z, du)$. The overall goal is to find a weight vector $\Lambda \in \mathbb{R}^{50}$ such that

$$H^\Lambda(u) := \sum_{j=1}^{50} \lambda_j h_j(u) \approx 1 \ \text{ for all } \ u \in S^2. \tag{8}$$

This is motivated by the trivial estimates

$$H_m^\Lambda S(Z) \leq \int_{S^2} H^\Lambda(u) S_2(Z, du) \leq H_M^\Lambda S(Z),$$

with $H_M^\Lambda := \max_{u \in S^2} H^\Lambda(u)$ and $H_m^\Lambda := \min_{u \in S^2} H^\Lambda(u)$, which imply

$$H_m^\Lambda \leq \frac{\int_{S^2} H^\Lambda(u) S_2(Z, du)}{S(Z)} \leq H_M^\Lambda. \tag{9}$$

The inequalities in (9) state that the relative worst case error is asymptotically bounded by

$$\max\{|1 - H_m^\Lambda|, |1 - H_M^\Lambda|\} = \|H^\Lambda - 1\|_\infty, \tag{10}$$

where $\|f\|_\infty = \max_{u \in S^2} |f(u)|$ denotes the maximum norm of a real bounded function f on the sphere. We will determine the minimum Λ_0 of the function $\Lambda \mapsto \|H^\Lambda - 1\|_\infty$. Using this optimal weight vector, (9) will imply that the relative worst case error is asymptotically less than 4%. Although the error bound of 4% is an asymptotic result, simulations suggest that it is also realistic for positive lattice distances $t > 0$, as the deviation of $\hat{S}(Z)$ from $\int_{S^2} H^\Lambda(u) S_2(Z, du)$ is typically much smaller than 4%.

4.1. **Construction of the coefficient vector.** Recall that we omit configurations of type one, and therefore the notion *configurations* is used in this section as a short term for configurations of type two, three, four or five. Recall from Section 3.2 that the support triangulation of $2 \times 2 \times 2$ configurations consists of 96 support triangles, such that any h-function restricted to a support triangle is linear. For each support triangle T there are exactly three h-functions which do not vanish on T; see Figure 6 for illustration. Minimizing the function in (10) does not uniquely determine the coefficient vector Λ. To be precise, the coefficients for configurations of type two, three and four are determined uniquely, whereas we are left with some choice for the coefficients for configurations of type five. First we will calculate the optimal coefficients for the restrictions of H^Λ to single support triangles. Using symmetry arguments this is then extended to determine the set of optimal coefficients for H^Λ.

As one can check with the help of Figure 6 there are only two types of support triangles in the sense that all other support triangles can be obtained from these two by rotation and reflection. Any support triangle that is obtained as intersection of supports of configurations of type two, three and four will be called *support triangle of the first type*. *Support triangles of the second type* are obtained as intersection of

supports of configurations of type two, three and five. H^Λ restricted to a support triangle $T := \{n \in S^2 \mid \langle p_i, n \rangle \geq 0 \text{ for } i = 1, 2, 3\}$ is of the form

$$H^\lambda(n) := \lambda_1 \langle p_1, n \rangle + \lambda_2 \langle p_2, n \rangle + \lambda_3 \langle p_3, n \rangle = H^{\lambda(a)}(n) = \langle a, n \rangle, \qquad n \in T,$$

where $\lambda = (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$ is a subvector of Λ, $a = \lambda_1 p_1 + \lambda_2 p_2 + \lambda_3 p_3$ and $\lambda(a) = (\lambda_1, \lambda_2, \lambda_3)$ iff $a = \lambda_1 p_1 + \lambda_2 p_2 + \lambda_3 p_3$. This is a slight abuse of notation as, strictly speaking, λ is a subvector of Λ of the form $(\lambda_{i_1}, \lambda_{i_2}, \lambda_{i_3})$. However, for the following arguments, the indices are irrelevant. Set $g_1(a) := \min_{n \in T} H^{\lambda(a)}(n)$ and $g_2(a) := \max_{n \in T} H^{\lambda(a)}(n)$. We are looking for the minimum over all $a \in \mathbb{R}^3$ of

$$\max_{n \in T} |H^{\lambda(a)}(n) - 1| = \max\{|g_1(a) - 1|, |g_2(a) - 1|\}$$

$$\geq \max\{1 - g_1(a), g_2(a) - 1\} =: G(a).$$

In this vein, we first determine the functions g_1 and g_2 explicitly, and show that G attains a unique minimum on the positive cone $D := \bigcup_{\alpha \geq 0} \alpha T$ spanned by T.

Lemma 4.1. *Let W be the 3×3 matrix, which has the vertices w_1, w_2, w_3 of T as columns. Then*

$$g_1(a) = \min_{n \in T} H^{\lambda(a)}(n) = \min_{i=1,2,3} \langle a, w_i \rangle.$$

If $a \in D$, then

$$g_2(a) = \max_{n \in T} H^{\lambda(a)}(n) = \|a\|, \tag{11}$$

otherwise

$$g_2(a) = \max_{\{i,j,k\}=\{1,2,3\}} \left(f_i \cos \varphi_i^* + \left(\frac{f_k - f_j \cos \varphi_i}{\sin \varphi_i} \right) \sin \varphi_i^* \right).$$

Here $f := W^T a$, the angles between the vertices of T are denoted by φ_i, $i = 1, 2, 3$, and

$$\varphi_i^* = \min \left\{ \varphi_i, \left(\arctan \frac{f_k - f_j \cos \varphi_i}{f_j \sin \varphi_i} \right)^+ \right\}.$$

The function G is convex on $\mathbb{R}^3$ and attains a unique minimum on D.

Proof. To determine the minimum of $H^{\lambda(a)}$ on T, consider an arbitrary point n in T. It can be written as a linear combination $n = \nu_1 w_1 + \nu_2 w_2 + \nu_3 w_3$, where $\nu_1, \nu_2, \nu_3 \geq 0$ and $\nu_1 + \nu_2 + \nu_3 \geq 1$. This shows

$$H^{\lambda(a)}(n) = \langle a, n \rangle \geq \min_{i=1,2,3} H^{\lambda(a)}(w_i).$$

Hence

$$g_1(a) = \min_{n \in T} H^{\lambda(a)}(n) = \min_{i=1,2,3} H^{\lambda(a)}(w_i) = \min_{i=1,2,3} \langle a, w_i \rangle,$$

which holds for all $a \in \mathbb{R}^3$. As

$$2H^{\lambda(a)}(n) = -\|n - a\|^2 + \|a\|^2 + 1,$$

the maximum of $H^{\lambda(a)}$ on T is attained in the point $\tilde{a} \in T$ with minimal distance from a. If $a \in D$, then $\tilde{a} = \frac{1}{\|a\|} a$ and

$$g_2(a) = \max_{n \in T} H^{\lambda(a)}(n) = H^{\lambda(a)}(\tilde{a}) = \|a\|.$$

This shows (11). If $a \notin D$, then $\tilde{a}$ is contained in a side of T, say in the arc with endpoints w_j and w_k, $j \neq k \in \{1, 2, 3\}$. Thus, for $i \in \{1, 2, 3\} \setminus \{j, k\}$, we have

$$\tilde{a} = \cos(\varphi) w_j + \sin(\varphi) w, \qquad 0 \leq \varphi \leq \varphi_i$$

with $\varphi_i = \arccos\langle w_j, w_k\rangle$, and

$$w = \frac{w_k - \langle w_j, w_k\rangle w_j}{\|w_k - \langle w_j, w_k\rangle w_j\|} = \frac{w_k - (\cos\varphi_i)w_j}{\sin\varphi_i}.$$

Elementary calculation gives that

$$H^{\lambda(a)}(\tilde{a}) = f_j \cos\varphi + \left(\frac{f_k - f_j \cos\varphi_i}{\sin\varphi_i}\right)\sin\varphi$$

attains its maximum at φ_i^* as given in the lemma.

It remains to show the claimed properties of G. The convexity of G follows immediately from the convexity of $-g_1$ and g_2. A standard compactness argument shows that G attains a minimum $a^* \in D$. As the functions g_1 and g_2 are positive homogeneous, we have

$$G(a^*) = 1 - g_1(a^*) = g_2(a^*) - 1. \tag{12}$$

Suppose that $a_1 \neq a_2$ are two minima of G in D. By convexity, G is constant (minimal) for all points in $[a_1 a_2] \subset D$. Applying (12) to a_1, a_2 and $(a_1 + a_2)/2$ yields

$$\|a_1 + a_2\| = \|a_1\| + \|a_2\|$$

and $\|a_1\| = \|a_2\|$. This implies that $a_1 = a_2$, which is a contradiction. $\square$

We numerically minimized G over D for a support triangle T_1 of first type and T_2 of second type using [12]. In both cases it turns out that the unique minimum of G on D lies in the interior of D and is therefore a strict global minimum of G on $\mathbb{R}^3$ by convexity. Furthermore,

$$\max_{n \in T} |H^{\lambda(a^*)}(n) - 1| = G(a^*) \tag{13}$$

in both cases, which implies that $\lambda = \lambda(a^*)$ is the unique minimum of $\lambda \mapsto \max_{n \in T} |H^\lambda(n) - 1|$ on $\mathbb{R}^3$.

The numerical values are

$$\min_{a \in D(T_1)} G^{T_1}(a) = 0.0398 \qquad \text{and} \qquad \min_{a \in D(T_2)} G^{T_2}(a) = 0.0241.$$

It is clear that the value 0.0398 is a lower bound for the asymptotic relative error $\min_\Lambda \|H^\Lambda - 1\|_\infty$. For the first type of support triangles the optimal coefficients are $\lambda^2 = 1.3579$, $\lambda^3 = 2.3519$ and $\lambda^4 = 0.9602$, where the upper index indicates the type of configuration the coefficient is associated to. By symmetry it follows that choosing all coefficients for configurations of type two, three and four like this, we obtain $\max_{\mathcal{T}_1} |H^\Lambda - 1| = 0.0398$, where $\mathcal{T}_1$ is the union of all support triangles of first type. The uniqueness of the minimum shows that any other choice of coefficients will increase the error.

For the coefficient λ^5 for configurations of type five we have some choice. Consider the function g_1 as given by Lemma 4.1 for a support triangle of second type with λ^2, λ^3 fixed. Solving this function for λ^5 yields the lower bound 1.6631 for the possible values of λ^5 that we can choose without increasing the asymptotic relative error. Solving also the expression for g_2 for λ^5 given in Lemma 4.1 one can check that for all $\lambda^5 \in [1.6631, 1.7452]$, we have $1 - g_1(a(\lambda)) = 0.0398$ and $0.0076 \leq g_2(a(\lambda)) - 1 \leq 0.0398$. Summarizing, the weights that lead to a minimal asymptotic worst case error of 0.0398 are

$$\lambda^2 = 1.3579, \quad \lambda^3 = 2.3519, \quad \lambda^4 = 0.9602, \quad \lambda^5 \in [1.6631, 1.7452]. \tag{14}$$
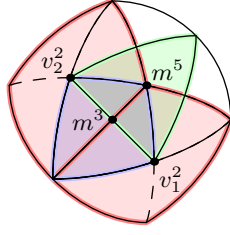
FIGURE 6. Support triangles and corresponding configurations.

Therefore we suggest to estimate the surface area of Z using $\hat{S}(Z)$ as given in (7) with weights $\Lambda = (\lambda_1, \ldots, \lambda_{50})$, where $\lambda_j = \lambda^{\tau(j)}$ for $j = 1, \ldots, 50$.

4.2. **Configurations of type one.** The above considerations attribute weight zero for configuration counts of type one. Alternatively, the dependence properties of h-functions can be used to assign positive weights to configurations of type one while correcting for this by decreasing some of the other weights. This gives rise to a one-parametric family $(\overline{\lambda}^1(r), \ldots, \overline{\lambda}^5(r))$, $r \in [0, 1]$, of weights, with the property that the maximal relative error of the corresponding estimator is asymptotically minimal (and independent of r). Proposition 3.4 yields

$$\lambda^5 h_{(B^5, W^5)} + \lambda^3 \left(h_{(B^3_{j_1}, W^3_{j_1})} + h_{(B^3_{j_2}, W^3_{j_2})} + h_{(B^3_{j_3}, W^3_{j_3})} \right)$$
$$= \lambda^5 r h_{(B^1, W^1)} + \lambda^5 (1 - r) h_{(B^5, W^5)}$$
$$+ (\lambda^3 - \lambda^5 r) \left(h_{(B^3_{j_1}, W^3_{j_1})} + h_{(B^3_{j_2}, W^3_{j_2})} + h_{(B^3_{j_3}, W^3_{j_3})} \right)$$

for all $r \in [0, 1]$. Hence,

$$\begin{aligned} \overline{\lambda}^1(r) = \lambda^5 r, \qquad \overline{\lambda}^2(r) = \lambda^2, \qquad \overline{\lambda}^3(r) = \lambda^3 - \lambda^5 r \\ \overline{\lambda}^4(r) = \lambda^4, \qquad \overline{\lambda}^5(r) = \lambda^5 (1 - r) \end{aligned} \tag{15}$$

yield the same asymptotic relative error for all $r \in [0, 1]$, where λ^2, λ^3, λ^4 and λ^5 are given by (14). In Table 2 we give all the weights that lead asymptotically to a minimal relative worst case error.

Type of configuration	Weight
one	$\overline{\lambda}^1 = sr$
two	$\overline{\lambda}^2 = 1.3579$
three	$\overline{\lambda}^3 = 2.3519 - sr$
four	$\overline{\lambda}^4 = 0.9602$
five	$\overline{\lambda}^5 = s(1 - r)$

TABLE 2. Weights for surface area estimator that minimize the asymptotic relative worst case error. The two parameters obey $0 \leq r \leq 1$ and $1.6631 \leq s \leq 1.7452$. The asymptotic relative error does not depend on the choice of r, but it may depend on s. However, its worst case bound is independent of s.

The dependence of the surface area estimator on r and s in a simulation example is illustrated in Figures 7 and 8. The first figure shows the estimated surface area of a unit ball, calculated from a random digitization with lattice spacing 0.1. (This means that approximately 4200 black points represent the digitized ball). For the estimation we use $\bar{\lambda}^1(r,s), \bar{\lambda}^2(r,s), \bar{\lambda}^3(r,s), \bar{\lambda}^4(r,s), \bar{\lambda}^5(r,s)$ and alternatingly fix $r \in [0,1]$ or $s \in [1.6631, 1.7452]$. The experiment is repeated 10 times with a randomized location of the digitizing lattice. The horizontal black line shows the true surface area 4π of the unit ball. Figure 8 is obtained in the same way, but digitizing a cylinder with radius 1 and height 2, which is not aligned with any coordinate direction. The lattice spacing is 0.055, which translates to approximately 38000 black points. For the ball the optimal choice for (r,s) appears to be $(0, 1.6631)$, whereas for the cylinder it is rather some value close to $(1, 1.7452)$. It is an interesting open question, whether there is a link between the accuracy of the estimate of the surface area and its variation for different values of (r,s).

4.3. **Isotropic objects and comparison with known weights.** Other approaches to surface area estimation based on weighted counts of $2 \times 2 \times 2$ configurations can be found in [10], [13] and [11]. In contrast to our method they also assign positive weights to certain *non-informative* configurations. These are negligible asymptotically and often also in practical applications. For instance, 97.13% of the 319,223 observed boundary configurations in a 3D rendering of a binarized MR-image of a human brain (in [10]) were informative. Hence, the weights for non-informative configurations are comparably immaterial, at least for sufficiently high resolution. It is therefore sensible to compare the weights for the five types of *informative* configurations. Both [10] and [13] determine the weights such that the estimator is asymptotically unbiased for isotropic objects. We can adapt our weights to fulfill this assumption by dividing them by

$$C(r,s) := \frac{\int_{S^2} H^{\Lambda(r,s)}(u) S_2(B^2, du)}{S(B^2)},$$

where $B^2 = \{x \in \mathbb{R}^3 | \|x\| \leq 1\}$ is the unit ball. By Proposition 3.4, the function $C(r,s) = C(s)$ only depends on s. We obtain for example

$$C(1.6631) = 1.0066, \qquad C(1.7036) = 1.0094, \qquad C(1.7452) = 1.0123.$$

We denote the weights published in [10] by λ_L^i, the ones in [13] by λ_S^i, and the ones published in [11] by λ_N^i; see Table 3. The considerable differences for certain types of configurations between the weights published in [10] and in [13] can mostly be explained by Proposition 3.4. Repeating the arguments that lead to the weights in Table 2, one can show that for $\square \in \{L, S, N\}$ the weight tuple $(\lambda_\square^1, \ldots, \lambda_\square^5)$ belongs to a two-parameter family $(\lambda_\square^1(r,s), \ldots, \lambda_\square^5(r,s))$ of weights with the property that the estimator's asymptotic error is independent of $r \in [0,1]$ and its worst case behavior does not depend on s. Using Lemma 4.1 it would be possible to determine the admissible range of values for s, but it is immaterial for comparison purposes, so we omitted this calculation. The parametric class is of the form as given in Table 2, but with different numerical constants. These constants and the parameters r and s are uniquely determined by the weights $\lambda_\square^1, \ldots, \lambda_\square^5$ and shown in Table 3. In order to put the weights in [11] into perspective, one has to note that here the authors do not require asymptotic unbiasedness for spherical shapes,
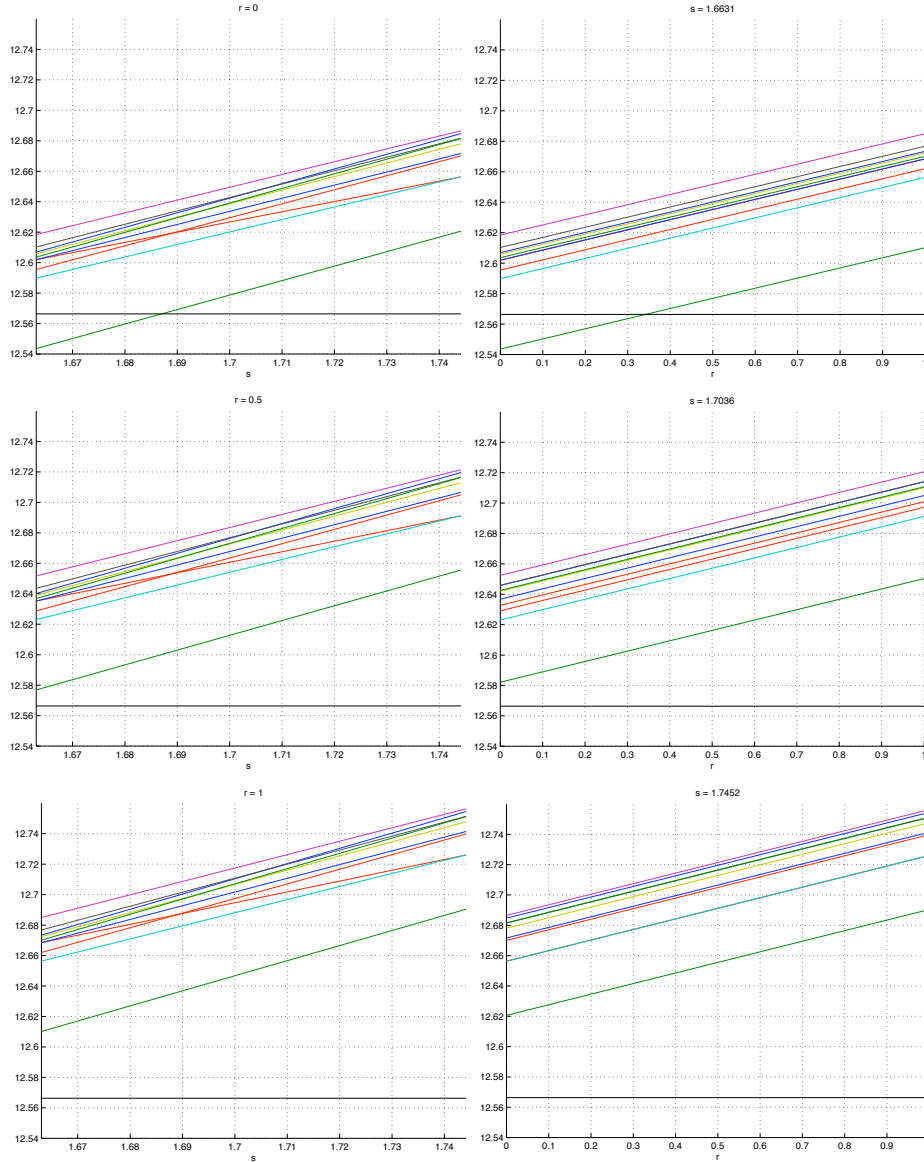
FIGURE 7. Estimation of surface area of a unit ball from a random dig-
itization with lattice spacing 0.1. The surface area estimator is based on
the weights $\overline{\lambda}^i(r,s)$ with $r = 0$, $r = 0.5$ and $r = 1$ as a function of s in
the graphs on the left. Then $s = 1.6631, 1.7036, 1.7452$ are fixed and the
results for all $r \in [0,1]$ are displayed. The second value chosen for s is
simply the mean of the upper and lower boundary of the interval. The
horizontal black line shows the true value of surface area 4π. Note that
maximal relative errors observed in 10 simulations for the different graphs
are 0.9%, 1.2%, 1.5%, 1.0%, 1.2%, 1.5% respectively.

in fact $(1/S(B^2)) \int_{S^2} H^{\Lambda_N}(u)S_2(B^2, du) = 1.0880 =: C_N$, which implies an asymp-
totic bias of more than 8% in the isotropic case. Table 3 shows that the weights
for configurations of type two and four, the values of s and the constant part of
the weight for configurations of type three are fairly similar in [10], [13] and the
approach presented in this paper. It is remarkable that although the weights in [11]
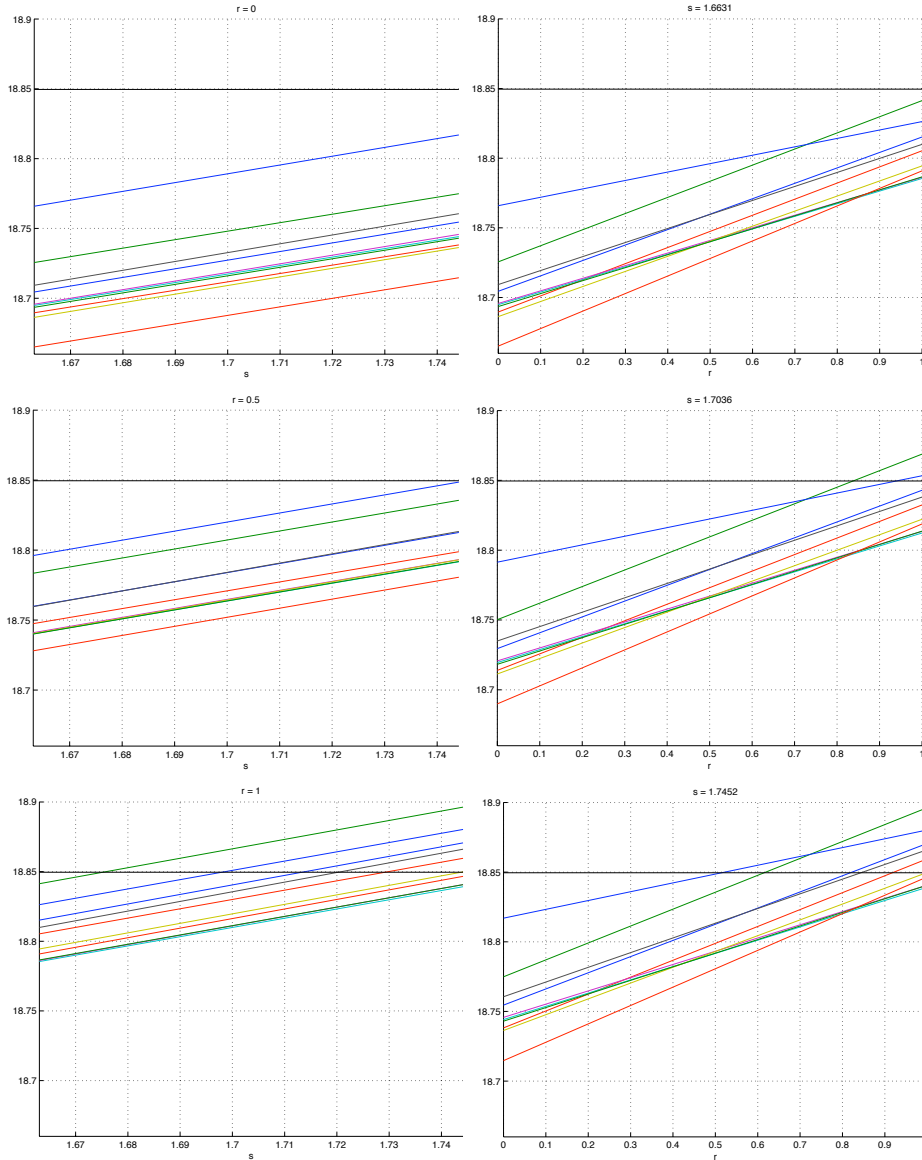
FIGURE 8. Estimation of surface area of a cylinder with radius 1 and height 2 (not aligned with any coordinate direction) from a random digitization with lattice spacing 0.055. As above, the surface area estimator is based on the weights $\overline{\lambda}^i(r, s)$ and the results are displayed for different fixed values of r and s. Again the horizontal black line shows the true value of surface area 6π. The maximal relative errors observed in 10 simulations were 0.4%, 0.3%, 0.2%, 0.4%, 0.3%, 0.2% respectively.

were obtained by a fairly simple approximation method, they are still close to the other three approaches.

The maximal asymptotic relative error for general shapes is 7.3% for the weights suggested in [10] and also for the weights suggested in [13], whereas it is 12.8% for the weights in [11]. In the first two cases the lower error bound is $H_m^{\Lambda_L} = H_m^{\Lambda_S} = 0.927$, whereas the upper bounds differ; they are $H_M^{\Lambda_L} = 1.0205$ and $H_M^{\Lambda_S} = 1.0230$ using the notation as in (9). For the set of weights λ_N^i we have $H_M^{\Lambda_N} = 1.1281$ and $H_m^{\Lambda_N} = 1$.

i	1	2	3	4	5	r	s
$\overline{\lambda}^i$	sr	1.358	$2.352 - sr$	0.960	$s(1-r)$	$[0,1]$	$[1.663, 1.745]$
λ_L^i	$1.272 = sr$	1.338	$1.108 = 2.380 - sr$	0.927	$0.421 = s(1-r)$	0.751	1.693
λ_S^i	$0.752 = sr$	1.318	$1.678 = 2.430 - sr$	0.927	$0.914 = s(1-r)$	0.451	1.666
$C_N^{-1}\lambda_N^i$	$0.399 = sr$	1.300	$2.114 = 2.513 - sr$	0.919	$1.194 = s(1-r)$	0.251	1.593

TABLE 3. Comparison of weights for surface area estimator suggested in [10], [13] and [11].

We proposed a two-parameter family of weights, which leads to an asymptotic minimal worst case error. As we are only assigning (positive) weights to *informative* configurations, one has to be aware that our results only hold for digitizations with good or very good resolution. If there is no a priori information on the shape of the objects considered, there is currently no natural choice for the parameters (r, s). Or, positively formulated, all choices of (r, s) yield good results. If, instead, some knowledge of the shape of the objects is available, it might be advisable to run simulations with similarly shaped objects with known surface area to find the best values for (r, s) in this particular setting.

5. ESTIMATION OF THE SURFACE AREA MEASURE

Let $Z \subset \mathbb{R}^3$ be a compact gentle set and (B, W) a $2 \times 2 \times 2$ configuration. Let N_t be the number of occurrences of (B, W) observed in the digitized set $Z \cap t(U + \mathbb{Z}^3)$. Recall that $t^2 N_t$ is an asymptotically unbiased estimator of the integral

$$\int_{S^2} h_{(B,W)}(n) S_2(Z, dn) \tag{16}$$

by Theorem 2.1. If (B, W) runs through the family of all informative $2 \times 2 \times 2$ configurations, (16) yields at most 50 non-trivial different integrals of $S_2(Z, \cdot)$, due to Proposition 3.4. Clearly, 50 integrals do not determine $S_2(Z, \cdot)$ uniquely. Therefore, a model will be imposed: We assume that $S_2(Z, \cdot)$ belongs to some class of measures parametrized by $\theta \in \Theta$, where $\Theta \subset \mathbb{R}^{50}$. The surface area measure with parameter θ will then be denoted by $S_2^\theta(Z, \cdot)$. Recall that we identified twins and ordered the obtained 58 classes of configurations according to the following scheme:

$$\underbrace{(B_1, W_1), \ldots, (B_8, W_8)}_{\text{type five}}, \underbrace{(B_9, W_9), \ldots, (B_{14}, W_{14})}_{\text{type four}}, \underbrace{(B_{15}, W_{15}), \ldots, (B_{38}, W_{38})}_{\text{type three}},$$
$$\underbrace{(B_{39}, W_{39}), \ldots, (B_{50}, W_{50})}_{\text{type two}}, \underbrace{(B_{51}, W_{51}), \ldots, (B_{58}, W_{58})}_{\text{type one}}.$$

Writing h_j for $h_{(B_j, W_j)}$ we define

$$I_j(\theta) := \int_{S^2} h_j(n) S_2^\theta(Z, dn) \tag{17}$$

for $j = 1, \ldots, 58$.

5.1. **Two simple models for the surface area measure.** The two models considered here are natural extensions of the models in the two-dimensional case discussed in [4]. Throughout this section configurations of type one are not considered as they are asymptotically redundant. For a further discussion of this matter see Sections 5.2 and 5.3.

A discrete model. Let M be the set of all maxima of h-functions. As the maxima of the h-functions of type one coincide with the maxima of the h-functions of type five, we have $M = \{m_1, \ldots, m_{50}\}$. The 50 directions contained in M are those directions where we can expect an outer tangent normal of Z with the highest probability given that the corresponding configuration has been observed. We therefore assume that $S_2^\theta(Z, \cdot)$ is a discrete measure supported by the set M, i.e. it belongs to the model class

$$\mathcal{D}(M) := \Big\{\sum_{j=1}^{50} \theta_j \delta_{m_j} \mid \theta \in \Theta\Big\},$$

where δ_v denotes the Dirac measure at $v \in \mathbb{R}^3$ and $\theta = (\theta_1, \ldots, \theta_{50}) \in \Theta = \mathbb{R}_+^{50}$. Assuming this model, (17) reads

$$I_j(\theta) = \theta_j \bar{h}_{jj}, \tag{18}$$

with $\bar{h}_{jj} := h_j(m_j)$, $j = 1, \ldots, 50$. As all scalars $\bar{h}_{jj}$ are positive, θ is *identifiable* in Θ (θ is uniquely determined by the 50 integrals in (18)), and

$$S_2^\theta(Z, \cdot) = \sum_{j=1}^{50} \frac{I_j(\theta)}{\bar{h}_{jj}} \delta_{m_j}.$$

A model with piecewise constant density. For this alternative model we assume that the surface area measure $S_2^\theta(Z, \cdot)$ has a piecewise constant density with respect to the ordinary surface area measure $\sigma_2(\cdot)$ on the sphere S^2. More precisely, we assume that the density is constant on each $T_j \subset S^2$, where $\{T_j\}$ is a suitable partition of S^2, i.e.

$$S_2^\theta(Z, \cdot) = \sum_{j=1}^{50} \theta_j \sigma_2(\cdot \cap T_j).$$

The partition is constructed by associating to each configuration (B_j, W_j) of type $\tau(j) \neq 1$ a set $T_j \subset S^2$, such that $\{T_j\}_{j=1,\ldots,50}$ forms a non-overlapping partition of S^2. The parameter space for this model is again $\Theta = \mathbb{R}_+^{50}$. For a configuration (B_j, W_j) consider the super-level set

$$\bar{T}_j := \{u \in S^2 \mid h_j(u) \geq \tfrac{1}{2} h_j(m_j)\}.$$

If this configuration is observed, then with high probability the set Z has an outer tangent normal in a direction where h_j takes a high value, hence it is sensible to assume that the normal lies in $\bar{T}_j$. The interiors of the sets $\bar{T}_j$, $j = 1, \ldots, 50$, are disjoint, but unfortunately $\{\bar{T}_j\}_{j=1,\ldots,50}$ is not a partition of S^2; see Figure 9 for illustration.

Of course, one could restrict the support of $S_2^\theta(Z, \cdot)$ to $\cup_j \bar{T}_j$ and work with the resulting model. As we want the model to include multiples of σ_2 (corresponding to isotropic sets Z), this is not an option though. We suggest to enlarge the sets $\bar{T}_j$ for configurations of type $\tau(j) = 3$ only, and to set $T_j = \bar{T}_j$ for configurations of type $\tau(j) \in \{2, 4, 5\}$. For a configuration (B_j, W_j) of type three, we define T_j as the
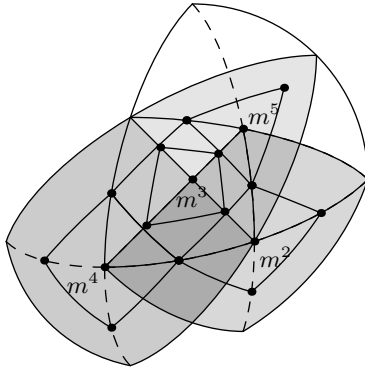
FIGURE 9. Sublevel sets $\{\bar{T}_j\}$ and partition $\{T_j\}$.

spherical convex hull of the midpoints of the arcs that are the edges of its support quadrilateral. Figure 9 illustrates this definition. Similarly to the discrete case we define $\tilde{h}_{jl} := \int_{T_l} h_j(n)\sigma_2(dn)$, $j = 1, \ldots, 50$. We obtain

$$I_j(\theta) = \sum_{l=1}^{50} \theta_l \tilde{h}_{jl}.$$

The so defined linear map from $\Theta = \mathbb{R}_+^{50}$ to $\mathbb{R}^{50}$, $\theta \mapsto (I_j(\theta))_{j=1}^{50}$ is injective as can be shown by an explicit calculation of the rank of the matrix $(\tilde{h}_{jl})_{j,l=1,\ldots,50}$, hence θ is identifiable in Θ.

5.2. **Statistical considerations.** This section describes how to estimate the surface area measure of $Z \subset \mathbb{R}^3$ from observations of $2 \times 2 \times 2$ configurations. The surface area measure is assumed to belong to one of the parametric classes introduced in Section 5.1, so $S_2(Z, \cdot) = S_2^{\theta^0}(Z, \cdot)$ with $\theta^0 \in \Theta = \mathbb{R}_+^{50}$. The parameter will be estimated by counting informative boundary configurations. Suppose that we have observed $N_{t,j}$ configurations of class $j \in T := \{1, \ldots, 58\}$ in the digitization of Z. Due to Theorem 2.1 we have, for $j \in T$,

$$t^2 N_{t,j} \approx n_j I_j(\theta) = n_j \int_{S^2} h_{(B_j, W_j)} \, dS_2^\theta(Z, \cdot), \tag{19}$$

for sufficiently small $t > 0$. Recall that the vector $n \in \{1, 2\}^{58}$ is defined in Section 4. Relations (19) suggest to estimate θ based on counts of *all* informative configurations. We will, however, not include configurations of type one in the estimation procedure. This is motivated by the following facts. Firstly, counts of type one configurations are asymptotically redundant. The right hand side of (19) for $\tau(j) = 1$ is a linear combination of corresponding integrals from other classes due to Proposition 3.4. Secondly, the asymptotic linear dependence just mentioned even holds for positive t in certain special cases. For flat surfaces a careful analysis of the arguments in [10] shows that Proposition 3.4 actually holds for the observed $N_{t,j}$: for each class of configurations j_1 of type one, there is a configuration j_5 of type five and three classes of configurations j_{31}, j_{32}, j_{33} of type three such that

$$n_{j_1}^{-1} N_{t,j_1} = n_{j_5}^{-1} N_{t,j_5} + n_{j_{31}}^{-1} N_{t,j_{31}} + n_{j_{32}}^{-1} N_{t,j_{32}} + n_{j_{33}}^{-1} N_{t,j_{33}}. \tag{20}$$

If we work with the discrete model for the surface area measure, then Z is a polygonal set, and (20) holds approximately for sufficiently high resolution. Finally, the

simulation studies in Section 5.3 give evidence that type one configurations do not contribute decisively, even for a model of a piecewise constant density.

As configurations of type one are collected in classes 51 to 58, we define the reduced index set $T_r := \{1, \ldots, 50\}$, which contains all indices for configurations of type five, four, three and two. (19) implies that the vector $\hat{w} := (N_{t,j})_{j \in T_r}$ is an empirical approximation for $w(\theta) = (w_j(\theta))_{j \in T_r}$ with

$$w_j(\theta) := t^{-2} n_j I_j(\theta), \qquad j \in T_r.$$

To determine θ, we therefore minimize an appropriately chosen distance of $\hat{w}$ and $w(\theta)$. Following [5] we chose the I_1-divergence from $\hat{w}$ to $w(\theta)$, which plays an important role in information theory. Recall that the I_1-*divergence* of two vectors $\hat{p} = (\hat{p}_1, \ldots, \hat{p}_k)$ and $p = (p_1, \ldots, p_k)$ in $\mathbb{R}_+^k$ (where p has non-zero components) is given by

$$I_1(\hat{p}, p) = \sum_{i=1}^{k} \left(\hat{p}_i \log \frac{\hat{p}_i}{p_i} - \hat{p}_i + p_i \right).$$

If $\hat{p}$ and p are probability vectors, $I_1(\hat{p}, p)$ coincides with the Kullback-Leibler divergence of $\hat{p}$ and p. Further details on this pseudo-distance can be found in [9] and [5]. Similar to [4], we set

$$L(\theta) := \sum_{j \in T_r} \left(N_{t,j} \log I_j(\theta) - t^{-2} n_j I_j(\theta) \right) \tag{21}$$

and use the fact that minimization of $I_1(\hat{w}, w(\theta))$, $\theta \in \Theta$, is equivalent to solving the problem

$$\begin{aligned} \text{maximize} \quad & L(\theta) \\ \text{subject to} \quad & \theta \in \Theta. \end{aligned} \tag{22}$$

Note that the I_1-divergence is positive definite. Hence, if there is a $\hat{\theta} \in \Theta$ such that $\hat{w} = w(\hat{\theta})$, then $\hat{\theta}$ is a solution of (22). The condition $\hat{w} = w(\hat{\theta})$ is equivalent to

$$I_j(\hat{\theta}) = \frac{t^2}{n_j} N_{t,j}, \qquad j \in T_r. \tag{23}$$

For the proposed model of a *discrete measure* (see Section 5.1) equation (23) yields the unique solution $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_{50}) \in \Theta = \mathbb{R}_+^{50}$ with

$$\hat{\theta}_j = \frac{t^2}{n_j \bar{h}_{jj}} N_{t,j}, \qquad j \in 1, \ldots, 50.$$

An estimator of the true surface area measure $S_2^{\theta^0}(Z, \cdot)$ is therefore given by

$$S_2^{\hat{\theta}}(Z, \cdot) = \sum_{j=1}^{50} \frac{t^2 N_{t,j}}{n_j \bar{h}_{jj}} \delta_{m_j}.$$

The vector $\hat{\theta}$ depends linearly on the configuration counts, so Theorem 2.1 implies that $\mathbb{E}\hat{\theta}$ converges to θ^0, as $t \to 0+$. Hence, $S_2^{\hat{\theta}}(Z, \cdot)$ is asymptotically unbiased for the true surface area measure $S_2^{\theta^0}(Z, \cdot)$, in the sense that

$$\lim_{t \to 0+} \mathbb{E}S_2^{\hat{\theta}}(Z, A) = S_2^{\theta^0}(Z, A), \qquad \text{for all measurable } A \subset S^2. \tag{24}$$

All measures involved here have a finite prescribed support, so (24) is equivalent to the statement that $\mathbb{E}S_2^{\hat{\theta}}(Z,\cdot)$ converges weakly to the mean normal measure of Z, as $t \to 0+$.

For the model of a *piecewise constant density*, let $\tilde{H} := (\tilde{h}_{jl})_{j,l \in T_r}$. Equation (23) yields the linear system

$$(\tilde{H}\hat{\theta})_j = \frac{t^2}{n_j}N_{t,j}, \qquad j \in T_r,$$

which does not necessarily have a positive solution for all possible values of $(N_{t,j})_{j \in T_r}$. Therefore it may be necessary to solve the given convex optimization problem (22) numerically.

Under both models, the obtained estimator $S_2^{\hat{\theta}}(Z,\cdot)$ need not satisfy the centroid condition

$$\int_{S^2} n\, S_2^{\hat{\theta}}(Z,dn) = 0, \tag{25}$$

even though the true surface area measure does satisfy it; see (2). If $S_2^{\hat{\theta}}(Z,\cdot)$ is the estimator in the discrete model, this condition holds asymptotically in mean,

$$\lim_{t \to 0+} \mathbb{E} \int_{S^2} n\, S_2^{\hat{\theta}}(Z,dn) = 0,$$

due to (24). In certain situations it is desirable that also $S_2^{\hat{\theta}}(Z,\cdot)$ satisfies the centroid condition. If, for instance, the set Z is to be analyzed by Monte Carlo simulations, one seeks to find a (random) set whose (specific) surface area measure coincides with the estimated $S_2^{\hat{\theta}}(Z,\cdot)$. Equation (2) shows that this is only possible, if the centroid condition holds. The suggested estimation procedure can easily be modified to yield an estimator satisfying (25). The optimal parameter $\hat{\theta}$ is again obtained solving (22), but now with

$$\Theta := \{\theta \in \mathbb{R}_+^{50} |\int_{S^2} n\, S_2^{\hat{\theta}}(Z,dn) = 0\}.$$

In both models, this introduces a linear constraint in (22). In general, even for the discrete measure, (22) must now be solved numerically. We do not include the centroid condition in the following simulation studies.

5.3. **Simulation results. Detection of anisotropy.** For detecting anisotropy or isotropy of particles we model the surface area measure $S_2^{\theta}(Z,\cdot)$ by a piecewise constant density as proposed above. The unit ball B^2 has surface area measure $S_2^{\theta^0}(B^2,\cdot)$ with parameter $\theta_j^0 = 1$ for all $j \in \{1,\dots,50\}$. We cannot expect equation (20) to hold even approximately for any positive resolution as B^2 is not locally flat. In order to justify the approach that omits configurations of type one in the estimation procedure we do the estimation twice: once like in the previous section, once including type one configurations. In this particular example the inclusion of type one configurations even worsens the estimator.

The lattice spacing was $t = 0.02$ in all simulations. We maximized $L(\theta)$ as given in (21) and obtained parameters $\hat{\theta}^1$. Including type one configurations, we also minimized the I_1-divergence between $\hat{W} = (N_{t,j})_{j \in T}$ and $W(\theta) = (w_j(\theta))_{j \in T}$ and obtained parameters $\hat{\theta}^2$. For both optimizations [12] was used. The largest deviations from the true value 1 are $\max_{j \in \{1,\dots,50\}} \hat{\theta}_j^1 = 1.0741$ and $\min_{j \in \{1,\dots,50\}} \hat{\theta}_j^1 = 0.9495$,
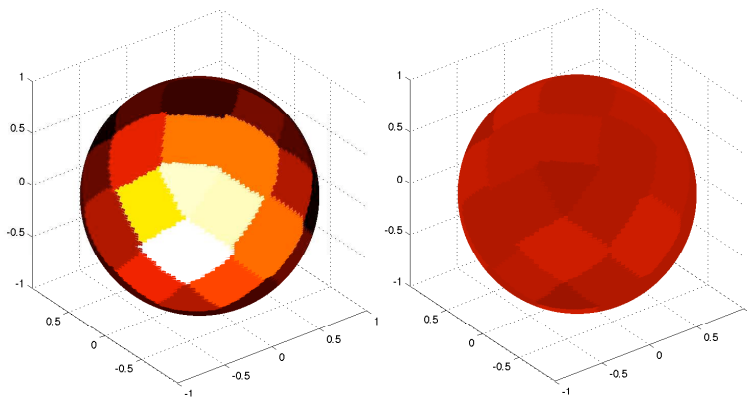
FIGURE 10. Surface area measure estimates of an oblate spheroid (left) and
a unit sphere (right) based on the model of a piecewise constant density.
The spheroid has half axis $a = 0.9$, $b = c = 1.1$. It is not aligned with the
coordinate axis. The lighter the color the larger the value of the coefficient
estimated for the corresponding spherical triangle or quadrangle.

but $\max_{j \in \{1,\dots,50\}} \hat{\theta}_j^2 = 1.0886$ and $\min_{j \in \{1,\dots 50\}} \hat{\theta}_j^2 = 0.7686$. The maximal overesti-
mations are comparable in both cases, but the underestimations differ substantially.
25 components of $\hat{\theta}^2$ are less than 0.95 and nine are larger than 1.05, whereas only
three components of $\hat{\theta}^1$ are larger than 1.05 and one is smaller than 0.95. We ob-
tained similar results for all simulated digitizations of balls. Therefore we suggest
excluding the counts of configurations of type one for the estimation procedure.

  We consider a spheroid which might be misjudged as isotropic, if only examined by
eyesight. The estimated parameters of its surface area measure based on the model
of a piecewise constant density clearly show its anisotropy though. We simulated an
oblate spheroid with half axis $a = 0.9$ and $b = c = 1.1$. To avoid alignment with the
coordinate axis, it was rotated around $(1, 1, 1)^T$ by an angle of $2\pi/7$. The coefficients
for the surface area measure were obtained by maximizing $L(\theta)$. The anisotropy of
Z is clearly visible in Figure 10. Dark color of a spherical polygon corresponds to
a small associated parameter value. The maximal component of $\hat{\theta}$ for the spheroid
is 1.7391, the minimal one is 0.7373. Strictly speaking, the surface area measure
of a spheroid with different half axis is not in the model of a piecewise constant
density, but for the purpose of anisotropy detection this model yields a reasonable
approximation. Naturally these simulation results on the detection of anisotropy are
not very general. We mention them here as we think that they can be extended to
systematic tools for detecting, quantifying or even statistically testing for anisotropy.
There are many important applications of the discrete model as well. We omit a
detailed discussion here as they are similar to the 2D case, which is described in [3].

REFERENCES

[1] D. Coeurjolly, F. Flin, O. Teytaud, L. Tougne, Multigrid convergence and surface area esti-
    mation, in: Theoretical Foundations of Computer Vision Geometry, Morphology, and Com-
    putational Imaging. LNCS, vol. 2616, Springer, Berlin, 2003.
[2] P. Gutkowski, E. B. V. Jensen, M. Kiderlen, Directional analysis of digitized three-dimensional
    images by configuration counts, J. Microsc. 216 (2004) 175–185.
[3] E. B. V. Jensen, M. Kiderlen, Directional analysis of digitized planar sets by configuration
    counts, J. Microsc. 212 (2003) 158–168.

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

[5] M. Kiderlen, A. Pfrang, Algorithms to estimate the rose of directions of a spatial fibre system, J. Microsc. 219 (2005) 50–60.

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

[7] R. Klette, A. Rosenfeld, Digital Geometry: Geometric Methods for Digital Picture Analysis, Morgan Kaufmann, San Francisco, 2004.

[8] R. Klette, H. Sun, Digital planar segment based polyhedrization for surface area estimation, in: C. Arcelli, L. Cordella, G. Sanniti di Baja (eds.), Visual Form 2001, LNCS, vol. 2059, Springer, Berlin, 2001.

[9] F. Liese, I. Vadja, Convex statistical distances, Teubner, Leipzig, 1987.

[10] J. Lindblad, Surface area estimation of digitized 3d objects using weighted local configurations, Image Vis. Comput. 23 (2005) 111–122.

[11] J. Lindblad, I. Nyström, Surface area estimation of digitized 3D objects using local computations, in: Discrete Geometry for Computer Imagery : 10th International Conference, DGCI 2002, Springer, Berlin, 2002.

[12] J. Löfberg, YALMIP: A Toolbox for Modeling and Optimization in MATLAB, in: Proceedings of the CACSD Conference, Taipei, Taiwan, 2004.

[13] K. Schladitz, J. Ohser, W. Nagel, Measuring intrinsic volumes in digital 3d images, DGCI (2006) 247–258.

[†]Department of Mathematics, ETH Zurich, 8092 Zurich, Switzerland, Tel. 0041 44 632 09 86.

E-mail address: johanna.ziegel@math.ethz.ch

[‡]Department of Mathematical Sciences, University of Aarhus, Ny Munkegade Build. 1530, DK-8000 Aarhus C, Denmark. Tel. 0045 89 42 35 72, Fax 0045 86 13 17 69.

E-mail address: kiderlen@imf.au.dk